

시각화 기능을 이용한 과학기술정보 검색기술 개발

Development Retrieval Technique of Science Technology Information Using Visualization Skill

여일연, 최호섭, 황미령, 심강섭*, 옥철영*
한국과학기술정보연구원, 울산대학교*

Yeo Il-Yeon, Choe Ho-Seop, Hwang Mi-Nyeong,
Sim Gang-Seop*, Ock Cheol-Young*
KISTI, Ulsan University*

요약

과학기술분야의 빠른 발전 속도와 함께 그에 연관된 전문용어의 출현이 끊임없이 이루어지고 있다. 이에 사용자가 찾고자 하는 용어의 정확한 의미를 구하는 일이 점점 어려워지고 있는 실정이다. 본 논문에서는 사용자의 검색방식을 의미적으로 접근 가능케 하고, 사용자가 요구하는 과학기술정보와 관련된 전문용어를 시각적으로 검색할 수 있는 검색 시각화 애플리케이션을 개발, 구현하였다.

Abstract

With a rapid developing speed of science technology field, related technical terminology is growing continuously. So it's increasingly difficult to search proper meaning of terminology. This paper tries to develop and implement a application that user can retrieve the technical terminology meaningly and visually.

I. 서론

1. 개발배경

하루가 다르게 발전하는 과학기술분야는 수많은 신기술들을 만들어 내고, 그에 파생되는 새로운 용어 또한 수없이 만들어지고 있다. 최근 인터넷의 빠른 발전은 사용자로 하여금 신기술과 새로운 전문용어에 보다 쉽고 빠르게 접근하도록 해주고 있지만, 사용자가 원하는 정확한 의미의 용어를 찾아내는 것이 그리 쉽지만은 않다. 본 논문에서는 사용자가 보다 쉽게 원하는 정보를 찾을 수 있도록 시각화 기술을 이용한 전문용어 검색도구를 소개하고자 한다. 본 도구는 하나의 용어에 대한 상위어와 하위어, 동의어, 유의어, 관련어 등의 정보를 제공함으로써 사용자가 용어에 대한 정보를 의미적으로 찾을 수 있게 해준다. 또한 검색 결과의 의미구조를 시각적으로 나타내 줌으로써 사용자의 접근성을 높여 주었다.

II. 본론

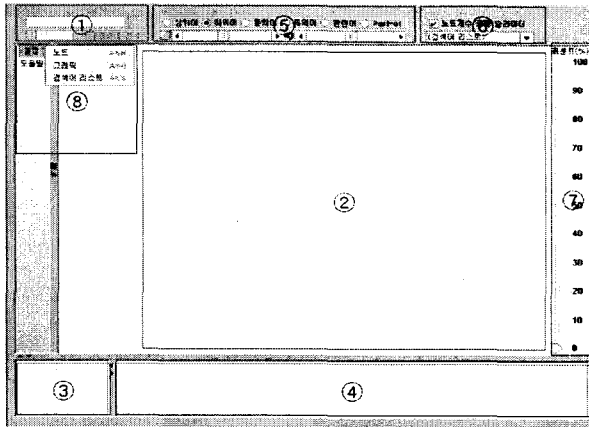
1. 프로그램 소개

과학기술리모콘이라 명한 검색 시각화 애플리케이션은 울산대학교 한국어처리연구실에서 구축한 언어자원을 바탕으로

touchgraph 기술을 이용하여 구현하였다[1]. 본 단락에서는 애플리케이션이 어떤 형식을 구성되어 있는지 알아보려고 한다.

1.1 화면 구성

그림 1에서 보는 바와 같이 애플리케이션은 크게 8개의 구역으로 나뉘어져 있다. ①번 구역은 사용자가 찾고자 하는 검색어를 입력하는 곳이고, ②번 구역은 사용자가 입력한 질의어에 해당하는 단어들 이 표시되는 창이다. ③번 구역은 ②번 구역에 표시되는 단어들 을 개별적으로 나타내는 구역이고, ④번 구역은 각각의 단어들에 대한 용어 뜻풀이가 표시되는 공간이다. ⑤번 구역은 ②번 구역에서 표시되는 용어들에 대한 조작을 할 수 있는 기능들을 모아놓은 곳으로, 선택된 단어들의 상위어, 하위어, 동의어, 유의어, 관련어 등을 찾아 볼 수 있으며 ②번 구역에서 표시되는 용어들의 display 형태를 조작할 수 있는 기능도 포함되어 있다. ⑥번과 ⑦번 구역은 찾고자 하는 용어가 많은 수의 연관된 단어를 가지고 있을 경우, 화면에 표시할 노드의 깊이를 조절하는 기능을 가지고 있다. ⑧번 구역은 애플리케이션에 대한 초기설정을 하는 곳으로 화면에 표시될 용어들의 초기 노드 개수와 표시되는 노드의 모양을 설정해 주고 이전 검색어 리스트를 몇 개까지 보여줄 것인지에 대한 설정을 하는 곳이다.



