HECON: Health Condition Evolution Ontology

Alba Catalina Morales Tirado[1[0000−0001−6984−5122]], Enrico Daga[1[0000−0002−3184−5407]], and Enrico Motta[1[0000−0003−0015−1952]]

KMi, The Open University, Milton Keynes, United Kingdom
{alba.morales-tirado,enrico.daga, enrico.motta}@open.ac.uk

Abstract. Health records contain extensive information, including conditions, test results, procedures, and appointments. These data reveal past medical observations that could allow drawing a picture of the current health state of a patient. Widely adopted clinical terminology taxonomies, such as SNOMED CT and the FHIR standard, facilitate the processing and exchange of electronic health records. However, despite these efforts, the task of estimating whether a particular condition is affecting a patient's health at a certain point in time is not supported yet.
This paper introduces HECON, the Health Condition Evolution Ontology that represents the evolution of health conditions over time. This representation enables reasoning on possible ongoing health issues derivable from patients' health records for the benefit of intelligent systems in the emergency domain. We describe the process for building the ontology and the application of HECON in a fire emergency scenario. We design the ontology following established ontology engineering practices, including Competency Questions and ontology reuse. Furthermore, we construct a Knowledge Graph from a database of extracted Health Evolution Statements and use it to validate the consistency and requirements of the HECON Ontology.

Keywords: Health events · Ontology · Knowledge Graph · Emergency Support · SNOMED CT.

1 Introduction

In recent years, rapid urbanisation and a growing population have increased the demand for better public services [2]. The implementation of intelligent systems that make use of healthcare data [18] provides opportunities to enhance healthcare service access, research, patient care and to improve emergency response operations. Widely adopted standards, such as FHIR[1], and terminological systems, such as SNOMED CT[2], facilitate the exchange and representation of Electronic Health Records (EHRs) and medical information [7,9,1,16], opening up new opportunities to use health-related data. For instance, the National Health

[1] FHIR - Fast Healthcare Interoperability Resources
[2] SNOMED CT - Clinical Terminology of clinical specialities and requirements.

Service (NHS) in the UK has adopted SNOMED CT as a standard vocabulary for recording health records [16], and FHIR to improve the exchange of information [15].

Electronic Health Records (EHRs) contain patients' exhaustive health information (for example, appointments, laboratory tests, surgeries, allergies), which can help draw a picture of the person's state of health at a given point in time. For instance, the analysis of standardised medical data (such as, EHRs using SNOMED CT terminology) could reveal chronic or ongoing issues, recent procedures or other health-related events that can impact on the current health condition[3] of a patient. Delivering this information to first responders represents an essential asset when making decisions and planning evacuation operations [13]. However, the lack of knowledge about clinical conditions' evolution over time represents a challenge when implementing intelligent systems that can automatically estimate a person's current state of health

This paper addresses these challenges by adopting a knowledge engineering approach. We present HECON[4], the Health Condition Evolution Ontology, as a formal model representing the evolution of health events over time. In particular, the process of recovering from a health situation is not limited to a fixed convalescence time. For example, conditions typically have a temporal span associated with their evolution, e.g., improving or worsening over a certain period or, alternatively, becoming chronic. Our ontology aims to support the identification of ongoing health events at a given point in time by representing conditions' evolution information. We link the health evolution data to the SNOMED CT taxonomy and extend it by aggregating information of the recovery time of conditions. We motivate the design rationale of the ontology by considering a fire emergency scenario as our reference application, which provides both a source for requirements and a validation setting. HECON is the result of previous research [12,13,14] to represent and reason on healthcare data and provide usable information to emergency services in the context of fire event.

This paper expands the model proposed in [13] by formally encoding its concepts using Web Ontology Language (OWL). Additionally, we use HECON as a framework to create a Knowledge Graph using the health evolution database built in [13].

Our main contributions are:

- A model for representing and reasoning about the evolution of health events over time: HECON - Health Condition Evolution Ontology.
- A Knowledge Graph that defines an abstraction of the available data about health evolution. Moreover, it also includes information that extends the clinical concepts descriptions provided by SNOMED CT taxonomy.
- A set of SPARQL queries for validation and guidance to users of both the ontology and the Knowledge Graph.

---

[3] FHIR terminology defines condition as *'A clinical condition, problem, diagnosis, or other event, situation, issue, or clinical concept that has risen to a level of concern.'* see *http://hl7.org/fhir/condition.html*

[4] HECON Ontology repository https://github.com/albamoralest/HECON-Ontology and BioPortal https://bioportal.bioontology.org/ontologies/HECON

The rest of the paper is organised as follows. Section 2, describes the background and motivating scenario. In Section 3, we review the literature on the use of healthcare ontologies, and health evolution databases. Section 4 details the methodology used to build the ontology and Section 5 states the knowledge requirements that HECON should address. Section 6 describes the development process of HECON. In Section 7 we present the Knowledge Graph and the evaluation of the ontology. Finally, in Section 8, we summarise the conclusions and describe future directions of the research.

2 Background and motivation

First responders attending an emergency should have an overview of the situation before performing any operation. One of their first tasks is to retrieve information about the type of event, the people involved and details of the premises. During a fire emergency, additional information such as people's health conditions or disabilities can assist first responders in making decisions for rescue operations [17]. This information usually is retrieved from people on the scene or fire wardens.

