How Information Retrieval Systems Construct and Amplify Immigration Narratives

Abstract. Information retrieval systems play a central role in how people access and understand information about complex social issues, including immigration. Yet little is known about how the datasets that underpin these systems represent migrants or structure public narratives about migration. In this paper, we investigate how immigration is framed within a widely used IR benchmark and how ranking models shape the visibility of those frames. Using MS MARCO as our data source, we curate immigration-related queries and annotate retrieved passages using a migration-specific framing taxonomy grounded in social-science research. Our goal is to identify which narratives dominate and to measure how different retrieval models influence their exposure. We find that legality and security frames are far more common than humanitarian or inclusive ones, and that neural reranking amplifies exclusionary portrayals compared to sparse retrieval.

1 Introduction

Information Retrieval (IR) systems play a pivotal role in how people interpret, access, and reason about information. However, a growing body of research shows that these systems often reproduce or amplify societal biases embedded in their underlying data and evaluation benchmarks [30,33]. This concern has motivated a range of studies that examine fairness and representational harm within information retrieval methods. For example, Rekabsaz and Schedl [34] analyzed gender bias neural rankers, showing that biased associations in training corpora can result in skewed retrieval rankings. Similarly, Bigdeli et al. [6,4,5] explored gender biases in relevance judgment datasets, pointing out systematic disparities in how male and female-related contents are labeled for relevance. Such studies are examples that show how information retrieval methods may implicitly encode existing social hierarchies and hence amplify such biases at scale.

Despite this progress, most existing bias-auditing frameworks in IR have mainly focus on demographic attributes that can be explicitly enumerated (e.g., gender categories). These approaches, while valuable, tend to overlook domain-specific issues, where prejudice and inequality are articulated through particular *frames*, rather than explicit labels. One such domain is immigration, where immigrants, refugees, and asylum seekers are portrayed through the lens of national security, border enforcement, and as something to be managed, rather than as vulnerable and marginalized groups [7] [13]. Immigrants constitute one of the world's most vulnerable groups, often portrayed as "illegal," "burdensome," or "threatening," rather than as rights-bearing individuals [29].

Given search systems, driven by information retrieval methods, are often users' first contact with a topic, how the content is framed can shape public perception [15]. The use of frames to analyze content has long been used by migration scholars to study migration discourse in news media [13], social media

[29], and parliamentary debates [35,11]. Card et al. [11] introduced a computational framework for detecting narrative frames in large-scale text corpora, including those related to migration, identifying recurring themes such as humanitarian, threat, and economic frames. Mendelsohn et al. [29] combined qualitative and computational techniques to analyze global migration narratives, revealing consistent moral and emotional structures across contexts. Eberl et al. [13] applied framing theory to European media coverage, showing that news narratives often reinforce securitization and illegality framings of migration, while Sánchez-Junquera et al. [35] highlighted the persistence of dehumanizing metaphors and "crisis" rhetoric in Mediterranean migration coverage. Yet, despite such studies, *frame analysis* has rarely been applied to the datasets that underlie information retrieval systems themselves, where such framings may quietly shape model behavior and user exposure, a gap our work intends to directly address.

Benchmark datasets such as MS MARCO [2] have become the de facto gold standards for training and evaluating information retrieval (IR) systems, yet they were never designed with attention to social meaning or representational bias. As a result, these datasets may silently encode the same cultural and political frames that shape broader public discourse on migration. In this work, we bring frame analysis into the domain of information retrieval to systematically audit how migration and migrants are represented within MS MARCO. Drawing on migration-studies scholarship, we construct a migration-specific framing taxonomy that captures recurring portrayals, such as humanitarian, threat, and economic frames, and employ large language models (LLMs) for large-scale annotation of relevant passages. This enables us to quantify how different framings appear, how they align with specific query types, and how retrieval and reranking models influence their visibility. Our analysis reveals that even ostensibly neutral queries can surface passages reflecting enduring stereotypes about immigrants, exposing subtle yet systematic biases in a central IR benchmark. Our contributions are threefold.

- We develop a frame-analysis methodology for bias auditing in IR datasets, integrating migration-studies theory with computational annotation to detect moral, political, and affective framings in retrieved content.
- We apply this methodology to MS MARCO, constructing a migration-specific framing taxonomy and using large language models (LLMs) to annotate and quantify how immigration is represented across queries and passages.
- We conduct a comparative exposure analysis showing that dense rerankers systematically amplify anti-immigrant framings relative to sparse baselines, revealing how retrieval architectures can shape the visibility of exclusionary narratives.

By bridging migration studies with fairness-oriented IR research [38,17], this paper advances a cross-disciplinary approach to auditing representational harms, encouraging a more reflexive, accountable, and socially informed future for information retrieval, especially as it pertains to the *global immigrant populations*.

