JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

JLIS.it

Indexing clinical documents for long-term preservation through an international coding system*

Erika Pasceri[(a)]
Maria Teresa Chiaravalloti[(b)]

a) University of Calabria, https://orcid.org/0000-0003-4440-2266
b) IIT-CNR, https://orcid.org/0000-0003-4695-2026

Contact: Erika Pasceri, erika.pasceri@unical.it; Maria Teresa Chiaravalloti, maria.chiaravalloti@iit.cnr.it
Received: 16 October 2024; **Accepted:** 03 March 2025; **First Published:** 15 May 2025

ABSTRACT
This paper describes the experience of mapping the document types defined by Medas company, a Digital Preservation Organization for healthcare institutions, to the Logical Observation Identifiers Names and Codes (LOINC) international standard, and consequently assesses its adequacy in representing specificities of the Italian context. Mapping operations were manually performed by LOINC Italia experts. The LOINC database was searched using the LOINC Search web browser and REgenstrief LOINC Mapping Assistant (RELMA) software, employing local names and synonyms of Medas document types. Out of 483 Medas document types, 144 fully match with a LOINC code; 211 were associated with a LOINC code which only partially covers the meaning; 128 are not covered by LOINC. Although the axes of LOINC Document Ontology are quite adequate for document types representation, an extension of Kind of document and Subject Matter Domain axes is desirable. Local Medas codes reflect the variety of clinical document types produced in the Italian healthcare domain and the results of this case study can be used as new terms submissions to enrich the LOINC standard, especially in relation to the specificities that characterize the Italian national context.

KEYWORDS
LOINC; Clinical documents; Long-term preservation; Mapping; LOINC Document Ontology.

* Authors all contributed in the realization of this paper, although Erika Pasceri specifically dealt with subsection "Introduction", "Material and Method" and Maria Teresa Chiaravalloti dealt with "Background and significance", "Results" and "Discussion" and together worked on "Conclusion".

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

Introduction

Long-term preservation is the activity aimed at protecting and safeguarding archives of documents and computer data over time. It is a mandatory task in Public Administration in Italy since the introduction of the Law Decree n. 82/2005 named *Codice dell'Amministrazione Digitale* (CAD). Long-term preservation systems have to guarantee authenticity, integrity, reliability, readability and availability of electronic documents over time (CAD, art. 44). Public Administration can choose to implement an inhouse or outsourcing long-term preservation service.

Actually, it introduced a gradual digitization of the information flows in public administrations by re-engineering processes and consequently promoting a new document approach from a structural point of view and about the semantic management of their content. In this context, health documents are of particular interest, both for the nature of the data they contain and for the requirements introduced by the advent of the Italian Electronic Health Record (EHR) system, namely Fascicolo Sanitario Elettronico (FSE).

FSE was established by the Law Decree n. 179/2012 and subsequently detailed by the Prime Minister Decree n. 178/2015, also in compliance with the European Directive 2011/24/EU concerning the application of patients' rights to cross-border healthcare. The FSE is a federated infrastructure, which enables interoperability services among regional EHR systems, such as, primarily, clinical documents search, retrieve and exchange.

FSE allows to overcome technological barriers of the regional and proprietary EHR systems to trace the entire medical history of a patient, so ensuring healthcare professionals the chance to access clinical documents produced by different stakeholders in order to provide efficient clinical care and pathways. Its architecture is based on a system of registries and repositories. Each Region has a registry, which indexes metadata of the clinical documents of the patients assisted in the Region itself (even if produced by facilities located in other Regions) and can choose to have one centralized repository or more distributed repositories to store documents produced by the healthcare facilities of the Region itself (even if indexed in the registry of another Region) (Frazier et al. 2001; Shapiro et al. 2005; De Pietro et al. 2011). Clinical data of the documents indexed in the FSE have to be represented by using coding and classification systems to ensure, eventually by transcoding, semantic interoperability among the various regional, national and European e-health solutions. Among minimum metadata required for indexing documents in the FSE there is the description of the type of clinical document, which has to be uniquely identified through codes belonging to the standard LOINC (Logical Observation Identifiers Names and Codes). The subset of LOINC codes currently used for this purpose is listed into the Italian Affinity Domain (AD) document.[1] Nonetheless, private organizations, which provide outsourcing services of document long-term preservation, find it insufficient compared to the need to have uniquely identified a great number of different clinical document types produced by healthcare facilities. Furthermore, the names of the clinical documents that they receive for long-term preservation are often assigned by the authors themselves or at most determined by institutional workflows, so they are most of the time not consistent when pooled.

