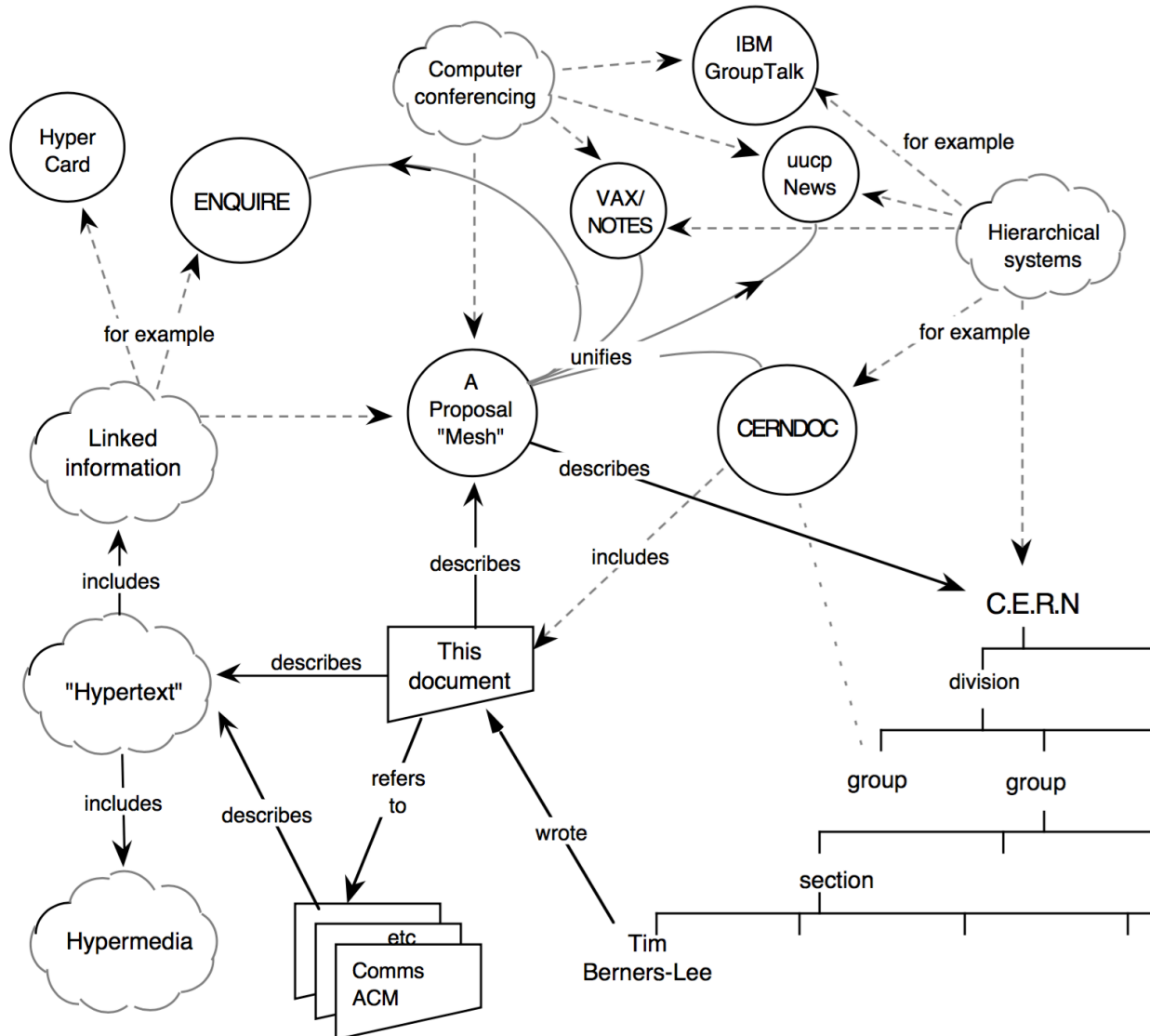


# Web

Lecture 23, Computer Networks (198:552)