We examine the scenario of a fire event in a large organisation in the UK. In a large organisation, an Access Control System (ACS) holds information of people entering the building as they use their magnetic cards to gain access, while visitors register (and obtain their cards) at the reception desk. The organisation adheres to general guidelines and procedures in case of an emergency; for example, the UK Government regulated guidance for emergency preparedness [22,23]. In relation to fire emergencies, the guidelines indicate that a person should notify any disability or health issue that may affect them during an evacuation to their line manager and discuss the elaboration of a Personal Emergency Evacuation Plan (PEEP), together with the Health & Safety Unit [22]. The PEEP is an elaborated plan that thoroughly explains the type of support a person should have to leave the premises, and it is tailored according to the person's capabilities. Crucially, the PEEP form collects employees' personal information about their health conditions and disabilities. However, several issues arise from the stated scenario:

- Although the procedures state the PEEP should be updated regularly, employees' most recent health events might not be recorded immediately, particularly in large organisations. As a result, emergency responders could question the completeness and accuracy of the information.
- Employees might choose not to disclose sensitive information or consider their health issues irrelevant, which leads to incomplete PEEP records. A clear example is an 'obstructive bronchitis' condition. Although it is not a permanent disability, typical symptoms include difficulty breathing and walking. If the issue is recent, symptoms are relatively acute and may lead to a vulnerable situation.
- Although general evacuation plans are put in place for visitors, it may be possible that guests do not provide sufficient information about their state of health; hence information might be incomplete.

Typically, if a fire starts, the Health & Safety Unit should recover the PEEPs to assist the emergency responders in identifying people in need of special assistance. However, a Smart City perspective could open opportunities to streamline this process. First, when the fire starts, the building's Access Control System (ACS) could identify the people in the building. Second, an intelligent system capable of automatically retrieving the information from the ACS and access health records from the National Health Service (NHS) could automatically analyse and extract updated information regarding people's ongoing health issues. Since health records hold granular information of a person's health condition, a challenging area of study is extracting helpful information that reports on a person's current medical status [12,13].

Hence, the analysis of health conditions is not a straightforward process. In [13] we introduced the concept of *Health Evolution Statement (HES)* to solve the problem of representing and reasoning over the person's state of health at a given point in time. The HES is an elaborated representation of the health recovery process, which describes two main characteristics of health evolution: the way it develops and its estimated duration. Specifically, the HES representation can express whether a condition improves or declines, how long it takes for a person to recover, and whether the condition is chronic. Crucially, a system using this model will automatically detect ongoing health events at the time of the emergency. To the best of our knowledge, no existing ontology is tailored to the representation of health evolution over time. The same applies to existing, reusable knowledge bases in healthcare.

3 Related Work

In this section, we revise related work in health evolution representation using ontologies, SNOMED CT applications and access to databases related to healthcare convalescence time.

Ontologies are largely used in applications in the Semantic Web field and are widely adopted in Knowledge Engineering in various domains. Uschold and Grüninguer [24] proposed the first guidelines for building ontologies. Furthermore, methodological frameworks like [10,19] also cover aspects of the ontology development process, life cycle, methods and techniques. For instance, the NeOn Methodology [21] proposes a more flexible approach with particular attention to 'ontology engineering by reuse' and collaborative development. Our methodology follows these established frameworks and adapts the different steps from each methodology according to our system and knowledge requirements in the emergency response domain. For instance, we follow [5,24] to identify the purpose of the ontology, the motivating scenario and to capture knowledge requirements expressed as Competency Questions. Moreover, we follow best practices of reusing established ontologies to minimise the addition of new ontology terms [21].

In the context of emergency support, well-structured information, readily available and pertinent to first responders' needs [17] is a valuable asset and a life-saving resource during an emergency event. The adoption of Electronic Health Records (EHRs) and related standards (SNOMED CT and FHIR) for collecting and exchanging healthcare data has leveraged the development of intelligent systems and ontologies for emergency support [4] that provide information fit for emergency services requirements. Research to date confirms that most knowledge-based applications are oriented to solving issues related to the semantic integration of heterogeneous healthcare data, particularly in crisis management scenarios [11]. Although these proposed systems focus on supporting emergency services using healthcare-related data, no applications have included the representation of knowledge about health events' evolution over time in the context of emergency response. This paper proposes an ontology as a formal representation of health evolution capable of identifying and reasoning upon conditions that could affect a patient's condition at the time of an emergency.

With respect to EHRs data handling and exchange, an important aspect is the standardised representation of clinical terminology. SNOMED CT, the Systematised Nomenclature of Medicine Clinical Terms is one of the most used schemes. It comprises medical terms such as disorders, procedures, body structures and other clinical findings. Other schemes such as ICD-10 (International Classification of Diseases) and LOINC (Logical Observation Identifiers, Names, and Codes) [1] focus on terminology for reporting and monitoring diseases and clinical measurement terms, respectively. However, these taxonomies do not include definitions or information about health evolution, revealing the need for reliable and structured information that describes health issues' convalescence time and recovery. As part of our research, we represent the database of health evolution built previously in [13] as a Knowledge Graph following a Linked Data approach. Furthermore, we expand the specification of SNOMED CT by directly using SNOMED URIs in the KG.

