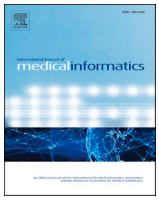
International Journal of Medical Informatics 190 (2024) 105545

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



#consented – A semantic consent code to facilitate consistent documentation and implementation of consent in collaborative medical research

Martin Bialke [a,*], Christopher Hampf [a], Arne Blumentritt [b], Frank-Michael Moser [b], Stefan Lang [c], Aileen Stehn [b], Ellen Sargsyan [d], Wolfgang Hoffmann [a], Monika Kraus [d]

[a] Institute for Community Medicine, Department Epidemiology of Health Care and Community Health, University Medicine Greifswald, Ellernholzstr. 1-2, 17475 Greifswald, Germany
[b] Trusted Third Party of the University Medicine Greifswald, Greifswald, Ellernholzstr. 1-2, 17475 Greifswald, Germany
[c] Gefyra GmbH, Otto-Hahn-Str. 9, 48161 Münster, Germany
[d] Research Unit of Molecular Epidemiology, Institute of Epidemiology, Helmholtz-Zentrum Munich, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

ARTICLE INFO

Keywords:
Medical data management
Terminology
Semantic
Informed consent
FHIR
Consent management
Patient rights
Opt-in
Opt-out
EU GDPR
Consent registry

ABSTRACT

Introduction: In German and international research networks different approaches concerning patient consent are applied. So far it is time-consuming to find out to what extent data from these networks can be used for a specific research project. To make the contents of the consents queryable, we aimed for a permission-based approach (Opt-In) that can map both the permission and the withdrawal of consent contents as well as make it queryable beyond project boundaries.

Materials and methods: The current state of research was analysed in terms of approach and reusability. Selected process models for defining consent policies were abstracted in a next step. On this basis, a standardised semantic terminology for the description of consent policies was developed and initially agreed with experts. In a final step, the resulting code was evaluated with regards to different aspects of applicability.

Results: A first and extendable version for a Semantic Consent Code (SCC) based on 3-axis (CLASS, ACTION, PURPOSE) was developed, consolidated und published. The added value achieved by the SCC was illustrated using the example of real consents from large national research associations (Medical Informatics Initiative and NUM NAPKON/NUKLEUS). The applicability of the SCC was successfully evaluated in terms of the manual semantic mapping of consents by briefly trained personnel and the automated interpretability of consent policies according to the SCC (and vice versa). In addition, a concept for the use of the SCC to simplify consent queries in heterogeneous research scenarios was presented.

Conclusions: The Semantic Consent Code has already successfully undergone initial evaluations. As the published 3-axis code SCC is an essential preliminary work to standardising initially diverse consent texts and contents and can iteratively be extended in multiple ways in terms of content and technical additions. It should be extended in cooperation with the potential user community.

1. Introduction

1.1. Challenges in consent-based medical research

According to the General Data Protection Regulation (GDPR) of the European Union (EU), the processing of personal data is only permitted under certain conditions (Art. 5 and Art. 6 GDPR). In fact, the processing of special categories of personal data, which include health data, is prohibited (Art. 9 (1) GDPR). Permission is based on a number of exceptions, one of which is the explicit consent (Art. 9 (2) lit. a GDPR). Recital 32 of the GDPR [1] focusses on an inclusion-based approach (opt-in principle) as requirements for patient consent. This 'opt-in

* Corresponding author.
E-mail addresses: martin.bialke@uni-greifswald.de (M. Bialke), christopher.hampf@uni-greifswald.de (C. Hampf), arne.blumentritt@uni-greifswald.de (A. Blumentritt), frank-michael.moser@med.uni-greifswald.de (F.-M. Moser), sl@Gefyra.de (S. Lang), aileen.stehn@outlook.de (A. Stehn), ellen.sargsyan@helmholtz-munich.de (E. Sargsyan), wolfgang.hoffmann@uni-greifswald.de (W. Hoffmann), monika.kraus@helmholtz-munich.de (M. Kraus).

https://doi.org/10.1016/j.ijmedinf.2024.105545
Received 18 March 2024; Received in revised form 4 July 2024; Accepted 8 July 2024
Available online 9 July 2024

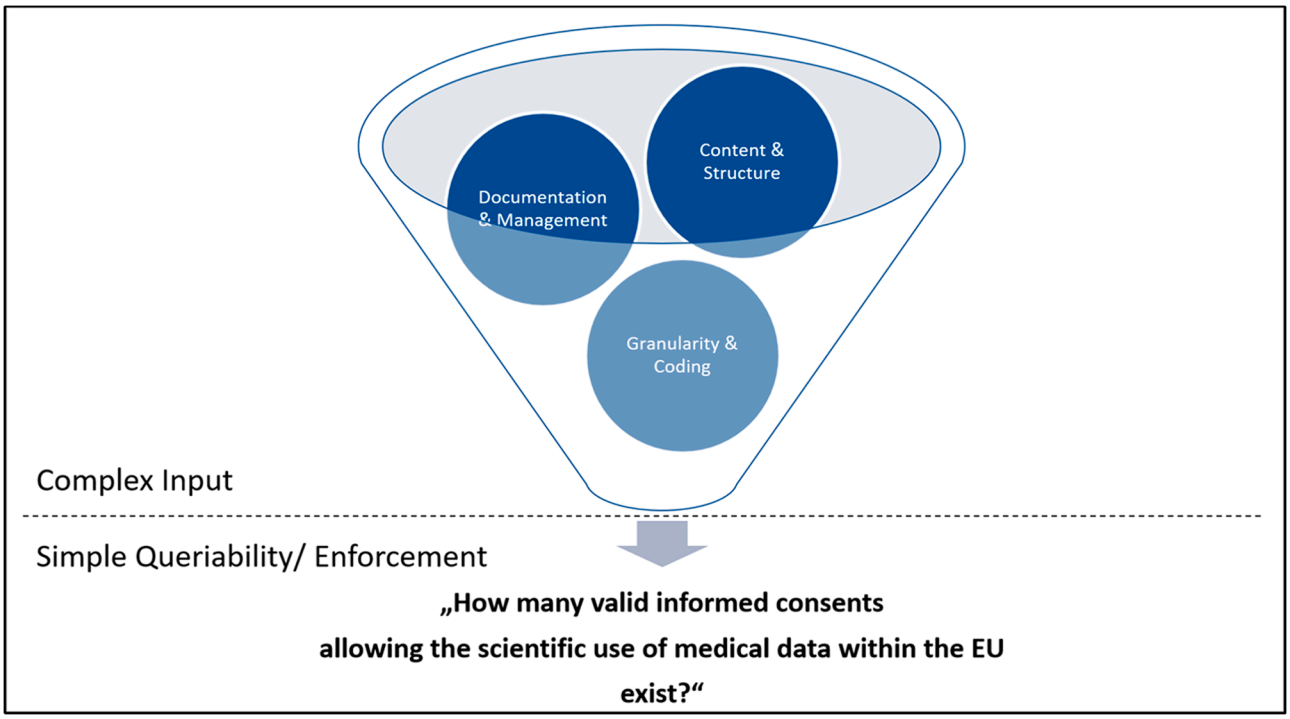


Fig. 1. The consent query challenge from a user point of view: Despite very complex input information, comprehensive consent enforcement must be as simple as possible.

principle' of consent is currently a common standard for the processing of medical data in the context of research within the EU.

