UNIVERSITY OF Southampton

History of Search Engines

COMP3220 Web Infrastructure
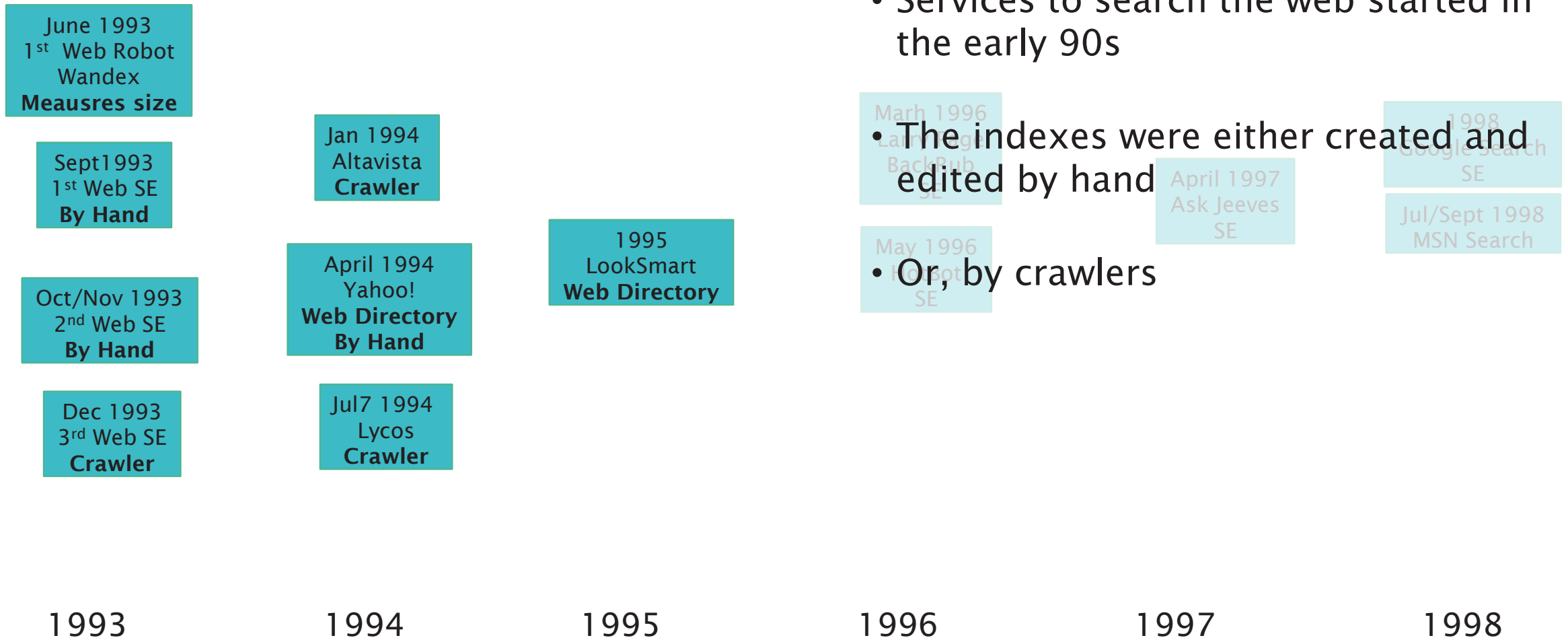
Dr Heather Packer – hp3@ecs.soton.ac.uk

