Mining for Meaning: using computational text analysis for social inquiry

Miriam Hurtado Bodell
Mining for Meaning:

using computational text analysis for social inquiry

Miriam Hurtado Bodell
At the Faculty of Arts and Sciences at Linköping University, research and doctoral studies are carried out within broad problem areas. Research is organized in interdisciplinary research environments and doctoral studies mainly in graduate schools. Jointly, they publish the series Linköping Studies in Arts and Sciences. This thesis comes from the Institute for Analytical Sociology at the Department of Management and Engineering.

Distributed by:
Department of Management and Engineering
Linköping University
SE-581 83 Linköping, Sweden

Miriam Hurtado Bodell
Mining for Meaning:
using computational text analysis for social inquiry

Edition 1:1
ISBN 978-91-8075-618-1 (PDF)
ISSN 0282-9800

Cover art: Official migration file of author's grandmother from The National Archives of Sweden. Private photo.

© Miriam Hurtado Bodell, 2024
Printed by LiU-Tryck, Linköping 2024

This work is licensed under a Creative Commons “Attribution 4.0 International” license.
ABSTRACT

People interpret their surroundings through associations, determining what they perceive as belonging or not belonging together. For instance, one individual may view immigrants as a beneficial addition to the domestic labor market, while another may perceive them as a threat to job opportunities for native citizens. Despite differing viewpoints on immigration, these individuals share a similar economic interpretation of immigration as a concept. Explaining how these interpretations develop and evolve is a fundamental and open question related to the social world.

For a long time, people's interpretations of the world have been hidden away in their minds, and researchers have primarily relied on surveys to try to measure them. However, individuals and groups leave behind traces of their understandings of the world in their communication and written expressions. Consequently, textual data hold immense potential for sociological research. This thesis pursues three primary objectives. First, to discuss the use of text data for social inquiry. Second, to introduce and explore intrinsically interpretable text models for sociological inquiry. Third, to explore rigorous ways of studying meaning and meaning-making in the Swedish immigration discourse using computational text analysis. The introductory chapter and four research articles presented in this thesis all speak to at least one of these aims.

Essay I addresses the question of how researchers can assess the data quality of a corpus to determine its suitability for addressing research questions. Drawing inspiration from survey research, this essay presents a general approach to evaluating the scientific value of a given text dataset. The framework outlined in this essay delineates potential errors that could affect the reliability and validity of any measures derived from a corpus, and offers methods for quantifying some of them.

Essay II presents a novel extension to standard word embedding models. Our extension gives researchers the ability to study how the meaning of words relates to pre-specified binary dimensions. We find that our proposed intrinsically interpretable model outperforms current standard approaches on classification tasks related to sentiment and gender. The methodology presented in Essay II will thus help sociologists to measure and test theories pertaining to binary concepts.

Essay III contributes to the ongoing discussions in sociology regarding the identification of more formal ways to measure aggregate-level meanings. This essay traces prevailing frames of immigration in Swedish national news media from the end of World War II until 2019, providing an unprecedented macro-level perspective on immigration frames. The analysis indicates that the framing of immigration in the Swedish media changes following periods of rupture rather than single events.

Essay IV delves into the mechanisms that influence changes in online discussions on Flashback following Jihadist terrorist attacks. We examine two mechanisms: changes in discussion content (within-individual change) and changes in the composition of discussion participants (compositional change). Our findings reveal that interpretations of immigration related to culture
and security become more prominent following terror attacks, and that both of the mechanisms examined play a role in shaping post-attack discussions.
Acknowledgments

Being a Ph.D. candidate during a global pandemic was sometimes lonely, but I’m very grateful to those who made this journey more enjoyable and provided invaluable support.

Firstly, I would like to express my gratitude to my supervisors. Marc Keuschnigg, your reassurance and belief in my abilities, especially during times of doubt, have been immensely important. Your advice to remain optimistic has been instrumental in my navigation of this process. I appreciate your empathy when life intervened during the writing of my dissertation, and your reminder that there are more important things in life than work.

Måns Magnusson, you were the first to recognize my potential as a researcher. Thank you for your unwavering belief in me from the outset of my Ph.D. and for your guidance throughout. As I wrote in the acknowledgments in my master thesis, you have been a consistent creative sounding board during this journey.

Peter Hedström, thank you for accepting me into the Ph.D. program despite my initial lack of understanding of analytical sociology. Your support and encouragement to use text-as-data have been invaluable. Thank you for always taking the time to read my work.

The Institute for Analytical Sociology is an inspiring and creative environment thanks to its exceptional researchers. Maria Brandén, as director of IAS your dedication to providing both professional and personal guidance in navigating academia is deeply appreciated. Sarah Valdez, your insightful questions during seminars have been truly inspirational. Etienne Ollion, your early recognition of my potential as a researcher is deeply appreciated, and I’m very grateful for your support. Jacob Habinek, your breadth of knowledge across sociological subfields, insightful comments, and generosity with your time have made you a great role model for all Ph.D. candidates at the IAS. Eduardo Tapia, your camaraderie and guidance in sociological inquiry have been much appreciated, the IAS has been better since you re-joined us. Karoly
Takac, your generous emphasis on the efforts of junior scholars is inspiring. Erik Rosenquist, I thank you for recognizing my expertise, and for reading my work even though our research does not share much in common. To Benjamin Jarvis, Carl Nordlund, Adel Daoud, Christian Stieglich, Abiel Sebhatu, Connor Jerzak, Karl Wennberg, Pontus Strimling, thank you for your questions, comments, and support during my work with this dissertation.

In addition to the faculty at the IAS, I have had the privilege of sharing my Ph.D. journey with a remarkable group of Ph.D. students, many of whom have already graduated.

Selcan Mutgan, seeing you navigate academia with incredible intellectual vigor has been truly inspiring. Our shared experiences in the US, and seeing you take on a new academic setting so fearlessly have left a lasting impression, and I am very grateful for the laughter we’ve shared.

Laura Fürsich, witnessing your growth as a researcher and seeing the acknowledgement you have received for your competence has been wonderful. Your thoughtful feedback on my research has been invaluable in shaping my work and has shown me how kindness plays an important role in forming great scholars.

Anastasia Menshikova, your curiosity and talent are evident, and I’m honored to have collaborated with you. I look forward to witnessing your future successes and am grateful for your unwavering support.

Àlex Giménez de la Prada, your depth of knowledge in sociology both intimidated and motivated me. Our lunch conversations were always enlightening. Niclas Lovsjö, thank you for your ability to keep things fun. Martin Arvidsson, your brilliance and perfectionism have challenged me in fruitful ways, and it has been a pleasure to work with you. Emanuel Wittberg, your encouragement has been a source of strength during challenging times. Elis Carlberg Larsson, our time together in the ARAB archives and our gossip sessions made difficult moments more bearable. Rodrigo Martínez Pena, José Luis Estévez Navarro, Hendrik Erz, Jesper Lindmaker, Maël Lecoursonnais, Cheng Lin, Flóra Samu, Erik Liss, and Alexandra Rottenkolber, thank you for your camaraderie, critical questions, and supportive comments during our Ph.D. meetings. Your presence has enriched my journey immeasurably.

The IAS owes much to Åsa Arnoldsson and Madelene Töpfner; without their crucial guidance in these final stages of my journey this dissertation would not have been completed, and my lunches would have been much less enjoyable without their presence.
I extend my gratitude to Simon Hengchen and Friedolin Merhout for their invaluable insights as external opponents during my seminars, which have significantly influenced this thesis.

To all members of the Mining for Meaning project, thank you for your feedback and support on the preliminary versions of the papers included in this thesis. Special thanks to Sophie Mützel for her collaboration on one of the papers; your insightful comments and encouragement were key in getting through a round of harsh reviews.

Without the assistance and guidance of the people working at the data lab at the National Library of Sweden, much of the empirical work of this dissertation would not have been possible.

I extend my gratitude to Amir Goldberg for welcoming me into his research group and fostering an environment of theoretical and analytical clarity. Thanks to all members of the Computational Cultural Lab for their engagement in my research and their invaluable feedback.

To my writing group, including Elida Ibrahim, Laura Fürsich, Elis Carlberg Larsson, and Serena Coppolino Perfumi, your feedback was essential during important stages of my work.

Life beyond academia has been crucial in sustaining me through this thesis journey. My friends have been a constant source of support and joy. Erika, your reminder to celebrate before things happen—it might be the only chance you get—has been invaluable. Thank you for your endless energy and optimism even during your own struggles. Lina, thank you for always coming to have lunch and coffee with me at the office when I couldn’t find the time to go anywhere else. Christina, you inspire me to choose to be happy instead of following the path-already-set. Malin, Wikan, Massoud, Ida, Henrik, Malin, Charlotte, Johan, thank you for allowing me to join late during dinner parties, and for reminding me that research is my work and not who I am. Oscar and Michael, thank you for going with me to Gotland, Malaga, and Ockelbo; these trips gave much-needed energy.

Finally, I would like to express my gratitude to my family, whose unwavering support has been instrumental to the completion of this dissertation. My sisters, Daniela Hurtado Bodell and Sofia Hurtado Bodell, your attempts to show interest in my research are deeply appreciated, and your reminder that it is only work even more appreciated. My mother, Christine-Charlotte Bodell—you have been an amazing support in all ways, always. To my grandfather, Olle Bodell, although he didn’t witness the completion of this journey, his pride in me endures as a guiding light. I also would like to thank my extended family: Felix, Axel, Helena, Janne, and Mette.
My deepest appreciation goes to my husband, Egil Martinsson, for his unwavering love and support throughout these years. Thank you for taking the time to read my work and for your thought-provoking questions. You never cease to amaze me, you enrich my life in countless ways, and I am eternally grateful to have you in my life.

Finally, I want to thank my son, Kåre Martinsson Bodell. Choosing to bring you into this world instead of completing my dissertation last year was undoubtedly the best decision I ever made. You have anchored me in ways I cannot fully articulate. While my academic pursuits may have been momentarily set aside, the love you have brought into our lives is immeasurable. Mamma älskar dig.

Miriam Hurtado Bodell
Stockholm, Sweden
March 2024

I also gratefully acknowledge the funding from the Swedish Research Council (2018–05170) which supported my time as a PhD student, as well as travel grants from Vitterhetsakademien. The computations and data storage were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2018-05973 and no. 2022-06725.
List of Essays

This thesis is based on the following essays, which are referred to in the text by their Roman numerals.


## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>viii</td>
</tr>
<tr>
<td>List of Essays</td>
<td>ix</td>
</tr>
<tr>
<td>Contents</td>
<td>xi</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td><strong>Text as data in sociology</strong></td>
<td>9</td>
</tr>
<tr>
<td>Computational text analysis</td>
<td>10</td>
</tr>
<tr>
<td>Common issues faced by sociologists using text as data</td>
<td>13</td>
</tr>
<tr>
<td><strong>Sociological approaches to meaning</strong></td>
<td>19</td>
</tr>
<tr>
<td>The meaning of meaning</td>
<td>19</td>
</tr>
<tr>
<td>Meaning and multi-level explanations of social outcomes</td>
<td>22</td>
</tr>
<tr>
<td><strong>Measuring meaning</strong></td>
<td>27</td>
</tr>
<tr>
<td>Measuring public culture</td>
<td>28</td>
</tr>
<tr>
<td>Measuring private culture</td>
<td>30</td>
</tr>
<tr>
<td><strong>The Swedish immigration discourse</strong></td>
<td>33</td>
</tr>
<tr>
<td>The history of immigration to Sweden</td>
<td>33</td>
</tr>
<tr>
<td>Actors in the Swedish immigration discourse</td>
<td>37</td>
</tr>
<tr>
<td>Data and ethical considerations</td>
<td>40</td>
</tr>
<tr>
<td><strong>Appended Essays</strong></td>
<td>47</td>
</tr>
<tr>
<td>Essay I: From Documents to Data: a Framework for Total Corpus Quality</td>
<td>47</td>
</tr>
</tbody>
</table>
Language enables individuals and groups of people to communicate with each other, and the earliest records of written language date back to 3200 BC (Charpin 2010). Ever since, people worldwide have written down the most mundane scribbles about everyday life and have documented life-altering events. A number of intersecting social forces have now catapulted large corpora, i.e. collections of texts, into becoming essential sensors of the social world, recording the conversations between actors in different societal spheres (van Loon 2022; Evans and Aceves 2016; Bail 2014). First, as a result of increasing literacy rates (Our World in Data 2018) and advances in access to communication technologies in different social contexts (Evans and Aceves 2016), texts are being produced with unprecedented speed by a substantial and progressively representative segment of populations around the world (van Loon 2022). This inclusive turn in the range of groups expressing themselves in text stands in stark contrast to historical records, which were mainly written by the wealthy, the educated, and the victors of war (Harris 2002; Mohr et al. 2020; Bode 2020). Second, more and more aspects of social life are playing out in ways that leave behind traces of text (Edelmann et al. 2020; Evans and Aceves 2016; Grimmer and Stewart 2013; van Loon 2022)—for example, direct messages between friends and colleagues in chats and emails (Goldberg et al. 2016), dating apps and websites (Bruch and Newman 2018), and discussions between opponents and allies on social media (Bail 2021). Third, archives and libraries across the globe are undertaking large-scale projects to digitize their collections (Bingham 2010; Pandey and Kumar 2020), breathing
new life into the voices of the past and making them available to researchers around the world.

