



PROYECTO FINAL CARRERA



---

**Módulo codificador de  
diagnósticos para HCE basado en  
estándares de interoperabilidad**

---

**Alumna:**

María Agustina GUILLEN

**Tutora:**

Biong. Eugenia Camila BERRINO

**Co-tutor:**

Ing. José Antonio PICÓN CHAPARRO

# Índice

<b>1. Resumen</b>	<b>1</b>
<b>2. Introducción</b>	<b>2</b>
2.1. Contexto . . . . .	2
2.2. Planteo del problema y necesidades . . . . .	3
2.3. Propósito . . . . .	6
2.4. Alcance . . . . .	7
2.5. Identificación de participantes . . . . .	7
<b>3. Marco Teórico</b>	<b>8</b>
3.1. Sistemas de información de salud . . . . .	9
3.1.1. Información Clínica . . . . .	9
3.1.2. Historia Clínica . . . . .	10
3.1.2.1. Historia Clínica Electrónica . . . . .	11
3.1.3. Diagnóstico . . . . .	11
3.2. Errores de codificación . . . . .	12
3.3. Interoperabilidad . . . . .	13
3.4. Terminologías . . . . .	14
3.5. SNOMED CT . . . . .	15
3.5.1. Beneficios . . . . .	17
3.5.2. Conceptos básicos y componentes . . . . .	18
3.5.2.1. Conceptos . . . . .	19
3.5.2.2. Descripciones . . . . .	19
3.5.2.3. Relaciones . . . . .	20
3.5.2.4. Conjuntos de referencia . . . . .	21
3.5.3. Expresiones SNOMED CT . . . . .	21
3.5.3.1. Expresiones precoordinadas . . . . .	22
3.5.3.2. Expresiones postcoordinadas . . . . .	22
3.5.4. Implementación . . . . .	22

3.5.5.	Modelo conceptual . . . . .	23
3.5.6.	Restricciones de búsqueda . . . . .	24
3.5.7.	Enlace terminológico . . . . .	26
3.6.	Estructura de la información clínica . . . . .	27
3.6.1.	Health Level 7 . . . . .	28
3.6.2.	Estándar FHIR . . . . .	29
3.6.2.1.	Recursos . . . . .	30
3.6.2.2.	Categorías . . . . .	32
3.6.2.3.	Campos . . . . .	35
3.7.	Implementación . . . . .	36
3.7.1.	Lenguaje de programación . . . . .	36
3.7.1.1.	Python Flask . . . . .	37
3.7.2.	HAPI FHIR . . . . .	37
3.7.3.	Versiones . . . . .	38
3.8.	Herramientas de software . . . . .	38
3.8.1.	Servicios web . . . . .	38
3.8.2.	JSON . . . . .	39
3.8.3.	API . . . . .	39
3.8.4.	Microservicios . . . . .	40
3.8.5.	Librería . . . . .	40
<b>4.</b>	<b>Desarrollo</b>	<b>43</b>
4.1.	Definición de campo diagnóstico . . . . .	43
4.2.	Arquitectura de software . . . . .	44
4.3.	Snomed suggestion . . . . .	45
4.3.1.	Métodos y bases utilizados . . . . .	47
4.3.2.	Procesamiento del lenguaje . . . . .	48
4.3.2.1.	Variaciones . . . . .	48
4.3.2.2.	Normalización . . . . .	49
4.3.2.3.	Errores ortográficos . . . . .	49

4.3.3.	ECL definida . . . . .	50
4.3.4.	Tipos de descripciones en resultados . . . . .	51
4.4.	Librería dsutils . . . . .	53
4.4.1.	Integración de sistemas . . . . .	53
4.4.2.	Funcionamiento . . . . .	53
4.4.3.	Elección de recurso y campos . . . . .	55
4.4.4.	Control de errores . . . . .	59
<b>5.</b>	<b>Resultados y métricas</b>	<b>59</b>
5.1.	Muestras . . . . .	59
5.2.	Uso ECL . . . . .	60
5.2.1.	Resultados de búsquedas . . . . .	64
5.2.1.1.	Especificidad . . . . .	67
5.2.1.2.	Resultados literales . . . . .	69
5.2.2.	Relevancia Resultados . . . . .	70
<b>6.</b>	<b>Discusión</b>	<b>72</b>
6.1.	Codificación . . . . .	72
6.1.1.	Comparación de resultados . . . . .	74
6.2.	Procesamiento del lenguaje natural . . . . .	74
6.2.1.	Diccionarios . . . . .	75
6.3.	Postcoordinación . . . . .	76
6.4.	Participación de profesionales . . . . .	77
6.5.	Aspectos legales y normativos . . . . .	79
<b>7.</b>	<b>Mejoras a futuro</b>	<b>80</b>
<b>8.</b>	<b>Conclusiones</b>	<b>81</b>
<b>A.</b>	<b>Anexo</b>	<b>83</b>
	<b>Bibliografía</b>	<b>85</b>

## Índice de figuras

1.	Clasificación de diagnósticos usando herramientas digitales actuales. . .	4
2.	Niveles de interoperabilidad. Imagen modificada, extraída de [9] . . .	14
3.	Beneficios de SNOMED CT. Imagen obtenida de [14] . . . . .	17
4.	Ejemplo de descripciones para un solo concepto. Imagen extraída de [14] modificada. . . . .	20
5.	Ejemplo de relaciones para un solo concepto. Imagen extraída de [14]	21
6.	Lista de conceptos de nivel superior. Imagen obtenida de [15] . . . . .	24
7.	Componentes de una restricción de expresiones. Imagen modificada, obtenida de [14]. . . . .	26
8.	Gramática composicional y lenguaje de restricciones. . . . .	27
9.	Recursos disponibles en la pagina de FHIR en un orden alfabético en inglés. Imagen obtenida de [20] . . . . .	31
10.	Categorías de la sección clínica. Imagen obtenida de [20]. . . . .	33
11.	Descripciones de tipos de datos en FHIR. Imagen modificada extraída de [20] . . . . .	36
12.	Archivos de la estructura de una librería. Imagen extraída de [33] . .	41
13.	Arquitectura de software del sistema implementado. . . . .	44
14.	Front representativo. . . . .	46
15.	Descripciones para el concepto 254837009 (F:F5N, S:Sinonimo) . . .	52
16.	README de la librería dsutils. . . . .	55
17.	Recurso ClinicalImpressione en JSON. Imagen extraída de [20] . . .	56
18.	Estructura del recurso ClinicalImpression en JSON. . . . .	58
19.	Conceptos codificados. Con ECL el porcentaje de codificación fue del 81 % mientras que sin ECL del 81,96 % . . . . .	61
20.	Resultado de conceptos devueltos por SNOMED para el termino “trauma” sin aplicar restricciones. . . . .	62
21.	Gráfico de dispersión de la cantidad de resultados de búsqueda con y sin restricciones de lenguaje. . . . .	63

22.	Resultados de búsquedas. . . . .	67
23.	Tipos de especificaciones. . . . .	68
24.	Histograma de posición de resultados “literales”. . . . .	70
25.	Términos codificables. Cantidad de términos relevantes. . . . .	72

## Índice de tablas

1.	Análisis de campos del recurso ClinicalImpression. Información obtenida de [20] . . . . .	57
2.	Campos seleccionados del recurso ClinicalImpression. . . . .	58
3.	Ejemplos de términos denominados como “resultado literal”. . . . .	64
4.	Ejemplos de textos cuyo termino aceptable es denominados como “resultado literal” y muestra al usuario el termino preferencial. . . . .	65
5.	Ejemplos de términos denominados como “Especificidad”. . . . .	65
6.	Ejemplos de abreviaturas de diagnósticos no reconocidas correctamente. . . . .	66
7.	Ejemplos de términos denominados como “Otros”. . . . .	66
8.	Ejemplos de especificaciones necesarias. . . . .	69

## Glosario

**API** Application Programming Interfaces. 39, 44

**ECL** Expression Constraint Language. 24, 25, 45

**FHIR** Fast Healthcare Interoperability Resources. 5, 29, 30, 37, 56

**FSN** Fully Specified Name. 19, 20, 52

**HC** Historia Clinica. 17, 18

**HCE** Historia Clinica Electronica. 11, 17, 26

**HTTP** HyperText Transfer Protocol . 39, 47

**IHSA** International Health Services Argentina . 2

**ISO** Organización Internacional de Normalización. 28

**JSON** JavaScript Object Notation. 39, 56, 58

**Refsets** conjuntos de referencia. 19, 21

**REST** Representational State Transfer. 39

**SIS** Sistemas de Informacion de Salud. 8

**SNOMED CT** Systematized Nomenclature of Medicine – Clinical Terms. 5, 7, 15,  
17–19, 21, 23, 24, 26, 27, 58, 60

## 1. Resumen

Este proyecto tiene como objetivo realizar un módulo codificador de diagnósticos. Para esto se investigan y desarrollan sistemas que respetan e implementan estándares de interoperabilidad SNOMED CT y FHIR, permitiendo su uso en la práctica médica.

El sistema se organiza en una arquitectura que permite la interacción modulable y escalable de los componentes del software. Se crea una API como punto de entrada para las solicitudes y respuestas, junto con la librería que estructura los datos necesarios. Además, se establece comunicación con una API de HAPI FHIR para crear y almacenar recursos en una base de datos.

Se desarrolla un sistema con facilidad de uso y practicidad que hace posible una selección precisa dentro de los términos de diagnóstico. Esta codificación permite realizar análisis posteriores como sistemas de soporte a la toma de decisiones, análisis con diferentes niveles de agregación de datos, previsión de insumos, entre otros.

Los diagnósticos tienen un impacto decisivo en la continuidad de la atención de un paciente. Una buena clasificación del diagnóstico lleva al tratamiento oportuno que puede mejorar la salud a una persona. Además, permite procesar y presentar estos datos de diferentes maneras para distintos fines. Por ejemplo, admite mayores oportunidades para el apoyo a la toma de decisiones e informes retrospectivos precisos para la investigación y la gestión.

El sistema fue diseñado para implementarse en el área médica de Emergencias (I.H.S.A). Se realizaron pruebas de validación con diagnósticos no reconocidos por los sistemas actuales de la empresa, logrando codificar correctamente el 81 % de 2805.

## 2. Introducción

### 2.1. Contexto

El proyecto se desarrolla en el marco de la empresa Emergencias (IHSA), una organización dedicada a proporcionar servicios de emergencias médicas extra-hospitalarias y orientada a brindar soluciones en la gestión de servicios de salud. [1]

A lo largo de los años, Emergencias desarrolló una variedad de sistemas informáticos para satisfacer necesidades específicas, lo que dio lugar a una diversidad de sistemas informáticos que llevan a una dificultad en la integridad y la interoperabilidad de los datos.

En la actualidad, no cuenta con una base centralizada donde se pueda consolidar toda la información relevante sobre la salud de los pacientes. Esto sumado a la falta de interoperabilidad entre los sistemas, presenta desafíos en términos de accesibilidad, precisión y eficiencia en la toma de decisiones clínicas.

Con el objetivo de superar estas limitaciones, Emergencias ha iniciado una iniciativa de transformación digital mediante la creación de un área especializada en este campo. Tiene como objetivo impulsar la implementación de soluciones tecnológicas innovadoras para optimizar los servicios de atención médica y mejorar la experiencia de los pacientes.

Dentro de esta estrategia de transformación digital, desarrolló una aplicación móvil que se utiliza en la atención domiciliaria, con un enfoque principal en el campo de la kinesiología. Esta aplicación permite a los profesionales de la salud registrar información relevante sobre los pacientes, incluyendo diagnósticos y otros datos clínicos. Esta propuesta contribuye en mejorar la accesibilidad y la eficiencia de la atención médica, brindando una herramienta digital para la gestión de la salud y el bienestar de las personas.

El proyecto se suma a esta iniciativa de transformación digital y está enfocado en la mejora de la interoperabilidad de la información médica en Emergencias, parti-

cularmente en el área de diagnósticos. El objetivo es asegurar que los diagnósticos registrados en la aplicación móvil y otros sistemas de la empresa se integren de manera eficiente en una Historia Clínica Electrónica interoperable.

Este proyecto va a afectar al siguiente flujo de trabajo existente en la Empresa:

1. Un paciente recibe atención médica de un médico externo a emergencias o está internado. Durante esta atención, se realiza un diagnóstico.
2. A partir del diagnóstico, en casos particulares el mismo es derivado a emergencias, con el objetivo de que al paciente se le brinde una atención/internación domiciliaria, y además se le asignan recursos. En este proceso, personal administrativo registra el diagnóstico brindado por el profesional de salud externo.
3. El profesional de salud de Emergencias atiende al paciente en su domicilio y confirma, corrige o agrega nuevos diagnósticos según sea necesario.

## 2.2. Planteo del problema y necesidades

Actualmente, Emergencias utiliza un sistema “Legacy” con una Clasificación Internacional de Enfermedades (CIE-10) reconocida mundialmente. Este sistema no proporciona una estructura adecuada para la clasificación y codificación de los diagnósticos para Emergencias, lo que dificulta la identificación precisa de los términos correspondientes.

Como resultado, los profesionales de salud y el personal administrativo que registra diagnósticos provenientes de médicos externos, a menudo clasifican los diagnósticos como “Otros diagnósticos” o utilizan texto libre para describir el diagnóstico real. Esto se ve en el siguiente gráfico:



Figura 1: Clasificación de diagnósticos usando herramientas digitales actuales.

La clasificación de los diagnósticos como “otros diagnósticos” en lugar de buscar el diagnóstico correspondiente en la CIE-10 puede deberse a varias razones. Una de ellas es la falta de familiaridad o conocimiento sobre los códigos y términos específicos de CIE-10. Esto puede dificultar la identificación precisa del diagnóstico correcto en la terminología estandarizada. Además, la falta de una estructura clara y fácil de usar en el sistema Legacy o la percepción de que es más rápido y fácil ingresar el diagnóstico utilizando texto libre puede influir en la elección de categorías genéricas. Esto se suma a la alta carga de trabajo de los administrativos.

Para abordar la problemática existente en Emergencias, es importante comprender los desafíos asociados al almacenamiento de información médica en lenguaje natural. Aunque el lenguaje natural es expresivo y fácil de utilizar para ingresar información, presenta limitaciones significativas en cuanto a la recuperación y análisis de datos. Los textos en lenguaje natural son poco útiles en comparación con el texto codificado, que ofrece una mayor capacidad para recuperar y analizar la información de manera eficiente. [2] [3]

En el contexto de Emergencias, la información de las notas clínicas y otros reportes médicos electrónicos se encuentra inaccesible para la mayoría de los métodos de análisis y recuperación de datos basados en computadoras. Esto limita la utilidad

clínica de los textos, ya que una gran cantidad de información valiosa se pierde debido a la dificultad de procesar la parte narrativa.

Para resolver estas limitaciones se eligió SNOMED CT por pedido de la empresa como la terminología principal para la codificación de diagnósticos en este proyecto. En Argentina, SNOMED CT se adoptó ampliamente como un estándar de codificación y terminología médica, lo que facilita la interoperabilidad y el intercambio de información clínica entre diferentes instituciones y profesionales de la salud. Al aprovechar estas ventajas, se busca garantizar una representación más precisa y completa de los diagnósticos, mejorar la comparación y el análisis de datos, y fomentar la colaboración y la investigación en el campo de la salud.

En la implementación del proyecto, se utiliza el Servidor SNOMED (SNOWSTORM) del Ministerio de Salud de la Nación para realizar las consultas de codificación. Esta elección se basa en la existencia de un servidor SNOMED operativo y funcional proporcionado por el Ministerio de Salud, lo que evita la necesidad de implementar un servidor propio. Al utilizar el servidor SNOMED existente, se garantiza una solución estable, evitando problemas y simplificando el proceso de consulta y codificación de diagnósticos.

A su vez se eligió FHIR que ya estaba siendo implementado por la empresa Emergencias como parte de su estrategia de transformación digital. La adopción de FHIR permite una mayor coherencia y compatibilidad con los sistemas existentes, facilitando la integración de la codificación de diagnósticos. También asegura una estructura de datos consistente y definida para el intercambio de información clínica, mejorando la interoperabilidad y el intercambio de datos en el ámbito de la salud.

En este contexto, el proyecto plantea los siguientes objetivos:

- Generar un codificador de diagnósticos interoperable basado en SNOMED CT y FHIR.
  - Comprender la importancia de SNOMED CT y FHIR como estándares de interoperabilidad en salud.

- Comprender la estructura de SNOMED CT y la API de SNOWSTORM con sus métodos correspondientes.
- Desarrollar una API que sea capaz de guardar el diagnóstico en FHIR, ya sea que se codifique el diagnóstico o no.
  - Establecer el recurso FHIR correcto para guardar diagnósticos.
  - Posibilitar el guardado de diagnóstico seleccionado en FHIR.
- Validar con ejemplos de la empresa.

Estos objetivos tienen como finalidad mejorar la precisión, la eficiencia y la calidad en el registro y análisis de la información médica, así como facilitar la interoperabilidad y el intercambio de datos en el contexto de Emergencias. Al cumplir con estos objetivos, se espera optimizar los procesos médicos, mejorar la atención al paciente y fortalecer los sistemas de información en el ámbito de la salud.

### 2.3. Propósito

El objetivo de este proyecto es el desarrollo de un modulo codificador de diagnósticos. Crear un sistema informático que reciba como entrada un diagnostico en texto libre y que devuelva como salida su codificación en un sistema de referencia. Todo esto respetando estándares terminológicos y sintácticos.

Se podrá realizar acciones posteriores como:

- La asignación de recursos sanitarios personalizada para cada paciente según su diagnostico.
- Visualización de la historia clínica integral de cada paciente.
- Análisis de datos y potencial ciencia de datos.
- Aceleración proyectos digitales de la corporación en términos de construir una HCE interoperable.

A su vez, se desarrolla un guardado cumpliendo los estándares de interoperabilidad. El guardado tendrá como entradas el identificador numérico y la descripción del concepto de SNOMED CT devuelto por la terminología o simplemente el texto libre en el caso de no poder ser codificado, junto a la identificación del paciente o simplemente el texto libre en el caso de no poder ser codificado, estado de la evaluación, fecha de la documentación y evaluación e identificación del profesional de salud que participa. La salida devuelve estos datos estructurados, cumpliendo formatos específicos según estándares. Esto ocurre independientemente de la codificación del texto.

## 2.4. Alcance

Definiendo los requisitos a nivel del proyecto, a cada texto libre ingresado como diagnóstico, se asigna un único código de SNOMED CT; esto representa un concepto médico, el cual debe ser ajustado para que corresponda al campo diagnóstico.

## 2.5. Identificación de participantes

Considerando participante a todo aquél que se beneficie del sistema propuesto, se tiene lo siguiente:

- Pacientes: Quien se beneficia gracias a un diagnóstico médico cuya identificación es precisa en términos de la enfermedad o afección. La adecuada determinación del tratamiento óptimo no solo mejora la calidad de vida del paciente y previene complicaciones graves, sino que también ayuda a optimizar los recursos y evitar gastos innecesarios.
- Personal de salud: Se favorece con un ingreso de diagnóstico efectivo, tanto en términos de eficiencia del tiempo como en la efectividad del registro del diagnóstico. Se reduce el riesgo de errores médicos al transmitir o recibir un diagnóstico con otros profesionales de la salud.
- Población: Se mejora la salud pública y previenen epidemias al facilitar la identificación y control de enfermedades. Un buen diagnóstico reduce la mortalidad

al detectar enfermedades en etapas tempranas, aumentando las posibilidades de éxito en el tratamiento. La evaluación de resultados y el uso de datos permiten mejorar tratamientos y terapias para enfermedades y afecciones.

- Establecimiento de salud: Se beneficia mejorando las tecnologías médicas existentes al proporcionar información precisa sobre la enfermedad o afección en cuestión. Contribuye en realizar análisis que agrupan afecciones similares, o basadas en los mismos patógenos. Se permite la toma de decisiones en base a los datos obtenidos. Asimismo, estimula la investigación y el desarrollo en el área de tecnología médica que puede llevar a nuevas tecnologías y avances en el campo.

Se tiene en cuenta que a implementación de tecnologías médicas en una institución de salud no solo beneficia a los pacientes, sino también tiene un impacto positivo en el aspecto económico, la planificación y la asignación de recursos. Al proporcionar información precisa, estimular la investigación y el desarrollo, optimizar los tiempos y mejorar la calidad de los registros médicos, estas tecnologías permiten una toma de decisiones más fundamentada, una atención más eficiente y una mejora general en la calidad de los servicios de salud.

### **3. Marco Teórico**

El marco de este proyecto desarrollará la descripción los sistemas de información clínica y de salud (SIS) y su registro en forma de historia clínica, así como también, las tendencias actuales en relación a los sistemas de información, la construcción, generación y representación del conocimiento médico, con especial énfasis en el proceso de diagnóstico. Además, se trazarán terminologías que se aplican para analizar los distintos sistemas de clasificación y nomenclaturas disponibles en la actualidad.

