

Week 9 • 데이터 저널리즘

Web Technologies

Joonhwan Lee

human-computer interaction + design lab.

오늘 다룰 내용

- Data Communication
- 인터넷과 인터넷 서비스
- 웹(WWW) 관련 기술의 발전 과정
- Web 2.0

1. Data Communication

Data Communication 이란...?

- ◆ Telecommunication의 subset
- ◆ Telecommunication is...
 - ◆ “any **process that permits the passage from a sender to one or more receivers** of information of any nature delivered in any usable form (printed copy, fixed or moving pictures, visible or audible signals, etc.) **by means of any electromagnetic system** (electrical transmission by wire, radio, optical transmission, guided waves, etc.).”
(Martin, James. Introduction to Teleprocessing. Englewood cliffs, NJ:Prentice-hall, 1972)

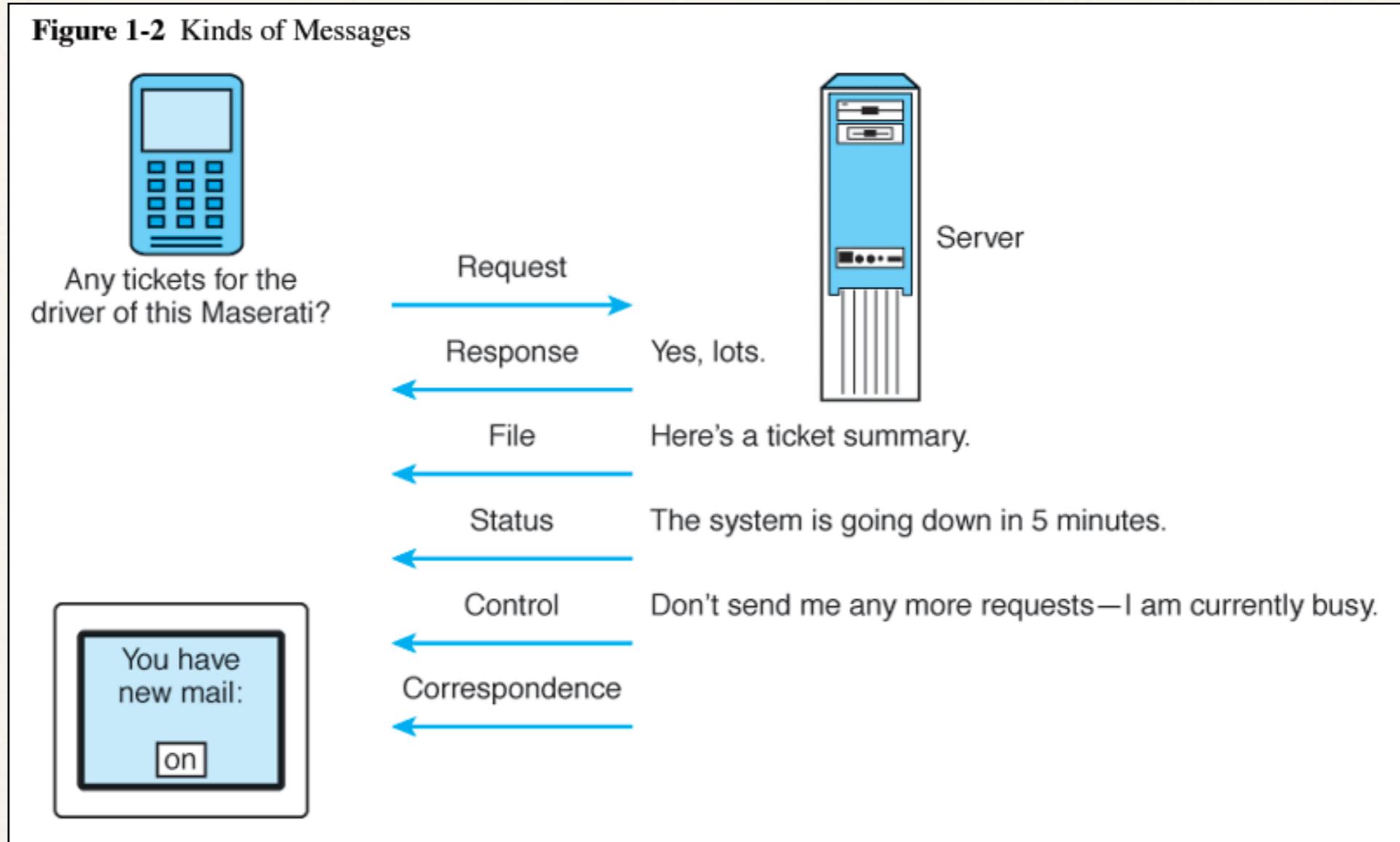
Data Communication 이란...?

- ◆ Data Communication...
 - ◆ “... can be defined as that **part of telecommunications that relates to computer systems, or the electronic transmission of computer data.** This definition *excludes the transmission of data to local peripherals such as disk, tape, and printers.*”
(Stamper, David A. Business Data Communications. Redwood city, CA: Benjamin/Cummings, 1989)
- ◆ Data communication 은 데이터가 네트워크를 통해 전송되기 때문에 보통 **networking** 이라고 불린다.
- ◆ 인터넷의 발명으로 data communication이 활성화.

Features of Communication

- ◆ 네가지 필수요소
 - ◆ Sender, receiver, medium, message
- ◆ Message의 종류
 - ◆ file
 - ◆ Request
 - ◆ Response
 - ◆ Status
 - ◆ Control
 - ◆ Correspondence
- ◆ Understandability
- ◆ Error Detection

Features of Communication

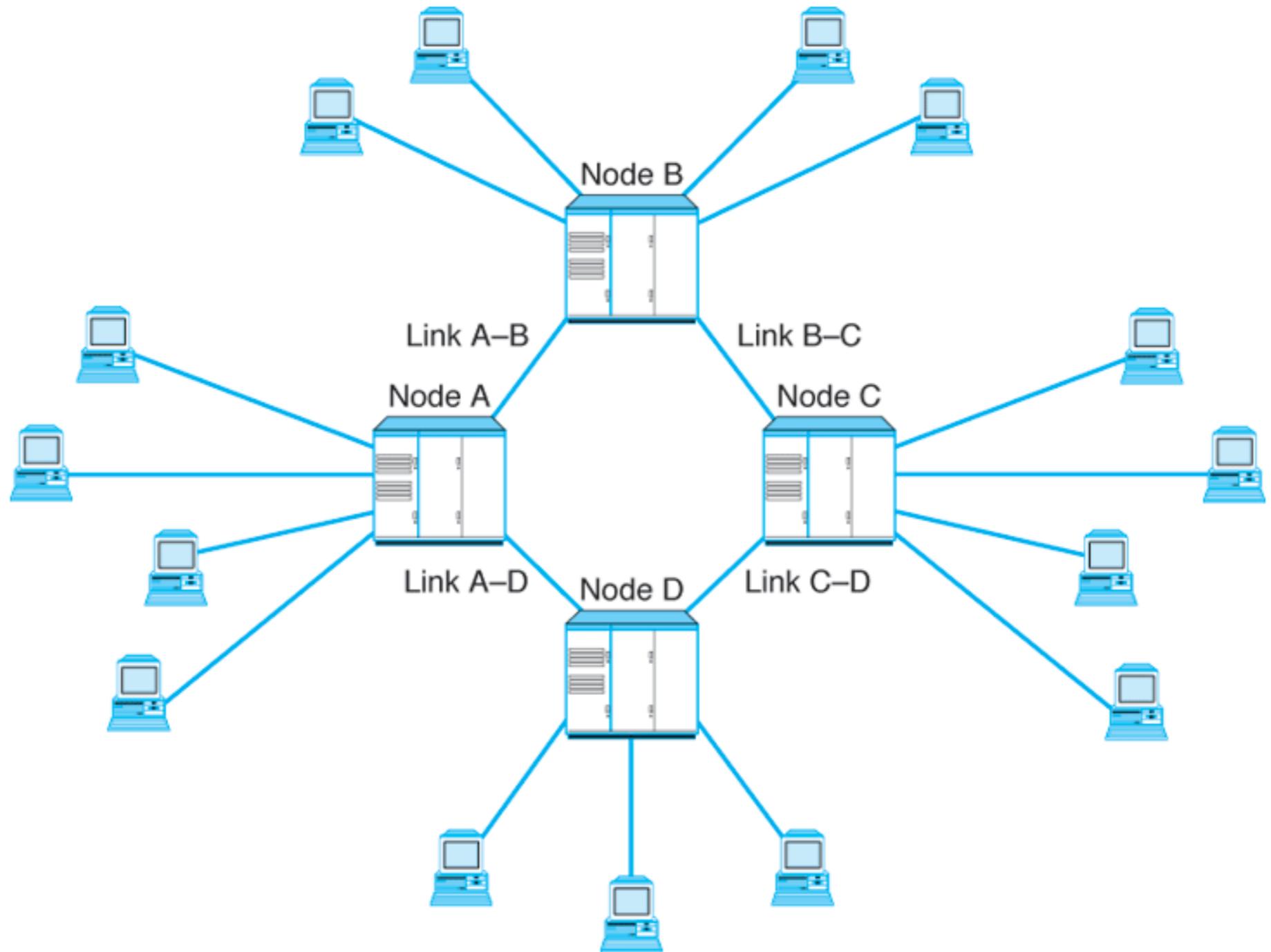


Key Data Communication Concepts

- ✦ **Session:** communication dialog between network users or applications
- ✦ **Network:** interconnected group of computers and communication devices
- ✦ **Node:** a network-attached computer
- ✦ **Link:** connects adjacent nodes
- ✦ **Path:** end-to-end route within a network
- ✦ **Circuit:** the conduit over which data travels
- ✦ **Packetizing:** dividing messages into fixed-length packets prior to transmission over a network's communication media
- ✦ **Routing:** determining a message's path from sending to receiving nodes

Key Data Communication Concepts

Figure 1-4 Links and Paths in a Network



History of Data Communication

- ✦ 1836: Telegraph (Samuel Morse)
- ✦ 1858-1866: Transatlantic cable
- ✦ 1876: Telephone (Alexander Graham Bell)
- ✦ 1957: USSR launches Sputnik
- ✦ 1962-1968: Packet-Switching networks
- ✦ 1969: **Birth of Internet**
- ✦ 1977: E-mail
- ✦ 1979: News Groups
- ✦ 1982: TCP/IP
- ✦ 1991: **Birth of WWW**