In conclusion, the analysis of the literature indicates that, despite a substantial body of research related to the use of healthcare data, there remains an evident opportunity to develop additional resources for the benefit of emergency teams tackling emergency events.

4 Methodology

In Section 2 we described the challenges found to extract relevant information from health records specifically to support emergency responders. Our efforts focus on providing an intelligent system with a model that allows the automatic evaluation of people's health conditions. In what follows, we describe the methodology used to develop the Health Condition Evolution Ontology (HECON), which is our proposed model for representing the evolution of health events over time and facilitating a system's reasoning on health records. We follow ontology engineering good practices and methodologies [24,19,21], and devise the following phases:

Abstracting the scenario. In this phase, we identify the scope and purpose of our ontology. The main goal of our ontology is to provide an Intelligent System with enough knowledge allowing the detection of ongoing health issues using health records as data source. We use the scenario described in Section 2 as a guide for identifying key ontology concepts and relationships that define 'health evolution' terms. Additionally, we identify data sources of condition evolution, for example, descriptions of health conditions.

Identify knowledge requirements and formulate Competency Questions (CQs). The second phase is dedicated to identifying the knowledge requirements of a system that accesses health records to assess a person's state of health. We express the requirements in the form of Competency Questions (CQs). The CQs will guide the ontology development process, providing the main framework to evaluate the expressiveness of the ontology and the completeness of the related knowledge graph (its ability to answer the questions on specific health records).

Ontology design and construction. This phase consists of two tasks: a) describe the concepts and relations that constitute the ontology, and b) use the vocabulary compiled in a) to express the ontology in a formal language. We build the ontology using Protégé ontology editor and specify its terms with the Web Ontology Language (OWL). This phase also contemplates the addition of the axioms expressing the relationships and constraints among the ontology entities. Similarly, during this phase, we identify the established ontologies for reuse and define how to integrate these concepts in HECON.

Knowledge Graph population The third phase is dedicated to building the Knowledge Graph (KG). First, we use the data sources identified in the first phase to build a structured database of Health Evolution Statements (HES) as detailed in our previous study in [13]. We use SPARQL Anything [3], to generate RDF data from HES database, store our KG and evaluate the CQs in the next phase.

Ontology validation. The last phase concerns to the evaluation of the ontology. We consider two aspects: the consistency of the ontology and the fulfilment of the CQs. We use the Hermit reasoner [20] provided in Protégé to check ontology consistency with respect to the semantics of OWL. In this phase we use the KG built in the previous phase and a dataset of Synthetic Health records [25]. We translate the CQs into SPARQL queries and use them to interrogate the ontology. We compare the results from the query execution with the expected values. CQ fulfilment indicates that the ontology satisfies the system requirements. In case of inconsistencies, we repeat the process to clarify the SPARQL queries or fix the ontology.

5 Requirements

In this section, we extract the knowledge requirements from the perspective of a system that handles the scenario proposed in Section 2.

In our scenario, we described a fire emergency in a large organisation. In the context of a Smart City, an intelligent system can leverage the organisation's Access Control System (ACS) to determine the people (employees and visitors) in the building at the moment of the emergency. The ACS provides the list of people in the building, and therefore, the healthcare provider can retrieve their electronic health records. The public health provider uses SNOMED

CT and FHIR as standards for accessing and exchanging EHRs. The intelligent system uses people's healthcare data to identify up-to-date health issues. The main question that our system should answer using the ontology is: *What is the current 'state of health' of a person at a certain point in time?* Building around the described scenario and this central question, we express the knowledge requirements as a set of Competency Questions (CQs) that our proposed ontology should answer. The CQs are expressed as follows:

- CQ1: What is the health evolution information of a given SNOMED concept?
- CQ2: How does a given condition evolve over time?
- CQ3: What is the pace at which a given SNOMED concept evolve?
- CQ4: What are the expected minimum and maximum recovery times for a given SNOMED concept?
- CQ5: If an emergency happens after the expected maximum recovery time, is a condition still ongoing?
- CQ6: If an emergency happens before the expected minimum recovery time, is a condition still ongoing?
- CQ7: If an emergency happens between the expected minimum and maximum recovery time, is a condition still ongoing?

The data used to build the database of health evolution is retrieved from different sources, identified in phase one of our methodology. These include two authoritative and reliable data sources: NHS England[5] and MAYO Clinic[6]. Consequently, we consider it essential to represent this knowledge by including the associated source and explaining the context on how information was generated. We express these requirements as follows:

- CQ8: What method generates specific health evolution information of a given SNOMED CT concept (e.g., user study, automatic knowledge extraction)?
- CQ9: What is the source of the health evolution information (e.g., authoritative sources, domain experts)?
- CQ10: What/Who is the organisation/person providing this information?
- CQ11: Is there additional information indicating the information's quality?

Although the example in this paper relates to a fire emergency, the ontology is generic and can support other types of emergencies, leaving the assessment of how the condition has to be handled to the emergency support system relying on it (see [13]). Handling privacy requirements is not part of the ontology, however our approach assures that EHR shared are adequate, relevant and limited to the intended use (data minimisation GDPR principle), these aspects are discussed in [12]. In addition, we consider relevant to describe the design constraints derived from the scope of our system and best ontology engineering practices:

- Reuse established ontologies when possible, minimising the addition of new ontology terms. In our case, to represent the time duration, we utilise the Time Ontology [6] and PROV Ontology [8] for provenance information.