Before Search Engines

- Exploit hyperlink structure

  – Personal homepages with links

  – Directories

- Word of mouth

  – Email, forums, Usenet
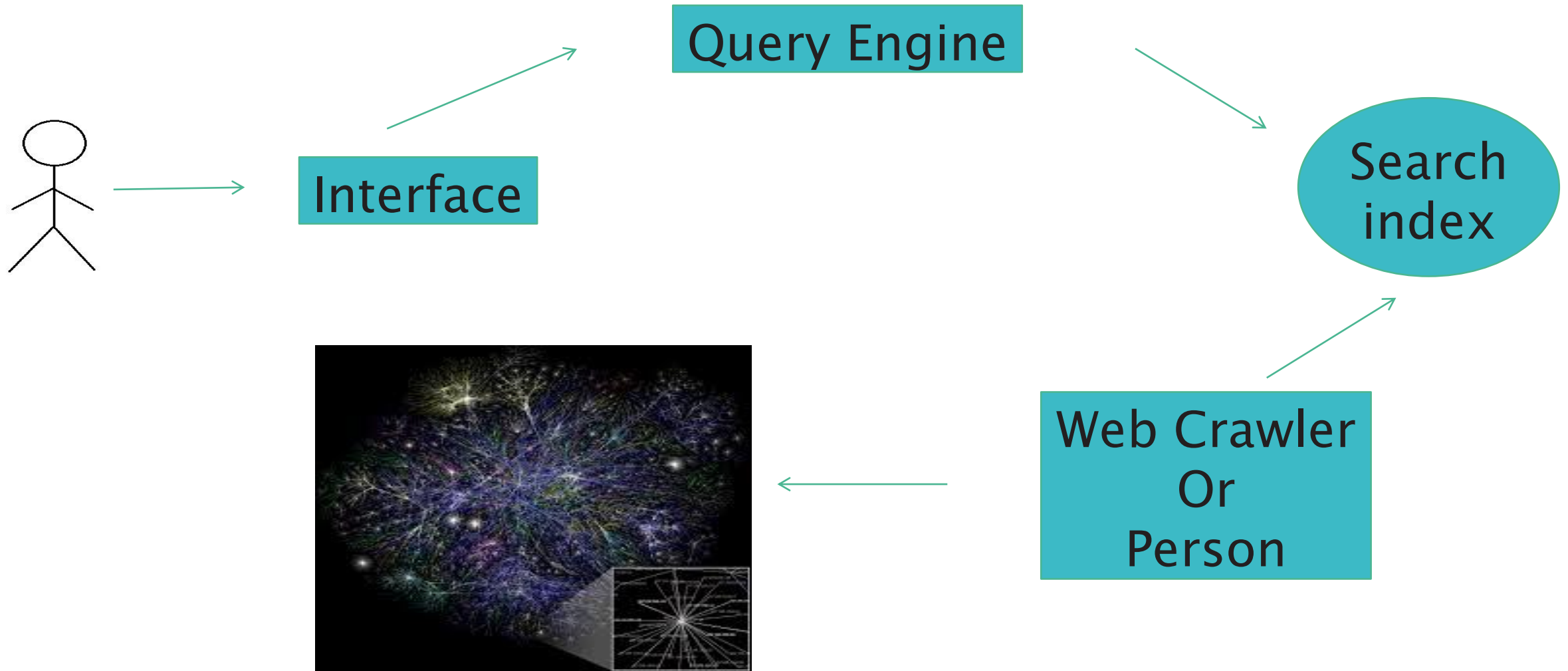
Timeline 90's – Hand vs Crawler

