

# 의학 용어의 효과적인 검색을 위한 SNOMED CT 브라우저의 구현

류우석\*

Implementation of a SNOMED CT Browser for Effective Searching of Clinical Terminology

Woo-Seok Ryu\*

요 약

SNOMED CT 표준 의학 용어 체계를 활용하여 진료기록을 작성하기 위해서는 용어 체계에 포함된 방대한 양의 용어들 중 적합한 용어를 빠른 시간에 선택하는 것이 필요하다. 기존의 검색 브라우저들은 검색 결과를 단순 목록으로 나열하므로 유사한 용어들을 사용자가 구분하여 선택하기 어려운 문제가 있다. 본 논문에서는 용어 체계의 특성을 고려하여 효과적인 용어의 검색이 가능한 브라우저를 제안한다. 제안하는 시스템에서는 간략화된 트리 형태로 검색 결과의 계층적 구조를 함께 표시하므로 검색 결과 용어들 중 원하는 용어를 빠르게 선택할 수 있다. 그리고 시스템 설계와 구현을 통해 제안하는 방법의 효과성을 입증한다.

## ABSTRACT

To write a medical record using SNOMED CT standard clinical terminologies, it is necessary to find and select an appropriate terminology from the huge volume of terminologies within short time. Using previous SNOMED CT search browsers, it is very difficult to select appropriate one from search results since they provide a simple list-up of similar candidate terminologies. This paper proposes a novel search browser which supports effect searching of clinical terminology by utilizing characteristics of SNOMED CT. The proposed system provides a simplified tree-view representing hierarchical structures of search results which enables fast selection of appropriate terminology from the search results. Design and Implementation of the system proves effectiveness of the proposed approach.

## 키워드

SNOMED CT, Clinical Terminology, Search Browser, Hierarchical View  
SNOMED CT, 의학 용어, 검색 브라우저, 계층 구조 조회

## 1. 서론

전자의무기록(EMR) 및 전자건강기록(EHR)에서 환자 검사, 진단, 치료 등을 일관성 있게 기록, 공유 및 교환을 하기 위해서는 표준화된 의학 용어 체계의 도입이 필요하다[1-2]. SNOMED CT(Systematized Nomenclature of Medicine-Clinical Terms)[3]는 약

40만 개의 의학적 의미(Concept, 이하 컨셉)를 이용하여 표준화된 환자 기록을 가능하게 하는 대표적인 의학 용어 체계이다. 이 용어 체계 자체는 명확한 구조를 가지고 있음에도 불구하고 내재된 컨셉의 수가 매우 방대함에 따라 실무에서는 SNOMED CT 브라우저라는 소프트웨어를 이용하여 임상적 상황에 가장 적절한 컨셉을 검색하고 이를 진료 기록에 활용한다

\* 교신저자 (corresponding author) : 부산가톨릭대학교 병원경영학과(wsryu@cup.ac.kr)  
접수일자 : 2015. 07. 30

심사(수정)일자 : 2015. 09. 13

게재확정일자 : 2015. 09. 23

[4]. 이 브라우저는 사용자로부터 의학 용어를 검색어로 입력받은 후 검색어에 해당하는 SNOMED CT 컨셉들을 검색하여 그 결과를 표시해 주는 일종의 검색 엔진으로 인터넷 웹 서비스, 모바일 앱, PC 프로그램 등 다양한 형태로 제공되고 있다[3-4].

브라우저의 기능은 문자열 기반의 컨셉 검색 및 검색 결과 조회, 컨셉의 상세 조회 등 비교적 간단하다. 하지만 SNOMED CT 용어 체계는 매우 복잡한 특성이 있다. 용어 체계에 포함된 컨셉 하나하나씩은 각각 고유한 의학적 의미를 가지고 있으며 숫자 형식의 식별자와 여러 개의 용어(Description)을 가지고 있다. 이때 하나의 컨셉이 여러 개의 용어를 가지는 이유는 동일한 증상이라도 이를 표현하는 용어가 국가별 등으로 서로 다를 수 있기 때문이다. 또한, 컨셉이 상당히 세분화된 만큼 의미적으로 유사한 컨셉이 많이 존재하며, 이에 따라 용어가 서로 동일하거나 매우 유사한 컨셉도 상당수 존재하는 특징이 있다[5].

