

# Providing ART-DECOR ValueSets via FHIR Terminology Servers – A Technical Report

Joshua WIEDEKOPF <sup>a, b, 1</sup>, Cora DRENKHAHN <sup>a, b</sup>, Hannes ULRICH <sup>a, b</sup>,  
Ann-Kristin KOCK-SCHOPPENHAUER <sup>b</sup> and Josef INGENERF <sup>a, b</sup>  
<sup>a</sup> *Institute of Medical Informatics, University of Lübeck, Lübeck, Germany*  
<sup>b</sup> *IT Center for Clinical Research, University of Lübeck, Lübeck, Germany*

**Abstract.** To ensure semantic interoperability within healthcare systems, using common, curated terminological systems to identify relevant concepts is of fundamental importance. The HL7 FHIR standard specifies means of modelling terminological systems and appropriate ways of accessing and querying these artefacts within a terminology server. Hence, initiatives towards healthcare interoperability like IHE specify not only software interfaces, but also common codes in the form of value sets and code systems. The way in which these coding tables are provided is not necessarily compatible to the current version of the HL7 FHIR specification and therefore cannot be used with current HL7 FHIR-based terminology servers. This work demonstrates a conversion of terminological resources specified by the Integrating the Healthcare Initiative in the ART-DECOR platform, partly available in HL7 FHIR, to ensure that they can be used within a HL7 FHIR-based terminological server. The approach itself can be used for other terminological resources specified within ART-DECOR but can also be used as the basis for other code-driven conversions of proprietary coding schemes.

**Keywords.** Health Information Exchange; Biological Ontologies; Terminology as Topic; Vocabulary, Controlled

## 1. Introduction

### 1.1. Background

Usage of a common set of terms for concepts is fundamental to the domain of medical informatics. However, since medicine is a very broad field with competing requirements in many domains, this need is present at many levels and in many systems throughout any technical system. For example, medical concepts in electronic patient records should ideally be represented using codes in appropriate reference terminologies like SNOMED CT [1] wherever possible. Furthermore, suitable terminology is a prerequisite when exchanging electronic documents within a regional or national health informatics infrastructure, encoding, for example, the type of document being exchanged, or the qualifications of the author.

---

<sup>1</sup> Corresponding Author: Joshua Wiedekopf, Institute of Medical Informatics and IT Center for Clinical Research, University of Lübeck, Ratzeburger Allee 160, 23564 Lübeck, Germany; E-Mail: j.wiedekopf@uni-luebeck.de.

Especially when coded data is exchanged across institutional boundaries, alignment towards a common set of codes and concepts is required to ensure that all stakeholders use machine-interpretable representation of those concepts, but also to ensure completeness according to the specified set. In this regard, the methods and protocols defined by the Integrating the Healthcare Enterprise (IHE) initiative are noteworthy as they define so-called Integration Profiles. These provide standards both for data exchange across institutions, and for coding schemes representing the concepts used within the scope of data exchange, following established standard terminology where appropriate [2,3].

Another very important challenge is the way the coding schemes are defined and represented on a technical and semantic level. The use of terminology servers that provide access to such concepts is a very useful and recommended approach [4,5]. Within the HL7 Fast Healthcare Interoperability Resources (HL7 FHIR) specification [6], the Terminology Module [7] provides for both a data model and an access layer for terminological resources. This specification has been proven useful in the development of terminology servers (TS) [4,5]. Additionally, IHE has defined an integration profile for the utilization of FHIR terminology resources from non-FHIR-native systems and applications, which is currently in trial use [8].

The aforementioned resources specified by the IHE initiative were created using a tool called ART-DECOR (“**A**dvanced **R**equirement **T**ooling using **D**ata **E**lements, **C**odes, **O**IDs and **R**ules”), which has been of particular importance to support the development of artefacts using HL7 Clinical Documents Architecture (HL7 CDA) [9,10]. The ART-DECOR system is critical for the creation of HL7 CDA templates and implementation guides, and is being used to great effect in e.g. the ELGA project in Austria [11], but it is not restricted to these applications.