June 1993
1st Web Robot
Wandex
Meausres size

Sept1993
1st Web SE
By Hand

Oct/Nov 1993
2nd Web SE
By Hand

Dec 1993
3rd Web SE
Crawler

Jan 1994
Altavista
Crawler

April 1994
Yahoo!
**Web Directory
By Hand**

Jul7 1994
Lycos
Crawler

1995
LookSmart
Web Directory

Marh 1996
Larry google
BackRub
SE

May 1996
HotBot
SE

April 1997
Ask Jeeves
SE

1998
Google Search
SE

Jul/Sept 1998
MSN Search

- Services to search the web started in the early 90s

- The indexes were either created and edited by hand

- Or, by crawlers

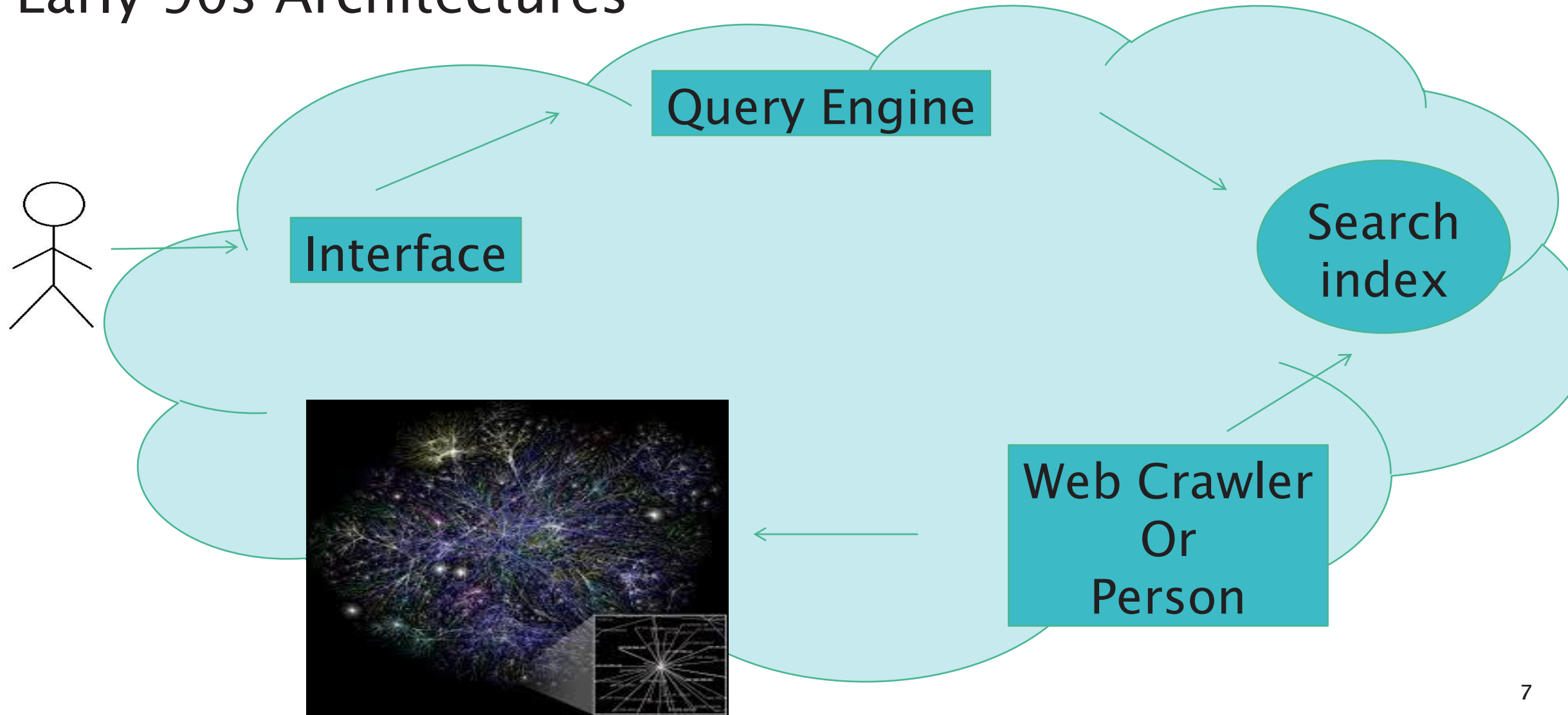1993   1994   1995   1996   1997   1998

Early 90s Architectures

Query Engine

Interface