http://carl.sandiego.edu/itmg360/internet_history.htm

2. 인터넷과 인터넷 서비스

인터넷

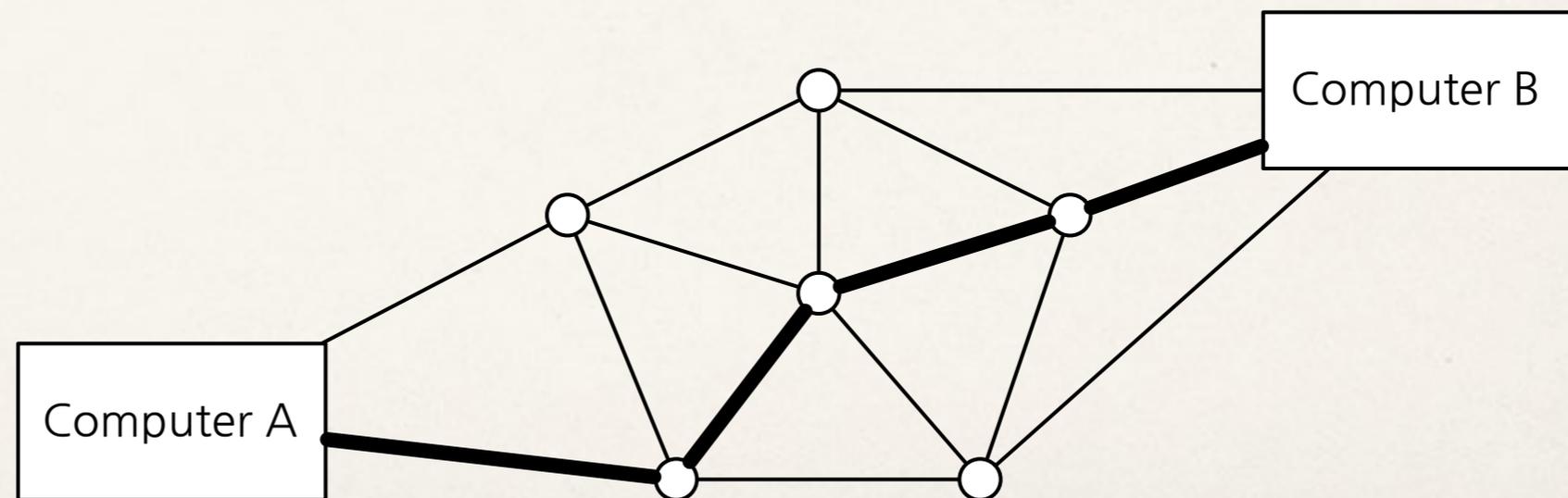
- ◆ 최근 가장 대표적으로 사용되는 정보 통신망
- ◆ Inter+Network 의 합성 단어 → 여러 네트워크들이 연결되어 형성된 네트워크
- ◆ 1969년 미 국방성 주도로 개발:
 - ◆ ARPANET (Advanced Research Projects Agency) 이라는 연구용 네트워크가 시초 → 차후 군사용 (MILNET) 으로 발전
 - ◆ 냉전 시기: 핵전쟁 등에서도 살아 남을 수 있는 네트워크 연구
 - 1960년대, 구 소련이 스푸트니크 (Sputnik) 인공위성 발사 (핵전쟁 위협 증가)
 - circuit switching 보다 packet switching 방식이 견고하고 생존력이 높다고 판단
 - packet switching 에 기반을 둔 ARPANET 개발 (UCLA 와 SRI 연결)
 - 유사시 군사 정보를 공유하는 것이 일차적 목적

데이터 망의 구성

- ◆ 데이터 전송방식에 따른 분류

- ◆ 서킷 (circuit) 스위칭

- ◆ 패킷 스위칭과 달리, 데이터를 전송하기 전, 시작점과 도착점 사이의 회선을 먼저 구성한 뒤, 데이터를 전송
 - ◆ 회선이 구성되면 데이터를 전송하지 않는 기간에도 회선을 유지해야 해서 전체적인 효율이 떨어질 수 있다
 - ◆ 설정된 회선은 외부의 간섭없이 안정적으로 통신 가능

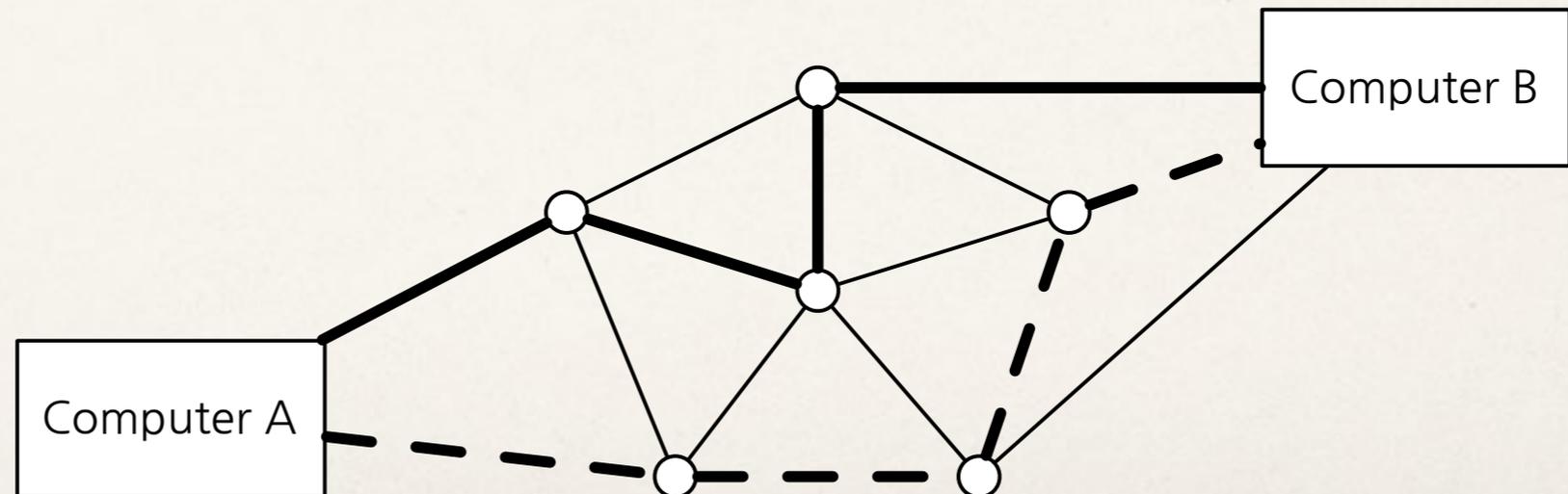


데이터 망의 구성

◆ 데이터 전송방식에 따른 분류

◆ 패킷 (packet) 스위칭

- ◆ 데이터를 패킷 단위로 분리하여 보내는 전송 방식
- ◆ 패킷을 보낼 때, 네트워크 트래픽이 덜 붐비는 경로를 찾아 데이터를 전송
- ◆ 가장 빠르게 보낼 수 있는 경로를 탐색해서 보내기 때문에 전송 효율이 높다
→ 매번 전송되는 패킷의 경로가 바뀔 수 있음
- ◆ http://www.pbs.org/opb/nerds2.0.1/geek_glossary/packet_switching_flash.html



인터넷

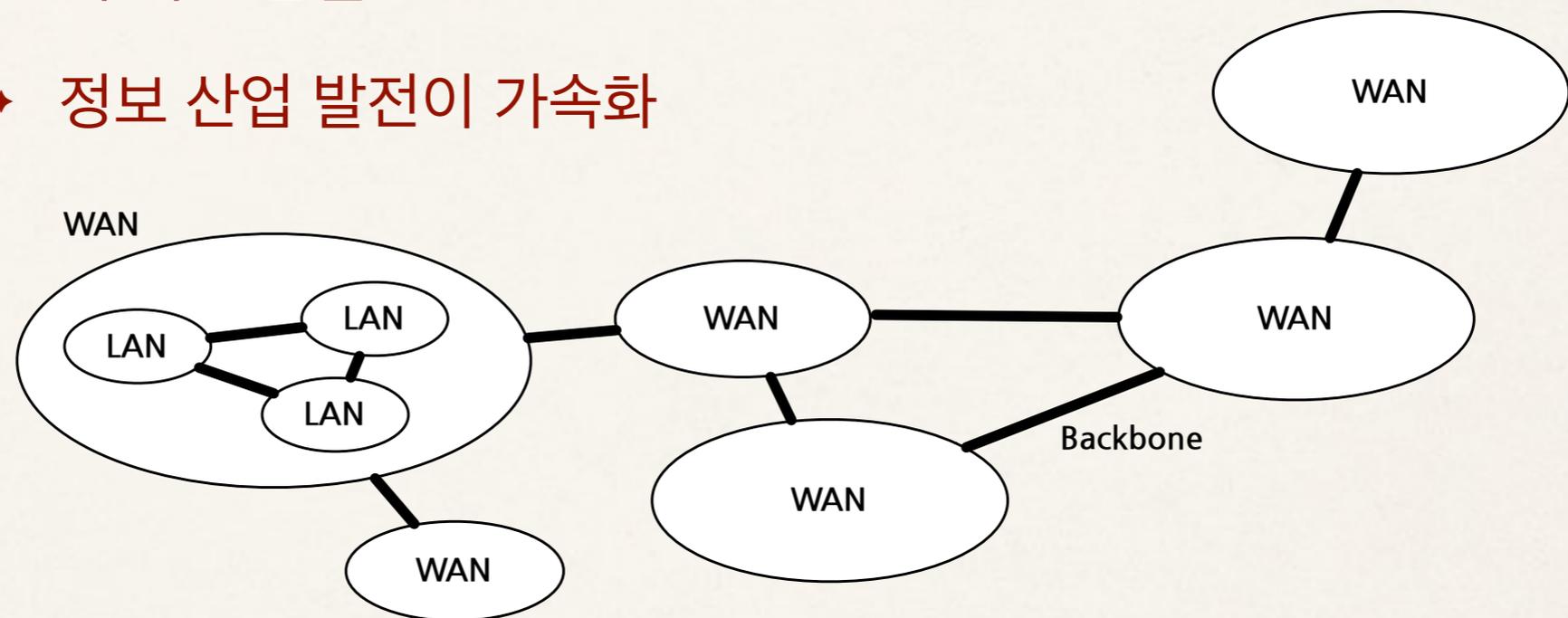
- ◆ 연결된 데이터의 전송을 위하여 IP (Internet Protocol) 이라는 전송규약 개발
 - ◆ TCP/IP
- ◆ Ethernet LAN 기술의 개발로 Ethernet (망) 과 TCP/IP (규약) 을 이용한 LAN 의 구축이 활발해 짐
 - ◆ 1980년대 중반, NSF 에서 NSFNET을 구축 (연구정보 공유)
 - ◆ 수많은 연구소/기관 등이 참여, 미국내 인터넷의 근간으로 활용
- ◆ 1990년대 Web (World Wide Web: WWW) 의 등장으로 정부기관이나 연구소를 벗어나 개인의 영역까지 인터넷 이 활성화
 - ◆ WWW 는 인터넷 서비스의 일종이지만 현재는 인터넷의 대명사로 통용

