# 한글 인터넷 정보검색 시스템의 기술방향

°이수현, 박동인
시스템 공학 연구소, 자연어 정보처리 연구부

# The State of the Art in Internet Information Retrieval System

SooHyun Lee, DongIn Park
Natural Language Information Processing, Systems Engineering Research Institute

요 약

빠르고 다양하게 늘어나는 정보 더미 속에서 사용자가 원하는 정보를 쉽고 정확하게 찾아주는 방법이 바로 정보 검색이다. 최근에는 인터넷을 이용하여 세계 각처에 흩어져 있는 데이터를 찾거나 웹 홈 페이지를 찾아 주는 인터넷 정보 검색 프로그램들이 각광을 받고 있다. 본 논문에서는 현재 국내외에서 상용되고 있는 인터넷 정보검색 시스템의 검색 방법에 따른 특성을 살펴 보고 국내에서 개발된 한국어 정보검색 시스템과 외국에서 개발된 정보 검색 시스템과의 기능을 비교함으로써, 한국어 정보검색 시스템의 처리 항목을 제시한다.

## 1. 서론

컴퓨터의 발달과 함께 정보도 세계화, 전문화, 다양화가 되고 있다. 그 때문에 현대 사회를 정보 사회라고도 부르고 있다. 최근에는, 정보의 바다로 불리우는 인터네트의 확산으로 보다 다양하고 많은 정보를 언제 어디서든지 쉽게 얻고 활용할 수 있는 세상이 되었다. 그러나, 아무리 좋은 정보라 할지라도 시간이 지난 정보는 이미 쓸모 없는 것이 되어 버리므로 필요할 때 빨리 수집하여야만 높은 부가 가치를 가진 정보가 되는 것이다. 이것은 빠른 시간에 필요한 정보를 보다 정확하게 얻어야 한다는 것을 의미한다.

세계에 흩어져 있는 정보 더미 속에서 사용자가 원하는 정보를 빠르고, 쉽게, 또 정확하게 찾아 주는 방법이 바로 정보 검색이다. 종래에는 자신이 찾고자 하는 정보가 어디에 있는지 정확하게 알고 있지 못하였기 때문에, 과학 기술 문헌 등의 문헌 데이터 베이스를 검색 전문가에게 검색을 의뢰하여 정보를 획득하는 것이 일반적이었다. 그런데, 최근에는 PC 통신, 인터네트, 전자 메일 등을 통하여, 수 많은 일반 사용자에게 바로 정보가 쏟아져 들어 옴으로서, 정보가 대형화되고, 검색하고 싶은 범위도 넓어짐으로 인해, 정형화된 데이터 베이스로부터 전문가에 의해 검색되는 정보 검색 시스템보다는 비전문가인 일반 사용자들이 직접 많은 정보를 보다 쉽고 간편하게 정리하고, 신속 정확하게 검색할 수 있는 시스템이 요구되어 지고 있다. 또한, 최근에는 비전문가가 인터넷을 이용하여 세계 각처에 흩어져 있는 데이터를 찾거나 웹 홈 페이지를 찾을 수 있는 인터넷 정보 검색 프로그램들이 각광을 받고 있으며, 이들 정보 검색 시스템의 사용 빈도는 lycos 의 경우 1 일 100 만건 이상, 알타비스타의 경우 초당 20 건에 이르고 있으며, 국내에서 개발된 검색 시스템만도 분당 10 건 정도로 검색에 이용되고 있는 실정이다.

그러나 필요 이상으로 많고, 정제되지 않은 정보는 오히려 사용자를 혼란에 빠뜨리기 쉽다. 즉, 사용자가 원하는 정보를 찾기 위하여, 여러번 질의어를 수정하면서 재검색을 함으로써 많은 노력과 시간을 허비하게 된다. 따라서, 많은 사람들이 보다 많은 정보를 보다 쉽고, 빠르고, 간결하고, 정확하게 얻기를 요구하게 되었다. 이러한 요구에 부응하여 최근에는 많은 정보검색 시스템들이 개발되어 상용화 되고 있다. 그러나, 대부분의 정보검색 시스템들이 외국에서 외국어 특히 영어의 특성에 맞게 개발되었기 때문에 외국의 정보 검색 시스템을 한국어에 그대로 사용하기는 곤란하다. 따라서, 몇몇 국내의 정보 검색 시스템들이 이러한 영어권의 정보 검색 시스템을 그대로 수용하는 실정에서 우리말 특성에 맞는 우리말 정보검색 시스템의 개발이 필요하게 되었으며, 최근에는 국내외로 지능형 정보검색 시스템에 관한 연구들이 활발히 진행되고 있고[1,2], 대학, 연구소 등에서도 우리말 정보 검색 시스템에 관한 연구가 활발히 진행되고 있다.[3,4] 더욱이, 한국어의 특성을 반영한 정보검색 시스템이 개발되어 현재 인터넷상에서 활용되고 있다.