Search index

Web Crawler Or Person

Early 90s Architectures

Query Engine

Interface

Search index

Web Crawler Or Person

Early 90s Architectures

Query Engine

Interface

Search index

Web Crawler Or Person
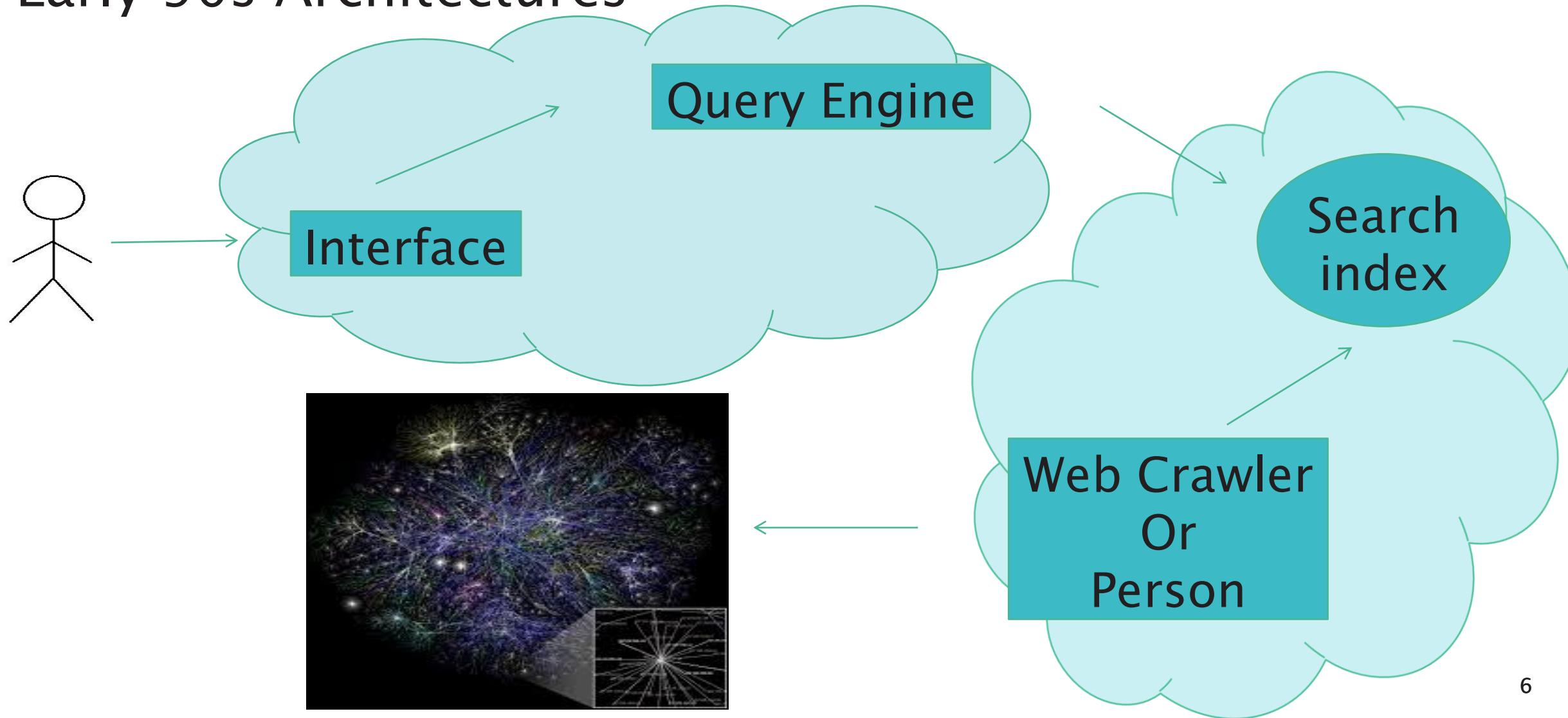
Web Crawlers

- **Wandex** 1993 – size of the web

- WebCrawler Dec 1993 - indexer

- *WebCrawler* 1994 - indexed entire web page

WebCrawler™

To search the WebCrawler database, type in your search keywords here. Type as many relevant keywords as possible; it will help to uniquely identify what you're looking for.

Pets

[Search]   ☒ AND words together

Number of results to return: 25 ▼

Yahoo!

