

# 모두가 아는 검색, 아무도 모르는 검색 알고리즘

– 검색서비스 중립성을 둘러싼 이론과 논점

강정수 · 디지털사회연구소 소장, 김수철 · 한양대학교 평화연구소 연구교수

## 검색엔진에 대한 다양한 시각들

검색엔진이란 어떤 기술인가? 검색엔진은 어떻게 작동하는가? 검색서비스 중립성은 왜 논란이 되는가? 검색엔진의 작동 원리에 대해서는 최근 많은 학계 연구자들과 기술 전문가, 시장 행위자들의 논의를 통해 많은 부분 밝혀진 상태다. 전부는 아니지만 구글이나 네이버 등의 검색 알고리즘도 일부 공개되어 공공의 지식이 되었다.

검색서비스는 크게 3가지로 구별할 수 있다.<sup>1)</sup> 첫째는 색인에 기초한 검색서비스다. 색인 검색서비스는 크롤러(crawler), 로봇(robots) 또는 스파이더(spider)로 불리는 자동화 프로그램을 통해 인터넷을 조사한다. 발견된 웹페이지는 검색서비스의 데이터베이스에 간략한 형태로 저장되며 이후 지속적으로 업데이트된다. 이 데이터베이스에 검색질의(search query)가 던져지면, 검색 알고리즘은 정해진 매개변수에 따라 검색질의에 대한 결과의 순서를 정한다. 검색결과는 개별결과를 목록형식으로 노출한다. 이때 개별결과는 하이퍼링크 값을 가지며 개별결과에 대한 간략한 설명을 포함한다. 두 번째는 이른바 웹카탈로그라 불리는 것으로 색인화 작업과 편집 과정을 거치는 검색서비스다. 웹페이지 수집과 저장은 컴퓨터 프로그램을 통한 자동화 과정을 통해 이루어지나, 검색질의에 대한 결과 목록은 검색서비스 사업자의 편집 행위를 통해 생산된다. 여기에는 인간 지능의 도움을 받은 분류가 검색결과의 정확도를 높일 수 있다는 가정이 전제되어 있다. 마지막으로 메타 검색서비스는 복수의 색인에 기초한 검색서비스에 검색질의를 보내 각 검색서비스로부터 받은 검색결과를 종합하여 빈도 등 정해진 매개변수에 따라 검색결과 목록을 작성한다.

1) Levene, M. (2010). *An Introduction to Search Engines and Web Navigation* (2nd ed). Hoboken, New Jersey: John Wiley & Sons. p. 150.

색인에 기초한 검색서비스가 작동하는 방식은 다시 세 단계로 구별할 수 있다.<sup>2)</sup> 첫째는 크롤링(crawling)이다. 크롤링이란 지속적으로 업데이트되는 웹페이지에 존재하는 수많은 문서나 정보들을 읽어 들이고 다운로드하는 과정을 말한다. 둘째는 색인화(indexing)다. 색인화는 앞의 크롤링을 통해 수집된 수많은 웹페이지들의 데이터베이스를 다양한 범주(카테고리)로 나누어 새로운 데이터를 생성하는 과정을 의미한다. 마지막으로 랭킹(ranking)이다. 랭킹이란 사용자가 검색엔진을 통해 검색을 실행했을 때 검색엔진에 의한 검색결과들의 순위를 매기는 과정을 말한다.

하지만 여기서 유념해야 할 점이 있다. 검색엔진은 사용자가 검색을 실행할 때마다 작동해 실시간으로 웹페이지 정보를 찾아내는 것이 아니라, 고유의 알고리즘을 통해 정기적으로 이미 크롤링되고 색인화된 콘텐츠를 검색한 결과들을 제시한다는 것이다. 즉, 검색결과들이 랜덤하게 제시되는 것이 아니라 특정 검색엔진의 알고리즘에 따라 이루어진다는 사실이다. 웹페이지의 상대적인 노출 정도와 검색결과들의 순서 배치를 결정하는 랭킹의 과정은 검색엔진 설계에 이미 포함되어 있다. 이는 검색엔진의 종류에 따라 검색결과가 제시되는 방식에 차이가 존재한다는 것을 의미한다. 이러한 측면에서 검색엔진이 하는 기본적인 일은 두 가지이다. 우선 관련성(relevance)이 있는 콘텐츠를 구별해내고, 관련성의 정도에 따라 웹페이지 검색결과들의 순서를 결정하는 것이다.

