

미래의 정보검색

동아대학교 컴퓨터공학과 고영중

검색의 미래에 대해서 논하기 전에 먼저 검색의 과거와 현재에 대해서 알아보겠습니다. 1994년 야후(Yahoo)가 등장하면서 전문가들이 웹 페이지의 내용을 잘 분류해주는 소위 전문가의 지능을 중심으로 움직이는 최초의 검색엔진이 있었습니다. 후속 검색엔진인 라이코스(Lycos)와 알타비스타(Altavista)는 키워드(keyword)들로 인터넷상에 있는 데이터들을 찾아내고자 했습니다. 이러한 두 가지 방식은 어정쩡한 방식으로 공존을 하다가 최근에는 완전히 통합된 형태를 가지게 됩니다. 구글(Google)은 기존의 방식과 전혀 다른 혁신적인 변화를 통해서 정보검색의 새 시대를 열었습니다. 전문가들이 아니라 홈페이지를 만드는 웹 마스터라는 일반인들의 링크들을 분석해서 그것을 토대로 인터넷의 순위를 매기는 “Page Rank”를 기반으로 한 검색엔진을 선보였습니다. “Page Rank”는 링크(link)를 가중치가 매겨진 투표의 개념으로 생각하여 링크가 많이 걸릴수록, 또 다른 사람들로 부터 많은 링크를 받은 누군가가 나를 링크해줄수록 높은 순위를 매겨주는 방식입니다. 즉, 웹 마스터들의 대중의 힘을 사용해서 정보검색의 성능을 높이는 방식이었습니다.

구글의 “Page Rank” 기술 다음의 기술은 무엇인가를 잠시 살펴보겠습니다. 이에 대해서 다음의 대표적인 5가지 이슈(issue)가 있었습니다: 검색 인터페이스(interface)의 변화, 검색의 개인화(personalization), 결과물의 자동 분류, 검색의 의도 파악, 인공지능. 많은 연구가 진행되고 있지만 현재까지 성공적인 서비스가 보여지지 못했고, 실패한 이유에 대해서 여러 가지 관점에서 논의되고 있습니다만, 아직도 정보검색을 개발하고 서비스하는 사람들에게는 관심을 가지고 연구하고 도전해야 할 많은 내용이 앞의 5가지에 숨어 있다고 하겠습니다.

정보검색의 미래에 대해서 우리가 관심을 가져야 하는 두 가지 경향이 있는데, 먼저 미래의 웹(웹 2.0)의 모든 고객들은 웹을 통해 마음껏 자유롭고 거리낄 것이 없이 정보를 제공하고, 공유한다는 특징이 있다는 것과 웹 2.0에서는 모든 정보가 구름과 같이 언제 어떻게 생겨나지도 파악하기 힘들고, 특정한 형태에 대해서 정의 내리려고 하면 쉽게 다른 모양으로 변해버리고, 시간이 지남에 따라서 처음에는 아무 의미를 갖지 않는 것처럼 보였던 것들도 나름대로 의미를 갖춘 모양으로 바꾸어가는 특징이 있다는 것이다. 결국으로는 구름과 같이 유동적인 웹 2.0이라는 물결 안에서는 한 사람, 한 사람이 다 자기만의 카드를 가지고 있어서 내가 들고 있는 카드가 대체 뭔지는 잘 모르겠으나 그것들을 다 모아봤더니 결국 우리가 다 같이 읽을 수 있는 어떤 형태의 모습이나오게 되는 카드섹션(card section)과 같은 특징이 있습니다. 미래의 검색은 기존에 정보검색자들이 추구해 왔던 효율, 전문가들이 공학적인 계산에 의해서 구축한 웹의 구조뿐만이 아니라, 대중의 힘이 작용하는 효과의 세계, 일반 사람들이 마음껏 비합리적이고, 비효율적으로 만들어도 아무튼 의미 있는 결과를 만들어 내는 웹2.0의 환경을 담아 낼 수 있어야 합니다. 이 두 가지는 분리되고 서로 배제되는 것이 아니라 만나고 연합해야 성공을 할 수 있고, 이들을 가장 효율적으

로 연합할 수 있는 지점이 블로그(blog)와 검색입니다. 그러므로, 이들을 중간 지점에서 잘 묶어 주고 적절하게 공존하면서, 어떻게 유의미한 결과물들을 만들어 낼 수 있느냐 하는 것에 대한 고민이 우리 시대에 주어진 화두이자, 정보검색이 앞으로 나아가야 할 방향이라고 할 수 있습니다. 다이퀘스트도 이러한 시대적 변화에 발맞추어 새로운 미래 정보검색 시장을 개척하는 회사로 거듭날 것이라 믿습니다.