인터넷

◆ 1996년 정보 고속도로 → 고속 백본 인터넷망 구축하면서 급속도로 발전

◆ 정부, 대학 및 교육기관, 개인, 심지어는 국가간도 인터넷을 통하여 서로 연결

◆ 정보 산업 발전이 가속화



◆ **Backbone Network (기간망)**

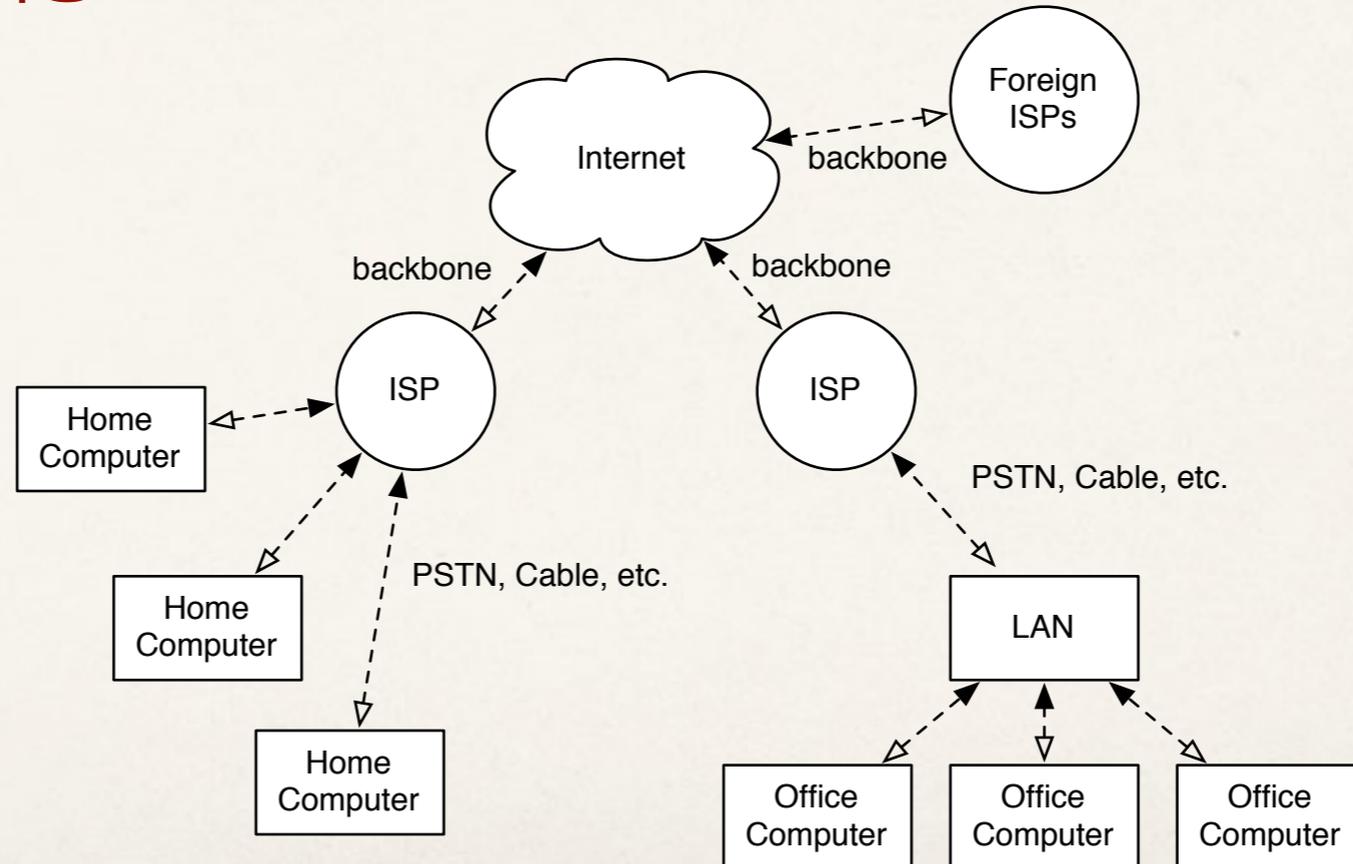
- ◆ 네트워크의 최상위 레벨 → 네트워크의 중심
- ◆ backbone 통신망을 통하여 서로 떨어진 LAN 이나 WAN을 연결
- ◆ backbone 통신망의 속도에 따라 LAN이나 WAN의 속도가 결정
- ◆ 128Mbps ~ 10Gbps 이상의 초고속 대역폭을 가짐 (주로 국가가 관리)

인터넷 인프라의 구성

- ◆ 인터넷: 크고 작은 네트워크의 집합체
- ◆ 개인 컴퓨터 혹은 소/중/대규모 LAN 으로 구성된 네트워크가 인터넷에 접속을 하려면 ISP (Internet Service Provider) 에 가입되어 있어야 함
- ◆ ISP: 초고속 인터넷 회선에 직접 연결되어 있어 인터넷 서비스를 제공하는 기관
 - ◆ 학술기관
 - ◆ 관공서
 - ◆ 상업적 인터넷 서비스 업체 (e.g., olleh, LG U+ 등)
- ◆ ISP에 가입한 사용자는 ISP의 호스트 컴퓨터에 접속 → 인터넷 서비스 이용

인터넷 인프라의 구성

- ◆ 초기 인터넷: PPP 접속, IP 세팅과 같은 복잡한 네트워크 환경 설정 필요
- ◆ 최근 → ISP 가 자동으로 인터넷에 접속할 수 있는 프로그램을 제공



인터넷의 구조

- ◆ 클라이언트-서버 모델 (Client-Server Model) 에 기반
 - ◆ 서버: 제공하고자 하는 서비스에 적합한 정보들을 자신의 하드디스크에 보관 → 서비스 요구에 따라 외부에 제공
 - ◆ 클라이언트: 사용자가 서버에서 제공되는 정보를 얻기 위해 서버에 접속하여 사용하는 컴퓨터

인터넷의 구조

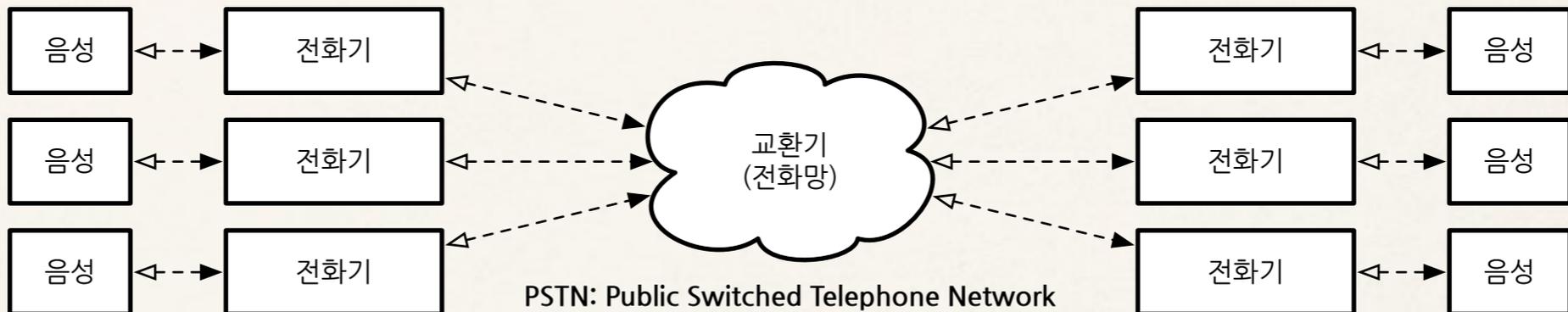
- ◆ 하나의 컴퓨터는 보통 여러가지 서버의 역할을 수행 (서버 프로그램 사용)
 - ◆ 웹서버: 아파치 (Apache) 등을 사용하여 웹 서비스 제공 → 웹 브라우저 필요
 - ◆ FTP서버: FTP 서버 프로그램으로 파일다운로드 서비스 제공 → FTP 클라이언트 필요
 - ◆ Email 서버: 메일 송수신 프로그램으로 메일 서비스 제공 → 이 메일 클라이언트 필요

우편망, PSTN, Computer Network

1. 우편망



2. 전화망



3. 컴퓨터 네트워크 (데이터 통신망)



인터넷의 전송방법과 주소체계

- ◆ TCP/IP 통신 규약
 - ◆ 다양한 운영체제 → 다양한 데이터 형식
 - ◆ 다양한 컴퓨터가 동일한 네트워크에 접속되기 위해서는 모든 컴퓨터에 동일한 형식으로 주소를 할당하고 (IP Address) 데이터를 주고 받기 위한 규약 (TCP/IP) 를 제정해야
 - ◆ 통신 프로토콜: 컴퓨터 간의 통신 규약 → 컴퓨터 사이에 정보를 전달하기 위해 필요한 규칙 및 약속의 집합

인터넷의 전송방법과 주소체계

- ◆ TCP/IP 통신 규약

- ◆ TCP: Transmission Control Protocol

- ◆ 데이터의 흐름을 제어하고 데이터가 정확한지 확인하는 역할
 - ◆ data flow control

- ◆ IP: Internet Protocol

- ◆ 데이터를 이동시킬 목적지를 지정
 - ◆ data destination control

인터넷의 전송방법과 주소체계

◆ TCP/IP 통신 규약

◆ TCP의 역할

- ◆ 데이터를 여러 개의 작은 조각으로 나누어 패킷 (Packet) 이라는 정보단위 생성
- ◆ 패킷이 제대로 전송되는 지 확인
- ◆ 패킷이 도착하면 패킷을 다시 원래의 순서대로 재복원

인터넷의 전송방법과 주소체계

◆ IP 주소 체계

- ◆ IP 의 역할: 주어진 패킷을 어떻게 목적지까지 보낼 것인가에 대한 전송 프로토콜
- ◆ 정보를 전송하기 위해 숫자로 된 주소체계를 가지고 있다
→ IP Address: 주소를 숫자로 표현한 것
(예: 66.249.89.99 → google.com)
- ◆ 숫자는 기억하기 어렵다 → domain name 의 등장
→ 66.249.89.99 = google.com
- ◆ IP 주소의 등록 및 관리 → NIC (Network Information Center) 라는 기구에서 관장
 - ◆ 각 대륙 별/국가 별로 산하 기관 존재 → 국내는 한국인터넷진흥원(KISA) 가 관장

인터넷의 전송방법과 주소체계

◆ IP 주소 체계

- ◆ IP 주소는 4개의 바이트 (Byte) 로 구성

- ◆ 각각의 Byte 는 . 으로 구분

- ◆ 1 Byte = 8 bits (= 2^8) = 256

- ◆ 표현할 수 있는 주소의 총합:

$$256 \times 256 \times 256 \times 256 = 4,294,967,296$$

- ◆ 현재 전 세계적으로 90% 정도 할당

- ◆ 대한민국에 할당된 IP Address: 73,266,688 (69,182,976 사용, 2009/8 기준)

- ◆ 전체 IP 중 1.71% 할당 (약 95% 정도 사용)

인터넷의 전송방법과 주소체계

- ◆ IP 주소 체계 (IPv4 vs. IPv6)
 - ◆ 4바이트 (32비트) 주소체계의 문제점
(255.255.255.255 → 8비트 X 4)
 - ◆ 컴퓨터와 모바일 혁명으로 IP 자원의 고갈이 예상
 - 개인이 사용하는 컴퓨터: 1대 이하에서 다수의 컴퓨터로
 - 모바일 네트워크 환경: 1대의 컴퓨터가 다수의 IP 를 사용
 - ◆ 유비쿼터스 시대의 도래 → 모든 기기에 인터넷 주소 할당필요
- ◆ IPv6 필요성 대두
 - ◆ 128비트 주소체계 (16바이트)
2001:0db8:85a3:08d3:1319:8a2e:0370:7334 → 16비트 X 8
 - ◆ 3.4×10^{38} 까지 표현 가능
(340,282,366,920,938,463,463,374,607,431,768,211,456)

인터넷의 전송방법과 주소체계

- ◆ 도메인 이름 (Domain Name)
 - ◆ IP 주소: 숫자로 구성되어 있어 기억하기 불편 → 기억할 수 있는 문자로 대체
 - ◆ 각 호스트 컴퓨터는 하나의 IP 주소와 도메인 이름을 가진다
 - ◆ 도메인 네임 서버 (Domain Name Server) 에서 도메인 이름을 IP 숫자로 변환해 주는 역할을 함 → 전화 중계소와 비슷한 역할
 - ◆ 국제인터넷주소자원관리기관 (Internet Corporation for Assigned Names and Numbers, ICANN) 에서 도메인 이름의 기술적 관리, IP 주소 할당 등을 담당

인터넷의 전송방법과 주소체계

- ◆ 도메인 이름 (Domain Name)
 - ◆ 도메인 이름은 계층적으로 구성되어 있다:
호스트이름 . 소속기관 . 단체성격 . 소속국가
(예: plaza.snu.ac.kr)
 - ◆ 미국은 소속 국가 표기가 없음 (국가 개념 없이 미국에서 인터넷 시작)
 - ◆ 단체 성격, 소속 국가의 도메인 이름은 국가나 기관이 관리
 - ◆ 도메인 네임의 구입 → “소속기관” 의 이름을 구입하는 형태
 - ◆ 호스트 이름은 소속기관에서 임의로 할당 가능
(예. media.daum.net, www.daum.net)