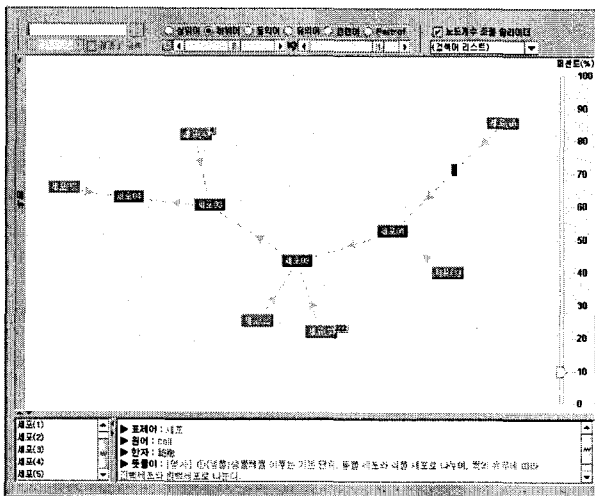
▶▶ 그림 1. 화면 구성

2. 검색

본 단락에서는 애플리케이션을 가지고 실제 검색을 해 봄으로써, 사용자가 찾고자 하는 용어의 의미를 시각적으로 접근해나가는 과정을 살펴보고자 한다.

2.1 일반검색

그림 2에서는 질의어를 ‘세포’로 하고 실제 애플리케이션에서 검색을 실행한 화면을 나타내고 있다. 그림에서 보는 바와 같이 애플리케이션의 ②번 구역에 6개의 ‘세포’라는 단어가 붉게 표시되어 나타난다. 각각의 용어는 서로 다른 의미를 가진 동형이의어를 나타내고 있다. 각 단어 옆에 푸른색으로 표시된 작은 숫자들은 그 용어의 하위어를 나타낸 것이다. ③번 구역에서는 각각의 용어를 다시 한번 표시해주고, ④번 구역에서 그 용어들에 대한 뜻풀이를 해주고 있다.

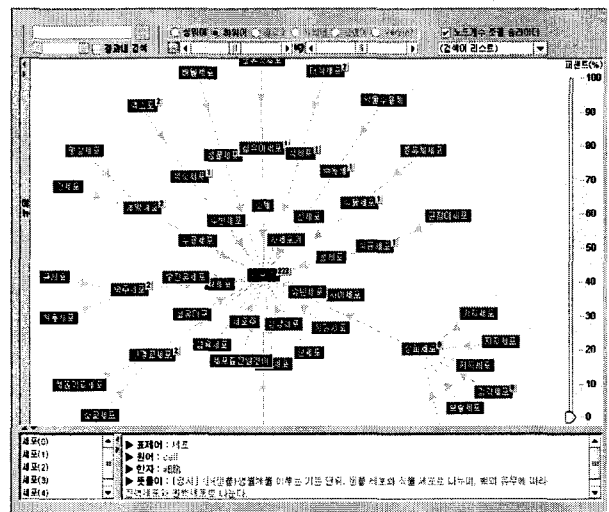


▶▶ 그림 2. 일반검색

2.2 하위어 검색

여기에서는 검색결과 얻어진 단어중 세포(3)에 해당하는 용

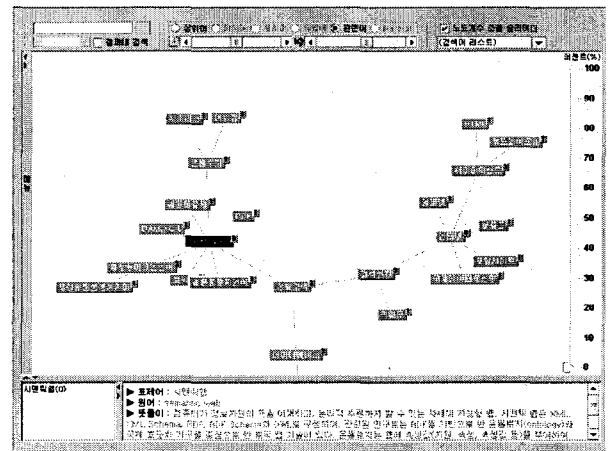
어를 살펴보기로 한다. 사전적 의미가 ‘(생물)생물체를 이루는 기본단위. 동물세포와 식물세포로 나누며, 핵의 유무에 따라 진핵세포와 원핵세포로 나눈다.’로 되어있는 용어는 222개의 하위어를 포함하고 있다. 애플리케이션에서 해당용어를 더블클릭하게 되면 그림 3에서 보는 바와 같이 222개의 하위어중 초기 설정되어있는 노드의 깊이만큼 나타내어 준다. 여기서 ⑥번과 ⑦번 구역의 노드개수 조절 슬라이더를 이용하여 표시되어지는 하위어의 수를 조절할 수 있다.



