Integrating the Healthcare Enterprise

IHE Europe

Guideline for interoperable XDS Affinity Domains

Metadata for exchange medical documents and images

Date:	June 29th , 2017
Version	1.0
Author:	IHE Europe – Co Authors
Email:	xds_metadata@googlegroups.com

Authors

From the International XDS Metadata Taskforce:

- Karima Bourquard, IHE-Europe
- Abderrazek Boufahja, IHE-Europe
- Pablo d'Alcantara, Abrumet, Belgium
- Samuel Danhardt, Agence eSanté, Luxembourg
- Didi Davis, Sequoia, USA
- Emma Harris, IHE-UK
- Marko Jalonen, Salivirta&Partners Oy
- Manuel Metz, ASIP Santé France
- Frank Oemig, Germany
- Charles Parisot, GE Healthcare
- Konstantin Hyppönen, Kela Finland
- Neil Robinson, IHE-UK
- Thor Schliemann, Danish Health Data Authority, Denmark
- Vincent van Pelt, Nictiz, the Netherlands

General information about IHE-Europe can be found at: http://ihe-europe.net

Information about the IHE IT Infrastructure domain can be found at: http://ihe.net/IHE_Domains.

CONTENTS

Executive Summary 4

Reading Guide 5

1. Introduction 6
   1.1. About the project 6
   1.2. Problem definition 6
   1.3. Organization 7
   1.4. Objectives 7
   1.5. Expected Benefits 8
   1.6. Target audience 8
   1.7. Scope 8
   1.8. Guiding principles 9
   1.9. Action Points 10

2. Information exchange in healthcare 10
   2.1. Interoperability 10
   2.2. Interoperability model 11

3. Categorization of documents and images 14
   3.1. An inventory of healthcare related documentation 14
      3.1.1. Explosion of health data 14
      3.1.2. Many types of documents 16
      3.1.3. Purpose of categorization 17
      3.1.4. Categorization challenges 18
   3.2. Organizing and structuring metadata` 20
      3.2.1. Approach 20
      3.2.2. Categorization model 22
      3.2.3. Categorization of images and textual documents 24
      3.2.4. Selecting the values for the attributes 29
   3.3. XDS Information Model 29

4. Harmonizing different initiatives 31
   4.1. Analysis of existing (proposals for) solutions 31
      4.1.1. Cross-Enterprise exchanges 31
      4.1.2. National initiatives 32
      4.1.3. European and global perspective 40
      4.1.4. The XDS metadata project 40
      4.1.5. Code Systems 41
   4.2. Discussion 43

5. IHE XDS.b and XDS.b-I metadata 44
   5.1. Metadata as classification tool 44
   5.2. Metadata as Query and Filtering tool 45
   5.3. Metadata as an end-to-end tool 45
   5.4. Domain specialization 46

6. Use of metadata 47
6.1. Use if metadata by IHE 47
6.2. Cross-domain interoperability 49
6.3. XDS metadata and interoperability 50
6.4. Linking the generic categorization model to XDS metadata set 51
7. Appendix A. XDS metadata-relevant elements for functional classification 55
7.1. XDS-I.b metadata elements 62
8. Appendix B. Filtering levels and XDS metadata 65
9. Appendic C. DICOM Categories 71
10. Appendix D. Radlex Playbook 72
11. Appendix E. LOINC 73
12. Appendix F. Links to relevant sites 74
13. Appendix G. Definition 76

Executive Summary

The XDS (Cross-Enterprise Document Sharing) and XDS-I profiles facilitate the secure, reliable and interoperable exchange of medical documents and images within XDS Affinity Domains, based upon the profile interoperability requirements that vendors must follow in their implementations. These requirements specify actors and transactions that enable software products from different vendors to cooperate and exchange information. In the case of XDS, they also specify the document metadata concepts that constitute the XDS document registry. For some of these metadata elements, the *values* that can be assigned to these concepts are defined in the profile specifications. However, for other metadata elements, these values have not been defined and are left to the implementing parties to assign.

Within an Affinity Domain, this may work out fine, because the participating healthcare organizations draw up their own set of metadata as they go along. But as the XDS communities mature, share a broader range of health information and increase in number, these communities are interested in becoming interconnected. This is the point when cross-community IHE profiles such as XCA (Cross-Community Access) support information exchange between XDS Affinity Domains possible. However the lack of a uniform definition of these metadata elements across communities becomes an obstacle for true interoperability. The current XDS metadata constraints set in the XDS and XCA profiles still leave too many degrees of freedom to enable seamless interoperability between Affinity Domains.

In countries where XDS networks are being set up, initiatives have risen to establish national metadata definitions of these hitherto not sufficiently defined metadata elements. In 2016, IHE representatives from 10 different countries (Austria, Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, the Netherlands, Switzerland, United Kingdom), IHE Europe and

the US have joined forces in the *International XDS Metadata Taskforce*, with the following goals:

1. To share experiences on XDS implementation in European countries;

2. To share views on how the metadata are used (use cases, terminologies, data);

3. To find a common way to harmonize the metadata

4. To build an implementation guide for those wanting to implement XDS metadata in future projects.

The first deliverable of the Taskforce is to present the white paper called "*Metadata for Exchange of information in XCA-XDS infrastructures*" in which the subject of document categorization is described and analyzed. After analysis of (draft and operational) versions of metadata sets from different countries, guiding principles based on best practices are being drawn up and a draft proposal for uniform value sets for metadata elements such as classCode, typeCode, eventCodeList (see next sections), healthCareFacilityTypeCode and practiceSettingCode are being proposed.

This work was presented at the 2017 European Connectathon to the IHE community in April 2017 during the IHE Symposium.

We, the taskforce, hope that this document can become the basis of an implementation guideline for interoperable XDS networks in Europe and beyond.

Reading Guide

This is version 0.91 of the ***Guideline on the use of metadata in Information Exchange between XDS Affinity Domains.***

A number of other documents are also referred to within this document, sometimes via hyperlinks.

The Guide to Interoperability between XDS Affinity Domains consists of the following chapters:

- ***Chapter 1*** *is an introduction to the project and its goals, organization, objectives, expected benefits, target audience, scope and approach;*
- ***Chapter 2*** *contains a short introduction to the concept of interoperability and to the interoperability model used in Europe validated by the eHealth network (eHN);*
- ***Chapter 3*** *contains a generic introduction to medical documents and their categorization, some theory on organizing and structuring metadata, and some guiding principles;*
- ***Chapter 4*** *focuses on the IHE XDS profile, specifically on the XDS metadata;*
- ***Chapter 5*** *looks at the different national initiatives and looks at patterns;*
- ***Chapter 6*** *provides recommendations for an international reference implementation of XDS metadata attributes;*
- ***Appendix A*** *Provides an overview of XDS and XDS-I metadata;*
- ***Appendix B*** *Links specific XDS metadata to levels of filtering;*
- ***Appendices C, D, E and F*** *list relevant terminology listings from DICOM, Radlex, SNOMED-CT and LOINC;*

- ***Appendix G*** *contains a list of relevant links;*
- ***Appendix H*** *is a list of relevant terms and abbreviations.*

1. Introduction

1.1. About the project

For the exchange of medical information between XDS environments, a shared set of document metadata is required. Although the XDS profile is a guideline for implementation, the profile still leaves degrees of freedom in the definition of the value sets that are used to fill in the metadata concepts defined in the profile definition. This degree of freedom becomes a problem when XDS Affinity Domains want to exchange documents with each other, because if they use different value sets, interoperability will partly be lost when multiple XDS Registries need to be queried. For cross-community information exchange, agreements must be made on how the value sets of all the document metadata concepts across the various Document Registries should be filled in.

This document is the result of an initiative that started in May 2016 by some members of the IHE community who had noticed that several countries in Europe and abroad are looking for a uniform set of specifications for the IHE XDS metadata set. They decided to set up The *International XDS Metadata Taskforce* on XDS metadata harmonization over Europe and the US, to

- Share experiences on XDS implementation in European countries;
- Share views on how the metadata are used (use cases, terminologies, data);
- To find a common way to harmonize the metadata;
- To build an implementation guide for those want to implement XDS metadata in their projects.

1.2. Problem definition

The XDS profile definition covers, amongst other things, the metadata information elements that describe the context of the medical documents and images within an XDS Affinity Domain. It defines the different concepts that constitute the metadata 'fields' of the XDS Registry. These concepts can be used to sort, filter and group the available documents for easier access by the end-user. However, for some of these metadata elements, it is left up to the implementer of an XDS Affinity Domain to fill in the possible values that some of these concepts may have. Some suggestions are made, but these leave enough room for different implementation choices.

As an example, for one of the XDS metadata elements, classCode, the XDS profile definition states that "… valid values for classCode attribute are specified by the policies of the creating entity". In other words, it is left up to the organization that sets up an XDS Affinity Domain to create their own value sets. And this poses a problem when multiple XDS Affinity Domains want to exchange information, because they most probably will have created different value sets

for this classCode metadata concept, and may not be able to sort, filter and group the available documents in a consistent and predictable way, thus resulting in gaps in interoperability.

Unless each Affinity Domains use the same (or compatible subsets) of metadata value sets, the possibilities for sorting, filtering and grouping will be seriously thwarted, reducing the possibility for cross-community information exchange. This problem has been recognized by different countries where XDS infrastructures have been deployed and where XDS Affinity Domains want to transparently exchange medical information with other XDS Affinity Domains. As a result, countries throughout Europe but also in the US have started initiatives to come to national agreements on how these different metadata value sets should be filled in.

In addition, even stand-alone XDS Affinity Domains have found that the level of skills needed to fill the definitions left open by the IHE XDS Profile to be much more advanced than they expected and realized that the risk of making metadata design mistakes, a significant challenge, as once defined, it is very difficult to change the definition of document metadata when a set of legacy shared documents exists.

As it happens, some of these countries are now (2016) at a point where they are working on draft or pilot implementation versions of national guidelines for XDS metadata, and other countries are thinking about doing something similar. This provides an excellent opportunity to join forces, exchange ideas and work out the possibilities to come to a shared vision on how these metadata can be defined in an internationally accepted implementation guideline.

1.3. Organization

The *International XDS Metadata Taskforce* consists of participants from Austria, Belgium, Denmark, Finland, France, Germany, Italy, Luxemburg, the Netherlands, Switzerland, the United Kingdom and the United States of America, as well as members from IHE International (see author list).

1.4. Objectives

Proper documentation, exchange and integration of healthcare information are key factors in modern-day healthcare. But getting the right information for the right purpose at the right time is vital too. Quick access to the kind of information that is needed at a specific point in time relies on the possibility to quickly select the available information for the task at hand. Information about the why, what, where, how, for and with whom provides context to the document or image. Also, functional and technical information about the document itself is relevant. This information is categorized in the metadata that accompanies the document when it is stored. The metadata are used by information systems to access and process efficiently these documents, such as to organize, sort and filter the medical documents, images, diagnostic study results, notes and summaries of a patient, to present healthcare professionals and other stakeholders with the right overview and quick access.

The subject of this paper is to create an overview of all these categories in a logical manner and to provide some guiding principles for a practical, logical and flexible classification system of health-related patient centric documents. This knowledge and these principles are then used to propose a refinement of the definition of the XDS metadata parameter elements.

1.5. Expected Benefits

The expected benefits of this endeavour can be summarized as follows:

- A well-defined categorization of healthcare related document, images and other related information carriers can be used in different fields in healthcare ICT. For IHE, XDS metadata can be used in different profiles (see chapter 4.2). But it can also be used in the HL7 domain for authorization on the level of document type (for example, in CCDA or FHIR Composition).
- A well thought through metadata set can be used as a point of departure by all who want to implement a document sharing environments based on XDS, MHD, XCA, XDR, XCDR, XDM profiles environment. Implementation costs and inefficiencies in the use of metadata go down because the metadata set is more robust and the same at each new install.
- Automatic interoperability between XDS Affinity Domains. Metadata filtering, sorting and grouping will work in a consistent and effective way across Affinity Domains.
- Software that uses XDS metadata can be optimized and built upon standard selection options due to the standardized metadata set;
- National extensions to document metadata can become a part of the implementation guideline to cater for specific national values;
- More 'off the shelf' implementations will lead to more cost-effective implementations of XDS Affinity Domains;
- Implementation guidelines form the basis for verifiable quality assessment of interoperability between different XDS infrastructures.

1.6. Target audience

This document is meant for executives, managers, information architects, analysts and technicians who are involved in setting up, designing and /or maintaining XDS Affinity Domains, and who are looking for answers to the issues that arise when different XDS Affinity Domains want to exchange information (through XCA). This document highlights management, organisational, healthcare and technical aspects, and can be consulted by these target groups.

1.7. Scope

This document provides an inventory, analysis, logical model and proposed implementation rules for the categorization of medical information carriers. It also gives propositions for specific XDS metadata elements, such as classCode, typeCode, eventCodeList, HealthcareFacilityTypeCode and others.

Out of scope are all aspects of governance, maintenance, versioning and so forth. These should be filled in by the IHE organization to fit in with the already existing procedures.

Current status

One of the first tasks of the taskforce has been to make an inventory of the different national proposals for an XDS metadata set. So far, input from Austria, the Netherlands, France, Switzerland, Finland, Germany and the UK have been gathered. These different proposals, together with information from other sources, have been used as the basis for comparison and inspiration.

This document is the first 'deliverable' of the Taskforce.

1.8. Guiding principles

After having looked at the different national initiatives towards national definition of the XDS metadata elements, it has become clear that a few guiding principles must be defined that delineate the guiding principles, boundaries and constraints. Here are the points of departure that we have used in the discussion about how the XDS metadata set should be built up:

- Real interoperability only exists at the implementation level.
- Categorization of the XDS metadata must reflect the usefulness for clinicians and other persons and applications that work within an XDS environment. The XDS metadata are mainly used by healthcare professionals, for quick selection (most likely relevant) of the desired documents and images. Therefore, categories must have medical and functional meaning to the end-user. Of course, some more 'technical' metadata attributes are needed as well.
- First agree upon the logical categorization axes, then create a model for all the different metadata attributes and accompanying value sets. The term "Axes" here relates to the fact orthogonality between these dimensions of the metadata, so that one intended "selection" is a unique set of "coordinates" (this is intended to remove ambiguity and simplifies the use of the metadata).
- When the logical model is finalized, look at the technical definition of the XDS profile and see how this logical model fits the XDS metadata elements. If some elements from the 'ideal' model do not fit in the XDS metadata elements, decide what to do with these elements. Discard them, try to fit them in the existing metadata set, or produce a IHE Change Proposal.
- Look whether existing international standards can be used as the basis for all value sets.
- Connect to other standards developing organizations (SDOs) such as HL7, LOINC, DICOM, SNOMED-CT when it is relevant, for consultation and cooperation.
- Allow for national/regional or local extensions of the proposed metadata value sets.
- Link all the XDS metadata elements to terminologies. This allows for better understanding and definition of the meaning of the concepts and values, and makes cross-country and cross-language interoperability possible.

- All choices for value sets should be argued and substantiated. Do not just make choices, but tell why these choices were made. This increases the acceptance of the metadata value sets.
- The XDS metadata set must be able to contain and describe any type of patient centric health related document.
- To prevent ambiguity as to what categories must be selected when the document is created (by the XDS Document Source), a clear definition of each category must be given. Decision trees and guidelines for the proper selection of metadata values are recommended where the selection may be a problem.
- New agreements regarding exchanges between Affinity Domains should be made at the level of the participating organizations. They must test the interoperability and exchangeability between XDS Affinity Domains using cross-community information exchange. These tests could become part of a Connectathon session.

1.9. Action Points

The taskforce has identified the following action points:

- Inventory of document and image types used in healthcare;
- Inventory of available standards related to the subject (see section 12) ;
- Inventory of existing proposals for national metadata sets specifications (see section 4);
- Definition of guiding principles (see section 1.8);
- Analysis and comparison of the inventory materials;
- Logical argumentation to come to a comprehensive document categorization (see section 3);
- Selection of fitting terminologies and terminology entries to link to metadata element concepts and values (see section 4.1.5);
- Definition of an implementation guideline for an international XDS metadata set;
- Review by third parties (out of scope of the task force);
- Publication of the implementation guideline by IHE International (next step after validation of the document);
- Possible IHE Change Proposals (next step after validation of the document).