### 3.1. Sistemas de información de salud

Los sistemas de información para el ámbito de la salud deben resolver una problemática muy compleja y particular. Se trata de una función que requiere la integración de habilidades de distintos ámbitos del conocimiento.

Para planificar y llevar a cabo las fases de un proyecto de software, se necesitan conocimientos de informática y habilidades prácticas. También ser capaz de comprender y tomar decisiones sobre los requisitos para la gestión de hardware y redes. En el área de la salud, se utilizan los datos con objetivos bien definidos, (médico asistencial; administrativo; investigación y gestión), que tienen necesidades distintas y por lo tanto requieren de un manejo de la información muchas veces incompatibles entre si cuando no son previstos. [4]

El flujo de información médica tiene una dinámica distinta que difiere según la situación y no se adhiere a los requisitos administrativos y con necesidades muy particulares.

La información demográfica y epidemiológica sobre las condiciones de salud y enfermedad de la población, así como indicadores de la estructura, proceso y resultados del sistema, son necesarios para la para la administración y dirección en el ámbito médico. Por lo tanto, para implementar una gestión efectiva que mejore los resultados de los pacientes y reduzca los costos, se requieren sistemas de información cuidadosamente considerados y bien resueltos.

Se requieren conocimientos sobre sistemas de clasificación y agrupamiento de problemas médicos; prácticas y fármacos; habilidades y manejo de estadística y epidemiología para el correcto ajuste de los riesgos y análisis científico de la información.

#### 3.1.1. Información Clínica

“Tradicionalmente se asoció el concepto de información clínica al de información exclusivamente médica y ambos al de historia clínica, que clásicamente se definió como la narración completa o parcial de la experiencia del médico en su relación

técnica con un paciente” [5]. Esta información se relacionaba con los estados patológicos del paciente, entendido éste como aquel individuo particular que precisa atención médica por ver alterado su estado de salud.

En la actualidad, se considera información clínica a todo dato, cualquiera que sea su forma, clase o tipo; que permite adquirir o ampliar conocimientos sobre el estado de salud de una persona, la forma de preservarla, mejorarla o recuperarla.

Información clínica es aquella que se genera como consecuencia de la asistencia médica recibida tanto en atención primaria, como especializada o socio-médica. Los diferentes niveles de atención que brindan servicio a un ciudadano generan información sobre su estado de salud. El concepto por lo tanto, engloba tanto la información clínica como la de estado de salud.[5]

### **3.1.2. Historia Clínica**

Agrupar en el tiempo los datos obtenidos por un profesional para asegurar la continuidad de los tratamientos de un paciente, es posible gracias a la historia clínica del mismo. Un componente crítico de la atención al paciente. Los detalles contenidos en la historia clínica de una persona son importantes para mantener la continuidad y verificar la evolución de la atención médica, determinar si un problema que motivó la búsqueda de atención puede solucionarse o no, e identificar nuevos problemas de salud a través de acciones de diagnóstico y tratamientos relacionados.[6]

A pesar de las numerosas iniciativas de estandarización en los últimos años, a veces es difícil obtener una imagen clara de la evolución del individuo a partir de la historia clínica, que también incluye síntomas, resultados de pruebas, consideraciones, planes terapéuticos y hallazgos clínicos. Principalmente de aquellos que sufren de múltiples enfermedades o problemas de salud.

Al considerar el contenido de la historia clínica, vale destacar que todo y cualquier atención en salud presupone el compromiso y la participación de muchos profesionales, por eso se recopilan múltiples datos de diferentes fuentes. Por otro lado,

los procedimientos realizados por los profesionales individualmente también generan otras tantas informaciones, que van a garantizar la continuidad del proceso de cuidado. Estas tantas y diferentes fuentes de datos generan consecuentemente una gran variedad de información, muchas veces inconexas.

Todos los datos derivados de la demanda de atención deben combinarse y ordenarse para crear un contexto que apoye y dirija la toma de decisiones sobre el tipo de tratamiento que debe recibir el paciente.

#### **3.1.2.1. Historia Clínica Electrónica**

La HCE es un repositorio para todos los datos de salud clínicos y administrativos recopilados a lo largo de la vida de una persona. Se pueden obtener numerosas ventajas de este formato de almacenamiento, incluido el acceso rápido a problemas e intervenciones de salud actuales; acceso al conocimiento científico actualizado, lo que conduce a una mejora en el proceso de toma de decisiones; mejora de la eficacia asistencial, que contribuiría a obtener mejores resultados de los tratamientos realizados y de la atención al paciente; posible reducción de costos; y optimización de recursos. [7]

#### **3.1.3. Diagnóstico**

La palabra diagnóstico deriva del griego DIA que significa a través o parte de, y GNOSIS que es conocer, es decir que es el arte de reconocer. Desde la antigüedad, el diagnóstico ha sido una tarea fundamental y el punto de base para el buen ejercicio de la actividad médica. No es posible un tratamiento adecuado sin un correcto diagnóstico.

La complejidad del proceso diagnóstico deriva de la extraordinaria variedad de modos de presentación de las entidades productoras de síntomas y esta complejidad es amplificada por el impacto del paciente sobre su expresión. El diagnóstico médico, se basa en observaciones clínicas y puede beneficiarse de técnicas auxiliares como el laboratorio clínico, imagenología, etc.

Desde la perspectiva de la lógica cuantitativa, el diagnóstico es un procedimiento para reducir la incertidumbre. Por lo tanto, el diagnóstico clínico es un proceso de clasificación o un intento de discernir la clase a la que pertenece un paciente en particular. Hay varias formas de identificar los procesos que dan como resultado un diagnóstico médico, incluido el reconocimiento de patrones, las estrategias de arborización, las estrategias de agotamiento y las estrategias hipotético-deductivas.

Por lo tanto, un paso fundamental en la práctica de las actividades de salud es el diagnóstico, ya que podría describirse como esencial y complejo. Básicamente se reduce a un procedimiento de clasificación.

Además, para diagnosticar a un paciente, se utiliza un conjunto jerárquico de conceptos, mediante el cual se designa una clase que representa una enfermedad.

### **3.2. Errores de codificación**

Es fundamental implementar estrategias de manejo de errores para en el proceso de registro y documentación de la información. Existe un manejo o control de errores que se refiere a los procedimientos de respuesta y recuperación de las condiciones de error presentes en una aplicación. Es el proceso compuesto por la anticipación, detección y resolución de errores de aplicación, errores de programación o errores de comunicación. El manejo de errores ayuda a mantener el flujo normal de ejecución del programa. [8]

Estos errores en el ingreso de información clínica pueden suponer una importante amenaza para la calidad y la seguridad en atención de salud, por lo que es necesario reflexionar sobre ellos. Para minimizar la cantidad de errores cometidos al momento de registrarlos, es necesario plantear un diseño propenso disminuirlos.

Con estrategias efectivas de manejo de errores de codificación se puede brindar una atención médica más precisa y segura, reduciendo el impacto negativo en los pacientes.

### 3.3. Interoperabilidad

La interoperabilidad se refiere a la capacidad de los sistemas de información de diferentes organizaciones o entidades para intercambiar datos y utilizarlos de manera coherente y efectiva. La interoperabilidad en salud implica la capacidad de los sistemas y aplicaciones de información de salud para comunicarse, compartir datos y utilizarlos de manera significativa en un entorno clínico. [9]

El principal beneficio es la disponibilidad de datos que se encuentran dispersos en numerosas instituciones o regiones para ayudar en las decisiones relacionadas con la salud del paciente y la eficiencia institucional.

Existen cuatro tipos de interoperabilidad: Técnica, Sintáctica, Semántica y Organizacional:

- Interoperabilidad Técnica: “Se centra en protocolos de comunicación y en la infraestructura necesaria para el intercambio de datos”. [9]
- Interoperabilidad Sintáctica: “Relacionada habitualmente con los formatos de los datos. Interviene en la estructura la información al momento de intercambiar”. [9]
- Interoperabilidad Semántica: “Se asocia con el vocabulario del contenido de la información intercambiada. Existe un entendimiento común entre los sistemas de códigos y el significado del contenido que se intercambia”. [9]
- Interoperabilidad Organizacional: “es la capacidad de las organizaciones para comunicar y transferir efectivamente los datos, a pesar de que este usando una variedad de sistemas de información sobre infraestructuras muy diferentes, a través de regiones geográficas y culturas distintas”. [9]

La interoperabilidad depende de que sean exitosos los niveles anteriores.

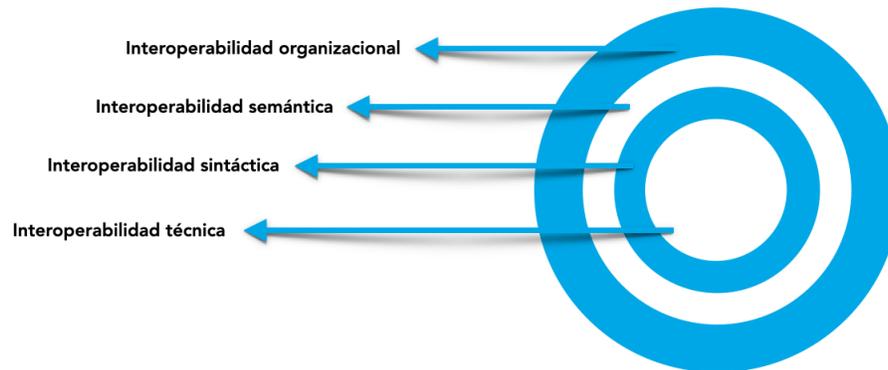


Figura 2: Niveles de interoperabilidad. Imagen modificada, extraída de [9]

La interoperabilidad semántica es un requisito previo en el contexto sanitario, para que los profesionales clínicos puedan compartir y acceder a toda la información de un paciente según sea necesario e incluso utilizarla con fines estadísticos y de investigación. [9]

Para construir sistemas que sean semánticamente interoperables necesitamos:

- Estándares terminológicos: proveen códigos específicos para conceptos clínicos. Ejemplos de estándares de terminologías médicas son SNOMED CT, LOINC, o CIE 10.

### 3.4. Terminologías

La Terminología de Interfaz es un vocabulario vital para los sistemas de información en salud. “La complejidad y diversidad encontrada en el acto asistencial representa una dificultad al momento de solicitar el ingreso de datos en los sistemas asistenciales. Por tal motivo, la Terminología de Interfaz tiene el rol fundamental de acercar al usuario un lenguaje que permita la rápida y eficiente extracción de información”. [10] Este objetivo se logra mediante una selección de términos que representan de manera adecuada y precisa lo que el profesional en salud desea expresar. La gran heterogeneidad de actores, situaciones, enfermedades e intervenciones

requiere una solución global y escalable, solo alcanzable a través de los servidores de terminología.

Los servidores terminológicos potencian la gestión de la información en los sistemas de salud. Permiten representar los datos clínicos de los pacientes a partir del ingreso de texto en lenguaje natural. Estos datos son asociados a códigos terminológicos que representan con gran detalle conceptos abstractos de la realidad asistencial. Los códigos pueden luego ser utilizados tanto para la interpretación por máquinas como para lograr un intercambio real de la información. La gestión del conocimiento por medio de terminologías clínicas permite controlar los datos contenidos en las bases de datos.

La vinculación de estas bases y los conceptos ingresados por los profesionales de la salud, mediante un vocabulario común, habilita la gran potencialidad de las herramientas de la tecnología de la información para mejorar la calidad de atención (soporte para la toma de decisiones). La utilización de una Terminología de Interfaz estandariza a nivel local la representación de conceptos en salud. La denominación común de aspectos relacionados a las ciencias biomédicas facilita la reutilización de datos con fines de investigación.

El ingreso de datos estandarizados en el SIS potencia la posibilidad de utilizar dicha información para la investigación, sin tener que recaer en la revisión manual o generación de reportes duplicados para cada estudio científico. El almacenamiento codificado de información biomédica en una única base de datos optimiza la recuperación de datos para su utilización según la necesidad.

### **3.5. SNOMED CT**

“SNOMED CT es un servicio de terminología clínica mantenida y actualizada por SNOMED Internacional, que conforma un estándar global para la representación de información en salud”. [11]

Se utiliza en sistemas de historia clínica electrónica, de prescripción electrónica,

laboratorios, encuestas sanitarias, registros de enfermedades y otros, con el objetivo de capturar de manera precisa y unívoca referencias a conceptos clínicos. Este nivel de detalle permite lograr una atención médica más segura y un sistema de salud que cuenta con información confiable y actualizada para la vigilancia y la toma de decisiones.

SNOMED CT posee más de 330.000 conceptos clínicos, divididos en jerarquías temáticas con síntomas, diagnósticos, antecedentes familiares, procedimientos quirúrgicos, estudios diagnósticos, sustancias, productos farmacéuticos, vacunas, dietas, dispositivos, descartables, microorganismos, anatomía, etc. Sus contenidos se actualizan semestralmente y se encuentra traducido al castellano en su totalidad.

Se destaca que SNOMED CT es una combinación de un sistema de clasificación y una ontología, lo que significa que no solo proporciona una lista de términos médicos, sino que también define las relaciones entre los conceptos y permite una representación estructurada de la información clínica.

Esto implica que las relaciones entre conceptos y atributos contribuyen al significado de cada concepto, los términos en la jerarquía de SNOMED CT pueden tener más de un “padre”. Esto también ocurre al realizar el mapeo de SNOMED con otros estándares internacionales, ya que los cambios en los conceptos y términos de SNOMED pueden requerir modificaciones en el mapeo. Asimismo, los cambios en el otro estándar pueden generar la necesidad de actualizar los mapas en SNOMED”. [12]

Las guías de aplicación de SNOMED CT ofrecen recomendaciones para utilizar y aplicar esta terminología en diferentes contextos[13]. Estas guías proporcionan orientación sobre cómo utilizar los conceptos y relaciones de SNOMED CT de manera efectiva, mapear terminologías locales y mantener la integridad de los datos clínicos. Las guías son desarrolladas por organizaciones especializadas y se actualizan regularmente para reflejar los avances y mejores prácticas en terminología clínica.

Argentina se ha incorporado a SNOMED Internacional en enero de 2018, como

el país número 31, desde entonces la terminología es de acceso gratuito en todo el territorio nacional. El Centro Nacional de Terminología en Salud, es el centro nacional de referencia para SNOMED Internacional y publica semestralmente las actualizaciones de la Edición Argentina de SNOMED CT, que incluyen contenido específico para el representar conceptos clínicos locales.[11]

### 3.5.1. Beneficios

Permite el desarrollo de contenido médico completo, de alta calidad en HC. Proporciona una forma estandarizada de representar frases registradas por el médico y permite su interpretación automática. Es un vocabulario controlado, validado clínicamente, con riqueza semántica, que facilita la evolución de la expresividad para cubrir los requerimientos a medida que surgen. [14]

Como vemos en la Figura 3, el uso efectivo de SNOMED CT beneficia a las poblaciones, a pacientes individuales y médicos; y apoya la atención médica basada en la evidencia.



Figura 3: Beneficios de SNOMED CT. Imagen obtenida de [14]

El uso de una HCE mejora la comunicación y aumenta la disponibilidad de información relevante. Si la información clínica se almacena de manera que permita la

recuperación basada en el significado, los beneficios aumentan considerablemente. Los beneficios adicionales van desde mayores oportunidades para el apoyo a la toma de decisiones en tiempo real hasta informes retrospectivos más precisos para la investigación y la gestión.

Por otro lado, cuenta con una gran amplitud en cuanto a especialidades. Permite procesar y presentar los mismos datos de diferentes maneras para distintos fines. Las HC pueden procesarse y presentarse de distintas formas para facilitar la atención directa de los pacientes, auditorías clínicas, investigación, epidemiología, gestión y planificación de servicios. Además, el alcance global de SNOMED CT reduce los efectos de límites geográficos.

Las estructuras de SNOMED CT permiten ingresar información utilizando sinónimos que se adaptan a las preferencias locales a la vez que se registra a información de manera coherente y comparable. Esta naturaleza jerárquica permite registrar la información con diferente nivel de detalle mediante la combinación de conceptos, cuando los conceptos disponibles no son suficientemente precisos. También admite la recuperación y reutilización selectiva de la información para cumplir con diferentes requerimientos en diversos niveles de generalización.

### **3.5.2. Conceptos básicos y componentes**

El conocimiento de las características, componentes y productos de SNOMED CT proporciona una base inicial sobre la cual construir una mayor comprensión. También informa las discusiones que conducen a decisiones sobre la adopción, implementación y uso de esta terminología.

Las características mas destacadas de SNOMED CT son las siguientes:

- Es la terminología clínica sanitaria multilingüe más completa del mundo.
- Es un recurso con contenido clínico completo y científicamente validado.
- Permite una representación coherente y procesable del contenido clínico en los registros electrónicos de salud.

- Está mapeado a otros estándares internacionales.
- Ya se utiliza en más de cincuenta países.

Cuando se implementa en aplicaciones de software, SNOMED CT se puede utilizar para representar información clínicamente relevante de manera consistente, confiable y completa como parte integral de la producción de información electrónica de salud. Esta implementación requiere una comprensión de la forma en que el contenido de SNOMED CT está representado.

El contenido de SNOMED CT se representa mediante tres tipos de componentes descriptos a continuación que se complementan con conjuntos de referencia (Refsets), que proporcionan características flexibles adicionales y permiten la configuración de la terminología para abordar diferentes requisitos.[14]

#### **3.5.2.1. Conceptos**

SNOMED CT contiene conceptos con significados únicos y definiciones formales basadas en lógica organizadas en jerarquías. Los conceptos de SNOMED CT representan pensamientos clínicos y tienen un identificador numérico único.

Dentro de cada jerarquía, los conceptos se organizan de lo general a lo más detallado. Esto permite registrar datos clínicos detallados y luego acceder o agregar a un nivel más general.

#### **3.5.2.2. Descripciones**

Las descripciones de SNOMED CT vinculan términos apropiados legibles por humanos con conceptos. Un concepto puede tener varias descripciones asociadas, cada una representando un sinónimo que describe el mismo concepto clínico. Se utilizan dos tipos de descripción para representar cada concepto: nombre completamente especificado (FSN) y sinónimo. Cada descripción tiene un identificador de descripción numérico único.

FSN representa una descripción única, no ambigua, del significado de un concepto.

La descripción completa no está destinada a ser visualizada en las historias clínicas. Su fin es desambiguar el significado de cada concepto. Esto es particularmente útil cuando se designan diferentes conceptos con la misma palabra o frase de uso frecuente. Cada concepto tiene una única (FSN) en cada idioma o dialecto.

Los conceptos tienen un sinónimo que está marcado en un idioma, dialecto o contexto de uso determinado. Esto se conoce como “preferido” y es una palabra o frase comúnmente utilizada por los médicos para nombrar ese concepto. Cualquier otro sinónimo se puede marcar como “aceptable”.

En la Figura 4 se puede observar un ejemplo de descripciones para un solo concepto.

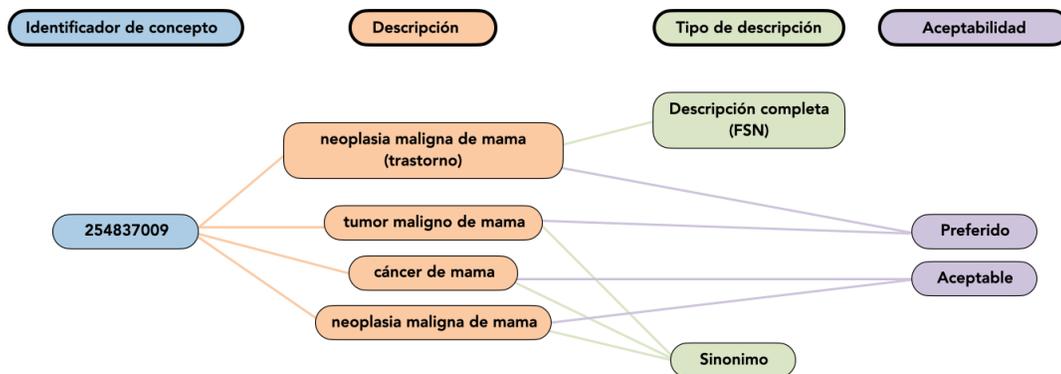


Figura 4: Ejemplo de descripciones para un solo concepto. Imagen extraída de [14] modificada.

### 3.5.2.3. Relaciones

Una relación representa una asociación entre dos conceptos. Estas relaciones proporcionan definiciones formales y otras propiedades del concepto. Un tipo de relación “*es una*” que relaciona un concepto con conceptos más generales. Estas son relaciones que definen la jerarquía de los conceptos de SNOMED CT. Cada relación tiene un identificador de relación numérico único y un tipo de relación (o atributo) como vemos en la Figura 5.