2 Theory of Change

Our theoretical grounding draws on a broader tradition that sees data systems as apparatuses of knowledge and power. Foucault described how regimes of truth emerge when power operates through techniques of observation and classification [40,25,19]. In the data system border, as Amoore and Bigo show, migrants are rendered legible through risk scores, datasets, and search results that precede their personhood [1,7]. Stuart Hall's notion of encoding/decoding further illuminates this asymmetry where once media audiences had space to negotiate meaning, algorithmic infrastructures collapse that interpretive agency. Algorithms encode dominant cultural logic, and targets of biased encoding have no meaningful input to the system [26].

This erasure of contestation produces real harm. In the context of migration, web search functions as a public epistemology, a shared means of knowing who migrants are and what migration means. When the top-ranked results disproportionately associate migration with crisis, security or criminality, they not only misinform the public but also shape the policy discourse and bureaucratic practices. Misrepresentation in search results can reinforce stigma, legitimize exclusionary measures, and deepens migrants' lived marginalization. As Nancy Fraser [20] reminds us, justice requires participatory parity; without the capacity to represent themselves within communicative structures, marginalized groups are denied both recognition and agency. Web search tools are one of the primary means of seeking information about different topics, including immigration. The infrastructure that supports such tools, namely information retrieval systems, does not just organize data; it mediates meaning [39]. When users submit a query (e.g. "illegal immigration," "refugee crisis," or "migrant crime"), they engage with algorithmic systems that filter, rank, and amplify particular narratives [41,15,9]. These systems, as scholars like Noble [30] and Eubanks [16] show, often replicate and legitimize pre-existing power relations by privileging dominant framings, while marginalizing migrant voices and perspectives.

Building on Goffman's [24] and Entman's [36] conception of framing as a selective process that highlights certain aspects of reality, we understand IR systems as technological framers. Because their design choices, including indexing, ranking, training data, and evaluation benchmarks, determine what counts as salient, credible, or relevant. Yet, unlike traditional media, these algorithmic systems operationalize frames at scale and without transparency [31]. Following Gamson [21], who viewed framing as a contestable terrain of meaning between social actors, we observe that in IR systems this contestation is foreclosed where migrants, refugees, and their advocates have little participatory agency in defining how they are represented, categorized, or retrieved. Our theory of change responds to this epistemic injustice through three interlinked stages:

(1) Diagnosis: Making Frames Visible. We begin by systematically auditing how migration and migrants are represented in benchmark IR data. Through LLM-assisted annotation and human validation, we identify recurrent framing tropes, including criminality, illegality, cultural threat, economic burden, or humanitarian pity, and quantify their prevalence. This diagnostic process trans-

forms the abstract problem of "bias" into a discursive map of how migration is constructed within the very data that power search and retrieval systems.
(2) Translation: From Exposure to Intervention. Visibility is only the first step. By translating diagnostic insights into concrete interventions, balanced query selection, frame-aware evaluation metrics, and dataset documentation, we aim to reshape the epistemic foundations of IR practice. This stage embodies Fraser's call for participatory representation, expanding fairness beyond technical neutrality to include the right to define oneself within the data. Through transparent and replicable audits, we enable researchers and practitioners to see whose realities are amplified and whose are erased.
(3) Transformation: Toward Epistemic and Institutional Change. At its core, this work challenges what Couldry and Mejias [12] term data colonialism, the extraction of human experience as data without consent or reciprocity. Migrants, whose digital traces are often instrumentalized for surveillance or control, are rarely participants in defining their data narratives. Our long-term vision aligns with Benjamin's [8] and Noble's [30] calls to confront "coded inequity" and reclaim agency over data infrastructures. Through the methodology of discursive auditing, we move beyond performance metrics to interrogate the sociopolitical imaginaries encoded in information-access pipelines [32].

3 Conceptual Framework

This section introduces frame analysis, the large-scale IR dataset we use in this paper, and the pipeline used to curate migration-related queries. We then outline the passage-retrieval and re-ranking setup, the frame-annotation scheme, and the procedures used to measure rank bias.

3.1 Frame Analysis

Framing theory conceptualizes how language selects certain aspects of reality to define problems, diagnose causes, make moral judgments, and suggest remedies [36,37]. The early foundational work by Goffman [24] conceptualized frames as 'interpretation schemas' that help individuals make sense of social reality. Entman [14] later defined framing as the process of selecting and emphasizing certain aspects of reality to promote a particular problem definition, causal interpretation, moral evaluation, or policy recommendation. Gamson [22,21] expanded this concept to show how the media do not just report facts, but structure how audiences interpret those facts. *Frames* serve as ready-made narratives that guide how the public perceives, discusses, and responds to complex topics such as migration, climate change, or inequality. How content is framed shapes public understanding by emphasizing certain aspects of reality (e.g., victimhood or threat) while downplaying others (e.g., economic contribution or human rights).

