

존 바텔,
『검색으로 세상을 바꾼 구글스토리』
 (랜던하우스중앙, 2005)

발제: 김 상 배

- o 구글(Google) 개관
 - 스탠포드대학 출신의 래리 페이지, 세르게이 브린이 1998년 9월 창업 --> 2001년 에릭 슈미트를 CEO로 영입 --> 구글 3인방
 - 구글(51%), 야후(24%), MSN(13%), AOL(5%) - 2005년 1/4분기 현재 세계 검색시장 점유율
 - 네이버(73.2%), 다음(9.4%), 야후코리아(7.6%), 애플스(5.5%), 구글(1.3%) - 국내 검색 서비스 점유율 상위 5위

- o 구글 성공의 기술적 측면
 - “지능이 있는 검색엔진을 창조한 기업은 구글뿐이다”(p.51) - "구글이 사용자들로부터 찬사를 받는 것은 검색이 잘된다는 단 한 가지 이유 때문“(p.215)
 - ‘의도의 데이터베이스(Database of Intention)’ --> 검색엔진의 목표는 당신의 진정한 의도의 맥락에서 정보를 찾아주는 것
 - 구글은 이러한 작업을 처리하기 위해서 17만 5,000개가 넘는 서버(2005년 1/4분기 현재)를 사용
 - 그런데 구글 컴퓨팅 시스템의 특징은 분산컴퓨팅을 채택하고 있다는 점 - 더 많은 컴퓨터를 투입하면 할수록 시스템이 더욱 견고해진다 - 부품 하나가 고장 난다고 해도 그냥 바꾸기만 하면 되기 때문에 별다른 문제가 없다 - 일종의 “network”형 cf. IBM과 후지쯔의 메인프레임 - 일종의 “organization”형
 - 자체 서버에다 고유의 운영체제(OS)를 개발하고 주문제작까지 하며 부품의 설계와 냉각, 적재방식을 특허로 등록하기까지 함 cf. 구글 서버의 OS는? MS or 리눅스?
 - 검색은 인간과 기계의 만남 --> 인공지능의 창조 --> 검색은 어떻게 이루어지나?
 - i) 크롤러(crawler) 또는 스파이더(spider): 웹 링크를 오가는 정보수집 전용 프로그램
 - ii) 인덱서(indexer): 문서들을 빨리 찾을 수 있게 자동으로 인덱스(데이터베이스, DB)를 만들어주는 프로그램 --> 크롤러의 작업을 통해서 인덱스는 계속 보완됨 --> 구글의 인덱싱 프로그램=구글봇(Googlebot)
 - iii) 런타임 시스템(runtime system) or 쿼리 프로세서(query processor): 사용자와 검색어를 태그(tag)가 붙여져 있는 인덱스에 연결하여 결과를 제공 - 사용자 인터페이스와 검색 소프트웨어 --> 검색엔진의 지능에 해당
 - 검색 소프트웨어 - 하이퍼텍스트의 초기 버전들은 역방향으로 링크를 추적할 수 없는 결함을 지니고 있었으나...

