

# 검색 포털들의 검색어 추천 서비스 분석 평가: 네이버와 구글의 연관 검색어 서비스를 중심으로\*

## Analysis and Evaluation of Term Suggestion Services of Korean Search Portals: The Case of Naver and Google Korea

박소연 (Soyeon Park)\*\*

### 초 록

본 연구에서는 주요 검색 포털들의 검색어 추천 서비스를 분석, 평가하였다. 이 연구에서는 네이버와 구글 코리아를 대상으로 추천되는 연관 검색어의 적합도 및 최신성을 평가하고, 연관 검색어의 개수 및 분포, 연관 검색어가 제공되지 않는 질의의 특징을 조사하였다. 또한 연관 검색어의 유형을 질의와 연관 검색어의 관계 측면에서 분석하고, 연관 검색어들 중 유해 검색어의 유형 및 특징, 비표준어의 유형 및 특징도 조사하였다. 마지막으로, 한글 질의와 영어 질의, 대중적인 질의와 전문적인 질의의 연관 검색어의 특징을 비교하였다. 연구 결과, 네이버가 구글보다 연관 검색어의 적합도와 최신성이 다소 높은 것으로 나타났다. 또한 구글과 네이버 모두 새로운 연관 검색어를 제시하기보다는 질의에 단어를 추가 또는 삭제하거나, 질의와 동일한 검색어나 동의어 검색어를 제공하는 경우가 많은 것으로 나타났다. 본 연구의 결과는 향후 포털들의 검색어 추천 서비스의 개선에 활용될 수 있을 것으로 기대된다.

### ABSTRACT

This study aims to analyze and evaluate term suggestion services of major search portals, Naver and Google Korea. In particular, this study evaluated relevance and currency of related search terms provided, and analyzed characteristics such as number and distribution of terms, and queries that did not produce terms. This study also analyzed types of terms in terms of the relationship between queries and terms, and investigated types and characteristics of harmful terms and terms with grammatical errors. Finally, Korean queries and English queries, and popular queries and academic queries were compared in terms of the amount and relevance of search terms provided. The results of this study show that the relevance and currency of Naver's related search terms are somewhat higher than those of Google. Both Naver and Google tend to add terms to or delete terms from original queries, and provide identical search terms or synonym terms rather than providing entirely new search terms. The results of this study can be implemented to the portal's effective development of term suggestion services.

키워드: 검색 포털, 검색어 추천, 연관 검색, 웹 검색, 정보 검색  
search portals, term suggestion, related search, web searching,  
information retrieval

\* 본 연구는 덕성여자대학교 2012년도 교내 연구비 지원에 의해 수행되었음.

\*\* 덕성여자대학교 문헌정보학과 교수(sypark@duksung.ac.kr)

■ 논문접수일자: 2013년 5월 28일 ■ 최초심사일자: 2013년 6월 4일 ■ 게재확정일자: 2013년 6월 13일  
■ 정보관리학회지, 30(2), 297-315, 2013. [http://dx.doi.org/10.3743/KOSIM.2013.30.2.297]

## 1. 서론

피씨 검색뿐만 아니라 모바일 검색이 급증하면서 국내 이용자들의 웹 검색이 일상화되는 상황이다. 웹 검색 환경의 특징으로는 웹 문서의 양이 천문학적으로 방대하다는 점을 들 수 있다. 또한 웹 이용자들의 검색 행태의 특징으로는 단순성을 들 수 있다. 즉, 웹 검색 분야의 많은 국내외 선행 연구들에서 공통적으로 발견한 것은 대부분의 웹 이용자들이 복잡한 검색식이나 연산자를 사용하지 않고, 한두 단어로 구성된 매우 짧고 단순한 질의를 입력한다는 점이다 (박소연, 2011; 이준호, 박소연, 권혁성, 2003; Sanderson, 2008; Silverstein, Henzinger, Marais, & Moricz, 1999; Spink, Wolfram, Jansen, & Saracevic, 2001). 한두 단어로 구성된 질의는 일반적이고, 특정성이 결여되기 쉽고, 모호할 수 있기 때문에 이용자의 정보 요구를 정확히 표현하기에는 한계가 있다. 이처럼 단순한 질의를 이용하여 방대한 문서 집합에서 적합 문서를 검색해야 한다는 데에 웹 검색의 어려움이 있다고 할 수 있다. 이러한 문제를 해결하기 위하여 국내외 대부분의 검색 포털들은 검색어 추천 기법을 도입하여 왔다.

검색어 추천은 “이용자가 보다 효율적으로 질의를 작성할 수 있도록 이용자가 입력한 초기 질의에 적합한 검색어들을 제안하는 정보 검색 기법의 일종”이다(Huang, Chien, & Oyang, 2003, p. 638). 즉, 검색어 추천은 특정 질의와 의미적 관련성이 긴밀한 검색어를 제안, 추천함으로써 이용자의 탐색 확장이나 변경을 지원하는 기법으로서, 웹 검색 엔진뿐만 아니라, 학술 데이터베이스나 전통적 검색 시스템 등에서도

오랜 기간 사용되어 온 중요한 서비스라고 할 수 있다. 웹 검색 엔진의 검색 성능을 향상시키고, 웹 이용자들의 질의 작성을 효율적으로 지원하고, 불필요한 재검색 수행을 줄이기 위하여, 효율적인 검색어 추천 서비스 개발이 매우 중요하다고 할 수 있다.

한편 국내외 선행 연구들의 경우, 웹 검색 엔진을 위한 검색어 추천 알고리즘 구축 및 제안에 집중되어 있으며, 제안된 검색어의 품질을 종합적으로 분석, 평가한 연구는 드문 실정이다. 이에 본 연구에서는 국내외 주요 검색 포털들인 구글과 네이버의 검색어 추천 서비스의 특징과 장단점을 조사하고자 한다. 좀 더 구체적으로 이 연구에서는 추천된 검색어, 즉 연관 검색어의 적합도, 최신성과 같은 품질을 분석하고자 한다. 둘째, 포털별로 제공되는 연관 검색어의 개수 및 분포, 연관 검색어가 제공되지 않는 질의의 특징을 조사하고자 한다. 셋째, 연관 검색어의 유형을 질의와 연관 검색어의 관계 및 연관 검색어들 간의 관계 등의 측면에서 분석하고자 한다. 넷째, 연관 검색어들 중 유해 검색어의 유형 및 특징, 비표준어의 유형 및 특징도 조사하고자 한다. 마지막으로, 한글 질의와 영어 질의, 대중적인 질의와 전문적인 질의 간의 연관 검색어의 특징을 비교, 분석하고자 한다. 국내에서는 검색어 추천 서비스가 연관 검색어로 통용되는 만큼, 이 논문에서도 추천된 검색어를 지칭하기 위하여 연관 검색어라는 용어를 사용하고자 한다.

이 연구의 결과는 향후 포털의 검색어 추천 서비스의 개선에 활용될 수 있을 것으로 기대된다. 즉 본 연구의 결과는 포털 업체들의 효과적인 검색어 추천 알고리즘 개발 및 인터페이스

개발에 중요한 자료로서 활용될 수 있을 것으로 기대된다.

## 2. 선행 연구

정보 검색 분야에서 검색어 추천 또는 제안과 관련된 국외 선행 연구는 오랜 기간 동안 수행되어 왔으며(Bates, 1981), 질의 확장(query expansion)과 밀접한 관련이 있다. 질의 확장은 이용자의 초기 질의를 정련하기 위한 기법으로, 이용자의 질의에 검색어나 구를 추가함으로써 실행되며, 검색 시스템에 의해 자동으로 실행되거나, 이용자에 의해서 적합성 피드백의 형태로 수작업으로 실행된다(Xu, Luo, Yu, & Xu, 2011). 전통적인 질의 확장 기법은 대개 문서 내의 출현 단어의 분석에 기반하는데 비해(Croft, Metzler, & Strohmman, 2010), 웹 검색 분야에서의 검색어 추천 기법은 문서 내 출현 단어뿐만 아니라, 질의에 대한 검색 결과 중 이용자들이 많이 조회한 결과와 같은 다양한 요소들에 근거하여 연관 검색어들을 추출한다는 점에서 차이가 있다.

