



Funded by
the European Union



UK Research
and Innovation

Deliverable number: D2.2



hybrids

Technical report on the State of the Art of Political Communication in Social Networks

A Comprehensive Survey of different Aspects of Political
Communication and its Effects on Social networks

Version 1.



Project Details

Project Acronym:	HYBRIDS
Project Title:	Hybrid Intelligence to monitor, promote and analyse transformations in good democracy practices
Grant Number:	101073351
Call	HORIZON-MSCA-2021-DN-01
Topic:	HORIZON-MSCA-2021-DN-01-01
Type of Action:	HORIZON-TMA-MSCA-DN
Project website:	https://hybridsproject.eu/
Coordinator	Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS)- Universidade de Santiago de Compostela (USC)
Main scientific representative:	Prof. Pablo Gamallo Otero, pablo.gamallo@usc.es
E-mail:	citius.kmt@usc.es , info@hybridsproject.eu
Phone:	+34 881 816 414

Deliverables Details

Number:	D2.2
Title:	Technical report on the state of the art on NLP applied to discourse analysis
Work Package	WP2: Public Discourse Analysis
Lead beneficiary:	FBK
Deliverable nature:	R- Document, report
Dissemination level:	PU-Public
Due Date (month):	30/04/2024 (M16)
Submission Date (month):	30/04/2024 (M16)
Keywords:	Social network analysis,

Abstract

D2.2. Technical report on the State of the Art of Political Communication in Social Networks provides an overview of different aspects of political communication on social networks. Specifically, it covers five topics of major interest across social network analysis and political discourse analysis: network representation of political narratives, political opinion and participation in social networks, implicit communication in political discourse on social media, mainly in negative terms, populist communication online and conspiracy theories. This document complements the content of D2.1. on “State of the art of NLP and AI methods for discourse analysis in the political domain”. Furthermore, it represents the theoretical background upon which HYBRIDS DCs’ activities will be based and that will inform the implementation of new tools and approaches for the analysis of political discourse online.

Deliverable Contributors

	Name	Institution	E-mail
Deliverable leader	Sara Tonelli	FBK	satonelli@fbk.eu
Contributing Authors	Martial Pastor	RU	martial.pastor@ru.nl
	Davide Bassi	CiTIUS-USC	davide.bassi@usc.es
	Siddharth Bhargava	FBK	sbhargava@fbk.eu
	Erik Marino	UEVORA	erik.marino@uevora.pt
	Katarina Laken	FBK	alaken@fbk.eu

History of Changes

Version	Date	Changes to previous version	Status
0.1	09/04/2024	First Draft	Draft
0.2	20/04/2024	Editing and merging contributions	Almost final
1.0	30/04/2024	Approved version to be submitted	Final

List of Acronyms

CMQ	Conspiracy mentality questionnaire
DC	Doctoral candidate
LDA	Latent Dirichlet allocation
LLMs	Large language models
ML	Machine learning
NER	Named entity recognition
NLP	Natural language processing
POS	Part of speech
SA	Speech act
SN	Social networks

Contents

1	Introduction	6
2	Network representations of political narratives and political discourse	6
2.1	Narrative graphs	9
2.1.1	Event graphs	11
2.2	Modeling political discourse	13
2.3	Graph Evaluation	14
2.4	Findings	15
2.5	Challenges and future research directions	16
3	Political Opinions and Public Participation on Social Networks	17
3.1	Understanding Political Opinions, Attitudes and Ideologies	18
3.1.1	Shaping Opinions and Ideologies	19
3.1.2	Political Alignment: A Brief Case Study	19
3.2	Role of Social Media in Political Participation	20
3.2.1	Levels of Political Participation	21
3.2.2	Moderating Political Participation	22
3.2.3	Electoral Campaigning and Social Media: A Brief Case Study	23
3.3	Studying Political Opinions and Participation	24
3.4	Challenges and future research directions	25
4	Rhetorics of Implicit Communication in political discourse on social media	26
4.1	Derailed Reflexivity	27
4.2	Implicit Communication	27
4.2.1	Unsaid/Unspoken Statements	27
4.2.2	Indirect Speech Acts	28
4.2.3	Enthymematic Communication	30
4.2.4	Stylistic analysis of Implicit communication	31
4.2.5	Parataxis	32
4.3	Challenges and future research directions	33
5	Populist Communication and its Community Impact	35
5.1	Populism as an Ideology	36
5.1.1	Case Study: From the Ideology to the Index	38
5.2	Populism as a Communicative Style	41
5.2.1	Case Study: The Populist Stylistic Toolkit	42
5.3	Populist Communication and Social Networks: a Vicious Circle	43
5.4	Challenges and Future research directions	46
6	The Role and Impact of Conspiracy Theories in Political Communication on Social Media	48
6.1	Theoretical Framework	48
6.2	Psychological Explanations: Cognitive biases	49

6.3 Social Media Addiction and Its Impact on Mental Health	51
6.4 Spread Mechanisms: Echo Chambers, Filter Bubbles, Machine Habitus and other alternative explanations	54
6.5 Impact of Conspiracies on Voting Behavior	55
6.6 How to measure the belief in Conspiracy Theories?	57
6.7 Case Study	58
6.8 Challenges and Future research directions	59
7 Conclusions	60

1 Introduction

One of the key aspects of political communication is *relationality* (Lejano, 2021) which typically involves all policy actors including citizens. And since political communication is no longer simply about territorial and image-based electoral strategy, but it is increasingly embracing digital media and in particular social media, this relationality often manifests itself in online interactions. In this document we therefore describe how political communication is shaped in social networks, what communication strategies it implies and what are the risks associated with it. In Section 2 we first introduce the notion of network, which can be modeled as a graph, with political actors or events being represented as nodes. The section then discusses how this formal representation can be applied to study interactions on social media, with some examples like narratives of conspiracy stories spreading online.

In Section 3 we present research on political opinions and public participation online, highlighting differences between different levels of participation and discussing the need for moderation. In Section 4 we focus on *implicit communication* in political discourse on social media, meant as a way to exploit ambiguity to subtly influence public opinion. Among the linguistic devices used in implicit communication we include unspoken statements, indirect speech acts and enthymematic communication. Finally, in Section 5 and 6 we discuss two risks associated with political discourse in social media: first, *populist communication*, which is amplified online and is expressed not only in terms of topics but also through a specific communicative style, Second, *conspiracy theories*, which take advantage of spreading mechanisms typical of social networks such as echo chambers and filter bubbles and can even influence voting behaviour.

Overall, this document presents a survey of the state of the art in political communication in online interactions, starting from a more theoretical perspective (e.g. how graphs can be modeled) to very specific risks associated with it, such as populism and conspiracy theories.

2 Network representations of political narratives and political discourse

This chapter provides an overview of research based on network representations of political narratives and discourse. A network (also called *graph*; the two terms will be used interchangeably

throughout this chapter) is a finite, nonempty set of elements called *nodes*¹, and a set of *edges*; each edge is a two element subset of the set of elements (Gould, 2012). In other words, each edge denotes a connection between two elements of the set of nodes. Edges and nodes can have further attributes such as types (edges/nodes), directions, and weights (edges).

There are several reasons why graphs are especially useful to represent political discourse. First of all, networks inherently express the idea of *relationality*, which is a key aspect of political narratives (Muller, 2014). Political narratives concern the representation of groups of people and their relations to each other (Piper et al., 2021). Moreover, political discourse is inherently driven by debate and actors having different motivations and goals. A network makes these relationships visible and makes it easy to identify communities (groups of nodes that are more related to each other than to the other nodes in the network), central nodes, and bridging nodes (nodes that connect a lot of nodes that otherwise would not have a connection to each other).

Representing discourse and/or narratives as a network allows researchers to quantify above-mentioned characteristics. Yousefi Nooraie et al. (2020) outline how social network analysis is an especially useful tool for approaches that not just mix quantitative and qualitative methods, but even fuse them. It allows researchers to make substantiated observations about the discourse as a whole and use qualitative, theory-driven approaches to drive the interpretation and choice of the metrics used. It can thus be implemented as a complementary computational approach to human close readings of text (Mohr et al., 2013).

Another advantage of networks is that they provide a single representation for information that is often spread out over hundreds or thousands of documents (articles, messages, posts, etc.) (Jing and Ahn, 2021). Single posts, news articles or messages on social media do usually not repeat a full narrative. This is especially the case of social media, where single posts rarely repeat the whole narrative, or when studying phenomena like conspiracy theories (Tangherlini, 2017). A storytelling event usually just repeats (and thereby reinforces) a small part of the narrative. Or, as Tangherlini et al. (2020) puts it, 'activates a subgraph comprising a selection of actants (nodes) and relationships (edges) from the narrative framework'. So in order to reconstruct the whole narrative, it is necessary to look at a large body of storytelling events (i.e. tweets, blogs, debates, posts) and aggregate the information that they convey. Moreover, network representations allow for a visual representation of narrative structure or debated topics, synthesizing information from lots of sources and making it insightful to both lay people and experts (Keith Norambuena and Mitra, 2021).

'Network approaches to political narratives and discourse' is not a singular, well-defined research field. Rather, researchers from very different fields and backgrounds have (seemingly largely independently) discovered that representing discourse as graphs is a fruitful avenue of research. This makes this topic extremely interesting and versatile, but it often makes it difficult to compare papers to each other. The literature shows a wide variety of types of networks employed to represent political discourse. Table 2 shows an overview of the papers cited in this research and what narrative and/or discourse elements/participants make up the edges and nodes in their graphs.

Narrative theory often emphasizes the different levels on which a narrative can be understood: story-internally or as embedded in a socio-cultural context and tradition (Piper et al., 2021;

¹or *vertices*

Paper	Data	Topic	Nodes	Edges
Abzianidze (2020)	Georgian newspaper articles	Nationalism	Actors sending or receiving nationalist discourse	n statements targeting each other
Ansah et al. (2019)	Twitter	Black Lives Matter	Events	An event evolution in the story
Ash et al. (2024)	US congress speeches	General	Coherent entity groups	Semantic relations
Bazoozbandi (2023)	Iranian political speeches	Authoritarian government discourse	Keywords	Co-occurrences
Bhattacharya (2020)	Official documents from German parliament	Domain-general	Politicians	(dis-)agreement
Botzer and Weninger (2023)	Reddit	Different subreddits discussing politics	Entities	Co-occurrences in a thread structure
Chen and Dluhošová (2023)	Historic Taiwanese media	Taiwanese ideologies	Keywords	Co-occurrences
Chong et al. (2021)	Various forums and SM	Conspiracy theories	Actants (linked entities)	Directed semantic relations
Introne et al. (2020)	Various forums	Vaccines	Narrative elements	Relations
Keith Norambuena and Mitra (2021)	Newspaper headlines	Covid-19	Events	Narrative coherence between events
Keith Norambuena et al. (2022)	Reddit	Cuba protests	Events	Likelihood of belonging to the same storyline
Khalili-Mahani et al. (2022)	Social media and news media	Elderly people & Covid-stress	Topics	Co-occurrences
Koch et al. (2023)	Ethnographic data from local politics	German biodiversity	Actors	Interactions
Larkins et al. (2024)	Ethnographic data	One Health project in Laos	Actors involved in the project	Interactions
Lauran et al. (2020)	Twitter	Scandal in Dutch agriculture	Twitter users	Interactions
Li et al. (2021)	News documents		Events	Temporal order, shared arguments, related arguments
Miani et al. (2022)	Text from web-pages	Conspiracies	Seeds, keywords, topics	Co-occurrences
Mohr et al. (2013)	Government documents	US National Security Strategies	Entities	Directed semantic relations
Muller (2014)	In-person political debate	Flemish internal politics	Debators	Degree of agreement
Pereira-Kohatsu et al. (2019)	Twitter	Hate speech	Users	Interactions
Segev (2020)	News documents; Twitter	Fake news; China	Words	Co-occurrences
Tangherlini et al. (2020)	Blog posts and news articles	Conspiracies	Keywords	Co-occurrences
Urman and Katz (2022)	Telegram	Far-right networks	Actants	Context-specific relationships
Yan and Tang (2023) https://hybridsproject.eu/	News documents	South China Sea	Telegram channels	Mentions
			Events	Temporal and thematic relations

Table 1: Table showing the types of networks built in the papers discussed in this section

Barthes and Duisit, 1975). Broadly speaking, this corresponds to the two main categories of networks used to represent political discourse (Botzer and Weninger, 2023), as showcased in Table 2. One line of inquiry aims to understand debates and policy decisions by representing stakeholders and debaters as nodes. Edges of such a network are often overlapping viewpoint, mentions, or citations (for example on Twitter). This approach, called *Discourse Network Analysis*, analyzes the discourse on the level of the people that participate in it. It is used to describe political debate in traditional, face-to-face scenarios (Koch et al., 2023), but when computationally constructed, it is especially useful to map debate on social media (Lauran et al., 2020; Urman and Katz, 2022). A network can also represent the narrative itself. These approaches often use the nodes to represent entities in the narrative; the edges are either relations between them (usually extracted using various parsing and semantic role labeling techniques from the standard NLP toolkit) or simple text co-occurrences. Another type of narrative graphs focuses on events. These types of graphs aim to synthesize information from a large quantity of news documents to extract a story graph, showing the key events in a coherent narrative (Keith Norambuena and Mitra, 2021; Yan and Tang, 2023; Ansah et al., 2019).

Although this document focuses on political discourse on social media, this first section broadens the scope of analysis to also include papers that deal with other types of political discourse when they represent interesting theoretical contributions to the topic. Also, we aim at describing the work that has been done on all platforms and modalities by researchers from different fields and backgrounds (from computer science to political science to linguistics), to showcase the full diversity of this line of research. Moreover, we adopt a broad definition of the construct of 'politics' that goes beyond state-centered politics. In particular, we follow the definition of 'politics' as formulated by Warren (1999): 'the subset of social relations characterized by conflict over goods in the face of pressure to associate for collective action, where at least one party to the conflict seeks collectively binding decisions and seeks to sanction decisions by means of power' (p. 218).

2.1 Narrative graphs

Narrative graphs are network representations of political narratives that focus on the *content* of the analyzed text, rather than the participants of the discourse. An influential article by Patterson and Monroe (1998) posits that 'narrative' refers to 'the ways in which we construct disparate facts in our own worlds and weave them together cognitively in order to make sense of our reality' (p. 315). Modern political science is placing increasingly more emphasis on the role of narratives as an important driving force in political discourse, as they are able to mobilize masses and push them to political action (Gabriel, 2021). Social media debates, where people constantly express their opinions, debate each other, and share ideas, become a primary space where political debate takes place.

Political discussions on social media often shape popular narratives, usually living in collective imagination. The actants are typically selected from a cultural 'pool' of actants, operating on stereotypes (Tangherlini et al., 2020). Discourse on social media also shapes narratives by setting the agenda: the mere fact that people are talking about something makes it more relevant to the political debates (Lauran et al., 2020).

Many studies aiming to create narrative graphs draw on literature on folklore and narrative

theory. Many theories of narrative posit that narratives consist of entities, their actions, and some context, although the exact definitions vary. For example, [Burke \(1945\)](#) defines narratives as consisting of actors, acts, and a scene, which is the basis for the framework implemented by [Mohr et al. \(2013\)](#). Similarly, [Tangherlini \(2017\)](#) and [Tangherlini et al. \(2020\)](#) use [Greimas \(1983\)](#)'s theory of narrative as a starting point, which posits that a narrative has actants, relationships between them, and a sequencing of those. Policy narratives have a more narrow definition. [Patterson and Monroe \(1998\)](#) note that political narratives necessarily revolve around human actors playing an active role in the story. [Jones and McBeth \(2010\)](#) define a policy narrative as having a context, a temporal and causal driving force, characters who are either heroes, villains, or victims, and a moral. Different approaches have different ways of reflecting these base elements of narratives in the narrative graph. We can roughly distinguish three categories of narrative graph, based on what their nodes represent: entities, keywords, or events.

One line of research focuses on *textual network analysis* or *quantitative narrative analysis* ([Chen and Dluhošová, 2023](#); [Bazoobandi, 2023](#); [Segev, 2020](#)). These approaches focus on keywords find their roots in poetics (the study of narratives) and discourse analysis. Computationally, they are often very simple; the focus is typically not on creating new, cutting-edge algorithms, but rather on using network approaches to draw conclusions that will be of interest to political scientists. The interpretation of the network in these cases is mainly a manual endeavour. An example is [Segev \(2020\)](#), who implements a computationally light method to model keyword co-occurrences. He shows how this approach is easily implemented to gain insight in both Twitter discourse and newspaper framing.

[Bazoobandi \(2023\)](#) analyzes speeches by Iran's Supreme Leader in order to test hypotheses about the rhetoric of authoritarian leaders. [Chen and Dluhošová \(2023\)](#) create a co-occurrence network of keywords to study the discourse of various political fractions in post-war Taiwan. They study the development of the semantic meaning of ideologically loaded words in itself, in order to discover which semantic fields exist and which political groups propagate them. Compared to [Bazoobandi \(2023\)](#), their approach is much more computationally complex, both for the way in which they identify keywords and concepts and for the clustering algorithms used to discover topics.

[Pereira-Kohatsu et al. \(2019\)](#) use the output of a hate speech detection classifier as the basis for a network visualization of terms occurring in tweets containing hate speech². They then employ a community detection algorithm to find clusters of terms. The authors just aim to create a tool for researchers studying hate speech; they do not go into the manual interpretation of their network themselves. They also create a graph plotting hateful users, where the nodes are users (red for senders of hate, blue for receivers) and the edges are constituted by the hateful interactions between them.

All narrative theories emphasize the importance of actants/actors for a narrative. Reflecting this, many papers use nodes to represent entities. The simpler approaches use named entity recognition to find relevant entities, and model the relationships between them as simple co-occurrences ([Botzer and Weninger, 2023](#)); more complex studies have nodes represent actants (entities occupying a semantic role in the narrative), and have the edges represent the relations

²note that hate speech is inherently political, as it targets people based on their membership of a (minoritized) social group

between them (Chong et al., 2021; Tangherlini, 2017; Ash et al., 2024). These papers usually make use of off-the-shelf named entity recognition algorithms. An exception is the work by Introne et al. (2020), who manually code forum posts on vaccines. They manually identify entities, events and actions, which are mapped to their narrative role in the story (of, for example, consequence, target, or actor). The entities for each role are then aggregated in a tree-like graph structure, where the nodes represent 'classes' of entities, actions or events and the edges the set/subset relationships between them.

Several papers use algorithms to automatically extract complex grammatical relations. An often-used technique is semantic role labeling (SRL) (Ash et al., 2024). SRL algorithms aim to identify spans in a clause that play semantic roles such as agent, patient, and actions (see Gildea and Jurafsky (2002)). Automatically extracting these roles allows researchers to create a network where the entities are the nodes and the semantic relations are the edges. Yan and Tang (2023) use dependency syntactic parsing in order to determine which argument belongs to which sub-event.

Especially when processing large amounts of text, the actions and entities need to be synthesized in order to create an interpretable network. This warrants the use of *entity linking*, i.e. mapping the extracted spans to a set of latent entities. This technique is used for different types of entity networks, whether linking entities together through semantic relations (Ash et al., 2024) or co-occurrences (Botzer and Weninger, 2023).

Although many papers use narrative graphs to discover commonly occurring themes across documents, other studies start with topic modeling techniques to build their network, especially in computationally less complex approaches. For example, Mohr et al. (2013) analyze USA national security documents by extracting actants using NER and relationships between them using POS tagging/syntactic analysis. The resulting acts are contextualized using LDA. This is an example of how a computationally simple analysis based on literary theory can yield interesting results for social sciences. Khalili-Mahani et al. (2022) study discourse on elderly people during the COVID-19 pandemic in news media, comments on news sites, and reactions on news articles on Facebook and Reddit. They use topic modeling to discover low-level topics, which are used as nodes in a network. The edges are based on the co-occurrences of topics in posts and news articles. They then use community detection algorithms (mostly the Louvain algorithm, developed by Blondel et al. (2008)) to identify larger, underlying themes or communities.

2.1.1 Event graphs

Some papers represent narratives not in terms of actants or keywords, but rather in terms of events. The aim of these studies is often not to gain new politicological insights. Instead, they aim to synthesize news stories from many different sources, to create a tool that allows users to easily understand unfolding (political) events and stories, such as the Covid-19 crisis (Keith Norambuena and Mitra, 2021), Black lives matter protest (Ansah et al., 2019), or the geopolitical conflict in the South China Sea (Yan and Tang, 2023). These types of approaches see narratives as chains of complex events. Complex events (called *topic themes* by Ansah et al. (2019)) are clusters of atomic events that all adhere to some common theme or storyline (Ansah et al., 2019; Yan and Tang, 2023).

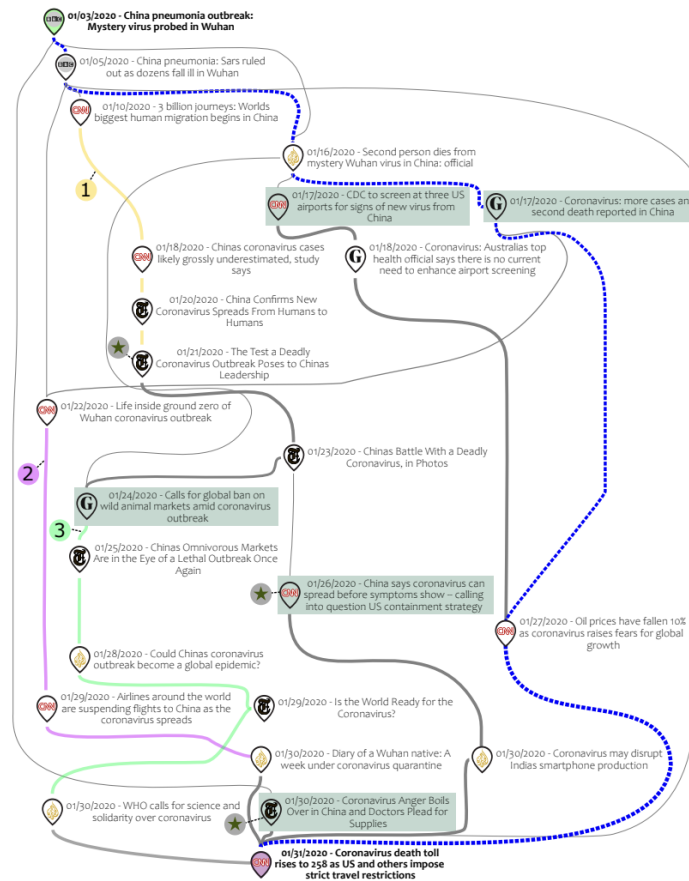


Fig. 5. Narrative Map for the Coronavirus Outbreak during January. Headlines in bold denote the starting event and ending event. The blue dashed line corresponds to the main storyline (the maximum likelihood path). The width of each edge depends on its weight (coherence). Events highlighted in green are representative landmarks (members of the first maximum antichain). The markers 1, 2, and 3 highlight some storylines that we reference throughout our discussion, we display them in distinctive colors. The star markers show events that share the theme of questioning or criticizing government responses.

Figure 1: Example of a directed acyclical graph used to reflect an event, as published in Keith Norambuena and Mitra (2021), p.17

Keith Norambuena and Mitra (2021) follow the framework of *information cartography*, developed by Shahaf et al. (2013), to represent political narratives about Covid-19. They depart from the theoretical definition that a narrative is a sequence of events with some underlying theme. Based on this assumption, they represent a narrative as a directed acyclical graph, where the nodes are events (based on newspaper headlines). Nodes are connected with edges whose weight is based on coherence metrics; their direction reflects the temporal sequencing of the events. This leads to a *narrative roadmap*, an intuitive way of visualizing narratives, as shown in Figure 1. Their goal is to find the most likely route between some starting and some ending event in order to explain different political narratives. In subsequent work, they incorporate community approval in social networks for a better narrative graph based on social media data (Keith Norambuena et al., 2022). In another paper, they combine their approach to narrative graphs with a semantic interaction framework (Keith Norambuena et al., 2023).

Li et al. (2021) focus on timeline summarization. They first construct an infinitely large graph and then extract events to construct the same graph with as little information loss as possible. Their approach can thus be seen as a way to compress event-centric knowledge graphs, an important step to make network visualizations of large datasets more manageable.

Studies that extract event-based knowledge graphs are rooted in research on event detection as a subfield of information extraction (Li et al., 2021). They tend to use computationally and mathematically more complex techniques than studies performing textual analysis. Yan and Tang (2023) perform event analysis to construct story graphs based on news documents. They use a graph neural network to detect events, and fine-tune a BERT model to extract the temporal relations between the identified events.

2.2 Modeling political discourse

Since the *argumentative turn* in political science, discourse plays a crucial role in the understanding of policy processes (Fischer and Forester, 1993). This view holds that policy is decided in a debate where different actors hold different views because they operate on different political paradigms that encompass sets of ideas, concepts and categories that give meaning to the social reality. The goal of each political actor is to have their own sets and ideas be accepted as hegemonic (Muller, 2014). The construction of these paradigms in discourse is known as *framing*.

Discourse network analysis understands policy debates as a negotiation of frames and meanings. Muller (2014) introduces *political discourse analysis*, which is a way to formally measure abovementioned discursive practices in policy making. This includes an *affiliation matrix*, a two-mode network that connects actor nodes with statement nodes by edges that express agreement or disagreement. Based on the assumption that shared opinions on political statement implies an affiliation between actors, this network can be used to compute political affinity between actors, which in turn results in an *actor co-occurrence network* and a related *conflict network*, that expresses the degree to which actors disagree. These networks are useful to identify coalitions and quantify the degree to which they differ from each other (inter-group dissimilarity) and the degree to which the actors in a coalition agree with each other (intra-group similarity).

Homophily is the effect that 'contact between similar people occurs at a higher rate than among dissimilar people' (McPherson et al., 2001b). In other words, if social network analysis

shows a cluster of actors that have a lot of interaction with each other, these people are often found to be similar to each other. Social networks that represent political actors as nodes often have edges that show how often they interact with each other. The assumption is thus made that the more groups or actors interact (or the more discourse they share) (i.e. the closer they are to each other in the network), the likelier they are to be ideologically close.

