Fairness in Open Web Search

New Challenges and Research Questions

Seán Healy
sean.healy@adaptcentre.ie

Late 19th century

U.S. federal law: "… messages received from any individual, company, or corporation, or from any telegraph lines connecting with this line at either of its termini, shall be impartially transmitted in the order of their reception, excepting that the dispatches of the government shall have priority."

John Stuart Mill – "The only purpose for which power can be rightfully over any member of a civilized community, against his will, is to prevent harm to others."

Open Web / Closed Web

1970s/80s

Welcome to the...

Electronic Mail Mailbox   <12,040> NEW ELECTRONIC MAIL MESSAGES!

[ ] **from:** Ray Tomlinson >> "Hit me back if you get this."
[ ]   **to:** Ray Tomlinson >> "OMG! I TOTALLY DID! WHAT IS THIS?"
[ ] **from:** Ray Tomlinson >> "It is like an electronic version of mai…"
[ ] **from:** intern   >> "Hey U guys need 2 see this funny kitte…"
[ ] **from:** Ray Tomlinson >> "HAHA! But that's really not what this is…"
[#] **from:** cookebpgdl1   >> "Cheap Phrama Only Today! \/l/\<,K/\ !!!!!!!!"
[#] **from:** no-reply   >> "Enhancing your rod will be the best wa…"
[#] **from:** real mom   >> "Local mom makes $22.80 in ONE WEEK!! "
[ ]   **to:** Ray Tomlinson >> "What's going on?"
[ ] **from:** Ray Tomlinson >> "I'm not sure. I can't stop it. :C"
[ ]   **to:** Ray Tomlinson >> "HA! That looked like face frowning side…"
[#] **from:** eqefq sweik   >> "Be a sperm-man with Shitloads of Sperms!"
[#] **from:** cialis4cheep   >> "you p3ni$ will be very gratefull!!"
[ ]   **to:** Ray Tomlinson >> "I hate you."
[ ] **from:** Ray Tomlinson >> "I hate me too."

🄾open 🄽new 🄼mark as delicious meat   1971-09-16T23:59:58.75

John Rawls:

Choose the social policy that maximises the life situations for the worst off members of society.

"Maximising the Minimum"

To: mary@ru.nl
Subject: Email undeliverable: james@aol.com
(different provider)

Open Web / Closed Web

1990s

"Information systems start small and grow. They also start isolated and then merge. A new system must allow existing systems to be linked together **without requiring any central control or coordination**." (Tim Berners-Lee, 1989)

Luck-egalitarian movement
(In response to Rawls' theory of justice)

"We take a middling stance: more egalitarian than utilitarianism, and less egalitarian than Rawls." – Equality of Opportunity (Roemer, 2002)

Open Web / Closed Web

1994

WWWW – World Wide Web Worm (1,500 daily queries)

1997

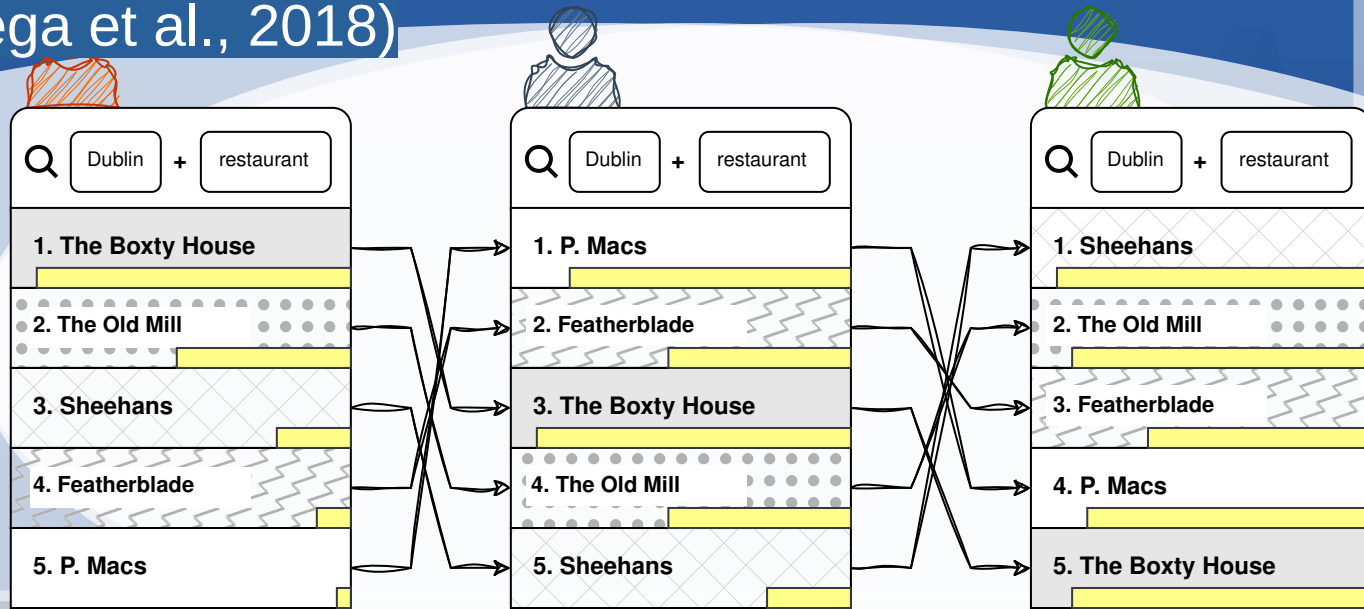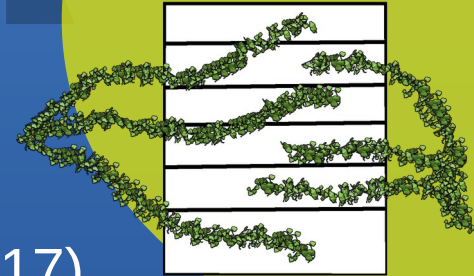**AltaVista (20 million daily queries)**