Some of these action points are not covered by this version of the guideline. Their planning is proposed

2. Information exchange in healthcare

2.1. Interoperability

The term ‘interoperability’ is comprehensive in that it describes all measures that need to be taken, by different stakeholders, to achieve the secure, reliable and efficient exchange of information. One of the definitions of interoperability is given below:

> Interoperability is the possibility of different autonomous, heterogeneous systems, equipment or other units (for example organizations or countries) to communicate with one another and interact. To achieve this, standards, protocols and procedures are needed to harmonize the different entities. (Wikipedia)

In order to achieve interoperability as defined below, other dimensions or layers should be taken into account because they support the implementation of the interoperability.

To make this clear, a European Commission project called Antilope[1], has presented an "interoperability multi layer model" that presents all the necessary layers from legal and regulatory layer, organisational and informational layers to application and infrastructure layers and to make this clear to all stakeholders involved in information exchange.

2.2. Interoperability model

The model is a refinement of the eHealth European Interoperability Framework. It is not a technical model, but explains the different aspects involved in deploying interoperability, where stakeholders on different levels have to come to terms and agreements. This model avoids technical terms and makes it clear that agreements must be made on and between all levels, and between all the parties involved. The Refined eHealth European Interoperability Framework, or ReEIF as it has been called, has been adopted in 2016 at the European level (eHealth Network[2]) as the standard model for interoperability.

[1] Antilope – see http://www.antilope-project.eu/wp-content/uploads/2013/05/D1.2a-Educational-material-presentation-v1_4.pdf

[2] eHealth Network – see http://ec.europa.eu/health/ehealth/docs/ev_20151123_co03_en.pdf

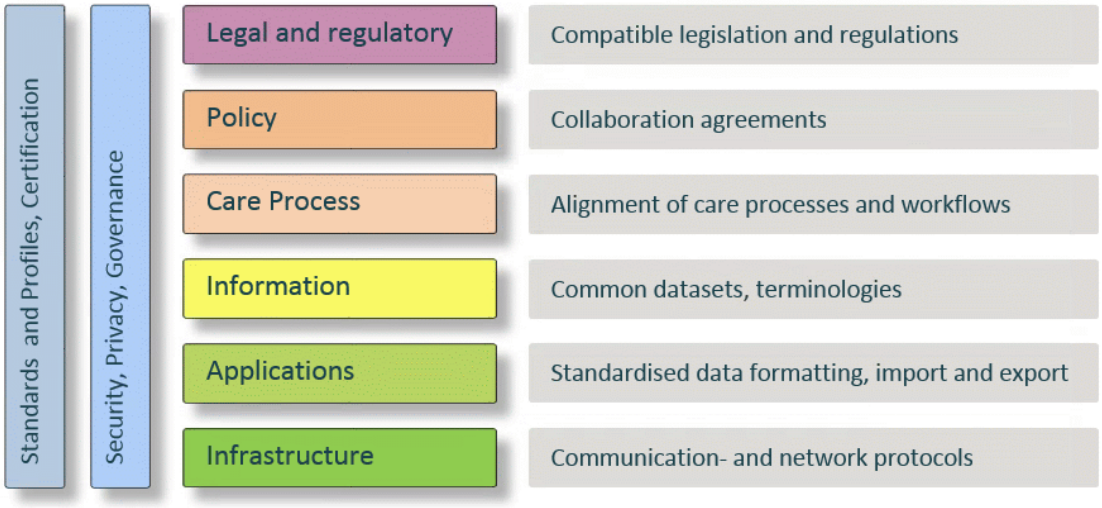


Figure 1 - eHN interoperability model (ReEIF)

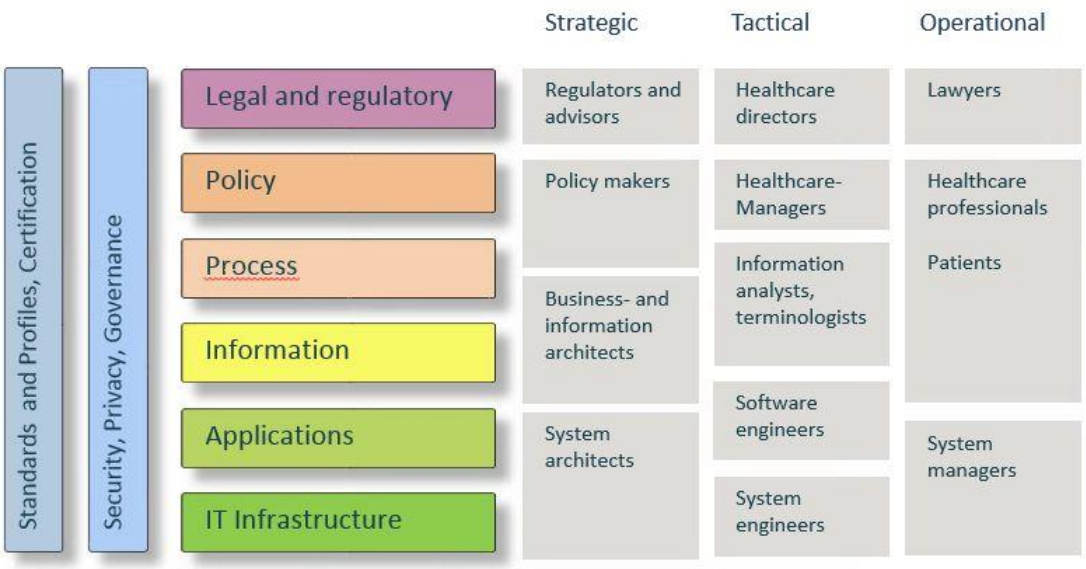


Figure 2 - Agreements between different stakeholders

The interoperability levels of the model are described briefly below:

Legislation and regulatory

Legislation and regulations indicate the limits which apply to the exchange of medical information.

Agreements are made at this level on the agreements on the implementation of the legislation and regulations.

Organization policy

In this layer, agreements are made at management level between cooperating organisations: organisation of the governance, cooperation contracts, framework and data processing agreements, as well as agreements on privacy and security, patient consent, uniform design of the infrastructure and so forth.

Care processes

At this level, use cases are described, workflows are defined, information transfer is harmonised and insight into the logistical processes of healthcare is made possible.

Information

At this level, the information to be transferred is defined: which information elements are needed to be exchanged and to which level of detail. Agreements must be made on the following matters:

- Dataset - Which information is transferred in a structured way? And which data elements and value sets are used for this?
- Terminology – this couples the concepts, and the values they may hold, with standardised terms.

Applications

At this level, the technical exchange format of the information is defined, such as HL7 CDA, FHIR or other formats. It also describes the way in which the documents and messages are exchanged between different ICT systems.

Infrastructure

This layer focuses on the infrastructure for the communication between systems in the different healthcare organisations. At this level, agreements are laid down on interoperable infrastructures and networks.

Security, privacy

Security and privacy are organised on all the levels mentioned above. At this level, agreements on how the legislation, norms and guidelines are to be implemented. It focuses on authentication and authorization, consents, the quality of the information and the safe transfer and storage of information.

Governance, testing

This level deals with the management, responsibilities and maintenance of the systems.

This interoperability model has proven to be very useful in explaining to different audiences that agreements and expertise are needed at different levels and on different fields of expertise that all these levels must be taken into account and are necessary to make interoperability possible.

3. Categorization of documents and images

3.1. An inventory of healthcare related documentation

Healthcare professionals write down medical and administrative information about their consultations with the patient, about their findings, plans, conclusions and actions, to properly document the information that is needed and used for optimal care.

In the pre-digital era, the medical record consisted mainly of a written synopsis of the patient's history and physical, notes on each of the consultations, diagnostic study- and surgery reports, together with referral- and discharge letters, laboratory results and other material. These were bound together to a big stack of papers, with tabs dividing the different information elements for quick access to the desired information categories. However, especially with chronic patients, these stacks would become unwieldly, and important information sometimes got lost in the paper mountain. Physicians often used the latest discharge letters as their starting point, together with the latest results from diagnostic studies. This saved time, but sometimes information got lost in the decision-making process, sometimes with unfavorable results.

In the digital world, new possibilities of selecting and presenting the right information at the right time in the right situation have arisen. Grouping together information from different moments, such as with laboratory results, timelines of the patient's history, conclusions from different studies combined in one overview are some of these possibilities. Like in the old paper records, tabs are used for quick access to the right type of information. The possibilities to browse through different records and documents, to look for certain keywords, or create different views on the available information, help the medical professional to work both thoroughly and efficiently.

3.1.1. Explosion of health data

The amount and diversity of healthcare documentation and exchange has grown considerably in the last 15 years, due to factors such as a growing and ageing population needing healthcare, new diagnostic-, imaging- and therapeutic technologies, increased information exchange between healthcare stakeholders, cheaper storage, faster exchange solutions and an increasing number of clinical-, administrative-, financial and legal quality improvements initiatives. The IDC and EMC's joint annual Digital Universe study[3] predicts an annual volume of 40,000 exabytes (or 40 zettabytes) in 2020. A 2012 Ponemon Institute survey found that 30% of the world's data storage resides in the healthcare industry. The amount of stored and exchanged information grows exponentially - in the last 5 years the amount of stored information has grown 10-fold, or 48% annually, due both to an increasing number of documents, documents of larger sizes and richer

[3] See: https://www.emc.com/infographics/digital-universe-2014.htm

types of document. Looking at the graph below, the growth rate is not expected to diminish in the foreseeable future.

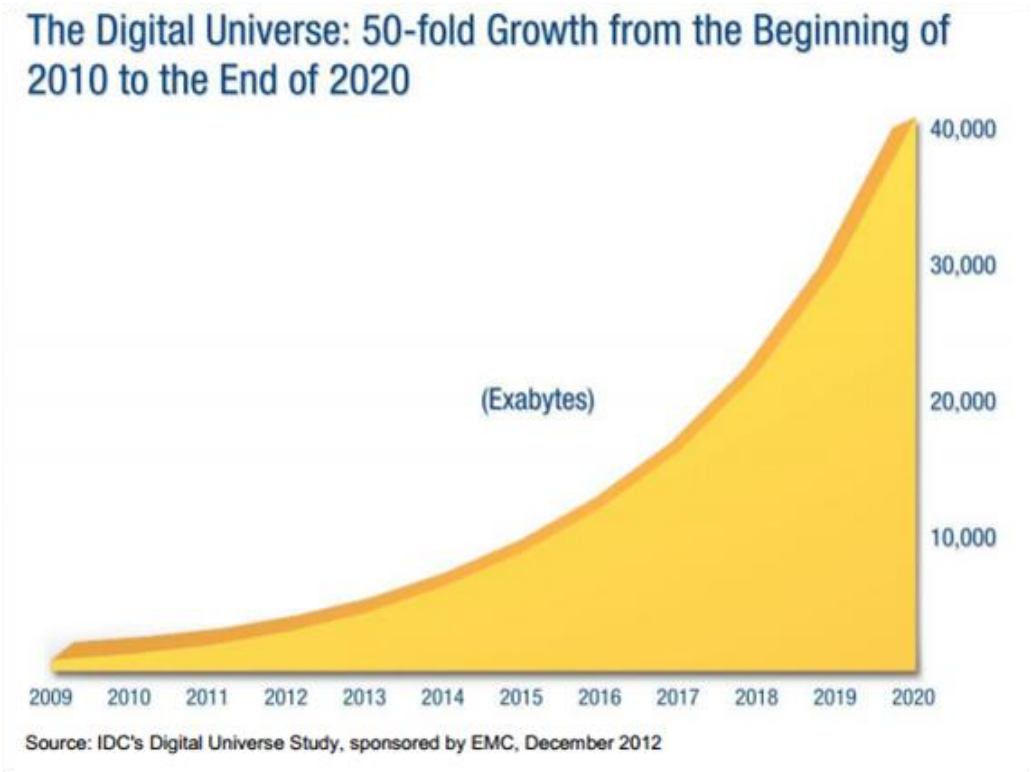


Figure 3 - 40 zettabytes = 40,000,000,000,000,000 terabytes generated and exchanged annually.

Note:

Powers of 1000: kilobyte · megabyte · gigabyte · terabyte · petabyte · exabyte · zettabyte

Average computer: 1 terabyte = 1.000.000 megabyte = 1.000.000.000.000 bytes (10^12).

In 2020, the number of bytes will be larger than the number of stars in the universe.

In daily healthcare practice, quick access to the desired information is important - and as the number of information carriers increase, an efficient and practical categorization system becomes essential. This system must enable an intuitive and deterministic way to get to the right information. In other words, it must make sense from the end user's perspective. And it must be flexible enough so that material can be found in different ways, using different (combinations of) search parameters.

Documents possess different characteristics that can be used for categorization: their purpose, subject, provenance (author, organization), intended recipient, organization, structure, technical and logical format, size, location of creation, care services,, date and time, relation to a workflow, status, confidentiality and so on. These characteristics can be used, often in

combinations, to filter, group and sort the available information. However, the health-related constraints require a patient centric, much more structured and deterministic approach to search (false negative matches are not acceptable). These constraints are not met by typical web search engines.

3.1.2. Many types of documents

There are many document types that can be distinguished. Sometimes these documents are linked to a specific task or workflow, some are specific to a department, a survey, research, diagnostic studies, et cetera. As an illustration, among the material we received from different participating countries, some lists of different image types were more than 4000 items long. A balance must be found between the expression power of the document type and the number of items in the list. When these lists become too long to be traversed quickly by the end-user, these lists must be analyzed to see whether they can be broken down into different factors, to create a set of more manageable lists. Another way of decreasing the number of list elements is to decrease the level of granularity: taxonomy should be created.

There are different ways to categorize documents (the term *document* in this white paper is used for any kind of file or object with information about a patient), for many different purposes. As an example, here is a high level, non-exhaustive overview of the kinds of patient centric documents that are being used in healthcare, from an end-user point of view:

- Medical
  - Referral letters
  - Discharge letters
  - Medication, Intake reports
  - History & physical reports
  - Consultation notes
  - Diagnostic test reports
  - Summaries
  - Surveys/Assessments
  - Multidisciplinary board reports
  - Genomics, bionics, proteomics information
- Workflow related
  - Orders, referrals
  - Diagnostic test requests
  - Appointments
  - Care plans
  - Workflow overviews and timelines
- Registrations
  - Quality assessments
  - Research (prospective and retrospective)
- Administrative
  - Demographics
  - Insurance / financial

    - Legal
    - Consents
- Images
    - Imaging studies (MRI, CT, ultrasound, PET, radiology)
    - Medical photographs
    - Microscopy / pathology studies
- Miscellaneous
    - Charts, tables, graphs, drawings, sound files, PDFs, datasheets, …

If you look at this list, at first glance it seems to make sense. But in practice it is often not so clear under which category a specific document falls. For this, each category also needs a clear definition and scope. Otherwise, ambiguities may lead to different choices of categorization between individuals. thus resulting in inconsistent search results.

As stated in the introduction, the subject of this guideline is to create an overview of all document categories in a logical manner and to provide guiding principles for the metadata attribute elements supporting a categorization of health-related documents that is practical, logical and flexible in parallel. But first, the challenges that surround this topic will be identified and explored.

