

# Information Retrieval



hussein suleman  
uct cs honours 2016

# Paper: SEO

---

- ❑ What is search engine optimization?
- ❑ 200 attributes and 1000s of machines?
- ❑ Why "V7ndotcom Elursrebmemb"?
- ❑ Web site vs. Web page?
- ❑ Why different results from different data centres?
- ❑ Size of sites, Pagerank, in-links, domain age, dmoz, yahoo?, del.icio.us
- ❑ What about localization?
- ❑ Other SEO techniques?

# Paper: CBIR

---

- ❑ What is CBIR?
- ❑ What are features of digital objects?
- ❑ Textual Annotation vs. Visual features?
- ❑ What goes into a feature vector?
- ❑ What problems are encountered with features?
- ❑ What problems are there in implementing CBIR systems?
- ❑ Possible applications?

# Paper: Brin&Page

---

- ❑ What were the primary concerns of Google?
- ❑ Web directory vs. search engine?
- ❑ Specifics:
  - What is the "random surfer"?
  - How is anchor text handled?
  - Why is the vector space model deemed poor?
  - What problems occur in Web-scale crawlers?
- ❑ In general ... what was interesting about this paper?