한편, 의뢰인이 정보검색 전문가에게 자연언어로 필요한 정보를 요구하듯이 사용자가 정보검색 시스템을 직접 조작하여 통상 사용하고 있는 자연언어로 정보를 요구할 수 있도록 하는 것이 앞으로 요구되어 질 정보검색 시스템일 것이다. 그러나, 컴퓨터로 하여금 사용자의 요구사항을 완전하게 파악하도록 하기 위하여는 질의어에 대한 완벽한 처리가 필연적이다.

본 논문에서는 현재 국내외에서 개발된 인터넷 정보검색 시스템의 특성을 살펴 보고 국내에서 개발된 시스템과 외국에서 개발된 시스템과의 기능을 검색 방법의 관점에서 비교함으로써 우리말의 특성이 무엇이며, 우리말 정보검색

시스템에서 고려하여야 할 기능들을 제시한다.

## 2. 정보 검색 시스템의 비교

### 2.1. 정보 검색 시스템의 분류

정보검색 시스템이란 사용자가 필요로 하는 정보를 수집, 분석하여 사용자가 찾기 쉬운 형태로 가공하여 두었다가 사용자의 요구에 따라 사용자가 원하는 정보를 제공하는 시스템을 말한다[5]. 정보 검색 시스템은 분류 방법에 따라 다양하게 분류될 수 있다. 대표적인 분류 방법으로는 정보의 유형에 따라 데이터 검색 시스템, 참조 정보 검색 시스템, 본문 검색 시스템, 질의 응답 시스템, 비디오 텍스 등의 5 가지로 분류하는 방법이 있으나[6], 정보 검색의 응용 사례에 따라서는 신문 기사 검색, 도서관 검색, 온라인 검색, 인터넷 검색, 전문 검색 등으로 나눌 수 있으며, 인터넷 정보 검색 시스템은 검색 방법에 따라 주제 검색과 주제어 검색의 두 가지로 나뉘어 진다. 주제 검색이란, 해당 검색 시스템이 이미 설정하여 놓은 주제 분류를 따라 깊이를 더해가면서 목적하는 최종 자료에 접근하는 방법을 말하며, 주제어 검색이란 사용자의 키워드를 갖고 직접 목적하는 자료에 접근하는 방법을 말한다. 본 논문에서는 주제어 검색 시스템을 중심으로 검색 방법을 비교한다.

### 2.2. 정보 검색 시스템의 비교 항목

본 절에서는 다음과 같이 검색 방법에 따른 19 가지의 비교 항목을 설정하여 각각의 정보검색 시스템에서 처리할 수 있는 기능의 유무를 비교한다.