이 장에서는 웹 검색 분야에서의 검색어 추천 관련 선행 연구에 초점을 맞추고자 한다. 검색어 추천과 관련된 많은 국외 연구들은 검색어 추천 방식 개발 및 구현에 집중되어 왔다(Beeferman & Berger, 2000; Chen, 2011; Huang, Chien, & Oyang, 2003; Liu, Miao, Zhang, Ma, & Ru, 2011; Xu et al., 2011; Zhang & Nasraoui, 2008). Huang, Chien, Oyang(2003)은 웹 검색 엔진의 질의 로그로부터 적합한 검색어를 추출하고 제안하는 방식을 제안하였다. 즉 이들은

특정 질의가 검색된 세션 내에서 함께 등장한 검색어들을 연관 검색어로 추천하는 방식을 제안하였으며, 약 200만 개의 질의가 포함된 대만의 검색 엔진 로그를 통하여 이 방식의 성능을 평가하였다. Zhang과 Nasraoui(2008)도 중국의 산업 분야 전문 검색 엔진인 Sina의 질의 로그를 대상으로 특정한 세션 내에서 이용자가 연속하여 입력하는 검색어들을 연관 검색어 후보군으로 추출한 후, 이 후보군에 포함된 검색어들 중 내용상 유사도가 높은 검색어들을 연관 검색어로 추천하는 알고리즘을 제안하였다. 오프라인 실험과 평가자들의 평가를 통하여 제안된 알고리즘이 효율적임을 발견하였다.

Liu 등(2011)은 이용자들이 특정 질의 입력 후 조회한 문서의 미리보기(snippet)에 등장한 단어들을 연관 검색어로 추천하는 알고리즘을(snippet click models) 두 가지 방식으로 제안하였다. 또한 이러한 알고리즘과 기존 검색 엔진의 알고리즘의 검색 성능을 비교한 결과, 새로운 알고리즘의 검색 성능이 더 높게 나타났다고 보고하였다. 이와 유사하게 Xu 등(2011)도 이용자들이 질의 입력 후 조회한 문서의 미리보기에 등장한 단어들을 추출한 후, 이 검색어들을 의미적 유사도에 따라 순위화한 후 연관 검색어로 추천하는 알고리즘을 개발하였다. 또한 이러한 알고리즘과 구글과 야후의 검색어 추천 방식을 비교하여, 제안된 알고리즘의 우수성을 보고하였다.

연관 검색에 관한 국내 연구는 전산학 분야에서 주로 수행되어 왔으며, 웹 검색 엔진에서 연관 검색어를 추출하기 위한 알고리즘을 제안하는 연구 주제에 집중되어 왔다. 예를 들어, 문상준과 최재걸(2004)은 네이버와 같은 검색 엔

진에서 수집된 검색어들 간의 상호 연관성을 찾기 위해 변형된 Apriori 알고리즘을 제안하였다. 기존의 데이터 마이닝 알고리즘이 일정 횟수 이상 검색된 라지 셋(large set) 검색어들에만 집중하는 반면, 이들의 연구에서는 일정 횟수 이하로 검색되는 검색어들인 스몰 셋(small set)의 연관 법칙을 찾기 위한 방법을 제시하였다. 박성주, 박하얀, 김용혁(2008)은 네이버를 대상으로 분산 객체 기반의 병렬 처리 방식으로 검색 결과 페이지들을 재분석하여 연관 검색어를 추천하는 방법을 제안하였다. 이정훈과 전서현(2009)은 질의 확장을 지원하기 위하여 검색 엔진 이용자가 방문하는 웹 문서로부터 추출된 용어들을 연관 검색어로 추천하고, 이렇게 수집된 검색어들의 유용성의 순위를 정하는 알고리즘을 제안하였다. 주길홍, 이주일, 이원석(2011) 역시 웹 환경에서 이용자가 조회한 문서들로 구성된 로그 데이터 분석을 통하여 연관 검색어를 추출하는 방법을 제안하였다. 이들은 Apriori 알고리즘을 사용하여 특정 주제어와 긴밀히 연관된 검색 주제어를 제시하는 기법인 로그 기반 키워드 추출 방법과 유사성이 높은 키워드들을 클러스터링하는 기법인 확산 기반 클러스터링 방법을 제안하였다. 김지혜, 장재연, 윤홍준, 김한준(2010)은 검색 포털들의 뉴스 기사를 대상으로 키워드 간의 연관성을 추출하기 위한 방식으로 변형된 tf-idf 방식을 제안하였다.

이처럼 국내외 선행 연구들에서는 웹 검색 엔진을 위한 검색어 추천 알고리즘을 제안하거나, 알고리즘의 검색 성능을 평가하기 위한 실험 연구가 주로 수행되어 왔다. 국내외 선행 연구들 중에서 웹 검색 포털들의 검색어 추천 서비스를 상세하고 종합적으로 분석, 평가한 연구는

찾아보기 어려운 실정이다. 이에 본 연구에서는 주요 검색 포털인 네이버와 구글의 검색어 추천 서비스의 특징과 품질을 분석, 평가하고자 한다.

### 3. 연구 방법

#### 3.1 자료 수집

본 연구를 수행하기 위하여서는 평가 대상 포털의 선택 및 연관 검색어 평가용 질의와 이 질의들에 대한 연관 검색어의 수집이 필요하다. 첫째, 평가 대상 포털들로, 구글, 네이버를 선택한 이유는 국내외 검색 포털 분야에서의 이들의 위상과 인지도 때문이다. 네이버는 2000년대 초반 이후 국내 검색 포털들 중 시장 점유율 조사, 방문자 수 조사, 검색 시간 점유율 조사 등에 있어서 지속적으로 1위를 차지하고 있다. 웹 사이트 평가 및 트래픽 분석업체인 Korean Click (<http://www.koreanclick.com>)의 조사에 따르면, 2013년도 4월 기준으로 네이버의 국내 검색 시장 점유율이 72.56%인 것으로 나타났다. 또한 구글은 전 세계 검색 엔진들 중 점유율 1위를 차지하고 있는 상황이다. 좀 더 구체적으로 이 연구에서는 구글 코리아를 대상으로 평가를 수행하였는데, 이는 구글 닷컴이 한글 질의에 대해서는 관련 검색어를 거의 추천하지 않고, 구글 닷컴과 구글 코리아의 검색어 추천 방식 및 인터페이스 등이 모두 동일하기 때문이다.

둘째, 질의를 수집하기 위하여 문헌정보학 전공 대학생들이 선택한 실제 질의들을 활용하였다. 즉, 2012년도 2학기 서울 소재 여자 사립 대학교의 문헌정보학과에서 개설한 “온라인 정보

검색론” 수강생들이 제공한 실제 질의 400개를 평가 대상으로 하였다. 이 수강생들은 모두 문헌정보학을 전공하거나(34명) 복수 전공하는(6명) 학생들로 3학년이 전체의 65%, 4학년이 35%였다. 이처럼 이용자들의 실제 질의를 선택한 이유는 검색 포털들이 일반에게 공개하는 질의들은 엔터테인먼트와 뉴스에 편중되어 다양한 질의에 대한 연관 검색어 평가에 제약이 따르기 때문이다. 즉, 네이버, 다음은 각각 인기 검색어 사이트(<http://searchc.naver.com/ntk/>)와 검색 트렌드 사이트(<http://trend.search.daum.net/SearchTrend/index.html>)를 운영, 공개하는데, 이 사이트들은 일정한 기간 동안 이용자들이 입력한 통합 검색 질의들 중 이슈 검색어, 인기 검색어, 급상승 검색어의 순위를 제공한다. 이렇게 공개되는 질의들은 엔터테인먼트와 뉴스 관련 검색어들에 집중되어서 경제, 의학, IT, 자연과학, 사회과학, 인문학과 같은 다양한 주제 분야의 질의들을 포함하지 못한다는 한계가 있다. 따라서 이 연구에서는 다양한 주제 분야와 다양한 난이도의 질의를 대상으로 연관 검색어를 평가하기 위하여 이용자들의 실제 질의를 활용하기로 하였다. 좀 더 구체적으로, 40명의 수강생들이 연관 검색 서비스 평가 용도로 본인이 관심 있는 질의 11개씩을 제출하였으며, 이때, 학문, 문화/예술/엔터테인먼트, 의학, 경제, 사회, 지역/여행, IT, 뉴스, 라이프스타일, 쇼핑 등 다양한 주제 분야에서 질의를 고루 선택하고, 대중적 질의와 전문적인 질의를 고루 선택하도록 안내하였다. 셋째, 이렇게 수집된 실제 질의들을 대상으로 중복 질의를 제거한 후, 2013년도 1월 9일과 10일 이틀 동안 구글, 네이버, 다음에서 제공하는 연관 검색어들을 수집하였다.