In the last years, several initiatives have been launched in Germany and Europe (e.g. BBMRI-ERIC [2], BBMRI-ERIC Colorectal Cancer Cohort [3], ECRIN-MDR [4]) to support collaboration and data sharing within research communities based on patient consent in accordance with ethics principles and data protection regulations. In Germany examples are Medical Informatics Initiative (MII) [5], Network of University Medicine (NUM) [6], German National Cohort (NAKO) [7] and the German Centre for Cardiovascular Research (DZHK) [8] among others. Due to this large number of research initiatives, the *content of consent documents is usually project-specific*. Texts, content and wording are usually elaborately coordinated with responsible parties [9]. Additionally, the consent documents *differ in their structure* [10]. Moreover, these patient consents are *documented and managed in different ways* in daily practice, ranging from efficient electronic solutions [11,12] to a pile of paper. Following the 'opt-in principle' (processing of personal data in the research context is excluded until it is explicitly permitted), we are aiming for the establishment of a permission-based approach that is as flexible as possible. In order to be enabled to query this consent information (from a more technical perspective), the semantic content of obtained consent information has to follow a set of precise semantic definitions. Each statement is permitting one aspect of data processing or data sharing. This inclusion-based approach shall not be mixed with

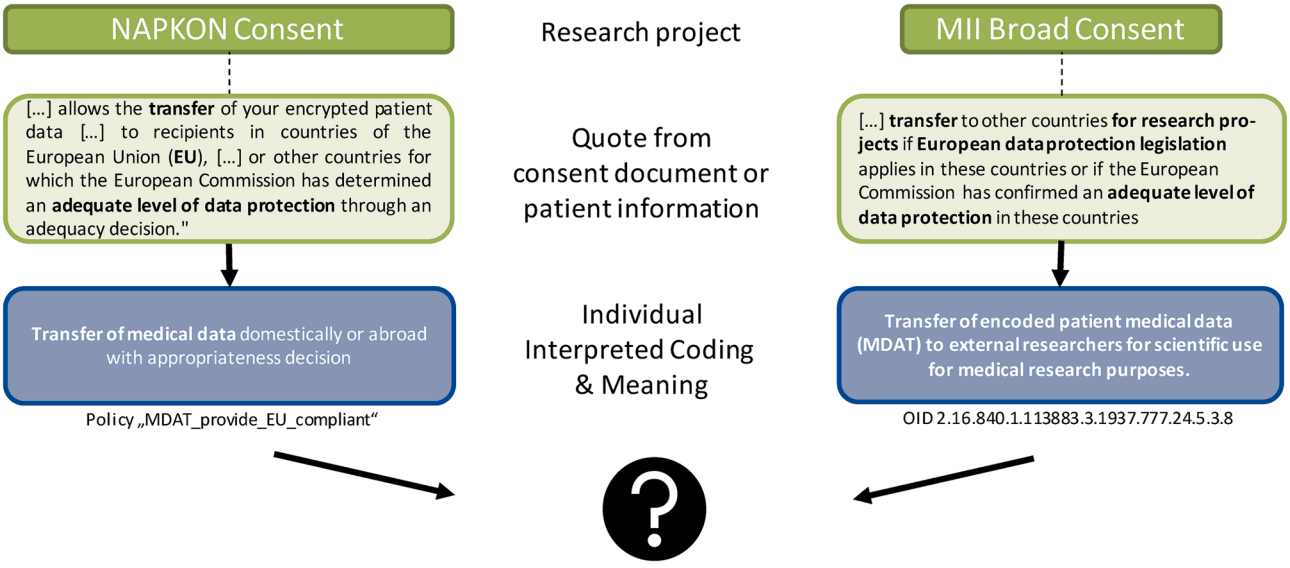


Fig. 2. Different query paths for consent content from two independent projects. A uniform and simple way to interpret, query and enforce the patients' consent is required.

exclusion-based mechanisms to prevent unintentional overlap of consent statements which might lead to invalid consent interpretation. A focus on purely permission-based consent and withdrawal processes strengthens the separation of permissible and impermissible data processing processes and simplifies implementation of the patient's will.

In summary, the challenge is to find a suitable compromise for the technical implementation of the patient's will (enforcement). From the user's point of view, this should enable simple and uniform retrieval of consent information (consent query), even if the necessary input information is based on various structured contents, different granularity of (technical) coding and a differently managed documentation (cf. Fig. 1). However, this simple and uniform queryability of consent information poses a practical problem due to the heterogeneity of methods, technical solutions and lack of standards and granularity in medical research (cf. Fig. 2).

2. Objectives

In order to describe consent content in human-readable as well as machine-readable manner a uniform and independent "semantic language" is required. Using related work as a starting point, this article focusses on the development of a generic way to semantically describe (encode) consent-related information in medical research projects. In addition, the applicability of the developed approach is evaluated and made available to the scientific community.

3. Materials and methods

3.1. Step 1: Literature review and related work

Based on a literature review, in a first step, different conceptual approaches to technically describe and encode consent-related information were assessed (cf. Table 1). Some of the approaches examined were developed for specific project scenarios and are therefore only transferable to other use cases to a limited extent. At the same time, selected approaches lack necessary granularity to correctly specify the patient's will. Inclusion and exclusion approaches are also combined, which in turn can lead to potential errors in interpretability. Further details on applicability, transferability and limitations can be found in Table 1. Additionally, it is generally noticeable that the term 'policy' has a different meaning within the different approaches. Thus, for this publication, we define the term 'policy' as a combination of semantic elements to describe a specific action, e.g. a single step of data processing.

3.2. Step 2: Semantic abstraction

Regarding Table 1, the approaches 5 (MII consent policies) and 6 (DZHK consent policies) show the necessary potential for further action. Both approaches use consent policies as a basis and are encoded in different ways. Within DZHK network query processes have already been established for consent management so that the content of consents can be checked by the Trusted Third Party after a data request has been

Table 1
Overview existing approaches to describe and encode consent-related information.

