Aalborg University

Master in Biomedical Engineering and Informatics

Master's Thesis

Project Group 10407

Preserving Contextual Information from Unstructured Free Text Documents Using NLP, SNOMED CT, and HL7 FHIR to Achieve Semantic Interoperability

Authors:

Johanne Krogsgaard Jensen
Thea Mentz Sørensen

Internal supervisor:

Louise Pape-Haugaard

AALBORG UNIVERSITY
STUDENT REPORT

Det Sundhedsvidenskabelige Fakultet
Niels Jernes Vej 10
9220 Aalborg Ø

AALBORG UNIVERSITET
STUDENTERRAPPORT

Titel:
Bevaring af kontekstuel information fra ustrukturerede fritekstdokumenter ved brug af NLP, SNOMED CT og HL7 FHIR for at opnå semantisk interoperabilitet

Projekt:
Kandidatspeciale i Sundhedsteknologi

Projektperiode:
Februar 2021 - Juni 2021

Projektgruppe:
21gr10407

Udarbejdet af:
Johanne Krogsgaard Jensen
Thea Mentz Sørensen

Intern vejleder:
Louise Pape-Haugaard

Oplagstal: Digitalt tilgængeligt
Sidetal: 85
Appendiks: 5
Afsluttet: 1. juni 2021

Synopsis:

Introduktion: Anvendelsen af fritekstdokumentation understøtter kliniske praksis, men der opstår udfordringer ved genanvendelse af fritekstdokumenter. Da det ikke er muligt at genanvende alt information, bør fokus være på bevarelse af den situationelle kontekst ved håndteringen af udfordringerne. Derfor vil dette studie udforske, hvordan den situationelle kontekst kan bevares når relevant information udtrækkes og struktureres fra fritekstdokumenter, for at opnå semantisk interoperabilitet.
Metode: Udskrivningsepikriser fra N2C2 2010 konkurrencen blev anvendt som datagrundlag, som sammen med en implementeringskontekst satte dette rammen for udviklingen. HL7 FHIR ressourcer, SNOMED CT udtryk og NLP systemet, cTAKES, blev anvendt til at strukturere, kode og udtrække information fra udskrivningsepikriserne. cTAKES blev justeret vha. en agil udviklingstilgang. Fokus for justeringerne var at inkludere mere kontekstuel information ved brug af post-koordinerede udtryk fra SNOMED CT. Dette var testet op imod en gold standard.
Resultat: De 21 validerede FHIR profiler indeholdte 95,5% af information fra udskrivningsepikriserne. Det justerede cTAKES havde en F-score på 0,120.
Konklusion: Den situationelle kontekstuelle information fra fritekstdokumenter kan bevares ved brug af HL7 FHIR og SNOMED CT. Derimod er automatiseret dataudtræk ved brug af cTAKES endnu ikke moden til klinisk anvendelse.

The Faculty of Medicine
Niels Jernes Vej 10
9220 Aalborg Ø

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Preserving Contextual Information from Unstructured Free Text Documents Using NLP, SNOMED CT, and HL7 FHIR to Achieve Semantic Interoperability

Project:
Master Thesis in Biomedical Engineering and Informatics

Project Period:
February 2021 - June 2021

Project Group:
21gr10407

Authors:
Johanne Krogsgaard Jensen
Thea Mentz Sørensen

Internal Supervisor:
Louise Pape-Haugaard

Circulation: Digital available
Number of total pages: 85
Number of appendix: 5
Submission date: June 1st 2021

Abstract:

Introduction: The use of free text documentation supports clinical practice but challenges arise when reusing free text documents. Since it is not possible to reuse all information within healthcare, a focus on preserving the situational context must be retained when handling the challenges. Therefore, the objective of this study was to explore how the situational context can be preserved when extracting and structuring relevant information from free text documents in order to obtain semantic interoperability.
Method: Discharge summaries from the N2C2 2010 challenge were used as the data foundation, which together with an implementation context set the scope for the development. HL7 FHIR resources, SNOMED CT expressions, and the NLP system cTAKES, were used to structure, encode, and extract information from the discharge summaries. cTAKES was adjusted using an agile development approach. The focus of the adjustments were to include more contextual information by using post-coordinated expressions from SNOMED CT, and these were tested against a gold standard.
Result: The 21 FHIR profiles contained 95.5% of information from the discharge summaries. The adjusted cTAKES had a F-score of 0.120.
Conclusion: The situational contextual information from free text documents can be preserved using HL7 FHIR and SNOMED CT. However, automatic information extraction using cTAKES, lack the maturity for clinical use.

Preface

This report communicates the study with the title 'Preserving Contextual Information from Unstructured Free Text Documents Using NLP, SNOMED CT, and HL7 FHIR to Achieve Semantic Interoperability'. This study constitutes the Master's Thesis authored by group 21gr10407, 4^{th} semester Master of Biomedical Engineering and Informatics at Aalborg University. The authors would like to thank Louise Pape-Haugaard from Aalborg University for inspiring supervision and counseling, and Maibrit Pape from Aalborg University Hospital for participating in an interview.
The authors of this study gained access to data sets from the N2C2 challenges [Department of Biomedical Informatics Harvard Medical School, 2018], but are not allowed to share the data with external partners. For this reason, all examples from the data foundation are given with inspiration from the data set.
In the study, several FHIR profiles are developed and adjustments are applied to the default cTAKES system. The profiles and source code are available for supervisor and external examiner from June 21^{st} till June 28^{st} 2021 on the following link: `https://github.com/JohanneJensen/21gr10407`.

Johanne Krogsgaard	Thea Mentz
Johanne Krogsgaard Jensen	Thea Mentz Sørensen

Reading Instructions

In the report, the problem domain and the objective are initially presented, followed by method and result for FHIR profiling, identification of current literature on NLP systems, and adjustments of the default cTAKES system. The methods and results are then discussed, followed by a conclusion.
The Harvard reference style is used for both active and passive references in the report. Active references are incorporated in a sentence, e.g. "In the article by Peterson et al. [2020] HL7 FHIR...". Passive references are in the end of a sentence, e.g. "...free text are unstructured [Meystre et al., 2017]." Passive references before a period covers the particular sentence, and passive references after a period covers the till the prior reference or paragraph. In the bibliography all references can be seen, sorted by family name of the first author.
Tables and figures are referred to using the chapter and an increasing number, e.g. figure 2.1 refers to the first figure in chapter 2.
Abbreviations are presented in parentheses after the word or phrase they describe. An overview of abbreviations used more than once throughout the report can be seen on page vi.

Abbreviations

BERT	Bidirectional Encoder Representations from Transformers
CDA	Clinical Document Architecture
CEN	European Committee for Standardization
cTAKES	Clinical Text Analysis and Knowledge Extraction System
EHR	Electronic Health Record
FHIR	Fast Healthcare Interoperability Resources
FN	False negatives
FP	False positives
HL7	Health Level 7
ISO	International Organization for Standardization
MIMIC-III	Medical Information Mart for Intensive Care III
MRCM	Machine Readable Concept Model
MTERMS	Medical Text Extraction, Reasoning, and Mapping System
NLP	Natural Language Processing
N2C2	National NLP Clinical Challenges
POS	Part-of-speech
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
TP	True positives
TUI	Type Unique Identifier
UMLS	Unified Medical Language System

Table of Contents

1 Introduction

Semantic interoperability involves seamless exchange of healthcare data between health information systems with an unambiguous understanding of data. It has the potential to save time for healthcare professionals and reduce errors [Benson and Grieve, 2016c]. Semantic interoperability allows healthcare data to be reused for secondary purposes like clinical decision support and research [Garde and Knaup, 2006]. In literature it is argued that it is not enough to achieve semantic interoperability, as some contextual information is missing despite its importance when reusing data. The ability to share information while maintaining the original perception of the context is called pragmatic interoperability. In addition to semantic interoperability, pragmatic interoperability rely on all information must be understood in the exact same way by the sender and receiver. [Asuncion and van Sinderen, 2010]

There are several reasons why interoperability of healthcare data in general is challenging including the complexity of information, various data types, variability of treatments, life span of information, demand for a holistic view of the patient, different users of the healthcare data, and patient safety [Garde and Knaup, 2006].

Healthcare data in an electronic health record (EHR) are stored as different data types e.g. narratives, documents, images, or laboratory results [Garde and Knaup, 2006]. These data types have varying degrees of structure where narratives and documents with free text are unstructured [Meystre et al., 2017]. In these unstructured data types, a lot of information is contained [Meystre et al., 2017; Chen et al., 2009; Meystre et al., 2008]. In order to achieve pragmatic interoperability of healthcare data, unstructured data need to be structured and therefore it is relevant to consider; why unstructured clinical data are necessary in healthcare, which challenges might arise when trying to achieve pragmatic interoperability of unstructured clinical data, and if it is possible to achieve at all.

Problem Analysis 2

This chapter represents an analysis of the problem domain and possible solutions. The methodological approach for collecting relevant knowledge for this chapter can be seen in appendix A.

2.1 The Need for Free Text Documentation

The unstructured data types described in section 1 on the previous page are convenient to use when presenting medical concepts or events in an EHR [Lin et al., 2015] and will henceforth be referred to at free text documents. In the opinion papers by González Bernaldo de Quirós et al. [2018] and Roberts [2017] it was questioned if it makes sense to structure all information about a patient. By doing so, the patient becomes a member of a population with similar measured values or diagnoses, which is valued for research purposes. On the contrary, is usage of free text documentation which allows to consider the patient in a holistic view, to treat the patient as an individual, and thereby making it possible to give them the optimum care. [González Bernaldo de Quirós et al., 2018; Roberts, 2017]

When documenting, healthcare professionals prefer to use free text written in natural language as it allows flexibility and efficiency [González Bernaldo de Quirós et al., 2018; Rosenbloom et al., 2011]. This is emphasized by the fact that around 70% of data available in an EHR are free text documents, and this data are elaborated by direct entry by healthcare providers, transcription of dictations, or use of speech recognition applications [Roberts, 2017; Meystre et al., 2008] By using natural language documentation the healthcare professionals can be expressive and describe impressions, thoughts, reasoning, level of concern, and uncertainties to those reviewing the free text document [Peterson et al., 2020; Rosenbloom et al., 2011]. Additionally, free text documents are often written under time pressure, which can result in parenthetical expressions, acronyms, and jargon being used, as well as eliding words. These elements contribute to increase the density of information in the free text documents. [Leaman et al., 2015] For these reasons, it is considered an advantage to use free text documents in medical areas with time pressure e.g. acute care making documentation flexible and efficient.

The context in which data are collected to an EHR is described through relations and situations that contributes to a meaningful understanding of the data [ISO/TC 215, 2005]. The context can be more or less explicitly described in an EHR, however it is crucial for the reliability when others try to interpret or reconstruct the meaning of the data. Though, when structuring the process of data collection a loss of contextual information occurs. [González Bernaldo de Quirós et al., 2018; Roberts, 2017; Patel et al., 2002; Ingenerf, 1999]

To summarize, it is important to retain the possibility for free text documentation to 1) ensure individualized care, 2) support workflow of healthcare professionals, and 3) secure that the contextual information is maintained when reviewing the data.

2.2 Free Text or Structured Data Entry

Structured data entry can be rich and convenient, but healthcare professionals cannot express a situation further than a pre-defined set of data entries [Roberts, 2017; Rosenbloom et al., 2011; Berg, 2001]. According to the articles by Jones et al. [2020]; Roberts [2017] and Meystre et al. [2008] a pre-defined set of characteristics and features are not sufficient to describe the following three areas; 1) considerations of e.g. personal, cultural, and social circumstances of the patient, 2) complex interplay e.g. between the patient's life, diseases, symptoms, and treatments, and 3) clinical considerations of e.g. intervention, pathology reports, imaging reports, and diagnoses.
Diagnoses and symptoms should according to the article by Meystre et al. [2008] be collected as free text. Though it has to be considered that this article is of an earlier date than the articles by Roberts [2017] and Jensen et al. [2012], who state that diagnoses and symptoms can be collected in a structured data format. Besides diagnoses and symptoms, structured data entry can be applied to medication Roberts [2017]; Jensen et al. [2012], radiological image data, and laboratory test results Jensen et al. [2012].
When using structured data entry, it has to be adjusted to the individual healthcare professional in order to avoid possible interruptions of their workflow [Galster, 2013a]. The observational study by Galster [2013a] showed that the structure of data collection performed by the healthcare professionals, was unique for almost every person. It was additionally found that there is a disagreement between the structure of the EHR, and the way healthcare professionals prefer to enter data. This disagreement can result in an interruption of workflow and possibly induced risks of an increased mental load on the healthcare professionals as well as an increased risk for medical errors. [Galster, 2013a; Greenhalgh et al., 2008] For this reason, it is important to have exhaustive knowledge about the workflow and implement flexibility before structuring data collection.

2.3 Obtain Information from Free Text Documents

Currently, there is a lot of relevant, and unused data stored in free text documents, which are fundamental to obtain if more is to be learn from this healthcare data [Jones et al., 2020; Viani et al., 2018]. Obtained relevant information from the free text documents can be reused for multiple purposes ranging from large scale research to quality assurance and patient management [Jones et al., 2020; Wong et al., 2018; Chen et al., 2009], also including possible identification of adverse medication events or clinical decision support [Wong et al., 2018].
According to an article by Peterson et al. [2020] the responsibility of structured data entry or extraction of relevant information from free text documents cannot be fully assigned to the healthcare professional. This is not a feasible option as it requires significant time and effort, which would be added upon an already full workload, why it must be handled through coding and information extraction [Peterson et al., 2020].
Through the structured literature search, see appendix A, it was found possible either to

structure the writing of the free text documents or extract data from free text documents, called semi-automatic annotation or information extraction, respectively [Lin et al., 2015]. Structuring of data can be used to enable interoperability, and thereby making it possible to share necessary patient data [Wong et al., 2018].
Semi-automatic annotation involves construction of a pre-defined lexicon, which is mapped to a standard medical terminology. The lexicon is used when a healthcare professional is documenting e.g. a patient's disease. The healthcare professionals are presented to a list of words or phrases based on the written characters, from which a fitting word or phrase can be chosen. The chosen word or phrase are mapped to a terminology binding in the information system, securing unambiguity of the it. [Peterson et al., 2020; Lin et al., 2015]
Natural language processing (NLP) is a subtype of information extraction and is often used to extract information for research [Peterson et al., 2020; Gaudet-Blavignac et al., 2018; Roberts, 2017; Lin et al., 2015; Chen et al., 2009; Lussier et al., 2001]. NLP can be used to extract and encode data from free text documents. It is, contrary to semi-automatic annotation, applied after the free text documents are elaborated. [Chen et al., 2009]

2.3.1 The Contextual Challenges

As mentioned in section 2.2 on the preceding page it is crucial to describe the context to enable reliable reuse of the collected healthcare data. Despite this, not all contextual information is collected and documented, which might be due to healthcare professionals being selective [Nelson, 1997] or have limited time to document [Galster, 2013b]. A situation where contextual information is shared but not documented is at a morning conference for healthcare professionals. Here many questions are asked concerning a patient e.g. who was the healthcare professional that made a certain decision, where did a given information collection take place, or under what circumstances was the information collected. [Galster, 2013b] This information concern both the organizational context e.g. the structure of the department and a common professional vocabulary [Nelson, 1997], as well as the situational context e.g. the experience of the healthcare professional and the patient's condition [Galster, 2013b]. For both the organizational and the situational context it is important to understand the circumstances under which the information is collected [Galster, 2013b].
As mentioned, not all contextual information is documented and when structuring data there are a loss of information, which makes it impossible to achieve pragmatic interoperability. Though, it is assumed possible to achieve semantic interoperability, as data from the situational context is partly documented and therefore structured.

2.3.2 The Linguistic Challenges

Several studies state multiple challenges which arise when attempting to extract relevant information from unstructured data [González Bernaldo de Quirós et al., 2018; Roberts, 2017; Leaman et al., 2015; Chen et al., 2009; Meystre et al., 2008; Lussier et al., 2001]. Free text documents are written to contain a high density of information [Leaman et al., 2015], why many abbreviations and acronyms are used [González Bernaldo de Quirós et al., 2018; Roberts, 2017; Leaman et al., 2015; Chen et al., 2009; Meystre et al., 2008; Lussier et al., 2001] along with local jargon and dialect [González Bernaldo de Quirós et al., 2018; Leaman et al., 2015]. Free text documents are additionally nuanced [Peterson et al., 2020; Roberts, 2017] and ambiguities can easily occur in jargon, abbreviation, or acronyms [González Bernaldo de Quirós et al., 2018; Roberts, 2017; Chen et al., 2009; Lussier et al., 2001]. Misspelling [Chen et al., 2009; Meystre et al., 2008; Lussier et al., 2001], incorrect grammar, and absence of words which potentially could be deduced based on the context [Leaman et al., 2015; Meystre et al., 2008], can occur as a result of the hectic situations where the information is collected. The hectic situations might also explain why attempts to do ad-hoc formatting are seen in clinical free text e.g. by the use of sections [Leaman et al., 2015; Meystre et al., 2008]. Further, the use of negations as well as words describing uncertainties are common [Peterson et al., 2020; Roberts, 2017; Chen et al., 2009]. Lastly, for free text documents originally written in an EHR, the possible characters are only limited to which can be typed or pasted, why it can be challenging to recognize in secondary systems [Leaman et al., 2015; Meystre et al., 2008].

2.4 Achieving Semantic Interoperability of Free Text

2.4.1 Standard Terminology

Information from free text documents has to be computer-readable to facilitate semantic interoperability across healthcare information systems [Chen et al., 2009]. To accommodate the mentioned difficulties about extracting relevant information from free text documents, the application of standard terminologies are proposed in the articles by Peterson et al. [2020] and Chen et al. [2009]. Using standard terminologies it is possible to encode words and phrases from the free text documents [Peterson et al., 2020; Roberts, 2017; Chen et al., 2009]. The creation of mapping between the words and phrases, and the standard terminologies is crucial. However it is still a challenge for semantic interoperability. [Peterson et al., 2020; Lussier et al., 2001]
Empirical work shows that words and phrases extracted from the free text documents often cannot be expressed using one code but needs multiple to capture more of the situational context [Peterson et al., 2020; Liu et al., 2012; Elkin et al., 2006]. This challenge can be addressed by the reference terminology Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) as it supports post-coordinated expressions [Gaudet-Blavignac et al., 2018]. This enables SNOMED CT to associate concepts and therefore it has properties similar to the natural language used in free text documents [Gaudet-Blavignac et al., 2018; González Bernaldo de Quirós et al., 2018]. Furthermore, SNOMED CT is multiple times mentioned as a promising standard terminology for mapping extracted words and phrases [Peterson et al., 2020; Gaudet-Blavignac et al., 2018; González Bernaldo de Quirós et al., 2018; Lin et al., 2015]. According to an article by Gaudet-Blavignac et al. [2018] binding post-coordinated SNOMED CT expressions to the extracted words and phrases is necessary in order to capture more of the situational context from free text documents.

2.4.2 Interoperability Standard

Using SNOMED CT allows handling of ambiguities and thereby partly the semantic issue when creating interoperability. But a standard for structuring data is necessary to enable sharing of data across systems. [Chen et al., 2009] Interoperability can be achieved when structuring the encoded words and phrases with a standard [Lin et al., 2015].

Four standards were identified through the structured literature search, see appendix A, each with a different purpose: European Committee for Standardization (CEN) 13606, openEHR, Health Level 7 (HL7) Clinical Document Architecture (CDA) [Lin et al., 2015], and HL7 Fast Healthcare Interoperability Resources (FHIR) [Bender and Sartipi, 2013].

2.4.2.1 Architectural Standards

CEN 13606 was later renamed International Organization for Standardization (ISO) 13606. Its purpose is to define a stable information architecture for exchanging data between information systems, why it is an architectural standard. [VeraTech for Health, 2021]
OpenEHR is an architectural standard as well which enables interoperability by defining clinical models and open specifications [OpenEHR International, 2021].
In order to obtain interoperability using the architectural standards ISO 13606 and OpenEHR, the information systems sharing the data need to follow the same architectural standard [VeraTech for Health, 2021; Beale, 2001].

2.4.2.2 Exchange Standards

HL7 CDA is an exchange standard which is usually implemented with the architectural standard HL7 v3. HL7 CDA structures the clinical data in different documents with varying degrees of structure, which can be used to exchange the healthcare data. [Benson and Grieve, 2016a] However, for HL7 CDA to facilitate interoperability it is depended on the sending and receiving system to follow the same architectural standard [Beale, 2001].
HL7 FHIR is an interoperability standard which further allows for clinical data to be structured. Additionally, this standard can be applied to existing systems as it is loose coupled to the system, which makes it flexible. [Benson and Grieve, 2016b]

2.5 Objective

It is necessary to retain the possibility to write free text documents for the healthcare professionals, as these are needed to 1) ensure individualized care, 2) support workflow of healthcare professionals, and 3) secure that the contextual information is maintained when reviewing the data. To obtain semantic interoperability of free text documents, it is necessary to extract, encode, and structure these while including the situational contextual information.
It was found that information in free text documents can both be structured using semi-automated annotation or NLP. It was found that the responsibility for structuring healthcare data cannot be assigned to the healthcare professionals, as they have an already full workload. Based on this the semi-automatic annotation was not selected, why NLP was chosen to extract relevant information from free text documents.
Additionally, the reference terminology SNOMED CT was found advantageous as it allows both pre- and post-coordinated expressions, and therefore can include situational context from the unstructured data, which is important to ensure mutual understanding. As mentioned both the architectural standards and HL7 CDA is dependent on the sending and receiving system to follow the same architectural standard. On the contrary, HL7 FHIR describes how to exchange and structure the healthcare data while loosely coupled to the systems. Therefore, HL7 FHIR was chosen to enable semantic interoperability in

collaboration with the reference terminology SNOMED CT.
The objective of this study is to explore how the situational context can be preserved when extracting and structuring relevant information from free text documents in order to obtain semantic interoperability.

3 Design Considerations

3.1 Design Premises

As mentioned in chapter 2 on page 2, inclusion of information concerning the situational context is crucial to ensure mutual understanding when sharing data. Creating semantic interoperability of free text documents while including situational contextual information can involve application of interoperability standard and standard terminology. HL7 FHIR and SNOMED CT were chosen as standards for these tasks, respectively. In addition, a NLP system was used to extract data from free text documents. The data foundation described in section 4.1.1 on page 11 was used to restrict and concretize the implementation and choices.

Working with international standards such as HL7 FHIR and SNOMED CT, some fundamental rules and guidelines must be followed in order to conform to the standard. For SNOMED CT these rules are called compositional rules and were followed when making pre- and post-coordinated expressions [IHTSDO, 2021b]. Additionally, the entire set of SNOMED CT expressions for a specific version was used. Rules for post-coordinated expressions were further based on the concept model [IHTSDO, 2021c] and Machine Readable Concept Model (MRCM) [IHTSDO, 2021d].
In HL7 FHIR these rules are expressed through a conformance layer, which contains rules about absolute requirements and prohibitions, as well as best practises. This layer was build, since the nature of FHIR resources are fairly loose, which make the standard very flexible. However, the interoperability between information systems is at risk if the FHIR resources are profiled very differently. [Health Level 7, 2021g] To support interoperability, the rules in the conformance layer was followed.
In the article by Peterson et al. [2020] HL7 FHIR and SNOMED CT expressions were used to structure and encode data from free text documents, which were extracted with a NLP system. In the article SNOMED CT was chosen since it allows post-coordinated expressions and can therefore link related expressions in order to better capture contextual information [Peterson et al., 2020]. This is in alignment with the reason for choosing SNOMED CT in this study. Furthermore, the Condition resource was selected to structure data in the article [Peterson et al., 2020]. When focusing on one FHIR resource it was investigated how data from the free text documents fits into the Condition resource. Another approach was to investigate which FHIR resources are necessary in order to describe data from the free text documents. In this way, the maturity of the FHIR resources can be investigated, through their ability to contain information from unstructured free text, as well as preservation of the situational context. The data foundation for the profiling can be seen in section 4.1.1 on page 11 and includes discharge summaries. They contain

highly heterogeneous information, and all information is important in order to understand the context [Spasic and Nenadic, 2020; Lenert et al., 2014]. This calls for the last mentioned approach, as it enables a more open-minded approach for selection of FHIR resources and the profiling of these, why this approach was chosen for this study.

In order to extract information from free text documents using a NLP system, an existing system was identified and used. By doing so, the focus was maintained on exploring how much of the situational context that could be preserved. Therefore, it was critical to initially identify an open-source, validated, and well-performing NLP system. Necessary elements were added to the NLP system, which e.g. includes rules for making post-coordinated expressions. The rules can both be applied to the output of the NLP system or added in the system's source code. The article by Kersloot et al. [2019] used extracted concepts and relationships between the concepts to identify relevant post-coordinated expressions from the output of the NLP system. In the article an additional system was developed including rules for encoding SNOMED CT expressions. This approach showed fairly good result, but the study only bound a few expressions from a standard terminology related to cancer. [Kersloot et al., 2019] The approach of adding rules to the system's source code was chosen, as it was undesirable to develop a new system and thereby maintaining the focus on including situational contextual information.

3.2 Implementation Context

Integration of healthcare data is necessary to support management of patient's health in different settings [Peng et al., 2020]. According to the article by Peng et al. [2020], the ability to create semantic interoperability is a precondition for data integration and utilization, and thereby the usage of data. Further, a precondition for semantic interoperability is an interoperability standard e.g. HL7 FHIR and a standard terminology e.g. SNOMED CT to ensure unambiguities, as argued in chapter 2 on page 2.
As mentioned, the data foundation is described in section 4.1.1 on page 11 and includes discharge summaries. According to the article by Lenert et al. [2014], discharge summaries were developed for the primary care physician to obtain knowledge about a patient after discharge from a hospital e.g. which procedures the patient underwent during admission or which discharge medication the patient was prescribed. To accommodate this, the intended use of the NLP system and the FHIR profiles was for primary care physicians, and the intended architecture for implementation, can be seen on figure 3.1 on the following page.
A cloud-based solution can be used to deliver specific services to existing EHR systems. The cloud-based solution can be used across multiple implementations within one organization or across several organizations. Additionally, it is considerably easier to implement a cloud-based solution than to implement new local applications if multiple users have to access the data. [McCallie Jr., 2016] If the proposed solution was to be implemented, it could be through a cloud-based solution. This would consist of the NLP module, the FHIR profiles, and a mapping between the two to ensure that the extracted information is assigned to fitting elements in a FHIR profiles, as illustrated on figure 3.1 on the next page. Mapping was not performed in this study, but is a crucial element in implementation.

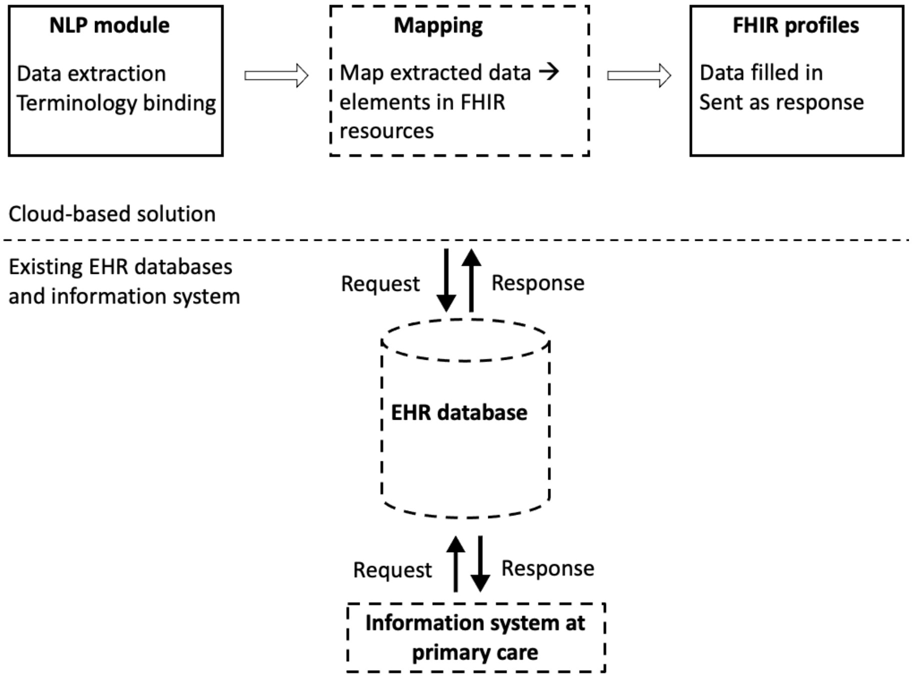

Figure 3.1. shows the overall architectural structure if the NLP system and HL7 FHIR profiles from this study were to be implemented. The horizontal dashed line indicates a separation of the applications, the dashed lined boxes indicates that the content of these are not handled, and the hard lined boxes indicates that the content was handled.