The LOINC codes identifying clinical documents' types are part of the Document Ontology

[1] More info at https://www.fascicolosanitario.gov.it/sites/default/files/public/media/AffinityDomainItalia_v2-3.pdf.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

(DO) of the standard, which was born from the initial effort of the Document Ontology Task Force (DOTF) (Dugas and Thun 2009), composed by members of both the US Veterans Health Administration and Health Level Seven (HL7), aiming at creating a system for clinical document naming. They considered a similar previous work of the European Committee for Standardization (CEN) resulting in the publication of ENV 13606-2, which is a domain term list for the electronic healthcare record communication. However, CEN document categories were still too generic and needed to be used in combination with local codes. On the contrary, DOTF aimed to produce a fully specified name. They started from the empiric analysis of over 2,000 documents and defined a structure based on six axes to have the chance of non-ambiguously combining concept names so that they are unique for document retrieval. The intent was to create a vocabulary to represent and communicate knowledge about clinical documents and define a set of relationships among its terms (Frazier et al. 2001).

Moving from these considerations, the objective of this work is mapping the document types managed by one of the major Digital Preservation Organization of clinical documents[2] towards the LOINC standard to assess the semantic coverage of the standard and consequently identify candidates for LOINC new terms submissions. Results of this work could be also used to enlarge the list of LOINC document type codes used to index documents into the FSE according to the AD.

Background and Significance

The LOINC DO is currently the domain reference standard, but it is worth mentioning similar effort of Brown et al. (2001) in creating a Document-naming Nomenclature (DNN) for the Computerized Patient Record System (CPRS) of the Veterans Affairs (VA) medical centers. They specified three document identifier categories to be populated with values from a nomenclature. Nonetheless their system was not precise enough in defining subspecialties.

Several studies investigated the LOINC DO coverage and usability to correctly identify clinical document types. Hyun et al. (2006) focused on the DO adequacy for representing nursing documents. Ninety-four document types were collected from the Eclipsys Clinical Information System (CIS) and the Medical Entities Dictionary (MED) of the New York Presbyterian Hospital. They were mapped considering primarily the specific subset of nurse terms in the DO but extending to others if the concept was present in LOINC but not specifically designed for nursing. As results 51.1% of terms were mapped to an exactly equivalent concept, 28.7% were classified as *not specified*, and 20.2% as *other* because representative values were not present in the Subject Matter Domain (SMD) axis of the DO. The study also showed the need for an extension and clarification of the Type of Service (TOS) axis to completely represent different subtypes of nursing documents. A limit of this study was the size of the sample documents coming from only two institutions, which could affect the generalizability of the results. Similarly, in Hyun et al. (2009), document collections from the Columbia University Medical Center campus (CUMC)

[2] The work presented in this paper was realized in collaboration with Medas Srl, an Italian company that offers innovative services and IT solutions in the healthcare domain. Among these, it carries out the long-term preservation of digital health documents.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

of New York Presbyterian Hospital (NYPH) were used to test the DO. The study evidenced the need for revisions of both the DO and the local clinical documents classification, as documents classified as *Not Specified* were too granular in their names and those classified as *Other* were not covered by the value list of the SMD axis. Chen et al. (2010) conducted a similar study on document names of two institutions (Fairview Health Services and Fletcher Allen Health Care). They used automated methods to identify exact matches, which resulted in 55% of the total amount of local document names (1,240). Authors opened a fundamental question: when mapping to an international standard, to what extent is it tolerable to lose granularity with respect to local names descriptions? And consequently, when is it appropriate to request the addition of new codes to the standard? Their position is that it depends on the use case (documents exchange, retrieval, etc.) and therefore they propose to maintain both local and standard names to «minimize information loss, preserve semantics, and enable flexible use. Moreover, they advanced the idea of a DO with a post-coordinated approach rather than the current pre-coordinated one. As DO is constantly updated, the results of studies from some years ago could have been different if conducted today, in any case they highlight some fundamental issues that are still detectable today, but which basically constitute the driving force for the improvement of DO itself. Such as in Li et al. (2011) where non-mapped notes are due to i) specificity of local items; ii) abbreviations and acronyms, which do not give clear evidence of what they represent; iii) missing DO values. While Wang et al. (2014) focused on values of Setting and Role axes, recommending specific additions to them, Rajamani et al. (2014a) addressed the Setting axis, with the aim of increasing DO representativeness of care settings other than hospital and office visits. The same authors analyzed also the Role axis finding lack of comprehensiveness, as DO values were either too broad or too specific (Rajamani et al. 2014b).