400개의 질의를 선택한 이유는 하루에 국내 검색 포털에 입력되는 통합 검색 질의의 수를 고려할 때, 표본 오차 95% 신뢰수준  $\pm 5\%$ 를 허용할 경우 필요한 표본의 크기가 384개로 통계학 문헌에서 제시되고 있기 때문이다(Arkin & Colton, 1963). 한편 400개 질의의 주제를 선행 연구에서(박소연, 이준호, 김지승, 2005) 도출된 주제 분류 체계에 따라 분석한 결과, 전체 질의 중 인문/사회과학 관련된 주제가 18.2%로 가장 많았고, 이어서 엔터테인먼트(14%), 라이프스타일(12.3%), 뉴스/미디어(10.7%), 사회/정치/행정/법/종교(8.7%), 건강/의학(5.9%), 지역/여행(5.9%), 컴퓨터/IT(4.9%), 문화/예술(4.5%) 등의 순으로 나타났다. 인문/사회과학 하위 범주로는 금융/경제/경영 관련 질의가 7.5%로 높게 나타났다. 도출된 주제 범주 중 엔터테인먼트, 라이프스타일 등은 대중적인 주제 범주로 인문/사회과학, 건강/의학 등은 전문적인 주제 범주로 통합될 수 있는데, 전체 질의 중 대중적인 주제는 약 58%로, 전문적인 주제는 42%로 나타났다. 400개 질의들 중 대중적인 질의의 예로는 “늑대소년”, “싸이”, “카카오톡”, “브레이킹 댄”, “아이폰 5”, “오바마”, “취업”, “무한도전”, “코레일 내일로”, “태블릿” 등을 들 수 있으며, 보다 전문적인 질의의 예로는 “정보 검색”, “기록 관리”, “장서 개발”, “재정 절벽”, “양적 완화”, “포괄 수가제”, “죄수의 딜레마”, “군중 심리”, “플라시보 효과”, “건설”, “안구 건조증” 등을 들 수 있다.

### 3.2 평가 기준

본 연구에서는 구글과 네이버의 연관 검색어

를 첫째 적합도 기준에 따라 평가하고자 한다. 연관 검색어의 적합도는 적합과 부적합으로 평가되었으며, 그 이유는 주제 분야 전문가인 교수들을 대상으로 파일럿 스터디를 수행한 결과, 포털들의 연관 검색어가 적합한 그룹과 전혀 부적합한 그룹으로 양분되는 경향이 뚜렷함을 발견하였기 때문이다. 즉, 검색 포털들이 추천하는 연관 검색어들은 질의의 의미를 설명하거나 질의와 형태나 내용이 중복되는 검색어 그룹과 질의와 전혀 무관하거나 내용상 오류가 있는 검색어 그룹으로 확연하게 양분되어서, 상세한 수준의 적합도 평가 기준이 불필요하다는 전문가들의 의견이 지배적이었다.

따라서 이 연구에서는 제공된 연관 검색어가 질의와 의미적으로 긴밀한 관계를 맺고 있어, 이용자가 연관 검색어를 이용하여 질의를 확장하거나 구체화하는데 도움이 될 수 있는 경우, 적합으로 평가하였다. 반면, 연관 검색어가 질의와 전혀 무관한 경우, 지나친 의미 확장이 발생한 경우, 유해 검색어가 노출된 경우, 오타가 발생한 경우, 내용상의 오류가 존재하는 경우, 질의의 일부만 포함되어 전체 질의와는 무관한 연관 검색어가 제공된 경우 등은 부적합으로 평가하였다. 좀 더 구체적으로, 연관 검색어가 질의와 의미적으로 대등한 관계, 또는 상하 관계를 맺고 있을 때, 적합으로 평가하였다. 예를 들어, 기업 관련 질의의 연관 검색어의 경우, 같은 분야의 경쟁사, 해당 기업이 생산한 프로그램이나 제품, 해당기업 계열사나 공장, 해당기업 정보, 해당기업 웹사이트, 카페 및 블로그 등은 적합으로 평가된 반면, 해당 기업과 전혀 무관한 업종의 기업은 부적합으로 평가되었다. 적합한 연관 검색어의 예로는 네이버에서 “슈뢰딩

거의 고양이”의 연관 검색어로 “양자역학”, “죄수의 딜레마”의 연관 검색어로 “게임 이론”, “아도르노”의 연관 검색어로 “프랑크푸르트 학파,” 농심의 연관 검색어로 “삼양” 및 “빙그레”, “야상곡”의 연관 검색어로 “쇼팽 야상곡”, “구글”의 연관 검색어로 “다음”, “야후”, “유튜브” 등이 제공된 예를 들 수 있다. 부적합한 연관 검색어의 예로는 네이버에서 “구글”의 연관 검색어로 “○○○ 합성 사진”, “토렌트”, “파일조”, “앤 해서웨이”, “야상곡”의 연관 검색어로 “비단옷을 입고 밤에 다닌다는 뜻으로, 성공은 했지만 보람이 없는 행동을 일컫는 사자성어”가 제공된 예를 들 수 있다. 적합 및 부적합 검색어의 상세한 예시는 4장 연구 결과 섹션에 제시되어 있다.

둘째, 최신성의 경우, 본 연구에 사용된 질의들이 2012년 12월에 수집되었다는 점을 고려하여, 연관 검색어의 내용의 시점이 2012년인 경우 최신성이 높은 검색어로, 연관 검색어의 내용이 1년 이상 경과한 경우, 즉 연관 검색어의 내용의 시점이 2011년인 경우 최신성이 낮은 검색어로 평가하였다.

셋째, 유해 검색어란 일반적인 검색어가 아니며, 음란성, 저작권 침해성, 비방성, 광고성 검색어 등과 같이 연관 검색어로 수집, 제공되기에 문제가 있는 검색어들을 의미한다. 위와 같은 유해 검색어의 유형은 선행 연구(윤여생, 유진호, 2012) 및 네이버의 “연관 검색어 삭제 기준”(NHN, 2013), 한국인터넷자율정책기구(KISO)(2013)의 “연관검색어 및 자동완성검색어에 대한 정책결정 15호”를 참고하여 도출되었다.

넷째, 비표준어는 표준 국어법을 준수하지 않는 검색어를 의미하며, 띄어쓰기 오류, 오타, 준

말이나 축약어 사용, 비속어나 외계어 사용 등의 유형으로 세분화하여 분석하였다.

마지막으로, 연관 검색어의 유형을 질의와의 관련성 측면에서 분석하였다. 이용자가 특정 연관 검색어를 클릭 시, 검색 엔진은 이 연관 검색어를 새로운 질의로 취급하여 재검색을 수행하게 된다. 즉, 검색 포털이 추천한 연관 검색어는 재검색 질의의 역할을 수행하게 되는데, 선행 연구에 따르면, 재검색 질의의 세부 유형으로는 검색어 추가 질의, 검색어 삭제 질의, 검색어 추가 및 삭제 질의, 동일 질의, 이전 질의와 중복된 검색어를 포함하지 않는 변경 질의를 들 수 있다(이준호, 박소연, 권혁성, 2003; Silverstein et al., 1999; Spink et al., 2001). 이 연구에서는 이러한 재검색 질의의 세부 유형을 참고하여 연관 검색어를 분석하였다.

이 연구에서는 이처럼 도출된 평가 기준에 대한 상세한 가이드라인을 작성하였으며, 문헌정보학과 전공자들로 구성된 세 명의 평가자들이 이러한 기준에 근거하여 연관 검색어 평가 작업을 2013년 1월 한 달 동안 수행하였다. 세 명의 평가자들은 나이, 성별, 전공 등이 모두 같은 동질적인 그룹이며, 한 달 이상 연구자로부터 연관 검색어 분석에 관한 교육을 받고, 파일럿 스터디를 수행하였기 때문에, 분석에 있어 평가자들의 주관이 개입할 여지는 매우 적다고 할 수 있다. 평가자들 간의 평가의 일치성은 95% 이상으로 매우 높았으며, 평가가 불일치하는 경우 재검토와 토론을 통하여 합의에 이르는 과정을 거쳤다. 평가자들은 동일한 질의와 연관 검색어에 대한 평가를 수행하였으며, 학술적인 질의의 연관 검색어 평가에 있어서는 각 분야 전문가인 교수들의 자문을 제공받았다. 연구의 일관성을 위

하여 동일한 질의를 네이버와 구글에 동시에 입력 후 노출되는 연관 검색어 결과에 대해 평가를 수행하였다. 본 연구에 수집된 데이터에 대한 기술 통계 및 추론 통계 분석을 위하여 IBM SPSS Statistics version 19가 활용되었다.