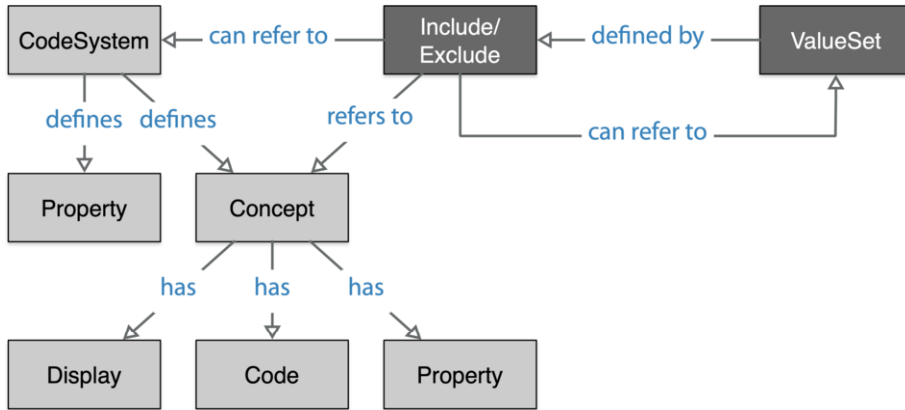
However, the support for HL7 FHIR-compatible terminology is currently still lacking. In particular, the dichotomy of code systems and value sets is not supported to the degree required for exporting complete HL7 FHIR artefacts. This hinders the utilization of the terminology in ETL processes that transform data into interoperable resources, since no terminology server can support the respective algorithm. Crucially, a FHIR Terminology Server is not only important for FHIR-based processes, but also within other pathways, e.g., those exporting openEHR templates.

Semantic interoperability, in addition to syntactic interoperability is a major focus of the medical informatics community, especially in the domain of secondary use of routinely captured data for research purposes. The resulting effort required for semantic interoperability is less the fewer systems need to reach consensus on the transported meaning. Hence, our experience in bridging the gap between an ART-DECOR, an established system central to the research-focused standardization initiatives in Germany, (such as the Medical Informatics Initiative), and the HL7 FHIR Terminology Services (TS) specification can benefit other scientists working within nation-wide standardization efforts. Especially with the strong focus on FHIR-based solutions in these projects, the availability of FHIR TS resources is extremely valuable.

### *1.2. Objective and Requirements*

The coding schemes were specified by IHE Germany in a way that is broadly compatible to the FHIR specification, but not using FHIR itself. The objective of this work is to make those resources compliant to the current version R4 of the FHIR specification, so that they can be processed with the available operations of a FHIR-compliant

terminology server. These resources shall be provided to the maintainers of the terminological artefacts, so that they can be included within ART-DECOR and be disseminated to users of the data models.



**Figure 1.** The relationship between CodeSystem (CS) and ValueSet (VS) as specified in HL7 FHIR R4. CS-specific elements have lighter shades, VS has darker shades. Both resources require associated metadata not shown here. CS define concepts, while VS refer to concepts defined in CS. Each concept is assigned code that is unique within the CS.

A commercial service by CSIRO called *Ontoserver* [4] was available on-site and used as the target TS. However, since the implemented method outputs standard-compliant FHIR resources, the converted terminology is compatible with all other standards-compliant FHIR TS implementations.

The challenge of this modeling approach is to make the resources as semantically identical as possible to the IHE source data and to minimize the discrepancy between the modeling of the available artifacts and the FHIR specification as much as possible.

## 2. State of the art

### 2.1. Related Work

The IHE ValueSets were specified using the proprietary ART-DECOR system [9,10]. This tooling platform is centered around the development of resources utilizing the HL7 CDA specification for medical documents. Because terminological modelling is an important part of the CDA specification, ART-DECOR can be utilized to also model and provide those kinds of resources.

Besides specifying artefacts for HL7 CDA, ART-DECOR is still widely used in the creation of implementation-agnostic data models [9,12,13], which can then be modelled using specifications like HL7 FHIR. This approach was used to great effect in the modelling of the core dataset of the Medical Informatics Initiative in Germany [14].