The limitation most complained by the analyzed previous studies is that the sample of document types used to evaluate the semantic coverage of the DO comes from one or at most two institutions, and therefore lacks a sufficiently broad representativeness. This study manages to overcome this limit as the documentary corpus made available by Medas comes from different types of healthcare facilities located throughout Italy.

More recently, a study used a deep learning model to automate the mapping of values to the DO axes (Zuo et al. 2021). The results showed improved performance compared to the manual mapping of the same corpus, which was used to annotate the framework. In the future it would be interesting to test the document corpus of this work through automatic mapping, in order to study methodologies for its refinement so that it could become a common practice to use the LOINC standard for the unambiguous identification of clinical document types.

Materials

The LOINC DO was created in the first 2000s from the joint effort of the LOINC committee and the HL7 document ontology task force, recognizing the need to uniformly name documents with the same content and distinguish those that could appear identical in the name but actually differ in the substance (Frazier et al. 2001). This is especially important to: i) exchange documents between different clinical information systems; ii) classify and retrieve documents coming from

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

different sources (as in the Italian EHR registries); iii) standardize naming of clinical documents templates. As for the LOINC codes for laboratory tests, these DO codes are not intended to substitute local names and codes but to come alongside them to ensure effective semantic interoperability.

The five axes of the DO codes are:

- *Kind of document*, which describes the basic structure of a document (e.g. note, certificate, etc.);
- *Type of service*, which identifies the type of service or activity delivered to the patient (e.g. consultation, admission evaluation, etc.);
- *Setting*, which is either the place or the situation where the document is produced (e.g. hospital, outpatient, etc.);
- *Subject Matter Domain (SMD)*, which represents the medical specialty or domain related to the document (e.g. cardiology, family medicine, etc.);
- *Role*, which is the medical level responsible for the document content (e.g. consultant, nurse, etc.).

Through the combination of values of those axes it is possible to model LOINC terms for clinical document types, whether they are more specifically medical or administrative. The LOINC DO class currently includes 3,620 records. The DO Committee constantly works to update and refine values of the axes in order to be representative of the expressive needs of the clinical practice.

The LOINC DO is also distributed as an OWL (Ontology Web Language) file[3] to give the users the chance to have a computable format for the hierarchy.

In Italy, LOINC DO codes are primarily used to identify clinical documents indexed into the EHR registries of the Regions. They represent the values of the metadata *document type* in the Italian AD. Because of this and considering their reference role as a standard in this domain, Medas, an Italian private company, which provides services of long-term preservation of clinical documents, decided to map its local document type code system to LOINC in order to evaluate the semantic coverage of the standard and the chance to adopt it to be compliant with the national EHR and speed up the processes of correct document type recognition, effective classification and efficient retrieve.

Medas created its own coding system to identify clinical documents received by healthcare providers for long-term preservation. It was necessary to avoid inconsistencies in data item entries. It is based on two hierarchical levels: Genus and Species. The Genus level is the general form in which the document is presented, it always has the same meaning regardless of the associated Species. The Genus' identifier is indicated in the first part of the code of each document type, e.g. the Genus code 090 is for *Report*. This is a speaking code, as the same number always corresponds to the same document type. The Species level indicates the specific type of the document and, differently from the Genus, it is not a speaking code, as the same number changes its meaning depending on the Genus to which it is associated. For example, 090.060 is the code for *Cardiology Report*. In this case, 060 is the Species code which indicates the medical specialty *Cardiology*. But the same Species code 060, if associated with another Genus, has a different meaning. For exam-

[3] For more details see: https://loinc.org/document-ontology/owl/.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

ple, 100.060 is the code for *Vaccination register*, where 100 is the Genus code for *Register* and 060 is the Species code for *Vaccination*. The total number of clinical document types encoded in the Medas local coding system is 483.