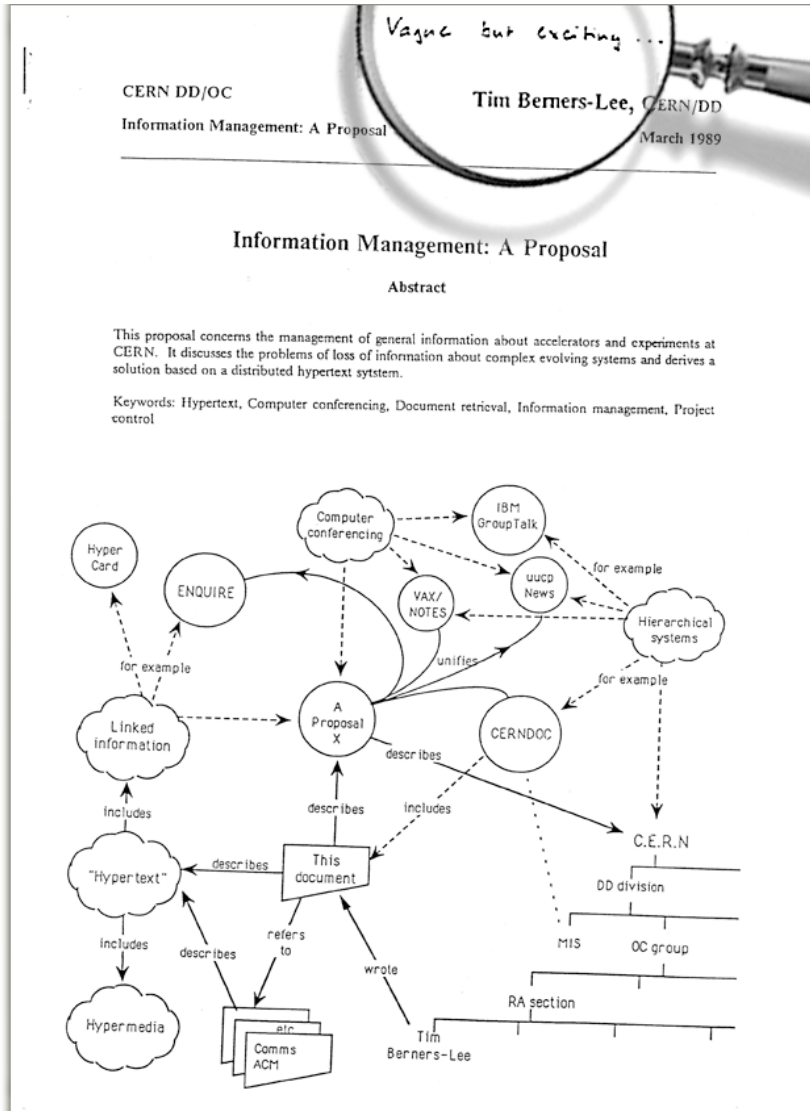
# Information management: A proposal



# Sir Tim Berners-Lee

<https://cds.cern.ch/record/369245/files/dd-89-001.pdf>

# Information management: A proposal



## “Vague, but exciting”

# Prescient ideas in the original proposal

- A distributed system linking content together
- Non-linear text navigation (“hypertext”)
- Access from different kinds of systems
- Access to different kinds of data and media
- The possibility of automated search and analysis
- Highly dynamic content and live links

# Structure of a web page: Rendered

A large, empty search input field with a light gray border and a subtle shadow. On the right side of the field is a small, colorful microphone icon, indicating voice search functionality.