기존의 SNOMED CT 브라우저들은 용어 체계의 특징을 고려하지 않고 용어 검색 시 결과에 해당하는 컨셉들을 단순한 목록 형태로 제시함에 따라 사용자가 검색 결과 컨셉들을 서로 구분하기가 매우 까다로운 문제가 있다. 그림 1을 예를 들면 “목 관절”을 의미하는 “Joint of Neck”을 검색어로 입력하였을 때 그 결과가 7건이 나타나는 것을 확인할 수 있다. 유사한 컨셉들이 단순 나열되어 있으므로 사용자가 그중 가장 적합한 것을 선택하기 위해서는 최악의 경우 검색 결과들을 모두 클릭해서 상세 조회를 해야 한다.

본 논문에서는 기존의 브라우저에서 발생하는 사용자의 불편을 최소화하고 효과적으로 검색하기 위한 SNOMED CT 브라우저 시스템을 설계하고 구현을 통해 이를 입증하고자 한다. 특히, 용어 검색 결과를 단순 목록으로 나타내는 대신 SNOMED CT 용어체계의 특징을 고려하여 계층화된 트리 형태로 제시함으로써, 사용자가 검색 결과 컨셉들을 보다 효과적으로 이해하고 원하는 컨셉을 선택할 수 있도록 한다.

본 논문의 구성은 다음과 같다. 2장에서는 효과적인 검색을 위한 브라우저의 요건을 정의하고, 3장에서 데이터베이스 설계 및 검색 결과 표시 알고리즘을 제안한다. 4장에서는 구현을 통해 브라우저의 유용성을 입증하고 마지막으로 5장에서 결론을 기술한다.

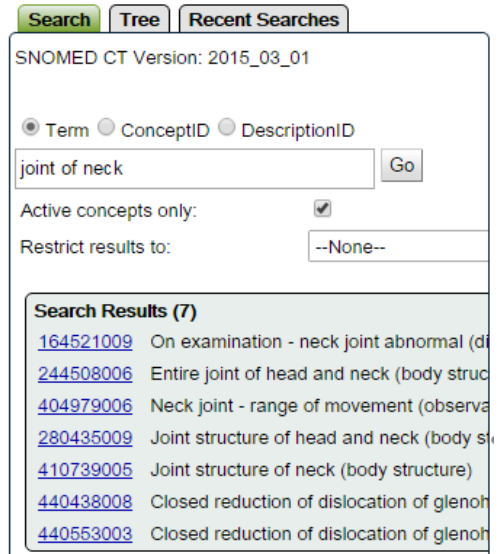


그림 1. NLM SNOMED CT 브라우저의 검색 결과 화면

Fig. 1 A screenshot of a search result of NLM SNOMED CT browser

## II. 브라우저 요건 정의

SNOMED CT 용어체계에서 의학적 의미를 가지고 있는 컨셉은 ID와 몇 개의 용어 이외에 자체적으로 추가적인 정보를 가지고 있지 않다. 해당 컨셉의 의미를 명확하게 식별하고 이해하기 위해서는 컨셉 간의 관계를 파악하는 것이 필요한데 그 중에서도 “is-a” 관계[6]를 이용하여 시각화하면 해당 컨셉의 의미를 보다 명확하게 이해할 수 있다[7-8].

그림 2는 SNOMED CT 컨셉의 “is-a” 계층 구조 예시이다[9]. “SNOMED CT Concept”은 컨셉 계층 구조상에서 가장 위에 있는 루트 컨셉이며 루트 컨셉에 직접적으로 연결된 최상위 컨셉(Top-Level Concept, TLC)은 “Clinical Finding”을 포함하여 19개가 존재한다. 그림에서 보면 “Bacterial pneumonia”라는 컨셉은 계층 구조를 통해 “Infective pneumonia”라는 것을 확인할 수 있고 이는 최종적으로 “Clinical Finding”이라는 것을 확인할 수 있다.