검색엔진 및 검색서비스를 둘러싼 논란들은 검색엔진을 어떤 종류의 기술로 보는지에 따라 다양하게 나타난다.

첫 번째 시각은 검색엔진을 기술 그 자체의 작동으로서 바라보는 것이다. 여기서 기술로 바라본다는 것의 의미는 검색 알고리즘의 작동방식에 집중하여 검색엔진이 매우 기계적



이고 자동화되어 객관적인, 과학적인, 혹은 가치중립적인 과정들로 이루어지는 것으로 보는 것이다. 여기에는 검색엔진을 기존의 미디어와는 구분되는 것으로 이해하려는 인식이 강하게 깔려있다. 그러나 검색엔진 및 검색서비스를 ‘객관적인 자동랭킹 시스템’이라고 볼 수 있을까?

2) Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and Isdn Systems*, 30(1-7), pp.107-117.

만약에 그렇다면 이른바 구글폭탄(Google Bomb)<sup>3)</sup> 현상은 어떻게 설명할 수 있을까? 구글 검색엔진에 최적화된 웹사이트 구축 서비스를 제공하는 수많은 마케팅 전문 회사들이 존재하는 상황에서 검색엔진을 객관적인 자동 랭킹 시스템으로 보는 것이 타당한가?

검색엔진 및 검색 알고리즘 관련 연구들에 따르면, 검색엔진과 검색서비스에서 인간의 지속적인 편집은 불가피한 것으로 나타난다.<sup>4)</sup> 특히 랭킹을 부여하거나 어떤 웹페이지에 가중치를 두거나 하는 편집, 개입 과정이 모두 검색 알고리즘을 통해 이루어진다는 것을 알 수 있다. 구글 검색 알고리즘의 한 부분으로 알려져 있는 페이지랭크(PageRank) 알고리즘이 대표적인데, 구글의 페이지랭크는 링크가 많이 걸려 있는 웹페이지가 더 중요한 페이지라는 대중성 원리에 바탕을 두고 외부로부터 더 많은 링크(inbound link)를 가지고 있는 웹페이지에 가중치를 부여함으로써 웹페이지의 가치를 평가하는 체계로 알려져 있다.<sup>5)</sup> 이러한 구글의 페이지랭크 알고리즘은 가치중립적인 과정이라기보다는 일정한 편향성(bias)을 갖고 있다. 또한 구글의 검색엔진은 더 오래된 사이트들, 상업적 사이트들, 그리고 미국에 기반을 두고 있는 사이트들을 과잉 대표하고 있다는 지적이 제기되어 왔다.<sup>6)</sup>

두 번째 시각은 검색엔진, 보다 정확하게는 검색서비스를 순수하게 하나의 상업적인 성격의 서비스로만 바라보는 시각이다. 이에 따르면 검색서비스를 제공하는 기업을 광고회사로, 이들이 가진 검색엔진은 일종의 광고엔진으로 볼 수 있다. 다만 이들이 만들어내는 상품(검색결과)

3) 구글 검색엔진은 특정 웹페이지의 내용을 수집할 뿐 아니라, 특정 웹페이지로 향하는 링크-이른바 인링크(Inlink)-와 그 링크와 연결된 텍스트를 함께 수집한다. 만약 '청와대'의 특정 웹페이지로 향하는 링크들의 텍스트에 '실패'라는 단어가 반복적으로 등장할 경우, 청와대의 해당 웹페이지에는 '실패'라는 단어가 등장하지 않아도 구글 검색서비스는 '실패'라는 검색어의 결과 값에 청와대의 해당 웹페이지를 보여준다. 1999년 마이크로소프트의 특정 웹페이지가 "more evil than satan himself" 검색어와 연결된 사건. 2003년 당시 미국 대통령 조지 부시가 "miserable failure"와 연결된 사건 등이 '구글 폭탄' 사례다. '구글 폭탄'은 이용자의 활동이 검색결과에 영향을 미칠 수 있음을 보여준 대표적인 근거로 평가받는다[Zeller, T. Jr. (2006, Oct 26). A New Campaign Tactic: Manipulating Google Data, *The New York Times*, URL: [http://www.nytimes.com/2006/10/26/us/politics/26googlebomb.html?\\_r=0](http://www.nytimes.com/2006/10/26/us/politics/26googlebomb.html?_r=0)]

4) Van Couring, E. (2009). Search Engine Bias: the structuration of traffic on the World-Wide Web. Dissertation, *The London School of Economics and Political Science*.