3.1.3. Purpose of categorization

Categorization of documents can be used for quick access to documents through selection of (a combination of) categories. You can look for date, author, subject, main category, events, leading to the document, document location, and other category attributes. But they can also be used to quickly aggregate certain documents, or to sort them in a certain order. Here is an overview of the different purposes for which document categories are used :

- Filtering: Four types of filtering can be distinguished: primary, secondary, tertiary and technical
    - Primary Filtering: attributes primarily used for selecting documents and sets of documents submitted together for sharing. This filtering may support a narrowly targeted query (looking for a specific or small set of documents) or a broad query intended to select a manageable set of likely relevant documents.
    - Secondary filtering: returned metadata attributes intended to be associated with the documents matching a primary query in order to enable a human (or application) to filter out among the returned candidate entries, the ones that are not relevant and need not be retrieved.
    - Third-level filtering: Once the relevant documents have been retrieved the content may be processed (aggregated, displayed, etc.) and relevant information extracted to enable a human or application to further select or directly access a targeted set of data. This third level is important for the querying user but is not included in the metadata table as metadata are not used for this third-level filtering.
    - Technical filtering: Metadata attributes critical for the operation of the queries, but generally not visible to the clinical user. They are used for integrity verification, performance management, configuration, etc.
- Additional mechanisms

- On folders[4] and groups of documents
- On 'technical' level

- Grouping
  - Categorization attributes can also be used for grouping of documents of a certain type together, for instance letters, notes, summaries et cetera. Grouping is a specific form of filtering based on metadata with the purpose of creating groups of certain types of documents. In many EHR user interfaces, tabs are used to quickly find the right type of document or image, and grouping is used to place the right documents under the right tabs.
- Sorting/Ordering
  - Category attributes have predefined value sets. The captions of these value sets can be shown in alphabetic order, or in another predefined order (for instance, showing the most frequently used options first)
  - Linking of the value set elements to terminologies allows for proprietary descriptions for sorting and finding documents and images
  - End-users should be able to use different kinds of sorting (date, department, subject, functional type, author, domain et cetera).

3.1.4. Categorization challenges

There are many aspects to consider in the quest for a user-friendly, logical, non-overlapping, non-ambiguous and flexible categorization system. And, where applicable, decision trees must be made to prevent ambiguous choices in selecting a category.

To achieve this, the following challenges must be met:

1. **Selection of logical, orthogonal categorization axes**
   If every type of document would be categorized into one category axis, the number of potential entries in that category axis would be as large as the number of different documents types. By breaking up the defining aspects of a document, the number of categories grows, but the lists of potential entries per category become much shorter. First, the number of category axes must be established and clearly defined. Until now, there is no standard that defines which axes should be used for categorization of medical documents and images. This issue is further discussed in chapter 3.2.1.

2. **Definition of value sets supporting each category**
   For each category axis, appropriate terminologies should be defined and managed. Then it is possible to bind these terminologies as value sets to the appropriate attributes.

3. **Linking to standards and terminologies**
   Find out whether code systems from existing standards and / or terminologies can be used to code the different value sets. For all category axis, the category and value sets must be linked to proper code systems, in best case terminologies. To be able to link these terms to

[4] Folders can be used to make a persistent record of a group certain documents that belong together or are in the same category.

definitions that can be translated internationally but also to enable linking the values to any desired personal, local, regional or national descriptions. The terminology code links all these descriptions and makes them useable for different cases or settings.

4. **Unambiguous selection of medical document categories**

As documents may have many purposes, subjects, roles and formats, it is hard to place some documents in one of the defined categories. For example, a discharge letter created in a multidisciplinary setting can be categorized under *discharge letters*, but also under *summaries* or under *multidisciplinary board reports*. Are individual quality reports, such as a PROM or PREM report, a *survey*, a *quality assessment* or a *research registration*? Unless you strictly define and clearly outline the boundaries of these categories, misclassification due to multiple potential alternatives will happen. One of the most important drawbacks of categorization is that people expect to find certain documents under certain categories, and may get rather upset when someone has stored the document under another category than they expected. To promote an unambiguous choice, a decision support system (following a decision tree) may be an option.

5. **Dealing with a multi-community, multi-enterprise, multidisciplinary world**

In healthcare organizations, the provenance of documents can be easily traced by looking at the author of the document and the department where the document was originated. And in a GP or a hospital environment, a *referral letter* is usually written by the GP and directed to a specialist. But in Health Information Exchange (HIE), where information can be found from different healthcare providers, different healthcare institutes, different regions and even different countries, the provenance and context of a document becomes crucial. Linking documents to processes, circumstances and location becomes increasingly important as scale of provenance and the number of documents grow.

6. **Deciding on the level of granularity**

If every type of document would be categorized into one category axis, the number of categories would be as large as the number of documents. But categorization does not mean that every type of document should have its own category. If the level of categorization does not contain too many elements, information can more quickly selected, but a larger number of matching documents may be returned.. Most of the times, it is necessary to use several categories, by breaking up the defining aspects that can be described about a document. The number of category axes grows, but the lists belonging to each category axis becomes much shorter. Such a multi-axis structuring appears powerful, as long as the axis are defined "orthogonally", or that the definition of the axis does not introduce ambiguities (i.e. any document can only be associated with a single point in that multi-axis coordinate system).

7. **Deciding on the number of value set elements per category**

Categories with value set lists that are too long become unpractical to use and error prone. To be practical they may have to be broken down into main- and subcategories. The critical number depends on the intuitive logic of these elements – the less ambiguous the list is, the longer the list can be. User interfaces ergonomic principles recommend that selection list should be rather short (e.g. 5-10 items maximum) and no more than one level of

subcategories be used. This places the ideal value set of categories to select in the 5-50 item range.

8. **Dealing with standardization in a hybrid situation of documents and images**
There are standards for images (DICOM, Radlex) that offer categorization possibilities. And there are standards for the description of documents, such as LOINC and SNOMED-CT. If the categorization model wants to encompass both types of documents (and more), which standard should be chosen? It is possible, desirable, to use two terminology systems in one category but need a rigorous maintenance process.

9. **Dealing with historically grown category lists**
Legacy categories often have grown empirically and historically and often not from a set of clear guiding principles. Also, over time, additions due to specific needs, research, change of governance and personal preferences have often 'polluted' the original setup. These historically grown lists must be analyzed with these factors in mind.

3.2. Organizing and structuring metadata`

Metadata provide the tools for a fast, user-friendly, intuitive, versatile, and flexible way of finding the desired information for the task at hand. Metadata attributes can be organized in a category classification model that defines all aspects that can be attached to a document, and the attributes that describe these aspects.

The complexity and extensiveness of a classification model should be in balance with the complexity and size of the filtering and sorting problem. Metadata are used to group, sort and/or filter documents belonging to just one patient. The adagio from a certain Mr. A. Einstein is applicable here: "make it as simple as possible, but not simpler". In other words, do not use too many metadata elements or excessive levels of detail - they do not improve usability. Only define metadata elements that are necessary for the defined use cases.

As Metadata is use for filtering of patient data for the purpose of exchange between sources of documents and consumers of documents, one critical design principle is that the metadata should be designed so that when used by the source in classifying any shared document, the filtering applied by the consumer of these shared documents shall never filter out relevant documents (No false negative allowed).

3.2.1. Approach

In this chapter, we describe the approach towards organizing and structuring the metadata, to create a functional, logical model for categorization of healthcare related information carriers. In chapter 4 we start looking at the XDS profile, comparing the results of our analysis in chapter 3 and the logical model with the XDS metadata definition, to see whether these results can be used to refine its definition of the XDs metadata.

Inventory and analysis of available document types

To see what types of metadata can be defined for medical documents, we have looked at lists of document and images types provided by healthcare organizations and national/regional healthcare ICT competence centers in different countries. Analysis and comparison of the different lists resulted in a first, high level list of functional categories (see chapter 3.1.2). This list has been completed with document types that may come from other sources, or expected to be added in the near future.

The resulting list gives us an impression of the kind of health and healthcare related documents, images and other files.

Categorization model

The list can be used as starting point for a logical categorization model. What kind of properties can be linked to documents? Agree upon the logical axes for the different categorization axes, and create a model for all the different metadata attributes and value sets.

Link to XDS metadata

After definition of the logical model, we compare this to the XDS metadata attributes and see whether any extra attributes would be desirable. If some elements from the 'ideal' model do not fit in the XDS metadata elements, decide what to do with these elements. Discard them, try to fit them in in the existing metadata set, or produce an IHE Change Proposal.

Links to standards and terminologies

Bind all the XDS metadata elements to terminologies. This allows for better understanding and definition of the meaning of the concepts and values, and makes cross-country and cross-language interoperability possible.

Look whether existing international standards can be used as the basis for each of the attributes and value sets.

Connect to other SDOs

Connect to other standards developing organizations such as HL7, LOINC, DICOM, SNOMED for consultation and cooperation. These organizations should be informed by the categorization. Cooperation and harmonization between these organizations on this subject (among others) would greatly enhance the power of standardization.

Make substantiated choices

All choices for attributes value sets should be discussed and well argued. Substantiated arguments improve transparency, and force the people to look critically at the process of choices.

Furthermore it helps to understand where to place a certain document and thus improves efficiency.

To prevent ambiguity as to what categories must be selected when a document is created (by the XDS Document Source), a clear definition of each concept must be given. Decision trees and guidelines for the proper selection of metadata values are recommended where the selection may be a problem.

Build in flexibility

- The XDS metadata set must be able to contain and describe any type of document in the intended (affinity) domain;
- Allow for national/regional or local extensions of the proposed metadata value sets;
- New agreements regarding exchanges between Affinity Domains should be made at the level of the participating organizations (see 'Connect to other SDOs' above).

Testing

Test interoperability of XDS metadata between XDS Affinity Domains using different scenarios for cross-community information exchange. These tests could become part of a Projectathon exercise.

3.2.2. Categorization model

The axes proposed at the beginning of chapter 3.2.1 follow the ancient Problem Analysis perspectives from Cicero and others (1st century BC):

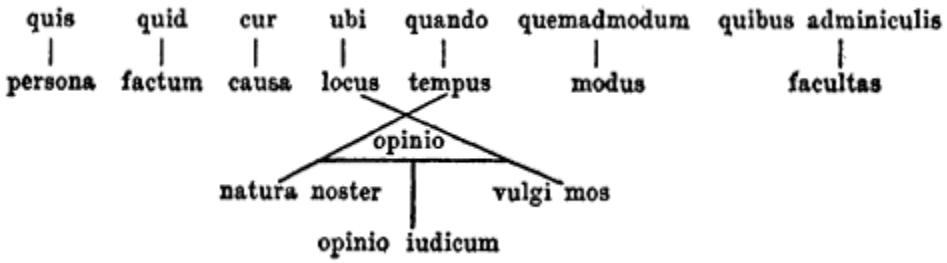


Figure 4 - who, what, when, where, why, in what way, by what means

Categorization of documents can be done by looking at the following parameters:

- Why
  - Purpose of the document (explicitly stated by the creator)
  - Link to order / request / workflow / care pathway / episode (in the context of the creator, it may be never used in the context of other workflows by consumers of the document)

- The event, clinical activity that this document was created for

- Who
  - By whom — author, creating application
  - Under the authority of whom — legal authenticator
  - For whom — intended recipient/ location
  - About whom — patient (patient ID), relatives

- What
  - Document functional type (main- and subtype)
  - Document template / structure
  - Document ID, technical format, MIME type, structure template, extension
  - Document size, version
  - Document language
  - Status (published, replaced, transformed, deprecated), confidentiality, integrity, availability

- When
  - Date and time of creation, last changed, last opened
  - Date and time of storage

- Where
  - domain, organization, location, department, where the document was originally created
  - Storage reference ID where document is stored

- How
  - In what way - method of creation
  - By what means - Modality or application that created the file
  - In the context of what procedure(s)
  - In the context of what workflow event

Another approach is to look at possible attributes for describing the different types of documents and images:

- Medical
  - Function — class and type (hierarchical subcategory of class)
  - Purpose — e.g.: Summary of an encounter, (Whole body MRI) for bone disease diagnosis
  - Organization — organization, specialism, department
  - Persons — related to the document
  - Modality — MRI, ultrasound, CT scan, SPECT et cetera
  - Body part — practical categorization; including laterality

- Event — event where was the document created
- Other — extra tags for selection and filtering

- General / administrative
  - Availability status — active or deprecated
  - Confidentiality — confidentiality tag of the document
  - Integrity — can be checked by a hash code
  - Language — language used in the document
  - Timestamp — date and time of creation, and of storage
- Technical
  - MIME type — document MIME type
  - Format — technical document format (e.g. .pdf, .xls)
  - Size — document (number of bytes)
  - Template type — document template
  - Unique ID — unique document identifier

NOTE: XDS categorization axes
In the XDS metadata definition, the metadata attributes are approached from the following axes: Patient Identity, Provenance, Security & Privacy, Descriptive, Object Lifecycle and Exchange.

3.2.3. Categorization of images and textual documents

A categorization system must be able to categorize any kind of documents, in the medical domain this means mostly image studies and textual and/or structured documents.

This poses a challenge for a categorization model: although some of the attributes can be used for both types, some are specific attributes to either images and to documents.

After studying the file descriptions gathered from study material from the countries participating in this project, we came to the following sets of attributes.

We start with the attributes that are being used for both images and documents:

3.2.3.1. All files

- Administrative
  - Availability — concept, active or deprecated
  - Confidentiality — confidentiality tag of the document
  - Event — in context of which type of event was the document created
  - Language — language used in the document
  - Organization — organization, specialism, department under whose responsibility the file was created
  - Persons — persons involved

- Patient
- Author
- Legal Authenticator

- Technical
  - Format — technical document format (e.g. .pdf, .xls)
  - MIME type — document MIME type
  - Size — document size (nr. of bytes)
  - Template ID — document template ID
  - Timestamp — date and time of
    - Creation
    - Authentication
    - Storage
    - Last viewed
  - Unique ID — unique document identifier

3.2.3.2. Images

Images are mostly categorized by the *procedures* they are used for. A procedure description usually consists of a combination of (some of) the following parameters:

- Modality — e.g.: MRI, ultrasound, CT scan, SPECT
- Technique — procedure specification, e.g. (MRI) angiography.
- Body part — e.g. arm, hand, upper body, neck
- Laterality — e.g. left, right, left and right
- Extra procedure — ,CT and biopsy
- Purpose — e.g.: (MRI) for diagnosis of bone disease
- Qualifiers — connecting qualifying terms like 'using', 'and', 'with'
- Other — e.g.: 'of', '(procedure)'

Below is a list of examples of diagnostic imaging procedures from a list of more than 10,000 (!) different procedures, including their SNOMED-CT codes. Because this list is 'flat' in the sense that all the attributes are in the category (except laterality), in a list such as this it is very hard not to make textual, logical errors, use two different descriptions for the same attribute value (for example: 'echography' and 'ultrasound scan'),

The 'Remarks' column highlights some explanation or inconsistencies, illustrating the problems that arise from such flat, large categorization lists. The color codes reflect the attributes described above.