The EHR database below the dashed line indicates existing, but unspecified databases able to contain both structured and unstructured data. The dashed box in the bottom of the figure indicates that several different primary care physicians could request a discharge summary as structured data from the information system used in primary care. Since the structured data were intended to be used by primary care physician, it was assumed that all information from the discharge summary was important for the primary care physician to get an overview of the patient's admission and discharge. For this reason, information of one patient contained in the FHIR profiles was intended to be gathered and send as a response to the request.

HL7 FHIR Profiling 4

4.1 Method

In the following section the method for profiling FHIR resources is presented. The choices made throughout this section followed HL7 FHIR conformance rules to support semantic interoperability, as described in section 3 on page 8.
The FHIR profiles were developed using the newest available version 4.0.1 of HL7 FHIR [Health Level 7, 2021f] on March 1^{st} 2021. The graphical user interface Forge R4 [Firely, Bos en Lommerplein 280, Amsterdam, The Netherlands] was chosen for development, as it supports HL7 FHIR version 4.0.1, has a build-in validator, is user-friendly, and generates the FHIR profiles as Extensible Markup Language (XML) or JavaScript Object Notation (JSON) files [Firely, 2020].

4.1.1 Data set

Through the structured literature search, see section 5.2 on page 45, three different open-source databases were identified, which all include free text documents. The first and most frequently used data sets are from National NLP Clinical Challenges (N2C2) formerly known as Informatics for Integrating Biology and the Bedside (I2B2). The N2C2 data sets consist of unstructured free text documents, which are annotated with a different focus depending of the challenge of the year [Department of Biomedical Informatics Harvard Medical School, 2018]. These data sets were used by 12 of the studies identified through the structured literature search. The second data set is from Medical Information Mart for Intensive Care III (MIMIC-III) and was used in one article. MIMIC-III is a large database with both structured and unstructured clinical data from 2001 to 2012 [Johnson et al., 2016]. Lastly, a data set from MTSamples was identified, which includes transcribed unstructured free text documents [MTHelpLine, 2021]. This data set was used in one article from the structured literature search.
Since the data sets from N2C2 was the most frequently used and was annotated, it was selected as data foundation. The N2C2 data set from the 2010 challenge concerned extraction of concepts including problems, treatments, and tests as well as relations between these concepts. The data set was in English, de-identified, and the information was extracted and annotated from discharge summaries from Beth Israel Deaconess Medical Center [Department of Biomedical Informatics Harvard Medical School, 2018]. Three randomly selected discharge summaries were used as the data foundation for the FHIR profiling to identify the necessary resources, and elements.

4.1.2 Selection of FHIR Resources

This section describes two ways for identifying relevant resources, based on the data foundation and the implementation context.

4.1.2.1 Based on the Data Foundation

From the three randomly selected discharge summaries, information that could be included in a FHIR resource was identified. Each discharge summary was analyzed as shown in the first column in table 4.1. Initially, relevant information was identified from the discharge summary and assigned to the best fitting FHIR resource, from the existing FHIR resources index [Health Level 7, 2021c]. The assignment was based on a comparison between the given information from the discharge summary, and the description of the given FHIR resource, the elements in each resource, and the possible extensions. If data were assigned to a given FHIR resource, this resource was selected for profiling.
Information not possible to assign to a FHIR resource was categorized as 'unassigned words and phrases'.

From discharge summary	Resource
This is a **female at 56 years**...	Patient
with a **history of coronary artery disease**.	Observation
She was walking in the garden on 2016-04-17...	unassigned
She was walking in the garden on **2016-04-17**...	Observation
... and **developed chest pain, which soon radiated to the left arm**.	Observation
On **2016-04-18**...	Procedure
...she had **surgery and an pacemaker was inserted**.	Procedure

Table 4.1. shows examples of selecting FHIR resources based on the information from the discharge summaries.

4.1.2.2 Based on the Implementation Context

Bundles
In order to comply with the implementation context, the information in the discharge summaries has to be shared as one package. This was handle using the Bundle resource. A Bundle resource is a container for a collection of resources, and the type of Bundle depends on the purpose. [Health Level 7, 2021a] The Bundle type 'document' can be build to represent a composition of selected FHIR resources. This type can represent FHIR resources as a coherent set of healthcare information. Furthermore, the Bundle type 'document' includes an immutable set of FHIR resources which is defined by developers. [Health Level 7, 2021d] The Bundle type 'document' was found appropriate to include for the implementation context and the immutable set should be composed of the FHIR resources selected based on the data foundation.
All documents follow the same structure which includes, 1) a Bundle resource with the type 'document', 2) a Composition resource, and 3) an immutable set of FHIR resources [Health Level 7, 2021d].
The Composition resource is fundamental when creating a Bundle resource of the type

'document', as it provides identification and defines the context for the document. Further, it contains key information such as a reference to the author of the document and patient, as well as provides structure to the document.

Implementation Guide
To support implementation of the FHIR profiles, the FHIR resource ImplementationGuide was profiled. This enables definition of the responsible persons, included profiles, and purpose of the developed profiles. [Health Level 7, 2021b]

4.1.3 Selection of Elements

All information assigned to a selected FHIR resource was initially gathered to obtain an overview of information from the three randomly selected discharge summaries. If information representing an element was present, the element was included in the profile. Examples of this approach can be seen in table 4.2.

From discharge summary	Element
56-years-old	Patient.birthDate
female	Patient.gender

Table 4.2. shows examples of selecting elements for a FHIR profile based on the information from the discharge summaries. In this example elements from the Patient resource was used.

Additionally, when profiling a Bundle resource with the type 'document' and a Composition resource, some elements are mandatory, as described in Health Level 7 [2021d]. These elements were included in the given FHIR profile.
When profiling the ImplementationGuide resource, it was desired to describe the following information:

- Machine and human readable titles
- Description of the developers by name or organization
- The status of the ImplementationGuide
- The usage of the ImplementationGuide
- Which profiles were supported by the ImplementationGuide

4.1.4 Conformance Rules

The conformance rules in the conformance layer apply to IsModifier, MustSupport, constrains, and cardinality [Health Level 7, 2021g]. The first three are indicated with a flag for the given element. IsModifier and constrains can be pre-defined in the FHIR resource, while MustSupport can be set by the developers [Health Level 7, 2021j].

4.1.4.1 IsModifier

An element marked with the flag IsModifier can potentially change the interpretation of the data in the entire FHIR profile e.g. if a patient is deceased or not [Health Level 7, 2021g]. All elements from the included FHIR resources with the flag IsModifier were included in the profile despite no data foundation. This was chosen in order to communicate information regarding if the interpretation of the profile was altered. If an element was included due to the flag IsModifier, but without a data foundation to support it, the cardinality of the given element was not changed, as described in case no. 3 in table 4.3 on the following page. No further IsModifier flags were added when profiling.

4.1.4.2 MustSupport

An element marked with the flag MustSupport has to be meaningfully supported by the sending and receiving systems, if information is included in the given element. The intended use of this flag is for implementation in a known context. [Health Level 7, 2021g] As the implementation context described in section 3 on page 8 did not specify a certain system, it was chosen not to use MustSupport.

4.1.4.3 Constraints

Constrains differs from the other flags, as this flag indicates a guideline, warning, or rule of any kind which apply to the element, e.g. if two elements are mutually exclusive [Health Level 7, 2021g]. All warnings, rules, and guidelines were followed and if it required inclusion of additional elements these elements were included in the FHIR profile without changing the cardinality, as described in case no. 4 in table 4.3 on the following page.
The rule 'element1.exists( ) or element2.exists( )' was applied in cases where two elements describe the same information from the data foundation, and it was assessed necessary to include information from at least one of them. An example could be a reason for having an encounter, where the reason could be either a reference to a Condition profile or a code describing the condition. In this case, the rule ensures that if a reference is not available, a code must be chosen and vice versa.

4.1.4.4 Cardinality

All elements in a resource have an initial set cardinalities, which can be either 0..1, 0..*, 1..1, or 1..* [Health Level 7, 2021g]. The initial minimum and maximum cardinality for each element was obeyed, to follow conformance rule.
A closed modelling approach was chosen when determining the cardinalities of the elements. A closed modelling approach encourage, contrary to the open modelling approach, that the cardinalities are restricted according to the data foundation. This results in FHIR profiles which fit to the data foundation, though they are often not forward compatible [Simplifier.net, 2021a]. The closed modelling approach has contributed to changing the cardinality in five cases presented in table 4.3 on the next page.

Case no. 1	
Condition	The initial minimum cardinality is 0 The information was assessed crucial to include based on the data foundation
Action	Change the minimum cardinality to 1
Case no. 2	
Condition	The initial maximum cardinality is * There was only one instance of information in the data foundation
Action	Change the maximum cardinality to 1
Case no. 3	
Condition	The element was included due to the flag IsModifier but without data foundation
Action	Maintain the initial cardinalities
Case no. 4	
Condition	The element was included due to a constrain but without data foundation
Action	Maintain the initial cardinalities
Case no. 5	
Condition	The initial minimum cardinality is 0 The element was not identified as relevant
Action	Change the maximum cardinality to 0

Table 4.3. shows the five cases, in which the cardinality of an element was changed, including which conditions should be present and which action should be taken for each element.

4.1.5 Data Types

An element in a profile can be one of the following data types; 1) simple e.g. string, boolean, or date, 2) general purpose e.g. CodeableConcept or HumanName, 3) meta-data used for the meta-resource, and 4) special purpose e.g. a reference or an extension [Health Level 7, 2021h]. In cases where an element had one data type this was respected, and in cases an element had multiple data types, they all were included.

If the data type of an element is a reference to another profile, one or more of the recommended references were chosen based on the data foundation assigned to the given element. Therefore, references to FHIR profiles beyond the scope of the data foundation were excluded. A diagram was elaborated to show the references for each profile. Additionally, a diagram to obtain a full overview of references between all selected FHIR profiles was elaborated.

If an element has the general purpose data type coding, it required a code and system represented as a code and an URL, respectively. The system can be either local from FHIR or an external e.g. SNOMED CT. [Health Level 7, 2021m] Since SNOMED CT was chosen as terminology for this study, only the newest available version of SNOMED CT was used as system. If the data type of an element is a coded data type e.g. a Code or a CodeableConcept, a ValueSet is most often required. A ValueSet is a subset of a code system and is used to constrain which data can be present in the given element. When binding to a ValueSet it is associated with four degrees of flexibility; required, extensible, preferred, and example. [Health Level 7, 2021m] In table 4.4 the cases for maintaining, extending, and replacing a ValueSet was presented.

Case no. 1	
Flexibility	Required
Condition	
Action	Maintain the ValueSet
Case no. 2	
Flexibility	Extensible, preferable, or example
Condition	Information from the data foundation was represented in the ValueSet
Action	Maintain the ValueSet
Case no. 3	
Flexibility	Extensible or preferred
Condition	Information from the data foundation was not included in the ValueSet
Action	Extend the ValueSet with a code representing the information from the data foundation
Case no. 4	
Flexibility	Example
Condition	Information from the data foundation was not included in the ValueSet
Action	Extend or replace the ValueSet to ensure the information from the data foundation was represented

Table 4.4. shows which conditions should be present for the ValueSet to be maintain, extended, or replaced for each if the four cases.

4.1.6 Slicing

Slicing can be used when an element has the maximum cardinality of more than one and when the data requires more than one value, e.g. several identifiers for a patient. To distinguish between the slices a discriminator must be determined which consists of a path and a type. The path was selected to be a key element, and the type was selected to be the 'value' of the slice, as this is the most commonly used type. [Health Level 7, 2021l]
The slicing was chosen to be open, which means that new slices could be added, as this allows as much information to be included in the FHIR profiles as possible. This contributes to increase semantic interoperability as it allows for additional information to be included in the element. In case the maximum cardinality was * slicing was used, and a path was defined.

4.1.7 Extension

The HL7 FHIR resources includes 80% of the most frequently used healthcare data. In order to adapt the resources to the data foundation, extension can be added to the profiles. [Health Level 7, 2021k]
An extension was only added in cases where the FHIR resource did not include sufficiently

descriptive elements. Only extensions from the registry Health Level 7 [2021i] were used, as this supports semantic interoperability.

4.1.8 Validation of the Developed Profiles

The developed FHIR profiles were validated using the Firely Terminal from Firely [Firely, Bos en Lommerplein280, Amsterdam, The Netherlands]. When using the Firely Terminal it was possible to communicated with a FHIR server and to validate the developed FHIR profiles by comparing them to the base profiles from HL7 FHIR version 4.0.1. Validation was performed after installing a JSON-package using the command 'fhir install hl7.fhir.r4.call 4.0.1', which describes the metadata of the profiles and the project dependencies. As shown on figure 4.1, the command 'fhir push' was used to push a FHIR profile to the stack, and the command 'fhir validate' was used to validate the FHIR profile in the Firely Terminal.
The validation of a FHIR profile could either result in the profile being 'Valid' or 'Invalid'. [Simplifier.net, 2021b] In the latter case corrections were made following the description of the error, in order to achieve soley valid FHIR profiles. An example of a successfully validated FHIR profile can be seen on figure 4.1.

```
C:\Users\Johanne Krogsgaard\Dropbox\AAU\ST10\FHIRtestmappe>fhir push PatientcTakes.StructureDefinition.xml
Pushed 1 resource(s) on the stack.

C:\Users\Johanne Krogsgaard\Dropbox\AAU\ST10\FHIRtestmappe> fhir validate
Result: VALID
```

Figure 4.1. shows a successful validation of a developed FHIR profile. This example includes the Patient profile.

4.2 Results

4.2.1 Selected Resources

The FHIR resources used for profiling were selected based on the information contained in discharge summaries number 51, 53, and 74. The selected FHIR resources can be seen in table 4.5 alongside with the number of times content from a discharge summary was represented in the given FHIR resource.

The Bundle resource was included based on the implementation context, but could additionally include information from the the three discharge summaries.
In total 24 words and phrases from the three discharge summaries were not possible to assign to a FHIR resource. The words and phrases concerned 1) events that had not occurred e.g. 'no need for a surgery', 2) formal information about the discharge summary e.g. who completed the discharge summary, or 3) unspecified information e.g. 'the patient is a gentleman'.

Resource	Record	Number of Mentions	Percentage
Data Containing Resources			
AllergyIntolerance	51	1	0.2%
	53	0	
	74	0	
Appointment	51	1	0.2%
	53	0	
	74	0	
CarePlan	51	9	6.5%
	53	7	
	74	20	
CareTeam	51	0	0.2%
	53	0	
	74	1	
Condition	51	6	2.5%
	53	4	
	74	4	
Device	51	0	0.7%
	53	2	
	74	2	
Encounter	51	10	5.8%
	53	12	
	74	10	
FamilyMemberHistory	51	1	0.4%
	53	0	
	74	1	
Location	51	0	0.4%
	53	0	
	74	2	
Medication	51	7	4.0%
	53	4	
	74	11	
MedicationAdministration	49	12	9.6%
	53	6	
	74	35	
MedicationDispense	51	4	4.5%
	53	0	
	74	19	
Observation	51	119	42.8%
	53	17	
	74	101	
Organization	51	0	0.2%
	53	0	
	74	1	
Patient	51	4	2.0%
	53	3	
	74	4	

continued on the next page

Resource	Record	Number of Mentions	Percentage
Practitioner	51	10	5.6%
	53	9	
	74	12	
Procedure	51	5	8.5%
	53	18	
	74	24	
RelatedPerson	51	0	0.2%
	53	0	
	74	1	
Implementation Resources			
Bundle	51	3	2.0%
	53	4	
	74	4	
Composition	51	0	0.0%
	53	0	
	74	0	
ImplementationGuide	51	0	0.0%
	53	0	
	74	0	
Unassigned Data			
Unassigned words and phrases	51	8	4.3%
	53	9	
	74	7	

Table 4.5. shows the FHIR resources identified as necessary to structure and exchange information from the three discharge summaries. The column 'Number of Mentions' indicates how often information from the randomly selected discharge summaries fit to the given FHIR resource. Additionally, a percentage was calculated to show how much information each resource represented.

4.2.2 Profiling

The elements selected for each FHIR profile can be seen in the corresponding tables below. The FHIR profiles were ordered alphabetically. The selection of elements was based on information from discharge summaries 51, 53, and 74.

4.2.2.1 AllergyIntolerance

The cardinality of the element AllergyIntolerance.category was changed from 0..* to 0..1, as it was desired only to have one category for each allergy or intolerance. The cardinality of the element AllergyIntolerance.code was changed from 0..1 to 1..1 to indicate that the allergy or intolerance must be given a code to avoid ambiguities.
The element AllergyIntolerance.code was bound to the ValueSet 'AllergyIntolerance Substance/Product, Condition and Negation Codes' (http://hl7.org/fhir/ValueSet/allergy intolerance-code), as it covered the information in the data foundation.
The AllergyIntolerance profile can be seen in table 4.6, and the elements referring to another profile can be seen on figure 4.2 on the next page.

Data	Element	Cardinality	Flag
	AllergyIntolerance.clinicalStatus	0..1	?! I
	AllergyIntolerance.verificationStatus	0..1	?!
Allergies: **Shellfish**	AllergyIntolerance.category	0..1	
	AllergyIntolerance.code	1..1	
	AllergyIntolerance.patient	1..1	

Table 4.6. shows the data foundation, cardinality, and flags for each of the included elements in the AllergyIntolerance profile.
?! = IsModifier, I = constrain

Figure 4.2. shows the reference from the AllergyIntolerance profile. The text on the arrow indicates the element that refers to the other profile.

The AllergyIntolerance profile was validated using the Firely Terminal and was found valid.

4.2.2.2 Appointment

The cardinality of the element Appointment.serviceCategory was changed from 0..* to 0..1, as it was desired to have only one category assigned to each appointment.
The element Appointment.participant.actor was restricted to refer a Patient, Practitioner, or Location profile. The element Appointment.reasonReference was restricted to refer a Condition, Observation, or Procedure profile.
The element Appointment.serviceCategory was bound to the ValueSet 'Service Category' (http://hl7.org/fhir/ValueSet/service-category) as this covered the information in the data foundation.
The elements Appointment.reasonReference and Appointment.reasonCode both concern a reason for having an appointment. The rule described in section 4.1.4.3 on page 14 was added to ensure that a reason for having an appointment was included. A rule was attached to the element Appointment.participant to indicated that either the element Appointment.participant.actor or Appointment.participant.type shall be present, why the element Appointment.participant.type was included.
To the elements Appointment.reasonReference and Appointment.reasonCode slices were added, which was chosen in order to enable that several reasons for having an appointment could be included. The paths were Appointment.reasonReference.identifier.value and Appointment.reasonCode.coding, respectively. The element Appointment.participant indicates that there always must be one or more participant attached to an appointment, why this element was sliced, using Appointment.participant.actor.identifier.value as the path. The element Appointment.participant.type indicates how an individual participates in the appointment. This element was sliced, using Appointment.participant.type.coding as the path.
The Appointment profile can be seen in table 4.7, and the elements referring to another profile can be seen in figure 4.3 on the next page.

Data	Element	Cardinality	Flag
	Appointment.status	1..1	?!
Call trauma clinic to **coordinate the study, result** and appointment dates	Appointment.serviceCategory	0..1	
	Appointment.reasonReference	0..*	I
	Appointment.reasonCode	0..*	I
Call trauma clinic to coordinate the study, result and **appointment dates**	Appointment.start	0..1	
	Appointment.end	0..1	
Call **trauma clinic** to coordinate the study, result and appointment dates	Appointment.participant	1..*	I
	Appointment.participant.actor	0..1	I
	Appointment.participant.type	0..*	I
	Appointment.participant.status	1..1	

Table 4.7. shows the data foundation, cardinality, and flags for each of the included elements in the Appointment profile.
?! = IsModifier, I = constrain.

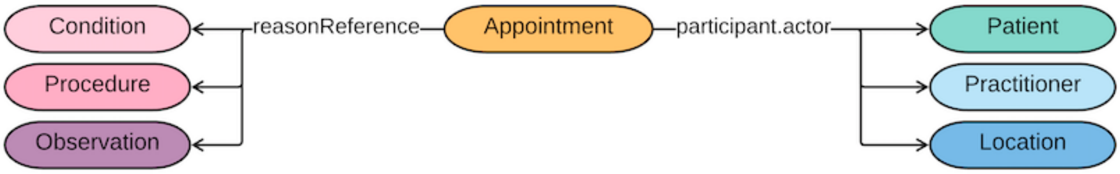

Figure 4.3. shows the references from the Appointment profile. The text on the arrow indicates the element that refers to the other profile.

The Appointment profile was validated using the Firely Terminal and was found valid.

4.2.2.3 CarePlan

The cardinality of the element CarePlan.contributor was changed from 0..* to 0..1, as only one person was identified in the data foundation to contribute to the care plan. The cardinality of the element CarePlan.addresses was changed from 0..* to 0..1, since it was found in the data foundation, that only one issue was addressed for each care plan. The cardinality of the element CarePlan.activity was changed from 0..* to 1..* to indicate that an activity must happen within a care plan. The cardinality of the element CarePlan.activity.detail was changed from 0..1 to 1..1 to indicate that it is important to have details of the activity assigned to a care plan. The cardinality of the element CarePlan.activity.detail.performer was changed from 0..* to 1..1 to indicate that it is important for the situational context to know who is responsible for the activity. Additionally, only one performer was identified for each activity in the data foundation, why the maximum cardinality was changed to 1.
The element CarePlan.subject was restricted to only refer the Patient profile, as no groups were mentioned in the data foundation. The element CarePlan.contributer was restricted to only refer a Patient, CareTeam, or Practitioner profile, as these entities were found to

contribute to the content of the care plan. The element CarePlan.activity.detail.performer was restricted to only refer a Practitioner profile since individual healthcare professionals were responsible for the care plan.
To the element CarePlan.activity slices were added to enable that several activities could be included in the care plan. The path was chosen to be CarePlan.activity.detail.code.
There were two ways to describe the details of an activity, where only the element CarePlan.activity.detail and its sub-elements was chosen to include. This was chosen since the element CarePlan.activity.reference did not capture the details described in the data foundation, e.g. did this element not allow to refer a performer.
The CarePlan profile can be seen in table 4.8, and the elements referring to another profile can be seen in figure 4.4.

Data	Element	Cardinality	Flag
	CarePlan.status	1..1	?!
he **should**	CarePlan.intent	1..1	?!
she will	CarePlan.subject	1..1	
in **4 weeks**	CarePlan.period	0..1	
physical therapy initiated rehabilitation	CarePlan.contributor	0..1	
physical therapy **initiated rehabilitation**	CarePlan.addresses	0..1	
follow up	CarePlan.activity	1..*	I
	CarePlan.activity.detail	1..1	I
	CarePlan.activity.detail.kind	0..1	
	CarePlan.activity.detail.code	0..1	
Avoid intensive exercise	CarePlan.activity.detail. doNotPerform	0..1	?!
Oncology Clinic	CarePlan.activity.detail.location	0..1	
follow up with **Dr. Lin**	CarePlan.activity.detail.performer	1..1	
	CarePlan.activity.detail.status	1..1	?!

Table 4.8. shows the data foundation, cardinality, and flags for each of the included elements in the CarePlan profile.
?! = IsModifier, I = constrain.

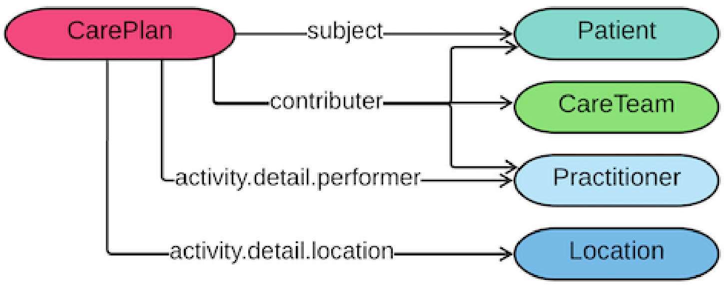

Figure 4.4. shows the references from the CarePlan profile. The text on the arrow indicates the element that refers to the other profile.

The CarePlan profile was validated using the Firely Terminal and was found valid.

4.2.2.4 CareTeam

The element CareTeam.subject was restricted to refer the Patient profile, since the data foundation only concerned individual patients.
The CareTeam profile can be seen in table 4.9, and the elements referring to another profile can be seen in figure 4.5.

Data	Element	Cardinality	Flag
	CareTeam.status	0..1	?!
physical therapy approved her for transfer	CareTeam.name	0..1	
physical therapy approved **her** for transfer	CareTeam.subject	0..1	
she attended exercise **multiple times**	CareTeam.period	0..1	

Table 4.9. shows the data foundation, cardinality, and flags for each of the included elements in the CareTeam profile.
?! = IsModifier.

Figure 4.5. shows the reference from the CareTeam profile. The text on the arrow indicates the element that refers to the other profile.

The CareTeam profile was validated using the Firely Terminal and was found valid.

4.2.2.5 Condition

The cardinality of the element Condition.code was changed from 0..1 to 1..1 to indicate the importance of encoding a condition in order to avoid ambiguities. The cardinality of the element Condition.bodysite was changed from 0..* to 0..1, as only one body site was found for each condition in the data foundation. The cardinality of the elements Condition.evidence, Condition.evidence.code, and Condition.evidence.detail were all changed from 0..* to 0..1 based on the data foundation.
The element Condition.subject was restricted to only refer a Patient profile, since no reference to a group was found in the data foundation. The element Condition.recorder was restricted to only refer a Practitioner profile, since it was found irrelevant to refer a PractitionerRole, Patient, or RelatedPerson profile in the data foundation. The element Condition.evidence.detail was restricted to refer a Procedure profile as this was covered by the data foundation.
The element Condition.code was bound to the ValueSet 'Condition/Problem/Diagnosis Codes' (http://hl7.org/fhir/ValueSet/condition-code). The element Condition.bodySite was bound to the ValueSet 'SNOMED CT Body Structures' (http://hl7.org/fhir/ValueSet/body-site). The element Condition.evidence.code was bound to the ValueSet 'Manifestation and Symptom Codes' (http://hl7.org/fhir/ValueSet/manifestation-or-symptom).
The Condition profile can be seen in table 4.10, and the elements referring to another profile can be seen in figure 4.6 on the following page.

Data	Element	Cardinality	Flag
	Condition.clinicalStatus	0..1	?! I
	Condition.verificationStatus	0..1	?! I
He had acute cardiac arrest	Condition.subject	1..1	
Fanny CR Dennis	Condition.recorder	0..1	
diagnosed **myocardial infarction**	Condition.code	1..1	
found in the **apex of heart**	Condition.bodySite	0..1	
status post	Condition.onset[x]	0..1	
by ultra sound scan	Condition.evidence	0..1	I
	Condition.evidence.code	0..1	I
	Condition.evidence.detail	0..1	I

Table 4.10. shows the data foundation, cardinality, and flags for each of the included elements in the Condition profile.
?! = IsModifier, I = constrain.

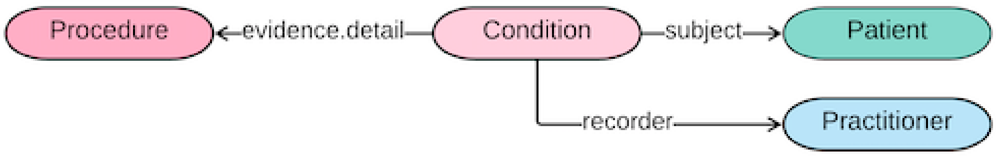

Figure 4.6. shows the references from the Condition profile. The text on the arrow indicates the element that refers to the other profile.

The Condition profile was validated using the Firely Terminal and was found valid.

4.2.2.6 Device

The cardinality of the element Device.deviceName was changed from 0..* to 1..1 to indicate that the name of the device is important as it ensures information regarding which device was used. This was done under the assumption that unique names were used for each device. In the data foundation each device was given one name, why the maximum cardinality was restricted to 1.
The Device profile can be seen in table 4.11, and the elements referring to another profile can be seen in figure 4.7 on the following page.