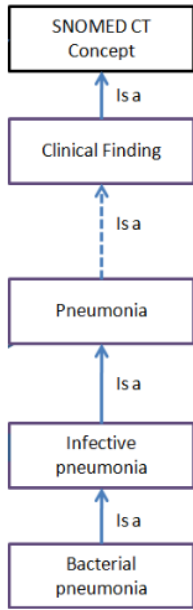


그림 2. SNOMED CT 개념의 계층 구조  
Fig. 2 Hierarchical structure of SNOMED CT concepts

즉, 계층 구조를 적절하게 이용하여 검색 결과를 구조화 하면 단순 목록에 비교하여 보다 직관적으로 검색 결과를 탐색할 수 있는 장점이 있다. 용어 검색 결과를 효과적으로 표시하기 위해서는 아래와 같은 브라우저 요건을 충족시키는 것이 필요하다.

- 검색 결과 개념의 계층 구조 도시 : 검색 결과를 단순 나열식으로 하는 대신 계층 구조를 함께 표시하여 개념의 의미를 직관적으로 이해할 수 있도록 한다.
- 검색 결과 개념들 간의 관계 도시 : 계층 구조를 표시할 때 검색 결과 개념들 간에 존재하는 다양한 관계(부모-자식 관계, 형제 관계, 조상-손자 관계 등)[10]를 식별할 수 있도록 한다.
- 간략화된 계층 구조 도시 : 조회의 용이성을 위해 검색 결과 트리를 최대한 간략화 하여 표시한다.

위 요건 중 세 번째 요건은 SNOMED CT 용어 체계의 특성상 필수적으로 고려되어야 할 부분이다. 계층 구조 상에서 루트 개념부터 각 개념까지 방문해야

하는 개념의 개수를 “거리”라고 정의하면 약 40만개의 개념들의 평균 거리가 10 이상이며 최대 29에 달하므로 계층 구조를 모두 검색 결과에 도시하면 검색 결과가 지나치게 복잡해질 수 있다. 그러므로 검색 결과 계층 구조에서 꼭 도시해야 하는 부분과 그렇지 않은 부분을 구분하여 선별적으로 계층 구조를 도시하는 것이 필요하다.

### III. 시스템 설계

2장에서 제시한 요건을 만족시키기 위해서 아래와 같은 테이블들을 포함하는 데이터베이스 설계를 제시한다.

- 베이스 테이블 : SNOMED CT 배포 버전에 포함되어 있는 개념, 용어, 및 관계를 각각 저장하는 테이블
- 경로 테이블 : 루트 개념부터 각 개념에 도달하기 위한 계층 구조에서의 경로 정보를 미리 검색하여 저장한 테이블

경로 테이블은 검색 결과를 실시간에 트리 형태로 표현하기 위해 각 개념들에 대한 경로 정보를 미리 검색한 후 별도의 테이블로 저장해 놓은 테이블이다. SNOMED CT 개념의 계층 구조는 다중 부모를 허용하는 특징이 있음에 따라 개념 하나의 경로가 여러 개 존재할 수 있다. 이 경우 각각 별도의 레코드로 저장함으로써 SNOMED CT 계층 구조상의 모든 경로 정보를 저장한다.

그림 3은 SNOMED CT 브라우저를 위한 데이터베이스 설계이다. 그림 3에서 Concept, Description, Relationship 테이블은 베이스 테이블이며 경로 테이블은 FullHierarchy라는 이름으로 설계되어 있다. FullHierarchy 테이블에서 id는 일련번호, conceptId는 개념의 식별자, FSN(Fully Specified Name)은 개념의 대표 용어, Depth는 계층 구조상에서 루트 개념부터 해당 개념까지의 거리, 그리고 topology는 루트 개념부터 해당 개념까지의 경로를 문자열 형태로 저장하는 필드이다.