[5] https://www.nhs.uk/conditions/

[6] https://www.mayoclinic.org/diseases-conditions

– Reuse standards in the health domain. We link the condition evolution information to SNOMED CT, a comprehensive collection of healthcare terminology. We also use the FHIR standard to represent the information of our synthetic dataset and facilitate the management of health-related data.

6 HECON's Ontology

In what follows, we describe the core concepts of the HECON Ontology and the integration of SNOMED CT clinical terms. Next, we explain how we incorporate Time [6] and PROV-O [8] Ontologies and their role in our model.

6.1 HECON core concepts

We define a number of ontological concepts to represent health evolution (Fig. 1). First, we define an **Health Evolution Statement** (HES) as an abstraction of the components that indicate the recovery process [13]; its main elements are: the types of health evolution, the velocity at which it changes, and its duration. We identify four main types of health evolution:

– **Improvement** - the evolution is favourable and indicates recovery of good health.
– **Decline** - the evolution is adverse and gradually becomes worse during time.
– **Permanent** - indicates a persistent condition.
– **Unaffected** - describes a health condition or SNOMED CT concepts with no effect on state of health, for example, blood test or administrative procedures.

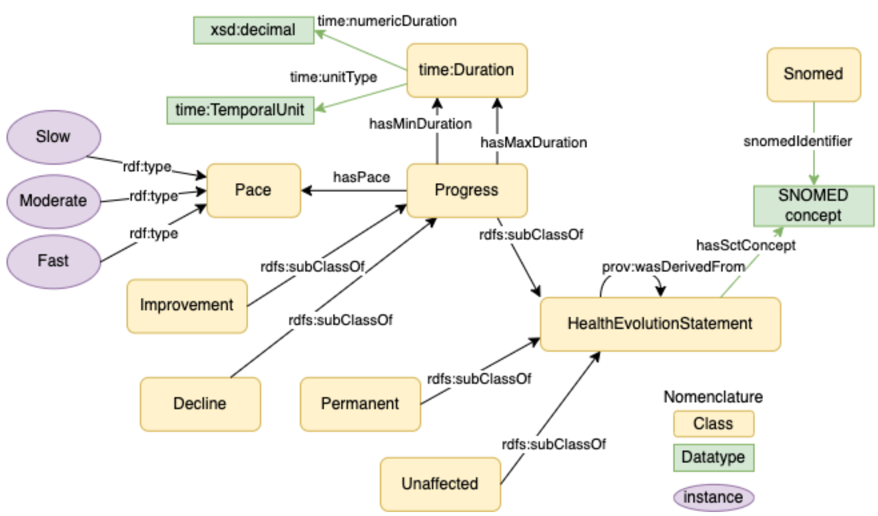

Fig. 1: HECON Ontology core concepts.

We use **Progress** to represent type of events that evolve over time. These types of events, **Improvement** and **Decline**, have pace and an estimate of their duration. For instance, a 'Fracture of ankle' event gets better after a period of time, its evolution type is 'Improvement'. The velocity at which a health event evolves is represented by **Pace** and could take values such as *FAST, MODERATE* and *SLOW*. The estimated duration is defined by *has minimum and maximum duration*. These two boundaries (min and max duration) express time extent, and could take any numerical value in *minutes, hours, days, weeks, months or years*; we represent these elements with **time:Duration** and properties **time:numericDuration**, **time:unitType** from the Time Ontology (see Table 1 for more examples). Health events are represented in health records using a standard terminology system; in our case, SNOMED CT scheme. Therefore, each expression of health evolution is linked (*has a SNOMED CT concept identifier*) to its corresponding term in the SNOMED CT taxonomy. We extend SNOMED CT by directly using SNOMED namespace URI ⟨**http://snomed.info/id/**⟩. For example, a health event such as 'Fracture of ankle' is represented as follows: http://snomed.info/id/16114001.

A system using HECON can predict whether a health condition is still ongoing by calculating the time that has passed between the date it was recorded (in a health record) and its estimated recovery date (the minimum and maximum duration). For example, an ankle fracture (see Table 1) that has not reached its minimum duration is more likely to impact a person's health condition than if the same event happened two months ago (max. duration). The type of health evolution also influences the impact on patient's state of health; a condition that improves is different from one that is permanent, as described in [13].

Table 1: Example of Health Evolution Statement for different conditions.

Health event		**Description**	**Health Evolution Statement (HES)**			
Snomed Id	Name	(*excerpts taken from web sources)	Type	Pace	minD	maxD
363732003	Addison's disease	*'Addison's disease symptoms usually develop slowly, often over several months.'*	Decline	Slow	6 months	1 year
16114001	Fracture of ankle	*'A broken ankle usually takes 6 to 8 weeks to heal, but it can take longer.'*	Improvement	Moderate	6 weeks	2 months
792004	Jakob Creutzfeldt disease	*'There's currently no cure for CJD, so treatment aims to relieve symptoms.'*	Permanent			

6.2 Provenance representation

Part of the requirements state the necessity to provide context on how the Health Evolution Statement (HES) was generated, the data sources used and the organisations or persons supporting this information (Fig. 2).