## 4. 검색어 추천 서비스 현황

국내의 검색 포털들이 도입하고 있는 검색어 추천 서비스의 예로는 Google의 “Search Related to”, Yahoo의 “Also try”, Bing의 “Related searches”, 네이버의 “연관 검색어”, 다음의 “관련”, 구글 코리아의 “관련 검색”, 네이버의 “관련 검색어” 등을 들 수 있다.

검색 포털들은 크게 세 가지 방식으로 연관 검색어를 추출하는데(Chen, 2011; Liu et al., 2011; NHN, 2013), 첫째, 한 세션 내에서 이용자가 특정 질의 입력 후 연속하여 입력하는 질의들을 연관 검색어들로 추출한다. 둘째, 특정 질의에 대한 검색 결과 중 이용자들이 많이 조회한 결과의 콘텐츠로부터 연관 검색어들을 추출한다. 셋째, 특정한 질의에 대한 백과사전이나, 블로그, 카페, 지식 문서 등의 콘텐츠로부터 연관 검색어들을 추출한다.

## 5. 연구 결과

### 5.1 연관 검색어 제공 현황

〈표 1〉은 이 연구에 사용된 질의에 대한 네이버와 구글의 연관 검색어 제공 현황을 보여준다.

〈표 1〉 포털별 연관 검색어 제공 현황

	구글	네이버
연관 검색어 제공 질의 수	380	387
연관 검색어 개수 평균	7.14	17
더보기 전		11.39
더보기 후		5.61
연관 검색어 개수 최빈치	10	20
더보기 전		11
더보기 후		0
연관 검색어 개수의 범위	0-10	0-20

본 연구에 사용된 400개의 질의 중 네이버에서는 총 387개의 질의에 대해 연관 검색어가 제공되었으며, 구글에서는 총 380개의 질의에 대해 연관 검색어가 제공되었다. 연관 검색어가 제공된 질의들에 대해 네이버에서는 평균 17개의 연관 검색어가, 구글에서는 평균 7.14개의 연관 검색어가 제공되어 네이버에서 구글보다 훨씬 많은 수의 연관 검색어를 제공하고 있었다. 네이버에서는 전체 질의 중 66.7%인 268개가 총 20개의 연관 검색어를 제공하고 있었고, 구글에서는 전체의 55%가 10개의 연관 검색어를 제공하고 있었다.

연관 검색어가 제공되지 않는 질의를 세부적으로 살펴 본 결과, 네이버에서는 영어 질의와 학문적인 질의에 대해 연관 검색어가 제공되지 않는 것으로 나타났다. 반면, 구글에서는 연관 검색어가 0건인 질의들이 대부분 한글 질의였으며, 주제별로 고루 분포되어 이들 간에 특별한 공통점이 발견되지 않았다. 또한 네이버의 경우, 대중적인 질의의 연관 검색어 개수 평균이 18.03개로 학문적인 질의의 연관 검색어 개수 평균인 15.59보다 높았으며, 독립 표본 t-test를 적용한 결과, 이 차이는 통계적으로 유의한

것으로 나타났다( $t(393) = 4.488, p(0.001)$ ). 반면, 구글의 경우, 대중적인 질의의 연관 검색어 개수 평균이 7.34개로 학문적인 질의의 연관 검색어 개수 평균인 6.87보다 약간 높았으나, 이 차이는 통계적으로 유의하지 않은 것으로 나타났다.

3장의 연구 방법에서 제시된 평가 기준에 따라 평가된 구글과 네이버의 연관 검색어의 적합도 평가 결과는 〈표 2〉와 같다.

〈표 2〉 적합도 평가 결과

	구글		네이버	
	빈도	%	빈도	%
적합 연관 검색어	2,322	81.4	5,698	83.5
부적합 연관 검색어	532	18.6	1,122	16.5
총 계	2,854		6,820	

〈표 2〉에 따르면, 네이버와 구글 모두 적합한 연관 검색어가 부적합한 연관 검색어보다 많고, 네이버가 구글보다 적합도가 높은 검색어의 비중이 다소 높은 것으로 나타났다. 대중적인 질의와 전문적인 질의 간의 적합도 분포의 차이를 분석하기 위하여 카이 제곱 검정을 적용한 결



과, 네이버에서 대중적인 질의와 전문적인 질의 간에는 검색어의 적합도 분포에 있어서 통계적으로 유의한 차이가 있었다( $\chi^2(1, N=6820) = 10.385, p<0.01$ ). 즉 대중적인 질의의 경우 적합 검색어의 비중이 84.8%로, 전문적인 질의의 적합 검색어 비중인 81.9%보다 높았으며, 이 차이는 통계적으로 유의미한 것으로 나타났다. 반면, 구글에서는 대중적인 질의와 전문적인 질의 간에 검색어의 적합도 분포에 있어서 유의한 차이가 없는 것으로 나타났다. 네이버에서는 연관 검색어를 더보기 전과 더보기 후로 구분하여 제공하는데, 더보기 전의 적합 검색어의 비중이 84.5%, 더보기 후의 적합 검색어의 비중은 81.6%로, 더보기 전과 더보기 후의 검색어의 적합도 분포에 있어서 통계적으로 유의한 차이가 있는 것으로 나타났다( $\chi^2(1, N=6820) = 9.281, p<0.01$ ).

### 5.2 연관 검색어 유형

네이버와 구글의 연관 검색어를 3장에서 논의된 재검색 질의 유형에 따라 분석한 결과는 <표 3>과 같다.

첫째, 연관 검색어들 중 가장 빈번히 등장하는 유형은 이용자가 입력한 질의에 검색어가 추

가된 경우로 네이버에서는 전체의 54%인 3,684개, 구글의 경우 전체의 78.2%인 2,232개로 나타났다. 예를 들어, 네이버에서 “국회도서관” 질의의 연관 검색어로 “국회도서관 이용”, “국회도서관 논문”, “국회도서관 대출”이 “공정 무역”의 연관 검색어로 “미국 공정무역”, “프랑스 공정무역”이 “2013년 공휴일”의 연관 검색어로 “2013년 공휴일 수”가 구글에서 “거식증”의 연관 검색어로 “거식증 증상”이 “날씨”의 연관 검색어로 “내일 날씨”가 제공된 경우를 들 수 있다.

둘째, 질의에서 검색어가 삭제된 경우가 네이버에서는 전체의 2.8%인 193개, 구글의 경우 전체의 6.1%인 173개로 나타났다. 예를 들어, “빅맥 지수”의 네이버 연관 검색어로 “빅맥”이 “고흐 전시회”의 네이버 연관 검색어로 “고흐”가 “여자 연예인”의 연관 검색어로 “연예인”이 “착한 사마리아인 법”의 구글 연관 검색어로 “사마리아인 법”, “착한 사마리아인”, “사마리아인”이 제공된 경우를 들 수 있다.