Within migration studies, frame analysis has become an increasingly prominent method to examine how migrants and migration are represented across news media [3,13], social media [29], as well as political discourse [11,35]. Recent advances in large-scale frame annotation have expanded the methodological toolkit for studying discourse at scale. Early efforts, such as the Media Frames

Corpus [10] and subsequent computational extensions [18,27] relied on manual or lexicon-based annotation schemes, often supported by supervised models fine-tuned on domain-specific datasets. While these approaches enabled systematic mapping of frames across large corpora, they remain resource-intensive and limited in generalizability due to the high cost of manual labeling and the strong domain dependence of framing vocabularies. Framing is inherently contextual with linguistic cues varying across issues, geographies, and cultural settings, making it difficult for models trained on one corpus to transfer effectively to another. To address these challenges, recent work has explored the use of large language models (LLMs) as zero- or few-shot annotators [23], offering scalable, flexible alternatives to fine-tuning. Building on this paradigm, our approach leverages LLMs to perform domain-specific frame detection in immigration discourse, capturing subtle moral and rhetorical cues that static models often overlook.

In order find relevant content to analyze, we first need to curate immigration-related queries and corresponding passages in our target dataset. The following subsections describes this process.

3.2 Dataset Curation from the MS MARCO Dataset

We base our experiments on the MS MARCO Passage Ranking dataset (`v2.1`) [2]. The dataset provides us with a corpus of roughly 1M anonymized search queries and 8.84M passages scraped from the web. MS MARCO *v1.1* provides training labels for relevant passages, but in our study we found too few immigration related queries to leverage training examples. Instead, we use *v2.1* and design strategies for (i) curating immigration related queries, and (ii) retrieving and ranking relevant passages.

Migration Query Curation. Prior studies commonly use of a list of immigration-related seed lexicons capturing issue-specific frames (e.g., policy, border/security, refugees/asylum, deportation, rights, integration, citizenship) [3,11,29,35]. We adopt a similar approach, starting from a set of keywords, filtering by semantic similarity, followed by LLM-based filtering, and finally manual curation.

Stage 1: Lexicon-based Filtering. We constructed a seed lexicon from the migration-studies literature, and further refined via snowball sampling across various dimensions. These dimensions include *general, refugee/asylum seekers, economic, legal, integration, human rights, healthcare, education, race/ethnicity, policy, and visa/applications*. We identified keywords related to each of these dimensions (e.g., economic dimension consists of phrases like 'immigrant unemployment', 'immigration economy', 'immigrant jobs', 'foreign workers', among others.). This was supplemented with some general 'immigration' keywords to ensure high coverage. The full list of seed lexicon is provided on our **GitHub repository**[1].

Stage 2: LLM-based Filtering. We then annotate all the extracted candidate queries from Stage 1 with two lightweight LLMs (Llama4 Scout and Phi-3), labeling each as either *immigration-related* or *irrelevant*. A query is considered

[1] `https://anonymous.4open.science/r/msmarco-bias-B5C7/README.md`

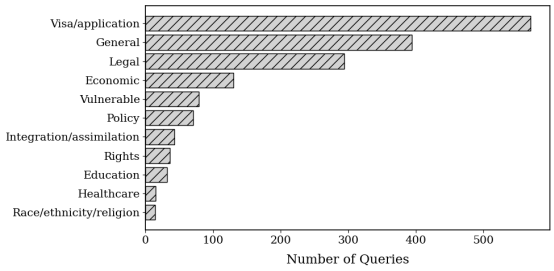

Fig. 1: Distribution of immigration-related queries across thematic categories.

immigration-related if it plausibly concerns movement of peoples across border, or its social, political, legal, and economic dimensions. We iteratively tuned prompts over three rounds; in each round, a simple random sample (N=300) was manually validated to make adjustments to the prompt, and refine inclusion and exclusion criteria to determine relevance. The final inclusion and exclusion criteria for retaining queries are as follows:

- Inclusion criteria: Policies and procedures (e.g., visas, asylum, sponsorship). Immigrant rights and integration (e.g., discrimination, healthcare access, settlement services). Personal/experiential content (e.g., *"how I moved to Australia"*). Public opinion or concerns (e.g., border control, demographic impacts).
- Exclusion criteria: Historical events not tied to current policy. Travel/tourism queries. Citizenship/ancestry topics not framed within immigration systems. Population/ethnicity statistics without migration context. Ambiguous cases were resolved in favor of queries implying settlement migration or discourse about immigration systems/communities.

Note. The final prompt is included on our GitHub repository.

Stage 3: Regular Expression Filtering Subsequent to LLM assessment, we completed another round of filtering using regular expressions to further leave out historical events are not considered relevant for the scope and purposes of the analysis (e.g. trial of tears, Chinese Exclusion Act, among others.). These entries were excluded because they reflect retrospective historical accounts rather than ongoing discussions about immigration. Their inclusion would have mixed historical narratives with present-day discourse and weakened the focus of our study, which aims to examine how current information retrieval datasets reproduce and amplify contemporary framings of immigration.

