SIGIR 2024
Washington, D.C.

Bias and Beyond

On Generative AI and the Future of Search and Society

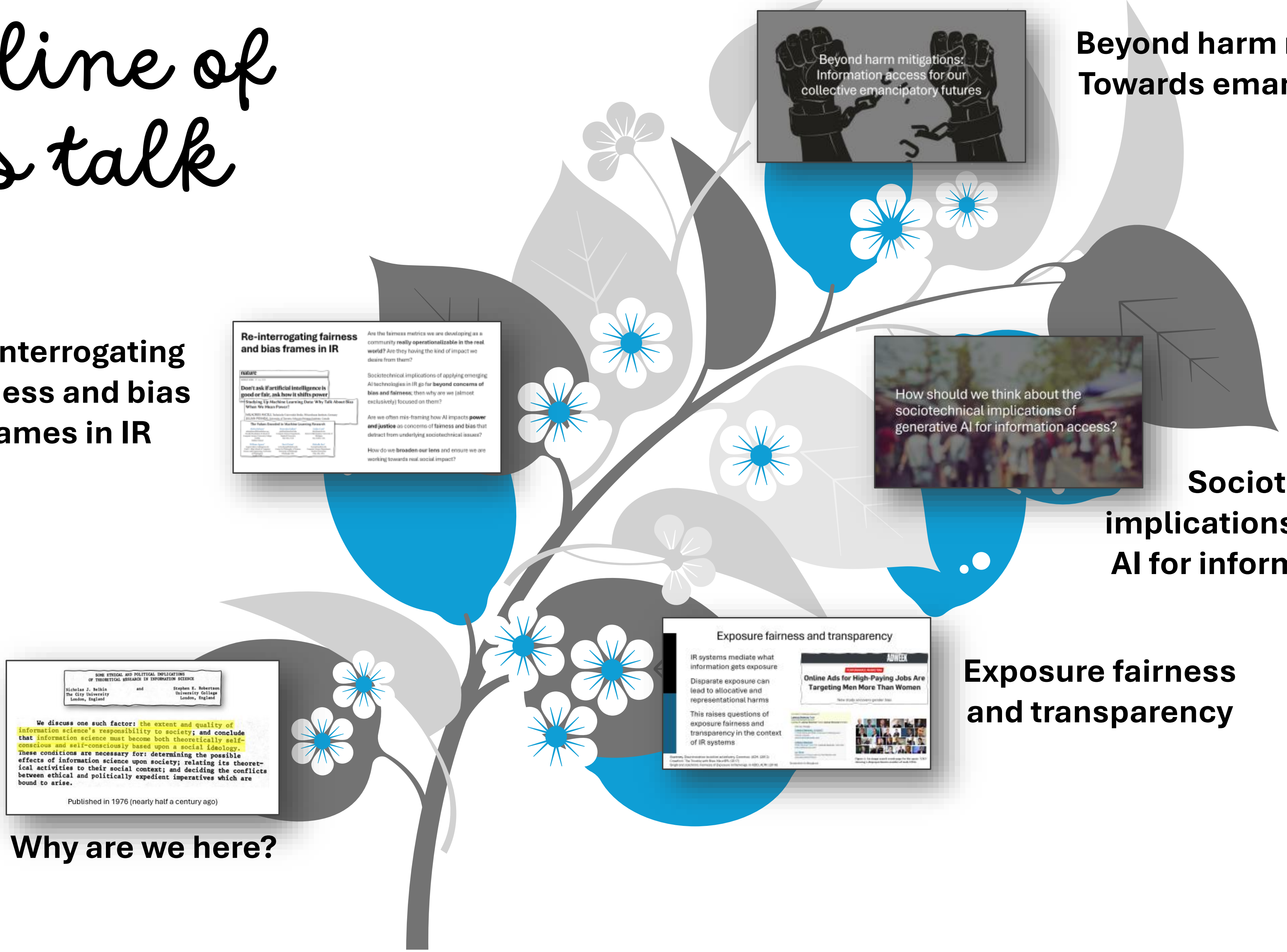
Bhaskar Mitra
Principal Researcher, Microsoft Research*
@UnderdogGeek bmitra@microsoft.com

* The views expressed in this talk are my own and do not reflect that of any institutions I am affiliated with.

Outline of this talk

Beyond harm mitigations: Towards emancipatory IR

Re-interrogating fairness and bias frames in IR



Sociotechnical implications of generative AI for information access

Exposure fairness and transparency

Why are we here?

SOME ETHICAL AND POLITICAL IMPLICATIONS
OF THEORETICAL RESEARCH IN INFORMATION SCIENCE

Nicholas J. Belkin
The City University
London, England

and

Stephen E. Robertson
University College
London, England

We discuss one such factor: the extent and quality of information science's responsibility to society; and conclude that information science must become both theoretically self-conscious and self-consciously based upon a social ideology. These conditions are necessary for: determining the possible effects of information science upon society; relating its theoretical activities to their social context; and deciding the conflicts between ethical and politically expedient imperatives which are bound to arise.

Published in 1976 (nearly half a century ago)

ARTICLE

Research Frontiers in Information Retrieval

Report from the Third Strategic Workshop on Information Retrieval in Lorne (SWIRL 2018)

Editors

J. Shane Culpepper, Fernando Diaz, and Mark D. Smucker

3 FACT IR: Fairness, Accountability, Confidentiality and Transparency in Information Retrieval

3.1 Description

IR is about connecting people to information. However, as with all software-based systems, IR systems are not free of human influence; they embed the biases of those that create, maintain and use them. Empirical evidence suggests that certain communities have differential access to information; in other words, their needs might not be equally well supported or certain information types or sources might be more or less retrievable or might not be well represented. In addition, as we increasingly rely on the outcome of IR systems such as search engines, recommender systems, and conversational agents for our decision making, there is a growing demand for these systems to be explainable. Such problems are related to many fundamental aspects of information retrieval, including information representation, information or answer reliability, information retrievability and access, evaluation, and others. While, traditionally, the IR community has been focused on building systems that support a variety of applications and needs; it is becoming imperative that we focus as much on the **human**, **social**, and **economic** impact of these systems as we do on the underlying algorithms and systems.

We argue that an IR system should be **fair** (e.g., a system should avoid discriminating across people), **accountable** (e.g, a system should be reliable and be able to justify the actions it takes), **confidential** (e.g., a system should not reveal secrets), and **transparent** (e.g., a system should be able to explain why results are returned). Judgment is needed sometimes to balance these four considerations (e.g., it is responsible to bias against unreliable sources). Other communities,

ARTICLE

FACTS-IR: Fairness, Accountability, Confidentiality, Transparency, and Safety in Information Retrieval

Editors

Alexandra Olteanu, Jean Garcia-Gathright, Maarten de Rijke, and Michael D. Ekstrand

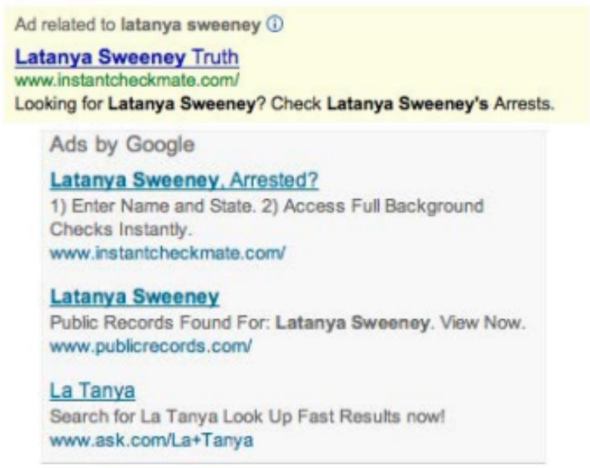
The purpose of the SIGIR 2019 workshop on Fairness, Accountability, Confidentiality, Transparency, and Safety (FACTS-IR) was to explore challenges in responsible information retrieval system development and deployment. To this end, the workshop aimed to crowdsource from the larger SIGIR community and draft an actionable research agenda on five key dimensions of responsible information retrieval: fairness, accountability, confidentiality, transparency, and safety. Such an agenda can guide others in the community that are interested in pursuing FACTS-IR research, as well as inform potential funders about relevant research avenues. The workshop brought together a diverse set of researchers and practitioners interested in contributing to the development of a technical research agenda for responsible information retrieval.

Exposure fairness and transparency

IR systems mediate what information gets exposure

Disparate exposure can lead to allocative and representational harms

This raises questions of exposure fairness and transparency in the context of IR systems



Screenshot of a Google ad.

Figure 2: An image search result page for the query "CEO" showing a disproportionate number of male CEOs.