Methods

Mapping aims to establish correspondences from concepts of a starting coding system, which is usually local and called *source* system, to concepts of a destination coding system, which is usually standardized at some level and called *target* system. This is the main way to implement the use of a clinical standardized coding system into a healthcare setting. The mapping is a non-trivial task because it is necessary to establish semantic correspondences, which refer to the content actually represented, beyond the names given to the clinical documents (Chiaravalloti 2022). This is not always simple if the standard is to be used by users who are different from the document producers, who actually knows its content, which may also change depending on the clinical flow in which it is integrated.

For the purpose of this work, mapping operations were manually performed by LOINC Italia experts. The work was approached by considering Medas document types according to their Genus in order to group together documents that are similar in their general meaning. This facilitated the maintenance of mapping coherence for documents that conceptually belong to the same semantic group and, instead, differ in the Species, mainly related to medical specialty. The search into the LOINC database was carried out through the LOINC Search web browser and the software RELMA (REgenstrief LOINC Mapping Assistant) using both local names and possible synonyms of the names of the Medas document types.

The LOINC codes candidates for mapping were primarily searched into the DO class. If they were not found there, the chance of identifying the target codes in other classes of the standard was then considered.

Similarly to previous abovementioned studies, we classified the type of matching between the source and the target codes as:

- *Adequate or Exact 1-1* to indicate terms with a full match between source and target systems;
- *Too broad* to indicate mappings to a LOINC concept, which is more generic than the source term;
    - *Many-to-1*: we used this subcategory tag when multiple local codes were mapped to the same LOINC code. In this case, the specificity expressed in the source coding system doesn't find semantic correspondence into the target coding system, which groups it into a single generic code;
- *Too specific* to indicate mappings to a LOINC concept, which is more specific than the source term;
    - *1-to-many*: we used this subcategory tag when one local code was mapped to multiple LOINC codes. In this case, the source coding system groups together concepts that are instead distinct in different codes of the target coding system.

Not covered to indicate no matching terms between source and target systems.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

In Table 1 some examples for each category are reported.

Type of matching	Medas	LOINC
Adequate (Exact 1-1)	090.100 Referto di chirurgia pediatrica (*Pediatric surgery Diagnostic study note*)	68794-7 Pediatric surgery Diagnostic study note
Too broad	090.140 Referto di dermatologia e venereologia (*Dermatology and venereology Diagnostic study note*)	78260-7 Dermatology Diagnostic study note
Many-to-1	090.090 Referto di chirurgia maxillo-facciale (*Maxillofacial Surgery Note*) AND 090.095 Referto di chirurgia orale (*Oral Surgery Note*)	34813-6 Oral and Maxillofacial Surgery Note
Too specific	090.070 Referto di chirurgia dell'apparato digerente (*Digestive system surgery Note*)	84073-6 Colon and rectal surgery Note
1-to-many	090.030 Referto di anestesia e rianimazione (*Anesthesiology and Critical care medicine Note*)	34750-0 Anesthesiology Note AND 34754-2 Critical care medicine Note
Not covered	090.040 Referto di biochimica clinica (*Clinical biochemistry Diagnostic study note*)	-

Table 1. Examples of mappings from Medas to LOINC codes for each defined type of matching.

The identified matchings were then validated by a domain expert from LOINC Italia, different from the one who performed the mapping.
After the validation phase, further meetings with the Medas archivists helped to understand the semantic area covered by Medas codes matched as *too broad*, *too specific* or *not covered* and therefore to determine whether they could be classified differently or they need to be required as a new term submission to the LOINC DO Committee.

Results

Out of the total of 483 clinical document types encoded in the Medas local system, their mappings to LOINC codes resulted in:

- 144 adequate mappings, with an exact correspondence of 1-1 between the source concept in Medas and the target concept in LOINC;
- 209 too broad mappings, where the target LOINC code only partially covers the meaning of the source concept in Medas. In these cases, there is a loss of information moving from the source to the target coding system;
- 2 too specific mappings, where the target LOINC code only partially covers the meaning of the source concept in Medas. In these cases, there is an excessive specialization of the information moving from the source to the target coding system;
- 128 Medas clinical document types were not mapped as no semantic correspondence of concepts could be found in LOINC.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

In percentages it means that 30% of the Medas local codes have a full match with one and only one LOINC code, 44% of them have a partial match, because in LOINC only broader or narrower concepts can be found, and finally 26% of the Medas codes do not have any match in LOINC as it was not possible to find a semantic representation of their meaning.
If we look in more detail at these percentages, it is interesting to note that among the Medas Genus codes that have the highest full match percentages there are *Letters*, *Prescriptions* and *Reports*. On the contrary, among the document types with less full match there are *Minutes* and *Declarations*, while the Medas Genus codes related to *Registers* and *Diaries* are not represented at all in LOINC (Table 2).