Description	SCT code	Remarks
Computed tomography of abdomen and pelvis	419394006	Combination of modality

Description	SCT code	Remarks
(procedure)		and two body regions
Computerised tomography of abdomen and pelvis with contrast (procedure)	310111000000101	Naming conflict with Computed tomography
CT of abdomen and pelvis	183881000000104	Naming conflict with Computed tomography
Computed tomography myelogram of lumbar region (procedure)	241592002	Combination of modality, special technique and body region
Computed tomography for measurement of brain volume (procedure)	443601000	Purpose of CT
Endoscopic retrograde cholangiopancreatography (procedure)	386718000	special technique. Body parts are included in the description (bile ducts and pancreas)
Fluoroscopic angiography of left side of heart (procedure)	420136008	Laterality is a separate code, options: left, right and right and left
Magnetic resonance imaging of left forearm (procedure)	241636007 + laterality	+ Laterality code: 7771000
Magnetic resonance imaging of right forearm (procedure)	241636007 + laterality	+ Laterality code: 24028007
Magnetic resonance imaging of left and right forearm (procedure)	241636007 + laterality	+ Laterality code: 51440002
Radionuclide myocardial perfusion study (procedure)	252432008	
Isotope study for gastrointestinal blood loss (procedure)	19252005	Purpose
Ultrasound scan of abdomen and pelvis (procedure)	418394000	
Intravascular ultrasound Doppler imaging of coronary artery using fluoroscopic guidance (procedure)	431747006	IVUS plus fluoroscopy
Echography of kidney (procedure)	306005 + laterality	Naming conflict with ultrasound scan
Diagnostic radiography of abdomen (procedure)	60654006	Inconsistent use of Diagnostic (most

Description	SCT code	Remarks
		procedures are diagnostic)
Plain chest X-ray (procedure)	399208008	Naming conflict of the term 'X-ray' with 'radiography'
X-ray of little finger (procedure)	418515004 + laterality	Body part too small for categorization?
X-ray of fingers (procedure)	418426008 + laterality	Body part still too small for categorization?
Radiography of hand (procedure)	49345004 + laterality	Naming conflict. Body region too large for categorization?
Fluoroscopic venography of lower limb (procedure)	418881009 + laterality	
Fluoroscopic vasography (procedure)	418623007	Conflict with fluoroscopic angiography?
Fluoroscopy of left upper limb (procedure)	419571009 + laterality	Inconsistency: the laterality is included in the procedure
Fluoroscopic antegrade pyelography (procedure)	418462009	Body part: kidney
Fluoroscopic angiography of renal transplant (procedure)	419139005	This code misses laterality!
Fluoroscopic micturating cystourethrography (procedure)	419245009	Body part(s)combined with procedure
Fluoroscopic angiography of carotid artery and insertion of stent (procedure)	418405008	Separate procedure
Magnetic resonance imaging angiography of head (procedure)	417936006	Procedure belonging to the main technique
Mammography (procedure)	71651007	Not clear which modality technique has been used. Body region included in name of procedure.
Prone stereotactic X-ray guided core needle biopsy of breast (procedure)	306381000000106	
Endoscopy of stomach (procedure)	386831001	
Bronchoscopy (procedure)	10847001	Bronchoscopy is a

Description	SCT code	Remarks
		special form of endoscopy
Percutaneous transluminal angioplasty of internal carotid artery using fluoroscopic guidance (procedure)	432039002	Therapeutic procedure using a modality technique
Computed tomography of liver (procedure)	241549007	standard
Computed tomography of liver with contrast (procedure)	429862006	with contrast
Computed tomography perfusion study of liver	448677000	technique related to the modality
Computed tomography and biopsy of liver (procedure)	418749009	with procedure not directly related to the modality
Computed tomography of transplanted liver (procedure)	911771000000107	with specification of body region
Computed tomography of transplanted liver with contrast (procedure)	911781000000109	idem, with contrast
Computed tomography triple phase study of liver (procedure)	438591004	Specialized technique
CT of head	408754009	Inconsistent use of modality naming
Positron emission tomography with computed tomography of liver using yttrium 90 microspheres (procedure)	699586000	'with' instead of 'and' used here to indicate that two modality techniques are used here

After studying this and other lists, some conclusions can be made:

1. Dividing the categorization in more than one category increases consistency and quality, decreases typing errors, make selection of the right categories easier through smaller lists.
2. Almost all the procedures in the English NHS list can be described using the attributes as described above
3. Because each of the attributes has relatively few elements, less codes are needed. Codes can be built up from a combination of the different attributes.
4. The description of one procedure can be built up from the constituting attributes, using combinations. For instance, Positron emission tomography and computed tomography of liver using yttrium 90 microspheres (procedure) can be built up using the schema: <modality> and <modality> of <body part> using < technique>. This formalism is known as precoordination of concepts.

5. The division into these attributes allows for a more consistent categorization, with fewer errors, ambiguities, different names for the same values, and easier maintenance.
6. Not all attributes needed to fully describe an imaging procedure may need to be promoted as document metadata. Likewise, the granularity of a given attribute may be to fine grained, and a coarser grain set of values may prove sufficient, simpler to use and less error prone (e.g. fine grained body part, versus coarser body regions).

3.2.4. Selecting the values for the attributes

For the selection of values that can be linked to the attributes from the previous chapter some basic guiding principles are proposed.

- Definition of the value sets must be substantiated;
- Use existing code systems from standards as much as possible;
- Avoid overlapping or ambiguous value set elements;
- Clearly define each value set element so that it can be distinguished from other elements in the value set list (avoid semantic overlaps);
- Provide clear guidelines for deciding in which category a certain document can be placed;
- Maintain balance between granularity and usability, based upon the selected use cases;
- Bind the metadata attributes and all value set elements to terminologies**.**

3.3. XDS Information Model

Appendix B will explain the relevant elements in detail. However, a high-level overview presenting the main classes and main attributes will be beneficial to understand how they can be used for querying. (The functional separation into registry and repository is left out for simplicity.)

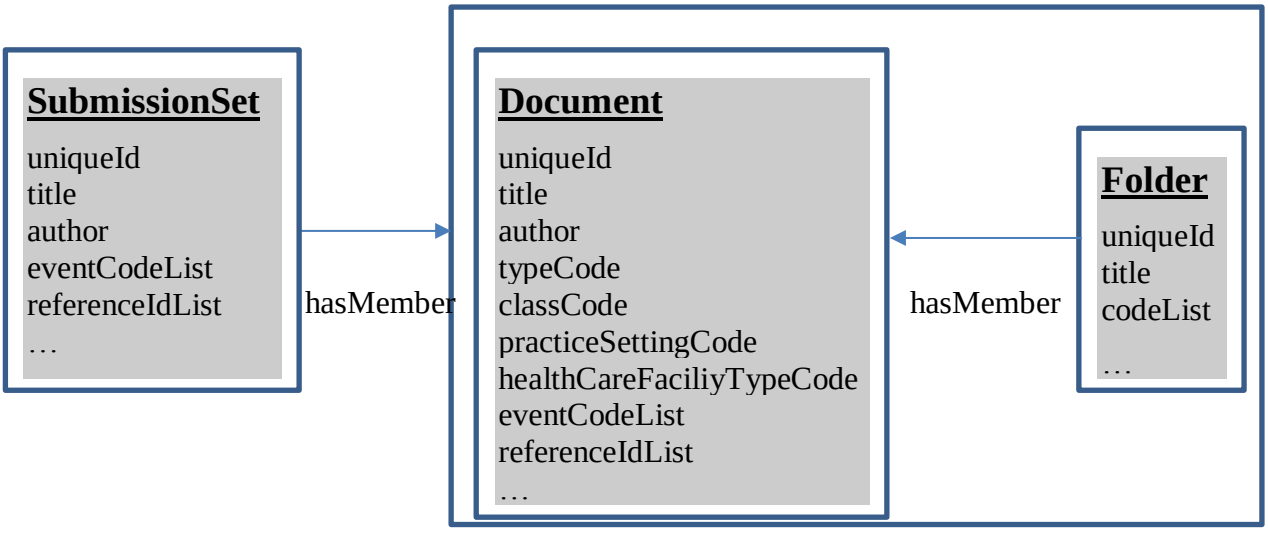


Figure 5 – XDS Information Model

Submission sets are used as a transactional parenthesis around documents and folders which are transferred into a patient record. A title is helpful for human readability, however, it cannot be used for querying because the semantics is unclear, not even counting possible typos. A unique identifier is necessary for identification and maintenance, but not directly helpful for querying because in most cases the internal identifier is unknown and not of any real interest.

Simple single-value attributes like typeCode and classCode are used to query and identify specific documents and are thus bound to the semantics as identified by their attributes. The same argument is valid for practiceSettingCode and healthcareFacilityTypeCode.

The use of eventCodeList and referenceIdList is determined by classCode and practiceSettingCode and therefore not free in its use. Furthermore, each document only allows for a single list of values. Hence, using these values for multiple purposes requires the use of pre-coordinated concepts possibly resulting in complicated constructs that are difficult to query.

An alternative concept available with XDS is the folder. A folder references multiple documents, and each document can be referenced by multiple folders. The XDS folder is therefore not comparable to a Windows folder, where hierarchies are used to establish semantics This way folder are a flat list representing a m:n relationship:

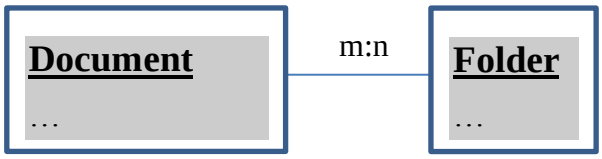


Figure 6 – XDS Document Folder Relationship

In combination with the primary attribute codeList folders can be used for different purposes in parallel:

- administrative documents
- simple grouping
- record-type specific grouping
- security markup
- …

An example of use could be to a set of folders that identify all documents that belong to a specific case record. Another set of folders could be used to identify another set of documents belonging to another case record. One code within the codeList would represent the fact that these folders are used to identify case records in general. A second code would then be used to represent a specific case record, e.g. to distinguish between a cancer case record and a diabetes case record. Further codes would then be used to introduce “folders” within a specific record.

For a query all documents could be retrieved for a specific case record ("retrieve all documents where folder.codeList contains {CASE_RECORD_TYPE, DIABETES}" using the folder structure for visualization ("grouped by folder.codeList displaying folder.title").

4. Harmonizing different initiatives

This section is organized in three parts and provides an overview of the existing use of XDS metadata:

- Analysis of existing (proposals for) solutions
  - Cross-enterprise exchange drives demand for more constraints
  - National initiatives
  - European and global perspective
  - The XDS metadata project
- Code systems
  - SNOMED-CT
  - LOINC
  - DICOM
  - Radlex Playbook
- Discussion

4.1. Analysis of existing (proposals for) solutions

4.1.1. Cross-Enterprise exchanges

It is now widely accepted that for effective and efficient delivery of care to citizens multiple entities must be involved. The scope of involvement is extremely wide:

- Acute Hospitals
- Community Healthcare providers
- Primary Care providers
- Mental Care services
- Private Healthcare providers
- Social Care providers
- Care Home providers
- Third Sector providers (Voluntary/Charity)
- Personal Careers
- Citizens

It is clear that with such a wide variety of organizations involved the language used to describe actions, people, interventions, can be equally diverse. If different organizations use different vocabularies to describe their content then any search based attempts to share data are likely to fail. This can lead to inevitable clinical risks. This has led to implementers of Health Information Exchanges (HIEs) to propose constraints of the allowed descriptors. The goal of search

technologies is to allow as much automated processing as possible. This requires that a coding structure is used to classify the elements and attribute values so they are computer processable.

There are currently a number of National initiatives which have all started from scratch. There is some commonality in the use of relevant part of SNOMED-CT and LOINC as coding structures for certain metadata elements.

4.1.2. National initiatives

There are a number of National Initiatives who are looking to define the metadata which will be used to search for clinical documents:

1. Swiss eHealth. - Website
   Have published a constrained model using XDS Metadata fields. Uses a mixture of National; SNOMED and LOINC codes
2. IHE-USA - Website
   Have published a National extension for XDS as to PCC-DAF (See Appendix B Page 62-67). Uses a mixture of National; SNOMED and LOINC codes
3. Denmark - Website
   Have published a constrained model using XDS Metadata fields. Uses a mixture of National; SNOMED and LOINC codes
4. France - Website
   Have published a constrained model using XDS Metadata fields. Uses largely National codes
5. United Kingdom
   Are developing a universal metadata charter. This will be developed to include an XDS instance. Have created a natural language metadata description and related this to the XDS metadata.
6. Germany – art-decor.org/art-decor
   Has instantiated a task force to create and maintain a solid and reasonable set of value sets for XDS metadata. The first formal ballot passed successfully. This set is formally a strong proposal for implementers.
7. Finland - Website (only in Finnish except for technical spec)
   Being developed as part of national medical imaging repository (XDS, XDS-I).
   Mostly national or nationally adapted coding systems. Metadata model relates to separate Finnish metadata model for CDA R2, see here (only in Finnish).

In the following table, comparison of metadata attributes is presented.

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
Subject (Patient)							Finnish national ID (SSN) Temporary ID (organization specific)		
	Patient ID	National Identifier	NHS Number	Use CDA value			Finnish national ID (SSN) Temporary ID (organization specific) IMMUTABLE	•National ID •Temporary ID (value extracted from CDA header)	
	Source Patient ID	Local Identifier	Local eg Hospital ID	Use CDA value	National Code set		National Code	•Other IDs (ex. local ID) (value extracted from CDA header)	Local ID
Author									
	Institution	Where the person publishing the document works when publishing the information.	National Code of Institution (ODS Code)	Use CDA value	National Code set		Repeating author structure for each author person and for each institutional author: service provider, service producer,	National ID for institution (value extracted from CDA header)	

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
							service unit, custodian		
	Person	The name of the person publishing the information	Name - (If NHS can have ODS code)	Use CDA value	National Code set		Repeating structure for each Person recorded	•National provider ID •Patient ID (for document authored by patients) (value extracted from CDA header)	
	Role	The job title of the person publishing the information	No agreement. Many local values. National Electronic Staff Record (ESR) could be used to provide standardization.	Not used	National Code set - Snomed	National Code set	Mandatory, based on enumerated value list	Not constrained, roles defined locally by users according to local organization (value extracted from CDA header)	
	Speciality	The Clinical/Care speciality of the person publishing the information.	Snomed code using "Under the Care" of values.		National Code set - Snomed	Snomed	Not recorded	National Code set (value extracted from CDA header)	Value set
	Contact	How to contact the Author. Eg Phone number; Email	No agreement.				Not recorded	No constraint	
Event									
	Speciality	The	Snomed code				eventCodeList	National Code	See above

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
		Clinical/Care specialities involved in creating or receiving/consuming the information.	using "Under the Care" of values.				includes specialty or the Finnish ~equivalent "View type", value from National Code set.	set (value extracted from CDA header)	
	Start time	These are used to describe the actual timing of the vent being reported. This might be for example :an operation; a visit; an interview or an assessment.	UTC time	Use CDA values			The date and time of the onset of the service episode or diagnostic study	UTC Time (value extracted from CDA header)	
	End Time		UTC time	Use CDA values			The date and time of the end of the service episode or diagnostic study	UTC Time (value extracted from CDA header)	
	Setting (Practice setting)	These can be used to record details of the clinical and physical setting. Such as for example: Orthopaedic out patient clinic; Community nursing visit in patients home	The clinical settings where Event occurred. Could be Primary Care; Pathology Lab; Radiology Clinic.	Not used	National Code set - Snomed	National Code set	National Code set	National Code set (value extracted from CDA header)	based on guidance from Bundesärztekammer and KBV
	Organization (Facility Type)		Organization Type: Could be Private Hospital; Acute Hospital.	DK IHE value	National Code set - Snomed		Finnish code set for patient registry	National Code set (value extracted from CDA header)	based on guidance from Bundesärztekammer and KBV
	Event type.	This list of	Use Snomed	Use CDA values			Modality (as per	• LOINC	Not used

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
	Clinical acts undertaken	codes represents the main clinical acts documented for the patient contact. Should include specialities of ALL clincial setting includinf recipients	codes.				XDS-I), Study code (National Code set), Anatomic region and specifier (National Code set), Specialty/View type (National Code set), CDA template Id	• ICD 10 • National Code Set (value extracted from CDA header)	
Document									
	Class	Class and type provide a hierarchical definition of the type of information set that is available. Type is the fine grained level and can be part of a single document class, class being used to create a coarser classification. Every document must have a type to be available for sharing.	IHE requires a 1:Many relationship for Class:Type. This cannot be supported by Snomed. Plan is to use Snomed and allow Many:Many nationally but constrain in local Afffinity Domain	DK IHE classCodes	National Code set	National Code set	National Code set	National Code set	Independent (orthogonal) axes for typeCode and classCode based on experience when examining long-lasting archives.
	Type		An Agreed set of typeCodes is under review for approval buyt the Royal Colledges (AMRoCo) via the PRSB.	DK IHE healthCare-Facility-TypeCodes	National Code set -LOINC	National Code set -LOINC	National Code set – redundant to Class (code)	• LOINC • National Code set (value extracted from CDA header)	

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
	Availability / Status	This can be used to track the history of a document or information set indicating whether a particular set is the latest collection or has been deprecated in favor of more recent information.		DK IHE value			Approved when entry received to registry, old versions set to deprecated automatically.	National Code Set extending XDS value set	
	Creation Time	This is the date/time when the record was created. This should always be in UTC coding to avoid confusion across geographies and calendar changes.	UTC time.	Use CDA value			Date and time of creation of XDS manifest or derived from CDA effectiveTime.	UTC Time (value extracted from CDA header)	
	Title	The document title.	If HL7 document (CDA or CDA on FHIR) use the document title.				Free text description. When entry based on CDA R2, use CDA R2 title	No constrain (value extracted from CDA header)	
	Description	Where appropriate a brief description of the content	Simple text value				Not recorded	No constrain	

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
		can be provided							
	Confidentiality	This should be used to set a broad control of access .	No agreement. Currently basic XDS set	Use CDA value	National Code set	National Code set	National Code set	National Code Set extending HL7 value set	Basic XDS set
Folder		Folders are used to tag documents						Not used	Use of base code set allows for all kind of tagging[5], eg. eCR, semantic grouping.
	codeList							Not used	
Provenance									
	Authenticator	Understanding the provenance of information can be crucial in attributing its utility. Where appropriate this can be related to and individual and linked to the organization making the information available.	Set by publishing organization				Not recorded	(value extracted from CDA header)	
	Publishing Organization		National Code of Institution (ODS Code)				Not recorded (semantically ~authorInstitution: custodian organization)	(value extracted from CDA header)	
Exchange/ Technical									

[5] A new analysis has confirmed the German approach to use folder structures because event code list is too restrictive when combining information for different purposes.