Papers performing discourse network analysis often focus on in-person political debates, taking an ethnographic approach (Koch et al., 2023; Larkins et al., 2024). This necessarily leads to rather small-scale studies, with most of the network analysis done manually. This is not the case for research studying debate on social media. These papers typically use networks where the nodes represent users and edges are usually interactions or mentions between the users (Urman and Katz, 2022; Luran et al., 2020; Willaert, 2023). Abzianidze (2020) takes an original approach, mining Georgian newspaper articles for (mentions of) nationalist discourse. She then plots them in a directed graph, where the nodes represent discourse actors mentioned in the newspaper articles and the edges are the amount of interactions (i.e. actors addressing each other).

Willaert (2023) look at Telegram channels spreading information (one to many style) on different topics that they hypothesize to be connected (war in Ukraine, crypto, far-right, Great Reset and stolen elections). They analyze messages by the acts they discuss, where 'act' is defined as a relationship between two pieces of content (in line with narrative theory). They vectorize messages based on these acts and using cosine distance they compare how much they shift or reinforce the narrative from the previous and following 10 messages in the channel. The authors use these to create behavioural profiles for the narrative in each channel for each of the five topics.

Community detection algorithms are often used to find groups of users with similar ideological positions. Luran et al. (2020) look at discourse on Twitter concerning a political scandal in Dutch agriculture. They employ community detection to find different groups of Twitter users talking about the scandal and use automated text analytics to characterize the narrative going on in each of the communities.

Networks can be used to study *direct affinity* (the interaction between two actors), but also *structural affinity*. This is the extent to which two nodes in the network behave similarly in relation to other nodes. The study of the structural affinity of two nodes is called structural equivalence analysis, introduced by Maoz et al. (2006). When applied to discourse network analysis, it can be used to detect structural similarities in discourse habits between actors. Abzianidze (2020) employs this technique in order to analyze nationalist debate in the Georgian public sphere. She found that in-group members (i.e. Georgian nationalists) are structurally similar to each other in the way they talk to and about ethnic minorities in Georgia. This means that the discourse structure as a whole sharpens the division between in- and out-group actors.

2.3 Graph Evaluation

A graph in itself is difficult to evaluate, and there is no clear agreed upon method among the papers discussed in this chapter to evaluate the goodness of a graph. Many papers (for example Segev (2020); Li et al. (2021); Keith Norambuena and Mitra (2021); Jing and Ahn (2021); Ash

et al. (2024) and Botzer and Weninger (2023)) use case studies as a way to show that their graphs can be used to produce sensible results: they implement their algorithm to some real-world problem, and if the results line up with some ground truth manually extracted from news sources, it is supposed to be a good graph (Chong et al., 2021).

Other papers use human evaluation (Chong et al., 2021). This is especially needed when the goal of the research is to use graphs in order to visualize large amounts of data in a way that is understandable for lay people wanting to interact with the topic. For example, Keith Norambuena and Mitra (2021) set up a list of carefully defined metrics, based on both their own work and earlier studies by other researchers, and ask Mechanical Turkers to evaluate their graphs.

Knowledge graphs, that use the synthesization of events to represent complex news stories, often emphasize coverage and coherence as metrics to evaluate their narrative graphs. Generally speaking, coherence is high if the events in a storyline have a high similarity; coverage means that the story graph covers as many diverse but relevant events as possible without lowering coherence (Yan and Tang, 2023). Keith Norambuena et al. (2022) add *community ratings* as a coherence metric: they use Reddit data for their study, and use the degree to which user upvotes align with the prediction of important events for their storyline. Ansah et al. (2019) evaluate their event graph using ROGUE (Lin and Hovy, 2003), a metric based on n-gram overlap that was originally developed for the evaluation of automatic summaries. The authors create a gold standard by manually collecting event reports from major news outlets and extracting keywords from the event descriptions. They then expect their model to produce at least one word matching the gold standard for each event in the graph.

2.4 Findings

This chapter so far has discussed the goal and design of different networks constructed to represent political discourse. This section aims to give a short overview of some practical or theoretical outcomes of these efforts, in order to illustrate the usefulness of narrative and discourse graphs for political science.

Narrative networks are often used to identify the most common narratives that are being spread in a certain information space or by a specific community. For example, Ash et al. (2024) represented political speeches from the US congress as narrative networks. Their analysis showed a clear reflection of historical events in political speeches; for example, the US invasion of Iraq sparked a cluster of narratives around God blessing the troops. Segev (2020) analyses large amount of news articles that mention fake news for the co occurrence of keywords. He finds different clusters of keywords, indicating different talking points, which he uses to synthesize the main narrative of the media in the fake news debate, namely the politicians (the villains) spreading false information through social media, the people (victims) falling for it, and journalists fighting back (the heroes). The identification of common narratives also helps researchers contrast the narrative of different political parties (Ash et al., 2024) or communities (Keith Norambuena et al., 2022; Botzer and Weninger, 2023).

There has been a considerable research effort to map conspiracist narratives on social media using narrative graphs. Like other narratives, conspiracy theories are rarely regurgitated as a whole. Instead, a small part is repeated in each storytelling event (in the case of social media this

can be for example a post or a thread). This has a self-confirming effect: the more often a specific part of a narrative is repeated, the more likely it is to appear in future storytelling events (Tangherlini et al., 2020). Tangherlini (2017) use graphs of actants to contrast the narrative structure of a real conspiracy with that of a conspiracy theory. They find considerable structural differences: where conspiracy theories tend to extend to all types of domains, the real conspiracy they studied remained limited to a very specific domain. Miani et al. (2022) find that conspiracist texts are more coherent with each other, but less coherent with themselves compared to mainstream/non-conspiracist text. This also suggests that conspiracist narratives tend to attach to the bigger narrative.

Chong et al. (2021) create a tool for contextualizing conspiracist stories on social media. Their goal is to create a tool that allows lay users to easily select a (potentially conspiracist) narrative and visualize the context surrounding it. They extract narrative graphs from conspiracy-related posts from various platforms and create an interface that allows the user to 'click around'. They find that new narratives revolving around conspiracist talking points often attach to popular, well-connected nodes, possibly to seem more appealing and connect to the larger narrative in order to be 'incorporated' by the overarching conspiracy theory.

Botzer and Weninger (2023) use entity graphs based on Reddit data to predict the direction conversations on the forum will take, in order to showcase broad patterns of political conversation on internet fora. They find that online discussions tend to go 'off tracks' as they continue, and that they tend to spiral back to a rather small set of topics, but these are not necessarily related to the starting topic.

Urman and Katz (2022) look at far-right telegram channels and represents them as nodes in a network, where the edges are mentions of other channels. The goal was to see how these channels are interconnected to other channels. This is an example of how networks can show how fringe ideologies (in this case the far right) are connected to the mainstream political discourse.

Graphs, due to them being mathematically well-defined, are also well-suited to serve as well-formed input for other text processing algorithms. Sawhney et al. (2020) use graphs to represent different types of context in political debates. This is then used as the input for a transformer-based neural architecture for political stance detection.

2.5 Challenges and future research directions

As shown in this chapter, networks are a fruitful way to represent political narratives and political discourse. With the volume of online information and discourse increasing, they are an important tool to synthesize information from different sources and make them insightful to the public. They are also useful tools for political scientists looking for a way to quantify and/or visualize observations about political narratives.

The research on networks representing political discourse is not a unified field. Researchers come from very different backgrounds, use rather different methodologies, and have different objectives. Different lines of research do not tend to cite each other (for example, papers on quantitative narrative analysis rarely build on knowledge from papers on event-centric knowledge graph, even though these lines of research clearly have overlapping goals). However, unified research effort could lead to interesting interdisciplinary collaborations between researchers. Linguists,

political scientists, and computer scientists are all involved in this field, each bringing their own expertise and knowledge. Unifying their efforts could lead to unexpected new insights and ideas.

Moreover, the lack of unification also means that, to the best of my knowledge, no efforts have been made to systematically compare and contrast network approaches to political discourse, for example in the form of a systematic literature review. This makes it difficult for researchers first wanting to implement a network-based methodology to find best practices. This is exacerbated by the different vocabulary different lines of research use to refer to the same or very similar things, especially regarding narrative graphs, making it hard to find potentially useful or similar research papers. As a methodology, narrative networks would also benefit from a more systematic approach to graph evaluation.

So far, I have discussed research that uses networks to represent either the narrative or the dynamics of the debate among users. But these levels are clearly related: the principle of homophily teaches that people with similar paradigms and frames tend to form social groups/alliances on social media, but we also know that debate in itself is how policy is determined (Muller, 2014). Social media offers a unique chance to make a network where both the debator and the debated are represented. Such a representation would theoretically allow us to quantify how political actors (including lay people that partake in political discussions online) shape the development of the narrative.

Some recent efforts have been made in this direction. Botzer and Weninger (2023) implement this idea by using networks to show the co-occurrences between entities in Reddit threads; users, then, are modeled as *moving through the graph* as their discussions jump from one entity to another. However, this is only scratching the surface of the possibilities.

Another way to expand the usefulness of narrative graphs is to create multilingual graphs. This would allow researchers to contrast political discourse between linguistic communities, and/or see how political narratives from different linguistic spaces influence each other. With the recent focus on multilingual NLP, this seems like a logical next step; however, to the best of my knowledge, no such attempts have yet been made.

3 Political Opinions and Public Participation on Social Networks

This chapter focuses on studying political expressions and political participation on social media platforms. The platforms have been seeing an exponential growth in terms of user population and diversity of application, and through them users communicate, participate and act on various political issues and tasks. In the first subsection, we underscore how political opinions and ideologies can be defined on these platforms, highlighting some influencing factors that shape how people make their opinions and express themselves. In the second subsection, online political participation is discussed focusing on the different levels of participation that can be observed. We also discuss how social media platforms can influence the access and visibility of political content through various recommendation algorithms and moderation strategies. In the third subsection, we briefly examine some of the political opinion datasets that have been collected in the past years, while the last section summarizes the chapter, highlighting the challenges and the

future work that may be done to pursue this research direction.

3.1 Understanding Political Opinions, Attitudes and Ideologies

Political scientists have been more successful in characterizing and accounting for public opinion than in specifying its role in the broader political process (Bartels, 2015). They have primarily collected and studied opinions in an attempt to analyse the politically relevant preferences and beliefs of the public, and evaluate the democratic competence of the citizens. Political opinions, can be then defined as “thoughts and statements held by private persons that governments find prudent to heed” (McDonagh, 1962). This is an expansive definition of the term, as opinions are in reality a property of an individual that gains power in a public sphere.³ ‘Political attitude’ is another expression commonly used in political opinion analysis. It is defined as the predispositions that help underline specific opinions (Erikson and Tedin, 2019). However, in many contexts opinions and attitudes can be used interchangeably.

There has been a resurgence of polarization of opinions in recent years due to the rise of mass politics through platforms like television, online news and the social media, as well as the intertwined evolution of modern democratic states and societies (Bartels, 2015). Political researchers have become interested in *what* attracts individuals to different ideological positions and *when* individuals think and make judgements about issues and candidates. This has brought back interest in understanding the belief systems that exist in public. Since the 1950s, there has been ongoing research to find a single overarching belief system that could guide political opinions (Converse, 2006). Ideologies are these belief systems or networks of inter-related ideas that can be shared, and are both descriptive and prescriptive in nature. Thus, they essentially provide an interpretation of how society currently is and also set out aspirations to how it ideally should be and how the political goals should be attained. In Figure 2, we observe the possible interactions between different stakeholders in influencing how an opinion is formed. Over the course of this chapter, each stakeholder in this political communication will be discussed in terms of their interaction and mutual influence.

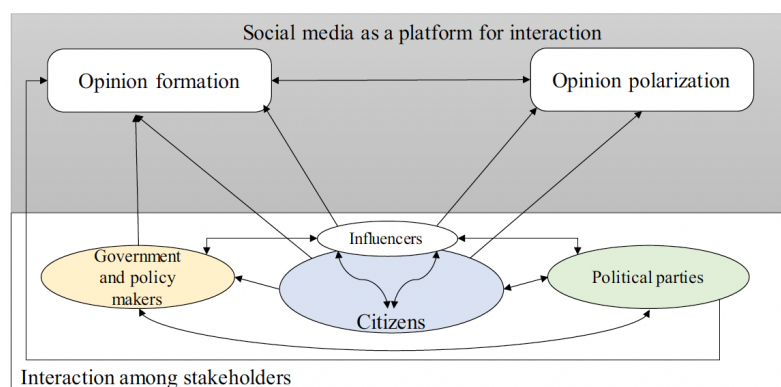


Figure 2: How opinions are formed and expressed on social media platforms, from Kushwaha et al. (2022)

³Within the context of this text, ‘political opinions’ and ‘public opinions’ terms will be used interchangeably. But the primary distinction between these terms can be noted in the context of their use.

3.1.1 Shaping Opinions and Ideologies

Interestingly, one would consider a person to think politically as they do, socially (Lazarsfeld et al., 1944). The social characteristics should help determine one's political preference. The *sociological approach* (Cox and McCubbins, 1993) holds that politics reflects the social conditions and processes. The authors state that demographic factors (or group influences) play a strong influencing role on the individual's opinions and even voting choices. The other school of thought, referred to as the *political approach*, holds that politics and the political thought are autonomous (Sartori, 1969). Here the authors claims that social characteristics, such as the demographic factors, cannot be really correlated with the political choices and actions made. For instance, in the political opinion survey conducted by (Lazarsfeld et al., 1944), it was seen that nearly 17% of the people in the "Strongly Democratic" group voted for "Republican". In conclusion, numerous empirical studies that have shown that demographic factors do have virtually some predictive power. But determining the value of strength of the association between these factors and the political factors is difficult to ascertain and no objective standard can be made for it. The interactions between the social characteristics and the political characteristics have been well documented in Weakliem (2003).

In addition to the political and social/demographic influences, some past works have studied the psychological and biological factors influencing how opinions and ideologies are shaped. Individual psychology plays a huge role in determining where people stand in the modern political landscape (Berinsky, 2011). There is also some correlation observed in the role of emotion in shaping one's political preferences (Berinsky, 2011). It has been observed that there are three primary emotional factors that tend to have the most political influences: fear, anger and enthusiasm. For instance, a sense of fear in situations of war or economic stress, tends to make people more alert and focused about the candidate's leadership and issue positions, and even reconsider their standing decisions on who they intend to support. Finally, an interesting connection has been observed in the role of biology in politics, also referred to as bio-politics. Berinsky (2011), for example, shows how genetics can shape one's political attitude through the heritability of personal traits and cognitive style. Being a controversial issue, the research is still in its early phase but it is certainly a promising and interesting area of research.

3.1.2 Political Alignment: A Brief Case Study

The adoption of social media for the purpose of mass politics has allowed researchers to study and monitor large volumes of opinions being generated from politically active participants in real time. These studies have revealed the need to also study the different positions or stances that these political actors take with regards to the different political issues and interests. Studying political alignment not only helps one to distinguish between these political actors based on their political viewpoints but also allows one to study conflicting signals at individual levels that can then be aggregated to help develop efficient political strategies. These can then be used for the purpose of mass political advertising, election campaigning and designing public policy. For instance, political parties can design their campaigns to target lower-traffic sites that are more popular among the politically-active users in order their return on advertising investments. These lower-traffic sites can be the sites that are more left-wing or ring-wing in nature, i.e. having a

polarized user-base. Thus this potentiality of the study has led to an increased focus on developing techniques and approaches for detecting and distinguishing between users having different political affiliations/viewpoints. Absence or improper implementation of these techniques would result in poorer aggregation over conflicting political signals and likely obscure the nuances most important for developing the political strategy (Conover et al., 2011). Consequently, there is a strong desire to build efficient and robust methods.

With regards to identifying relevant research and methodologies, some key concepts have been described in the past including political polarization (Kushwaha et al., 2022; Németh, 2023), political stance detection (Stefanov et al., 2020; Barberá et al., 2015; Barberá and Rivero, 2015; Glandt et al., 2021), political (ideology) alignment or leaning (Conover et al., 2011; Ribeiro et al., 2018), opinion manipulation (Silva, 2016) and opinion diffusion (Bredereck and Elkind, 2017). In this chapter, we will also include studies focusing on social media platforms. While these terms have their distinct definitions within the context of identifying political viewpoints, research methodologies are quite similar, though not the same, as will be briefly discussed in the following subsections.

3.2 Role of Social Media in Political Participation

Social media have actively expanded the citizens' ability to participate in civil and political activities, also birthing new modes of political actions and reactions. Social media have now become a primary means of initiating political dialogues and expression, organizing political rallies and protest on various political issues, and taking public actions (Waeterloos et al., 2021). Traditionally political information would be mediated to the public through journalists and media houses, and would thus be biased to the opinion of the latter. The public would generally be the consumers of information, only getting to express their opinions through some interviews, opinion polls and surveys. Social media have effectively changed this scenario, providing the public with a platform that gives them a stronger voice and better representation in the political world. This has effectively changed the landscape of political expression and participation.

Cyber-pessimists view the Internet and digital media as having potentially widened participatory inequalities, reinforcing the Democratic divide. Morozov (2014) highlighted that often political activity on social media occurs to impress one's digital community, rather than being due to one's commitment to the ideas and politics. Internet censorship and digital surveillance has also eroded the public's ability to have open and free political participation and collective action (Stoycheff et al., 2020; Chan et al., 2024). In contrast, the cyber-optimistic emphasize the democratic potential of digital technologies, especially social media, since they have lowered the costs for political participation, allowing for more voices to be heard (Xenos et al., 2014). By comparing online participation (on social media) and offline participation (such as protesting and rallies), it has been seen that online participation has been less costly, less time consuming and less risky. It has also been observed, that nowadays online participation leads or "spills over" to offline participation (Vitak et al., 2011). Thus online participation or activity can be considered a possible predictor of offline participation. Activities such as 'online news consumption', 'online political expressions' or just 'the general frequency of the digital media use' can be studied to identify online political participation and make predictions on the sentiment and stance of the public with regards

to different political issues. The general consensus is that online political participation should be seen as a proper political action (Gil de Zúñiga et al., 2014; Erkulwater, 2012). Indeed, two main levels of political participation can be identified on the social media platforms: active participation and passive participation.

3.2.1 Levels of Political Participation

Active participation includes participants that primarily generate or lead the lines of political expression and political actions on social media platforms. These can be political actors such as politicians, the electoral candidates/campaigners, and the government representatives, whose purpose is to represent and promote their political party and ideology, to address the public regarding their concerns and interests, and to communicate the public policies and actions to be implemented. But active political participation in the social media context can also include another sub-group, often referred to as the “social media influencers”, who do not necessarily have to play a political representative role in society but tend to use their influencing skills and platforms to help promote certain ideologies and political expressions. These “third party actors” generally play an important role acting as the intermediate communicators and diffusers of information between the main political representatives and the general public. While these politically active third parties can have a level of independence to their political actions and participation (Raymond et al., 2022), there could also be a financially or socio-politically backed “hidden” agenda working to help promote certain perspectives and narratives in the online society. Furthermore, in these sub-group networks, we can also often identify “opinion leaders” that represent the community of participants that act as the central node in political communication, and have the power to express, represent and promote certain ideas and actions on the general public. Often these “opinion leaders” can also be seen as “hyperactive users” (Matuszewski and Szabó, 2023). Hyperactive users act as useful pathways for efficiently disseminating the original messages beyond the followers of opinion leaders. These accounts have significant importance when it comes distributing attention on social media (Papakyriakopoulos et al., 2020)

The other level of participation is *passive participation* that usually represents the groups that tend to primarily consume political content on social media platforms. Their level of interaction with the political content can be seen as minimal as compared to an active or hyperactive participant on the platform. They are usually the target of political actors, as they represent the larger section of users on social media platforms and are the voter banks that political actors want to influence and gain support from. Studying passive participants on these platforms can help to develop better and efficient election campaign schemes and reach large groups of audience to promote one’s political agenda and action. It has been observed that the majority of users on the online social networks follow extreme value distribution (Lerman and Ghosh, 2010; Benevenuto et al., 2009). The majority of these user bases stay passive, participating with very low frequency, while the hyperactive users over-proportionally exert their political opinions and attitudes on the platform (Papakyriakopoulos et al., 2020). Hyperactive accounts usually have their own and authentic followers thereby enabling them to re-transmit the messages from the opinion leaders to the broader audience. Papakyriakopoulos et al. (2020) noted that active and hyper-active users are able to influence the way initial political information is disseminated to passive users, thereby

directly influencing political actors. Also for controversial topics, frequent in the political context, they are seen as a complex social contagion (Mønsted et al., 2017; Romero et al., 2011).

The process by which people learn about the political world and acquire their opinions and views from the others is referred to as *political socialization* (Jost et al., 2009). The nature and manner of the generated content is heavily influenced by who the consumers are, but interestingly the consumers themselves are heavily influenced by the content made visible to them. Political socialization on social networks takes on another layer of influence altogether. It is important to remember that social networks are essentially businesses that operate on the task of selling advertisements on their platforms. Their business model operates on seeking to maximize profit through building up and retaining an active (in terms of time and user activity) user-base. The aim is to transform the social engagement to profits through services such as subscriptions, advertising and product marketing. To maintain this user activity and user interest, social networks have designed recommendation algorithms that run constantly, ensuring that each user is displayed content specific to their tastes and preferences. These algorithms are adjusted to each user in terms of content popularity and user preferences. Users are heavily influenced by the content that is displayed to them through these algorithms. These algorithms act as a filter, a censor, and a source of information manipulation, bringing in their own bias and agenda to the platform. Apart from using recommendation algorithms to maintain the user activity and interest on the platform, social networks also use gamification tactics and tools to increase user engagement, motivation and activity (Hassan and Hamari (2020)). They gamify the experience on these platforms by adding possibilities to like, share and comment on the various posts and events on the platform. Users can be rewarded with improved visibility and social status on these platforms, which can further increase their engagement, motivation and enjoyment on these platforms. This has often been implemented during electoral campaigning as discussed in Section 3.2.3. Thus, understanding these groups and the interactions between them is essential for conducting political research, as the interactions between these groups heavily influences how political campaigning and opinion diffusion occur on these platforms.

3.2.2 Moderating Political Participation

The rise of inappropriate content like misinformation, spam and hate speech has become a prominent issue on social media platforms leading, to the pollution of the platform's environment. Thereby, political expression and participation have often seen many levels of moderation, censorship, control and surveillance observed and managed by various authorities such as the platform companies and/or the regional or national political powers. Introducing moderation and control on social media platforms can be useful to ensure that there is equal and fair access to the true and politically correct content and statements. As mentioned earlier in the section (Stoycheff et al., 2020; Chan et al., 2024), this control and surveillance can also be misused to control the access as well as the level of visibility of some political opinions and narratives when in conflict with the party with power and vested interest.

The majority of moderation methods focus on content moderation (Chancellor et al., 2016; Jhaver et al., 2018), wherein the generated content is reviewed before or after submission on the platform, and detection algorithms are used to flag any content that disagrees with the social

and safety guidelines of each platform. There are many algorithms being implemented with the purpose of hate speech detection, claim and fact checking, misinformation detection and possible spam (like machine-generated text or bots) (Anjum and Katarya, 2024; Altay et al., 2023). Such content would often be censored, removed and the content creator would be alerted of the issue. In case of repeated offences, the content creator can be flagged and face legal charges and have their account suspended or terminated.

Alternatively, user moderation such as user bans can be seen as a possible moderation strategy in terms of personal orientation, severity (temporary vs. permanent bans) and transparency (Myers West, 2018; Zhang et al., 2022b). Social media also use platform recognition strategies to control user visibility and prestige. By using concepts of 'badges and recommendations', they effectively control and motivate the users to contribute only the correct and verifiable knowledge on the platform (Ma et al., 2022). In this form of a carrot-and-stick strategy, the platforms are able to effectively ensure that users comply with the platform's rules and norms. Similarly, 'demonetization' tactics can be used to discourage users to produce and promote content that disagrees with the platform's rules and regulations (Zappin et al., 2022). In case such types of content are detected, a platform can deactivate the ability of content creators to make money off of that particular content. There have also been cases where these tactics have resulted in reducing the content creator's freedom of speech and expression (Graupmann et al., 2012; Rosenberg and Siegel, 2018). The study conducted by Zhang et al. (2022b) analyzed the effectiveness and impact of user bans on a Chinese social platform called Zhihu (similar to Quora in purpose). They observed that individual user moderation as opposed to content moderation leads to immediate change in the users' content generation behavior. Likewise, they observed that user bans tend to have a heterogeneous effect dependent on the level of user's status and recognition on the social platform in question. Finally, they have observed that platform governance does not necessarily decrease participation likelihood and noticed an increase in low-quality content generation after a user-ban was observed. Thus user-bans can unintentionally induce stronger user reactance and lead to further platform pollution. Incorporating multiple moderation strategies holistically, such as user recognition tactics, can potentially help reduce inappropriate content and even boost content quality.

3.2.3 Electoral Campaigning and Social Media: A Brief Case Study

Of the key areas of research focus for the political scientists is examining whether public opinions do in fact play an influential role on public policy. Several pathways have been discovered that have linked public opinions to public policy such as electoral accountability, interest and political groups, politicians' views and ideology, and the political parties (Stier et al., 2018). Elections have been shown to be one of the major mechanisms through which the public have been able to make an impact and influence who would form the office, employing the candidates that reflect their own views and ideologies, and removing the ones that failed to do so. The political candidates in turn attempt to curry favors with the voters by trying to align themselves with the general public sentiment and attitude. This behaviour can be further backed by the observation that the overall responsiveness and participation often sees a major increase around the election period (Waeterloos et al., 2021; Vitak et al., 2011). With social media now being an active platform for

political discussions, exchanges and information mitigation, studying the their role in electoral campaigning is important and useful for understanding opinion propagation and manipulation.