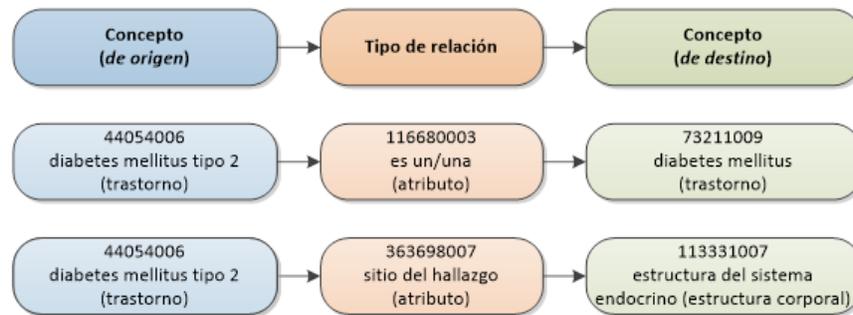


Figura 5: Ejemplo de relaciones para un solo concepto. Imagen extraída de [14]

### 3.5.2.4. Conjuntos de referencia

Los Refsets son un enfoque estándar flexible utilizado para admitir una variedad de requisitos para la personalización y mejora de SNOMED CT. Estos incluyen la representación de subconjuntos, preferencias de idioma para el uso de términos particulares y mapeo desde o hacia otros sistemas de código. Cada conjunto de referencia tiene un identificador de concepto numérico único.[14]

### 3.5.3. Expresiones SNOMED CT

SNOMED CT proporciona un mecanismo que permite representar frases clínicas, incluso cuando un solo concepto de SNOMED CT no captura el nivel de detalle requerido. Esto es importante ya que permite capturar una amplia gama de significados clínicos en un registro, sin requerir que la terminología incluya un concepto separado para cada combinación detallada de ideas que potencialmente necesiten ser registradas. El software de aplicación que admite el uso de expresiones permite registrar, recuperar y analizar información clínica detallada.

Las expresiones son una combinación estructurada de uno o más identificadores de conceptos utilizados para representar una idea clínica de manera lógica, procesable automáticamente. Se representan con la gramática de composición de SNOMED CT, que es una sintaxis liviana para representar las expresiones de la terminología.

La lógica en la que se basa el modelo conceptual de SNOMED CT permite reconocer y comparar representaciones alternativas de una misma información o de

informaciones similares.

Las expresiones clínicas que utilizan conceptos de SNOMED CT pueden ser de dos tipos: expresiones precoordinadas, que utilizan un único identificador de concepto y expresiones poscoordinadas, que contienen más de un identificador.

#### **3.5.3.1. Expresiones precoordinadas**

Las expresiones precoordinadas son expresiones que representan el significado de conceptos individuales que están predefinidos en SNOMED CT. Además del identificador de concepto único y las descripciones, cada concepto también tiene una definición lógica formal representada por un conjunto de relaciones definitorias con otros conceptos.[14]

#### **3.5.3.2. Expresiones postcoordinadas**

Las expresiones que contienen dos o más identificadores de concepto se denominan expresiones postcoordinadas. La postcoordinación combina conceptos y permite agregar más detalles al significado representado por un solo concepto. Una expresión postcoordinada no es solo una lista de identificadores de conceptos, sino que sigue un conjunto de reglas que imitan la forma en que se usan los atributos y valores para definir los conceptos de SNOMED CT. En la poscoordinación se aplica aunque exista un único concepto para representar el significado requerido. Sin embargo, la verdadera utilidad de la poscoordinación es representar una frase clínica aunque el concepto exacto no esté presente en SNOMED CT. [14]

#### **3.5.4. Implementación**

SNOMED CT es solo parte de la solución para responder a los requerimientos de registros clínicos electrónicos efectivos. Una terminología por sí sola no hace nada. Para obtener algún beneficio de una terminología es necesario implementarla y utilizarla como parte de una aplicación. El diseño de la aplicación de software en la cual se utiliza y los objetivos y motivaciones de los usuarios son los factores clave que determinarán el éxito. [14]

La capacidad de SNOMED CT para facilitar la recuperación de datos basada en la información médica y el significado requiere una cuidadosa consideración del contexto real, en términos de alcance del uso, estructura del registro, ingreso y recuperación de datos y comunicación.

En el proceso de implementación de SNOMED CT, una herramienta valiosa que se utiliza es Browser Snowstorm. Es una plataforma desarrollada por la comunidad de SNOMED que ofrece diversas funcionalidades para consumir SNOMED CT de manera eficiente.

Snowstorm es un servidor que contiene el modelo terminológico y disponibiliza una API para que entre otras cosas, consultar esa terminología e implementarla en un contexto clínico. Sin embargo, es responsabilidad del propietario de cada instancia de Snowstorm elegir y configurar el modelo de SNOMED CT, incluyendo su nivel de actualización y extensiones específicas del país o región.

### 3.5.5. Modelo conceptual

El modelo conceptual de SNOMED CT especifica la manera de definir los conceptos de la terminología mediante una combinación de lógica formal y reglas editoriales. Las reglas del modelo conceptual especifican los conceptos del nivel superior debajo de los cuales se disponen los conceptos en la jerarquía de subtipos y los tipos de relaciones permitidas entre conceptos en ramas específicas de la jerarquía.

En el extremo superior de la jerarquía de SNOMED CT se encuentra el concepto raíz, “concepto de SNOMED CT”. Todos los conceptos descienden de este concepto raíz a través de al menos una secuencia de relaciones “*es un/una*”. Esto significa que el concepto raíz es un supertipo de todos los otros conceptos y éstos, a su vez, son subtipos de él.

Los subtipos directos del concepto raíz se denominan “conceptos del nivel superior”. Se utilizan para nombrar las ramas principales de la jerarquía. Cada uno de estos conceptos del nivel superior, junto con sus numerosos subtipos descendientes,

forma una de las ramas principales de la jerarquía de SNOMED CT y contiene tipos de conceptos similares. A medida que se desciende por las jerarquías, los conceptos incluidos en ellas son cada vez más específicos. La siguiente lista muestra los conceptos del nivel superior.[14]

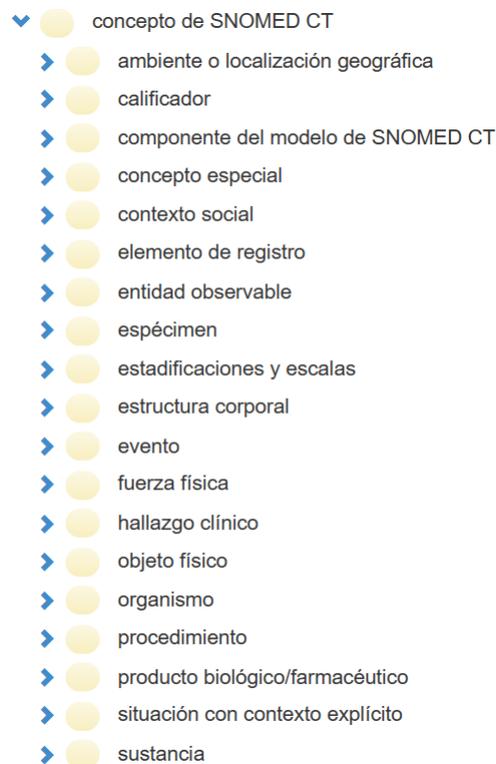
- 
- ▼ ● concepto de SNOMED CT
    - ▶ ● ambiente o localización geográfica
    - ▶ ● calificador
    - ▶ ● componente del modelo de SNOMED CT
    - ▶ ● concepto especial
    - ▶ ● contexto social
    - ▶ ● elemento de registro
    - ▶ ● entidad observable
    - ▶ ● espécimen
    - ▶ ● estadificaciones y escalas
    - ▶ ● estructura corporal
    - ▶ ● evento
    - ▶ ● fuerza física
    - ▶ ● hallazgo clínico
    - ▶ ● objeto físico
    - ▶ ● organismo
    - ▶ ● procedimiento
    - ▶ ● producto biológico/farmacéutico
    - ▶ ● situación con contexto explícito
    - ▶ ● sustancia

Figura 6: Lista de conceptos de nivel superior. Imagen obtenida de [15]

### 3.5.6. Restricciones de búsqueda

Con el propósito de lograr que el sistema de salud cuente con tecnologías de la información que permitan registrar información en forma primaria durante el contacto con el paciente, a través de sistemas interoperables que permitan compartir la información entre niveles y jurisdicciones, la organización SNOMED International desarrollo el Lenguaje de Restricciones de Expresiones (ECL) de SNOMED CT. [16]

ECL es una sintaxis formal que tiene como función representar restricciones de expresiones de SNOMED CT. Las restricciones de expresiones son reglas computables

que sirven para definir subconjuntos delimitados de ideas clínicas. Estas pueden ser utilizadas para restringir los valores válidos de un elemento de datos, por ejemplo en una historia clínica electrónica.

La integración de terminología es el vínculo entre el modelo de información y la terminología propiamente dicha. Estos vínculos pueden utilizarse para recortar o restringir el conjunto de valores posibles que puede asumir un determinado elemento del modelo de datos, o bien pueden definir el significado de un artefacto del modelo de información a través del uso de la terminología.

Esto es de suma importancia a la hora de dar soporte a algunas funciones del sistema de información, entre ellas la recuperación de información y consultas (queries) y la interoperabilidad semántica.

El lenguaje formal ECL permite la integración de terminología; a través del uso de reglas intensionales restringe el conjunto de códigos posibles para un modelo de información determinado.

Las ECL pueden utilizarse en su versión larga, que incluye los términos asociados a los Concept ID, haciéndola legible para el ojo humano. O puede directamente al momento de implementar utilizarse la versión corta, que solo contiene los ID de concepto.[16]

Algunos de los componentes de una restricción de expresiones los podemos ver en la siguiente figura:

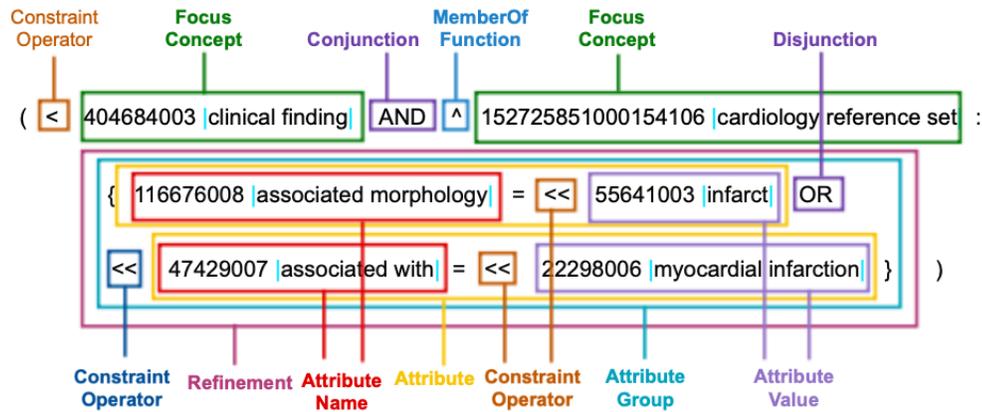


Figura 7: Componentes de una restricción de expresiones. Imagen modificada, obtenida de [14].

### 3.5.7. Enlace terminológico

Tal como vimos en las secciones anteriores, para conseguir un alto grado de interoperabilidad es imprescindible llevar a cabo un enlace terminológico entre los modelos de información y las terminologías.[17]

Existen dos tipos de enlace terminológico: [18]

- **Enlace semántico:** que proporciona significado inequívoco a las estructuras de información contenidas en el modelo de información mediante un enlace entre un elemento del modelo de información y un término precoordinado o postcoordinado de la terminología.
- **Enlace de contenido:** que restringe el conjunto de posibles valores codificados de la terminología, susceptibles de ser asociados a un elemento del modelo de información.

La mejor opción para llevar a cabo el enlace de contenido requiere un mecanismo de definición de subconjuntos de conceptos terminológicos, como HCE de SNOMED CT descrito en la sección 3.5.6.

Gracias a este lenguaje es posible definir los subconjuntos de conceptos médicos

de manera intencional, es decir, mediante una expresión susceptible de ser evaluada para calcular los conceptos que forman el subconjunto.

Es importante recordar que la gramática composicional SNOMED CT se emplea para definir conceptos postcoordinados (i.e. conceptos que no existen en la terminología pero que pueden ser requeridos en un momento dado), mientras que el lenguaje de restricciones se utiliza para definir subconjuntos de conceptos precoordinados, es decir, conceptos que existen como tales en la terminología. Esto lo vemos en la Figura 8.

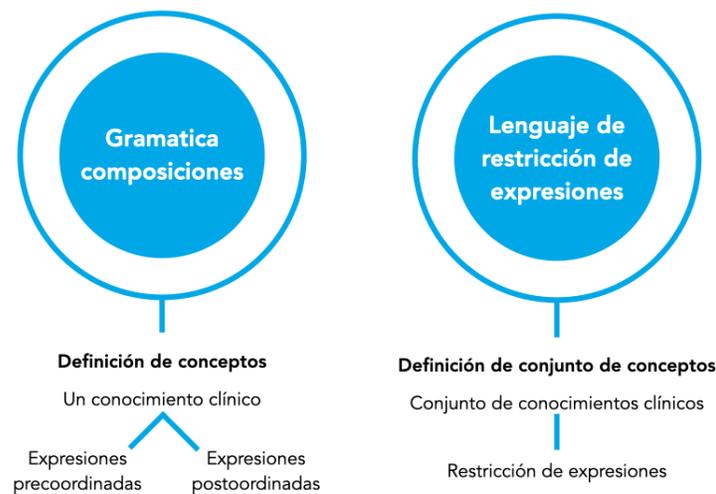


Figura 8: Gramática composicional y lenguaje de restricciones.

### 3.6. Estructura de la información clínica

Una limitación importante es la diversidad de puntos de vista relacionados con la estructura de la información clínica y la superposición entre los modelos de información y la terminología. También hay diferentes opiniones sobre el diseño de la aplicación, diferentes requisitos para la recopilación de información clínica y diferentes perspectivas sobre las estructuras de registros y los métodos de entrada de datos apropiados para diferentes casos de uso.

SNOMED International está trabajando con otros organismos de normalización,

incluida la Organización Internacional de Normalización (ISO) y Health Level 7 (HL7). Además, está desarrollando estrategias colaborativas, que exploran la relación entre la terminología y la información clínica estructurada. El objetivo es garantizar que el papel de SNOMED CT, como componente clave de la información, y los sistemas clínicos se entienda como parte de los esfuerzos generales hacia la armonización y la interoperabilidad.[14]

En términos generales, SNOMED CT podría considerarse un diccionario con la capacidad de consultar exactamente el significados de los conceptos médicos, que a diferencia de uno ordinario, sólo ofrece una definición para cada término.

Para complementar lo anterior el estándar terminológico requiere de uno sintáctico, que brinde directrices sobre como organizar la información, en qué tipo de formato y bajo qué etiquetas específicas.

Para ello, requiere un modelo de información, un estándar sobre cómo se deben diseñar y rellenar los registros y los mensajes. En esta instancia, es donde participa HL7 FHIR, se combinan dos estándares líderes mundiales de terminología y modelo de información. Al utilizar conjuntamente SNOMED CT y FHIR, los sistemas pueden comunicar de forma estandarizada un significado claro y inequívoco, que se comprende automáticamente en todo el mundo.

### **3.6.1. Health Level 7**

Health Level Seven es una “Organización de Desarrollo de Estándares” para el ámbito de salud. Fundada en 1987, sin fines de lucro, opera a nivel internacional y su misión es proveer estándares globales para los dominios: clínico, asistencial, administrativo y logístico, con el fin de facilitar el intercambio electrónico de información sanitaria. En definitiva, su misión es lograr una interoperabilidad clínica real entre los diferentes sistemas de información presentes en las organizaciones de salud.

La palabra “Health”(salud) refiere al área de trabajo de la organización y las palabras “Level seven” (nivel siete) hacen referencia al último nivel del modelo de

comunicaciones para sistemas abiertos (OSI) de la Organización Internacional para la Estandarización (ISO).[19]

### 3.6.2. Estándar FHIR

Se puede decir que FHIR es el último estándar de interoperabilidad clínico desarrollado por HL7. Es un estándar de especificación para el intercambio electrónico de información sanitaria. Es una especificación de plataforma que define un conjunto de capacidades utilizadas en todo el proceso de la atención médica, en todas las jurisdicciones, y en muchos contextos diferentes.[20]

FHIR facilita la eliminación de las barreras entre los distintos sistemas y la entrega de la información necesaria, a quien la necesite, casi en tiempo real.

FHIR permite el acceso a datos distribuidos en sistemas, bases de datos y dispositivos en tiempo real, como una “web para la sanidad”. Gracias a las aplicaciones basadas en el, los profesionales sanitarios en cualquier situación pueden consultar exactamente el conjunto de información que necesitan, como el historial de medicación, las alergias o los resultados de las pruebas. Esto ayuda a evitar la carga de tener que rebuscar en cientos de documentos para llegar a la información que necesitan.

El propósito principal de FHIR es abordar la interoperabilidad con modelos de datos expresivos y bien estructurados y mecanismos de intercambio de datos simples y eficientes. Además, FHIR se alinea con los siguientes principios arquitectónicos:

- Reutilización y composición: los recursos FHIR están diseñados teniendo en cuenta la regla 80/20: se centran en el 20% de los requisitos que satisfacen el 80% de las necesidades de interoperabilidad.
- Escalabilidad: la alineación de las API de FHIR con el estilo arquitectónico de REST garantiza que todas las transacciones sean sin estado, lo que reduce el uso de la memoria.
- Rendimiento: los recursos de FHIR son ajustados y optimizados para el inter-

cambio a través de la red.

- Usabilidad: los recursos FHIR son entendidos por expertos técnicos y personas no técnicas por igual.
- Fidelidad de los datos: FHIR está fuertemente tipado y tiene mecanismos incorporados para la vinculación y validación de la terminología clínica.
- Implementación: una de las fuerzas impulsoras de FHIR es la necesidad de crear un estándar con alta adopción en comunidades de desarrolladores dispares.

Mientras que los fundamentos de las especificaciones de FHIR son relativamente sencillas, se puede presentar una dificultad a la hora de implementar una solución basada en FHIR.

#### 3.6.2.1. Recursos

FHIR parte del concepto fundamental de recurso, la unidad básica de interoperabilidad y más pequeña que tiene sentido intercambiar. Los recursos son representaciones de conceptos del mundo sanitario: paciente, médico, problema de salud, etc. Los recursos se basan en estructuras XML o JSON que utilizan un protocolo REST.

Se definen diferentes tipos de recursos que se pueden usar para intercambiar y/o almacenar datos a fin de resolver una amplia gama de problemas relacionados con la atención médica, tanto clínicos como administrativos.[20]

Desde una perspectiva clínica, las partes más importantes de la especificación FHIR para comprender son los Recursos. Se entienden como “formularios” impresos que reflejan diferentes tipos de información clínica y administrativa que pueden capturarse y compartirse. La especificación FHIR define una “plantilla de formulario” genérica para cada tipo de información clínica.

Podemos definir un recurso como una entidad que: [21]

- Tiene una identidad única (ubicación *url*) asignada en el servidor.

- Pertenece a un tipo de recurso entre los definidos por la especificación FHIR.
- Contiene un conjunto de datos estructurados, en función del tipo de recurso.
- Tiene una versión que cambia si se modifica el contenido del recurso.