Attribute Type	Attribute	Description	UK	Denmark	Switzer-land	USA	Finland	France	Germany
	Mime Type	The technical description of the type of information	MIME type of the document in the Repository. Shall have a single value.		National Code set		Enumerated values for specific document types.	Standard mime types	Standard mime types with additions
	Entry ID	In order that the correct information is retrieved from a repository linked to the registry which hold the metadata there must be a unique and validated link.	System generated	System generated			System generated	System generated	System generated
	Repository ID		System generated	System generated			System generated	System generated	System generated
	Format	The formatCode element describes the format of the document with respect to achieve the agreed level of semantic interoperability.	No agreement. Likely to move to FHIR resource description.	DK IHE value	National Code set		As per XDS-I or CDA R2 templateId. For semantic interoperability.	National Code Set extending IHE format for some CDA documents	IHE format codes with national additions

4.1.3. European and global perspective

Whilst many of the metadata definition initiatives share common use of international coding standards such as SNOMED-CT and LOINC they all require significant use of local coding structures. This is an inevitable consequence of the need to provide pragmatic approaches which can be used with documents and information resources already in place. Within the underlying ethos of XDS is a requirement that the document created is immutable. This will require that metadata coding structures are able to maintain compatibility with the underlying document. This is particularly important for structured documents such as HL7 CDA level 3. Reference to the coding applied at the time of registration will be required.

4.1.4. The XDS metadata project

The IHE-Europe metadata project is seeking to bring together many of the National initiatives to harmonize the use of International coding standards such as SNOMED-CT and LOINC and provide a model to include many of the metadata attributes currently coded using local coding structures.

There are however significant limitations to this approach as not all international coding systems are able to describe the nature of the relationships within the XDS metadata. For example XDS uses a mono-hierarchy (1:many) to describe the levels of document naming Class:Type (see the approach of US DAF appendix A and France). SNOMED-CT is built on a poly-hierarchy so is unable to support the Class:Type relationship used by XDS. This is despite the fact that Type is very well described by the SNOMED-CT concept of a "Record Artefact".

It is important to utilize internationally recognized coding standards wherever possible if international cross border sharing of clinical information is to be supported. It does need to be recognized, however, that more than 90% of sharing needs are within a single national environment such that projects should not be restrained because of a need to define and use local coding approaches.

Key areas where XDS Metadata has not been systematically applied are:

Clinical Events: The eventCode attribute provides an opportunity to record key elements of the "Clinical Encounter" being reported. This is the CDA interpretation but is not the XDS semantics. Currently there has been little structured use of this. Both LOINC and SNOMED-CT are able to support structured and codified content and this needs to be a focus developed form this project. Before any LOINC and SNOMED codes can be used, the semantic of this event-code list need to be well understood.. Use of SNOMED is growing internationally in the field of clinical documentation and reporting concepts. But SNOMED has not been designed with Metadata use cases in mind.

4.1.5. Code Systems

4.1.5.1. SNOMED-CT

SNOMED-CT or SNOMED Clinical Terms is a systematically organized computer processable collection of medical terms providing codes, terms, synonyms and definitions used in clinical documentation and reporting. SNOMED CT is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world. The primary purpose of SNOMED-CT is to encode the meanings that are used in health information and to support the effective clinical recording of data with the aim of improving patient care. SNOMED-CT provides the core general terminology for electronic health records. SNOMED-CT comprehensive coverage includes: clinical findings, symptoms, diagnoses, procedures, body structures, organisms and other etiologies, substances, pharmaceuticals, devices and specimens.

SNOMED-CT is maintained and distributed by SNOMED International, an international non-profit standards development organization, located in London, UK. SNOMED International is the trading name of the International Health Terminology Standards Development Organization (IHTSDO), established in 2007.

SNOMED-CT supports consistent information interchange and is fundamental to an interoperable electronic health record. It provides a consistent means to index, store, retrieve, and aggregate clinical data across specialties and sites of care. It also helps in organizing the content of electronic health records systems by reducing the variability in the way data are captured, encoded and used for clinical care of patients and research. https://en.wikipedia.org/wiki/SNOMED_CT - cite_note-4.

SNOMED-CT can be used to directly record clinical details of individuals in electronic patient records. It also provides the user with a number of linkages to clinical care pathways, shared care plans and other knowledge resources, in order to facilitate informed decision-making, and to support long-term patient care. The availability of free automatic coding tools and services, which can return a ranked list of SNOMED-CT descriptors to encode any clinical report, could help healthcare professionals to navigate the terminology.

SNOMED-CT is a terminology that can cross-map to other international standards and classifications. Specific language editions are available which augment the international edition and can contain language translations, as well as additional national terms. For example, SNOMED-CT-AU, released in December 2009 in Australia, is based on the international version of SNOMED CT, but encompasses words and ideas that are clinically and technically unique to Australia.

4.1.5.2. LOINC

Logical Observation Identifiers Names and Codes (LOINC) is a database and universal standard for identifying medical laboratory observations. First developed in 1994, it was created and is maintained by the Regenstrief Institute, a US nonprofit medical research organization. LOINC was created in response to the demand for an electronic database for clinical care and management and is publicly available at no cost.

It is endorsed by the American Clinical Laboratory Association and the College of American Pathologists. Since its inception, the database has expanded to include not just medical laboratory code names but also nursing diagnosis, nursing interventions, outcomes classification, and patient care data sets.

LOINC applies universal code names and identifiers to medical terminology related to electronic health records. The purpose is to assist in the electronic exchange and gathering of clinical results (such as laboratory tests, clinical observations, outcomes management and research). LOINC has two main parts: laboratory LOINC and clinical LOINC. Clinical LOINC contains a subdomain of Document Ontology which captures types of clinical reports and documents.[1][2]

Several standards, such as **IHE** or **HL7**, use LOINC to electronically transfer results from different reporting systems to the appropriate healthcare networks. However, the health information enclosed is identified by a multiplicity of code values that may vary according to the entity producing those results. This has obvious disadvantages to the healthcare network that may need to adopt different codes to access and manage information coming from multiple sources. Managed care providers, for example, often have negotiated contracts that reimburse episodes of care and unique coding to trigger automated claim payment. Mapping each entity-specific code to its corresponding universal code can represent a significant investment of both human and financial capital.

A universal code system will enable facilities and departments across the world to receive and send results from their areas for comparison and consultation and may contribute toward a larger public health initiative of improving clinical outcomes and quality of care.

LOINC is one of the standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information. In 1999, it was identified by the HL7 Standards Development Organization as a preferred code set for laboratory test names in transactions between health care facilities, laboratories, laboratory testing devices, and public health authorities.[3]

4.1.5.3. DICOM

Digital Imaging and Communications in Medicine (**DICOM**) is a standard for handling, storing, printing, and transmitting information in medical imaging. It includes a file format definition and a network communications protocol. The communication protocol is an application protocol that uses TCP/IP to communicate between systems. DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format. The National Electrical Manufacturers Association (NEMA) holds the copyright to this standard.[1] It was developed by the DICOM Standards Committee, whose members[2] are also partly members of NEMA.[3]

DICOM enables the integration of medical imaging devices – like scanners, servers, workstations, printers, network hardware, and picture archiving and communication systems (PACS) – from multiple manufacturers. The different devices come with DICOM Conformance Statements which clearly state which DICOM classes they support. DICOM has been widely adopted by hospitals and is making inroads in smaller applications like dentists' and doctors' offices.

DICOM is known as NEMA standard PS3, and as ISO standard 12052:2006 "Health informatics -- Digital imaging and communication in medicine (DICOM) including workflow and data management".

4.1.5.4. Radlex Playbook

RadLex Playbook is a project of the Radiological Society of North America (RSNA), and constitutes a portion of the RadLex ontology. Playbook aims to provide a standard system for naming radiology procedures, based on the elements which define an imaging exam such as modality and body part. By providing standard names and codes for radiologic studies, Playbook is intended to facilitate a variety of operational and quality improvement efforts, including workflow optimization, charge master management, radiation dose tracking, enterprise integration and image exchange.

Historically, departments and institutions have adopted or developed idiosyncratic codes and names for radiology exams, which may have been internally generated or vendor-dependent. This approach led to limited exam interoperability. At its core, Playbook is a set of standardized codes and names which may be used in place of (or alongside) historical codes, in systems which track imaging procedures. Such systems include PACS, reporting applications, RIS, physician order entry systems and electronic medical records.

Playbook currently addresses imaging exams at the level of radiology orderables (i.e. studies which a referring physician may request through an order entry system). Depending on institutional practice, such orderables may be less specific than the exams actually performed. For example, “CT abdomen/pelvis with contrast” is less specific than “CT abdomen/pelvis with contrast, liver protocol.”

Access RadLex Playbook on the web at http://playbook.radlex.org where a graphical search interface is available, as well as a set of downloadable spreadsheets.

4.2. Discussion

It is clear that harmonization of metadata will improve the ability to successfully exchange clinical information, however, if local requirements are not acknowledged and supported such initiative is destined to fail.

Where IHE can assist the process is to support the establishment of an agreed approach to metadata and an agreed definition of the use that coded or text descriptions are used. IHE-UK claim to have defined an approach that would be internationally applicable:

IHE-UK Universal Metadata Charter

An integrated approach to information is a key enabler for delivering effective patient care, whilst improving efficiency. A wide range of information is typically available in diverse organizations and systems. Integration therefore depends on location and selection of that information.

There are a number of current, separate projects which address availability of registries or indexes of information. However there has been little progress in identifying standard characteristics – metadata – to ensure that clinicians can support patients in a consistent way.

IHE-UK proposes that signatories to this charter support the development, use and availability of a common clinical document metadata standard.

Metadata consists of a set of common attributes with agreed semantics and scope according to the following guidance:

- Each attribute will identify its source;
- Where an international coding system is available and suitable this should be given precedence;
- Where international coding is not available or suitable, a national coding system should be used;
- Patients are identified by the applicable national standard;
- Each attribute will define the scope within which it is expected to be used and understood.

5. IHE XDS.b and XDS.b-I metadata

The main source of information and a substantive design concerning Metadata can be found in the IHE IT Infrastructure Technical Framework, Volume 3. According to this document, *"Metadata encodes the properties of Documents, the environments they come from, circumstances of their submission, terms to reference in queries, and grouping with other Documents".*

5.1. Metadata as classification tool

XDS metadata provides slots to record the context of the document: who, when, where, what… IHE classifies the metadata in 6 non-exclusives categories with different objectives:

- Patient Identity: who is the subject of the document
- Provenance: where does the document comes from, in human and technical terms
- Security & Privacy: ensures conformance to different security and privacy regulation
- Descriptive: clinical value of the document, finding a balance between too much information which duplicate the document or disclose information, and too few information which does not allow the inclusion of the document in clinical or reporting workflows
- Object Lifecycle: informs on the states and lifecycle of the document
- Exchange: technical information to allow push and pull transfers of documents

Each metadata attribute can serve one or several purposes, for instance the patient identifier can be seen as belonging to the categories Patient Identity (who is the subject of the document), Provenance (the patient can be the author of the document) and Security & Privacy (manage access rights of the patient).

The Descriptive category, being the most versatile to adapt to different use-cases, is probably also the one giving the implementer most freedom to define the list of codes. The Provenance category may also show differences between implementations.

Appendix B describes all metadata (and their possible use for filtering, see also next paragraph), while appendix A presents a more detailed explanation on metadata which may be particularly useful for classification.

5.2. Metadata as Query and Filtering tool

XDS repositories can contain thousands of documents. Thus, it is mandatory for XDS infrastructures to propose mechanisms to retrieve the needed information for a specific context (e.g. all prescriptions made during the last year by a general practitioner). XDS metadata are designed for this purpose. Appendix B suggests a query and filtering framework using XDS metadata. We shall also point out that XDS does not make any assumption about who chose the query criteria and how it was entered. Obviously, software can propose drop-down lists, date filters and such fields to allow the user to define his own query. Additionally, the software can bring some sort of IA to automatically propose and trigger queries (e.g. displaying to the GP the last 5 prescriptions when writing a new one) or even support workflows and Clinical Decision Support systems (e.g. the CPOE system from the hospital detects potential risk on a patient and triggers targeted queries to a regional XDS infrastructure to complement the information and refine the protocol).

5.3. Metadata as an end-to-end tool

Metadata are not a classification or a query tool. They are two sides of the same coin and closely related. There is no separation of concern among metadata, for example the practiceSettingCode is the same concept whether it's considered as a classification term or as a query term. As such, implementers shall keep both points of view in mind when defining the value sets or considering the usage. The whole document lifecycle shall thus be analyzed: who will send document(s), who will want to search for it/them, which clinical or public health questions will the XDS infrastructure support?

A poor classification leads to poor query capabilities, but a good classification without smart query possibilities may lead to lower-than-expected adoption of the system or even generates additional workload. The opposite is also applicable: even with the smarter tool offering the right filters at the right time, if the documents are not classified with coherent and adequate metadata, the desired information may not be found.

Metadata are at the heart of XDS registries and built upon those dual and complementary facets, usages and refinements can emerge. The next chapter will detail how the generic definition of metadata can be specialized to bring value in specific domains like Radiology. Another application of metadata is with the DSUB (Document Metadata Subscription) profile which relies on metadata to define trigger filters when subscribing to notifications. As an illustration, the following metadata are used:

- Document Entry
    - patientID

- classCode
- typeCode
- referenceIdList
- practiceSettingCode
- healthcareFacilityTypeCode
- eventCodeList
- confidentialityCode
- formatCode
- author

- SubmissionSet
  - patientId
  - sourceId
  - author
  - intendedRecipient

5.4. Domain specialization

When the XDS Profile was designed, it appeared clearly that certain domains of health information may require the use of more than the "generic" metadata attributes defined in the XDS Profile. This lead to the definition of a set of metadata attributes that could be added in a structured way:

- Each additional attribute concept is defined by a unique coded value, that qualifies the metadata value covered and specified with an appropriate value set;
- Although not explicitly stated such metadata extensions, are considered as means to further refine a query. If certain documents entries in an XDS Registry do not have such values, the filtering should be less discriminatory;
- Two such metadata attributes have been defined:
  - eventCodeList. This extension concept is explained in the XDS Profile: "When defining the value sets for eventCodes, they should not conflict with the values inherent in the classCode, practiceSettingCode or typeCode as such a conflict would create an ambiguous situation.";
  - referenceIdList. Enables the linkage of a document to an identifier of a workflow where this document has been created;
- It was decided to convey these in an attribute called eventCodeList, because some of these attributes could be related to the clinical event that produces these additional metadata. This where the XDS definition has departed from the CDA definition, although these attributes have the same name.