Opinion polls have always played a major role in electoral campaigns. Misrepresentations of public opinions have had a direct effect on the political campaigning. Voters' ideology and support can often be predicted from their political identities, such as party identification, and their social identities, such as race, ethnicity, and religion. Voters can also be more influenced by media reporting about the economy than from the economy itself, like for example in 1992 when Bill Clinton benefited from the outlook of weak economy even though recession was over [Doherty and Gimpel \(1997\)](#).

3.3 Studying Political Opinions and Participation

Public Opinion Research developed rapidly from the appearance of sample surveys in the mid-1930s that allowed for a systematic process of studying the opinion generation, diffusion and manipulation. Some notable early works include the works of [de Tocqueville \(1969\)](#), [Dacey \(1914\)](#), and Lord Bryce's take on public opinion ([Wilson, 1939](#)). These past research have helped shape the modern sense and definition of the public opinion ([Osborne and Rose, 1999](#)). In the present day, online forums are becoming the increasingly central channel through which political communication is organized but also observed by the research community ([Stier et al., 2018](#)). However, understanding political opinions is generally difficult due to several reasons, from nuanced language and ambiguous sentiment to complex content. Traditional sources of political opinions such as news articles ([Zhang et al., 2022a](#)), debate transcripts ([Iyyer et al., 2014](#); [Chen et al., 2017](#)) and posts from official sources such as news agencies, government officials and politicians are easier to analyse as they follow structured representations, with clear and concise language and grammar. In contrast, analyzing opinions from social media users presents various challenges such as lack of brevity of the text, absence of contextual information, and the use of non-textual content such as emoticons, abbreviations, hashtags, and multi-modal content. Also, acquiring labels for social media data is quite difficult with manual annotation methods often being too expensive and impractical ([Li and Goldwasser, 2019](#)). Lastly, defining the political ideologies and political opinions of the different users is also quite challenging due to the varying influencing factors to be accounted for, as discussed at the beginning of this Section ([Berinsky, 2011](#)).

To mitigate the challenge of acquiring clean social media data and preparing the datasets for further analysis, many tools and techniques have been developed. Survey works on social media data creation and analysis such as [Sapountzi and Psannis \(2018\)](#); [Beheshti et al. \(2018\)](#); [Hosseinzadeh et al. \(2023\)](#) highlight some useful strategies that can be employed in helping prepare and clean social media data sets. With regards to the challenge of detecting political opinions and political ideologies, many supervised approaches have been proposed that use text content on a user's timeline ([Pla and Hurtado, 2014](#); [Birmingham and Smeaton, 2011](#)) and their network structures ([Volkova and Bachrach, 2015](#)) to identify the political statements and leanings of the users. In a research study by [Zhang et al. \(2023\)](#), the authors discussed using a hybrid strategy, involving political hashtags and the social network graphs, to identify political ideologies on the X platform. With regards to designing low-cost and resource-efficient annotation strategies, [Samih and Darwish \(2021\)](#); [Elfardy and Diab \(2016\)](#) have proposed using label propagation methods.

This is a semi-supervised method that starts with a seed list of labeled users and propagates the labels to the other users in the network that have similarity to the labeled users based on account and activity. This approach can be biased toward users with extreme views and thus careful selection of the thresholds and the initial seeds is required along with post-annotation checks for quality control. Unsupervised techniques for stance detection were proposed by Hasan and Ng (2012); Baly et al. (2020), where they mapped users to a lower-dimensional space based on user-user similarity and identified core sets of users representing different stances. This approach has been shown to be effective for polarizing topics. Another recent research direction is deploying large language models to perform specific stance detection tasks using these social media data sets (Linegar et al., 2023). Some interesting and important research works that summarize relevant political data sets and annotation schemas include ALDayel and Magdy (2021); Schaefer and Stede (2021); Lai et al. (2020); Bauwelinck and Lefever (2020).

3.4 Challenges and future research directions

In this section, we have discussed how political opinions and ideologies are important aspects to study in order to understand political expression and participation. We emphasised some of the influencing factors that can help predict or determine the nature of the opinion and the ideology that an individual might hold or stand by. We then looked into how individuals participate in the political communities on social media in section 3.2, either as active or passive participants. Identifying the different political entities and their roles in social and political networks can help to build better political strategies for electoral campaigning, devising public policy and political decision-making. Interactions between the different groups on social media is further controlled by the use of recommendation algorithms and other social gamification tactics on these platforms that govern access, nature and visibility of different political expressions and political ideologies. Also, there has been an overwhelming influx of inappropriate information that has polluted the platform environment. In Section 3.2.2, moderation strategies were discussed as well as the possible biases or hidden agenda behind them. Finally in Section 3.3 we presented some social media datasets developed for opinion analysis in the political domain that have been used by the research community.

Since much of the user activity on social media is not normally distributed, one of the main challenges to classifying political opinions on social networks is to train robust statistical and neural architectures on these extremely skewed and unbalanced datasets (Hernández Santiago et al., 2012). There is a high risk that due to the skewed opinion distributions and data, the trained algorithms will fail to learn and propagate or recommend unbiased and coherent political information to the users. This can lead to algorithmic manipulation in the political exchanges that occur on these platforms (Bucher, 2012). Often the recommender systems on these platforms come with a social influence bias. They have the power to alter public opinion (Krishnan et al., 2014; Gearhart et al., 2020). In order to mitigate this risk, better transparency in the algorithms is required (Reisach, 2021). Also some effort should be devoted to create balanced datasets that are accurate representations of the public opinion and distribution.

Another challenge to be mitigated is the influence observed from users who are remarkably more active than the others, as explained in Section 3.2.1. These users have the power to distort

the recommendation capabilities of trained models, through the likes of adversarial attacks (Kantopoulos et al., 2020) and social network manipulations, referred to as graph poisoning. This influences the visibility of certain posts and opinions on the platform, and thus lead to manipulated and propagandist information being shared to the general public (Olaniran and Williams, 2020). In graph poisoning, network structures can be altered to ensure certain information become invisible or be over-represented, consequently manipulating the recommendation systems (Zügner et al., 2018; Li et al., 2023). Proper moderation strategies need to be designed that can help to bring in some regulation to the activities of these hyperactive users and allow for better exchange of information on these platforms.

Finally, automated content (and user) moderation is also an issue of concern for the political community (Gorwa et al., 2020). Though the intention is to promote platform responsibility, safety and security on the global stage, these approaches can unintentionally result in increased opacity in understanding the ethical and legal practices of the platforms, complicate maintenance of fairness and justice on the platform, and sometimes even exacerbate the problem by creating tensions between the users and the platform itself. Also, often the social media platform's policies define the hate speech solely with respect to the content of the message. Wilson and Land (2021) highlighted that it is important to understand the hate speech with reference to the political realities and power asymmetries that exist in the society in order to ensure that the moderation policies are better structured and robust to the real-world. Thus, there is a need to develop efficient and user-friendly moderation strategies to ensure that the platforms operate in a fair and democratic manner allowing users to be able to ethically and legally express their political opinions and viewpoints.

4 Rhetorics of Implicit Communication in political discourse on social media

In this section, we present existing works related to *implicit communication*, which we define as something that is suggested or hinted at, rather than being directly expressed. We therefore refer to the interpretation of utterances whose meaning may not be overt (Yus, 1999), while we exclude from this analysis non-verbal cues such as gestures and poses, which are not detectable on social media.

Implicit communication has become pervasive in political discourse as it predominantly revolves around the propagation of negative and disruptive ideas aimed at manipulating reflexive concerns. In the current landscape of politics, where impression management reigns supreme, even negative publicity serves as a tool for garnering attention. Implicit communication emerges as the arena where these strategies unfold, exploiting ambiguity to influence public opinion and maintain a grip on power. This dynamic fosters a culture where information that matters is sidelined, perpetuating a cycle of superficial discourse and societal polarization.

4.1 Derailed Reflexivity

Reflexivity can be defined as an individual's capacity to imagine what is happening in other people's minds. Our participation and evolution as active members of society primarily depends on our ability to adjust to the thoughts and potential reflections of others.

Reflexivity is being studied more and more by discourse analysts as it delineates a discursive pattern that is aligned with the deliberate assessment of various behaviors or types of communication within the public sphere to determine their acceptability or appropriateness. Ultimately, what [Verschueren \(2021\)](#) points out, is that when discursive activity only deals with reflexive concerns, it tends to be characterized “less by the intrinsic value of goals to be achieved than by considerations of what is necessary to obtain or maintain a position of power enabling one to achieve goals in the first place, whatever their value.” The outcome of this are policies that do not directly tackle societal issues. Instead, they aim to tackle constructions rooted in what a considerable portion of the population might desire or be easily persuaded of. In a surprising juxtaposition, these policies are presented as highly democratic.

This is where reflexivity goes off the rails. It derails in the sense that there is an overuse of the power of reflexivity to support or attack expectations for persuasive purposes. Moreover, an inevitable consequence of this is that political debate becomes a matter of impression management, where making an impression —given the popular idiom “there is no such thing as bad publicity” — either negative or not, is all that matters.

While we could concentrate on discourse tailored to meet specific social constructs and expectations, in this section, we aim to emphasize a particular form of rhetoric focused on communicating disruptive ideas. This rhetoric primarily centers on the transgression of social expectations and of reflexive concerns. As we will see, it is achieved not only by addressing reflexive concerns but also by employing a particular style of writing on social media that is slippery, subject to interpretation and relatively easy to hide behind or distance oneself from.

4.2 Implicit Communication

Implicit statements are perfectly suited for the rapid escalation of scandals, facilitated by a media-savvy performance of politics that seems to become more important than the political process itself ([Grande, 2000](#)). They have gained popularity among political actors not only for their ability to create pragmatic-stylistic effects like humor or sarcasm but also because they enable the author to avoid outright claiming what is implied.

While implicitness is commonly associated with unspoken claims or statements deliberately left unsaid, it can also pertain to a certain stylistic mode of expression characterized by the absence of explicit connectives or syntactic elements that link independent discourse segments (single claims or independent clauses).

4.2.1 Unsaid/Unspoken Statements

For [Wodak \(2015\)](#), engaging in the manipulation of ambiguity and implicitness corresponds to a specific discursive strategy that she refers to as *calculated ambivalence*, where controversial claims are made by making use of implicit communication cues. Her analysis of the concept

originates from her study of how certain politicians address multiple and contradictory audiences through a single, intricately layered message (Engel and Wodak, 2013).

As part of the cases studies that she introduces in her book *The Politics of Fear*, she mentions how the leader of the FPÖ, HC Strache, shared a caricature on Facebook that reinterpreted an American caricature from 1962 in a way that clearly referenced antisemitic caricatures from the Nazi era, which were regularly published in the 1930s in the notorious German newspaper “Der Stürmer”. Following the expected scandal caused by the explicit antisemitic elements of the caricature, numerous newspapers in Austria and Germany published editorials and news reports on the incident. Strache was also interviewed on television on 20 August 2012; initially, he denied altering the original caricature, then denied that the stars visible on the banker’s cufflinks were Stars of David, and finally categorically denied any resemblance to antisemitic caricatures. What emerges from this is that readers could either believe that any resemblance to an antisemitic caricature was entirely coincidental, or they could interpret the crooked nose and specific cufflinks as implying antisemitic meanings. This strategy enables the politician in question to assert victimhood when facing accusations of racism or antisemitism from opposing parties and certain media outlets. Consequently, the incident is often dramatized and exaggerated, with the FPÖ/Strache asserting that they have been unjustly accused of promoting racist or antisemitic sentiments.

In addition to its potential to elicit a viral response on social media (propelled by the fact that if a media communication chooses not to cover it, they may be perceived as endorsing it), calculated ambivalence most importantly achieves the successful implicit conveyance of a controversial message.

The persuasive strength of leaving statements unsaid lies in exploiting our attention and cognitive faculties to draw our own conclusions. Implied meaning allows us to independently reproduce controversial reasoning, giving us the impression that such thinking is inherent to our own beliefs, as if we arrived at these conclusions naturally without external persuasion. This is effectively achieved through specialized communications that juxtapose unrelated statistics, presented as facts, to create a perceived link between them. A common scenario involves promoting the mistaken belief that there is a connection between immigration and criminal behavior (Philippe and Valette, 2023). In various forms of communication such as radio broadcasts, social media posts, and television interviews, politicians like FN’s Marie Le Pen, Eric Ciotti from LR, and even Emmanuel Macron have been observed using two distinct sets of statistics: one set pertains to the total number of immigrants in France, while the other relates to the disproportionately higher number of immigrants in prison. They use these statistics to ostensibly ‘prove’ that immigrants are more inclined to engage in criminal activities.

While we do not delve into the inherent biases resulting from discriminatory treatment by the judiciary and police, which could explain the second statistic in detail here, it is important to note that the final conclusion “immigrants are potential criminals” is never explicitly stated. Instead, the conclusion is left for us to infer inevitably.

4.2.2 Indirect Speech Acts

Indirect speech acts have been used as linguistic tools to describe and identify implicit forms of communication. In a discourse analysis study focused on Twitter posts, de Rijk (2020) employs

Speech Act theory (Searle, 1979) to annotate tweets supporting a conspiracy theory.

The motivation behind this study stemmed from the notion that the automatic recognition of misinformation or disinformation is hindered by certain fundamental communicative features, such as sarcasm or implied meaning. Therefore, it is relevant to investigate whether indirect communication plays a role in misinformation and disinformation dissemination, and to what extent. Speech Act Theory categorizes every utterance into three forces: 1) locution, which is the literal meaning of what is said, 2) illocution, which refers to what the utterance accomplishes, and 3) perlocution, which denotes the effect or outcome of the utterance. For instance, in the statement "Peter, you are standing on my foot," the locutionary force asserts the factual state of affairs. The illocutionary force would typically entail a request for Peter to move his foot, now that he is aware of the situation. The perlocutionary act would then be Peter indeed relocating his foot elsewhere.

de Rijk (2020) conducted a qualitative analysis of 130 tweets using the dataset compiled for the MediaEval 2020 FakeNews Task (Pogorelov et al. (2020)). The analysis focused on coding for direct and indirect Speech Acts (SAs). Only tweets that supported a conspiracy theory (e.g., claims that corona is a coverup for 5G deaths or that 5G causes corona) were included in the results section. Indirect SAs were specifically coded when they were identified as the primary communicative force in the tweet. For instance, the tweet provided below was coded for 'Indirect SA: concluding' as the implied meaning was deemed to be the primary interpretation. Without this implied meaning, the assertions made in the tweet would be considered loose statements.

(1) *Anyone else curious about the majority of deaths in china seem to be the same areas they rolled out their stand alone 5G just a couple months ago. Verry few deaths being reported in other areas in comparison. #5G #CoronavirusOutbreak #COVID19c#5gamechanger*

The tweet in question begins with a question in the first sentence, indicated by the syntactic structure ('direct SA: asking'), despite the absence of punctuation. This is followed by two assertions ('direct SA: asserting'). However, the intended relationship between these sentences is not explicitly stated through coherence markers like meaningful connectives. The last sentence does include a lexical cue phrase ('in comparison') that indicates the user's intention to highlight a relationship between different areas. However, since the causal relationship is not explicitly stated, the act of concluding that these assertions are causally related is considered an indirect SA.

After the annotation, it is revealed that out of the total tweets analyzed, 59 of them, constituting 65.6% of the dataset, contained indirect SA. These indirect SA encompassed various forms such as suggestions, conclusions, invitations, and descriptions, indicating the diverse ways in which individuals communicate indirectly through online platforms. Once again, the author illustrates that indirect Speech Acts are employed to refrain from explicitly asserting the existence of a conspiracy. This is frequently achieved by implying a causal link, as previously seen with the use of statistics, or by citing another source, allowing the users to distance themselves from directly making the conspiracy claim.

4.2.3 Enthymematic Communication

Enthymemes are shortened versions of rhetorical syllogisms where one or more premises are not explicitly stated. Scholars have expanded on Aristotle's definition of enthymemes by acknowledging that they can involve multiple premises or conclusions provided by the audience, leading to various potential meanings. As noted by Blair (2012), enthymemes are arguments that require audience participation to fill in the missing parts. This style of argumentation can be highly effective because enthymemes are not merely based on propositions; they represent a viewpoint grounded in syllogistic premises and emotionally resonant ideas or images that encourage a passionate alignment with the speaker's perspective.

The main difference here with other forms of implicit communication techniques mentioned above, is that enthymematic structure relies on the power of evocation. Rather than communicating information, they encourage the audience to both act and believe, relying heavily on evoking emotions and sensations.

For instance the logical argument that "all fully human individuals are considered equal/all white individuals are fully human/all white individuals are equal" and its converse, which implies that non-Whites are not fully human, is not openly accepted in contemporary society. However, Jackson (2006) observes that White supremacist ideologies have subtly permeated through the effective deployment of enthymemes. For instance, when law enforcement officers are cleared of misconduct in cases involving the death of a non-White suspect based solely on their own version of events, Jackson (2006) argues that although not explicitly stated, the widely accepted major premise is that "white individuals are trustworthy witnesses." When combined with the minor premise that the officer in question is white, the conclusion that "the officer is a reliable witness" is tacitly accepted as factual.

Enthymemes are most persuasive when commonly understood, but, because the audience supplies the unstated premises, different audiences may deduce very different meanings from the same enthymeme. As such, it is possible that different people uttering the same enthymematic phrase are assuming different premises. In another example, Mercieca (2018) notes how a tweet from former President Bill Clinton championing "diversity" represented to many his commitment to "fundamental American values," whereas to the alt-right it was evidence that he embraced an anti-White, anti-American, anti-male ideology because "diversity is a Commie leftist code word."

According to Hawdon et al. (2022), enthymematic communication has been highly effective within active online hate groups. The authors observe through survey data that the closeness to a particular online community increases the likelihood of engaging with hateful content. This is because social media platforms enable the gathering of internet users around clusters of loosely connected and unstated ideas and this is where enthymemes are most powerful. It lies in the fact that the unstated premises are widely recognized and accepted. This strategic use of rhetoric enables extremists to present themselves as less extreme and connect with the emotions of potentially sympathetic audiences, ultimately expanding their support base. This means that online extremism can strike a chord with individuals who would otherwise be repelled by the overtly hateful ideas forming the basis of the argument, as the most extreme and hateful premises are left unspoken but are nonetheless understood, whether consciously or subconsciously.

4.2.4 Stylistic analysis of Implicit communication

In the previous sections we have investigated rhetorical strategies that rely on implicitness to articulate controversial ideas and potentially hateful content. Here, instead of concentrating solely on the content and the types of disruptive claims being put forth, we investigate the stylistic characteristics of implicit communication. Particularly, the use of micro-level elements such as discourse connectives and, to some extent, syntax. Although the style of communication we delve into here is frequently observed in language, we will observe that it operates interdependently with content, especially in the context of persuasive discourse on social media.

In [de Rijk \(2020\)](#)'s analysis of conspiracy-related tweets discussed in the section on Indirect Speech acts, she also demonstrates that tweets involving indirect SAs frequently omit discourse connectives and lexical cue phrases. She subsequently advances the idea that the distribution of linguistic characteristics associated with Indirect Speech Acts in her corpus becomes a central aspect of investigation and that the absence of explicit discourse connectives could be leveraged to examine misinformation.

In fact, the limited use of discourse connectives has also been a significant area of focus in discourse analysis. In an experimental study exploring the discursive structures that contribute to text coherence and ease of comprehension, [Kamalski et al. \(2008\)](#) discovered that texts with minimal usage of discourse connectives were more easily processed by the participants. Their study shows that *forewarning* factors, such as using syntactic and explicit cues to link textual segments, can result in resistance in engaging with content. While non-implicit methods of establishing coherence aid in reading comprehension in informative texts, they observe that in persuasive texts, the presence of connectives can trigger resistance in the reader, as they become aware of the persuasive intent. This is what psychologists refer to as the *forewarning effect*.

While politicians may strategically navigate around this forewarning effect, [Pastor et al. \(2024\)](#) note that social media comment sections generally tend to use fewer explicit discourse connectives. By making use of discourse coherence relations and discourse parsing systems, the authors observe that distinct patterns of coherence relations emerge from comment sections compared to regular opinion pieces and news articles. Consistent with [de Rijk \(2020\)](#) analysis of Indirect Speech Acts, the predominant patterns of coherence relations that emerge involve implicit connectives. These particular relations are referred to as JOINT relations : A JOINT relation unites two loosely connected segments of texts without establishing any clear semantic or pragmatic relations ([Stede et al., 2017](#)). Yet there is still a coherent link as they contribute to the overall text function.

The study then progresses to analyze various grammatical characteristics of JOINT relations and their relevance in persuasive texts. For instance JOINT relations can be described by *asyndetic connections*, which involve the juxtaposition of sentences where there are no shared grammatical or lexical characteristics.

(2) [Keep stirring that pot.] ← JOINT → [Divide and concur.]

Here, even though the two sentences are unrelated, the connection is made clear: we comprehend that this comment implies a particular theory where immigration serves the unclear con-

quering interests of the government. Asyndetic constructions such as this one typically get the reader to speculate on the possible connections between the sentences, even if they might be none. Another typical formation involving JOINT relations can be characterized by *connection by sequences*, where segments of text linked share common grammatical characteristics and imply a temporal or causal relation, depicting real world events.

(3) [Dry them off,] ← JOINT → [give them a hot meal,] ← JOINT → [and re-send them a Muslim country.]

In connections based on sequences a micro-narrative is often presented. We typically observe that the tense of the verbs remains consistent, and the sequence follows a regular chronological order. In examples where three or more sentences/clauses are juxtaposed by JOINT relations, the last element of the sequence often carries a climactic effect. It is likely to be introduced by the discourse marker, which aims to emphasize an inevitable conclusive point.

By using these stylistic processes, the persuasive intent is not taken too seriously. Therefore, the authors of such tweets appear to have no clear motives, which allows them to make controversial claims that are perceived as jokes while distancing themselves from the claims. This type of communication, which has solidified into a particular style used for discussing controversial ideas and statements, can be effectively described with the rhetorical concept of *parataxis*, as we will introduce in the following section.

4.2.5 Parataxis

Parataxis has long been recognized as a rhetorical strategy since the earliest days of political discourse, from the drafting of political texts in ancient Athens to the public speaking traditions in ancient Rome. While some associate it with a poetic vision through subconscious evocations (Adorno et al. (1995)), others describe it as a simplistic and decadent form of discourse, specifically tailored for an illiterate populace (Antoine (1899)). Despite its long history and the debate surrounding its purpose, there has been consensus on its definition: it is a literary technique that involves placing loosely connected short and simple sentences side by side.

Most importantly, Theye and Melling (2018) emphasize that this type of paratactic communication reflects the distinctive persona of a politician and seeks to position them socially in order to gain greater acceptance in specific contexts. More specifically, this is achieved through how parataxis is often perceived as “straight talk,” which can encompass a wide variety of small-scale transgressions, such as exaggerations, boastful language, inaccuracies, and digressions. Indeed, linguistic precautions such as “political correctness” are often seen to restrict freedom of expression, and politicians who appear to speak candidly are setting new discursive standards that enable people to identify with them. As stated in Theye and Melling (2018), “Though this straight talk style might seem like the strategy through which candidates can win elections, politicians’ elevation of political correctness to exigence also gives the style an entelechial endpoint”.

In addition to the social positioning effect facilitated by this specific style, parataxis is also ideal for simplifying complex issues and effectively communicating controversial ideas. Similar to what we saw in the previous section, the following paratactic structure evokes the sense of a build up towards a conclusive endpoint:

(4) [Put everybody in a refugee camp for few years.] ← JOINT → [allow only legit refugees in,,] ← JOINT → [send everyone else back.] ← EVALUATION → [Problem solved.]

In (4), the comment constructs a narrative with a consequential structure based on syntactic parallelism. The first three text segments all begin with a verb and present the same clause structure, while the concluding statement uses the present perfect tense, indicating a perfective aspect. What is noteworthy to highlight is that there is a distinct break in the stylistic flow of coherence, which serves to isolate, with stylistic emphasis, the final sentence. In this case, this type of buildup implies a sequential flow with a natural conclusive point, aiming to suggest that the solution to immigration is straightforward and that actual policies are evasive.

In a similar way to the forewarning effect, [Olbrechts-Tyteca \(1969\)](#) have observed what they refer to as *charge of device*. They describe it as “the impression that a device is being used when the speaker seems to adopt rules or techniques which, because they are too uniform or too far-fetched, do not seem to fit the object in an altogether natural manner.” Conversely, parataxis is perceived as unburdened by any charge of device, as it generates a rapid sequence of loosely connected statements that appear spontaneous and sincere. In political discourse, this fosters the perception that complex problems can be resolved with a one-line solution, often presented as a natural conclusion arising from a consequential flow of implicitly connected statements.

4.3 Challenges and future research directions

Through the application of rhetorical strategies operating at the content and the stylistic level, we have seen how politicians aiming to achieve virality can effectively navigate sensitive and controversial topics on social media. We have also described how this approach allows them to maintain plausible deniability, while forging connections with specific audiences. For each topic discussed in this section, we envision the following research directions:

- **Indirect Speech Acts:** Speech act theory has proven to be an effective tool for annotating and comprehending implicit communication, particularly in the context of conspiracy-related content. For indirect speech acts, [de Rijk](#) suggests that studies should broaden their scope beyond conspiracy-related content and investigate the extent to which Indirect SAs may be relevant in general misinformation. It is also suggested that NLP approaches should differentiate between information and disinformation by incorporating discourse features based indirect speech act annotations.
- **Enthymematic Communication:** [Hawdon et al. \(2022\)](#) found that enthymematic communication thrives in online hate groups, as individuals feel closer to these communities and are more likely to engage with hateful content. Social media platforms facilitate the dissemination of unstated ideas, making enthymemes powerful tools for extremists to connect with sympathetic audiences by presenting themselves as less extreme. This strategic use of rhetoric allows online extremists to expand their support base, reaching individuals who may otherwise be repelled by overtly hateful ideas, as the most extreme premises remain unspoken but understood. To understand how enthymematic communication functions and how evocation becomes effective in certain contexts, the authors suggest that further re-

search should examine how online hate activism evolves in response to specific political events. This approach would aid in understanding whether online hate activism originates primarily from the affordances of social media or if it is influenced by external factors. They also propose the idea that multiple surveys, similar to the one they conducted, should be made to identify patterns and determine if the results can be generalized. One generalizable finding from their results could be the systematic connection between online hate activism and individuals' personal sense of community and proximity to groups such as religious organizations, associations, or even family.