The use of text to capture how individuals or societies interpret their surroundings has a long tradition within the social sciences (Lasswell and Leites 1949), most importantly perhaps in the form of content analysis. Early forms of content analyses relied on counting explicit features of the text, i.e., the use of specific "symbols" such as actors or meaningful categories, to test hypotheses about language use and to answer questions about what is being said by whom (Franzosi 2008; Carley 1993; Franzosi 1994). However, as text data become increasingly available, new tools for harvesting the information they contain are rapidly being developed. Interdisciplinary work combining knowledge from computer science, statistics, mathematics, and linguistics has produced important advances in natural language processing (NLP) and machine learning (ML) centered around learning, understanding, and producing human languages (Hirschberg and Manning 2015). The increasing availability of texts and the rapid methodological advances for computationally analyzing their content at scale are yielding unprecedented opportunities for sociologists to "answer old questions in new ways" (Bearman 2015; Lazer et al. 2020).

Against this background, the research in this thesis has three main objectives: (1) to delve into the potential data-related challenges that sociologists face when analyzing increasing volumes of text data, (2) to introduce and explore the potential of intrinsically interpretable computational methods for sociological research, and (3) to empirically explore the use of computational text analysis as a means of producing insights into meaning and meaning-making. Whenever possible, I use the Swedish immigration discourse 1 as a case study.

In relation to all three objectives, I focus in particular on how researchers within analytical sociology can use text data and the methods being developed to analyze large-scale repositories for empirical studies of meaning. Analytical sociology has three core characteristics that distinguish it from other areas of quantitative sociology: the production of multi-level mechanistic explanations 2.

---

1The term "discourse" has been criticized as problematic due to its many uses across disciplines (Baker 2006). Linguists typically use the term discourse to refer to spoken or written language use, but the term can also include non-verbal communication such as body language or photography (Fairclough 1993). In this thesis, I follow the sociological tradition in which "discourse" is predominantly viewed as a social practice (Van Leeuwen and Wodak 1999; Wodak 2009). This perspective highlights the interdependency between individual discursive events (e.g., speech and texts) and the social environments, structures, and institutions in which they are embedded. Discursive acts then play a role in producing, constructing, restoring, reproducing, and transforming social environments, while simultaneously being affected by the context in which they occur.
via macro-micro-macro links; a focus on agents rather than factors; and empirical realism. First, in analytical sociology, macro-level phenomena can only be explained when the mechanism(s) producing them are well-defined (Hedström 2005; Hedström and Ylikoski 2014). A social mechanism is defined as consisting of “entities (or parts) whose activities and interactions are organized to be responsible for the phenomena” (Glennan and Illari 2017). Most importantly, mechanistic explanations should go beyond describing the social world in order to reduce the risk of spurious patterns being mistaken for causality while at the same time increasing the precision and clarity with which the social world is explained (Hedström and Swedberg 1996). Second, *individuals* are the entities whose (inter)actions are ideally used to explain social phenomena (Coleman 1986). As such, analytical sociology emphasizes that all macro-level social phenomena, e.g., political discourse, polarization, or segregation, are produced by the actions of individuals (Udehn 2002). Analytical sociology’s third and final defining characteristic is an unwavering focus on realism. This characteristic further distinguishes the discipline from other quantitative social sciences and instructs researchers to reject the practice of making simplistic assumptions in order to obtain tractable models (Jarvis et al. 2021). It also means that empirical research should consider all known complexities about agents’ behaviors and social systems.

However, the lack of large-scale data covering large parts of the populations of interest has made it challenging to close the gap between theoretical and empirical work in analytical sociology. Much empirical work has relied on small-scale survey data, which provide only snapshots of what a population is doing or believing at specific points in time. Moreover, many such small-scale datasets lack information about social networks and interactions. Therefore, finding large-scale and representative data sources for use in empirical research is essential for living up to the high ambitions of producing and empirically testing mechanistic explanations and modelling the emergence of macro-level phenomena from individual-level behaviors. I argue that text data hold exceptional potential for sociological research programs, such as analytical sociology, that seek to provide insightful population-level descriptions and explain the micro-level dynamics that bring them about.

Importantly, sociologists can analyze text both at the aggregate and individual levels. Collections of texts written by meso-level actors, such as political parties or the media, are one example of aggregate-level text data. Researchers can

---

2The exact definition of a social mechanism is still a matter of debate; see, for example, Table 1.1 in Hedström and Bearman (2009) or Table 1 in Hedström and Ylikoski (2010) for some competing definitions.
also create such corpora by combining texts written by individuals to represent collectives, for example, by analyzing the posts on a social media platform in the aggregate. On the other hand, I use the term individual-level text data to describe collections of texts produced by an individual, such as posts written on social media, personal letters, or diary entries, where the unit of analysis remains the individual actor. In aggregate-level analyses, sociologists can use texts as a social sensor of a given cultural context that provides measures of how concepts, objects, or events are talked about and how they are collectively understood. Aggregate-level text analysis provides insightful and thick descriptions of a current cultural context, and allows for the study of, for example, when and how shared understandings change and how different corpora e.g., social media and traditional media, inter-connect. However, aggregate-level text analysis yields little to no possibility of producing causal explanations of how the observed macro-level features come about. Sociologists who collect and analyze individual-level texts can extract micro-level features of what and whom people are paying attention to, and what they think and feel about what is happening around them. Extracted measures of micro-level features are especially interesting for analytical sociologists, as they can be used for mechanistic multi-level explanations of macro outcomes.

While analytical sociologists study a vast array of social phenomena, the empirical work of this thesis centers around the concept of meaning, i.e., interpretations of the social world’s material or abstract constructs (Strauss and Quinn 1997; Goldberg and Stein 2018; Goldberg and Singell 2023). My work views meaning as relational, suggesting that the interpretation of an object, concept, event, or group emerges through associations to other representations (Mohr 1998; Patterson et al. 2007; Kirchner and Mohr 2010; Goldberg and Stein 2018). Understanding how meanings emerge and change over time is fundamental to explaining the social world and the ways people act within it (Mohr et al. 2020). A multi-level conceptualization of meaning suggests that meaning is both a macro-level outcome of social systems of interacting individuals and a micro-level feature of individual actors (Lizardo 2017). At the macro-level, shared meaning structures and their emergence in different cultural contexts is a social phenomenon for analytical sociologists to explain. Similarly, opportunities for any given action can be restricted or aided by such cultural meanings and norms. On the other hand, meaning at the micro-level can be used to develop mechanistic explanations of how such shared understandings and norms emerge. Analytical sociologists can also use micro-level meanings in explanations of individual-level behaviors. This thesis explores how meanings can be measured as they manifest themselves through the written word at both the aggregate and individual level, and how such
measures can be used to produce thick descriptions and multi-level explanations of the dynamics of the Swedish immigration discourse.

As mentioned above, the research in this thesis has three main objectives. The first is to discuss the use of text data for social inquiry. As text data acquire a more prominent role in empirical research, sociologists must consider the consequences that decisions taken in the course of data collection and data preprocessing may have on their results. Compared to other data sources, there are considerable gaps in the literature regarding the types of errors that may affect text data, how such errors affect the estimates produced using text data, and how to deal with these challenges.

Second, the work of this thesis makes a methods-focused contribution by introducing and exploring intrinsically interpretable text models that are particularly suited to sociological inquiry. I argue that intrinsically interpretable models have the potential to combine inductive (unsupervised) and deductive (supervised) approaches to test social theories. Intrinsically interpretable models can be used as a theory-based approach to modeling text, as they allow researchers to connect the output to social theories before modeling, while simultaneously not requiring researchers to come up with strict definitions of theoretical concepts before inductively discovering patterns in large-scale text data. As the toolkit from natural language processing expands, sociologists need to consider which novel instruments best fit the empirical applications of interest, and how to incorporate social theory using non-conventional methods. For example, while research has shown large language models with billions of parameters to have unparalleled performance in many NLP tasks such as document classification and language translation, sociologists need to consider how to link their outputs to sociologically relevant theoretical concepts (Bonikowski and Nelson 2022) and make sure they are reproducible and transparent (Nelson 2019). One of my contributions is to add to the literature by extending NLP methods towards improved interpretability.

Third, the work in this thesis explores how sociologists can use computational text analysis to rigorously study meaning and meaning-making at both an aggregate and individual level. Essay III and Essay IV produce insights into how different actors have discussed and interpreted immigration in Sweden over time. My work explores when and how collectively shared interpretations have shifted in meaningful ways, and showcases how to use individual-level text data for mechanistic explanations of aggregate-level measures of meaning. The first two essays of the thesis primarily contribute to the first two aims, making methodologically oriented contributions, while the latter two develop and apply language models to produce substantive empirical contributions.
Introduction

Essay I: *From Documents to Data: A Framework for Total Corpus Quality* presents a general approach to evaluating the scientific value of a given text dataset. The presented framework provides a map of potential errors that can impact the reliability and validity of any measure extracted from a corpus, and showcases ways of quantifying such errors. Here I address the first objective of the thesis, namely, the question of how researchers can know whether the data quality of a corpus is good enough to answer their research question.

Essay II: *Interpretable Word Embeddings via Informative Priors* presents a novel extension to the increasingly popular word embedding method. Our extension gives researchers the ability to study how the meaning of words relates to pre-specified binary dimensions, e.g., sentiment and gender. By presenting a novel, intrinsically interpretable model that outperforms its inductive precursor on multiple tasks and corpora, we show how increasing interpretability can improve the usability of word embedding models for sociologists interested in measuring and testing specific social theories. Essay II thus contributes to the second objective of the thesis.

Essay III: *Seeded Topic Models in Digital Archives: Analyzing Interpretations of Immigration in Swedish Newspapers, 1945–2019*, adds to the ongoing discussions in sociology regarding the identification of more formal ways to measure aggregate-level meanings. We trace dominant framings of immigration within the Swedish news media from the end of World War II until 2019. Essay III contributes to the second and third aims of the thesis by making a case for intrinsically interpretable models as a means of measure meaning, and by showcasing a never-before-seen comprehensive macro-level view of frames of immigration.

Essay IV: *Mechanisms of change: what explains shifts in online immigration discourse after terror attacks?*, explores the mechanisms that bring about changes in online discussions in the wake of Jihadist terrorist attacks. We test two mechanisms, namely: (1) changes in how people talk about an issue (within-individual change) and (2) changes in who is participating in the discussion (compositional change). Relying on seeded topic modeling of a corpus of posts written between 2010 and 2020 by anonymous users on Sweden’s most well-frequented online discussion forum Flashback, we study how online discussions about immigration shifted following 37 Jihadist terrorist attacks in Sweden and other European countries. Essay IV contributes to the thesis’s third aim and highlights how sociologists can utilize the increasing access to large-scale text data from individuals to produce the type of mechanistic explanations sought by analytical sociologists.

Together, these four essays contribute to methodological and theoretical discussions linked to the advances made in computational text analysis within...
the social sciences and the empirical literature on the use of text-as-data to study public discourses, particularly on immigration. The thesis introduces new methods for sociologists, which allow for the more transparent and replicable use of computational text analysis, and new tools to facilitate the discussion of the strengths and weaknesses associated with available data. Furthermore, the thesis provides insights into why and how the meanings of immigration produced by the Swedish media and the online public have shifted over time in Sweden. These results describe a cultural backdrop that has witnessed increasingly restrictive immigration policies and the rise of populist-right parties, and they also tackle the challenge of using individual-level data to study the mechanisms that bring about aggregate-level changes. I hope that my work takes an important step towards both integrating new methodologies into sociology and testing their potential to contribute to a new wave of empirical work that will strive to develop mechanistic explanations of the social phenomena detected in text data.

In the remainder of this introductory chapter, I first present current discussions on how computational text analysis can be used for sociological inquiry and the challenges associated with using text as data. I then introduce the theoretical foundations of meaning and meaning-making used in my work. In the fourth chapter, I synthesize the theories and ideas presented in the previous chapters to discuss how analytical sociologists can use text data to develop and test multi-level explanations of meaning and meaning-making. The subsequent chapter presents the domain of study for all the empirical work presented in this thesis, the Swedish immigration discourse. This chapter also discusses the ethical considerations associated with the thesis's work and data collection. This is followed by a summary of the appended essays. Finally, I discuss the results and conclusions from my thesis as a whole and potential future directions for this line of research.
Text as data in sociology

Viewing collections of text as an aggregate- or individual-level sensor suggests that sociologists can use texts as data from which they can extract information beyond the words on the page. More concretely, this means that aggregate-level texts contain traces of latent cultural features, such as cultural meanings, collectively shared emotions, and cultural biases, that are present in the context in which the texts have been written. For example, word usage, tonality, and topical associations all encode latent cultural features. Likewise, individual-level texts can contain traces of the author’s interpretations, opinions, beliefs, or sentiments that have seeped into the text either deliberately or unconsciously. Considering texts as an individual- or aggregate-level sensor allows us to infer information about the social actors who wrote the texts as well as to make inferences about their social environments (for examples of empirical work using texts as a social sensor, see Mohr 1998; Golder and Macy 2011; Mohr et al. 2020; DiMaggio et al. 2013; Bail et al. 2017; Fligstein et al. 2017; Törnberg and Törnberg 2016; Nelson 2021; Goldberg et al. 2016; Kozlowski et al. 2019).

Figure 1 presents a visualization of this idea that the target of inference is not the text data itself, but the texts’ authors or populations of authors. The arrow pointing back from the final analysis—either a description or an inference task—highlights that the target of analysis is the group or individual(s) that have written the texts and not the texts themselves. Figure 1 also shows how aggregate- and individual-level analyses may interact. Sociologists can infer macro-level characteristics by combining the texts written by multiple people and analyzing
Text as data in sociology

Figure 1: Representation of the use of text as a sensor. Researchers can choose to analyze all texts produced in a social context, extracting features at an aggregate level, or they can separate the texts written by different individuals to obtain micro-level features. Sociologists can use the extracted latent features either to produce insightful descriptions or as an input for additional statistical inference.

them as one corpus. Aggregate-level data can, of course, also comprise text data that are produced by entities from which individuals are indistinguishable, for example, manifestos of political parties. By contrast, if the texts produced by different individuals (for example, the orange and green actors in Figure 1) are kept as separate units of analysis, sociologists can use them to make inferences about individual-level characteristics.