Google Search

I'm Feeling Lucky

# Structure of a web page: Source code

```
decoration:underline}.hRvfYe #fsettl a:hover{text-decoration:underline}.hRvfYe a:hover{text-decoration:underline}.fbar p{display:inline}.fbar a,#fsettl{text-decoration:none;white-space:nowrap}.fbar{margin-left:-27px}.Fx4vi{padding-left:27px;margin:0 !important}.In26Ec{padding:0 !important;margin:0 !important}#fbarcnt{display:block}.smiUbb img{margin-right:4px}.smiUbb a,.M6hT6 #swml a{text-decoration:none}.fmulti{text-align:center}.fmulti #fsr{display:block;float:none}.fmulti #fuser{display:block;float:none}#fuserm{line-height:25px}#fsr{float:right;white-space:nowrap}#fsl{white-space:nowrap}#fsettl{background:#fff;border:1px solid #999;bottom:30px;padding:10px 0;position:absolute;box-shadow:0 4px 0px rgba(0,0,0,0.2);box-shadow:0 2px 4px rgba(0,0,0,0.2);text-align:left;z-index:104}#fsettl a{display:block;line-height:44px;padding:0 20px;text-decoration:none;white-space:nowrap}.p2Kmc{padding:8px 16px;margin-right:10px}.B4GxFc{margin-left:166px}#fbar{background:#f2f2f2;border-top:1px solid #e4e4e4;line-height:40px;min-width:980px}.yBWvle{margin-left:166px}.fbar p,.fbar a,#fsettl,#fsettl a{color:#5f6368}.fbar a:hover,#fsettl a:hover{color:#333}.fbar{font-size:small}#fuser{float:right}.smiUbb{margin-left:166px;line-height:15px;color:#5f6368;}</style><style>#fsl{margin-left:30px}#fsr{margin-right:30px}.fmulti #fsl{margin-left:0}.fmulti #fsr{margin-right:0}</style><div class="EvHmz hRvfYe" id="fbar"><div class="fbar"><div class="b2hzT"> <style>.b0KoTc{color:rgba(0,0,0,.54);padding-left:27px}.b0KoTc .UnWzvd{margin-left:13px}.Q8LRLc{font-size:15px}.b0KoTc{margin-left:30px;text-align:left}</style> <div class="b0KoTc"> </div> </div><span id="fsr"><a class="Fx4vi" href="https://www.google.com/intl/en_us/policies/privacy/?fg=1" ping="/url?sa=t&source=webhp&url=https://www.google.com/intl/en_us/policies/privacy/%3Ffg%3D1&ved=0ahUKEwi74_rVjvLeAhVhkeAKHYVjCK0Q8awCCA0">Privacy</a><a class="Fx4vi" href="https://www.google.com/intl/en_us/policies/terms/?fg=1" ping="/url?sa=t&source=webhp&url=https://www.google.com/intl/en_us/policies/terms/%3Ffg%3D1&ved=0ahUKEwi74_rVjvLeAhVhkeAKHYVjCK0Q8awCCA0">Terms</a><span style="display:inline-block;position:relative"><a class="Fx4vi" href="https://www.google.com/preferences?hl=en" id="fsettl" aria-controls="fsettl" aria-expanded="false" aria-haspopup="true" role="button" onclick="fsettl.toggle()" ping="/url?sa=t&source=webhp&url=https://www.google.com/preferences?hl=en" id="fsettl" aria-controls="fsettl" aria-expanded="false" aria-haspopup="true" role="button" onclick="fsettl.toggle()" ping="/url?
```

- HTML: text content and (basic) formatting
  - `<a href="...">link text</a>`
- JavaScript: content rendering
  - `document.getElementById("x")`
- CSS: styling of an entire page
  - `link { color: red }`
- Media: images, videos, etc.