- **Stylistic analysis of Implicit communications:** Coherence relations and the specific cues through which they are indicated have been identified as key elements in describing content dealing with persuasion techniques. This communication style often aims to implicitly convey controversial content through pragmatic-stylistic effects. Future research could explore the extent to which the presented discursive strategies are prevalent in comment sections. Additionally, it could further evaluate the performance of discourse parsers in analyzing these structures to determine the feasibility of automating such analyses.
- **Parataxis:** Parataxis, demonstrated through syntactic parallelism and a sequential buildup, creates a stylistic break that emphasizes the final statement, implying a straightforward solution to immigration. We have seen how this rapid sequence of loosely connected statements fosters a perception of sincerity in political discourse, suggesting simple resolutions to complex issues, presented as natural conclusions. Future research should investigate the cognitive processing mechanisms involved in audience reception of paratactic rhetoric in political contexts, aiming to understand how this rhetorical strategy influences perceptions of authenticity, credibility, and persuasiveness. Specifically, it should further contextualize parataxis with concepts such as the forewarning effect and charge of device.

5 Populist Communication and its Community Impact

From a historical standpoint, populism manifests as a fluctuating trend. With cyclical waves, it surfaced on the political sphere to fade away and come back. Fig.3 illustrates the historical evolution of populism over the past century, showing fluctuations in the proportion of countries ruled by populist leaders (Funke et al., 2023). Since 1916, year of the first populist president in Argentina, there have been two main peaks: during the Great Depression of the 1930s and in the 2010s, while the 1980s marked a low point. Since the 1990s, populism has seen a resurgence, with a notable peak in 2018 attributed to the rise of a new populist right-wing movement in Europe and beyond.

Funke et al. (2023), through a longitudinal cross-study among 50 countries, showed how populist leadership incurs significant economic costs, resulting in a pronounced long-term decline in consumption and output. Hadiz and Chryssogelos (2017) linked populism with damages on international relations. Vittori (2022) emphasizes how populism engenders democracies, especially by neglecting minority rights. Finally, Egelhofer et al. (2022); Moffitt and Tormey (2014a) highlight how populism contributes to the erosion of trust in media and institutions. This, in turn, can lead to an increase spread of disinformation (Ehrke et al., 2023), with negative effects on community's well being (Kemei et al., 2022).

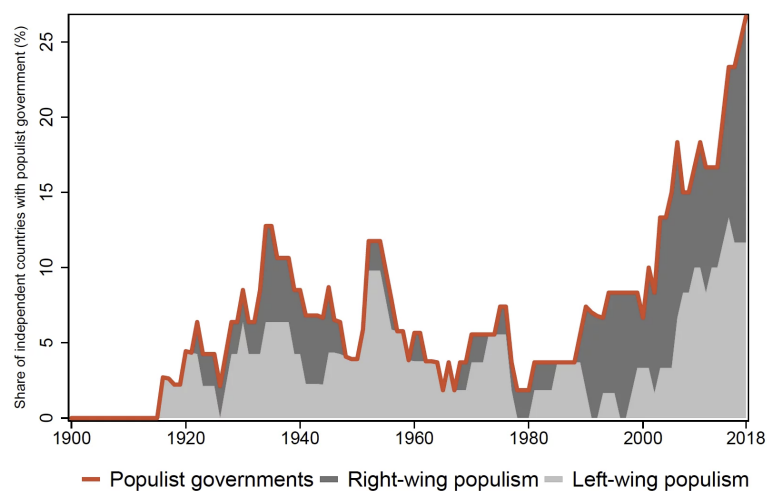


Figure 3: Populists in power, (Funke et al., 2023).

The development of social networks (SNs henceforth) has created a particularly fertile environment for the growth and spread of populism. Hopster (2021) identifies four social network's populist affordance to explain this elective affinity:

- SNs allow citizens and politicians to **circumvent editorial filters**. This, in turn, empowers populist leaders to bypass traditional editorial gatekeepers, thus amplifying uncommon and extreme political voices (Dittrich, 2017; Engesser et al., 2017; Owen, 2020).
- SNs **algorithms prioritize content that maximizes user engagement**, such as views and interactions, fostering an 'attention economy' where sensational and populist messages gain visibility. This system benefits populist communication styles, characterized by urgency

and appeals to crisis, by promoting content that resonates widely and has viral potential, regardless of newsworthiness (Bennett and Seyis, 2023).

- SNs platforms feature low-level affordances that naturally support populist communication styles, including Twitter's **preference for brief, bold messages** and Facebook's promotion of emotional interactions (Oliver and Rahn, 2016). These platforms' algorithms further amplify personalized and emotive content from populist leaders, enhancing visibility and engagement. This environment not only encourages a populist style among all political actors but also contributes to a broader shift in political discourse toward sensationalism, paralleling a rise in tabloid-style engagement across the media landscape (Hameleers et al., 2017).
- SNs offer **tools for politicians to gauge and respond to the public's general will**, providing real-time data on public opinion. This allow them to align closely with the popular will, as seen in populist parties' support for direct democracy initiatives (Gerbaudo, 2018a).

Taken together, the findings regarding the negative effects of populism, and the potentialities offered to it by SNs, amplify concerns about populism in contemporary governance. The scientific community, recognizing the issue, has demonstrated a growing interest in elucidating this phenomenon (Hunger and Paxton, 2022). However, despite the extensive studies, the concept of populism remains elusive and contested among scholars, highlighting a lack of consensus on its core characteristics (Hunger and Paxton, 2022).

Consequently, this section presents a synthesis of the two primary theoretical two lenses through which populism communication and its influence can be examined: as an *ideology* (Section 5.1, or a *rhetorical style* (Section 5.2). Our overall goal is thus to observe how populism communication operates and impacts communities through SNs.

This exploration is further enriched by the presentation of two case studies — one for each approach — providing concrete examples that illuminate the distinct but interconnected facets of populist discourse: revealing how the phenomenon is manifested both in the substantive messages content and in the modality through which they're delivered. Finally, building on these theoretical foundations, Sec.5.3 examines the impact of populist discourses on SNs users by deepening the interplay between populist-specific communication styles, audience reactions, and the unique affordances of SNs for populist discourse.

5.1 Populism as an Ideology

In this section we present the ideological perspective on populism and how it can be operationalized and applied to real-world scenarios. An ideology can be understood as an abstract-codified sets of principles that serve as ideal types, explaining the coherent, often complex belief systems of groups or individuals (Müller, 2023).

Engesser et al. (2017) organized populist ideology around 5 core themes which, in turn, can be connected through specific "narrative rules". Fig.4 provides an at-glance understanding of this conceptualization. The "Populist", "Sovereignty", "Elite", "Others, and "People, are the five core themes, while the different arrows represent possible interconnections.

Indeed, populists claim to advocate for “People’s” sovereignty, which they allege has been deprived by the “Elite” to favor “the Others”. For this reason the establishment politicians are seen as betraying the people to favor their “rivals”. The populist’s goal of people sovereignty’s restoration is narrated, on one side, as a challenge to the “Elite”. On “the Others” side, it works jeopardizing the “usurpers sovereignty” through exclusion practices, or hindering their inclusion⁴. This “ideal” structure can be understood as a matrix whose “empirical products” manifest as a multitude of different contents. Consequently, the study of populism as an ideology is strongly oriented towards the observation of the content dimension.

In this respect, [Mudde \(2004\)](#) characterize populism as a “thin ideology,” because of populist ability to amalgamate with various ideologies (e.g., conservative, socialist) and a wide range of political theories and thoughts ([Hunger and Paxton, 2022](#); [Mudde, 2004](#)). The “thin” morphological structure of populist ideology, being limited to this narrative architecture, does not offer a wide-ranging or exhaustive array of solutions to the political issues societies face. To fill these “gaps”, populism draws from other ideologies, where it gets sometimes assimilated into (e.g., nativism prominent in Europe, socialism in Latin America) ([Snieckute, 2020](#)).

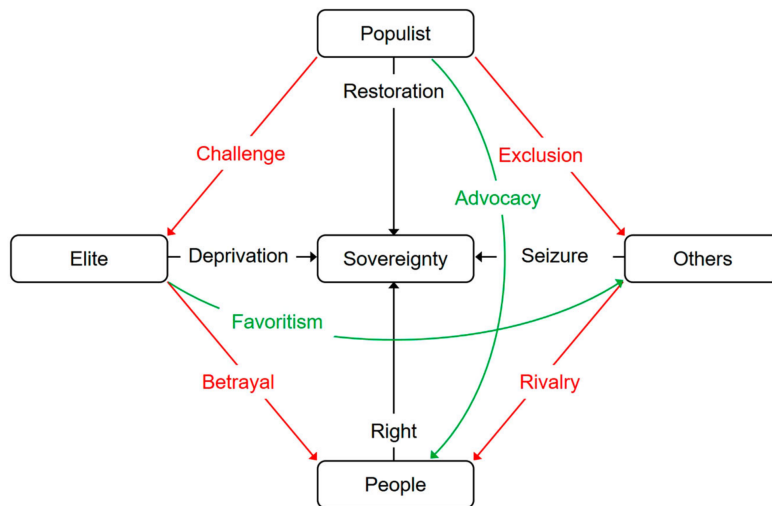


Figure 4: Heuristic model of populist ideology from [Engesser et al. \(2017\)](#).

The ideological approach has strongly impacted academic circles due to its potential in providing a general understanding of populism, providing scholar with a common reference to orient and share their research efforts ([Pauwels, 2014](#)). Being focused on examining content produced by populist actors, this approach entails uncovering the presence of key populist elements, interwoven according to the logic of populism (as depicted in Fig.4). Moreover, as showed by [Müller \(2023\)](#) the ideological conceptualization of populism, being coherent with how populist actions, behaviors and speeches, can provide useful explanatory insights and suggestions for research hypothesis.

⁴Additional explanation and examples of each “nucleus” of the populist ideology are provided in the following section (see Tab.2).

5.1.1 Case Study: From the Ideology to the Index

As outlined above, despite originating from the same “narrative architecture”, to “fill” their narrative populist discourse draws specific contents from the political-cultural context in which they are placed. To study populist content in Social Networks, [Mazzoleni and Bracciale \(2018\)](#) operationalized the ideological definition of populism in a three dimensions taxonomy. Each dimension corresponds to a core topic of the ideology, namely: the People, the Others and the Elite. The first two were further unpacked in sub-categories to increase the precision of the observation. Tab.2 describes this taxonomy: each category or sub-category is briefly explained and accompanied with an example from the text analyzed by [Mazzoleni and Bracciale \(2018\)](#). They studied Facebook posts of 5 Italian politicians: three considered by the literature as populist, namely Luigi di Maio- ‘Movimento 5 Stelle’ (283 posts), Silvio Berlusconi - ‘Forza Italia’ (66 posts), Matteo Salvini- ‘Lega Nord’ (213 posts), and two non-populist, namely Matteo Renzi - ‘Partito Democratico’ (156 posts) and Giorgia Meloni - ‘Fratelli d’Italia’ (216 posts). A total of 934 posts was collected, covering a time span between October 2016 to October 2017.

Relying on the fact that it can be very difficult to identify specific actors as sheer populists or nonpopulists, the authors created a “Populist Index” to assess the degree of populist ideology expressed by politicians on Facebook. The assumption was that delineating populism into various levels (light, moderate, and bold) can elucidate leaders’ communicative decisions in accordance with their objectives and strategies. This four-level index focuses on whether core populist themes are present, not the individual occurrences of each theme.

1. Zero-level Populism (0): no populist themes appear.
2. Soft Populism (1): indicates the presence of at least one theme.
3. Moderate Populism (2): signifies the combination of at least two distinct themes
4. Bold Populism (3): signifies the presence of all major populist themes within a single post

Overall, results showed that ‘Soft Populism’ is the communicative trait of Italian mainstream parties (‘Partito Democratico’, ‘Forza Italia’), ‘Moderate Populism’ is the communicative characteristic of the opposition (‘Fratelli d’Italia’, ‘Lega Nord’ and ‘Movimento 5 Stelle’) and ‘Bold Populism’ is characteristic of extreme-right parties. This trend arguably relates to the intimate core of populism: the ‘Us vs Them’ narrative, in fact, works perfectly as a tool of attack and contraposition that fosters the ‘hyper-mediatisation’ of the populist messages and fits with opposition parties’ aims ([Mazzoleni and Bracciale, 2018](#)). An additional interesting result is that populist contents are present across the different actors analyzed. Also non-populist political actors like Matteo Renzi use fragments of populist ideology in their posts. The authors trace back these results to an ‘endemic populism’, i.e. an ‘overflow’ of populist discourse into the conventional political discourse. This, in turn, advocates for a strategic use of populist ideology and, thus, for its characterization as ‘thin ideology’ ([Mazzoleni and Bracciale, 2018](#)). Despite using a slightly different taxonomy, the study reached similar conclusions analyzing 4 parties Twitter posts during the 2019 Spanish election campaign, finding a cross-political leaning trend in using populist contents.

[Mazzoleni and Bracciale \(2018\)](#) provide an in-depth analysis of the outcomes of their study. Here, we selectively presented only the most pertinent ones to underscore the potential of em-

playing an ideological perspective in scrutinizing the dissemination of populism on social media platforms.

Category	Description & Examples
The People	<p>Sovereignty: Populists depict “the people” as the ultimate democratic sovereign betrayed by the elites. This allows the populist politicians to depict themselves as the advocates of the people’s rights. <i>Example: “I think that it will be a good day for the Italians the day when they can recover their sovereignty and can choose their governments and when we will no longer have an executive of lobbies and Merkel but a government of Italy and the Italians.” - Giorgia Meloni, 24/11/16</i></p> <p>Class: Populists frame people as underprivileged citizens distinct from elites in terms of economic status, education, and access to power. <i>Example: “Here we are witnessing the massacre of a generation now supposed to split the world. The members of parliament don’t give a toss about all this because they get high pensions at the age of 65 for having worked for just four and a half years.” - Luigi di Maio, 2/2/2017</i></p> <p>Nation/Ethnic Group: “The people” refers to a national or ethnic community with emphasis on native population membership as the key distinction. <i>Example: “Illegal immigrants are enjoying hotel accommodation, while Italians are homeless.” - Matteo Salvini, 14/11/2016</i></p> <p>Cultural: “The people” signifies “our people,” defined by shared values, religion, history, and customs rather than formal citizenship or ethnicity. <i>Example: “I would NOT concede even half a square metre to the Islamic communities until they sign black on white that WOMEN have the same RIGHTS as men. Is that understood? THIS IS HOW IT WORKS AT OUR HOME!” - Matteo Salvini, 6/6/2017</i></p>
The Elite	<p>Opposition to elites and the establishment, emphasizing the divide between “us” (the common people) and “them” (the perceived adversaries led by the ruling elite). This divide is central to populist discourse, which often targets various elite groups (such as politicians, the media, or economic powers) as the enemy. <i>Example: “In the U.S. the people won against the elites, the oligarchies, high finance, and the dominant ideology. The same will happen in Italy on 4 December: Italian citizens will not listen to radical-chic people.” - Giorgia Meloni, 9/9/2016</i></p>
The Others	<p>Dangerous Others: Populist actors pinpoint certain groups within the populace as internal foes, marking and ostracizing these segments from the broader community. <i>Example: “Three bastard fake refugees, supported by Italians’ money, raped a 14-year-old girl in Trieste... Enough with out of control immigration, enough with being fooled by illegal immigrants passed off as refugees.” - Giorgia Meloni, 28/7/2017</i></p> <p>Authoritarianism: This is when political figures advocate for strict or undemocratic policies to counter individuals or groups perceived as threats to the unity of the people. <i>Example: “Chemical castration for paedophiles and rapists, whether Italian or foreign, as in many civilized countries: you do it once and you’ll never do it again!” - Matteo Salvini, 1/3/2017</i></p>

Table 2: Populist Content Taxonomy, according to Mazzoleni and Bracciale (2018).

NLP and ML Methods for Populist Content Assessment The presented case study leveraged manual annotation. Given the burgeoning volume of online interactions, relying solely on this approach is unfeasible. For this reason, in Table 3 we provide a synthesis of NLP methods applied to the automatic detection of populist contents. All these different methods have the potential to make more efficient the analysis of online produced text. To our knowledge, only Gründl (2022) actually applied theme to social networks.

- **Dictionary-Based:** involves assembling specialized dictionaries. It operationalizes concepts by counting occurrences of specific words chosen to indicate the presence of populist discourse.
 - *Pro:* it offers perfect reliability due to its automated nature, ensuring consistency across analyses without human coder bias.

- *Contra*: validity depends on the content of the dictionaries, which can be subjective. Choosing words for dictionaries involves human bias, potentially affecting the method's objectivity.
- **Supervised Machine Learning (ML)**: This method employs annotated data to train algorithms to identify patterns indicative of populist discourse. Supervised ML relies on labeled datasets to learn the associations between textual features and the concepts being studied.
 - *Pro*: Capable of capturing complex patterns and nuances beyond pre-defined dictionaries, thus potentially enhancing the predictive accuracy and the validity of the analysis.
 - *Contra*: Its efficacy is contingent on the quality and representativeness of the training data; any biases in the labeled dataset may be learned and perpetuated by the model.
- **Unsupervised Machine Learning (ML)**: Utilizes algorithms to discover latent structures and patterns within unlabelled data, categorizing text based on the inherent similarities without prior annotations.
 - *Pro*: Capable of revealing unexpected patterns and themes in the data, providing insights without preconceived notions of what constitutes populist discourse.
 - *Contra*: The lack of labeled data can lead to difficulties in validating and interpreting the clusters or patterns identified, potentially resulting in less definitive conclusions about the presence of populism (Jankowski and Huber, 2023).

Author(s)	Method	Sampling Unit(s)	Coding Unit	Themes/Index
Gründl (2022)	Dictionary-Based	Twitter and Facebook posts	Words	1.People 2.Anti-Elite
Oliver and Rahn (2016)	Dictionary-Based	2016 US primary campaign speeches	Words and phrases	1.People 2.Anti-elite
Rooduijn and Pauwels (2011)	Dictionary-Based	Party manifestos	I. Paragraph II. Word	1. People 2.Anti-elite
Pauwels (2011)	Dictionary-Based	Party manifestos and magazines	Words	N/A
Bonikowski and Gidron (2016)	Dictionary-Based/ Holistic Grading	US presidential campaign speeches	Words (n-grams)	Unspecified
Aslanidis (2018)	CBSTA ⁵	Grassroots speeches	Semantic triplets	1.People 2.Anti-Elite

Continued on next page

⁵Clause-Based Semantic Analysis (CBSTA) doesn't fall in any of the categories. Utilizing a semantic grammar, original text is reformulated into a set of clauses known as semantic triplets. This approach produces perfectly comparable coding units in any language by relying on the elementary syntactic components of language: subject, verb, and (optionally) object.

Table 3 – *Continued from previous page*

Author(s)	Method	Sampling Unit(s)	Coding Unit	Themes/Index
Di Cocco and Monechi (2022)	Supervised ML	Political Manifestos	Sentence	1.People 2.Anti-Elite
Ulinskaitė and Pukelis (2021)	Unsupervised ML	Debates + Political Manifestos	Paragraph	1.People 2.Anti-Elitism
Bonikowski et al. (2022)	Unsupervised ML	USA Political Speeches	N/A	1.Anti-Elité 2.The Oth- ers(Exclusion)

Table 3: Scholarship on quantitative content analysis of populism.

Adapted from [Aslanidis \(2018\)](#).

5.2 Populism as a Communicative Style

For effective transmission through SNs, the substance of populist ideology must be suitably encapsulated, suggesting that beyond its ideological underpinnings, populism's distinct communicative style is crucial for comprehending its impact ([Hunger and Paxton, 2022](#); [Zarefsky and Mohammed, 2020](#)). In this section we present how populism can be studied from a stylistic point of view, as well as how this approach can be integrated with the ideological one.

[Bourdieu \(1991\)](#) defines style as “a set of systematic differences, apprehended syncretically”. This social distinctiveness understanding of style is particularly appropriate for populism. In fact, its peculiar communicative performance precisely hinges on emphasizing distinctions to articulate socio-cultural identities, both by manifesting proximity to “the people” and distinction from the political establishment; thereby becoming a critical element of their allure ([Ekström et al., 2018](#)).

On a methodological level, this research venue, encompasses the analysis of features like populist emotional, simplistic or the colloquial tone, as well as rhetoric and argumentative strategies ([Koller et al., 2023](#); [McDonnell and Ondelli, 2022](#); [Moffitt and Tormey, 2014b](#)) (see Sec.5.2.1 for a nuanced list).

Stylistic and ideological perspective are not mutually exclusive, since they can be integrated. At the same time, as stressed by [Engesser et al. \(2017\)](#); [Ernst et al. \(2019a\)](#); [Bracciale et al. \(2021\)](#), conducting rigorous analysis integrating these two approaches requires maintaining clear conceptual boundaries, not confusing between each other. Thus, drawing from [Engesser et al. \(2017\)](#), we distinguish them in terms of “focuses”. The ideological approach understands populism as a set of ideas, this implies focusing on the *content* of the populist messages, i.e. answering to the question “*What*”. The stylist approach, on the other hand, conceives populism as mode of representation, thus is focused on the *form* of populist communication, i.e. answering the question “*How*”.

Adopting this perspective would require to harness all the spectrum of possible communicative means, spanning from linguistic and argumentative aspects of media and social networks interac-

tions, to prosodic and visual elements of public debates and interviews. Venizelos (2023)'s study on Trump communication performance exemplifies this comprehensive approach by integrating these elements with ethnographic inquiries, thereby offering an exhaustive analysis of Trump's charismatic and transgressive style, as well as its impact on the audience.

However, such a fine-grained approach is hardly scalable to the worldwide dimension of populism. For this reason, in this section we will focus mainly on studies aimed at analyzing the linguistic dimension of populist style, mainly in social networks. This choice relies on two factors. Firstly, because Social Networks provides the best conditions for populist development (see also (Ernst et al., 2019a)). Secondly, the digitization of interaction offered by social networks allows to apply ML (machine learning) and AI (artificial intelligence) algorithms to automatically carry out the analysis, improving (exponentially) the efficiency and the scope of the analysis. Moreover, social networks provide a unique environment to observe how the populist narrative style impacts on the audience in terms of, for example, engagement, transmission and popularity.

5.2.1 Case Study: The Populist Stylistic Toolkit

To describe this approach, we present Ernst et al. (2019a), a work that aims at providing solid foundations for the identification of populist-specific stylistic features by interweaving theoretical reflection with empirical data.

To operationalize the populist communicative style, the authors started by searching the literature to identify the stylistic features theoretically attributed to populist actors. This process led them to distinguish six populist specific stylistic devices (see Tab.4 for additional descriptions): Negativism (Álvares and Dahlgren, 2016; Block and Negrine, 2017; Bracciale and Martella, 2017; Engesser et al., 2017), Crisis Rhetoric (Neville-Shepard, 2024; Wirz et al., 2019), Emotional Tone (Block and Negrine, 2017; Bos and Brants, 2014; Bracciale and Martella, 2017; Engesser et al., 2017; Hameleers and Schmuck, 2017), Absolutism (Bos and Brants, 2014; Engesser et al., 2017; Hawkins, 2009), Patriotism (Block and Negrine, 2017; Rydgren, 2017), Colloquial Style (Alber-tazzi and McDonnell, 2007; Bracciale and Martella, 2017; Engesser et al., 2017; Moffitt, 2016) and Intimization (Bracciale and Martella, 2017; Stanyer, 2013). These specific stylistic devices, in turn, can be organized in three macro-categories, as shown in Tab.4.

Then, the research analyzed and compared the communication strategies of a wide range of political actors across six countries (Switzerland, Germany, France, Italy, the United Kingdom, and the United States) on Facebook, Twitter, and televised political talk shows, during a three-month period in 2015. In total 2,067 social media and talk show statements from 31 political parties with different stance were gathered and analyzed.

Also in this case they proceeded with manual annotations. In addition to the stylistic features, the research kept track also of the presence of populist ideological key-elements. We highlight how, despite being organized in a different way, the used taxonomy encompasses the same elements of the one presented above (Sec.5.1).

Overall, 38% of the total statements contained at least a populist feature (content or stylistic). Interestingly, populist contents and stylistic features are weakly correlated ($r = .173$, $p < .01$), and the style elements (31.3%) are used significantly more, $t(2066) = 17.35$, $p < .001$, than populist key messages (13.6%).

Additionally, the research aimed at testing two additional hypothesis. H1 assumed that the proportions of populist key messages and related style elements are higher on Facebook and Twitter (social networks) than on political talk shows (television). To test this hypothesis the researchers employed multilevel modeling. This methodology facilitated a two-level data analysis: level one comprised statements by political actors, nested within political parties at level two. The analysis focused on the prevalence of populism in these statements, defining it as the dependent variable. Independent variables consisted of dummy variables for communication channels (Facebook and Twitter), using talk shows as the baseline. The results showed how statements on Facebook and Twitter significantly predict populism-related communication. This demonstrates that the degree of populism-related communication is higher on both social media platforms than political talk shows. Moreover, this pattern is identical for style elements in isolation as well. H2, on the other hand, assumed a a greater proportion of populist key messages and related style elements than moderate parties. Similarly, to test this hypothesis, party extremism was used as key variable in the multilevel modeling. The results supported also this hypothesis, even if with slight less strength: party extremism significantly increases combined populist communication but shows only mild trend when considering contents and style separately.

Another aim of the paper was to prove the theoretical consistency of the proposed stylistic features taxonomy with empirical data. To test this hypothesis a principal component factor analysis with the seven style elements was conducted (varimax rotation, Kaiser normalization). The factor analysis identified three distinct dimensions (51.4% explained variance, $p < .001$) without any substantial cross-loadings. “Negativity” comprises negativism and crisis rhetoric; “Emotionality” includes emotional tone, absolutism, and patriotism; and “Sociability” is composed of colloquialism and intimidation.