En la Figura 9 se ven los recursos existentes:

Alphabetical			
<b>A-C:</b>	<b>D-I:</b>	<b>L-P:</b>	<b>Q-Z:</b>
<ul style="list-style-type: none"> <li>Account 2</li> <li>ActivityDefinition 4</li> <li>ActorDefinition 1</li> <li>AdministrableProductDefinition 2</li> <li>AdverseEvent 2</li> <li>AllergyIntolerance 3</li> <li>Appointment 3</li> <li>AppointmentResponse 3</li> <li>ArtifactAssessment 1</li> <li>AuditEvent 4</li> <li>Basic 3</li> <li>Binary <b>N</b></li> <li>BiologicallyDerivedProduct 2</li> <li>BiologicallyDerivedProductDispense 0</li> <li>BodyStructure 1</li> <li>Bundle <b>N</b></li> <li>CapabilityStatement <b>N</b></li> <li>CarePlan 2</li> <li>CareTeam 2</li> <li>ChargeItem 1</li> <li>ChargeItemDefinition 1</li> <li>Citation 1</li> <li>Claim 2</li> <li>ClaimResponse 2</li> <li>ClinicalImpression 1</li> <li>ClinicalUseDefinition 2</li> <li>CodeSystem <b>N</b></li> <li>Communication 2</li> <li>CommunicationRequest 2</li> <li>CompartmentDefinition 3</li> <li>Composition 4</li> <li>ConceptMap 3</li> <li>Condition (aka Problem) 5</li> <li>ConditionDefinition 0</li> <li>Consent 2</li> <li>Contract 1</li> <li>Coverage 4</li> <li>CoverageEligibilityRequest 4</li> <li>CoverageEligibilityResponse 4</li> </ul>	<ul style="list-style-type: none"> <li>DetectedIssue 2</li> <li>Device 2</li> <li>DeviceAssociation 0</li> <li>DeviceDefinition 1</li> <li>DeviceDispense 0</li> <li>DeviceMetric 1</li> <li>DeviceRequest 1</li> <li>DeviceUsage 1</li> <li>DiagnosticReport 3</li> <li>DocumentReference 4</li> <li>Encounter 4</li> <li>EncounterHistory 0</li> <li>Endpoint 2</li> <li>EnrollmentRequest 0</li> <li>EnrollmentResponse 0</li> <li>EpisodeOfCare 2</li> <li>EventDefinition 0</li> <li>Evidence 1</li> <li>EvidenceReport 0</li> <li>EvidenceVariable 1</li> <li>ExampleScenario 1</li> <li>ExplanationOfBenefit 2</li> <li>FamilyMemberHistory 2</li> <li>Flag 1</li> <li>FormularyItem 0</li> <li>GenomicStudy 0</li> <li>Goal 2</li> <li>GraphDefinition 2</li> <li>Group 3</li> <li>GuidanceResponse 2</li> <li>HealthcareService 4</li> <li>ImagingSelection 1</li> <li>ImagingStudy 4</li> <li>Immunization 5</li> <li>ImmunizationEvaluation 1</li> <li>ImmunizationRecommendation 1</li> <li>ImplementationGuide 4</li> <li>Ingredient 2</li> <li>InsurancePlan 0</li> <li>InventoryItem 0</li> <li>InventoryReport 0</li> <li>Invoice 0</li> </ul>	<ul style="list-style-type: none"> <li>Library 4</li> <li>Linkage 0</li> <li>List 4</li> <li>Location 5</li> <li>ManufacturedItemDefinition 2</li> <li>Measure 4</li> <li>MeasureReport 4</li> <li>Medication 4</li> <li>MedicationAdministration 2</li> <li>MedicationDispense 2</li> <li>MedicationKnowledge 1</li> <li>MedicationRequest 4</li> <li>MedicationStatement 4</li> <li>MedicinalProductDefinition 3</li> <li>MessageDefinition 1</li> <li>MessageHeader 4</li> <li>MolecularSequence 1</li> <li>NamingSystem 4</li> <li>NutritionIntake 1</li> <li>NutritionOrder 2</li> <li>NutritionProduct 1</li> <li>Observation <b>N</b></li> <li>ObservationDefinition 1</li> <li>OperationDefinition <b>N</b></li> <li>OperationOutcome <b>N</b></li> <li>Organization 5</li> <li>OrganizationAffiliation 1</li> <li>PackagedProductDefinition 2</li> <li>Parameters <b>N</b></li> <li>Patient <b>N</b></li> <li>PaymentNotice 4</li> <li>PaymentReconciliation 4</li> <li>Permission 0</li> <li>Person 4</li> <li>PlanDefinition 4</li> <li>Practitioner 5</li> <li>PractitionerRole 4</li> <li>Procedure 4</li> <li>Provenance 4</li> </ul>	<ul style="list-style-type: none"> <li>Questionnaire 5</li> <li>QuestionnaireResponse 5</li> <li>RegulatedAuthorization 2</li> <li>RelatedPerson 5</li> <li>RequestOrchestration 4</li> <li>Requirements 1</li> <li>ResearchStudy 0</li> <li>ResearchSubject 0</li> <li>RiskAssessment 2</li> <li>Schedule 3</li> <li>SearchParameter 5</li> <li>ServiceRequest 4</li> <li>Slot 3</li> <li>Specimen 2</li> <li>SpecimenDefinition 1</li> <li>StructureDefinition <b>N</b></li> <li>StructureMap 4</li> <li>Subscription 3</li> <li>SubscriptionStatus 2</li> <li>SubscriptionTopic 2</li> <li>Substance 2</li> <li>SubstanceDefinition 1</li> <li>SubstanceNucleicAcid 0</li> <li>SubstancePolymer 0</li> <li>SubstanceProtein 0</li> <li>SubstanceReferenceInformation 0</li> <li>SubstanceSourceMaterial 0</li> <li>SupplyDelivery 1</li> <li>SupplyRequest 1</li> <li>Task 3</li> <li>TerminologyCapabilities 1</li> <li>TestPlan 0</li> <li>TestReport 1</li> <li>TestScript 4</li> <li>Transport 1</li> <li>ValueSet <b>N</b></li> <li>VerificationResult 1</li> <li>VisionPrescription 3</li> </ul>

Figura 9: Recursos disponibles en la página de FHIR en un orden alfabético en inglés. Imagen obtenida de [20]

A todos los recursos de esta especificación se les asigna un “Nivel de madurez”,

conocido como FMM. El nivel FMM puede ser utilizado por los implementadores para juzgar qué tan avanzado y, por lo tanto, estable es un recurso. Se definen los siguientes niveles:

- Borrador (0): Este nivel es sinónimo de Borrador.
- FMM 1: Esta sustancialmente completo y listo para su implementación.
- FMM 2: El recurso ha sido probado y admite con éxito la interoperabilidad entre al menos tres sistemas desarrollados independientemente.
- FMM 3: El artefacto ha sido verificado por el grupo de trabajo, cumpliendo con las Pautas de Calidad de Recursos de Conformidad.
- FMM 4: El artefacto ha sido probado en todo su alcance, publicado en una publicación formal e implementado en múltiples proyectos prototipo.
- FMM 5: el artefacto se ha publicado en dos ciclos formales de publicación y se ha implementado en al menos 5 sistemas de producción independientes en más de un país.
- Normativo: El artefacto ahora se considera estable.

#### **3.6.2.2. Categorías**

Los recursos se clasifican en seis secciones: clínica, base, financiero, fundación y especializado.

Se va profundizar en la sección clínica que incluye cinco categorías que se muestran en la Figura 10.

Summary	Diagnostics	Medications	Care Provision	Request & Response
<ul style="list-style-type: none"> <li>• AllergyIntolerance 3</li> <li>• AdverseEvent 2</li> <li>• Condition (Problem) 5</li> <li>• Procedure 4</li> <li>• FamilyMemberHistory 2</li> <li>• ClinicalImpression 1</li> <li>• DetectedIssue 2</li> </ul>	<ul style="list-style-type: none"> <li>• Observation <b>N</b></li> <li>• DocumentReference 4</li> <li>• DiagnosticReport 3</li> <li>• Specimen 2</li> <li>• BodyStructure 1</li> <li>• ImagingSelection 1</li> <li>• ImagingStudy 4</li> <li>• QuestionnaireResponse 5</li> <li>• MolecularSequence 1</li> <li>• GenomicStudy 0</li> </ul>	<ul style="list-style-type: none"> <li>• MedicationRequest 4</li> <li>• MedicationAdministration 2</li> <li>• MedicationDispense 2</li> <li>• MedicationStatement 4</li> <li>• Medication 4</li> <li>• MedicationKnowledge 1</li> <li>• Immunization 5</li> <li>• ImmunizationEvaluation 1</li> <li>• ImmunizationRecommendation 1</li> <li>• FormularyItem 0</li> </ul>	<ul style="list-style-type: none"> <li>• CarePlan 2</li> <li>• CareTeam 2</li> <li>• Goal 2</li> <li>• ServiceRequest 4</li> <li>• NutritionOrder 2</li> <li>• NutritionIntake 1</li> <li>• VisionPrescription 3</li> <li>• RiskAssessment 2</li> <li>• RequestOrchestration 4</li> </ul>	<ul style="list-style-type: none"> <li>• Communication 2</li> <li>• CommunicationRequest 2</li> <li>• DeviceRequest 1</li> <li>• DeviceDispense 0</li> <li>• DeviceAssociation 0</li> <li>• DeviceUsage 1</li> <li>• BiologicallyDerivedProductDispense 0</li> <li>• GuidanceResponse 2</li> <li>• SupplyRequest 1</li> <li>• SupplyDelivery 1</li> <li>• InventoryItem 0</li> <li>• InventoryReport 0</li> </ul>

Figura 10: Categorías de la sección clínica. Imagen obtenida de [20].

Con un enfoque en la codificación de diagnósticos, se describen los recursos relevantes para este proyecto. Dentro la categoría Diagnósticos se encuentran:

- **Observation:** los usos del recurso de observación incluyen: signos vitales, datos de laboratorio, resultados de imágenes, hallazgos clínicos, mediciones del dispositivo, herramientas de evaluación clínica, características personales, antecedentes sociales y características básicas. En esencia, Observation permite expresar un par nombre-valor o una colección estructurada de pares nombre-valor.
- **Media:** recurso multimedia contiene fotos, vídeos y grabaciones de audio. Se utiliza con medios adquiridos o utilizados como parte del proceso sanitario.
- **DiagnosticReport:** tiene información sobre el informe de diagnóstico en sí, y sobre el tema y, en el caso de las pruebas de laboratorio, la muestra del informe. También puede referirse a los detalles de la solicitud y los detalles de las observaciones atómicas o las instancias de imagen. Las conclusiones del informe se pueden expresar como un blob de texto simple, datos codificados estructurados o como un informe adjunto con formato completo, como un PDF.
- **Specimen:** cubre sustancias utilizadas para pruebas de diagnóstico y ambientales. El enfoque del recurso de especímenes es el proceso para recolectar, mantener y procesar el espécimen, así como el lugar donde se originó el espécimen.

men. Esto es distinto del uso de la sustancia, que solo se utiliza cuando estos otros aspectos no son relevantes.

- **Body Structure:** contiene detalles sobre la ubicación anatómica de una muestra o parte del cuerpo, incluida información del paciente, identificadores, así como descripciones de texto e imágenes. Prevé la adición de calificativos como lateralidad y direccionalidad a la ubicación anatómica para aquellos casos de uso en los que no sea posible la coordinación previa de los códigos. El recurso `BodyStructure` admite el registro y seguimiento de una ubicación o estructura anatómica en un paciente fuera del contexto de otro recurso.
- **ImagingStudy:** proporciona información sobre un estudio de imágenes DICOM, y la serie y los objetos de imágenes en ese estudio. También proporciona información sobre cómo recuperar esa información.
- **QuestionnaireResponse:** proporciona una lista completa o parcial de respuestas a un conjunto de preguntas completadas al responder a un cuestionario. Algunos de los ejemplos de cuestionarios incluyen: antecedentes médicos, enfermedades familiares, historia social, cuestionarios de investigación y formularios de informes de casos y formularios.
- **MolecularSequence:** está diseñado para describir una secuencia atómica que contiene el resultado de la prueba de secuenciación de alineación y múltiples variaciones.

A continuación se analiza la categoría de `Summary`. La misma cuenta con siete recursos como se ve en la Figura 10, dentro de estos se analizó el siguiente:

- **ClinicalImpression:** es un resumen clínico de información y/o una opinión formada, que es el resultado del proceso de evaluación clínica. La impresión clínica puede conducir a una declaración de una condición sobre un paciente. Este recurso está destinado a ser utilizado para cubrir evaluaciones clínicas que simplemente dan como resultado una impresión registrada como una sola nota de texto en el “registro” del paciente o evaluaciones que otros se asocian con

un registro cuidadoso y detallado de la evidencia recopilada y el razonamiento que conduce a un diagnóstico diferencial.

### 3.6.2.3. Campos

Cada recurso se compone de una estructura de datos, una serie de campos, que aportarán la información determinada. Estos campos se podrán completar con un tipo de dato concreto. La especificación FHIR define un conjunto de tipos de datos que se utilizan para los elementos de un recursos. Hay cuatro categorías de tipos de datos: [19]

- **Simple:** elementos individuales con un valor primitivo. Tienen solo un valor no cuentan con elementos adicionales como hijos.
- **Complejos:** grupos de elementos reutilizables. Cuenta con elementos secundarios.
- **Metadatos:** grupo utilizado para transmitir datos que describen otros datos.
- **De uso especial:** definidos en otra parte de la especificación para usos específicos.

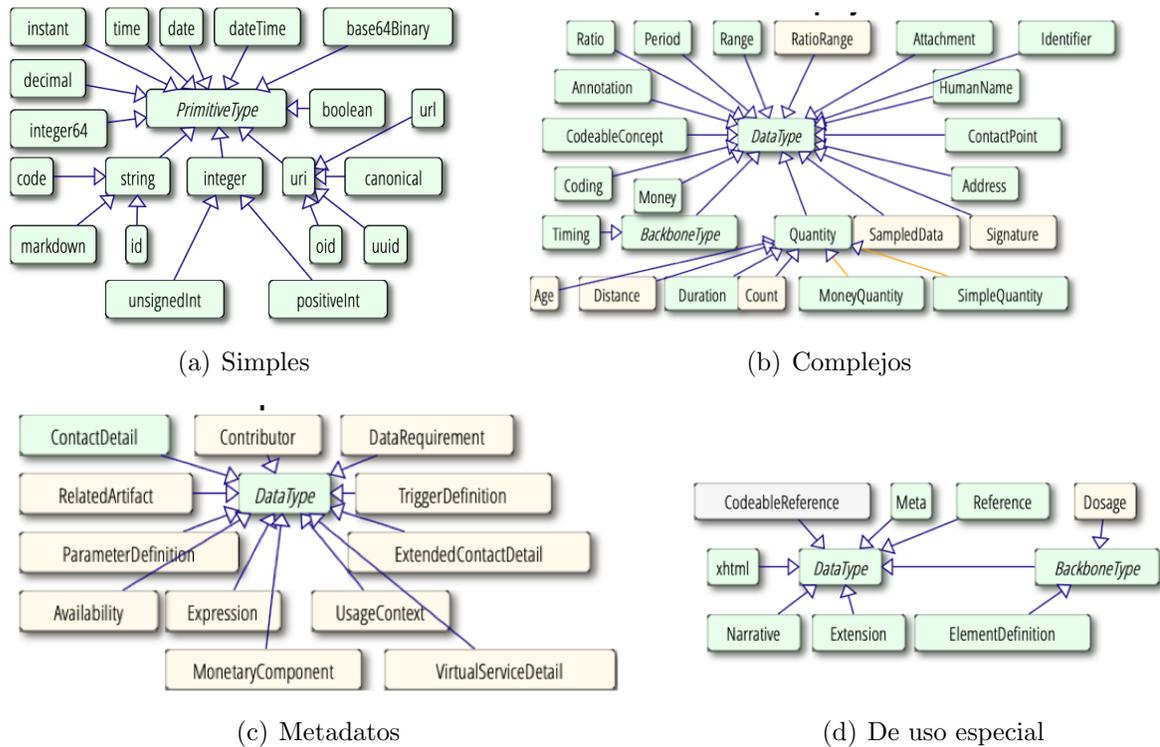


Figura 11: Descripciones de tipos de datos en FHIR. Imagen modificada extraída de [20]

### 3.7. Implementación

La documentación de FHIR indica 2 formas de realizar o utilizar los recursos: crear una programación propia para su implementación o utilizar el código y librerías existentes sobre FHIR.

#### 3.7.1. Lenguaje de programación

El lenguaje de programación utilizado es Python. Por su simplicidad, facilidad para crear Servicios Web Rest y viabilidad para utilizar JSON.

Python es un lenguaje de programación interpretado, orientado a objetos de alto nivel y con semántica dinámica. Su sintaxis hace énfasis en la legibilidad del código, lo que facilita su depuración y, por tanto, favorece la productividad. Se trata de un

lenguaje de programación multiparadigma, ya que soporta parcialmente la orientación a objetos, programación imperativa y, en menor programación funcional. Es un lenguaje interpretado, dinámico y multiplataforma.[22]

#### **3.7.1.1. Python Flask**

Flask es un microframework escrito en Python y desarrollado para simplificar y hacer más fácil la creación de Aplicaciones Web.

Un framework web es una arquitectura de software que contiene herramientas y librerías utilizadas para desarrollar una aplicación web de forma rápida y eficiente. Representa una colección de bibliotecas y módulos que permiten a los desarrolladores de aplicaciones web escribir aplicaciones sin preocuparse por detalles de bajo nivel como protocolo.

Proporciona solo funcionalidades básicas que incluyen validación de formularios, manejo de carga, mapeadores relacionales de objetos, autenticación abierta, etc. Usando estos, uno puede construir sitios web a pequeña y gran escala. Aunque Flask no incluye una capa de abstracción de base de datos ni funcionalidades adicionales de forma nativa, ofrece extensiones que permiten implementar estas características según las necesidades del proyecto. [23]

Es un microframework, pero eso no significa que toda su aplicación deba estar dentro de un solo archivo de Python. Puede y debe usar muchos archivos para programas más grandes, para manejar la complejidad. Micro significa que el marco Flask es simple pero extensible.[24]

#### **3.7.2. HAPI FHIR**

Existen distintas aplicaciones comerciales en el ámbito clínico que soportan FHIR de manera nativa, pero para una solución custom que se adapte a nuestras aplicaciones, podemos utilizar una librería especializada. HAPI FHIR es una librería Java para Clientes y Servidores que implementa la especificación HL7 FHIR, de código abierto y disponible bajo licencia Apache Software License 2.0.

HAPI FHIR define un modelo de clase para cada recurso y tipo de datos definido en la especificación FHIR, que a su vez podemos codificar en XML o JSON, para su intercambio a través de las APIs REST.

Esta interfaz de programación de aplicaciones HL7 ha existido durante mucho tiempo como una biblioteca de referencia para la incorporación de HL7 v2 en aplicaciones escritas en Java. La comunidad HAPI se ha centrado en la creación de una biblioteca FHIR que permita consumir y exponer las API de FHIR. El proyecto es rico en características que incluyen: el servidor, excelente documentación, una comunidad de apoyo y versiones frecuentes (corrección de errores, adición de características y ampliación del soporte para nuevos recursos y estándares de FHIR). [25]

Uno de los escenarios de uso descritos de HAPI destacado es:[26]

- Utilizar el servidor HAPI FHIR en una aplicación para permitir que las aplicaciones externas accedan o modifiquen los datos de su aplicación.

### 3.7.3. Versiones

Antes de intentar usar FHIR es necesario determinar qué versión de FHIR desea admitir en su aplicación. Por lo general esta sería la última versión, pero si se busca interactuar con una aplicación que ya existe, se debe considerar la misma versión implementada por esa aplicación.[27]

## 3.8. Herramientas de software

Los conceptos de esta sección facilitan la interacción entre diferentes sistemas y la integración de funcionalidades. Cada uno de ellos se utiliza en diferentes contextos dentro del desarrollo del software.

### 3.8.1. Servicios web

Un servicio web es una tecnología que maneja un conjunto de protocolos y estándares que sirven para intercambiar datos entre aplicaciones.

REST es cualquier interfaz entre sistemas que use HTTP para obtener datos o generar operaciones sobre esos datos en todos los formatos posibles, como XML y JSON. Es una alternativa en auge a otros protocolos de intercambio de datos como SOAP (Simple Object Access Protocol), que disponen de una gran capacidad, pero también mucha complejidad.[28]

### 3.8.2. JSON

JSON es un formato ligero de intercambio de datos. Leerlo y escribirlo es simple para humanos, mientras que para las máquinas es simple interpretarlo y generarlo. Está basado en un subconjunto del Lenguaje de Programación JavaScript. Es un formato de texto que es completamente independiente del lenguaje pero utiliza convenciones que son ampliamente conocidos por los programadores de distintas familia de lenguajes.

JSON está constituido por dos estructuras. Por un lado, una colección de pares de nombre/valor en varios lenguajes esto es conocido como un objeto, registro, estructura, diccionario, tabla hash, lista de claves o un arreglo asociativo. Por otro, una lista ordenada de valores que en la mayoría de los lenguajes, esto se implementa como arreglos, vectores, listas o secuencias.[29]

### 3.8.3. API

El término API es una abreviatura de Application Programming Interfaces, que en español significa interfaz de programación de aplicaciones. Se trata de un conjunto de definiciones y protocolos que se utilizan para desarrollar e integrar el software de las aplicaciones, permitiendo la comunicación entre dos aplicaciones de software a través de un conjunto de reglas.

Establece cómo un módulo de un software se comunica o interactúa con otro para cumplir una o muchas funciones, pudiendo llegar a ser auténticos kits de herramientas. Cuando esto pasa, una aplicación puede enviar una solicitud con una estructura particular, y esta estructura determina cómo responde el servicio o el software. [30]

Una API se basa en tecnologías web como HTTP y el protocolo RESTful, pudiendo utilizar formatos XML, JSON o RDF para la representación de datos. De manera general, una URI cuenta con un formato dónde se distingue la dirección del servidor, el tipo de recurso, y su identificador.[21]

Las principales interacciones con la API FHIR utilizarán los métodos HTTP habituales:

- POST: Creación de recurso.
- PUT: Modificación de recurso.
- PATCH: Modificación parcial de un recurso.
- GET: Lectura de un recurso.
- DELETE: Eliminación de un recurso.