	Full match	**Partial match**	**No match**
Report	76%	5%	19%
Declaration	22%	22%	56%
Letter	80%	20%	0%
Prescription	73%	0%	27%
Minutes	53%	6%	41%
Registry	0%	0%	100%
Diary	0%	0%	100%

Table 2. Percentages of full, partial and no match for some of the Medas Genus.

In searching for correspondences between the Medas Genus and the LOINC codes, differences inevitably emerged deriving from the contexts in which the two coding systems were born. In fact, legislation has a great impact on the definition of types of clinical documents. Codes from the Medas system therefore do not always find semantic correspondence in LOINC, because they represent peculiar aspects of the Italian context. Conversely, many of the LOINC codes were created to reflect the clinical document types typical of the North American context. For these specific cases coding can therefore only take place with the creation of LOINC codes customized for the purpose. An example of this is the Genus *Certificate*. The Medas document types belonging to it are 21, and 7 of them full match with a LOINC code, while the remainder have been classified as *too broad* matchings as they can be either mapped to the LOINC 80571-3 *Certificate* or related to the Kind of Document *Note*, thus not allowing to specify information about the punctual function of a certificate. A certificate is, in fact, defined as a "written declaration, issued by a public authority or even, in specific cases, by a private individual, which expresses knowledge of the existence or truth of a fact or expresses the result of an evaluation",[4] which means that it has a more specific and official value than a simple clinical note. Either way the specificity of the information would be lost, therefore those concepts are candidates for new LOINC codes' submissions.
Another special case is the Genus *Consent*. In the Medas coding systems it contains 133 clinical document types, but only 6 of them have a full match in LOINC. Although the source codes are of such a large number, the LOINC codes retrieved by typing the word *consent* are only 19. This

[4] The translation is derived from https://www.treccani.it/vocabolario/certificato/.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

discrepancy is due to the Italian legislative system (Chiaravalloti et al. 2015), which is greatly focused on patients' rights and privacy protection, requiring consent documents for almost every medical procedure or event. Furthermore, there are document types not only to express consent in a positive form, but also to express the denial of consent or the modification of previously given consent. There is thus an information gap between the source and the target coding systems. In Table 3, some examples of mapping *many-to-1* for the Genus *Certificate* are reported. In them, moving from the Medas coding system to LOINC there is a loss of information because Medas document types are narrower than the concepts conveyed by the existing LOINC codes. Therefore, in this case the choice is whether to require LOINC DO Committee to create new codes to maintain the specificity of the source information or whether to use the existing generic LOINC codes as a sort of higher-level categories, but with the knowledge that they act only as containers of *n* more specific document types.

DocumentTypeLabel	**LOINCDocumentCode**	**LOINCDocumentName**
Carotid stenosis consent	64293-4	Procedure consent document
MRI consent		
CT consent		
PET CT consent		
Consent for surgery	61358-8	Surgical operation consent document
Obesity surgery consent		
DBP review consent		
Bifemoral bypass consent		

Table 3. Examples of mappings many-to-1 for the Medas Genus Consent.

Regarding the Medas Genus *Report*, it is intended as "the written report issued by the doctor who subjected a patient to a clinical or instrumental examination".[5] To map this type of clinical documents, it was initially necessary to disambiguate the meaning of the LOINC document types *Note*, *Diagnostic study note* and *Report*. After also a discussion with members of the LOINC DO Committee, it was established that *Note* is the most generic term to indicate a report resulting from a medical visit, while *Diagnostic Study Note* is used when the document is the result of a diagnostic study, therefore more specifically created with the support of medical devices. *Report*, instead, has a more limited use in English because it concerns what is reported because it has been originally expressed in another way such as the Laboratory report. Out of the 87 *Report* types in the Medas coding system, 60 have been mapped to a LOINC code. Nonetheless, it was not always possible to attribute a 1:1 correspondence between the source and target concepts. Some of the most particular cases detected are presented below. There are cases of 2:1 correspondence, in which two different types of clinical documents of the source coding system are mapped towards the same type of clinical document in the target coding system, which is the LOINC standard. Some examples are:

