International Studies Quarterly (2022) **66**, sqac044

Combining Computational and Archival Methods to Study International Organizations: Refugees and the International Labour Organization, 1919–2015

RESEARCH NOTE

William L Allen
Department of Politics and International Relations, University of Oxford, Oxford, UK

and

Evan Easton-Calabria
Feinstein International Center, Tufts University, Boston, USA; Refugee Studies Centre, University of Oxford, Oxford, UK

Researchers studying international organizations have access to growing and varied archives due to digitization efforts. While developments in computational methods confer efficiency gains for examining these materials at scale, they raise concerns about their validity when applied to interpretive tasks in historical settings. In response, we present a general and flexible workflow that uses simple computational techniques from linguistics to enhance archival researchers' interpretive skills and sensitivity to historical contexts. These techniques also identify patterns that can serve as evidence of causal mechanisms when embedded within strong research designs and theoretical expectations. Then, we demonstrate our mixed-method approach by applying it to a dataset of International Labour Organization's (ILO) annual reports spanning ninety-three years. Examining the ILO's engagement with refugees as described in these documents, we identify key moments during which refugees have been particularly salient for this organization, and the emergence of new issues on its high-level policy agenda.

Los investigadores que estudian las organizaciones internacionales tienen acceso a una cantidad de archivos cada vez mayor y más variados gracias a los esfuerzos de digitalización. Si bien los avances en los métodos informáticos permiten mejorar la eficiencia para examinar estos materiales a escala, plantean dudas sobre su validez cuando se aplican a tareas de interpretación en entornos históricos. En respuesta, presentamos un flujo de trabajo general y flexible que utiliza técnicas informáticas sencillas de la lingüística para mejorar las habilidades relacionadas con la interpretación de los investigadores de archivos y su sensibilidad a los contextos históricos. Estas técnicas también identifican patrones que pueden servir como prueba de mecanismos causales cuando se integran en diseños de investigación y expectativas teóricas sólidos. A continuación, demostramos nuestro enfoque de métodos mixtos aplicándolo a un conjunto de datos de informes anuales de la Organización Internacional del Trabajo (OIT) que abarcan 93 años. Al examinar el compromiso de la OIT con los refugiados, tal y como se describe en estos documentos, identificamos los momentos clave en los que los refugiados han sido especialmente importantes para esta organización, así como la aparición de nuevas cuestiones en su agenda política de alto nivel.

Les chercheurs étudiant les organisations internationales ont accès à des archives de plus en plus nombreuses et variées grâce aux efforts de numérisation. Bien que les évolutions des méthodes informatiques permettent de gagner en efficacité pour l'examen global de ces données, elles soulèvent des préoccupations quant à leur validité lorsqu'elles sont appliquées à des tâches d'interprétation dans des contextes historiques. En réponse à cela, nous présentons un flux opérationnel général flexible qui repose sur des techniques informatiques simples issues de la linguistique pour améliorer les compétences interprétatives des chercheurs en archives et leur sensibilité aux contextes historiques. Ces techniques permettent également d'identifier des modèles qui peuvent servir de preuves de mécanismes causaux lorsqu'ils sont intégrés à de solides conceptions de recherche et attentes théoriques. Nous illustrons ensuite notre approche par méthodes mixtes en l'appliquant à un jeu de données constitué de rapports annuels s'étendant sur 93 ans de l'Organisation internationale du travail (OIT). Nous avons examiné l'implication de l'OIT auprès des réfugiés décrite dans ces documents et nous avons identifié des moments clés durant lesquels les réfugiés ont été particulièrement importants pour cette organisation ainsi que l'apparition de nouveaux sujets dans son programme politique à haut niveau.

William L. Allen is a British Academy Postdoctoral Fellow at the Department of Politics and International Relations, University of Oxford. He focuses on political behavior, political communication, and migration.

Evan Easton-Calabria is Senior Researcher at the Feinstein International Center, Tufts University, and a Research Associate at the Refugee Studies Centre, University of Oxford. She focuses on refugee livelihoods and self-reliance, and the role of local actors in refugee responses.

Authors' note. This work was supported by the UK Economic and Social Research Council (grant number ES/J500112/1), the British Academy (grant number PF21\210066), and the Swiss Agency for Development and Cooperation/Refugee Studies Centre, University of Oxford. Earlier versions of this paper were presented at the 2018 conference "Researching the United Nations and Other International Organizations: Rethinking Methods of Investigation" (June 18–20, Geneva, Switzerland), the 2020 European Consortium for Political Research (ECPR) General Conference (August 24–28, online), and an internal seminar at the Centre on Migration, Policy, and Society (COMPAS), University of Oxford. We would like to thank the participants and discussants at those events for their useful feedback, as well as the anonymous reviewers and *International Studies Quarterly* editorial team for further constructive comments. The documents underlying this article are available on the International Labour Organization Digital Repository at www.labordoc.ilo.org. The data underlying this article are available on the ISQ Dataverse, at https://dataverse.harvard.edu/dataverse/isq.

Allen, William L, and Evan Easton-Calabria. (2022) Combining Computational and Archival Methods to Study International Organizations: Refugees and the International Labour Organization, 1919–2015. *International Studies Quarterly*, https://doi.org/10.1093/isq/sqac044

Introduction

Researchers examining international organizations often draw upon historical documents produced or held by these bodies (e.g., Bradley 2021). Archives potentially contribute valuable knowledge about these organizations' past activities, which escapes approaches solely involving contemporaneous materials such as interviews or surveys (McGarr 2020). For example, discursive analyses can link documents' contents with their historical contexts to reveal political strategies (Wodak 2001). Archival materials can also address questions involving identifying effects of events or developing and testing mechanisms that link causes with outcomes (Beach 2016).

Organizations' documents have become easier to access as archives have been digitized to varying degrees, such as the International Labour Organization Digital Repository, the World Bank Groups' Digital Collections, and the Total Digital Access to the League of Nations Archives Project.[1] Responding to this volume and variety of data, social scientists increasingly turn to computational methods and tools to study historical events at scale and in multiple languages (Lucas et al. 2015; Cirone and Spirling 2021; Trubowitz and Watanabe 2021). Yet, despite conferring remarkable efficiency gains, these methods—when applied in isolation—attract concerns about their validity for addressing the range of questions and concepts that scholars of international politics bring to archival materials (van Atteveldt, van der Velden, and Boukes 2021; Baden et al. 2022). More broadly, they raise questions about how to incorporate knowledge about historical contexts and constraints into research designs that increasingly feature quantitative elements (Dennison 2021).

In response, we present and demonstrate an iterative approach to studying large amounts of digitized archival data that increasingly typify those made available by international organizations. Specifically, we arrange existing computational and archival methods to create a general and flexible workflow for making sense of relatively large amounts of documentary materials. Despite sharing some features with corpus-assisted discourse studies (Partington, Duguid, and Taylor 2013), it remains distinctive in how it specifically engages with, and aims to enhance, the interpretive skills and goals of archival researchers who use repositories of domain- and organization-specific documents. We also argue that this approach, coupled with well-justified research designs and theoretical expectations, can identify evidence related to causal mechanisms (Beach 2016). This is particularly the case for rare yet significant usages of terms that may be difficult to find in huge amounts of text solely through close reading (Quinn et al. 2010).