Con estas interacciones se realizan peticiones al servidor y se obtiene una respuesta de el.

#### **3.8.4. Microservicios**

Los microservicios son un enfoque arquitectónico y organizativo para el desarrollo de software donde el software está compuesto por pequeños servicios independientes que se comunican a través de APIs bien definidas. Los propietarios de estos servicios son equipos pequeños independientes.

Las arquitecturas de microservicios hacen que las aplicaciones sean más fáciles de escalar y más rápidas de desarrollar. Esto permite la innovación y acelera el tiempo de comercialización de las nuevas características. [31]

#### **3.8.5. Librería**

En programación una librería responde al conjunto de funcionalidades que permiten al usuario llevar a cabo nuevas tareas que antes no se podían realizar.[32]

Las librerías de Python responden al conjunto de implementaciones que permiten codificar este lenguaje, con el objeto de crear una interfaz independiente.

Se forman por códigos reutilizables y se utilizan para facilitar el desarrollo de aplicaciones, proporcionando funciones y métodos que se pueden integrar en el código de la aplicación.

Cada una de las librerías de Python disponen de diferentes módulos, con funciones específicas y que también varían en función del sistema operativo con el que se trabaje.

Existen librerías de Python básicas o bibliotecas estándar de Python que vienen ya junto a Python.

Para crear una librería de Python se necesitan los siguientes archivos.[33]

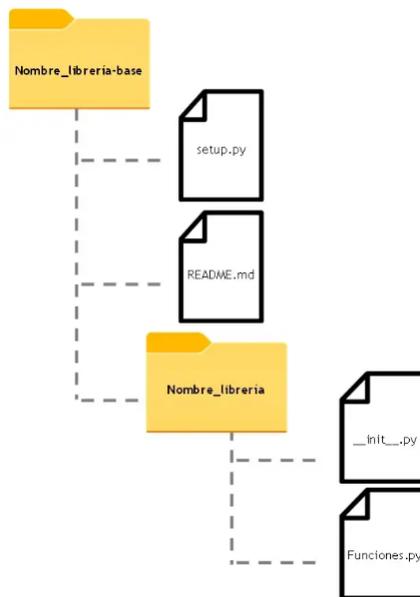


Figura 12: Archivos de la estructura de una librería. Imagen extraída de [33]

Cada archivo incluye:

- `Funciones.py`: Archivo que se utiliza para almacenar las funciones. No hay límite en la cantidad de ficheros de funciones, especialmente si nos ayudan

a estructurar nuestro código de manera adecuada. El nombre del archivo no necesariamente debe ser “funciones”.

- `setup.py`: En este documento se concretan las instrucciones necesarias para instalar la librería como las librerías de las que se depende y otros requisitos previos de instalación.
- `README.md`: Aunque es opcional, se recomienda crear este archivo, ya que proporciona información sobre nuestra librería: como instalarla, dependencias, ejemplos de uso, etc. El formato en el que se escribe este fichero es Markdown.
- `_init_.py`: Es el fichero fundamental para la estructura. Al existir un fichero con este nombre, Python reconoce una librería. Contiene las importaciones correspondientes del fichero (o ficheros) de funciones.

## 4. Desarrollo

### 4.1. Definición de campo diagnóstico

Como primer paso en el desarrollo de este proyecto, se delimitaron claramente los conceptos que se considera correcto incluir en el campo “diagnósticos”. Se tomó como referencia la definición de diagnóstico como la identificación del estado de salud o enfermedad que lleva a acciones o tratamientos adecuados posteriores.

Se tomó un conjunto de datos y se los analizó para tener una noción de los tipos de diagnósticos que se utilizaban y de la forma en que se escribían.

De estos el 38 % de los mismos eran considerados como diagnósticos según la definición. Sin embargo, dentro de este porcentaje menos de la mitad fueron reconocidos por la terminología de SNOMED. Debido principalmente a las abreviaturas, sinónimos o errores en la escritura de los textos. De esta manera se pudieron comprender las necesidades locales y el uso de “jergas” que utilizan los profesionales al interactuar con estos sistemas. Proporcionando una idea del escenario de diagnóstico necesario y de las necesidades particulares de los usuarios de la aplicación.

Dentro del 52 % restante, se analizaron textos que contienen información de síntomas o signos que no son diagnósticos, pero que a menudo se utilizan incorrectamente como tales. Estos síntomas y signos son relevantes para el proceso de diagnóstico médico pero no son suficientes para asignar recursos y determinar necesidades de manera precisa. Ejemplos de estos síntomas incluyen dolor, sensibilidad o sensaciones (como tos, dolor, bostezos, ronquidos, fiebre, entre otros).

Por otro lado, los textos que describían procedimientos quirúrgicos que también eran muy frecuentemente escritos como diagnósticos sí fueron considerados para incluir en el campo diagnóstico. Un procedimiento quirúrgico se basa en el diagnóstico médico del paciente, por lo cual el procedimiento quirúrgico aplicado se puede relacionar con el diagnóstico que se está tratando. A partir de ellos se puede predecir el diagnóstico y se puede generar acciones como un posterior tratamiento adecuado.

Por ejemplo, el termino “reemplazo total de cadera” es un procedimiento quirúrgico que se realiza para tratar las condiciones que causan dolor y limitación de la movilidad. A partir del mismo se infiere que el paciente esta tratando esta condición.

El campo diagnostico para Emergencias permite diagnósticos estrictos y procedimientos quirúrgicos.

## 4.2. Arquitectura de software

En esta sección se describe la estructura fundamental y organización del sistema, incluyendo la forma en que los componentes del software interactúan entre sí.

La arquitectura establece las bases para el diseño, desarrollo, implementación y mantenimiento del software. Permite una mayor modularidad, escalabilidad y facilidad de mantenimiento, ya que cada componente es independiente y se puede modificar sin afectar a los otros.

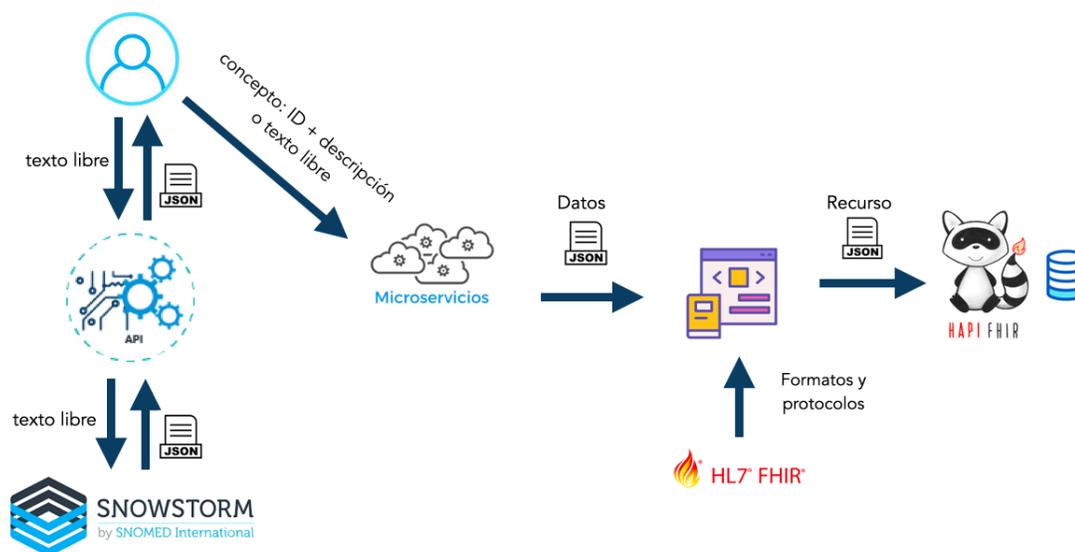


Figura 13: Arquitectura de software del sistema implementado.

La API “Snomed Suggestion” es el punto de entrada que recibe las solicitudes del usuario y devuelve las respuestas correspondiente del servicio de terminología.

Cuando el usuario elige una opción, el resultado se manda al microservicio. Luego el microservicio se comunica con la librería “ds utils” y devuelve los datos necesarios para estructurarlos junto a la respuesta de la API.

Finalmente se comunica con una API de HAPI FHIR que crea el recurso y lo almacena en su base de datos.

### 4.3. Snomed suggestion

Se desarrolló una API denominada “Snomed suggestion”. Esta API utiliza la terminología de SNOMED CT. Aplica las restricciones de búsqueda (contiene la ECL) y principalmente, codifica el concepto. Recibe como entrada un diagnóstico en texto libre y que devuelve como salida conceptos de selección específicos que corresponden con la definición de campo diagnóstico establecida. Además, incluye un adecuado procesamiento del lenguaje natural que se aplica al texto libre ingresado.

Para esto, se utiliza una funcionalidad presente en muchas aplicaciones y sitios web, donde los usuarios pueden ingresar un texto en un campo de búsqueda, seleccionar el botón de ‘buscar’ y luego se les muestran los resultados relacionados con ese texto. Las partes detalladas del funcionamiento son las siguientes:

- Interfaz de búsqueda: La aplicación proporciona una interfaz de búsqueda donde los usuarios pueden ingresar su texto.
- Ingreso del texto de búsqueda: El usuario escribe el diagnóstico que desea buscar en el campo de texto.
- Selección del botón “Buscar”: Una vez ingresado el texto, se selecciona el botón “Buscar” que inicia el proceso de búsqueda.
- Procesamiento de la búsqueda: La aplicación toma el texto ingresado por el usuario y realiza la búsqueda de los resultados en el Browser de Snowstorm.
- Presentación de los resultados: Si el texto ingresado es encontrado, despliega en una lista las opciones de conceptos que están relacionados y que se encuentran

en el rango de la ECL definida.

- Interacción con los resultados: El usuario examina la lista de resultados presentados, puede elegir uno de ellos haciendo clic en él o puede dejar el texto original seleccionando el botón “Dejar texto”. Esto último resulta útil si el usuario decide que la búsqueda no es relevante o desea realizar una búsqueda diferente.



Figura 14: Front representativo.

Como resultado, si existe un código para el texto ingresado este se codifica automáticamente. Luego se ingresa otro diagnóstico y se repite el proceso. Si no codifica el término o no se muestran sugerencias correctas se deja el contenido escrito ya que las sugerencias no son obligatorias.

### 4.3.1. Métodos y bases utilizados

Esta API consume a su vez una API de Snowstorm que proporciona información sobre la estructura interna de SNOMED. Para acceder a ella se necesita la *url* base de la API y un método HTTP adecuado. Además utiliza headers para indicar las preferencias de respuesta.

La *url* base de la API es la dirección principal que identifica el punto de entrada para todas las solicitudes a la API. En este caso:

```
“https://snowstorm.msal.gov.ar:443/MAIN/concepts?activeFilter=true&ecl=(<<64572001 OR <<387713003)&term=”
```

Donde los componentes son:

- **“https://”**: Indica el uso del protocolo HTTPS para realizar la comunicación segura con el servidor
- **“snowstorm.msal.gov.ar:443”**: es el dominio de la API Snowstorm del Ministerio de Salud de Argentina y el número de puerto utilizado para la comunicación segura a través de HTTPS.
- **“/MAIN/concepts”**: Indica que se está accediendo al endpoint “concepts” dentro del contexto “MAIN” en la API Snowstorm.
- **“?activeFilter=true&ecl=(<<64572001 OR << 387713003)&term=”** son los parámetros de la solicitud que se envían junto con la *url* y se separan con ‘&’. En este caso, se establece el parámetro “activeFilter” como “true” que devuelve conceptos activos y se define la expresión de consulta ECL que delimita los conceptos a obtener. El parámetro “term” contiene la información del texto libre que representa al diagnóstico.

Los “headers” que proporcionan información adicional sobre la solicitud:

- **“Accept:application/json”**: Especifica el tipo de contenido que se acepta en la respuesta del servidor. En este caso en formato JSON.

- **“Accept-Language:es”**: Indica una respuesta en español.

Los métodos HTTP especifican el tipo de acción que se realiza en la API. En este caso es “GET” ya que se realiza una solicitud para obtener información de la API Snowstorm.

### 4.3.2. Procesamiento del lenguaje

Se cuenta con herramientas de procesamiento del lenguaje natural para la gestión terminológica. Se implementan funciones y herramientas léxicas que facilitan la búsqueda de términos.

#### 4.3.2.1. Variaciones

Con el objetivo de optimizar el uso y la codificación, se incorporo un diccionario rico en significado y con gran variedad léxica que permite aumentar la capacidad de reconocimiento de textos.

Se creo especialmente para la situación donde va a utilizarse y esta adaptado al modelo del servidor de terminología utilizado (SNOMED). Fue necesaria la participación de los usuarios en su desarrollo.

Se forma a partir una lista estándar de 2805 diagnósticos de la institución ingresados como texto libre y la participación de los médicos de la misma. Se define la jerga local, abreviaturas y acrónimos utilizados y variantes léxicas. Este diccionario es dinámico y siempre va creciendo de acuerdo a las necesidades de los usuarios.

Este diccionario fue rigurosamente corregido y revisado por médicos de la empresa para asegurar su validez. Se encuentra detallado en el Anexo (Apéndice A).

Por otro lado, para mejorar la calidad de las opciones encontradas se aplica al texto ingresado funciones que contemplan las diferentes variaciones morfológicas de las palabras, como de género, número, inflexiones, derivaciones.

La variación de número como la singularización es una de las mas necesarias para la posterior codificación. Presenta cierta dificultad ya que existen muchas reglas que

están relacionadas con las letras finales de la palabra así como excepciones.

#### **4.3.2.2. Normalización**

El texto ingresado por los usuarios es procesado para eliminar diferencias inherentes al formato, necesita cierta “normalización” antes de ser usado. Como el objetivo es de crear un programa eficaz y eficiente, para normalizar el texto se utilizó un mecanismo llamado “expresiones regulares”. Las expresiones regulares permiten buscar patrones bien definidos para luego substituirlos, separarlos o eliminarlos.

Esto comprende la eliminación de signos de puntuación, espacios de más, artículos y preposiciones. Además, se realiza una conversión de todo a minúsculas ya que típicamente los componentes léxicos son compactados como minúsculas y así facilita su manipulación.

Por último se suprime del texto ingresado todos aquellos caracteres como el guión del medio (-), guión bajo (\_), barras (/), corchetes ([ ]), comillas (“).

Aunque algunas de estas modificaciones pueden cambiar el sentido literal de una frase, no cambian el concepto a asignar.

#### **4.3.2.3. Errores ortográficos**

Cuando se utiliza el buscador de SNOMED para buscar un término específico, este se basa en la coincidencia exacta entre la consulta realizada y los términos almacenados en su base de datos. Esto significa que si se introduce una mala escritura, como una letra incorrecta o un orden incorrecto de las letras, es probable que el buscador no pueda encontrar el término deseado.

Si se cometen errores ortográficos, el buscador no tiene la capacidad de interpretar la intención del usuario y realizar correcciones automáticas. Por lo tanto, juega un papel muy importante la detección de estos errores. El usuario tiene completa libertad al momento de ingresar el texto. Esto aumenta las posibilidades de ingreso de faltas de ortografía.

Para solucionar esto se utiliza un algoritmo que se alimenta de un diccionario de palabras correctamente escritas. Este diccionario contiene los términos mas utilizados y mas propensos a errores de los diagnósticos analizados.

El corrector ortográfico mide la distancia entre dos palabras y busca en el diccionario aquellas que la menor para el reemplazo.

El diccionario tiene la posibilidad de actualizarse constantemente a lo largo del tiempo según las necesidades y de forma manual o semiautomática.

### 4.3.3. ECL definida

A partir de la investigación bibliográfica, revisando reglamentos, manuales y normas (Secc.3) entendiendo los sistemas de información de salud, la terminologías clínicas y de la definición del campo diagnóstico acordada; se define la ECL que delimita los subconjuntos de conceptos disponibles en la búsqueda.

Cuando se implementa el lenguaje de restricciones de expresiones de SNOMED CT, los factores que hay que tener en consideración dependen del tipo de tareas que se quieren llevar a cabo, en este caso, codificar diagnósticos.

Existen recomendaciones del Centro Nacional de Terminología para aplicar restricciones de expresiones en una historia clínica electrónica en distintos campos. La recomendación para el campo Diagnostico-Problema es incluir todos los conceptos descendientes de tres jerarquías de nivel superior: “Hallazgo clínico”, “Evento” y “Situación con contexto explícito” [34]:

```
(<< 243796009 | situación con contexto explícito (situación)| OR << 272379006 |evento (evento)| OR << 404684003 | hallazgo clínico (hallazgo))
```

Se partió de esta sugerencia y se adapto con el fin de que el subconjunto de ideas clínicas se corresponda al criterio de un diagnostico para Emergencias.

Se descartó las jerarquías de “Evento” que contiene sucesos que afectan a la salud o la atención médica. Así también como “Situación con contexto explícito” que

define conceptos que incluyen información de contexto. Ambos subniveles contienen conceptos que corresponden a otros campos.

Por otro lado, se analizó en profundidad el subnivel de “Hallazgo clínico”. SNO-MED define dentro del nivel superior de hallazgo clínico:

- Trastorno: siempre y necesariamente un estado clínico anormal. Tiene un proceso patológico subyacente.[14]
- Hallazgos: observaciones, juicios o evaluaciones normales/anormales de pacientes. [14]

Considerando que un diagnóstico de forma estricta cumple la primera definición, se define la siguiente restricción restricción que incluye todos los conceptos identificados con ‘(trastornos)’:

ECL:  $\llcorner$  64572001 | enfermedad (trastorno) |

Como para Emergencias además de diagnósticos estrictos se consideran procedimientos quirúrgicos en el campo diagnóstico (Secc.4.1), se amplió este rango agregando subniveles del nivel superior “Procedimiento”. Este representa actividades que se llevan a cabo durante la atención de la salud. Se tomó un subnivel hijo:

- Procedimiento quirúrgico (387713003): Procedimiento que implica alteración intencional no transitoria de las estructuras del cuerpo.[14]

La ECL final que define el subconjunto de concepto para el campo diagnósticos incluyendo diagnósticos estrictos y procedimientos quirúrgicos:

ECL:  $\llcorner$  64572001 | enfermedad (trastorno) | OR  $\llcorner$  387713003 | procedimiento quirúrgico |

#### 4.3.4. Tipos de descripciones en resultados

SNOMED International aconseja, para la implementación de interfaces de usuario en sistemas clínicos, utilizar las descripciones activas del idioma o dialecto utilizado para encontrar el mejor concepto. Una vez encontrado se debe presentar al usuario el

término preferido. De esta forma funciona la búsqueda de Snowstorm en la mayoría de las funciones de la API, y en este caso corresponde a un endpoint de conceptos.

Al seguir esta recomendación los resultados que se muestran al usuario son solo sinónimos preferidos. Estos corresponden a un sinónimo de la descripción completa (FSN) de un concepto que se marca como ‘preferido’, es una palabra o frase utilizada frecuentemente por los médicos para nombrar a ese concepto.

Por ejemplo, para el concepto 254837009 tenemos las siguientes descripciones:

Term	Aceptabilidad (ES)
F ★ neoplasia maligna de mama (trastorno)	Preferido
S ★ tumor maligno de mama	Preferido
S ✓ cáncer de mama	Aceptable
S ✓ neoplasia maligna de mama	Aceptable

Figura 15: Descripciones para el concepto 254837009 (F:FSN, S:Sinonimo)

Cuando el usuario ingresa “cáncer de mama” el concepto se encuentra como un termino sinónimo aceptable y devuelve su termino sinónimo preferido “tumor maligno de mama”. Por lo tanto, en algunos casos el termino ingresado no corresponde a este termino preferido y se debe seleccionar el sinónimo.

Por otro lado, la exploración de conceptos en el buscador de SNOMED International presenta tanto el término coincidente como el FSN. Se utiliza la búsqueda de descripciones, endpoint de descripciones. Podría usarse para obtener las descripciones que coincidan con el termino ingresado, pero no es correcta para una búsqueda en una interfaz de usuario clínica. Además no es muy recomendable por la documentación para búsquedas en la mayoría de los contextos ya que tiene pocos parámetros y filtros. Se destaca que no puede filtrar usando ECL lo que presenta una limitación para lo requerido en el proyecto.

En conclusión, aunque no coincida con el sinónimo buscado se utiliza la búsqueda

por endpoint de conceptos recomendada por la documentación. Ofrece más flexibilidad para realizar búsquedas y aplicar filtros adicionales. Además los términos devueltos suelen contener información completa y contextualizada.

## 4.4. Librería dsutils

### 4.4.1. Integración de sistemas

Es necesario entender que datos se manejan, como se estructuran y como es posible intercambiar su información en cuanto a los servicios externos a utilizar.