[5] The translation is derived from https://www.treccani.it/enciclopedia/referto_%28Dizionario-di-Medicina%29/.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

- *Cardiology report* and *Cardiovascular system disease report* have been both mapped to the LOINC code 75425-9 *Cardiology Diagnostic study note;*
- *Oral surgery report* and *Maxillofacial surgery report* have both been mapped to the LOINC code 34813-6 *Oral and Maxillofacial Surgery Note*.

In other cases, it was necessary to carry out mappings towards broader LOINC terms compared to the concepts conveyed by the Medas document types. This type of association inevitably leads to a loss of information in the transition from the local to the standard coding system. Some examples are:

- *Dermatology and Venereology report* was mapped to the LOINC code 78260-7 *Dermatology Diagnostic study note*, thus losing the information relating to the missing SMD Venereology;
- *Anesthesia and resuscitation report* can be mapped either to the LOINC code 34754-2 *Critical care medicine Note* or to the LOINC code 34750-0 *Anesthesiology note*. In both cases the information relating to one of the two mentioned medical branches is missing.

Some types of Medas *Report* have not been mapped because the corresponding specialist branch is missing in the LOINC DO. It should be represented in the SMD axis; therefore, these could be addition proposals for the standard in order to have greater semantic coverage of the reference domain. Some examples are: *Clinical Biochemistry*, *Medical Physics*, *Thermal Medicine*, *Tropical medicine*, *Health statistics and biometrics*, *Senology*, *Orthodontics*, *Neurophysiopathology*, *Orthopedics and Traumatology*, *Clinical Psychology*, *Neuropsychiatry*, and *Child Neuropsychiatry*.
It is to be noticed that the case of *Forensic medicine report* as *forensic medicine* is present as a possible value for the SMD axis, but the only document type to be associated is *Referral note* therefore the representation of the *report* concept associated with this specialist branch is missing.

Discussion

The abovementioned mapping results show that LOINC is the most suitable standard for uniquely identifying clinical document types, as it was possible to find a semantically representative LOINC code for most of the concepts expressed in the Medas coding system. Out of a total of 483 document types analyzed, it was possible to map 144 of them with a full match of the concepts in the source and target coding systems, and 211 of them with a partial match.
However, although the DO is continuously updated to be representative of emerging clinical needs, its semantic coverage can be improved to adapt to national contexts other than the US master one. In this work we identified some additions to the DO axes that can be submitted to the LOINC DO Committee. Specifically, possible extensions for the Kind of Document axis are *Register*, *Diary*, *Index*. The SMD axis could be extended to include *Venereology*, *Clinical Biochemistry*, *Medical Physics*, *Thermal Medicine*, *Tropical Medicine*, *Health Statistics and Biometrics*, *Orthodontics*, *Neurophysiopathology*, *Orthopedics and Traumatology*, *Clinical Psychology*, *Senology*, *Neuropsychiatry*, and *Child Neuropsychiatry*. Finally, for the Type of Service axis a possible concept to submit is *Dispensation*. The process of creating a new LOINC term after submission is quite fast, which makes the standard particularly flexible in adapting to new scientific discoveries and the needs of the users' community (Cardillo et al. 2024). Moreover, this moment is particularly