Considering text as a sensor—both at the aggregate and individual level—also requires us to think about what a collection of what texts constitutes a sensor of. For example, if the corpus consists of documents comprising political party manifestos, researchers can use it to capture the ideas of political actors. Similarly, journalistic culture and media framing can be captured by analyzing a collection of newspaper articles. Simply put, we can only make inferences about the cultural context in which the texts were written or the individuals who have written the texts. In this dissertation, I use texts written by Swedish journalists and texts produced by that part of Sweden’s population which actively participates in online discussions as a social sensor of how the media and the online public interpret and talk about immigration.

Computational text analysis

Assuming that texts can be viewed as a social or individual-level sensor, how can we capture their signals to infer the cultural context in which the texts have been produced or the individuals who have written them? To do this, tools that deal with large-scale and unstructured text data are essential. The work in this thesis relies on methods developed in the vibrant fields of natural language processing.
NLP) and machine learning (ML). NLP and ML are both interdisciplinary fields that combine computer science, mathematics, and linguistics, and in which methods are increasingly quickly being developed to perform tasks such as classification, translation, and prediction (Hirschberg and Manning 2015; Jordan and Mitchell 2015; Evans and Aceves 2016). In this thesis, the term “computational text analysis” refers to analyses that utilize these methods (DiMaggio 2015). Answering social scientific research questions using computational text analysis usually involves different tasks. I use the categorization suggested by Grimmer et al. (2022) to summarize the tasks for which sociologists may use computational text analysis: discovery, measurement, and inference.

**Discovery**

Discovery tasks aim to aid the researcher in finding new ways to organize data or to confirm that the expected category, concept, or pattern exists within these data (Evans and Aceves 2016; Grimmer et al. 2022). Because text data are extraordinarily high-dimensional\(^3\), they may be represented in almost infinite ways. Consider, for example, the English language, which according to the Merriam-Webster Dictionary (2021) consists of roughly 470,000 words. This means that an English language text that is 50 words long can, in theory, be represented by \(470,000^{50}\) possible combinations of words. The high dimensionality of text makes it nearly impossible to be aware of all the relevant information that texts hold before data are collected, which leaves a great deal of room for inductive explorations of this type of material. The process of reorganizing and summarizing data in different ways can be labelled discovery.

Discovery tasks often rely on models that identify categories in data completely inductively, for example, by identifying clusters in the data or by reducing the dimensionality of data to make such patterns easier to visualize (Evans and Aceves 2016). These methods that do not require ascribing labels to learning patterns are often called unsupervised methods (McFarland et al. 2013; Hastie et al. 2009a). Researchers can create different views of data by using different methods to identify or optimize different features of texts. Hence, discovery can help researchers to generate novel insights about the data that, to begin with, are not visible to the human eye.

While the discovery phase largely remains undescribed in most published research, it is increasingly recognized that is essential to identify and transparently

---

\(^3\)The high-dimensional nature of text also results in issues associated with the “curse of dimensionality” (Bellman 1964). This “curse” states that as the dimensions of the data increase, an (exponentially) increasing amount of data is needed to analyze the space (Hastie et al. 2009b)
Text as data in sociology

communicate the starting points for deductive analysis in order to explain why specific patterns appear (Nelson 2020).

Measurement

The excitement surrounding computational text analysis within the social sciences is closely linked to the perception that these methods yield new and more rigorous ways of measuring theoretical concepts than those allowed by traditional data sources (Bail 2014; Mohr et al. 2020; Evans and Aceves 2016; Bhatt et al. 2022). Measurement tasks involve translating texts into a latent representation of a theoretical construct (Grimmer et al. 2022). Many methodologies exist for measurement tasks, including both unsupervised and supervised methodologies. As described previously, unsupervised methods can inductively find patterns in data. The identified pattern can then post hoc be interpreted as reflecting a specific theoretical construct. Supervised methods are instead models in which a training stage typically relies on labelling data to establish predefined categories in the data set (Evans and Aceves 2016). Supervised methods, therefore, require researchers to a priori define the theoretical construct of interest and indicate how it should be operationalized in the data. The recent advances in large language models (LLMs) blur the line between unsupervised and supervised modeling approaches. These models have demonstrated the ability to classify, summarize, and measure sociologically relevant concepts in texts without the need for labeled data (see Ziems et al. 2024, for a summary). However, both traditional method families are used to extract measures that provide thick and systematic descriptions on a large scale, which can be used as dependent or as explanatory variables in the final task of social inquiry—inference.

Inference

Both discovery and measurement are tasks that are centered around the available data, whereas inference is about generating insights about a broader context or population, or—in other words—results that extend beyond the texts themselves (Hofman et al. 2021). Inference is crucial for most social scientists who wish to predict or explain social phenomena.

Typically, inference follows measurement tasks. Sociologists can use these measures to test social theories at the aggregate or individual level. At the aggregate level, researchers may be interested in learning how a specific group or society understands important events, or how different groups in society influence one another’s worldviews, for example, by studying how different
groups frame minorities or current political issues at different points in time (Törnberg and Törnberg 2016; DiMaggio et al. 2013; Fligstein et al. 2017; Garg et al. 2018; Card et al. 2022). At the individual level, researchers may instead be interested in testing different causal mechanisms by which agents’ perceptions influence their behaviors, or how the interaction between individuals impacts the interpretations of objects that emerge in a group.

**Common issues faced by sociologists using text as data**

The three objectives of this thesis require an understanding of the limitations of the approaches that are currently employed when using text as data. I have identified three challenges to the use of text data in sociology: (1) ensuring reliability and validity, (2) incorporating social theory, and (3) evaluating models.

**Reliability and validity**

Having more data does not resolve issues associated with poor data quality. Concerns have previously been raised about the need to reflect on data quality and the effects that data curation may have, particularly when vast data repositories are used to gain insights about the social world (Nelson 2019; Ignatow 2020; Nguyen et al. 2020; Bonikowski and Nelson 2022; Yung et al. 2023). When using text data in empirical work, we must ensure—as with any other type of data—that the data are reliable and valid. These criteria pertain to the data themselves, before they are used as an input in modeling, and the challenges of reliability and validity therefore differ from the third challenge—evaluating models—which arises when texts are then actually used as data. How can researchers judge the scientific value of their available corpus in a structured way? How can they understand the strengths and weaknesses of a specific corpus and ensure reliable and valid results? Without developing standards for answering these questions, users of text as data in empirical research run the risk of drawing faulty conclusions solely as a result of data quality issues.

The need to evaluate the appropriateness of data before using it for research might seem like a trivial insight. To be sure, the lack of theoretical discussions of the data quality of text data stands in contrast to the situation regarding other types of data that are traditionally used in the social sciences. Inspired by the evaluative frameworks developed by survey methodologists (Andersen et al. 1979; Lyberg 2012; Biemer and Lyberg 2003; Groves and Lyberg 2010), I use the concept of data quality to discuss the “fitness for use” of a particular corpus (Biemer and Lyberg 2003). Essay I proposes such a framework by cataloging and
Text as data in sociology
categorizing dimensions through which a corpus’s fitness for use can be evaluated. The proposed Total Corpus Quality framework identifies three dimensions that impact the reliability and validity of a corpus for social inquiry: total corpus error, corpus comparability, and corpus reproducibility.

Data quality problems must be recognized and examined, rather than remaining largely ignored within the computational text analysis community. The first essay of this thesis extends previous work in this area by providing a generally applicable framework to discuss and measure different quality dimensions of a given corpus used for social inquiry. Hence, my work provides a stepping stone for future discussions on data quality, identifying the types of errors that may affect text data and measures extracted from these data, and discussing important practical suggestions for evaluating the impact of different types of errors.

Incorporating social theory

Incorporating social theory is one of the most fundamental challenges to the use of computational text analysis. I argue that “intrinsically interpretable” models hold great potential for the augmentation of sociological expertise and the testing of social theories. I use the term intrinsically interpretable to distinguish models that provide an output that can be understood in light of theory without relying on post hoc interpretations of the output produced by unsupervised and black-box models. Introducing extensions to standard unsupervised methods that address the challenges of incorporating social theory into computational models constitutes an essential contribution of this thesis.

Intrinsically interpretable models have so far not received much attention from sociologists who use computational text analysis. These models differ from purely unsupervised methods for analyzing texts. Unsupervised methods search for patterns in data inductively, which means that they lack a channel through which sociologists can directly test social theories and quantify interpretative uncertainty (DiMaggio 2015). Hence, relying on unsupervised methods makes it difficult to directly connect one’s modeling to the testing of sociological theories.

Instead, researchers commonly study the output of unsupervised models and try to connect it to social theory post hoc. This form of post hoc interpretation means that explanations of results in light of theory are provided after the model has been trained (Choudhary et al. 2022; Madsen et al. 2022). For example, the seminal paper by DiMaggio et al. (2013) used an unsupervised model to detect latent topics related to art grants in US newspapers. After running the model, the authors ascribed some of the topics meaningful interpretations in the context of art funding, using these selected topics to describe the development of political
debates on “culture wars” over time. More recent mixed-method approaches, which combine an inductive discovery process with model selection via post hoc interpretations and validations, have been promoted by different researchers (Nelson 2020; Grimmer et al. 2022) and utilized to produce impressive empirical work (for example Nelson 2019; Nelson 2021; Kozlowski et al. 2019).

Even given the success of these methods, there is nothing in the modeling procedure that ensures that the model will recognize the concepts of interest that researchers have in mind a priori in relation to the analysis. I would argue that this reliance on post hoc interpretations may lead researchers to reformulate their research questions or change their chosen theoretical framing if a model does not identify theoretically informed patterns. Changing one’s research question or revising theory in the light of one’s results involves a departure from the deductive research model traditionally used by sociologists. The deductive approach typically means that researchers use theory to formulate testable hypotheses, after which they collect data and, finally, analyze it (Gorski 2004). Others have gone further in their criticism, suggesting that studies that employ the inductive approach are inherently atheoretical (Brandt and Timmermans 2021; Pääkkönen and Ylikoski 2021).

To avoid post hoc reasoning, researchers have used supervised models. Traditionally, supervised models rely on hand-coded samples, on which they are trained, to learn about specific features of a text. In the supervised tradition, researchers use theory to decide what concepts the model should be trained to detect and to create training data. Recently, a class of complex nonlinear models—sometimes referred to as deep learning models, or “black-box” models—has been gaining momentum in social science. These involve an increased use of pre-trained transformer models (Bonikowski et al. 2022; Vicinanza et al. 2023; Card et al. 2022; Le Mens et al. 2023; Kovács et al. 2023), which are then typically fine-tuned to fit a specific research question. For example, Bonikowski et al. (2022) used theory to label which texts were about “nationalism” and “populism,” which they then used to fine-tune a model to automatically identify these concepts in other texts.

It is worth thinking about the implications for social theory associated with a reliance on these “black-box” models. The term “black-box” signals that the models make it difficult to interpret single parameters to learn a sociological theory-like mapping between the dependent and explanatory variables and to explain how and why predicted relationships hold in a manner that humans can easily understand (Rudin 2019). This is in stark contrast to the traditional statistical analytical frameworks that are familiar to most sociologists, which estimate the relationship between the outcome of interest and each explanatory
variable included in the model (Breiman 2001; Daoud and Dubhashi 2021). The magnitude and direction of the relationships between explanatory and dependent variables typically constitute the main focus of interest when testing social theories and explaining social phenomena. To be clear, while I would argue that the use of “black-box models” may be problematic in relation to the testing of social theories about the relationship between input and output variables, the increasing evidence of their usefulness for e.g., data annotation (Do et al. 2022; Gilardi et al. 2023; Ziems et al. 2024) and the generation of synthetic data samples and survey answers (Argyle et al. 2023; Kim and Lee 2023) means that they remain relevant to sociological research.

Viewed from a different perspective, “black-box models” are needed to capture the very complex relationships between input and output variables. The strong performance boost provided by transformer models is one example of the way increasing complexity works well as a means of modelling some aspects of the social world. Even if “black boxes” are indeed needed to map relations in the social world, it is worth considering increasing the extent to which such relationships might be made explicit and mapped to social theories.

I view intrinsically interpretable models as one promising addition to the ways in which sociologists can incorporate social theory. In NLP and ML, the term “interpretable” often refers to understanding the ins and outs of how a model works, rather than explaining what aspect of the social world produces the result in question (Rudin 2019)—the main objective for most social science. Thus, efforts to increase the interpretability of models within the NLP and ML communities focus on identifying what part of the input was most important in generating a predicted outcome.4

In the social sciences, interpretability typically refers to explaining a phenomenon in such a way that people are able to understand it. Sociologists are at the same time interested in explaining what produces a specific observed relationship between the dependent and explanatory variables in a model. These two views of interpretability can be complementary, i.e., increasing our understanding of how models work can say something about relationships in the social world and vice versa. Essay II improves on the NLP and ML type of interpretability by isolating which part of a model captures dimensions of interest. The model is also relevant for sociologists because it gives them the ability to study how the meaning of words shifts on a pre-specified binary dimension,  

4For example, LIME (Ribeiro et al. 2016) and SHAP (Lundberg and Lee 2017) are two model-agnostic algorithms that are applied post hoc to classification models to explain what part of the input has been most important in producing the class prediction. Both models identify the words in a text or the pixels in an image that are most important in generating a prediction.
which can help to answer questions about cultural change. Essays III and IV use a semi-supervised extension of topic models, which improves their social scientific interpretability and allows us to construct a research design to explore how the media framing of immigration shifts over time, as well as how and why online discussions about immigration shift around the time of deadly Jihadist terrorist attacks.