No.	Existing approach	Example(s)	Scope and limitations
1	Document-related	HL7 Consent Policy Rule Codes [13]	Each code focusses entire consent documents or rule sets, e.g. "Illinois Consent by Minors to Medical Procedures" [13]. Lack of granularity for the intended purpose.
2	Categorical	Standard Use Condition Consent Codes [14]	General application categories that allow the definition of inclusion and exclusion scenarios, e.g. "user-specific restrictions". Scope of granularity and combination of permission and denial is not suitable for the intended purpose.
3	Action-related	FHIR Consent Actions [15]	Defined set of permissible common data processing procedure actions, e.g. collect, access, etc. Limited granularity of permissible actions and applicable rules are unsuitable for the intended purpose.
4	Detailed policy-sets combining inclusion (permit) and exclusion (deny) criteria	IHE APPC Specification [16]	Focus on access control rules combining inclusion and exclusion scenarios and specific use cases/workflows, e.g. "Withhold consent to disclose to a specific provider organization". Scope of granularity and combination of permission and denial is unsuitable for the intended purpose.
5	Mapping of specific consents (per research project) based on explicit permission using individual consent policies (variant 1)	Consent Policies, identified by unique object identifiers (OIDs) [17]	Mapping of consent content to OIDs unambiguously refers to specific application context and document version (here: MII Broad Consent [10]). These MII OIDs focus on explicit permission. Their meaning is linguistically and semantically aligned. Re-use of this specific static mapping outside the MII scope is deliberately not intended.
6	Mapping of specific consents (per research project) based on explicit permission using individual consent policies (variant 2)	Consent policies identified by structured unique naming (TTP Policy Codes) [18] as applied in *research projects* DZHK/NAPKON/NUM NUKLEUS [8,19] and Trusted Third Party (TTP) of the University Medicine Greifswald	Structured identifiers and their meaning are project-specific and focus on explicit permission. Semantic assignment to consent texts requires considerable prior knowledge and expertise. Within the research project (e.g. DZHK [8]), however, they are suitable for mapping different study consents. Re-use outside the specific research project not envisaged. Nevertheless, naming of consent policies might provide orientation for similar use cases.
7	Meta data description for records	Focus on data use conditions allowing usage and access for research, as implemented in the project "Leipzig Health Atlas" [20] aligned with works of the MII [21]	Allocation of fine-grained rights for e.g. academic research, within the EU, with biomaterial. Simultaneous exclusion of certain matters, e.g. to recontact the patient concerned. Combination of inclusion and exclusion scenarios not suitable for intended purpose.

sent by the data transfer office (unit for receiving requests for data transfer and for actual transfer of data) and before data sets are transferred to data scientists and researchers. DZHK consent policies were adapted for the patient consents of NUM cohort platform NAPKON [22] and similarly applied. Within the Medical Informatics Initiative, an analogous policy-system already existed, with different terms and slightly different interpretation of the systematic based on unique digital object identifiers. These approaches are not arbitrarily interoperable.

As a next step, the DZHK policies and the MII polices had to be semantically aligned to each other and were checked for overlapping and essential components or 'comparable building-blocks'. This comparison was done by checking all the policies of the projects while the names and the contents were allocated and semantically compared. Although, the approaches used were based on policies already consisting of different building blocks with an individual meaning but allocated to similar processing and transfer operations, it was possible to identify noteworthy overlaps of these building blocks.

3.3. Step 3: Drafting a 3-axis-code

The next step was to categorise the individual building blocks of the policies. Each of these categories group thematically related policy building blocks and can be used as the axis of a 3-axis code to describe the semantic meaning of consent policies in terms of content (cf. Table 2).

This basic idea of an axis-based code was discussed with an expert representative of the "Technical Committee for terminologies of HL7 Germany" regarding content and technical suitability for the intended application. After this initial agreement on the proceeding, a comprehensive Semantic Consent Code (SCC) was developed to close the gap of a generic way to semantically describe consent-related information in medical research. Technically, the SCC consists of several code systems and value sets using the HL7 FHIR standard (cf. supplemental files).

3.4. Step 4: Evaluation of the applicability

3.4.1. Manual applicability

Selected examples from practice were mapped using the Semantic Consent Code (SCC) to gather initial impressions and first technical experiences in applying SCC to individual consent policies, consent modules, and entire consent templates. Based on that, the extent to which a manual use of SCC approaches is suitable for practical application was investigated. The evaluation focussed on user-friendliness and comprehensibility of the designed terminology. Two students, both not familiar with either the study or the SCC, received the following documents independent from each other: the patient information and consent forms of NAPKON with biomaterial collection and an excel table with a description of SCC (analogue to the later code system) as well as an overview of these three SCC categories.

The informed consent forms were translated manually to SCC based on a first version of the SCC codes provided to the participating students (cf. [23–26]). The complexity of the translated documents was very high regarding a large number of contents to be depicted individually within a consent section. This was in line with typical challenges and a realistic approach. The objectives were to generally review the usability and self-explanatory nature of the code and to identify the main difficulties in using the current version of the Code for users not involved in its creation. It was also possible to make an initial assessment of the comparability and readability of the translation of the same patient documents into code by independent people. Both students translated the written text into the codes and commented and specified questions and difficulties related to each step of the translation.

3.4.2. Technical applicability

The technical evaluation of SCC focused on testing the automated interpretability of selected consent policy terminologies. This includes both structured policy identifiers (NUM NAPKON/NUKLEUS, TTP Policy Codes) and the OID (cf. Table 1) approaches used within the MII (v1.6d and v.1.6f).

SCC can be considered as a descriptive language to describe the meaning of the respective consent policies. As many research projects naturally develop their own codes for similar content, a solely string-based interpretation of a structured consent policy might only be of limited benefit. This varying description of similar content can be perceived as a "dialect" in terms of language and implementation.

The interpretability of the consent policies was evaluated based on a comprehensive set of assertions implemented with the JUnit-Testing-Framework [27]. Based on string-comparisons most of the consent policies could already be translated (at least partially) automatically. Remaining unknown parts of structured policy names, so-called semantic syllables with unmapped SCC meaning, were identified. In a second step, project-specific SCC dialects were introduced, to optimise these translation results. The multilingual definition of dialects was subsequently integrated into these test sequences.

Table 2
To semantically describe and encode the meaning of consent policies, a semantic code based on three axes (class, action, purpose) and one optional axis (actor) is applied.

Axis	Method of allocation	Description	Type of word	Example
class	data type in a separate but integrated data management system, e.g. patient identifying information or specimen Subclasses can be used to address relevant subsets of this data type, e.g. analysis data based on specimen	class combines data or material types and subtypes "Which data type is addressed with the policy?"	acronym	person identifying information (PII or IDAT) biomaterial (BIOMAT)
action	description of permissible activity resulting from a Consent Policy	action combines permitted processing steps and activities "What am I allowed to do?"	verb	collect
purpose	specification of the application context of usage and/or scope of coverage of a Consent Policy	purpose combines further information relevant for – or – regarding the context "Why/What for/Where from/Where to/For whom/How/…?"	adjective and/or noun	timely_restricted
actor	extendible categorisation of requestor of a consent request	specification of the Actor "Who is asking?"	acronym	DTU (data transfer unit)