A number of IHE profiles specifying profiles for the content of the documents being exchanged have chosen to require their extensions, in order to ensure a consistent use their specialized metadata attributes placed in eventCodeList. Two examples are:

- In the case of cross-enterprise workflows (XDW profile),:

the eventCodeList is required to contain: “For a Workflow Document, one code of this list shall be used to define the overall status of the workflow. This code shall have one of the following two values:

- code: urn:ihe:iti:xdw:2011:eventCode:open codingScheme: 1.3.6.1.4.1.19376.1.2.3
- code: urn:ihe:iti:xdw:2011:eventCode:closed codingScheme: 1.3.6.1.4.1.19376.1.2.3 (See Section 5.4.5.7.)”

the referenceIdList is required to contain: “the workflow identifier. Only a single value shall be sent in this list. Only the CXi.1 and CXi.5 components shall be used:

- CXi.1 shall contain same value as XDW.WorkflowDocument.workflowInstanceId
- CXi.5 shall contain urn:ihe:iti:xdw:2013:workflowInstanceId.

- In the case of cross-enterprise sharing of imaging information (XDS-I profile):

  The eventCodeList metadata attribute shall contain:

  a code from the DICOM Content Mapping Resource (DICOM PS3.16) Context Group CID 29 for each distinct acquisition modality with which images were acquired in the study.

  code(s) from the DICOM Content Mapping Resource (DICOM PS3.16) Context Group CID 4 for each anatomic region in the study.

The referenceIdList metadata attribute shall be populated by the XDS-I Imaging Document Source with the Accession Number and assigning authority of the Order Filler for each Order associated with the Imaging Document, if the Accession Number is known.

6. Use of metadata

6.1. Use if metadata by IHE

The profiles using the metadata in the IT Infrastructure domain are the following:

Profiles defining metadata:

- XDS.b Cross-Enterprise Document Sharing
- XDR Cross-Enterprise Document Reliable Interchange
- XCA Cross-Community Access
- XDM Cross-Enterprise Document Media Interchange
- XDS-I.b Cross-enterprise Document Sharing for Imaging
- XCA Cross-Community Access
- XCA-I Cross-Community Access for Imaging
- XDR-I Cross-Enterprise Document Reliable Interchange of Images
- XCF Cross-Community Fetch
- MPQ Multi-Patient Queries
- DSUB Document Metadata Subscription

- DSG Document Digital Signature
- XDS Metadata Update

Profiles defining content and values sets

Although the concept of metadata finds its root in the IT Infrastructure domain, a lot of other domains and profiles rely on metadata. The following list provides non-exhaustive examples of profiles and how they relate to metadata.

- IT Infrastructure (ITI)
  - XDW (Cross-Enterprise Document Workflow Content Profile)
    - See examples above.
    - FormatCode
  - BPPC (Basic Patient Privacy Consents)
    - classCode
    - eventCodelist
    - formatCode

- Pathology and Laboratory Medicine (PaLM)
  - APSR (Anatomic Pathology Structured Reports)
    - Restricts the formatCode to 21 possibilities representing different organs
    - XDSDocumentEntry.eventCodeList precise the usage: index anatomic pathology reports by reportable conditions (e.g. certain types of tumors...)
    - Gives freedom to the affinity domain: *"Metadata values in an XDSSubmissionSet with names identical to those in the XDSDocumentEntry may be inherited from XDSDocumentEntry metadata, but this is left to affinity domain policy and/or application configuration."*
  - XD-LAB (Sharing Laboratory Reports)
    - Precise the usage of XDSDocumentEntry.formatCode
    - Gives indication on XDSDocumentEntry.eventCodeList but only for non-human subjects and reportable conditions, the rest is left to the Affinity Domain
- Cardiology
  - CRC (Cath Report Content), EPRC-IE (Electrophysiology Implant/Explant Report Content)
    - Restricts the formatCode to 1 possibility
- Dental
  - SEDI (Secure Exchange of Dental Information)
    - Defines cardinality of certain DocumentEntry Metadata and precise certain usage, for example eventCodeList
- Eye Care
  - EC-Summary (Eye Care Summary Record)
    - Restricts the formatCode, practiceSettingCode, typeCode (28691-5), …

- Patient Care Coordination
  - EDR (Cross Enterprise Sharing of Medical Summaries Integration Profile), XDS-MS (Cross-Enterprise Sharing of Medical Summaries), EDR (Emergency Department Referral)
    - Provides extensive information on the generation of the DocumentEntry metadata from the CDA document, but gives more freedom for the SubmissionSet
- Pharmacy
  - MTP (Medication Treatment Plan), PML (Pharmacy Medication List)
    - Restricts the formatCode to 1 possibility
- Quality, Research and Public Health
  - FP (Family Planning)
    - Restricts the formatCode to 1 possibility
- Radiology
  - XDS-I.b (Cross-enterprise Document Sharing for Imaging)
    - Provides extensive information and list of values for DocumentEntry Metadata
    - Gives freedom to the Affinity Domain: "The coding system of the Radiology Imaging Requested Procedure Code will be designated by the XDS Affinity Domain and shared by all Imaging Document Sources in the XDS Affinity Domain. *(In other words: depends on the Affinity Domain)"*

Even if some of those profiles are still Trial Implementation, we can see that the topic of metadata is a broad one impacting several domains and use cases.

6.2. Cross-domain interoperability

From the previous list, we recognize as well that important metadata are left to the discretion of the Affinity Domain, which is problematic for cross-domain (cross-community) interoperability. For instance, the classCode, which represents the type of document, is not strictly limited: *"Valid values for classCode attribute are specified by the policies of the creating entity […] about 10 to 100 entries […] XDS specifies that the XDS Affinity Domain will establish this list"*.

As an example, let's imagine the following situation. A national ePrescription project need detailed granularity about the type of prescription: radiology exam, drugs, devices, treatment… On another side, a regional public health repository needs to classify those prescription only between their source (hospital or general practitioner). The communication between those two infrastructures will be complicated as the mapping through the classCode and typeCode may not be possible and more detailed decision rules will be needed.

The following list provides some categories of difficulties regarding metadata when establishing cross-domain communication:

- The granularity of the classification is different (classification of documents, profession and specialty of the author, practice setting…). In this situation, a unidirectional mapping may be possible (from the more detailed to the more generic) but the other way round is not possible. This may be the case for example for overlapping codification or specialization (ex. pediatry and neuropediatry settings)

- The concern is different: some actors prefers a medical classification when other actors prefers an administrative classification
- The patient is involved: when an infrastructure involves the patient, it may be required to use a more accessible language which impacts on the metadata
- The terminology used is different: different affinity domains may choose different source terminologies to use in the metadata (e.g. ICD vs SNOMED)
- The use of local coding: in some cases, the source terminology may not be complete for a specific need and a local code may be defined

Note that there is a possibility to update metadata once the document is submitted (IT Infrastructure Technical Framework Supplement "XDS Metadata Update", currently in Trial Implementation). However, this is only possible within an Affinity Domain.

Having reviewed the metadata and the possible blockers for cross-domain interoperability, the following list presents the main metadata where Affinity Domains can define values and which alignement for for cross-domain exchangeis critical.

- DocumentEntry
  - authorRole
  - authorSpeciality
  - classCode
  - confidentiality Code
  - eventCodeList
  - healthcareFacilityTypeCode
  - practiceSettingCode
  - referenceIdList
  - typeCode
- Folder (if used)
  - codeList
- SubmissionSet
  - contentTypeCode

6.3. XDS metadata and interoperability

In the IHE XDS profile, secure, reliable and interoperable exchange of medical images and documents between different XDS infrastructures and Affinity Domains is not described. In the Netherlands there are more than 10 regional organizations that use XDS infrastructures for sharing images, reports, letters, summaries, assignments, workflows and other patient-related information between healthcare organizations within these XDS infrastructure. But since patients have a tendency not to stay within regional boundaries, there is a real need for the exchange of medical information between different XDS infrastructures. IHE has defined a profile for the connection and seamless accessibility of information from other XDS infrastructures, Cross-Community Access (XCA). However, more agreements have to be laid down, at all levels of interoperability, to enable true interoperability between these different XDS Affinity Domains

guarantee for example a comparable degree of security, the uniform approach towards the recording of patient consent, compatible metadata, reliable infrastructures and controllable quality of information transfer in each region.

6.4. Linking the generic categorization model to XDS metadata set

In the following table, the generic attributes set from chapter 3.2.2 is mapped to the IHE XDS metadata elements:

Axis	Generic attribute	XDS metadata attribute
Why	Purpose of the document	<purpose> folderCodeList
	Link to order / request / workflow / care pathway / episode	referenceIdList
	The event, clinical activity that this document was created for	eventCodeList folderCodeList
Who	By whom	author
	Under the authority of whom	legalAuthenticator
	For whom	<intended recipient>
	About whom	patientId
		sourcePatientId
What	Document class	classCode
	Document functional main type	typeCode
	Document functional sub type	title
	Document language	languageCode
	Document ID	entryUUID
	Document original ID	uniqueId
	Document unique filename	URI

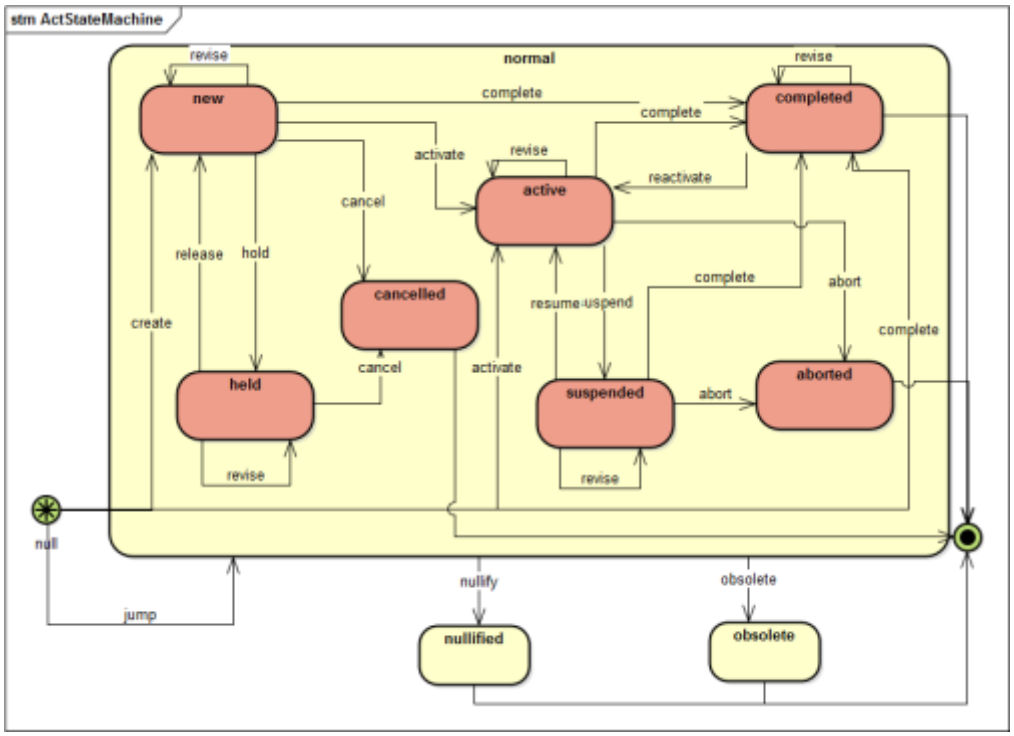
	Document MIME type	mimeType
	Document structure template	formatCode
	Document size	size
	Document version	<document version>
	Document lifecycle status	<document lifecycle status>
	Document availability	availabilityStatus
	Document confidentiality	confidentialityCode
	Document integrity	hash
	Comments for this document	comments
When	Date and time of creation	creationTime
	Start time of procedure / event	serviceStartTime
	End time of procedure / event	serviceStopTime
	Date and time of storage in this domain	<date and time of storage in this domain>
	Date and time of last opened / viewed	<when was this document last viewed>
Where	Domain where the document was originally created	homeCommunityId
	Organization where the document was originally created	authorInstitution
	Organization type of the location where the document was created	healthcareFacilityTypeCode
	Department type where the document was created	authorInstitution
	Specialty that created the document	practiceSettingCode authorSpecialty
	Storage reference ID of this document	repositoryUniqueId
How	In what way - method of creation	objectType
	Compliance to the attribute set requirements	limitedMetadata

An Excel spreadsheet with more information is available in the document set belonging to this white paper.

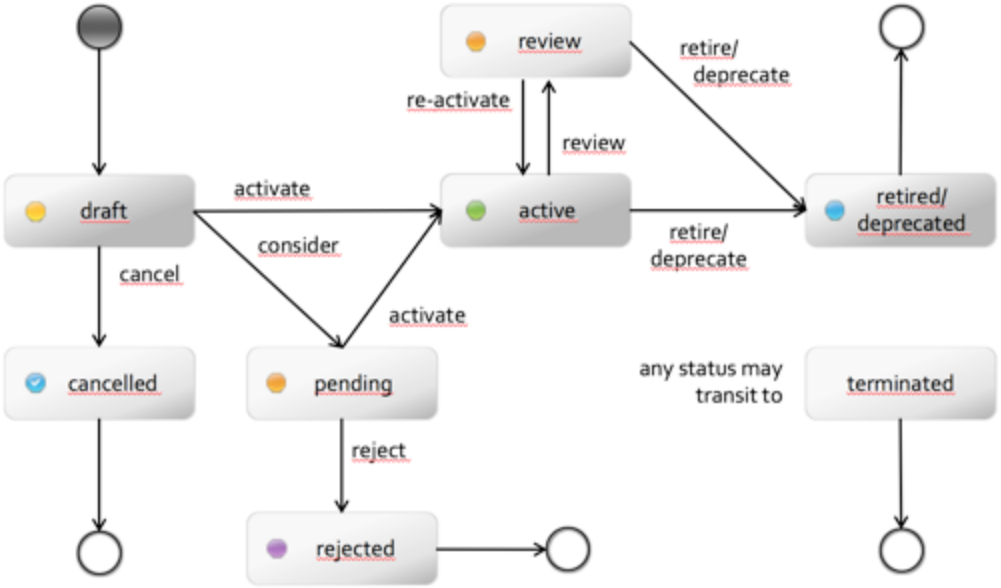
Some remarks on the attributes that are not included in the XDS metadata definition (these terms are in chevrons):

<purpose>	why was this document created? Link to the Problem?
<intended recipient>	who was the original intended recipient of this document (if any)?
<document version>	what is the version number of this document?
<document lifecycle status>	in what state is the document?

There are different document state schemas. The HL7 RIM model has the following states for and Act:



Another model has been devised for the ART-DECOR ecosystem:



There are other status definitions as well, but first a decision has to be made whether this attribute is useful in the XDS domain and what statuses should be allowed for this attribute.

<date and time of storage in this domain>	When was this document stored in this domain?
<when was this document last viewed>	When was this document last accessed?

7. Appendix A. XDS metadata-relevant elements for functional classification

The XDS metadata elements which we suggest to be used to classify DocumentEntry are described below. The texts are taken from the IHE IT Infrastructure Technical Framework, Volume 3 Rev. 13.0 and IHE Radiology Technical Framework, Volume 3 Rev. 15.1. Some descriptions have been summarized for clarity reasons, please consult the corresponding text for all details.