Se tiene en cuenta un concepto importante, la reutilización de código, es la práctica de usar códigos existentes para una nueva función o software.

Para reutilizar el código, este debe ser de alta calidad. Y eso significa que debe ser seguro y confiable. Desarrollar un software que cumpla con estos requisitos fue uno de los desafíos. Desarrollar sistemáticamente componentes y marcos de software reutilizables de alta calidad es difícil por lo que reutilizar el código es un gran objetivo.

En un entorno ideal, un desarrollador podría acceder a bibliotecas de código estables y de confianza. Luego podrían reutilizar el código de esas bibliotecas como bloques de construcción dentro de su aplicación.

Bajo este concepto se creó un librería que es fácilmente ampliable y se adaptado a la nueva aplicación, no presenta defectos ni problemas que afecten a la fiabilidad, la seguridad o la protección.

### 4.4.2. Funcionamiento

Se creo una librería de Python privada, de uso exclusivamente corporativo. La cual será consumida desde diferentes microservicios que corresponden a la arquitectura de software de emergencias. La misma se llama “dsutils”. Para garantizar su privacidad se creo como un repositorio de la empresa dentro de Github.

Esta librería es la encargada de facilitar un programa capaz de almacenar el dato

del diagnóstico luego de ser codificado. Por lo tanto, es la encargada de codificar el diagnóstico en formato JSON, en un tipo de recurso valido para FHIR.

Toma como entrada el concepto seleccionado el usuario formado por el código numérico y el termino preferido, o simplemente el texto libre en el caso de no poder ser codificado.

Además, requiere de una serie de metadatos, como la fecha en la que fue codificado el diagnóstico, la fecha en la que se documento, el estado de la evaluacion e informacion del paciente y del profesional de salud involucrado.

Entre las funciones del sistema se encuentran la validación de los datos ingresados según su tipo esperado y la construcción del formato de salida.

Como salida, la API crea la estructura un recurso del tipo ClinicalImpression, y lo almacena como formato JSON en una base de datos. La arquitectura de base de datos que se utiliza para este propósito es relacional, semi estructurada, y basada en la arquitectura propuesta para un servidor de FHIR por HAPI FHIR (Secc:3.7.2).

La Figura 16 contiene un readme, que describe la información que contiene la librería, como instalarla y como usar la misma.

```
pip install git+https://github.com/IHSA-TDG/dsutils

## Usage

... python
import dsutils
from dsutils import *

# Imports necesarios
import dsutils.ehr
from dsutils.ehr.resource import Resource
import dsutils.ehr.clinicalimpression
from dsutils.ehr.clinicalimpression.clinicalimpression import clinicalImpression

# Datos ejemplo
data = {
    "status": "completed",
    "effectiveDateTime": "2022-12-06T22:33:00+11:00",
    "date": "2022-12-06",
    "id_patient_beneficiary": 9,
    "id_provider": 6,
    "snomed_code": 1076621000119101,
    "display": 'tendinitis de rodilla derecha (trastorno)'
}

# Armar la estructura del recurso clinicalimpression con data
recurso = clinicalImpression(data)
recurso

# Crear el recurso clinicalimpression - Verificacion con un codigo
# [201] Created
# [200] OK
# [202] Accepted
# [504] Gateway Timeout
# [505] HTTP Version Not Supported
# [400] Bad Request

recurso.create_resource()

## Contribuciones

Célula de Data Science - Equipo GEN - Emergencias
```

Figura 16: README de la librería dsutils.

#### 4.4.3. Elección de recurso y campos

Previo a la decisión de utilizar el recurso Clinicalimpression, se evaluaron diferentes opciones entre los recursos de FHIR.

Se analizo en profundidad los campos y tipos de datos que conforman a cada recurso de la sección clínica (Sec.3.6.2.2).

Los recursos de la categoría “Diagnostic” dentro de la sección “Clinical”, se utiliza para representar información relacionada con el diagnóstico médico que incluyen informes de pruebas diagnósticas, estudios de imágenes, observaciones clínicas y muestras para análisis. Se ve que no están destinados a ser utilizados para registrar el diagnóstico descrito en la Sección 3.1.3 sobre un paciente o sujeto.

Por consiguiente, se analizó categoría de “Summary” y dentro de la misma se selecciono el recurso ClinicalImpression para ser utilizado en campo diagnostico.

En FHIR, una evaluación es típicamente un instrumento o herramienta utilizada para recopilar información sobre un paciente.

El recurso ClinicalImpression de la categoría Summary es un registro de una evaluación clínica que se realiza para determinar qué problemas pueden afectar al paciente. A partir de esto se planifican los tratamientos o estrategias de manejo que son mejores para manejar la condición de un paciente.

En resumen, ClinicalImpression puede conducir a una declaración de una condición sobre un paciente; hace referencia a una impresión como un resumen clínico de información y/o una opinión formada.

En la Figura 18 se ve la estructura del recurso ClinicalImpression con JSON.

```

{
  "resourceType": "ClinicalImpression",
  // from Resource: id, meta, implicitRules, and language
  // from DomainResource: text, contained, extension, and modifierExtension
  "identifier": [{ Identifier }], // Business identifier
  "status": "<code>", // R! preparation | in-progress | not-done | on-hold | stopped | completed | entered-in-error | unknown
  "statusReason": { CodeableConcept }, // Reason for current status
  "description": "<string>", // Why/how the assessment was performed
  "subject": { Reference(Group|Patient) }, // R! Patient or group assessed
  "encounter": { Reference(Encounter) }, // The Encounter during which this ClinicalImpression was created
  // effective[x]: Time of assessment. One of these 2:
  "effectiveDateTime": "<dateTime>",
  "effectivePeriod": { Period },
  "date": "<dateTime>", // When the assessment was documented
  "performer": { Reference(Practitioner|PractitionerRole) }, // The clinician performing the assessment
  "previous": { Reference(ClinicalImpression) }, // Reference to last assessment
  "problem": [{ Reference(AllergyIntolerance|Condition) }], // Relevant impressions of patient state
  "changePattern": { CodeableConcept }, // Change in the status/pattern of a subject's condition since previously assessed, such as worsening, improving, or no change
  "protocol": ["<uri>"], // Clinical Protocol followed
  "summary": "<string>", // Summary of the assessment
  "finding": [{ // Possible or likely findings and diagnoses
    "item": { CodeableReference(Condition|DocumentReference|Observation) }, // What was found
    "basis": "<string>" // Which investigations support finding
  }],
  "prognosisCodeableConcept": [{ CodeableConcept }], // Estimate of likely outcome
  "prognosisReference": [{ Reference(RiskAssessment) }], // RiskAssessment expressing likely outcome
  "supportingInfo": [{ Reference(Any) }], // Information supporting the clinical impression
  "note": [{ Annotation }], // Comments made about the ClinicalImpression
}

```

Figura 17: Recurso ClinicalImpression en JSON. Imagen extraída de [20]

ClinicalImpression esta publicado en la compilación actual. Su nivel de madurez se encuentra en “FMM 1”, indica que se encuentra sustancialmente completo y listo para su implementación.

Por otra parte, se definieron los campos a utilizar, basados en la justificación y enfoque de la estructura de FHIR necesaria para un diagnóstico.

En la Tabla 1 podemos analizar los campos del recurso ClinicalImpression, con el tipo de dato y una descripción breve de cada uno.

Nombre	Tipo	Descripción y restricción
resourceType	Resource	Recurso dominante
identifier	Identifier	Identificador
status	code	preparation in-progress not-done on-hold  stopped completed entered-in-error unknown
statusReason	CodeableConcept	Motivo del estado actual
description	string	Por qué/cómo se realizó la evaluación
subject	Reference(Paciente   Group)	Paciente o grupo evaluado
encounter	Reference(Encounter)	Encuentro durante el cual se creó esta Impresión Clínica
effectiveDateTime	dateTime	Cuando se evaluó
date	dateTime	Cuándo se documentó la evaluación
performer	Reference(Practitioner   PractitionerRole)	El clínico que realiza la evaluación
previous	Reference(ClinicalImpression)	Referencia a la última evaluación
problem	Reference(Condición   AlergiaIntolerancia)	Impresiones relevantes del estado del paciente
changePattern	CodeableConcept	Cambio en el estado/patrón de un sujeto desde que se evaluó previamente
protocol	uri	Protocolo clínico seguido
summary	string	Resumen de la evaluación
finding	Element BackboneElement	Hallazgos posibles o probables y diagnósticos
prognosisCodeableConcept	CodeableConcept	Estimación del resultado probable
prognosisReference	Reference(RiskAssessment)	Evaluación de riesgos que expresa el resultado probable
supportingInfo	Reference(Any)	Información que respalda la impresión clínica
note	Annotation	Comentarios realizados

Tabla 1: Análisis de campos del recurso ClinicalImpression. Información obtenida de [20]

Para determinar como guardar la descripción provista por el profesional, primero se definieron los campos del recurso a utilizar. Existen tres campos obligatorios definidos por el recurso elegido.

Los otros campos definidos, fueron analizados de acuerdo a los formatos establecidos y las necesidades de Emergencias; y según la justificación y enfoque de la estructura de FHIR necesaria para un diagnóstico. En la siguiente tabla vemos los mismos:

Uno de los principales campos de este recurso es “finding”. Este incluye un campo codificable “codeable concept”, que hace referencia a un código definido por un siste-

Nombre	Tipo
ClinicalImpression	Recurso
status	preparation in-progress not-done on-hold  stopped completed entered-in-error unknown
subject	Paciente o grupo evaluado
effectiveDateTime	Cuando se evaluó
dateTime	Cuando se documento la evaluación
performer	Clínico que realiza la evaluación
finding	Hallazgos posibles o probables y diagnósticos

Tabla 2: Campos seleccionados del recurso ClinicalImpression.

ma terminológico, donde se guarda el diagnóstico junto con su código de SNOMED CT y su correspondiente url.

También incluye un campo que corresponde a un mensaje de texto “text” que representa el significado previsto del usuario y permite texto libre. Se utiliza en el caso de que el diagnóstico no fuera identificado como tal o no se reconozca el texto ingresado.

Considerando esto, creó una la instancia JSON, reemplazando el contenido de los valores según las reglas y la descripción de contenido que se encuentran en el valor para cada elemento.

Se crea una estructura de recurso en JSON como la siguiente:

```
{
  'resourceType': 'ClinicalImpression',
  'status': 'completed',
  'effectiveDateTime': '2022-12-06T22:33:00+11:00',
  'date': '2022-12-06',
  'subject': {
    'reference': 'https://servidor-fhir-deployment.emergencias.com.ar/fhir/Patient?identifier=https://www.emergencias.com%7C9'
  },
  'performer': {
    'reference': 'https://servidor-fhir-deployment.emergencias.com.ar/fhir/Practitioner?identifier=https://www.emergencias.com%7C6'
  },
  'finding': {
    'coding': [
      {
        'system': 'http://snomed.info/sct',
        'code': '1076621000119101',
        'display': 'tendinitis de rodilla derecha (trastorno)'
      }
    ]
  },
  'text': 'tendinitis de rodilla derecha (trastorno)'
}
```

Figura 18: Estructura del recurso ClinicalImpression en JSON.

#### 4.4.4. Control de errores

Para ayudar a que el código fluya y se ejecute sin problemas ni interrupciones se usaron métodos de manejo de errores. Si se producen errores en cualquier línea de código, el control de errores se encarga de ellos y, a continuación, el código reanuda la ejecución.

Schema es una biblioteca para validar estructuras de datos de Python, como las obtenidas de archivos de configuración, formularios, servicios externos o línea de comandos análisis. Si los datos son válidos se continua con normalidad. Por otro lado, si los datos no son válidos, los corrige o genera un mensaje de error.

Se consideran errores los datos que no se corresponden con la especificación de FHIR. En ella se define un conjunto de tipos de datos que se utilizan para los elementos de recursos. En este caso `datetime`, `date`, `string` o `int`.

Existen otros errores incluyen una mala conexión con la API de FHIR que realiza el guardado en HAPI FHIR. También errores de tiempo de espera cuando la API de FHIR o el servidor HAPI FHIR tardan demasiado en responder. Cuando estos ocurren se muestran mensajes de advertencia a los usuarios para ayudar comprender y resolver el error.

## 5. Resultados y métricas

El testeo se realizó sobre los textos ingresado por usuarios de diferentes áreas, definidos por clasificadores existentes de la empresa como ‘Otros diagnósticos’.

### 5.1. Muestras

De los textos existentes para el testeo se procesaron únicamente aquellos que se consideran “codificable”, que son diagnósticos en si. Para esto se realizó un arreglo de los textos existentes.

Se excluyen los textos de lugares, prestaciones, pagos, familias, planes e informa-

ción de postoperatorios. Los correspondientes a antecedentes, secuelas o anomalías también fueron delimitados. Estos textos fueron considerados “no codificables” y no se tuvieron en cuenta para el testeo.

Por otro lado, como el objetivo de la codificación es que a cada texto libre se le asigne un código de SNOMED CT, que represente un solo concepto médico. Se distinguen textos que contienen mas de un concepto médico y que no pueden ser representados según esta regla. Se realiza una previa separación de los términos con el fin de no perder muestras y analizar la mayor cantidad de casos posibles. Se dividieron en textos independientes los diagnósticos múltiples separados por caracteres como el signo mas (+), comas (,), barras (/), guiones(-) o con ‘y’ como conjunción.

Con estas consideraciones el set de datos de testeo final esta compuesto por 2805 diagnósticos codificables.

## 5.2. Uso ECL

Dentro de los 2805 textos definidos para el testeo, se codifican aquellos que se reconozcan y existan en SNOMED CT. Se implementa la ECL definida previamente para restringir las opciones devueltas por la codificación.

El gráfico de la Figura 19 siguiente muestra la comparación entre la codificación con y sin la aplicación de ECL. La cantidad de textos codificados es muy parecida ambos casos: con el uso de ECL se han codificado 2290 textos y sin ECL se han codificado 2317 textos.

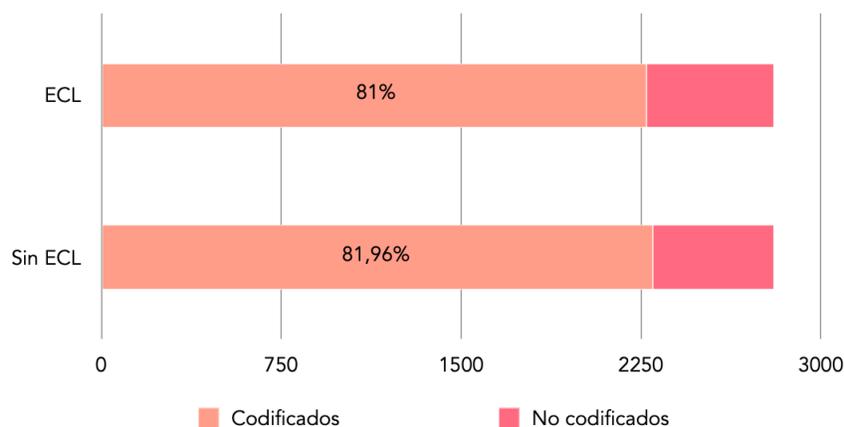


Figura 19: Conceptos codificados. Con ECL el porcentaje de codificación fue del 81% mientras que sin ECL del 81,96%

La principal diferencia que se observa al utilizar restricciones de lenguaje se refleja en los resultados que devuelve la terminología. Con el término “trauma” se distingue notablemente esto en su búsqueda sin ECL obteniendo los conceptos de la Figura 22. En ella se distinguen conceptos que hacen referencia a “trastornos”, “hallazgos”, “calificadores”, “escala de evaluación”, “régimen/tratamiento”, “ocupación”, “medio ambiente” o “fuerza física”. Al aplicar la ECL se limitan estas opciones a las categorizadas como “trastornos” o como “procedimientos”, evitando la selección de términos que corresponden a otros dominios.

- "lesión traumática (trastorno)"
- "traumático (calificador)"
- "riesgo de lesión traumática (hallazgo)"
- "traumatólogo (ocupación)"
- "lesión traumática contusa (trastorno)"
- "índice de trauma (escala de evaluación)"
- "puntaje de traumatismo (escala de evaluación)"
- "traumatismo fetal (trastorno)"
- "traumatismo obstétrico en el feto (trastorno)"
- "traumatismo de la vía aérea (trastorno)"
- "traumatismo dental (trastorno)"
- "traumatismo directo (calificador)"
- "centro de traumatología (medio ambiente)"
- "trauma coclear (trastorno)"
- "lesión traumática de glúteo (trastorno)"
- "tratamiento de trauma psicológico (régimen/tratamiento)"
- "cirujano especialista en traumatología (ocupación)"
- "lesión traumática orbitaria (trastorno)"
- "lesión traumática del hígado (trastorno)"
- "choque traumático (trastorno)"
- "agente traumático (fuerza física)"
- "traumatismo oclusal (trastorno)"
- "trauma cerebral (trastorno)"
- "lesiones múltiples (trastorno)"
- "ampolla de causa traumática (trastorno)"
- "evento traumático (evento)"
- "úlceras traumáticas (trastorno)"
- "lesiones traumáticas múltiples (trastorno)"
- "apnea traumática (hallazgo)"

Figura 20: Resultado de conceptos devueltos por SNOMED para el termino “trauma” sin aplicar restricciones.

Los términos que son codificados por la terminología pero que al aplicar la ECL no se muestran se analizaron en detalle. Se vio que algunos se definían por SNOMED como “hallazgos” como en el caso de ‘dislalia’, ‘trastorno deglutorio’ o ‘hemianopsia’ y por lo tanto entran en la restricción aplicada. Otros eran procedimientos como ‘toilette quirúrgico’ que no se clasificaban dentro del subnivel “procedimientos quirúrgicos” requerido.

En otros casos no se encontraba el texto como tal en la terminología, pero si formando parte de textos pertenecientes a otro subniveles. Por ejemplo los resultados

para el texto ‘cáncer prostático’ son ‘sospecha de cáncer prostático (situación)’, ‘clasificación de Walter Reed del cáncer prostático (estadificación tumoral)’ o ‘sistema de puntaje de Gleason para cáncer prostático (escala de estadificación)’

En resumen, el uso de ECL es beneficioso para limitar las opciones devueltas por la codificación, mejorar la precisión y relevancia de los resultados y evitar conceptos no deseados. Adapta la codificación al contexto y requisitos del campo diagnóstico obteniendo resultados más alineados.

Por otro lado, para comparar la cantidad de resultados de búsqueda de un mismo término con ECL y sin ECL se realizó un gráfico de dispersión.

El gráfico de la Figura 21 permite observar la cantidad de resultados en la búsqueda de cada uno de los términos para la búsqueda con ECL (eje x) y sin ECL (eje y).

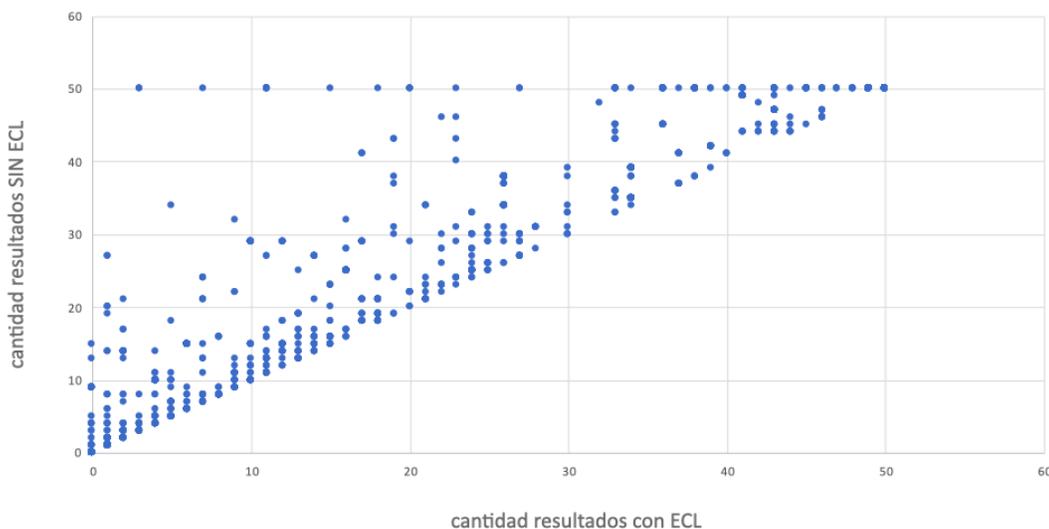


Figura 21: Gráfico de dispersión de la cantidad de resultados de búsqueda con y sin restricciones de lenguaje.

Los casos en los que para un mismo diagnóstico se observó la misma cantidad de resultados con y sin ECL, se ven en los puntos del gráfico cercanos a una línea diagonal que va desde el origen hasta el punto máximo de ambos ejes. Los puntos que

se encuentran en la parte superior de la diagonal del gráfico indican una variabilidad en la relación, Mostrando una mayor cantidad de resultados en la búsqueda sin ECL.