Internet Yellow Pages 1994
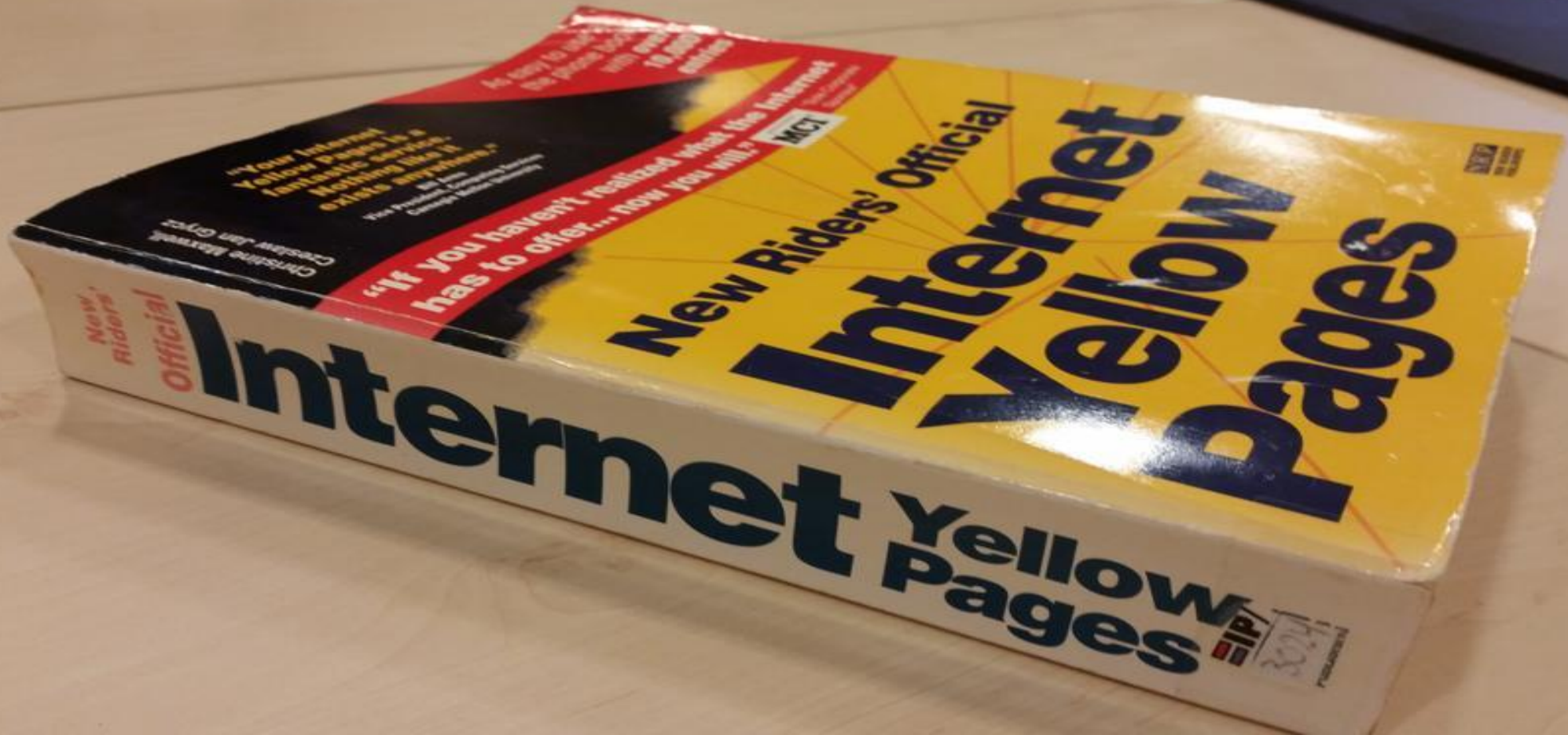
Website - Telegraph 1994

Electronic Telegraph   15 November, 1994   THE FRONT PAGE

electronic Telegraph   The Front Page

Tories block open inquiry on MPs' cash

LABOUR failed last night to force the Commons privileges committee to hold hearings into "cash-for-questions" allegations against Tory MPs in public rather than private. The Government used its majority to block the Labour demand for an open investigation, despite warnings from some senior Conservatives that the row over standards in public life was damaging Parliament's standing.

Rebels threaten to behead Britons

THREE young British tourists have been kidnapped by Kashmiri rebels who threatened to behead them if Indian authorities do not release nine Islamic militants from jail.