NLP and ML Methods for Populist Style Assessment We emphasize how, also in this case, although the study focused on social media it leveraged manual annotation methodologies. To our knowledge, no prior study employed computational methods to assess these rhetorical means in populist speeches or online content. At the same time we highlight how the communicative means listed by [Ernst et al. \(2019a\)](#) refers to well-known persuasive and discursive techniques which, in turn, have been extensively studied by the NLP community. Thus, in Tab.5, we map the populist stylistic devices identified by [Ernst et al. \(2019a\)](#) onto corresponding persuasion and discursive techniques previously examined by the NLP community. Subsequently we compile a collection of NLP and ML research articles that offer models for the automated identification of these techniques. This constitutes a useful starting base for the development of future research on populist communication in social media.

5.3 Populist Communication and Social Networks: a Vicious Circle

The two previous sections addressed the communicative strategies characterizing the populist performance. We stressed how populist communication is instrumentally exploited to manifest proximity to “the people” and distinction from the political establishment. In this sense, populist communicative style challenges current codes of “doing politics”. This “distinctiveness” of populism has led scholars to assess how the mechanisms by which it is realized impact on democracy

Table 4: Conceptualization and Operationalization of Populism-Related Style Elements, adapted from Ernst et al. (2019a).

Dimension	Stylistic Devices	Underlying Style Element & Categories
Negativity	Negativism	Negatively labeling targets by attributing negative characteristics (e.g. criminal, lazy, malevolent) or condemning actions/situations with negative outcome.
	Crisis rhetoric	Portraying a situation/development as a crisis using exaggerations, emergency rhetoric or declaring a scandal.
Emotionality	Emotional tone	Sharing positive and negative emotions or revealing feelings.
	Absolutism	Using an assertive tone and lacking relativizing words. Tendency to paint world in black and white, presenting something as the only conceivable option.
Sociability	Patriotism	Emphasizing of the superiority of own country by referencing an idealized and utopic heartland.
	Colloquialism	Preference for a simple dialect, colloquial or vulgar language, and use of nicknames to reach the ordinary people.
	Intimization	Recounting personal and intimate details about personal life.

Table 5: Computational Methods for the Automatic Detection of Populist Stylistic Devices

Dimension	Stylistic Devices	Persuasion/Discursive Technique: Papers
Negativity	Negativism	"Name-Calling/Labeling"; "Doubt"; "Tu quoque": (San Martino et al., 2019, 2020a,b; Habernal et al., 2018; Nikolaidis et al., 2023; Piskorski et al., 2023a,b); "Ad Hominem": Delobelle et al. (2019)
	Crisis rhetoric	"Exaggeration"; "Appeal to Fear"; "Appeal to Strong Emotions": (San Martino et al., 2019, 2020a,b; Nikolaidis et al., 2023; Piskorski et al., 2023a,b); "Appeal to time": (Nikolaidis et al., 2023; Piskorski et al., 2023a,b)
Emotionality	Emotional tone	"Loaded Language": (Nikolaidis et al., 2023; Piskorski et al., 2023a,b; San Martino et al., 2019, 2020a,b)
	Absolutism	"Black&White Fallacy"; "Thought-Terminating Cliché"; "Causal Oversimplification": (San Martino et al., 2019, 2020a,b; Nikolaidis et al., 2023; Piskorski et al., 2023a,b); "False Dilemma/No Choice": (Nikolaidis et al., 2023; Piskorski et al., 2023a,b)
Sociability	Patriotism	"Flag-Waving": (San Martino et al., 2019, 2020a,b; Nikolaidis et al., 2023; Piskorski et al., 2023a,b)
	Colloquialism	"Slang": (Pei et al., 2019); "Informal Language": (Bugnard et al., 2021)
	Intimization	"Intimacy": (Yuan and Chen, 2023; Pei and Jurgens, 2020)

(van Klink and van der Geest, 2020).

Mouffe (2018); Laclau (2005) contend that populism has a “dual nature,” serving both as a democratic corrective by bringing to light issues ignored by the elite, and as a threat by undermining institutions that protect minority rights. Nonetheless, considering the comprehensive examination in prior sections, our viewpoint aligns more closely with the argument presented by Müller (2017). This perspective suggests that populism’s inherent logic distorts democratic processes, ultimately corroding them from within. This is evident in both Sec.5.1 and Sec.5.2, where populism is shown to rely on an unequivocal exclusion of “the others,” fostering a polarized societal divide (Snieckute, 2020).

In this regard, Zanotti (2022) emphasizes that populist politicians purposefully escalate affective polarization in the polity to sustain the mobilization of their supporters. This continued mobilization is not achieved through programmatic alignment but rather through feelings of in-group and out-group discrimination. While a degree of ideological polarization is necessary for a healthy democratic representation (Zanotti, 2019; Morgan, 2011), McCoy et al. (2018) point out how, recently, polarization has been entailed mainly by strong emotional components rather than ideological ones. Affective polarization, in turn, refers to an *hostility* towards the opposing political parties, reducing the objective pluralism of society into a single dimension of “us VS them”. Accordingly, this logic drives the distinctions between groups to graft on mutually exclusive identities and individuals are forced to chose a side or be labelled as “the others”.

Furthermore, the emotional dimension of this “in-group/out-group” dichotomy results in the delegitimization of the out-group, raising substantial issues as it may lead to a reduction in citizens’ willingness to interact with opposing political viewpoints and recognize others’ democratic rights. Typically, this pattern turns political discourse into acrimonious debates and, in its extreme manifestation, as illustrated by the 2021 USA riots, it can precipitate a refusal to accept electoral outcomes Zanotti (2022); McCoy et al. (2018). This dynamic is outlined as “Path A” in Fig.5.

Turning our attention to SNs, we detail at the beginning of Section 5 four affordances of SNs that cater to populist messaging. The examination of populist communication strategies reveals a clear alignment between their emotive and divisive traits and the goals and communication techniques preferred by SNs. Thus, the latter end up by rewarding populist actor adopting this style, generating a self-reinforcing process as depicted in Fig.5 “Path B”.

Finally, we highlight an additional mechanism. As depicted in Fig.5 “Path C”, online controversies, fueled by populist driven polarization, constitute an additional advantage point for populist actors’ online activity. Venturini and Munk (2021) highlights that, in the fierce battle for public attention on SNs, controversies confer a competitive edge. The cognitive mechanisms inherent to controversies, such as judgment, evaluation, and decision-making, facilitate the elicitation of strong emotional responses (Storbeck and Clore, 2008), thereby capturing individuals’ attention. This enhanced attentiveness, in turn, increases the likelihood of controversies being disseminated widely across social networks (Berger and Milkman, 2012; Guadagno et al., 2013). De Bruycker and Rooduijn (2021); Ernst et al. (2019b); Blassnig et al. (2019) showed how populist actors strategically use controversies and contentious subjects to direct user attention and leverage the capabilities of SNs for their benefit, aligning with the alt-right’s maxim “*conflict is attention and attention is influence*” (Marantz, 2020).

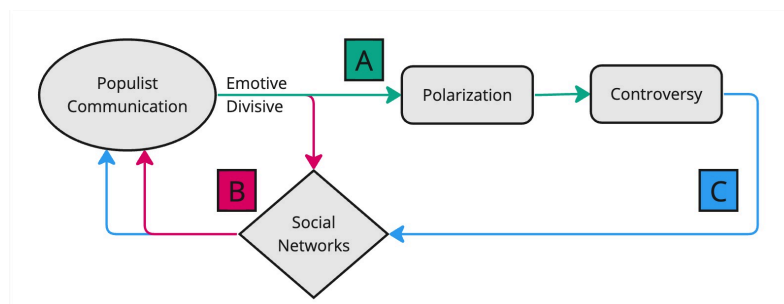


Figure 5: Populist Communication and Social Networks Interactions

5.4 Challenges and Future research directions

In this section we examined the propagation mechanisms of populism and its multifaceted impacts, particularly emphasizing the crucial role of social networks. The analysis encompassed both the ideological and communicative style perspectives. In Sec.5.1, the ideological perspective, focusing on the “what” of populist messages, helped identify the core themes of populist messages and their interrelations, revealing a strong reliance on divisive arguments that separate “the (voting) people” from “the others” and “populist politicians” from the “establishment actors.” Then, through a case study (Sec.5.1.1), we demonstrated how to apply this approach in practice. The results revealed an “endemic populism”, i.e. the spread of populist rhetoric across actors, regardless of their political stance. This bolsters the perspective of populism as a ‘thin ideology,’ employed by politicians more as a strategic tool than as a foundation for setting political agendas.

Subsequently, in Sec.5.2, the focus shifted to the conveyance, the “how” of these divisive narratives, highlighting the strategic use of rhetorical devices to amplify negative sentiments towards opponents while nurturing positive emotions towards the “in-group.” Sec.5.2.1 further exemplifies this approach through a case study that demonstrates how to apply and integrate both the ideological and stylistic perspectives of populism. Moreover, the study highlighted how populist actors significantly prefer the use of SNs over “traditional media”. Discussing this analytical approach shed light on how the strategies populism employs to secure support ends up by deepening societal cleavages. This last point was then further addressed by presenting a critical analysis of the synergistic relationship between populist communication and SNs. We outlined how this interaction contributes to a “vicious circle” where populist strategies and SNs’ dynamics feed into each other, exacerbating the challenges to democratic discourse and governance.

It is important to note that there is scant evidence to suggest that social media and social network companies deliberately aim to promote populism. On the contrary, these companies often advocate a neoliberal perspective, which is frequently contested by populist movements (Gerbaudo, 2018b). Therefore, the relationship between social media companies and populism can more accurately be characterized as an unintentional alignment, particularly with respect to the ideological leanings of these corporations. However, it is crucial to recognize that, on the other side of this elective affinity, populist leaders actively harness the structural features of social media platforms. This generates a symbiotic dynamic: social media environments provide opportunities for populist expressions, and simultaneously, populism generates content that ben-

efits the platforms by fostering increased user engagement and SNs popularity Engesser et al. (2017). Therefore, it is unlikely that social media companies will adjust their algorithms against populism. This shared interest may solidify populism's global political presence. Hence, given the aforementioned considerations and the significant ramifications stemming from populism, as elucidated by Nannini et al. (2024), two imperative actions should be undertaken:

Firstly, there is a pressing need for institutions to engage in collaborative efforts aimed at monitoring and regulating the interactions between political entities and stakeholders within SNs. Such cooperation is essential for ensuring transparency, accountability, and the preservation of democratic principles in the face of emerging challenges posed by digital platforms.

Secondly, it falls upon the scientific community to undertake the responsibility of generating knowledge and developing tools that enable institutions to enact their regulatory measures effectively. By leveraging scientific expertise and advancements, institutions can enhance their capacity to navigate the complexities of the digital landscape and mitigate the risks associated with misinformation, polarization, and the manipulation of public opinion.

Concerning possible research directions that can be pursued by the research community in the future, we identify the following:

- **Multi modality:** the present survey was limited to textual messages, yet contemporary interactions on SNs predominantly utilize multi-modal forms, including images, audio, and video. Therefore, it is imperative for future research to develop methodologies and theoretical frameworks that also encompass these aspects.
- **Impact Assessment:** we described how populist communication, when interacting with SNs can amplify societal divisions eroding democratic engagement. Based on these observations, it is imperative to pursue the following future research avenues:
 - **Evaluating User Responses to Populist Strategies:** Investigating how users react to varying degrees of populist communicative strategies is crucial. This entails exploring the relationship between the intensity of populist messaging and its reception among different user demographics, aiming to understand the effectiveness and reach of such strategies.
 - **Cross-Platform Comparative Studies:** Given the varied nature of social media platforms, conducting comparative analyses across these platforms can reveal how different environments influence the spread and impact of populist narratives. This research could uncover platform-specific dynamics that either amplify or mitigate the effects of populist communication.
- **Populist Counter-Measures:** The significance of mitigating the effects of populist communication necessitates the exploration of effective counter-strategies. Traditional approaches, such as fact-checking, which seeks to debunk populist messages to influence public perception of the politicians involved, have been critiqued for their limited efficacy and methodological challenges, particularly in verifying future-oriented claims (Plug and Wagemans, 2020).

In response, we propose two innovative methodologies that could guide future research aimed at neutralizing the adverse impacts of populism:

- Argumentative Checking: this method evaluates the logical coherence between the premises and conclusions of an argument, thereby providing insights into the argument's quality (Plug and Wagemans, 2020).
- Rhetorical Checking: this approach conducts a comprehensive evaluation of persuasive communication, including the examination of the speaker's character appeals and emotional arguments (van Klink and van der Geest, 2020).
- **Ethical Issues:** in the context of handling user-generated data, ethical concerns assume a significant role. It is imperative to ensure that any analysis and prospective sharing of datasets strictly adhere to the privacy rights of the individuals involved. An ELSEC (Ethical, Legal, Social, Economic, and Cultural) approach is therefore crucial in AI ethics and data protection. It provides a holistic framework that acknowledges the complex interplay of these factors, ensuring that AI technologies respect diverse societal values, legal requirements, economic considerations, and cultural contexts, thereby fostering responsible and inclusive AI development.
- **Computational Methods:** We acknowledge that the proposed strategies and future research directions for analyzing and countering populist communication extensively rely on computational techniques. In Section 5.1.1 and 5.2.1 we have already outlined a repertoire of resources for examining the stylistic and substantive facets of populism. At the same time, we highlight two prospective research pathways:
 - Large Language Models (LLMs) have the potential to offer sophisticated insights into the nuances of language employed by politicians, particularly for identifying stylistic elements like colloquialism and intimidation. Thus, future research should employ also these models to analyze populism.
 - A pervasive challenge with LLMs and, more in general, deep learning approaches (mainly used for the stylistic features detection) is their opaque nature, which compromises their explainability. This aspect becomes critically important when such methods are employed by institutions to implement and assess regulations.
Consequently, future research should also incorporate hybrid methodologies: marrying deep learning's advanced and flexible capabilities with shallow learning interpretability (Azevedo et al. (2024)). Such a combination can significantly enhance models explainability, promoting institutions transparent controversy detection.

6 The Role and Impact of Conspiracy Theories in Political Communication on Social Media

6.1 Theoretical Framework

This chapter examines how conspiracy theories influence political communication on social media. It defines conspiracy theories in the context of politics and explores the psychological biases that make people susceptible to them. The section also looks at how social media platforms,

through mechanisms like echo chambers and filter bubbles, may amplify the spread of such theories, potentially leading to increased polarization. It, then, discusses the impact of these theories on voting behavior and political beliefs, offering the Great Replacement Theory case study as an example. With no claim to exhaustiveness, it provides an overview of the complex interplay between conspiracy theories, social media, and political communication.

Conceptualizing Conspiracy Theories: What constitutes a conspiracy theory within the context of political communication? Many definitions of conspiracy theory beliefs have been conceptualized. In this context we will use [Aronovitch \(2009\)](#) definition, which is concise, precise and straight-forward, as well as established in the literature ([Brotherton et al., 2013](#)).

"A conspiracist belief can be described as the unnecessary assumption of conspiracy when other explanations are more probable" ([Aronovitch, 2009](#)).

6.2 Psychological Explanations: Cognitive biases

Having established a theoretical framework for understanding conspiracy theories, we now delve into the cognitive biases that underpin individual susceptibility to these narratives. This section explores how cognitive biases make us vulnerable to conspiracy theories. It covers the most relevant cognitive biases, which can lead conspiracy theories to seem more plausible. By understanding these psychological mechanisms, we try to understand why conspiracy theories are appealing and how they can skew our perception of reality. Conspiracy beliefs are partially driven by psychological motives, like the desire for control; and social motives, such as the need to maintain a positive self-image and group identity ([Kim et al., 2022](#)). [Douglas et al. \(2019\)](#) reviewed the extensive and interdisciplinary research on conspiracy theories, addressing three main questions: the factors driving conspiracy beliefs, how these theories are spread, and their societal impacts. Findings indicate that conspiracy beliefs are influenced by psychological, political, and social factors, and are spread through both traditional and social media. What is fascinating is that all of us, as humans, can occasionally fall into the trap of conspiracy theories, often without explicitly intending to: it is a human evolutionary mechanism that has frequently been to our advantage, especially when we lived in small hunter-gatherer groups ([Brotherton, 2015](#)). As human societies have evolved, our cognitive characteristics have not changed, given that we modern humans are biologically quite identical to our ancestors who lived in the savannah in close-knit tribes ([Henshilwood et al., 2002](#)). What one perceives as true often stems more from their pre-existing beliefs than from direct observations of truth ([Koehler, 1993](#); [Witt, 2011](#)). Our beliefs about events, such as election fraud or treason, are not usually based on personal witness but rather on unreliable interpretations from fallible sources ([Uscinski and Enders, 2023](#)). What is clearly false to one might be clearly true to another. The more people find a conspiracy theory convincing, the more this demonstrates that its perceived falseness is not universally apparent ([Douglas et al., 2019](#)). Conspiracy theories act as a means to justify deep-rooted biases and irrational fears ([Douglas et al., 2017](#)). In the following part of the section we will explore the most relevant cognitive biases that influence the building of stable conspiratorial beliefs.

Intentionality bias: The intentionality bias refers to the tendency of people to ascribe intention, purpose, or volition to the actions of others, even in situations where such attributions may not be true ([Brotherton and French, 2015](#)). This cognitive bias has allowed us to survive in contexts of

uncertainty, such as when traversing a forest at night and hearing the sound of a breaking twig: assuming this noise is caused by the intent of a bear or a rival clan leads to a higher chance of survival by putting our body on alert, even when this assumption proves to be unfounded (Brotherton, 2015). In our complex societies, this bias leads individuals to assume that others act with deliberate intent, rather than considering alternative explanations such as chance, external pressures, or unconscious processes. It can influence how we interpret the behavior of others, often leading to misattributions of motives and intentions.

False consensus effect: The false consensus effect is the tendency to attribute to others one's own opinions, beliefs, and preferences, assuming that others think like oneself (Ross et al., 1977). This presumed agreement is often based on unfounded statistics and can lead to a perception of vast consensus that does not actually exist. This error can lead an individual or a group to believe that their opinions are more widespread among the public than they actually are.

Proportionality bias: The proportionality bias is the inclination to believe that significant events must have significant causes (Leman and Cinnirella, 2007). This cognitive bias plays a key role in people's propensity to embrace conspiracy theories (Van Prooijen and Van Dijk, 2014). When something significant occurs, we tend to assume that it must have been caused by something equally significant (Brotherton, 2015). It is less intuitive for humans assuming that small causes can lead to huge consequences.

Confirmation bias: The confirmation bias is the tendency to search for, interpret, favor, and recall information in a way that confirms or supports one's prior beliefs or values (Wason, 1960). Individuals exhibit this bias by choosing information that backs up their views while disregarding opposing ones, or when they interpret specific facts to bolster their own opinions (Knobloch-Westerwick et al., 2020).

Dunning-Kruger effect: The Dunning-Kruger effect is a psychological phenomenon where people with limited skills in a certain area often think they are more competent than they actually are, while those who are highly skilled tend to underestimate their abilities (Dunning, 2011). This happens because individuals who are not very skilled in a particular field lack the necessary understanding to judge their own incompetence accurately, leading them to overestimate what they can do (Kruger and Dunning, 2002). On the other hand, those who are very skilled are more conscious of what they do not know or cannot do, which makes them undervalue their own abilities. This concept also relates to why people with less knowledge about specialized subjects, like vaccine biotechnology or migration studies, might be more susceptible to believing in conspiracy theories (Motta et al., 2018). They find it harder to differentiate between what is true and what is not, making them more likely to fall for false claims.

Fundamental attribution error: The fundamental attribution error is a cognitive bias which states that people overestimate the influence of personal traits and underestimate the role of situational factors when explaining others' behaviors (Harvey et al., 1981). In the context of conspiracy beliefs, this error can lead individuals to attribute malevolent intentions or extreme competence to those believed to be behind conspiracies, overlooking the complexity of situations and the possibility of accidents or coincidences (Clarke, 2019). This tendency to favor personal attributions can fuel conspiracy theories by simplifying complex events into actions driven by powerful individuals or groups with specific agendas.

Projection mechanism: Projection is a psychological defense mechanism where individuals at-

tribute their own unacceptable thoughts, feelings, or motives to someone else (Holmes, 1968). In the context of conspiracy beliefs, projection plays a role by leading individuals to assume that others are engaging in deceitful, manipulative, or secretive behaviors because they themselves may harbor such tendencies or fears (Douglas and Sutton, 2011). This can contribute to the formation and persistence of conspiracy theories, as people project their own anxieties, mistrust, or paranoia onto external entities or groups, viewing them as the orchestrators of complex plots without substantial evidence. Paradoxically, individuals who are prone to believe in conspiracy theories might, when given power or authority, exhibit behaviors that are secretive, manipulative, or conspiratorial (Brotherton, 2015). This can be seen as an extension of the projection mechanism, where their deep-seated fears and mistrust towards authority or the establishment, nurtured through belief in various conspiracies, can influence their own actions when they find themselves in positions of power. Or, as Hofstadter (2012) would say: "The enemy seems to be a projection of the self: both the desirable traits and the undesirable traits are projected onto them. The paranoid mindset often involves paradoxically emulating the perceived (and usually nonexistent) enemy".

6.3 Social Media Addiction and Its Impact on Mental Health

In the previous section, we explored how belief in conspiracy theories is a psychological phenomenon common to all humans, making each of us susceptible to it. These cognitive predispositions not only shape our perception of reality but are also amplified by our interactions with social media, leading to a discussion on how social media addiction further entrenches conspiracy beliefs. Clearly, the more someone is already predisposed ("conspiracist ideation" by Swami et al. (2011)) to believe in a particular conspiracy theory, the more likely medias (and social medias too) will further reinforce their belief in this worldview (Enders et al., 2021). In the following sections, we will examine factors that contribute to the proliferation of conspiracy messages on social media: the mental health issues that social medias induce, the disinhibition on the internet, and the mechanisms by which news spreads on social platforms. Addiction to social media platforms may increase the exposure to conspiracy theories and make users more susceptible to their influence.

In recent years, the use of social media has consistently grown. By 2022, over 4.59 billion people were using social media worldwide, a number that is projected to increase to almost six billion by 2027, according to Statista (2022). Marshall McLuhan, often regarded as the father of communication studies, saw media as an extension of human nervous system (McLuhan, 1994). At the time he wrote, social media did not exist; however, his vision can now be considered undoubtedly predictive of the current landscape. We have literally connected our nervous system to the internet, with reward feedback mechanisms and stimuli that are specifically designed and constantly improved through machine learning. This makes having a healthy, non-abusive relationship with social media very challenging.

Social media addiction is a complex phenomenon influenced by a variety of factors, including individual dispositions (such as attachment styles and time perspective), motivational aspects (like the need for social gratification and self-presentation), neurobiological changes, decision-making processes, learning behaviors, and the influence of technology and social networks (Sun

and Zhang, 2021).

The financial success of social media platforms is significantly tied to their ability to foster habitual use among their users (Anderson and Wood, 2021). By offering rewarding experiences and embedding design cues that promote automatic engagement, these platforms ensure users keep coming back, often without a conscious decision to do so (Lindström et al., 2021). Rewards such as social recognition and content enjoyment play a key role in motivating users to engage repeatedly, leading to the formation of strong usage habits (Fareri and Delgado, 2014). Additionally, the platforms' designs are crafted to trigger these habits automatically, securing continuous user engagement and contributing to the platforms' financial achievements. Montag et al. (2019) explore how smartphone apps employ six sophisticated mechanisms to hook users, seamlessly blending psychological insights and economic strategies. The endless scrolling feature, by offering an unending stream of content, taps into the human desire for continuous engagement, effectively distorting users' perception of time (Alter, 2017). Similarly, the endowment and mere exposure effects are based on the principle that users attribute more value to their digital creations and are more inclined to enjoy content with which they frequently interact, making it challenging for them to disengage (Zajonc, 2001). Social pressure is exploited through features like read receipts, creating a sense of urgency and expectation for prompt communication (Montag et al., 2019). Apps further refine user engagement by employing algorithms that tailor content to individual preferences, ensuring that the feed remains irresistibly relevant and engaging, a mechanism that leads to filter bubbles, that we will further explore (Pariser, 2011). Social comparison and reward mechanisms, exemplified by likes and comments, offer a digital form of social validation that encourages continuous app usage. Lastly, the Zeigarnik and Ovsiankina effects highlight users' propensity to remember and return to unfinished tasks, a phenomenon that games exploit by designing challenging levels that entice users to keep playing or even spend money to progress (Zeigarnik, 1927; Rickers-Ovsiankina, 1928). Together, these mechanisms form a potent formula designed to maximize user engagement and time spent within apps, raising questions about the ethical implications of such design choices and the need for regulatory or alternative business models to mitigate addictive behaviors, especially in the context of disinformation.

Sun and Zhang (2021) reviewed 55 empirical studies of social media addiction and found that the excessive use of social media is linked to several adverse outcomes, including diminished productivity, poor quality social relationships, and lower overall satisfaction with life. This overindulgence in social media contributes to decreased work performance, less healthy interactions within social networks, sleep disturbances, and a general decline in life contentment. Furthermore, it is associated with negative emotions such as jealousy, anxiety, and depression. D'Arienzo et al. (2019) reviewed 32 papers published between 2000 and 2018 and found a strong link between forms of insecure attachment, including both anxiety and avoidance, and increased, problematic usage of the internet and social media. It indicates that individuals with insecure attachment tendencies tend to engage more intensively with social media platforms, suggesting this use as a substitute to fill the void of affection and connection they lack from their immediate social circles, such as family and peers. Nonetheless, the meta-analysis by Cheng et al. (2021) revealed that the prevalence of social media addiction varies significantly based on the classification scheme used, and is significantly higher in collectivist cultures (31%) compared to individualist ones (14%). These findings highlight the impact of classification criteria and cultural

context on the understanding of social media addiction prevalence rates globally. Haidt (2022) argues that the surge in adolescent depression, anxiety, and self-harm, particularly among young girls, around 2012-2013 coincides with the rapid increase in social media use, suggesting a direct correlation between social media engagement and mental health deterioration. He points out that while many researchers and technologists might view social media's impact through a dose-response model — implying that its effects are proportional to the amount of usage — this perspective fails to account for the broader, more insidious impacts on social dynamics, self-perception, and mental well-being (Odgers and Jensen, 2020). Haidt criticizes the reductionist approach that treats social media like sugar, where harm is seen as a result of excessive consumption, arguing instead that social media fundamentally alters peer relationships, family interactions, and daily life in ways that are particularly damaging to adolescents (Nesi et al., 2018). His analysis suggests that beyond fostering addictive behaviors, social media platforms contribute to a pervasive environment of comparison, exclusion, and relational aggression, which is especially harmful to the psychological development and mental health of young users (Haidt and Allen, 2020).