▶▶ 그림 3. 하위어 검색

2.3 관련어 검색

본 애플리케이션은 해당용어의 관련어를 보여줌으로써 사용자가 더 많은 정보를 얻을 수 있도록 하였다.

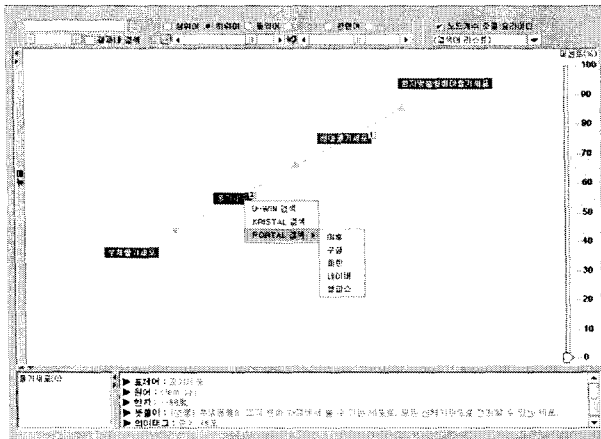


▶▶ 그림 4. 관련어 검색

2.4 포털 검색

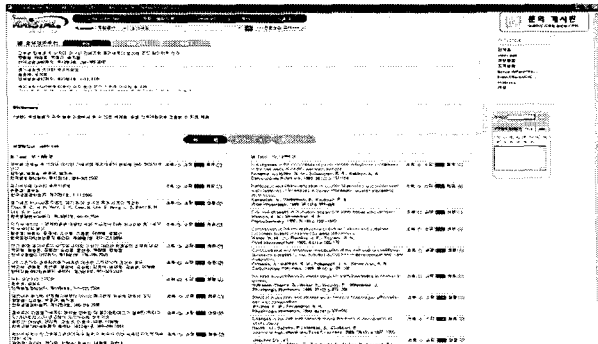
그림 5에서는 그림 3에서 보인 하위어중 ‘출기세포’에 대한 검색결과를 보인 것이다. 그림에서 보는 바와 같이 본 애플리

케이션을 통하여 줄기세포에 관한 정보뿐만 아니라 이와 관련된 용어의 정보까지 얻을 수 있다. 또한 오른쪽 마우스 클릭을 통하여 한국과학기술정보연구원의 데이터베이스를 검색할 수 있고, 상용 포털 기관의 검색엔진을 통한 결과도 얻을 수 있도록 하였다.

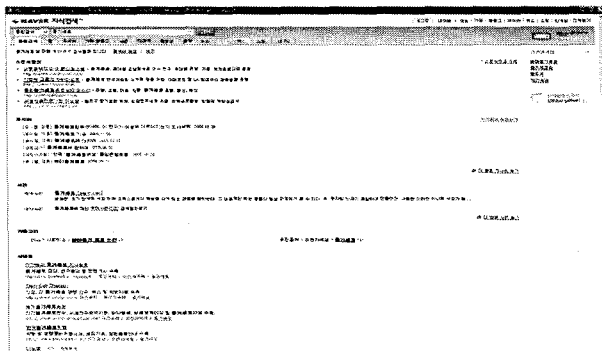


▶▶ 그림 5. 포털 검색

그림 6은 KRISTAL 검색엔진을 이용하여 한국과학기술정보연구원 DB를 검색한 결과를 나타내고 있고, 그림 7은 상용 포털 사이트인 네이버에서 검색한 화면을 보여주고 있다.



▶▶ 그림 6. KRISTAL 검색



▶▶ 그림 7. PORTAL 검색

III. 결 론

급격히 변화하는 시대의 흐름에 맞춰 과학기술의 발전은 눈부시게 이루어지고 있다. 새로운 기술, 용어, 이슈 등은 일반 사용자가 따라잡기 힘들 정도로 늘어만 가고 있다. 상용 포털 기관을 이용한 정보의 검색은 너무도 많은 내용을 포함하고 있기 때문에 그 속에서 사용자가 원하는 정보를 찾아내는 일 또한 쉽지만은 않다. 본 논문에서는 사용자가 보다 쉽게 정보의 의미를 구할 수 있도록 하는 시각화 애플리케이션을 소개하였다. 본 애플리케이션을 통하여 사용자가 원하는 정보의 의미를 마우스 클릭을 통하여 쉽게 시각적으로 찾을 수 있을 것이다. 또한 관련된 하위어, 동의어, 상위어 등을 통하여 사용자의 지식욕구를 보다 충실히 채워줄 것이라 생각된다.

■ 참고 문헌 ■

[1] <http://www.touchgraph.com>