Selected Query Subset. Out of 13,144 candidate queries, 1,142 (8.6%) were retained as relevant, corresponding to 0.1% of total queries in MS MARCOv2.1. We retained the top two categories by semantic distance with respective representative keywords. Figure1 shows the frequency of queries for each category.

Table 1: Representative examples of framing categories in immigration-related passages, showing illustrative excerpts, and interpretive rationale.

Frame	Example Passage	Justification
C1: Equal Citizens / Anti-Xenophobia	"DACA has been a lifeline... permits allow them to work and defers deportation."	Centers legal protection and inclusion; emphasizes rights to work and protection from removal.
C2: Humanitarian / Suffering Victims	"Deportations are tearing apart American families and communities."	Emphasizes human suffering and family separation.
C3: Security / Criminality Threat	"Illegal Immigration Is Not A Victimless Crime..."	Constructs immigrants as lawbreakers who harm the public and threaten national security.
C4: Economic / Resource Threat	"Illegal immigration costs America some $148 billion annually..."	Portrays immigrants as an economic drain and labor threat.
C5: Cultural / Symbolic Threat	"Sweden opened its doors to Muslim immigrants... pays a high price."	Invokes civilizational clash and erosion of national values.
C6: Illegality Essentialism	"Lawmakers might deny [citizenship] to children born in the U.S. to illegal immigrants."	Frames citizenship as legal purity and exclusion.
C7: Dehumanizing Metaphors	"Trump reportedly said all Haitians have AIDS."	Reduces nationalities to disease and primitiveness.
C8: Racial / Ethnic / Religious Stereotypes	"Japan refuses to surrender its culture to Islam."	Constructs Islam as incompatible and inferior.
C9: Restrictionism / Securitized Border	"Trump's order to withhold federal funds from 'sanctuary cities'..."	Advocates coercive border enforcement and compliance.
C10: Labor Commodification	"Employers can specify how long they plan to employ the 457 visa holder."	Treats migrant labor as flexible and disposable.

Among the annotated queries, visa and application-related questions were by far the most common (570), followed by general (394) and legal (294) queries. Categories such as economic (130), vulnerable populations (79), and policy (70) appeared less frequently. Socially oriented topics like integration/assimilation (43), rights (36), education (32), healthcare (15), and race/ethnicity/religion (14) were relatively rare. Since a query could belong to multiple categories, these frequencies capture overlapping dimensions rather than mutually exclusive themes.

3.3 Passage Retrieval and Re-ranking

For the retained set of queries, we used BM25 as our retriever, and retrieved 1,000 passages for each query resulting in 944,662 unique passages after deduplication. For reranking the retrieved passages by BM25, we used *BAAI/bge-reranker-v2-m3* [28], a reranker built on *xlm-roberta-base* that achieved high scores in benchmarks for retrieval tasks in MSMarco. For each query, we kept the top-10 ranked passages retrieved by BM25 and re-ranked by BGE. This resulted in a final set of *14,645* unique passages. We study the bias exposure of these two rankers in the next section, after frame annotation and validation.

3.4 Framing Taxonomy and Annotation

Framing research distinguishes between issue-generic and issue-specific frames. Issue-generic frames, such as those in the Media Frames Corpus [10], capture

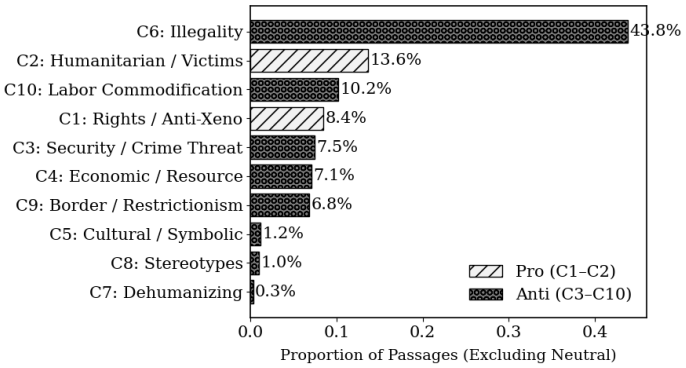

Fig. 2: Distribution of biased frames across passages (excluding neutral).

generalizable dimensions like economic, security, morality, and fairness, allowing comparisons across diverse policy domains. In contrast, issue-specific frames are tailored to a single topic and focus on how particular actors or groups are represented. For immigration, scholars like Benson [3] and Mendelsohn et al. [29]) categorize portrayals of migrants as heroes, victims, or threats, which more directly reflect the moral and political stakes of migration discourse. Building on these traditions, subsequent analyses [13,11] have identified recurring patterns such as economic burden, cultural threat, criminality, and dehumanization.