The first part of this note explains the workflow, focusing on three key features: it uses simple and established techniques to identify textual patterns across an entire dataset as well as within subsets of data; it preserves qualitative access to the data that facilitate more focused, interpretive inquiry; and it operates serially and iteratively whereby knowledge gained in one step informs the next step. In the second part, we demonstrate this mixed-method approach by applying it to a dataset of ninety-three annual reports, comprising over eighteen million words, produced by the International Labour Organization (ILO) between 1919 and 2015.[2] These materials, to the best of our knowledge, have not been collectively examined as windows into the ILO. We specifically chose to examine the salience of refugees during this period because prior scholarship indicates that the ILO's mission and work have involved this population (Kott and Droux 2013). Yet, attempting to systematically examine the history of the ILO's work with refugees created a situation of high search costs for uncertain results—not an atypical position for researchers using archives. Our analysis reveals three key moments of refugees' salience in the ILO's high-level agenda over the last century, which is of qualitative interest to understanding how the organization has operated. It also identifies evidence of certain terminologies' recency. These findings open further avenues for research into international organizations.

How Computational and Archival Methods Can Inform Each Other

Addressing Interpretive and Causal Questions Using Archival Methods and Materials

Commonly understood as identifying, selecting, and analyzing data from archival records, "archival methods" encompass several techniques and approaches that investigate texts "created at some point in the relatively distant past" (Ventresca and Mohr 2017, 805). These include content analysis of historical sources to explicitly theoretical readings of history, and involve both inductive and deductive reasoning (Kipping, Wadhwani, and Bucheli 2014). Often, researchers use these techniques to address interpretive questions, such as making sense of beliefs, actions, and motivations that are latent within texts (Bevir and Rhodes 2012). While doing so, they scrutinize the people and conditions that generated the documents, being attuned to reading "along," "with," and "against" the grains of archives (Stoler 2002) to contextualize and justify conclusions arising from close reading.

In contrast, social scientists have increasingly turned to archival materials with expressly causal tasks in mind, such as finding evidence of historical events' *effects* or of *mechanisms* that link proposed causes to an observed outcome (Rohlfing 2012). Depending on researchers' explicit theorization of what constitutes evidence of such effects or mechanisms—"fingerprints" in the empirical record (Beach 2016)—documents can be sources of information that either strengthen or weaken confidence in proposed claims. For example, if a theorized mechanism involves an organization prioritizing refugees in its plans, then finding repeated mentions of that population in its public record may be an important piece of trace evidence that supports the mechanism.

However, in both interpretive and causal modes, selecting observations from archival material presents challenges. Inclusion criteria may be unknowingly yet systematically biased either in terms of what investigators determine to be "interesting" or "relevant," or in what is practically feasible to collect (Tirabassi 2010; Cirone and Spirling 2021). Moreover, evaluating the typicality of those materials—whether they represent continuity or divergence with established patterns—is difficult without knowing their relation to other parts of the archive.[3]

[1] See https://www.worldbank.org/en/archive/digital-collections; and https://libraryresources.unog.ch/leagueofnationsarchives/Digitized/LeagueArchives.

[2] As explained later, some reports either are missing or were not published.

[3] More fundamentally, empirical materials from which to select observations may be missing altogether due to organizations' inability or unwillingness to produce observable fingerprints—especially on subjects perceived to be either unimportant (and unworthy of documentation) or sensitive (Gonzalez-Ocantos and LaPorte 2021). This problem differs from practical issues of archival digitization, where documents may be human- but not machine-readable.

Some approaches under the banner of "text as data" attempt to identify patterns using computationally and statistically intensive methods (Grimmer and Stewart 2013). These patterns may include sentiment (positive or negative tonality), topics (indicated by co-occurring words), or events (Trubowitz and Watanabe 2021). These approaches present impressive gains in efficiency, particularly when it comes to processing large amounts of text (Barberá et al. 2021). Yet, their empirical validity for certain tasks is still debated. On one hand, these methods potentially provide entry points into archives by answering "what is this corpus about?" (Murakami et al. 2017). On the other hand, recent work observes how widely used computational methods (including dictionaries and machine learning) fall short of humans' performance when it comes to identifying sentiment and conceptually meaningful themes (Brookes and McEnery 2018; van Atteveldt, van der Velden, and Boukes 2021). This is partly due to the tendency of computational text analysis toward developing more specialized tools that neither fully reflect social scientists' deep knowledge of how language operates nor generate valid measures of complex constructs as demanded in specific social settings (Baden et al. 2022).

While digital humanities scholars and historians have productively reflected on using quantitative textual analysis to aid interpretive work (Anderson 2007; Klein, Eisenstein, and Sun 2015), these methodological insights have not been fully applied to either archival data or the domain of international organizations (although see Ongenaert and Joye 2019). We address this gap by demonstrating how simple tools can augment the workflow and materials of social scientists engaging in archival research. If testing approaches to text analysis involves examining "their ability to perform some useful social scientific task" (Grimmer and Stewart 2013, 270), we argue that three features are important for archival contexts: reliably identifying relevant portions of a corpus for further examination, while acknowledging multiple ways of conceptualizing and ascertaining relevancy; measuring the typicality of an observation against some reference; and comprehensively accessing qualitative examples for disambiguation and closer analysis.

Combining Computational and Archival Methods: An Iterative Approach

Our approach, arranging established computational and corpus linguistic methods (Biber and Reppen 2015) alongside archival methods, addresses these needs. Moreover, it does so flexibly—both in terms of the precise techniques involved and in terms of the order of steps. Our workflow shares some features with other approaches to studying language such as corpus-assisted discourse studies (Partington, Duguid, and Taylor 2013), which have also been used to identify representations of migrants (e.g., Taylor 2014). Yet, it remains distinctive in how it foregrounds historical knowledge and sensitivities in addressing the range of questions that social scientists bring to archival materials (Dennison 2021).

Our workflow comprises three techniques. First, *frequency analysis* counts all instances of a given word or group of words, and then displays these frequencies either in terms of the whole corpus or differentiated by subsets of texts ("subcorpora"). Usually, to enable comparison of frequencies arising from differently sized subcorpora, these frequencies are normalized into occurrences per thousand or million words (McEnery and Hardie 2011). For example, identifying shifts in the frequencies of terms such as "refugee" or "migration" over time could indicate changes in organizational attention to these populations. Our general workflow is agnostic as to how researchers determine what constitutes meaningful shifts for their specific situation. Possible options include relying on visual inspection for deviations from a selected baseline, quantitatively identifying discrete periods in time series that mark shifts in frequencies (Killick and Eckley 2014), or focusing on spikes that exceed predetermined statistical boundaries of normal variation (Gabrielatos et al. 2012).