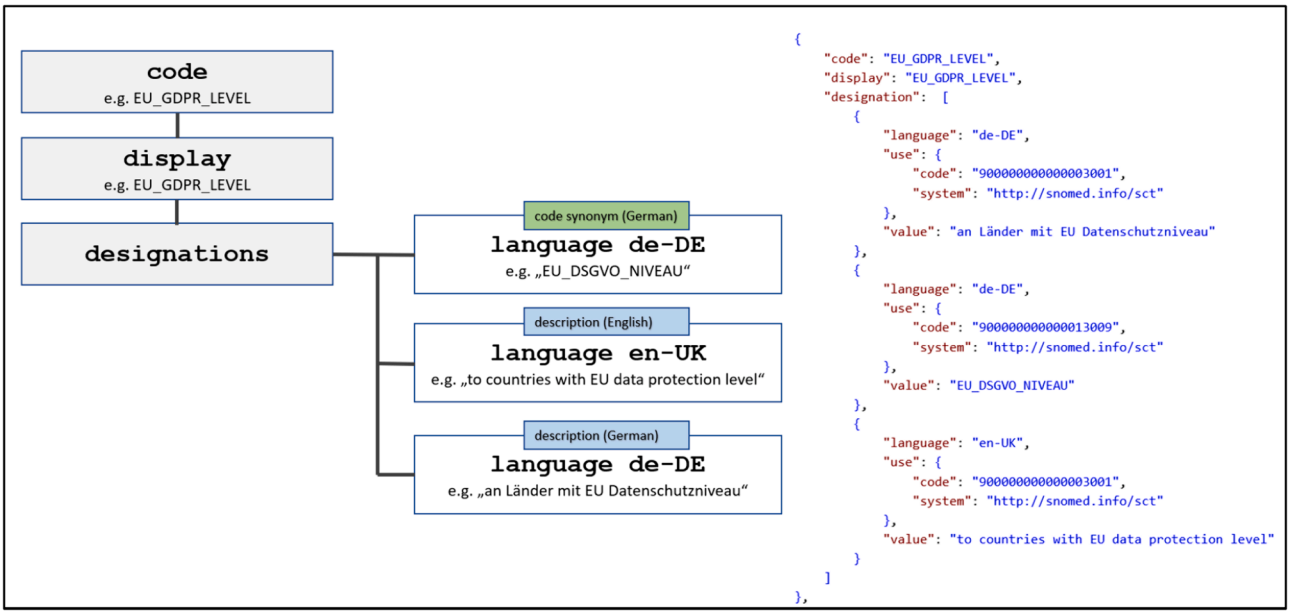


Fig. 3. Technical structure and FHIR-based definition (excerpt in JSON-Format) of a SCC-component (e.g. “EU_GDPR_LEVEL”: the English code is semantically identical with the German representation).

As a last evaluation step, and to highlight the benefits of the SCC, a concept for using the SCC in the context of a potential research infrastructure scenario was developed.

4. Results

4.1. Semantic consent code (SCC)

SCC consists of three mandatory and one optional axis. Each of these axes is represented by a separate FHIR code system. Since code systems usually evolve over time, a value set was additionally created for each code, which at the present time simply contains all individual codes of the respective axis. This creates the necessary technical bindingness and unambiguous referencing of SCC.

Each code focusses primarily on a machine-processability. Thus, a description that precisely highlights the semantic meaning is assigned to each code. The formulations of these descriptions are based on existing and agreed formulations from MII and DZHK (also used for NAPKON) definitions of policies (cf. Fig. 3).

At the same time, these descriptions were abstracted in such a way that their meaning remains the same, but the formulation can now be used in a more general context. Placeholders with references to the other code axes (e.g. [CLASS]) are also included to enable a later aggregation of semantic building blocks, so that a human-readable variant of the resulting semantics can easily be calculated. These descriptions are available in both English and German as shown in Fig. 3 via so-called designations. In addition, a German equivalent (code synonym DE) of the primarily English code has been included in the FHIR specifications to simplify subsequent mapping processes for the user.

The resulting code systems are expandable and represent a first draft (version 1.0) in terms of content and applicability. The list of codes for the axes ‘CLASS’ [23], ‘ACTION’ [24], ‘PURPOSE’ [25] and the optional axis ‘ACTOR’ [26] are available online and attached as supplementary files to this publication. Table 3 lists the codes of the ’ACTION’ axis as a representative for the components of the Semantic Consent Code.

4.2. Semantic modelling of informed consents

SCC primarily focuses on application scenarios with machine processing, e.g. for automated validity queries of consent and consent policies. Modelling of entire consent templates (e.g. MII Broad Consent with 7 modules and 27 policies) is possible [28]. In order to demonstrate

Table 3
List of Codes for Axis ‘Action’ (v1.0). FHIR representations of the full SCC axes ‘CLASS’ [23], ‘ACTION’ [24], ‘PURPOSE’ [25] and the optional axis ‘ACTOR’ [26] are available online.

Code	Description
provide	transfer of [CLASS]
transfer_ownership	transfer the ownership of [CLASS] from the participant to the controller/data processor/carrier of the biobank
view	insight into [CLASS]
collect	collection of new [CLASS]
analyse_genetic	obtaining genetic data from [CLASS]
inform	information
store_process	storage and processing of [CLASS]
contact	contact participant (if necessary, regarding [ACTION|PURPOSE])
use	provision of [CLASS] for use/analyses/for evaluation purposes [PURPOSE]
process	processing of [CLASS]
merge	merging and, if necessary, transferring of [CLASS]
query	query with/from [PURPOSE]
supplement	supplementary/complementary collection of additional [CLASS] with / from [PURPOSE]
link	linking data of the data subject to [PURPOSE]

the comparability of two different consent policies based on SCC, two actual consents of two research projects were mapped as examples using SCC: (1) one of the NAPKON cohort consents (consent not available online) [29] and (2) the Broad Consent of the MII (version 1.6d) [10]. These are separate medical research projects with different goals and non-congruent patient information (both in structure and linguistic formulation). In NAPKON, the semantic description of consent policies is achieved based on a unique name (with corresponding textual explanation). In the MII, instead, unambiguously assigned object identifiers (OID) are applied to describe a consent policy semantically [17], also with a free-text explanation.

To be able to check the validity of consent, e.g. in order to share and transfer data, **before SCC** it was necessary:

- to know the relevant passage in the respective consent text
- to know precisely the form of coding of the consent policy assigned to this text passage, both in NAPKON and MII
- to query these policies separately using an appropriate technical implementation
- to merge the result of the query appropriately.

With SCC a uniform possibility to query both variants simultaneously was achieved. Fig. 4 shows an example of the translation of consent content in text form describing the possibility to transfer data to third countries under certain circumstances to SCC. For translation the values 'MDAT' (which is described with 'medical data'), 'provide' (which is described with 'transfer of') and 'EU_GDPR_LEVEL' (which is described with 'to countries with EU data protection level') were combined. In this way, the meaning "transfer of medical data to countries with EU data protection level" can be modelled in a uniform manner. As a result, the consent coding approaches NAPKON and MII have become comparable and compatible in a way the exact knowledge of the original written text of informed consent forms must not be known for querying the fact if individually consented data may be sent to countries with EU data protection level.