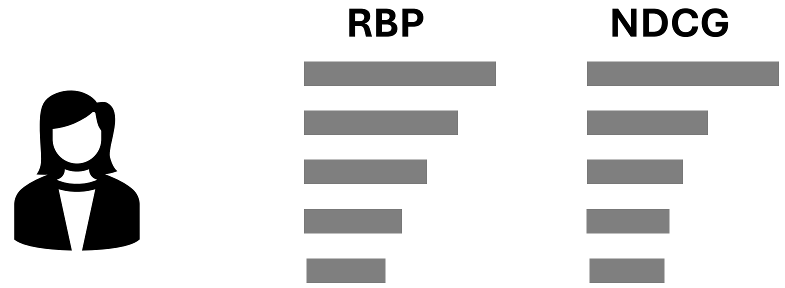
Sweeney. Discrimination in online ad delivery. Commun. ACM. (2013)
Crawford. The Trouble with Bias. NeurIPS. (2017)
Singh and Joachims. Fairness of Exposure in Rankings. In KDD, ACM. (2018)

Formalizing search exposure using user browsing models

User browsing models are simplified models of how users inspect and interact with retrieved results

It estimates the probability of inspecting a particular item in a ranked list

For example, consider the RBP user model…



Probability of exposure at different ranks according to NDCG and RBP user browsing models

$$p_{RBP}(\underbrace{\epsilon}_{\text{exposure event}}|\underbrace{d}_{\text{an item}},\underbrace{\sigma}_{\text{a ranked list of items}}) = \underbrace{\gamma}_{\text{patience factor}}^{(\underbrace{\rho_{d,\sigma}}_{\text{rank of the item in the ranked list}}-1)}$$

Diaz, Mitra, Ekstrand, Biega, & Carterette. Evaluating Stochastic Rankings with Expected Exposure. In Proc. CIKM, 2020.

Stochastic ranking and expected exposure metric

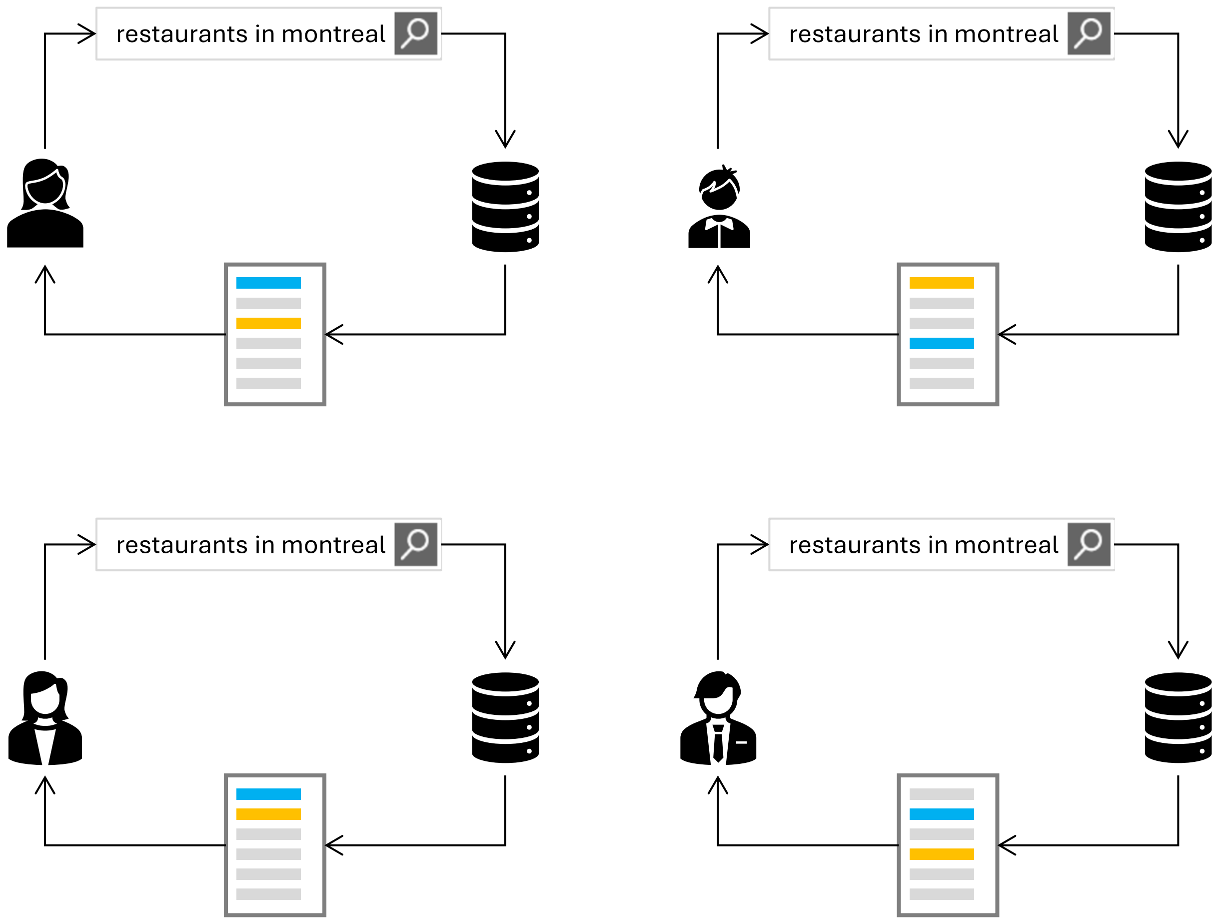
Stochastic ranking can distribute exposure more fairly across items of similar relevance and minimize *rich getting richer* effects

Expected exposure of document d under a ranking policy π_q is:

$$\varepsilon_d = \mathbb{E}_{\sigma \sim \pi_q}[\mu(d|\sigma)]$$

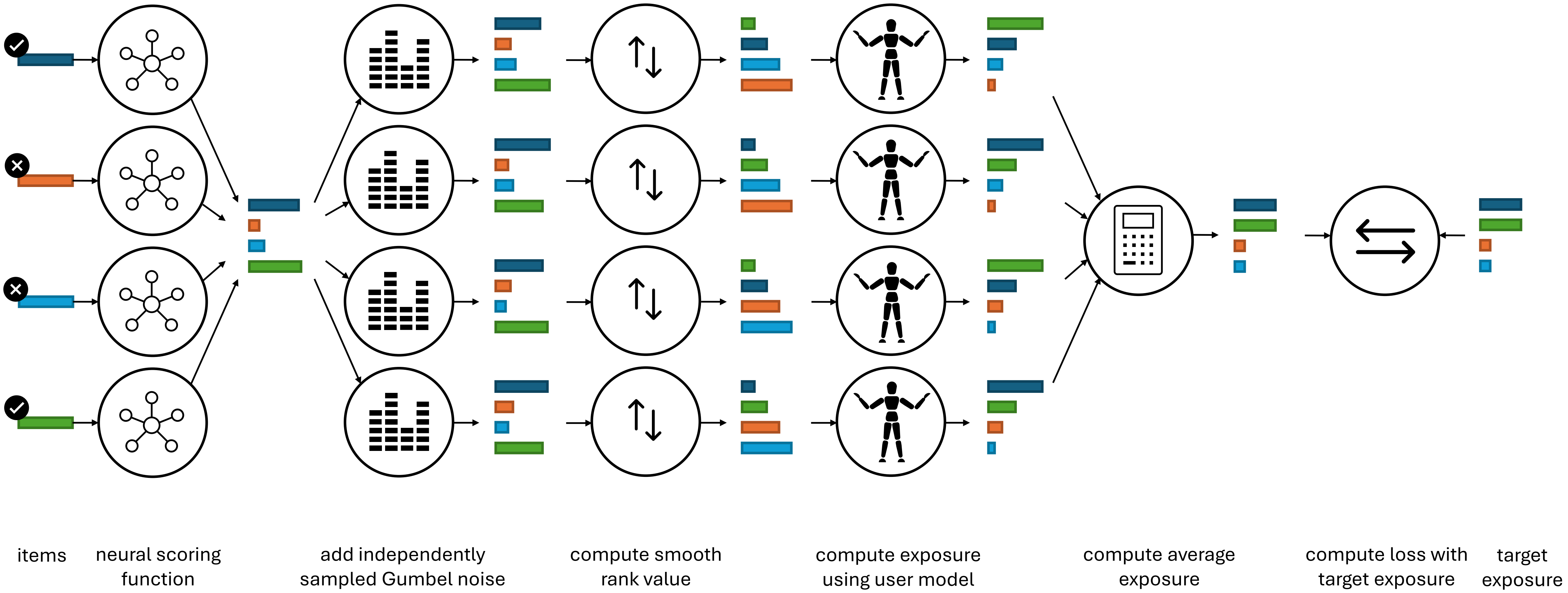
Deviation between expected and target exposure can be computed as:

$$\mathrm{EE}(\pi, q) = \|\varepsilon - \varepsilon^*\|_2^2$$



Diaz, Mitra, Ekstrand, Biega, & Carterette. Evaluating Stochastic Rankings with Expected Exposure. In Proc. CIKM, 2020.

Optimizing for target exposure



Diaz, Mitra, Ekstrand, Biega, & Carterette. [Evaluating Stochastic Rankings with Expected Exposure](). In Proc. CIKM, 2020.