5) 가장 많은 링크, 즉 대중성에 가중치를 부여하는 페이지랭크의 특징은 평판(reputation)의 가치를 중시하는 정동경제(affective economy) 혹은 인자본주의의 특성을 가장 잘 실현한 기술이라는 지적도 있다[이항우 (2014). 구글의 정동경제: 사용자 정동 노동의 동원과 전용. <경제와 사회>, 제102권, 208-236면].

6) 이러한 비판들은 구글 검색엔진에 의해 제시된 콘텐츠에 대한 접근성에 주목하면서 이론과 실제 사이의 간극을 지적한다. 구글 검색서비스의 모든 카테고리에서 소수 거대 사이트들이 링크 패턴을 지배하면서 사용자들 대부분이 그 사이트들로 연결되는 "구글위계체제(Googlearchy)"가 만들어지고 있다는 지적이다.[Hindman, M. Tsioutsouliklis, K. & Johnson, J. A. (2003, July 28). "Googlearchy": How a few heavily-linked sites dominate politics on the web. Retrieved from <http://www.princeton.edu/~mhindman/googlearchy--hindman.pdf>]. 이는 대부분의 사용자들이 실제로 아주 단순하고 그리 세밀하지 않은 검색 단어들을 사용하며 검색결과의 첫 페이지를 넘어서는 잘 들여다보지 않기 때문이다. Inrona, D. L. & Nissenbaum, H. (2000). Shaping the Web: Why the politics of search engines matter, *The information Society*, 16, pp. 169-185.; Lawrence, S., & Giles, C. L. (1999). Accessibility of information on the World Wide Web. *Nature*, 400, pp. 107-109.; Mowshowitz, A. & Kawaguchi, A. (2002). Assessing bias in search engines. *Information Processing & Management*, 38, pp. 141-156.; Vaughan, L. & Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes. *Information Processing & Management*, 40(4), pp. 693-707.

을 사용자들에게 무료로 제공하는 대신, 쿠키를 이용해 사용자의 웹 활동을 추적하여 생성되는 다양한 사용자 정보를 광고주들에게 판매해 이윤을 얻는다는 점에서 일반적인 광고회사와 차이가 있다. 이러한 측면에서 보면 검색엔진 기술은 시장경제의 논리에 기반을 둔 비즈니스의 일부일 뿐이며 검색엔진의 질을 결정하는 가장 중요한 요소는 시장에서의 성공이라는 결론을 내릴 수 있다.

하지만 검색엔진 기술 및 기술의 변화와 발전을 시장의 성공이라는 기준만으로 충분히 설명할 수 있을까? 시장에서의 성공이 만들어내는 검색엔진의 편향성은 없을까? 특히 순수하게 검색엔진에 의해서만 제공된 검색결과 콘텐츠(organic results)와 광고 콘텐츠(스폰서링크, 프로모션링크 등)가 뚜렷하게 구분하지 않아 불공정거래 문제가 제기되고 있는 상황에서 더 많은 고객, 더 높은 매출액, 더 많은 점유율 등 시장의 성공이 검색엔진의 더 높은 품질을 의미한다는 설명에 과연 우리는 얼마나 만족할 수 있을지 의문이다.

세 번째는 검색엔진을 미디어로 바라보는 시각<sup>7)</sup>이다. 이는 TV, 종이신문 대신 인터넷이 정보와 문화 그리고 광고의 주요 원천이 되면서 전통적 미디어에 대한 공공적 규제 기준이 온라인 공간에도 적용될 필요가 있다는 시각과 연결되어 있다. 검색엔진이 지식정보시대에서 핵심적 역할을 차지하게 되었으므로 검색서비스 시장에 대한 공적 규제가 불가피하다는 것이다. 전통적 미디어 연구에서 미디어를 민주주의 작동에 필수적인 요소로 보고, 표현의 자유와 시민의 알 권리 보장을 위해 미디어에 대한 규제가 정당화될 수 있다고 본 것처럼 검색서비스에 대한 규제도 같은 맥락에서 정당화될 수 있다는 것이다.

그러나 검색엔진, 검색서비스를 전통적인 미디어와 동일한 것으로 볼 수 있을까? 민주주의, 시민의 알권리, 공공성, 공정성을 앞세워 검색 알고리즘의 투명성을 강조하고 나아가 구글이나 네이버와 같은 검색서비스의 알고리즘을 공적지식화하는 것이 가능할까? 또 이는 바람직한 것인가?