4.3. Evaluation of the manual applicability

First evaluative steps focussed on the user-friendliness and comprehensibility of the designed terminology. The translated consent form in SCC-format (see 3.4 step 4) resulted in eight modules to which the study participant could give consent. The module corresponding to the general privacy paragraph was by far the largest, comprising 19 code combinations. The documents were also translated by the Ethics Coordination Office of the cohort studies into SCC terms, which correspond to the code already used in the productive implementation. Differences in translation between the students and the Ethics Coordination Unit were calculated on the basis of the number of different code combinations. In the most comprehensive and difficult to interpret module, the General Privacy Paragraph, there was still a 58 % agreement between the students and the Ethics Coordinator's translation, while in the other seven modules an average agreement of 83 % was achieved. In addition, lessons learned during the evaluation process were extracted into a dual-language user manual [30], which was attached to this publication as a supplemental file.

4.4. Evaluation of the technical applicability

Fig. 5 shows the results of the technical evaluation (see 3.4 step 4). The MII policies (v1.6d: total of 27; v1.6f: total of 33) and the "TTP Policy Codes" [18] (total of 75) could be 100 % automatically translated into SCC using the dialect approach as described above. 71 of 105 policies could be translated all-over the 133 highly specific consent templates of NAPKON/NUKLEUS. The approach enables a quite intuitive

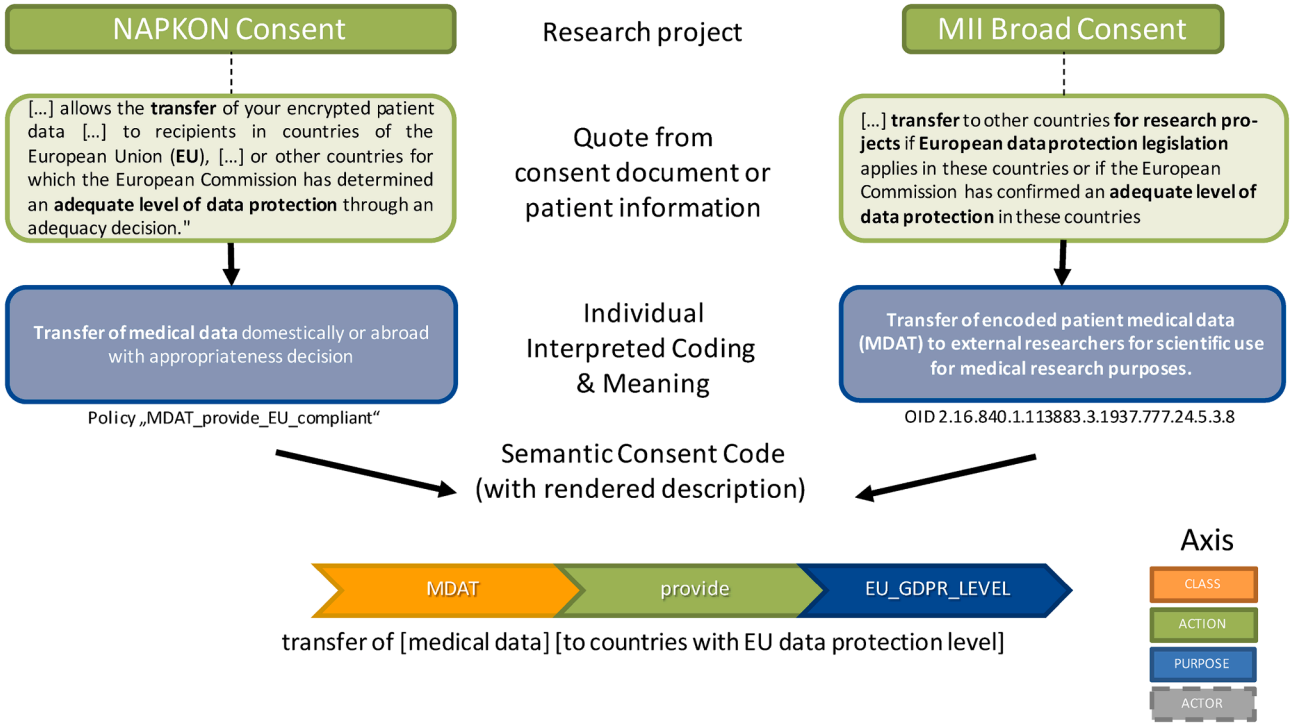


Fig. 4. Semantic modelling example of SCC for application in different research projects (NAPKON, MII). For better comprehensibility, the NAPKON policy "MDAT_provide_EU_compliant"– originally mapped in German – has been translated into English.

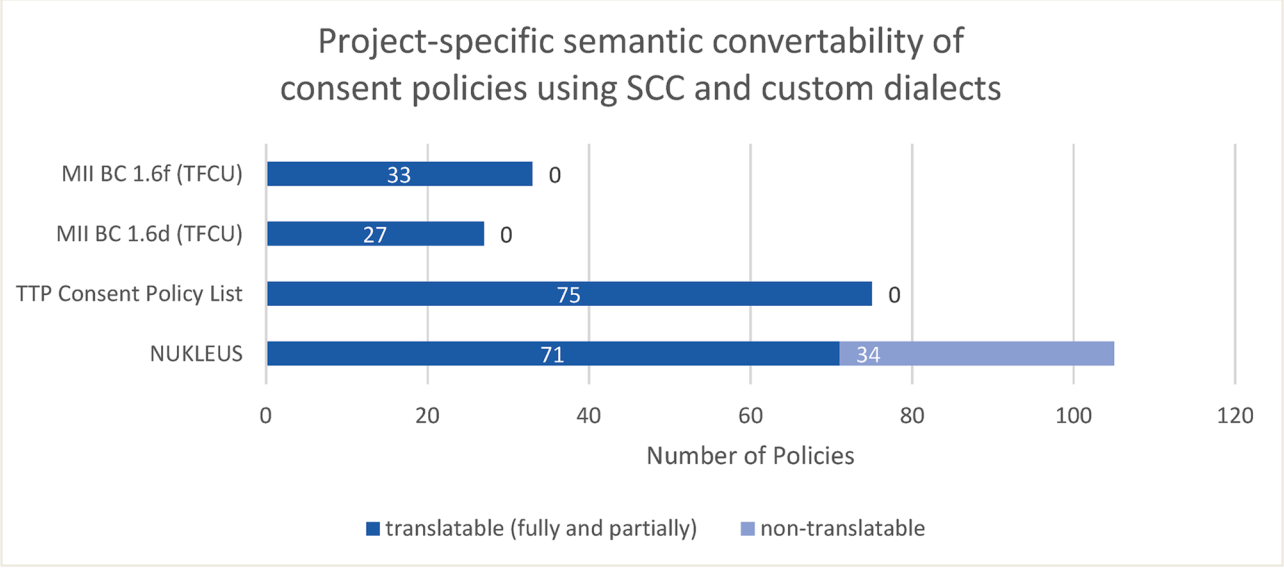


Fig. 5. Results of a first evaluation of SCC focusing on the technical applicability.

mapping between different language encodings. Further details on the definition of SCC dialects and the carried-out results of the conversion tests, which form the basis for Fig. 5, are attached to this publication as supplemental files.