셋째, 이용자의 질의와 연관 검색어가 동일한 경우가 네이버에서는 전체 연관 검색어의 1.9%, 구글에서는 2.7%로 나타났다. 질의와 동일한 연관 검색어는 질의에 포함된 검색어들만으로 구성된 연관 검색어를 의미하며, 검색어 순서가

<표 3> 연관 검색어 유형

	구글		네이버	
	빈도	%	빈도	%
이용자 질의에 검색어가 추가됨	2,232	78.2	3,684	54
질의에서 검색어가 삭제됨	173	6.1	193	2.8
질의와 동일함	76	2.7	132	1.9
질의와 동의어임	50	1.8	114	1.7
질의를 완전히 변경함	323	11.3	2,697	39.5

변경된 경우, 불용어 제거 후 이전 질의와 동일한 검색어들을 포함하는 경우를 의미한다. 예를 들어, “재정 절벽” 질의의 네이버 연관 검색어로 “재정절벽이란”, “공정무역”의 네이버 연관 검색어로 “공정무역이란”, “착한 사마리아인 법”의 구글 연관 검색어로 “착한 사마리아인의 법”이 제공된 경우를 들 수 있다. 특히 특정 개념이 질의로 사용된 경우, 질의 뒤에 “이란” 또는 “란”과 같이 조사만 추가된 무성의한 연관 검색어가 이에 해당한다. 또한 이용자 질의의 동의어가 연관 검색어로 제공된 경우가 네이버에서는 전체의 1.7%, 구글에서는 1.8%로 나타났다. 네이버에서 “중국인”의 연관 검색어로 “중국 사람”이 “크리스마스”의 연관 검색어로 “성탄절”이 구글에서 “오바마”의 연관 검색어로 “버락 오바마”가 “수능”의 연관 검색어로 “수학능력시험”이 제공된 경우가 이에 해당한다.

마지막으로, 이용자의 질의와 중복되지 않은 새로운 연관 검색어가 제공되는 경우는 네이버에서는 전체의 39.5%였고, 구글에서는 11.3%에 불과하였다.

또한, 연관 검색어에 대한 내용 분석 결과, 구글과 네이버 모두 연관 검색어들 간에 상당히 많은 중복이 발생하였으며, 전체 연관 검색어들 중 상호 중복되는 검색어들은 네이버의 경우 전체의 32%인 2,192개, 구글의 경우 전체의 33%인 948개로 나타났다. 이러한 검색어들의 예로는 “국회도서관”의 네이버 연관 검색어인 “riss”, “riss4u”, “리스포유”처럼 동의어가 반복되는 경우를 들 수 있다. 좀 더 구체적으로, 첫째, 연관 검색어들의 형태가 동일한 경우가 네이버는 전체 검색어의 4.3%인 294개, 구글은 3.4%인

97개로 나타났다. 조사 등 불용어 제거 후 동일한 검색어가 제공되는 경우, 검색어의 순서만 변경된 경우, 동일한 검색어를 한글과 영어로 표기한 경우, 동일한 외래어에 대한 표기 방식의 차이가 있는 경우 등이 이에 해당한다. 둘째, 연관 검색어의 내용이 유사한 경우가 네이버의 경우 전체의 12.7%인 863개, 구글의 경우 전체의 17%인 492개였으며, 이러한 검색어들의 예로는 “홍루몽”의 네이버 연관 검색어인 “홍루몽 내용”, “홍루몽 줄거리”, “요요 현상”의 네이버 검색어인 “요요 현상 예방법”, “요요 현상 막는 법”을 들 수 있다. 셋째, 날짜나 국가명과 같은 일정한 패턴이 반복되는 검색어들이 네이버는 전체의 15.2%인 1,035개, 구글은 전체의 12.6%인 359개였으며, “아시아 시리즈”의 네이버 연관 검색어인 “2009 아시아 시리즈”, “2008 아시아 시리즈”, “2007 아시아 시리즈”, “수능 예상 등급컷”의 구글 연관 검색어인 “2006 수능 등급컷”, “2007 수능 등급컷”, “2008 수능 예상 등급컷”, “2009 수능 예상 등급컷” 등을 예로 들 수 있다.

이처럼 네이버와 구글의 연관 검색어를 상세히 살펴보면, 질의와 밀접한 관련성이 있는 새로운 연관 검색어를 제시하기보다는 질의에 단어를 추가 또는 삭제하거나, 질의와 동일한 검색어나 동의어 검색어를 제공하는 등 질의를 약간 변형한 경우가 대부분인 것을 알 수 있다. 이러한 현상은 구글에서 더 심각한 것으로 나타났다. 또한 연관 검색어들 간의 관계에 있어서는 연관 검색어가 상호 중복되는 경우가 많아, 실질적으로 이용자의 질의 확장이나 변경에 기여하는 바가 크지 않을 것으로 추정된다. 상호 중복되는 연관 검색어들이 제공되거나 질의와 동

일하거나 유사한(중복되는) 연관 검색어가 제시되는 현상은 연관 검색어의 서비스의 품질을 저하시킬 수 있으므로, 이에 대한 개선이 요청된다.

### 5.3 최신성 평가

구글과 네이버의 최신성 평가 결과, 최신성이 명백히 결여된 연관 검색어의 비중이 각각 전체의 3.7%와 1.2%로 나타났다. 구글에서는 특정 질의에 대해 연도별 일정이 연관 검색어로 제공되는 경우가 많았는데, 예를 들어, “한국 시리즈” 질의에 대한 가장 최신 일정으로 “2009 한국 시리즈 일정”이, “토익 시험”에 대한 최신 일정으로 “2009 토익 시험 일정”이, “수능 예상 등급 컷” 질의에 대한 최신 일정으로 “2006 수능 등급컷”, “2007 수능 등급컷”, “2008 수능 등급컷”, “2009 수능 등급컷”이 노출되었다. 또한 “국회 도서관” 질의에 대해서는 이미 서비스가 종료된 “엠포스 국회도서관”이, “박스 오피스 순위” 질의에 대해서 “2009 박스 오피스”가 연관 검색어로 노출되었다. 이를 통해, 구글에서는 상당수의 연관 검색어들에 대해 2009년 이후 업데이트가 수행되지 않고 있는 것으로 추정된다. 이 외에도 구글은 네이버에 비해 내용 면에서 최신성이 떨어지는 연관 검색어가 많이 등장하였다.

예를 들어, 하나의 질의가 두 개 이상의 뜻을 가진(복수 개의 주제를 포괄하는) 다의어(多義語)인 경우에, 구글에서는 최신 주제가 반영되지 못한 연관 검색어들이 제공되고 있었다. 예를 들어, “싸이”라는 질의에 대해 가수 “싸이”와 관련된 연관 검색어는 전혀 제공되지 않

고, “싸이 월드”에 대한 연관 검색어만 제공되고 있었고, “브라우니”라는 질의에 대해서도 개그 콘서트의 브라우니에 대한 연관 검색어는 제공되지 않고, 브라우니 케이크와 관련된 연관 검색어만 제공되고 있었다. 반면, 네이버에서는 다의어 질의에 대해 가장 최신 주제만 연관 검색어로 부각되는 경향이 있었다.

### 5.4 유용한 연관 검색어와 내용상 오류가 포함된 연관 검색어

연관 검색어의 내용을 상세히 분석한 결과, 이용자에게 도움이 될 수 있는 유용한 연관 검색어들의 사례를 네이버에서 주로 발견할 수 있었다. 첫째 유형으로 질의가 전문 용어인 경우, 용어의 개념이나 정의를 제공하는 연관 검색어들을 들 수 있다. 예를 들어, 네이버의 경우, “중도가자”의 연관 검색어로 “세계 최고 금속 활자”, “가장 오래된 금속 활자”, “RFID”의 연관 검색어로 “전파를 이용해 먼 거리에서 정보를 인식하는 기술로 태그와 판독기가 필요”, “MFCS”의 연관 검색어로 “적하 목록 취합 시스템” 등이 제공되고 있었다. 둘째 유형으로 질의에 대한 답변을 제공하는 연관 검색어를 들 수 있으며, 예를 들어, 네이버에서 “오늘의 작가”에 대한 연관 검색어로 이 상의 수상자인 “정연두”, “장연순”, “서용선”, “윤광조” 등이 제공되고 있었다. 마지막으로 질의와 관련된 분야나 개념, 인물을 소개하는 유형으로, 네이버에서 “슈뢰딩거의 고양이”의 연관 검색어로 “양자역학”, “최수의 딜레마”의 연관 검색어로 “게임 이론”, “아도르노”의 연관 검색어로 “프랑크푸르트 학파”, “바움 가르트너”의 연관 검색어

로 “성층권 스카이다이빙”이 제공된 예를 들 수 있다. 이러한 연관 검색어가 제공됨으로써, 이용자가 질의의 의미나 개념을 찾기 위한 별도의 검색을 수행할 필요가 없어지므로 효율적이라고 볼 수 있다.