Data	Element	Cardinality	Flag
patient **had** ostomy bag	Device.status	0..1	?!
Her **pacemaker**	Device.deviceName	1..1	
	Device.deviceName.name	1..1	
	Device.deviceName.type	1..1	
Her pacemaker	Device.patient	0..1	

Table 4.11. shows the data foundation, cardinality, and flags for each of the included elements in the Device profile.
?! = IsModifier.

Figure 4.7. shows the reference from the Device profile. The text on the arrow indicates the element that refers to the other profile.

The Device profile was validated using the Firely Terminal and was found valid.

4.2.2.7 Encounter

The cardinality of the element Encounter.subject was changed from 0..1 to 1..1 to indicate that every encounter must be associated with a subject.
The element Encounter.subject was restricted to only refer a Patient profile, as no group was mentioned in the data foundation. The element Encounter.reasonReference was restricted to only refer a Condition, Observation, and Procedure profile.
The element Encounter.serviceType was bound to the ValueSet 'Service Type' (http://hl7.org/fhir/ValueSet/service-type). The element Encounter.hospitalization.dischargeDisposition was bound to the ValueSet 'Discharge Disposition' (http://hl7.org/fhir/ValueSet/encounter-discharge-disposition).
The elements Encounter.reasonReference and Encounter.reasonCode both concern reasons for having an encounter. The rule described in section 4.1.4.3 on page 14 was added to indicate that either one of the elements or both must be present.
The element Encounter.location was sliced with the element Encounter.location.location.identifier.value as the path, as patients were transferred between different locations during their hospitalization. The elements Encounter.reasonReference and Encounter.reasonCode were both sliced, as multiple reasons for having an encounter were identified. The elements Encounter.reasonReference.identifier.value and Encounter.reasonCode.coding were used as paths, respectively. Lastly, the element Encounter.appointment was sliced, as it was found that multiple appointments could be referred by one encounter. Encounter.appointment.identifier.value was used as path.
The Encounter profile can be seen in table 4.12 on the following page, and the elements referring to another profile can be seen in figure 4.8.

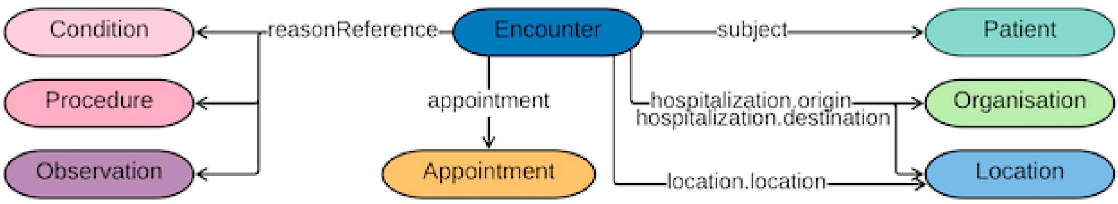

Figure 4.8. shows the references from the Encounter profile. The text on the arrow indicates the element that refers to the other profile.

Data	Element	Cardinality	Flag
	Encounter.status	1..1	?1
	Encounter.class	1..1	
Admission Date: **Discharge** Date:	Encounter.period	0..1	
patient was admitted	Encounter.subject	1..1	
Followup Instructions:	Encounter.serviceType	0..1	
Call **Oncology Clinic** in two weeks with side effects	Encounter.location	0..*	
	Encounter.location.location	1..1	
Call Oncology Clinic **in two weeks** with side effects	Encounter.location.period	0..1	
	Encounter.location.status	0..1	
Call Oncology Clinic in two weeks with **side effects**	Encounter.reasonReference Encounter.reasonCode	0..* 0..*	I I
Call Oncology Clinic	Encounter.appointment	0..*	
DISCHARGE STATUS: **To hospice**	Encounter.hospitalization	0..1	
	Encounter.hospitalization. dischargeDisposition	0..1	
transferred to the **intensive care unit**	Encounter.hospitalization.destination	0..1	
transfer from intensive care	Encounter.hospitalization.origin	0..1	
transferred to	Encounter.hospitalization.admitSource	0..1	

Table 4.12. shows the data foundation, cardinality, and flags for each of the included elements in the Encounter profile.
?! = IsModifier, I = constrain.

The Encounter profile was validated using the Firely Terminal and was found valid.

4.2.2.8 FamilyMemberHistory

The cardinality of the element FamilyMemberHistory.condition was changed from 0..* to 1..1 to indicate that a condition for the given family member must be noted. Only one condition was identified for a family member, why the maximum cardinality was restricted to 1.
The element FamiliMemberHistory.condition.code was bound to the ValueSet 'Condition/Problem/Diagnosis Codes' (http://hl7.org/fhir/ValueSet/condition-code). In the data foundation, it was described that a family member had no present diagnoses. Therefore, the ValueSet was extended with the code '103330002 | No diagnosis |' from the code system 'http://snomed.info/sct'. Lastly, the element FamilyMemberHistory.relationship was bound to the ValueSet 'V3 Value SetFamilyMember' (http://terminology.hl7.org/ValueSet/v3-FamilyMember).
The FamilyMemberHistory profile can be seen in table 4.13, and the elements referring to another profile can be seen in figure 4.9 on the following page.

Data	Element	Cardinality	Flag
	FamilyMemberHistory.status	1..1	?!
FAMILY HISTORY: Presence of **diabetes**	FamilyMemberHistory.condition	1..1	
	FamilyMemberHistory.condition.code	1..1	
	FamilyMemberHistory.relationship	1..1	
	FamilyMemberHistory.patient	1..1	

Table 4.13. shows the data foundation, cardinality, and flags for each of the included elements in the FamilyMemberHistory profile.
?! = IsModifier.

Figure 4.9. shows the reference from the FamilyMemberHistory profile. The text on the arrow indicates the element that refers to the other profile.

The FamilyMemberHistory profile was validated using the Firely Terminal and was found valid.

4.2.2.9 Location

The cardinality of the element Location.name was changed from 0..1 to 1..1 to indicate that a name of the location should be noted. By indicating this, it is clear where the patient is e.g. transferred to.
The Location profile can be seen in table 4.14, and since no references was included in the profile, no figure was presented.

Data	Element	Cardinality	Flag
	Location.status	0..1	?!
transferred to **intensive care unit**	Location.name	1..1	

Table 4.14. shows the data foundation, cardinality, and flags for each of the included elements in the Location profile.
?! = IsModifier.

The Location profile was validated using the Firely Terminal and was found valid.

4.2.2.10 Medication

The cardinality of the element Medication.ingredient was changed from 0..* to 1..1, to indicate that at least one ingredient must be noted for each medication. Since only one ingredient was identified for each medication in the data foundation, the maximum cardinality was restricted to 1.
The element Medication.ingredient.itemReference was restricted to refer a Medication profile, as there was found no need in the data foundation for a reference to a Substance profile.
The element Medication.form was bound to the ValueSet 'SNOMED CT Form Codes'

(http://hl7.org/fhir/ValueSet/medication-form-codes).
The Medication profile can be seen in table 4.15, and the elements referring to another profile can be seen in figure 4.10.

Data	Element	Cardinality	Flag
	Medication.status	0..1	?!
15 mg	Medication.amount	0..1	
given **IV**	Medication.form	0..1	
Acetylsalicylic acid	Medication.ingredient	1..1	
	Medication.ingredient.item[x]	1..1	

Table 4.15. shows the data foundation, cardinality, and flags for each of the included elements in the Medication profile.
?! = IsModifier.

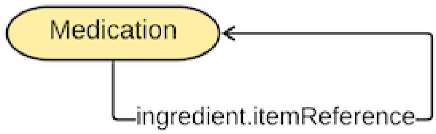

Figure 4.10. shows the reference from the Medication profile. The text on the arrow indicates the element that refers to the other profile.

The Medication profile was validated using the Firely Terminal and was found valid.

4.2.2.11 MedicationAdministration

The cardinality of the elements MedicationAdministration.reasonReference and MedicationAdministration.reasonCode was changed from 0..* to 0..1 as the data foundation only mentioned one reason for administering medication.
The element MedicationAdministration.subject was restricted to only refer a Patient profile, as the data foundation concern individual patients. The element MedicationAdministration.reasonReference was restricted to only refer a Observation or Condition profile, as this was sufficient to support the data foundation.
The element MedicationAdministration.medication[x] with the data type CodeableConcept was bound to the ValueSet 'SNOMED CT Medication Codes' (http://hl7.org/fhir/ValueSet/medication-codes). The element MedicationAdministration.reasonCode was bound to the ValueSet 'SNOMED CT Medication As Needed Reason Codes' (Value Set http://hl7.org/fhir/ValueSet/medication-as-needed-reason) instead of the suggested ValueSet 'Reason Medication Given Codes' (http://hl7.org/fhir/ValueSet/reason-medication-given-codes), since it was more suitable for the information in the data foundation. The element MedicationAdministration.dosage.route was bound to the ValueSet 'SNOMED CT Route Codes' (Value Set http://hl7.org/fhir/ValueSet/route-codes).
The elements MedicationAdministration.reasonReference and MedicationAdministration.reasonCode both concern a reason for administering medication to a patient. The rule described in section 4.1.4.3 on page 14 was added to indicate that either one of the elements or both must be present.

The MedicationAdministration profile can be seen in table 4.16, and the elements referring to another profile can be seen in figure 4.11.

Data	Element	Cardinality	Flag
	MedicationAdministration.status	1..1	?!
She took daily **insulin**	MedicationAdministration. medication[x]	1..1	
She took daily insulin	MedicationAdministration.subject	1..1	
for **10 days**	MedicationAdministration.effective[x]	1..1	
as needed for knee-pain	MedicationAdministration. reasonReference	0..1	I
	MedicationAdministration. reasonCode	0..1	I
One (1) tablet PC	MedicationAdministration.dosage	0..1	I
	MedicationAdministration.dosage. dose	0..1	
oral medication	MedicationAdministration.dosage. route	0..1	

Table 4.16. shows the data foundation, cardinality, and flags for each of the included elements in the MedicationAdministration profile.
?! = IsModifier, I = constrain.

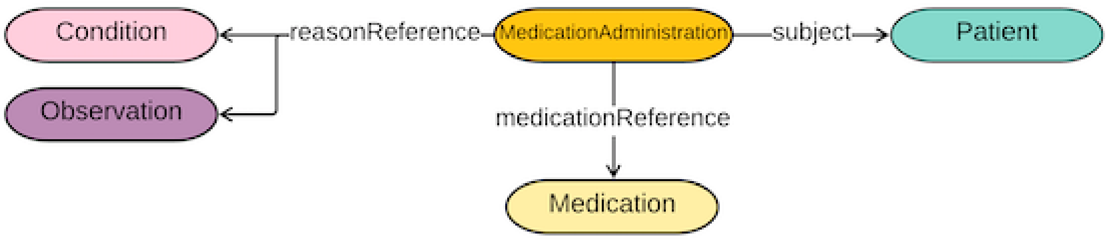

Figure 4.11. shows the references from the MedicationAdministration profile. The text on the arrow indicates the element that refers to the other profile.

The MedicationAdministration profile was validated using the Firely Terminal and was found valid.

4.2.2.12 MedicationDispense

The cardinality of the element MedicationDispense.dosageInstruction was changed from 0..* to 0..1, since there was only identified one instruction per medication in the data foundation.
The MedicationDispense profile can be seen in table 4.17, and the elements referring to another profile can be seen in figure 4.12 on the next page.

Data	Element	Cardinality	Flag
	MedicationDispense.status	1..1	?!
	MedicationDispense.medication[x]	1..1	
Disp :* **5 tablets** *	MedicationDispense.quantity	0..1	
One (1) tablet **PC**	MedicationDispense.dosageInstruction	0..1	

Table 4.17. shows the data foundation, cardinality, and flags for each of the included elements in the MedicationDispense profile.
?! = IsModifier.

Figure 4.12. shows the reference from the MedicationDispense profile. The text on the arrow indicates the element that refers to the other profile.

The MedicationDispense profile was validated using the Firely Terminal and was found valid.

4.2.2.13 Observation

The cardinality of the element Observation.subject was changed from 0..1 to 1..1 to indicate that a subject must be referred when documenting an observation. The cardinality of the element Observation.category was changed from 0..* to 0..1, since it was found appropriate to bind one category to each observation. The cardinality of the elements Observation.interpretation and Observation.performer was changed from 0..* to 0..1, as only one interpretation and one performer was found for each observation, respectively.
The element Observation.subject was restricted to only refer a Patient profile due to the data foundation. The element Observation.performer was restricted to only refer a Practitioner and RelatedPerson profile, since observations were performed by healthcare professionals and related persons in the data foundation.
The element Observation.code was bound to the ValueSet 'LOINC Codes' (http://hl7.org/fhir/ValueSet/observation-codes). The element Observation.bodySite was bound to the ValueSet 'SNOMED CT Body Structures' (http://hl7.org/fhir/ValueSet/body-site). The element Observation.interpretation was bound to the ValueSet 'Observation Interpretation Codes' (http://hl7.org/fhir/ValueSet/observation-interpretation), which was extended with a code to indicate a patient was stable. The added code was '58158008 | Stable |' from the code system 'http://snomed.info/sct'.
The rule attached to the elements Observation.value[x] and Observation.dataAbsentReason indicates that if no value is present, then a reason for the absence shall be given in the element Observation.dataAbsentReason, why this element was included.
The Observation profile can be seen in table 4.18, and the elements referring to another profile can be seen in figure 4.13 on the following page.

Data	Element	Cardinality	Flag
	Observation.status	1..1	?!
He was awake	Observation.subject	1..1	
Left arm in cast	Observation.bodySite	0..1	
rise in **HbA1c values**	Observation.category	0..1	
	Observation.code	1..1	
Glucose: **15** mmol/L	Observation.value[x]	0..1	I
her glucose was **stable**	Observation.interpretation	0..1	
history of hypoglycemia	Observation.effective[x]	0..1	
her **husband** reported dizziness and paleness	Observation.performer	0..1	
	Observation.dataAbsentReason	0..1	I

Table 4.18. shows the data foundation, cardinality, and flags for each of the included elements in the Observation profile.
?! = IsModifier, I = constrain.

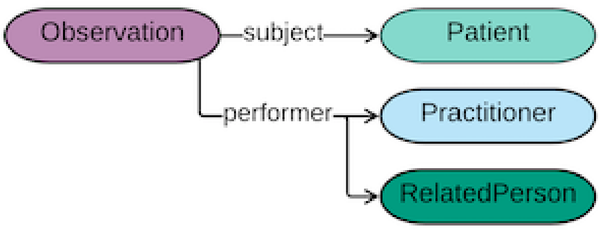

Figure 4.13. shows the references from the Observation profile. The text on the arrow indicates the element that refers to the other profile.

The Observation profile was validated using the Firely Terminal and was found valid.

4.2.2.14 Organization

The cardinality of element Organization.type was changed from 0..* to 0..1, as only one type was assigned to each organization in the data foundation.
The element Organization.type was bound to the ValueSet 'Organization Type' (http://hl7.org/fhir/ValueSet/organization-type).
Due to a rule stating that the Organization profile should at least include an identifier or a name, the elements Organization.name and Organization.identifier were included.
The element Organization.identifier was sliced, using Organization.identifier.value as path.
The Organization profile can be seen in table 4.19 on the next page, and since no references was included in the profile, no figure was presented.

Data	Element	Cardinality	Flag
	Organization.active	0..1	?!
discharged from **hospital**	Organization.type	0..1	
	Organization.identifier	0..*	I
	Organization.name	0..1	I

Table 4.19. shows the data foundation, cardinality, and flags for each of the included elements in the Organization profile.
?! = IsModifier, I = constrain.

The Organization profile was validated using the Firely Terminal and was found valid.

4.2.2.15 Patient

The element Patient.link.other was restricted to only include a reference to a Patient profile, due to the data foundation. The element Patient.link was sliced using the element Patient.link.other.identifier.value as path.
The Patient profile can be seen in table 4.20, and the elements referring to another profile can be seen in figure 4.14.

Data	Element	Cardinality	Flag
	Patient.active	0..1	?!
Sex	Patient.gender	0..1	
Date of Birth	Patient.birthDate	0..1	
	Patient.deceased[x]	0..1	?!
	Patient.link	0..*	?!
	Patient.link.other	1..1	
	Patient.link.type	1..1	

Table 4.20. shows the data foundation, cardinality, and flags for each of the included elements in the Patient profile.
?! = IsModifier.

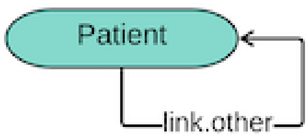

Figure 4.14. shows the reference from the Patient profile. The text on the arrow indicates the element that refers to the other profile.

The Patient profile was validated using the Firely Terminal and was found valid.

4.2.2.16 Practitioner

The cardinality of the element Practitioner.identifier was changed from 0..* to 1..1, to indicate that it should be possible to identify a practitioner. Since only one identifier was presented in the data foundation, the maximum cardinality was changed to 1. The cardinality of the element Practitioner.name was changed from 0..* to 0..1, since each

practitioner had one name. The cardinality of the element Practitioner.qualification was changed from 0..* to 1..1 to indicate that each practitioner must be described with one qualification. The maximum cardinality was changed to 1, since each healthcare professional was associated with one qualification in the data foundation.
The element Practitioner.qualification.code was bound to the ValueSet 'v2 table 0360, Version 2.7' (http://terminology.hl7.org/ValueSet/v2-2.7-0360).
The Practitioner profile can be seen in table 4.21, and since no references was included in the profile, no figure was presented.

Data	Element	Cardinality	Flag
73-984	Practitioner.identifier	1..1	
Leonora Colmor	Practitioner.name	0..1	
MD	Practitioner.qualification	1..1	
	Practitioner.qualification.code	1..1	

Table 4.21. shows the data foundation, cardinality, and flags for each of the included elements in the Practitioner profile.
?! = IsModifier

The Practitioner profile was validated using the Firely Terminal and was found valid.

4.2.2.17 Procedure

The cardinality of the element Procedure.code was changed from 0..1 to 1..1 to indicate that each procedure must be given a code, describing which procedure was or is about to be performed. The cardinality of the elements Prodcedure.reasonCode and Procedure.reasonReference was changed from 0..* to 0..1, since there was identified one reason for each procedure. The cardinality of the element Procedure.performer was changed from 0..* to 0..1, as one performer performed a procedure in the data foundation. The cardinality of the element Procedure.bodySite was changed from 0..* to 0..1, as only one body site was found for each procedure in the data foundation.
The element Procedure.subject was restricted to only refer a Patient profile based on the data foundation. The element Procedure.reasonReference was restricted to only refer a Condition profile. The element Procedure.performer.actor was restricted to only refer a Practitioner and RelatedPerson profile.
The element Procedure.category was bound to the ValueSet 'Procedure Category Codes (SNOMED CT)' (http://hl7.org/fhir/ValueSet/procedure-category) and the element Procedure.code was bound to the ValueSet 'Procedure Codes (SNOMED CT)' (http://hl7.org/fhir/ValueSet/procedure-code). The element Procedure.bodySite was bound to the ValueSet 'SNOMED CT Body Structures' (http://hl7.org/fhir/ValueSet/body-site).
The elements Procedure.reasonCode and Procedure.reasonReference can both contain a reason for having a procedure. The rule described in section 4.1.4.3 on page 14 was added as it was important to know the reason for the procedure.
The element Procedure.complicatedDetail was sliced to enable one or more conditions to follow a procedure, with the path Procedure.complicatedDetail.identifier.value. The element Procedure.focalDevice was sliced, since information from the data foundation showed that several devices was used during the same procedure. The path was Procedure.focalDevice.manipulated.identifier.value.

The Procedure profile can be seen in table 4.22, and the elements referring to another profile can be seen in figure 4.15.

Data	Element	Cardinality	Flag
medication was stopped **after surgery**	Procedure.status	1..1	?!
coronary artery bypass on 2016-07-24	Procedure.category	0..1	
	Procedure.code	1..1	
son gave **her** cardiac massage	Procedure.subject	1..1	
acute myocardial infarction	Procedure.reasonCode	0..1	I
	Procedure.reasonReference	0..1	I
coronary artery bypass on **2016-07-24**	Procedure.performed[x]	0..1	
Dr. William Osler performed the procedure.	Procedure.performer	0..1	
	Procedure.performer.actor	1..1	
Abdomen clear for fluids	Procedure.bodySite	0..1	
diagnosed coronary artery disease by ultra sound	Procedure.complicatedDetail	0..*	
temporary pacemaker was **removed** successfully	Procedure.focalDevice	0..*	
	Procedure.focalDevice.action	0..1	
temporary pacemaker was removed successfully	Procedure.focalDevice.manipulated	1..1	

Table 4.22. shows the data foundation, cardinality, and flags for each of the included elements in the Procedure profile.
?! = IsModifier, I = constrain

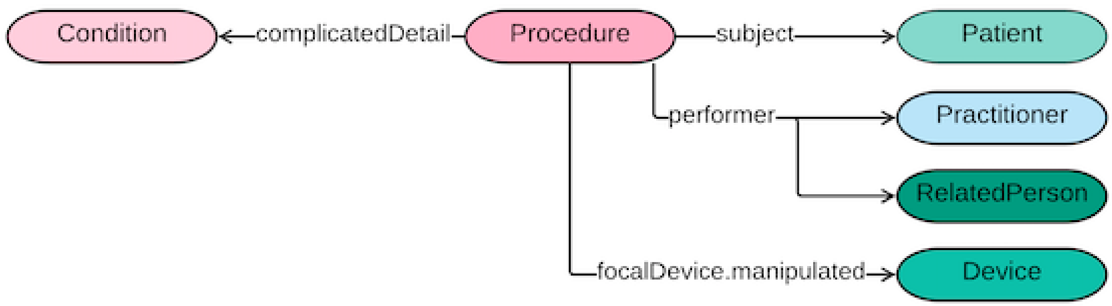

Figure 4.15. shows the references from the Procedure profile. The text on the arrow indicates the element that refers to the other profile.

The Procedure profile was validated using the Firely Terminal and was found valid.

4.2.2.18 RelatedPerson

The cardinality of the element RelatedPerson.relationship was changed from 0..* to 0..1, as a relationship between two people found in the data foundation, could be described with one relation.

The RelatedPerson profile can be seen in table 4.23, and the elements referring to another profile can be seen in figure 4.16.

Data	Element	Cardinality	Flag
	RelatedPerson.active	0..1	!?
His wife	RelatedPerson.patient	1..1	
His wife	RelatedPerson.relationship	0..1	
	RelatedPerson.gender	0..1	

Table 4.23. shows the data foundation, cardinality, and flags for each of the included elements in the RelatedPerson profile.
?! = IsModifier

Figure 4.16. shows the reference from the RelatedPerson profile. The text on the arrow indicates the element that refers to the other profile.

The RelatedPerson profile was validated using the Firely Terminal and was found valid.

4.2.2.19 Implementation

Bundles

The cardinality of the elements Bundle.identifier and Bundle.entry.resource was changed from 0..1 to 1..1, since the Bundle.type is a 'document', why an identifier and the included resources must be present.
The element Bundle.resource was restricted to refer a Composition profile, since the Bundle type is a 'document'.
The Bundle profile can be seen in table 4.24, and the elements referring to another profile can be seen in figure 4.17.

Data	Element	Cardinality	Flag
	Bundle.identifier	1..1	I
	Bundle.type	1..1	
TUE 2016-04-12 7:03	Bundle.timestamp	0..1	
	Bundle.entry	0..1	I
	Bundle.entry.resource	1..1	
Signed electronically by : DR. Jose Kines	Bundle.signature	0..1	

Table 4.24. shows the data foundation, cardinality, and flags for each of the included elements in the Bundle profile.
?! = IsModifier, I = constrain

Figure 4.17. shows the reference from the Bundle profile. The text on the arrow indicates the element that refers to the other profile.

The element Composition.subject was restricted to only refer a Patient profile, since the discharge summaries in the data foundation concerned one patient. The elements Composition.event.detail, Composition.section.focus, and Composition.section.entry were restricted to refer the FHIR resources identified to be necessary since they represent the data foundation, see table 4.5 on page 19. The elements Composition.attester.party, Composition.author, and Composition.section.author were restricted to only refer a Practitioner profile, since practitioners had authored the discharge summaries.
The element Composition.event was sliced using Composition.event.detail.identifier.value as the path. The element Composition.event.detail was sliced to allow several events to be documented under the same service using Composition.event.detail.identifier.use as path. The element Composition.section was sliced to allow division of the composition into multiple section, and Composition.section.entry was sliced to allow multiple references to support the given section. The paths were Composition.section.entry.identifier.value, and Composition.section.entry.identifier.use, respectively. The elements Composition.attester, Composition.author, and Composition.section.author contained information about who were responsible for the discharge summary, using the paths Composition.attester.party. identifier.value, Composition.author.identifier.value, and Composition.section.author.identifier.value, respectively.
The Composition profile can be seen in table 4.25, and the elements referring to another profile can be seen in figure 4.18 on the following page.

Mandatory	**Element**	**Cardinality**	**Flag**
Yes	Composition.identifier	0..1	
No	Composition.status	1..1	?!
No	Composition.type	1..1	
Yes	Composition.subject	0..1	
Yes	Composition.encounter	0..1	
No	Composition.date	1..1	
Yes	Composition.author	1..*	
No	Composition.title	1..1	
No	Composition.attester	0..*	
No	Composition.attester.mode	1..1	
Yes	Composition.attester.party	0..1	
Yes	Composition.custodian	0..1	
No	Composition.event	0..*	
Yes	Composition.event.detail	0..*	
No	Composition.section	0..*	I
Yes	Composition.section.author	0..*	
Yes	Composition.section.focus	0..1	
Yes	Composition.section.entry	0..*	I

Table 4.25. shows if the element is mandatory since the Bundle type is 'document' [Health Level 7, 2021d], the cardinality, and flags for each of the included elements in the Composition profile.
?! = IsModifier, I = constrain

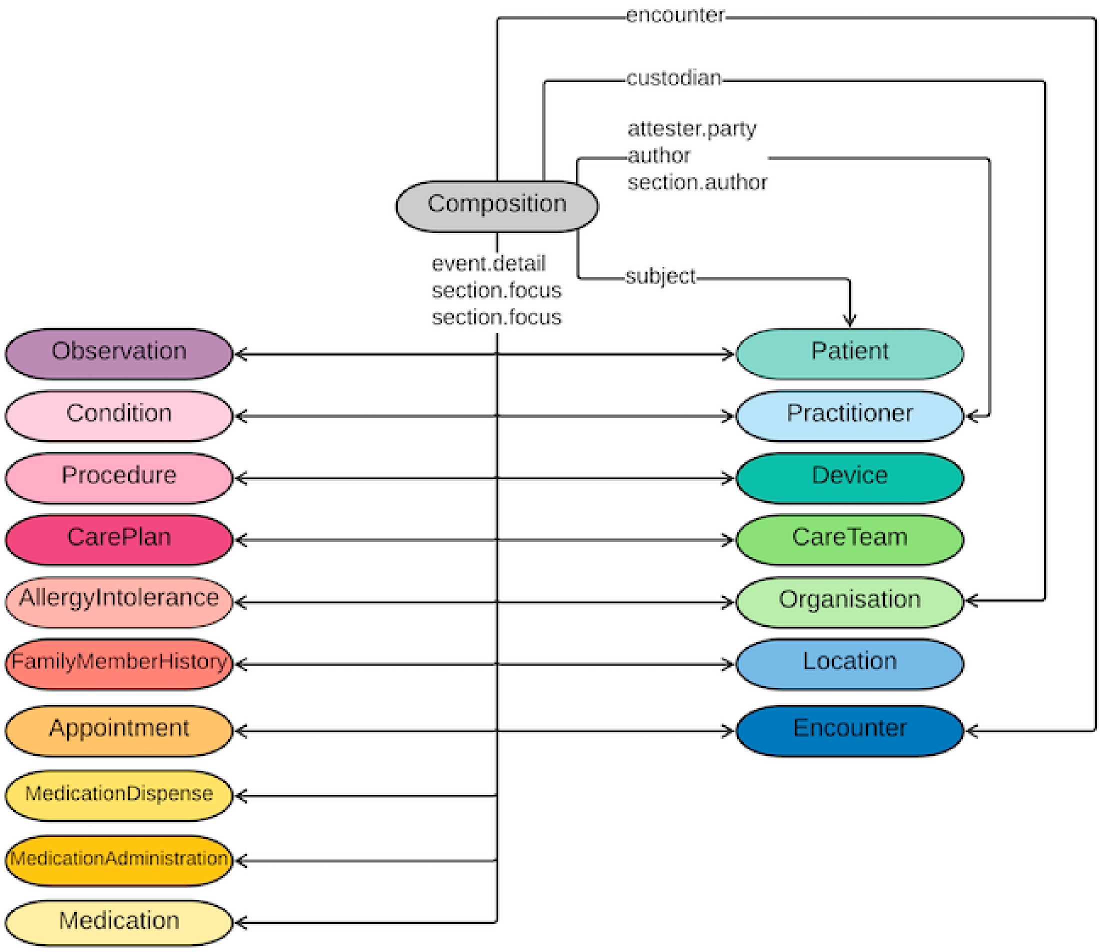

Figure 4.18. shows the references from the Composition profile. The text on the arrow indicates the element that refers to the other profile.