Se observa positivamente que en caso de que se encuentren resultados correctos aplicando ECL se brindan menos opciones al usuario, mejorando el tiempo que se consume para encontrar la opción correcta.

### 5.2.1. Resultados de búsquedas

Se tomó una muestra aleatoria de 360 términos, de los datos codificados utilizando la ECL para hacer inferencias más precisas; analizando en detalle y con distintos enfoques los conceptos devueltos. No se incluyen conceptos iguales en este análisis.

El texto ingresado y representado de forma idéntica o casi textual en los conceptos devueltos, se denominó “resultado literal”. Algunos ejemplos se ven en la siguiente figura:

Texto ingresado	Resultados de búsqueda
bronquiolitis	bronquiolitis
irc	insuficiencia renal cronica
fractura de columna	fractura de columna vertebral
infección en sitio qx	infección de sitio quirúrgico

Tabla 3: Ejemplos de términos denominados como “resultado literal”.

Los resultados literales que se encuentran en SNOMED CT como términos aceptables pero que en los resultados se muestra el sinónimo preferencial también se consideran resultados textuales. Se destaca que estos corresponden al 2,7% de la muestra aleatoria tomada.

Texto ingresado	Resultados de búsqueda
cancer de pulmón	tumor maligno de pulmón
hallux valgus	deformidad en valgo del primer dedo del pie
síndrome CHARGE	síndrome de coloboma, malformación cardíaca, atresia de coanas, retraso de crecimiento y desarrollo, trastornos genitourinarios y malformaciones auriculares

Tabla 4: Ejemplos de textos cuyo termino aceptable es denominados como “resultado literal” y muestra al usuario el termino preferencial.

En los textos que devolvían opciones que contenían el termino ingresado pero con una descripción mas general o de manera mas detallada se analizo la “especificidad”. En estos casos no se representaba de forma textual al texto ingresado por el exceso o falta de especificaciones.

Texto ingresado	Resultados de búsqueda
nefrostomía izquierda	nefrostomía
artritis de pie	- artritis de ambos pies - artritis pie derecho - artritis de pie izquierdo

Tabla 5: Ejemplos de términos denominados como “Especificidad”.

En otras ocasiones las abreviaturas de diagnósticos no eran reconocidas por la terminología como tal o por el diccionario implementado. También se consideran aquellas abreviaciones mal escritas que no son corregidas por el diccionario de ortografía debido a su escasa longitud de caracteres. En estos casos, SNOMED CT puede devolver conceptos que contienen los mismos caracteres de la abreviatura, pero no se corresponden con el significado buscado.

Se presentan algunos ejemplos en la Figura 6 del texto “aco” que podría corresponder a “acv” pero que no fue corregido , lo cual genera resultados que no se corresponden con lo buscado. Asimismo, en el caso de “drt”, que podría ser una abreviatura de ”depresión resistente al tratamiento”, no se reconoce correctamente. Estos casos demuestran cómo las abreviaciones mal escritas o no reconocidas por el diccionario pueden llevar a interpretaciones incorrectas o resultados no deseados.

Texto ingresado	Resultados de búsqueda
aco	- acondroplasia - útero acervical - acondrogénesis
drt	- acidosis tubular renal con sordera nerviosa progresiva

Tabla 6: Ejemplos de abreviaturas de diagnósticos no reconocidas correctamente.

Cuando el texto era reconocido pero las opciones de conceptos representaban términos muy diferentes al buscado se marcaron como “Otros”. Se mostraban opciones de diagnósticos pero sin relación al texto ingresado como se ve en la Figura 7.

Texto ingresado	Resultados de búsqueda
ulceras vasculares	livedo reticularis idiopática con ulceraciones de verano
hemorragia de vías digestivas altas	hemorragia del tracto gastrointestinal superior concomitante con gastritis y debida a ella
disfunción valvular	displasia valvular mixomatosa ligada al cromosoma X asociada a filamina A

Tabla 7: Ejemplos de términos denominados como “Otros”.

Teniendo en cuenta estos enfoques se obtuvieron las siguientes métricas:

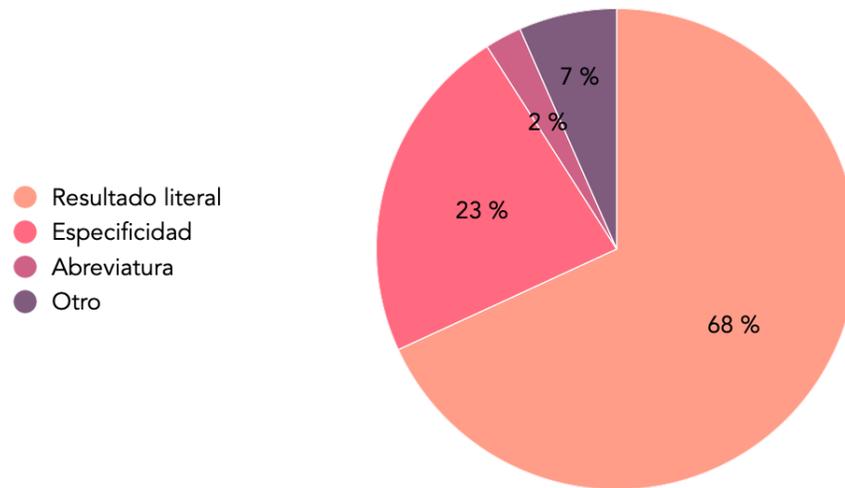


Figura 22: Resultados de búsquedas.

Se logró un porcentaje del 68 % de reconocimiento literal, es decir, las opciones devueltas coincidieron exactamente con el término ingresado. Esto demuestra la capacidad del sistema para comprender y responder de manera precisa a los textos de entrada. Además, tiene un beneficio significativo, ya que permite a los usuarios obtener resultados directos y relevantes sin la necesidad de realizar ajustes o modificaciones en sus consultas.

Sin embargo, en un 24 % de los casos, las opciones presentaban una falta o exceso de especificidad, lo que no representaba de forma textual al texto original. También se identificó un 2 % de casos en los que las abreviaturas de diagnóstico no fueron reconocidas correctamente. Estas métricas destacan la importancia de mejorar la especificidad y el reconocimiento de abreviaturas para proporcionar opciones relevantes y precisas al usuario.

Además, hubo un 7 % de casos en los que las opciones de diagnóstico no guardaban relación con el texto ingresado y se clasificaron como “Otros”.

#### 5.2.1.1. Especificidad

Se observó que un 23 % de textos no presentaban resultados literales debido a la

especificaciones presentadas por SNOMED CT dentro de los términos delimitados por la ECL. Se realizó un análisis en profundidad de estos textos, representado por la Figura 23.

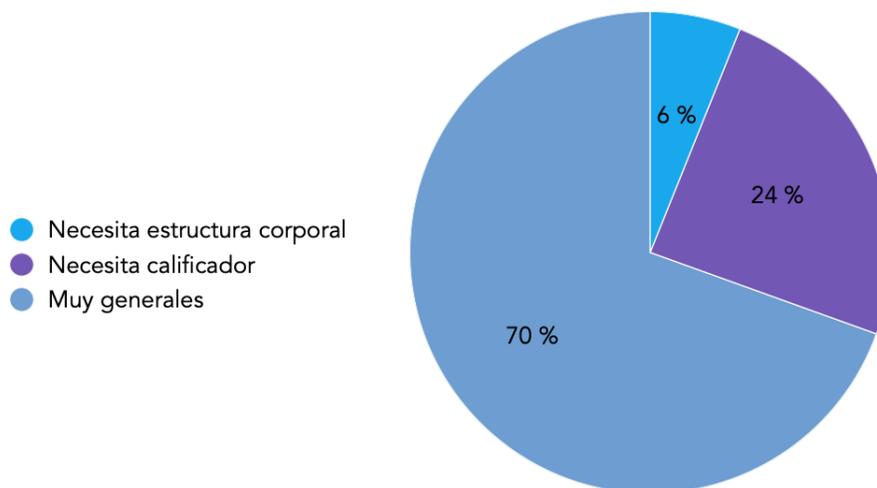


Figura 23: Tipos de especificaciones.

Un 70% de estos textos carecían de detalles, siendo muy generales para ser representados textualmente. En este caso se presenta al usuario términos con más detalle. En un principio, esto puede ser positivo ya que obliga al usuario a seleccionar un diagnóstico que proporcione información más clara y precisa desarrollando un plan de tratamiento más específico y efectivo. Sin embargo la información extra de los conceptos puede no ser adecuada a lo buscado y se debería dejar el concepto sin codificar.

El resto se debía al caso contrario, la terminología devolvía términos muy generales cuando se buscaban conceptos con descripciones que proporcionen más información. Un 24% requería de una calificación adicional como la gravedad (aguda o severa), grado de lesión (1,2,3), nivel de importancia (principal o secundario), espacio (interno o externo), entre otros. Mientras que el 6% se requería información del lugar o de la estructura anatómica del diagnóstico que no era especificada.

Ejemplos de estos casos los podemos ver en la Figura 8.

Epecificidad	Texto ingresado	Resultados de búsqueda
Necesita estructura corporal	hematoma subdural	hematoma
Necesita calificación	nefrostomía izquierda	nefrostomía
Muy generales	artritis de pie	- artritis de ambos pies - artritis pie derecho - artritis de pie izquierdo

Tabla 8: Ejemplos de especificaciones necesarias.

En base a este análisis se destaca la carencia de la especificidad necesaria para representar adecuadamente la información buscada así también como términos con resultados demasiado generales. Ambos casos limitan la representación de los conceptos buscados.

Como conclusión, se ve la necesidad de equilibrar la especificidad y la generalidad de los términos utilizados en SNOMED CT considerando la implementación de expresiones postcoordinadas. Considerando las necesidades del usuario para obtener resultados más adecuados.

El problema de obtener resultados de diagnósticos demasiado generales se puede solucionar de manera más sencilla mediante la interacción entre el profesional de la salud y el sistema. Al proporcionar información precisa, el profesional ayuda a refinar los resultados y encontrar un término adecuado para el diagnóstico buscado.

#### 5.2.1.2. Resultados literales

Para los resultados textuales se pudo analizar la búsqueda de un concepto entre todas las opciones.

Como el objetivo era encontrar un concepto determinado entre los conceptos de una lista, se utilizó una métrica de posición absoluta.

La posición absoluta es útil para evaluar la eficacia de la búsqueda sin tener en cuenta la importancia relativa del término en comparación con otros términos de la lista. Sirve para indicar qué tan bien se realizó la búsqueda de un término.

Una posición absoluta baja en los resultados define que los conceptos se encuentran entre las primeras posiciones. Si el término aparece cerca del principio de la lista, indica que la búsqueda fue efectiva y se encontró el término de manera eficiente. En estos caso la experiencia de usuario es óptima y se minimizan los tiempos de ingreso del campo.

Los resultados de esta métrica se representan en el histograma de la Figura 24, el eje vertical se indica el numero de posición y en el horizontal la frecuencia de términos encontrados.

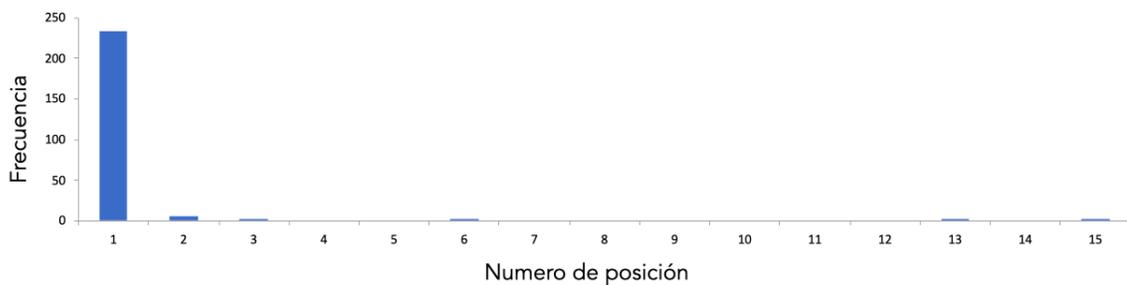


Figura 24: Histograma de posición de resultados “literales”.

Un 94% de términos se encuentran entre la primera y un 5% en la segunda posición de la lista. En algunos casos muy particulares aparecen en otras posiciones. El concepto buscado de “injerto de pie” aparece en la posición 16, ya que las primeras posiciones se encuentran conceptos relacionados con de “injerto de piel”.

Se ve que los primeros dos resultados son los más relevantes para esta búsqueda.

### 5.2.2. Relevancia Resultados

Se puede considerar que existe cierto grado de redundancia cuando un subconjunto significativo de términos representa una proporción sustancial del conjunto de datos.

En este caso, el 47,3% de los términos se representan en un subconjunto formado por 9 términos:

- Fracturas
- Procedimientos de reemplazo (cadera o rodilla)
- Cáncer
- Artritis o artrosis
- Diabetes
- Infecciones (urinarias o de sitio quirúrgico)
- EPOC
- Neumonía
- Bronquitis o bronquiolitis

Este porcentaje indica una concentración significativa de información en esos términos específicos y sugiere una presencia considerable de términos redundantes.

Estos términos tienen una alta relevancia y son de gran importancia para comprender y analizar el conjunto de datos. El gráfico de la Figura 25 proporciona una visión más profunda y detallada de los aspectos destacados en el campo de diagnóstico y tratamiento de los mismos.

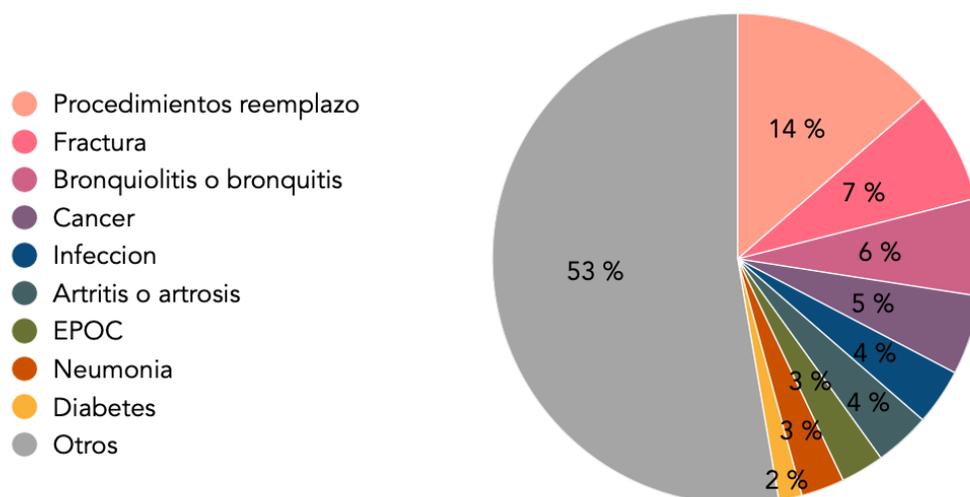


Figura 25: Términos codificables. Cantidad de términos relevantes.

Es importante asegurar la codificación de estos términos frecuentes de manera precisa y consistente. Ya que si estos términos son codificados, podemos asegurar que casi la mitad de los textos van a estar cubiertos.

## 6. Discusión

### 6.1. Codificación

El 81 % de los diagnósticos definidos como “codificables” fueron representados correctamente por SNOMED CT. Destacando que los resultados se correspondían a la definición de diagnósticos, debido a la restricción de lenguaje aplicada.

Se obtuvo una eficiente búsqueda donde los conceptos más relevantes eran presentados en la primera y segunda posición. Generando mayor visibilidad del término y facilitando la selección del usuario. La implementación realizada contribuye a una experiencia de usuario positiva y satisfactoria. En un 68 % de las búsquedas se logra una representación casi literal.

Es importante destacar que, aunque el sistema logró un alto porcentaje en las

métricas de codificación, considerando las de “Resultados de búsqueda” (Secc. 5.2.1) hubo un 24 % de casos en los que las opciones proporcionadas presentaban falta o exceso de especificidad y un 2 % de casos en los que las abreviaturas de diagnóstico no fueron reconocidas correctamente. Esto mostró la necesidad de mejorar la especificidad y el reconocimiento de abreviaturas en el sistema. Al hacerlo, se proporcionarían opciones más relevantes y precisas al usuario.

Mejorar la especificidad garantiza que las opciones devueltas se ajusten de manera más precisa a las necesidades del usuario, evitando opciones demasiado generales o detalladas. Del mismo modo, un mejor reconocimiento de abreviaturas asegura que se comprenda correctamente el significado detrás de esas abreviaturas, evitando opciones que no estén relacionadas con el diagnóstico buscado.

La capacitación de los profesionales de salud en el uso adecuado del sistema de codificación puede contribuir a mejorar estas métricas. Esta puede enfocarse en el uso consistente del sistema, la variación del nivel de especificidad y la eliminación de abreviaturas. Junto a la retroalimentación inmediata y el seguimiento de resultados, se permiten ajustar y mejorar la forma en que se ingresan los diagnósticos, logrando una mayor precisión en la representación de los mismos.

Por otro lado, el porcentaje restante de textos no reconocidos incluye principalmente:

- Términos que están correctamente escritos y definidos pero que no se encuentran definidos por SNOMED. Términos que están incluidos en SNOMED pero que contienen detalles o descripciones no reconocidas o filtradas (63 %).
- Términos que contienen más de un concepto y que no fueron separados para el testeó (14 %).
- Términos que no corresponden al campo diagnóstico (8 %).
- Términos mal escritos (3 %).
- Siglas y descripciones de diagnósticos son de uso poco habitual o específico de

alguna especialidad; así como siglas de menos de tres caracteres no reconocidas (12 %).

### 6.1.1. Comparación de resultados

Teniendo en cuenta la codificación de conceptos clínicos utilizando SNOMED CT, que este proyecto logró un porcentaje del 81 %. Al comparar estos resultados con los obtenidos en otros trabajos [2],[35], se observa que uno de ellos obtuvo porcentaje ligeramente inferior, con un 79 % de precisión en la codificación de conceptos clínicos en su implementación de servicios terminológicos en una red de atención ambulatoria. El otro reportó un porcentaje de codificación del 85 % al utilizar SNOMED CT para la codificación de una lista de problemas.

Este proyecto se encuentra en un rango medio de los porcentajes de codificación, es importante en cuenta que las implementaciones se centran en un contextos específicos y puede haber diferencias en la metodología, características del data sets utilizados y otros factores que puedan influir en los resultados.

Pero lo que si se puede analizar es a SNOMED CT como terminología en la codificación clínica. Se ve que SNOMED CT demostró su eficacia y utilidad en diferentes contextos y aplicaciones clínicas. Estos resultados respaldan la idea de que SNOMED CT es una opción adecuada para la representación de conceptos clínicos en diversos sistemas de información de salud.

## 6.2. Procesamiento del lenguaje natural

El procesamiento de lenguaje natural presenta limitaciones en su aplicabilidad a dominios acotados y áreas específicas. En este proyecto, específicamente enfocado en la atención domiciliaria y la kinesiología, es importante reconocer que los resultados y conclusiones obtenidos pueden estar restringidos a este contexto particular. Por lo tanto, no se pueden generalizar directamente a otros dominios o áreas de la salud. Esto implica que los modelos y algoritmos de procesamiento de lenguaje natural desarrollados y evaluados en este proyecto podrían no ser adecuados para

su aplicación en diferentes entornos o especialidades médicas.

Por otro lado, el corrector ortográfico implementado basado en un diccionario predefinido, no es capaz de corregir las abreviaturas escritas de manera incorrecta o los errores ortográficos múltiples. Estos errores pueden tener un impacto significativo en la calidad y precisión de los resultados del procesamiento de lenguaje natural.

Cuando se encuentran abreviaturas mal escritas o palabras con más de un error ortográfico, se dificulta la correcta identificación y procesamiento de los términos. Esto puede llevar a interpretaciones erróneas o inexactas de los datos, lo que afecta la fiabilidad de los resultados obtenidos. Por lo tanto, es importante abordar esta limitación implementando técnicas de corrección ortográfica y estableciendo pautas claras para la introducción de datos, con el fin de mitigar estos problemas y mejorar la calidad de los resultados del procesamiento de lenguaje natural.

### **6.2.1. Diccionarios**

Es importante resaltar las ventajas significativas de implementar un diccionario términos de adaptabilidad, facilidad de actualización y mantenimiento.

Una de las fortalezas del diccionario radica en su capacidad de adaptarse a diversas situaciones y contextos específicos. Puede ser ajustado y personalizado para satisfacer las necesidades particulares de los usuarios. Esto permite optimizar su rendimiento y precisión según los requisitos específicos de cada aplicación o dominio de conocimiento.

Además, la facilidad de actualización es un atributo valioso del diccionario. Al ser dinámico y estar en constante crecimiento, puede expandirse y actualizarse para incluir nuevos términos, variaciones léxicas o cambios en el uso de palabras existentes. Esto garantiza que el diccionario se mantenga actualizado y relevante, adaptándose a los avances y evoluciones del lenguaje, así como a las necesidades de los usuarios.