- 자신들의 사이트를 링크하고 있는 백링크(Backlink) 찾기 - 래리 페이지의 백링크(BackRub) 프로젝트 - annotation이나 citation과 유사한 문제 - 랭킹시스템으로 발전 - 이를 검색에 적용
 - 즉, 특정 사이트로 이어진 링크의 수와 그 링크를 담고 있는 사이트에 연결된 링크의 수 모두를 고려하는 알고리즘을 개발 --> 페이지랭크(PageRank)라고 불림 --> 더 좋은 검색을 창출 --> “냉정한 알고리즘에 기초한 랭킹 체계”(p.143)
- o 구글 성공의 비즈니스 측면
- 검색을 광고에 연결시키지 않은 채 수익을 창출하는 모델 - 광고주보다는 사용자를 우선시하는 방법의 모색 - 구매의사를 가진 사람을 도와주는 검색엔진
 - 방문객들이 클릭해서 광고주들의 사이트에 들어갈 때만 돈을 지불하는 모델 - 고투닷컴의 빌 그로스의 방식에 기원 - 클릭당 지불방식 - 유료검색 광고 네트워크를 창출 --> 구글의 애드워드(AdWords) - 키워드 검색광고
 - 의도에 기초한 마케팅 cf. 성과당 지불방식
 - 애드센스(AdSense) - 콘텐츠 연동광고: 소비자들의 취향에 따른 검색에 따라 움직이는 것이 아니라 사이트의 콘텐츠에 따라 움직임
- o 구글 권력(1)
- 정보화시대의 컴퓨터 활용에 반드시 필요한 인터페이스(p.31) --> 인터넷의 모든 산업이 검색으로 연결 - 블로그, 이메일, 뉴스 등
 - 구글뉴스: 다른 사이트들의 기사를 찾아서 읽게 해주는 서비스 - 9.11 이후탄생
 - cf. Gateway: 인터넷에서 원하는 정보를 찾을 때 꼭 거쳐야하는 gateways --> --> 컴퓨터 OS(MS 윈도우), 웹브라우저(인터넷 익스플로러), 도메인이름체계(DNS), 검색엔진 or 포털사이트(구글, 야후, 네이버) 등 cf. platform으로서의 권력
 - “구글은 전세계 거의 모든 나라 경제의 공급과 수요를 중개하는 플랫폼을 제공하기를 원한다”(p.379) - 일종의 플랫폼 파워 - 구글 그리드(Google Grid): 구글이 가지고 있는 모든 서비스를 통합한 거대한 플랫폼 - 차세대 컴퓨팅의 플랫폼
 - 승자독식의 power law 법칙이 작용하는 세계의 표준설정(standards setting)
 - 웹2.0시대의 플랫폼 cf. Wintelism(PC시대의 플랫폼)과의 유사점과 차이점 cf. 상대적 폐쇄성의 모델 - 모노허브형
 - 구글아키(Googlearchy)? "Rule by Google" --> archy의 번역어? 지배, 정체(政體), 치(治) --> 구글지배 cf. Archy의 뜻에는 Jef Raskin이라는 사람이 디자인한 휴먼-컴퓨터 인터페이스 시스템의 이름이기도 함 --> 아마도 구글아키를 조어할 때 참조된 듯함
- o 구글 권력(2)
- 검색결과 알고리즘 변경이 가져오는 권력
 - 구글댄스(Google Dance): 구글의 주기적인 알고리즘 업데이트에 붙여진 별명
 - 검색결과에서 더 높은 순위를 차지하기 위해서 조작된 페이지, 즉 스팸 제거의 목적으로 수행 --> 소위 검색엔진 마케팅 및 최적화 작업 종사자 견제 --> 블랙햇과 화이트햇을 가려내기 --> 그러나 구글은 알고리즘을 철저하게 감추고 그 과정을

공개하지 않음

- 그런데 구글은 “검색자가 보고 싶어 하는 것은 상업적인 사이트가 아니라 리뷰일 뿐이라고 판단” - 순수하게 조직적인 목록과 돈을 지불한 목록을 어떻게 구분할 것인가? cf. 선의(善意)의 독단? cf. 상표법 문제
- 구글에 대한 신뢰의 문제와 연결 cf. 구글에 대한 반독점 소송의 가능성?
- 구글이 지닌 의도의 순수성에 대한 의심 - ‘플로리다(허리케인에 비유된 구글댄스 중의 하나)’는 에드워즈를 사게 하기 위한 전략이었다? ex) 2bigfeet.com의 사례
- 구글의 비의도성 --> “구글은 악덕 검색엔진 최적화 업체와 제휴 프로그램을 제거하려는 시도를 한 것인데, 불행하게도 다수의 악의 없는 구경꾼들이 희생양이 되었던 것”(p.265) --> 비의도적 권력
- cf. Castells이 말하는 일종의 programmer로서의 권력, S. Strange의 구조적 권력

o 구글 권력(3)

- 검색에 의해서 창조되고 저장되는 데이터는 누가 소유하는가? cf. 감시권력
- 프라이버시 문제 - MapQuest의 사례 - 전화번호와 우리집 지도 - 역(逆) 디렉토리 지역검색 - Gmail과 광고의 사례 - 구글 데스크톱 서치 프로그램 (GDS)
- 공공기록에 대한 전자적 접근의 문제 - 특히 9.11 이후 미국애국법안 - 감시 (surveillance)적 지식국가와 구글의 신뢰성
- 중국의 정보통제 사례 - Great Firewall - 인터넷의 탈집중성 vs. 국가의 통제

o 구글권력(4)