We use PROV-O basic classes and properties to represent the origin of each HES. A SNOMED CT concept can be linked to none or many Health Evolution Statements. Each HES is a subclass of **prov:Entity** that was generated (**prov:wasGeneratedBy**) using one or many approaches or activities - **prov:Activity**. We describe three possible activities:

Knowledge extraction - the health evolution data was built using knowledge acquisition techniques for extracting information from unstructured data sources.

Therefore, every HES generated have data properties **confidence** and **support**. For example, the HES ‘Improvement Moderate from 8 days to 2 months’ for Fracture of ankle (sct:16114001) was generated by the ’KnowledgeExtraction’ activity, and its support value is ’0.0016’.

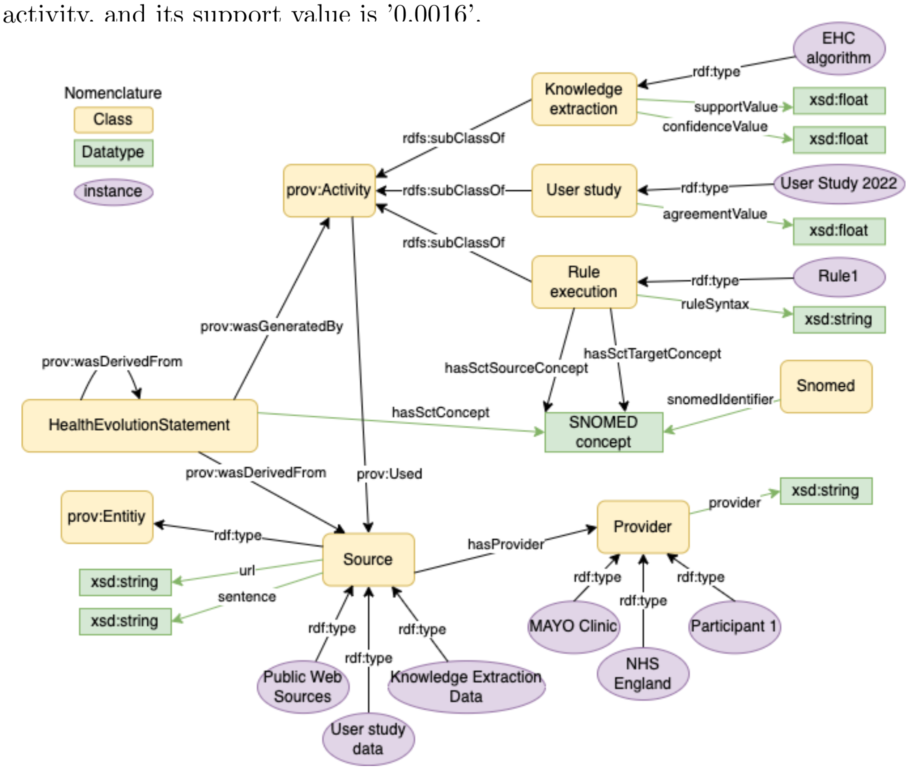

Fig. 2: HECON Ontology provenance representation.

Rule execution - the health evolution data was generated as a result of a knowledge completion process (as described in detail in [13]). We apply a set of rules (based on the relationships and attributes in SNOMED CT taxonomy) that indicate a ’target’ SNOMED CT concept that can inherit the same HES as another SNOMED CT ’source’ concept that already has a HES. The relation between **RuleExecution** and SNOMEC CT concept is expressed as *hasSctTargetConcept* and *hasSctSourceConcept* respectively. We complement information about this activity with the data property *ruleSyntax*. For instance, the HES ’Improvement Slow from 2 months to 6 months’ for **Bacterial sinusitis** (sct:703470001) was generated by the ’RuleExecution’ activity. It has as source concept ’Sinusitis’ (sct:36971009) and target concept Bacterial sinusitis (sct:703470001).

User study - in this case, the HES is the result of a manual evaluation that checks the accuracy of the current database or collects new HESs. This type of activity is carried out with the support of domain experts (e.g., doctors, nurses, emergency responders). This activity also provides data on the agreement between compiled responses in terms of an *agreeementValue*.

An activity, for example, Knowledge Extraction, can use one or multiple **Source**s. A source (**prov:Entity**) at the same time *has a* **Provider**, this could be an organisation or a person, for example: NHS England, or a domain expert. Some sources optionally include details of public URL or exact text that generated the HES, expressed with data properties **url** and **sentence**. Following the previous example of 'Fracture of ankle', we can express that it was derived from 'Public Web Sources' and its provider is 'MAYO Clinic'. It is possible to represent all the possible Sources used by a certain activity with the property *prov:Used*.

7 Knowledge Graph

In this section, we give a brief description of the data source collection, and then we describe the process followed to build the Knowledge Graph.

First, as data input, we use the database of Health Condition Evolution (HES) we built in the context of our previous study [13]. The data was collected from reliable public sources (MAYO Clinic and NHS England). We relied on knowledge acquisition techniques such as Machine Learning (ML) and built a semi-automatic supervised classification pipeline to extract information from unstructured data. The database is a collection of Health Condition Statements; each statement is a structured representation of how a health event evolves. For example, the expression: *'...it takes 2 to 3 weeks to recover completely.'* is represented as *Improvement Moderate from 2 weeks to 3 months*. Each HES is linked to a health condition represented as a SNOMED CT concept identifier. A condition, now formally a SNOMED CT concept, could be connected to one or more HES. Additionally, the database contains information about the sources used to generate each HES. The database of HES is the result of the various knowledge acquisition pipelines (processes described in [13]), the output is a set of CSV files which we use to generate the KG.