1997 – Present

Google (billions of daily queries)

"Our study of search engines suggests that they systematically exclude certain sites and certain types of sites in favor of others."
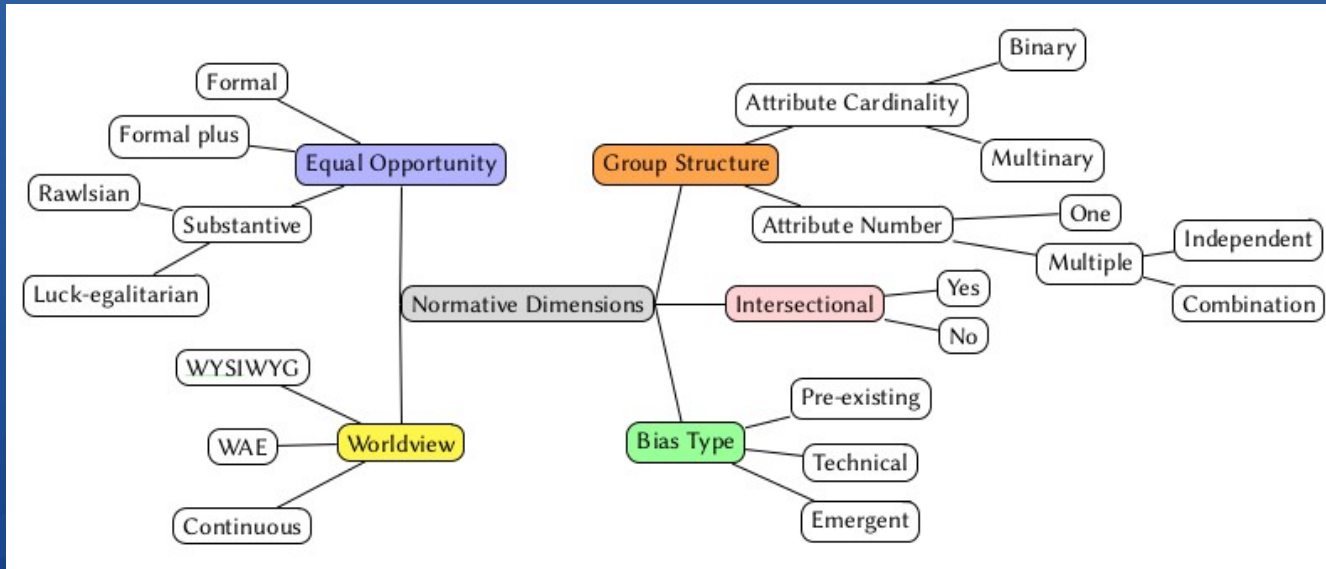
(Introna and Nissenbaum, 2000)

Fairness in Plain Old Search

- *"Fairness Through Awareness"* (Dwork et al. 2012)
  - Individual fairness vs. Group fairness
- *"FA*IR: A Fair Top-k Ranking Algorithm"* (Zehlike et al. 2017)
- *"Gender Shades"* (Buolamwini and Gebru, 2018)
- "Equity of Attention (Biega et al., 2018)
- **"Amortized fairness"**

Dublin + restaurant	Dublin + restaurant	Dublin + restaurant
1. The Boxty House	1. P. Macs	1. Sheehans
2. The Old Mill	2. Featherblade	2. The Old Mill
3. Sheehans	3. The Boxty House	3. Featherblade
4. Featherblade	4. The Old Mill	4. P. Macs
5. P. Macs	5. Sheehans	5. The Boxty House

Fairness in Plain Old Search

- *"Fairness in Ranking: A Survey" (Zehlike, Stoyanovich, Yang, 2021)*
  - *Summary: What "kind of fairness" is everyone doing?*
  - ***Spoiler**: We're all doing their own interpretation of fairness.*

Fairness in Plain Old Search

- *Possible reasons for this massive variability in pursued fairness goals:*
  - *<u>Projection?</u> "I thought everyone had the same interpretation of fairness as me."*
      Meritocratic vs. Egalitarian vs. Luck-egalitarian vs. Open web.
  - *<u>Colonialism?</u> "Everyone should have the same interpretation as me."*
  - *<u>Band-wagon effect?</u> "This was X's interpretation."*
  - ***<u>Use-case?</u>** "This is the appropriate interetation in my studied field."*

Fairness in ~~Plain Old Search~~ Open Web Search

Research questions:
- What is the appropriate interpretation of fairness within <u>web</u> search?
- Can fairness even be achieved in an <u>open</u> web search setting?

How does web search differ?

- In web search, user search queries are far more variable.
- The corpus size can be very large, so computational efficiency is crucial.
- Web search is used heavily on a daily basis by users.
  - Decreases in quality can have a much larger impact.

How does open web search differ?

» Free, open and unbiased access to information – we have lost these core principles in web search and urgently need to restore them. This is why openwebsearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction. «

https://openwebsearch.eu/

Who was paying for search?

» Free, open and unbiased access to information – we have lost these core principles in web search and urgently need to restore them. This is why openwebsearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction. «

https://openwebsearch.eu/

Fairness without Surveillance

- What happens if we don't have identity data?
- What happens if the identity data is incomplete?
- What if the search engine is an unowned distributed system?
- How can we avoid emergent bias in these instances?
- How can we ensure our efforts bring the world closer together?

How can we ensure our efforts bring the world closer together?

First research steps…

Contact:
sean.healy@adaptcentre.ie
Mastodon:
mastodon.ie/seanh

librecoir.com