Both the Bundle and Composition profiles were validated using the Firely Terminal and were found valid.

ImplementationGuide

The cardinality of the element ImplementationGuide.fhirVersion was changed from 1..* to 1..1 as only one version of HL7 FHIR was used for the developed profiles. The cardinality of the element ImplementationGuide.useContext was changed from 0..* to 0..1, under the assumption that one implementation guide was used in one context e.g. one hospital department.
The element ImplementationGuide.definition.resource.reference was restricted to refer the FHIR resources identified to be necessary in table 4.5 on page 19, in order to accommodate the context and usage of this implementation guide.
To the element ImplementationGuide.name there was a warning, which concern that the name should be usable as an identifier to be used by machine processing application.
Slicing was used for the element ImplementationGuide.definition.resource to allow all identified profiles to be included in the implementation guide. The path ImplementationGuide.definition.resources.reference.identifier.value was used.
The ImplementationGuide profile can be seen in table 4.26 and the elements referring to another profile can be seen in figure 4.19 on the following page.

Concern	Element	Cardinality	Flag
Unique Identifier	ImplementationGuide.url	1..1	
Machine readable title	ImplementationGuide.name	1..1	I
Human readable title	ImplementationGuide.title	0..1	
Status of the ImplementationGuide	ImplementationGuide.status	1..1	?!
Described the developers	ImplementationGuide.publisher	0..1	
Usage of the ImplementationGuide	ImplementationGuide.description	0..1	
	ImplementationGuide.useContext	0..1	
	ImplementationGuide.packageId	1..1	
	ImplementationGuide.fhirVersion	1..1	
Profiles supported by the ImplementationGuide	ImplementationGuide.definition	0..1	I
	ImplementationGuide.definition. resources	1..*	
	ImplementationGuide.definition. resources.reference	1..1	

Table 4.26. shows the which field the element concern, cardinality, and flags for each of the included elements in the ImplementationGuide profile.
?! = IsModifier, I = constrain

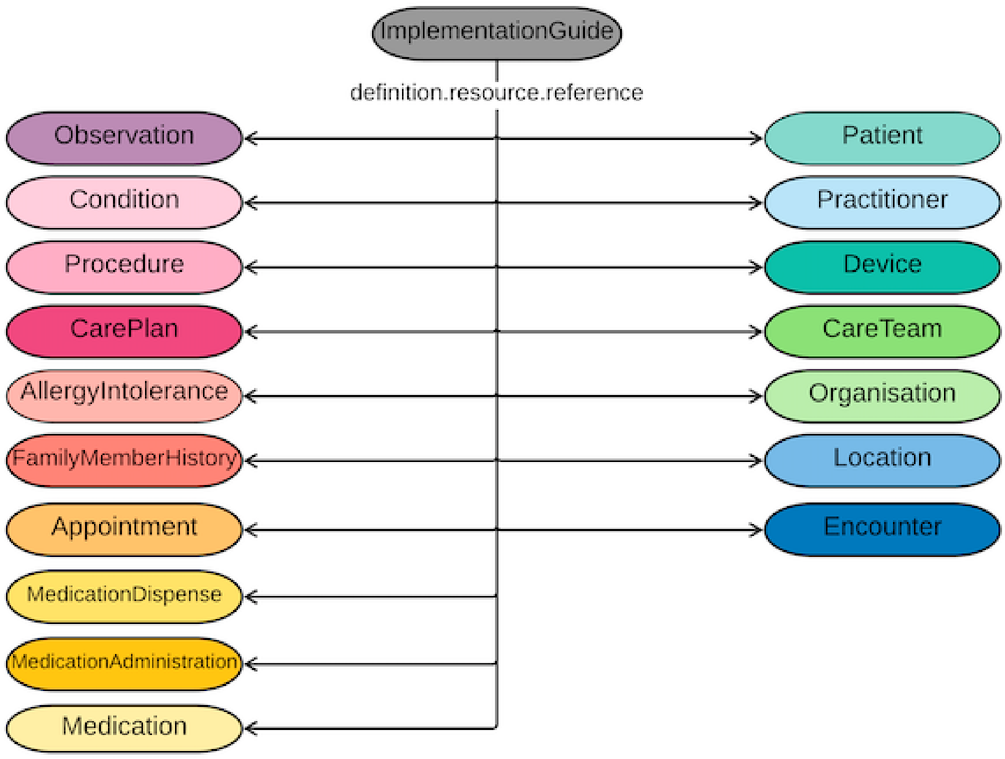

Figure 4.19. shows the reference from the ImplementationGuide profile. The text on the arrow indicates the element that refers to the other profile.

The ImplementationGuide profile was validated using the Firely Terminal and was found valid.

4.2.3 References between FHIR Profiles

Based on the selected elements in the FHIR profiles described above, figure 4.20 on the next page shows the references between all the developed FHIR profiles. The figure illustrates that when including the situational contextual information from discharge summaries and structuring it using FHIR profiles, a lot of information lies within the dependencies between the FHIR profiles.

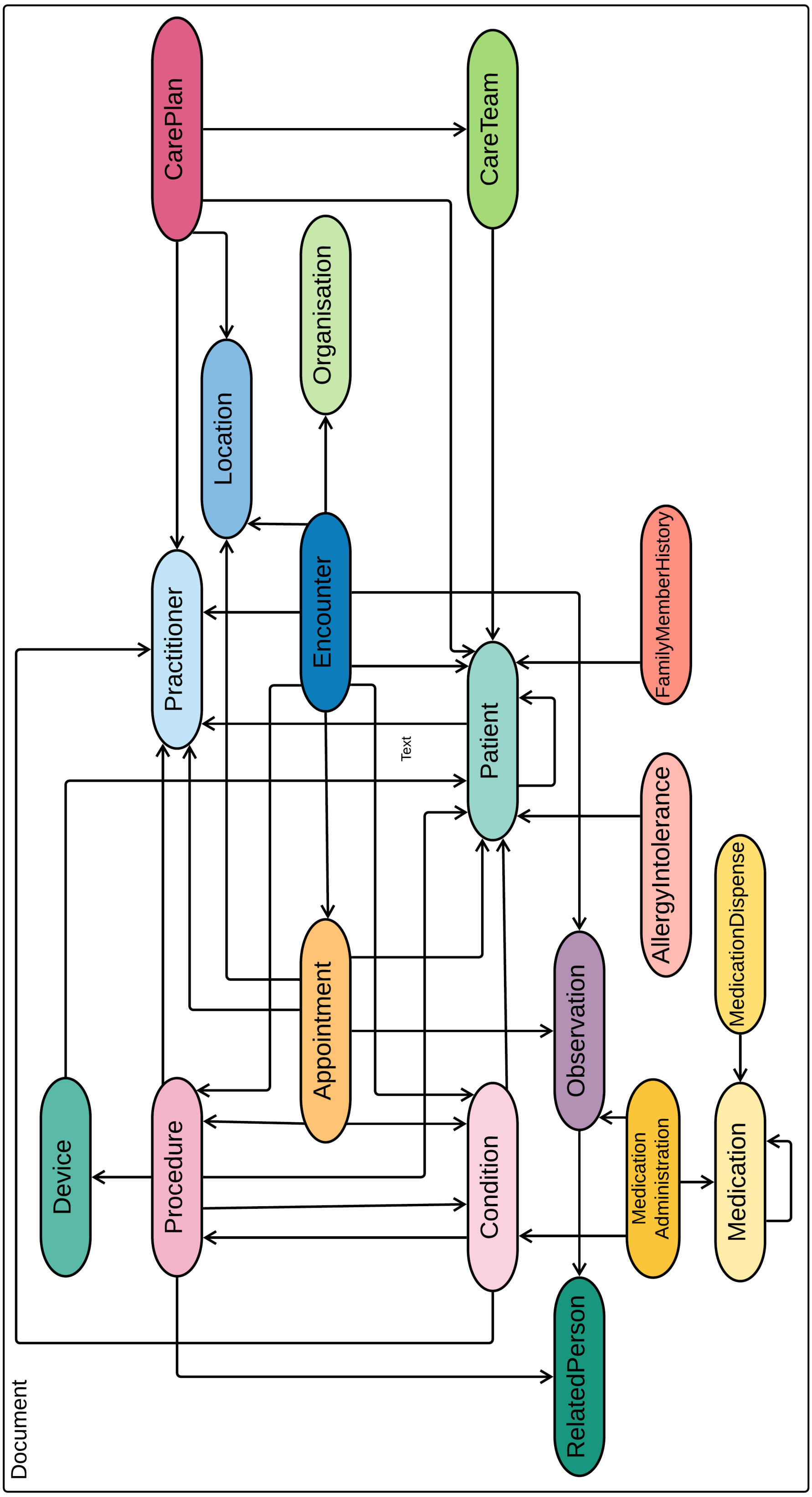

Figure 4.20. shows the references between all the developed FHIR profiles, based on information from discharge summary number 51, 53, and 74. The arrow points towards the profiles being referred e.g. an element in the Device profile refers the Patient profile.

Identifying Existing NLP Systems 5

5.1 Method

As mentioned in section 2.3 on page 3, studies show that NLP can be used to extract relevant information from free text documents. To identify existing NLP systems a structured literature search was conducted.

5.1.1 Structured Literature Search

Through an initial exploratory literature search the article by Meystre et al. [2008] was identified. This article thoroughly surveyed the current status of NLP systems used on free text documents from 1995 to 2008 Leaman et al. [2015]; Meystre et al. [2008]. Therefore, the current structured literature search includes the article by Meystre et al. [2008] and limits the search period to include articles from 1/1/2008 to 19/2/2021.
A block search was conducted in the databases PubMed and Embase, as there is a retrieval rate of 92.8% when searching in these two databases according to the article by [Bramer et al., 2017]. In table 5.1 on the next page and 5.2 on page 43 the search terms for PubMed and Embase can be seen, respectively. The three columns were combined using the boolean operator 'AND', and the rows in each column were combined using the boolean operator 'OR'. As it can be seen in the tables, free text words in the columns 'Clinical Notes' and 'NLP' must appear in the title or abstract of the article, where the MeSH and Emtree terms must appear as a major topic. This was chosen to ensure the relevance of the articles. Search terms in the column 'Technical NLP Terms' may appear in all fields which was valid for both the free text words, MeSH terms and Emtree terms. The search terms in this column was found through the initial exploratory search of NLP terms, methods, and sub-elements.

Clinical Notes	NLP	Technical NLP Terms
MeSH as Major Topic		**MeSH**
"Narrative Medicine"[MeSH Major Topic]	"Natural Language Processing"[MeSH Major Topic] "Data Mining"[MeSH Major Topic] NOT "Multifactor Dimensionality Reduction"[MeSH Major Topic]	
Free Text Title/Abstract		**Free Text All fields**
Clinical document*	Data mining	Annotat*
Clinical narrative*	Information extraction	Concept extract*
Clinic* note*	Medical language processing	Concept Mapp*
Free text	Natural language processing	Concept Unique Identifier
Medical document*	Natural language understanding	Dependency pars*
Medical narrative*	Text mining	Encoder
Medical note*		Lemmatization
		Named Entity Recognition
		Negation
		Relation type select*
		Section detector
		Segmentation
		Semantic pars*
		Sentence detector
		Sentence Splitt*
		Stemming
		Syntactic pars*
		Tokeniz*

Table 5.1. shows the search terms for the block search conducted in PubMed. The free text in the columns 'Clinical Notes' and 'NLP' must appear in the title and/or abstract. The free text in the column 'Technical NLP Terms' may appear in all fields. '[MeSH Major Topic]' indicates that the MeSH term must appear as a major topic.

Clinical Notes	**NLP**	**Technical NLP Terms**
Emtrees as Major Topic		**Emtree**
‘Medical documentation’/exp/mj	’Natural language processing’/exp/mj	’Syntactic processing’/exp
	’Data mining’/exp/mj	’Negation’/exp
	‘Data extraction’/exp/mj	’Named entity recognition’/exp
Free Text Title/Abstract		**Free Text All fields**
Clinical document*	Data mining	Annotat*
Clinical narrative*	Information extraction	Concept extract*
Clinic* note*	Medical language processing	Concept Mapp*
Free text	Natural language processing	Concept Unique Identifier
Medical document*	Natural language understanding	Dependency pars*
Medical narrative*	Text mining	Encoder
Medical note*		Lemmatization
		Named Entity Recognition
		Negation
		Relation type select*
		Section detector
		Segmentation
		Semantic pars*
		Sentence detector
		Sentence Splitt*
		Stemming
		Syntactic pars*
		Tokeniz*

Table 5.2. shows the search terms for the block search conducted in Embase. The free text in the columns ’Clinical Notes’ and ’NLP’ must appear in the title and/or abstract. The free text in the column ’Technical NLP Terms’ may appear in all fields. ’/exp’ indicates that the Emtree includes all sub-Emtrees. ’/exp/mj’ indicates that the Emtree term must appear as a major topic.

The articles retrieved from the two databases were combined and duplicates were removed. The remaining articles were initially screened based on title and abstract, followed by a reading of the articles in full length. For the articles to be included they had to fulfill the following inclusion and exclusion criteria.
Inclusion criterion:

- I1: The article concerned information extraction on electronic free text documents

Exclusion criteria:

- E1: The article used non-human data
- E2: The article presented no arguments for the purpose or influence of the sub-element(s) in the NLP system that performed the information extraction
- E3: The article concerned de-identification of data
- E4: The article used a non-English data set
- E5: The article was not available in full length
- E6: The article was not written in English or Danish

5.1.2 Assessment of the Included Articles

The included articles were further assessed. Two different types of studies were identified among the included articles; systematic reviews and development studies. Systematic reviews were assessed using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [Moher et al., 2009], and based on the assessment the articles were assigned to be low, medium, or high. The development studies were assessed based on three criteria; 1) reliability concerning e.g. the reproducibility of the study, 2) internal validity concerning e.g. the amount of data included in the study, and 3) external validity concerning e.g. the generalizability of the study. Each criterion was assessed to be either low, medium, or high referring to how much the article fulfilled the given criterion. An overall assessment of the article was then made based on the grade from the three criteria. If an article had a low overall assessment the article was excluded.

5.1.3 Selection of NLP system

The selection of a NLP system included; 1) identification of relevant sub-elements in a NLP system, 2) identification of existing systems, and 3) selection of a NLP system based on three formulated criteria.
Initially, identification of relevant sub-elements was based on sub-elements in the included articles e.g. tokenizer or negation detector. The purpose of identifying the sub-elements was to identify the important functions a NLP system must contain. From the 51 included articles at least 15 articles must include a sub-element, to ensure that the sub-element was commonly used. Only sub-elements mentioned in the articles were used in the identification of relevant sub-elements. The sub-elements that fulfilled the above mentioned criteria must be a part of the NLP system used, otherwise the functionality of the sub-element were applied to the source code.
Secondly, in order to select a NLP system an overview was initially given of the NLP systems identified from the included articles. In this overview the main system was presented, the related F-score to compare the performance of the system, and it was noted whether the system was open-source.
Lastly, selection of a NLP system was based on three criteria; 1) the NLP system was open-source, 2) the NLP system must be used in several articles, and 3) the NLP system had the highest possible F-score. A F-score represents the performance of a binary classifier [Gobbel et al., 2014] and was used to compare the performance of the NLP systems in several of the identified articles. In articles where the F-score was not calculated but the presented results that made it possible to calculate the F-score, this calculation was based on the following equation from the article by Gobbel et al. [2014]:

$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5.1}$$

In articles where the NLP system was tested on several data sets, a mean of the F-scores was calculated, using equation 5.2.

$$F - score = \frac{1}{n}\sum_{i=1}^{n} F_i = \frac{F_1 + F_2 + \cdots + F_n}{n}, \tag{5.2}$$

where n is the number of data sets.

In articles where multiple systems were developed or compared in the article, the F-score for each system was presented. If a F-score was not possible to calculate this was indicated. If a NLP system was named in the article this name was used, otherwise no name was given to the system.

5.2 Results

5.2.1 Included Articles

The structured literature search resulted in 56 articles after including the article by Meystre et al. [2008] to ensure information about NLP systems before 2008.
After the assessment of the 56 articles, 5 were excluded due to a low overall assessment, which resulted in 51 articles being included for selection of a NLP system, see figure 5.1. The articles with a medium or high overall assessment, can be seen in table 5.3 on page 47.

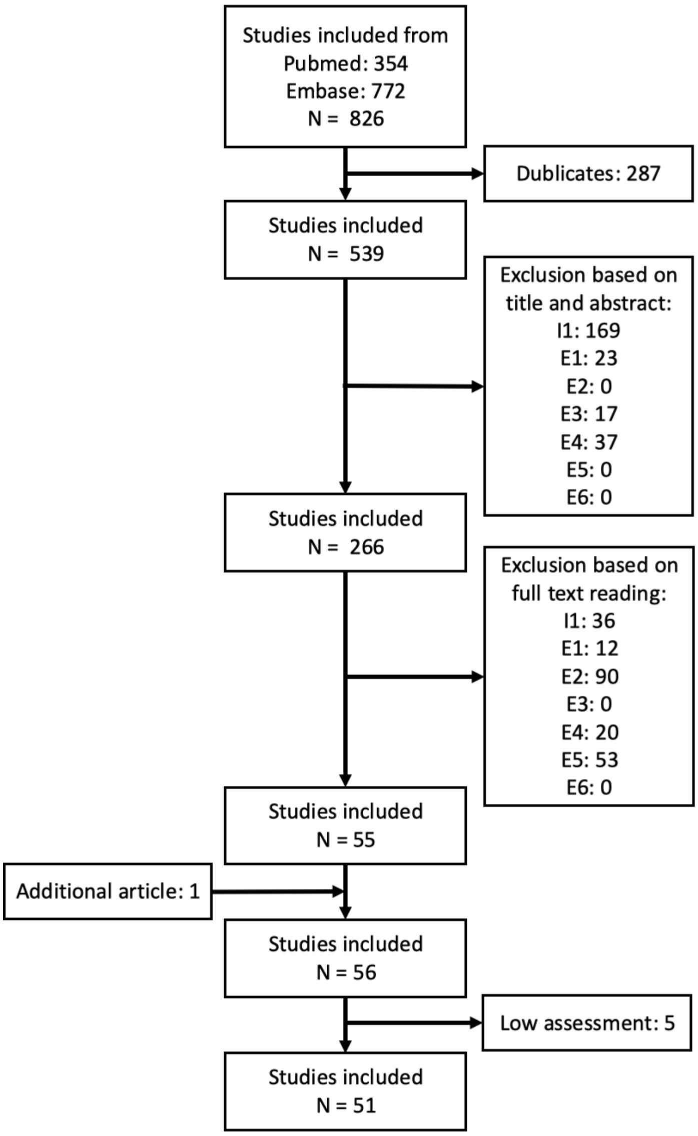

Figure 5.1. shows the number of included and excluded articles from the structured literature search.

Reference	Reliability	Internal Validity	External Validity	Overall Assess-ment
Development studies				
Bill et al. [2014]	Medium	Medium	High	Medium
Bozkurt et al. [2019]	Medium	Medium	High	Medium
Cai et al. [2019]	Medium	Low	Medium	Medium
Cheng et al. [2010]	Medium	Medium	Medium	Medium
Coden et al. [2009]	Medium	Low	Medium	Medium
Cote et al. [2017]	Low	Medium	Medium	Medium
Denny et al. [2009]	Low	High	Low	Meduim
Divita et al. [2014]	High	High	High	High
Doan et al. [2010]	High	High	Medium	High
Epstein et al. [2013]	Medium	High	Medium	Medium
Fu et al. [2020]	Medium	High	Low	Medium
Gao et al. [2015]	High	High	Medium	High
Garla et al. [2011]	High	Medium	Low	Medium
Gobbel et al. [2014]	Low	Medium	Medium	Medium
Goss et al. [2014]	Medium	Low	Medium	Medium
Guan and Devarakonda [2019]	Medium	Low	Medium	Mediu
Hamon and Grabar [2010]	Low	High	Medium	Medium
Hao et al. [2016]	High	Medium	Low	Medium
Iqbal et al. [2017]	High	Medium	Medium	Medium
Jindal and Roth [2013]	Medium	High	Medium	Medium
Kersloot et al. [2019]	High	Low	Medium	Medium
Knoll et al. [2019]	Medium	High	Medium	Medium
Kovačević et al. [2013]	Medium	Medium	Medium	Medium
Kulshrestha et al. [2020]	Medium	High	Medium	Medium
Li et al. [2020]	Medium	High	Medium	Medium
Liu et al. [2019a]	Medium	Medium	High	Medium
Liu et al. [2019b]	Medium	Low	Medium	Medium
Lou et al. [2020]	Medium	Medium	High	Medium
Martinez et al. [2014]	Medium	Low	Medium	Medium
McCart et al. [2013]	Medium	Medium	Medium	Medium
Meystre et al. [2010a]	High	Medium	Medium	Medium
Meystre et al. [2010b]	Medium	Medium	Medium	Medium
Mishra et al. [2019]	Medium	Medium	Medium	Medium
Moon et al. [2019]	Medium	Low	Medium	Medium
Nassif et al. [2009]	Medium	Low	Medium	Medium
Oleynik et al. [2019]	Medium	Low	Medium	Medium
Peterson et al. [2020]	High	Medium	Medium	Medium
Qiu et al. [2018]	Medium	Medium	Medium	Medium
Savova et al. [2010]	High	Medium	High	High
Sevenster et al. [2012]	Medium	Medium	Low	Medium
Shah et al. [2019]	Medium	Medium	Low	Medium
Tao et al. [2018]	Medium	Medium	Medium	Medium
Topaz et al. [2019]	Medium	Medium	Medium	Medium
Trivedi et al. [2019]	Medium	Medium	Medium	Medium
Wei et al. [2019]	Low	Medium	Medium	Medium

Continued on next page

Reference	Reliability	Internal Validity	External Validity	Overall Assessment
Yadav et al. [2013]	Medium	Medium	Low	Medium
Yang et al. [2020a]	Medium	Medium	Medium	Medium
Yang et al. [2020b]	Medium	Low	Medium	Medium
Zheng et al. [2012]	Medium	Low	Medium	Medium
Zhou et al. [2011]	High	Medium	Medium	Medium
Systematic review				
Meystre et al. [2008]				Medium

Table 5.3. shows the assessment of each article included from the structured literature search.

5.2.2 Sub-elements

In table 5.4 sub-elements that fulfill the criteria stated in section 5.1.3 on page 44 was presented. A full overview of which systems from the structured literature search mentioned particular sub-elements can be seen in appendix B.

Sub-element	Number of references
Tokenization	37
Sentence splitting	31
Negation detection	22
Normalization	18

Table 5.4. shows the sub-elements mentioned in more than 15 articles included from the structured literature search.

5.2.3 Existing Systems

In table 5.5 on page 50 an overview of the identified systems can be seen. The table was ordered by the used NLP system and F-score.

The results show that the best performing systems were pre-defined Bidirectional Encoder Representations from Transformers (BERT) and Medical Text Extraction, Reasoning, and Mapping System (MTERMS). Since these systems are not open-source they are not possible to use in this study.

ConText was the best performing open-source system, followed by Clinical Text Analysis and Knowledge Extraction System (cTAKES). All the identified sub-elements were included in cTAKES but not in ConText, and cTAKES was used in more articles than ConText. For these reasons, cTAKES [The Apache Software Foundation, 1000 N West Street, Wilmington, U.S.A.] was selected as the NLP system for this study.

Reference	Extracted information	NLP system	F-score	Open-source
NLP systems used in several articles				
Garla et al. [2011]	Hepatic decompensation	cTAKES	0,886 †	Y
Kersloot et al. [2019]	Oncology related terms	cTAKES	0,845 †	Y
Savova et al. [2010]	Semantically viable information	cTAKES	0,824	Y
Mishra et al. [2019]	Cranio facial and dental informations	cTAKES	0,81	Y
Liu et al. [2019a]	Phenotypes	cTAKES	0,726 †	Y
Zheng et al. [2012]	Coreference resolution	cTAKES	‡	Y
Liu et al. [2019a]	Phenotypes	MedLEE	0,712 †	Y
Sevenster et al. [2012]	Breast tumors	MedLEE	0,521 †	Y
Yadav et al. [2013]	Blunt facial trauma	MedLEE	‡	Y
Meystre et al. [2008]	Pneumonia	MedLEE	‡	Y
Wei et al. [2019]	Medicine	BERT	0,941	N
Wei et al. [2019]	Medicine	BERT	0,919	N
Guan and Devarakonda [2019]	Medicine	BERT	0,810	N
Zhou et al. [2011]	Diabetes and Cardiovascular system	MTERMS	0,906	N
Goss et al. [2014]	Allergies	MTERMS	0,876	N
Iqbal et al. [2017]	Diseases	ConText	0,900†	Y
Meystre et al. [2008]	Negations and event	ConText	0,842	Y
Denny et al. [2009]	Enlogated QT-segment	KnowledgeMap Concept Identifier	0,998	N
Meystre et al. [2008]	Cardiology reports	KnowledgeMap Concept Identifier	‡	N
NLP systems used in one articles				
Bozkurt et al. [2019]	Tumors	-	0,982	N
Hao et al. [2016]	Diabetes	Valx	0,98	Y
Hao et al. [2016]	Breastcancer	-	0,97	N
Trivedi et al. [2019]	Breastcancer	-	0,966	N
Knoll et al. [2019]	Unknown	-	0,96	Y
Gobbel et al. [2014]	Conceptualizing data in groups	RapTAT	0,95	Y
Cai et al. [2019]	Nummeric values	EXTEND	0,949 †	Y

Continued on the next page

Reference	**Extracted information**	**NLP system**	**F-score**	**Open-source**
Yang et al. [2020a]	Relations	-	0,948	N
Epstein et al. [2013]	Medication and food allergies	-	0,939	N
Kovačević et al. [2013]	Temporal information	-	0,894	Y
Peterson et al. [2020]	Events	ClinicalBERT	0,890	Y
Tao et al. [2018]	Medicine prescription	FABLE	0,878	N
Kovačević et al. [2013]	Event recognition	-	0,872 †	N
Coden et al. [2009]	Cancer characteristics	MedTas/P	0,859 †	N
McCart et al. [2013]	Identifying falls	-	0,850	N
Topaz et al. [2019]	Identifying falls	NimbleMiner	0,850	Y
McCart et al. [2013]	Identifying falls	-	0,849	N
Gao et al. [2015]	Mammographi	-	0,848 $\star$ †	Y
Doan et al. [2010]	Medicine	MedEx	0,821	N
Martinez et al. [2014]	Cancer tumors	-	0,805†	N
Qiu et al. [2018]	Breast and lung cancer	-	0,804	N
Cheng et al. [2010]	Cancer Tumors	-	0,800 †	N
Martinez et al. [2014]	Cancer tumors	-	0,791†	N
Jindal and Roth [2013]	Temporal information	HeidelTime	0,790	Y
Li et al. [2020]	Ischemic stroke	-	0,790	Y
Hamon and Grabar [2010]	Medicine	-	0,780	N
Martinez et al. [2014]	Cancer tumors	-	0,774 †	N
Meystre et al. [2010b,a]		Medicine	0,770	N
Martinez et al. [2014]	Cancer tumors	-	0,756 †	N
Oleynik et al. [2019]	Unknown		0,753	Y
Qiu et al. [2018]	Breast and lung cancer	-	0,728	N
Jindal and Roth [2013]	Events	-	0,710	N
Liu et al. [2019a]	Phenotypes	MetaMapLite	0,682	Y
Yang et al. [2020b]	Concepts and relations	-	0,654	N
Oleynik et al. [2019]	Unknown	-	0,590	Y
Oleynik et al. [2019]	Unknown	-	0,575	Y
Bill et al. [2014]	Family and observations	BioMediCUS	0,554	0
Divita et al. [2014]	Unknown	SPECIALIST	0,531	Y
Liu et al. [2019a]	Phenotypes	ClinPhen	0,461†	N
Lou et al. [2020]	Cancer surveillance	-	0,451	N
Oleynik et al. [2019]	Unknown	-	0,427	Y

Continued on the next page

Reference	**Extracted information**	**NLP system**	**F-score**	**Open-source**
Lou et al. [2020]	Cancer surveillance	-	0,387	N
Lou et al. [2020]	Cancer surveillance	-	0,381	N
Cote et al. [2017]	Brain	-	‡	Y
Fu et al. [2020]	Temporal information	-	‡	Y
Fu et al. [2020]	Temporal information	-	‡	Y
Fu et al. [2020]	Temporal information	-	‡	Y
Meystre et al. [2008]	Drug information and side effects	Linguistic String Project	‡	N
Meystre et al. [2008]	Negations in mammografi reports	NegExpander	‡	N
Meystre et al. [2008]	Negation detection	NegEx	‡	Y
Meystre et al. [2008]	Negation detection	NegFinder	‡	N
Meystre et al. [2008]	Temporal relations	TimeText	‡	N
Meystre et al. [2008]	Emergency department	MMTx	‡	Y
Meystre et al. [2008]	Lung scans, pneumonia and cental venous catheter	SymText	‡	N
Meystre et al. [2008]	Pathology	SPIN	‡	N
Meystre et al. [2008]	Cancer	caTIES	‡	Y
Meystre et al. [2008]	Diagnosis and smoking status	HITEx	‡	Y

Table 5.5. shows the NLP systems used in the articles included in the structured literature search, the associated F-score, and weather the system is open-source. The column 'Extracted information' gives a headline of which type of information was extracted. In the column 'NLP system' the used NLP system was mentioned. Initially the NLP systems used in several articles were presented, followed by NLP systems only used in one article.
† = Average was calculated, ⋆ = Calculated from precision and recall, ‡ = not available nor possible to calculate, Y = yes, N = no

Adjustments of cTAKES 6

6.1 Method

As mentioned in section 3 on page 8, adjustments were added to the source code of the selected NLP system, cTAKES, and is henceforth referred to as the adjusted cTAKES.