4.5. Suggestion of a technical concept for the application of SCC

To illustrate the benefits of the SCC approach, a potential use case with heterogeneous consent information is utilised, in which researchers aim to determine the number of valid patient consents for data exchange within the EU. To estimate the number of available medical datasets, they use a research support system, which here is representative of any software system with corresponding interfaces. The researchers have no knowledge of the type of consent information that needs to be evaluated to answer the search query and are primarily interested in the most precise answer possible. Their query is interpreted by a conversion adapter as semantic consent code (1) and converted into a search query (FHIR Query Language) (2). The search query is transferred (3) to a corresponding system (federated Consent Registry), which is connected to different data sources with consent information, each of them uses different ways of describing consent information (4a, 4b, 4c). The Consent Registry can extract the necessary information (4) from the connected data sources using appropriate converters and interpreters (SCC to OID, SCC to Policy, SCC to Custom Study Consent). The partial results are aggregated and transmitted to the researcher (5) in FHIR-Format.

The approach (cf. Fig. 6) shall illustrate the additional value of SCC and might be supplemented and adapted in various directions. For example, the Consent Registry is not limited to the forms of supported consent coding shown and can be extended (almost) as necessary. The shown conversion adapter can also be part of the Research Support System or a separate software solution.

5. Discussion

5.1. Complexity and limitations

In the first version of SCC, we focussed on a data-centred point of view (e.g. medical data, specimen) to reduce complexity. This required a non-overlapping and semantic definition of the individual terms and

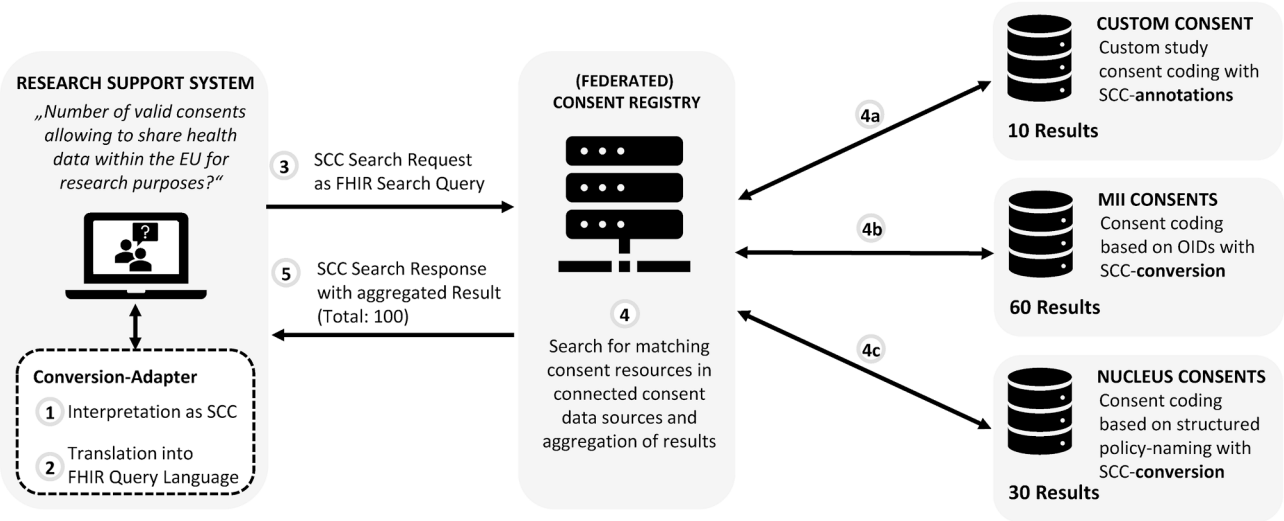


Fig. 6. Conceptual application of the Semantic Consent Code in a multi-project scenario to simplify aggregation of heterogeneous consent information.

elements of the codes. Further, axes can intuitively be combined and a human-readable description based on prepared sentence components for simplified understanding can be calculated. Although numerous typical content queries to be used in research networks can already be addressed, at the moment of publication not all query-relevant eventualities could be considered in the current version of SCC due to a very high complexity of disease definitions and already existing standardization codes (e.g. ICD [31], DSM [32]).

Nevertheless, the SCC is flexible and can be applied in detail to supplement existing terminologies that lack precision, particularly in the HL7 standard. With regard to the specification of actions, SCC has overlaps with the capabilities of HL7 [15]. The SCC extends these capabilities while maintaining compatibility. Depending on the use case, SCC can be used for modelling monolithic consent documents [13], but also for the fine-grained annotation of FHIR consent resources with the help of provision elements.

Existing consent codes for the formulation of "standard use conditions" [14] do not provide for use with infrastructural queries (cf. Fig. 6). The Consent Codes specified [14] only represent application purposes in a very coarsened form, which makes specific transferability to other research projects more difficult. SCC can be utilised (using the dialect approach described under 4.5) to achieve compatibility with other existing approaches and to enable the prerequisite for standardised technical queryability.

In addition to IHE APPC [16], the latest IHE PCF specification [33] also focuses primarily on the specification of roles (e.g. Consent Recorder, Consent Registry) and consent-related communication processes. SCC would offer the opportunity to specify the communicated content and thus, in addition to roles and processes, to add clear specifications for the consent-related content interchanged between the relevant actors. Technical compatibility already exists anyway due to the common technical basis (FHIR Consent).

The focus of SCC is medical research and is applied to mark and supplement relevant data (genome data, various analysis data, health insurance data). SCC offers previously unavailable flexibility and query granularity and enables combined queries for the utilization of datasets, as exemplified by NUM, MI-I, and DZHK. Thus, SCC fills a gap in comparison to existing approaches.

5.2. Usability

Individual aspects of data processing can be described specifically and comprehensibly using the composition of codes from the three axes. When translating the consent content into the codes, the wording of the description must be carefully chosen, because the interpretation of definitions in detail can vary between different interpreters. In order to simplify the process of translating/interpreting the consent content into SCC and for pragmatic application, we provide a publicly available user manual [30] to support a most intuitive translation process.

6. Conclusions

A first approach to map consent information from different application scenarios in terms of a Semantic Consent Code (SCC) was developed and evaluated. The extensible 3-axis code is an essential preliminary work to standardising originally different consent texts and content, with a focus on international applicability.

With its extension options, the SCC offers a general possibility to support the linkage to data supported by opt-out solutions for specific contexts, such as data related to the Health Data Utilisation Act [34] and electronic health records in Germany, and will be able to support not yet specifically defined tasks of data providers in the European Health Data Space [35]. Secondary languages can be easily integrated, as already exemplified with a German translation for a simplified applicability in German research projects. The resulting code systems are freely available to interested scientists and have already successfully undergone first evaluations.