구글의 영업 비밀을 공적 규제 혹은 지식의 다양성 보장, 검색 독점에 의한 지식 독점에 대한 규제라는 명목으로 완전히 공개할 경우, 다른 수많은 검색서비스들은 시장의 성공을 위해 구글

---

7) 검색서비스 사업자에게 검색순위 알고리즘은 핵심 경쟁요소로, 검색 순위 알고리즘은 그 기본 원칙을 제외하고는 공개되지 않는다. 따라서 이용자 개인은 검색서비스에서 보이는 결과물에 대한 중요도 또는 현실 관련성 등에 대한 평가 또는 피드백을 수행할 수 없고 검색서비스 사업자는 검색 순위 알고리즘에 기초하여 특정 정보 및 콘텐츠에 대한 이용자의 관심을 집중시키거나 회피시킬 수 있다[Moffat, R. V. (2009). *Regulating Search*, *Harvard Journal of Law & Technology*, 22(2), p. 476.] 이와 같은 맥락에서 검색서비스는 이용자에게 정보에 대한 접근권을 제공하고 정보 및 콘텐츠를 매개로 이용자와 공급자를 중재하는 게이트키퍼(gatekeeper)로서 기능한다[앞의 Levene (2010), p. 64; Vogl, P. & Barrett, M. (2010). *Regulating the Information Gatekeepers*, *Communications of the ACM*, 53(11), p. 67]. 결과적으로 검색서비스 사업자는 수집, 평가 및 처리, 가공 그리고 표현 등 일련의 과정을 통해 정보 및 콘텐츠의 이용자와 공급자에 대한 영향력을 확보하게 된다. 여기서 아날로그 시대에서 정보 및 콘텐츠 공급자가 게이트키퍼로서 가졌던 여론 형성의 힘이 검색서비스로 이동하고 있음을 확인할 수 있다. 검색서비스는 디지털사회 또는 정보사회에서 정보중개의 역할뿐 아니라 자유로운 정보교환과 표현의 자유에 '잠재적인' 위협요소로 기능할 수 있다.

의 검색 알고리즘을 따르게 될 가능성이 높아져 검색 다양성이 아닌 검색서비스의 획일화로 나아갈 수 있다는 주장도 있다.<sup>8)</sup> 이처럼 검색서비스를 전통 미디어와 동일시하여 공적 규제의 필요성을 강조하는 시각은 자칫 무분별한 규제와 개입으로 인해 본래의 목적과는 다른 결과를 가져올 가능성이 존재한다.

## 검색엔진은 어떤 기술인가?

검색엔진과 서비스 기술의 민주화는 검색서비스 기술을 작동하게 하는 조건들과 이 기술의 사용방식에 대해 더욱 세밀한 성찰과 토론을 요구한다. 검색엔진에 대한 문제를 검토하는 이론적 분석틀로는 ‘기술의 사회적 구성론’과 ‘기술 실천론’을 들 수 있다.

기술의 사회적 구성론은 우리 주변에 있는 다양한 종류의 기술이 사회적 과정과 사회적 행위를 통해 설명될 수 있다고 보는 관점이다. 즉, 어떻게 기술이 갈등적 상황을 극복하고 닫힘(closure), 또는 안정화(stabilization)를 성취하게 되는지의 과정을 설명하고자 한다. 예를 들어 자전거나 전등과 같은 고정된 기술적 발명품이나 제품이 시장에서 사용(채택) 또는 퇴출을 거쳐, 상업적 성공 혹은 실패로 기술 발전이 마무리되는지의 과정을 설명하는 것이다.

그러나 검색엔진은 정보기술 서비스(Information Technology Service)이자 미디어 기술로서 자전거나 전등 기술의 발전과정과는 다른 시각에서 이해될 필요가 있다. ‘기술 실천론’은 현대 기술들이 인공물이 아닌 IT 서비스와 같은 미디어 기술인 경우가 많다는 특성에 대한 인식에서 출발한다. 즉, 현대 정보기술은 기술적 논쟁의 종결을 의미하는 기술적 닫힘이나 안정화의 문제가 언제나 변화 가능성을 갖고 열려 있다는 점에 주목하는 것이다. 기술 실천론적 관점에서 보면 검색엔진 기술은 이 기술을 통해 다양한 행위자들(예를 들어 이용자들, 개발자들, 기획자들 등)이 지속적으로 상호작용한다는 특징을 지닌다. 즉, 집단에 따라 검색엔진을 이해하는 관점에 차이가 존재하며 검색엔진을 정의하는 방식이나 운영방식 및 필요한 규제가 무엇인지에 대한 인식도 모두 상이하다. 따라서 검색서비스를 둘러싼 사회적 관계와 상호작용을 다양한 맥락에서 살펴볼 필요가 있다.