인터넷의 전송방법과 주소체계

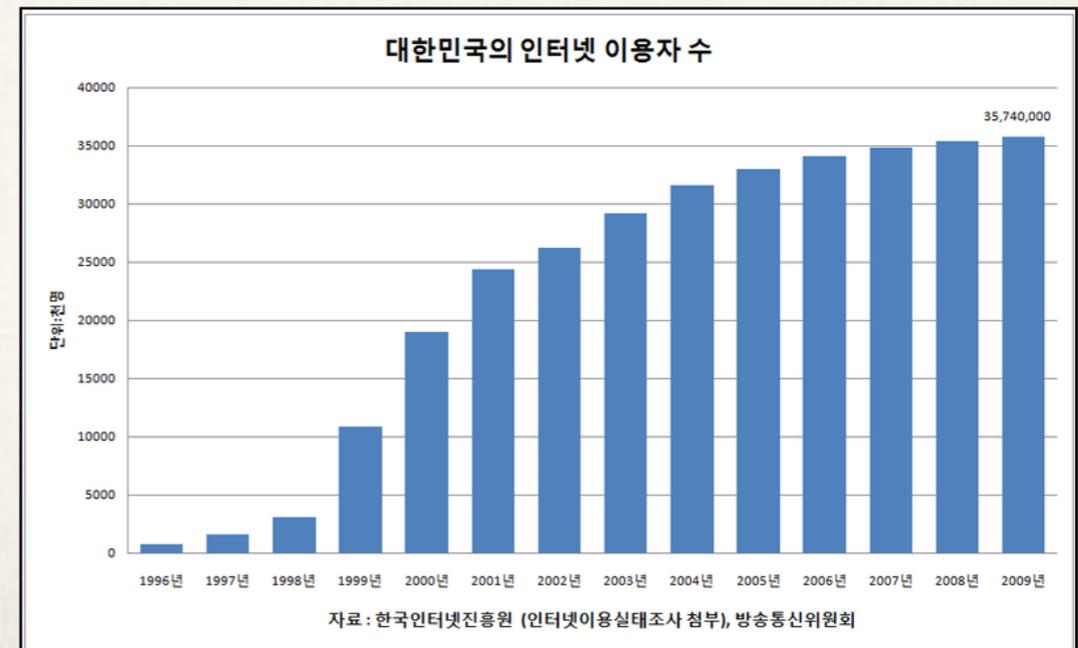
◆ 도메인 이름 체계

기관명	도메인 (미국)	도메인 (기타)
교육기관	edu	ac
사업체	com	co
공공기관	org	or
정부기관	gov	go
네트워크 관련기관	net	ne
연구기관	-	re
국제기구	int	-
국방관련 기관	mil	-

국가	최상위 도메인
Korea	kr
Korea N.	kp
Japan	jp
United Kingdom	uk
France	fr
Germany	de
Canada	ca
Tuvalu	tv

인터넷의 확산

- ◆ 인터넷에 연결되어 있는 컴퓨터의 수
 - ◆ 1969년 ARPANET: 4대
 - ◆ 1996년 1천만대 → 2001년 1억대 → 2008년 5억대
- ◆ 인터넷 사용자
 - ◆ 1996년 1600만 명 → 1998년 1억 명
→ 2008년 약 13억 명 (전세계 20%)
- ◆ 국내의 인터넷 사용인구
 - ◆ 1999년 1000만 명
→ 2006년 3400만 명 (국민의 70%)

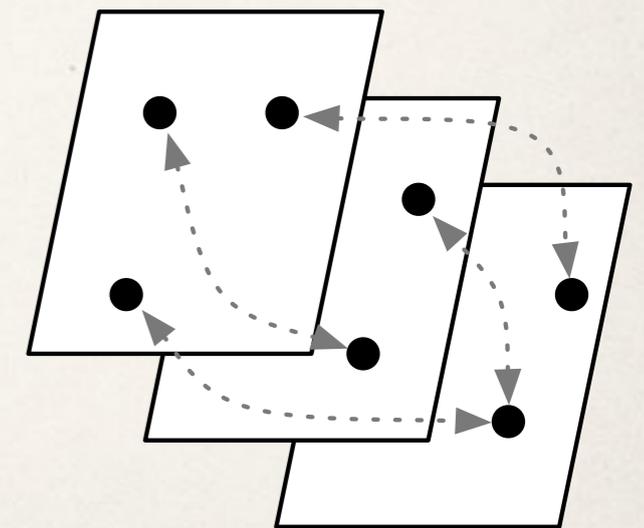


인터넷 서비스

- ◆ 초기의 인터넷 → 교육 및 연구기관 종사자의 정보 전달 수단
 - ◆ 파일 전송: FTP (File Transfer Protocol)
 - ◆ 원격 접속: Telnet
 - ◆ 전자 우편: Email
... 등이 주로 사용
- ◆ 인터넷을 통한 다양한 형태의 정보 전달이 가능해지고, 다양한 계층의 사용자가 늘어나면서 여러가지 새로운 서비스들이 등장

World Wide Web

- ◆ 웹의 기본 개념은 하이퍼 텍스트 (hypertext) 에서 시작
- ◆ 하이퍼 텍스트와 하이퍼 미디어
 - ◆ 하이퍼 텍스트: 텍스트 위주의 문서가 링크로 서로 연결되어 있는 것
 - ◆ 하이퍼 미디어: 텍스트 뿐 아니라 이미지, 그래픽, 사운드, 동영상 등의 정보가 링크로 서로 연결되어 있는 것
 - ◆ 순차적인 연결 구조를 가진 전통적인 책 (text) 과는 달리 하이퍼 텍스트에서는 정보 조각들이 비순차적으로 서로 연결 되어 있다
 - 노드(note): 텍스트 정보의 단위
 - 링크(link): 노드들을 연결하는 포인터



World Wide Web

◆ 하이퍼미디어의 발전 과정

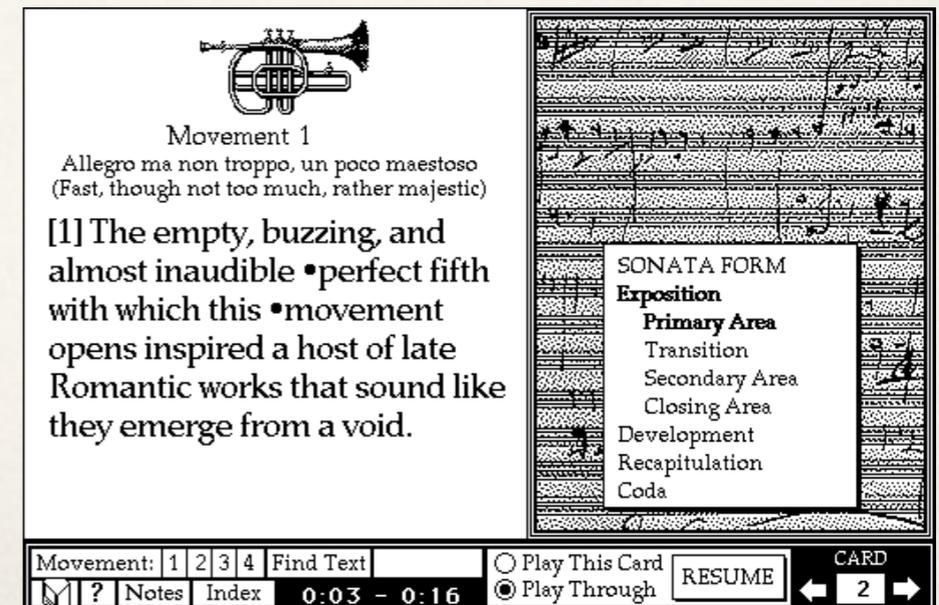
- ◆ 1945년 Vannevar Bush 가 제안한 Memex 에서 최초로 하이퍼텍스트 개념이 소개 → 실제 구현되지는 못함
- ◆ 1965년 Ted Nelson 이 제안한 Xanadu 시스템에서 처음으로 하이퍼텍스트라는 용어 사용 → 세상의 모든 문헌을 하나의 하이퍼텍스트로 연결하겠다는 계획
- ◆ 1977년 MIT의 Andrew Lippman이 Aspen Movie Map 이라는 최초의 하이퍼미디어 시스템을 선보임



<http://www.youtube.com/watch?v=Hf6LkqgXPMU>

World Wide Web

- ◆ 1987년 Apple사에서 HyperCard 개발
 - ◆ 가장 대중적인 하이퍼 미디어 저작도구
 - ◆ 카드에 기반하여, 카드들을 링크로 연결
 - ◆ 하이퍼미디어를 이용한 레퍼런스 등이 등장 (백과사전, 음악사전 등)
- ◆ 하이퍼 텍스트는 원래 텍스트와 이미지의 연결 구조만 지칭 → 컴퓨터의 발전으로 노드가 그래픽, 사운드, 애니메이션, 비디오 등으로 개념 확장
→ 하이퍼 미디어 등장



World Wide Web

- ◆ 하이퍼미디어 시스템의 개념을 도입하여, 인터넷에서 정보를 교환하기 위한 시스템 → 월드 와이드 웹 (WWW)
- ◆ 웹의 특징
 - ◆ 서비스 프로토콜로 HTTP (Hyper Text Transfer Protocol) 을 사용
 - ◆ TCP/IP 상에서 인터넷 어플리케이션이 서비스 되기 위한 규약
 - ◆ HTML (Hyper Text Markup Language) 표준으로 문서를 작성
 - ◆ URL(Uniform Resource Locator)라는 표준 주소표기 방식 사용
 - ◆ 인터넷의 다양한 서비스에 접근하기 위한 수단

World Wide Web

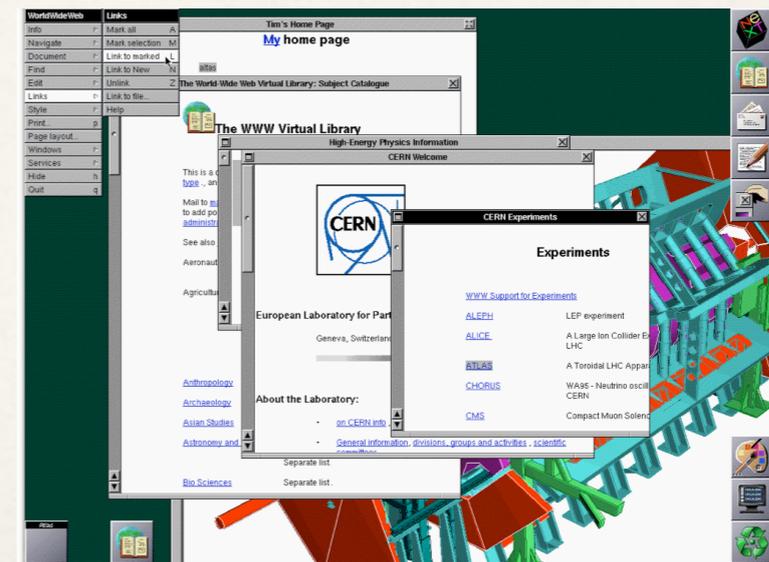
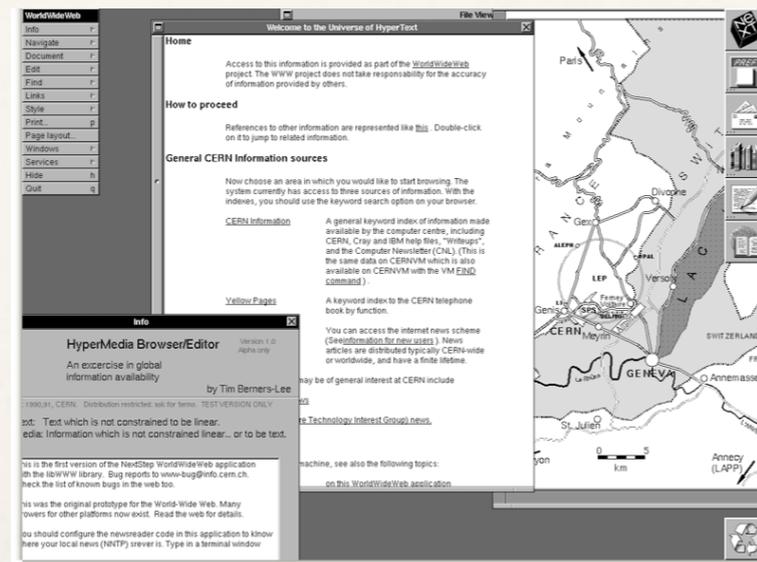
- ◆ URL(Uniform Resource Locator)
 - ◆ 네트워크 상에서 자원이 어디 있는지를 알려주기 위한 규약
 - ◆ 인터넷 뿐 아니라 기타 네트워크에서도 사용
 - ◆ 주소와, 연결을 하기 위한 프로토콜을 같이 제공
 - ◆ 예: `http://www.google.com` → `http`: 프로토콜, `google.com`...: 주소
`ftp://ftp.google.com` → `ftp`: 프로토콜, `ftp.google`...: 주소
 - ◆ 자원의 위치도 URL을 통해 제공
 - ◆ `프로토콜://주소/자원의 경로` → `http://www.google.com/test.jpg`