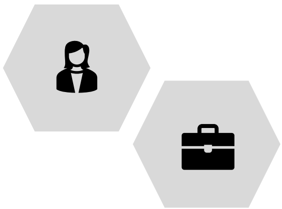
Exposure fairness is a multisided problem

It is important to ask not just whether specific content receives exposure, but *who* it is exposed to and in *what* context

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

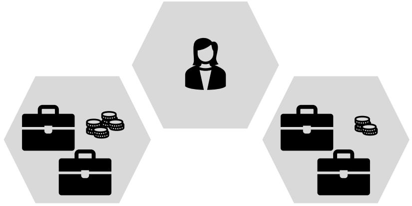
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

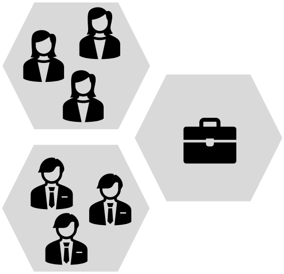
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

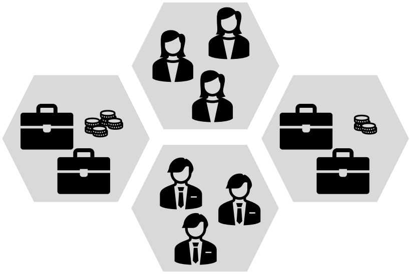
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

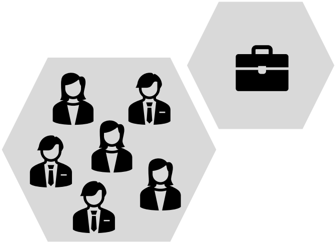
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

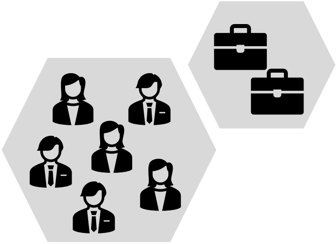
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

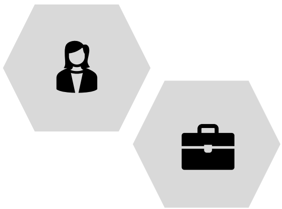
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

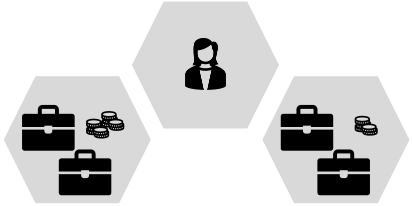
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

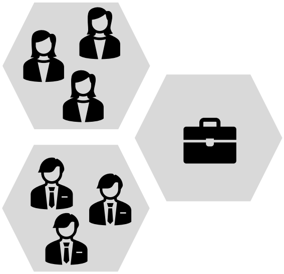
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

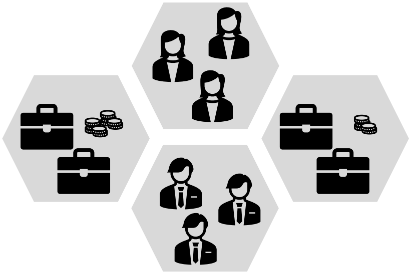
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

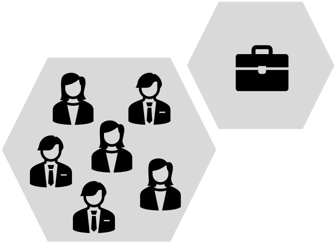
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

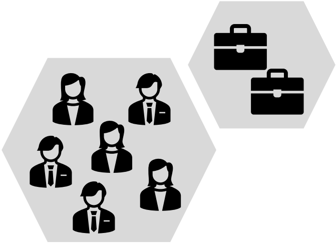
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

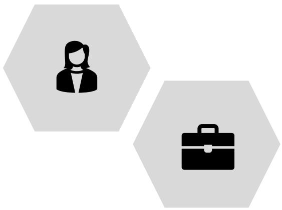
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

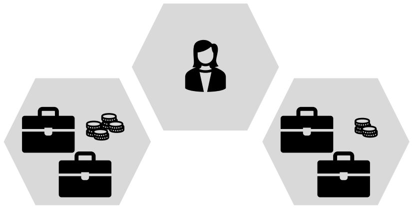
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

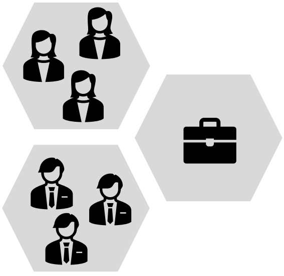
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

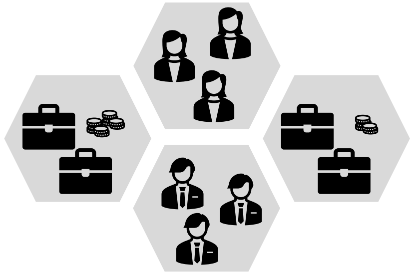
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

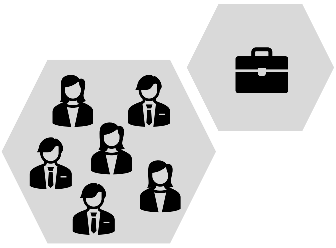
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

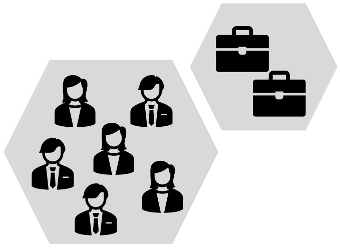
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

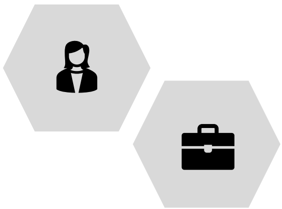
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

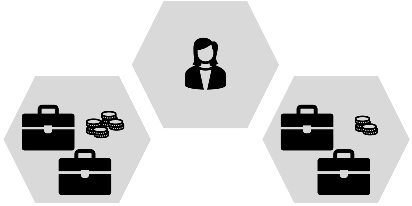
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

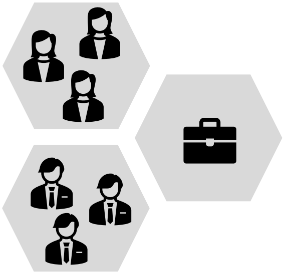
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

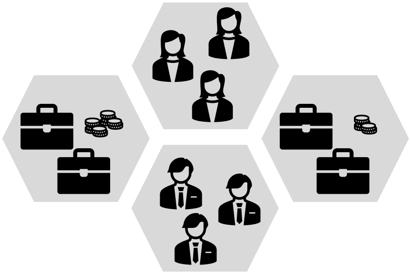
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

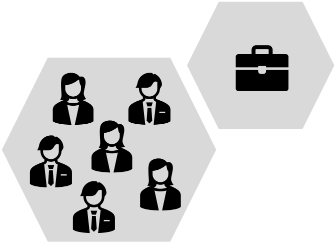
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

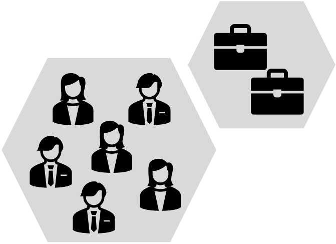
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

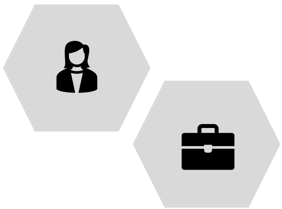
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

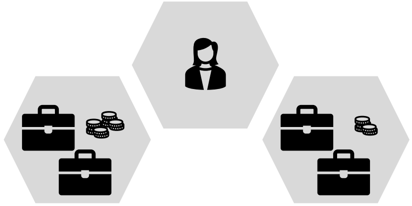
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

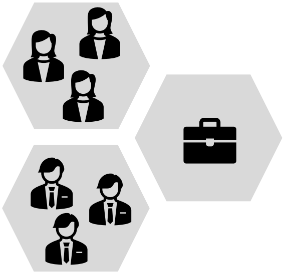
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

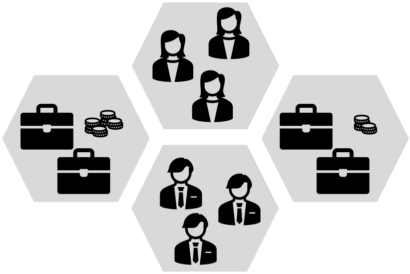
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

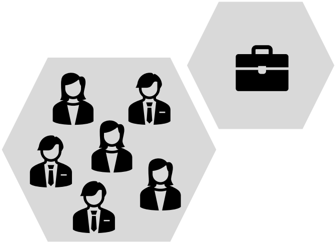
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

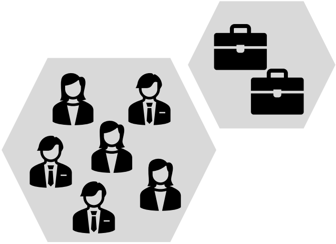
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

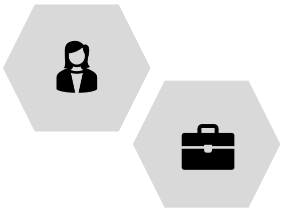
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

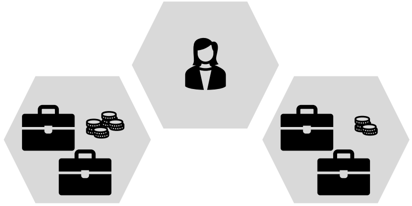
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