Model evaluation

The complexity of language makes many methods of computational text analysis difficult to validate (Grimmer et al. 2022). Validation becomes even more complicated when sociologists employ methods that were initially developed for other tasks, and that are used in different ways to answer social inquiries, such as the use of prediction models to make inferences. Nevertheless, all text analysts need to consider the question: how can we know how well our analytical strategy works? Developing standards to answer this question continues to represent a major methodological challenge.

Today, most sociologists rely on the standards developed within the NLP and ML communities, which commonly rely on comparing the output from algorithms with human judgment, letting humans annotate a so-called "gold standard." A gold standard contains labels related to the desired output that are produced by one or several experts. Examples include labels on the topics discussed in a document, the sentiment of a social media post, or the political leaning of speech. Manual annotation is typically conducted on texts that are not used to train the model, i.e., a hold-out test dataset. By comparing labels predicted by the model with the human-annotated gold standard, we learn how well a model works on previously unseen data, i.e., its out-of-sample accuracy (Barberá et al. 2021). Recent developments show that combining human annotation with computational text models constitutes a promising approach for producing labeled data at scale (Nelson et al. 2021; Do et al. 2022). However, many theoretical concepts in social science are complex and can be understood in many different ways, such as social status or political ideology. Deciding whether a text contains a specific concept can therefore be difficult even for humans (Baden et al. 2022). The complexity of many annotation tasks often leads to a low level of agreement between independent human coders (Bais et al. 2019) and negatively affects the likelihood that the same coder will assign similar labels at different times (Belur et al. 2021). The reliance on human-annotated gold standards thus ignores the fact that human coding is error-prone (Hopkins and King 2007), especially in relation to highly complex theoretical constructs. Still, I partly rely on comparing model
output with human annotations in Essay III and IV. In Essay III, we also propose some innovative ways to validate seeded topic models, for example studying the thematic patterns among words that the model inductively identified as being about focal topics.

To conclude, the three challenges discussed here are often not noted in empirical work, which makes them difficult to locate and resolve. Highlighting these issues, rather than sweeping them under the carpet, is essential in order to stimulate more transparent and more rigorous research in the sociological application of computational text analysis.
Most people understand “meaning” as equivalent to its dictionary definition: “the meaning of something is what it expresses or represents” (Cambridge Dictionary 2022). In the wake of the intersecting forces of the computational revolution outlined in the introduction, the interest in measuring meaning has been gaining momentum among cultural sociologists in recent years (Mohr et al. 2020; Taylor and Stoltz 2020; Stoltz and Taylor 2021; Bhatt et al. 2022; Lix et al. 2022; Nelson 2019; Arseniev-Koehler and Foster 2022; Bonikowski and Nelson 2022; Nelson 2021; Best and Arseniev-Koehler 2023; Goldberg and Singell 2023). This section begins with a definition of how the concept is used in this thesis.

**The meaning of meaning**

Meaning conceptualizes the idea that objects, subjects, and actions can be assigned different interpretations by different individuals, groups, and societies through the act of meaning-making (Mohr et al. 2020). In this thesis, I consider meaning as both a multi-level concept and as a relational concept.

**A multi-level concept**

Over recent decades, cultural sociologists and other scholars of meaning have been heavily influenced by cognitive scientific discoveries in their theorizing about meaning and how meaning-making occurs (DiMaggio 1997; Strauss and
Sociological approaches to meaning

Quinn 1997; Cerulo et al. 2021; Goldberg and Singell 2023). In the cognitive view, meaning is something that is primarily intrinsic to people’s minds, making it potentially difficult to measure outside the brain and its cognitive functions. Given its interest in how meanings can be measured using text data, and in how and why individuals’ and groups’ talk about immigration changes over time, this thesis moves beyond considering meaning through a purely cognitive lens and adopts a multi-level view of meaning.

The multi-level view of meaning suggests that interpretations of objects, subjects, actions, and events exist internally within individuals and externally to any single person as systems of meaningful associations that emerge dynamically through interactions between people and institutions. As a proponent of the multi-level view of meaning, Lizardo (2017) refers to this twofold nature of meaning as “public culture” and “private culture”. Public culture consists of meaning structures such as frames and narratives that are produced and expressed by groups, organizations, or other extra-individual entities, for example, the media (Arseniev-Koehler and Foster 2022; Lizardo 2017). Private culture, on the other hand, operates at a declarative and non-declarative level within individuals (Vaisey 2009; Lizardo 2017). Individuals’ interpretations, attitudes, and worldviews are part of the declarative private culture they can express through symbols such as language. Those types of interpretations of the world that cannot easily be expressed but that instead operate automatically within the cognitive or bodily functions of individuals—for example, cognitive schemas and associations—constitute the non-declarative individual-level meanings that are also part of the multi-level view.

Considering meaning as a multi-level concept makes it fit well with the idea that texts at both the aggregate and individual levels are of interest for sociological inquiry. I utilize the dual nature of meaning in the empirical essays presented in this thesis; while public culture is the main focus of Essay III, which describes when and how media framings of immigration develop and shift over time, Essay IV instead tests the use of posts that people write online, viewed as a source of declarative private culture, to explain changes in forum-level interpretations.

5 Although this view of meaning emphasizes meaning at the cognitive level within individuals, it also lays the foundations for considering meaning at a collective level, since people can share cognitive schemas. When people share a schema, such schemas are sometimes called cultural schemas rather than cognitive schemas. Boutyline and Soter (2021) define cultural schemas as “socially shared representations deployable in automatic cognition".
A relational concept

The empirical work in this thesis also operationalizes meaning in relational terms (Emirbayer 1997; Mohr 1998). The relational perspective suggests that the meaning of an object, concept, or people arises in its associations with other objects, concepts, or people (Goldberg 2011; Goldberg and Stein 2018; Hunzaker and Valentino 2019; Goldberg and Singell 2023). The relational view means that “[...] meanings do not inhere in symbols (words, icons, gestures) but that symbols derive their meaning from the other symbols with which they appear and interact” (DiMaggio et al. 2013). This signals that meaning does not reside within the substance of single objects or symbols but emerges as objects or symbols are associated with other objects or symbols (Kirchner and Mohr 2010). We learn about the meaning of things and ideas by mapping their similarities and dissimilarities. Importantly, different individuals and groups will hold different views on such patterns of relations, and these views may change over time as a result of distinct social processes.

Figure 2: Conceptualization of “meaning” as the association between the concept A and the other concepts B, C, D, and E among three different actors. A tie symbolizes that the actor believes that two concepts are related. The width of the tie signals the degree to which two concepts are related.

Following this relational view, both Essays III and IV study how different groups in society discuss and fill the concept of “immigration” with meaning by studying how immigration is associated with other concepts at different times. Figure 2 illustrates the way the meaning of immigration is conceptualized in the empirical essays of this thesis. The figure shows how the meaning of “immigration” (A) is dependent on its relation to the concepts “economy” (B), “human rights” (C),
Sociological approaches to meaning

"security" (D), and “culture” (E) for three different actors. Throughout this work, I consider both the existence of ties between concepts (nodes)—signaling that two concepts “go together” (Goldberg and Stein 2018)—and the strength of the association as essential features that make up meaning. In Figure 2, actors 1 and 3 share ties between concepts, but the strength of these ties differs. Actor 2, on the other hand, associates concepts in very different way to the other actors. Actor 1 associates immigration mostly with security, culture, economy, and lastly, human rights. Actor 2 mainly associates immigration with human rights and culture and, to a lesser extent, with security. Actor 2 does not relate immigration to the economy at all. Finally, actor 3 associates immigration most strongly with the economy and security, followed by human rights and culture. In this example, actors 1 and 3 interpret immigration primarily in economic terms.

While actors 1 and 3 do not fully attach the same meaning to immigration, their interpretations of immigration are more similar to one another than they are to that of actor 2. If the ties and their widths are identical across actors, the actors share an understanding of the concept under scrutiny. When many members of a social group or society share a meaning structure, it is sometimes called a shared understanding (Goldberg 2011) or a cultural meaning (Strauss and Quinn 1997; d’Andrade et al. 1984). To be clear, Figure 2 represents a simplistic image of meaning, focusing on one concept (immigration) and its relations to other concepts. In reality, each of the other concepts included in the figure—economy, human rights, security, and culture— are in turn defined by concepts that may be unrelated to the concept of immigration and, therefore, not present in the figure. Much like the social world itself, meaning does not exist in a vacuum. Instead, the meaning of all concepts, events, and people in the social world comprises a myriad of relations, creating a single extensive network of associations.

Lastly, this thesis considers meanings as more general than attitudes because meanings only signal associations and not the evaluation of these associations. Others have suggested that giving something meaning by associating it with other things precedes the process of assigning an evaluation to the object or practice (Goldberg and Stein 2018). Conceptually, however, attitudes can be incorporated into Figure 2 by adding signed associations between an actor and the concepts presented in the figure.

For illustrative purposes, the concepts associated with immigration that are used in the two essays differ slightly from those included in Figure 2. Also, Essay III does not study meaning at the individual level.
Meaning and multi-level explanations of social outcomes

I argue that the multi-level view of meaning adopted in this thesis allows analytical sociologists to consider using the concept to develop mechanistic explanations of social outcomes. For example, meaning can be theorized as both a driver in people’s physical and social environments and an intra-individual force (Lizardo 2023). Public culture—the interpretations available to individuals in the outside world—may guide actions by helping to form internalized meaning structures and habits (Swidler 1986; Bourdieu 1990; Alexander 2006). For example, Arseniev-Koehler and Foster (2022) have argued that the way the media frame what it means to be “fat” is crucial in shaping people’s cognitive schemas by connecting the concept of “fat” to other concepts in news reporting. Public culture also depends on people’s private culture and behaviors (Lizardo 2017). Both undeclared private culture—internal states—and declared private culture—what people express about their internal states—have long been central in explaining how individuals in social systems motivate and justify the actions that generate macro-level social outcomes (Weber 2013; Tucker 1965; Vaisey 2009).

When understood in a multi-level perspective, I believe that meaning can be incorporated in multiple useful ways into a (cognitively) extended version of the explanatory scheme often referred to as the “Coleman boat” (Coleman 1986), which is used by analytical sociologists to represent the interdependency between macro- and micro-dynamics. The two upper layers of Figure 3 represent the traditional non-extended version of Coleman’s boat.

The first arrow in Figure 3 illustrates how macro-level meaning structures at time \(t\) impact individuals. The impact of public culture on individuals does not have to be direct but may be filtered through an information environment, for example, the news media.

The second arrow illustrates the mechanism through which agents’ private culture can impact their behaviors. The traditional Coleman boat does not allow for distinguishing between declarative and non-declarative meanings. I therefore propose adding a cognitive layer to Figure 3 to better align the multi-level view of meaning with multi-level explanations (more on this below).

The third arrow captures the mechanism through which the actions of other individuals may have consequences for other agents. This feedback loop is a standard extension of the original mechanisms used in analytical sociology (Ylikoski 2016).

The fourth arrow represents the transformational mechanism, capturing how the actions of individuals aggregate to form macro-level meaning structures (or other social outcomes of interest) at time \(t + 1\).
Sociological approaches to meaning

Public culture

Public culture exists at a meta-individual level, placing it in the upper part of the boat, which represents macro-level societal structures. Public culture has two potential functions in a multi-level explanation of social outcomes in Figure 3, either as (1) a macro-level structure that impacts agents’ behaviors (node 1) or (2) as the social outcome of interest (node 4). The social movement literature is rich with examples of how elements of public culture impact individual behavior. Much of the work in this literature seeks to explain how the frames used by social movements mobilize actors into participatory behaviors (for overviews and examples see Benford and Snow 2000; Farrell 2013; Dunivin et al. 2022). Similarly, scholars of religion engage in discussions of how moral institutions impact individuals (Wuthnow 2008).

Other literature has instead focused on aggregate-level meanings as the outcome of interest. For example, some studies on cultural change have used survey data to examine how and why cultural narratives and views change as a result of individual-level dynamics (Vaisey and Lizardo 2016; Vaisey and Kiley 2021). Instead of explaining why public culture changes using micro-level dynamics, Essay III provides insightful descriptions and macro-level explanations...
as to why the publicly available meaning structures produced by the media, i.e., frames, change over time.

**Private culture**

Both declarative and non-declarative private culture can also be incorporated into multi-level explanations of social outcomes. Declarative private culture comprises the worldviews that an individual can articulate. Declarative private culture is what individuals themselves use to motivate or justify their behaviors, and thus constitutes part of the characteristics of agents found in node (ii) in Figure 3. For example, Vaisey and Lizardo (2010) studied how declarative private culture explains friendship formation. Essay IV relates to this literature by studying shifts in the way actors describe interpretations of immigration in online discussions (declarative private culture) following Jihadist terrorist attacks.