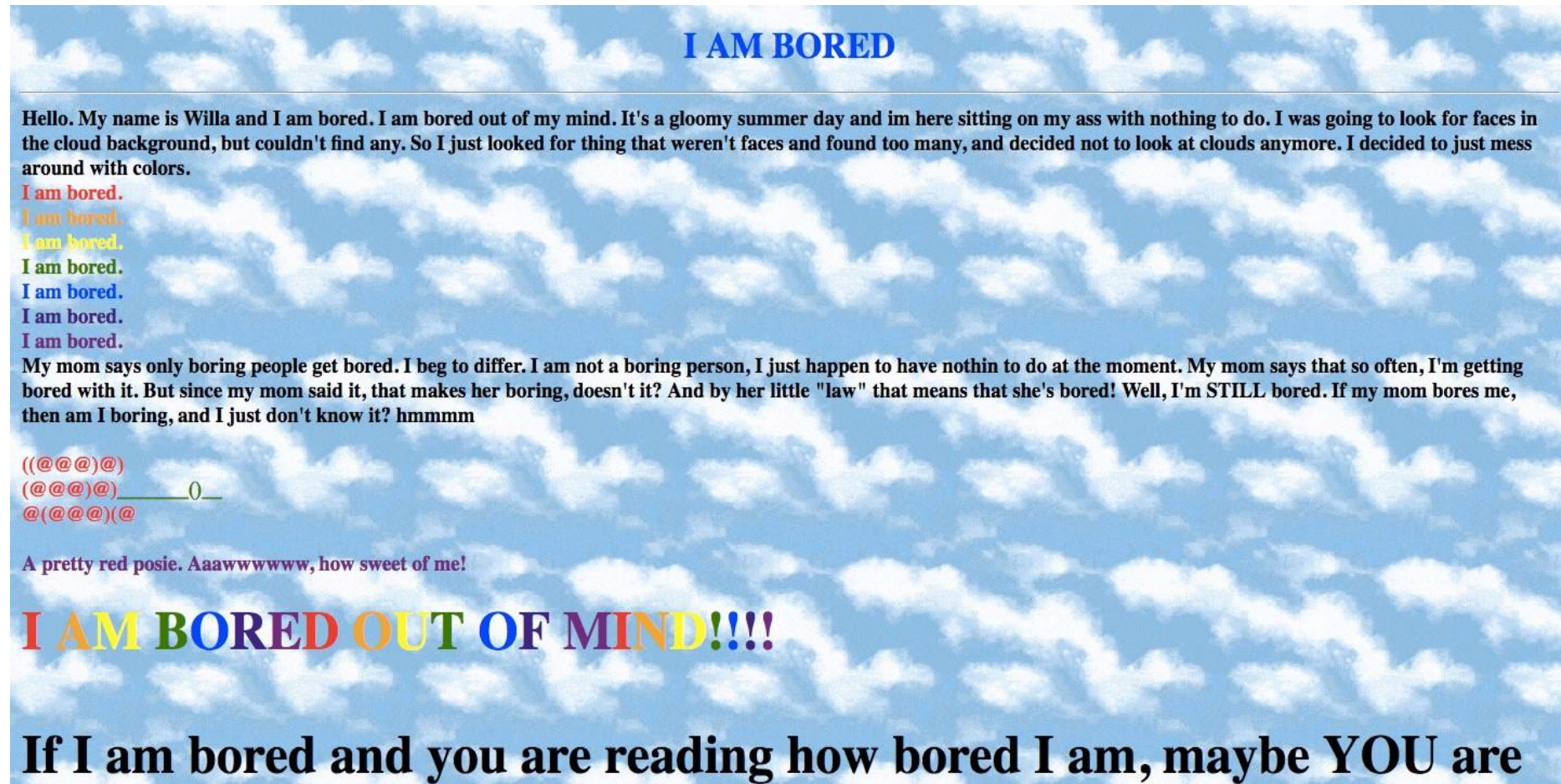# Website – New York Times 1995

Website – BBC News 1997

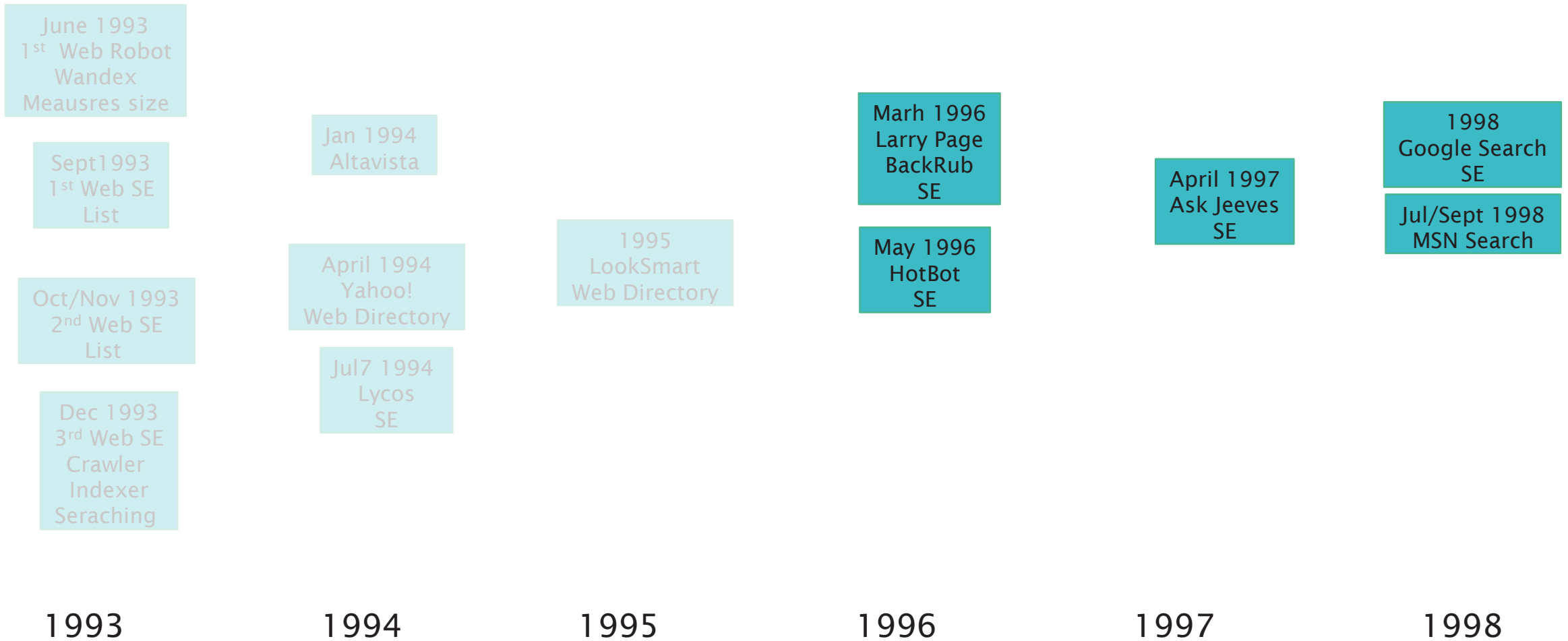