Based on initial results for the technical applicability of SCC, we believe that the SCC approach offers potential for the creation of a "consent query language" for patient consent in the context of medical research. The next steps should therefore focus on the practical implementation of the application concept outlined in Fig. 6. This includes the identification of a suitable pilot project, the technical implementation of the outlined converters/interpreters and the realisation of a (federated) consent registry based on the FHIR standard for consent in medical research [36] on a national and international level.

7. Summary Table

What challenges have been faced in the field of consent-based medical research to date?

- Due to a large number of national and international research initiatives, the content and structure of consent documents are usually project specific. The documentation and semantic annotation of these consent information follows varying existing technical and semantical approaches.
- A simple and uniform queryability of consent information poses a practical problem. The challenge is to find a suitable compromise for the technical implementation of the patient's will to enable simple and uniform retrieval of consent information (consent query), even if the necessary input information is highly complex and heterogeneous.

What potential added value does the presented semantic consent code offer users and researchers?

- We propose a Semantic Consent Code (SCC) for the semantic description of consent content based on HL7 FHIR, which is capable of supplementing existing terminologies.
- The SCC facilitates a standardised semantic description of the meaning of consent content and thus makes it uniformly machine readable. This is possible regardless of the individual research project and the structure of the consent document.
- The SCC has already been evaluated in terms of human readability and handling (guideline), machine interpretability and potential integration into research infrastructures.

List of Abbreviations.

Abbreviation, Description

APPC, Advanced Patient Privacy Consent
DZHK, German Centre for Cardiovascular Research
FHIR®, Fast Healthcare Interoperable Resources
GDPR, EU General Data Protection Regulation
gICS®, generic Informed Consent Service
HL7®, Health Level Seven International
IDAT, Person-identifying information (unofficial)
IHE, Integrating the Healthcare Enterprise
MDAT, Medical/health data
MII, Medical Informatics Initiative
MII BC, MII Broad Consent
NAKO, Research Project "German National Cohort"
NAPKON, German National Pandemic Cohort Network
NUKLEUS, Clinical Epidemiology and Study Platform of the Network of University Medicine
NUM, Network of university medicine (NUM) to support COVID and pandemic research at national level in Germany
OID, Object Identifier
PII, Person-identifying information (official)
RDP, Routine Data Platform
SCC, Semantic Consent Code
SÜP, Cross-Sectoral platform (Sektorenübergreifende Plattform)

Ethics approval and consent to participate: Not applicable.

Consent to publish: Not applicable.

Availability of data and material: The referenced FHIR Code systems are publicly available from https://simplifier.net/ths-greifswald and part of the supplemental files of this publication in JSON-format. The SCC manual [30] and details of the technical evaluation of the SCC are attached as supplemental files as well.

8. Authors' contributions

Drafting of the manuscript: MB, MK. Creation of figures and tables: MB, MK. FHIR Profiling and implementational consulting: SL, MB, CH, FMM, AB. Semantic Tests: MB, FMM, CH, AB. Guideline and Evaluation: MK, AS, ES, MB. Technical prototyping: FMM, MB, CH. Revision of the manuscript: MB, CH, AB, FMM, SL, AS, ES, WH, MK. All the authors approved the final version of the manuscript.

Funding

The NUM-RDP project and the NUM-NUKLEUS infrastructure (both: grant number 01KX2121) are funded by the Federal Ministry of Education and Research of Germany (BMBF). The funders had no role in the design of the study, the collection, analysis or interpretation of data, or the writing of the manuscript.

Credit authorship contribution statement

Martin Bialke: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Conceptualization. **Christopher Hampf:** Writing – review & editing, Writing – original draft, Conceptualization. **Arne Blumentritt:** Writing – review & editing. **Frank-Michael Moser:** Writing – review & editing, Validation, Software, Methodology, Formal analysis, Conceptualization. **Stefan Lang:** Writing – review & editing, Methodology, Conceptualization. **Aileen Stehn:** Writing – review & editing, Validation. **Ellen Sargsyan:** Writing – review & editing, Validation. **Wolfgang Hoffmann:** Writing – review & editing. **Monika Kraus:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Since this work has been developed over many years on the basis of the registries, cohorts and studies of the DZHK and the NAPKON STUDY GROUP, which were carried out during day-to-day work, we would like to express our special thanks to them. The mentioned paragraphs of the patient documents originate from one of the NAPKON-cohorts.

Thorsten Rottmann, Dirk Langner, Lars Geidel and Thomas Bahls also carried out preliminary work many years ago and colleagues from the DZHK's Clinical Research Platform, especially to be mentioned Tabea Scharfe through her role in the DZHK transfer office, have shared their ideas for their respective areas.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2024.105545.