8) Granka, A. L. (2010). The politics of search: a decade retrospective. *The Information Society*, pp. 364-374.

## 검색서비스와 사회적 상호작용

### 1) 시장 및 비즈니스 측면

검색서비스 시장은 간접 네트워크 효과<sup>9)</sup>로 복수의 시장주체가 결합된 다면시장(multi-sided markets)으로 정의할 수 있다.<sup>10)</sup> 특정 검색서비스를 이용하는 이용자가 많으면 많을수록, 웹페이지나 웹사이트를 검색서비스의 수집 대상으로 제공하는 정보·콘텐츠 사업자가 증가한다. 또 검색서비스에 색인된 정보와 콘텐츠가 증가하면 증가할수록, 더욱 더 많은 이용자가 해당 검색서비스를 이용하게 된다. 이는 결과적으로 광고주에게 해당 검색서비스에 대한 경제적 가치를 증가시키고, 늘어난 검색서비스의 광고매출은 검색사업자가 서비스를 진화시키고 확장시킬 가능성을 높인다.

이러한 간접 네트워크 효과는 검색서비스 비즈니스 모델의 중요한 구조적 특성이다. 검색서비스의 광고매출 전망은 이용자 및 정보·콘텐츠 규모에 종속되는 반면 이용자와 정보·콘텐츠 규모는 광고매출 전망에 종속되지 않는다. 이러한 비대칭성으로 인해 검색서비스 사업자는 광고주에게만 수수료를 요구한다. 즉, 검색서비스의 간접 네트워크 효과는 소비자 집단에 기초한 규모의 경제(demand-side economies of scale)를 가능케 하고, 소비자측면에서 발생하는 규모의 경제는 중개자인 검색서비스 사업자의 독점화 경향을 유발한다.<sup>11)</sup>

또한 검색서비스는 경험재의 성격을 가지고 있다. 개별 이용자가 검색서비스를 소비한 이후에야 그 유익을 결정할 수 있기 때문이다. 그리고 대부분의 이용자는 검색결과의 정확성과 품질을 적절하게 판단할 수 있는 능력을 가지고 있지 않는 경우가 많아 검색사업자와 이용자 사이에 정보비대칭(information asymmetries)이 나타날 수 있다.

인터넷에서 폭발적으로 증가하는 데이터를 수집, 분류, 평가, 표현하는 과정과 (순위)알고리즘을 개발하고 개선하는 데에는 많은 경제적·기술적 비용이 발생한다. 이는 시장진입의 장애물로 작용하게 된다. 따라서 소수의 검색사업자가 시장 우위를 선점하면 신규 사업자의 시장진입이 어려워지는 것이다.

9) 동일 소비자 집단의 규모가 증가할수록, 해당 집단에 포함된 개인의 유익이 증가하는 것을 직접 네트워크 효과라고 부른다. 전화, 이메일 등에서 직접 네트워크 효과가 나타난다. 반면 간접 네트워크 효과는, 이중 소비자 집단 사이에 발생하는 네트워크 효과를 의미한다. 예를 들어 신용카드 소비자 집단의 규모가 증가할수록, 신용카드 가맹점의 유익이 증가한다. 역으로 신용카드 가맹점 규모가 늘어날 때 신용카드 소지자의 유익이 커진다. 신용카드 소지자와 신용카드 가맹점이라는 이중 소비자 집단 사이에는 간접 네트워크 효과가 존재한다. 그리고 간접 네트워크 효과가 존재하는 이중 소비자 집단을 연결하는 시장주체를 플랫폼 사업자라고 부른다.

10) Evans, S. D. (2003). The Antitrust Economics of Multi-Sided Platform Markets, *Yale Journal on Regulation*, 20(2), p. 325.