While both public culture and declarative private culture can fit into the traditional two-layered version of Coleman’s boat, non-declarative meaning structures (such as schemas) cannot. To fully fuse the theoretical multi-level view of meaning with the idea of multi-level explanations, the traditional two-layered model of explanation must be extended with a cognitive layer. Figure 3 includes this cognitive extension. Arrow 1a symbolizes the mechanism by which characteristics and experiences of the agent impact the accessibility of specific cognitive schemas. Arrow 2a, in turn, represents the mechanism that activates specific schemas, and arrow 3a shows how schema activation affects the future cognitive schemas of the individual. The final arrow, 4a, shows the direct link between action schemas and the other non-declarative forms of culture. I argue that arrow 4a captures the link between non-declarative culture and action, while arrow 2 captures the impact of declarative culture on action (Lizardo 2017). This division means that while actors can reason about and discuss how materialized forms of meanings affect behavior (arrow 2), actors cannot easily do so with regard to the cognitive processes that impact on action (arrow 4a) since these operate at an unconscious cognitive level.\(^7\)

---

\(^7\)Figure 3 may raise questions if the relationship between meaning and individual-level action is causal. As is highlighted by the division between declarative and non-declarative culture, much action is habitual and performed with no particular motivation (Vaisey 2009; Miles 2015). This division also signals that even when actions are unconscious, they are driven by cognitive aspects of meaning. However, actors will most likely not be able to articulate the relationship between meaning and action other than by using meaning as a post hoc justification. Researchers may potentially not be able to use text data to produce realistic multi-level explanations in these cases.
I have so far discussed how analytical sociologists could theoretically incorporate the multi-level view of meaning into multi-level explanations. To empirically engage with processes of meaning-making, analytical sociologists must be equipped with the data and methods to measure meaning. Much empirical work in sociology has historically relied on self-reported survey data (Lazer et al. 2009; Keuschnigg et al. 2018). It is well-known that survey data are riddled with bias, especially when people are asked to explain past behavior or socially sensitive issues (Fisher 1993; Krumpal 2013). Furthermore, neither survey data nor the traditional statistical models used to analyze such data are well-suited to establishing micro-macro links, in part as a result of the limited possibilities for collecting (large-scale) networked data, partly because of the need to assume independence between observations. Analytical sociologists are therefore looking for alternative data sources and methodologies that support the study of mechanistic explanations. For example, large-scale observational data, such as population-level register data, are used to study the macro-level consequences of micro-level decision making in various social contexts (Bruch and Mare 2006; Mutgan 2021; Arvidsson et al. 2021; de la Prada and Tapia 2022). However, while population registers may contain information about revealed preferences—i.e., actions—they do not include information about how individuals or groups interpret their surroundings and talk about what is happening in the world. I would therefore argue that sociologists interested in meaning-making must find alternative sources to measure meaning.
Measuring meaning

This thesis argues that text data constitute a promising data source for (1) measuring macro-level features and producing thick aggregate-level descriptions of shared interpretations about the world in different groups and societies. Moreover, individual-level text data hold even greater promise for analytical sociologists, since they can be used to (2) extract declared micro-level measures of what people pay attention to and how they interpret what is going on around them, which opens up new avenues for empirically testing social theories about meaning-making.

With respect to individual-level data, there is no consensus regarding whether text data include declarative or non-declarative culture (or both). In the empirical work presented in this thesis, I rely on a non-cognitive interpretation of language (Arseniev-Koehler 2022). The non-cognitive interpretation of language suggests that the way individuals and groups talk can be used to measure public culture and private declarative meaning structures. This indicates that texts are most useful for measuring the macro and micro layers of meaning. The non-cognitive perspective on text data used in this thesis aligns well with the bulk of the empirical research focused on measuring meaning using texts, which relies on the idea that declarative meanings—i.e., what people say about how they view or feel about the world—are encoded into texts (Vaisey 2009; Kozlowski et al. 2019; Arseniev-Koehler 2022).

Moreover, I argue that considering only non-cognitive constructions of meaning resonates well with the emphasis on realism in analytical sociology. The pursuit of realistic explanations has led to a shift away from “intentional explanations” (Hedström 2021; Arvidsson 2022). Moving away from intentional explanations allows analytical sociologists to focus on what people say about their perceptions of the world—i.e., declarative private culture—rather than trying to unlock the cognitive processes that are in play. This does not mean that non-declarative private culture is unimportant for multi-level explanations. Instead, focusing on public and declarative private culture allows researchers to explore meaning structures that are at least partly observable when using the data and methods discussed in this thesis.

Measuring public culture

Most empirical research using text data describes long-term cultural developments. In other words, text data have successfully been used to map measures of public culture (Garg et al. 2018; Caliskan et al. 2017; Kozlowski et al. 2019; Card et al. 2022; Rule et al. 2015; Michel et al. 2011; Nelson 2020; Voyer et al. 2022; Lorenz-Spreen et al. 2019; Best and Arseniev-Koehler 2023; Boutyline et al. 2023; van
Loon and Freese 2023). These studies provide insightful accounts of how the content and tonality found in various corpora—ranging from political speeches to historical literature—have shifted over long periods, and they have enriched analytical sociology by describing cultural landscapes at different times.

Currently, there are two distinct “families” of computational text methods that are employed to measure macro-level meaning structures in a relational sense of the term; these focus either on document structure or on semantic similarities (van Loon 2022). Methods that focus on document structure consider how words are clustered within the bounded unit of a document as a means of measuring the meaning of latent high-level abstractions, such as themes of related content. Topic models are one of the most common methods used to extract meaning from document structure (Blei et al. 2003). Topic models have been used to study the aggregate-level meanings of various themes by studying their co-occurrence with other themes. Examples include the meaning of art in US newspapers (DiMaggio et al. 2013), financial policies among the Federal Open Market Committee (Fligstein et al. 2017), and differences and similarities in the political logic that guided the first- and second-wave women’s movements in Chicago and New York City (Nelson 2020).

The second family of methods relies on semantic similarity. Methods of this kind produce a numerical representation of words, which captures their meaning. In contrast to document similarity methods, semantic similarity methods focus on the meaning of individual words rather than higher-order concepts such as themes. The most common methods used to extract meaning in this way are word embedding models (Mikolov et al. 2013; Pennington et al. 2014) and contextualized word embedding models (Devlin et al. 2018). The former estimate one numeric vector per word in the vocabulary, whereas contextualized word embedding models generate separate vectors for the same word depending on its context. Extending word embeddings to account for context is believed to allow more complex and precise patterns of meanings to emerge (Bonikowski et al. 2022). These models have been used to study the public meaning of gender online and in historical literature (Garg et al. 2018; Caliskan et al. 2017), social class in historical literature (Kozlowski et al. 2019), and intersectional experiences in North American slave narratives (Nelson 2021).

The main benefit of using document structure methods when measuring macro-level meanings is their simplicity when it comes to training new models on the specific corpora related to a sociological inquiry. In contrast, most empirical research using semantic similarity models relies on pre-trained models. Pre-trained language models are typically trained on large corpora of texts before
Measuring meaning

being released to the scientific community for download. Sociologists then fine-tune the pre-trained models to their specific sociologically relevant task. For example, Bonikowski et al. (2022) used this approach to identify the presence of populism in political speeches. Hence, while semantic similarity methods may capture more complex measures of public culture, it is difficult to evaluate the extent to which potential biases in the corpora used for training may impact the final results. Furthermore, there tends to be a lag in the availability of pre-trained large language models for low-resource languages, creating a gap in what sociologists of public culture are able to analyze. On the other hand, document structure methods measure meaning in a more simplified manner but can easily be applied in many domains and cultural contexts. I do not therefore view either of these approaches as constituting the best overall approach to measuring macro-level meaning; instead the appropriateness of the models depends on the sociological inquiry of interest. I explore both methods as means of measuring and describing shifts in public culture. Essay III measures macro-level features to describe how the salient frames in the media discourse about immigration have shifted over the last 75 years. In contrast, Essay II proposes an alternative word embedding model that measures the meaning of words on a pre-specified binary dimension using informative priors. We use this method to describe how the sentiment and degree to which words are considered male or female shifted in different spheres over time.

While most previous research using aggregate-level texts provides descriptions of public culture, they can also be used to analyze macro dynamics. Examples include investigations of the role that different social groups play in shifting shared understandings (Russell Neuman et al. 2014; Barberá et al. 2019), and of how disruptive events impact collective sentiments (Flores 2017; Garcia and Rimé 2019). While analytical sociologists ideally want to study macro-micro-macro dynamics, I would argue that recording when and how shifts in cultural landscapes occur can constitute a first step towards developing mechanistic explanations of cultural change and stability. Sociologists can use the extracted macro-level features as dependent variables in regression-type analytical strategies or as empirically calibrated macro-level outcomes in agent-based models, which constitute the primary tool of many analytical sociologists (Macy and Willer 2002; Hedström and Manzo 2015; Manzo 2014).

8During the writing of this introductory chapter, new large language models have been released that allow researchers to train models using limited computational resources. This type of model may change the level of reliance on pre-trained language models in sociology.
Measuring private culture

The second avenue by which computational text analysis can add value to analytical sociology is by extracting individual-level features of the way people talk about what is happening around them. The way people talk is likely to contain traces of what people think, believe, and pay attention to (Benoit 2020; Bonikowski and Nelson 2022). For example, people use words related to a set of psychologically meaningful categories to a differing extent depending on personality traits (Tausczik and Pennebaker 2010; Chung and Pennebaker 2018); politicians of different ideologies talk differently about the same issues (Card et al. 2022; Fuhse et al. 2020; Bonikowski and Gidron 2016); and journalists can report on the same news in different ways, depending for example on the political leaning of a given newspaper (Roggeband and Vliegenthart 2007; Gentzkow and Shapiro 2010; Bleich et al. 2015). Analytical sociologists can thus use individual-level text data as a sensor for declarative private culture, i.e., the way individuals perceive people, objects, and events.9

The most promising observational data source for measuring micro-level features and testing mechanistic explanations is found in digital trace data (Jarvis et al. 2021). Digital trace data contain information about transactions and conversations that are left online or captured by sensor devices out in the real world (Agarwal et al. 2008). The digital traces left online often take the form of panel data, which allow for longitudinal studies of many individuals. Importantly, digital trace data also often contain information about network structures, i.e., who are interacting with each other online. Network structure allows us to incorporate information about how individuals in social systems potentially influence one another, adding additional realism to our explanations. At the present time, however, there are still relatively few studies that link computational text analysis and individual-level data to produce mechanistic explanations (see Karell and Freedman 2020; Czymara et al. 2022; Bonikowski et al. 2022, for

9I do not argue that the individual’s emotions, beliefs, or worldviews are perfectly represented in texts. Instead, I recognize that any such measures extracted from data are a noisy signal based on how the person talks about how she/he thinks and feels about the world. Furthermore, when people meet and interact, they may present themselves as being different from their “true selves” in order to be perceived as socially acceptable (Goffman 1956). Similar ideas also exist in relation to online interactions, where people can either try to recreate their offline selves or adopt an online persona (Bullingham and Vasconcelos 2013). The pros and cons of using “ready-made data” (e.g., existing textual sources) and “custom-made data” (e.g., survey data) to study the social world is discussed in-depth in Salganik (2019). I would argue that textual resources that include texts written by a large number of individuals, and that are recorded with high granularity, are unlikely to mostly contain people who are constantly presenting a facade in such a way as to make extracted measures completely invalid and unreliable.
Measuring meaning

some notable exceptions). Essay IV adds to the relatively sparse literature that has used individual-level text data to study the mechanisms that drive aggregate-level changes in online discussions following unexpected events. The use of individual-level digital trace panel data from Sweden’s largest online discussion forum allows us to test whether observed changes to collective interpretations of immigration are driven by how people talk or who is talking.

Even if digital trace data are abundant\(^\text{10}\), most available methodologies require large amounts of data focused on the unit of analysis to obtain reliable results. More concretely, when you are interested in learning about individuals, large numbers of texts produced by these individuals must typically be available for most computational methods to work well. This means that even if people go online and write posts regularly, there may still not be enough data to produce valid results. This limitation may be one reason why studies exploring individual-level dynamics are still relatively scarce. To avoid issues associated with having too little data at the individual level, previous research has typically studied aggregations of individuals that represent specific groups, comparing, e.g., the different emotional responses among high- and low-level emotional users on social media (Garcia and Rimé 2019), sentiments among political parties (Card et al. 2022), or pro-government and anti-government Chinese online communities (Stine and Agarwal 2020). Other scholars have suggested that large language models can be used to generate samples of texts and survey answers that represent different respondents in order to increase the number of observations (Argyle et al. 2023; Chu et al. 2023; Kim and Lee 2023).

To conclude, text data and the computational methods used to analyze them have great potential to contribute to analytical sociology and the literature on meaning and meaning-making. But the adoption of new methodologies within a discipline will require continued exploration in order to find natural synergies and develop best practices. The alliance between these data and methods and analytical sociology is still in its early stages, and the work presented in this thesis contributes only to inching somewhat closer to a solution to some of the common and specific challenges that need to be resolved in order to fully fulfill the promise of this alliance.

\(^{10}\)During the writing of this dissertation, digital data have in many cases become more difficult for researchers to obtain. One clear example is X (previously Twitter) closing its research API. Although it has become more challenging to obtain social media data, some platforms still allow for the collection of these data.
The Swedish immigration discourse

Immigration causes macro-level shifts in the ethnic composition of a population. At the same time, at the micro-level, it might change what your neighbors or colleagues look like, what types of food restaurants serve, or what languages you hear on the bus on the way to and from work. These types of large and small changes may all spark a reaction among the people who experience them. Over the last few decades, immigration has attracted an increasing amount of attention in Western European societies, including in the form of discussions of immigration among political parties and politicians (Green-Pedersen and Krogstrup 2008; Krzyżanowski 2018), journalists and the news media (Eberl et al. 2018), and citizens across Europe (Åkerlund 2020). The literature suggests that collective meaning-making intensifies in periods of high immigration, as different individuals and groups in society struggle to define what immigration is and how it impacts host societies (Eberl and Galyga 2021; Galyga and Lind 2021; Gessler and Hunger 2022). The immigration discourse constitutes an important testing ground for empirical studies of meaning-making based on text data at both the aggregate and individual levels. Using aggregate-level data focused on public culture, the empirical work presented in this thesis maps how the national media reporting on immigration has changed over the last seven decades, while individual-level data from social media are used to test the dynamics that unfold in online immigration debates following disruptive events.
The Swedish immigration discourse

The history of immigration to Sweden

During the 20th century, Sweden experienced a demographic transformation, with the population share of foreign-born individuals increasing from 0.7% in 1900 to 20% in 2021 (Statistics Sweden 2019; Statistics Sweden 2022b). Figure 4 shows the number of people who immigrated to Sweden between 1945 and 2021. The figure shows that immigration to Sweden has increased continuously since the end of World War II, with peaks in the early 1970s, around 1990, in the mid-2000s, and with an unprecedented number of immigrants arriving in the years leading up to and during the so-called European “refugee crisis.” However, the type of immigration represented by the numbers in Figure 4 has varied greatly over time, as have migration policies and the political discourse on migration. What follows is a decade-by-decade review of Swedish migration history starting in 1945.