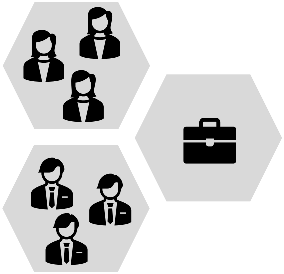
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

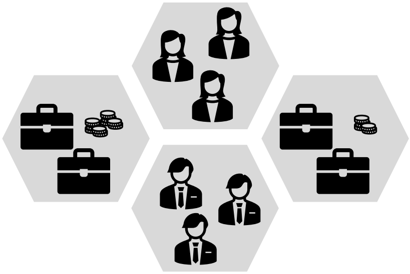
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

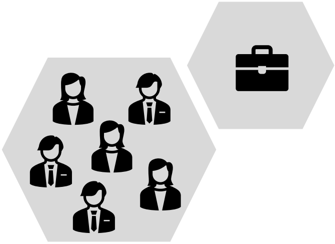
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

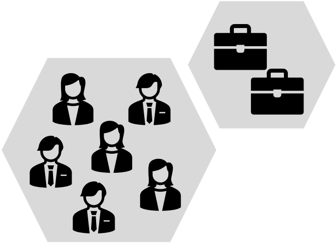
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

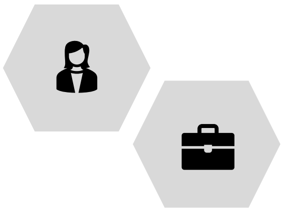
Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

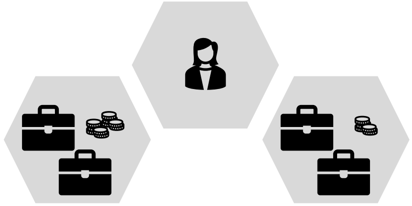
Exposure fairness is a multisided problem

Take the example of a job recommendation system



Individual-user-to-Individual-item (II-F)

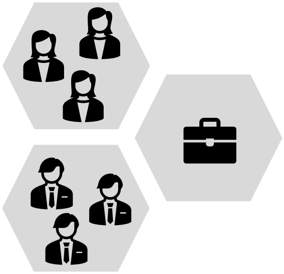
Are Individual items under/over-exposed to Individual users?



Individual-user-to-group-of-items (IG-F)

Are groups of items under/over-exposed to individual users?

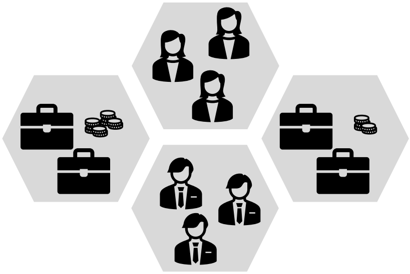
E.g., a specific user being disproportionately recommended low-paying jobs.



Group-of-users-to-Individual-item (GI-F)

Are Individual items under/over-exposed to groups of users?

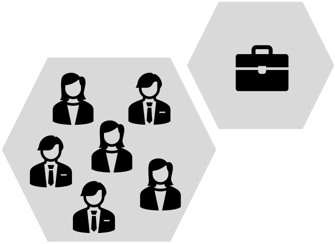
E.g., a specific job being disproportionately recommended to men and not to women and non-binary people.



Group-of-users-to-group-of-items (GG-F)

Are groups of items under/over-exposed to groups of users?

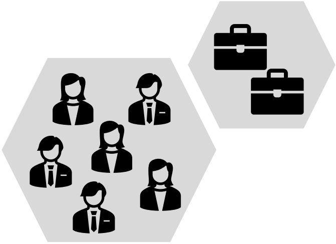
E.g., men being disproportionately recommended high-paying jobs and women low-paying jobs.



All-users-to-Individual-item (AI-F)

Are Individual items under/over-exposed to all users overall?

E.g., a specific job being disproportionately under-exposed to all users.



All-users-to-group-of-items (AG-F)

Are groups of items under/over-exposed to all users overall?

E.g., jobs at Black-owned businesses being disproportionately under-exposed to all users.

Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.

Group-aware search success

Towards fairness of quality of service

Different groups may search for different queries **and** may have different information intents for the same query

Group-aware search success metrics consider the probability that search results satisfy all groups, **not** just success on average

Expected exposure can be used to develop group-aware search success metrics

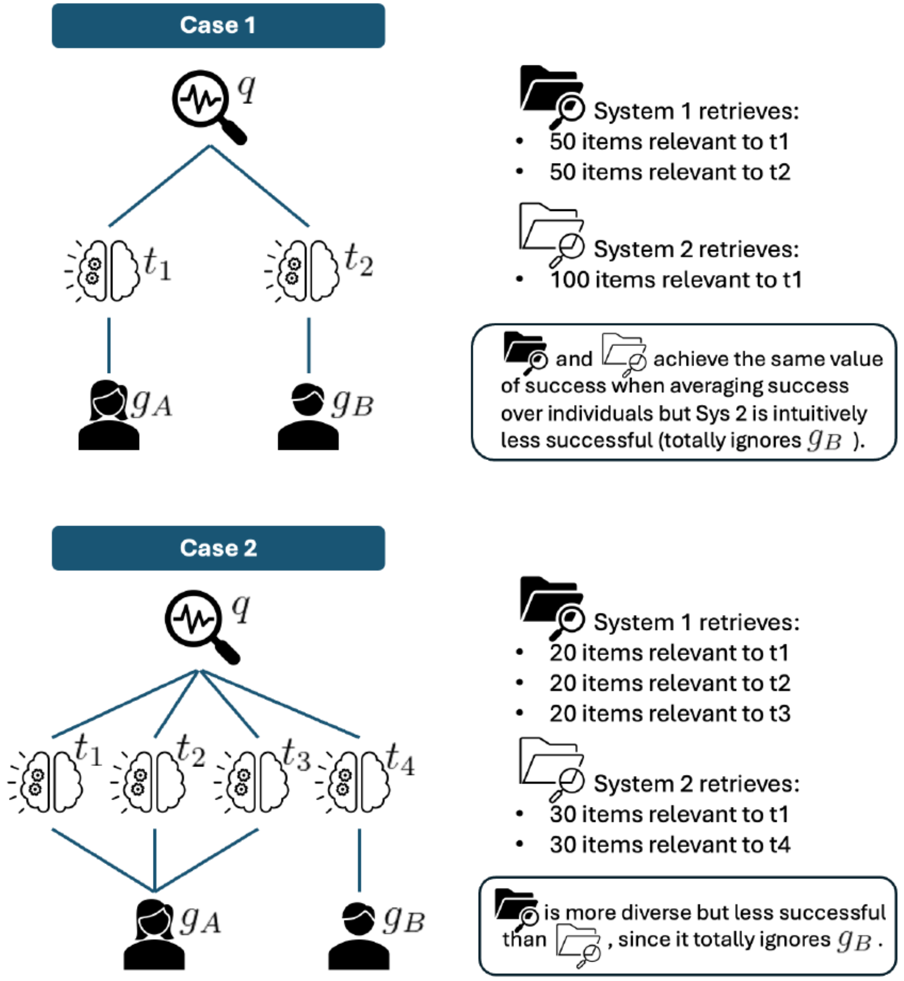
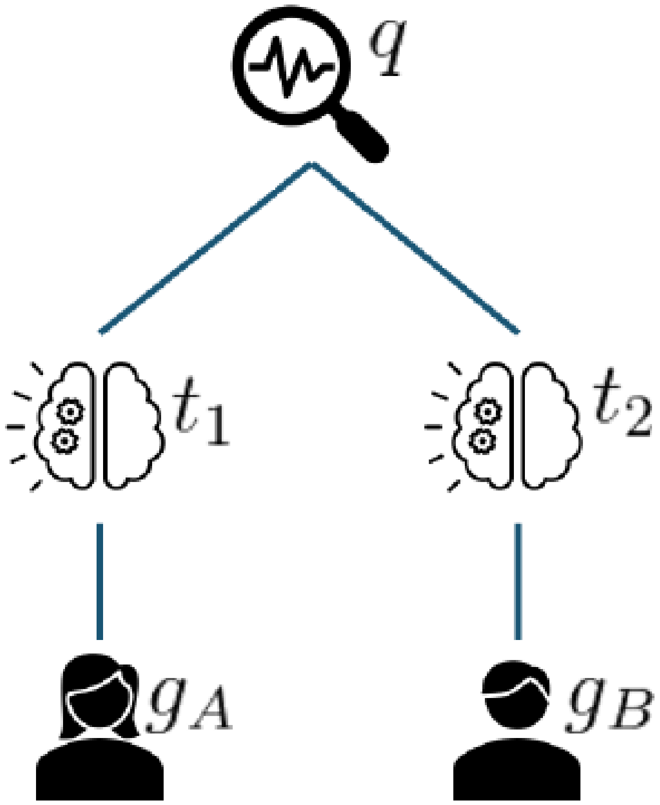


Figure 1: Two motivation examples to show that previous search success measures cannot distinguish certain nuances. Each edge in the figure between the query (q) and intent (t) carries equal weight, signifying that the query is uniformly relevant to the connected intents. Similarly, the edges linking the user group (g) to the intent (t) have equal weight within each group, indicating that members of the group have a uniform level of interest in the associated intent.