1. 절단 기능 : 임의의 글자로 시작하거나 임의의 글자로 끝나는 단어를 검색하는 기능으로서 다음과 같이 3 가지로 나뉠 수 있다.
   전방 절단 : 임의의 글자로 끝나는 단어.
   예) *처리 => 정보처리, 기호처리, 수식처리
   후방 절단 : 임의의 글자로 시작하는 단어.
   예) sing* => singing, singer,...
   한글* => 한글날, 한글처리, 한글공학
   중간 절단 : 임의의 글자로 시작해서 임의의 글자로 끝나는 단어.
   예) 한글*처리 => 한글정보처리, 한글입력처리
2. 구 : 구 단위 입력에 대한 처리
   예) computer system
3. 대소문자 구분 : 한국어 처리에는 필요하지 않는 기능이나, 영어의 경우에는 이를 구별하는 기능이 필요하다. 만약, 한국어 검색 시스템에서 영어 검색을 위해서는 이 기능이 고려되어야 할 것이다.
4. 불리언 연산 : 불리언 연산자로는 AND, OR, NOT 의 3 가지가 있다. Alta Vista 를 예로 들면 '+', '-' 기호를 이용하여 검색어가 존재하느냐, 존재하지 않느냐를 밝히는데, 이것은 엄밀한 의미의 연산자가 아니다.
5. 인접 연산 : 두개의 단어가 어느 정도의 거리로 인접되어 있는 경우를 처리하는 기능
   NEAR : 순서를 무시한 인접 연산. 좌우 단어의 선후에 관계없이 좌우에 입력한 두 단어가 인접해 있는 모든 레코드가 검색된다.
   ADJ : 순서를 고려한 인접 연산. 반드시 연산자 좌측에 위치한 단어가 우측에 위치한 단어 보다 먼저 나타나야 한다.
6. 입력 오류 정정 : 입력된 질의어에 오타가 포함되어 있을 경우, 자동적으로 정정하는 기능.
   예) telivision => television,
   성장율 => 성장률
7. 유의어 확장 : 질의어와 유사한 표현으로 확장하여 검색 하는 기능.
   예) 검색 => 탐색, 수색,
   data => information
8. 시소러스 : 시소러스를 참조하여 검색.
9. Concept : 개념을 이용하여 검색.
10. Field 검색 (제한 검색) : 특정 키워드를 레코드 전체를 통하여 검색하는 것이 아니라 Title 필드로 제한해서 검색한다든지, Head 필드로 제한하여 검색하는 기능.
11. 전문 검색 (full-text) : 문서의 일부만이 아니라 전체 문서를 검색하는가.
12. 자연 언어 질의 : 자연 언어로 된 질의어를 처리하는 기능.
    예) 검색에 관한 연구 ?
    가장 빠른 퍼스널 컴퓨터를 생산하는 곳은?
13. 축약, 기호 : 축약어, 기호를 처리하는 기능. 예를 들면, KT 혹은 한통이라는 질의어가 입력되었을 때, 한국통신이라는 단어도 찾는 기능.
14. 질의어 자동 번역 :한국어 검색어를 영어 등으로 번역하여 검색하는 기능. 다국어 검색 시스템 혹은 한국어 검색 시스템에서는 필수적이라 할 수 있겠다.
    예) 정보 => data, information
15. 복합 명사 : 예) 정보검색시스템
16. 동사 활용 : 동사의 활용을 해석하여, 원형, 활용형에 대하여 확장해서 검색하는 기능.
    예) retrieve => retrieval, retrieved
    만들다 => 만들면, 만드니
17. 본문 요약 : 검색 결과로부터 각각의 내용을 요약하여 출력하는 기능.
18. Ranking : 검색 결과를 우선 순위를 계산하여 높은 순위부터 출력한다.
19. Sort : 검색 결과를 사용자 편의에 따라 정렬하여 출력한다. (날짜순, 크기순 등등)

### 2.3. 외국의 정보검색 시스템의 비교

외국에서의 정보검색에 관한 연구는 이미 오래전부터 이루어져 왔으며, 기술 수준도 상당한 레벨에 올라와 있는 실정이다. 특히, 여러가지의 검색 모델과 인덱싱 방법에 관한 많은 연구가 발표되어 왔으며, 최근에는 자연언어 인터페이스, 본문요약, 자동문서분류 등에 관한 연구가 활발히 진행되고 있다. 이러한 환경과 더불어 현재 외국에서 개발되어 사용되어 지고 있는 인터넷 정보검색 시스템은 수십가지에 이르며, 대부분의 검색 시스템에서는 검색 모델로서 불리언 모델을 채택하고 있다. 각각의 시스템들이 갖고 있는 기능은 유사한 점이 많고 약간의 차이를 나타낼 뿐이다. 본 절에서는 현재 사용되어 지고 있는 인터넷 정보검색 시스템들 가운데 12 개의 대표적인 것들을 모아 2.2 절의 비교 항목에 따라 각 시스템들의 검색 방법을 비교하여 표 1 에 보인다.

국외에서 개발된 대부분의 검색 시스템들의 공통점은 기본적으로 전후 절단 기능과 불리언 연산 기능, 전문검색, 자연언어 질의어 처리, ranking, sort 등의 기능을 갖고 있으며, 불리언 연산은 AND, OR, NOT 의 기본 연산과 함께 확장된 복잡한 연산 처리 기능을 갖고 있다. 각 시스템 간의 차이점은 구단위 처리 기능, 대소문자 기능, 인접 연산 기능, Field 검색 기능의 유무에 있을 뿐이며, 현재 각 시스템에서 탑재하고 있는 기능들은 상당한 수준에 올라 있고, 상당히 복잡한 연산도 가