11) Evans, S. D. & Schmalensee, R. (2013). The Antitrust Analysis of Multi-Sided Platform Businesses, in Blair, R. & Sokol, D. (eds.), *Oxford Handbook on International Antitrust Economics*, URL: <http://ssrn.com/abstract=2185373>

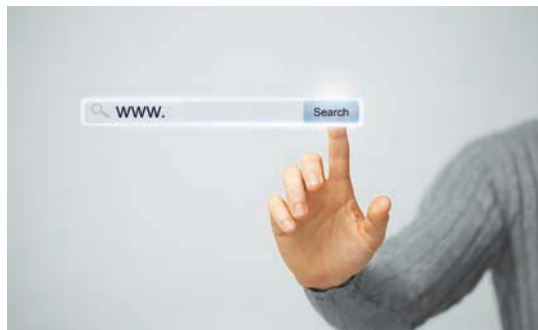
한편, 네이버와 구글 등은 검색서비스 외에도 지도, 블로그, 커뮤니티, 쇼핑 등 검색서비스와 연결될 수 있는 다양한 서비스를 제공한다. 검색사업자는 이러한 정보·콘텐츠 서비스가 자사 검색서비스의 검색(순위)알고리즘에서 높은 평가를 받을 수 있도록 검색엔진을 최적화할 수 있다.<sup>12)</sup> 이 경우 검색서비스 외의 영역에서도 신규 사업자에게 심각한 경쟁제약이 발생할 수 있다.

## 2) 법적·정치적 측면

검색사업자의 시장지배력 강화는 정치적 갈등 및 법적 규제 논란과도 연결된다. 2012년 이후 국내에서 제기되고 있는 ‘검색 중립성’ 논쟁이 대표적이다. 미래창조과학부는 2013년 10월 4일 ‘인터넷 검색서비스 발전을 위한 권고안(이하 권고안)’을 발표했다. 권고안의 내용은 ① 검색원칙 공개, ② 광고와 검색결과와의 명확한 구분, ③ 검색결과에서 이용자가 분명하게 인지할 수 있도록 자사 서비스를 구분하여 표기, ④ 미래창조과학부, 검색서비스 사업자, 전문가 등이 참여하는 정책자문기구를 설치해 권고안의 이행 및 개선 등의 방안 연구 등이다. 또한 권고안은 검색서비스의 공정성 및 투명성을 높여 이용자의 권익증진과 인터넷 생태계의 지속 가능한 발전을 도모하는 데 그 목적을 두고 있다.

검색서비스는 정보망에 접근할 수 있도록 하고 동시에 정보망에서 정보를 발견할 가능성을 결정하는 정보중개서비스다. 전기통신사업법, 정보통신망 이용촉진 및 정보보호 등에 관한 법 등은 정보통신서비스 제공자에게 물리적 망과 가상 플랫폼의 맥락에서 중립성을 요구하고 있다. 그러나 중립성이라는 개념의 정의, 중립성을 확보하기 위한 요구사항에 대한 논쟁이 지속되고 있다.<sup>13)</sup>

검색중립성은 정보 및 콘텐츠에 대한 검색결과가 순수하게 중요도(relevance)에 따라 배치되어야 하며 검색결과에 검색서비스 사업자의 편집 정책이 영향을 미쳐서는 안 된다는 원칙을 의미한다.<sup>14)</sup>



12) Dakanalis, D. & van Rooijen, A. (2011). EU: Google Under Antitrust Scrutiny, *Computer Law Review International*, 1, p. 29.

13) 김보라미 (2012). 국내 망중립성 정책의 문제점. 방송통신위원회의 정책을 중심으로. 망중립성이용자포럼 편, 121-155면. <망중립성을 말하다>. 서울: 블로터앤미디어. 검색서비스가 정보사회에서 가지는 중요성에 비해, 검색중립성이라는 개념이 법률적 및 기술적 논쟁의 대상이 된 것은 매우 최근의 일이다. 국내의 경우, 2012년 9월 민주당 신경민 의원실에서 주최한 “포털의 검색중립성 토론회”와 2013년 4월 새누리당 경제민주화 실천모임에서 개최한 “대형포털의 불공정거래. 무엇이 문제인가?” 토론회 등에서 검색중립성이 처음으로 정치권 및 학계의 논의대상으로 등장했다.

14) Raff, A. (2009. Dec. 28.). Search, but You May Find, *The New York Times*, URL: [http://www.nytimes.com/2009/12/28/opinion/28raff.html?\\_r=0](http://www.nytimes.com/2009/12/28/opinion/28raff.html?_r=0)