# Number of Websites June 1994 - Dec 1995
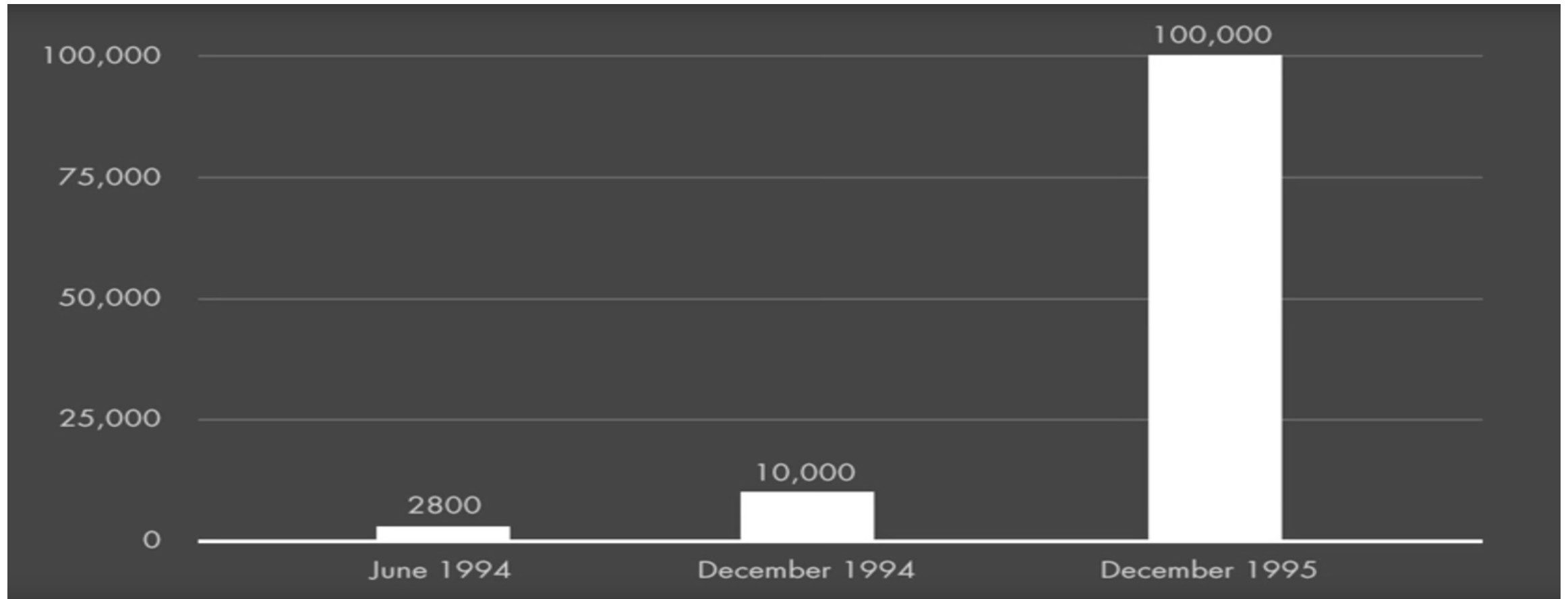
Search Engines Issues Mid '90s