Haolun, Mitra, & Craswell. Towards Group-aware Search Success. In Proc. ICTIR, 2024.

Case 1



System 1 retrieves:

- 50 items relevant to t1
- 50 items relevant to t2

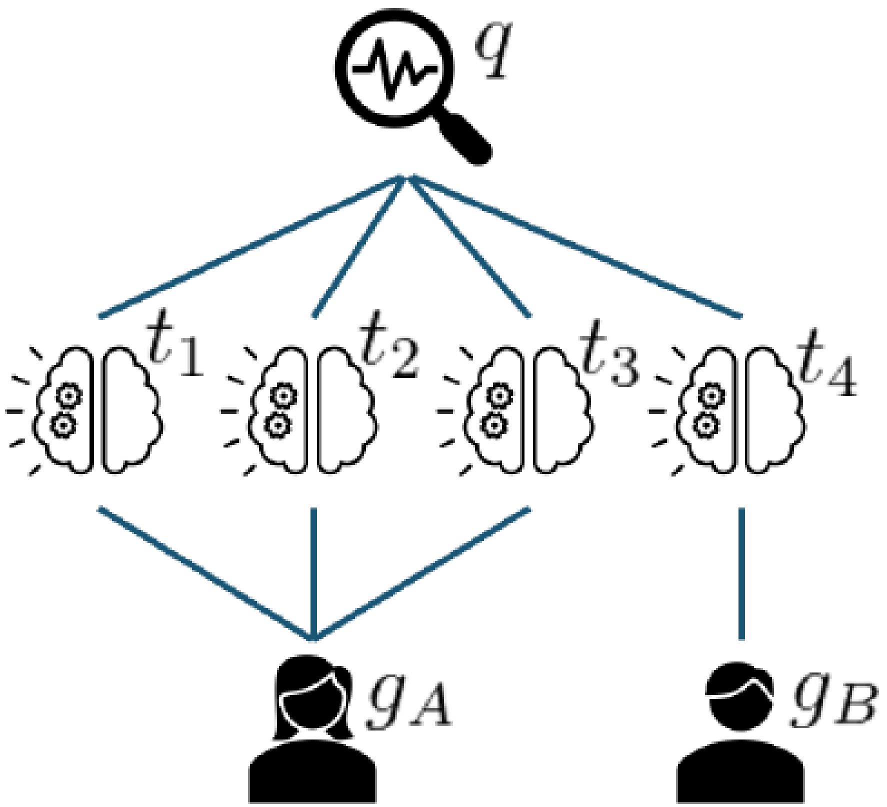
System 2 retrieves:

- 100 items relevant to t1

and achieve the same value of success when averaging success over individuals but Sys 2 is intuitively less successful (totally ignores g_B).

Haolun, Mitra, & Craswell. Towards Group-aware Search Success. In Proc. ICTIR, 2024.

Case 2



System 1 retrieves:

- 20 items relevant to t1
- 20 items relevant to t2
- 20 items relevant to t3

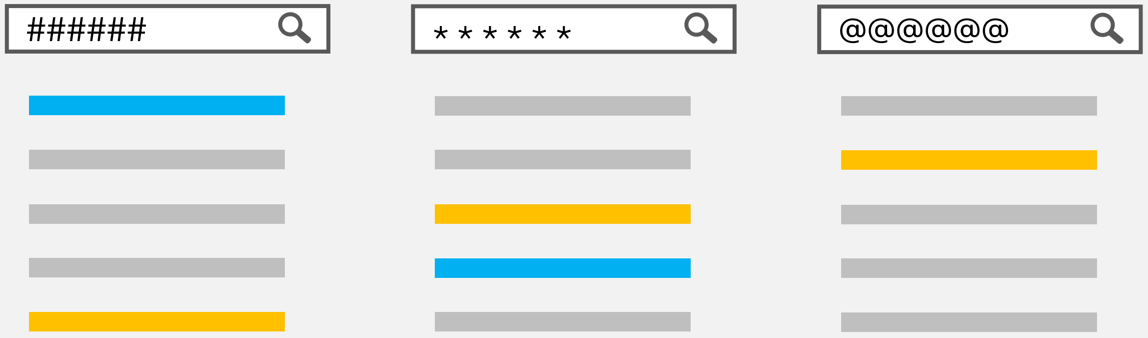
System 2 retrieves:

- 30 items relevant to t1
- 30 items relevant to t4

is more diverse but less successful than , since it totally ignores g_B .

Haolun, Mitra, & Craswell. Towards Group-aware Search Success. In Proc. ICTIR, 2024.

Exposure transparency: What query exposes me (or my documents)?

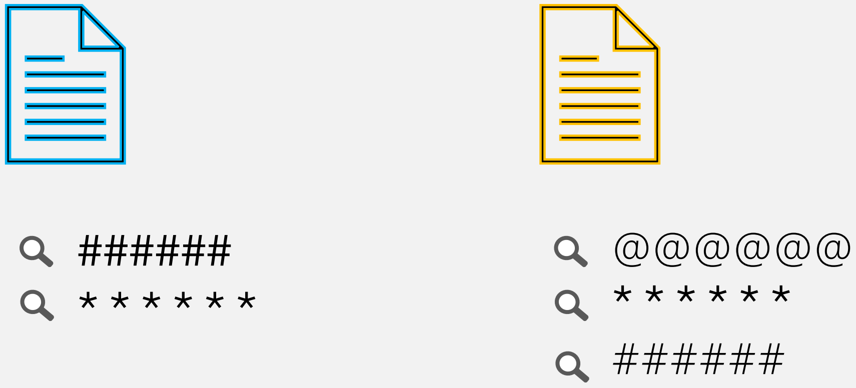


Document retrieval

Given a user-specified query, the document retrieval system retrieves a list of documents from a collection ranked by their estimated relevance to the query

Exposing query retrieval

Given a document and a specified document retrieval system, the exposing query retrieval system retrieves a list of queries from a log ranked by how prominently the document is exposed by the query when searched against the document retrieval system



Li, Li, Mitra, Biega, & Diaz. Exposing Query Identification for Search Transparency. In Proc. Web Conference, 2022.

Re-interrogating fairness and bias frames in IR

nature

WORLD VIEW | 07 July 2020

Don't ask if artificial intelligence is good or fair, ask how it shifts power

Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?

MILAGROS MICELI, Technische Universität Berlin, Weizenbaum Institute, Germany
JULIAN POSADA, University of Toronto, Schwartz Reisman Institute, Canada

The Values Encoded in Machine Learning Research

Abeba Birhane*
abeba@mozillafoundation.org
Mozilla Foundation & School of Computer Science, University College Dublin
Dublin, Ireland

Pratyusha Kalluri*
pkalluri@stanford.edu
Computer Science Department, Stanford University
Palo Alto, USA

Dallas Card*
dalc@umich.edu
School of Information, University of Michigan
Ann Arbor, USA

William Agnew*
wagnew3@cs.washington.edu
Paul G. Allen School of Computer Science and Engineering, University of Washington
Seattle, USA

Ravit Dotan*
ravit.dotan@berkeley.edu
Center for Philosophy of Science, University of Pittsburgh
Pittsburgh, USA

Michelle Bao*
baom@stanford.edu
Computer Science Department, Stanford University
Palo Alto, USA

Are the fairness metrics we are developing as a community **really operationalizable in the real world?** Are they having the kind of impact we desire from them?

Sociotechnical implications of applying emerging AI technologies in IR go far **beyond concerns of bias and fairness**; then why are we (almost exclusively) focused on them?

Are we often mis-framing how AI impacts **power and justice** as concerns of fairness and bias that detract from underlying sociotechnical issues?

How do we **broaden our lens** and ensure we are working towards real social impact?

Information retrieval research is undergoing transformative changes

What AI makes plausible

Generative AI may enable new ways in which we access information, but we are only starting to understand and grapple with their broader implications for society

»

What IR research should we do?