6.1.1 Data set

In order to train and test the adjusted cTAKES, a subset of the N2C2 2010 challenge data set described in section 4.1.1 on page 11 was used.
In section 2.2 on page 3 it was stated that diagnoses and symptoms can be in a structured format which corresponds to the extracted concepts concerning problems from the 2010 data set. As the adjusted cTAKES was a proof of concept, a random sample of ten discharge summaries was selected, three for training and seven for testing.

6.1.2 Gold Standard

6.1.2.1 Elaboration

It was desired to have a gold standard data set, as this made it possible to evaluate the performance of the adjusted cTAKES. The extracted problems in the data foundation, from now on referred to as concepts, were annotated by the N2C2 organization [Department of Biomedical Informatics Harvard Medical School, 2018]. It was chosen to bind the concepts to SNOMED CT expressions from version United State Edition 2021-03-01. The binding was performed on both the training and test data set. Prior to binding SNOMED CT expressions, all concepts and the associated sentence and headline were extracted.
A concept that had not yet occurred in the discharge summary e.g. something the patient should be aware of after discharge, was excluded, since it was unknown if the problem would occur or not.

The gold standard was elaborated in four steps; 1) identification of information assessed to be related and relevant for the concept by a healthcare professional, 2) outline a guideline for identifying related, relevant information, 3) outline rules for binding SNOMED CT expressions, and 4) binding each concept and related, relevant information in a SNOMED CT expressions.
The first step concerned assessment of the concepts. The SNOMED CT expressions, which healthcare professionals choose are different from those developers would choose de Keizer et al. [2008]. For this reason and since the authors of this study did not have a background within practical healthcare, a third party helped to assess the understanding of the concepts and related, relevant information from the sentence and headline. The

third party was Anaesthetic Nurse at Aalborg University Hospital, Maibrit Pape. The procedure for the interview can be seen in appendix C. Concepts from three of the ten randomly selected discharge summaries were assessed.
The second step concerned outlining a guideline, which was done using visual inspection of the related, relevant information. If the type of information was mentioned as relevant multiple times it was included in the guideline.
The third step concerned outlining rules for making SNOMED CT expressions. They should describe the concept and related, relevant information from the sentence and headline. The allowed attributes were identified in the concept model described by IHTSDO [2021c] chapter 6. To identify which SNOMED CT expressions were allowed to be bound to a given attribute, the rules stated in the MRCM described by IHTSDO [2021d] were used.
The fourth step concerned binding each concept and related, relevant information to a SNOMED CT expressions. Based on the guideline and the rules described in the second and third step, a concept was bound to a pre- or post-coordinated expression, the procedure for this can be seen on figure 6.1 on the next page.
Both the training and test data set were bound using this method, why the training and test data set henceforth include SNOMED CT expressions.

6.1.2.2 Evaluation

After binding the SNOMED CT expression in the fourth step, it was evaluate to what extend it was possible to follow the guideline. This was calculated using the three discharge summaries assessed by the third party. Each concept was assessed to be a 'full overlap', or to contain 'too much information' or 'too little information' compared to which information was assessed to be related and relevant by the third party. The category 'too much information' was relevant since it was assumed that the authors could assess additional information to be related and relevant based on the guideline. The category 'too little information' was relevant, since SNOMED CT is not exhaustive and that not all refinements was allowed in post-coordinated expressions.

An overview of the training and test data set was given including the number of 1) pre-coordinated expression, 2) post-coordinated expressions, 3) unintelligible concepts, and 4) concepts that were not possible to bind using SNOMED CT. Additionally, an average and percentage was calculated for each category, to show similarities and differences between the two data sets. A two-tailed t-test, with a significance level of 0.05 was used to calculate the differences between the two data sets. The hypothesis stated that the mean of the two groups are equal, while assuming equal variance.

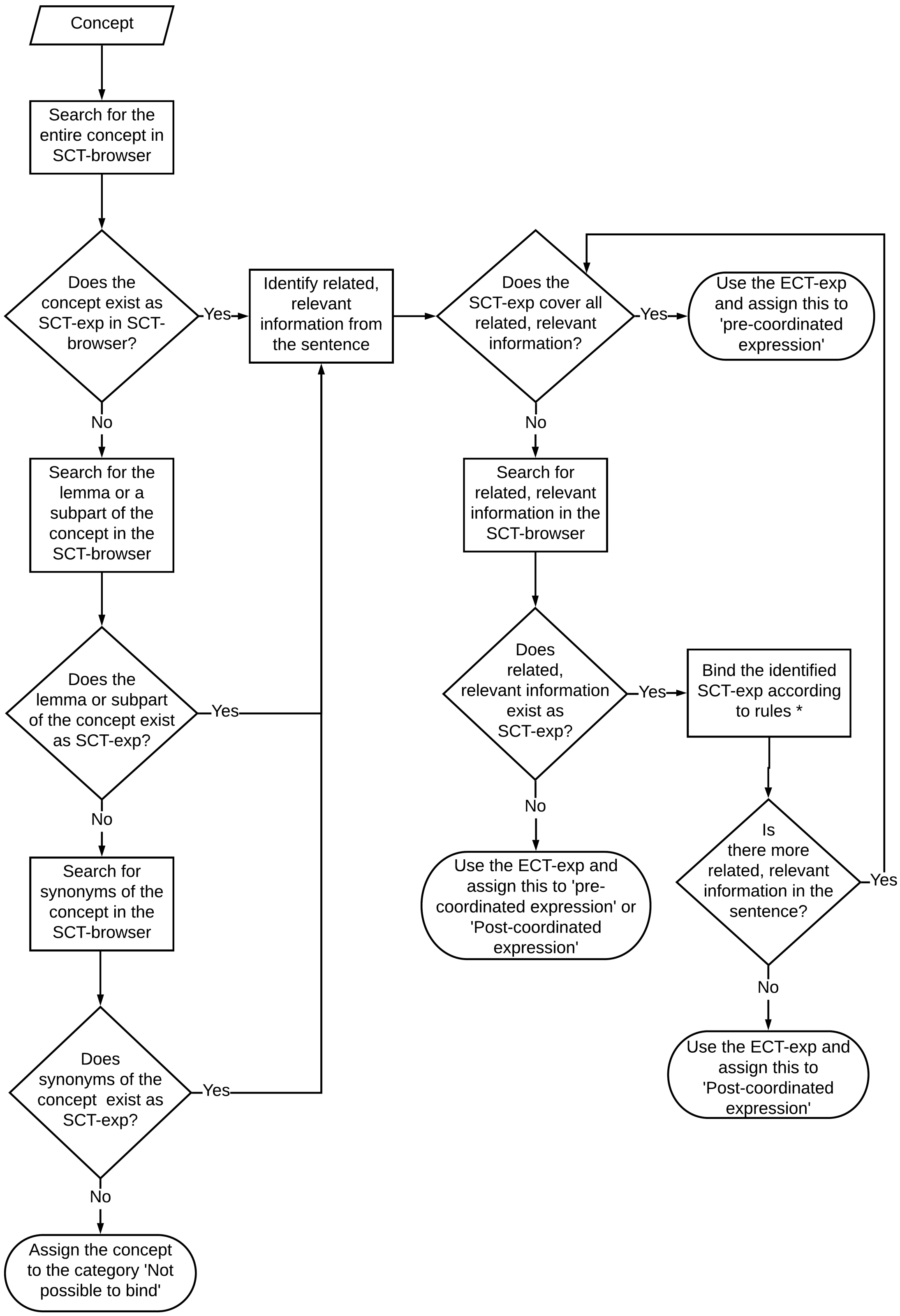

Figure 6.1. shows the procedure for identifying pre- or post-coordinated expressions for the concepts and related, relevant information from the discharge summaries.
SCT = SNOMED CT, SCT-exp = SNOMED CT expression, * = refer to the rules describe in this section, SCT-browser = SNOMED CT browser IHTSDO [2021a].

6.1.3 Agile Methodology

An agile methodology was used to support the test-driven adjustment of cTAKES, to ensure a synergy between the adjustments and outcome. The seventh principle of the agile methodology states that working software is the primary measure for progress [Hunt, 2006]. Therefore, the focus was on delivering working software frequently. To achieve this, mile stones were establish for the adjustments.
The mile stones were set in order to deliver working software in several steps. In figure 6.2 an overview of the milestones can be seen. The first mile stone was to install the default version of cTAKES as this was the foundation for adjustments. In the structured literature search it was found that cTAKES could included the necessary sub-elements, see table 5.4 on page 47. A milestone was set to ensure, that they all are included in the pipeline. Incorporation of rules for pre-coordinated expressions and post-coordinated expressions with one refinement was then identified as the next mile stones. The last mile stone was to incorporate rules for post-coordinated expressions with two refinements. Two refinements was found sufficient, as this is a proof of concept.
Another focus of the agile approach is that documentation have to be supportive [Hunt, 2006], why the adjustments made to the default cTAKES pipeline and source code were documented by using e.g. flowcharts.

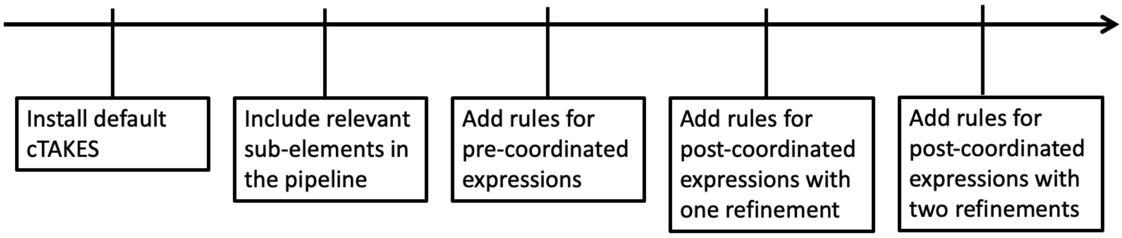

Figure 6.2. shows the timeline of the mile stones for applying adjustments to the default cTAKES.

6.1.4 Data Analyses

In order to incorporate rules for generating pre- and post-coordinated expressions into the adjusted cTAKES four analyses were made on the SNOMED CT expressions in the training data set.

6.1.4.1 Identification of Top Levels

Development of SNOMED CT post-coordinated expressions followed the concept model defined by IHTSDO [2021c] as mentioned in the third step in section 6.1.2 on page 51. These rules are based on the 19 SNOMED CT top levels and state which top levels can be combined into post-coordinated expressions using the attributes [IHTSDO, 2021c]. Therefore, it was important to analyse the training data in order to identify which SNOMED CT expression from the default cTAKES belonged to which top level as well as which top levels were needed to make rules based on the training data. The analysis used the output from the default cTAKES to divide the extracted SNOMED CT expressions into top levels.

6.1.4.2 Pre-coordinated Expressions

Pre-coordinated expressions can belong to any of the 19 top levels [IHTSDO, 2021c]. Therefore, it was identified which top levels were represented as pre-coordinated expressions in the training data set. This was done in order to limit the number of pre-coordinated expressions generated by the adjusted cTAKES and thereby reduce the number of incorrectly identified SNOMED CT expressions.

6.1.4.3 Translation of Qualifier Values

An analysis of the qualifier values was conducted aiming at identifying which words or phrases in the sentence and headline from the training data set could be translated to a given qualifier value. This was important as the adjusted cTAKES only bound words or phrases which are an exact textual match. Not using the results from this analysis would cause undefined synonyms not being bound e.g. could the absence of a disease be indicated by a preceding 'no', but it would not be detected.

6.1.4.4 Prioritization of Refinements

In the training data set an unlimited number of refinements could be added to a focus concept when making post-coordinated expression. If the adjusted cTAKES had to bind every possible combination of SNOMED CT expressions it would result in a high number of incorrectly identified SNOMED CT expressions. However, since this is a proof of concept the adjusted cTAKES included no more than two refinements. Therefore, two analyses were conducted, to identify how frequent top levels were represented in the refinements in the post-coordinated expressions. The first analysis concerned frequency of top level in any refinement. This was done to ensure that the generated post-coordinated expression included the most frequent top level available, hence the information assumed to be most related and relevant to describe the focus concept. Based on the frequency, the top levels were prioritized in the first refinement. The second analysis was conducted to prioritize the top levels for the second refinement. The combination of the refinements in a post-coordinated expression was counted. For instance, if a top level of a refinement was 'Procedure' it was counted at what frequency each of the 19 top levels were represented in the additional refinements.

6.1.5 Test of the Adjusted cTAKES

To test the performance of the adjusted cTAKESthe results from the adjusted cTAKES were compared to the gold standard. The results were divided into true positives (TP), false positives (FP), and false negatives (FN) defined as in table 6.1 on the next page which was based on the article by Kersloot et al. [2019]. From these categories precision, recall, and F-score were calculated to obtain a measure for the performance of the adjusted cTAKES.

Category	Description
TP	SNOMED CT expression identical to the gold standard
FP	SNOMED CT expression differs from the gold standard
FN	Concept not bound to a SNOMED CT expression

Table 6.1. shows the three categories of the results from the adjusted cTAKES and their associated description. Descriptions were made with inspiration from the article by Kersloot et al. [2019].

Precision, also called the positive predicted value, represents the ability of the adjusted cTAKES to correctly annotated concepts compared to all annotations [Kersloot et al., 2019], which was calculated with equation 6.1:

$$Precision = \frac{TP}{TP + FP} \tag{6.1}$$

Recall, also called the sensitivity, represents the ability of the adjusted cTAKES to correctly annotate the concepts [Kersloot et al., 2019], which was calculated with equation 6.2:

$$Recall = \frac{TP}{TP + FN} \tag{6.2}$$

The F-score represents the performance of a binary classifier based on precision and recall [Kersloot et al., 2019], and was calculated using equation 5.1 on page 44.

6.2 Results

6.2.1 Gold Standard

Discharge summary no. 32, 67, and 81 composed the training data set, and discharge summary no. 13, 14, 33, 36, 38, 74, and 80 composed the test data set. The characteristics for the training and test data sets can be seen in table 6.2 and 6.3 on the next page, respectively. A statistical comparison of the data sets can be seen in table 6.4 on the following page.

	32	67	81	Average ± SD	Percent [%]
Pre-coordinated expressions	17	8	23	16.0 ± 7.5	27.1
Post-coordinated expressions	55	13	36	35.7 ± 21.2	60.5
Unintelligible words	6	0	0	2.0 ± 3.5	3.4
Not possible to bind	9	1	6	5.3 ± 4.0	9
In total	87	22	68	59.0 ± 33.4	100

Table 6.2. shows the number of pre- and post-coordinated expressions, unintelligible concepts, and concepts not possible to bind for each discharge summary in the training data set. For each of the four categories the average, the standard deviation, and the percentage can be seen.

	13	14	33	36	38	74	80	Average ± SD	Percent [%]
Pre-coordinated expressions	17	21	28	6	16	9	10	15.3 ± 6.9	31.1
Post-coordinated expressions	23	37	50	13	40	18	26	29.6 ± 13.9	60.2
Unintelligible words	2	0	0	0	0	0	0	0.3 ± 0.8	0.6
Not possible to bind	1	1	8	0	11	2	5	4.0 ± 4.2	8.1
In total	43	59	86	19	67	29	41	49.1 ± 23.1	100

Table 6.3. shows the number of pre- and post-coordinated expressions, unintelligible concepts, and concepts not possible to bind for each discharge summary in the test data set. For each of the four categories the average, the standard deviation, and the percentage can be seen.

The t-test showed that none of the categories in the two data sets were statistical significant ($p > 0.05$), see table 6.4.

	Train	Test	p-value
Pre-coordinated expressions	16.0	15.3	0.89
Post-coordinated expressions	26.0	29.6	0.60
Unintelligible words	2.0	0.3	0.22
Not possible to bind	5.0	4.0	0.65

Table 6.4. shows the average number of pre-coordinated expressions, post-coordinated expressions, unintelligible concepts, and concepts not possible to bind for the training and test data sets, as well as the p-value for each category.

Discharge summary no. 67 from the training data set and no. 14 and 80 from the test data set were assessed by Anaesthetic Nurse at Aalborg University Hospital, Maibrit Pape. In table 6.5 the agreement between the selected SNOMED CT expressions and the third party can be seen.

The guideline for identifying related, relevant information was elaborated and contained five observations; 1) indication of whether the concept was a part of the patient's anamnesis, most often indicated with 'history of', 2) if the concept was negated, e.g. 'no abdominal pain', 3) an administered medication associated with the concept, e.g. 'Vancomycin given for Corynebacterium', 4) if a procedure caused the problem described in the concept, e.g. 'hyponatremia following a surgery', and 5) if there are associated diseases or symptoms.

	14	**67**	**80**	**Percent [%]**
Full overlap	44	17	24	70.2
Too much bound	6	0	4	8.3
Too little bound	8	4	7	15.7
Not bound	1	1	5	5.8
In total	59	22	40	100

Table 6.5. shows the agreement between the SNOMED CT expressions and the related, relevant information selected by the third party. The agreement was made based on concepts extracted from discharge summary no. 67 from the training and no. 14 and 80 from the test data set.

6.2.2 Documentation of the Adjusted cTAKES

An overview of the adjustments added to the default cTAKES can be seen in appendix D. The pipeline for the adjusted cTAKES followed the pipeline defined in the file 'AggregatePlaintextFastUMLSProcessor.xml' with cTAKES' status and negation detector added, which can be seen on figure 6.3 on the next page.

By default cTAKES uses a dictionary based on Unified Medical Language System (UMLS), which includes several terminologies, such as RxNorm, SNOMED CT, and Logical Observation Identifiers, Names, and Codes (LOINC). Since SNOMED CT was chosen as standard terminology in section 2.5 on page 6, an UMLS dictionary only including SNOMED CT USA edition 2020 was used, as this was the latest version available in English for download on 30/3/2021.

As a part of the included dictionary the adjusted cTAKES used a term consumer to determine which span of words in a sentence were eligible for terminology binding. It was chosen to use the 'PrecisionTermConsumer.java', as this included the longest overlapping semantic span of words e.g. 'knee pain' instead of 'knee' and 'pain'. In this way more of the situational contextual information was preserved in a single SNOMED CT expression.

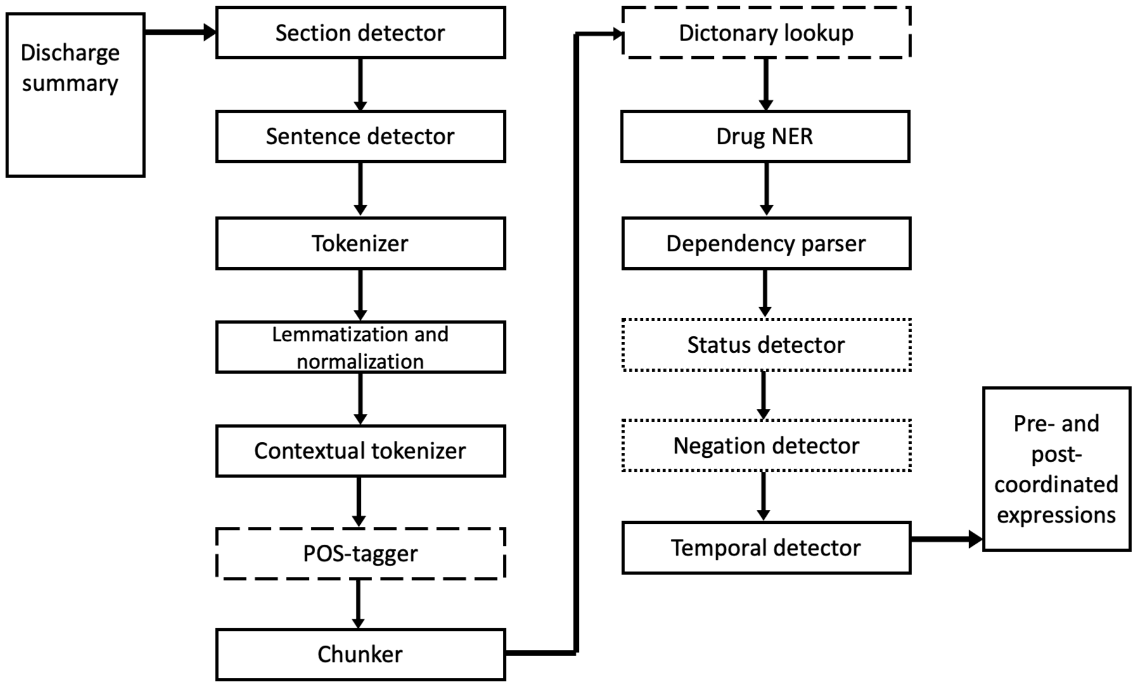

Figure 6.3. shows the pipeline for executing the adjusted cTAKES. The boxes with the dashed line indicate in which sub-elements adjustments were made. The boxes with the dotted line indicate the added sub-elements.
POS = Part-of-speach, NER = Named Entiry Regognition.

6.2.2.1 Adjustments in POS-tagger

In the dashed box with part-of-speech (POS) tagging on figure 6.3, each extracted word was assigned to a word class, e.g. noun, verb, adjective. The default POS-tag list was extended to include the POS-tags in the following list, in order to accommodate that non-noun words in a sentence could be included into the SNOMED CT expressions.

- Preposition or subordinating conjunction (IN)
- Noun, singular or mass (NN)
- Noun, plural (NNS)
- Proper noun, singular (NNP)
- Proper noun, plural (NNPS)
- Verb, base form (VB)
- Verb, past tense (VBD)
- Verb, gerund or present participle (VBG)
- Verb, past participle (VBN)
- Verb, non-3rd person singular present (VBP)
- Verb, 3rd person singular present (VBZ), [Treebank, 2003]

6.2.2.2 Adjustments in Dictionary Lookup

In the dashed box with dictionary lookup on figure 6.3, rules and supporting source code were added to enable pre- and post-coordinated expression to be generated.

Identification of SNOMED CT Top Levels

Not all of the 19 SNOMED CT top levels were represented in the SNOMED CT expressions in the training data set. To identify the top levels used in the training data set, the analysis described in section 6.1.4.1 on page 54 was used. In the list below the identified top levels were presented. Some of the top levels presented in the list below were in fact sub-levels of a top level, but for convenience they are henceforth called top levels, since they contained important information in relation to the implemented rules.

- Attribute (sub-level of SNOMED CT Model Component)
- Body Structure
- Clinical Finding
- Morphologic Abnormality (sub-level of Body Structure)
- Observable Entity
- Organism
- Person (sub-level of Social Context)
- Procedure
- Qualifier Value
- Situation with Explicit Context
- Substance, [IHTSDO, 2021c]

As the SNOMED CT expressions in default cTAKES were assigned to the extracted words and phrases based on an UMLS dictionary, the SNOMED CT expressions were not pre-categorized into the 19 SNOMED CT top levels defined in IHTSDO [2021c] chapter 6. However, it was found that each SNOMED CT expression was associated with a Type Unique Identifier (TUI) code which represents semantic groups defined by UMLS [Gu et al., 2016]. In the adjusted cTAKES, the TUI codes were used to assign the SNOMED CT expressions to a top level, except for the top level 'Attribute'. The choice of which TUI code belonged to which top level can be seen in table 6.6.

Top Level	**TUI**
Body Structure	T017, T018, T021, T022, T023, T024, T025, T026, T029, T030, T031
Clinical Finding	T019, T025, T026, T029, T034, T037, T046, T047, T048, T049, T050 T051, T055, T184, T191
Morphologic Abnormality	T020, T190
Observable Entity	T038, T039, T041, T042, T052, T201
Organism	T005, T007, T204
Person	T016, T099, T100, T101
Procedure	T056, T058, T059, T060, T061, T065
Qualifier Value	T033, T077, T079, T080, T081, T082
Situation with Explicit Context	-
Substance	T103, T104, T109, T110, T111, T114, T115, T116, T118, T119, T120, T122, T123, T124, T125, T126, T127, T129, T130, T131, T167, T192, T196, T197

Table 6.6. shows which TUI code was assigned to which top level. Findings were based on the training data set.
- = no TUI code identified for the top level.

Based on the description of the TUI codes and the training data set, no TUI code was identified for the top level 'Situation with Explicit Context', why this top level was not used further in the adjusted cTAKES.
This analysis showed that it was to some extend possible to divide the SNOMED CT expressions into 10 top levels using TUI codes. Therefore, the TUI codes were extracted and used to assign each SNOMED CT expression to a top level.

Rules for Pre-coordinated Expressions
Pre-coordinated expressions generated by the adjusted cTAKES were based on SNOMED CT expression identified in one sentence, as it was not possible to use information from the headline. The headline was not used, as the default part of cTAKES only detected one section per discharge summary, hence only one headline.
The analysis of all pre-coordinated expressions described in section 6.1.4.2 on page 55, showed that only SNOMED CT expressions from the top levels 'Clinical Finding' and 'Morphologic Abnormality' were used. Pre-coordinated expressions from these two top levels accounted for 27.1% of the SNOMED CT expression in the training data set, see table 6.2 on page 57. Therefore, the adjusted cTAKES generated pre-coordinated expressions from all SNOMED CT expressions in these two top levels.

Rules for Post-coordinated Expressions
As with the pre-coordinated expressions, post-coordinated expressions were based on SNOMED CT expressions identified in one sentence. The adjusted cTAKES bound a focus concept and up to two refinements in one post-coordinated expression. Based on the training data set it was found that only SNOMED CT expressions from the top level 'Clinical Finding' were used as focus concepts in post-coordinated expressions, which was implemented in the source code.
The first refinement bound by the adjusted cTAKES, was prioritized based on the frequency of the top level, as described in the first analysis in section 6.1.4.4 on page 55. The prioritization of the top levels used in refinements can be seen in table 6.7.

First Top Level	Frequency
Qualifier Value	56
Clinical Finding	33
Procedure	22
Body Structure	20
Person	6
Morphologic Abnormality	2
Substance	1
Organism	1
Observable Entity	1

Table 6.7. shows the prioritization of the top levels used as the first refinement based on how frequent the top level was represented in the training data set.