Second, we used SPARQL Anything [3], a system that allows to generate RDF data from different types of formats with plain SPARQL CONSTRUCT queries[7]. Next, we store RDF data in TTL data format. Finally, we use the Blazegraph database to store our KG and evaluate the CQs by executing SPARQL queries on the data.

7.1 Data statistics

The resulting Knowledge Graph contains 5,446,510 triples and 33,828 unique SNOMED CT concepts. A SNOMED CT concept is linked to one or more HESs using the corresponding SNOMED CT identifier. Table 2 presents the summary statistics for each OWL class. The number of SNOMED CT concepts represented in our KG is approximately 10% of the total of concepts in the SNOMED CT taxonomy. We are currently working on a user study, with domain experts advice, to validate the current knowledge acquisition pipeline.

[7] All the queries and results are available in the HECON repository

Table 2: Core. Number of instances per class.

	HES	Improve	Decline	Permanent	Unaffected	time:Duration
Total Entities	355,535	230,943	44,331	73,818	6,443	8

In terms of provenance information, the activities that generate the HESs are: the *Health Evolution Extraction Algorithm* (a Knowledge Extraction activity) and the *Knowledge Completion Process* (a Rule Execution activity). A User Study activity will be added in the future. A total of 1,264 SNOMED CT concepts have at least one HES generated by the Health Evolution Extraction Algorithm activity, whereas 96,245 SNOMED CT concepts have at least one HES generated by the Knowledge Completion Process activity (Rule Execution). Table 3 presents the summary statistics for each class. Additionally, two organisations are represented: NHS England and MAYO Clinic.

Table 3: Provenance. Number of instances per class.

	Activity	Source	Knwl. Ext	Rule Exc.
Total Entities	358,323	2,576	2,805	355,518

7.2 Evaluation

Our objective is to formally represent how health events evolve over time and, crucially, meet the knowledge requirements of an intelligent system that automatically detects ongoing health issues at the time of an emergency, for example, a fire event. We built the HECON Ontology following the list of knowledge requirements collected in Section 5.

We use Protégé and the Hermit reasoner to check the formal consistency of the ontology. The final KG represents a large amount of data, therefore, to make the evaluation practicable, we based the consistency check on a random data sample. In order to evaluate the expressivity of the ontology and the completeness (fitness for use) of the related KG for our task, we encode the Competency Questions (CQs) listed in Section 5 into SPARQL queries. We use the same dataset of patients' synthetic health records used in [13] and compare the outcome of the queries against the expected results. The dataset is described using FHIR and SNOMED CT standards, emulating UK national information standards requirements and therefore we make sure our approach is compliant with real implementations. In what follows, we group closely related CQs to give a description of the evaluation process and the results.

Competency Questions one to four (CQ1 to CQ4). These competency questions inquire about the duration, pace and type of change characterising the evolution of a condition. As described in Section 6 the Health Evolution Statement is a compilation of three pieces of information: type of event (classes: Improvement, Decline, Permanent or Unaffected), pace (Slow, Moderate or Fast) and time range (minimum and maximum duration). By retrieving the type of event (CQ2), the pace (CQ3) and the expected recovery time (CQ4), we can satisfy CQ1, as shown in Fig.3. For example, if we find in the health records an event such as

'Broken leg', the answer will be a list of HES linked to the associated SNOMED CT concept, as shown in Fig. 4. These results will be used by an intelligent system to assess the validity of the given condition.

```
SELECT DISTINCT ?snmdIdentifier ?typeEvent ?pace ?maxDuration ?minDuration
FROM <http://conrad.kmi.open.ac.uk/hecon/kg/HES>
WHERE {
  VALUES (?snmdIdentifier) { (sct:46866001)  }# leg fracture
  ?s rdf:type ?typeEvent ;
     hecon:hasSctConcept ?snmdIdentifier ;
     prov:wasGeneratedBy ?anActivity_IRI .
  ?anActivity_IRI rdf:type hecon:KnowledgeExtraction .
  FILTER ( ?typeEvent != hecon:HealthEvolutionStatement) .
  OPTIONAL { ?anActivity_IRI hecon:confidenceValue ?confidenceVal } .
  OPTIONAL {?s hecon:hasPace ?pace ;
       hecon:hasMaxDuration ?maxDuration ;
       hecon:hasMinDuration ?minDuration ; }
} ORDER BY ?snmdIdentifier DESC ( ?confidenceVal )
```

Fig. 3: Competency Questions CQ1 to CQ4 coded as a SPARQL query.

snmdIdentifier	typeEvent	pace	maxDuration	minDuration
sct:46866001	hecon:Improvement	hecon:MODERATE	kg:1/YEAR	kg:6/MONTH
sct:46866001	hecon:Improvement	hecon:SLOW	kg:6/MONTH	kg:2/MONTH
sct:46866001	hecon:Improvement	hecon:MODERATE	kg:6/MONTH	kg:2/MONTH
sct:46866001	hecon:Improvement	hecon:MODERATE	kg:2/MONTH	kg:8/DAY

Fig. 4: Query outcome for CQs:1-4. SNOMED CT cnpt 46866001 Leg Fracture.