표 1. 국외 정보 검색 시스템의 검색 방법 비교

| 기능 \ 엔진 | Info Seek | ALTA VISTA | OPEN TEXT | DEJA NEWS | LISZT | Lycos | Web Crawler | EINET Galaxy | INKTOMI | Veronica | Yahoo | EXCITE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 절단기능 | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● |
| 구 (phrase) | ● | ● | ● | ● | ● | × | × | × | × | × | × | × |
| 대소문자 구분 | ● | ● | × | × | ● | × | × | × | × | × | × | × |
| AND | ◎ | ● | ● | ● | ● | ● | ● | ● | ◎ | ● | ● | ◎ |
| OR | ◎ | ● | ● | ● | ● | ● | ● | ● | ◎ | ● | ● | ◎ |
| NOT | ◎ | ● | ● | ● | ● | × | ● | ● | ◎ | ● | × | × |
| NEAR | ● | ● | ● | ● | × | ◎ | × | × | × | × | × | × |
| ADJ | ● | × | ● | × | × | × | × | × | × | × | × | × |
| 전문검색(full text) | ● | ● | ● | ● | ● | × | × | ● | ● | × | × | ● |
| 자연언어질의 | ● | ● | × | ● | × | ● | ● | ● | ● | × | × | ● |
| 동사활용 | × | × | × | × | × | ● | × | × | × | × | × | × |
| 입력오류정정 | × | × | × | × | × | × | × | × | × | × | × | × |
| 유의어확장 | × | × | × | × | × | × | × | × | × | × | × | × |
| 복합명사 | × | × | × | × | × | × | × | × | × | × | × | × |
| 축약 | × | × | × | × | × | × | × | × | × | × | × | × |
| Thesaurus 제공 | × | × | × | × | × | × | × | × | × | × | × | × |
| Concept 검색 기능 | × | × | × | × | × | × | × | × | × | × | × | × |
| Field 검색 | × | ● | ● | ● | × | × | × | × | × | × | × | × |
| 본문요약 | × | × | × | × | × | × | × | × | × | × | × | × |
| Ranking | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● |
| Sort | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● |

● : 기능 제공
◎ : 비슷한 기능. +,- 등의 기호를 사용
× : 기능 제공 안 됨

능하다. 본문 요약, 시소러스, 개념 검색 등은 대학 등에서 많은 연구가 진행되고 있다.

### 2.4. 국내의 정보검색 시스템의 비교

국내에서는 수십년 전부터 도서관의 도서 관리 및 검색에 관한 시스템이 개발되었으며, 최근에 들어서는 우리말 특성을 반영한 검색 시스템들이 출현하고 있으나, 아직 단어 나열에 의한 검색이 대부분으로서, 일정 수준의 한국어 처리를 통한 시스템은 2 가지에 불과하다. 여기에서는 이들을 중심으로 설명한다.

국내의 정보검색 시스템에서 공통적으로 처리되는 기능은 구단위 처리와 불리언 연산, 전문검색, 복합명사처리 등이고, 인접어 연산, 개념, Field, 본문 요약, sort 등의 기능은 전무한 상태이다. 한국어 처리에 있어서, 한국어는 대소문자 구분 기능이 필요 없으므로 이에 대한 처리는 고려하지 않아도 무방하나, 영어 입력을 고려할 때에는 그에 대한 처리가 필요할 것이다. 또한, 일부 시스템에서 한글 확장 기능으로서, 오류 정정, 유의어 발음 확장 기능 등을 탑재하고 있으나, 진정한 의미의 형태소 해석 등을 통한 처리가 아니라, 사전 대조식 처리로 인하여 그에는 한계가 있으며, 입력 문장을 해석하여 검색하는 자연언어 질의 처리가 미흡하다. 현재, 국내의 정보검색 연구에서 가장 큰 문제점으로는 복합명사 처리, 자연언어 입력 처리, 순위 인접 연산 처리 등을 들 수 있으며, 이들에 관한 연구가 각계에서 진행되고 있으며,