Second, *keyword analysis* identifies terms that are central to a corpus or subcorpus. Typically, this involves comparing how frequently a word appears in one set of texts to how frequently it appears in another set. Although there are multiple keyness measures, each with advantages and disadvantages (McEnery and Hardie 2011), the point of including this technique is that words more unique to one of the sets may be useful candidates for exploring what that set might be about (Taylor and Marchi 2018).

Such qualitative investigation can be facilitated by the third technique, *concordance analysis*, which displays some of the content surrounding an observation of a word or phrase of interest to facilitate the kinds of close reading already familiar to historical researchers. If keyword or frequency analysis suggests that "immigration" is particularly important, researchers could use concordances to read these usages in context to interpret their meanings (Veeramoothoo 2022). Moreover, concordances can disambiguate words with multiple senses: the word "asylum" could be used in the context of migration—as in "asylum-seekers"—or mental health. In a study about refugees, concordance analysis would enable identifying and retaining the former usages while dropping the latter set.

Next, we arrange these techniques so that the outputs of one step become inputs for subsequent steps, as summarized in figure 1. This sequential roadmap with multiple pathways differs from approaches such as "triangulation" (Baker and Egbert 2016), which involve comparing analyses from parallel procedures. Rather, it echoes the current practice of concatenating methods to enhance validity in specific settings (Lee et al. 2017; Baden et al. 2022). As such, it can accommodate a range of queries about the presence of terms of interest (e.g., particular populations, places, actions; what we term "pathway A") or characteristics of the archive (e.g., which kinds of terms distinguish a given subcorpus; what we term "pathway B").

Both pathways begin by recognizing how choices about document selection produce datasets that are informed by researchers' own historical knowledge and interests.[4] Moreover, creating machine-readable corpora for computational analyses involves practical matters of digitization: not all available methods for optical character recognition (OCR) may perform equally well on different types of documents, particularly in languages other than English (Hegghammer 2021; see footnote 3). Next, if the goal is to examine usages of particular terms, the following step along pathway A involves ascertaining the frequency of those terms to reveal areas of the materials that contain potentially relevant information.[5] These areas might correspond to time periods, document types, messengers, or an intersection of dimen-

[4] Here, historical researchers' sensitivity to the conditions under which archives were constructed—and the types of documents they contain and omit—is valuable (Dennison 2021; see footnote 3).

[5] To be clear, depending on the question, large frequencies may not always indicate qualitative "relevance." Finding examples of less common "minority discourses" (Baker 2006) is an important part of historians' approach to reading "against" the archival grain.

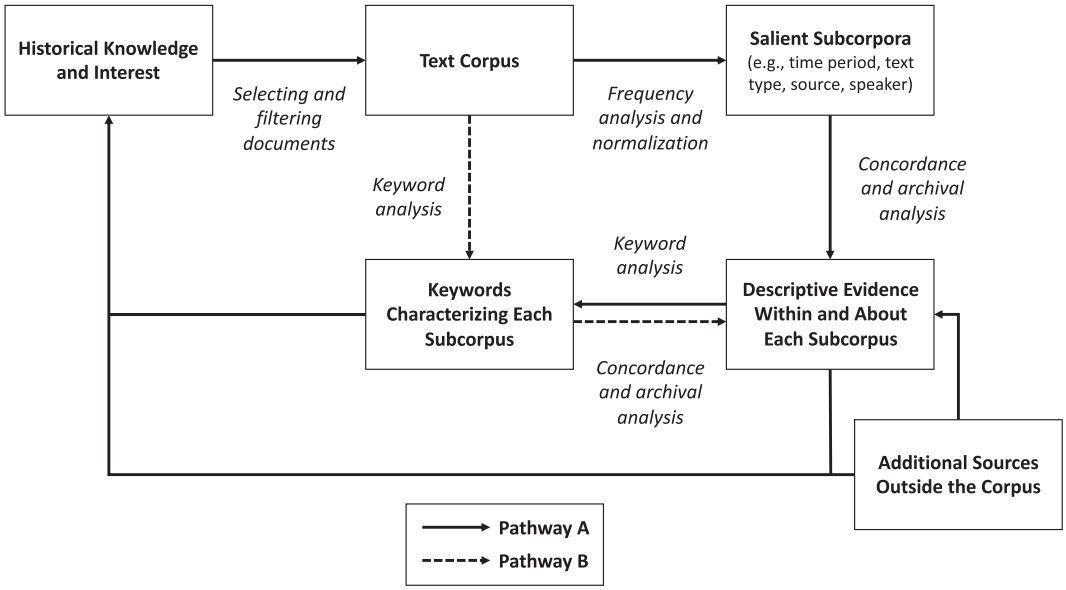

Figure 1. A general approach to documentary analysis that combines computational and archival methods.

sions. Then, researchers can retrieve specific mentions of the terms through concordance analysis and interpretive analysis to generate a fuller qualitative picture: for example, comparing how an organization portrays refugees in different periods might indicate shifts in priorities.

In contrast, initial research questions may involve identifying which features characterize a given subcorpus: which issues, for example, are more prominent in recent years compared to earlier periods? In this case, the next step along pathway B involves conducting keyword analyses among broad subcorpora to identify terms that are good candidates for further concordance and archival analysis. The intuition behind keyword analysis is that words more frequently present in one body of texts (a "focus" corpus) compared to another body of texts (a "reference" corpus) potentially signal something distinctive about the former texts. Consequently, these terms and their surrounding text are worth investigating in greater depth. However, linguists have debated about which statistical techniques most appropriately capture and express this difference (Gabrielatos 2018). Working from the observation that "models that are less sophisticated in the use of language may provide more useful analysis of texts" (Grimmer and Stewart 2013, 270), for this demonstration we use a simple comparison of proportions (Kilgarriff 2009),

$$\frac{f_a + N}{f_b + N},$$

where f_a expresses a word's frequency normalized per million words in the focus corpus, f_b expresses that same word's normalized frequency in the reference corpus, and N is a parameter set by the researcher.[6]

What advantages does assembling these computational and archival methods into a general approach afford over either set of methods in isolation? First, for researchers working in interpretive traditions, it provides an efficient means of directing attention to parts of the archive that may be more promising and relevant for deeper qualitative analysis. Frequency analysis, for instance, produces a simple yet informative way of ascertaining both the presence and the concentration of words of interest—or indeed the *lack* of those words. Of course, it does not deal with preceding questions about *which* terms should be selected: this choice could itself be informed by keyword analysis that identifies terms more unique to a portion of the archive. Moreover, it assumes that researchers have linked their outcome of interest to some observable textual forms.

Yet, our approach does not claim that the quantitative elements used in isolation provide ready-made outcomes for every research setting. Indeed, part of the frequency analysis may involve searching for multiple relevant terms. Rather, it responds to the reality that researchers face practical challenges to their time and attention: while close reading provides high degrees of conceptual familiarity, it also incurs high costs in terms of the number of texts that researchers can handle (Quinn et al. 2010). This is where computational and corpus linguistics offer useful bridges between highly automated techniques and traditional archival research by providing "efficient ways of detecting patterns in large amounts of text, while preserving enough direct access to documents themselves that enables researchers to interpret these patterns" (Allen and Blinder 2018, 221).