World Wide Web

◆ 웹브라우저의 발전 → WWW의 발전 과정

◆ Tim Berners-Lee: 인터넷의 아버지

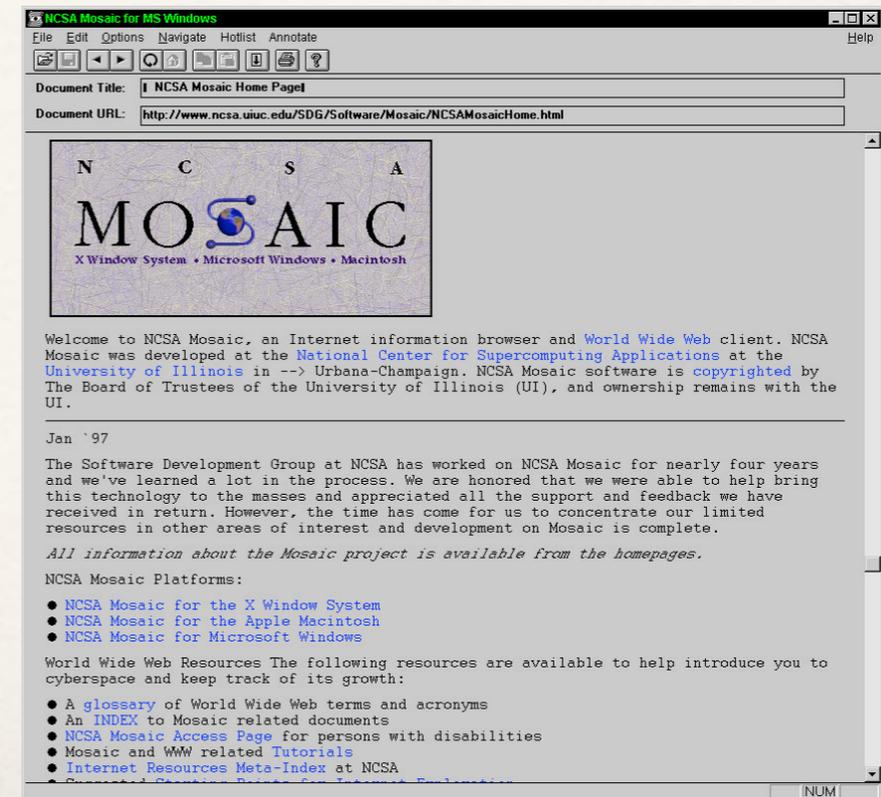
- ◆ CERN에서 Enquire를 개발 (WWW 개념의 시초)
- ◆ 1990, 최초의 하이퍼텍스트 브라우저와 편집기 개발 (NeXT 시스템)
- ◆ 1991년 8월 6일 최초의 웹사이트 런치
<http://info.cern.ch/>
- ◆ 1994년 W3C (World Wide Web Consortium) 창립



World Wide Web

◆ Mosaic 브라우저

- ◆ 1993년 UIUC (일리노이 대학) 의 NCSA 연구센터에서 Marc Andreessen과 Eric Bina가 개발 → 멀티미디어 환경을 지원하는 최초의 웹브라우저
- ◆ 편리한 GUI 제공으로 사용이 편리하고, 다양한 멀티미디어 지원
- ◆ 무료 소프트웨어로 공개:
윈도, 유닉스, 맥 등 다양한
환경 지원
→ 웹의 보급화에 큰 기여



World Wide Web

- ◆ 브라우저 전쟁: Netscape Navigator와 Internet Explorer
 - ◆ 1994년 Marc Andreessen과 Jim Clark 이 Netscape 을 설립, Mosaic을 개량한 브라우저 Navigator를 출시 → 웹의 보급과 함께 인터넷 비즈니스 시대 도래
 - ◆ 1994년 Microsoft가 Mosaic 소스코드에 기반한 Spyglass 사의 브라우저를 라이선싱, Internet Explorer로 시장에 출시
 - ◆ 초기에는 Netscape Navigator 에 시장을 잠식 당했으나, 운영체제의 독점력을 이용, 무료 끼워주기를 한 결과 Navigator 를 물리치고 시장을 장악
 - ◆ 윈도의 데스크탑과 integration → 탐색기 == IE
 - ◆ Unix, Mac OS 등의 플랫폼 지원 X → 2000년 반독점 금지법에 저촉된다는 판결

World Wide Web

- ◆ 브라우저 전쟁: Internet Explorer와 모질라 재단, 기타 브라우저
 - ◆ MS와의 시장경쟁에서 밀리던 Netscape은 시장에서 사라지기 직전 소스코드의 공개를 결정 (1998년) → 모질라 (Mozilla) 재단 설립
 - ◆ 2002년 Mozilla 1.0 브라우저 발표 → 다양한 브랜치의 모질라 브라우저 탄생
 - ◆ 2004년 Firefox 브라우저 발표
 - CSS 2, CSS 3 등의 새로운 웹 표준 기술 지원
 - 탭기능, Plugin/Extension 기능의 강화로 브라우저의 확장성 강조
 - MS IE 시장을 급속도로 잠식

World Wide Web

- ◆ 브라우저 전쟁: Internet Explorer와 모질라 재단, 기타 브라우저
 - ◆ 사파리 (Apple)
 - ◆ MS의 맥용 IE 개발 중단, Netscape 의 몰락 등을 계기로 독자 브라우저 개발 필요성 증대
 - ◆ 오픈소스 진영의 KHTML 엔진 (Konquerer 브라우저) 을 기반으로 Webkit 이라는 새로운 엔진 개발 → 오픈소스
 - ◆ Chrome (Google)
 - ◆ Webkit에 기반한 브라우저 → 빠른 속도, 기능 확장 등으로 점유율 상승
 - ◆ 웹기반의 어플리케이션 플랫폼으로 발전 중

3. 웹(WWW) 관련 기술의 발전 과정

월드와이드웹 (WWW)

- ◆ 콘텐츠를 전달하는 수단으로 시작
 - ◆ 콘텐츠를 저작할 수 있는 기술의 개발 과정
 - ◆ HTML etc.
 - ◆ 다양한 멀티미디어 콘텐츠를 지원하는 방향으로 기술 발전
 - ◆ images, movie clips, interactive media, etc.
- ◆ 다른 인터넷 서비스를 웹 환경으로 흡수하며 성장
 - ◆ Client side programming
 - ◆ Server side programming

HTML의 탄생

- ◆ HTML: Hyper Text Markup Language
 - ◆ 1989년 경, Tim-Berners Lee 에 의해 개발
 - ◆ HTML 은 SGML 을 확장한 형태로 시작 됨
 - ◆ SGML: Standard Generalized Markup Language (ISO 8879:1986 SGML) 은 국제 표준 마크업 랭귀지
 - ◆ 마크업 랭귀지: 데이터를 구조화 하기 위한 언어로 프로그래밍 언어와 구별 함
 - ◆ IBM의 GML 이 원류

SGML

- ◆ 문서의 각 구성 요소가 논리적인 구조를 갖도록 문서의 계층적 구조를 설정할 수 있는 방법 제시
- ◆ 미리 정의된 태그를 선택 사용함으로써 문서 작성 중에 구성요소가 어떻게 보일 것인지, 어떤 성격을 가질 것인지를 지정
- ◆ 태그는 DTD (Document Type Definition) 에 의해 미리 그 규칙이 정의 됨

```
Document: Bungler OED           At: "<entry>"
<entry>
  <hwsec>
    <hwgp>
      <hwlem>bungler</hwlem>
      <pron>b<I>ʊ</I>'ŋɡlɚ</pron>. </hwgp>
      <vfl>Also <vd>b</vd> <vf>bongler</vf>,
      </vfl>
      <etym>f. as prec. + <xra><xlem>-ER</xlem>
    </hwsec>
    <sen>One who bungles; a clumsy unskilful
    <quot>
      <qdat>1533 </qdat>
      <auth>MORE </auth>
      <wk>Answ. Poyson, 3k. </wk>Wks. (1557)
      <qtxt>He is euen but a very bungler.
    </quot>
  </entry>
```

DTD와 그 사용 예

```
<!DOCTYPE people_list [  
  <!ELEMENT people_list (person)*>  
  <!ELEMENT person (name, birthdate?,  
gender?,  
  socialsecuritynumber?)>  
  <!ELEMENT name (#PCDATA)>  
  <!ELEMENT birthdate (#PCDATA)>  
  <!ELEMENT gender (#PCDATA)>  
  <!ELEMENT socialsecuritynumber (#PCDATA)>  
>  
<people_list>  
<person>  
  <name>Fred Bloggs</name>  
  <birthdate>2008-11-27</birthdate>  
  <gender>Male</gender>  
</person>  
</people_list>
```

HTML 의 간단 역사

◆ HTML 1.0 ~ HTML 2.0

- ◆ 1989년 ~ 1991년 경

- ◆ 다양한 포맷으로 문서를 작성하기 보다는, 텍스트 데이터를 HTML을 이용해 구조화하고 각각의 페이지를 하이퍼 링크로 연결하는데 관심을 가짐.

- ◆ 1.0은 SGML 과는 다른 측면에서 개발되었으나, 차후 2.0 부터 SGML 의 기본 구조를 따르기 시작

- ◆ Tim-Berners Lee 에 의해 개발

HTML 의 간단 역사

◆ HTML 3

- ◆ 1995 ~ 1997
- ◆ 브라우저 전쟁 시기 (Netscape vs. IE)
- ◆ 브라우저에 따른 다양한 태그의 난립 → W3C 가 표준을 제정하고, 3.2 버전을 권고안으로서 발표 (대다수의 넷스케이프 비주얼 마크업 태그를 채택)
- ◆ 현재 사용되는 태그의 기본안이 거의 확정

HTML 의 간단 역사

◆ HTML 4

- ◆ 1998 ~ 1999
- ◆ 브라우저 전쟁 끝남 (IE의 주도)
- ◆ W3C 가 웹표준을 제시
- ◆ 데이터 구조와 스타일을 분리하기 위해 HTML 과 CSS 로 분리 → 이전 버전에서 채택되었던 넷스케이프의 비주얼 스타일 태그들이 사용되지 않게 됨.
- ◆ Strict, Transitional, Frameset 등의 문서형태가 제공

HTML의 간단 역사

◆ XHTML 1.0

- ◆ 2000 ~
- ◆ 다양한 디바이스에서 웹이 사용됨에 따라 확장성이 중요
- ◆ HTML 4 를 기반으로 하여 XML 의 확장성을 도입
(새로운 태그를 사용자가 정의해서 사용 가능)
- ◆ HTML 문법에 비해 엄격 (e.g., 태그 사용 규칙 등)
- ◆ W3C → HTML 의 버전 업데이트는 더 이상 없다고 발표
 - ◆ XHTML 의 문제점 노출
 - ◆ XHTML 과 HTML4 간의 문법적인 혼란, 브라우저 지원 문제 등 (IE 8)

HTML 의 간단 역사

- ◆ XHTML 의 특징 (엄격한 문법 체크)
 - ◆ 종료 태그가 반드시 있어야 한다 (`<p> ~ </p>` → O)
 - ◆ 잘못된 중첩을 허용하지 않는다
(`<p>random <i>text</p></i>` → X)
 - ◆ 모든 요소(element)와 속성(attribute)은 소문자로 표기되어야 한다
 - ◆ 속성의 값(value)은 항상 따옴표로 감싸주어야 한다
(`<p align = "center">`)
 - ◆ 비어있는 태그도 종료되어야 한다
(`
, <hr>` → `
, <hr />`)
- ◆ 느슨한 문법 구조인 HTML 에 비해 엄격하게 체크하기 때문에 다양한 브라우저에서 정확하게 문서가 표시될 수 있다 → 단, 브라우저가 표준을 따랐다면...