표 2. 국내의 정보검색 시스템의 검색 방법 비교

| 기능 \ 엔진 | 심마니 | 정보 탐정 | 서치아일랜드 | 까치네 | 코씨크 |
|---|---|---|---|---|---|
| 절단기능 | × | ● | ◎ | ◎ | ◎ |
| 구 | ● | ● | ● | × | × |
| 대소문자 구분 | × | ● | × | × | × |
| AND | ● | ◎ | ● | × | × |
| OR | ● | ◎ | ● | × | × |
| NOT | ◎ | × | × | × | × |
| NEAR | × | × | × | × | × |
| ADJ | × | × | × | × | × |
| 전문 검색 (full text) | ● | ● | × | × | × |
| 자연언어질의 | × | ● | × | × | × |
| 동사활용 | ◎ | × | × | × | × |
| 입력 오류 정정 | ● | × | ● | × | × |
| 유의어 확장 | ● | × | × | × | × |
| 복합 명사 | ● | ● | × | × | × |
| 축약 | × | × | × | × | × |
| 질의어자동번역 | ● | × | × | × | × |
| Thesaurus 제공 | × | ● | × | × | × |
| Concept 검색 기능 | × | × | × | × | × |
| Field 검색 | × | × | × | × | × |
| 본문 요약 | × | × | × | × | × |
| Ranking | ● | ● | × | × | × |
| Sort | × | × | × | × | × |

한편으로는 시소러스, 개념에 의한 검색에 관한 연구도 시작 단계에 들입한 실정이다.

### 2.5. 국내외 정보검색 시스템의 비교

국내의 정보검색 시스템은 절단기능, 구단위 처리기능, 불리언 연산, 인접 연산, field 검색 등, 거의 전반에 걸쳐서 국외 시스템들에 비하여 기능의 수준이 상당히 미약하며, 전문검색과 자연언어 질의처리에 있어서는 외국의 경우에는 의미 해석을 통한 처리를 함으로서 상당한 수준의 질의 처리를 하고 있으나, 국내의 경우는 조사, 부사 등의 불용어를 제거하는 수준이다. 전문검색과 자연언어 질의를 처리하기 위하여는 형태소해석뿐만 아니라, 구문, 의미해석 기술이 관건이며, 국외에서는 수십년전부터 자연언어 처리에 관한 연구가 진행되어, 형태소해석은 이미 완성 단계에 왔다고도 볼 수 있다.

## 3. 한글 정보검색 시스템의 기능

2 장에서는 기존의 검색 시스템을 비교하였다. 3 장에서는 먼저 한국어의 특성에 대하여 간단히 살펴 보고, 한국어 정보검색 시스템의 설계시 고려하여야 할 기능들에 대해서 살펴본다.

### 3.1. 한국어의 특성

한국어는 영어에 비하여 동사활용이 복잡하고, 어순이 자유로운 등 영어에서 찾아 볼 수 없는 여러가지 특징적인

현상이 있다. 일반적으로 한국어만이 갖고 있는 특성은 형태적 특성과 통사적 특성의 2가지로 나누어서 살펴 볼 수 있다[7].

✧ 한국어의 형태적 특성
- 한국어는 어근에 파생 접사가 덧붙어서 품사가 바뀌거나 어미가 차례로 여럿이 결합하여 단어를 이룬다.
  예) 덧붙이시었겟더군요
- 한국어는 용언의 불규칙 현상이 발달하였다.
- 한국어는 복합명사의 띄어 쓰기가 자유롭다.

✧ 한국어의 통사적 특성
- 어순의 차이 : 한국어는 문장 구성 요소의 자리 옮김이 자유롭다.
- 문법적 형태소는 반드시 어근이나 어간 뒤에 온다. 즉, 조사는 체언 뒤에 붙어 쓰이며, 활용 어미는 용언의 어간 뒤에 쓰인다.
- 수식 관계 : 한국어의 이유 구문은 앞에서 뒤로 수식한다.
- 주어, 목적어 등의 성분이 생략 가능하다.

등을 들 수 있다.

### 3.2. 한국어 정보검색 시스템의 처리기능

위에서 열거한 한국어의 특성은 영어와의 차이점이 되며, 한국어 정보검색 시스템에서 한국어의 처리를 위하여 고려되어야 할 것들이다. 특히, 한국어의 특성에만 나타남으로서, 한국어 정보검색 시스템에서 처리되어야 할 기능에 다음과 같은 것을 들 수 있으며,

- 복합명사처리
- 자연언어 질의처리
- 용언활용처리
- 입력오류정정
- 유의어 확장
- 영어의 한글표기 처리기능
- 한글-영어-숫자 혼용표기 처리기능
- 축약어 처리
- 질의어 자동번역
- 본문요약

이밖에 개념검색, 자동문서분류기능, 시소러스 등이 있다. 한국어 처리에 있어서 특히 고려하여야 할 사항을 다음에 설명한다.