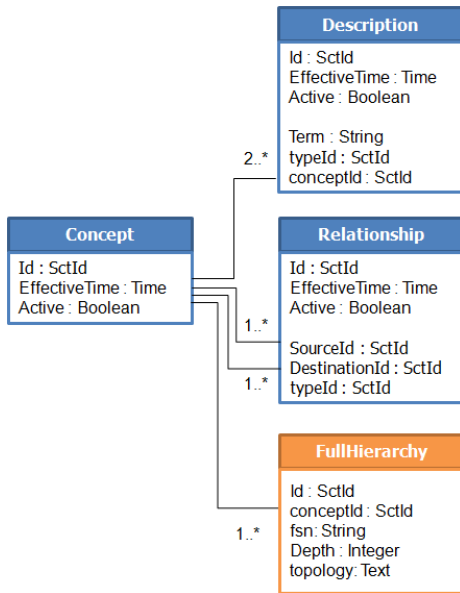


그림 3. 데이터베이스 스키마 설계  
 Fig. 3 Database schema design

위 테이블 설계를 기반으로 한 용어 검색 알고리즘은 다음과 같은 순서로 수행된다.

- (Step 1) 입력받은 문자열에 대해 Description 테이블의 Term 필드에서 Full-Text 검색을 시행하여 매칭되는 Concept ID의 집합을 생성
- (Step 2) Concept ID 집합에 대해 FullHierarchy 테이블에서 ID별로 경로 정보를 검색하여 컨셉 경로 집합을 생성
- (Step 3) 경로 집합을 정합하여 하나의 검색 결과 트리를 생성
- (Step 4) 생성된 검색 결과 트리를 화면에 표시

위 알고리즘 중 Step 3은 두 컨셉의 경로가 유사할 시 계층 구조상에서 동일한 컨셉들은 하나로 통합하는 것을 의미한다. 예를 들어 두 검색 결과 컨셉 I, J의 계층구조가 각각 A->B->C->I, A->B->C->J인 경우 A->B->C까지는 계층 구조가 동일하므로 A->B->C->I, J로 트리를 정합하는 것을 의미한다. 또한 중복되는 계층 구조를 모두 화면에 표시하는 것은 결과의 복잡성을 유발하므로 Step 4에서 최대한

간략화 하여 표시한다. Step 3과 Step 4의 구체적인 방법은 4장 구현을 통해 제시한다.

#### IV. 브라우저의 구현

본 논문에서 제안하는 SNOMED CT 브라우저의 구현 환경은 다음과 같다. 2014년 1월에 배포된 SNOMED CT 용어 체계 자료[11]에 포함된 베이스 테이블을 MySQL Database에 구축하였고, FullHierarchy 테이블을 생성하는 애플리케이션은 자바를 이용하여 구현하였다. 용어 검색을 수행하는 웹 서버 애플리케이션은 Tomcat 8.0 환경에서 JSP와 서블릿을 이용하여 구현하였으며, JQuery 및 Javascript, HTML5를 이용하여 웹 클라이언트를 구현하였다.

구현된 SNOMED CT 브라우저는 용어에 대한 풀 텍스트 검색을 지원하며 검색 결과는 아래와 같은 세 가지 형식으로 표시한다.

- 리스트 뷰 : 기존의 브라우저와 동일하게 결과 컨셉들을 목록 형태로 표시
- 트리 뷰 : 검색 결과 컨셉들을 포함하는 컨셉 계층 구조를 모두 표시. 이 방법은 검색 결과의 상대적 비교가 용이한 장점이 있으나, 검색 결과 건수가 많거나 계층 구조에서의 거리가 먼 경우 트리가 매우 복잡하게 나타남
- 짧은 트리 뷰 : 기본적으로 트리 뷰 형식으로 표시되 검색 결과에 직접적으로 영향을 미치지 않는 중간 노드들을 생략하여 표시하는 방법

그림 4는 구현한 SNOMED CT 브라우저에서 “joint of neck”을 입력한 검색 결과를 짧은 트리 뷰로 표시한 화면의 예시이다. 그림에서 보는 바와 같이 7건의 검색 결과는 각각 최상위 컨셉(Top-Level Concepts)의 하위 트리에 짧은 글꼴로 표시되어 있다. 검색 결과에 최상위 컨셉(TLC)가 함께 표시되고 있으며, 특히 검색 결과 중 “Joint structure of head and neck” 컨셉은 다른 검색 결과인 “Entire joint of head and neck”과 “Joint structure of neck”의 부모 컨셉이며 나머지 두 컨셉은 서로 형제 컨셉인 것을 검색 결과에서 바로 확인할 수 있다. 그림 1의 검색

결과와 비교하여 볼 때 컨셉들을 서로 구분하기가 매우 용이하며 컨셉들 간의 관계 또한 즉시 확인할 수 있는 장점이 있다.