Through an iterative process, we synthesized these typologies by first, manually inspecting the data, and subsequently doing three iterations of LLM annotation, human validation, and prompt refinement. Our final framework consists of ten frames (C0–C10), which integrate both issue-generic dimensions (economic, cultural, legal, and security) and immigration-specific portrayals (humanitarian, restrictionist, dehumanizing, and labor-focused). It captures valence of frames, i.e., neutral (C0) frames, positive (C1, C2) frames, and negative (C3-C10), while maintaining sufficient granularity to capture what nature of biases are most prominent. Table 1 shows the different frames we identified, and corresponding examples and valence. Figure 2 shows the distribution of passages by frame.

We ran a quantized version of Llama-4-Scout-17B locally using Ollama to annotate frames based on our taxonomy of frames. The prompt, and code we used to do this are provided in our GitHub repository. We allowed for multiple anti-immigrant labels in the annotation for C3-C10, but only one for C0, C1, and C2. This constraint is needed to ensure the LLM does not label something as both neutral and biased. Similarly, it prevents biased frames from overlapping with unbiased frames, allowing us to compare how often biased frames are exposed in high ranks compared to pro-immigrant frames.

The results of the LLM-based frame annotation reveal a highly uneven distribution across the ten frame categories. The vast majority of passages (C0: 12,618) were classified as *Neutral / Informational / Historical*. Considering these are passages on the web, and a majority of our queries were about visa applications and other information seeking queries, this is unsurprising.

3.5 Human Validation

To validate the results of the LLM, we provided two human annotators the same prompt we provided the LLM, and asked them to qualitatively assess the quality of the annotations. Both annotators were given 30 examples (except C5, C7, and C8 which had less) from each of the frames, with Annotator A inspecting the first 6 frames (C0 - C5) and Annotator B the rest (C6 - C10). We judged the LLM on how well it was able to determine valence- neutral, negative, or positive frames, and how well it performed in identifying individual frames. Based on the human validation labels, we found the LLM performs reasonably well at identifying neutral (100% accuracy, 30/30), pro-immigration (80%, 45/56), and anti-immigration (82%, 71/86) frames, making our analysis of rank bias in the next section reasonably justified.

When it came to identifying individual frames the results were mixed. *C0: Neutral* frames were correctly labeled on all 30 examples. *C1: Equal Citizens/Anti-xenophobia* frames identified by LLM were also quite accurate (92%) with the annotator unsure about 4 out of 30 examples, and only 2 where they disagreed with the LLM labels. The two examples reveal the difficulty of the task. The annotator agreed that the passage expressed sympathy towards immigrants, but still encouraged reporting illegal immigrants. This makes the passage also a legitimate candidate for C6, or C9.

For C2, the annotators observed 9 out of 30 examples to be historical accounts, not directly related to present day movement of people (which is a criteria we use for separating relevant vs irrelevant passages). The LLM identified sympathetic tone towards immigrants, in a sense, but not in the way we intended. This underscores some key limitations in our work, which we discuss further in the final limitations section of this paper. Among the biased frames, the annotators had some disagreements with the LLM about certain labels such as C3, C5, and C9, where the annotator agreed they belonged to one or more of the frames but took issue with certain passages belonging to one over the other. This is another limitation we acknowledge, and for future work, we plan to more clearly define criteria for these particular frames.

3.6 Distribution and Political Context of Framing Patterns

Building on the validated frame annotations, we next examine how different framings are distributed across the corpus and how they reflect broader political contexts. Our focus here is on non-neutral frames, that is, passages carrying positive or negative evaluative meaning. Among these, *Illegality Essentialism* (C6, 1,012 passages) emerged as the most frequent, followed by *Security and Criminality Threat* (C3, 172 passages). This distribution points to a dominance of legality-oriented discourse, where immigration is often represented through themes of law enforcement, criminality, and border control. The term *illegal* appeared in 12.9% (N = 1,893) of all passages, showing how strongly this framing permeates the dataset. In contrast, cultural and racialized framings (C5–C8) were relatively rare, with only a small number of passages invoking symbolic, ethnic, or dehumanizing portrayals.

To further contextualize these patterns, we examined the presence of political figures and institutions within the annotated passages. Mentions of *Donald Trump* were particularly frequent, revealing the political context embedded in the corpus. While most references to Trump appeared within neutral or informational content (C0, 3.0%), their share was substantially higher within *Security/Criminality Threat* (C3, 27.3%), *Illegality Essentialism* (C6, 25.3%), and *Economic/Resource Threat* (C4, 25.6%) frames. These patterns suggest that discussions of the Trump presidency were commonly associated with restrictive or punitive narratives about immigration. One annotator noted that several passages labeled as *Humanitarian/Suffering Victims* (C2) also expressed fear or vulnerability linked to Trump-era policies, such as in the example "the denial of a hardship waiver would mean exposure of undocumented status to the US government... leaving immigrants feeling very vulnerable in the age of Trump."