반면, 잘못된 정보를 전달하여 이용자를 오도할 수 있는 연관 검색어도 존재하였다. 직지심경은 1972년에 공식적으로 발견되었는데, 네이버에서 “직지심경”의 연관 검색어로 “1974년에 발견된 세계 최초의 금속 활자”가 노출되었다. 또한 싸이 강남스타일은 빌보드 1위에 등극하지 못하였는데 네이버에서 “싸이 강남스타일”의 연관 검색어로 “싸이 강남스타일 빌보드 1위”가 노출되었다. 구글의 경우, “사이코패스” 질의에 대한 연관 검색어로 사이코패스와 무관한 인물들의 이름이 노출되는 경우가 있었다. 이 밖에도 존재하지 않는 연관 검색어의 예로, “Sherlock”의 네이버 연관 검색어인 “sherlock4”, “열정락서”의 연관 검색어인 “청춘락서”를 들 수 있고, 사실이 아닌 연관 검색어의 예로 “조석 마음의 소리”의 연관 검색어인 “조석 수능만점”을 들 수 있다. 또한 암, 정신분열증, 거식증과 같은 질병 명에 그 질병과 무관한 연예인 이름이 연관 검색어로 제시되는 경우도 있었으며, 이는 특정 개인에 대한 명예 훼손에 해당된다고 할 수 있다.

### 5.5 유해 검색어의 유형 및 특징

유해 검색어란 정상적인 검색어가 아니며, 음란성, 저작권 침해성, 비방성, 광고성 검색어 등과 같이 연관 검색어로 수집, 제공되기에 문제가 있는 검색어들을 의미한다. <표 4>는 네이버

와 구글의 연관 검색어들에 포함된 유해 검색어의 유형 및 분포를 보여준다.

<표 4> 유해 검색어의 유형 별 분포

유해 검색어 유형	구글		네이버	
	빈도	%	빈도	%
광고성	121	4.2	236	3.5
저작권 침해성	109	3.8	61	0.9
선정성	45	1.6	47	0.7
비방성	8	0.3	37	0.5
총계	283	9.9	381	5.6

네이버와 구글의 전체 연관 검색어들 중 유해 검색어는 각각 5.6%, 9.9%로 나타났다. 유해 검색어 중 가장 큰 비중을 차지하는 유형은 네이버와 구글 모두 광고성 검색어였다. 음식명, 제품명과 같이 광고가 예상될 수 있는 질의뿐만 아니라, 광고와 무관한 인명, 지명 관련 질의에도 광고성 검색어가 등장하였다. 전자의 예로는 “감자탕”, “발열 내의”, “케이프 코트”에 대한 광고성 검색어들을 들 수 있으며, 후자의 예로는 “카페라떼 효과”의 구글 검색어인 “매일 카페라떼”, “미학”의 구글 검색어인 “미학 스튜디오”, “이효리” 질의의 네이버 검색어인 “샤넬 점보백”, “부산”의 네이버 검색어인 “W모텔,” 안구 건조증, 폐암과 같은 질병명에 대한 연관 검색어로 특정 병원명이 노출된 경우 등을 들 수 있다. 구글의 경우 광고성 검색어가 등장하는 질의의 주제 범위가 네이버보다 더 광범위하였다.

둘째, 드라마나 영화, 컴퓨터 소프트웨어 검색 시, 저작권 침해의 소지가 큰 “토렌트”, “온디스크”, “파일조”와 같은 불법 다운로드 사이트나 “보기”, “다운로드”, “... crack codec”과

같은 저작권 침해성 검색어도 연관 검색어로 자주 노출되었다. 셋째, 네이버와 구글에서 모두 선정적인 음란성 연관 검색어의 문제가 심각한 것으로 나타났다. 네이버에서 인명 검색, 특히 여자 인명 검색 시, 노출 또는 신체 일부와 관련된 연관 검색어가 상시적으로 등장하였다. 심지어 네이버에서 선정성과 무관한 “구글” 검색 시에도, “○○ 노출”, “○○○ 합성 사진”처럼 질의의와는 무관한 선정적인 연관 검색어가 등장하였다. 구글의 경우 음란성 검색어가 등장하는 질의의 범위가 네이버보다 광범위하여서 인명 뿐만 아니라, 일본, 필리핀과 같은 지명 검색 시에도 음란성 검색어가 등장하였다. 마지막으로, 특정 대상을 비하하는 비방성 검색어의 예로는 네이버에서 “일본” 검색 시 노출되는 “일본 지진 축하”, “중국인” 검색 시 노출되는 “중국 인육 상설시장”, “중국인 인육사건”, “중국 인신 매매” 등을 들 수 있다.

한편 네이버에서는 연관 검색어 삭제 기준을 설정하여(NHN, 2013), 연관 검색어가 개인 정보를 노출하는 경우, 사생활 침해나 명예 훼손에 해당하는 경우, 저작권을 명백히 침해하는 경우, 불법 정보나 선정적인 정보를 노출하는 경우, 오타, 욕설을 포함하는 경우, 상업적인 용도 등으로 남용되는 경우 등에는 해당 검색어를 삭제 또는 제한할 수 있다고 명시하고 있다. 그럼에도 불구하고, 이처럼 유해 검색어가 빈번하게 등장하는 것은 네이버에서 연관 검색어 품질 관리 및 필터링을 체계적으로 수행하지 못하고 있음을 시사한다. 구글의 경우 연관 검색어에 대한 규정이나 기준 자체가 공표되어 있지 않기 때문에, 유해 검색어의 문제가 더 심각하다고 할 수 있다.

유해 검색어의 문제점으로 특정 개인이나 집단, 국가에 대한 사생활 침해나 명예 훼손, 저작권 침해 목적의 불법 정보 유통 등을 들 수 있다. 검색 결과에 유해 검색어가 포함될 경우 검색의 성능이 저하되고 이용자의 만족도가 감소될 수 있다. 따라서 향후 연관 검색어에 대한 체계적인 품질 관리를 통해 유해 검색어 노출을 예방할 수 있는 방안이 모색되어야 할 것이다. 특히 네이버는 광고성 검색어에 대한, 구글은 광고성 및 저작권 침해성 검색어에 대한 방안을 강구해야 할 것으로 보인다.

### 5.6 비표준어의 유형 및 특징

비표준어란 표준 국어법을 준수하지 않는 검색어들을 의미하며, 그 세부 유형으로는 띄어쓰기 오류, 오타, 준말이나 축약어 사용, 비속어나 외계어 사용 등을 들 수 있다. <표 5>는 네이버와 구글의 연관 검색어들에 포함된 비표준어의 유형 및 분포를 보여준다.

<표 5> 비표준어의 유형 별 분포

비표준어 유형	구글		네이버	
	빈도	%	빈도	%
띄어쓰기 오류	1,110	38.9	3,158	46.3
오타	9	0.3	21	0.3
준말, 축약어	3	0.1	26	0.4
속어, 외계어	11	0.5	12	0.2
총계	1,133	39.8	3,217	47.2

표준 국어법과 관련하여 가장 흔히 발견되는 오류는 띄어쓰기 오류로, 네이버의 경우 전체 연관 검색어 중 46.3%인 3,138개가, 구글의 경우 전체 연관 검색어 중 38.9%인 1,110개가 띄

어쓰기 오류인 것으로 나타났다. 구글의 띄어쓰기 오류의 예로는 “서브프라임모기지론”, “제주도여행싸게가는법”, “운전면허시험관리공단”, “사회적기업지원센터”, “나눔로또당첨번호확인.” 네이버의 예로는 “주택금융공사모기지론”, “IBK 월복리자유적금”, “누가말을걸어도뒤돌아보지마세요”, “windows7tablet”, “ktttablet”, “delltablet”, “hpttablet” 등을 들 수 있다. 둘째, 네이버의 “caffé”, 구글의 “짜이월드”, “ㅋ코레일”, “다음음카페”, “다음카페”, “psy d”와 같이 명백한 오타도 발견되었으며, 이러한 오타의 비중은 네이버와 구글에서 모두 전체 연관 검색어의 0.3%로 나타났다.

