
의학용어의 구조 검색을 지원하는 SNOMED CT 브라우저 시스템

류우석*

*부산가톨릭대학교

A SNOMED CT Browser System Supporting Structural Search of Clinical Terminology

Wooseok Ryu*

*Catholic University of Pusan

E-mail : wsryu@cup.ac.kr

요 약

SNOMED CT 브라우저는 SNOMED CT 의학 용어 체계에 포함된 용어들을 검색하는 검색 브라우저이다. 이 용어들은 서로 다양한 관계를 통해 구조화되어 있는 특징이 있는데 기존의 브라우저들은 그 구조를 이용하지 않고 단지 문자열 매칭에 의한 결과 목록만을 제시하는 문제가 있다. 본 논문에서는 검색 결과를 서브그래프 형태로 표시함으로써 용어의 구조 검색을 가능하게 하는 브라우저 시스템을 제안하고 이를 구현하였다. 구현된 시스템은 문자열 기반 검색, 트리 기반 검색 결과 구조화, 컨셉 조회 히스토리 등의 기능을 포함하는 특징이 있다.

ABSTRACT

SNOMED CT browser is a search browser which searches and browses terminologies include in SNOMED CT. These terminologies shows a structural form using a variety of relationships. However, previous browsers merely lists up substring-matched search results, rather than using structural characteristics. This paper proposes and implements a browser system which shows a sub-graph of search results enabling structural search of the results. The implementation includes searching of terminologies based on substring-matching, tree-based graphical organization of the search results, and history of concept views.

키워드

SNOMED CT, Search Browser, Comparative Search

1. 서 론

SNOMED CT 브라우저 시스템은 표준화된 종합적 의학용어 체계인 SNOMED CT에 포함된 40만 개 이상의 의학 용어(컨셉)를 검색하고 브라우저하기 위한 S/W이다[1]. 기존에 공개된 브라우저는 PC 애플리케이션, 웹 서비스, 모바일 앱 등

다양한 형태로 개발되어 제공되고 있으나, 용어 검색 시 컨셉의 ID 또는 용어 문자열 매칭을 통한 검색을 수행하고 그 결과를 목록으로 나열하여 표시한다는 측면에서는 공통적인 특성을 지니고 있다[2]. 하지만, SNOMED CT 용어체계는 의학 용어 간의 유사성, 중복성 및 모호성 문제를 내포하고 있으므로 단순히 나열식으로 표시하는

방법으로는 검색 사용자가 원하는 컨셉을 빠르게 찾기가 매우 어려운 문제가 있다[3].

본 논문에서는 SNOMED CT 용어 검색을 수행 후 검색 결과에서 사용자가 원하는 빠르게 선택할 수 있도록 검색 결과를 구조화하여 제시하는 새로운 형태의 브라우저 시스템을 제안하고 이를 구현을 통해 입증한다.

II. 기본 아이디어

40여만 건 이상의 컨셉으로 구성된 SNOMED CT 용어 체계에서 컨셉 각각은 독립적이지 않으며 다양한 관계(Relationship)들로 서로 간에 연결되어 있다. SNOMED CT에서는 "is-a", "part of", "severity", "finding site" 등 60개 이상의 관계가 정의되어 있고 컨셉의 의미를 부여하기 위해 다른 컨셉들과 다양한 관계로 연결되는 특징이 있다. 그 중 "is-a" 관계는 부모/자식 관계를 의미하는 것으로서 "SNOMED Clinical Terms"이라는 루트 컨셉의 자식으로 19개의 최상단 컨셉(Top Level Concepts, TLC)이 정의되어 있으며 나머지 40여만 개의 컨셉들은 그 의미에 따라 TLC의 자손으로 서로 연결되어 있다[4].

"is-a" 관계는 컨셉의 의미를 비교/구분하기 위한 가장 기본적인면서도 중요한 관계이며, 그 관계를 이용하여 브라우저에서 용어 검색을 구조화하면 검색 결과 간의 의미를 더욱 빠르게 구분할 수 있다[5]. 그러므로 본 논문에서는 "is-a" 관계에 근거하여 검색 결과를 구조화할 수 있도록 시스템을 구현한다.

III. 시스템 구현

본 논문에서 제안하는 SNOMED CT 브라우저 시스템의 구현 환경은 다음과 같다. 데이터는 SNOMED CT 용어 체계의 2014년 1월 배포본 자료를 이용하였으며 이를 MySQL 데이터베이스에 구축하였다. 브라우저 시스템은 자바 기반의 웹 애플리케이션으로 구현하였으며 Tomcat 8.0 웹 애플리케이션 서버에서 구동되도록 자바 서블릿과 JSP, JQuery, HTML5를 이용하여 구현하였다. 본 시스템은 선행 연구[6]에서의 구현 결과를 확장하여 검색 결과 상세 조회, 상세 조회 히스토리 기능을 추가 구현하였으며, 검색 인터페이스를 개선했다.