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

timely, as announced during the LOINC 2024 conference, the LOINC DO will undergo a comprehensive review process.
Beyond the lack of a specific LOINC code for each document type, the problem relating to the concepts matching can be considered broader and it regards, for example, the organizational logic used in the different contexts. Therefore, this type of work cannot be separated from a preliminary analysis of the starting context nor to be uniform for all organizations dealing with the long-term preservation of clinical documents. The strategy that proved to be successful in our case for solving some mapping issues was the direct comparison with those who developed the source coding system. This allowed us to identify the most suitable solution, which in some cases was to identify a suitable LOINC code using synonyms, in others to propose the concept for a new submission to the LOINC DO Committee as it was not represented in the standard.
The study also offered the possibility of analyzing synonymy in LOINC and above all the richness of synonyms currently present in the standard database. They are important because they are further, often valuable, search keys to arrive at the target concept. It should be noted that for the documents section of LOINC there is not a wealth of synonyms comparable to that present for the laboratory and clinical sections. On the one hand, this could be an indication of a certain uniformity of the context from which the DO values were created, but on the other hand it could constitute an obstacle in finding the codes in contexts, such as the Italian one, in which there is a high linguistic variability in naming the types of clinical documents.
Furthermore, the study also made it possible to verify the names of the clinical documents used in the Medas coding system, reconciling those that are different in form but equal in substance. It happens, in fact, that in a post-coordinated coding system, such as the Medas one, control over the expansion of all the branches of the hierarchy can be lost. Therefore, repetitions of concepts can be created, expressed through synonyms, due to the different names of the clinical documents used in the local systems of the healthcare facilities, which use the long-term preservation services offered by Medas. However, this aspect can be converted into an advantage, considering that the results of this mapping towards an international standard could lay the foundations for the development of guidelines to regulate naming practices within single healthcare facilities or at a regional and national level. It could mean reorganizing existing naming conventions or primarily regulating them at the time of the creation of the document management system.

Conclusion

This paper deals with the process of aligning the names of the types of clinical documents placed in long-term preservation with the international reference standard for the domain, namely LOINC. This activity is carried out through mapping operations that connect the semantically equivalent concepts of the source coding system to those of the target or destination coding system. The case study described was created in collaboration with Medas, a leading Italian company in providing long-term preservation services for clinical documents. The results obtained show how the Medas local coding system allows a certain flexibility through its structuring into Genus and Species and is well represented through the LOINC codes. The concepts that could not be mapped to the international standard, however, mainly concern specific areas of the Italian context and missing

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

values in LOINC. It was thus possible to evaluate the semantic coverage of the standard and consequently identify concepts to candidate for LOINC new terms submissions.
Regarding research perspectives, the database of the validated mappings of the Medas coding system towards LOINC could represent a knowledge base for supporting mapping operations of clinical document types defined by other companies in the healthcare domain. The common reference to a standard could, in fact, create uniformity where a high degree of linguistic variability remains in determining the names of clinical documents. It should move from a more general reflection on a rationalization of the document types in the document management system, considering that the transition from analog to digital does not mean "translating" existing documents' types into the corresponding digital ones, but a "rethinking" of the entire process is necessary in order to streamline not only the single document types but above all the documents' flows. Furthermore, results of this work could be used to enrich the list of LOINC document type codes, which are the metadata values to index clinic document types into the Italian FSE according to the AD. From an international perspective, the use of a standard like LOINC represents an advantage in terms of cross-border clinical interoperability within the exchange of clinical documents among EU countries and beyond.

Acknowledgements

We would like to thank Medas Company (particularly Umberto Ferri and Carlo Alzati) for supporting and sharing their data used in this study. Details about Medas coding system are not shown in this study because of property rights, therefore only aggregated analysis and significative examples are used to highlight the case study.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

References*

Brown, Steven H., Micheal Lincoln, Shawn Hardenbrook, Olga N. Petukhova, S. Trent Rosenbloom, Paul Carpenter, and Peter Elkin. 2001. "Derivation and Evaluation of a Document-Naming Nomenclature." *Journal of the American Medical Informatics Association* 8 (4): 379–390. https://doi.org/10.1136/jamia.2001.0080379.

Cardillo, Elena, Maria Chiaravalloti, and Erika Pasceri. 2024. "Covid-19 Impact on Standard Coding Systems Update." In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: HEALTHINF*, 490–497. https://doi.org/10.5220/0012387600003657.

Chen, Elizabeth S., Genevieve B. Melton, Mark E. Engelstad, and Indra Neil Sarkar. 2010. "Standardizing Clinical Document Names Using the HL7/LOINC Document Ontology and LOINC Codes." In *AMIA Annual Symposium Proceedings*, November 13, 2010, 101-105.

Chiaravalloti, Maria Teresa, Mario Ciampi, Erika Pasceri, Mario Sicuranza, Giuseppe De Pietro, and Roberto Guarasci. 2015. "A Model for Realizing Interoperable EHR Systems in Italy." In *Proceedings of the 15th International HL7 Interoperability Conference*, Prague, Czech Republic, February 9-11, 2015, 13-22.