HTML 의 간단 역사

◆ HTML 5

- ◆ 2009 ~
- ◆ 현재 표준화 작업 진행 중
- ◆ 새로운 브라우저 전쟁 시대 도래 (IE vs. Webkit vs. Gecko etc)
- ◆ XHTML 의 문제점을 노출로 HTML 버전의 발전 필요성 증대
- ◆ 웹 어플리케이션 개발의 필요성 증대
 - HTML 기술이 콘텐츠의 전달 플랫폼이 아닌 어플리케이션 구동 플랫폼으로 진화할 필요성
 - WHATWG (Web Hypertext Application Technology Working Group) 결성, 새로운 스펙 구성 시작

웹 구조의 이해

- ◆ 월드와이드웹 (World Wide Web: WWW) 이란?
 - ◆ 인터넷에 연결된 컴퓨터들을 통해 사람들이 정보를 공유할 수 있는 전세계적인 정보 공간
 - ◆ 웹은 인터넷에서 사용되는 여러 서비스 중 하나 (예. 메일, FTP 등)
 - ◆ 하이퍼텍스트로 연결된 페이지들을 웹서버를 통해 서비스 하고 웹 브라우저를 통해 모니터로 출력하는 형태
 - ◆ 사용자는 하이퍼링크를 따라 다른 문서로 이동할 수 있다.
 - ◆ 그물망 처럼 링크로 연결되어 있다고 해서 web 이라 부름
 - ◆ 일련의 웹페이지가 모여 있는 공간을 웹사이트로 지칭

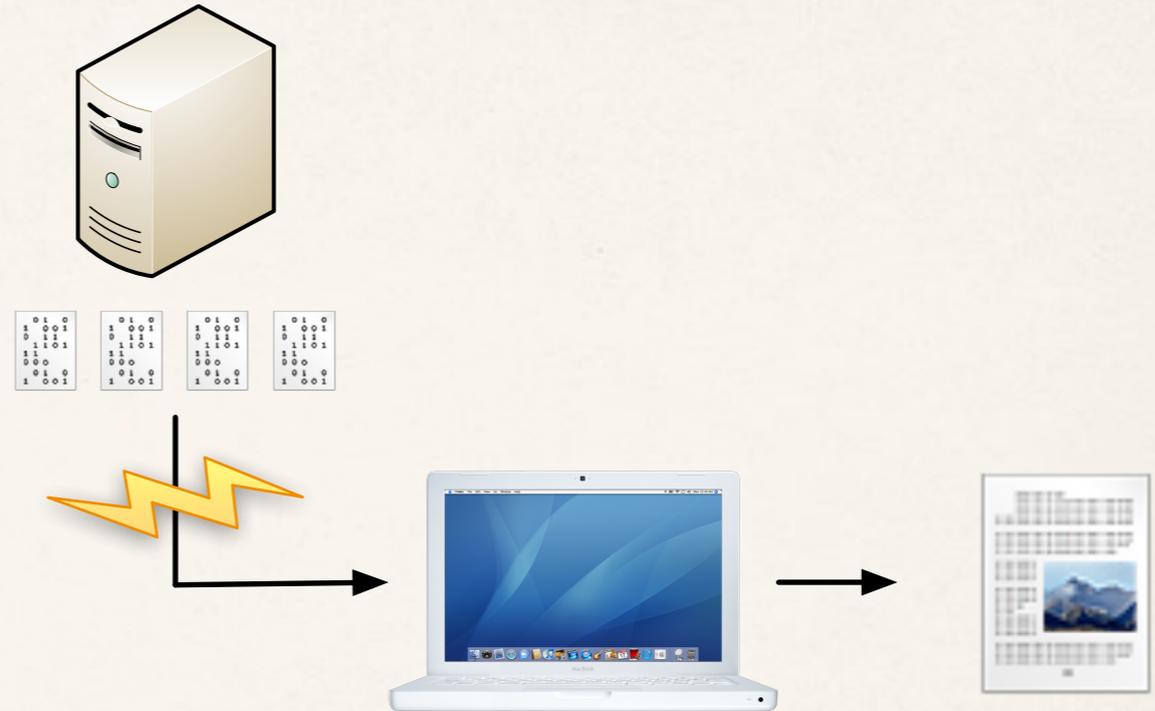
웹 구조의 이해

◆ 웹페이지

- ◆ 웹페이지는 HTML 이라는 언어로 만들어진 일종의 프로그램
- ◆ 웹페이지는 텍스트, 이미지, 동영상, 인터랙티브 미디어 (플래시) 등을 담을 수 있다
- ◆ 웹 브라우저가 HTML 의 내용을 해석하고 사용자가 볼 수 있는 형태로 렌더링 함
 - ◆ 렌더링 엔진에 따라 브라우저 분류: Trident(Internet Explorer), Webkit (Safari, Chrome), Gecko (Firefox)



웹 구조의 이해



◆ 웹 서버

- ◆ 일련의 웹페이지를 저장하고 사용자에게 제공해 주는 공간
- ◆ 프로토콜 (규약, 절차)에 따라 웹 주소 해석 (http)
 - ◆ <http://www.example.com/path/file.html>
 - ◆ GET /path/file.html HTTP/1.1
Host: www.example.com
 - ◆ /var/www/path/file.html (리눅스 서버의 경우)

웹 서버의 역할

- ◆ 웹 페이지를 클라이언트 사용자 (접속자)에게 제공
- ◆ 프로토콜 제공
 - ◆ HTTP 1.1, HTTP 1.2 등
- ◆ 사용자 접속 load 컨트롤
- ◆ 접속 기록 관리

- ◆ 안정된 웹사이트 운영

웹 서버의 종류

Vendor	Product	Websites Hosted (mill)	Percent	note
Apache	Apache	111	54%	PHP, RoR, JAVA
Microsoft	IIS	50	24%	ASP
Igor Sysoev	nginx	16	8%	
Google	GWS	15	7%	
lighttpd	lighttpd	1	0%	RoR

웹 관련 기술

- ◆ 서버 사이드 : 웹서버

- ◆ 서버에서 자료를 보관, 요청에 따라 클라이언트로 전달

- ◆ 클라이언트 사이드 : 브라우저

- ◆ 서버에 보관된 자료를 요청하여 클라이언트 (브라우저) 에 표시 (렌더링)
- ◆ 자료는 HTML 로 기술되어 있어야 하고, HTML 로 만들어진 자료를 브라우저가 받아 내부 규약에 맞게 화면에 렌더링

웹 관련 기술

◆ 서버 사이드 : 웹서버

◆ 자료의 저장 - 하드디스크 vs. 데이터 베이스

- ◆ 하드디스크: 단순히 자료를 서버의 물리적인 위치에 저장
- ◆ 데이터베이스: 자료를 구조화 시켜 효율적으로 검색, 참조가 가능한 형태로 데이터베이스에 저장 (Oracle, MySQL, MS SQL 등)

◆ 자료의 처리 방법

- ◆ static: <http://www.naver.com/test.html> → 서버에 저장된 자료 단순 호출
- ◆ dynamic: <http://www.youtube.com/watch?v=B4GBqb5jRZ0> → DB에 동영상 아이디를 전달해 주고 해당 자료 호출

웹 관련 기술

◆ 서버 사이드 : 웹서버

- ◆ dynamic 하게 자료를 호출하기 위해서는 서버 쪽에서 DB와 연동하여 사용자의 request 를 처리하는 작업이 필요 → Server-side scripting
- ◆ 사용자의 request에 따라 데이터를 가공하여 새로운 웹페이지 작성 → 응답
 - ◆ ASP
 - ◆ Java via JavaServer Pages (*.jsp)
 - ◆ PHP (*.php)
 - ◆ Perl (*.pl)
 - ◆ Python (*.py)
 - ◆ Ruby (*.rb)

웹 관련 기술

◆ 클라이언트 사이드 : 브라우저

- ◆ 웹브라우저가 점차 다양한 인터넷 서비스 및 데스크탑 어플리케이션의 기능을 대체
- ◆ 브라우저의 HTML 렌더링 기능만으로는 이러한 요구를 수용하기 어려움
 - ◆ 동영상: 다양한 포맷의 동영상을 웹 브라우저 안에서 재생
→ 동영상 파일은 웹브라우저가 기본으로 제공하는 미디어 형태가 아님
 - ◆ 게임: 인터랙티브한 게임을 웹 브라우저 안에서 플레이
→ dynamic interaction 을 위해서는 데스크탑 어플리케이션 수준의 어플리케이션 구동 플랫폼이 필요
- ◆ 플러그인/익스텐션으로 브라우저의 기능 보강 (e.g., Quicktime, Flash, Java 등)

웹 관련 기술

◆ 클라이언트 사이드 : 브라우저

- ◆ Flash, Java 등의 등장 → 브라우저가 지원하지 않는 형태의 멀티미디어를 실행하는 플랫폼 → RIA 로 발전
- ◆ RIA (Rich Internet Application) 의 등장
 - ◆ 전통적인 데스크탑 어플리케이션의 기능과 특징을 웹 상에서 구현한 어플리케이션
 - ◆ 브라우저의 플러그인 형태로 설치 됨 → 해당 어플리케이션이 로딩되면 플러그인을 통해 어플리케이션 구동 플랫폼이 실행되고 그 위에서 RIA 가 실행
 - ◆ Adobe Flash / Adobe Air
 - ◆ Microsoft Silverlight
 - ◆ Java / JavaFX

웹 관련 기술

◆ 클라이언트 사이드 : 브라우저

- ◆ Client-side scripting: 웹 페이지 내에서 사용자의 요구에 의해 콘텐츠를 다이내믹하게 변경하는 기술
- ◆ Server-side scripting 이 동적으로 HTML 및 데이터를 만들어 준다면, Client-side scripting 은 이미 만들어진 HTML 및 데이터를 동적으로 구성하도록 해줌
 - ◆ JavaScript: Netscape 에서 개발, 발전
 - ◆ VBScript: Microsoft 의 Visual Basic 언어를 사용, IE 에서 사용.