To avoid an excessive number of incorrect identified post-coordinated expressions, the second refinement was prioritized in respect to each top level, as described in the second analysis in section 6.1.4.4 on page 55. The prioritization can be seen in table 6.8 on the following page.

First Top Level	Second Top Level	Instances
Body Structure	Qualifier Value	6
	Procedure	4
	Body Structure	2
	Stand alone refinement	8
Clinical Finding	Qualifier Value	7
	Procedure	5
	Clinical Finding	6
	Stand alone refinement	10
Morphologic Abnormality	Qualifier Value - After	2
	Stand alone refinement	1
Observable Entity	Stand alone refinement	1
Organism	Qualifier Value	1
	Stand alone refinement	-
Person	Qualifier Value	4
	Stand alone refinement	-
Procedure	Clinical Finding	5
	Qualifier Value	5
	Body Structure	4
	Stand alone refinement	7
Qualifier Value	clinical Finding	7
	Body Structure	6
	Procedure	5
	Person	4
	Morphologic Abnormality	2
	Organism	1
	Stand alone refinement	29
Substance	Stand alone refinement	1

Table 6.8. shows the prioritization of the top levels used as the second refinement in respect to the first refinement. The results were identified in the training data set.
'Stand alone refinement' = the number of times a top level in a refinement was used without any other refinements.

For the top level 'Clinical Finding' in the first refinement the most frequent combined top level in the second refinement was an additional 'Clinical Finding'. However, it was further observed that five out of the six clinical findings in the second refinement were the SNOMED CT expression '365854008 | History finding |'. This particular SNOMED CT expression describes a finding from the patient's anamnesis, and in the training data set this SNOMED CT expression was bound based on information in the headline and not the sentence. As the headline was not used, the SNOMED CT expression '365854008 | History finding |' was never used. Therefore, the prioritization was changed and followed table 6.8.

The process of generating the pre- and post-coordinated expressions can be seen on figure 6.4 on the following page while the combination of the top levels and associated attributes can be seen in appendix E.

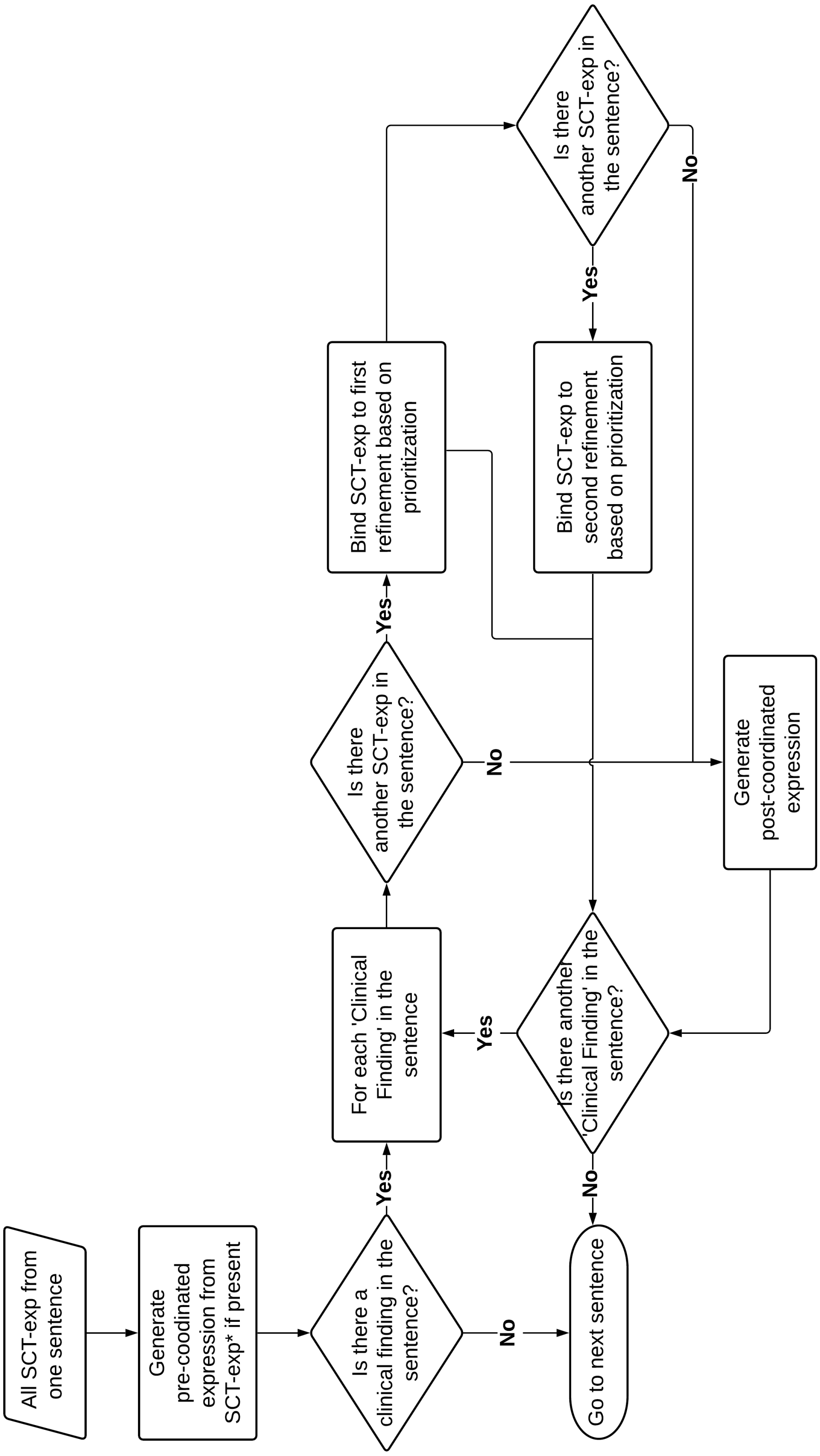

Figure 6.4. shows a flowchart of the source code for making the pre- og post-coordinated expressions in the adjusted cTAKES. SCT-exp = SNOMED CT expression, * = SNOMED CT expression from the top levels 'Clinical Finding' or 'Morphologic Abnormality'.

Determine SNOMED CT Expressions for Attributes

Since only SNOMED CT expressions from the top level 'Clinical Finding' were used as focus concepts in post-coordinated expressions only attributes possible to bind to 'Clinical Finding' were used, based on the concept model [IHTSDO, 2021c]. The process for determining attributes can be seen on figure 6.5 while the pre-defined set of SNOMED CT expressions for attributes can be seen in appendix E.

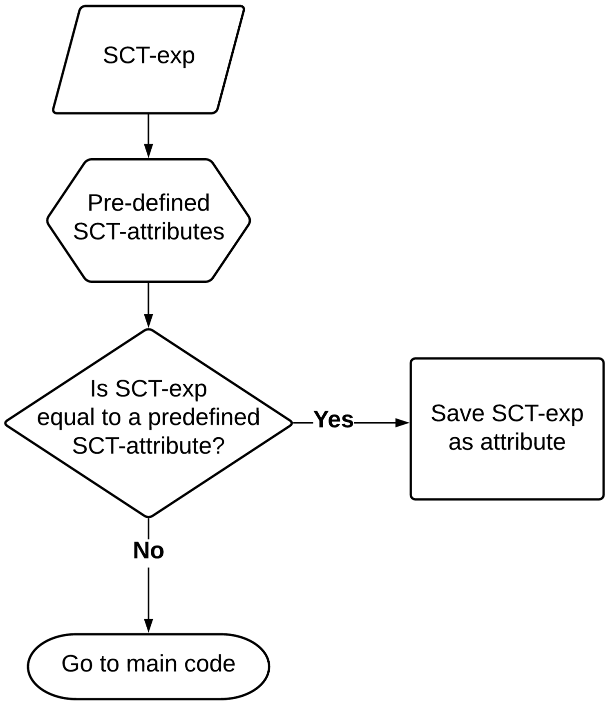

Figure 6.5. shows a flowchart of the process for determining if a SNOMED CT expression is an attribute.
SCT-exp = SNOMED CT expression, SCT-attribute = SNOMED CT expressions from the top level 'Attribute'.

Translation of Qualifier Values

As described in section 6.1.4.3 on page 55 synonyms for qualifier values were identified in addition to the identified SNOMED CT expressions assigned to the top level 'Qualifier Value' based on TUI codes. The synonyms were used the to determine if a SNOMED CT expression covered a word or phrase representing a qualifier value. The process of translating synonyms into qualifier values can be seen on figure 6.6 on the next page, and the results can be seen in table 6.9 on page 66.

As the synonym 'No' was originally used to indicate two qualifier values, a deeper analysis was conducted. This showed that the synonym 'No' was used once to describe '264868006 | No growth |' and 15 times to describe '2667000 | Absent |'. Therefore, the synonym 'No' was used to describe '2667000 | Absent |' in the adjusted cTAKES resulting in the qualifier value '264868006 | No growth |' to be excluded.

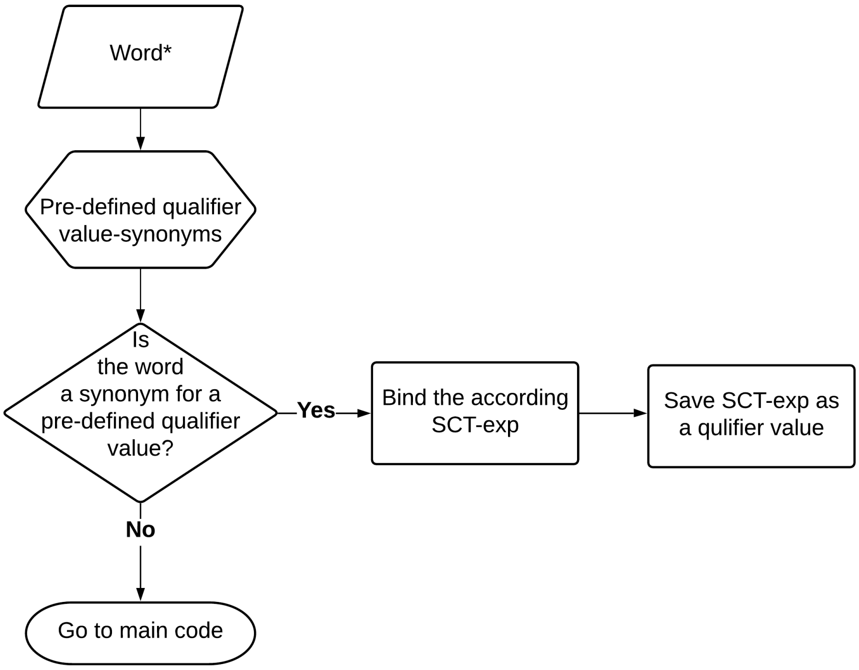

Figure 6.6. shows a flowchart of the process for determining whether a word or phrase is a synonym for a qualifier value.
Word* = the word or phrase which a SNOMED CT expression covers, SCT-exp = SNOMED CT expression.

Qualifier Value	Synonym
2667000 | Absent |	Denies Negative No Non Without
373933003 | Acute onset |	Spontaneous
62459000 | Chronic persistent |	Persistent
448371000124103 | Clinically undetermined |	Like Likely Possible Presumably
263730007 | Continual |	Weeks
1250004 | Decreased |	Decreased Less
303114002 | Early neonatal period |	Preterm Infants
54328002 | Indifferent |	Stable
255604002 | Mild |	Mild
6736007 | Moderate |	Moderate
264868006 | No growth |	-
425404009 | Slightly |	Slightly
255507004 | Small |	Small
425323003 | Sudden onset AND short duration |	Immediately
257556004 | Surgery |	Surgery
260360000 | Very high |	Very

Table 6.9. shows which synonyms were used in the sentence or headline, in order to identify qualifier values for the post-coordinated expressions.

6.2.3 Performance of the Adjusted cTAKES

A comparison of the SNOMED CT expressions in the gold standard and generated by the adjusted cTAKES was conducted for both the training and test data set. The SNOMED CT expressions from the adjusted cTAKES were generated by running the Apache GUI Collection Processing Engine for Windows. The rate of TP, FP, and FN as well as the precision, recall, and F-score can be seen in table 6.10.
In discharge summaries no. 36 and 80 from the test data set, no TP were found. Additionally, 96.4% and 96.8% of the SNOMED CT expressions generated by the adjusted cTAKES were incorrect identified SNOMED CT expressions for the training and test data set, respectively.

	FN	FP	TP	Precision	Recall	F-score
Training Data Set	69	63	22	0.259	0.242	0.250
Test Data Set	157	136	20	0.128	0.113	0.120

Table 6.10. shows the results of the comparison between the SNOMED CT expressions in the training and test data set and generated by the adjusted cTAKES.

Discussion 7

Profiling was performed on 21 FHIR resources which contained 95.7% of information from the three randomly selected discharge summaries. All the developed FHIR profiles were found valid. A gold standard was made, where concepts and relevant information was bound to pre- (31.1%) and post-coordinated SNOMED CT expressions (60.2%) in the test data set. The default cTAKES was adjusted and used to extract information from discharge summaries. Both pre- and post-coordinated SNOMED CT expressions were extracted from the discharge summaries, resulting in a F-score on 0.120.
It is considered how the complexity of healthcare data, the maturity of HL7 FHIR, and whether the implementation context affects the structuring of healthcare data. The results of data extraction were affected by the complexity of the healthcare data, why it might be worth considering if it makes sense to try structuring free text documents.

7.1 The Complexity of Healthcare Data

Knowledge within the clinical domain, can according to the article by Garde and Knaup [2006] be described as; broad as it constantly expands, deep as existing knowledge becomes more detailed, and complex as new interactions are found. Therefore, is it only possible to develop and implement a fully relevant system for a limited period. This is problematic, since the system has to be up to date to ensure that patients at all times are provided with the most appropriate care. [Garde and Knaup, 2006] Due to this, there must be a set of requirement to define which information is relevant at the given time. In the developed FHIR profiles, this requirement included restricting the minimum cardinality from 0 to 1, as described in table 4.3 on page 15. A minimum cardinality of 1 requires the element to contain data when exchanging the FHIR profile, though it cannot be assured that the data are valid. To ensure this, definition of a FHIRPath is required which describes the path to the correct data element in a database. As information can be incomplete, changing the minimum cardinality to 1 can become a challenge, especially when the profiles are used in different settings, e.g. different hospital departments. [Health Level 7, 2021g]
Since interoperability expands [Dinh-Le et al., 2019], it might not be optimal to make FHIR profiles for single usage, but rather for multiple usages. This would require a common understanding of which information is of interest and which is not. The article by Dinh-Le et al. [2019] investigated the challenges associated with health wearables. It was found that some of these challenges concern interoperability, integration of data into EHR, and handling the tremendous amount of data. Further, it was stated that there should be established requirements for the data that should be integrated, such as data analysis or selection of the important information [Dinh-Le et al., 2019]. This underlines the need for data requirements when exchanging data in different settings of healthcare.

As earlier mentioned, discharge summaries are developed for the primary care physicians [Lenert et al., 2014], and they are used as the primary medium for communication between hospitals and primary care physicians [Spasic and Nenadic, 2020]. Therefore, they are essential in order to ensure patient safety and the continuity of care after discharge [Spasic and Nenadic, 2020]. Discharge summaries typically include; date of admission and discharge, reasons for hospitalization, findings within different specialities, performed tests and test results, conditions, medical circumstances or changes, etc. [Spasic and Nenadic, 2020]. The developed FHIR profiles include the majority of the data mentioned above and are developed based on the data foundation. Another approach could be to develop FHIR profiles based on existing standards for discharge summaries. The standardization organization MedCom has elaborated a standard for Danish discharge summaries [Medcom, 2021; MedCom, 2019]. The information in the MedCom standard differs from the discharge summaries used in the data foundation in regard to the structure and content. An example is that information about the receiver should included a name, department, organization, and address if available, as well as a mandatory receiver identifier. [MedCom, 2019] This shows that when following a standard, there are other requirements for the data, and in this case it would result in more information to be included in the FHIR profiles. According to the articles by Spasic and Nenadic [2020] and O'Leary et al. [2009] the structure and content of discharge summaries vary greatly between institutions and clinicians, which indicates a lack of international agreement. Compared to the data foundation, international agreement upon the structure and content of the discharge summaries, would set up new requirements for the data. Furthermore, it would most likely result in the developed FHIR profiles being broader, hence enable usage at multiple hospitals following the same standard for discharge summaries. This is in contrast to the developed profiles, since they are not applicable for usage at multiple hospitals.

7.2 Maturity of HL7 FHIR

The articles by Peterson et al. [2020] and Hong et al. [2018], both mapped data extracted with a NLP system to a FHIR resource. The article by Hong et al. [2018] extracted medication information and mapped it to the MedicationStatement resource, where the article by Peterson et al. [2020] extracted information about diagnoses and mapped it to the Condition resource. The article by Peterson et al. [2020] claims that the applied approach can be used for mapping to any resource. Despite this, both articles are quite narrow in their development and no implementation context was considered. This is in contrast to the implementation context defined in section 3 on page 8 as the data foundation was used to identify relevant FHIR resources for the profiling. By doing so, it permitted an evaluation of the maturity of the FHIR resources. In total, 95.7% of information from the discharge summary no. 51, 53, and 74 was possible to assign to elements in FHIR resources. This shows that HL7 FHIR can enable interoperability of information from discharge summaries. However, it leaves 4.3% of the information to be unassigned to a FHIR resource, including orientation of patients and formalities.

Information about orientation of a patient is important when making quality assurance of healthcare [Jünger and Nagel, 2019]. Additionally, involvement of patients in the decision making process can positively affect their health, as they possibly become more willing to follow treatments, why this is a fundamental element in future healthcare [Jünger and

Nagel, 2019; Stacey et al., 2017]. This information is important for the primary care physician, in order to know whether the patient is e.g. biased for or against the treatment [Stacey et al., 2017]. Therefore, it can be assumed that the healthcare professionals are motivated to document when they have oriented a patient about a situation. A solution to share this information using HL7 FHIR, could be to add an extension in the HL7 FHIR registry.

In the discharge summaries some formalities were identified e.g. who dictated the discharge summary. This information was assumed to be important situational contextual information, as the dictating healthcare professional likely plays a central role in the care of the patient. However, it was not possible to bind this information in a FHIR profile. The responsibility of the provided care must be assigned to a person Pickard [2019]. According to a commentary by Pickard [2019] responsibility in healthcare can be assigned to a person, when the person knows what they are doing, knows the alternatives, and has sufficient control of their actions. This is valid at the time of the action and in a period afterwards [Pickard, 2019]. This argues that all information about a healthcare professional's responsibilities should be possible to document in an EHR and include in a FHIR profile.

7.3 The Implementation Context

In section 2.4.2 on page 5 it was mentioned that both HL7 CDA and HL7 FHIR can be used as exchange standards. The architectural standard HL7 v3 has the ability to exchange messages, which contains CDA documents [Rinner and Duftschmid, 2016]. In the study by Rinner and Duftschmid [2016], mapping from CDA documents to FHIR resources was investigated, which was successful for six FHIR resources. Both the ability to exchange messages and documents were adapted by HL7 FHIR [Rinner and Duftschmid, 2016], by using the Bundle types 'message' or 'document', respectively [Health Level 7, 2021e,d].

The type of bundle was chosen to be a 'document', as this was suitable for the intended implementation context, described in section 3 on page 8. The Bundle type 'document' was chosen under the assumption that all primary care physicians need the same information from the discharge summaries.

Using the Bundle type 'message', a request is send from a source application asking for specific content from a destination application. The FHIR resources in a message depends on the request and is therefore not an immutable set of FHIR resources, as in the Bundle type 'document'. [Health Level 7, 2021e,d] Using the Bundle type 'message' allows the user to request data structured in different FHIR resources for every request, whereas the immutable set of FHIR resources in the Bundle type 'document' could result in too much or too little data being exchanged [Health Level 7, 2021e,d]. As described in section 3 there is a trade-off between the flexibility and degree of interoperability, when working with HL7 FHIR.

Based on the above mentioned, it could be stated that the Bundle type 'message' supports flexibility, to a higher degree than the Bundle type 'document'. Though, the Bundle type 'document' can be defined for usage in different settings, hence ensuring interoperability of information when requirements for data are followed. For instance, specifying a document for several hospital departments since a surgical department need a different set of information than a medical department. Therefore, the Bundle type 'document' could be used in the defined implementation context.

7.4 The Semantic Challenges

As mentioned in section 2.4.1 on page 5, post-coordinated expressions from SNOMED CT could be used to include additional situational contextual information than the pre-coordinated expressions. This suggestion is supported by the findings in the gold standard, since it included 60.5% and 60.2% post-coordinated expressions in the training and test data set, respectively. The agreement between the selected SNOMED CT expressions and the related, relevant information identified by Anaesthetic Nurse at Aalborg University Hospital, Maibrit Pape was 70.2%. This shows that post-coordinated expressions are in fact needed to describe information in discharge summaries in clinical practice. The downside of these post-coordinated expressions appears in the poor results from the adjusted cTAKES. Future work could be to apply logic in order to enable the adjusted cTAKES to choose the correct post-coordinated expression. If this is applied the adjusted cTAKES would be one step closer to implementation in clinical practice, hence improving the process of structuring free text documents.

The F-score obtained by the adjusted cTAKES was poor compared to the identified studies using cTAKES, in table 5.5 on page 50. Several influential factor can affect these results, including but not limited to; the division of SNOMED CT expressions into SNOMED CT top levels and the lack of extracted data.
The division of the SNOMED CT expressions into the 19 SNOMED CT top levels was based on TUI codes. This process was crucial for the adjusted cTAKES as the rules for generating SNOMED CT expressions was based on the concept model and therefore use the top levels [IHTSDO, 2021c,d]. A concern is that the TUI codes do not correspond to one top level, but often describe information that could be assigned to multiple top levels [Gu et al., 2016]. This affected the division negatively as some of the extracted SNOMED CT expressions were incorrectly assigned to a top level. A possible solution could be to obtain a more comprehensive mapping between SNOMED CT and UMLS. This would most likely result in more of the situational contextual information to be divided thus bound correctly.
In the adjusted cTAKES minimal changes were made to the existing source code, and mainly new rules and a class were added as described in section 6.2.2 on page 58. The output of the default part of the adjusted cTAKES is challenged by misspellings, synonyms, and abbreviations [Mishra et al., 2019; Zheng et al., 2012; Savova et al., 2010]. These are found to be commonly present in free text documents as mentioned in section 2.3.2 on page 4. Misspellings were not handled, but synonyms were partly handled by including the UMLS dictionary with synonyms in the adjusted cTAKES. An alternative to include an extensive number of synonyms is to implement a module for detection of misspellings and synonyms as described in the article by Kersloot et al. [2019]. This module assigned SNOMED CT expressions to extracted concepts based on a 5% error margin without additional FPs [Kersloot et al., 2019]. The abbreviations e.g. 'colon can' was incorrectly bound to the SNOMED CT expression '71854001 | Colon structure |' as 'can' is not a abbreviation for 'cancer' in the dictionary. The use of abbreviations in the discharge summaries are probably a part of the explanation for the low TP and FP rate, since too few clinical findings is extracted.
Another issue is the number of clinical findings, which represent the patient's anamnesis. In the discharge summaries, this is often indicated by the headline, but rules for including

headlines were not implemented in the adjusted cTAKES.
If the challenges concerning misspellings, synonyms, abbreviations, and identification of the patient's anamnesis were resolved, it would allow for more of the situational context to be included in the post-coordinated expressions. Though, it is questionable whether this would result in a better performance of the adjusted cTAKES. Therefore, it should be further investigated before drawing any conclusion.

7.5 Structuring of Healthcare Data

Despite the poor results from the adjusted cTAKES, some elements did contribute to improve the performance, hence extracting and encoding the discharge summaries. In the test data set 20 (0.2%) TP SNOMED CT expressions were identified. This was 18 pre-coordinated expressions and two post-coordinated expressions with a qualifier values as the only refinement. The agreement in pre-coordinated expressions is due to the fact that all SNOMED CT expressions from the top levels 'Clinical Finding' and 'Morphologic Abnormalities' were generated as pre-coordinated expressions. For the post-coordinated expressions the translation of the specific words and phrases from the discharge summaries into qualifier values resulted in a controlled choice of SNOMED CT expressions compared with SNOMED CT expressions in the remaining top levels. This positive effect is also seen in the result from the study by Kersloot et al. [2019] where the NLP system had an F-score of 0.845. In the study, three pre-defined pre-coordinated expressions and one pre-defined post-coordinated expression were bound to free text documents and were evaluated against a gold standard [Kersloot et al., 2019]. Therefore, an approach to improve the results of the adjusted cTAKES, could be to extend the process of making a controlled set of SNOMED CT expressions for each top level. However, clinical information are complex and ever changing [Garde and Knaup, 2006]. This makes it difficult to maintain an up to date controlled set of SNOMED CT expressions, which would be necessary in order to include the situational contextual information and obtain better performance.

To the authors current knowledge none of the identified NLP systems in table 5.5 on page 50 are used in clinical practice. This indicates the immaturity of the NLP systems and their ability to extract and encode the correct information. Semi-automatic annotation could possibly accommodate this hurdle, since the healthcare professional chose the best fitting pre-defined code for a given word or phrase. The study by de Keizer et al. [2008] investigated the match between reasons for transferring a patient to an intensive care unit. The reasons were collected as free text documents and pre-defined codes from a terminology system. Healthcare professionals wrote a free text document and chose one or more fitting codes to describe the same reasons. The results showed that only 11% of the reasons were an exact match, 79% did partly match, and 10% did not match. [de Keizer et al., 2008] These results could represent semi-automated annotation, and it was suggested that healthcare professionals should be involved when selecting a vocabulary for the results to improve [de Keizer et al., 2008]. However, it might not be sufficient to solve the challenges of extracting and encoding information from free text documents. The article by [Peterson et al., 2020] states that the work of structuring healthcare data should not be forced onto the healthcare professional, as they have an already full work load. Therefore, it is worth considering if NLP or semi-automatic annotation is a solution and whether it make sense to structure all information. According to the article by Greenhalgh et al. [2008] when

standardizing and structuring data in systems very carefully in one area it is likely to cause problems elsewhere in the system, since not all parts are equally structured. Combining the immaturity of NLP systems and difficulties structuring one part of a system, it might not be optimal to structure all data from free text documents.
An alternative could be to semi-structure free text documents by making the headlines more reusable and shareable, while leave the text be. In the article by Galster [2013a] it is stated that a healthcare professional uses a lot of time and mental capacity on collecting all relevant information if the information is scattered around the EHR. It is suggested to structure the information in smaller views, only presenting the necessary information for a given task [Galster, 2013a]. In the article by Galster [2013a] information is scattered around the EHR, whereas information is very dense in discharge summaries. Though, it is assumed that the idea of smaller views could be applicable for discharge summaries, since primary care physicians do not have to read the entire discharge summary. This could be obtained by using the semi-structured approach suggested above.

Conclusion 8

The objective of this study is to explore how the situational context can be preserved when extracting and structuring relevant information from free text documents in order to obtain semantic interoperability.
It is possible to preserve information about the situational context from free text documents, though some limitations are identified. The main limitation is that the clinical problems are not extracted and encoded sufficiently by the adjusted cTAKES which results in a poor performance. However, within the boundaries of the defined implementation context it is possible to structure the information using FHIR resources and achieve a semantic, unambiguous understanding of the clinical problems using SNOMED CT expression.

Based on the design considerations, an open-minded approach for selection of FHIR resources is chosen, which results in development of 20 FHIR profiles and one implementation guide. In total, the FHIR profiles can hold 95.7% of information from the discharge summaries. For the developed FHIR profiles to be useful in different settings, national or international standards shall be followed to set up requirements for which information shall be accessible to ensure interoperability.
A gold standard is elaborated with help from a healthcare professional, where clinical problems from the N2C2 data set are bound to pre- and post-coordinated expressions. Since 60.2% of the expressions from the test data set are post-coordinated expressions, it indicates that these are necessary to use in order to include the situational context.
The adjusted cTAKES is distinguishable different from the identified NLP systems, as it is adjusted to include the situational context through pre- and post-coordinated expressions. It is trained on three discharge summaries, but due to the poor performance, is shall be trained on more data before drawing conclusions on the overall performance.
In the literature it is found that there exist several NLP systems to extract and structure free text documents, though none of them are used in clinical practice. This indicates the difficulties of extracting information from free text documents well enough for clinical practise. In terms of the discharge summaries a suggestion can be to use a semi-structured approach with more explicit categorization of the information. Another suggestion can be to apply logic to the adjusted cTAKES, and learn the system to select an appropriate SNOMED CT expression.