### 5.7 영어 질의의 연관 검색어

네이버에서 영어 질의의 경우 연관 검색어가 제공 안 되는 경우가 많으며, 연관 검색어가 제공되는 경우에도 부적합한 연관 검색어의 비중이 매우 높았다. 예를 들어, “Steve Jobs”라는 질의에 대한 네이버의 연관 검색어로 “the”, “were”, “was”, “up”, “with”와 “Obama”라는 질의에 대한 네이버의 연관 검색어로 “who”, “too”, “the”, “of course”, “was”, “right”가 제공되는데, 이들은 질의와 무관한 부적합한 연관 검색어들이라고 할 수 있다. 이처럼 영어 질의 검색 시 불용어, 전치사, 관사, 부사 등이 노출되는 것은 이 질의들이 등장하는 문서에서 빈번하게 등장하는 단어들을 검색 엔진이 기계적으로 추출하기 때문인 것으로 보인다.

네이버에서 영어 질의에 대한 부적합 검색어의 비중이 29.7%로, 한글 질의의 부적합 검색어 비중인 14.8%보다 두 배 이상 높았으며, 한글

질의와 영어 질의 간의 검색어 적합도 분포의 차이를 분석하기 위하여 카이 제곱 검정을 적용한 결과, 이 차이는 통계적으로 유의미한 것으로 나타났다( $\chi^2(1, N=6820)=105.274, p<0.001$ ). 반면, 구글에서는 한글 질의와 영어 질의 간에 검색어의 적합도 분포에 있어서 유의한 차이가 없는 것으로 나타났다.

또한 네이버에서는 한글 질의의 연관 검색어 개수 평균이 17.68개로 영어 질의의 연관 검색어 개수 평균인 12.84개보다 현저히 높았으며, 이 차이를 비교하기 위하여 독립 표본 t-test를 적용한 결과, 한글 질의와 영어 질의 간에 제공되는 연관 검색어의 수에 있어서 통계적으로 유의한 차이가 있는 것으로 나타났다( $t(398)=6.133, p<0.001$ ). 반면, 구글에서는 한글 질의의 연관 검색어 개수 평균이 7.01개, 영어 질의의 연관 검색어 개수 평균이 7.95개였으며, 이 차이는 통계적으로 유의하지 않은 것으로 나타났다.

## 6. 결론

본 연구에서는 주요 검색 포털들인 네이버와 구글의 연관 검색어 서비스의 특징 및 품질을 조사, 분석하였다. 좀 더 구체적으로 이 연구에서는 구글과 네이버의 연관 검색어의 적합도, 최신성을 평가하고, 연관 검색어의 개수 및 분포, 연관 검색어가 제공되지 않는 질의의 특징을 조사하였다. 둘째, 연관 검색어의 유형을 질의와 연관 검색어의 관계 측면에서 분석하였다. 셋째, 연관 검색어들 중 유해 검색어의 유형 및 특징, 비표준어의 유형 및 특징도 조사하였다. 마지막으로, 한글 질의와 영어 질의들 및 대중

적인 질의와 전문적인 질의의 연관 검색어의 특징을 비교, 분석하였다.

연구 결과, 첫째, 네이버가 구글보다 2배 이상 많은 수의 연관 검색어를 제공하고 있었으며, 네이버가 구글보다 연관 검색어의 적합도와 최신성이 다소 높은 것으로 나타났다. 둘째, 구글과 네이버 모두 새로운 연관 검색어를 제시하기보다는 질의에 단어를 추가 또는 삭제하거나, 질의와 동일한 검색어나 동의어 검색어를 제공하는 경우가 많은 것으로 나타났다. 또한 연관 검색어들이 상호 중복되는 경우가 많아, 실질적으로 이용자의 질의 확장이나 변경에 기여하는 바가 크지 않을 것으로 판단된다. 셋째, 구글에서 네이버보다 유해 검색어의 비중이 더 높은 것으로 나타났으며, 두 포털 모두 유해 검색어 중 가장 큰 비중을 차지하는 유형은 광고성 검색어였으며, 이어서 저작권 침해성, 선정성 검색어 순으로 나타났다. 네이버와 구글에서 인명 검색 시 선정성, 비방성 검색어가 항상 등장하였으며, 특히 여자 인명 검색 시, 선정성 연관 검색어가 상시적으로 등장하였다. 또한 제품명, 음식명과 같은 쇼핑 관련 질의에서는 광고성 검색어가 항상 등장하였다. 넷째, 네이버의 경우, 대중적인 질의의 연관 검색어의 적합도가 학문적인 질의의 적합도보다 높았으며, 한글 질의의 연관 검색어의 적합도가 영어 질의의 적합도보다 현저히 높았으며, 이러한 차이는 통계적으로 유의한 것으로 나타났다. 반면, 구글의 경우, 대중적인 질의와 학문적인 질의 간에, 그리고 한글 질의와 영어 질의 간에 연관 검색어의 적합도의 분포에 있어서 통계적으로 유의한 차이가 없었다. 네이버의 경우, 영어 질의의 검색 시 질의와 무관한 전치사, 관사, 부사 등의

불용어가 주로 노출되었으며, 이로 인해 한글 질의보다 영어 질의에서 부적합한 연관 검색어의 비중이 현저히 높았다. 마지막으로, 표준 국어법과 관련하여 가장 흔한 오류는 띄어쓰기 오류로 네이버의 경우 전체 연관 검색어의 46%, 구글은 전체의 39%로 나타났다.

본 연구의 조사 결과, 검색 포털들의 연관 검색어 서비스에 있어서 개선이 필요한 분야는 다음과 같다. 첫째, 질의와 중복되는 무성의한 연관 검색어를 배제하고, 연관 검색어들 간의 과도한 중복을 제거함으로써 연관 검색어의 효율성을 제고할 필요가 있다. 네이버에서는 현재 연관 검색어를 더보기 전과 더보기 후로 구분하여 제공하는데, 더보기 후에서 제공되는 연관 검색어들이 더보기 전에서 제공되는 연관 검색어보다 적합도가 낮았으며, 이 차이는 통계적으로 유의하였다. 따라서 연관 검색어들이 과도하게 중복된 상태에서 더보기 후에 연관 검색어를 추가하는 방식보다는 중복되는 연관 검색어를 제거하고, 더보기 후를 삭제하여 이용자의 불필요한 클릭을 배제하는 것이 더 효율적일 것으로 판단된다. 둘째, 포털의 영향력을 고려할 때, 이용자를 오도할 수 있는 부정확한 연관 검색어를 철저히 관리할 필요가 있다.

셋째, 광고성, 음란성, 저작권 침해성, 비방성 검색어와 같은 유해 검색어들을 제외할 수 있는 방안이 요구된다. 이 연구에는 전문적인 질의가 많이 포함되어 유해 검색어의 비중이 상대적으로 낮은 것으로 분석된다. 그러나 대부분의 웹 검색이 연예인, 뉴스, 쇼핑 등의 분야에 집중되는 현실을 고려할 때, 실제 웹 검색 환경에서는 유해 검색어의 비중이 이 조사 결과보다 훨씬 높을 것으로 예상된다. 인터넷 쇼핑물의 매출을

높이기 위한 마케팅 방법으로 연관 검색어 활용을 제안하는 단행본이(최재봉, 2012) 출판된 것은 연관 검색어의 상업적 성격에 대한 사회적 인식을 반영한 사례라고 할 수 있다. 넷째, 비표준어에 대한 품질 관리가 요청된다. 특히 구글과 네이버 모두 띄어쓰기 오류가 전체 연관 검색어의 절반 정도를 차지하므로, 이에 대한 정정 작업이 요구된다. 다섯째, 네이버의 경우, 영어 질의의 연관 검색어에 대한 품질 관리가 필요하다. 현재 영어 질의에 대한 연관 검색어에는 질의와 무관한 불용어들이 큰 비중을 차지하고 있어서 적합도를 저하시키는 요인으로 작용하고 있다. 또한 학문적이고 전문적인 질의의 연관 검색어에 대해서도 품질 관리가 요구된다. 마지막으로, 구글은 연관 검색어에 대한 주기적인 업데이트를 통하여 연관 검색어 서비스의 최신성을 도모할 필요가 있는 것으로 보인다.

한편 네이버에서는 이용자에게 도움을 줄 수 있는 유익한 연관 검색어도 일부 제공하고 있었는데, 연관 검색어 서비스의 품질을 향상시키기 위하여 이러한 유형의 연관 검색어를 활성화시키는 것이 바람직할 것으로 보인다.