4. 웹 2.0

웹 2.0

◆ 웹 2.0의 탄생 배경

- ◆ 인터넷 관련 산업의 고속성장 → 2000년 대 초, 닷컴(.com) 버블의 붕괴
- ◆ 많은 인터넷 업체 들이 도산
 - ◆ kosmo.com - 인터넷 주문 즉시 1시간 내 배송 (자전거)
 - ◆ IPO 에 집착, 수익 모델 없었음.
 - ◆ 기술 없이 회사 가치에 근거한 비합리적인 투자에 의해 회사가 운영
- ◆ 구글, 아마존, 이베이 → 닷컴 붕괴에서도 살아남은 기업
 - ◆ 기술적 우위
 - ◆ 안정된 수익모델 / 고객 확보

웹 2.0

- ◆ 2004년 10월 O'Reilly 사가 주최한 컨퍼런스에서 처음 사용
 - ◆ 10년 간 발생한 웹의 환경 변화와 발전 방향을 정의 → 웹 2.0이라 지칭
- ◆ 웹 2.0
 - ◆ 닷컴 붕괴의 대명사
 - ◆ 살아 남은 기업의 특징을 분석하기 위한 용어
 - ◆ 웹 비즈니스에서 살아 남기 위한 일종의 지침
 - ◆ 기술적인 버전 업의 의미가 아님

웹 2.0

◆ 웹 2.0 시대 인터넷 기업의 특징

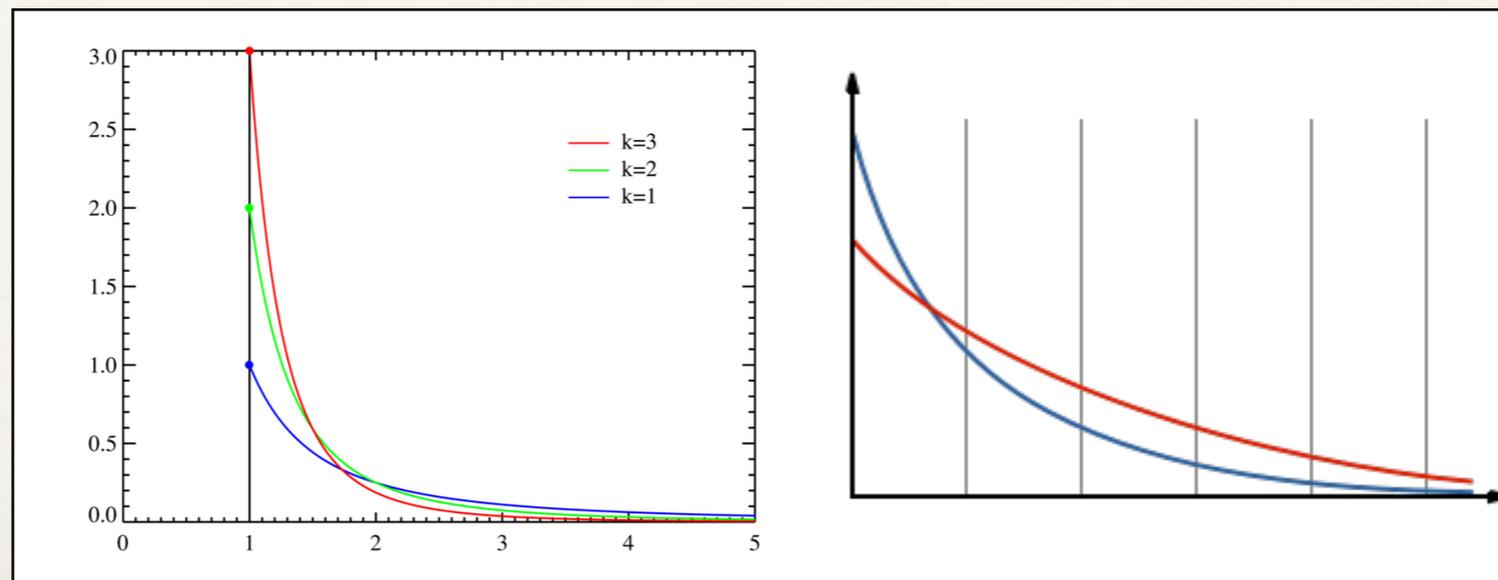
◆ 파레토 법칙이 적용되지 않는다

→ 전체 결과의 80%가 전체 원인의 20%에서 일어나는 현상

→ 대부분의 오프라인 기업: 팔리는 20%의 상품이 전체 매출의 80%를 차지한다

◆ Amazon.com

→ 20%의 인기 상품의 매출보다 하위 80%의 평범한 상품의 매출 합계가 더 크다 (롱테일 (Long Tail) 법칙)



웹 2.0

- ◆ 웹 2.0 비즈니스가 롱테일의 특성을 나타내는 이유
 - ◆ 검색과 필터링
 - ◆ 수요와 공급의 연결을 용이하게 해줌
 - ◆ 고객이 원하는 것을 정확하게 찾을 수 있는 시스템 구축 → display model 과 비교
 - ◆ 사용자 참여를 통해 정보를 구조화 → 태그, 키워드, 추천
 - ◆ Aggregation (자동 수집)
 - ◆ RSS 수집 → 기존의 방식으로는 얻을 수 없던 광범위한 정보의 접근이 가능
 - ◆ 소규모의 정보 제공자가 대규모 뉴스 포털과 대등한 위치에서 경쟁 → 롱테일 가능성

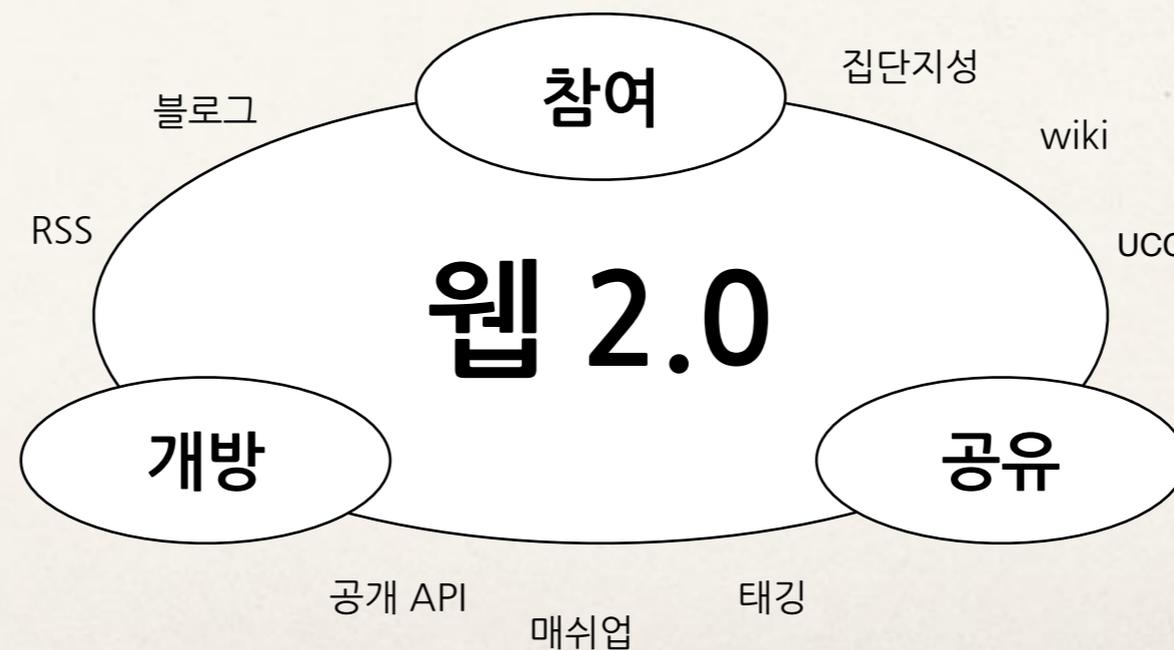
웹 2.0

◆ 롱테일 비즈니스

- ◆ 기술 혁신을 통해 “꼬리” 부분을 머리와 비교해 보다 두터워지고 길어질 수 있도록 도와주는 서비스
 - ◆ 콘텐츠의 생산: 블로그, GarageBand 와 같은 저작툴, 팟 캐스팅
 - ◆ 콘텐츠의 공유: 블로그, 메타블로그 (RSS), Wiki, UCC
 - ◆ 콘텐츠의 추천, 검색: 태그, 키워드
 - ◆ 콘텐츠의 홍보: AdSense 와 같은 광고 도구 (소규모 광고주들에게 어필)

웹 2.0의 기본 개념

- ◆ 웹 2.0 시대에 인터넷 사용방법의 큰 변화
 - ◆ 플랫폼으로서의 웹
 - ◆ 집단 지성 (Collective Intelligence) 을 활용한 콘텐츠의 제공, 공유



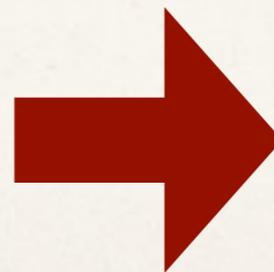
웹 2.0의 기본 개념

- ◆ 웹 2.0 콘텐츠 생성

- ◆ 개인의 참여와 서로간의 상호 작용으로 이루어 짐
- ◆ 사용자 집단의 능동적인 참여와 공유에 의해서 더욱 가치가 증가
- ◆ UCC (User Created Contents), UGC (User Generated Contents)

- ◆ 사용자 → 콘텐츠의 공급자인 동시에 소비자 → Prosumer

- ◆ 블로그
- ◆ 위키피디아
- ◆ 플릭커, 유튜브



태깅 → 집단 지성에 의해 콘텐츠를 분류, 소비촉진

웹 2.0의 기본 개념

◆ 웹 2.0 → 개방적

- ◆ 전통적인 웹사이트는 자신의 데이터를 폐쇄적으로 운영
- ◆ 웹 2.0 → 데이터를 독점하지 X, 모든 사람이 공유하여 사용할 수 있는 플랫폼 제공
- ◆ 뉴스 포털: 정보를 포털 안에 묶어 두는 독점 형태
- ◆ 블로그, 메타블로그: 여러 사람들에 의해 생성된 콘텐츠를 수집하고 유통
→ RSS 의 사용
- ◆ 공개 API: 서비스를 공유하기 위한 개방형 포맷
 - ◆ 구글, 야후 등은 자신의 데이터베이스와 서비스의 API 를 제공, 다른 어플리케이션에서 활용하도록 함 → mashup

웹 2.0 기술

- ◆ 웹 2.0 은 특정한 기술의 표준을 지칭하는 것이 아님
- ◆ XML 로 표현된 콘텐츠를 다양한 기술을 통해 표현하는 것이 웹 2.0 의 기술
- ◆ 대표적인 웹 2.0 기술
 - ◆ 블로그, RSS
 - ◆ 태깅
 - ◆ 공개 API (Open API)
 - ◆ Mashup
 - ◆ AJAX 등

웹 2.0 기술

◆ 블로그와 RSS

- ◆ 블로그 = Web + Log 의 합성어
- ◆ 자신의 콘텐츠를 웹에 기록하여 다른 사람과 공유하는 서비스
- ◆ RSS feed 를 통해 업데이트된 블로그 정보를 구독할 수 있다
- ◆ RSS: Really Simple Syndication, Rich Site Summary
 - ◆ 웹 사이트 간의 콘텐츠를 교환하기 위한 XML 기반의 표현형식
 - ◆ 텍스트와 이미지 위주의 블로그 사이트의 정보를 교환하는데 사용
 - ◆ 최근, 음악/동영상 콘텐츠 교환에도 사용됨 → podcast

웹 2.0 기술

◆ Social Tagging and Foxonomy

- ◆ 이전의 웹사이트: 카테고리에 의해 정보를 분류 (야후 디렉토리)
- ◆ 소셜태깅: 사용자가 키워드를 임의의 태그로 붙이고 이들을 검색하는 기법
- ◆ Foxonomy = Folk + Taxonomy (사용자에 의한 + (정보) 분류방법)
- ◆ 웹사이트나 이미지 등 다양한 자료에 원하는 만큼의 태그를 붙일 수 있다.
 - ◆ Flickr.com, del.icio.us 등이 대표적인 사이트
- ◆ Tag Cloud: 사용자들이 달아놓은 태그를 한눈에 보이도록 하는 비주얼라이제이션