Second, for researchers looking for evidence of causal effects or mechanisms, our approach can generate observations about language use that likely could not be readily or reliably found through close reading alone. For example, by identifying all instances of certain words in documents (via frequency analysis), and showing how their presence is unique to either a particular subset of texts or a given point in time (via keyword analysis), researchers can be more confident in their claims of having found evidence of causal mechanisms that are unique to an organization (which may correspond to trace evidence) or a time period (as in sequential evidence). This would be a difficult case to make without some sense of the entirety of the archive. Of course, this assumes that the researcher has sufficiently justified *how* a given pattern indicates the causal effect or mechanism in question: we do not claim that the workflow does this theoretical work by virtue of counting words. Rather, by enabling analysis of an entire body of assembled documents, linguistic techniques contextualize a specific usage to highlight how representative or distinctive it is. As a result, they provide ways of generating statements about language use

[6] Choosing a higher N produces keyness scores that favor more common words, while a smaller N favors rarer words. For demonstration purposes, we use $N = 1$ to illustrate keyword analysis as a general step within our approach, rather than advocate for this specific scoring approach.

that go beyond the limited and possibly idiosyncratic views, interests, and abilities of researchers.

In summary, we argue that our approach demonstrates features that mutually benefit computational and archival researchers. On the one hand, simple computational techniques such as establishing terms' frequencies or keyness enable archival researchers to both identify potential subsets of documents for deeper analysis and generate evidence about the distribution of a term across an entire archive. This helps address critiques of archive-based research as lacking rigor or justification for document selection (Freshwater 2003). On the other hand, through close reading and inclusion of non-digitized (or non-digitizable) materials beyond an immediate corpus, archival methods enable computational researchers to acknowledge historical contexts and constraints on descriptive processes as well as causal mechanisms that are important for building and testing theories (Dennison 2021). Besides validating existing computational solutions—an important outcome in itself—interpretive approaches support observing variation and contingency, as well as discovering concepts (Longo and Zacka 2019). These objectives are important means by which scholars of international organizations can unsettle established ways of thinking to reveal new questions for empirical testing.

Refugees in the International Labour Organization, 1919–2015

Next, we demonstrate our approach by applying it to nearly a century of archival documents from the ILO. Multiple historians have examined the ILO's organizational activities (e.g., Van Daele et al. 2010) to reveal how its priorities have expanded beyond initial concerns about social justice and decent work to include themes such as women in development (Boris 2014). More generally, it has been an international platform for policymaking, setting labor standards, and furthering related agendas through its unique tripartite structure comprising governments, employers, and labor unions (Maul 2012). Yet, with few exceptions (Garnier 2014), similar attention has not been given to the ILO's role in international refugee governance despite forced migration constituting a pressing global issue and the organization being involved in global efforts to support refugees' access to employment (ILO 2019). Therefore, we aimed to explore whether and how refugees figured in the organization's high-level discussions: what has been ILO's involvement in refugee assistance historically?

Dataset and Assumptions

ILO official bulletins between 1919 and 2015 comprise two series: series A reports the activities of the ILO as well as official texts and conventions formally adopted by the ILO's International Labour Conference at its annual meeting, while series B reports on the ILO's Committee on Freedom of Association on specific labor disputes in member countries. Since we are interested in identifying how and to what extent refugees have featured in the ILO's broad policies and strategy, we focus on series A. The dataset contains 93 bulletins, comprising 18,108,690 words. Bulletins between 1940 and 1943 are missing, and the series A bulletin for 2009 was not machine-readable.[7] We used the Sketch Engine version 2.36.5 (Kilgarriff et al. 2014), a web-based platform for computational linguistic analysis and lexicography, to organize and study the dataset. Importantly for our demonstration, while the techniques we use are not software-specific, we acknowledge how different tools (and versions of them) can produce different results by virtue of programming and preprocessing decisions that are not always transparent to users (Anthony 2013; Hickman et al. 2022).

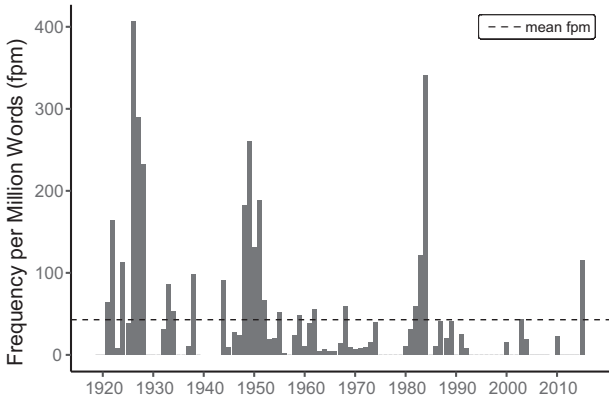

Figure 2. Normalized frequencies of "refugee" or "refugees" in ILO bulletins, 1919–2015.

We make key assumptions about the texts that have implications for the scope of our analysis. First, we assume they are comprehensive summaries of collectively agreed ILO activities. Since these bulletins are aggregate policy statements produced as part of official ILO procedures, they are distinct representations of the organization and its public workings. Second, we assume that these are windows into the ILO bureaucracy at a relatively high level. Unlike reports or minutes from field offices, these bulletins generally focus on strategic aspects of labor issues that have implications for the organization's operations at an international level, including its relationships with other international bodies. Third, we assume that these bulletins serve as an authoritative record of proceedings, achievements, and agendas at given points in time. The value of this authoritativeness cannot be understated, especially in the context of international organizations that rely on different kinds of authority to justify their existence and decisions (Barnett and Finnemore 2004). Even so, these documents provide a necessarily limited view onto the ILO and a particular subset of actors.

The Changing Salience of Refugees in the History of the ILO

Where, if at all, do the terms *refugee* or *refugees* appear in this material? (Throughout, italicized words refer to the words themselves rather than the concept or population.)[8] Answering this question is the first step along pathway A in our model. Figure 2 displays how often these words appeared in each annual bulletin, normalized by dividing the raw frequencies by the number of words in each document and then multiplying this percentage by 1,000,000.

The results clearly show how the salience of *refugee(s)* has changed over the last century. Over the period, these terms appear on average about forty-three times per million words. However, there are three periods that have exceeded this number: 1921–1928 (with some exceptions), 1948–1952, and 1982–1984. Each of these periods had a peak within it: 1926 (407 mentions per million words), 1949 (260 mentions per million words), and 1984 (341 mentions

[7] Although the 2009 bulletin does not appear in our quantitative analysis, we read it as part of the qualitative archival analysis. This illustrates how computational and interpretive techniques augment each other.