여기서 검색결과를 중요도에 따라 분류하는 검색 (순위)알고리즘은 검색결과에 차별성을 부여하여 다양한 검색서비스 사업자를 구별하는 기술이다. 차별화된 검색결과는 이용자들이 특정 검색서비스를 선택하는 이유이기도 하다. 따라서 검색서비스 사업자별로 정보 및 콘텐츠 중요도를 평가하고 종합하는 기술의 차이<sup>15)</sup>는 부정되어서는 안 된다. 따라서 검색중립성은 망중립성과 달리 절대적 객관성을 요구하지 않는다. 그러나 ① 검색최적화 등 검색오용(abusing)에 대한 대응의 적절성, ② 검색결과에서 음란물, 허위정보 등의 삭제 및 접근제한 요청에<sup>16)</sup> 대한 대응의 적절성, ③ 실시간 급상승 검색어, 연관 검색어 등을 둘러싼 검색결과 편집 가능성, ④ 검색광고 등 유료 검색어의 검색결과 상위노출의 적절성, ⑤ 검색결과에서 검색서비스 사업자의 수직 서비스(vertical services, 이른바 '자사 서비스') 우대의 타당성 문제<sup>17)</sup> 등 검색중립성과 관련한 내외부적 관여에 대해서는 많은 논쟁이 있다.

### 3) 사회문화적 측면

검색서비스에 사회적 책임을 요구하는 이유로 힌만(Hinman)은 네 가지를 제시한다.<sup>18)</sup> 첫째, 검색서비스는 정보 접근에 있어 결정적 역할을 담당한다. 검색서비스가 존재하지 않는다면 웹에 접근이 불가능하게 되고, 웹의 유용성 또한 상실된다. 둘째, 정보 접근은 책임감 있는 시민의식 형성에 주요한 역할을 담당한다. 민주사회에서 시민은 정확한 정보에 대한 접근권이 있을 때 숙의된 결정을 내릴 수 있기 때문이다. 셋째, 검색서비스는 교육 분야에서 중요한 역할을 한다. 많은 학생들이 도서관을 방문하기보다 검색서비스를 통해 정보와 지식에 접근하고 있다. 넷째, 검색서비스는 이윤극대화를 추구하는 사기업에 의해 운영된다. 따라서 이윤을 추구하는 기업 목표와 검색서비스가 가지는 공공의 이해는 충돌할 수밖에 없으며 사회적 갈등 해결을 위한 노력이 요구된다.

검색서비스 사업자의 사회적 책임과 공공의 이해 간의 조화 가능성 및 타당성 여부 등은 끊임 없는 논란의 대상이다. 이용자의 인터넷 이용시간 증가와 함께 인터넷을 통한 정보 소비는 여

15) 구글(Google)의 검색 순위알고리즘인 페이지랭크(PageRank)는 최적의 검색결과를 첫 3개에 표현하는데 그 우수성이 있다는 평가를 받고 있다. 그리고 이 우수성은 다른 경쟁 검색서비스보다 구글을 우위에 위치시킨 결정적 요인이다. 구글의 페이지랭크는 검색결과에 색인되는 정보 및 콘텐츠를 명백하게 '차별'하고 있다.

16) 명예훼손 소지가 있는 인터넷 게시물에 대한 접근을 제한하는 임시조치는 외부로부터 검색서비스 사업자에게 가해지는 편집 압력으로 해석할 수 있다. 또한 정보 및 콘텐츠 공급자에 의해 이뤄지는 검색최적화는 검색서비스에 대한 외부 간섭으로 분류할 수 있다. 이처럼 검색서비스 외부에서 검색결과에 가해지는 영향을 '외부 영향력 또는 외부 간섭(external interference)'이라 부른다[Kühling, J. & Gauß, N. (2007). Suchmaschinen - eine Gefahr für den Informationszugang und die Informationsvielfalt?, *Zeitschrift für Urheber- und Medienrecht: ZUM*, 51(12), pp. 881-889.].

17) ①과 ②를 제외하고 ③부터 ⑤까지 검색결과에 대한 검색서비스 사업자 영향력 행사를 '내부 영향력 또는 내부 간섭(internal interference)'이라 부른다[앞의 Kühling & Gauß, 2007, p. 883].

18) Hinman, M. L. (2005). Esse est indicato in Google: Ethical and Political Issues in Search Engines, *International Review of Information Ethics*, 3, pp. 21-22.

론 형성과 관련해 더욱 큰 의미를 얻고 있다.<sup>19)</sup> 검색서비스는 전통 미디어와 같이 스스로 정보나 여론을 생산하고 전달하지는 않는다. 그러나 검색서비스 및 이와 연결된 제2시장, 제3시장<sup>20)</sup>은 다양한 정보 및 콘텐츠에 이용자의 관심을 배분하는 기능을 담당하고 있다. 특히 이용자의 관심을 검색결과 상위에 위치한 정보 또는 콘텐츠로 유도할 수 있기 때문에 검색결과 첫 페이지라는 제한된 공간을 선점하려는 경쟁이 강화될수록 검색서비스의 이용자 관심 배분 결정력은 강화된다.