# Structure of a web page: DOM

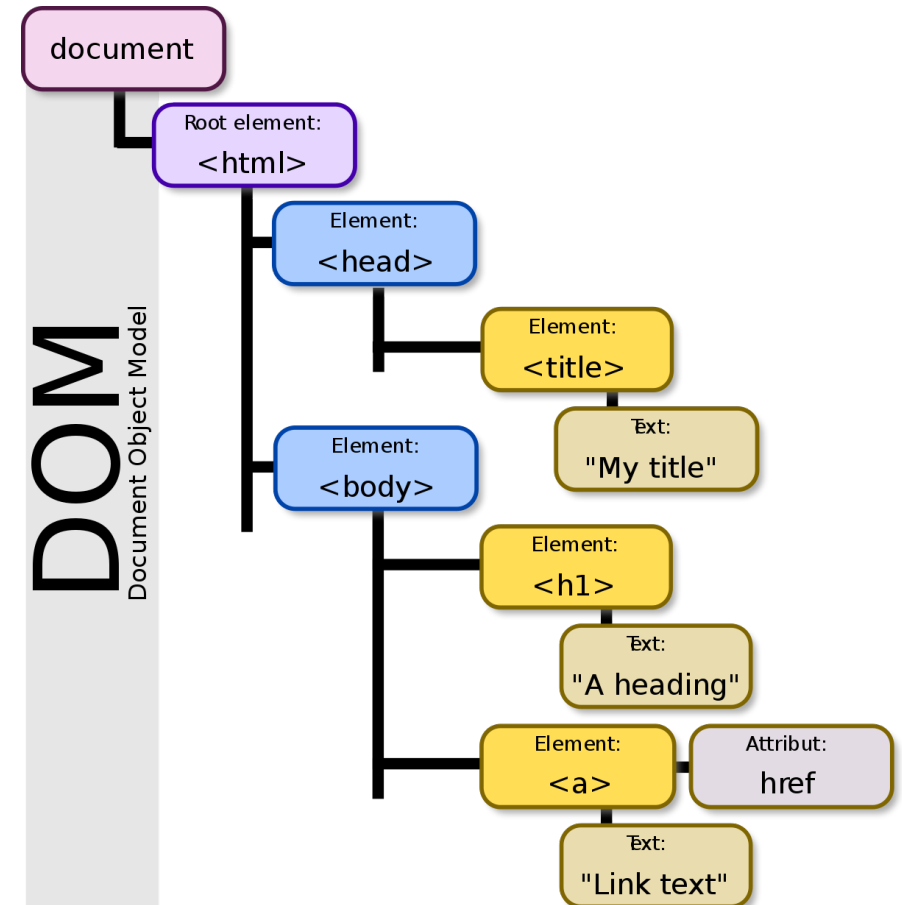
## Document object model

Hierarchical representation of elements

*Lexical* dependencies

Interface for processing, not rendering

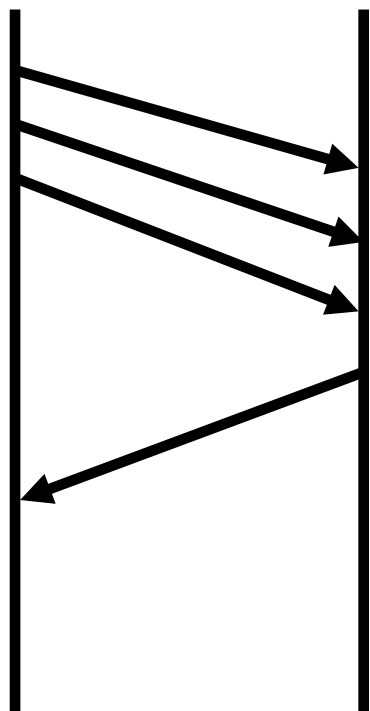
Common language for content  
*specification* (HTML) and *formatting* (JS,  
CSS)



# Structure of a web page: The network

Client  
(browser)

Origin/Server  
(google.com)



**HTTP  
requests**

**HTTP  
response**

The screenshot shows the HttpDetect (EffeTech HTTP Sniffer) window. The main table lists the following data:

No.	Time	Client (IP:PORT)	Server (IP:PORT)	URL	File Size	Status
69	Ju...	192.168.1.3 :3057	www.iffetech.com :80	/	35860	FIN, 200
70	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/style.css		FIN, 304
71	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/logo_main.jpg		FIN, 304
72	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/chinese_edition.gif		FIN, 304
73	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/space.gif		FIN, 304
74	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/arrow_small.gif		FIN, 304
75	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/award_tucows_4ratel...		FIN, 304
76	Ju...	192.168.1.3 :3060	www.etherdetect.com...	/images/logo_ms.gif	628	FIN, 200
77	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_ibm.gif	1217	FIN, 200
78	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/award_FileHungry_5s...		FIN, 304
79	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/award_softwareseeke...		FIN, 304
80	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_mit.gif	259	FIN, 200
81	Ju...	192.168.1.3 :3060	www.etherdetect.com...	/images/logo_ms.gif		Requested
82	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/award_webaward2002e...		FIN, 304
83	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/ed_small.gif	24269	FIN, 200
84	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_cornell.gif	2027	FIN, 200
85	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_detail.gif	1026	FIN, 200
86	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_reuters.gif	1822	FIN, 200
87	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_demo.gif	1013	FIN, 200
88	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_demo.gif	1013	FIN, 200

The bottom section of the window displays the HTTP Request Header and HTTP Response Header for the selected request (No. 77).

**HTTP Request Header**

GET /images/logo\_ibm.gif HTTP/1.1  
Accept: \*/\*  
Referer: http://www.iffetech.com/  
Accept-Language: zh-cn  
Accept-Encoding: gzip, deflate  
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)  
Host: www.etherdetect.com  
Connection: Keep-Alive

**HTTP Response Header**

HTTP/1.1 200 OK  
Date: Sat, 07 Jun 2003 13:32:07 GMT  
Server: Apache/1.3.27  
Last-Modified: Mon, 14 Apr 2003 14:11:33 GMT  
ETag: "bdae-4c1-3e9ac195"  
Accept-Ranges: bytes  
Content-Length: 1217  
Keep-Alive: timeout=5, max=100