Online we are usually more uninhibited The addictive nature of social media not only impacts mental health but also fosters a unique form of online behavior, characterized by disinhibition. Online individuals often exhibit a greater degree of openness and freedom compared to face-to-face interactions, a phenomenon extensively explored by Suler (2004), who identifies a complex interplay of six key factors: dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and the minimization of authority. These elements work together to create an online environment where individuals feel more liberated to express themselves without the immediate repercussions or social constraints they face offline. Dissociative anonymity allows people to detach their online actions from their real-world identity, encouraging openness and risk-taking by reducing fear of real-world consequences. Invisibility eliminates the physical presence, removing social inhibitions and prompting more candid or daring interactions. Asynchronicity gives users time to contemplate and craft their responses carefully, lessening the immediate impacts of their words and allowing for more deliberate or unrestrained communication (Wu et al., 2017). Solipsistic introjection blurs the line between one's thoughts and what is communicated online, fostering a sense of intimacy or connection that might not exist in person. Dissociative imagination encourages users to see their online behavior as part of a separate, fictional realm, leading to a detachment from real-life responsibilities and norms. Lastly, the minimization of status and authority on the internet creates a more egalitarian space for expression, free from the usual hierarchical constraints. However, the degree to which individuals experience online disinhibition can vary widely, influenced by personal traits, emotions, needs, the specific social platform and the strength of their self-regulatory mechanisms (Cheung et al., 2020). Suler (2004) also posits that this disinhibition should not be seen as uncovering a hidden "true self" but rather as a shift to a different facet of the self, characterized by unique emotional and cognitive patterns distinct from those in offline interactions. Rather than viewing online disinhibition as revealing a person's "true self," Suler (2004) proposes understanding it as a shift to a different "constellation" within the self-structure, characterized by different clusters of affect and cognition than those typically engaged in person-to-person interactions. In other words, a series of different masks (Goffman, 2016). This perspective challenges the notion of a singular, true self,

suggesting instead that identity is multifaceted and context-dependent. This is in line with the socio-anthropological tradition that support the idea of the contextual negotiation of social identity which refers to the dynamic process by which individuals adapt and modify their identities based on the social, cultural, and situational contexts they find themselves in (Butler, 2003; Kottak, 2017). These different identities get influenced by the different context they are exposed to and interact with, following specific schemes. If exposed to conspiracy theories, people could fall into the so called rabbit-hole (Sutton and Douglas, 2022). In the next section we will explore the motivations about why this could happen, which are the main explanations experts have given so far and why some of them are starting to vacillate.

6.4 Spread Mechanisms: Echo Chambers, Filter Bubbles, Machine Habitus and other alternative explanations

This online disinhibition, facilitated by social media addiction, and the cognitive biases, play an important role in the mechanisms through which conspiracy theories spread, including echo chambers and filter bubbles. Traditionally scholars have outlined the concepts of **Echo Chambers** and **Filter Bubbles** as central to understanding the dynamics of online political discourse and polarization (Terren and Borge-Bravo, 2021). Echo chambers refer to situations where individuals are primarily exposed to opinions and information that reinforce their existing beliefs, facilitated by both their personal choices and algorithmic recommendations on social media platforms (Sunstein, 2001, 2018). This concept is intricately linked with homophily: the tendency for individuals to associate and engage with others who are similar to them (McPherson et al., 2001a). This tendency is exacerbated by selective exposure, where individuals prefer ideologically aligned information to avoid cognitive dissonance (Smith et al., 2008; Frey and Wicklund, 1978), and confirmation bias (Nickerson, 1998), as well as the other cognitive biases we mentioned previously. Filter bubbles, on the other hand, are primarily driven by algorithmic recommendation on social media platforms (Pariser, 2011). These algorithms, by design, tailor the content feed to the user's past online behavior, significantly shaping the diversity of content to which they are exposed (Nguyen et al., 2014). This customization, while intended to enhance user experience, supposedly limits exposure to a variety of viewpoints, further increasing ideological polarization by reducing users' likelihood to encounter ideologically cross-cutting news content (Spohr, 2017). Terren and Borge-Bravo (2021) systematically reviewed 55 studies on echo chambers on social media, identifying patterns across their foci, methods, and findings. They found that research outcomes are significantly influenced by conceptual and methodological choices, with studies based on digital trace data more likely to report echo chambers compared to those relying on self-reported data. In fact, there is still no common consensus regarding the echo chamber mechanism. Recent research challenges the prevailing notion that social media's filter bubbles are the primary cause of societal polarization (Törnberg, 2022). Studies have found little evidence of ideological isolation online, suggesting that individuals are more likely to encounter opposing viewpoints in digital spaces than in their real-life interactions (Bakshy et al., 2015). In reality, people's real-life social circles — comprising friends, family, colleagues, and neighbors — tend to be more ideologically similar than their online counterparts (Gentzkow and Shapiro, 2011), according once again to the homophilia principle. This suggests that the so-called filter bubble is

more prevalent in physical spaces rather than online. The problem, as highlighted by recent findings, is not the existence of filter bubbles per se but the overwhelming amount of disagreement and conflicting information that individuals encounter on social media (Törnberg, 2022). Human beings, who evolved to foster cooperation and social cohesion within relatively homogeneous groups, struggle to process the vast diversity of opinions and worldviews presented in the digital town square (Tajfel et al., 1971). This leads to social sorting, where individuals categorize others based on their opinions, creating an "us versus them" mentality that undermines societal cohesion and rational discourse (McCoy et al., 2018). Social media platforms, driven by engagement algorithms, exacerbate this issue by amplifying anger and controversy, which are highly engaging (Fan et al., 2014). This results in the most extreme and controversial opinions gaining visibility, further polarizing public discourse (Baumeister et al., 2001). The online environment simplifies and distorts disagreements, making opposing viewpoints appear not just different but morally reprehensible, further eroding the possibility of civil discourse and mutual understanding (Törnberg, 2022). It is no doubt that the online sphere plays a crucial role in the spreading of disinformation, with scholars defining it a "conspirosphere", referring to a broad and international web of websites and blogs dedicated to promoting these views (Reichstadt, 2015). This sphere includes diverse political and ideological groups but is united by a common distrust of official narratives and a tendency to attribute historical events to secret plots.

The concept of habitus, as developed by Pierre Bourdieu, refers to the deeply ingrained habits, skills, and dispositions that individuals acquire through living in their particular social environments (Bourdieu, 2018). These are not just superficial customs but are deeply rooted in the practices, perceptions, and understandings of individuals, shaping their worldviews and actions without the need for conscious deliberation. Habitus forms through the socialization processes of family, education, culture, and personal experiences, creating a durable yet dynamic set of dispositions that guide behavior and thinking (Bourdieu, 2020). Extending this concept to the context of technology and artificial intelligence, **Machine Habitus** represents the idea that algorithms and AI systems can similarly develop sets of dispositions or patterns through their "experiences" — in this case, the data they process and the interactions they have with humans (Airoldi, 2021). Just as human habitus is shaped by social conditions and experiences, machine habitus is shaped by the digital environments in which they operate, the data they are fed, and the objectives set by their human creators. Machine habitus, therefore, encapsulates how AI systems internalize and reflect the biases, values, and norms present in the data they are trained on (Noble, 2018). It suggests that these systems can exhibit patterns that mimic human biases, not because they possess consciousness or intent, but because they are processing inputs derived from human behaviors and societal norms. For instance, an AI trained on historical hiring data may "learn" and perpetuate the gender biases implicit in that data, not through any understanding of gender but simply by replicating the patterns it has been fed (O'neil, 2017).

6.5 Impact of Conspiracies on Voting Behavior

Influence on Political Opinions and Voting Behaviour: The social media dynamics not only shape public discourse but have tangible effects on political landscapes, influencing voting behavior and political ideologies. The conspiracist belief is not a new phenomenon of our digital

era (van Prooijen and Douglas, 2017). Humans have always witnessed plots and actual conspiracies. At the same time conspiracy theories, as a way of perceiving the reality, have always existed (Brotherton, 2015). Throughout history, there have been instances of conspiratorial actions. Recognizing these acts is not a matter of paranoia. Political behavior inherently involves strategic planning, and many strategies rely on maintaining secrecy for a period of time to be effective. Consequently, it is often accurate, albeit sometimes with slight exaggeration, to characterize these secret actions as conspiratorial (Hofstadter, 2012). The unique aspect of the conspiratorial style is not merely that its proponents identify conspiracies or plots scattered throughout history, but that conspiracy theorists view an extensive or enormous conspiracy as the main driving force behind historical events (Hofstadter, 2012). The paranoid individual eventually can assume the role of a combative leader. Unlike a typical politician, who views social conflicts as issues to be resolved through mediation and compromise, the conspiracist leader sees these conflicts as battles between absolute good and absolute evil. Because the adversary is considered wholly evil, the paranoid believes that this enemy must be completely eradicated — if not from the world, then at least from the power (Hofstadter, 2012). "The paranoid style has had a consummatory triumph [...] in Germany. It is a common ingredient of fascism, and of frustrated nationalism" (Eysenck, 1966). Hofstadter already in 1964 wrote that "The modern right wing feels dispossessed: America has been largely taken away from them [...] the old competitive capitalism has been gradually undermined by socialist and communist schemers". The paranoid tendency is aroused by a confrontation of opposed interests which are (or are felt to be) totally irreconcilable, and thus by nature not susceptible to the normal political processes of bargain and compromise (Hofstadter, 2012).

Oliver and Wood (2014) investigate the relationship between conspiracy theory endorsement and political beliefs, showing that the attraction to magical thinking and Manichean narratives predominantly influences these beliefs. This goes in contrast with previous beliefs that stated that political ignorance or conservatism were the main causes of belief in such theories (Sutton and Douglas, 2020). This suggests that conspiracy theories can profoundly shape political opinions by aligning with pre-existing cognitive orientations and narrative preferences. Van Prooijen et al. (2015) research across the United States and the Netherlands indicates a U-shaped relationship between the extremity of political ideologies — both left and right — and the propensity to endorse conspiracy theories about diverse political topics (Krouwel et al., 2017). This association is mediated by the belief in simplistic solutions to complex societal issues among extremists on both ends of the spectrum, not simply due to extreme attitudes in general. The core finding suggests that political extremism correlates with conspiracy beliefs through a structured thinking style aimed at comprehending societal events. The negative emotions people feel during crises — such as fear, uncertainty, and a sense of loss of control — prompt a need to better understand the situation. This, in turn, raises the chances of seeing conspiracies in social contexts (van Prooijen and Douglas, 2017). Conspiratorial thinking acts as a cognitive shortcut: a strategy for dealing with informative uncertainty (Kovic and Fuchsli, 2018). Kim et al. (2022) showed that the feelings of control and social validation, which may come from electoral victories, can counteract some of the psychological factors that fuel conspiracy beliefs. Voting and election outcomes might address these underlying motives by providing a sense of agency (existential motive) and validating one's group identity (social motive), potentially reducing the allure of con-

spiracy theories. Interestingly, Kim (2022) found that conspiracy theories can have a mobilizing effect, especially when they signal that the audience is a victim of the conspiracies, that these conspiracies are not insurmountable, and that culpable elites can be identified and potentially challenged. Nevertheless, while it is true that conspiracy theories can influence positive political behaviors, like voting, this benefit is undermined if the individuals swayed by such theories end up supporting parties that oppose constitutional values (Van Prooijen et al., 2015). It is not a case that in US 2016 elections individuals with a higher tendency towards conspiracy thinking were more likely to have voted for Donald Trump, a candidate who can hardly be described as moderate and attached to constitutional values (Lamberty et al., 2018). Christner (2022) found a strong correlation between conspiracist beliefs involving government participation, low political trust and the populist attitudes of anti-elitism and sovereignty. This result is confirmed also by Salvati et al. (2022), who found that voters of populist parties were more likely to endorse conspiracy theories and exhibit higher levels of Right-Wing Authoritarianism.

6.6 How to measure the belief in Conspiracy Theories?

Support for conspiracy theories is not only based on the political attitude or the logical assessment of the evidence for each particular conspiracy claim. Instead, it seems that there are consistent personal differences in the overall inclination to accept conspiracy-based explanations for events. This characteristic is known as “conspiracist ideation” (Swami et al., 2011). Scholars proposed different ways to understand the conspiratorial attitude of people. Recognizing the complexity and the stable individual differences in conspiracist ideation, Brotherton et al. (2013) propose a new measurement tool, the Generic Conspiracist Beliefs (GCB) scale, to assess individual tendencies towards conspiracist ideation more accurately than previous methods. The GCB scale reflects a monological belief system characterized by a small number of generic assumptions about conspiratorial activity in the world. The scale offers a reliable tool for measuring generic conspiracist ideation across various contexts. Instead of employing a lengthy questionnaire with multiple items or questions to gauge a person’s belief in various conspiracy theories, Lantian et al. (2016) proposed a single-item conspiracy theory scale, which simplifies the measurement to a single item. It is a measurement tool designed to assess an individual’s tendency to believe in conspiracy theories using just one question or statement. In the context of the study by Lantian et al. (2016), the single-item scale asked participants to rate their agreement with a statement regarding their perception of whether authorities often hide the truth about significant events. The purpose of using a single-item scale is to offer a quick, efficient, and straightforward method for researchers to assess conspiracy belief levels, especially in situations where time or space on a survey is limited. Despite its simplicity, the scale was validated through several studies to ensure it reliably and accurately measures the general propensity towards conspiracist ideation. The single item used to measure belief in conspiracy theories was:

“I think that the official version of the events given by the authorities very often hides the truth.”

Participants were asked to rate their agreement with this statement on a scale from 1 (Completely false) to 9 (Completely true), providing a concise measure of their general tendency to believe in conspiracy theories. This single-item approach was validated across multiple studies to ensure its effectiveness in capturing the essence of conspiracist ideation (Lantian et al., 2016).

Moreover, [Bruder et al. \(2013\)](#) proposed the Conspiracy Mentality Questionnaire (CMQ), a concise, 5-item tool developed to measure an individual's general predisposition towards believing in conspiracy theories, independent of specific events or narratives. It aims to quantify the extent to which a person subscribes to the view that significant, undisclosed forces or groups manipulate societal events, political decisions, and public perception from behind the scenes. The questionnaire assesses beliefs in secretive monitoring by government agencies, undisclosed important global events, hidden motives behind political decisions, connections between seemingly unrelated events due to secret activities, and the influence of clandestine organizations on political outcomes. Validated across different cultures and languages, the CMQ offers a robust, one-dimensional construct for researchers studying the psychological and social underpinnings of conspiracist ideation globally. The idea is that the items in the CMQ all tap into this single underlying tendency, suggesting that individual differences in conspiracy mentality can be captured along a single continuum, from low to high susceptibility to conspiracist ideation. In the context of these measurements it is always important to keep a grain of salt, in order to distinguish between a sane skepticism and critical thinking and actual conspiracist ideation.

6.7 Case Study

Legitimizing Extreme Right Discourse Through Conspiracy Beliefs

The functionality of conspiracy beliefs, particularly in legitimizing negationism and enhancing the discourse of the extreme right, represents a critical aspect of their political utility ([Pezzini and Terracciano, 2022](#)). By fostering narratives that question historical accuracies and deny factual evidences, such as the Holocaust or climate change, these beliefs create a foundation for revisionist history that serves the interests of the extreme right. Negationism is closely related to “historical revisionism”, which seeks to rewrite twentieth-century history from a perspective more favorable to the extreme right ([Tateo, 2005](#)). Plus, negationism and conspiracy theories share a common narrative structure that includes the rejection of established facts, the belief in hidden manipulations, and the inversion of victimhood ([Reichstadt, 2015](#)). The extreme right's adoption of conspiracy theories goes beyond mere political rhetoric; it strategically utilizes these narratives to validate its ideologies, legitimize their opinions, and undermine democratic norms and institutions ([Betz, 2016](#)). This legitimization process benefits from the psychological appeal of conspiracy theories, which offer simplified explanations for complex social issues, thereby attracting individuals seeking clarity amidst uncertainty. For the extreme right, conspiracy theories also provide a mechanism to project blame onto out-groups, reinforcing us-versus-them mentalities that are central to their divisive agenda.

Furthermore, the strategic dissemination of conspiracy beliefs through social media platforms amplifies their reach and impact, contributing to the normalization of extreme right discourse in public conversations ([Smith et al., 2022](#)). This propagation – that is not limited to the digital media – not only broadens the audience for such narratives but also enhances the perceived legitimacy of the extreme right's ideological positions, making it a significant concern for the preservation of democratic values and social harmony.

Great Replacement Theory and the Extreme Right

[Pape \(2021\)](#) conducted a study on the 2021 January 6th insurrection and revealed that the ma-

majority of those involved were middle-aged, employed, with 40% holding positions in white-collar professions. This is particularly interesting in the contemporary US context, in which all the economic indicators state a huge economic growth, with unemployment decreasing to record levels as well as racial inequalities (USAFacts, 2024). The main motivator for participation in the insurrectionist movement was the fear among white individuals that Black and Latin people might gain more rights than them (Pape and Ruby, 2021; Smith et al., 2022). This fear ties back to the “Great Replacement” conspiracy theory, which claims that the populations of traditionally white countries with low birth rates are being deliberately replaced by people of color and immigrants (Ekman, 2022). This theory claims that such changes pose risks to the safety and economic well-being of white families, blaming - among the others - feminists, “social justice warriors”, leftist, “the Globalist elite”, “cultural Marxists”, and Jews for these shifts. Thus, it appears that fears of losing white dominance are shaped more by narratives and deep-seated beliefs within far-right circles than by tangible socioeconomic conditions (Marwick et al., 2022). In fact, the decreasing of traditional racial and gender inequalities may be perceived as a threat by the white upper class (Craig and Richeson, 2014; Myers and Levy, 2018). And to legitimize and rationalize such fear, the Great Replacement Theory comes to hand. These findings suggest that concerns over perceived threats, rather than actual economic hardships, may be more influential in choosing the political side (Marwick et al., 2022). Investigating conspiratorial beliefs is essential, as they significantly shape citizens’ political views, often more profoundly than traditional socioeconomic indicators. In fact, the perception of demographic replacement may fuel Islamophobia and right-wing extremism (Obaidi et al., 2022). To conclude, the Thomas theorem’s comes to hand: “If men define situations as real, they are real in their consequences” (Thomas and Thomas, 1938).

6.8 Challenges and Future research directions

The case study of the Great Replacement Theory illustrates the need to pursue a number of different research directions, including revisiting past Internet models and investigating the relationship between conspiracy theories and the rise of far-right parties. Drawing on Durkheim (2023)’s analogy of society as a biological organism, where harmonious interaction within and between its various segments fosters a healthy societal structure, we propose revisiting past Internet models to enhance societal cohesion in the digital age. Durkheim emphasized the importance of specialized, integrated roles within society’s sub-groups for overall societal health. Today, the broad connectivity enabled by social media often leads to polarization and a loss of nuanced conversation, disrupting this delicate balance.

To counteract these challenges, we suggest drawing inspiration from earlier Internet models — such as bulletin boards, forums, and blogs. These platforms cultivated smaller, more cohesive communities, fostering a diversity of opinions within a structured environment that more closely mirrored real-life social dynamics than the expansive networks typical of contemporary social media platforms. By promoting engagement within smaller, interest-based communities, we can reintroduce the complexity and context to our digital interactions, akin to Durkheim’s vision of a well-functioning society. This approach not only aims to reduce social sorting but also encourages a more tolerant and empathetic online discourse, thereby mitigating the decontextualization and polarization prevalent in today’s digital exchanges.

Another potential future direction could involve regulatory changes requiring the public release of the algorithms underlying filter bubbles, which distort users' perception of reality. A step forward has been made by the EU with the recent AI Act, but more efforts should be made.

A critical area for further investigation is the link between the endorsement of conspiracy theories about migration and the ascent of far-right political parties. This inquiry is vital because such parties often base their platforms on incorrect information and distorted narratives, which misrepresent reality. If these parties gain power democratically, the policies they enact could range from being ineffective — addressing non-existent issues — to causing significant harm, potentially violating international law and human rights. This danger stems directly from the policies' foundation on disinformation.

In conclusion, this chapter explored the complex interplay between conspiracy theories, social media, and their impact on political behavior and societal division. We began by defining conspiracy theories and explaining their psychological allure, highlighting how social media platforms amplify these narratives, reinforcing existing biases and contributing to polarization. Through examining mechanisms like echo chambers and filter bubbles, we showed how digital environments foster the spread of conspiracies, influencing political attitudes and potentially leading to extremist ideologies. We discussed the real-world consequences of these dynamics, including their effect on voting behavior and the rise of far-right movements, illustrated by the case study of the Great Replacement Theory. The chapter wants to stress the importance of addressing these challenges through strategies like legislative regulation, further research in the field, hybrids AI models, digital literacy, and platform design changes, aiming to foster a healthier online discourse.

7 Conclusions

In this document, we cover different research topics related to WP2 “Public Discourse Analysis”, focusing in particular on political communication in social networks. The document encompasses the work of five DCs in the doctoral network (DC 1 – 5), i.e. those that are mainly focused on the application of human and social sciences to study different aspects of disinformation and abusive language in the political domain. Indeed, this deliverable complements the content of D2.1. on “State of the art of NLP and AI methods for discourse analysis in the political domain”, which was more technology-oriented, while D2.2. provides a more theoretical overview of the topic with an emphasis on the negative effects and risks related to misuse of social media use in political communication. We cover five topics of major interest across social network analysis and political discourse analysis: network representation of political narratives (Section 2), political opinion and participation in social networks (Section 3), implicit communication in political discourse on social media, mainly in negative terms (Section 4), populist communication online (Section 5) and conspiracy theories (Section 6). This document represents the theoretical background upon which HYBRIDS DCs' activities will be based and that will inform the implementation of new tools and approaches for the analysis of political discourse online.

Psychological Mechanisms Built-in Social Media/Messenger Apps and/or Freemium Games	Example/Illustration
Endless scrolling/streaming	As soon as one video is at the end on a website such as YouTube, the next video begins with either a similar content or the second episode of a TV show and so forth. By this, viewers get more and more absorbed, which makes it hard to stop watching.
Endowment effect/ mere-exposure effect	Every time players visit the app platform and invest more time in the construction of the virtual world, it will get harder for them to detach from the game or even delete the app. The endowment effect might be both explained by ownership and loss aversion. Also, of importance is the mere exposure effect describing that the more often you are exposed to a certain (neutral) thing or application (here a game), the more you like it.
Social pressure	Illustration from a WhatsApp feature: If a user sends a message to a friend, the sender is presented with two gray ticks, which means that the message has successfully arrived at the recipient's phone. If the recipient reads the message, the grey ticks turn blue. As both sides know about these rules, social pressure emerges. Both parties likely expect a fast answer, above all, if the message apparently has been read.
Show users of an app what they like	Facebook has a great interest in studying the behavior of each person at perfection and in much detail, so that at best only such information is presented in the 'Newsfeed' which is most interesting for the user. Otherwise, people could get bored and close the browser window.
Social comparison and social reward	Perhaps one of the most prominent features of social reward mechanisms in social media is the iconic 'thumbs up'. A 'thumbs up' ('Like') demonstrates either positive social feedback on one's own post or gives another person such a feedback.
Zeigarnik effect/ Ovsiankina effect	<p>The Zeigarnik effect refers to better remembering of tasks, where a person has been interrupted. Rickers-Ovsiankina then showed that such interrupted tasks are more likely to be finished later on (even if one is not forced to do this).</p> <p>Illustration: Some levels in Freemium games are very hard to solve and in case of Candy Crush Saga it is even mentioned that a "super hard level" is coming up. As some of these levels are "super hard" to solve (rumor has it that it is even impossible at first try), players easily loose several of those free lives ending up with no energy to finish this "super hard level". Being now really attracted by the game, this results in emotional strain which consequently provokes people to spend extra money to buy additional lives/gaming energy, because the next level is only a couple of minutes away.</p>

Figure 6: Features designed to extend engagement with social media applications and Freemium games, according to Montag et al. (2019).