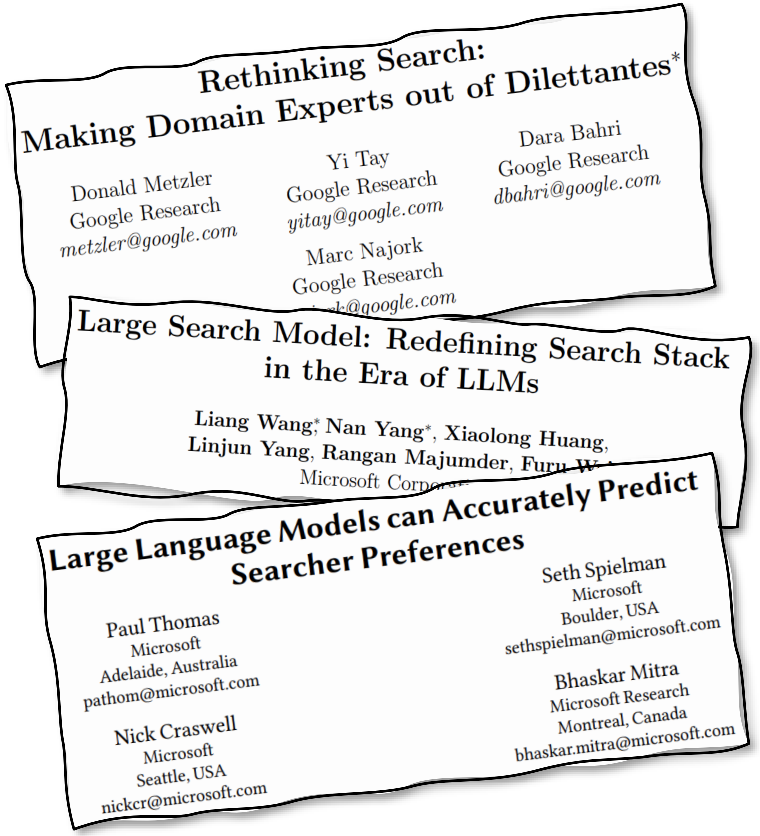
«

What the world needs

Our world is facing a confluence of forces pushing us towards precarity (e.g., global conflicts, pandemics, and climate change) and we need robust access to reliable information in this critical time

Generative AI for information access

The tale of two research perspectives


Rethinking Search:
Making Domain Experts out of Dilettantes*

Donald Metzler
Google Research
metzler@google.com

Yi Tay
Google Research
yitay@google.com

Dara Bahri
Google Research
dbahri@google.com

Marc Najork
Google Research

Large Search Model: Redefining Search Stack in the Era of LLMs

Liang Wang*, Nan Yang*, Xiaolong Huang,
Linjun Yang, Rangan Majumder, Furu

Large Language Models can Accurately Predict Searcher Preferences

Paul Thomas
Microsoft
Adelaide, Australia
pathom@microsoft.com

Nick Craswell
Microsoft
Seattle, USA
nickcr@microsoft.com

Seth Spielman
Microsoft
Boulder, USA
sethspielman@microsoft.com

Bhaskar Mitra
Microsoft Research
Montreal, Canada
bhaskar.mitra@microsoft.com


On the Dangers of Stochastic Parrots:
Can Language Models Be Too Big? 🦜

Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI

Angelina McMillan-Major

Situating Search

Chirag Shah
chirags@uw.edu

Emily M. Bender
ebender@uw.edu
University of Washington
Seattle, Washington, USA

SOCIOTECHNICAL IMPLICATIONS OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR INFORMATION ACCESS

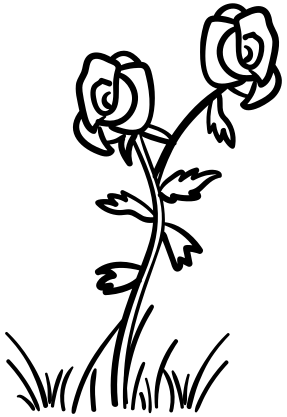
Bhaskar Mitra
Microsoft Research
Montréal, Canada
bmitra@microsoft.com

Henriette Cramer
PaperMoon AI
San Francisco, USA
henriette.cramer@gmail.com

Olya Gurevich
PaperMoon AI

Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?

CHIRAG SHAH, University of Washington, Seattle, USA
EMILY M. BENDER, University of Washington, Seattle, USA



Helps realize new information access modalities; reimagines the information retrieval stack; predicts relevance as well as anyone besides the original searcher

Disrupts information ecosystems; increases misinformation; concentrates power; reproduces historical marginalizations; accelerates climate change

How should we think about the sociotechnical implications of generative AI for information access?

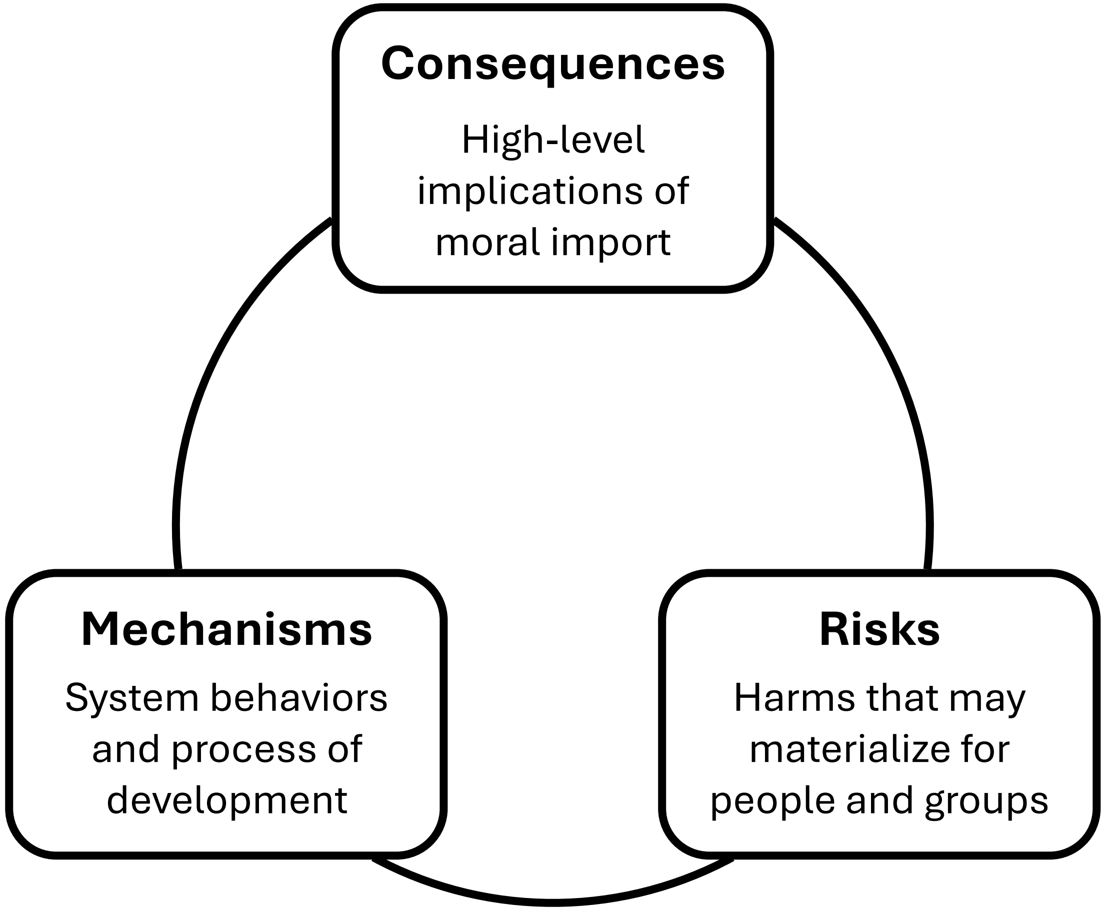
Consequences-Mechanisms-Risks (CMR) framework

Consequences motivate viewing the changes introduced by the technology through a systemic lens

Mechanisms contribute to consequences and risks; represent sites for actionable mitigation

Risks ground any investigation or mitigation to actual potential harms on people

Identified consequences, mechanisms, and risks can be mapped to each other



Gausen, Mitra, & Lindley. A Framework for Exploring the Consequences of AI-Mediated Enterprise Knowledge Access and Identifying Risks to Workers. In Proc. FAccT, 2024.

Sociotechnical implications of generative AI for information access

Table 1: Overview of the consequences for information access from generative AI, the related mechanisms introduced by these AI technologies, and corresponding risks.

Consequences	Mechanisms	Risks
Information ecosystem disruption (§2.1.1)	Content pollution (§2.1.1.1)	Risks to society: democracy, health and wellbeing, and global inequity (§2.2.1)
	The "Game of telephone" effect (§2.1.1.2)	
	Search engine manipulation (§2.1.1.3)	
	Degrading retrieval quality (§2.1.1.4)	
	Direct model access (§2.1.1.5)	
	The paradox of reuse (§2.1.1.6)	
Concentration of power (§2.1.2)	Compute and data moat (§2.1.2.1)	
	AI persuasion (§2.1.2.2)	
	AI alignment (§2.1.2.3)	
Marginalization (§2.1.3)	Appropriation of data labor (§2.1.3.1)	
	$\boxed{\text{Bias amplification (§2.1.3.2)}}$	
	AI doxing (§2.1.3.3)	
Innovation decay (§2.1.4)	Industry capture (§2.1.4.1)	Risks to IR research (§2.2.2)
	Pollution of research artefacts (§2.1.4.2)	
Ecological impact (§2.1.5)	Resource demand and waste (§2.1.5.1)	Risks to environment (§2.2.3)
	Persuasive advertising (§2.1.5.2)	

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Consequences of generative AI for information access

Information ecosystem disruption

Significantly changing how different actors and stakeholders in the online information ecosystem operate on their own and how they relate to each other

Concentration of power

Worsening inequities in how power and control are distributed within our society and different communities

Marginalization

Relegating certain individuals and groups to the margins of society and corresponding discrimination

Innovation decay

Constraining scientific explorations to specific narrow directions while throttling progress in other areas of information access research

Ecological impact

Worsening anthropogenic climate change

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Mechanisms of information ecosystem disruption

The paradox of reuse

Websites like Wikipedia and StackExchange power online information access platforms, which in turn reduce the need to visit those websites.