![Figure 4: Immigration to Sweden 1945–2021, counts in the thousands.](image)

1945–1950: Following World War II, immigration to Sweden increased rapidly. First, refugees arrived from Sweden’s neighboring countries, which in contrast to Sweden had been directly affected by the war. Second, as rebounding industries cried out for more workers (Hansen 2001), labor immigration started to expand.

1950s: As the result of a law introduced in 1954, which formally created a single common labor market for the Nordic countries, most early labor immigrants came from neighboring countries, particularly Finland (Byström and Frohnert 2017). In the late 1950s, labor migrants increasingly also arrived from Yugoslavia, Greece, Italy, and Turkey.
1960s: During the 1960s, political attention slowly started to turn to the costs of immigration, resulting in a first set of immigrant policies in the late 1960s that regulated the conditions for immigrants and ethnic minorities in Sweden (Hansen 2001; Dahlström 2004). Politicians increasingly introduced legislation in 1967/1968 that discussed potential requirements for obtaining residence and work permits before entering the country (Lundh and Ohlsson 1999).

1970s: With the arrival of the first large groups of non-European refugees—mainly from Chile and Iran—and an increase in family-based immigration, the proportion of refugees among the immigrant population increased from 10% to 30% (Hansen 2001). However, most immigrants in Sweden at the time were still from the other Nordic countries (Byström and Frohnert 2017). The Swedish labor market contracted in the wake of the oil crises of 1971 and 1973, and the demand for foreign labor decreased. In response to the economic downturn, the strong labor unions adopted policies to restrict work visas, which slowed down labor migration considerably (Johansson 2008; Byström and Frohnert 2017).

1980s: The number of refugees continued to increase throughout the 1980s, increasing the workload of the agencies responsible for processing asylum applications (Byström and Frohnert 2017). The increased number of refugees led to the launch of a “Whole of Sweden” strategy (Hela Sverige-strategin), which aimed to distribute refugees more equally across the country. The strategy met with resistance in some municipalities, leading to local referendums about whether or not refugees should be allowed to settle there (Qvist 2012). Towards the end of the decade, refugee policies became more restrictive after the fall of the Berlin wall, as politicians feared an influx of refugees from the former Soviet Union. A decision to restrict immigration was taken on December 13th, 1989, which is therefore often referred to as the “Lucia decision” (Luciabeslutet). Some have described the Lucia Decision as marking the definitive end of a historical political consensus regarding immigration policy (Byström and Frohnert 2017).

1990s: The conservative refugee policies of the previous decade only lasted two years, and were followed by a large influx of refugees during the wars in Yugoslavia in the early 1990s. For the first time in history, an anti-immigration party, New Democracy (Ny Demokrati) gained seats in the national parliament (Eger and Valdez 2019; Dahlström 2004). Their entrance on the political stage further eroded the previously broad political
The Swedish immigration discourse

consensus in support of integration and multiculturalism (Demker and Malmström 1999). However, New Democracy’s presence in the national parliament only lasted for one term. In 1995, Sweden became a member of the European Union following a narrow victory for the yes-vote in a 1994 referendum (Gilljam et al. 1996).

2000s: EU membership came with common EU immigration reforms and free movement within the Schengen area in 2001, opening up the Swedish labor market to migrants from other EU countries (Swedish Migration Agency 2001; Andersson et al. 2010). Sweden continued to receive more asylum seekers per capita than any other EU country during the first decade of the 21st century (Eurostat 2021).

2010s: The largest immigrant group in Sweden during the 2010s comprised refugees from Syria. During the 2015 “refugee crisis,” Sweden received some of the highest numbers of asylum seekers per capita worldwide (Skodo 2018). The government halted the influx of refugees in mid-November by re-introducing border identity checkpoints (Peterson 2020). In 2016, to further reduce the number of arriving refugees, the government passed a more restrictive interim immigration law. The Sweden Democrats (Sverigedemokraterna), a populist-right party, entered the Swedish parliament in 2010 (Valmyndigheten 2010). All other political parties—irrespective of their placement on the traditional left-right scale—avoided the Sweden Democrats when they first entered parliament, treating them as a pariah (Lindvall et al. 2020). However, as the immigration question became increasingly important to Swedish voters (Martinsson and Andersson 2019), some other parties warmed to the Sweden Democrats. The 2010s have been described as the decade when Swedish “exceptionalism”—in terms of liberal migration policies and low levels of anti-immigrant attitudes—was transformed into a situation more typical of European countries in general (Demker and Odmalm 2022).

2020s: Following the continuation of restrictive immigration policies, family reunification became the most common migration type in the early 2020s, followed by labor migration and, lastly, asylum seekers (Fores 2022). With Russia starting a war with Ukraine in early 2022, there was an upsurge in the number of asylum seekers from Ukraine (Statistics Sweden 2022a). In the same year, the Sweden Democrats became the second largest party in the national election (Valmyndigheten 2022). With their support, the Moderate party, the Christian Democrats, and the Liberal Party formed
a new government, officially changing the previous status of the Sweden Democrats as a pariah.

**Actors in the Swedish immigration discourse**

Different social groups can attach different meanings to the same concept. Societal spheres, and the individuals within them, contribute to producing shared understandings of how immigration is perceived and talked about. This thesis focuses on two societal spheres: the traditional media and the online public who are active on social media.

**The traditional media**

The traditional media—i.e., outlets that existed before the internet, and most importantly for this thesis, the printed news media—are an integral part of the national conversation. The media highlight events, movements, and ideas that are currently on the agenda and decide what are the most important aspects of current affairs, thereby choosing the topics that are relevant to the public debate (Helbling 2014). The journalists who write the printed news curate the news flow by linking issues and events in an attempt to make sense of what is currently happening for their audience. The traditional media may also move beyond describing reality, and have the capacity to “make events” by assembling and promoting certain interpretations of current affairs (Molotch and Lester 1974). Media content is a unique source of information beyond personal experience about what is happening in the world (Czymara and Dochow 2018). As such, the media have a unique position in society, functioning as both a producer—that activates people’s discussions (King et al. 2017)—and an interpreter of reality at the national level. Moreover, the media play an essential role in both priming existing schemas and developing new schemas within individuals via selective re-telling (DiMaggio et al. 2013).

Previous empirical research related to meaning-making has predominantly drawn on theories of the power of the traditional news media to impact how others talk about events, groups, and concepts—for example, agenda-setting theory (McCombs and Shaw 1972) and framing theory (Goffman 1974; Entman 1993). The body of work on agenda-setting and framing explores the media’s ability to dictate what the public focuses on at a given time, and how well the media are able to disseminate interpretations of what is happening in the world (Czymara and Dochow 2018).
The Swedish immigration discourse

There has been an increase in the use of text data to study the agenda-setting and framing capacities of different actors in the public discourse, with a strong emphasis on elite actors such as the media and political parties (Helbling 2014; Eberl and Galyga 2021; Greussing and Boomgaarden 2017; Bleich and Veen 2021; Heidenreich et al. 2019; Czymara and Klinger 2022; Card et al. 2022; Webson et al. 2020). The combined findings of this literature clearly show that immigration has become a salient topic in the European media (Strömbäck et al. 2021) and that coverage of immigration is often negative and conflict-centered (Eberl et al. 2018). Moreover, previous studies on the Swedish media’s reporting on immigration have shown that immigrants are in general portrayed as victims, and that links are often made to racism (Hovden and Mjelde 2019). However, since the early 2000s, anti-immigrant sentiments have been becoming a normalized part of the media discourse (Berry et al. 2016). During the last decade, the focus has been directed at refugees and asylum seekers rather than labor migrants or other types of immigrants (Strömbäck et al. 2017).

Part of the existing literature has focused on media reporting during the European “refugee crisis” of 2015. The results of this literature are mixed. Some studies suggest that the Swedish media highlighted humanitarian issues and included more positive migration frames than the media in other European countries (Heidenreich et al. 2019). Others have found that humanitarian framings were less common than in other European countries and that the economic effects of refugee reception were particularly salient (Berry et al. 2016). Moreover, Sweden probably experienced a more drastic increase in negative sentiment in the news during the refugee crisis than other European countries (Eberl et al. 2019).

Besides shifts in the media imagery during periods of high immigration, interpretations of immigration may depend on differences in internal organizational factors and cultures at different news organizations (Achtenhagen and Raviola 2009). Previous studies suggest that newspapers with different political leanings and target audiences differ in how they portray immigration and immigrants (Theorin 2021; Ekman 2019).

Essay III adds to all of the literature mentioned above, as it tracks how the media have portrayed immigration from 1945 until 2019. This work complements much of the previous literature by providing a vibrant description of the framing of immigration in the Swedish media. Moreover, Essay III provides an overview of when the framing of immigration in the Swedish media has undergone substantial shifts, thus adding to previous historical accounts of important turning points in the Swedish immigration discourse. While we do not cover the tonality of news coverage, our results suggest that news reporting is closely related to the
type of immigration that is dominant at different points in time. We detect a trend towards immigration becoming politicized in national newspapers, perhaps indicating that a more diverse set of views about immigration has become part of the public discourse. Seeking to identify turning points in the way the media have framed immigration, the results show that while the refugee crisis did indeed increase the amount of media attention focused on the topic of immigration, it did not produce a shift in how immigration was framed. Finally, the results in Essay III show that the identified trends in the way the traditional media frame immigration are common to all four of Sweden’s largest newspapers. While the four newspapers represent a diverse set of political leanings, they cover the topic of immigration in very similar ways, and they have shifted the way they frame immigration almost simultaneously.

The online public

The second discursive sphere investigated in this thesis comprises the public as represented on online platforms. I use the active participants in online discussions as a social sensor of what the general public is currently talking about. In Sweden, roughly 89% of the population aged 16–85 use the internet on a daily basis (Statistics Sweden 2021). Many of these individuals are active on social media platforms that leave traces of their activities, interpretations, and feelings about current developments (Swedish Internet Foundation 2021). While the power of the traditional media over the public interpretation of current developments has long been a focus for research, more recent developments suggests that what people talk about online can also impact elite discourses (Russell Neuman et al. 2014; Jiang 2014; Ragas et al. 2014; Conway et al. 2015). This “reverse agenda-setting” (Russell Neuman et al. 2014) suggests that mapping the online discourse is important for studies of meaning-making.

Previous research that has studied the meaning of immigration among the European public has mainly focused on attitudes using survey data. Drawing on group threat theory (Blumer 1958; Quillian 1995), contact theory (Allport et al. 1954), or the salience hypothesis (Blalock 1967), many such studies have investigated the relationship between immigrant group size, local economic conditions, and the rise and fall of anti-immigration attitudes. Empirical studies have focused on these relationships across different cultural contexts (Hjerm 2007) and eras (Olzak and Shanahan 2014), and for various minority groups (Dixon 2006). This literature indicates that attitudes toward immigrants in developed economies range from reluctant acceptance to outright rejection (Ceobanu and Escandell 2010). Compared with other European countries,
The Swedish immigration discourse

Swedes support, Swedes typically report relatively positive attitudes toward immigration (Strömbäck et al. 2021). However, attitudes toward related issues such as multiculturalism are heavily skewed within the population, depending on political affiliation. In 2022, 71% of left-center voters reported being favorable toward multiculturalism, but only 28% of the right-conservative-liberal voters (Ahmadi and Cetrez 2022). Empirical studies linking attitudes on immigration and voting behavior also suggest that anti-immigration attitudes affect the propensity for voting for far-right populist parties (Halikiopoulou and Vlandas 2020; Stockemer et al. 2021). Essays III and IV draw inspiration from this literature to produce expectations regarding the different interpretations of immigration in the Swedish context.

One of Sweden’s largest online communities is an discussion forum named Flashback (Swedish Internet Foundation 2021). Previous research on Flashback has found that Muslims are discussed as an outgroup that is in conflict with Swedish culture and values (Törnberg and Törnberg 2016). Frames on Flashback during the “refugee crisis” have been suggested to have been more emotional than those found in the traditional news media during the same period (Yantseva 2020). Furthermore, Flashback plays a crucial role in relation to radical right-wing political violence, e.g., with posts encouraging the burning of refugee centers (Wahlström and Törnberg 2021).

While Flashback is infamous for having members who express anti-immigration sentiment, previous research has found that Muslims are also portrayed as a cultural threat on other important social media platforms in Sweden, such as Facebook (Merrill and Åkerlund 2018; Ekman 2019). Similarly, studies of the Swedish Twitterverse indicate that far-right discourse is (re)produced by influential users on the platform (Åkerlund 2020).

Essay IV adds to the empirical literature on the Swedish online population, yielding insights into the dynamics of online meaning-making. We test two mechanisms that may contribute to changes in discussions about immigration following unexpected events: (1) changes in how individuals talk about an issue and (2) changes in the composition of those participating in discussions. Using a seeded topic modeling approach, we show that immigration starts to be discussed more as a cultural and security issue than an economic one following Jihadist terrorist attacks. The quick adoption on the platform of a culturally focused discourse about immigration is in line with previous research showing that the Swedish online population primarily focuses on the impact of immigration on Swedish cultural values and norms. Moreover, we show that both within-individual and compositional changes are equally important for shaping immediate and long-term shifts in online discussions.
Data and ethical considerations

While the methods discussed in this thesis provide much promise for empirical research, the ethical risks associated with how they may allow researchers to extract potentially sensitive data warrant further discussion (Franzke et al. 2020).