- ‘구글주스(Googlejuice)’ - 구글 검색결과에서 더 높게 나오게 해주는 힘을 뜻하는데 다른 사이트에서 많이 링크해줄수록 힘이 커진다 cf. 네트워크 외부성
- 검색의 힘은 바로 “꼬리” 부분에 놓여 있다 - 불특정 다수로서의 롱테일을 엮는 권력 - 애드센스 cf. 네트워킹의 권력
- 좋은 질문을 만드는 것보다 더 위대한 창조행위는 없다 cf. 어젠다 세팅 파워
- 옹호광고(advocacy ad) 공과를 통해 정치, 경제, 사회와 관련된 문제를 거론함으로써 광고주의 생각을 표명하면서 여론을 환기시켜 대중의 동의를 얻으려는 광고 - 구글은 이러한 광고의 구매를 반대

o 구글의 corporate governance

- 대부분의 엔지니어링 위주의 기업체에서 전형적으로 나타나는 서열제적 요소 - 소위 “2+1시스템” --> 위계(hierarchy)에 기초한 전통적인 경영구조
- 이중의결(dual class) 주주구조: 창립자들과 고위 임원들이 일반 주주들보다 훨씬 많은 지배력을 가짐 - 30% 주식 소유, 그러나 10배의 voting power
- 기업문화: “Don't Be Evil(악(惡)해지지 말자)” - 세계의 정보를 조직해 널리 접근할 수 있게 해서 유용하게 하자
- “구글은 전통적인 주식판매방식뿐만 아니라 기업지배구조, 투자자와의 의사소통, 그리고 경영구조 등에 대한 월가의 전통적인 접근법을 상당히 모욕“(p.338)
- 기업공개 이후 구글은 핵심인 검색, 광고상품, 그리고 ‘20퍼센트’(지메일, 구글뉴스, 오쿠트닷컴), ‘10퍼센트’(구글맵스, 구글어쓰)라는 기능별 핵심그룹으로 구성

- “본질적으로 구글은 검색과 광고라는 두 가지 핵심사업을 시행하는 부서에 집중했고, 앞으로 핵심사업으로 바뀔지 알 수 없는 프로젝트를 추구하는 그룹은 좀 더 느슨히 운영”(p.358) --> 꼭 GE와 같은 구조라고 지적
- 구글의 새로운 인재 리쿠루팅 관행 - 엘리트주의 접근

o 구글 vs. 야후

- 회사의 기본적 성격이 구글은 기술회사임에 반해, 야후는 미디어회사임
- 검색에 대한 접근법(철학)의 차이가 가장 중요
 - 구글은 공공연하게 편집적인 지시를 거의 내리지 않음 - 내용적으로 혹은 편집적으로 이끌어지는 회사가 된다는 것에 대한 거부감 - 주로 기술을 통해서 문제를 해결 - 영리한 알고리즘과 완전한 컴퓨터를 추구 - 구글은 알고리즘의 실패 시에만 검색에 인간이 개입 - 처리하기 어렵고 반복적인 수학문제를 해결하는 매우 영리한 “알고리즘”의 개발에서 시작한 회사의 배경과 관련
 - 야후는 특정결과가 인터넷보다는 편집적 판단에 따라 구성, 추천검색기능 등 - 야후는 명백하게 검색자들을 자체 편집 서비스로 몰아가는데 그 서비스가 검색의 의도를 만족시킬 수 있다고 믿음 - 기본적으로 미디어 기업을 추구 - 재능을 가진 인간이 문제 해결의 절대적 필요요소 - 프로세스는 자동화되었지만 인간이 참여하는 전통적인 구조 - 야후는 인터넷에 대한 “편집가이드”에서 시작
- 기업문화의 차이
 - 구글의 회의문화(늦는 것이 다반사) vs. 야후의 회의문화(정시 시작)
 - cf. 다른 종류의 정보화시대의 권력을 보는 듯 - 즉 구조적 권력 vs. 구성적 권력

o 인터넷 검색시장의 산업구조

- 수직적 통합과 수평적 통합 --> “궁극적으로 수직적 통합이 인터넷 미디어 업계에 적합한 구도가 아닐지 모르지만, 현재까지는 수직적 통합이 적합한 구도로 작용”(p.201)
- MS나 야후와 같은 거대 미디어 기업들이 각종 검색 기능을 가진 회사들을 인수 합병 --> 구글은 역경을 딛고 일어서 빠른 기간 내에 성공

o 구글에 대한 대항담론?

- PC를 터미널로 하는 검색 vs. 포스트PC 기반 검색
 - 유비쿼터스 검색과 시간을 초월한 검색
- 무료로 소스가 공개되는 오픈소스검색엔진 --> 너치(Nutch)
- 비(非) ASCII권의 언어공동체를 배경으로 한 검색서비스
 - 지식검색 vs. 지식문화 --> 네이버 지식In의 사례 cf. 일본의 OK웨이브
- 구글은 일종의 taxonomy인가? 그렇다면 구글에 대항하는 folksonomy 모델은? 하이퍼텍스트, UCC, 소셜북마크
- 태깅이 되어 있는 메타데이터(meta-data)를 기반으로 한 검색 --> 시멘틱 웹 --> 태그의 표준화 문제가 큰 걸림돌 cf. 웹파운틴
- 블로그나 RSS(Real Simple Syndication)과의 관계