References

- Aaronovitch, D. (2009). *Voodoo histories: the role of the conspiracy theory in shaping modern history*. Jonathan Cape London, England, London, England.
- Abzianidze, N. (2020). Us vs. them as structural equivalence: Analysing nationalist discourse networks in the georgian print media. *Politics and Governance*, 8(2):243–256.
- Adorno, T. W., Tiedemann, R., and NicholSEN, S. W. (1995). Notes to literature. *Journal of Aesthetics and Art Criticism*, 53(3).
- Airoldi, M. (2021). *Machine habitus: Toward a sociology of algorithms*. John Wiley & Sons.
- Albertazzi, D. and McDonnell, D. (2007). *Twenty-first century populism: The spectre of Western European democracy*. Springer.
- ALDayel, A. and Magdy, W. (2021). Stance detection on social media: State of the art and trends. *Information Processing & Management*, 58(4):102597.
- Altay, S., Berriche, M., Heuer, H., Farkas, J., and Rathje, S. (2023). A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field. *Harvard Kennedy School Misinformation Review*.
- Alter, A. (2017). *Irresistible: The rise of addictive technology and the business of keeping us hooked*. Penguin.
- Anderson, I. A. and Wood, W. (2021). Habits and the electronic herd: The psychology behind social media's successes and failures. *Consumer Psychology Review*, 4(1):83–99.
- Anjum and Katarya, R. (2024). Hate speech, toxicity detection in online social media: A recent survey of state of the art and opportunities. *International Journal of Information Security*, 23(1):577–608.
- Ansah, J., Liu, L., Kang, W., Kwashie, S., Li, J., and Li, J. (2019). A graph is worth a thousand words: Telling event stories using timeline summarization graphs. In *The World Wide Web Conference*, pages 2565–2571.
- Antoine, F. (1899). De la parataxe et de l'hypotaxe dans la langue latine. *Revue des études anciennes*, 1(1):27–46.
- Ash, E., Gauthier, G., and Widmer, P. (2024). Relatio: Text semantics capture political and economic narratives. *Political Analysis*, 32(1):115–132.
- Aslanidis, P. (2018). Measuring populist discourse with semantic text analysis: an application on grassroots populist mobilization. *Quality & Quantity*, 52(3):1241–1263.
- Azevedo, B. F., Rocha, A. M. A., and Pereira, A. I. (2024). Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Machine Learning*, pages 1–43.
- Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

- Baly, R., Da San Martino, G., Glass, J., and Nakov, P. (2020). We Can Detect Your Bias: Predicting the Political Ideology of News Articles. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4982–4991, Online. Association for Computational Linguistics.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10):1531–1542.
- Barberá, P. and Rivero, G. (2015). Understanding the Political Representativeness of Twitter Users. *Social Science Computer Review*, 33(6):712–729.
- Bartels, L. M. (2015). Attitudes, Political and Public Opinion. In *International Encyclopedia of the Social & Behavioral Sciences*, pages 193–196. Elsevier.
- Barthes, R. and Duisit, L. (1975). An introduction to the structural analysis of narrative. *New literary history*, 6(2):237–272.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4):323–370.
- Bauwelinck, N. and Lefever, E. (2020). Annotating topics, stance, argumentativeness and claims in Dutch social media comments: A pilot study. In *Proceedings of the 7th Workshop on Argument Mining*, pages 8–18, Online. Association for Computational Linguistics.
- Bazoobandi, S. (2023). Populism, jihad, and economic resistance: Studying the political discourse of iran's supreme leader. *Digest of Middle East Studies*, 32(4):321–339.
- Beheshti, A., Vaghani, K., Benatallah, B., and Tabebordbar, A. (2018). CrowdCorrect: A Curation Pipeline for Social Data Cleansing and Curation. In Mendling, J. and Mouratidis, H., editors, *Information Systems in the Big Data Era*, pages 24–38, Cham. Springer International Publishing.
- Benevenuto, F., Rodrigues, T., Cha, M., and Almeida, V. (2009). Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*, pages 49–62, Chicago Illinois USA. ACM.
- Bennett, A. and Seyis, D. (2023). The Online Market's Invisible Hand: Internet Media and Rising Populism. *Political Studies*, 71(3):655–675. Publisher: SAGE Publications Ltd.
- Berger, J. and Milkman, K. L. (2012). What makes online content viral? *Journal of marketing research*, 49(2):192–205.
- Berinsky, A. J., editor (2011). *New Directions in Public Opinion*. Routledge, New York.
- Bermingham, A. and Smeaton, A. (2011). On Using Twitter to Monitor Political Sentiment and Predict Election Results. In Bandyopadhyay, S. and Okumura, M., editors, *Proceedings of the Workshop on Sentiment Analysis Where AI Meets Psychology (SAAIP 2011)*, pages 2–10, Chiang Mai, Thailand. Asian Federation of Natural Language Processing.

- Betz, H.-G. (2016). Against the 'green totalitarianism': Anti-islamic nativism in contemporary radical right-wing populism in western europe. In *Europe for the Europeans*, pages 33–54. Routledge.
- Bhattacharya, C. (2020). Gatekeeping the plenary floor: discourse network analysis as a novel approach to party control. *Politics and Governance*, 8(2):229–242.
- Blair, J. A. (2012). The rhetoric of visual arguments. In *Defining visual rhetorics*, pages 41–61. Routledge.
- Blassnig, S., Engesser, S., Ernst, N., and Esser, F. (2019). Hitting a nerve: Populist news articles lead to more frequent and more populist reader comments. *Political Communication*, 36(4):629–651.
- Block, E. and Negrine, R. (2017). The populist communication style: Toward a critical framework. *International journal of communication systems*, 11:178–197.
- Blondel, V., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical mechanics: Theory and Experiment*, 2008.
- Bonikowski, B. and Gidron, N. (2016). The populist style in american politics: Presidential campaign discourse, 1952–1996. *Social Forces*, 94(4):1593–1621.
- Bonikowski, B., Luo, Y., and Stuhler, O. (2022). Politics as usual? measuring populism, nationalism, and authoritarianism in u.s. presidential campaigns (1952–2020) with neural language models. *Sociological Methods & Research*, 51(4):1721–1787.
- Bos, L. and Brants, K. (2014). Populist rhetoric in politics and media: A longitudinal study of the netherlands. *European Journal of Communication*, 29(6):703–719.
- Botzer, N. and Weninger, T. (2023). Entity graphs for exploring online discourse. *Knowledge and Information Systems*, 65(9):3591–3609.
- Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.
- Bourdieu, P. (2018). Distinction a social critique of the judgement of taste. In *Inequality*, pages 287–318. Routledge.
- Bourdieu, P. (2020). Outline of a theory of practice. In *The new social theory reader*, pages 80–86. Routledge.
- Bracciale, R., Andretta, M., and Martella, A. (2021). Does populism go viral? How Italian leaders engage citizens through social media. *Information, Communication & Society*, 24(10):1477–1494.
- Bracciale, R. and Martella, A. (2017). Define the populist political communication style: the case of italian political leaders on twitter. *Information, communication & society*, 20(9):1310–1329.

- Bredereck, R. and Elkind, E. (2017). Manipulating Opinion Diffusion in Social Networks. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 894–900, Melbourne, Australia. International Joint Conferences on Artificial Intelligence Organization.
- Brotherton, R. (2015). *Suspicious minds: Why we believe conspiracy theories*. Bloomsbury Publishing.
- Brotherton, R. and French, C. C. (2015). Intention seekers: Conspiracist ideation and biased attributions of intentionality. *Plos one*, 10(5):e0124125.
- Brotherton, R., French, C. C., and Pickering, A. D. (2013). Measuring belief in conspiracy theories: The generic conspiracist beliefs scale. *Frontiers in psychology*, page 279.
- Bruder, M., Haffke, P., Neave, N., Nouripanah, N., and Imhoff, R. (2013). Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy mentality questionnaire. *Frontiers in psychology*, 4:43078.
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7):1164–1180.
- Bugnard, A., Piguat, M., Santacroce, N., and Gros-Daillon, J. (2021). Informal language detection. application on quotes found in news articles. *ADA. Applied Data Analysis*.
- Burke, K. (1945). *A grammar of motives*, volume 177. Univ of California Press.
- Butler, J. (2003). Gender trouble. *Continental feminism reader*, pages 29–56.
- Chan, M., Yi, J., and Kuznetsov, D. (2024). Government Digital Repression and Political Engagement: A Cross-National Multilevel Analysis Examining the Roles of Online Surveillance and Censorship. *The International Journal of Press/Politics*, 29(2):371–393.
- Chancellor, S., Pater, J., Clear, T., Gilbert, E., and Choudhury, M. (2016). #thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. pages 1199–1211.
- Chen, A. C.-H. and Dluhošová, T. (2023). Concepts in contexts: Discourse-based semantic networks of ideologies in taiwan (1945–1949). *Asian Studies Review*, 47(1):173–193.
- Chen, W., Zhang, X., Wang, T., Yang, B., and Li, Y. (2017). Opinion-aware Knowledge Graph for Political Ideology Detection. pages 3647–3653.
- Cheng, C., Lau, Y.-c., Chan, L., and Luk, J. W. (2021). Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. *Addictive behaviors*, 117:106845.
- Cheung, C. M. K., Wong, R., and Chan, T. K. H. (2020). Online disinhibition: conceptualization, measurement, and implications for online deviant behavior. *Ind. Manag. Data Syst.*, 121:48–64.

- Chong, D., Lee, E., Fan, M., Holur, P., Shahsavari, S., Tangherlini, T., and Roychowdhury, V. (2021). A real-time platform for contextualized conspiracy theory analysis. In *2021 International Conference on Data Mining Workshops (ICDMW)*, pages 118–127. IEEE.
- Christner, C. (2022). Populist attitudes and conspiracy beliefs: exploring the relation between the latent structures of populist attitudes and conspiracy beliefs. *Journal of Social and Political Psychology*, 10(1):72–85.
- Clarke, S. (2019). Appealing to the fundamental attribution error: Was it all a big mistake? In *Conspiracy Theories*, pages 129–132. Routledge.
- Conover, M. D., Goncalves, B., Ratkiewicz, J., Flammini, A., and Menczer, F. (2011). Predicting the Political Alignment of Twitter Users. In *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*, pages 192–199, Boston, MA, USA. IEEE.
- Converse, P. E. (2006). The nature of belief systems in mass publics (1964). *Critical Review*, 18(1-3):1–74.
- Cox, G. W. and McCubbins, M. D. (1993). *Legislative Leviathan: Party Government in the House*.
- Craig, M. A. and Richeson, J. A. (2014). More diverse yet less tolerant? how the increasingly diverse racial landscape affects white americans' racial attitudes. *Personality and Social Psychology Bulletin*, 40(6):750–761.
- De Bruycker, I. and Rooduijn, M. (2021). The people's champions? populist communication as a contextually dependent political strategy. *Journalism & Mass Communication Quarterly*, 98(3):896–922.
- de Rijk, L. (2020). You said it? how mis-and disinformation tweets surrounding the corona-5g-conspiracy communicate through implying. In *MediaEval*.
- de Tocqueville, A. (1969). *Democracy in America*. Doubleday.
- Delobelle, P., Cunha, M., Cano, E. M., Peperkamp, J., and Berendt, B. (2019). Computational ad hominem detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pages 203–209.
- Di Cocco, J. and Monechi, B. (2022). How populist are parties? measuring degrees of populism in party manifestos using supervised machine learning. *Political Analysis*, 30(3):311 – 327.
- Dicey, A. V. (1914). *Lectures on the Relation Between Law & Public Opinion in England: During the Nineteenth Century*. Transaction Publishers.
- Dittrich, P.-J. (2017). Social networks and populism in the eu: Four things you should know. *Policy paper*, 19(2):1–20.
- Doherty, K. M. and Gimpel, J. G. (1997). Candidate Character vs. the Economy in the 1992 Election. *Political Behavior*, 19(3):177–196.

- Douglas, K. M. and Sutton, R. M. (2011). Does it take one to know one? endorsement of conspiracy theories is influenced by personal willingness to conspire. *British Journal of Social Psychology*, 50(3):544–552.
- Douglas, K. M., Sutton, R. M., and Cichocka, A. (2017). The psychology of conspiracy theories. *Current Directions in Psychological Science*, 26:538 – 542.
- Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., and Deravi, F. (2019). Understanding conspiracy theories. *Political psychology*, 40:3–35.
- Dunning, D. (2011). The dunning–kruger effect: On being ignorant of one’s own ignorance. In *Advances in experimental social psychology*, volume 44, pages 247–296. Elsevier.
- Durkheim, E. (2023). The division of labour in society. In *Social Theory Re-Wired*, pages 15–34. Routledge.
- D’Arienzo, M. C., Boursier, V., and Griffiths, M. D. (2019). Addiction to social media and attachment styles: a systematic literature review. *International Journal of Mental Health and Addiction*, 17:1094–1118.
- Egelhofer, J. L., Boyer, M., Lecheler, S., and Aldering, L. (2022). Populist attitudes and politicians’ disinformation accusations: effects on perceptions of media and politicians. *Journal of Communication*, 72(6):619–632.
- Ehrke, F., Grommisch, G., Busch, E. P., and Kaczmarek, M. C. (2023). Populist attitudes predict compliance-related attitudes and behaviors during the covid-19 pandemic via trust in institutions. *Social Psychology*.
- Ekman, M. (2022). The great replacement: Strategic mainstreaming of far-right conspiracy claims. *Convergence*, 28(4):1127–1143.
- Ekström, M., Patrona, M., and Thornborrow, J. (2018). Right-wing populism and the dynamics of style: a discourse-analytic perspective on mediated political performances. *Palgrave Communications*, 4(1):1–11.
- Elfardy, H. and Diab, M. (2016). Addressing Annotation Complexity: The Case of Annotating Ideological Perspective in Egyptian Social Media. pages 79–88.
- Enders, A., Uscinski, J., Seelig, M. I., Klofstad, C. A., Wuchty, S., Funchion, J., Murthi, M., Premaratne, K., and Stoler, J. (2021). The relationship between social media use and beliefs in conspiracy theories and misinformation. *Political Behavior*, 45:781 – 804.
- Engel, J. and Wodak, R. (2013). “calculated ambivalence” and holocaust denial in austria. In *Analysing Fascist Discourse*, pages 73–96. Routledge.
- Engesser, S., Fawzi, N., and Larsson, A. O. (2017). Populist online communication: Introduction to the special issue.
- Erikson, R. S. and Tedin, K. L. (2019). *American Public Opinion: Its Origins, Content, and Impact*. Routledge, New York, 10 edition.

- Erkulwater, J. (2012). Political Participation over the Life Cycle. *The Unheavenly Chorus: Unequal Political Voice and the Broken Promise of American Democracy*, pages 199–231.
- Ernst, N., Blassnig, S., Engesser, S., Büchel, F., and Esser, F. (2019a). Populists prefer social media over talk shows: An analysis of populist messages and stylistic elements across six countries. *Social media+ society*, 5(1):2056305118823358.
- Ernst, N., Esser, F., Blassnig, S., and Engesser, S. (2019b). Favorable opportunity structures for populist communication: Comparing different types of politicians and issues in social media, television and the press. *The International Journal of Press/Politics*, 24(2):165–188.
- Eysenck, H. J. (1966). The paranoid style in american politics and other essays, by richard hofstadter. *Political Science Quarterly*, 81(4):645–646.
- Fan, R., Zhao, J., Chen, Y., and Xu, K. (2014). Anger is more influential than joy: Sentiment correlation in weibo. *PloS one*, 9(10):e110184.
- Fareri, D. S. and Delgado, M. R. (2014). Social rewards and social networks in the human brain. *The Neuroscientist*, 20(4):387–402.
- Fischer, F. and Forester, J. (1993). *The argumentative turn in policy analysis and planning*. Duke University Press.
- Frey, D. and Wicklund, R. A. (1978). A clarification of selective exposure: The impact of choice. *Journal of Experimental Social Psychology*, 14:132–139.
- Funke, M., Schularick, M., and Trebesch, C. (2023). Populist leaders and the economy. *American Economic Review*, 113(12):3249–3288.
- Gabriel, Y. (2021). Narrative ecologies in post-truth times: nostalgia and conspiracy theories in narrative jungles? In R.A.W. Rhodes, R. and Hodgett, S., editors, *What Political Science Can Learn from the Humanities: Blurring Genres*, pages 33–56. Springer Nature.
- Gearhart, S., Moe, A., and Zhang, B. (2020). Hostile media bias on social media: Testing the effect of user comments on perceptions of news bias and credibility. *Human Behavior and Emerging Technologies*, 2.
- Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4):1799–1839.
- Gerbaudo, P. (2018a). Social media and populism: an elective affinity? *Media, culture & society*, 40(5):745–753.
- Gerbaudo, P. (2018b). Social media and populism: an elective affinity? *Media, Culture and Society*, 40(5):745–753.
- Gil de Zúñiga, H., Molyneux, L., and Zheng, P. (2014). Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships. *Journal of Communication*, 64:612–634.

- Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational linguistics*, 28(3):245–288.
- Glandt, K., Khanal, S., Li, Y., Caragea, D., and Caragea, C. (2021). Stance Detection in COVID-19 Tweets. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1596–1611, Online. Association for Computational Linguistics.
- Goffman, E. (2016). The presentation of self in everyday life. In *Social Theory Re-Wired*, pages 482–493. Routledge.
- Gorwa, R., Binns, R., and Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1):2053951719897945.
- Gould, R. (2012). *Graph theory*. Courier Corporation.
- Grande, E. (2000). Charisma und komplexität. *Leviathan*, 28(1):122–141.
- Graupmann, V., Jonas, E., Meier, E., Hawelka, S., and Aichhorn, M. (2012). Reactance, the self, and its group: When threats to freedom come from the ingroup versus the outgroup. *European Journal of Social Psychology*, 42:164–173.
- Greimas, A. (1983). *Structural Semantics: An Attempt at a Method*. University of Nebraska Press.
- Gründl, J. (2022). Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*, 24(6):1481–1499.
- Guadagno, R. E., Rempala, D. M., Murphy, S., and Okdie, B. M. (2013). What makes a video go viral? an analysis of emotional contagion and internet memes. *Computers in Human Behavior*, 29(6):2312–2319.
- Habernal, I., Wachsmuth, H., Gurevych, I., and Stein, B. (2018). Before name-calling: Dynamics and triggers of ad hominem fallacies in web argumentation. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, volume 1, page 386–396, Long Papers.
- Hadiz, V. R. and Chryssogelos, A. (2017). Populism in world politics: A comparative cross-regional perspective. *International Political Science Review*, 38(4):399–411.
- Haidt, J. (2022). Why the past 10 years of american life have been uniquely stupid. *The Atlantic*, 11:2022.
- Haidt, J. and Allen, N. (2020). Scrutinizing the effects of digital technology on mental health.
- Hameleers, M., Bos, L., and de Vreese, C. H. (2017). The appeal of media populism: The media preferences of citizens with populist attitudes. *Mass Communication and Society*, 20(4):481–504.

- Hameleers, M. and Schmuck, D. (2017). It's us against them: A comparative experiment on the effects of populist messages communicated via social media. *Information, Communication & Society*, 20(9):1425–1444.
- Harvey, J. H., Town, J. P., and Yarkin, K. L. (1981). How fundamental is "the fundamental attribution error"? *Journal of personality and social psychology*, 40(2):346.
- Hasan, K. and Ng, V. (2012). Predicting Stance in Ideological Debate with Rich Linguistic Knowledge. pages 451–460.
- Hassan, L. and Hamari, J. (2020). Gameful civic engagement: A review of the literature on gamification of e-participation. *Government Information Quarterly*, 37(3):101461.
- Hawdon, J., Costello, M., Bernatzky, C., and Restifo, S. J. (2022). The enthymemes of supporting President Trump: explaining the association between structural location, supporting the president, and agreeing with online extremism. *Social Science Computer Review*, 40(1):24–41.
- Hawkins, K. A. (2009). Is Chávez populist? measuring populist discourse in comparative perspective. *Comparative political studies*, 42(8):1040–1067.
- Henshilwood, C., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G., Mercier, N., Sealy, J., Valladas, H., Watts, I., and Wintle, A. (2002). Emergence of modern human behavior: Middle stone age engravings from south africa. *Science*, 295:1278 – 1280.
- Hernández Santiago, J., Cervantes, J., López-Chau, A., and Lamont, F. G. (2012). Enhancing the Performance of SVM on Skewed Data Sets by Exciting Support Vectors. In Pavón, J., Duque-Méndez, N. D., and Fuentes-Fernández, R., editors, *Advances in Artificial Intelligence – IBERAMIA 2012*, pages 101–110, Berlin, Heidelberg. Springer.
- Hofstadter, R. (2012). *The paranoid style in American politics*. Vintage.
- Holmes, D. S. (1968). Dimensions of projection. *Psychological bulletin*, 69(4):248.
- Hopster, J. (2021). Mutual affordances: the dynamics between social media and populism. *Media, culture & society*, 43(3):551–560.
- Hosseinzadeh, M., Azhir, E., Ahmed, O. H., Ghafour, M. Y., Ahmed, S. H., Rahmani, A. M., and Vo, B. (2023). Data cleansing mechanisms and approaches for big data analytics: A systematic study. *Journal of Ambient Intelligence and Humanized Computing*, 14(1):99–111.
- Hunger, S. and Paxton, F. (2022). What's in a buzzword? a systematic review of the state of populism research in political science. *Political Science Research and Methods*, 10(3):617–633.
- Introne, J., Korsunskaya, A., Krsova, L., and Zhang, Z. (2020). Mapping the narrative ecosystem of conspiracy theories in online anti-vaccination discussions. In *International Conference on Social Media and Society*, pages 184–192.

- lyyer, M., Enns, P., Boyd-Graber, J., and Resnik, P. (2014). Political Ideology Detection Using Recursive Neural Networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1113–1122, Baltimore, Maryland. Association for Computational Linguistics.
- Jackson, M. (2006). The enthymematic hegemony of whiteness: The enthymeme as antiracist rhetorical strategy. *JAC*, pages 601–641.
- Jankowski, M. and Huber, R. A. (2023). When correlation is not enough: Validating populism scores from supervised machine-learning models. *Political Analysis*, 31(4):591–605.
- Jhaver, S., Ghoshal, S., Bruckman, A., and Gilbert, E. (2018). Online Harassment and Content Moderation: The Case of Blocklists. *ACM Transactions on Computer-Human Interaction*, 25:1–33.
- Jing, E. and Ahn, Y.-Y. (2021). Characterizing partisan political narrative frameworks about covid-19 on twitter. *EPJ data science*, 10(1):53.
- Jones, M. D. and McBeth, M. K. (2010). A narrative policy framework: Clear enough to be wrong? *Policy studies journal*, 38(2):329–353.
- Jost, J. T., Federico, C. M., and Napier, J. L. (2009). Political Ideology: Its Structure, Functions, and Elective Affinities. *Annual Review of Psychology*, 60(1):307–337.
- Kamalski, J., Lentz, L., Sanders, T., and Zwaan, R. A. (2008). The forewarning effect of coherence markers in persuasive discourse: Evidence from persuasion and processing. *Discourse Processes*, 45(6):545–579.
- Kantartopoulos, P., Pitropakis, N., Mylonas, A., and Kyllilis, N. (2020). Exploring Adversarial Attacks and Defences for Fake Twitter Account Detection. *Technologies*, 8(4):64.
- Keith Norambuena, B. F. and Mitra, T. (2021). Narrative maps: An algorithmic approach to represent and extract information narratives. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3):1–33.
- Keith Norambuena, B. F., Mitra, T., and North, C. (2022). Characterizing social movement narratives in online communities: The 2021 cuban protests on reddit. *arXiv preprint arXiv:2208.04465*.
- Keith Norambuena, B. F., Mitra, T., and North, C. (2023). Mixed multi-model semantic interaction for graph-based narrative visualizations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, pages 866–888.
- Kemei, J., Alaazi, D. A., Tulli, M., Kennedy, M., Tunde-Byass, M., Bailey, P., Sekyi-Otu, A., Murdoch, S., Mohamud, H., Lehman, J., et al. (2022). A scoping review of covid-19 online mis/disinformation in black communities. *Journal of Global Health*, 12.
- Khalili-Mahani, N., Sawchuk, K., Elbaz, S., Hebblethwaite, S., and Timm-Bottos, J. (2022). A social-media study of the older adults coping with the covid-19 stress by information and communication technologies. In *International Conference on Human-Computer Interaction*, pages 346–364. Springer.

- Kim, S., Stavrova, O., and Vohs, K. D. (2022). Do voting and election outcomes predict changes in conspiracy beliefs? evidence from two high-profile us elections. *Journal of Experimental Social Psychology*, 103:104396.
- Kim, Y. (2022). How conspiracy theories can stimulate political engagement. *Journal of Elections, Public Opinion and Parties*, 32(1):1–21.
- Knobloch-Westerwick, S., Mothes, C., and Polavin, N. (2020). Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication research*, 47(1):104–124.
- Koch, L., Gorris, P., Prell, C., and Pahl-Wostl, C. (2023). Communication, trust and leadership in co-managing biodiversity: A network analysis to understand social drivers shaping a common narrative. *Journal of Environmental Management*, 336:117551.
- Koehler, J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes*, 56:28–55.
- Koller, V., Borza, N., Demata, M., Filardo-Llamas, L., Gustafsson, A. W., Kopf, S., Miglbauer, M., Reggi, V., Šarić, L., Brylla, C. S., et al. (2023). *Voices of Supporters: Populist parties, social media and the 2019 European elections*. John Benjamins.
- Kottak, C. P. (2017). *Cultural anthropology: Appreciating cultural diversity*. McGraw-Hill.
- Kovic, M. and Föchslin, T. (2018). Probability and conspiratorial thinking. *Applied Cognitive Psychology*, 32:390–400.
- Krishnan, S., Patel, J., Franklin, M. J., and Goldberg, K. (2014). A methodology for learning, analyzing, and mitigating social influence bias in recommender systems. In *Proceedings of the 8th ACM Conference on Recommender Systems*, pages 137–144, Foster City, Silicon Valley California USA. ACM.
- Krouwel, A., Kutiyski, Y., Van Prooijen, J.-W., Martinsson, J., and Markstedt, E. (2017). Does extreme political ideology predict conspiracy beliefs, economic evaluations and political trust? evidence from sweden. *Journal of Social and Political Psychology*, 5(2):435–462.
- Kruger, J. and Dunning, D. (2002). Unskilled and unaware—but why? a reply to krueger and mueller (2002).
- Kushwaha, A. K., Kar, A. K., Roy, S. K., and Ilavarasan, P. V. (2022). Capricious opinions: A study of polarization of social media groups. *Government Information Quarterly*, 39(3):101709.
- Laclau, E. (2005). Populism: What's in a name. *Populism and the Mirror of Democracy*, pages 103–114.
- Lai, M., Cignarella, A. T., Hernández Farías, D. I., Bosco, C., Patti, V., and Rosso, P. (2020). Multilingual stance detection in social media political debates. *Computer Speech & Language*, 63:101075.