References to *Immigration and Customs Enforcement (ICE)* followed a similar trend. They appeared most often in *Restrictionism/Securitized Border* (C9, 22.2%) and *Security/Criminality Threat* (C3, 10.5%) frames, emphasizing ICE's symbolic association with enforcement and surveillance. Together, these findings indicate that political actors and agencies serve as framing anchors within the dataset, reinforcing legality-centric and enforcement-oriented representations of migration. The prominence of such associations suggests that benchmark IR corpora like MS MARCO do not simply mirror the informational landscape of the web but also reproduce its political hierarchies, shaping the narrative boundaries from which retrieval models learn and generalize.

4 Rank Exposure of Biased Frames

In this section, we analyze how prominently non-neutral frames are ranked for immigration related query. Exposure to biased content at top of search results reinforce existing biases. Across all queries, neutral or informational passages dominate the top-ranked results, comprising roughly 43% of retrieved content for both systems. Yet when focusing on evaluative framings, *anti-immigrant* passages are consistently more prevalent than *pro-immigrant* ones. In the BM25 baseline, **5.2%** of top-10 passages express anti-immigrant or threat-based sentiment, compared to only **1.8%** that portray migrants positively or emphasize rights and inclusion. After reranking with BGE, this imbalance widens where anti-immigrant passages rise to **5.7%**, while positive ones decline to **1.5%**.

We further perform a Wilcoxon signed-rank test, which confirms that this difference is statistically significant ($W = 13{,}883.5$, $p = 0.0003$), indicating that semantic reranking systematically increases the visibility of exclusionary or threat-oriented framings across queries. Although these changes appear modest in absolute terms, their consistency across the dataset suggests that reranking may inadvertently amplify anti-immigrant narratives.

A closer look at rank position reveals that such bias is *front-loaded*. Among the top three passages per query, the share of anti-immigrant framings rises from 10.5 percent under BM25 to 11.7 percent in BGE, an increase of roughly 1.2 percentage points. Similar patterns hold for the top-five and top-ten results, with anti-immigrant content consistently higher after reranking ($\Delta T. \approx 0.9$–1.0

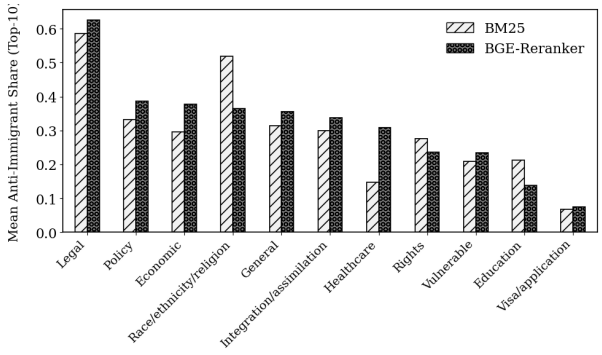

Fig. 3: Mean anti-immigrant share across query categories for BM25 and BGE.

pp). These small but systematic differences at the very top of the ranking indicate that reranking not only reproduces but also front-loads exclusionary narratives, thereby shaping initial information exposure and potential user interpretation.

Figure 3 shows anti-immigration frames by query category. The highest baseline levels of anti-immigrant framing appear in *legal* and *policy*-related queries, where BM25 already retrieves a large share of threat-oriented passages (*58.6%* and *33.2%*, respectively). After reranking, these proportions rise further to *62.5%* in *legal* and *38.6%* in *policy* queries, indicating a systematic amplification of exclusionary framings in institutional and regulatory contexts. *Economic* queries exhibit the most pronounced increase, from *29.5% to 37.8% (+8.3 pp)*, suggesting that discourses around migrants as economic actors remain susceptible to burden- or competition-based narratives.

By contrast, some thematic areas show a reduction in anti-immigrant content. Queries centered on *race, ethnicity, or religion* display a sharp decline (from *51.9% to 36.5%*, -15.4 pp), while *rights*-focused and *education*-related queries also decrease modestly (-3.9 pp and -7.5 pp). These shifts suggest that semantic reranking may surface more balanced or inclusive material when the query itself explicitly invokes equality or social-justice framings. Other categories occupy the middle ground. *Healthcare* queries nearly double in anti-immigrant content (from 14.7% to 30.8%, +16 pp), and *integration/assimilation* queries increase slightly (from 29.9% to 33.7%, +3.8 pp). These results reveal a pattern where the BGE reranker *amplifies bias selectively*, heightening exclusionary framings in domains tied to economics, legality, and public-service access while attenuating them in overtly rights-based or demographic-sensitivity contexts. This behaviour points to the need for *domain-specific auditing* of retrieval systems, as aggregate fairness metrics alone can obscure substantial disparities across issue areas.