All metadata attributes:

Attribute	Description		Source	Section
author	The humans and/or machines that authored the document. This attribute contains the sub-attributes: authorInstitution, authorPerson, authorRole, authorSpecialty and authorTelecommunication.	See Section 4.2.3.2.1	ebRIM Classification	4.2.3.2.1
availabilityStatus	The lifecycle status of the DocumentEntry	Predefined URN	XML attribute	4.2.3.2.2
classCode	The code specifying the high-level use classification of the document type (e.g., Report, Summary, Images, Treatment Plan, Patient Preferences, Workflow).	Code	ebRIM Classification	4.2.3.2.3
comments	Comments associated with the document.	String	ebRIM Description	4.2.3.2.4
Confidentiality-Code	The code specifying the level of confidentiality of the document.	Code	ebRIM Classification	4.2.3.2.5
creationTime	The time the author created the document.	DTM	ebRIM Slot	4.2.3.2.6
entryUUID	A globally unique identifier used to identify the entry.	UUID	XML attribute	4.2.3.2.7
eventCodeList	This list of codes represents the main clinical acts, such as a colonoscopy or an appendectomy, being documented.	Code	ebRIM Classification	4.2.3.2.8
folderCodeList	A list of codes used to semantically identify a set of documents for a specific purpose.	Code	ebRIM Classification	
formatCode	The code specifying the detailed technical format of the document.	Code	ebRIM Classification	4.2.3.2.9

hash	The hash of the contents of the document.	SHA1 hash	ebRIM Slot	4.2.3.2.10
healthcareFacility TypeCode	This code represents the type of organizational setting of the clinical encounter during which the documented act occurred.	Code	ebRIM Classification	4.2.3.2.11
Home-CommunityId	A globally unique identifier for a community.	OID URN	home XML attribute	4.2.3.2.12
languageCode	Specifies the human language of character data in the document.	String	ebRIM Slot	4.2.3.2.13
Legal-Authenticator	Represents a participant within an authorInstitution who has legally authenticated or attested the document.	XCN	ebRIM Slot	4.2.3.2.14
limitedMetadata	Indicates whether the DocumentEntry was created using the less rigorous requirements of metadata as defined for the Metadata-Limited Document	4.2.3.2.29	ebRIM Classification	4.2.3.2.29
mimeType	MIME type of the document.	MIME type	XML attribute	4.2.3.2.15
objectType	The type of DocumentEntry (e.g., On-Demand DocumentEntry).	UUID	XML attribute	4.2.3.2.30
patientId	The patientId represents the subject of care of the document.	CX	ebRIM External-Identifier	4.2.3.1.3
Practice-SettingCode	The code specifying the clinical specialty where the act that resulted in the document was performed (e.g., Family Practice, Laboratory, Radiology).	Code	ebRIM Classification	4.2.3.2.17
referenceIdList	A list of identifiers related to the document.	CXi	ebRIM Slot	4.2.3.2.28
Repository-UniqueId	The globally unique identifier of the repository where the document can be accessed.	OID	ebRIM Slot	4.2.3.2.18
serviceStartTime	The start time of the service being documented.	DTM	ebRIM Slot	4.2.3.2.19
serviceStopTime	The stop time of the service being documented.	DTM	ebRIM Slot	4.2.3.2.20
size	Size in bytes of the document.	Integer	ebRIM Slot	4.2.3.2.21
sourcePatientId	The sourcePatientId represents the subject of care's medical record identifier (e.g., Patient Id) in the local patient identifier domain of the creating entity.	CX	ebRIM Slot	4.2.3.2.22
sourcePatientInfo	This attribute contains demographic information of the patient to whose medical record this document belongs.	CX	ebRIM Slot	4.2.3.2.23

title	The title of the document.	UTF-8	ebRIM Name	4.2.3.2.24
typeCode	The code specifying the precise type of document from the user perspective (e.g., LOINC code).	Code	ebRIM Classification	4.2.3.2.25
uniqueId	Globally unique identifier assigned to the document by its creator.	OID	ebRIM External-Identifier	4.2.3.1.3
URI	The URI for the document.	URI	ebRIM Slot	4.2.3.2.27

More information on the most relevant XDS metadata attributes:

1525

XDS DocumentEntry metadata element	**Info**
classCode	The code specifying the high-level use classification of the document type (e.g., Report, Summary, Images, Treatment Plan, Patient Preferences, Workflow). **Description** Valid values for classCode attribute are specified by the policies of the creating entity. It is recommended that the creating entity draws these values from a coding scheme providing a coarse level of granularity (about 10 to 100 entries). For example, XDS specifies that the XDS Affinity Domain will establish this list. **Coding** There shall be exactly zero or one attribute for any Document. Example for Class Code classification may be found with a specific value set: http://wiki.ihe.net/index.php/XDS_classCode_Metadata_Coding_System
confidentialityCode	The code specifying the security and privacy tags of the document. **Description** IHE recommends, but does not require, the HL7 Privacy and Security Classification System (HCS). The use of this method is up to the policy domain such as the XDS Affinity Domain or other Trust Domain where all parties including sender and recipients are trusted to appropriately tag and enforce.

	Coding Depends on the system used
eventCodeList	This list of codes represents the main clinical acts, such as a colonoscopy or an appendectomy, being documented. The eventCodelist allows for defining specific metadata for documents with a specific classCode and practice setting. For example for a classCode: IMAGES and a practiceSettingCode: Radiology, eventCodeList is required to be included in the metadata if known by the Imaging Document Source. In other words, it is "promoted" from an optional (O) attribute in XDS to a "required if known" (R2) attribute in XDS-I.b. **Description** This list of codes represents the main clinical acts, such as a colonoscopy or an appendectomy, being documented. In some cases, the event is inherent in the typeCode, such as a "History and Physical Report" in which the procedure being documented is necessarily a "History and Physical" act. An event can further specialize the act inherent in the typeCode, such as where it is simply "Procedure Report" and the procedure was a "colonoscopy". When defining the value sets for eventCodes, they should not conflict with the values inherent in the classCode, practiceSettingCode or typeCode as such a conflict would create an ambiguous situation. **Coding** There may be zero or more codes for any Document
formatCode	The code specifying the detailed technical format of the document. **Description** Along with the typeCode, it should provide sufficient information to allow potential consumer to know if it will be able to process the document. The mimeType indicates the base format; the formatCode indicates the detailed-level technical format. The formatCode shall be sufficiently specific to ensure processing/display by identifying a document encoding, structure and template (e.g., for a CDA Document, the fact

	that it complies with a CDA schema, possibly a template and the choice of a content-specific style sheet). Format codes may be specified by multiple organizations. Format codes defined by the ITI domain shall have names with the prefix urn:ihe:iti: Format codes defined by other IHE domains shall have names with the prefix urn:ihe:'domain initials': Format codes defined by non-IHE domains should be a valid unique URN. **Coding** There shall be zero or one formatCode for any Document
healthcareFacility-TypeCode	This code represents the type of organizational setting of the clinical encounter during which the documented act occurred. **Description** In some cases, the setting of the encounter is inherent in the typeCode, such as "Diabetes Clinic Progress Note". healthcareFacilityTypeCode shall be equivalent to or further specialize the value inherent in the typeCode; for example, where the typeCode is simply "Clinic Progress Note" and the value of healthcareFacilityTypeCode is "private clinic". **Coding** There shall be zero or one healthcareFacilityTypeCode for any Document
mimeType	MIME type of the document **Description** MIME type of the document in the Repository **Coding** Max length is unbounded. Shall have only a single value.
objectType	The objectType attribute reflects the type of DocumentEntry

	Description There are two DocumentEntry types: Stable Document Entry and On-Demand Document Entry. A Stable Document Entry contains metadata about an already created document available for retrieval is a Document Entry and is designated by setting objectType equal to the UUID for Stable. An On-Demand DocumentEntry contains metadata which can be used to create an on-demand document which collects the latest, most recent available information at the time of retrieval. It is designed by setting an objectType equal to the UUID for on-demand. **Coding** Max length is unbounded. The format of the objectType value is UUID.
practiceSettingCode	The code specifying the clinical specialty where the act that resulted in the document was performed (e.g., Family Practice, Laboratory, Radiology) **Description** It is suggested that the creating entity draws these values from a coding scheme providing a coarse level of granularity (about 10 to 100 entries). Note: it will not be a very detailed list of medical specialties, but "non-overlapping" specialties. For example you want to have "pediatry" and neurology, but not "neuropediatry". **Coding** There shall be zero or one practiceSettingCode for any Document
referenceIdList	This list contains zero or more Identifiers **Description** These Identifiers may be internal or external identifiers, e.g., Identifiers may be Accession Numbers, Order Numbers, Referral Request Identifiers, XDW Workflow Instance Identifiers, etc. **Coding** The referenceIdList contains Identifiers CX encoded. May have multiple values. Max length for each value is 256 characters.

Title	Represents the title of the document **Description** Clinical documents often do not have a title, in such case the classCode (e.g., a "consultation" or "progress note") is often used as the title. In that case, the title is usually omitted. The title can be very useful as a second-level filtering criteria on a list of documents, browsed by a clinician. **Coding** The format of DocumentEntry.title shall be any string of length less than 128 characters.
typeCode	The code specifying the precise type of document from the user's perspective. **Description** It is recommended that the creating entity draw these values from a coding scheme providing a fine level of granularity such as LOINC. **Coding** There shall be zero or one typeCode for any Document

XDS Folder element	Info
Title	Represents the title of the folder. **Description** The title can be very useful to understand the semantic meaning of this gourping. **Coding** The format of Folder.title shall be any string of length less than 128 characters.
codeList	A list of codes identifying the semantics of a set of documents. These values are to be drawn from a vocabulary. Typically, a set of codes is used together to express a specific semantics, allowing for separating a

	specific folder from any other. **Coding** This attribute shall have at least one code to be meaningful.

XDS SubmissionSet metadata element	Info
contentTypeCode	**Description** The code specifying the type of clinical activity that resulted in placing these DocumentEntries, Folders, and/or Associations in this SubmissionSet. These values are to be drawn from a vocabulary defined by the creating entity that contributed the SubmissionSet. **Coding** There shall be zero or one attribute for any Submission Set

7.1. XDS-I.b metadata elements

The XDS-I.b profile is dedicated to the Radiology domain and is widely used. In this domain, health professionals have specific needs when querying for documents, for example, they might want to search by modality or body part[6].

As a consequence, the metadata are detailed as follows. It is interesting to see how the general definition can be adapted to fit a dedicated domain: here, codes are taken from DICOM and a mapping is proposed between DICOM attributes and metadata.

XDS-I.b DocumentEntry metadata element	**Info**
creationTime	This attribute value shall be populated by the XDS-I Imaging Document Source to record the date and time at which the clinical content conveyed in the shared document is created. If the published document is a DICOM object or is transformed from a DICOM information object, this attribute value should be taken from the tags Instance Creation Date (0008,0012) and Instance Creation

[6] Folder structure can also be identified (see section 3.3)

	Time (0008,0013) of the DICOM object.
eventCodeList	This attribute is required to be included in the metadata if known by the XDS-I Imaging Document Source. In other words, it is "promoted" from an optional (O) attribute in XDS to a "required if known" (R2) attribute in XDS-I.b. This attribute shall be populated by the Imaging Document Source to describe both the Acquisition Modality and Anatomic Region. The values comes from DICOM (see profile for more details), for example, for a lung CT study which contains 3 CT acquisition series, one SR series, and one Segmentation series, will have a two entries in eventCodeList: a single entry for Acquisition Modality using the code triplet "(CT, DCM, Computed Tomography)" and an entry for Anatomic Region using the code triplet; "(T-28000, SRT, Lung)".
formatCode	This attribute shall be populated by the XDS-I Imaging Document Source as follows: • Shall use "1.2.840.10008.5.1.4.1.1.88.59" (DICOM KOS SOP Class UID) as the Format Code Value and "1.2.840.10008.2.6.1" (DICOM UID Registry UID) as the Format Coding Scheme OID for a DICOM Manifest document. • Shall use "urn:ihe:rad:TEXT" for a CDA Wrapped Text Report • Shall use “"urn:ihe:rad:PDF" for a PDF Report • Shall use "urn:ihe:rad:CDA:ImagingReportStructuredHeadings:2013" for a CDA Imaging Report with Structured Headings unless overridden by a requirement in a Content Profile (such as IHE Cardiology CIRC or CRC).
mimeType	This attribute shall be populated by the Imaging Document Source from one of the following values: • “application/dicom” for a DICOM Manifest document • “text/xml” for a CDA Wrapped Text Report • “text/xml” for a CDA Imaging Report with Structured Headings • “application/pdf” for a PDF Report
practiceSettingCode	This attribute shall be populated by the Imaging Document Source to describe the high-level imaging specialty such as (R-3027B, SRT, “Radiology”), (R-3026B, SRT, “Pathology”), or (R-30248, SRT, “Cardiology”). The list of acceptable values is constrained by the organization managing the XDS Registry (i.e., the XDS Affinity Domain). It is strongly recommended to use the values from the DICOM Content Mapping Resource (PS3.16) Context Group CID 7030.

	Note: if the affinity domain is used for other types of documents, consistency should be maintained in order to avoid conflict.
referenceIdList	This attribute shall be populated by the XDS-I Imaging Document Source with the Accession Number and assigning authority of the Order Filler for each Order associated with the Imaging Document, if the Accession Number is known. An Imaging Document may be referenced by multiple Accession Numbers, so this attribute may contain multiple values, e.g., for images from a grouped acquisition or when images are acquired and read in different facilities.
serviceStartTime	This attribute shall be populated by the Imaging Document Source for a date and time that indicates the imaging service start time. As an example, the Imaging Document Source could take the value of Study Date (0008,0020) and Study Time (0008,0030) from the associated DICOM study
typeCode	This attribute shall be populated by the XDS-I Imaging Document Source from a code in the Procedure Code Sequence (0008,1032) of the performed procedure with which the document is associated. The coding system of the Radiology Imaging performed Procedure Code will be designated by the XDS Affinity Domain and shared by all XDS-I Imaging Document Sources in the XDS Affinity Domain.
uniqueID	This attribute shall contain the Document unique ID generated by the XDS-I Imaging Document Source. It shall be an ISO OID. For a DICOM Manifest document, this attribute value shall be the same as the SOP Instance UID of the corresponding DICOM KOS object. For a CDA Imaging Report with Structured Headings or a CDA Wrapped Text Report, this value shall be formulated from the HL7 CDA R2 header as follows: ClinicalDocument/id@root.ClinicalDocument/id@extension

8. Appendix B. Filtering levels and XDS metadata

Table 1 Metadata Overview Table – DocumentEntry provides the list of the metadata that may be associated with documents stored and shared within an XDS affinity domain, across XCA communities or stored on portable media created using the XDM Profile.

This Appendix provides guidance for implementers of interoperable applications leveraging the Document Sharing metadata for queries. The list should be seen as a suggestion - in practice, all XDS metadata elements can be used for grouping, filtering and sorting purposes, depending on the actual need of the person who is using an XDS Document Consumer actor.

The table qualifies the various metadata elements into four types of usage:

- Primary Filtering: Metadata attributes primarily used for querying documents and submission sets (Registry Stored Query). This may be a narrowly targeted query (looking for a specific or small set of documents) or a broad query intended to select a manageable set of likely relevant documents.
- Secondary filtering: Metadata attributes intended to be returned with the matches of a primary query and allow a human (or application) to filter, out among the returned candidates, the ones that are not relevant and need not be retrieved.
- Third-level filtering: Once the relevant documents have been retrieved the content may be processed (aggregated, displayed, etc.) and relevant information extracted. This third level is not included in the metadata table as metadata are not used for this third-level filtering.
- Technical filtering: Metadata attributes critical for the operation of the queries, but generally not visible to the clinical user. They are used for integrity verification, performance management, configuration, etc.