검색서비스를 둘러싼 사회문화적 갈등의 또 다른 측면은 검색결과에 대한 사회적 간섭과 그 타당성이다. 사회적 간섭은 내부간섭(internal interference)과 외부간섭(external interference)으로 구분된다. 내부간섭에는 검색결과에서 자사 서비스를 우대하거나 자체적인 유해 콘텐츠 필터링 등이 있다. 외부간섭의 대표적 사례는 검색엔진 최적화(Search Engine Optimization)<sup>21)</sup>, 검색엔진 마케팅(Search Engine Marketing) 등 정보제공자에 의한 것과, 임시조치나 정부의 권고안과 같이 공적기관에 의한 것이 있다.<sup>22)</sup>

검색엔진 최적화 수단은 웹페이지 또는 웹사이트가 크롤러 등 수집 프로그램에 효과적으로 파악되고 분류될 수 있도록 지원하는 긍정적인 면만 있는 것이 아니다. 여기에는 크로킹(Cloaking, 검색 크롤러가 데이터를 수집하는 웹페이지와 이용자가 방문하는 웹페이지를 다르게 만드는 방법), 도어웨이-페이지(Doorway-Pages, 검색순위만을 위해 존재하는 웹페이지로서 이용자가 방문하였을 경우 재전송을 통해 다른 웹페이지로 연결하는 방법), 키워드 스테핑(Keyword-Stuffing, 웹페이지와 관련 없는 키워드를 나열하는 방법) 등처럼 부적절하게 검색서비스와 이용자를 현혹하는 기법 등도 포함되어 있다. 이렇게 검색결과를 교란하는 부정적 의미의 검색엔진 최적화에 대해 일반적으로 검색사업자는 해당 웹페이지 등을 검색결과에서 제외하는 조치를 취한다. 끊임없이 새로워지는 부정적 검색엔진 최적화 기법(외부 간섭)에 대응하기 위한 검색서비스의 방법론(내부 간섭)이 진화하고 있는 것이다.

19) Pasquale, F. (2010). Beyond Innovation and competition: The need for qualified transparency in internet intermediaries, *Northwestern University Law Review*, 104(1), p. 110.

20) 네이버, 다음, 구글의 뉴스 서비스 및 블로그, 지도 서비스 등이 제2시장 또는 제3시장으로 분류될 수 있다.

21) 검색(엔진)최적화는, 정보 및 콘텐츠를 담은 웹페이지 또는 웹사이트가 검색(순위)알고리즘에 높게 평가받도록 지원하고 이를 통해 해당 웹페이지 및 웹사이트가 검색결과에서 상위에 위치하는데 도움을 주는 다양한 수단을 말한다.

22) 최근 정부 여당에 의한 포털의 뉴스 검색서비스의 공정성에 대한 문제제기도 이러한 맥락에서 이해될 수 있다.

## 정보기술 거버넌스의 조건들

이렇게 검색서비스를 둘러싼 시장경제적 문제, 법적·정치적 충돌, 사회문화적 갈등은 검색서비스와 관계를 맺고 있는 다양한 이해관계자의 상호작용에 영향을 미쳐 특정한 구조, 작동방식을 형성하고 있다. 그리고 다양한 기술 실천 행위자들에 의해 검색서비스 기술의 성격과 시장의 특성이 끊임없이 변화하고 있다.

검색엔진이나 검색서비스는 단순한 기술이 아니다. 서비스와 관련된 다양한 행위자들 사이의 경쟁과 갈등, 미디어 및 문화 콘텐츠 제공자, 검색광고 제공자, 입법부 및 행정부 등 다양한 시장 주체들의 협상과 합의를 내포하는 기술이다. 따라서 검색엔진, 검색서비스는 단순히 성공적인 비즈니스 모델의 문제가 아닌 기술 거버넌스의 대상이자 산물로 바라볼 필요가 있다.

특히 최근 넷플릭스, 구글, 유튜브 등과 같은 해외 서비스들의 국내 진출이 본격화되고 있는 상황에서 바람직한 정보기술 거버넌스에 대한 논의는 급변하고 있는 글로벌 IT 서비스의 발전과 시장 상황의 변화, 국가와 시민사회(이용자)의 적절한 대응방식 등에 대한 깊은 토론을 필요로 하고 있다. 앞서 제시한 검색서비스 기술에 대한 이론적 틀과 방법들이 향후 검색중립성 문제에 대한 상호 협력적 논의에 활용되기를 바란다. 