While topic-specific patterns reveal where anti-immigrant bias is most pronounced, they do not capture how such content is positioned within ranked results. Since users are disproportionately influenced by higher-ranked items, even small shifts in ranking order can substantially affect what information is seen and trusted. We quantify bias in *exposure*, i.e., which frames users are likely to see at higher ranks using complementary metrics. Following [36,13,11], where

prior work compared exposure of positive vs. negative portrayals (or other groupings) in early ranks, we use two metrics to compare exposure of biased frames for BM25, and BGE rankers. To evaluate whether anti-immigrant passages are not only more frequent but also ranked more prominently, we computed two exposure-sensitive fairness metrics, namely Average Rank Bias (ARaB) [34] and Normalized Fairness in Ranked Results (NFaiRR) [33].

The ARaB metric quantifies the average rank difference between anti- and pro-immigrant passages (lower values indicating fairer exposure), while NFaiRR measures overall parity in visibility between groups (values closer to 1 indicate greater fairness). While the earlier top-k analysis highlighted visible differences in the proportion of biased passages at fixed ranks, ARB and NFaiRR provide rank-sensitive measures that capture cumulative exposure disparities across all retrieved results. Across all queries, both measures reveal that neural reranking (BGE ranker) amplifies exposure bias. The ARaB metric for the BGE reranker (0.396) is roughly ten times higher than that of the BM25 baseline (0.039), indicating that anti-immigrant passages are consistently ranked earlier and thus granted greater visibility. Correspondingly, NFaiRR declines from 0.123 under BM25 to 0.098 under BGE, reflecting a measurable reduction in exposure parity between positive and negative framings. We find that these results demonstrate that the reranker not only increases the overall prevalence of anti-immigrant passages but also front-loads them in the ranking, reinforcing unequal visibility and amplifying exclusionary narratives at the point of first user contact.

5 Concluding Remarks

This paper presented a framework for examining bias in information retrieval datasets using migration framing. By drawing on research in migration studies and applying large-scale frame analysis to MS MARCO, we showed how retrieval datasets can shape the way immigrants are represented. The results demonstrate that current retrieval systems may reinforce common political and legal framings of migration. This work highlights the importance of considering how information retrieval systems participate in shaping public understanding of immigration and of incorporating social context into their evaluation and design.

This study has several limitations that point to avenues for future work. First, while our annotation pipeline achieved promising coverage, its validation was limited where only two annotators independently reviewed a subset of 100 examples to assess frame quality and consistency. A more rigorous evaluation using larger samples, inter-annotator agreement metrics, and additional domains would strengthen the reliability of the frame labels. Second, our retrieval experiments were conducted using a lightweight BGE reranker, which, although suitable for initial testing, does not capture the full range of possible retrieval architectures or parameter settings. Future work should evaluate whether these findings generalize across stronger or differently optimized rerankers. Finally, while we employed large language models in zero- and one-shot annotation settings to reduce manual effort, fine-tuned models trained on a sufficiently large, high-quality set of human-labeled examples may yield more stable and domain-adapted frame predictions.

References

1. Amoore, L.: Cloud ethics: Algorithms and the attributes of ourselves and others. Duke University Press (2020)
2. Bajaj, P., Campos, D., Craswell, N., Deng, L., Gao, J., Liu, X., Majumder, R., McNamara, A., Mitra, B., Nguyen, T., et al.: Ms marco: A human generated machine reading comprehension dataset. arXiv preprint arXiv:1611.09268 (2016)
3. Benson, R.: Shaping immigration news: A French-American comparison. Cambridge University Press (2013)
4. Bigdeli, A., Arabzadeh, N., SeyedSalehi, S., Zihayat, M., Bagheri, E.: Gender fairness in information retrieval systems. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 3436–3439 (2022)
5. Bigdeli, A., Arabzadeh, N., Seyedsalehi, S., Zihayat, M., Bagheri, E.: Understanding and mitigating gender bias in information retrieval systems. In: European Conference on Information Retrieval. pp. 315–323. Springer (2023)
6. Bigdeli, A., Arabzadeh, N., Zihayat, M., Bagheri, E.: Exploring gender biases in information retrieval relevance judgement datasets. In: European Conference on Information Retrieval. pp. 216–224. Springer (2021)
7. Bigo, D.: Security and immigration: Toward a critique of the governmentality of unease. Alternatives **27**(1_suppl), 63–92 (2002)
8. Boccio, R.: Race after technology: Abolitionist tools for the new jim code by ruha benjamin. Configurations **30**(2), 236–238 (2022)
9. Boyd, D., Golebiewski, M.: Data voids: Where missing data can easily be exploited (2018)
10. Card, D., Boydstun, A., Gross, J.H., Resnik, P., Smith, N.A.: The media frames corpus: Annotations of frames across issues. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). pp. 438–444 (2015)
11. Card, D., Chang, S., Becker, C., Mendelsohn, J., Voigt, R., Boustan, L., Abramitzky, R., Jurafsky, D.: Computational analysis of 140 years of us political speeches reveals more positive but increasingly polarized framing of immigration. Proceedings of the National Academy of Sciences **119**(31), e2120510119 (2022)
12. Couldry, N., Mejias, U.A.: The costs of connection: How data is colonizing human life and appropriating it for capitalism. In: The costs of connection. Stanford University Press (2019)
13. Eberl, J.M., Meltzer, C.E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Boomgaarden, H.G., Schemer, C., Strömbäck, J.: The european media discourse on immigration and its effects: A literature review. Annals of the International Communication Association **42**(3), 207–223 (2018)
14. Entman, R.M.: Framing: Towards clarification of a fractured paradigm. McQuail's reader in mass communication theory **390**, 397 (1993)
15. Epstein, R., Robertson, R.E.: The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. Proceedings of the national academy of sciences **112**(33), E4512–E4521 (2015)
16. Eubanks, V.: Automating inequality: how high-tech tools profile, police, and punish the poor. (No Title)
17. Fang, Y., Singh, A., Tao, Z., et al.: Fairness in search systems. Foundations and Trends® in Information Retrieval **18**(3), 262–416 (2024)