The empirical work in this thesis relies on two primary data sources: the Swedish National Newspaper Corpus 1945–2019 and posts from the anonymous Swedish online discussion forum Flashback. In this section, I will discuss the ethical considerations associated with these two corpora based on three dimensions: the existence of human subjects in the data, anonymity, and platform terms of service and copyright. Traditionally, human subjects have been individuals who are represented in data, and who risk harm if these data are misused. In the era of “passive big data”—i.e., large-scale observational data—defining and protecting human subjects has been shown to be a non-trivial task (Metcalf and Crawford 2016) that requires thoughtful consideration of each data source used in empirical studies. Moreover, anonymity has long been one of the guiding principles of ethical research, constituting one of the primary safeguards against causing harm to human subjects (Mai 2016). Text data that can ensure human subjects’ anonymity are therefore preferable from an ethical perspective. Lastly, besides causing harm to human subjects, there is a legal aspect to the collection and use of data (Parry and Mauthner 2004), and I therefore also want to discuss the terms of service (TOS) associated with the data used in this thesis.

Swedish national newspaper corpus 1945–2019

The Swedish Nation Newspaper Corpus 1945-2019, digitized by the National Library of Sweden, consists of all articles printed in Sweden’s four largest newspapers: Aftonbladet, Dagens Nyheter, Expressen, and Svenska Dagbladet, between 1945 and 2019. These four newspapers represent a broad range of political leanings, from left-leaning (Aftonbladet) to moderate (Dagens Nyheter, Expressen) and right-leaning (Svenska Dagbladet). Moreover, Dagens Nyheter and Svenska Dagbladet are broadsheet newspapers, while Aftonbladet and Expressen are tabloids.

The raw corpus is processed to include only the content produced by journalists, for example, reporting on current events and opinion pieces. In contrast, other parts of the newspapers, such as commercial ads and crossword puzzles, have been removed. As such, the texts included in the corpus have been written by journalists in their professional role, and in the knowledge that the texts will be read by the public and that they will be archived. Before going to print, the texts have been edited and approved by the editor-in-chief to make sure that the
The Swedish immigration discourse

content is in alignment with the views of the newspaper. The processed corpus contains roughly 29.3M documents, i.e., newspaper texts written by journalists in one of the four newspapers, comprising almost 1.6B tokens.

Human subjects: In a newspaper corpus, the human subjects that may be at risk of harm can be either (1) the journalists who write the texts included in the corpus or (2) the people who are covered in the news. While journalists have written the texts used for analysis, the news content undergoes an editorial process that largely detaches any potentially sensitive opinions held by journalists from the textual content. Based on the same reasoning, there is likely to be no added risk of harm to other human subjects present in the texts, for example, people interviewed in the newspaper, since media houses have their own ethical standards focused on protecting these individuals in their reporting (Fairfield and Shtein 2014).

Anonymity: The human subjects at risk of harm in the newspaper corpus—journalists and people covered by the news—are mostly not anonymous. Most journalists who write texts in newspapers sign their articles with their names. However, in the newspaper corpus used in this thesis, there is no meta-data regarding which journalists have written specific news articles. This means that it is almost impossible to trace a text to a particular journalist, creating near anonymity for the human subjects who produced the analyzed material. Moreover, most people covered in the news are—or become—public figures, making it unlikely that any information available in the corpus risks harming them beyond the information that is already available to the public.

Platform terms of services and copyright: The issue of copyright infringement is highly relevant when using the Swedish National Newspaper Corpus 1945-2019, since the newspapers own their published material. However, to store its cultural heritage, Sweden has laws—dating back to the 17th century—that require the collection of all printed (and nowadays also electronic) material by the National Library of Sweden. Because a governmental institution stores the material—and is responsible for facilitating national computational research—it is considered to fall within the remit of the fair use for research doctrine (National Library of Sweden 2010). The use of this data still comes with stipulations, namely (1) original material cannot be included verbatim in publications, and (2) the data cannot leave the National Library of Sweden. Following these rules allows me to use the data without legal or ethical concerns.
Online discussion forum Flashback

Flashback is an online discussion forum that promotes itself as allowing “real freedom of speech” and that requires the total anonymity of its users. Flashback is an anonymous online environment with more than 1.4 million registered users, who write around 50 million posts per year. Little is known about the individuals who are active on the forum, although it is believed that males are over-represented (Westerlund et al. 2015). Information about themselves is one of the few things users cannot speak about on the forum. The forum also emphasizes its lack of interest in storing personal information about its users (Flashback Forum 2022). For this thesis, texts have been collected from the Politics subforum of the platform, a place where users self-select to discuss topics related to politics. The collected material contains the users’ views on various societal and political issues.

Human subjects: Research using digital trace data, such as the posts written on Flashback, could potentially be considered human-subject research. As posts online are written by individual users, who are non-public figures to a large extent, researchers could cause harm by extracting sensitive information about users’ political ideologies or interpretations of the world, whose use in research they have not actively approved. However, the anonymity of the users—imposed by the forum—makes it challenging to track the expressed opinions of specific physical individuals. In the absence of such mapping between users and physical human subjects, the risk of exposing sensitive information about individuals is limited.

Anonymity: Anonymity is an interesting and complicated issue when dealing with digital trace data. The users are, as previously discussed, anonymous. Their anonymity means that no meta-data that include personal information is available on the website. The only user identifiers are their chosen usernames. While usernames are meant to create anonymity, there are no rules against using the same pseudonyms across platforms, thus creating a possibility for the identification of individuals across platforms. To protect individuals’ anonymity, it is good practice not to include direct quotes from texts on the forum and to avoid making cross-platform comparisons of content.

Platform terms of services and copyright: Flashback does not have any openly available Terms of Service (TOS). However, general rules for how to behave on the platform are available. The rules of conduct prohibit
The Swedish immigration discourse

all attempts to unveil other users’ identities and to link avatars and pseudonyms on Flashback to other forums and social media platforms. The rules do not mention any issues regarding the use of the platform for research or collecting data.

Reinforcing privacy and preventing harm

The lack of human subjects at risk of harm in the newspaper material makes these data relatively straightforward to use from an ethical perspective. Research using newspaper texts may shed light on patterns in reporting that are nearly impossible for typical readers to recognize. Laying bare previously hidden patterns in news reporting, such as the prevalence of a specific frame when reporting on immigration, or the use of harmful stereotypes, might trigger hostile emotions towards newspapers or individual journalists. While this is a somewhat far-fetched concern, it is worth considering how research based on large-scale text analysis might result in new ethical problems.

Ethically, using a social media corpus is more complicated. Texts written online may contain potentially sensitive information regarding individuals’ political opinions, which they may not realize can be used for research. However, I do not have access to any personal information stored about the users, not even IP addresses, making it almost impossible to link “real-life individuals” to users on the platform. There may be a minimal risk that researchers deanonymize users on Flashback by comparing pseudonyms or usernames across platforms or via their network ties. Deanonymization may also occur when using techniques for authorship attribution. As such, reducing the risk of deanonymization is crucial, although no such efforts constitute part of my own research agenda. Risk reduction entails not publishing or sharing any posts at the individual level or tracking specific individuals over time.

Recently, some have begun to challenge the focus on protecting individual human subjects since “big data analytics” typically traces behaviors and opinions among groups rather than at the individual level (Mittelstadt 2017). Analysis of clusters of individuals commonly entails inferring individual-level characteristics, behaviors, or views from aggregate-level patterns among group members. In turn, this may lead to faulty assumptions being attached to a group, causing indirect harm to human subjects based on group membership. For example, inferring individuals’ behaviors or views based on a category that is created post hoc or on one’s position within a social network may result in biases and stereotypical estimations that do not capture the variance between individual identities. Making inferences based on group-level characteristics means that
even if data do not include human subjects, research may cause group-level harm. The idea of group-level harm makes it important to extend future discussions on human subjects and anonymity to groups when dealing with large-scale data.

To ensure that we would conduct all data collection and data use ethically, we applied for and were granted ethical approval from an Ethical Review Board.
Appended Essays

Essay I: From Documents to Data: a Framework for Total Corpus Quality

As large corpora of digitized text become increasingly available, researchers are rediscovering the potential of studying social and cultural phenomena using text data. While textual corpora promise to enrich our knowledge of the social world, the avoidance of problems related to data quality remains a challenge in this field of empirical research. Like a messy kitchen hidden behind a closed door while an impressive dinner is presented to guests, scientists have largely ignored potential issues with data quality, even as the body of empirical work using text data to describe and explain phenomena in the social world continues to grow. Hence, evaluating the quality of a corpus will become a pivotal issue for future social scientific inquiries. In Essay I, we propose a conceptual framework for total corpus quality that incorporates three crucial dimensions: total corpus error, corpus comparability, and corpus reproducibility. These dimensions impact the validity and reliability of inferences drawn from text data. In addition, our framework provides insights that will help in evaluating and improving studies based on large-scale textual analyses.

The contribution of this essay is two-fold. The paper provides researchers with a conceptual framework to discuss data quality and the potential trade-offs between different quality dimensions given a specific research question. Second, the paper provides practical examples of how researchers can quantify errors that
Appended Essays

may affect quality. It also discusses which error types are quantifiable and which are not, providing guidance to other researchers on how to understand both qualitative and quantitative dimensions that contribute to the overall evaluation of corpus quality. These two contributions will hopefully spur further discussion, and the development, of standards for evaluating whether a text data source is good enough to answer a specific research question.

Essay II: Interpretable Word Embeddings via Informative Priors

Essay II builds on a simple idea: we should be able to enhance the interpretability of the increasingly popular computational text analytical method family word embeddings, and thus their usefulness for social inquiry. Our paper relies heavily on previous work on probabilistic word embeddings, which can easily be extended to incorporate informative priors. Informative priors guide the model to recognize patterns in data that follow a predefined logic. In this paper, we test how well a set of models that rely on different informative prior specifications can capture two binary concepts: sentiment and gender.

The results of this methodological paper show that the use of informative priors can outperform other commonly used approaches—which rely on vector algebra—in identifying interpretable dimensions in the vector space identified by word embedding models. Our novel use of informative priors provides a new tool for inferring meaningful elements that relate to theoretical constructs of interest to social inquiry. For example, Yrjänäinen and Magnusson (2022) have already generalized our flexible model framework to study how the meaning of words and binary concepts develops over time and in different groups.


Essay III adds to the discussion of how sociologists can find a more formal way to measure meaning using text data, focusing on the publically available interpretations of current affairs produced by the national news media. We argue that there is great potential to measure meaning using seeded topic models. While seeded topic models have existed for many years, they have only rarely been used by sociologists and other social scientists. This paper focuses attention on their potential as a middle ground between unsupervised and supervised approaches. Seeded topic models increase the transparency and replicability of unsupervised models without losing their scalability and potential for inductively expanding our understanding of data. Seeded topic models allow researchers to *a priori*
indicate what topics the model should search for by providing the model with a set of seed words. Equipped with seed words supplied by the researcher, the model inductively generates topics that crystallize around concepts based on word co-occurrence.

Using newspaper texts written in Swedish national newspapers between 1945 and 2019, we map how the media have framed the topic of immigration over 75 years. We use a change point model to identify turning points in the media discourse on immigration and thereby inductively partition our extracted time series into stable eras of meaning. The eras identified by our inductive approach overlap with the partitioning of Swedish migration discourse typically employed by historians, with some clear exceptions. For example, while the 2015 “refugee crisis” indicated a watershed moment regarding the extent to which immigration held the attention of the Swedish public, it did not produce a shift in how the Swedish media portrayed immigration. Our combined results highlight the potential of using models that allow for augmenting social scientific knowledge in the modeling procedure.

**Essay IV: Mechanisms of change: what explains shifts in online immigration discourse after terror attacks?**

The central question addressed in Essay IV is: how do online discussions change in the wake of specific events? While the empirical literature based on digital trace data is growing, such studies often only map discursive trends at the aggregate level. In this paper, we test for the presence of two processes suggested by scholars who have studied long-term cultural change: within-individual and compositional change, to see whether they also can be used to explain the short-term and long-term dynamics that drive online discussions. Within-individual change involves shifts in the way people talk about a given concept. On the other hand, compositional change would suggest that aggregate-level changes in how immigration is discussed online are driven by who is participating in these discussions rather than by changes in the way individuals discuss a topic.

Using text data from Sweden’s most well-frequented online discussion forum, Flashback, we explore if and how Jihadist terrorist attacks change the content of online discussions about immigration. The results show that Flashback users start associating immigration more with culture- and security-focused themes in the wake of Jihadist terrorism, regardless of whether the attack occurred inside or outside Sweden. We also find evidence that within-individual changes are the main driver in producing the observed aggregate-level shifts, although compositional change also plays a crucial role. With this essay, we add to the
empirical literature that seeks to identify the processes that generate dramatic shifts in online discussions following widely recognized events. We also highlight the potential of text data as a promising new avenue for dissecting aggregate-level changes in online conversations and disentangling the individual-level processes that produce them.

Authorship contributions

**Essay I:** This essay was co-written with Sophie Mützel and Måns Magnusson. *Published in Socius.* Authorship contributions: (1) Conceptualization and idea development: MHB, SM, MM, (2) Data work and analysis: MHB, and (3) Writing: MHB (lead), SM, MM. All authors were actively involved in all decisions throughout the work.

**Essay II:** This essay was co-written with Martin Arvidsson and Måns Magnusson. *Published in the Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP).* Authorship contributions: (1) Conceptualization and idea development: MHB, MA, MM, (2) Data work and analysis: MHB, MA (lead), and (3) Writing: MHB, MA. All authors were actively involved in all decisions throughout the work.

**Essay III:** This essay was co-written with Marc Keuschnigg and Måns Magnusson. *Under review.* Authorship contributions: (1) Conceptualization and idea development: MHB, MK, MM, (2) Data work and analysis: MHB, and (3) Writing: MHB, MK. All authors were actively involved in all decisions throughout the work.