References

[1] intersoft consulting. Recital 32 General Data Protection Regulation, Conditions for consent. https://gdpr-info.eu/recitals/no-32/ Accessed 01 Mar 2018.
[2] BBMRI-ERIC. BBMRI-ERIC Homepage. https://www.bbmri-eric.eu/ (2024). Accessed 06 May 2024.
[3] BBMRI-ERIC. Project-Page Colorectal Cancer Cohort - ADOPT BBMRI-ERIC. https://www.bbmri-eric.eu/scientific-collaboration/colorectal-cancer-cohort/ (2024). Accessed 15 May 2024.
[4] ecrin metadata repository. ecrin Clinical Research Metadata Repository. https://newmdr.ecrin.org/ (2024). Accessed 11 May 2024.
[5] S. Semler, F. Wissing, R. Heyder, German Medical Informatics Initiative, Methods Inf. Med. 57 (S 01) (2018 Jul) e50–e56, https://doi.org/10.3414/ME18-03-000.
[6] H.U. Prokosch, T. Bahls, M. Bialke, J. Eils, C. Fegeler, J. Gruendner, et al., The COVID-19 Data Exchange Platform of the German University Medicine, Stud. Health Technol. Inform. 294 (2022) 674–678, https://doi.org/10.3233/SHTI220554.
[7] German National Cohort (GNC) Consortium, The German National Cohort: aims, study design and organization, Eur. J. Epidemiol. 29 (5) (2014 May) 371–382, https://doi.org/10.1007/s10654-014-9890-7.
[8] J. Hoffmann, S. Hanß, M. Kraus, J. Schaller, C. Schäfer, D. Stahl, et al., The DZHK research platform: maximisation of scientific value by enabling access to health data and biological samples collected in cardiovascular clinical studies, Clin. Res. Cardiol. (2023), https://doi.org/10.1007/s00392-023-02177-5.
[9] M. Kraus, M. Nauck, D. Stahl, A. Blumentritt, G. Anton, H. Wichmann, et al., Studienübergreifende Harmonisierung datenschutzrechtlicher und ethischer Gesichtspunkte in Patientenunterlagen: Eine Praxisanalyse (German only), in: G. Richter, W. Loh, A. Buyx, S. Graf von Kielmansegg (Eds.), Datenreiche Medizin und das Problem der Einwilligung, Springer, Berlin, 2021.
[10] S. Zenker, D. Strech, K. Ihrig, R. Jahns, G. Müller, C. Schickhardt, et al. Data protection-compliant broad consent for secondary use of health care data and human biosamples for (bio)medical research: Towards a new German national standard. J. Biomed. Informat. 2022 131 (July 2022, 104096) 1532–0464. https://doi.org/10.1016/j.jbi.2022.104096.
[11] H. Rau, L. Geidel, M. Bialke, A. Blumentritt, M. Langanke, W. Liedtke, et al., The generic Informed Consent Service gICS®: implementation and benefits of a modular consent software tool to master the challenge of electronic consent management in research, J. Transl. Med. 1 (2020 Dec 18) 287, https://doi.org/10.1186/s12967-020-02457-y.
[12] ths-greifswald.de. New gICS integration module for Thieme eConsent Pro (german). https://www.ths-greifswald.de/live-demo-neues-integrationsmodul-fuer-gics-in-thiemes-e-consentpro-vorgestellt/ (2022). Accessed 1 Jun 2023.
[13] HL7 International. HL7 Terminology 2.0, CodeSystem: Consent PolicyRule Codes. https://terminology.hl7.org/2.0.0/CodeSystem-consentpolicycodes.html Accessed 10 Jul 2023.
[14] S. Dyke, A. Philippakis, J. Rambla De Argila, D. Paltoo, E. Luetkemeier, B. Knoppers, et al., Consent Codes: Upholding Standard Data Use Conditions, PLoS Genet. 12 (1) (2016), https://doi.org/10.1371/journal.pgen.1005772.
[15] HL7 Terminology (THO). ValueSet: Consent Action Codes. https://build.fhir.org/ig/HL7/UTG/ValueSet-consent-action.html (2023). Accessed 12 Jul 2023.
[16] IHE ITI Technical Committee. Advanced Patient Privacy Consent (APPC) - Draft for public comment. https://ihe.net/uploadedFiles/Documents/ITI/IHE_ITI_Suppl_APPC_Rev1.0_PC_2016-05-27.pdf (2016). Accessed 03 Feb 2023.
[17] MII Task Force "Consent Implementation". ART-DECOR MII Consent Policy ValueSet. https://art-decor.org/decor/services/RetrieveValueSet?id=2.16.840.1.113883.3.1937.777.24.11.36&language=en-US (2021). Accessed 05 May 2021.
[18] University Medicine Greifswald. FHIR Representation Consent Policy ValueSet 'Policy'. https://www.ths-greifswald.de/gics/policies (2023). Accessed 01 Dec 2023.
[19] D. Krefting, G. Anton, I. Chaplinskaya-Sobol, S. Hanss, W. Hoffmann, S. Hopff, et al., The Importance of Being FAIR and FAST - The Clinical Epidemiology and Study Platform of the German Network University Medicine (NUKLEUS), Stud. Health Technol. Inform. 302 (2023) 93–97, https://doi.org/10.3233/SHTI230071.
[20] FAIR4Health Platform portal. IACS data set. https://health-atlas.de/data_files/571?graph_view=tree Accessed 01 Jun 2023.
[21] Taskforce Metadata of the Medicalinformatics Initiative. Metadata schema for the description of datasets in the data integration centres of the MII core dataset (Internal document, not available for public use). 2021.
[22] M. Schons, L. Pilgram, J. Reese, M. Stecher, G. Anton, K. Appel, et al., The German National Pandemic Cohort Network (NAPKON): rationale, study design and baseline characteristics, Eur. J. Epidemiol. 37 (8) (2022) 849–870, https://doi.org/10.1007/s10654-022-00896-z.
[23] University Medicine Greifswald. FHIR representation for Axis 'CLASS' (v1.0). (2024) https://www.ths-greifswald.de/fhir/CodeSystem/gics/ConsentPolicyClass Accessed 30 Apr 2024.
[24] University Medicine Greifswald. FHIR representation for Axis 'ACTION' (v1.0). (2024) https://www.ths-greifswald.de/fhir/CodeSystem/gics/ConsentPolicyAction Accessed 30 Apr 2024.
[25] University Medicine Greifswald. FHIR representation for Axis 'PURPOSE' (v1.0). (2024) https://www.ths-greifswald.de/fhir/CodeSystem/gics/ConsentPolicyPurpose Accessed 30 Apr 2024.
[26] University Medicine Greifswald. FHIR representation for Axis 'ACTOR' (v1.0). https://www.ths-greifswald.de/fhir/CodeSystem/gics/ConsentPolicyActor (2024). Accessed 30 Apr 2024.

[27] The JUnit Team. JUnit Test Framework. https://junit.org/junit5 (2024). Accessed 01 Feb 2024.
[28] M. Bialke, T. Bahls, L. Geidel, H. Rau, A. Blumentritt, S. Pasewald, et al., MAGIC: Once Upon a time in consent management - A FHIR tale, J. Transl. Med. 16 (2018) 256, https://doi.org/10.1186/s12967-018-1631-3.
[29] NAPKON SÜP Study. Study Synopsis (German). (2022) https://napkon.de/wp-content/uploads/2023/02/SUEP_Studiensynopse.pdf Accessed 21 Jun 2023.
[30] Kraus M. Semantic Consent Code - Dual language User Manual (GER, EN). (2024) https://ths-greifswald.de/scc/manual Accessed 30 Jun 2024.
[31] World Health Organization. WHO classification of diseases. https://www.who.int/standards/classifications/classification-of-diseases Accessed 12 Jul 2023.
[32] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Diagnostic and Statistical Manual of Mental Disorders. https://dsm.psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787 Accessed 12 Jul 2023.
[33] IHE International. Privacy Consent on FHIR (PCF) - 1.0.0-comment - ballot. https://profiles.ihe.net/ITI/PCF/1.0.0-comment/index.html (2023). Accessed 01 Jul 2024.
[34] Federal Ministry of Health. The Health Data Utilisation Act (Gesundheitsdatennutzungsgesetz (GDNG)). https://www.bundesgesundheitsministerium.de/service/gesetze-und-verordnungen/detail/gesundheitsdatennutzungsgesetz.html (2023). Accessed 08 May 2024.
[35] European Parliament. EU Health Data Space: more efficient treatments and life-saving research. https://www.europarl.europa.eu/news/en/press-room/20240419IPR20573/eu-health-data-space-more-efficient-treatments-and-life-saving-research (2024). Accessed 08 May 2024.
[36] HL7 Germany. Official Implementation Guide of the HL7 Working Group Consent Management (Version 1.0.1). https://ig.fhir.de/einwilligungsmanagement/1.0.1/ (2022). Accessed 24 Nov 2022.