Competency Questions five to seven (CQ5 to CQ7). For CQs five to seven, we encode a SPARQL query that retrieves the maximum and minimum duration of a condition in a health record (as shown in Fig.5a). For this example, we set the date the fire started as '2019-10-16' and obtain the number of days in between the start of the health event and the fire event. The query also retrieves the minimum and maximum duration, which allows the intelligent system to infer if the health event is still ongoing, as detailed in [13]. We perform the test multiple times with different reference dates; the queries and results are available in the HECON repository.

Competency Questions eight to eleven (CQ8 to CQ11). Finally, we consider CQs that are related to provenance information. We encode the query (as shown in Fig.5b) and retrieve the Activity (CQ8), the Source (CQ9) and the Provider (CQ10) that support the value of HES for a given SNOMED CT concept. The query results are displayed in Fig. 6.

In summary, these results show that the HECON Ontology allows us to represent and access knowledge about the evolution of conditions recorded in health records. We selected a random sample of health records and queried their HES using HECON and the KG. We were able to replicate the results obtained in [13], therefore meeting our requirements successfully.

activity	sourceName	provider	confidenceVal	supportVal
hecon:KnowledgeExtraction	Public Web Sources	kg:NHS	0.008264462	1.93E-04
hecon:KnowledgeExtraction	Public Web Sources	kg:NHS	0.001618123	1.93E-04
hecon:KnowledgeExtraction	Public Web Sources	kg:NHS	0.001152074	1.93E-04
hecon:RuleExecution	Knowledge Extraction Data	kg:NHS		

Fig. 6: Query outcome for CQs:8-11. SNOMED CT concept 36971009-Sinusitis.

```sparql
SELECT  ?patientId ?snmdIdentifier ?reasonDescription ?typeEvent ?pace ?startDate
?fireEvent ?days ?minDurationValue ?minDurationUnit ?maxDurationValue ?maxDurationUnit
FROM <http://conrad.kmi.open.ac.uk/hecon/kg/EHR>
WHERE {
  VALUES (?patientId)
  { ( <http://kmi.open.ac.uk/emergency/hr/618e8418-9eb9-42f2-9905-617649e99c13> ) }
  {?id fhir:Patient.identifier ?patientId ;
        fhir:Condition.period ?blankNode ;
        fhir:Condition.codeableConcept ?reasonBlank .
    ?blankNode fhir:Period.start ?startDate ;
        fhir:Period.end ?endDate .
    ?reasonBlank skos:inScheme ?snmdIdentifier ;
        fhir:CodeableConcept.text ?reasonDescription .
  }.{
    GRAPH <http://conrad.kmi.open.ac.uk/hecon/kg/HES>{
      ?s rdf:type ?typeEvent ;
         hecon:hasSctConcept ?snmdIdentifier ;
         prov:wasGeneratedBy ?anActivity_IRI .
      ?anActivity_IRI rdf:type hecon:KnowledgeExtraction .
      FILTER ( ?typeEvent != hecon:HealthEvolutionStatement) .
      OPTIONAL { ?anActivity_IRI hecon:confidenceValue ?confidenceVal } .
      OPTIONAL {?s hecon:hasPace ?pace ;
                   hecon:hasMaxDuration ?maxDuration_IRI ;
                   hecon:hasMinDuration ?minDuration_IRI .
                ?maxDuration_IRI time:numericDuration ?maxDurationValue ;
                                 time:unitType ?maxDurationUnit .
                ?minDuration_IRI time:numericDuration ?minDurationValue ;
                                 time:unitType ?minDurationUnit . }
    }
  }BIND ( "2019-10-17"^^<http://www.w3.org/2001/XMLSchema#dateTime> AS ?fireEvent
  BIND ((?fireEvent)-(?startDate) AS ?days).
}ORDER  BY DESC ( ?confidenceVal )
```

(a) CQs: 5-7.

```sparql
SELECT DISTINCT ?activity ?sourceName ?provider ?confidenceVal ?supportVal
WHERE {
  VALUES (?snmdIdentifier) { (sct:36971009) } # sinusitis
  ?s rdf:type ?typeEvent ;
     hecon:hasSctConcept ?snmdIdentifier ;
     prov:wasGeneratedBy ?anActivity_IRI ;
     prov:wasDerivedFrom ?aSource_IRI .
  ?anActivity_IRI rdf:type ?activity .
  ?aSource_IRI rdf:type hecon:Source ;
               rdfs:label ?sourceName ;
               hecon:hasProvider ?provider .
  FILTER ( ?typeEvent != hecon:HealthEvolutionStatement
         && ?activity != prov:Activity ) .
  OPTIONAL { ?anActivity_IRI hecon:confidenceValue ?confidenceVal ;
           hecon:supportValue ?supportVal .}
  OPTIONAL {?anActivity_IRI hecon:hasSourceSnmdConcept ?sourceConcept ;
           hecon:hasTargetSnmdConcept ?targetConcept ;
           rdfs:label ?ruleName }.
  OPTIONAL {?s hecon:hasPace ?pace ;
     hecon:hasMaxDuration ?maxDuration ;
     hecon:hasMinDuration ?minDuration ; }
} ORDER BY ?snmdIdentifier DESC ( ?confidenceVal )
```

(b) CQ: 8-11.