Chiaravalloti, Maria Teresa. 2022. "A proposal for a LOINC automatic mapping support tool." *AIDAinformazioni* 3-4:47-56.

Ciampi, Mario, Giuseppe De Pietro, Christian Esposito, Mario Sicuranza, and Paolo Donzelli. 2012. "On Federating Health Information Systems." In *Proceedings of the International Conference on Green and Ubiquitous Technology*, Bandung, Indonesia, June, 30–July 1, 2012, 139-143.

De Pietro, Giuseppe, and Mario Ciampi. 2011. "L'interoperabilità del FSE: il progetto OpeninFSE." *E Health Care Innovazione e Tecnologia Sanità Elettronica* 3: 8-14.

Dugas, Martin, Sylvia Thun, Thomas Frankewitsch, and Kai U. Heitmann. 2009. "LOINC Codes for Hospital Information Systems Documents: A Case Study." *Journal of the American Medical Informatics Association* 16 (3): 400-403. https://doi.org/10.1197/jamia.m2882.

Frazier, Pavla, Angelo Rossi-Mori, Robert H. Dolin, Liora Alschuler, and Stanley M. Huff. 2001. "The Creation of an Ontology of Clinical Document Names." *Studies in Health Technology and Informatics* 84 (Pt 1): 94-98.

Hyun, Sookyung, Rosemary Ventura, Stephen B. Johnson, and Suzanne Bakken. 2006. "Is the Health Level 7/LOINC Document Ontology Adequate for Representing Nursing Documents?" *Studies in Health Technology and Informatics* 122: 527-531.

Hyun, Sookyung, Rosemary Ventura, Stephen Johnson, and Suzanne Bakken. 2009. "Iterative Evaluation of the Health Level 7—Logical Observation Identifiers Names and Codes Clinical Document Ontology for Representing Clinical Document Names: A Case Report." *Journal of the American Medical Informatics Association* 16 (3): 395-399. https://doi.org/10.1197/jamia.M2821.

* Websites were last accessed on 23 November 2024.

JLIS.it vol. 16, no. 2 (May 2025)
ISSN: 2038-1026 online

DOI: 10.36253/jlis.it-632

Italia. *Legge 22 dicembre 2017, n. 219, "Norme in materia di consenso informato e di disposizioni anticipate di trattamento."* Gazzetta Ufficiale n. 12, January 16, 2018.

Li, Li, C. Paul Morrey, and David Baorto. 2011. "Cross-Mapping Clinical Notes between Hospitals: An Application of the LOINC Document Ontology." In *AMIA Annual Symposium Proceedings*, October 22, 2011, 777-783.

Rajamani, Sripriya, Elizabeth S. Chen, Mari E. Akre, Yan Wang, and Genevieve B. Melton. 2015. "Assessing the Adequacy of the HL7/LOINC Document Ontology Role Axis." *Journal of the American Medical Informatics Association* 22 (3): 615-620. https://doi.org/10.1136/amiajnl-2014-003100.

Rajamani, Sripriya, Elizabeth S. Chen, Yan Wang, and Genevieve B. Melton. 2014. "Extending the HL7/LOINC Document Ontology Settings of Care." In *AMIA Annual Symposium Proceedings*, November 14, 2014, 994-1001.

Shapiro, Jason S., Suzanne Bakken, Sookyung Hyun, Genevieve B. Melton, Cara Schlegel, and Stephen B. Johnson. 2005. "Document Ontology: Supporting Narrative Documents in Electronic Health Records." In *AMIA Annual Symposium Proceedings*, 684-688.

Wang, Yan, Serguei Pakhomov, Justin L. Dale, Elizabeth S. Chen, and Genevieve B. Melton. 2014. "Application of HL7/LOINC Document Ontology to a University-Affiliated Integrated Health System Research Clinical Data Repository." In *AMIA Joint Summits on Translational Science Proceedings*, April 7, 2014, 230-234.

Zuo, Xu, Jianfu Li, Bo Zhao, Yujia Zhou, Xiao Dong, Jon Duke, Karthik Natarajan, George Hripcsak, Nigam Shah, Juan M. Banda, Ruth Reeves, Timothy Miller, and Hua Xu. 2021. "Normalizing Clinical Document Titles to LOINC Document Ontology: An Initial Study." In *AMIA Annual Symposium Proceedings*, January 25, 2021, 1441-1450.