[8] Prior to 1951, there was no legal definition of "refugee." Instead, populations based on nationality were classified as refugees. In this study, we followed classification practices as the ILO used them in the eras we researched.

per million words). Drawing upon prior studies of refugee governance (e.g., Pincock, Betts, and Easton-Calabria 2020), we could expand our analysis to consider other related terms: *asylum-seekers* or variations of *humanitarian* for example. Repeating the frequency analysis showed how *asylum-seeker(s)* only appeared ten times in the entire archive. Meanwhile, *humanitarian(ism)* tended to appear mostly in the post-World War II period, although its normalized frequencies never exceeded thirty-five mentions per million words. In a larger project investigating forced migration more broadly, inclusion of these terms would complement a direct investigation of *refugee(s)*.

Even so, by suggesting variation in refugees' visibility on the ILO's high-level agenda, these findings open substantive questions that are well suited for archival research in an interpretive mode. For example, what might account for these shifts, and what did they signify? Next, we draw upon the archival materials—both within and beyond the ILO bulletin dataset—to examine these three moments in greater depth.

1919–1945: Technical "solutions" to refugee "problems." After World War I, states used their still-developing abilities to care for needy populations via humanitarian efforts generally, and international coordination and organization toward refugee issues specifically (League of Nations 1923). The formal commission on the League of Nations (hereafter "the League") was established on January 25, 1919, which included the creation of the ILO. One of the League's main foci, shared by multiple countries, was addressing widespread unemployment among both citizens and migrants (Holborn 1939). This informed broader international debates on social issues and solutions that eventually manifested in the first international Labour Conference in 1919, which aimed to create and adopt mechanisms for overseeing international labor standards (Sauthier 2013).

In 1921, the League created the High Commission for Refugees (hereafter the "High Commission") to address in part the perceived threat that displaced people posed to European stability. By 1925, the ILO had taken over refugee operations from the League to primarily address "the employment, emigration and settlement of refugees" (ILO 1928). Then, between 1925 and 1929, it initiated and maintained a successful "employment-matching scheme" for over 50,000 refugees. This scheme was premised on European countries' need for foreign employment and the availability of skilled refugees for jobs (ILO 1928).

Why was this development significant? Katy Long (2013, 8) observes how during this time "refugee protection was constructed around the twin facets of migration—movement and employment—in the first decades of the international refugee regime." As a result, refugees were considered labor migrants rather than subjects of humanitarian need. Moreover, states had a limited view of their responsibilities toward citizens and refugees, which was premised largely on perceived capabilities to work rather than an inherent right to protection. This was congruent with the nature of assistance offered to refugees, which placed them into employment or helped them become "productive" (League of Nations 1934, 66). The League's doctrine of refugee self-reliance persisted until the start of World War II, although when the 1930s depression hit, the ILO struggled to resettle refugees regardless of their prior training and capacity to work. This is reflected in the steep drop-off of mentions of refugees in figure 2. Despite this, it envisioned itself as becoming the main agency for migration and displacement once the war ended—an ambition that carried over into the following decades.

1945–1951: International competition over the "ownership" of refugee assistance. After World War II, there were millions of migrants throughout Europe. The International Refugee Organisation (IRO) had been formed in 1943 to address "surplus populations" in Europe, but was dissolved after the war. This left a gap among existing international institutions equipped to address migration. As a result, a series of international meetings and conferences throughout the late 1940s and early 1950s occurred, in which the ILO played a major part. During this time, the ILO made several public statements and policy announcements about migration, possibly to assert its ownership over the issue. For example, in 1944 the ILO issued the Declaration of Philadelphia, which included a commitment to focus on "migration for employment and settlement" (ILO 1944, III(c)). Five years later, it announced the 1949 Migration for Employment Convention (No. 97), which included a "Model Agreement on Temporary and Permanent Migration for Employment, including Migration of Refugees and Displaced Persons," demonstrating the ILO's renewed attention to forced migrants after World War II.

This attention culminated in 1951 at the ILO's General Migration Conference (the "Naples Conference"). During the meeting, the ILO proposed to become the main international organization dealing with migration. However, the United States, Australia, and Canada—fearing the loss of control over their national migration policies and lacking faith in ILO's capacity to manage such a monumental task—rejected the proposal (Elie 2010). What is more, the United Nations High Commissioner for Refugees (UNHCR) was also formed in 1951, taking over from the earlier League High Commission. These developments firmly diminished the ILO's role as a major international organization serving refugees.

However, the decreasing salience of refugees for the ILO was not just due to interorganizational conflict. Whereas refugee assistance prior to World War II was concentrated in Europe, the postwar years saw shifts in attention toward countries in the process of decolonization—and the resulting refugees. For example, in the early 1960s, UNHCR began operating in Africa and instituted a policy of providing basic material assistance: food, shelter, and agricultural tools. In contrast to the loans and minimal assistance of the interwar years, this assistance was provided without expectation of repayment, contributing toward a framework for entitlements that is still associated with refugee assistance (Loescher, Betts, and Milner 2008). As a result, refugees were increasingly cast as beneficiaries of aid rather than employable labor—as victims rather than employees (Easton-Calabria 2015)—which may have reduced the attractiveness of ILO involvement in refugee issues.

1979–1984: From competition to cooperation in refugee assistance programs. Although the ILO continued undertaking isolated projects with refugees, it made a significant breakthrough at the end of 1970s and early 1980s. Despite aid-based interventions of UNHCR and other refugee assistance actors in freshly postcolonial settings in Africa—and more funding than ever before—these refugee settlements were largely failing to foster refugee self-reliance. Several factors contributed to the widespread failure of refugees' economic self-reliance: poor soil that impeded agricultural production, policies of encampment that restricted refugees' access to local economies, and the presence of poorly trained authoritarian settlement administrators (Easton-Calabria 2015). Consequently, most encamped refugees remained partially dependent on foreign assistance (Refugee Policy Group 1985).

Table 1. Keywords characterizing the 1980–2015 period in the ILO corpus

Rank	*Keyword*	*Keyness score*	*Frequency (1980–2015)*	*Frequency (1919–1979)*
1	*sustainable*	177.08	668	5
2	*globalization*	171.33	486	0
3	*gender*	123.31	349	0
4	*hiv*	122.26	346	0
5	*dialogue*	105.11	1,339	53
6	*partner*	102.79	848	29
7	*myanmar*	98.78	279	0
8	*issues*	87.26	393	9
9	*global*	86.58	908	41
10	*sectoral*	84.56	618	24
11	*restructuring*	83.09	343	7
12	*recognize*	80.42	723	33
13	*authorize*	72.85	532	24
14	*organize*	68.88	271	6
15	*organization*	65.51	5,575	440

Notes: Keyness scores calculated using normalized (per million) frequencies in the focus (1980–2015) and reference (1919–79) subcorpora, with an *N* parameter of 1. The focus set contains 2,853,334 tokens, while the reference set contains 15,255,360 tokens (as reported by the Sketch Engine).