Examples. LLMs training on content from these websites that they later regurgitate without attribution. LLM-powered conversational search systems deemphasize source websites reducing the clickthrough relative to the classic ten-blue-links interface.

Other mechanisms

Content pollution. Enabling low-cost generation of derivative low-quality content at unprecedented scale that pollute the web.

The "Game of telephone" effect. LLMs inserted between users and search results shifts the responsibility of information inspection and interpretation to the LLM.

Search engine manipulation. E.g., prompt injection attacks.

Degrading retrieval quality. E.g., Minimizing click feedback signals.

Direct model access. Open access models pose challenges for content moderation.

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

On technological power concentration

Annual change in global risk perceptions over the short term (2 years)

Biggest increase in ranking

12th	Technological power concentration	↑17
1st	**Misinformation and disinformation**	↑15
5th	**Interstate armed conflict**	↑9
18th	Insufficient public infrastructure and services	↑9

Source
World Economic Forum Global Risks Perception Surveys 2022-2023 and 2023-2024.

meaningfully confront the core problem that AI presents, and one of the most difficult challenges of our time: **the concentration of economic and political power in the hands of the tech industry—Big Tech in particular.**

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Mechanisms of concentration of power

Compute and data moat. Only a handful of (typically private sector) institutions own and control the compute and data resources for training and deployment of generative AI models. Availability of “open access” models don’t fundamentally challenge the predominant vision of what AI looks like today, which would require dismantling the data and compute moat itself and turning them into public infrastructure.

AI persuasion. A process by which AI systems alter the beliefs of their users. E.g., application of LLMs for hyper-personalized hyper-persuasive ads.

AI alignment. Approaches such as reinforcement learning from human feedback (RLHF) presupposes some notions of desirable values to be determined and enforced by platform owners.

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Mechanisms of marginalization

Appropriation of data labor

Includes the uncompensated appropriation of works by writers, authors, programmers, and peer production communities like Wikipedia and under-compensated crowdwork for data labeling that have been instrumental in the development of these technologies.

AI for me, data labor for thee. AI data labor dynamics reinforces structures of racial capitalism and coloniality, employs global labor exploitation and extractive practices, and reinforces the global north and south divide.

Other mechanisms

Bias amplification. AI models reproduce and amplify harmful biases and stereotypes from their training datasets leading to allocative and representational harms.

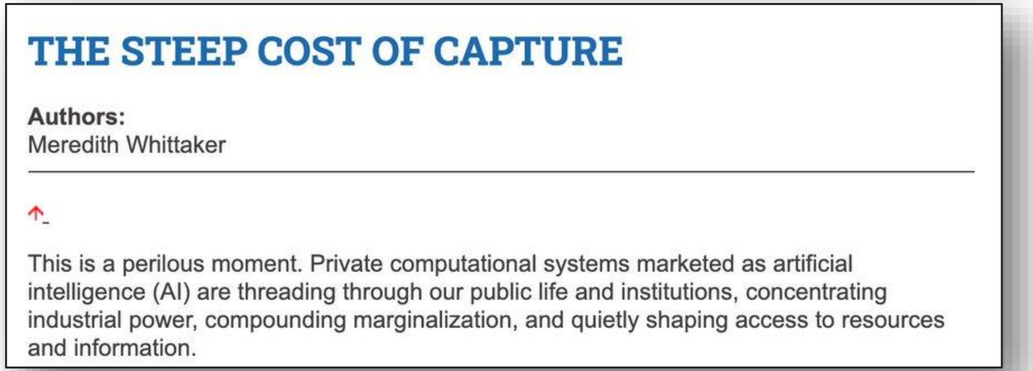
AI doxing. AI models may leak private information about people present in their training data or be employed to predict people's sensitive information based on what is known about them publicly.

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Mechanisms of...

Innovation decay

Industry capture. Profit-driven goals inordinately influences scientific exploration and dissuade investments in research not immediately monetizable or which challenges the status quo.


THE STEEP COST OF CAPTURE

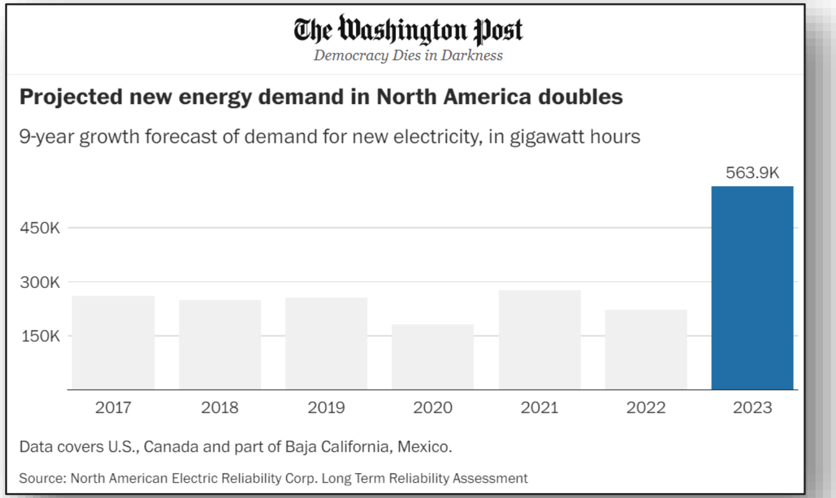
Authors:
Meredith Whittaker

This is a perilous moment. Private computational systems marketed as artificial intelligence (AI) are threading through our public life and institutions, concentrating industrial power, compounding marginalization, and quietly shaping access to resources and information.

Pollution of research artefacts. Misapplications of LLMs in scholarly publications and reviewing may negatively impact IR scholarship.

Ecological impact

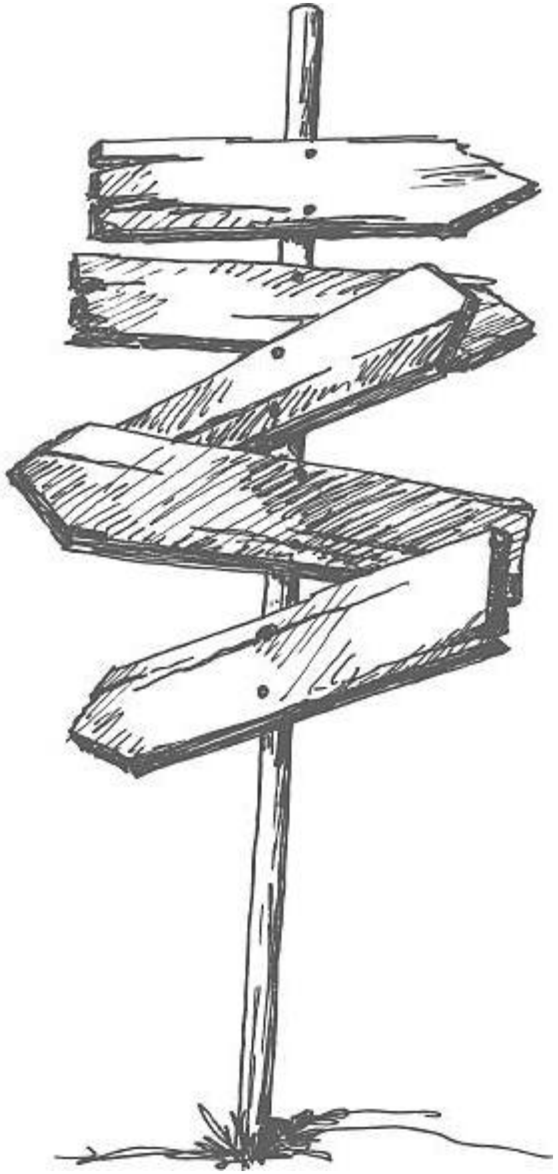
Resource demand and waste. Increasing demand for electricity and water, and electronic wastes.



Persuasive advertising. Could supercharge climate change disinformation and promote environmentally unfriendly business models like fast-fashion.

Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.

Beyond harm mitigations: Information access for our collective emancipatory futures



Sociotechnical imaginaries

"Visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology"

~Jasanoff and Kim (2015)

Whose sociotechnical imaginaries are granted normative status and what myriad of radically alternative futures are we overlooking?

How does increasing dominance of established for-profit platforms over academic research influences and/or homogenizes the kinds of IR systems we build?

What would information access systems look like if designed for futures informed by feminist, queer, decolonial, anti-racist, anti-casteist, and abolitionist thoughts?

Mitra. Search and Society: Reimagining Information Access for Radical Futures. ArXiv preprint, 2024.