ART-DECOR is capable of exporting one of the two kinds of resources required by the HL7 FHIR (FHIR ValueSets) specification. However, the second part is fundamental, because of the dichotomy between code systems and value sets. The relationship between

those two kinds of resources, as well as the support by the ART-DECOR system, is presented in Figure 1.

Other approaches for providing terminological resources compliant to the FHIR specification do exist. This includes manual content creation tooling, such as CSIRO's *Snapper* application [15]. On the other hand, creating FHIR resources programmatically is also possible, utilizing appropriate software libraries [5,16]. A third approach is used for complex reference terminologies like SNOMED CT, which are exposed by TS using the FHIR Terminology API, but with custom indexes and implementations [4].

We are not aware of other approaches to extracting concepts to FHIR resources from the ART-DECOR platform. Furthermore, the data available from ART-DECOR are normative, i.e., other representations and sources might not be entirely consistent with the standard dataset.

## 2.2. Shortcomings

The shortcoming of the existing terminological modelling within the IHE ValueSets [3] using ART-DECOR lies in the missing support of the dichotomy of code system definition and use-case specific subsets, being fundamental to the HL7 FHIR specification of terminological resources.

The HL7 FHIR terminology module [7] revolves around the two resource types *CodeSystem* and *ValueSet*. The resource type *CodeSystem* (CS) is used to define concepts with associated codes and properties. These concepts are the atoms of coded information within FHIR resources and definitions, and every concept has a code that is unique within the CS. Together with the canonical identifier of the CS in the form of a Uniform Resource Identifier (URI), and the version of the resource, the code-canonical-version tuple should be globally unique.

Concepts have associated display labels, formal definitions, and can be assigned both standard and custom properties. Of note is the possibility to specify code hierarchies within the CS in two ways: First, concepts can have nested child concepts, which only allows for monohierarchical relationships where every child concept has exactly one parent. This approach suffices for classifications like ICD-10-GM. The second approach uses properties defined on the concepts, so that polyhierarchical relationships are also possible, which are required for more complex terminologies like SNOMED CT.

The resource type *ValueSet* (VS) defines the second part of the dichotomy. A VS is a use-case dependent subset of at least one CS. For example, consider the diagnosis classification ICD-10-GM. When utilizing a FHIR TS e.g., to provide coding assistance to end-users, different subsets, like "diabetes complications" or "cardiovascular diseases" could be used depending on the use. Since VS can include codes from multiple other CS, or even other VS, they support reference terminologies like SNOMED CT, with the possibility of adding custom codes where necessary.

Especially when utilizing FHIR profiles to model jurisdiction-specific use cases, this dichotomy is essential because FHIR requires that coded elements always be bound to a VS, not directly to a CS. Binding appropriate terminology to the resource definitions is one of the most important part in profiling FHIR, since ideally, all participants in data exchange using the profile in question should be able to encode all codable information in the same way. For example, when profiling the FHIR resource type *Condition* for use within a cardiovascular research study, the attribute *Condition.code* has to be bound to an appropriate subset of cardiovascular codes from the space of all available codes for conditions, allowing for representing the exact diagnosis of the patient without

information loss. However, the fundamental properties, like hierarchies or annotations of the code system are not and should not be present in the ValueSet according to the specification, because they exist independent of the use case in questions.

As such, it is not possible to extract FHIR CS from the system natively, even though the export of FHIR VS is possible, as shown in Figure 1. When exporting the VS to FHIR and loading them into a FHIR TS, some operations on the VS are not possible, depending on the implementation of the FHIR operation specification.

Ontoserver is not able to successfully perform the *expansion* operation on VS where an underlying CS is not present in the index. This operation is however fundamental when working with VS, because it enumerates all the concepts present in that VS, with an optional filter. Especially when VS are defined intensionally, i.e. using filter operations on the contained CS, rather than extensionally, i.e. by picking codes by hand, the full list of contained codes can only be retrieved when querying the expansion of that resource [5].

### 3. Concept

To address the observed shortcomings of the existing modelling, the concepts defined in these VS must be extracted to form new CS instances. Since exporting FHIR VS is supported by ART-DECOR (but requires manual and error-prone parametrization by the maintainers for every resource), VS do not have to be extracted or altered.