Bibliography

Asuncion and van Sinderen, 2010. Camlon H. Asuncion and Marten J. van Sinderen. *Pragmatic interoperability: A systematic review of published definitions.* IFIP Advances in Information and Communication Technology, 326, 164–175, 2010. doi: 10.1007/978-3-642-15509-3_15.

Beale, 2001. Thomas Beale. *Archetypes: Constraint-based Domain Models for Future-proof Information Systems.* OOPSLA 2002 workshop on behavioural semantics, pages 1–69, 2001.

Bender and Sartipi, 2013. Duane Bender and Kamran Sartipi. *HL7 FHIR: An agile and RESTful approach to healthcare information exchange.* Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems, pages 326–331, 2013. doi: 10.1109/CBMS.2013.6627810.

Benson and Grieve, 2016a. Tim Benson and Grahame Grieve. *CDA – Clinical Document Architecture*, 2016a.

Benson and Grieve, 2016b. Tim Benson and Grahame Grieve. *Principles of FHIR*, 2016b.

Benson and Grieve, 2016c. Tim Benson and Grahame Grieve. *Why Interoperability Is Hard*, 2016c.

Berg, 2001. Marc Berg. *Implementing information systems in health care organizations: Myths and challenges.* International Journal of Medical Informatics, 64(2-3), 143–156, 2001. doi: 10.1016/S1386-5056(01)00200-3.

Bill et al., 2014. R Bill, S Pakhomov, E S Chen, T J Winden, E W Carter and G B Melton. *Automated extraction of family history information from clinical notes.* AMIA Symposium, 2014, 1709–1717, 2014.

Bozkurt et al., 2019. Selen Bozkurt, Emel Alkim, Imon Banerjee and Daniel L. Rubin. *Automated Detection of Measurements and Their Descriptors in Radiology Reports Using a Hybrid Natural Language Processing Algorithm.* Journal of digital imaging, 32 (4), 544–553, 2019. doi: 10.1007/s10278-019-00237-9.

Bramer et al., 2017. Wichor M. Bramer, Melissa L. Rethlefsen, Jos Kleijnen and Oscar H. Franco. *Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study.* Systematic Reviews, 6(1), 245, 2017. doi: 10.1186/s13643-017-0644-y.

Cai et al., 2019. T Cai, L Zhang, N Yang, K K Kumamaru, F J Rybicki, T Cai and K P Liao. *EXTraction of EMR numerical data: an efficient and generalizable tool to EXTEND clinical research.* BMC medical informatics and decision making, 19(1), 226, 2019. doi: 10.1186/s12911-019-0970-1.

Chen et al., 2009. Elizabeth S. Chen, Francine L. Maloney, Eugene Shilmayster and Howard S. Goldberg. *Laying the groundwork for enterprise-wide medical language processing services: architecture and process.* AMIA Symposium, 2009, 97–101, 2009.

Cheng et al., 2010. Lionel T.E. Cheng, Jiaping Zheng, Guergana K. Savova and Bradley J. Erickson. *Discerning tumor status from unstructured MRI reports–completeness of information in existing reports and utility of automated natural language processing.* Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology, 23(2), 119–132, 2010. doi: 10.1007/s10278-009-9215-7.

Coden et al., 2009. Anni Coden, Guergana Savova, Igor Sominsky, Michael Tanenblatt, James Masanz, Karin Schuler, James Cooper, Wei Guan and Piet C. de Groen. *Automatically extracting cancer disease characteristics from pathology reports into a Disease Knowledge Representation Model.* Journal of Biomedical Informatics, 42(5), 937–949, 2009. doi: 10.1016/j.jbi.2008.12.005.

Cote et al., 2017. D Cote, J Senders, A Karhade, S Gupta, N Lamba, B Hancock, T Smith and O Arnaout. *Natural language processing for the automated quantification of brain metastases in radiology free text reports.* Neuro-Oncology, 19, vi44–vi45, 2017. doi: 10.1093/neuonc/nox168.

de Keizer et al., 2008. Nicolette F. de Keizer, Ferishta Bakhshi-Raiez, Evert de Jonge and Ronald Cornet. *Post-coordination in practice: Evaluating compositional terminological system-based registration of ICU reasons for admission.* International Journal of Medical Informatics, 77(12), 828–835, 2008. doi: 10.1016/j.ijmedinf.2008.05.002.

Denny et al., 2009. J C Denny, R A Miller, L R Waitman, M A Arrieta and J F Peterson. *Identifying QT prolongation from ECG impressions using a general-purpose Natural Language Processor.* International Journal of Medical Informatics, 78(SUPPL. 1), 34–42, 2009. doi: 10.1016/j.ijmedinf.2008.09.001.

Department of Biomedical Informatics Harvard Medical School, 2018. Department of Biomedical Informatics Harvard Medical School. *N2C2 NLP Research Data Sets*, 2018. URL `https://portal.dbmi.hms.harvard.edu/`.

Dinh-Le et al., 2019. Catherine Dinh-Le, Rachel Chuang, Sara Chokshi and Devin Mann. *Wearable health technology and electronic health record integration: Scoping review and future directions.* JMIR mHealth and uHealth, 7(9), 2019. doi: 10.2196/12861.

Divita et al., 2014. G Divita, Q T Zeng, A V Gundlapalli, S Duvall, J Nebeker and M H Samore. *Sophia: A Expedient UMLS Concept Extraction Annotator.* AMIA Symposium, 2014, 467–476, 2014.

Doan et al., 2010. S Doan, L Bastarache, S Klimkowski, J C Denny and H Xu. *Integrating existing natural language processing tools for medication extraction from discharge summaries.* Journal of the American Medical Informatics Association, 17(5), 528–531, 2010. doi: 10.1136/jamia.2010.003855.

Elkin et al., 2006. Peter L Elkin, Steven H Brown, Casey S Husser and Brent A Bauer. *Evaluation of the Content Coverage of SNOMED CT: Ability of SNOMED Clinical Terms to Represent Clinical Problem Lists - ProQuest*. Mayo Clinic Proceedings, 81, 2006.

Epstein et al., 2013. Richard H. Epstein, Paul St Jacques, Michael Stockin, Brian Rothman, Jesse M. Ehrenfeld and Joshua C. Denny. *Automated identification of drug and food allergies entered using non-standard terminology*. Journal of the American Medical Informatics Association, 20(5), 962–968, 2013. doi: 10.1136/amiajnl-2013-001756.

Firely, 2020. Firely. *Firely Docs 3.5.1 documentation*, 2020. URL `https://docs.fire.ly/`.

Fu et al., 2020. J T Fu, E Sholle, S Krichevsky, J Scandura and T R Campion. *Extracting and classifying diagnosis dates from clinical notes: A case study*. Journal of Biomedical Informatics, 110, 2020. doi: 10.1016/j.jbi.2020.103569.

Galster, 2013a. Gert Galster. *Contextualisation of clinical information from fragmented health records*. Studies in health technology and informatics, 194, 41–47, 2013a. doi: 10.3233/978-1-61499-293-6-41.

Galster, 2013b. Gert Galster. *Production-contextual clinical information*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8175 LNAI, 248–258, 2013b. doi: 10.1007/978-3-642-40972-1_19.

Gao et al., 2015. Hongyuan Gao, Erin J. Aiello Bowles, David Carrell and Diana S.M. M Buist. *Using natural language processing to extract mammographic findings*. Journal of Biomedical Informatics, 54, 77–84, 2015. ISSN 1532-0464. doi: 10.1016/j.jbi.2015.01.010.

Garcelon et al., 4 2020. Nicolas Garcelon, Anita Burgun, Rémi Salomon and Antoine Neuraz. *Electronic health records for the diagnosis of rare diseases*. Kidney International, 97, 676–686, 2020. doi: 10.1016/j.kint.2019.11.037.

Garde and Knaup, 2006. Sebastian Garde and Petra Knaup. *Requirements engineering in health care: The example of chemotherapy planning in paediatric oncology*. Requirements Engineering, 11(4), 265–278, 2006. doi: 10.1007/s00766-006-0029-6.

Garla et al., 2011. V Garla, V L Re III., Z Dorey-Stein, F Kidwai, M Scotch, J Womack, A Justice and C Brandt. *The Yale cTAKES extensions for document classification: Architecture and application*. Journal of the American Medical Informatics Association, 18(5), 614–620, 2011. doi: 10.1136/amiajnl-2011-000093.

Gaudet-Blavignac et al., 2018. Christophe Gaudet-Blavignac, Vasiliki Foufi, Eric Wehrli and Christian Lovis. *Automatic annotation of French medical narratives with SNOMED CT concepts*. Studies in Health Technology and Informatics, 247, 710–714, 2018. doi: 10.3233/978-1-61499-852-5-710.

Gobbel et al., 2014. Glenn T. Gobbel, Ruth Reeves, Shrimalini Jayaramaraja, Dario Giuse, Theodore Speroff, Steven H. Brown, Peter L. Elkin and Michael E. Matheny. *Development and evaluation of RapTAT: A machine learning system for concept mapping of phrases from medical narratives.* Journal of Biomedical Informatics, 48, 54–65, 2014. doi: 10.1016/j.jbi.2013.11.008.

González Bernaldo de Quirós et al., 2018. Fernán González Bernaldo de Quirós, Carlos Otero and Daniel Luna. *Terminology Services: Standard Terminologies to Control Health Vocabulary.* Yearbook of medical informatics, 27(1), 227–233, 2018. doi: 10.1055/s-0038-1641200.

Goss et al., 2014. Foster R. Goss, Joseph M. Plasek, Jason J. Lau, Diane L. Seger, Frank Y. Chang and Li Zhou. *An evaluation of a natural language processing tool for identifying and encoding allergy information in emergency department clinical notes.* AMIA Symposium, 2014, 580–588, 2014.

Greenhalgh et al., 2008. Trisha Greenhalgh, Henry W.W. Potts, Geoff Wong, Pippa Bark and Deborah Swinglehurst. *Tensions and Paradoxes in Electronic Patient Record Research: A Systematic Literature Review Using the Meta-narrative Method.* The Milbank quarterly, 86(4), 529–32, 2008. doi: 10.1111/j.1468-0009.2008.00538.x.

Gu et al., 2016. H Gu, Y Chen, Z He, M Halper, L Chen and Huanying Gu. *Quality Assurance of UMLS Semantic Type Assignments to SNOMED CT Concepts with the Use of SNOMED CT Hierarchies HHS Public Access.* Methods Inf Med, 55(2), 158–165, 2016. doi: 10.3414/ME14-01-0104.

Guan and Devarakonda, 2019. H Guan and M Devarakonda. *Leveraging Contextual Information in Extracting Long Distance Relations from Clinical Notes.* AMIA Symposium, 2019, 1051–1060, 2019.

Hamon and Grabar, 2010. Thierry Hamon and Natalia Grabar. *Linguistic approach for identification of medication names and related information in clinical narratives.* Journal of the American Medical Informatics Association, 17(5), 549–554, 2010. doi: 10.1136/jamia.2010.004036.

Hao et al., 2016. Tianyong Hao, Hongfang Liu and Chunhua Weng. *Valx: A System for Extracting and Structuring Numeric Lab Test Comparison Statements from Text.* Methods of information in medicine, 55(3), 266–275, 2016. doi: 10.3414/ME15-01-0112.

Health Level 7, 2021a. Health Level 7. *Resource Bundle - Content*, 2021a. URL `https://www.hl7.org/fhir/document-example-dischargesummary.xml.html`.

Health Level 7, 2021b. Health Level 7. *Resource ImplementationGuide - Content*, 2021b. URL `https://www.hl7.org/fhir/implementationguide.html`.

Health Level 7, 2021c. Health Level 7. *Resource Index*, 2021c. URL `https://www.hl7.org/fhir/resourcelist.html`.

Health Level 7, 2021d. Health Level 7. *FHIR Documents*, 2021d. URL `https://www.hl7.org/fhir/documents.html`.

Health Level 7, **2021e**. Health Level 7. *Messaging using FHIR Resources*, 2021e. URL https://www.hl7.org/fhir/messaging.html.

Health Level 7, **2021f**. Health Level 7. *Home - FHIR v4.0.1*, 2021f. URL http://hl7.org/fhir/.

Health Level 7, **2021g**. Health Level 7. *Conformance-rules - FHIR v4.0.1*, 2021g. URL https://www.hl7.org/fhir/conformance-rules.html.

Health Level 7, **2021h**. Health Level 7. *Data Types*, 2021h. URL https://www.hl7.org/fhir/datatypes.html.

Health Level 7, **2021i**. Health Level 7. *Extensibility Registry*, 2021i. URL https://www.hl7.org/fhir/extensibility-registry.html.

Health Level 7, **2021j**. Health Level 7. *Flag - FHIR v4.0.1*, 2021j. URL https://www.hl7.org/fhir/flag.html.

Health Level 7, **2021k**. Health Level 7. *Principles*, 2021k. URL https://www.hl7.org/fhir/overview-arch.html#principles.

Health Level 7, **2021l**. Health Level 7. *Slicing*, 2021l. URL https://www.hl7.org/fhir/profiling.html#slicing.

Health Level 7, **2021m**. Health Level 7. *Terminologies*, 2021m. URL https://www.hl7.org/fhir/terminologies.html.

Hong et al., **2018**. Na Hong, Andrew Wen, Feichen Shen, Sunghwan Sohn, Sijia Liu, Hongfang Liu and Guoqian Jiang. *Integrating Structured and Unstructured EHR Data Using an FHIR-based Type System: A Case Study with Medication Data.* AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2017, 74–83, 2018.

John Hunt. Agile Methods and the Agile Manifesto. In *Agile Software Construction*, pages 9–30. Springer London, London, 2006. doi: 10.1007/1-84628-262-4_2.

IHTSDO, **2021a**. International Health Terminology Standards Development Organisation IHTSDO. *SNOMED CT Browser*, 2021a. URL https://browser.ihtsdotools.org/?perspective=full{&}conceptId1=404684003{&}edition=MAIN/2021-01-31{&}release={&}languages=en.

IHTSDO, **2021b**. International Health Terminology Standards Development Organisation IHTSDO. *Compositional Grammar - Specification and Guide*, 2021b. URL https://confluence.ihtsdotools.org/display/DOCSCG/Compositional+Grammar+-+Specification+and+Guide.

IHTSDO, **2021c**. International Health Terminology Standards Development Organisation IHTSDO. *SNOMED CT Starter Guide*, 2021c. URL https://confluence.ihtsdotools.org/display/DOCSTART/SNOMED+CT+Starter+Guide.

IHTSDO, **2021d**. International Health Terminology Standards Development Organisation IHTSDO. *SNOMED CT MRCM Maintenance Tool*, 2021d. URL https://browser.ihtsdotools.org/mrcm/?branch=MAIN%2F2021-01-31.

Ingenerf, 1999. Josef Ingenerf. *Telemedicine and terminology: Different needs of context information.* IEEE Transactions on Information Technology in Biomedicine, 3 (2), 92–100, 1999. doi: 10.1109/4233.767084.

Iqbal et al., 2017. Ehtesham Iqbal, Robbie Mallah, Daniel Rhodes, Honghan Wu, Alvin Romero, Nynn Chang, Olubanke Dzahini, Chandra Pandey, Matthew Broadbent, Robert Stewart, Richard J.B. B Dobson and Zina M. Ibrahim. *ADEPt, a semantically-enriched pipeline for extracting adverse drug events from free-text electronic health records.* PLoS ONE, 12(11), 2017. doi: 10.1371/journal.pone.0187121.

ISO/TC 215, 2005. ISO/TC 215. *ISO/TR 17119:2005(en), Health informatics - Health informatics profiling framework*, 2005. URL `https://www.iso.org/obp/ui/#iso:std:iso:tr:17119:ed-1:v1:en`.

Jensen et al., 2012. Peter B Jensen, Lars J Jensen and Søren Brunak. *Mining electronic health records: towards better research applications and clinical care.* Nature Reviews, 13, 2012. doi: 10.1038/nrg3208.

Jindal and Roth, 2013. Prateek Jindal and Dan Roth. *Extraction of events and temporal expressions from clinical narratives.* Journal of Biomedical Informatics, 46 (SUPPL.), 2013. doi: 10.1016/j.jbi.2013.08.010.

Johnson et al., 2016. Alistair E.W. Johnson, Tom J. Pollard, Lu Shen, Li Wei H. Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi and Roger G. Mark. *MIMIC-III, a freely accessible critical care database.* Scientific Data, 3(1), 1–9, 2016. doi: 10.1038/sdata.2016.35.

Jones et al., 2020. Kerina H. Jones, Elizabeth M. Ford, Nathan Lea, Lucy J. Griffiths, Lamiece Hassan, Sharon Heys, Emma Squires and Goran Nenadic. *Toward the development of data governance standards for using clinical free-text data in health research: Position paper.* Journal of Medical Internet Research, 22(6), 2020. doi: 10.2196/16760.

Jünger and Nagel, 2019. Jana Jünger and Eckhard Nagel. *Improvement of patient orientation and patient safety in health care – from competency-based and interprofessional medical education to cross-sector care.* GMS Journal for Medical Education, 36(3), 1–7, 2019. doi: 10.3205/zma001240.

Kersloot et al., 2019. Martijn G. Kersloot, Francis Lau, Ameen Abu-Hanna, Derk L. Arts and Ronald Cornet. *Automated SNOMED CT concept and attribute relationship detection through a web-based implementation of cTAKES.* Journal of biomedical semantics, 10(1), 14, 2019. doi: 10.1186/s13326-019-0207-3.

Knoll et al., 2019. B C Knoll, E A Lindemann, A L Albert, G B Melton and S V S Pakhomov. *Recurrent Deep Network Models for Clinical NLP Tasks: Use Case with Sentence Boundary Disambiguation.* Studies in health technology and informatics, 264, 198–202, 2019. doi: 10.3233/SHTI190211.

Kovačević et al., 2013. Aleksandar Kovačević, Azad Dehghan, Michele Filannino, John A. Keane and Goran Nenadic. *Combining rules and machine learning for* *extraction of temporal expressions and events from clinical narratives.* Journal of the American Medical Informatics Association, 20(5), 859–866, 2013. doi: 10.1136/amiajnl-2013-001625.

Kulshrestha et al., 2020. Sujay Kulshrestha, Dmitriy Dligach, Cara Joyce, Marshall S Baker, Richard Gonzalez, Ann P O'Rourke, Joshua M Glazer, Anne Stey, Jacqueline M Kruser, Matthew M Churpek and Majid Afshar. *Prediction of severe chest injury using natural language processing from the electronic health record.* Injury, 52(2), 205–212, 2020. doi: 10.1016/j.injury.2020.10.094.

Leaman et al., 2015. Robert Leaman, Ritu Khare and Zhiyong Lu. *Challenges in clinical natural language processing for automated disorder normalization.* Journal of Biomedical Informatics, 57, 28–37, 2015. doi: 10.1016/j.jbi.2015.07.010.

Lenert et al., 2014. Leslie A. Lenert, Farrant H. Sakaguchi and Charlene R. Weir. *Rethinking the discharge summary: A focus on handoff communication.* Academic Medicine, 89(3), 393–398, 2014. doi: 10.1097/ACM.0000000000000145.

Li et al., 2020. M D Li, M Lang, F Deng, K Chang, K Buch, S Rincon, W A Mehan, T M Leslie-Mazwi and J Kalpathy-Cramer. *Analysis of Stroke Detection during the COVID-19 Pandemic Using Natural Language Processing of Radiology Reports.* AJNR. American journal of neuroradiology, 2020. doi: 10.3174/ajnr.A6961.

Lin et al., 2015. Ching Heng Lin, Nai Yuan Wu, Wei Shao Lai and Der Ming Liou. *Comparison of a semi-automatic annotation tool and a natural language processing application for the generation of clinical statement entries.* Journal of the American Medical Informatics Association, 22(1), 132–142, 2015. doi: 10.1136/amiajnl-2014-002991.

Liu et al., 2019a. Cong Liu, Casey N. Ta, James R. Rogers, Ziran Li, Junghwan Lee, Alex M. Butler, Ning Shang, Fabricio Sampaio Peres Kury, Liwei Wang, Feichen Shen, Hongfang Liu, Lyudmila Ena, Carol Friedman and Chunhua Weng. *Ensembles of natural language processing systems for portable phenotyping solutions.* Journal of Biomedical Informatics, 100, 2019a. doi: 10.1016/j.jbi.2019.103318.

Liu et al., 2019b. Feifan Liu, Richeek Pradhan, Emily Druhl, Elaine Freund, Weisong Liu, Brian C. Sauer, Fran Cunningham, Adam J. Gordon, Celena B. Peters and Hong Yu. *Learning to detect and understand drug discontinuation events from clinical narratives.* Journal of the American Medical Informatics Association, 26(10), 943–951, 2019b. doi: 10.1093/jamia/ocz048.

Liu et al., 2012. Hongfang Liu, Kavishwar Wagholikar, Stephen Tze and Inn Wu. *Using SNOMED-CT to encode summary level data - a corpus analysis*, 2012.

Lou et al., 2020. Robert Lou, Darco Lalevic, Charles Chambers, Hanna M. Zafar and Tessa S. Cook. *Automated Detection of Radiology Reports that Require Follow-up Imaging Using Natural Language Processing Feature Engineering and Machine Learning Classification.* Journal of digital imaging, 33(1), 131–136, 2020. doi: 10.1007/s10278-019-00271-7.

Lussier et al., 2001. Yves A Lussier, Lyudmila Shagina' and Carol Friedman. *Automating SNOMED Coding Using Medical Language Understanding: A Feasibility Study.* Journal of the American Medical Informatics Association, 2001.

Martinez et al., 2014. David Martinez, Graham Pitson, Andrew MacKinlay and Lawrence Cavedon. *Cross-hospital portability of information extraction of cancer staging information.* Artificial Intelligence in Medicine, 62(1), 11–21, 2014. doi: 10.1016/j.artmed.2014.06.002.

David P. McCallie Jr. Chapter 28 - Emerging Clinical Decision Support Technology for the Twenty First Century. In *Healthcare Information Management Systems: Cases, Strategies, and Solutions: Fourth Edition*, pages 477–491. 2016. ISBN 9783319207650. doi: 10.1007/978-3-319-20765-0.

McCart et al., 2013. J A McCart, D J Berndt, J Jarman, D K Finch and S L Luther. *Finding falls in ambulatory care clinical documents using statistical text mining.* Journal of the American Medical Informatics Association, 20(5), 906–914, 2013. doi: 10.1136/amiajnl-2012-001334.

MedCom, 2019. MedCom. *Den gode XML udskrivningsepikrise.* pages 1–41, 2019. URL `http://svn.medcom.dk/svn/releases/Standarder/Dengodeepikrise/XML/XDIS01/Dokumentation/XDIS01.pdf`.

Medcom, 2021. Medcom. *Medcom*, 2021. URL `https://www.medcom.dk`.

Meystre et al., 2008. S. M. Meystre, G. K. Savova, K. C. Kipper-Schuler and J. F. Hurdle. *Extracting information from textual documents in the electronic health record: a review of recent research.* Yearbook of medical informatics, pages 128–144, 2008. doi: 10.1055/s-0038-1638592.

Meystre et al., 2010a. S M Meystre, J Thibault, S Shen, J F Hurdle and B R South. *Automatically detecting medications and the reason for their prescription in clinical narrative text documents.* Studies in Health Technology and Informatics, 160(PART 1), 944–948, 2010a. doi: 10.3233/978-1-60750-588-4-944.

Meystre et al., 2010b. S M Meystre, J Thibault, S Shen, J F Hurdle and B R South. *Textractor: A hybrid system for medications and reason for their prescription extraction from clinical text documents.* Journal of the American Medical Informatics Association, 17(5), 559–562, 2010b. doi: 10.1136/jamia.2010.004028.

Meystre et al., 2017. S. M. Meystre, C. Lovis, T. Bürkle, G. Tognola, A. Budrionis and C. U. Lehmann. *Clinical Data Reuse or Secondary Use: Current Status and Potential Future Progress.* Yearbook of medical informatics, 26(1), 38–52, 2017. doi: 10.15265/IY-2017-007.

Mishra et al., 2019. Rashmi Mishra, Andrea Burke, Bonnie Gitman, Payal Verma, Mark Engelstad, Melissa A. Haendel, Ilias Alevizos, William A. Gahl, Michael T. Collins, Janice S. Lee and Murat Sincan. *Data-driven method to enhance craniofacial and oral phenotype vocabularies.* Journal of the American Dental Association (1939), 150(11), 933–939, 2019. doi: 10.1016/j.adaj.2019.05.029.

Moher et al., 2009. David Moher, Alessandro Liberati, Jennifer Tetzlaff and Douglas G. Altman. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. PLoS Medicine, 6, 2009. doi: 10.1371/journal.pmed.1000097.

Moon et al., 2019. Sungrim Moon, Sijia Liu, Christopher G. Scott, Sujith Samudrala, Mohamed M. Abidian, Jeffrey B. Geske, Peter A. Noseworthy, Jane L. Shellum, Rajeev Chaudhry, Steve R. Ommen, Rick A. Nishimura, Hongfang Liu and Adelaide M. Arruda-Olson. *Automated extraction of sudden cardiac death risk factors in hypertrophic cardiomyopathy patients by natural language processing*. International Journal of Medical Informatics, 128, 32–38, 2019. doi: 10.1016/j.ijmedinf.2019.05.008.

MTHelpLine, 2021. MTHelpLine. *Transcribed Medical Transcription Sample Reports and Examples - MTSamples*, 2021. URL `https://mtsamples.com/`.

Nassif et al., 2009. Houssam Nassif, Ryan Woods, Elizabeth Burnside, Mehmet Ayvaci, Jude Shavlik and David Page. *Information extraction for clinical data mining: A mammography case study*. ICDM Workshops 2009 - IEEE International Conference on Data Mining, pages 37–42, 2009. doi: 10.1109/ICDMW.2009.63.

Nelson, 1997. Donald A. F. Nelson. *A Defined Minimum Data Set: Will it Work for Direct Patient Care?* Computers for Nursing, 15(2), 43–47, 1997.

O'Leary et al., 2009. Kevin J. O'Leary, David M. Liebovitz, Joseph Feinglass, David T. Liss, Daniel B. Evans, Nita Kulkarni, Matthew P. Landler and David W. Baker. *Creating a better discharge summary: Improvement in quality and timeliness using an electronic discharge summary*. Journal of Hospital Medicine, 4(4), 219–225, 2009. doi: 10.1002/jhm.425.

Oleynik et al., 2019. Michel Oleynik, Amila Kugic, Zdenko Kasáč and Markus Kreuzthaler. *Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification*. Journal of the American Medical Informatics Association, 26(11), 1247–1254, 2019. doi: 10.1093/jamia/ocz149.

OpenEHR International, 2021. OpenEHR International. *What is OpenEHR?*, 2021. URL `https://www.openehr.org/about/what_is_openehr`.

Patel et al., 2002. Vimla L. Patel, José F. Arocha and André W. Kushniruk. *Patients' and physicians' understanding of health and biomedical concepts: Relationship to the design of EMR systems*. Journal of Biomedical Informatics, 35(1), 8–16, 2002. doi: 10.1016/S1532-0464(02)00002-3.

Peng et al., 2020. Cong Peng, Prashant Goswami and Guohua Bai. *A literature review of current technologies on health data integration for patient-centered health management*. Health Informatics Journal, 26(3), 1926–1951, 2020. doi: 10.1177/1460458219892387.

Peterson et al., 2020. Kevin J. Peterson, Guoqian Jiang and Hongfang Liu. *A corpus-driven standardization framework for encoding clinical problems with HL7 FHIR*. Journal of Biomedical Informatics, 110, 2020. ISSN 15320464. doi: 10.1016/j.jbi.2020.103541.

Pickard, 2019. Hanna Pickard. *Responsibility in healthcare: What's the point?* Journal of Medical Ethics, 45(10), 650–651, 2019. doi: 10.1136/medethics-2019-105829.