Recommendations for re-centering IR on societal needs

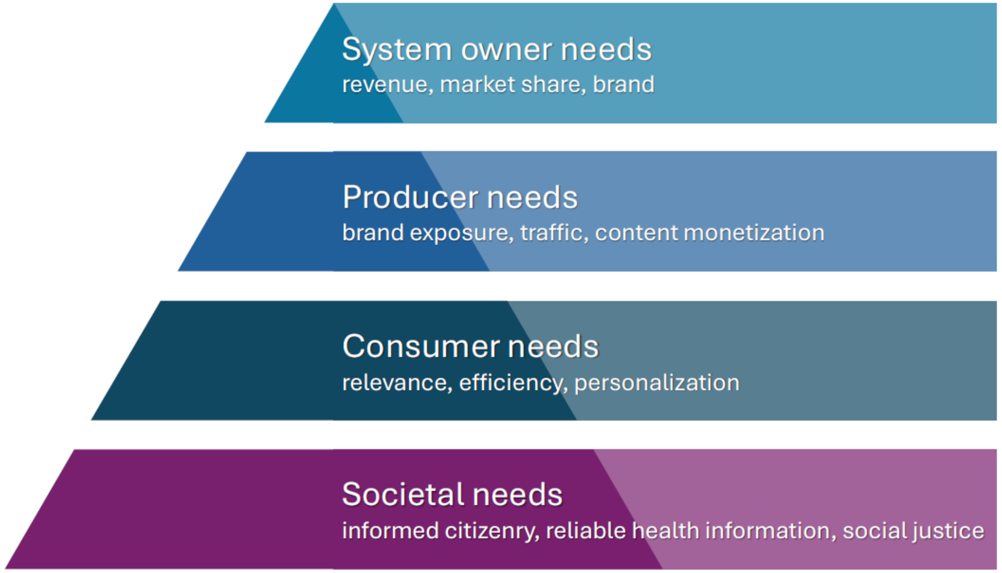


Figure 1: Hierarchy of IR stakeholder needs. More fundamental and critical needs are at the bottom of the pyramid. This figure is inspired by Maslow's Hierarchy of Needs [246] and Siksika (Blackfoot) way of life [302].

Explicitly articulate a **hierarchy of stakeholder needs** that places societal needs as the most critical concern for IR research and development

Dismantle the artificial separation between fairness and ethics research in IR and the rest of IR research; Move away from **reactionary mitigation strategies** for emerging technologies to **proactively design IR systems for social good**

Develop **sociotechnical imaginaries** based on prefigurative politics and **theories of change**

Mitra. Search and Society: Reimagining Information Access for Radical Futures. ArXiv preprint, 2024.

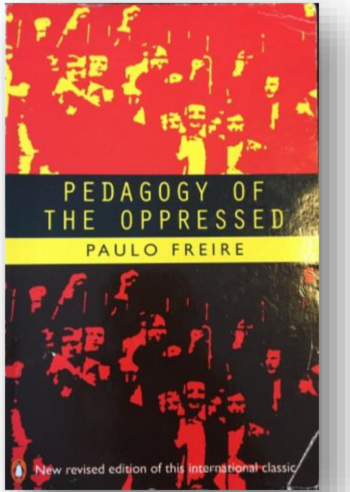
Reimagining IR through the lens of prefigurative politics

Figure 2: An image search result page for the query "CEO" showing a disproportionate number of male CEOs.

Instead of trying to algorithmically fix under-representation of women and people of color in image search results for occupational roles, can we reclaim that digital space as a site of resistance and emancipatory pedagogy by allowing **feminist**, **queer**, and **anti-racist** scholars, activists, and artists to create experiences that teach the history of these movements and struggles?

Open WebSearch.eu

Can **emancipatory** and **anti-capitalist** perspectives motivate us to reimagine search and recommender systems as decentralized and federated?



Can we translate Freire’s **emancipatory** pedagogy to strategies for anti-oppressive information access? Can search result pages support dialogical interactions between searchers that leads to knowledge production and better digital literacy?

Mitra. Search and Society: Reimagining Information Access for Radical Futures. ArXiv preprint, 2024.

Who gets to participate?

This is a call for **collective struggle of solidarity** with social scientists, legal scholars, critical theorists, activists, and artists; **not for technosolutionism**.

To challenge the homogeneity of the future imaginaries saliently bound by colonial, cisheteropatriarchal, and capitalist ways of knowing the world, **we need broad and diverse participation from our community**.

Inclusion of people without inclusion of their history, struggles, and politics is simply tokenism and epistemic injustice; we should **go beyond Diversity and Inclusion (D&I), and enshrine as our goal Justice, Equity, and Diversity & Inclusivity (JEDI)**.

Mitra. Search and Society: Reimagining Information Access for Radical Futures. ArXiv preprint, 2024.

Why *are* we here?

SOME ETHICAL AND POLITICAL IMPLICATIONS
OF THEORETICAL RESEARCH IN INFORMATION SCIENCE

Nicholas J. Belkin
The City University
London, England

and

Stephen E. Robertson
University College
London, England

We discuss one such factor: the extent and quality of information science's responsibility to society; and conclude that information science must become both theoretically self-conscious and self-consciously based upon a social ideology.

Our work should be in recognition of the responsibilities of information access technologies and research to society, but we should be motivated by pluralistic sociotechnical imaginaries informed by the diverse history and struggles of our peoples

ARTICLE

FACTS-IR: Fairness, Accountability, Confidentiality, Transparency, and Safety in Information Retrieval

Editors
Alexandra Olteanu, Jean Garcia-Gathright, Maarten de Rijke, and Michael D. Ekstrand

The workshop brought together a diverse set of researchers and practitioners interested in contributing to the development of a technical research agenda for responsible information retrieval.

We should recognize that this research agenda needs to essentially be sociotechnical and requires us to explicate our values and visions for our desired futures as a community

Concluding thoughts

Hope this **sparks many passionate conversations** and debates; **radicalizes us to work on issues of social import** and reflect on why we do what we do; encourages us to **prioritize praxis** (research activities and reflection directed at structural change) **over proxies** (e.g., optimizing for SOTA / leaderboard rankings that do not translate to scientific or social progress); and inspires us to build technology not just out of excitement for technology, but **as an act of radical love for all peoples and the worlds we share**.

"If you have come here to help me you are wasting your time, but if you have come because your liberation is bound up with mine, then let us work together."

– Lilla Watson
and other members of an
Aboriginal Rights group in Queensland

References

1. Mitra. Search and Society: Reimagining Information Access for Radical Futures. ArXiv preprint, 2024.
2. Mitra, Cramer, & Gurevich. Sociotechnical implications of generative artificial intelligence for information access. Preprint of chapter for an upcoming edited book, 2024.
3. Haolun, Mitra, & Craswell. Towards Group-aware Search Success. In Proc. ICTIR, 2024.
4. Gausen, Mitra, & Lindley. A Framework for Exploring the Consequences of AI-Mediated Enterprise Knowledge Access and Identifying Risks to Workers. In Proc. FAccT, 2024.
5. Cortiñas-Lorenzo, Lindley, Larsen-Ledet, & Mitra. Through the Looking-Glass: Transparency Implications and Challenges in Enterprise AI Knowledge Systems. ArXiv preprint, 2024.
6. Diaz & Mitra. Recall, Robustness, and Lexicographic Evaluation. ArXiv preprint, 2023.
7. Raj, Mitra, Craswell, & Ekstrand. Patterns of Gender-Specializing Query Reformulation. In Proc. SIGIR, 2023.
8. Bigdeli, Arabzadeh, Seyedsalehi, Mitra, Zihayat, & Bagheri. De-biasing Relevance Judgements for Fair Ranking. In Proc. ECIR, 2023.
9. Larsen-Ledet, Mitra, & Lindley. Ethical and Social Considerations in Automatic Expert Identification and People Recommendation in Organizational Knowledge Management Systems. In proc. FAccTRec Workshop on Responsible Recommendation at RecSys, 2022.
10. Haolun, Mitra, Ma, & Liu. Joint Multisided Exposure Fairness for Recommendation. In Proc. SIGIR, 2022.
11. Li, Li, Mitra, Biega, & Diaz. Exposing Query Identification for Search Transparency. In Proc. Web Conference, 2022.
12. Wu, Ma, Mitra, Diaz, & Liu. A Multi-Objective Optimization Framework for Multi-Stakeholder Fairness-Aware Recommendation. In TOIS, 2022.
13. Cohen, Du, Mitra, Mercurio, Rekabsaz, & Eickhoff. Inconsistent Ranking Assumptions in Medical Search and Their Downstream Consequences. In Proc. SIGIR, 2022.
14. SeyedSalehi, Bigdeli, Arabzadeh, Mitra, Zihayat, & Bagheri. Bias-aware Fair Neural Ranking for Addressing Stereotypical Gender Biases. In Proc. EDBT, 2022.
15. Neophytou, Mitra, & Stinson. Revisiting Popularity and Demographic Biases in Recommender Evaluation and Effectiveness. In Proc. ECIR, 2022.
16. Diaz, Mitra, Ekstrand, Biega, & Carterette. Evaluating Stochastic Rankings with Expected Exposure. In Proc. CIKM, 2020.

“The exercise of imagination is dangerous to those who profit from the way things are because it has the power to show that the way things are is not permanent, not universal, not necessary.”

– Ursula K. Le Guin

Thank you for listening!

@UnderdogGeek bmitra@microsoft.com