Table 1 Metadata Overview Table – DocumentEntry (Vol.3 4.2.3.2)

XDS METADATA ATTRIBUTE	ATTRIBUTE DEFINITION	INTENDED USE	QUERY KEYS NEED VALUE SETS
METADATA USE FOR BROAD SEARCHES			
patientId	The patientId represents the subject of care of the document. It contains the Health ID with its two parts: Authority Domain Id (OID enforced by the Registry) An Id in the above domain issued by the PDQ Supplier Actor.	Primary Query	No
serviceStartTime	Represents the start time the service being documented took place (clinically significant, but not necessarily when the document was produced or approved).	Primary Query	No
serviceStopTime	Represents the stop time the service being documented took place. Same details as serviceStartTime	Primary Query	No

classCode	The code specifying the particular kind of document. Shall have a single value. Coded with a coarse level of granularity.	Primary Query	Yes
practiceSettingCode	The code specifying the clinical specialty where the act that resulted in the document was performed (e.g., Intensive care, Laboratory, Radiology). Coarse level of granularity. Has a single value.	Primary Query	Yes
METADATA FOR TARGETED SEARCH			
healthcareFacility-TypeCode	This code represents the type of organization where the clinical encounter during which the documented act occurred. The value chosen in the value set needs to avoid conflict with the value used in the typeCode, as such a conflict would create an ambiguous situation. Has a single value.	Primary Query	Yes
availabilityStatus	An XDS Document shall have one of two availability statuses: Approved: available for patient care Deprecated: obsolete	Primary Query	No
confidentialityCode	The code specifying the level of confidentiality of the document. Has one or more values.	Primary Query	Yes
uniqueId	The globally unique identifier assigned by the document creator to this document. This unique identifier may be used in the body of other documents to reference this document. The structure and format of this Id is consistent with the document content Interoperability Specification, in particular with the formatCode attribute. Has a single value.	Primary Query	No
eventCodeList	This list of codes represents the main clinical "key words" for queries specific to certain document content (e.g. test panel code for laboratory results).The value chosen shall not conflict with the values selected in the classCode, practiceSettingCode or typeCode, as such a conflict would create an ambiguous situation. This value list may have zero or more values.	Primary Query (second level-Use Case specific)	Yes
referenceIdList	This list contains zero or more Identifiers. These Identifiers may be internal or external identifiers, E.g., Identifiers may be Accession Numbers, Order Numbers, Referral Request Identifiers, XDW Workflow Instance Identifiers, etc.	Primary Query (second level-Use Case specific)	No
METADATA FOR FILTERING QUERY RESPONSES BEFORE RETRIEVING			
title	Represents the title of the document and is encoded in UTF-8.	Secondary Filtering	No
typeCode	The code specifying the precise kind of document (e.g., Pulmonary History and Physical, Discharge Summary, Ultrasound Report). Shall have a single value.	Secondary Filtering	Yes

author	Represents the humans and/or machines that authored the document and contains the following sub-attributes: authorInstitution authorPerson authorRole authorSpeciality authorTelecommunication	Secondary Filtering	Yes+No
comments	Comments associated with the Document. Free-form text.	Secondary Filtering	No
creationTime	Represents the time the author created the document in the Document Source. Shall have a single value	Secondary Filtering	No
SPECIAL PURPOSE METADATA			
entryUUID	This globally unique identifier is primarily intended for use as a document registry management identifier. It is not meant to be an external reference (outside of the Document Registry) for documents (e.g., in links within other documents).	Technical	No
formatCode	Code globally uniquely specifying the format of the document. Along with the typeCode, it provides sufficient information to allow any potential Document Consumer to know if it will be able to process/display the document by identifying an encoding, structure and template	Technical	No
hash	Hash key of the document itself. This value is computed by the Document Repository and used by the Document Registry for detecting tampering or the improper resubmission of documents .Has a single value.	Technical	No
homeCommunityId	A globally unique identifier for a community. Configured in every document source, consumer, repository, or registry actor to enable cross community access between multiple XDS affinity domains.	Technical	No
mimeType	MIME type of the document in the Repository. Shall have a single value.	Technical	No
repositoryUniqueId	The globally unique identifier of the repository where the document is stored, assigned by each Document Repository. Has a single value.	Technical	No
size	Size in bytes of the byte stream of the document that was provided in the [ITI-42] Provide and Register – Transaction and stored by the XDS Document Repository.	Technical	No
languageCode	Specifies the human language of character data in the document. The values of the attribute are language identifiers as described by the IETF (Internet Engineering Task Force) RFC 3066. Has a single value.	Secondary Filtering	Yes
sourcePatientId	The sourcePatientId represents the subject of care medical	May Not Be Used	N/A

	record Identifier (e.g., Patient Id) in the local patient Identifier Domain of the Document Source. If used, it contains two parts: Authority Domain Id An Id in the local domain (e.g., Patient Id). It is only intended as an audit/checking mechanism and has occasional use for Document Consumer Actors.		
sourcePatientInfo	This attribute should contain demographics information of the patient to whose medical record this document belongs, as the Document Source knew it at the time of Submission. It is only intended as an audit/checking mechanism and has occasional use for Document Consumer actors.	May Not Be Used	N/A
legalAuthenticator	Represents a participant who has legally authenticated or attested the document within the authorInstitution. Legal authentication implies that a document has been signed manually or electronically by the legalAuthenticator. This attribute may be absent if not applicable. If present, shall have a single value	May Not Be Used	N/A
URI	The URI for the document	Technical	No
objectType	The type of DocumentEntry (e.g. On-Demand DocumentEntry)	?	Yes
limitedMetadata	Indicates whether the DocumentEntry was created using the less rigorous requirements of metadata as defined for the Metadata-Limited Document Source	Technical	Yes

A SubmissionSet plays the role of an "envelope" within which zero or more documents have to be placed for submission and registration. Such a concept is clinically important, when it represents semantics, not only about the concurrent sharing of a document set, but also about the clinical significance of their grouping (e.g. a hospital discharge summary along with attached laboratory and cardiology reports). The metadata attributes related to a SubmissionSet and their use for queries is presented in Table 2 Metadata Overview Table – SubmissionSet.

Table 2 Metadata Overview Table – SubmissionSet (Vol.3 4.2.3.3)

XDS METADATA ATTRIBUTE	ATTRIBUTE DEFINITION	INTENDED USE	PRIMARY QUERY CODES NEED VALUE SET
availabilityStatus	A SubmissionSet has one of two availability statuses: Approved: available for patient care	Primary Query	No

	Deprecated: this document has been replaced If present, has a single value.		
contentTypeCode	The code specifying the type of clinical activity that resulted in placing these documents in this SubmissionSet. Has a single value.	Primary Query	Yes
entryUUID	This globally unique identifier is primarily intended for use as a document registry management identifier. It is not meant to be an external reference (outside of the Document Registry) for documents (e.g., in links within other documents).	Primary Query	No
intendedRecipient	Represents the organization(s) or person(s) for whom the SubmissionSet is intended.	Primary Query	No
patientId	The patientId represents the medical record identifier of subject of care whose longitudinal record is being maintained. Has a single value.	Primary Query	No
sourceId	OID identifying the instance of the Document Source that contributed the SubmissionSet. When a "broker" is involved in sending submission sets from a collection of client systems, it should use a different source ID for submissions from each separate system to allow for tracking.	Primary Query	No
submissionTime	Point in Time at the Document Source when the SubmissionSet was created and issued for registration to the Document Registry. Has a single value.	Primary Query	No
uniqueId	Globally unique identifier for the submission-set instance assigned by the Document Source in OID format. Has a single value.	Primary Query	N/A
title	Represents the title of the SubmissionSet .If present, has a single value.	Secondary Filtering	No
comments	Comments associated with the SubmissionSet. Free form text with an XDS Affinity Domain specified usage.	Secondary Filtering	No
author	Represents the humans and/or machines that authored the document. This attribute contains the following sub-attributes: authorInstitution authorRole authorSpeciality authorPerson authorTelecommunication	Secondary Filtering	No
homeCommunityId	A globally unique identifier for a community.	Technical	No
limitedMetadata	A flag that the associated SubmissionSet was created using the less rigorous metadata requirements as defined for the Metadata-Limited Document Source	Technical	Yes

1575

Table 3 Metadata Overview Table – Folder (Vol.3 4.2.3.4)

XDS METADATA ATTRIBUTE	ATTRIBUTE DEFINITION	INTENDED USE	PRIMARY QUERY CODES NEED VALUE SET
availabilityStatus	A Folder has one of two availability statuses: • Approved: available for patient care • Deprecated: this document has been replaced If present, has a single value.	Primary Query	No
CodeList	This list of codes represents the main clinical "key words" for queries specific to certain document content (e.g. test panel code for laboratory results). This value list may have zero or more values. In order to use folders without influencing each other more than one code is recommended.	Primary Query	Yes
entryUUID	This globally unique identifier is primarily intended for use as a document registry management identifier. It is not meant to be an external reference (outside of the Document Registry) for documents (e.g., in links within other documents).	Primary Query	No
lastUpdateTime	Point in Time at the Document Source when the Folder was created and issued for registration to the Document Registry. Has a single value.	Primary Query	No
patientId	The patientId represents the subject of care of the document. It contains the Health ID with its two parts: • Authority Domain Id (OID enforced by the Registry) • An Id in the above domain issued by the PDQ Supplier Actor.	Primary Query	No
uniqueId	Globally unique identifier for the folder instance assigned by the Document Source in OID format. Has a single value.	Primary Query	N/A
title	Represents the title of the Folder. If present, has a single value.	Secondary Filtering	No
comments	Comments associated with the Folder. Free form text with an XDS Affinity Domain specified usage.	Secondary Filtering	No
homeCommunityId	A globally unique identifier for a community.	Technical	No
limitedMetadata	A flag that the associated Folder was created using the less rigorous metadata requirements as defined for the Metadata-Limited Document Source	Technical	Yes

9. Appendic C. DICOM Categories

DICOM data object consists of a number of attributes, including items such as name, ID, etc., and also one special attribute containing the image pixel data.

Below is the table DICOM CID 29 table: Acquisition Modality

DICOM Acquisition Modality

Value	Description
AU	Audio
BI	Biomagnetic Imaging
CD	Color flow Doppler
CR	Computed radiography
CT	Computed tomography
DD	Duplex Doppler
DG	Diaphanography
DSA	Digital Subtraction Angiography
DX	Digital Radiography
ECG	Electrocardiography
EPS	Cardiac Electrophysiology
ES	Endoscopy
GM	General Microscopy
HC	Hard Copy
HD	Hemodynamic Waveform
IO	Intra-Oral Radiography
IVUS	Intravascular Ultrasound
LS	Laser surface scan
MG	Mammography
MR	Magnetic Resonance
NM	Nuclear Medicine
OCT	Optical Coherence Tomography
OP	Ophthalmic Photography

OPM	Ophthalmic Mapping
OPR	Ophthalmic Refraction
OPV	Ophthalmic Visual Field
OT	Other
PR	Presentation State
PET	Positron Emission Tomography - PET
PX	Panoramic X-Ray
REG	Registration
RF	Radio Fluoroscopy
RG	Radiographic imaging (conventional film/screen)
RTDOSE	Radiotherapy Dose
RTIMAGE	Radiotherapy Image
RTPLAN	Radiotherapy Plan
RTRECORD	RT Treatment Record
RTSTRUCT	Radiotherapy Structure Set
SEG	Segmentation
SM	Slide Microscopy
SMR	Stereometric Relationship
SR	SR Document
ST	Single-photon emission computed tomography (SPECT)
TG	Thermography
US	Ultrasound
XA	X-Ray Angiography
XC	External-camera photography

10. Appendix D. Radlex Playbook

1. http://www.rsna.org/RadLex_Playbook.aspx
2. <more info needed> contact has been made with Ken Wang (kcwang@gmail.com) and Christopher Carr (ccarr@rsna.org

11. Appendix E. LOINC

LOINC has defined document types. FHIR has selected a subset of these document types as *class* valueset for the Composition resource (http://hl7.org/fhir/composition.html):

The FHIR Class Codes have several weaknesses:

- numerous overlaps between values: Legal and driver's license, Legal and contract, etc.
- very biased towards administrative information
- has photographic image, but not medical image
- etc.

Composition class codes: http://hl7.org/fhir/valueset-doc-classcodes.html

Composition type codes: http://hl7.org/fhir/valueset-doc-typecodes.html

Class codes (LOINC)

Code	Display
LP173387-4	Administrative note
LP173388-2	Against medical advice note
LP173389-0	Agreement
LP173390-8	Certificate
LP173394-0	Consent
LP173403-9	Contract
LP193873-9	Driver license
LP173404-7	Health insurance card
LP173405-4	Health insurance-related form
LP173406-2	Health record cover sheet
LP173407-0	Legal document
LP181089-6	Request
LP173409-6	Advance directive
LP173410-4	Do not resuscitate
LP173412-0	Living will
LP173413-8	Rescinded advance directive

Code	Display
LP173414-6	Diagram
LP173415-3	Flowsheet
LP181112-6	Form
LP181116-7	Instructions
LP181119-1	Action plan
LP173118-3	Discharge instructions
LP173416-1	Legal
LP173417-9	Letter
LP173418-7	Note
LP173419-5	Adverse event note
LP173420-3	Alert
LP181207-4	Order
LP181204-1	Prescription
LP156982-3	Photographic image
LP173421-1	Report
LP183503-4	Case report
LP183502-6	Registry report

12. Appendix F. Links to relevant sites

IHE XDS Metadata:
http://wiki.hl7.de/index.php?title=IG:Value_Sets_f%C3%BCr_XDS
http://wiki.hl7.de/index.php?title=IHE_DE_Cookbook
http://art-decor.org/art-decor/decor-valuesets--ihede-

IHE Document Class Codes:

http://wiki.ihe.net/index.php/XDS_classCode_Metadata_Coding_System

IHE Format Codes :

http://wiki.ihe.net/index.php?title=IHE_Format_Codes&redirect=no

De INT versie has a direct link:

http://art-decor.org/art-decor/decor-project--ihexds-

NL version:

http://decor.nictiz.nl/art-decor/home

-> DECOR -> Nictiz -> IHE XDS:

http://decor.nictiz.nl/art-decor/decor-project--ihexds-

ABRUMET version:

https://gazelle.ehealth.brussels/art-decor/decor-valuesets--abrumetxds-

ASIP Santé version:

Mapping between CDA header and XDS metadata:
http://esante.gouv.fr/sites/default/files/CI-SIS_ANX_LIENS-CDA-METADONNEES-XDS_V1.3.1.0.pdf
http://esante.gouv.fr/sites/default/files/asset/document/ci-sis_contenu_volet-structuration-minimale_v1.3.2.1.pdf
http://esante.gouv.fr/sites/default/files/asset/document/ci-sis_services_volet-partage-documents-sante_v1.3.2.1.pdf
http://esante.gouv.fr/sites/default/files/HIS-IF-Content_Layer-Common_Rules_and_Templates_for_CDA_Headers_Module_V1.1.0_R.pdf
http://esante.gouv.fr/sites/default/files/HIS-IF-Service_Layer-Document_Sharing_Module_v1.1.0_R.pdf

Implementation guide
"Guide to Interoperability between XDS Affinity Domains 2015".

AUSTRIA version:
APPC - mandatory for XDS in national EHR ELGA
http://www.bura.at/austrian-pacs-procedure-code-appc/appc-the-code/
http://www.bura.at/austrian-pacs-procedure-code-appc/

XDS metadata specifications :

http://www.elga.gv.at/fileadmin/user_upload/Dokumente_PDF_MP4/CDA/Implementierungsleitfaeden/Implementierungsleitfaeden_2.06/Implementierungsleitfaden_XDS_Metadaten_V2.06.pdf

Germany version:

http://wiki.hl7.de/index.php?title=IG:Value_Sets_f%C3%BCr_XDS

http://wiki.hl7.de/index.php?title=IHE_DE_Cookbook

http://art-decor.org/art-decor/decor-valuesets--ihede-

FHIR

Composition resource:

http://hl7.org/fhir/composition.html

Composition type codes:

http://hl7.org/fhir/valueset-doc-typecodes.html

Composition class codes:

http://hl7.org/fhir/valueset-doc-classcodes.html

13. Appendix G. Definition

Concept	Description	Source
Granularity	The level of detail considered in a model or decision making process. The greater the granularity, the deeper the level of detail.	HITCH D6.4 Final Report
Interoperability	The possibility of two ICT systems to exchange information. There are more types of interoperability. *Syntactic* interoperability is defined as the possibility of two systems to interact, *semantic* interoperability as the possibility of different systems to automatically interpret the information that is being exchanged. To achieve semantic interoperability, both sides must refer to a common information exchange reference model	