**Essay IV:** This essay was co-written with Anastasia Menshikova. *Unpublished manuscript.* Authorship contributions: (1) Conceptualization and idea development: MHB and AM, (2) Data work and analysis: MHB (lead), AM, and (3) Writing: MHB and AM. Both authors were actively involved in all decisions throughout the work.
Discussion and Conclusions

Analytical sociology and other social sciences are experiencing a computational turn with a newfound wealth of available data sources, the methodological advances needed to model this data, and the computational power to do so at scale. This thesis has been written in the midst of these developments, and focuses specifically on the promises and pitfalls associated with using text data and the methods needed to analyze these data. The aims of the thesis are three-fold. The first aim reflects the need to take a step back from the initial excitement of having an apparently incredible new source for social inquiry, and to ask and answer fundamental questions about how we can ensure that text data allow us to draw valid conclusions about the social world. The second aim is also methodological, and involves exploring how sociologists can adapt the algorithms developed in other fields to become more attuned to the central demands of social science research, where interpretability constitutes a central focus. The third aim is more specific and focuses on the use of computational text analysis to extract measures of the meaning of immigration within the Swedish public discourse, and to study when and how the meaning of immigration has shifted. By measuring the ways people have talked about immigration, we can better understand how different segments of the Swedish population have understood both long-term and short-term changes related to, for example, the demographic transformation that the country has undergone over recent decades and the impact of unexpected events such as terrorist attacks.
Discussion and Conclusions

Essay I focuses solely on the first of these three aims, and suggests that the existing empirical literature that has used text-as-data has mostly ignored the fact that text data—like all other forms of data—can be erroneous. Data errors can impact the bias or variance of any estimates extracted from the data in question. This means that empirical work using text as data will suffer from "garbage-in, garbage-out" (Geiger et al. 2020) problems, regardless of the extent to which data grows or how fast new state-of-the-art methods are developed. Essay I provides a conceptual framework that can be used by social scientists in thinking about how choices that are made in the research process, from data collection to the final analysis, can impact the results. Before starting with the empirical work presented in Essays III and IV, we evaluated the corpus in accordance with the framework presented in Essay I. Learning about the strengths and weaknesses of the Swedish National Newspaper Corpus 1945-2019 and the corpus of posts from Flashback has been an essential basis for deciding what types of analysis were appropriate given the data. While Essay I discusses what error types may be present in a specific corpus, and also dimensions of data quality, it leaves one crucial question unanswered: how should insufficient data quality best be remedied? More work is needed to develop best practices to improve data quality. One natural starting point for future work would be to draw inspiration from the work that has been conducted to remedy quality issues in survey data.

Essay II contributes mainly to the second aim of the thesis by providing a simple extension to the standard word embedding model to increase its interpretability and usefulness in testing sociological theory. Giving sociologists the possibility to focus in on a particular predefined element of meaning makes it easier to learn about when and why cultural associations appear and change. I believe that the new method proposed in Essay II to measure specific binary cultural associations holds great potential as an approach to sociological inquiry. However, using the method proposed in Essay II in subsequent empirical projects, I have also identified some limitations of our proposed model. Most importantly, many concepts in sociology are not easily reduced to binary dimensions, for example, status, race, and class. Extending the model to also identify non-binary concepts would further increase the applicability of the methodology in empirical research. Moreover, others have started to explore the potential of using word embedding models to learn about intersectionality (Nelson 2021). By incorporating priors to estimate multiple interpretable dimensions, the method presented in Essay II may be of great use to this emerging field. Essays III and IV also show the potential for using intrinsically interpretable methodologies in empirical research by using seeded topic models, an extension of the traditionally inductive LDA topic model that leads to part of the output being estimated using
a more deductive approach that utilizes informative priors. By using intrinsically interpretable extensions of an unsupervised methodology, the two empirical essays presented in this thesis highlight the usefulness of using sociological theory and knowledge to augment the modeling procedure rather than relying only on post hoc theorizing.

While this thesis argues that intrinsically interpretable models should be used more in research using text as data, more empirical sociological studies are instead moving towards the use of "black-box" large language models. While uncannily impressive at generating textual content and detecting patterns in text data, spurring discussions on the transformative power of these methods both inside and outside academia (Thorpe 2023; King and ChatGPT 2023; Bubeck et al. 2023; Ziem et al. 2024), I believe it is important to remember why and how we use such methods. In conferences, seminars, and discussions with other researchers, I have noticed a tendency for people to ask questions of the kind “Why are you not using the newer method X? Have you heard about the new method Y?” People probably raise such questions because methodological developments are occurring so rapidly, and we should always consider why we choose one method over another. However, I believe that this question also captures a cultural schism between the natural language processing community and the social sciences. Instead of focusing on the substantive research question of interest, the focus has shifted to using the newest method. As computational text analysis matures as a discipline, I hope that sociologists will remain focused on making their choice of method on the basis of the ability of alternative methods to answer social inquiries, rather than their novelty. Moreover, as large language models continue to make their way into empirical research, my hope is that sociologists will continue to focus on issues related to reliability, transparency, model evaluation, and the incorporation of theory. Concentrating on the sociological questions of interest rather than the need to adopt the ever-growing toolbox of the NLP community will spur more work in the direction outlined above. Or stated more sharply, sociologists should remember that “a method is only as good as the theoretical insights it is able to generate for the purposes of advancing scientific knowledge” (Bonikowski and Nelson 2022).

The final aim of this thesis, that of using text data to explore aggregate- and individual-level meaning-making in relation to immigration in Sweden, constitutes the focus of Essays III and IV. Together, these two essays add to our understanding of how the ways different actors in the Swedish immigration discourse talk about immigration are formed and have changed. Essay III shows that the media react to real-life developments in the field of immigration, framing the issue differently over time. In the aftermath of World War II, the dominant
media framing of immigration viewed the issue as a human rights-related problem. The traditional media then slowly became more focused on framing immigration on the basis of an economic perspective during the 1950s and 1960s, when labor migration increased rapidly. Similarly, we can see that immigration then started to be framed as a political issue, signaling the way immigration began to become politicized in the 1970s. Today, the framing of immigration as a political issue has come to predominate in the traditional media. We also find that single events, such as terrorist attacks, only shift how the media report on immigration immediately following the event. The results of Essay IV corroborate this finding. Our results suggest that online discussions about immigration mostly returned to their pre-event state within a week following the Jihadist terrorist attacks. However, culture-centered interpretations of immigration remain more salient than they were before the attacks for several weeks. Taken together, the results of Essay IV suggest that current members of the Swedish public are quick to start talking about immigration as a cultural issue, and that they continue to focus on topics such as religion and national identity when discussing immigration for much longer than they do on alternative ways of interpreting immigration. In combination, the findings of Essays III and IV paint a picture that corroborates previous findings showing that “Swedish exceptionalism,” i.e., the societal dominance of a strongly humanitarian-focused view of migration, is a thing of the past (Krzyżanowski 2018; Rydgren and Van der Meiden 2019).

The results of Essay IV also relate to discussions about the mechanisms that might produce aggregate-level changes in the way topics are discussed and understood by members of the public. Previous research in cultural sociology has found that people's private culture—such as attitudes on politicized issues such as immigration—are relatively stable over time, or at least that persistent changes are rare (Vaisey and Lizardo 2016; Kiley and Vaisey 2020; Underwood et al. 2022). This would suggest little change over time in the way the same individual talks about immigration. However, recent empirical evidence has shown that significant life course transitions, such as entering parenthood, can result in persistent shifts in private culture (Lersch 2023). Rather than focusing on individual-level life course transitions, Essay IV explores how unexpected events impact whether and how individuals discuss immigration. The results suggest that individuals do indeed change the way they discuss immigration online in the wake of Jihadist terrorist events, focusing more on cultural and security issues rather than economic issues. However, the findings indicate that such individual-level changes are not sufficient to completely account for aggregate-level shifts. The new users who join online discussions of immigration following Jihadist terrorist attacks are also crucial in shaping post-event conversations.
The results of Essay IV thus highlight how shifts in public discourses about immigration can be brought about by multiple individual-level mechanisms that work together. More studies exploring the mechanisms that bring about aggregate-level changes in social media behavior are needed to understand the conditions under which different mechanisms dominate. Such studies could also help to guide policy-makers in understanding how policy impacts reactions among the public. If interpretations about the world are sensitive to unexpected events, policies intended to improve public perceptions of immigration can be undermined by forces that are beyond political control, such as terrorist attacks.

Because Essays III and IV study similar phenomena among different social actors, it is helpful to consider them in relation to each other. These essays show how both the traditional media and the general public react quickly and substantially to what is happening in the world around them. Essay IV shows that economic themes dominated online discussions of immigration in uneventful times, whereas security- and culture-related themes became more salient following unexpected events. However, Essay III shows that over the same period of time, the salience of the economic framing of immigration in the traditional news media has been in decline, with more space being given to frames related to security and political issues. In the traditional news media, content only gets published if it meets with an editor’s factual and ethical approval. At the same time, posts written on online forums such as Flashback remain mostly unmoderated. More research is needed to disentangle whether the difference in theme salience between the two forms of media is due to differences in worldviews and the types of structural constraints and possibilities mentioned above.

While Essays III and IV look at how two parts of society react to events, they do not disentangle the dynamics that are in operation between these different social spheres. Does the public use news coverage when forming their interpretations of what is happening, or do the media pick up what the public is discussing and amplify already emerging meanings? For example, the results in Essay IV indicate that the online public adopts a culture-focused way of discussing immigration in the wake of Jihadist terrorist attacks. At the same time, the supplementary event analysis in Essay III suggests that the traditional news media do not start to frame immigration through a more cultural lens than that used before the attacks. If it is not the traditional media, what is it that influences the online discourse to become more dominated by cultural themes? More studies into the causal interdependencies that affect the ways different groups in society discuss immigration are needed to enrich our understanding of the dynamics of meaning-making.
Further, the way members of different social groups discuss immigration might depend on their latent perceptions of immigrants (Gorodzeisky and Semyonov 2009; Hellwig and Sinno 2017; Gorodzeisky and Semyonov 2019), sometimes referred to as “imagined immigration” (Blinder 2015). The concept of imagined immigration suggests that different people consider different types of immigrants when they think about immigration. For example, some may have expatriates with high-paid jobs in mind when discussing immigration, whereas others may primarily be thinking about refugees—these differences in imagination result in people expressing different attitudes or debating the topic differently. For example, Yantseva (2023) has found that immigrants are more associated with costs and threats to society than refugees using text collected from Swedish Facebook. The empirical essays in this thesis largely ignore potential issues related to imagined immigration by studying the meaning of immigration as a single concept rather than exploring the importance attached to specific migrant types, such as refugees, labor migrants, asylum seekers, or minority groups such as Muslims, the Sámi people, or Eastern Europeans. I nonetheless view the aggregate measures used in this thesis as a valuable starting point for testing how computational text analysis can be used in empirical studies focused on the meanings of immigration.

The empirical essays presented in this thesis open up for discussions regarding the extent to which the way people talk about something reflects the meaning they attach to the object of debate. To answer these questions, more work is needed that explores the alignment between how people express themselves and how they “truly” understand the world. One step in this direction would be to combine data collected from the same individuals in different settings, i.e., via surveys, implicit association tests, neural images, and texts written in a natural context. A data collection effort of this kind would allow researchers to find what meanings people attach to objects or events under different conditions. Some promising work towards understanding the relationship between values and perceptions on the basis of research that pairs survey data with digital trace data has recently begun to develop (Lu et al. 2024).

After writing this thesis, my answer to the question of whether text and computational text analysis has a future in analytical sociology is an unequivocal yes. While none of the essays in this thesis live up to the analytical sociologist’s dream of mapping out a macro-micro-macro process and fully identifying the mechanisms that produce a social phenomenon, I have taken a number of crucial steps towards this goal. Firstly, by focusing on the fundamental question of data quality and interpretable modeling, empirical research using text data can move closer to living up to the analytical sociological ideal of rigor and
realism in explanation. Furthermore, Essay IV showcases the potential of using measures of people’s private culture to explain behaviors that produce aggregate-level outcomes. A move towards increasingly studying meaning-making using individual-level text data is crucial to demonstrating the continued promise of this novel data source. Increasing the use of individual-level text data would allow analytical sociologists to study the meaning-to-action link in more depth, an essential step in the formulation of mechanistic explanations of social phenomena. For example, the insights from the research presented in this thesis could inform future work focused on combining expressions of private culture with behavioral data, linking people’s perceptions of specific neighborhoods extracted from text data to the housing decisions that generate residential segregation. The work in this thesis should be viewed as a point of departure for future studies by analytical sociologists using text as data.


Bibliography


issue attention and agenda setting by legislators and the mass public using social media data.” In: American Political Science Review 113.4, pp. 883–901.


Bibliography


Bibliography


63
Bibliography


Bibliography


Galyga, Sebastian and Fabienne Lind (2021). “Linguistic features of migration coverage in European mass media.” In: *Media and Public Attitudes Toward*


Bibliography


Bibliography

Hopkins, Daniel and Gary King (2007). Extracting systematic social science meaning from text. url: https://tinyurl.com/5a953s56.


Bibliography


Nelson, Laura K (2019). “To measure meaning in big data, don’t give me a map, give me transparency and reproducibility.” In: *Sociological Methodology* 49.1, pp. 139–143.

Bibliography


the far right electorate.” In: Journal of Ethnic and Migration Studies 47.15, pp. 3409–3436.


76


Bibliography

Essays

The essays associated with this thesis have been removed for copyright reasons. For more details about these see:

https://doi.org/10.3384/9789180756181
Mining for Meaning: using computational text analysis for social inquiry

Miriam Hurtado Bodell