- Scale

  - How to rank pages to give the best results

- Spamming

  - Web searches favoured web pages with high keyword density

- Keywords

  - Spelling mistakes

Timeline 90's

June 1993
1st Web Robot
Wandex
Meausres size

Sept1993
1st Web SE
List

Oct/Nov 1993
2nd Web SE
List

Dec 1993
3rd Web SE
Crawler
Indexer
Seraching

Jan 1994
Altavista

April 1994
Yahoo!
Web Directory

Jul7 1994
Lycos
SE

1995
LookSmart
Web Directory

Marh 1996
Larry Page
BackRub
SE

May 1996
HotBot
SE

April 1997
Ask Jeeves
SE

1998
Google Search
SE

Jul/Sept 1998
MSN Search

| 1993 | 1994 | 1995 | 1996 | 1997 | 1998 |

BackRub

- Utilised PageRank

  - More advanced than previous indexing

- Used back links

The Original Google Storage

- Larry Page needed a large amount of diskspace to test PageRank

- 10 4GB hard drives

Google

Google 1997-8

The Evolution of Search Engine Features

- Search Engine algorithms are constantly changing

- They strive to improve their:

  - Efficiency

  - How to order the results

    - Freshness

    - Relevancy

  - How to improve user interaction

  - Additional features

Google Adwords 2000 Funding

Relevancy

- During 9/11 it was apparent that search engines did not cater of time-sensitive queries

- This lead to Google developing Google News

- The freshness of search engine's index became more important

Google 2012

- In 1999, it took Google one month to crawl and build an index of about 50 million pages.*

- In 2012, the same task was accomplished in less than one minute.*

*Mitchell, Jon. "How Google Search Really Works." Readwrite. February 29, 2012.

Search Verticals

- In a bid for content and market share, Search Engines develop other search verticals:

  – Books

  – Travel

  – Finance

  – Shopping

  – Scholar

Mobile Search

- By 2015 mobile accounted for more than half of all searches

- Searches need to be context aware (location)

- Limited screen and bandwidth means relevancy is critical

- Boost mobile-friendly pages in ranking algorithm

Outcome

- Search engines were built out of necessity

- Search engines features where developed due to demand or world events