While ART-DECOR provides FHIR VS for the resources in question, the native export format was used as the input of the system, since the FHIR resource is generated based on maintainer input and does not include all of the data available in the native format.

It is assumed that the range of defined codes for each referenced CS is enumerated across all the VS within the respective ART-DECOR project. CS may be referenced by more than one VS, so that already-converted CS need to be updated with new concepts if they were already processed in another VS. Conflicts between the display labels of a concept or between the metadata of the CS may occur and need to be addressed, relying on user input. Since some IHE VS feature a code hierarchy, this should be reflected in the output CS. For the references to be functional within the VS, the output CS should carry the identifiers and canonical URLs available in the CS metadata of the VS.

A very important first step lies in the identification of CS that are already available from other sources. Those include CS as specified by HL7 International (which are generally available in FHIR R4 format), those by the DICOM standard used in digital imaging [17], or by the National Association of Statutory Health Insurance Physicians (KBV) in Germany [18], as well as standard terminologies like SNOMED CT and LOINC. Those code systems should be excluded from conversion and the available representation should be validated for compatibility with the IHE VS.

The exclusion of available CS could be based on the Object Identifier (OID) assigned to the resource and referenced in the IHE VS. Because of the hierarchical organization of these identifiers, based on organizational affiliation, a prefix-based matcher can effectively filter those resources during the conversion. For example, HL7 International has been assigned the root OID 2.16.840.1.113883, so that all OIDs of terminological resources start with that prefix and can be reasonably assumed to be already available in FHIR R4 format. The same procedure of finding and excluding

organizational root OIDs is applicable for the other CodeSystems that are elsewhere defined.

Next, the validity of the output resources must be verified manually and using type checkers utilizing the machine-readable FHIR specification. During the development process, intermediate outputs as well as the external dependencies need to be uploaded in an automated fashion to the TS deployment, so that the respective operations can be called and verified.

## 4. Implementation

### 4.1. Solution

The proposed system was implemented as a console application implemented in Python 3.x. Python was chosen in part due to its powerful dictionary manipulation tools, which is extremely useful for processing the VS in the ART-DECOR-native JSON representation. The main dependency of the script is the package *fhir.resources* [16], which is used to ensure the validity of the output resources with regards to FHIR R4.

To facilitate script-driven conversion, all options of the program are exposed as (optional) arguments on the command line. Some arguments, namely the input location and the output directory path, are prompted interactively if not provided via command line arguments. The input of the script can either be provided by URL or as a reference to a local file. In both cases, the input is provided from ART-DECOR by means of an HTTP API.

Once the ValueSet is loaded into memory, metadata are extracted and empty CS are created as appropriate with the respective identifier, canonical URL and other metadata, if not excluded by the OID matcher described above. Next, the concepts are enumerated and added to the respective CS instance. Since codes may follow a hierarchy, the attribute *level* is observed. If this value is not equal to zero, the concept has a (single) parent concept that must be identified. While these parent codes were first chosen interactively from the subset of codes at the next-higher level within the same code system, it was discovered that the order of codes within the ART-DECOR is stable, and hierarchy is modelled by the order of those codes in combination with the *level* attribute. This selection can thus be entirely automated, but this behavior can be overridden with a command line flag if desired, restoring the manual matching.

The *designation* property and the deprecated status are mapped over from the ART-DECOR output as required. After conversion of the respective VS, the output directory is checked for a previous version of the respective CS. If found, the “new” instance is merged into the previous one, while checking for conflicts in the display labels and parent relationships. The updated code system is written out in JSON format.

To verify the validity of the output, shell scripts were written using Bash to upload the VS and dependencies and to verify that the expansion operation is successful. To this end, the shell tools *cURL* [19] and *Firely Terminal* [20] were used for HTTP operations and FHIR operations. The *jq* tool enabled filtering JSON files using a simple query language [21]. This automation suite greatly simplified the development and validation process.

## 4.2. System in Use

The proposed system was successfully used to transform the second ballot version of the IHE Germany ValueSets to HL7 FHIR R4, and successfully used on an Ontoserver instance. CS available from other sources were identified and uploaded. Sources and resources are available to the public and can be freely used and modified<sup>2</sup>.