Fig. 5: SPARQL queries for CQs

8 Conclusions and Future work

In this paper, we presented HECON - Health Condition Evolution Ontology, a formal model for representing and reasoning on health events evolution over time. We presented our approach as an important tool to address the knowledge requirements of an intelligent system that automatically estimates a person's current state of health in the context of a fire emergency.

An essential contribution of our work refers to the Knowledge Graph of Health Condition Evolution. We identified that structured information about health condition's evolution is not readily accessible. By following a Linked Data approach we provide a structured database of conditions' evolution over time, which is easy to access and fit to out system requirements.

Here, we focused on the ontology and the knowledge graph, leaving the quality evaluation of the knowledge acquisition processes to future work. The review and validation of the automatic data construction with domain experts' help are essential; it will extend SNOMED CT coverage and ultimately improve the identification of ongoing health events.

References

1. Bodenreider, O., Cornet, R., Vreeman, D.: Recent Developments in Clinical Terminologies - SNOMED CT, LOINC, and RxNorm. Yearb Med Inform (2018)
2. Cook, D.J., Duncan, G., Sprint, G., Fritz, R.L.: Using Smart City Technology to Make Healthcare Smarter. Proceedings of the IEEE (4), 708–722 (Apr 2018)
3. Daga, E., Asprino, L., Mulholland, P., Gangemi, A.: Facade-X: An Opinionated Approach to SPARQL Anything. In: Volume 53: Further with Knowledge Graphs (2021), http://oro.open.ac.uk/78973/
4. Galton, A., Worboys, M.: An Ontology of Information for Emergency Management. In: Proc. of ISCRAM. p. 10 (2011)

5. Gruninger, M., Fox, M.S.: Methodology for the Design and Evaluation of Ontologies. Workshop on Basic Ontological Issues in Knowledge Sharing p. 10 (1995)
6. Hobbs, J.R., Feng, P.: Time ontology in OWL (Sep 2006), https://www.w3.org/TR/owl-time/
7. Lamprinakos, G., Mousas, A., Boufis, A., Karmiris, P., Mantzouratos, S., Kapsalis, A., Kaklamani, D., Venieris, I.: Using FHIR to develop a healthcare mobile application. In: Proceedings of the 4th ICWMCH (2014)
8. Lebo, T., Sahoo, S., McGuinness, D., Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S., Zhao, J.: PROV-O:The PROV ontology (2013)
9. Leroux, H., Metke-Jimenez, A., Lawley, M.J.: Towards achieving semantic interoperability of clinical study data with FHIR. Jrnl. of Biomedical Semantics (2017)
10. López, F.: Overview Of Methodologies For Building Ontologies p. 13 (1999)
11. McNeill, F., Bental, D., Missier, P., Steyn, J., Komar, T., Bryans, J.: Communication in Emergency Management through Data Integration and Trust: an introduction to the CEM-DIT system. In: Proc. of ISCRAM. p. 12 (2019)
12. Morales Tirado, A.C., Daga, E., Motta, E.: Effective Use of Personal Health Records to Support Emergency Services. In: International Conference on Knowledge Engineering and Knowledge Management (2020)
13. Morales Tirado, A.C., Daga, E., Motta, E.: Reasoning on Health Condition Evolution for Enhanced Detection of Vulnerable People in Emergency Settings. In: Proceedings of the 11th on KCAP (2021)
14. Morales Tirado, A.C., Daga, E., Motta, E.: CONRAD - Health Condition Radar: an Intelligent System for Emergency Support. In: 5th Workshop on Semantic Web solutions for large-scale biomedical data analytics (2022)
15. NHS Digital services: FHIR UK Core, https://digital.nhs.uk/services/fhir-uk-core
16. NHS Digital services: National requirements for SNOMED CT, https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct
17. Nunavath, V., Prinz, A., Comes, T.: Identifying First Responders Information Needs: Supporting Search and Rescue Operations for Fire Emergency Response. IJISCRAM (2016)
18. Patsakis, C., Papageorgiou, A., Falcone, F., Solanas, A.: s-Health as a driver towards better emergency response systems in urban environments. In: IEEE International Symposium on MeMeA Proceedings (2015)
19. Presutti, V., Daga, E., Gangemi, A., Blomqvist, E.: eXtreme Design with Content Ontology Design Patterns. Proceedings Workshop on Ontology Patterns (2009)
20. Shearer, R., Motik, B., Horrocks, I.: HermiT: A Highly-Efficient OWL Reasoner (2008)
21. Suarez-Figueroa, M.C., Gomez-Perez, A., Fernandez-Lopez, M.: The NeOn Methodology for Ontology Engineering, p. 9–34 (2012)
22. UK Government: Fire safety risk assessment: means of escape for disabled people (2007), https://www.gov.uk/government/publications/fire-safety-risk-assessment-means-of-escape-for-disabled-people
23. UK Government: Guidance emergency response and recovery (Oct 2013), https://www.gov.uk/guidance/emergency-response-and-recovery
24. Uschold, M., Gruninger, M.: Ontologies: principles, methods and applications. The Knowledge Engineering Review (02) (Jun 1996)
25. Walonoski, J., Kramer, M., Nichols, J., Quina, A., Moesel, C., Hall, D., Duffett, C., Dube, K., Gallagher, T., McLachlan, S.: Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. Journal of the American Medical Informatics Association (2018)