As the number of refugees worldwide rose from three million in 1977 to over ten million in 1982 (Loescher 2001, 202), UNHCR increasingly focused on promoting refugee self-reliance (UNGA 1985). At the time, UNHCR sought collaboration with several multilateral actors and UN agencies including the ILO, in part following commitments made at the 1981 and 1984 International Conferences on Assistance to Refugees in Africa (ICARA I and II) (Gorman 1986, 287). Recognizing the mounting costs of long-term aid to protracted refugee populations, as well as the need to engage refugees in income generation, UNHCR signed a Memorandum of Understanding (MoU) with the ILO in 1983. Concordance analysis of *refugee(s)* in the 1983 and 1984 bulletins reveals the dynamics around this event. The ILO (1983, n. p.) justified the memorandum by observing that, with respect to assistance to and standards of treatment for refugees, "there are areas of common concern in which the present co-operation between the two organizations should be further strengthened and developed." It went on to stake out its fields of competencies in this new relationship: employment creation, income-generating projects, and vocational training (ILO 1984, 76).

In turn, ILO projects were featured in a 1985 UN General Assembly report on UNHCR's work. One section discusses the scope and scale of cooperation between UNHCR and other members of the UN system:

> Following the signing of the Memorandum of Understanding between the International Labour Organisation (ILO) and UNHCR in 1983, and the secondment of an expert in income-generating activities by ILO to the Specialist Support Unit of UNHCR, co-operation between the two organizations further increased, especially in the fields of small-scale enterprise development and income-generating activities. (UNGA 1985, paragraph 184)

This arrangement had significant impacts, including the creation of several large-scale refugee assistance programs in Pakistan, Sudan, and Somalia (ILO 1991). Reports making the case for greater employment among refugees (e.g., ILO 1984) accompanied shifts in UNHCR's own approaches toward refugees from care-and-maintenance to self-reliance, with ILO repeatedly invoked in discussions of refugee self-reliance, livelihoods, and income generation.[9]

Using Keywords to Identify Distinctive Language during 1980–2015

Up to this point, we have focused on documenting shifts in the salience of refugees within the ILO's high-level agenda. Following pathway A in our workflow, frequency analysis provided windows onto specific moments when refugees were especially visible, enabling us to strategically apply more intensive archival methods. But what if we were interested in identifying distinctive terminologies in certain parts of the dataset?

This exploratory question, following pathway B, requires us to turn to keyword analysis. Drawing upon historical studies into the periodicity of refugee livelihoods over the twentieth century (Easton-Calabria 2022), we created three subcorpora containing bulletins from 1919–1939, 1944–1979, and 1980–2015.[10] Then, we conducted keyword analysis using 1980–2015 as the focus subcorpus and 1919–1979 as the reference subcorpus. Table 1 shows the twenty-five most frequent keywords appearing in the 1980–2015 subcorpus, sorted by the keyness score detailed earlier, and excluding names and titles.[11]

While these words are, by design, more distinctive of the focus subcorpus, some nouns are unique to this period (*globalization*, *gender*, *HIV*). Depending on the research design and theoretical mechanisms in a project, these patterns could function as trace evidence of how certain issues have risen on policy agendas—an observation that would be difficult to make over this period using close reading alone. Furthermore, other keywords might inspire further qualitative investigation into links comprising causal mechanisms, depending on the research questions at hand. For example, some keywords (*dialogue*, *partner*, *restructuring*) might speak to how the ILO has sought to achieve its goals. Meanwhile,

[9] ILO projects regarding refugees and self-sufficiency were discussed in the UN General Assembly report, which is notable as an annual update of UNHCR's work.

[10] For demonstration purposes, we justified these cutoffs based on readings of refugee governance over the twentieth century.

[11] Although proper nouns can signal important aspects, such as actors' presence, we excluded them since we were primarily interested in shifts in language use.

examining the highest ranked keyword (*sustainable*) could reveal how and in which contexts it has been applied to relatively recent ILO activities. Attending to these kinds of patterns potentially reveals insights into the self-presentation and operations of international organizations (Moretti and Pestre 2015).

Discussion

Social scientists who examine aspects of international organizations often rely on archives and archival materials. Some have interpretive goals in mind, while others seek to find evidence of causal effects and mechanisms (Cirone and Spirling 2021). Facing larger quantities of digitized (and digitizable) texts, researchers clearly need techniques that are "transparent, tweakable, and sensitive to language-specific differences" (Baden et al. 2022, 15), which enable locating potentially relevant areas of archives or documents and placing those usages in context. Our mixed-method workflow draws upon linguistic techniques to guide attention toward these areas as it identifies patterns that can inform and direct subsequent investigation. Moreover, it is general and flexible enough to accommodate a range of questions that social scientists bring to archives. We demonstrated this by applying it to a dataset of ILO documents that illustrates the ILO as an important yet under-recognized actor in refugee assistance. Empirically, we uncovered three key moments during which refugees have been salient for the ILO, as well as identified keywords that are distinctly associated with the ILO's more recent history. These findings open further research addressing both interpretive and causal questions.

First, shifts in global demand for labor prominently feature in each of the moments we studied. For example, while the ILO was able to resettle thousands of refugees as workers in the 1920s, this was not the case after the 1929 depression, whether due to rising European restrictionism or broader unemployment. Meanwhile, the ILO's reduced role in refugee resettlement following World War II was partly due to interorganizational struggles over issue ownership as well as declining numbers of refugees to be resettled as "workers" in the first place. Finally, the ILO's resurgence on the refugee issue in the 1980s occurred in the context of the 1981 and 1984 global oil crises, when countries such as the United States experienced a recession and may have sought to reduce donor obligations. One possibility, therefore, is that the ILO was engaged in part to foster long-term refugee populations' self-reliance to reduce aid budgets, a topic explicitly discussed at the time by UNHCR (UNGA 1985, paragraph 115).

Second, organizational strategies might have contributed to the shifting construction of refugees. It is well documented how international organizations express priorities in sometimes subtle ways, including by redefining their missions and objectives (Barnett and Finnemore 2004). The kinds of refugee assistance offered during the interwar years effectively treated refugees as labor migrants, providing the ILO with a natural role in delivering these initiatives. However, after World War II and into the 1950s, refugees were increasingly viewed as victims in need of humanitarian protection and assistance rather than employment. Although the ILO remained involved in isolated refugee projects in regions in sub-Saharan Africa in the 1960s and 1970s, it appears that the ILO became more involved in refugee assistance again when self-reliance became a dominant theme in the sector.

Taken together, these findings suggest some dynamics linking the ILO with refugee assistance, which warrant further investigation: other international organizations considering refugees as workers, alignment with the ILO's own views, and greater interagency collaboration on issues including livelihoods. If this is the case, the ILO's recent re-involvement with refugees makes sense, as it comes at a moment when refugee self-reliance is considered to be paramount (Aleinikoff 2015) and international poverty reduction initiatives such as the Sustainable Development Goals aim to include everyone, including refugees, in promoting decent work and other human rights. While more research and theorization need to formally spell out the mechanisms at play, it is clear that the ILO's engagement with refugees appears to be recurrent rather than novel. Moreover, the apparent increase in relative mentions of *refugees* within the 2015 bulletin, seen in figure 1, merits further investigation: perhaps the so-called 2015 European refugee "crisis" marked the beginning of another significant moment for the ILO in refugee assistance (Allen et al. 2018).