현재 대부분의 검색 포털들은 자동화된 방식으로 연관 검색어를 추출하고 있는 상황이다. 즉 한 세션 내에서 이용자가 특정 질의와 함께 검색하는 질의들이나, 특정 질의에 대한 검색 결과 중 이용자들이 클릭한 결과 문서로

부터 연관 검색어를 자동으로 추출하는 방식을 취하고 있다. 이 밖에도 특정 질의에 대한 백과사전이나 블로그, 카페, 지식 문서 등으로부터도 연관 검색어를 추출하고 있다. 그러나 이 연구에서 발견된 연관 검색어의 복합적인 문제점들을 해결하기 위해서는 자동화된 방식과 더불어 수작업으로 연관 검색어를 추출, 관리하는 방안이 고려되어야 할 것으로 보인다. 특히, 전문적이고 학문적인 성격의 질의에 대해서는 전문가의 자문을 구하는 방안도 고려될 수 있을 것이다.

향후 유무선 상의 웹 검색의 이용은 꾸준히 증가할 것으로 예측되므로, 연관 검색에 관한 지속적인 연구가 요청된다. 첫째, 로그 분석을 통하여 이용자들의 연관 검색어 이용 행태를 분석하고, 이를 연관 검색 서비스 및 인터페이스 개선에 반영하는 것이 필요할 것이다. 또한 심층 면접이나 설문 조사를 통하여 연관 검색어에 대한 이용자 만족도 및 개선이 필요한 사항을 조사하여 연관 검색어 서비스 개선에 반영할 필요가 있을 것이다. 한편 이 연구에서는 대학생들의 관심 질의들에 대한 연관 검색어들을 대상으로 분석과 평가를 수행하였다. 향후 연구에서는 성격이 다른 이용자 집단의 질의들 및 실제로 이용자들이 주로 검색하는 질의들에 대한 평가 작업도 필요할 것으로 보인다. 마지막으로, 후속 연구에서는 본 연구에서 제시한 방법론에 대한 검증 및 보완 작업이 요청된다.



## 참 고 문 헌

- 김지혜, 장재영, 윤홍준, 김한준 (2010). 키워드 관련도를 이용한 뉴스기사의 연관 검색 기법. 한국정보과학회 학술발표논문집, 37(1C), 53-57.
- 문상준, 최재걸 (2004). 검색어의 연관법칙. 한국정보과학회 학술발표논문집, 31(2), 10-12.
- 박성주, 박하연, 김용혁 (2008). 검색엔진의 검색 결과 분석을 통한 병렬처리 기반의 연관 검색어 추천 기법. 한국정보과학회 학술발표논문집, 35(2C), 214-217.
- 박소연 (2011). 웹 검색 행태의 추이 및 변화 분석. 한국문헌정보학회지, 45(1), 377-393.  
<http://dx.doi.org/10.4275/KSLIS.2011.45.1.377>
- 박소연, 이준호, 김지승 (2005). 클릭 로그에 근거한 네이버 검색 질의의 형태 및 주제 분석. 한국문헌정보학회지, 39(1), 265-278.
- 윤여생, 유진호 (2012). 불법유해정보 법·제도 동향 분석. 정보보호학회지, 22(3), 25-36.
- 이정훈, 전서현 (2009). 검색어 질의 확장을 위한 추천 질의 랭킹 시스템. 한국정보과학회 학술발표논문집, 36(2C), 210-214.
- 이준호, 박소연, 권혁성. (2003). 질의 로그 분석을 통한 네이버 이용자의 검색 행태 연구. 정보관리학회지, 20(2), 27-41.
- 주길홍, 이주일, 이원석 (2011). 효율적인 문서 검색을 위한 연관 키워드 추출 및 확산 클러스터링 방법. 한국정보기술학회 논문지, 9(6), 155-166.
- 최재봉 (2012). 인터넷 쇼핑몰 마케팅 비법. 서울: 애플북스.
- 한국인터넷자율정책기구(KISO) (2013). 연관검색어 및 자동완성검색어에 대한 정책결정 15호.  
 Retrieved from [http://www.kiso.or.kr/decision/decision\\_view.html](http://www.kiso.or.kr/decision/decision_view.html).
- Arkin, H., & Colton, R. (1963). Tables for statisticians. New York: Barnes & Noble Inc.
- Bates, M. J. (1981). Search techniques. Annual Review of Information Science and Technology, 16, 139-169.
- Beeferman, D., & Berger, A. (2000). Agglomerative clustering of a search engine query log. Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 407-416. <http://dx.doi.org/10.1145/347090.347176>
- Chen, L. (2011). Term suggestion with similarity measure based on semantic analysis techniques in query logs. Online Information Review, 35(1), 9-33.
- Croft, W. B., Metzler, D., & Strohman, T. (2010). Search engines: Information retrieval in practice. Boston: Pearson Education. 임해창, 임희석, 한경수, 박소영 공역 (2012). 검색엔진: 최신정보검색론. 서울: 휴먼사이언스.

- Huang, C., Chien, L., & Oyang, Y. (2003). Relevant term suggestion in interactive web search based on contextual information in query session logs. *Journal of the American Society for Information Science and Technology*, 5(7), 638-649. <http://dx.doi.org/10.1002/asi.10256>
- Liu, Y., Miao, J., Zhang, M., Ma, S., & Ru, L. (2011). How do users describe their information need: Query recommendation based on snippet click model. *Expert systems with applications*, 38, 13847-13856. <http://dx.doi.org/10.1016/j.eswa.2011.04.188>
- NHN. (2013). 연관검색어는 어떤 서비스인가요? Retrieved from <http://help.naver.com/ops/faq/faqPrint.nhn?faqId=24796>.
- Sanderson, M. (2008). Ambiguous queries: Test collections need more sense. *Proceedings of the 31st Annual international ACM SIGIR conference on research and development in information retrieval*, 499-506. <http://dx.doi.org/10.1145/1390334.1390420>
- Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1), 6-12.
- Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the Web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 226-234.
- Xu, Z., Luo, X., Yu, J., & Xu, W. (2011). Mining Web search engines for query suggestion. *Concurrency and Computation: Practice and Experience*, 23, 1101-1113. <http://dx.doi.org/10.1002/cpe.1689>

<p>• 국문 참고문헌에 대한 영문 표기 (English translation of references written in Korean)</p>
--

- Choi, Jae Bong (2012). *Marketing secrets for Internet shopping mall*. Seoul: AppBooks.
- Joo, Kil-Hong, Lee, Joo-Il, & Lee, Won-Suk (2011). An associated keywords extraction and a spread clustering methods for an efficient document searching. *Journal of Korean Institute of Information Technology*, 9(6), 155-166.
- Kim, Ji-hye, Jang, Jae-Young, Yune, Hongjune, & Kim, Han-joon (2010). A relationship search in news articles using a keyword association frequency. *Proceedings of the Korea Computer Congress 2010*, 37(1C), 53-57.
- Lee, Joon-Ho, Park, Soyeon, & Kwon, Hyuk-Sung (2003). Information seeking behavior of the NAVER users via query log analysis. *Journal of the Korean Society for Information Management*, 20(2), 27-41.

- Lee, Jung-Hun, & Cheon, Suh H. (2009). Recommendation query ranking system for the search query expansion. *Proceedings of the 36th KIISE Fall Conference*, 36(2C), 210-214.
- Mun, Sang-June, & Choi, Jonathan (2004). ARMS: Association rule for small set. *Proceedings of the 31th KIISE Fall Conference*, 31(2), 10-12.
- NHN. (2013). About related search term service. Retrieved from <http://help.naver.com/ops/faq/faqPrint.nhn?faqId=24796>.
- Park, Soyeon (2011). Trends and changes of Web searching behavior. *Journal of the Korean Society for Library and Information Science*, 45(1), 377-393.  
<http://dx.doi.org/10.4275/KSLIS.2011.45.1.377>
- Park, Soyeon, Lee, Joon-Ho, & Kim, Ji Seoung (2005). An analysis of query types and topics submitted to Naver. *Journal of the Korean Society for Library and Information Science*, 39(1), 265-278.
- Park, Sung-Ju, Park, Ha-yan, & Kim, Yong-Hyuk (2008). A parallel processing method of relative keyword recommendation through analyzing search results. *Proceedings of the 35th KIISE Fall Conference*, 35(2C), 214-217.
- Yun, Yeo Saeng, & Yu, Jin Ho (2012). Recend trends on laws and policies regarding illegal & harmful information. *Journal of the Korea Institute of Information Security and Criptology*, 22(3), 25-36.