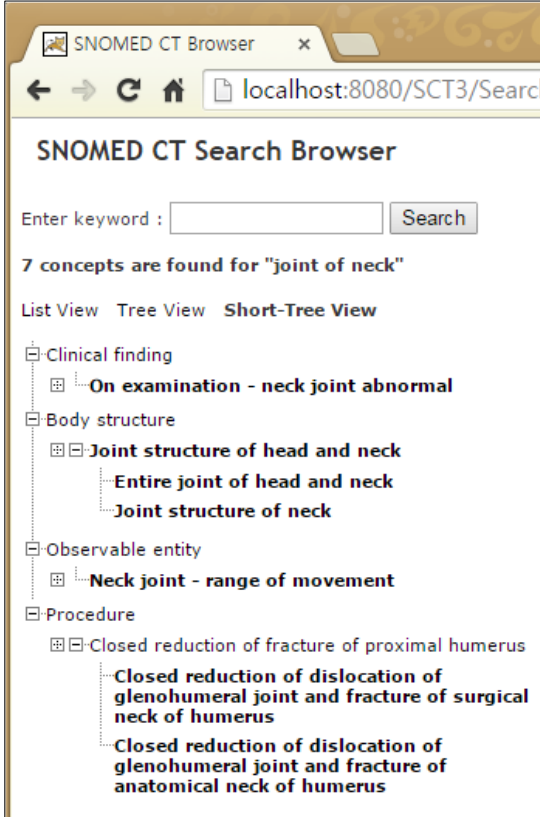


그림 4. SNOMED CT 브라우저 검색 결과 화면

Fig. 4 A screenshot of a search result of SNOMED CT browser

짧은 트리 뷰의 또 다른 장점은 검색 결과 트리 화면 중 계층구조상에 존재하지만 검색 결과가 아닌 컨셉들은 별도의 아이콘 등으로 최대한 축약하여 표시한다는 점이다. 그림 4에서 “Body structure” 컨셉과 “Joint structure of head and neck”의 사이에 작은 아이콘이 존재하는데 이를 클릭하면 그림 5와 같이 계층구조 상에서 생략된 컨셉들이 이탤릭 글꼴로 모두 표시된다. 이를 통해 “Body structure”와 “Joint structure of head and neck” 사이에 생략된 컨셉들이 8개가 있음을 확인할 수 있다. 이때, 생략되는 컨셉의

조건은 자식 컨셉이 한 개 이면서 그 자식 컨셉이 검색 결과에 포함되지 않은 경우이다. 이와 같이 검색 결과에 직접 영향을 미치지 않는 컨셉들을 생략하고 필요시 조회 가능하게 함으로써 검색 결과 컨셉들의 내재된 의미를 비교적 간단하면서도 보다 명확하게 파악할 수 있는 장점이 있다.