Es importante tener en cuenta las limitaciones del diccionario. La ambigüedad en el reconocimiento de textos representa una limitación significativa al implementar

un diccionario con un único significado por palabra. Esta ambigüedad surge cuando una palabra tiene múltiples significados, lo que dificulta determinar el significado exacto en un contexto específico. Existen casos que pueden llevar a interpretaciones erróneas y confusiones en la comprensión del texto.

Otra limitación es la adaptación según la palabra previa en el texto. En algunos casos, es necesario considerar el género, número o relación con la palabra anterior para interpretar correctamente el significado de una palabra en un contexto dado. Si el diccionario no tiene la capacidad de realizar estas adaptaciones, pueden producirse errores en la comprensión del texto, lo que afecta la precisión del reconocimiento y codificación de textos.

Para evitar esto en el diccionario implementado en el caso de variaciones de género, se toma una forma abreviada aceptada que sigue manteniendo la base de la palabra y conserva su significado reconocible pero que no distingue en masculino o femenino. Sin embargo, es importante tener en cuenta que el uso de abreviaciones puede no ser adecuado en todos los contextos y considera la abreviación podría generar confusiones.

Por último, se destaca que la interfaz ayudaría en caso de que el término mostrado no sea el correcto, permitiendo que el texto ingresado se mantenga. Esta funcionalidad permite al usuario decidir si el término presentado es el correcto.

### **6.3. Postcordinación**

Teniendo en cuenta que en el proyecto realizado solo se utiliza la precoordinación por su alcance definido, se vio en la Sección 5.2.1.1 la dificultad de codificar conceptos más abarcativos y términos detallados.

Durante el proceso de codificación, se encontró que algunos conceptos y términos que presentaban mayor nivel de detalle o abarcaban un espectro más amplio de información no eran representados de manera precisa. Esto puede resultar en una pérdida de información importante o en una representación menos precisa de la

complejidad de los datos de entrada.

En este contexto se analiza en cuenta la herramienta de gramática composicional de SNOMED CT que brinda una definición de conocimientos clínicos a través de más de un concepto utilizando una postcoordinación.

La postcoordinación permite combinar términos de manera más flexible y detallada, lo cual resulta especialmente relevante en el contexto de la kinesiología y la atención domiciliaria, donde los casos clínicos pueden presentar situaciones complejas y variables.

La falta de postcoordinación puede llevar a una pérdida de sutileza y especificidad en la descripción de los casos clínicos estudiados. Al no utilizar esta técnica, se limita la capacidad de establecer relaciones más precisas y capturar los matices y particularidades de cada situación clínica. Esto puede tener un impacto en la interpretación de los resultados obtenidos y dificultar la obtención de conclusiones sólidas.

Para abordar esta limitación y mejorar la calidad y nivel de detalle en futuros proyectos, es recomendable considerar la incorporación de la postcoordinación de conceptos. Al hacerlo, se permitiría una descripción más completa y precisa de los casos clínicos, teniendo en cuenta su complejidad y variabilidad. Esto facilitaría la comprensión y el análisis de los resultados, así como la identificación de patrones o tendencias en los datos recopilados.

#### **6.4. Participación de profesionales**

La adopción exitosa del sistema en el contexto de la atención domiciliaria y la kinesiología tiene importantes implicaciones para los profesionales involucrados. La participación activa y la aceptación de los profesionales de la salud son fundamentales para lograr una implementación efectiva del sistema. Sin embargo, también se reconocen las posibles barreras para la adopción, como la resistencia al cambio, la falta de familiaridad con la tecnología y las preocupaciones sobre la privacidad de

los datos.

Para abordar estas barreras, es recomendable implementar programas de capacitación adecuados que proporcionen a los profesionales la formación necesaria para comprender y utilizar eficazmente el sistema. Estos programas deben abordar tanto los aspectos técnicos como los beneficios clínicos del sistema, con el objetivo de que los profesionales se sientan cómodos y confiados al utilizarlo en su práctica diaria.

Asimismo, una comunicación clara y efectiva es esencial para fomentar una mayor adopción y aceptación del sistema. Los profesionales de la salud deben recibir información clara sobre los beneficios del sistema, cómo se integrará en su práctica y cómo se abordarán las preocupaciones de privacidad y seguridad. Además, es importante que los puedan brindar retroalimentación y participar en un diálogo continuo con los desarrolladores y proveedores del sistema, de modo que puedan compartir sus experiencias y contribuir a mejoras futuras.

En cuanto a las ventajas de la adopción del sistema, se destaca una posible mejora en la eficiencia, ya que contribuye para automatizar tareas rutinarias y reducir la carga administrativa de los profesionales, permitiéndoles dedicar más tiempo a la atención directa de los pacientes. Además, el acceso a conocimientos actualizados y evidencia basada en la práctica a través del sistema ayudaría a mejorar la precisión y la calidad de la atención, así como la personalización de los tratamientos.

Otra ventaja destacada es la posibilidad de una mayor colaboración y continuidad de la atención, ya que el sistema facilita la comunicación y el intercambio de datos entre los profesionales de la salud que trabajan en la atención domiciliaria.

Por todo esto, la participación y aceptación de los profesionales de la salud son esenciales para lograr una implementación efectiva del sistema en la atención domiciliaria y la kinesiología. Su compromiso, disposición a aprender y adaptarse, y su capacidad para comprender y comunicar los beneficios del sistema son factores determinantes en el éxito de la adopción y en la mejora de la calidad de la atención proporcionada.

## 6.5. Aspectos legales y normativos

En este proyecto, si bien no se llevó a cabo un estudio exhaustivo de los aspectos legales y normativos, se realizó un análisis preliminar basado en trabajos en áreas similares.[36]

En este análisis, se identificaron diversas consideraciones legales y normativas que deberían ser abordadas para garantizar el cumplimiento de la ley en una aplicación de salud. Dado que este proyecto se va a implementar en una aplicación de este tipo, se podría ampliar y profundizar el análisis de los aspectos legales y normativos mencionados a continuación.

Entre las áreas identificadas, se destaca la importancia de proteger los datos personales sensibles, de acuerdo con las leyes de protección de datos aplicables, como la Ley Nacional 25.326 y la Ley de CABA 1.845. Además, se debe tener en cuenta la Ley de CABA 5.669, que establece la obligatoriedad de disponer de una Historia Clínica Electrónica y regula su uso en situaciones de emergencia.

También obtener el consentimiento informado de los pacientes para recopilar y utilizar sus datos médicos, y garantizar la confidencialidad y seguridad de dicha información, siguiendo las disposiciones del secreto profesional y las leyes de protección de datos. Además, se deben implementar medidas de seguridad adecuadas como accesos restringidos y claves de acceso, para asegurar la integridad de los datos, según lo establecido en la Ley de CABA 5.669.

Se deben tener en cuenta los derechos del paciente, reconocidos por la Ley Nacional 26.529 y la Ley 153 de CABA, que incluyen la titularidad de su historia clínica, la voluntad y confidencialidad, y la garantía de su intimidad.

Finalmente, es esencial considerar la interoperabilidad de los datos de la Historia Clínica Electrónica, de acuerdo con el artículo 27 de la Ley 5.669.

## 7. Mejoras a futuro

La inclusión de diccionarios de siglas, abreviaturas y términos médicos; y ampliación el diccionario de ortografía mejoran el reconocimiento de textos pero presenta limitaciones (Secc. 6.2) . Existen muchas otras formas en que se pueden mejorar como incluir aprendizaje automático o uso de modelos de redes neuronales que permitan a los sistemas de reconocimiento de texto aprender y adaptarse mejor a partir de grandes cantidades de datos. Otra forma puede ser la integración de mas técnicas de procesamiento del lenguaje natural, que permitiría comprender mejor el contexto y el significado del texto que se está buscando.

Además se podría agregar un campo que contenga detalles del diagnostico para agregar especificaciones adicionales. Esto utilizando la postcoordinación de conceptos de SNOMED CT que combina conceptos y permite agregar más detalle al significado representados por un solo concepto. Por ejemplo, incorporando conceptos definidos como ‘calificador’ para los complementar a los textos que requieren mayor especificación, principalmente para las calificaciones como gravedad de la lesión, grado, nivel de importancia y espacio. Para se esto se debe evaluar las reglas de atributos y valores correspondientes.

Esta postcoordinación también podría ser aplicada y tener un gran aporte para representar los conceptos que no están presentes de forma exacta en SNOMED CT.

Una posible propuesta a futuro del codificador podría ser agregar funciones de sugerencias. Al de ingresar un texto que no sea un diagnóstico se podría advertir al usuario y sugerir el campo correspondiente del mismo. De esta forma seria posible mejorar la experiencia del usuario ayudando al usuario a definir el campo correcto.

Por ultimo, la adopción exitosa del sistema de procesamiento de lenguaje natural depende de la participación activa y la aceptación de los profesionales de la salud. Por esto se recomienda implementar programas de capacitación, medidas de seguridad de datos y una comunicación efectiva.

## 8. Conclusiones

Se desarrollo un sistema capaz de cumplir con la codificación de diagnósticos y su posterior guardado. Se creó una herramienta de acuerdo a los estándares de interoperabilidad permitiendo la integración con otros sistemas. A lo largo del proyecto, desarrollado en Emergencias desde octubre de 2022 y junio de 2023, se cumplen los objetivos para ser implementado en la empresa otorgando un impacto positivo.

El proyecto esta principalmente formado por una API que brinda la terminología de SNOMED CT y una librería privada que estructura los datos siguiendo especificaciones de HL7 FHIR. Ambos sistemas permiten actualizaciones en el tiempo permitiendo la capacidad de adaptarse a las necesidades de cada momento.

La interoperabilidad es clave para una atención médica eficaz ya que mejora la calidad de la atención, reduce los costos y mejora la prestación de servicios, lo que lleva a una atención con mejores resultados. Por lo tanto, la interoperabilidad debe ser una prioridad en áreas donde se requiere el intercambio de información.

La implementación del servicio de terminología de SNOMED CT otorga una amplia y gran variedad de conceptos, y demostró una buena codificación de diagnósticos. Sin embargo, se destaca que la codificación es una estrategia con muchas limitaciones relacionadas al procesamiento del lenguaje. La mejora de esto depende en gran medida de los ajustes de normalización y de los diccionarios que contienen sinónimos, abreviaturas y jerga local.

El otro punto fue la especificación FHIR que proporcionó una excelente estructura dada por el recurso ClinicalImpression. Esto facilitó el modelado de los datos y permitió la adaptación del recurso según la información necesaria y las especificaciones institucionales. Se tiene en cuenta que este estándar se encuentra en continuo desarrollo por lo que la información obtenida en la actualidad puede verse modificada en un futuro debido a las actualizaciones de versiones. Es importante trabajar con la versión implementada.

Con respecto a los resultados, en el testeo se logró codificar el 81 % de los textos que no eran reconocidos por los sistemas utilizados actualmente. Se contribuye a una experiencia de usuario positiva y satisfactoria.

En próximos pasos queda una evaluación en un entorno productivo para asegurar que el sistema esté preparado y sea capaz de cumplir con las expectativas y requisitos establecidos antes de su implementación a gran escala.

Finalmente, el desarrollo de este proyecto implica la adquisición de conocimiento para poder seguir desarrollando en un futuro otros campos y culminar en una historia clínica interoperable integral.

## A. Anexo

Diccionario	
acv	accidente cerebrovascular
bq	bronquitis
bql	bronquiolitis
brq	bronquitis
brql	bronquiolitis
ca	cancer
cardiorrenal	cardio renal
celular	de células
coxidinia	coxiálgia
crva	cirugía de reemplazo valvular aórtico
ctrol	control
dbt	diabetes
dcha	derecha
dcp	desproporción cefalopélvica
demente	demencia
der	derecha
dif	dificultad
dl	dolor
eap	edema agudo de pulmón
enf	enfermedad
exploradora	exploratoria
fqp	fibrosis quística pulmonar
fx	fractura
genitoplastia	cirugía para reasignación de género
gonalgia	dolor de rodilla
hqx	herida quirúrgica
hsa	hemorragia subaracnoidea
hta	hipertensión arterial
icc	insuficiencia cardíaca crónica
insf	insuficiencia
insuf	insuficiencia
intraencefalica	encefálico
ippb	infección de piel y partes blandas
ira	infección respiratoria aguda
irc	insuficiencia renal crónica
itu	infección del tracto urinario
izq	izquierda
lumboartrosis	artrosis lumbar
luxofractura	fractura luxación
lx	luxación
meniscopatía	desgarro meniscal
miles	resección abdominoperineal
mmii	miembros inferiores
mts	metástasis
nac	neumonía adquirida en la comunidad
osteosintesis	fijación
peritoneal	del peritoneo

poli	múltiple
politrauma	múltiple trauma
politraumatismo	múltiple trauma
precoz	prematura
protésica	prótesis
puerperales	postparto
qx	quirúrgico
rabdomio sarcoma	rabdomiosarcoma
reconstrucción	reparación
rt	reemplazo total
rtc	reemplazo total de cadera
rtr	reemplazo total de rodilla
saf	síndrome antifosfolipídico
sahos	síndrome apnea obstructiva del sueño
sde	síndrome
sdme	síndrome
smd	síndrome
sme	síndrome
sng	sonda nasogástrica
tea	trastorno del espectro autista
tep	tromboembolia pulmonar
tgd	trastorno generalizado del desarrollo
tqt	traqueotomía
trast	trastorno
tto	tratamiento
tvp	trombosis venosa profunda

## Bibliografía

- [1] Grupo IHSA – *Expertos en soluciones integrales de salud y bienestar*, es, <https://ihsa.com.ar/>, Accessed: 2023-7-13.
- [2] C. T. Casanelli, H. Navas, S. Benitez et al., “Implementación de servicios terminológicos en una red de atención ambulatoria,” 2008.
- [3] W. Panko, J. Silverstein y T. Lincoln, “Technologies for extracting full value from the electronic patient record,” en *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, IEEE, 1999, 9-pp.
- [4] G. B. de Quirós Fernán, S. Enrique, L. Daniel et al., “Desarrollo e implementación de una Historia Clínica Electrónica de Internación en un Hospital de alta complejidad,” *1ras. Jornadas de Sistemas de Información en Salud. Hospital Italiano de Buenos Aires*, 2003.
- [5] J. A. F. Mota y J. N. Fariña, “LA INFORMACIÓN CLÍNICA Y DE SALUD,” *Informes SEIS*, pág. 77,
- [6] J. C. Wyatt y P. Wright, “Design should help use of patients’ data,” *The Lancet*, vol. 352, n.º 9137, págs. 1375-1378, 1998.
- [7] R. Dick y E. Steen, *Institute of Medicine, The Computer Based Patient Record: An Essential Technology for Health Care*, Institute of Medicine, 1991.
- [8] W. i. E. H.-D. Techopedia, *Error Handling*, en, <https://www.techopedia.com/definition/16626/error-handling>, Accessed: 2023-3-1, mayo de 2017.
- [9] F. Campos, *Principios de interoperabilidad en salud y estándares (Spanish Edition)*, 2018.
- [10] *Hospital Italiano de Buenos Aires*, <https://www.hospitalitaliano.org.ar>, Accessed: 2023-3-31.
- [11] *¿Qué es Snomed CT?* es, <https://www.argentina.gob.ar/salud/terminologia/snomed-ct>, Accessed: 2022-11-22, jul. de 2021.
- [12] S. Vergara, *Terminología estándar en salud: SNOMED CT como vocabulario clínico estructurado*, es, <https://www.itdo.com/blog/terminologia-estandar-en-salud-snomed-ct-como-vocabulario-clinico-estructurado/>, Accessed: 2023-7-14, mar. de 2022.

- [13] *INICIO*, en, <https://www.snomed.org/?lang=es>, Accessed: 2023-7-14.
- [14] D. Markwell, *SNOMED CT Starter Guide*, <https://confluence.ihtsdotools.org/display/DOCSTART/SNOMED+CT+Starter+Guide+>, Accessed: 2022-11-20.
- [15] SNOMED International, *SNOMED International Browser*, <https://browser.ihtsdotools.org/?perspective=full&conceptId1=404684003&edition=MAIN/2023-03-31&release=&languages=en>, Accessed: 2023-4-9.
- [16] L. Bird, *SNOMED CT expression constraint language - SNOMED CT languages project group - SNOMED confluence*, <https://confluence.ihtsdotools.org/display/SLPG/SNOMED+CT+Expression+Constraint+Language>, Accessed: 2022-11-23.
- [17] I. Berges, J. Bermudez y A. Illarramendi, “Binding SNOMED CT terms to archetype elements,” *Methods of information in medicine*, vol. 54, n.º 01, págs. 45-49, 2015.
- [18] A. Muñoz, A. Romero Gutiérrez, G. Marco Cuenca et al., “Manual práctico de interoperabilidad semántica para entornos sanitarios basada en arquetipos,” 2013.
- [19] *Health Level Seven International - Homepage*, <https://www.hl7.org/>, Accessed: 2022-11-25.
- [20] *Modules - FHIR v4.3.0*, en, <https://www.hl7.org/fhir/modules.html>, Accessed: 2022-11-25.
- [21] Amanda, *Intercambio de datos clínicos con HL7 FHIR e integración con Red Hat Fuse (Apache Camel)*, es, <https://www.chakray.com/es/utilizar-hl7-fhir-intercambio-informacion-clinica-integracion-apache-camel/>, Accessed: 2023-1-19, sep. de 2021.
- [22] *Scikit-Learn, herramienta básica para el Data Science en Python*, es, <https://www.master-data-scientist.com/scikit-learn-data-science/>, Accessed: 2023-7-13.
- [23] *Welcome to flask — flask documentation (2.3.X)*, en, <https://flask.palletsprojects.com/en/2.3.x/>, Accessed: 2023-6-22.
- [24] *What is flask python*, en, <https://pythonbasics.org/what-is-flask-python/>, Accessed: 2022-12-6.

- [25] Y. K. R. Sánchez, S. A. Demurjian y M. S. Baihan, “Achieving RBAC on RESTful APIs for mobile apps using FHIR,” en *2017 5th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (Mobile-Cloud)*, IEEE, 2017, págs. 139-144.
- [26] M. A. Hussain, S. G. Langer y M. Kohli, “Learning HL7 FHIR using the HAPI FHIR server and its use in medical imaging with the SIIM dataset,” en *Journal of digital imaging*, vol. 31, n.º 3, págs. 334-340, 2018, ISSN: 0897-1889. DOI: 10.1007/s10278-018-0090-y. dirección: <http://dx.doi.org/10.1007/s10278-018-0090-y>.
- [27] *HAPI FHIR modules - HAPI FHIR documentation*, en [https://hapifhir.io/hapi-fhir/docs/getting\\_started/modules.html](https://hapifhir.io/hapi-fhir/docs/getting_started/modules.html), Accessed: 2022-12-13.
- [28] *API REST: qué es y cuáles son sus ventajas en el desarrollo de proyectos*, es, <https://www.bbvaapimarket.com/es/mundo-api/api-rest-que-es-y-cuales-son-sus-ventajas-en-el-desarrollo-de-proyectos/>, Accessed: 2022-12-6, mar. de 2016.
- [29] *JSON*, <https://www.json.org/json-es.html>, Accessed: 2022-12-6.
- [30] Y. Fernández, *API: qué es y para qué sirve*, es, <https://www.xataka.com/basics/api-que-sirve>, Accessed: 2023-1-19, ago. de 2019.
- [31] *Microservicios*, es, <https://aws.amazon.com/es/microservices/>, Accessed: 2023-7-13.
- [32] *La biblioteca estándar de Python*, es, <https://docs.python.org/es/3/library/>, Accessed: 2023-1-13.
- [33] *Cómo crear librerías en Python*, es, <https://tutorialesinformatica.com/programacion/como-crear-librerias-en-python/>, Accessed: 2023-1-16, mar. de 2022.
- [34] R. P. M. de HCE, *SNOMED CT - Restricciones de búsqueda (ECL)*, [https://www.argentina.gob.ar/sites/default/files/2021/07/propuesta\\_ecl\\_hce\\_ips\\_otros\\_-\\_julio\\_2021.pdf](https://www.argentina.gob.ar/sites/default/files/2021/07/propuesta_ecl_hce_ips_otros_-_julio_2021.pdf), Accessed: 2023-7-13.
- [35] L. Osornio, S. Montenegro, G. Marti et al., “Codificación múltiple de una lista de problemas utilizando la CIAP-2, CIE-10 y SNOMED CT,” *INFORMEDICA 2004*, 2004.

- 
- [36] C. Carengo y M. d. l. M. Scalese, “Diseño de E-Health App para acceso a datos clínicos relevantes de pacientes sin facultades de responder el interrogatorio médico mediante validación de su identidad con reconocimiento facial,” 2020.