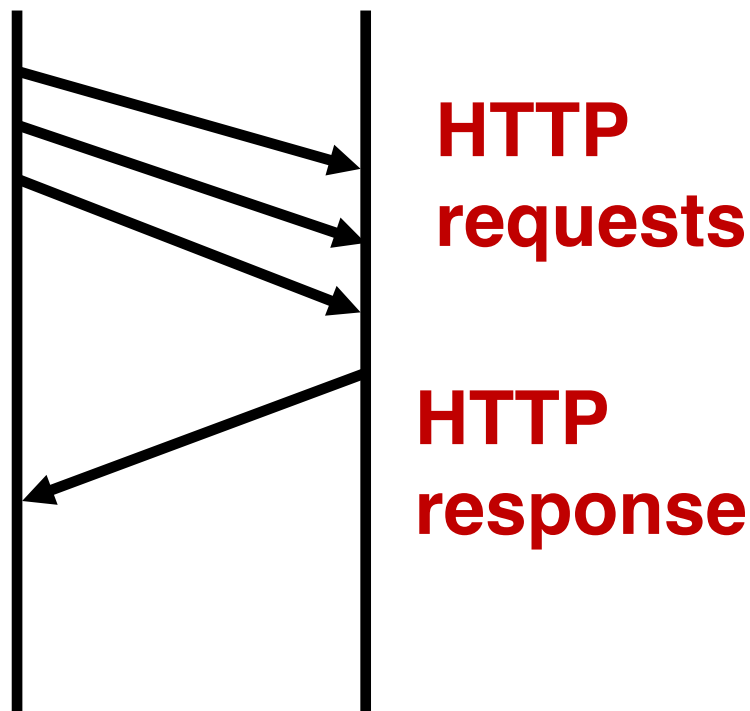
Ready Buffer: 3% URLs: 95 Packets: 393



# Structure of a web page: The network

Client  
(browser)

Origin/Server  
(google.com)



Requests based on DOM or  
JS (xmlHttpRequest/AJAX)

The screenshot shows the HttpDetect (EffeTech HTTP Sniffer) interface. It displays a list of HTTP requests and responses, with the selected request (No. 77) showing its details in the bottom panels.

No.	Time	Client (IP:PORT)	Server (IP:PORT)	URL	File Size	Status
69	Ju...	192.168.1.3 :3057	www.iffetech.com :80	/	35860	FIN, 200
70	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/style.css		FIN, 304
71	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/logo_main.jpg		FIN, 304
72	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/chinese_edition.gif		FIN, 304
73	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/space.gif		FIN, 304
74	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/arrow_small.gif		FIN, 304
75	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/award_tucows_4ratel...		FIN, 304
76	Ju...	192.168.1.3 :3060	www.etherdetect.com...	/images/logo_ms.gif	628	FIN, 200
77	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_ibm.gif	1217	FIN, 200
78	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/award_FileHungry_5s...		FIN, 304
79	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/award_softwareseeke...		FIN, 304
80	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_mit.gif	259	FIN, 200
81	Ju...	192.168.1.3 :3060	www.etherdetect.com...	/images/logo_ms.gif		Requested
82	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/award_webaward2002e...		FIN, 304
83	Ju...	192.168.1.3 :3058	www.iffetech.com :80	/images/ed_small.gif	24269	FIN, 200
84	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_cornell.gif	2027	FIN, 200
85	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_detail.gif	1026	FIN, 200
86	Ju...	192.168.1.3 :3061	www.etherdetect.com...	/images/logo_reuters.gif	1822	FIN, 200
87	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_demo.gif	1013	FIN, 200
88	Ju...	192.168.1.3 :3059	www.iffetech.com :80	/images/flag_demo.gif	1013	FIN, 200

HTTP Request Header	HTTP Response Header
GET /images/logo_ibm.gif HTTP/1.1 Accept: /*/* Referer: http://www.iffetech.com/ Accept-Language: zh-cn Accept-Encoding: gzip, deflate User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1) Host: www.etherdetect.com Connection: Keep-Alive	HTTP/1.1 200 OK Date: Sat, 07 Jun 2003 13:32:07 GMT Server: Apache/1.3.27 Last-Modified: Mon, 14 Apr 2003 14:11:33 GMT ETag: "bdae-4c1-3e9ac195" Accept-Ranges: bytes Content-Length: 1217 Keep-Alive: timeout=5, max=100

Ready Buffer: 3% URLs: 95 Packets: 393

# Polaris: Faster Page Loads using Fine-Grained Dependency Tracking

Ravi Netravali et al., NSDI '16

(ACK: Material by Ravi Netravali)