Of note is one CS referenced within the *IHEXDSformatCodeDE* VS, called *DICOM UID*. This refers to the registry of OIDs used in Service-Object Pair classes within the DICOM standard, which is not available in HL7 FHIR format so far. However, it is freely available in the HTML version of the DICOM specification [17]. This representation was edited in a web browser using development tools to remove unrelated HTML markup. After saving the edited website to a file, it was further cleaned-up using a spreadsheet application and exported to a plain-text CSV file. This file was loaded into the terminology authoring toolkit Snapper [15] to populate a new CodeSystem instance.

The system worked well on all VS in version 2 of the IHE Germany VS, and was also tested with those defined by the Medical Informatics Initiative Core Data Set [14] for verification of the system's portability to other initiatives and use cases. In testing the transformation, some errors and discrepancies were identified in the IHE VS and reported as ballot comments for the upcoming third version.

Hence, a TS that was populated with the resources can provide all the operations specified by the FHIR R4 Terminology Package to consumers and can facilitate the use of these terminologies by other programs.

All in all, 17 VS were contained in the IHE VS version 2, from which 16 CS were extracted. The VS reference four CS defined elsewhere, of which one is the DICOM OID registry. Additionally, the VS reference two other VS by DICOM.

## 5. Lessons learned

The given requirements have been met appropriately. All the VS within the IHE Germany ValueSets for XDS have been converted to FHIR R4-compliant CS resources. The implementation in Python was straightforward, which illustrates the developer-focused approach of HL7 FHIR in action. In particular, the use of a FHIR structures package within Python would not have been necessary but made sure that the resources generated are indeed compliant with the current specification.

Another noteworthy lesson that has been learned is the ease of calling terminology services from automated test scripts. This automation using command line tools and scripting further ensured the compliance and correctness, while reducing the implementation effort.

Regarding the manual creation of the DICOM UID registry CodeSystem, the use of spreadsheet software and a specialized content creation tool for terminology resources (CSIRO's *Snapper* [15]) was an effective choice for quickly transforming the HTML specification, which is much more user-friendly than implementing scripts to achieve the same thing from code. This demonstrates that tooling for creating FHIR resources from a range of specification and distribution formats is available.

---

<sup>2</sup> <https://github.com/itcr-uni-luebeck/artdecor-vs-to-cs>

## 6. Conclusion

Our work demonstrates that conversion of terminological resources in proprietary formats to the dominant standard, HL7 FHIR, is feasible and sensible. As long as these resources are provided in a format that can be made to conform to the CodeSystem/ValueSet dichotomy, a FHIR representation should capture the same information in an easy-to-use specification framework.

However, this approach necessitates the propagation of changes in the FHIR resources to the maintainers, and that changes in the upstream by the same maintainers result in new version of FHIR resources. Ideally, as the medical informatics community increasingly embraces the FHIR specification, the dominant format of terminology should also become FHIR. As the need for support of terminology servers is widely recognized throughout the Medical Informatics research community [1,4,7,22], the FHIR specification serves as a solid and tested foundation for the implementations of such servers, and crucially also enables the use of these services independent of the broader FHIR specification [4]. Hence, when FHIR is supported by content creation tools, great care should be taken that all concepts and artefacts referenced by these projects are available in a suitable representation. In the meantime, syntactic mappings such as the one described in this paper can fill the gap and make terminology usable in the modern age of semantic healthcare interoperability.

## Declarations

*Conflict of Interest:* The authors declare that there is no conflict of interest.

*Contributions of the authors:* JW implemented the approach described. CD and HU supported the development of the algorithm. The manuscript was written by JW. All authors have read and revised the manuscript.

*Acknowledgement:* This work was supported by the HiGHmed project within the Medical Informatics Initiative, funded by the Federal Ministry of Education and Research (grant number 01ZZ1802Z).