#### 3.2.1. 복합명사처리

단어와 단어가 결합하여 만드는 복합명사, 복합동사 등은 단어간의 조합에 따라 무한히 표현할 수 있으며, 결합 순서에 따라 의미가 달라지는 등 많은 변화가 있다. 사용자가 "정보검색" 이라는 질의어로 검색을 하려 할 때, 복합 명사 처리가 이루어지지 않는다면, "정보검색"이라는 단어가 들어 있는 문서만이 검색될 것이고, "정보의 검색", "정보를 검색하는" 등의 표현으로 쓰여진 문서는 검색되지 않을 것이다. 그리고, 복합 명사를 일일이 사전에 기술한다는 것은 불가능한 일이다. 복합 명사 처리시 고려하여야 할 사항을 다음에 든다.

- 똑같은 두개의 단어가 결합하여 생성되는 복합 명사도 결합 순서에 따라 의미가 변한다.
  예) 정보검색 --- 검색정보
- 순서고려 인접연산에 따른 복합명사 연산
- 의미적으로 일치하는 복합 명사
  예) 정보를 찾아 본다 -> 정보검색
- 조사에 따른 명사의 선후교체와 의미변화
  화일에 관한 속성 => 화일속성
  장치에 의한 처리 => 처리장치
- 띄어 쓰기
  정보검색 = 정보 검색

#### 3.2.2. 자연언어 질의처리

현재 연구되어 지고 있는 자연언어 질의처리는 대부분 단문에 대한 연구이며, 앞으로 연구되어질 자연언어 질의는 다음과 같이 여러개의 문장으로 이루어진 복합문의 처리에 관한 것이다. 이를 처리하기 위하여는 우선 질의어를 단문으로 분리하여, 각각의 단문으로 부터 keyword를 추출한 다음, keyword를 재조합하면서 검색을 반복함으로서 원하는 정보를 얻을 수 있다.

예 1) 세계 주요 도시 곳곳에 있는 해외 지점과 인터넷를 통해 업무 연락을 하거나 자료를 주고 받고 싶다. 중동 지역에 있는 튜니시아(Tunisia) 지점에서 인터넷를 접속하려고 하는데 튜니시아에서 인터넷 접속 서비스를 제공하는 회사는 ?

예 2) 아시아지역 국가들 중에서 한국은 월드컵 결승전에 몇번이나 참가하였으며, 또 한국이 처음으로 참가한 월드컵 결승전은 언제 어느 나라에서 개최한 대회였는가?

## 4. 결론

본 논문에서는 국내에서 개발된 정보검색 시스템과 국외에서 개발된 정보검색 시스템의 기능을 비교하는 한편, 한국어의 특성에 따른 영어와의 차이점에 대하여 살펴 보았으며, 한국어 정보 검색 시스템에서 처리되어야 할 기능에 대하여 고찰하였다. 국내에서 개발된 정보검색 시스템은 국외에서 개발된 정보검색 시스템에 비하여 비슷한 기능이 많이 부가되었으나, 처리 능력은 절반 정도의 수준에 머물고 있으며, 아직 연구가 계속되어야 할 부분이 많다. 특히, 복합명사처리, 자연언어 질의 처리등을 유연하게 하기 위해서는 형태소해석, 구문해석, 의미해석 기술에 덧붙여 문맥해석 기술이 발전되어야 할 것이다. 나아가서 시소러스, 개념검색, 본문요약에 관한 연구도 점진적으로 이루어져야 할 것이다.

## 5. 참고 문헌


[1] 한국과학기술원, "지능형 정보검색에 관한 연구", 1995.

[2] 박영몽, 이정태 외, "지식 기반의 정보 검색 시스템", 한국 정보 과학회 논문지, Vol.21, No.11, pp.2090-2098, 1994.

[3] 이은철, 이종혁 외, "한국어 텍스트 검색 시스템 KIRT 의 구현", 한국 정보 과학회 봄 학술 발표 논문집, Vol.19, No.1, pp.639-642, 1992.

[4] 이승률 외, "자연어 질의 정보 검색 시스템의 비주제어 탐색 방법을 통한 성능 개선", 한글 및 한국어 정보 처리 학술 대회, pp.374-377, 1994.

[5] 정영미, "정보 검색론", 구미무역 출판부, 1993.

[6] 김영택, "자연 언어 처리", 교학사, 1994.

[7] 남기심, 고영근, "표준 국어 문법론", 탑출판사 (1991).