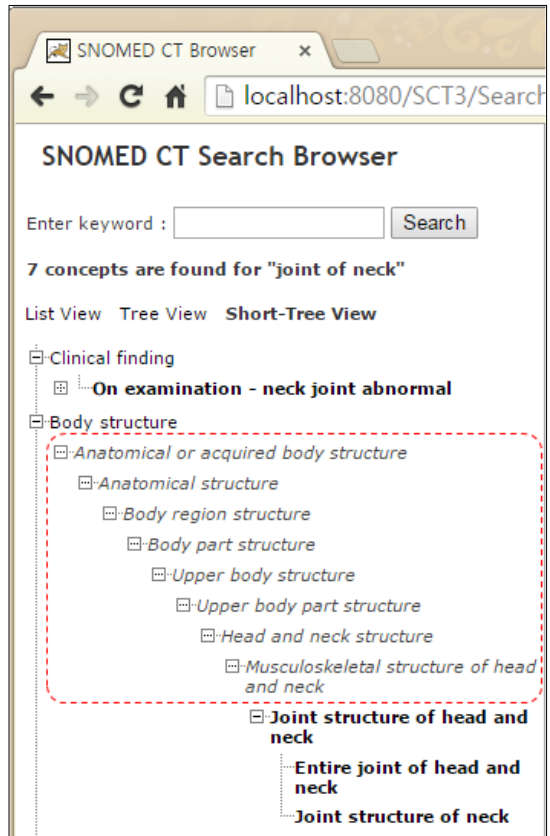


그림 5. SNOMED CT 브라우저 검색 결과의 확장

Fig. 5 An extended view of the search result of SNOMED CT browser

## VI. 결론

본 논문에서는 SNOMED CT 용어 체계의 계층 구조적 특성을 고려하여 용어 검색시 그 결과를 간략화된 트리 형태로 재구성하여 제시하는 새로운 형태의 SNOMED CT 브라우저를 제안하고 이를 구현을 통

해 입증하였다. 이를 통해 사용자는 기존의 브라우저와 비교하여 검색 결과에 나타나는 컨셉들을 보다 빠르게 이해하고 효과적으로 선택함으로써 효율적으로 진료 기록을 수행할 수 있다. 향후 연구로는 본 연구 성과를 임상 과정에 적용하고 진료의사, 의무기록사 등의 전문가 관점에서의 시스템 효용성을 임상 현장에서 평가하는 것이다.

**감사의 글**

이 논문은 2015년도 부산가톨릭대학교 교내연구비에 의하여 연구되었음

**References**

[1] K. Donnelly, "SNOMED-CT: The Advanced Terminology and Coding System for eHealth," *Stud Health Technol Inform*, vol 121, 2006, pp. 279-290.

[2] H. Lee and J. Oh, "A Study on the Health Screening Solution by Using Electronic Medical Record," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 10, no. 7, 2015, pp. 825-830.

[3] <http://www.ihtsdo.org/snomed-ct/>

[4] J. Rogers and O. Bodenreider, "SNOMED CT: Browsing the Browsers," In *Proc. Int. Conf. Knowledge Representation in Medicine*, Lausanne, Switzerland, 2008, pp.30-36.

[5] S. Lusignan, T. Chan, and S. Jones, "Large complex terminologies: more coding choice, but harder to find data - reflections on introduction of SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms) as an NHS standard," *Informatics in Primary Care*, vol. 19, no. 3, 2011, pp.3-5.

[6] W. Ryu, "Requirement Analysis of Search Browser for Efficient Searching of Clinical Terminology," *J. of the Korea Institute of Information and Communication Engineering*, vol. 18, no. 11, 2014, pp. 2691-2696.

[7] J. Lee, Y. Kim, and D. Kim, "Segmentation and Visualization of Human Anatomy using Medical Imagery," *J. of the Korea Institute of*

*Electronic Communication Sciences*, vol. 8, no. 1, 2013, pp. 191-197.

[8] Y. Oh, K. Park, and E. Kim, "Analysis on the Effectiveness of Algorithm Visualization System for Structured Programming Language Education," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 7, no. 1, 2012, pp. 45-51.

[9] <http://www.snomed.org/tig.pdf>

[10] W. Ryu, "A Restructuring Method for Search Results of SNOMED CT Browser," *J. of Embedded Systems and Applications*, vol. 10, no. 3, 2015, pp. 165-170.

[11] <http://www.nlm.nih.gov/research/umls/>

**저자 소개**



**류우석(Woo-Seok Ryu)**

1997년 부산대학교 컴퓨터공학과 졸업 (공학사)

1999년 부산대학교 대학원 컴퓨터공학과 졸업(공학석사)

2012년 부산대학교 대학원 컴퓨터공학과 졸업(공학박사)

2013년~현재 부산가톨릭대학교 병원경영학과 조교수

※ 관심분야 : 의료정보, 의학용어, U-Health, 빅데이터