## References

- [1] Drenkhahn C, Ingenerf J. Referenzterminologie SNOMED CT. Forum der Medizin\_Dokumentation und Medizin\_Informatik 2020;22:81–4.
- [2] IHE Deutschland e.V. ART-DECOR – XDS Value Sets für Deutschland, Version 2 2018. <https://art-decor.org/art-decor/decor-valuesets--ihede-> (accessed April 2, 2021).
- [3] IHE Deutschland e.V. Value Sets für Aktenprojekte im deutschen Gesundheitswesen. 2018.
- [4] Metke-Jimenez A, Steel J, Hansen D, Lawley M. Ontoserver: a syndicated terminology server. J Biomed Semant 2018;9. <https://doi.org/10.1186/s13326-018-0191-z>.
- [5] Benson T, Grieve G. Principles of Health Interoperability. Third edition. Springer International Publishing; 2016. <https://doi.org/10.1007/978-3-319-30370-3>.
- [6] Bender D, Sartipi K. HL7 FHIR: An Agile and RESTful approach to healthcare information exchange. Proc. 26th IEEE Int. Symp. Comput.-Based Med. Syst., IEEE; 2013, p. 326–31.
- [7] HL7 International. Terminology Module – HL7 FHIR R4 Specification 2019. <https://www.hl7.org/fhir/terminology-module.html> (accessed March 24, 2021).



- [8] IHE - ITI Technical Committee. Sharing ValueSets, Codes and Maps (SVCM). IHE Wiki 2020. [https://wiki.ihe.net/index.php/Sharing\\_ValueSets,\\_Codes\\_and\\_Maps\\_\(SVCM\)](https://wiki.ihe.net/index.php/Sharing_ValueSets,_Codes_and_Maps_(SVCM)) (accessed July 4, 2021).
- [9] Heitmann K-U, FHL P, Germany H, Team DE. ART-DECOR: 'An open-source tool bridging the chasm between clinicians and health IT.' HL7 Int News 2014;1-4.
- [10] Boufahja A, Heitmann K, Poiseau E. Model-based Validation of HL7 CDA R2 Documents and Implementation Guides Using Gazelle ObjectsChecker and ART-DECOR®. 2016.
- [11] Ott S, Rinner C, Duftschmid G. Expressing Patient Selection Criteria Based on HL7 V3 Templates Within the Open-Source Tool ART-DECOR. *Stud Health Technol Inform* 2019;260:226-33.
- [12] Ammon D, Bietenbeck A, Boeker M, Ganslandt T, Heckmann S, Heitmann K, et al. Der Kerndatensatz der Medizininformatik-Initiative - Spezifikation in ART-DECOR und als FHIR-Profiles. *HL7-Mitteilungen* 2019;43:25-8.
- [13] Bild R, Bialke M, Buckow K, Ganslandt T, Ihrig K, Jahns R, et al. Towards a comprehensive and interoperable representation of consent-based data usage permissions in the German medical informatics initiative. *BMC Med Inform Decis Mak* 2020;20. <https://doi.org/10.1186/s12911-020-01138-6>.
- [14] Ganslandt T, Boeker M, Löbe M, Prasser F, Schepers J, Semler S, et al. Der Kerndatensatz der Medizininformatik-Initiative: Ein Schritt zur Sekundärnutzung von Versorgungsdaten auf nationaler Ebene. *Forum Med-Dok Med-Inform* 2018;20:17-21.
- [15] CSIRO. Snapper. 2021.
- [16] Islam N. fhir.resources – Python package, version 6.1.0. 2021. <https://pypi.org/project/fhir.resources/#history>.
- [17] Medical Imaging & Technology Alliance. DICOM S3.16 2021a – Content Mapping Resource – DICOM Specification 2021. <http://dicom.nema.org/medical/dicom/current/output/html/part16.html> (accessed April 4, 2021).
- [18] Kassenärztliche Bundesvereinigung. Schlüsseltabellen 2021. <https://applications.kbv.de> (accessed April 4, 2021).
- [19] Stenberg D. cURL, version 7.68.0. 2020. <https://curl.se>.
- [20] Firely. Firely Terminal, version 2.0.0. 2021. <https://fire.ly/products/firely-terminal/>.
- [21] Dolan S. jq, version 1.6. 2018. <https://stedolan.github.io/jq/>.
- [22] HL7 International, Object Management Group. HL7 Version 3 Standard: Common Terminology Services (CTS), Release 2. 2015.