While empirically tracing and testing these possibilities is beyond the scope of this note, acknowledging their presence highlights how our workflow can enhance and augment the sensibilities of archival researchers. It contributes a way of linking established tools together to improve the rigor and significance of international studies research that uses archives. As more documents become available and accessible, researchers face practical and theoretical challenges in making sense of these materials. In response, our intention has been to provide a roadmap for scholars doing historical and archive-based research.

Supplementary Information

Supplementary information is available at the *International Studies Quarterly* data archive.

References

Aleinikoff, T. Alexander. 2015. "From Dependence to Self-Reliance: Changing the Paradigms in Protracted Refugee Situations." Migration Policy Institute, Washington, DC.

Allen, William, Bridget Anderson, Nicholas Van Hear, Madeleine Sumption, Lena Rose, Jennifer Hough, Rachel Humphris, and Sarah Walker. 2018. "Who Counts in Crises? The New Geopolitics of International Migration and Refugee Governance." *Geopolitics* 23 (1): 217–43.

Allen, William, and Scott Blinder. 2018. "Media Independence through Routine Press–State Relations: Immigration and Government Statistics in the British Press." *The International Journal of Press/Politics* 23 (2): 202–26.

Anderson, Margo. 2007. "Quantitative History." In *The SAGE Handbook of Social Science Methodology*, edited by William Outhwaite and Stephen P. Turner, 246–64. London: SAGE.

Anthony, Laurence. 2013. "A Critical Look at Software Tools in Corpus Linguistics." *Linguistic Research* 30 (2): 141–61.

Baden, Christian, Christian Pipal, Martijn Schoonvelde, and Mariken A.C.G. van der Velden. 2022. "Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda." *Communication Methods and Measures* 16 (1): 1–18.

Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. London: Bloomsbury Academic.

Baker, Paul, and Jesse Egbert. 2016. *Triangulating Methodological Approaches in Corpus Linguistic Research*. New York: Routledge.

Barberá, Pablo, Amber E. Boydstun, Suzanna Linn, Ryan McMahon, and Jonathan Nagler. 2021. "Automated Text Classification of News Articles: A Practical Guide." *Political Analysis* 29 (1): 19–42.

Barnett, Michael, and Martha Finnemore. 2004. *Rules for the World: International Organizations in Global Politics*. Ithaca, NY: Cornell University Press.

Beach, Derek. 2016. "It's All about Mechanisms: What Process-Tracing Case Studies Should Be Tracing." *New Political Economy* 21 (5): 463–72.

Bevir, Mark, and R.A.W. Rhodes. 2012. "Interpretivism and the Analysis of Traditions and Practices." *Critical Policy Studies* 6 (2): 201–208.

Biber, Douglas, and Randi Reppen, eds. 2015. *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press.

Boris, Eileen. 2014. "Mothers, Household Managers, and Productive Workers: The International Labor Organization and Women in Development." *Global Social Policy* 14 (2): 189–208.

Bradley, Megan. 2021. "Realising the Right of Return: Refugees' Roles in Localising Norms and Socialising UNHCR." *Geopolitics* November 1–28. https://doi.org/10.1080/14650045.2021.1994399.

Brookes, Gavin, and Tony McEnery. 2018. "The Utility of Topic Modelling for Discourse Studies: A Critical Evaluation." *Discourse Studies* 21 (1): 3–21.

Cirone, Alexandra, and Arthur Spirling. 2021. "Turning History into Data: Data Collection, Measurement, and Inference in HPE." *Journal of Historical Political Economy* 1 (1): 127–54.

Dennison, Tracy. 2021. "Context Is Everything: The Problem of History in Quantitative Social Science." *Journal of Historical Political Economy* 1 (1): 105–26.

Easton-Calabria, Evan. 2015. "From Bottom-Up to Top-Down: The 'Pre-History' of Refugee Livelihoods Assistance from 1919 to 1979." *Journal of Refugee Studies* 28 (3): 412–36.

———. 2022. *Refugees, Self-Reliance, Development: A Critical History*. Bristol: Bristol University Press.

Elie, Jérôme. 2010. "The Historical Roots of Cooperation between the UN High Commissioner for Refugees and the International Organization for Migration." *Global Governance* 16 (3): 345–60.

Freshwater, Helen. 2003. "The Allure of the Archive." *Poetics Today* 24 (4): 729–58.

Gabrielatos, Costas. 2018. "Keyness Analysis: Nature, Metrics and Techniques." In *Corpus Approaches to Discourse: A Critical Review*, edited by Charlotte Taylor and Anna Marchi, 225–58. London: Routledge.

Gabrielatos, Costas, Tony McEnery, Peter J. Diggle, and Paul Baker. 2012. "The Peaks and Troughs of Corpus-Based Contextual Analysis." *International Journal of Corpus Linguistics* 17 (2): 151–75.

Garnier, Adèle. 2014. "Arrested Development? UNHCR, ILO, and the Refugees' Right to Work." *Refuge: Canada's Journal on Refugees* 30 (2): 15–25.

Gonzalez-Ocantos, Ezequiel, and Jody LaPorte. 2021. "Process Tracing and the Problem of Missing Data." *Sociological Methods & Research* 50 (3): 1407–35.

Gorman, Robert F. 1986. "Beyond ICARA II: Implementing Refugee-Related Development Assistance." *International Migration Review* 20 (2): 283–98.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–97.

Hegghammer, Thomas. 2021. "OCR with Tesseract, Amazon Textract, and Google Document AI: A Benchmarking Experiment." *Journal of Computational Social Science* 5: 861–82.

Hickman, Louis, Stuti Thapa, Louis Tay, Mengyang Cao, and Padmini Srinivasan. 2022. "Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations." *Organizational Research Methods* 25 (1): 114–46.

Holborn, Louise W. 1939. "The League of Nations and the Refugee Problem." *The Annals of the American Academy of Political and Social Science* 203 (1): 124–35.

ILO. 1928. "Refugee Problems and Their Solution." *International Labour Review* 1768–85.

———. 1944. "Official Bulletin." International Labour Organization, Geneva, Switzerland.

———. 1983. "Memorandum of Understanding Between the Director-General of the International Labour Organization and the United Nations High Commissioner for Refugees." Official Bulletin, International Labour Organization, Geneva, Switzerland.

———. 1984. "Tradition and Dynamism among Afghan Refugees: A Report on Income-Generating Activities for Afghan Refugees in Pakistan." International Labour Organization, Geneva, Switzerland.

———. 1991. "Assessment of Selected ILO Projects Concerning Migrants, Refugees and Displaced Persons." GB.251/Op/2/1. International Labour Organization, Geneva. Accessed January 7, 2022. http://www.ilo.org/public/english/standards/relm/gb/docs/gb251/op-21.htm#Heading6.