웹 2.0 기술

- ◆ Open API & Mashup

- ◆ API: Application Programming Interface

- ◆ 프로그래밍을 할 때, 데이터나 서비스에 손쉽게 접근하도록 제공해 주는 인터페이스

- ◆ Open API

- ◆ 특정 사이트에서 제공하는 데이터, 서비스에 접근할 수 있도록 공개되어 제공하는 인터페이스

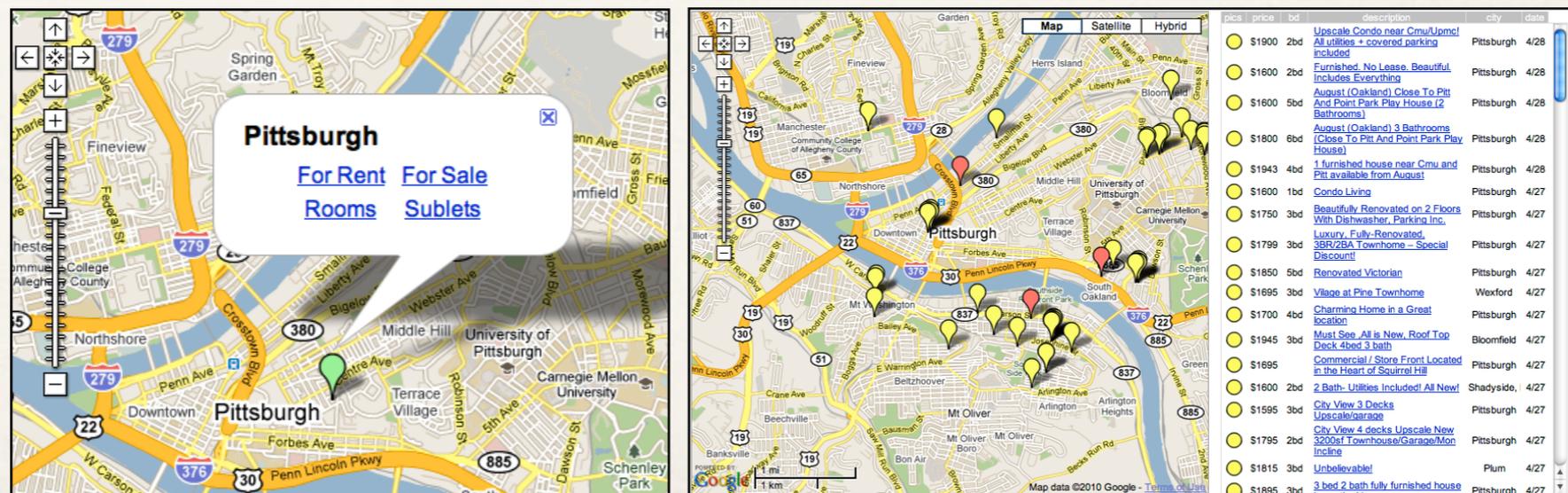
- ◆ 예: 구글맵 api → 구글의 지도 정보를 손쉽게 접근할 수 있게 해준다

```
var map = new GMap2(document.getElementById("map_canvas"));  
map.setCenter(new GLatLng(37.4419, -122.1419),  
13);
```

웹 2.0 기술

◆ Open API & Mashup

- ◆ Mashup: 두가지 이상의 웹서비스를 조합하여 새로운 웹 서비스를 제공하는 방법
- ◆ 최초의 매쉬업 서비스: HousingMaps (<http://www.housingmaps.com/>)



The screenshot displays the HousingMaps interface. On the left, a map of Pittsburgh is shown with a search overlay that includes the text "Pittsburgh" and links for "For Rent", "For Sale", "Rooms", and "Sublets". On the right, a list of property listings is displayed with columns for price, bedrooms, description, city, and date. The listings include details such as "Upscale Condo near Cmu/Upmc", "August (Oakland) Close To Pitt", and "Luxury Fully-Renovated 3BR/2BA Townhome - Special Discount!".

price	bed	description	city	date
\$1900	2bd	Upscale Condo near Cmu/Upmc! All utilities + covered parking included	Pittsburgh	4/28
\$1600	2bd	Furnished, No Lease, Beautiful, Includes Everything	Pittsburgh	4/28
\$1600	5bd	August (Oakland) Close To Pitt And Point Park Play House (2 Bathrooms)	Pittsburgh	4/28
\$1800	6bd	August (Oakland) 3 Bathrooms (Close To Pitt And Point Park Play House)	Pittsburgh	4/28
\$1943	4bd	1 furnished house near Cmu and Pitt available from August!	Pittsburgh	4/28
\$1600	1bd	Condo Living	Pittsburgh	4/27
\$1750	3bd	Beautifully Renovated on 2 Floors With Dishwasher, Parking Inc.	Pittsburgh	4/27
\$1799	3bd	Luxury Fully-Renovated, 3BR/2BA Townhome - Special Discount!	Pittsburgh	4/27
\$1850	5bd	Penn Lincoln Pkwy	Pittsburgh	4/27
\$1695	3bd	Village at Pine Townhome	Wexford	4/27
\$1700	4bd	Charming Home in a Great location	Pittsburgh	4/27
\$1945	3bd	Must See - All is New, Roof Top Deck, 4bed 3 bath	Bloomfield	4/27
\$1695		Commercial / Store Front Located in the Heart of Squirrel Hill	Pittsburgh	4/27
\$1600	2bd	2 Bath - Utilities Included! All New!	Shadyside	4/27
\$1595	3bd	City View 3 Decks Upscale Garage	Pittsburgh	4/27
\$1795	2bd	City View 4 decks Upscale New 3200sf Townhome/Garage/Mon Incline	Pittsburgh	4/27
\$1815	3bd	Unbelievable!	Plum	4/27
\$1895	3bd	3 bed 2 bath fully furnished house	Pittsburgh	4/27

대표적인 웹 2.0 서비스

- ◆ 구글 (google.com)

- ◆ 1998년 스탠포드의 Lawrence Page 와 Sergey Brin 이 설립
- ◆ 당시 최고의 검색엔진이었던 yahoo.com 검색보다 나은 검색

기술을 제공

- ◆ PageRank 개념 적용

- ◆ 해당 페이지를 링크로 참조하는 페이지들의 중요도를 가중치로 계산하는 방식의 검색 알고리즘 제공

- ◆ AdSense 도입 → 주 수입원

- ◆ 페이지의 콘텐츠를 검색하여 관련된 기업의 광고를 보여주는 광고기술 (검색기 반)

→ 롱테일 비즈니스에 큰 도움

대표적인 웹 2.0 서비스

◆ iTunes 와 팟캐스팅

- ◆ iTunes: 대표적인 콘텐츠 유통 플랫폼
- ◆ RSS 를 멀티미디어에 접목하여 사용자가 제작한 음악/영상 콘텐츠의 유통을 시도
- ◆ 개인 방송국, UCC 의 새로운 장을 열다

대표적인 웹 2.0 서비스

- ◆ 위키피디아 (wikipedia)
 - ◆ 위키 (wiki)
 - ◆ 일종의 공개 게시판
 - ◆ 다양한 사용자가 참여하여 웹페이지의 내용을 자유롭게 추가할 수 있는 커뮤니케이션 도구 (수정/삭제/관리가 자유로움)
 - ◆ wikipedia 를 통해 사용자 참여 백과사전으로 진화
 - ◆ 집단 지성으로 작성되는 UCC의 대표적인 사례
 - ◆ 콘텐츠의 양, 질적인 면에서 브리태니카에 필적 → 신속함은 당연히 브리태니카가 보다 우월

대표적인 웹 2.0 서비스

- ◆ YouTube, Flickr
 - ◆ UCC (UGC) 의 대표적인 사이트
 - ◆ 사용자가 직접 작성한 동영상, 사진을 공유
 - ◆ 국내에서도 TV팟 등이 비슷한 서비스를 제공
- ◆ SNS (Social Network Service)
 - ◆ 사용자들이 커뮤니티를 형성하여 서로 참여하는 사이트
 - ◆ MySpace, Facebook, Twitter, digg 등이 대표적인 사이트
 - ◆ 커뮤니티와 차별화 해야 (커뮤니티 → 그룹 중심, SNS → 개인 이 네트워크 형성)

클라우드 컴퓨팅

- ◆ 데이터가 개인 컴퓨터에 저장되는 것이 아니라 인터넷 상의 서버에 영구적으로 저장되고 데스크탑, 노트북, 휴대용 기기 등과 같은 클라이언트에서 서비스 형태로 제공되는 컴퓨팅 기법
- ◆ 어딘가에서 필요한 컴퓨팅 자원이 할당되어 작업을 수행할 수 있다
 - ◆ 예: google docs, calendar 등 → 문서를 어느 곳에서나 액세스 가능
- ◆ 최근의 모바일 환경과 관련하여 차세대 컴퓨팅 기술로 각광을 받고 있다.
 - ◆ 구글 CEO 에릭 슈미츠 → 웹 3.0은 구름 속에 있는 웹 어플리케이션의 조합
 - ◆ 어플리케이션, 데이터 → 서비스
 - ◆ 매시업 등으로 더욱 풍부한 서비스 가능

클라우드 컴퓨팅

- ◆ 비판: Ricard Stallman

- ◆ “자신의 컴퓨팅 활동을 오직 웹 기반 서비스만 쓰는 것은 자신의 통제권을 잃는 것이다. 그것은 상용 소프트웨어를 쓰는 것 만큼 나쁜 일이다. 컴퓨팅에 있어 자유를 제공하는 소프트웨어를 이용해서 여러분의 PC에서 여러분의 작업을 하라. 여러분이 상용 소프트웨어를 쓰거나 누군가의 웹서버에 있는 애플리케이션 기반 사용하지 말아야 하는 이유이다.”

- ◆ 데이터에 대한 통제권을 잃는 것을 우려 (대기업에서 정보 통제권 가짐)

- ◆ 보안의 문제

- ◆ 안정성 문제 → 인터넷 장애와 같은 인프라 장애 → 데이터 액세스 불가

Web 2.0과 정보검색

- ◆ 사용자 참여를 통한 집단지성의 반영

- ◆ James Surowiecki, The wisdom of crowds

- ◆ 대중들이 일부 뛰어난 천재들보다 뛰어나기 위한 조건

- ◆ 첫째는 다양성(diverse)이며, 둘째로는 집단이 탈 중앙화(decentralized)되어 있어야 하고, 셋째로는 구성원의 의견이 정리되고 모여져 하나의 결정을 만들어낼 수 있는 방법론(way of summarizing people's opinions into one collective verdict) 을 가지고 있어야 하며, 마지막으로 다른 구성원의 의견에 영향 받지 않도록 구성원이 상호 독립적이어야 한다.

- ◆ 위키피디아, del.icio.us 등과 같이 사용자들이 참여를 통해 정보를 직접 만들고, 운영하고, 검증하며, 색인, 분류 등의 과정을 통해 참여형 정보검색 서비스를 만들게 될 것

Web 2.0과 정보검색

- ◆ 사용자 개인에게 최적화된 정보검색 서비스 제공
 - ◆ 데스크탑 시대와 달리 모바일은 개인정보를 기반으로 정보를 필터링하는 것이 수월
 - ◆ 모든 정보가 온라인을 통해 유통
 - ◆ digital footprints
 - ◆ social network
 - ◆ agent 기반의 개인화 정보검색 서비스의 제공 → Siri

Web 2.0과 정보검색

- ◆ 사용자 인터페이스의 진화
 - ◆ 현재의 정보검색은 키워드 또는 질의문의 입력을 리스트의 형태로 받아보는 단방향커뮤니케이션 형태
 - ◆ 정보검색에서 사용자의 의도나 추가적인 정보를 추출하기 위해 대화형 인터페이스나 태깅 등의 방식이 적용될 것으로 예상 됨
→ Siri
 - ◆ 검색 결과의 태깅, 추천 등을 통해 사용자 주관적인 정보의 계열화가 가능한 인터페이스가 도입 → google +1 등이 시도 중

Questions?