- Lamberty, P. K., Hellmann, J. H., and Oeberst, A. (2018). The winner knew it all? conspiracy beliefs and hindsight perspective after the 2016 us general election. *Personality and Individual Differences*, 123:236–240.
- Lantian, A., Muller, D., Nurra, C., and Douglas, K. M. (2016). Measuring belief in conspiracy theories: Validation of a french and english single-item scale. *International Review of Social Psychology*, 29(1):1–14.
- Larkins, A., Vannamahaxay, S., Puttana, V., Chittavong, M., Southammavong, F., Mayxay, M., Boyd, D., Bruce, M., and Ash, A. (2024). Scaling up one health: A network analysis in lao pdr. *One health*, 18:100661.
- Lauran, N., Kunneman, F., and Van de Wijngaert, L. (2020). Connecting social media data and crisis communication theory: a case study on the chicken and the egg. *Journal of Risk Research*, 23(10):1259–1277.
- Lazarsfeld, P. F., Berelson, B., and Gaudet, H. (1944). *The People's Choice*. The People's Choice. Duell, Sloan & Pearce, Oxford, England.
- Lejano, R. P. (2021). Relationality: an alternative framework for analysing policy. *Journal of Public Policy*, 41(2):360–383.
- Leman, P. J. and Cinnirella, M. (2007). A major event has a major cause: Evidence for the role of heuristics in reasoning about conspiracy theories. *Social Psychological Review*, 9(2):18–28.
- Lerman, K. and Ghosh, R. (2010). Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks. *ICWSM*, 10.
- Li, C. and Goldwasser, D. (2019). Encoding Social Information with Graph Convolutional Networks for Political Perspective Detection in News Media. In Korhonen, A., Traum, D., and Màrquez, L., editors, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2594–2604, Florence, Italy. Association for Computational Linguistics.
- Li, M., Ma, T., Yu, M., Wu, L., Gao, T., Ji, H., and McKeown, K. (2021). Timeline summarization based on event graph compression via time-aware optimal transport. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6443–6456.
- Li, Y., Chen, C., Zheng, X., Zhang, Y., Gong, B., Wang, J., and Chen, L. (2023). Selective and collaborative influence function for efficient recommendation unlearning. *Expert Systems with Applications*, 234:121025.
- Lin, C.-Y. and Hovy, E. (2003). Automatic evaluation of summaries using n-gram co-occurrence statistics. In *Proceedings of the 2003 human language technology conference of the North American chapter of the association for computational linguistics*, pages 150–157.
- Lindström, B., Bellander, M., Schultner, D. T., Chang, A., Tobler, P. N., and Amodio, D. M. (2021). A computational reward learning account of social media engagement. *Nature communications*, 12(1):1311.

- Linegar, M., Kocielnik, R., and Alvarez, R. M. (2023). Large language models and political science. *Frontiers in Political Science*, 5.
- Ma, D., Li, S., Du, J. T., Bu, Z., Cao, J., and Sun, J. (2022). Engaging voluntary contributions in online review platforms: The effects of a hierarchical badges system. *Computers in Human Behavior*, 127:107042.
- Maoz, Z., Kuperman, R. D., Terris, L., and Talmud, I. (2006). Structural equivalence and international conflict: A social networks analysis. *Journal of Conflict Resolution*, 50(5):664–689.
- Marantz, A. (2020). *Antisocial: Online extremists, techno-utopians, and the hijacking of the American conversation*. Penguin.
- Marwick, A., Clancy, B., and Furl, K. (2022). Far-right online radicalization: A review of the literature. *The Bulletin of Technology & Public Life*.
- Matuszewski, P. and Szabó, G. (2023). The role of hyperactive Twitter accounts in the diffusion of political information. *Policy Studies*, pages 1–26.
- Mazzoleni, G. and Bracciale, R. (2018). Socially mediated populism: the communicative strategies of political leaders on Facebook. *Palgrave Communications*, 4(1).
- McCoy, J., Rahman, T., and Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, 62(1):16–42.
- McDonagh, E. C. (1962). Public opinion and american democracy. By V. O. Key, jr. New york: Alfred a. Knopf, 1961. 566 pp. *Social Forces*, 41(1):90–91.
- McDonnell, D. and Ondelli, S. (2022). The language of right-wing populist leaders: Not so simple. *Perspectives on Politics*, 20(3):828–841.
- McLuhan, M. (1994). *Understanding media: The extensions of man*. MIT press.
- McPherson, M., Smith-Lovin, L., and Cook, J. (2001a). Birds of a feather: Homophily in social networks. *Review of Sociology*, 27:415–444.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001b). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.
- Mercieca, J. (2018). America is in the middle of a battle over the meaning of words like ‘diversity’. *The Conversation*.
- Miani, A., Hills, T., and Bangerter, A. (2022). Interconnectedness and (in) coherence as a signature of conspiracy worldviews. *Science Advances*, 8(43):eabq3668.
- Moffitt, B. (2016). *The Global Rise of Populism: Performance, Political Style, and Representation*. Stanford University Press, Stanford.
- Moffitt, B. and Tormey, S. (2014a). Rethinking populism: Politics, mediatisation and political style. *Political studies*, 62(2):381–397.

- Moffitt, B. and Tormey, S. (2014b). Rethinking Populism: Politics, Mediatization and Political Style. *Political Studies*, 62:381–397.
- Mohr, J. W., Wagner-Pacifici, R., Breiger, R. L., and Bogdanov, P. (2013). Graphing the grammar of motives in national security strategies: Cultural interpretation, automated text analysis and the drama of global politics. *Poetics*, 41(6):670–700.
- Mørsted, B., Sapieżyński, P., Ferrara, E., and Lehmann, S. (2017). Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PLOS ONE*, 12(9):e0184148.
- Montag, C., Lachmann, B., Herrlich, M., and Zweig, K. (2019). Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. *International journal of environmental research and public health*, 16(14):2612.
- Morgan, J. (2011). *Bankrupt representation and party system collapse*. Penn State University Press.
- Morozov, E. (2014). The rise of data and the death of politics. *The Observer*.
- Motta, M., Callaghan, T., and Sylvester, S. (2018). Knowing less but presuming more: Dunning-kruger effects and the endorsement of anti-vaccine policy attitudes. *Social Science & Medicine*, 211:274–281.
- Mouffe, C. (2018). *For a left populism*. Verso Books.
- Mudde, C. (2004). The populist zeitgeist. *Government and Opposition*, 39(4):541–563.
- Muller, A. (2014). Het meten van discourscoalities met discoursnetwerkanalyse: Naar een formele analyse van het politieke vertoog. *Res Publica*, 56(3).
- Müller, J. F. (2023). An epistemic account of populism. *Episteme*, pages 1–22.
- Müller, J.-W. (2017). *What is populism?* Penguin UK.
- Myers, D. and Levy, M. (2018). Racial population projections and reactions to alternative news accounts of growing diversity. *The ANNALS of the American Academy of Political and Social Science*, 677(1):215–228.
- Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20(11):4366–4383.
- Nannini, L., Bonel, E., Bassi, D., and Joshua, M. M. (2024). Beyond phase-in: Assessing impacts on disinformation of the eu digital services act. *AI Ethics*.
- Németh, R. (2023). A scoping review on the use of natural language processing in research on political polarization: Trends and research prospects. *Journal of Computational Social Science*, 6(1):289–313.

- Nesi, J., Choukas-Bradley, S., and Prinstein, M. J. (2018). Transformation of adolescent peer relations in the social media context: Part 1—a theoretical framework and application to dyadic peer relationships. *Clinical child and family psychology review*, 21:267–294.
- Neville-Shepard, M. (2024). Masks and emasculation: Populist crisis rhetoric and the 2020 presidential election. *American Behavioral Scientist*, 68(1):97–111.
- Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., and Konstan, J. A. (2014). Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web, WWW '14*, page 677–686, New York, NY, USA. Association for Computing Machinery.
- Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2:175 – 220.
- Nikolaidis, N., Stefanovitch, N., and Piskorski, J. (2023). On experiments of detecting persuasion techniques in polish and russian online news: Preliminary study. In *Proceedings of the 9th Workshop on Slavic Natural Language Processing 2023 (SlavicNLP 2023)*, pages 155–164.
- Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. In *Algorithms of oppression*. New York university press.
- Obaidi, M., Kunst, J., Ozer, S., and Kimel, S. Y. (2022). The “great replacement” conspiracy: How the perceived ousting of whites can evoke violent extremism and islamophobia. *Group Processes & Intergroup Relations*, 25(7):1675–1695.
- Odgers, C. L. and Jensen, M. R. (2020). Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions. *Journal of Child Psychology and Psychiatry*, 61(3):336–348.
- Olaniran, B. and Williams, I. (2020). Social Media Effects: Hijacking Democracy and Civility in Civic Engagement. *Platforms, Protests, and the Challenge of Networked Democracy*, pages 77–94.
- Olbrechts-Tyteca, L. (1969). The new rhetoric: a treatise on argumentation. *Notre Dame, London*, 117.
- Oliver, J. E. and Rahn, W. M. (2016). Rise of the trumpenvolk: Populism in the 2016 election. *The ANNALS of the American Academy of Political and Social Science*, 667(1):189–206.
- Oliver, J. E. and Wood, T. J. (2014). Conspiracy theories and the paranoid style (s) of mass opinion. *American journal of political science*, 58(4):952–966.
- O’neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Osborne, TSD. and Rose, N. (1999). Do the Social Sciences Create Phenomena? The example of public opinion research. *British Journal of Sociology*, 50(3):367–396.

- Owen, D. (2020). The new media's role in politics/the age of perplexity: Rethinking the world we knew. *Open Mind*. URL: <https://www.bbvaopenmind.com/en/articles/the-new-media-s-role-in-politics>.
- Papakyriakopoulos, O., Serrano, J. C. M., and Hegelich, S. (2020). Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media*, 15:100058.
- Pape, R. (2021). Understanding american domestic terrorism: Mobilization potential and risk factors of a new threat trajectory. *Univ. Chicago, Chicago, IL, USA, Apr, 6*.
- Pape, R. A. and Ruby, K. (2021). The capitol rioters aren't like other extremists. *The Atlantic*, 2.
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. penguin UK.
- Pastor, M., Oostdijk, N., and Larson, M. (2024). The contribution of coherence relations to understanding paratactic forms of communication in social media comment sections. In *JADT 2024: Proceedings of the 17th international conference on statistical analysis of textual data (to appear)*.
- Patterson, M. and Monroe, K. R. (1998). Narrative in political science. *Annual review of political science*, 1(1):315–331.
- Pauwels, T. (2011). Measuring populism: A quantitative text analysis of party literature in belgium. *Journal of Elections, Public Opinion and Parties*, 21(1):97–119.
- Pauwels, T. (2014). *Populism in Western Europe: Comparing Belgium, Germany and the Netherlands*. Routledge.
- Pei, J. and Jurgens, D. (2020). Quantifying intimacy in language. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5307–5326, Online. Association for Computational Linguistics.
- Pei, Z., Sun, Z., and Xu, Y. (2019). Slang detection and identification. In *Proceedings of the 23rd conference on computational natural language learning (CoNLL)*, pages 881–889.
- Pereira-Kohatsu, J. C., Quijano-Sánchez, L., Liberatore, F., and Camacho-Collados, M. (2019). Detecting and monitoring hate speech in twitter. *Sensors*, 19(21):4654.
- Pezzini, I. and Terracciano, B. (2022). Negationist rhetoric and post-truth figures. transversal research in social media in the covid-19 era. *Estudos Semióticos*, 18(2):135–158.
- Philippe, A. and Valette, J. (2023). Immigration et délinquance: réalités et perceptions. *Lettre du CEPII*, (436).
- Piper, A., So, R. J., and Bamman, D. (2021). Narrative theory for computational narrative understanding. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 298–311.

- Piskorski, J., Stefanovitch, N., Da San Martino, G., and Nakov, P. (2023a). Semeval-2023 task 3: Detecting the category, the framing, and the persuasion techniques in online news in a multi-lingual setup. In *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 2343–2361.
- Piskorski, J., Stefanovitch, N., Nikolaidis, N., Da San Martino, G., and Nakov, P. (2023b). Multi-lingual multifaceted understanding of online news in terms of genre, framing, and persuasion techniques. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3001–3022.
- Pla, F. and Hurtado, L.-F. (2014). Political tendency identification in Twitter using sentiment analysis techniques. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 183–192.
- Plug, H. J. and Wagemans, J. H. (2020). From fact-checking to rhetoric-checking: extending methods for evaluating populist discourse. In *Vox Populi*, pages 236–252. Edward Elgar Publishing.
- Pogorelov, K., Schroeder, D. T., Burchard, L., Moe, J., Brenner, S., Filkukova, P., and Langguth, J. (2020). Fakenews: Corona virus and 5g conspiracy task at mediaeval 2020. In *MediaEval*.
- Raymond, C., Waller, I., and Anderson, A. (2022). Measuring Alignment of Online Grassroots Political Communities with Political Campaigns. *Proceedings of the International AAAI Conference on Web and Social Media*, 16:806–816.
- Reichstadt, R. (2015). Extending the domain of denial: conspiracism and negationism. *Diogenes*, 62(3-4):48–55.
- Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, 291(3):906–917.
- Ribeiro, F., Henrique, L., Benevenuto, F., Chakraborty, A., Kulshrestha, J., Babaei, M., and Gum-madi, K. (2018). Media Bias Monitor: Quantifying Biases of Social Media News Outlets at Large-Scale. *Proceedings of the International AAAI Conference on Web and Social Media*, 12(1).
- Rickers-Ovsiankina, M. A. (1928). *Ripresa delle azioni interrotte*. PhD thesis, Libreria editrice Julius Springer.
- Romero, D. M., Meeder, B., and Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th International Conference on World Wide Web, WWW '11*, page 695–704, New York, NY, USA. Association for Computing Machinery.
- Rooduijn, M. and Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6):1272–1283.
- Rosenberg, B. D. and Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, 4(4):281–300.

- Ross, L., Greene, D., and House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of experimental social psychology*, 13(3):279–301.
- Rydgren, J. (2017). Radical right-wing parties in europe: What’s populism got to do with it? *Journal of Language and Politics*, 16(4):485–496.
- Salvati, M., Giacomantonio, M., Pellegrini, V., De Cristofaro, V., and Leone, L. (2022). Conspiracy beliefs of italian voters for populist parties: The moderated mediational role of political interest and ideological attitudes. *Acta Psychologica*, 223:103508.
- Samih, Y. and Darwish, K. (2021). A Few Topical Tweets are Enough for Effective User Stance Detection. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2637–2646, Online. Association for Computational Linguistics.
- San Martino, G., Barrón-Cedeño, A., Wachsmuth, H., Petrov, R., and Nakov, P. (2020a). Semeval-2020 task 11: Detection of propaganda techniques in news articles. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, page 1377–1414.
- San Martino, G., Shaar, S., Zhang, Y., Yu, S., Barrón-Cedeño, A., and Nakov, P. (2020b). Prta: A system to support the analysis of propaganda techniques in the news. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, page 287–293.
- San Martino, G., Yu, S., Barrón-Cedeño, A., Petrov, R., and Nakov, P. (2019). Fine-grained analysis of propaganda in news article. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, page 5636–5646.
- Sapountzi, A. and Psannis, K. E. (2018). Social networking data analysis tools & challenges. *Future Generation Computer Systems*, 86:893–913.
- Sartori, G. (1969). Politics, Ideology, and Belief Systems. *The American Political Science Review*, 63(2):398–411.
- Sawhney, R., Wadhwa, A., Agarwal, S., and Shah, R. (2020). Gpols: A contextual graph-based language model for analyzing parliamentary debates and political cohesion. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4847–4859.
- Schaefer, R. and Stede, M. (2021). Argument Mining on Twitter: A survey. *it - Information Technology*, 63(1):45–58.
- Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge University Press.
- Segev, E. (2020). Textual network analysis: Detecting prevailing themes and biases in international news and social media. *Sociology Compass*, 14(4):e12779.

- Shahaf, D., Yang, J., Suen, C., Jacobs, J., Wang, H., and Leskovec, J. (2013). Information cartography: creating zoomable, large-scale maps of information. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1097–1105.
- Silva, A. (2016). Opinion Manipulation in Social Networks.
- Smith, N., Maddox, A., Southerton, C., and Baker, S. (2022). Conspiracy. *M/C Journal*.
- Smith, S. M., Fabrigar, L., and Norris, M. E. (2008). Reflecting on six decades of selective exposure research: Progress, challenges, and opportunities. *Social and Personality Psychology Compass*, 2:464–493.
- Snieckute, M. (2020). Values in populism and argumentative counter-strategies: the case of Viktor Orban. In *Vox Populi*, pages 219–235. Edward Elgar Publishing.
- Spohr, D. (2017). Fake news and ideological polarization. *Business Information Review*, 34:150 – 160.
- Stanyer, J. (2013). *Intimate politics: Publicity, privacy and the personal lives of politicians in media saturated democracies*. John Wiley & Sons.
- Statista (2022). Number of worldwide social network users. Accessed: 2024-03-25.
- Stede, M., Taboada, M., and Das, D. (2017). Annotation guidelines for rhetorical structure. *Manuscript. University of Potsdam and Simon Fraser University*.
- Stefanov, P., Darwish, K., Atanasov, A., and Nakov, P. (2020). Predicting the Topical Stance and Political Leaning of Media using Tweets. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 527–537, Online. Association for Computational Linguistics.
- Stier, S., Bleier, A., Lietz, H., and Strohmaier, M. (2018). Election Campaigning on Social Media: Politicians, Audiences, and the Mediation of Political Communication on Facebook and Twitter. *Political Communication*, 35(1):50–74.
- Storbeck, J. and Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and personality psychology compass*, 2(5):1824–1843.
- Stoycheff, E., Burgess, G. S., and Martucci, M. C. (2020). Online censorship and digital surveillance: The relationship between suppression technologies and democratization across countries. *Information, Communication & Society*, 23(4):474–490.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3):321–326.
- Sun, Y. and Zhang, Y. (2021). A review of theories and models applied in studies of social media addiction and implications for future research. *Addictive behaviors*, 114:106699.
- Sunstein, C. (2018). *# Republic: Divided democracy in the age of social media*. Princeton university press.

- Sunstein, C. R. (2001). *Republic.com*. Princeton university press.
- Sutton, R. M. and Douglas, K. M. (2020). Conspiracy theories and the conspiracy mindset: implications for political ideology. *Current Opinion in Behavioral Sciences*, 34:118–122.
- Sutton, R. M. and Douglas, K. M. (2022). Rabbit hole syndrome: Inadvertent, accelerating, and entrenched commitment to conspiracy beliefs. *Current Opinion in Psychology*, page 101462.
- Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S., and Voracek, M. (2011). Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *British Journal of Psychology*, 102(3):443–463.
- Tajfel, H., Billig, M. G., Bundy, R. P., and Flament, C. (1971). Social categorization and intergroup behaviour. *European journal of social psychology*, 1(2):149–178.
- Tangherlini, T. R. (2017). Toward a generative model of legend: Pizzas, bridges, vaccines, and witches. *Humanities*, 7(1):1–19.
- Tangherlini, T. R., Shahsavari, S., Shahbazi, B., Ebrahimzadeh, E., and Roychowdhury, V. (2020). An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, pizzagate and storytelling on the web. *PloS one*, 15(6):e0233879.
- Tateo, L. (2005). The Italian extreme right on-line network: An exploratory study using an integrated social network analysis and content analysis approach. *Journal of Computer-Mediated Communication*, 10(2):JCMC1022.
- Terren, L. T. L. and Borge-Bravo, R. B.-B. R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, 9.
- Theye, K. and Melling, S. (2018). Total losers and bad hombres: The political incorrectness and perceived authenticity of Donald J. Trump. *Southern Communication Journal*, 83(5):322–337.
- Thomas, W. I. and Thomas, D. S. (1938). *The child in America*. .
- Törnberg, P. (2022). How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences*, 119(42):e2207159119.
- Ulinskaitė, J. and Pukelis, L. (2021). Identifying populist paragraphs in text: A machine-learning approach. *arXiv preprint arXiv:2106.03161*.
- Urman, A. and Katz, S. (2022). What they do in the shadows: examining the far-right networks on Telegram. *Information, communication & society*, 25(7):904–923.
- USAFacts (2024). Economy data and statistics.
- Uscinski, J. E. and Enders, A. M. (2023). What is a conspiracy theory and why does it matter? *Critical Review*, 35(1-2):148–169.
- van Klink, B. and van der Geest, I. (2020). How to confront the populist challenge? In *Vox Populi*, pages 253–269. Edward Elgar Publishing.

- van Prooijen, J. and Douglas, K. M. (2017). Conspiracy theories as part of history: The role of societal crisis situations. *Memory Studies*, 10:323 – 333.
- Van Prooijen, J.-W., Krouwel, A. P., and Pollet, T. V. (2015). Political extremism predicts belief in conspiracy theories. *Social psychological and personality science*, 6(5):570–578.
- Van Prooijen, J.-W. and Van Dijk, E. (2014). When consequence size predicts belief in conspiracy theories: The moderating role of perspective taking. *Journal of Experimental Social Psychology*, 55:63–73.
- Venizelos, G. (2023). Donald trump in power: Discourse, performativity, identification. *Critical Sociology*, 49(4-5):647–667.
- Venturini, T. and Munk, A. K. (2021). *Controversy mapping: A field guide*. John Wiley & Sons.
- Verschueren, J. (2021). *Complicity in discourse and practice*. Routledge.
- Vitak, J., Ellison, N., and Steinfield, C. (2011). The Ties That Bond: Re-Examining the Relationship between Facebook Use and Bonding Social Capital. In *Proceedings of the 44th Hawaii International Conference On System Sciences*, pages 1–10.
- Vittori, D. (2022). Threat or corrective? assessing the impact of populist parties in government on the qualities of democracy: A 19-country comparison. *Government and Opposition*, 57(4):589–609.
- Volkova, S. and Bachrach, Y. (2015). On Predicting Sociodemographic Traits and Emotions from Communications in Social Networks and Their Implications to Online Self-Disclosure. *Cyberpsychology, Behavior, and Social Networking*, 18(12):726–736.
- Waeterloos, C., Walrave, M., and Ponnet, K. (2021). Designing and validating the Social Media Political Participation Scale: An instrument to measure political participation on social media. *Technology in Society*, 64:101493.
- Warren, M. E. (1999). What is political? *Journal of theoretical politics*, 11(2):207–231.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, 12(3):129–140.
- Weakliem, D. L. (2003). Public Opinion, Political Attitudes, and Ideology. In Janoski, T., Alford, R. R., Hicks, A. M., and Schwartz, M. A., editors, *The Handbook of Political Sociology*, pages 227–246. Cambridge University Press, 1 edition.
- Willaert, T. (2023). A computational analysis of telegram’s narrative affordances. *Plos one*, 18(11):e0293508.
- Wilson, F. G. (1939). James Bryce on Public Opinion: Fifty Years Later. *The Public Opinion Quarterly*, 3(3):420–435.
- Wilson, R. A. and Land, M. K. (2020/2021). Hate Speech on Social Media: Content Moderation in Context. *Connecticut Law Review*, 52:1029.

- Wirz, D. S., Wettstein, M., Schulz, A., Ernst, N., Schemer, C., and Wirth, W. (2019). How populist crisis rhetoric affects voters in Switzerland. *Studies in Communication Sciences*, 19(1):69–83.
- Witt, J. (2011). Action's effect on perception. *Current Directions in Psychological Science*, 20:201–206.
- Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. Sage.
- Wu, S., Lin, T.-C., and Shih, J.-F. (2017). Examining the antecedents of online disinhibition. *Inf. Technol. People*, 30:189–209.
- Xenos, M., Vromen, A., and Loader, B. D. (2014). The great equalizer? Patterns of social media use and youth political engagement in three advanced democracies. *Information, Communication & Society*, 17(2):151–167.
- Yan, Z. and Tang, X. (2023). Narrative graph: Telling evolving stories based on event-centric temporal knowledge graph. *Journal of Systems Science and Systems Engineering*, 32(2):206–221.
- Yousefi Nooraie, R., Sale, J. E., Marin, A., and Ross, L. E. (2020). Social network analysis: An example of fusion between quantitative and qualitative methods. *Journal of Mixed Methods Research*, 14(1):110–124.
- Yuan, M. and Chen, C. (2023). Lazybob at semeval-2023 task 9: Quantifying intimacy of multilingual tweets with multi-task learning. In *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 927–933.
- Yus, F. (1999). Misunderstandings and explicit/implicit communication. *Pragmatics*, 9(4):487–517.
- Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current directions in psychological science*, 10(6):224–228.
- Zanotti, L. (2019). *Populist polarization in Italian politics, 1994-2016: an assessment from a Latin American analytical perspective*. PhD thesis, Universiteit Leiden.
- Zanotti, L. (2022). Populism and polarisation. In *The Populism Interviews*, pages 79–84. Routledge.
- Zappin, A., Malik, H., Shakshuki, E. M., and Dampier, D. A. (2022). YouTube Monetization and Censorship by Proxy: A Machine Learning Prospective. *Procedia Computer Science*, 198:23–32.
- Zarefsky, D. and Mohammed, D. (2020). The rhetorical stance of populism. *Vox populi: Populism as a rhetorical and democratic challenge*, pages 17–28.
- Zeigarnik, B. V. (1927). *Über das Behalten von erledigten und unerledigten Handlungen*. J. Springer.
- Zhang, C., Zhou, Z., Peng, X., and Xu, K. (2023). *DoubleH: Twitter User Stance Detection via Bipartite Graph Neural Networks*.

- Zhang, W., Feng, S., Chen, Z., Lei, Z., Li, J., and Luo, M. (2022a). KCD: Knowledge Walks and Textual Cues Enhanced Political Perspective Detection in News Media. In Carpuat, M., de Marneffe, M.-C., and Meza Ruiz, I. V., editors, *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4129–4140, Seattle, United States. Association for Computational Linguistics.
- Zhang, X., Wei, Z., Du, Q., and Zhang, Z. (2022b). Social Media Moderation and Content Generation: Evidence from User Bans. *SSRN Electronic Journal*.
- Zügner, D., Akbarnejad, A., and Günnemann, S. (2018). Adversarial Attacks on Neural Networks for Graph Data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, pages 2847–2856, New York, NY, USA. Association for Computing Machinery.
- Álvares, C. and Dahlgren, P. (2016). Populism, extremism and media: mapping an uncertain terrain. *European Journal of Communication*, 31(1):46–57.