———. 2019. "Crisis Migration." Accessed July 15, 2022. https://www.ilo.org/global/topics/labour-migration/policy-areas/crisis/lang—en/index.htm.

Kilgarriff, Adam. 2009. "Simple Maths for Keywords." Proceedings of the Corpus Linguistics Conference, Liverpool. Accessed February 6, 2020. https://www.sketchengine.eu/wp-content/uploads/2015/04/2009-Simple-maths-for-keywords.pdf.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. "The Sketch Engine: Ten Years On." *Lexicography* 1 (1): 7–36.

Killick, Rebecca, and Idris Eckley. 2014. "Changepoint: An R Package for Changepoint Analysis." *Journal of Statistical Software* 58 (3): 1–19.

Kipping, Matthias, R. Daniel Wadhwani, and Marcelo Bucheli. 2014. "Analyzing and Interpreting Historical Sources: A Basic Methodology." In *Organizations in Time: History, Theory, Methods*, edited by Marcelo Bucheli and R. Daniel Wadhwani, 305–29. Oxford: Oxford University Press.

Klein, Lauren F., Jacob Eisenstein, and Iris Sun. 2015. "Exploratory Thematic Analysis for Digitized Archival Collections." *Digital Scholarship in the Humanities* 30 (suppl_1): i130–41.

Kott, Sandrine, and Joëlle Droux. 2013. *Globalizing Social Rights: The International Labour Organization and Beyond*. Basingstoke/New York: Palgrave Macmillan.

League of Nations. 1923. "Near East Refugees, Western Thrace Refugee Settlement." *Report by Dr. Nansen, High Commissioner for Refugees. To League of Nations, Communicated to the Council*. April 22, 1923, Geneva, Switzerland.

———. 1934. "Human Welfare and the League." O.LNU/1934(13). League of Nations Union.

Lee, Tak Yeon, Alison Smith, Kevin Seppi, Niklas Elmqvist, Jordan Boyd-Graber, and Leah Findlater. 2017. "The Human Touch: How Non-Expert Users Perceive, Interpret, and Fix Topic Models." *International Journal of Human-Computer Studies* 105: 28–42.

Loescher, Gil. 2001. *The UNCHR and World Politics: A Perilous Path*. Oxford: Oxford University Press.

Loescher, Gil, Alexander Betts, and James Milner. 2008. *The United Nations High Commissioner for Refugees (UNHCR): The Politics and Practice of Refugee Protection into the Twenty-First Century*. London: Routledge.

Long, Katy. 2013. "When Refugees Stopped Being Migrants: Movement, Labour and Humanitarian Protection." *Migration Studies* 1 (1): 4–26.

Longo, Matthew, and Bernardo Zacka. 2019. "Political Theory in an Ethnographic Key." *American Political Science Review* 113 (4): 1066–70.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* 23 (2): 254–77.

Maul, Daniel. 2012. *Human Rights, Development and Decolonization: The International Labour Organization, 1940–70*. Basingstoke: Palgrave Macmillan.

McEnery, Tony, and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

McGarr, Kathryn J. 2020. "The Importance of Historical Perspective and Archival Methods in Political Communication Research." *Political Communication* 37 (1): 110–16.

Moretti, Franco, and Dominique Pestre. 2015. "Bankspeak: The Language of World Bank Reports." *New Left Review* 92 (2): 75–99.

Murakami, Akira, Paul Thompson, Susan Hunston, and Dominik Vajn. 2017. "'What Is This Corpus About?' Using Topic Modelling to Explore a Specialised Corpus." *Corpora* 12 (2): 243–77.

Ongenaert, David, and Stijn Joye. 2019. "Selling Displaced People? A Multi-Method Study of the Public Communication Strategies of International Refugee Organisations." *Disasters* 43 (3): 478–508.

Partington, Alan, Alison Duguid, and Charlotte Taylor. 2013. *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*, Vol. 55. Amsterdam: John Benjamins Publishing.

Pincock, Kate, Alexander Betts, and Evan Easton-Calabria. 2020. *The Global Governed? Refugees as Providers of Protection and Assistance*. Cambridge: Cambridge University Press.

Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54 (1): 209–28.

Refugee Policy Group. 1985. "Older Refugee Settlements in Africa: Final Report." Refugee Policy Group, Washington, DC.

Rohlfing, Ingo. 2012. *Case Studies and Causal Inference: An Integrative Framework*. Basingstoke: Palgrave Macmillan.

Sauthier, Ingrid Liebeskind. 2013. "Modern Unemployment: From the Creation of the Concept to the International Labour Office's First Standards." In *Globalizing Social Rights: The International Labour Organization and Beyond*, edited by Sandrine Kott and Joëlle Droux, 67–84. New York: Springer.

Stoler, Ann Laura. 2002. "Colonial Archives and the Arts of Governance." *Archival Science* 2 (1): 87–109.

Taylor, Charlotte. 2014. "Investigating the Representation of Migrants in the UK and Italian Press: A Cross-Linguistic Corpus-Assisted Discourse Analysis." *International Journal of Corpus Linguistics* 19 (3): 368–400.

Taylor, Charlotte, and Anna Marchi, eds. 2018. *Corpus Approaches to Discourse: A Critical Review*. New York: Routledge.

Tirabassi, Katherine E. 2010. "Journeying into the Archives: Exploring the Pragmatics of Archival Research." In *Working in the Archives: Practical Research Methods for Rhetoric and Composition*, edited by Alexis E. Ramsey, Wendy B. Sharer, Barbara L'Eplattenier and Lisa S. Mastrangelo, 169–80. Carbondale, IL: Southern Illinois University Press.

Trubowitz, Peter, and Kohei Watanabe. 2021. "The Geopolitical Threat Index: A Text-Based Computational Approach to Identifying Foreign Threats." *International Studies Quarterly* 65 (3): 852–65.

UNGA. 1985. "Report of the United Nations High Commissioner for Refugees." Supplement 12 (A/40/12). United Nations General Assembly Official Records, Fortieth Session.

van Atteveldt, Wouter, Mariken A.C.G. van der Velden, and Mark Boukes. 2021. "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms." *Communication Methods and Measures* 15 (2): 121–40.

Van Daele, Jasmien, Magaly Rodríguez, Geert van Goethem, and Marcel van der Linden, eds. 2010. *Essays on the International Labour Organization and Its Impact on the World during the Twentieth Century*. Bern: Peter Lang.

Veeramoothoo, Saveena (Chakrika). 2022. "Social Justice and the Portrayal of Migrants in International Organization for Migration's World Migration Reports." *Journal of Technical Writing and Communication* 52 (1): 57–93.

Ventresca, Marc J., and John W. Mohr. 2017. "Archival Research Methods." In *The Blackwell Companion to Organizations*, edited by Joel A.C. Baum, 805–28. Wiley.

Wodak, Ruth. 2001. "The Discourse-Historical Approach." In *Methods of Critical Discourse Analysis*, edited by Ruth Wodak and Michael Meyer, 63–94. London: SAGE.