구현된 시스템은 크게 두 가지 모듈로 구성된다. 첫 번째 모듈은 검색 모듈로서 입력된 검색어에 해당하는 컨셉을 검색한 후 검색 결과를 구조화 하여 표시하는 모듈이다. 이때 구조화하여 표시하는 방법으로 List View, Tree View, Short-Tree View의 세 가지 형태[6]를 모두 구현하였다. List View는 그림 1과 같이 검색 결과를

19개의 TLC(Top Level Concept)로 구분하여 각각을 목록으로 표시하는 방법이다. 이 방법은 검색 결과를 TLC 기준으로 크게 구분하여 구조화함으로써 비교적 단순하게 표현할 수 있는 반면에 TLC 내의 컨셉들은 여전히 목록으로 표현되는 특징이 있다. 즉, 그림 1에서 검색 결과로 표시된 세 개의 컨셉은 "Clinical Finding"의 하위 컨셉인 것은 확인 가능하나 이들이 서로 어떤 관계인지를 확인하지는 못한다.



그림 1. 검색 결과에 대한 List View

Tree View는 검색 결과 컨셉들에 대해 계층 구조상의 모든 조상 컨셉들을 함께 표시하는 방법이다. 이는 검색어에 매칭되는 컨셉들 간의 계층구조상의 관계를 그림 2와 같이 검색 결과에 트리 형식으로 함께 표시하는 방법으로 3건의 검색 결과 컨셉들 간의 관계를 한눈에 확인할 수 있다. 즉, List View와 비교해볼 때 직관적인 이해가 매우 높으나, 검색 결과의 계층구조가 복잡할 경우 검색 결과상에 표시되는 정보의 양이 매우 많아져서 가독성이 오히려 떨어지는 문제가 발생한다.

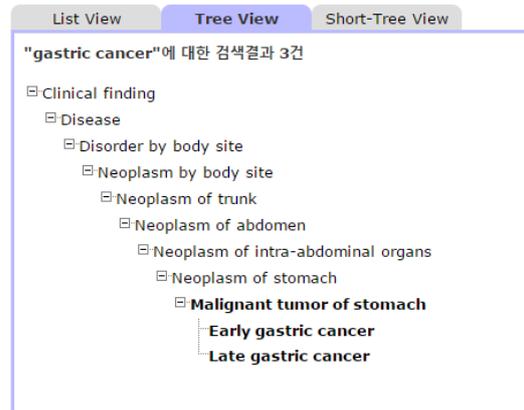


그림 2. 검색 결과에 대한 Tree View

Short-Tree View는 Tree View에서 표시하는 조상 컨셉들 중 검색 결과와 직접적인 관계가 없는 컨셉들을 화면에 생략하여 표시하는 방법이다. 그림 2에서 표시되었던 Disease, Disorder by body site 등의 컨셉들은 3건의 검색 결과의 관계를 비교하는데 직접적인 관련이 없으므로 그림 3

과 같이 간단한 아이콘으로 생략하여 표현한 방법이다. 이는 Tree View와 동일하게 계층구조상의 관계를 함께 표시하여 직관적인 이해를 높이면서도 검색 결과에 표시되는 정보의 양을 줄임으로써 결과 조회의 편의성을 향상시킨 방법이다.



그림 3. 검색 결과에 대한 Short-Tree View

두 번째 모듈은 상세 조회 모듈로서 이는 검색 모듈에서 표시된 컨셉들을 클릭하였을 때 선택한 컨셉에 대한 상세 정보를 검색하여 제공하는 모듈이다. 이 모듈을 통해 제공되는 상세 정보는 Concept Status Information, Descriptions, Parent Concepts, Child Concepts, (from/to) Relationships 등이 있다. 이를 통해 컨셉에 대한 상세 정보를 조회할 수 있으며, 필요시 링크를 통해 연관된 다른 컨셉의 정보를 조회할 수 있다. 그리고, 이 모듈에서는 컨셉 상세 조회 히스토리 기능을 추가로 제공하고 있는데, 이는 상세 조회를 선택한 여러 컨셉들을 서로 쉽게 비교해 보는 특징이 있다.

IV. 결 론

본 논문에서는 SNOMED CT 용어 체계의 구조적 특징을 이용하여 용어 검색시 검색 결과 컨셉들의 계층적 구조를 함께 표시하는 브라우저 시스템을 구현하였다. 사용자는 이 시스템을 통해 부모-자식 관계, 형제 관계 등 검색 결과 컨셉들 간의 관계를 즉시 확인할 수 있으므로 보다 효과적으로 검색 결과를 브라우징 할 수 있다. 또한 히스토리 기능을 통해 컨셉들을 서로 쉽게 비교해 볼 수 있다. 향후 연구로서 임상 과정에서의 활용성 평가를 통해 제안한 시스템의 효과성을 검증하는 것이 필요하다.

참고문헌

[1] Donnelly K., "SNOMED-CT: The Advanced Terminology and Coding System for eHealth," Stud Health Technol Inform, vol 121, pp. 279-290, 2006.
 [2] IHTSDO, "SNOMED CT Browser", http://ihtsdo.org/fileadmin/user_upload/doc

/browsers/browsers.html.

[3] Lusignan S., Chan T., and Jones S., "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, pp. 3-5, 2011.
 [4] IHTSDO, "SNOMED CT Technical Implementation Guide", <http://www.snomed.org/tig.pdf>, 2014.
 [5] Ryu W., "A Restructuring Method for Search Results of SNOMED CT Browser," Journal of Embedded Systems and Applications, Vol 10, No 3, pp. 165-170, 2015.
 [6] Ryu W., "Implementation of a SNOMED CT Browser for Effective Searching of Clinical Terminology," Journal of the Korean Institute of Electronic Communication Sciences, Vol. 10, No. 9, pp. 1059-1064, 2015.