Qiu et al., 2018. John X. Qiu, Hong Jun H.-J. Yoon, Paul A. Fearn and Georgia D. Tourassi. *Deep Learning for Automated Extraction of Primary Sites from Cancer Pathology Reports.* IEEE Journal of Biomedical and Health Informatics, 22(1), 244–251, 2018. doi: 10.1109/JBHI.2017.2700722.

Rinner and Duftschmid, 2016. Christoph Rinner and Georg Duftschmid. *Bridging the gap between HL7 CDA and HL7 FHIR: A JSOn based mapping.* Studies in Health Technology and Informatics, 223, 100–106, 2016. doi: 10.3233/978-1-61499-645-3-100.

Roberts, 2017. Angus Roberts. *Language, structure, and reuse in the electronic health record.* AMA Journal of Ethics, 19(3), 281–288, 2017. ISSN 23766980. doi: 10.1001/journalofethics.2017.19.3.stas1-1703.

Rosenbloom et al., 2011. S. Trent Rosenbloom, Joshua C. Denny, Hua Xu, Nancy Lorenzi, William W. Stead and Kevin B. Johnson. *Data from clinical notes: A perspective on the tension between structure and flexible documentation.* Journal of the American Medical Informatics Association, 18(2), 181–186, 2011. doi: 10.1136/jamia.2010.007237.

Savova et al., 2010. Guergana K. Savova, James J. Masanz, Philip V. Ogren, Jiaping Zheng, Sunghwan Sohn, Karin C. Kipper-Schuler and Christopher G. Chute. *Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications.* Journal of the American Medical Informatics Association, 17(5), 507–513, 2010. doi: 10.1136/jamia.2009.001560.

Sevenster et al., 2012. Merlijn Sevenster, Rob Van Ommering and Yuechen Qian. *Automatically correlating clinical findings and body locations in radiology reports using MedLEE.* Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology, 25(2), 240–249, 2012. doi: 10.1007/s10278-011-9411-0.

Shah et al., 2019. A D Shah, E Bailey, T Williams, S Denaxas, R Dobson and H Hemingway. *Natural language processing for disease phenotyping in UK primary care records for research: a pilot study in myocardial infarction and death.* Journal of biomedical semantics, 10, 20, 2019. doi: 10.1186/s13326-019-0214-4.

Simplifier.net, 2021a. Simplifier.net. *Best Practices.* 2021a. URL `https://simplifier.net/guide/profilingacademy/Best-practices`.

Simplifier.net, 2021b. Simplifier.net. *Get started with Simplifier*, 2021b. URL `https://simplifier.net/guide/ProfilingAcademy/GetstartedwithSimplifier`.

Spasic and Nenadic, 2020. Irena Spasic and Goran Nenadic. *Clinical Text Data in Machine Learning: Systematic Review.* JMIR medical informatics, 8(3), e17984, 2020. doi: 10.2196/17984.

Stacey et al., 2017. D Stacey, F Légaré, K Lewis, Barry Mj, Bennett Cl, Eden Kb, A Lyddiatt, R Thomson and L Trevena. *Decision aids for people facing health* *treatment or screening decisions (Review)*. Cochrane Database of Systematic Reviews, 2017. doi: 10.1002/14651858.CD001431.

Tao et al., **2018**. Carson Tao, Michele Filannino and Özlem Uzuner. *FABLE: A Semi-Supervised Prescription Information Extraction System*. AMIA Symposium, 2018, 1534–1543, 2018.

Topaz et al., **2019**. Maxim Topaz, Ludmila Murga, Katherine M. Gaddis, Margaret V. McDonald, Ofrit Bar-Bachar, Yoav Goldberg and Kathryn H. Bowles. *Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approaches*. Journal of Biomedical Informatics, 90, 2019. doi: 10.1016/j.jbi.2019.103103.

Treebank, **2003**. Penn Treebank. *Alphabetical list of part-of-speech tags used in the Penn Treebank Project*, 2003. URL `https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html`.

Trivedi et al., **2019**. Hari M. Trivedi, Maryam Panahiazar, April Liang, Dmytro Lituiev, Peter Chang, Jae Ho Sohn, Yunn Yi Y.-Y. Chen, Benjamin L. Franc, Bonnie Joe and Dexter Hadley. *Large Scale Semi-Automated Labeling of Routine Free-Text Clinical Records for Deep Learning*. Journal of digital imaging, 32(1), 30–37, 2019. doi: 10.1007/s10278-018-0105-8.

VeraTech for Health, **2021**. VeraTech for Health. *ISO 13606 Standard - EHR Interoperability*, 2021. URL `http://www.en13606.org/information.html`.

Viani et al., **2018**. Natalia Viani, Lorenzo Chiudinelli, Cristina Tasca, Alberto Zambelli, Mauro Bucalo, Arianna Ghirardi, Nicola Barbarini, Eleonora Sfreddo, Lucia Sacchi, Carlo Tondini and Riccardo Bellazzi. *Automatic processing of anatomic pathology reports in the Italian language to enhance the reuse of clinical data*. Studies in Health Technology and Informatics, 247, 715–719, 2018. doi: 10.3233/978-1-61499-852-5-715.

Wei et al., **2019**. Qiang Wei, Zongcheng Ji, Yuqi Si, Jingcheng Du, Jingqi Wang, Firat Tiryaki, Stephen Wu, Cui Tao, Kirk Roberts and Hua Xu. *Relation Extraction from Clinical Narratives Using Pre-trained Language Models*. AMIA Symposium, 2019, 1236–1245, 2019.

Wong et al., **2018**. Adrian Wong, Joseph M. Plasek, Steven P. Montecalvo and Li Zhou. *Natural Language Processing and Its Implications for the Future of Medication Safety: A Narrative Review of Recent Advances and Challenges*. Pharmacotherapy, 38(8), 822–841, 2018. doi: 10.1002/phar.2151.

Yadav et al., **2013**. Kabir Yadav, Efsun Sarioglu, Meaghan Smith and Hyeong Ah H.-A. Choi. *Automated outcome classification of emergency department computed tomography imaging reports*. Academic Emergency Medicine, 20(8), 848–854, 2013. doi: 10.1111/acem.12174.

Yang et al., **2020a**. Xi Yang, Jiang Bian, Ruogu Fang, Ragnhildur I. Bjarnadottir, William R. Hogan and Yonghui Wu. *Identifying relations of medications with adverse* *drug events using recurrent convolutional neural networks and gradient boosting.* Journal of the American Medical Informatics Association, 27(1), 65–72, 2020a. doi: 10.1093/jamia/ocz144.

Yang et al., 2020b. Xi Yang, Hansi Zhang, Xing He, Jiang Bian and Yonghui Wu. *Extracting Family History of Patients From Clinical Narratives: Exploring an End-to-End Solution With Deep Learning Models.* JMIR medical informatics, 8(12), e22982, 2020b. doi: 10.2196/22982.

Zheng et al., 2012. Jiaping Zheng, Wendy W. Chapman, Timothy A. Miller, Chen Lin, Rebecca S. Crowley and Guergana K. Savova. *A system for coreference resolution for the clinical narrative.* Journal of the American Medical Informatics Association, 19(4), 660–667, 2012. doi: 10.1136/amiajnl-2011-000599.

Zhou et al., 2011. L. Zhou, Joseph M. Plasek, Lisa M. Mahoney, Neelima Karipineni, Frank Chang, Xuemin Yan, Fenny Chang, Dana Dimaggio, Debora S. Goldman and Roberto A. Rocha. *Using Medical Text Extraction, Reasoning and Mapping System (MTERMS) to process medication information in outpatient clinical notes.* AMIA Symposium, 2011, 1639–1648, 2011.

Method to Obtain Current Knowledge A

To obtain knowledge for chapter 2 on page 2 three methods were used, 1) exploratory search to obtain a knowledge base and vocabulary, 2) a structured literature search in terms of a block search to investigate of structuring free text documents in order to achieve semantic interoperability, and 3) chaining search to further explore information from the included articles. The block search is further described in the following section.

A.1 Structured Literature Search

The purpose of the structured literature search was to investigate structuring of free text documents in order to achieve semantic interoperability. The block search was conducted in the databases PubMed and Embase, as a retrieval rate of 92.8% was found according to the article by Bramer et al. [2017] when searching in these two databases, which was assumed sufficient. The search terms used in PubMed and Embase, can be seen in table A.1 and A.2, respectively. The two columns are combined using the boolean operator 'AND', where the rows in each column are combined using the boolean operator 'OR'.

Clinical notes	**Structuring**	**Interoperability**
	MeSH terms	
Narrative Medicine[MeSH]		Health Information Interoperability[MeSH]
		Data Curation[MeSH]
	All fields	
Clinical document*	Structuring	Interoperability
Clinic* note*	Data structur*	Interoperable
Clinical narrative*	Organizing	Reuse
Clinical dicta*		
Medical document*		
Medical note*		
Medical narrative*		
Medical dicta*		
Free text		

Table A.1. shows the search terms used in the structured literature search conducted in PubMed.

Clinical notes	**Structuring**	**Interoperability**
	Emtree terms	
Medical documentation		Data interoperability
	All fields	
Clinical document*	Structuring	Interoperability
Clinic* note*	Data structur*	Interoperable
Clinical narrative*	Organizing	Reuse
Clinical dicta*		
Medical document*		
Medical note*		
Medical narrative*		
Medical dicta*		
Free text		

Table A.2. shows the search terms used in the structured literature search conducted in Embase.

The articles from the two databases were combined and duplicates were removed. The remaining articles were initially screened based on title and abstract, which was followed by a reading of the articles in full length. The following inclusion and exclusion criteria had to be respected for an article to be included.

Inclusion criteria:

- I1: The processed data are free text documents.

Exclusion criteria:

- E1: The study does not describe the possibilities and/or limitations for interoperability of the processed data.
- E2: Not available in full length.
- E3: The language is different from Danish or English.

Inclusion and exclusion of articles can be seen in figure A.1, resulting in inclusion of eight articles. These articles were then assessed based on several criteria, which can be seen in section A.2.

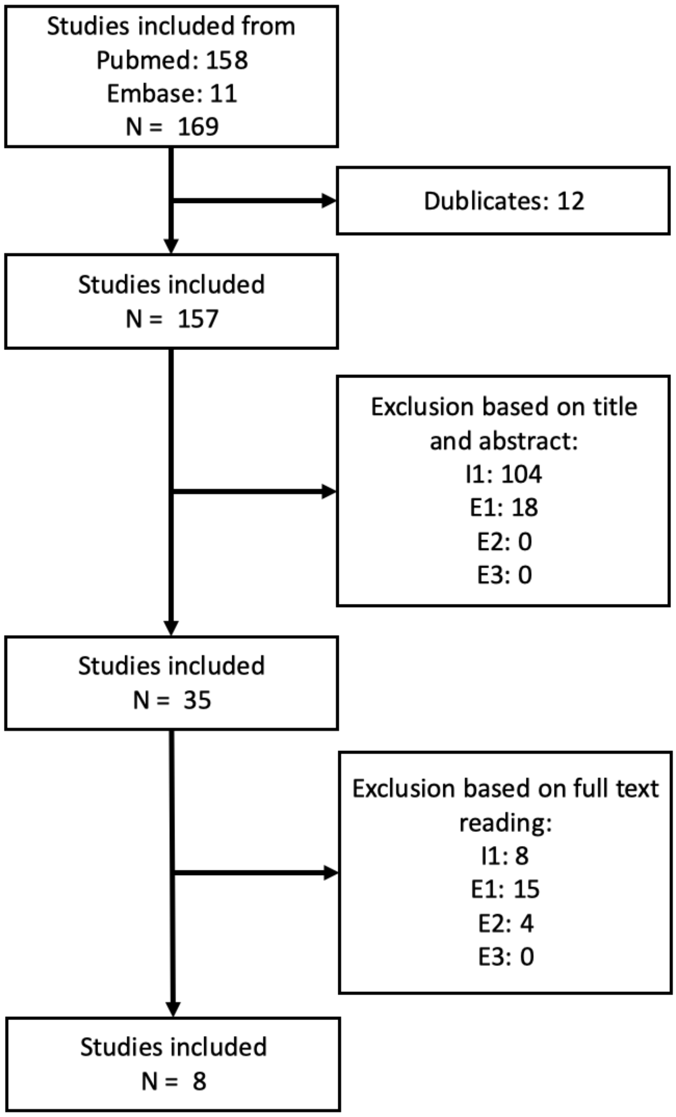

Figure A.1. shows the number of included and excluded articles from the structured literature search.

A.2 Assessment of Included Articles

The included articles were of three different types: review, opinion papers, and development studies. The review was assessed using PRISMA [Moher et al., 2009], where the development studies were assessed using relevance, reliability, internal validity, and external validity. Opinion papers are at the bottom of the evidence hierarchy and are therefore not assessed, but will get a low assessment. The results of the assessment can be seen in table A.3, which shows that the articles range from low to high. All articles are used in the Problem Analysis in chapter 2. Statements from articles with a low assessment will not stand alone, but have their statement supported by one or more articles. Statements from articles with a medium or high assessment can stand alone.

Article	Relevance	Reliability	Internal Validity	External Validity	Overallassessment
		Review			
Garcelon et al. [2020]	-	-	-	-	Medium
		Development Study			
Chen et al. [2009]	Medium	Medium	High	Medium	Medium
Gaudet-Blavignac et al. [2018]	High	High	Medium	High	High
Lin et al. [2015]	High	High	High	High	High
Peterson et al. [2020]	High	Medium	High	Medium	High
		Opinion Paper			
González Bernaldo de Quirós et al. [2018]	-	-	-	-	Low
Lussier et al. [2001]	-	-	-	-	Low
Roberts [2017]	-	-	-	-	Low

Table A.3. shows the assessment of each article or paper included from the structured literature search.

Sub-elements in NLP Systems B

Table B on the following page shows the sub-elements mentioned in the articles identified trough the structured literature search, see chapter 5.1.

Reference	**Tokenization**	**Sentence splitting**	**Negation detection**	**Normalization**	**Term. mapping**	**POS-tagging**	**Section splitting**	**Remove characters**	**Context detection**	**Lower casing**	**NER**	**Lemmatization/stemming**	**Spell check**	**Temporal parser/tagger**	**Concept identification**	**Remove stop words**	**Chunking**	**Relation identification**	**Disambiguation**	**Word splitting**	**Shallow parser**	**CUI**	**Dependence parsing**	**Abbreviation**	**Acronyms**	**Context assertion**	**Syntactic parsing**
Bill et al. [2014]							x																				
Bozkurt et al. [2019]							x																				
Cai et al. [2019]	x			x	x																						
Cheng et al. [2010]			x																								
Coden et al. [2009]	x	x	x			x									x			x			x						
Cote et al. [2017]	x							x	x	x		x				x				x							
Denny et al. [2009]			x																								
Divita et al. [2014]				x											x												
Doan et al. [2010]		x	x		x								x														
Epstein et al. [2013]		x			x			x					x			x											
Fu et al. [2020]				x																							
Gao et al. [2015]	x	x	x				x																				
Garla et al. [2011]	x	x	x			x	x				x	x					x										
Gobbel et al. [2014]	x				x	x				x		x				x											
Goss et al. [2014]	x	x	x	x	x		x				x			x					x					x			
Guan and Devarakonda [2019]		x	x																								
Hamon and Grabar [2010]	x	x	x		x	x	x				x	x							x	x							

Continued on next page

Reference	Tokenization	Sentence splitting	Negation detection	Normalization	Term. mapping	POS-tagging	Section splitting	Remove characters	Context detection	Lower casing	NER	Lemmatization/stemming	Spell check	Temporal parser/tagger	Concept identification	Remove stop words	Chunking	Relation identification	Disambiguation	Word splitting	Shallow parser	CUI	Dependence parsing	Abbreviation	Acronyms	Context assertion	Syntactic parsing
Hao et al. [2016]		x	x	x	x								x													x	
Iqbal et al. [2017]	x	x				x																					
Jindal and Roth [2013]			x																								
Kersloot et al. [2019]	x	x	x		x	x	x					x		x			x	x									
Knoll et al. [2019]	x																										
Kovačević et al. [2013]	x	x	x	x	x	x								x			x					x					
Kulshrestha et al. [2020]	x	x		x				x	x				x			x						x					
Li et al. [2020]			x									x															
Liu et al. [2019a]	x	x			x				x		x							x									
Liu et al. [2019b]	x			x					x	x		x															
Lou et al. [2020]	x							x		x						x											
Martinez et al. [2014]	x		x	x				x																			
McCart et al. [2013]	x			x				x		x																	
Meystre et al. [2010b]	x	x				x	x		x					x					x								
Meystre et al. [2010a]																											
Mishra et al. [2019]	x	x				x	x				x						x										
Moon et al. [2019]	x	x	x												x												

Continued on next page

Reference	Tokenization	Sentence splitting	Negation detection	Normalization	Term. mapping	POS-tagging	Section splitting	Remove characters	Context detection	Lower casing	NER	Lemmatization/stemming	Spell check	Temporal parser/tagger	Concept identification	Remove stop words	Chunking	Relation identification	Disambiguation	Word splitting	Shallow parser	CUI	Dependence parsing	Abbreviation	Acronyms	Context assertion	Syntactic parsing
Nassif et al. [2009]			x										x		x												x
Oleynik et al. [2019]	x	x	x					x	x	x						x											
Peterson et al. [2020]															x			x					x				
Qiu et al. [2018]	x							x		x																	
Savova et al. [2010]	x	x		x		x					x										x						
Sevenster et al. [2012]		x		x																							
Shah et al. [2019]			x		x			x		x										x							
Tao et al. [2018]	x	x		x		x								x				x									
Topaz et al. [2019]			x					x	x	x																	
Trivedi et al. [2019]	x				x				x				x												x		
Wei et al. [2019]	x	x																									
Yadav et al. [2013]					x								x														
Yang et al. [2020b]	x	x		x							x				x												
Yang et al. [2020a]	x	x		x																							
Zheng et al. [2012]	x	x				x					x									x	x		x				
Zhou et al. [2011]	x	x	x	x	x		x		x		x			x			x		x								
	32	27	21	16	14	13	11	10	10	9	9	7	7	7	6	6	5	5	5	4	3	2	2	1	1	1	1

Table B.1. shows the sub-elements used in NLP systems developed in the included article from the structured literature search. POS-tagging = Part-of-speech tagging. Term. mapping = Terminology mapping. NER = Named Entity Recognition. CUI = Concept unique identifier.

Procedure for Validating SNOMED CT expressions C

This appendix describes the procedure for interviewing a healthcare professional regarding relevant, related information from the data set, as described in section 6.1.2 on page 51.

C.1 Introduction to the Interview

Before starting the interview, the healthcare professional is presented to the project and the purpose of the interview, as described below.

- Presentation to the project, including:
  - The objective of the study
  - The implementation context
  - The role of the NLP system
  - The data set including the extracted concept, sentence, and headline
- The purpose of the interview:
  - To obtain an guideline of which information is most often identified as relevant and related.
  - In practise this means that the healthcare professional should make clear, which information from a given sentence and headline are important to describe the concept.

C.2 Procedure of the Interview

The procedure for the interview is described below. Three discharge summaries were assessed.

- One of the authors (TM) read the concept out loud.
- The healthcare professional reads the associated sentence and headline.
- The healthcare professional says if non, one, or multiple words or phrases are important information to describe the concept.
- The mentioned information are underlined by an author (JK).
- In case a concept is not possible for the authors to understand, the healthcare professional is asked if he/she is familiar with any synonyms or alternative descriptions. If so, the synonyms or alternative descriptions are noted.

C.3 Post-processing of the Interview

After the interview, a guideline was made based on the three discharge summaries assessed by the healthcare professional. The elaboration of the guideline was performed using visual inspection of the underlined words and phrases. Information mentioned multiple times were included in the guideline. This was followed by assigning the pre- and post-coordinated expressions to the concepts in all ten discharge summaries, following the approach described in section 6.1.2 on page 51. Lastly, it was evaluated to what extend, it was possible to bind the concepts and related, relevant information to the concept. If related, relevant information from the guideline was not underlined in the sentence, but this was added by the authors, it was assigned to the category 'Too much bound'.

- Full overlap: The concept and all underlined words or phrases are presented in the SNOMED CT expression.
- Too much bound: The concept and all underlined words or phrases are presented in the SNOMED CT expression, and more are added due to the guideline.
- Too little bound: The concept is presented in the SNOMED CT expression, but too few or no refinements were found in the SNOMED CT expression.
- Not bound: The concept was not possible to understand or not possible to bind.

Overview of the Adjusted cTAKES D

Source code was added to the file 'AbstractCasTermAnnotator.java' as it enabled access to the temporary memory of the adjusted cTAKES. Here the SNOMED CT expressions generated by the default part of cTAKES were located. After the source code from the default part of the adjusted cTAKES was processed two adjustments were added in this file. Initially, all SNOMED CT expressions were assigned a sentence number, based on the index of the SNOMED CT expression. Further each SNOMED CT expression was assigned to a SNOMED CT top-level.

A class 'PostTermAnnotator.java' was created, which contained the functions for generating pre- and post-coordinated expressions as well as functions to determine if a word or phrase was an attribute, or a qualifier value. An overview of the added source code in the adjusted cTAKES can be seen on figure D.1 on the next page.

Three sub-processes were defined on figure D.1 on the following page, which described the process of determining the attributes and qualifier values, as well as making the pre- and post-coordinated expressions, as described in figure 6.5 on page 64, 6.6 on page 65, and 6.4 on page 63, respectively.

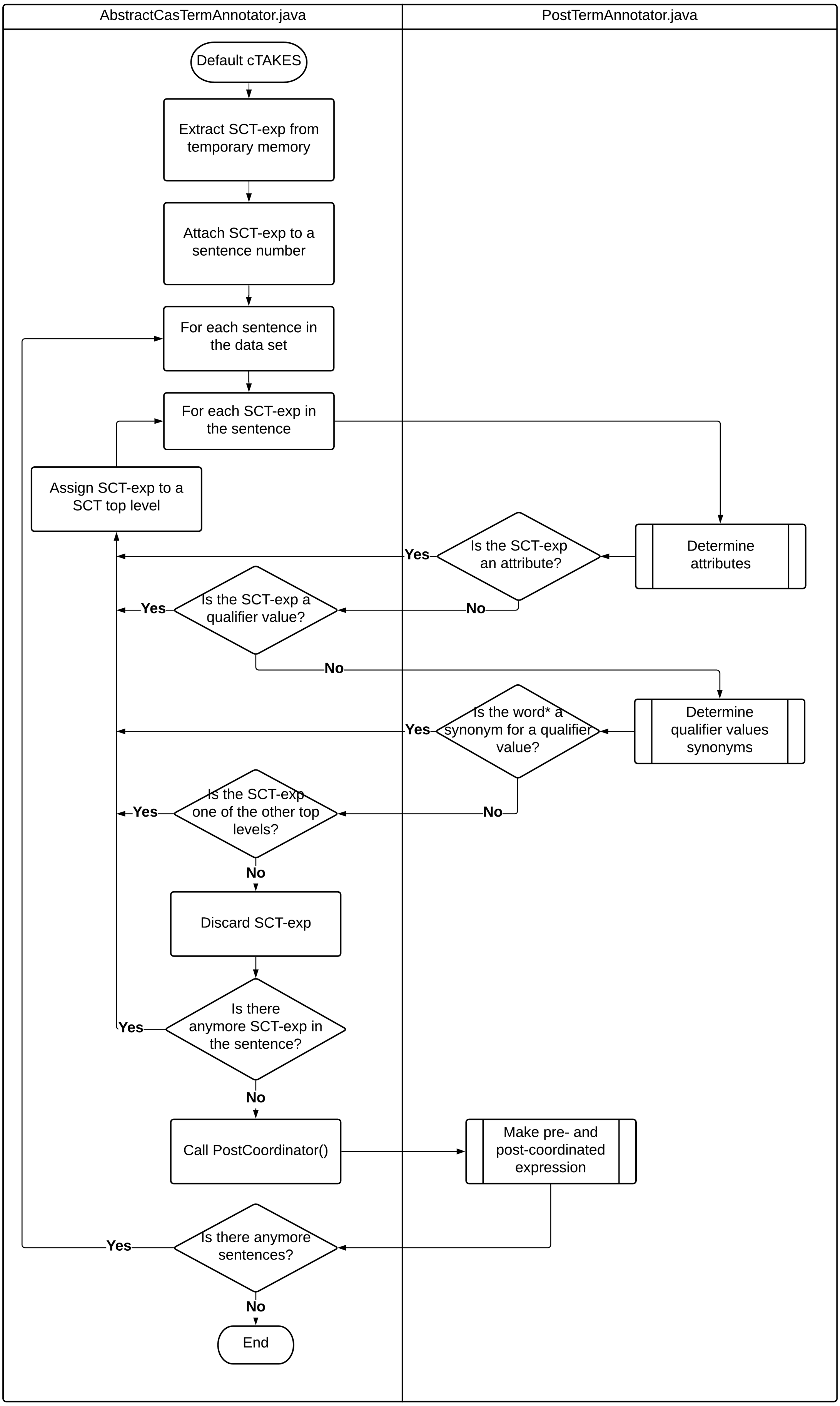

Figure D.1. shows the flowchart of the source code for adjusting cTAKES in the existing class 'AbstractCasTermAnnotator.java' and the created class 'PostTermAnnotator.java'. SCT-exp = SNOMED CT expression, SCT top level = SNOMED CT top level, word* = the word which a SNOMED CT expressions covers.

E Attributes and Qualifier Values

The pre-defined SNOMED CT expressions for attributes were based on which attributes were allowed to be bound to a focus concept from the top level 'Clinical Finding' according to the concept model described in IHTSDO [2021c] chapter 6. Based on the concept model described in IHTSDO [2021d], the attributes as well as the top or sub level which the attribute can be combined with in a refinement were identified and can be seen in table E.1.

Attribute	**Top Level**
255234002 | After |	404684003 |Clinical finding| 71388002 |Procedure|
116676008 | Associated morphology |	49755003 |Morphologically abnormal structure|
47429007 | Associated with |	404684003 |Clinical finding| 272379006 |Event| 410607006 |Organism| 78621006 |Physical force| 260787004 |Physical object| 71388002 |Procedure| 105590001 |Substance|
246075003 | Causative agent |	410607006 |Organism| 373873005 |Pharmaceutical / biologic product| 78621006 |Physical force| 260787004 |Physical object| 105590001 |Substance|
263502005 | Clinical course |	288524001 |Courses|
42752001 | Due to |	404684003 |Clinical finding| 272379006 |Event| 71388002 |Procedure|
246456000 | Episodicity |	288526004 |Episodicities|
419066007 | Finding informer |	420158005 |Performer of method| 444018008 |Person with characteristic related to subject of record| 419358007 |Subject of record or other provider of history|
418775008 | Finding method |	71388002 |Procedure|
363698007 | Finding site |	442083009 |Anatomical or acquired body structure|
363713009 | Has interpretation |	263714004 |Colors| 260245000 |Finding values|
363714003 | Interprets |	386053000 |Evaluation procedure| 108252007 |Laboratory procedure| 363787002 |Observable entity|
246454002 | Occurrence |	282032007 |Periods of life|
246112005 | Severity |	272141005 |Severities|

Table E.1. shows which attributes can be used in a refinement for clinical findings, and which top and sub levels the attributes can be combined with in the refinement.

Qualifier values can only be bound using specific attributes [IHTSDO, 2021c]. In order to combine the right qualifier value and attribute in a refinement, an analysis of the training data was conducted. All post-coordinated expressions were extracted from the training data set. Refinements containing a qualifier value were investigate, and the qualifier as well as the associated attribute were noted. This resulted in the combinations of qualifier values and attributes seen in table E.2.

Attribute	Qualifier Value
255234002 | After |	257556004 | Surgery |
263502005 | Clinical course |	263730007 | Continual | 425323003 | Sudden onset AND short duration | 373933003 | Acute onset | 62459000 | Chronic persistent |
246112005 | Severity |	255604002 | Mild | 6736007 | Moderate | 260360000 | Very high |
363713009 | Has interpretation |	2667000 | Absent | 448371000124103 | Clinically undetermined | 21250004 | Decreased | 54328002 | Indifferent | 64868006 | No growth | 425404009 | Slightly | 255507004 | Small |
246454002 |Occurrence|	303114002 | Early neonatal period |

Table E.2. shows which qualifier values were used in the post-coordinated expressions in the training data set and which attribute was paired with which qualifier value.