18. Field, A., Kliger, D., Wintner, S., Pan, J., Jurafsky, D., Tsvetkov, Y.: Framing and agenda-setting in russian news: a computational analysis of intricate political strategies. arXiv preprint arXiv:1808.09386 (2018)
19. Foucault, M., et al.: The ethic of care for the self as a practice of freedom (1987)
20. Fraser, N.: Rethinking the public sphere: A contribution to the critique of actually existing. Democracy"." In Habermas and the Public Sphere, edited by C. Calhoun pp. 109–142 (1999)
21. Gamson, W.A.: Talking politics. University of Cambridge (1992)
22. Gamson, W.A., Modigliani, A.: Media discourse and public opinion on nuclear power: A constructionist approach. American journal of sociology **95**(1), 1–37 (1989)
23. Gilardi, F., Alizadeh, M., Kubli, M.: Chatgpt outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences **120**(30), e2305016120 (2023)
24. Goffman, E.: Frame analysis: An essay on the organization of experience. Harvard university press (1974)
25. Gordon, C., Foucault, M.: selected interviews and other writings 1972- 1977 (1980)
26. Hall, S., Hobson, D., Lowe, A., Willis, P.: Culture, media, language: Working papers in cultural studies, 1972-79. Routledge (2003)
27. Kwak, H., An, J., Ahn, Y.Y.: A systematic media frame analysis of 1.5 million new york times articles from 2000 to 2017. In: Proceedings of the 12th acm conference on web science. pp. 305–314 (2020)
28. Li, C., Liu, Z., Xiao, S., Shao, Y.: Making large language models a better foundation for dense retrieval (2023)
29. Mendelsohn, J., Budak, C., Jurgens, D.: Modeling framing in immigration discourse on social media. arXiv preprint arXiv:2104.06443 (2021)
30. Noble, S.U.: Algorithms of oppression: How search engines reinforce racism. In: Algorithms of oppression. New York university press (2018)
31. Pasquale, F.: The black box society: The secret algorithms that control money and information. In: The black box society. Harvard university press (2015)
32. Raji, I.D., Smart, A., White, R.N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., Barnes, P.: Closing the ai accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the 2020 conference on fairness, accountability, and transparency. pp. 33–44 (2020)
33. Rekabsaz, N., Kopeinik, S., Schedl, M.: Societal biases in retrieved contents: Measurement framework and adversarial mitigation of bert rankers. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 306–316 (2021)
34. Rekabsaz, N., Schedl, M.: Do neural ranking models intensify gender bias? In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 2065–2068 (2020)
35. Sánchez-Junquera, J., Chulvi, B., Rosso, P., Ponzetto, S.P.: How do you speak about immigrants? taxonomy and stereoimmigrants dataset for identifying stereotypes about immigrants. Applied Sciences **11**(8), 3610 (2021)
36. Scheufele, B.: Framing: Toward clarification of a fractured paradigm: von robert m. entman (1993). In: Schlüsselwerke: Theorien (in) der Kommunikationswissenschaft, pp. 115–127. Springer (2022)
37. Scheufele, D.A.: Framing as a theory of media effects. Journal of communication **49**(1), 103–122 (1999)

38. Seyedsalehi, S., Bigdeli, A., Arabzadeh, N., AlMousawi, B., Marshall, Z., Zihayat, M., Bagheri, E., et al.: Understanding and mitigating gender bias in information retrieval systems. Foundations and Trends® in Information Retrieval **19**(3), 191–364 (2025)
39. Söderberg, J.: Media technologies: Essays on communication, materiality, and society. (2014)
40. Tarbet, D.W.: Discipline and punish: The birth of the prison. (1978)
41. Trielli, D., Diakopoulos, N.: Search as news curator: The role of google in shaping attention to news information. In: Proceedings of the 2019 CHI Conference on human factors in computing systems. pp. 1–15 (2019)