

Prevalent Discord

Exploring and estimating the prevalence of the type of user disagreement on news media Facebook posts discussing the Colombian peace process

(2020-2022)

Luis Felipe Villota Macías

January 2024

Abstract

This thesis is dedicated to exploring and understanding public reactions within negotiated peace settlements based on social media data. Concretely, to modeling public opinion and sentiment within the context of the Colombian peace process using a curated dataset of N= ~1.3 million user comments expressing discord on 15,509 Facebook posts, throughout three years (2020-2022). A critical period embracing unprecedented sociopolitical events such as the COVID-19 health emergency, the waves of the *Estallido social* and the rise to power of the first leftist president in the country. This information was facilitated thanks to the research initiative Agonistic Algorithms from the PUSHPEACE project at Lund University's Department of Political Science.

Based on specialized literature, predictive modeling with a binary logistic regression strategy was employed to discern if, on aggregate, the user comments on a post enabled by news media entities were predominantly antagonistic or not. This approach considered an array of predictors encompassing linguistic features, temporal indicators, engagement metrics, and contextual elements extracted from the Facebook posts. The results indicate limited explanatory capabilities of the exploratory model. Yet, it performed with moderate predictive accuracy on unseen data (64% of overall correct classifications). Regarding the particular status of prevalent antagonism, the model correctly identified this category in 8 out of 10 cases. The covariate referring to location of the publisher of the post emerged as the most influential factor. Despite the limitations, the results suggest that Bogotábased post publishers carry a higher likelihood of eliciting prevalent user antagonism in comments, compared to posts enablers from other locations.

Key words: Agonistic peace, antagonism, big data analytics, Colombia, discord,

Facebook, peace process, public opinion and sentiment, social media

Word count: 20,770

Acknowledgements

A mis padres, Irene y Luis Fernando; y a mis hermanos, María y Jose.

I am deeply grateful for being able to complete this puzzle.

As part of the first cohort of the MSc in Social Scientific Data Analysis at Lund University, my experience has been transformative and rewarding. I would like to especially thank the researchers in the PUSHPEACE project in the Department of Political Science, who not only provided the fundamental data and theoretical inputs for this thesis, but also exceptional prospects in my learning journey. I would like to extend my gratitude to family, friends, teachers, thesis advisors and medical team. My heartfelt appreciation also to Anny, who has walked with me since Medicon.

List of figures and tables

- Figure 1. Flowchart describing the research background.
- Figure 2. Diagram of the empirical material divided into four primary sets of information containing the comments (in millions) of Facebook users across posts published by the different public fan pages.
- Figure 3. The complete Annotation Model. Adapted from the Annotation Manual within the files of *PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook (2022)*. The dashed contours indicate the elements that were not applied in the end by AA.
- Figure 4. Summary plot of target variable "label_antagonism" in Set C with the frequency of its three levels (without removal of NAs). In Set D, user comments showing "not disagreement" are filtered out and hence, this level is dropped.
- Figure 5. Cumulative of user comments and unique FB posts made per year in Set D (yet to be cleaned).
- Figure 6. Final proportions of type of user disagreement in clean Set D N= 1,349,964 user comments.
- Figure 7. Histograms of clean Set D.
- Figure 8. Variable importance plot (normalized scores from vip() function)
- Figure 9. Plot displaying the distribution (in percentages) of prevalence status across the most important predictor.
- Table 1. Share of antagonism at page-level: "head and tail" of all the unique FB pages ran by news media entities in clean Set D (63); ranked by the share of antagonistic comments received from users in total; in the context of page popularity, activity and attention received on aggregate (from Jan. 2, 2020 to Oct. 19, 2022).
- Table 2. Top 4 unique FB posts (and their original attachment title) with the most cases of an antagonistic comments from users
- Table 3. Comparison of all models implemented, alongside their performance indices (rounding up)

Abbreviations

AA Agonistic Algorithms

CELAC Comunidad de Estados Latinoamericanos y Caribeños

CERAC Centro de Recursos para el Análisis de Conflictos

CEV Comisión para el Esclarecimiento de la Verdad, la Convivencia y la No Repetición

CINEP Centro de Investigación y Educación Popular

ELN Ejército de Liberación Nacional

EPL Ejército Popular de Liberación

FARC-EP Fuerzas Armadas Revolucionarias de Colombia, Ejército del Pueblo

FB Facebook

JEP Jurisdicción Especial para la Paz

M-19 Movimiento 19 de Abril

MOE Misión de Observación Electoral

OACP Oficina del Alto Comisionado para la Paz

PA 2016 Peace Agreement, signed on Sept. 26

PAM Peace Accords Matrix Barometer Initiative

PMI Plan Marco de Implementación

PPP Programa Por la Paz

RPA Revised 2016 Peace Agreement, signed on Nov. 24

RUV Registro Único de Víctimas

SIVJRNR Sistema Integral de Verdad, Justicia, Reparación y No Repetición

UBPD Unidad de Búsqueda de Personas Desaparecidas

Table of contents

A	lbstract	i
A	Acknowledgements	ii
L	ist of figures and tables	iv
Δ	bbreviations	χ.
1		
	1.1 Background of the study	7
	1.2 Research aim, objective and question	11
2	Context review	13
	2.1 A glance at the armed conflict's roots in Colombia	15
	2.2 The 2016 Colombian peace process	19
3	Theoretical framework	28
	3.1 Social media in a war-to-peace transition	29
	3.1.1 News stories framing: insights from war and peace journalism	34
	3.1.2 The lenses of agonistic peace	3 /
4	Methodology	41
	4.1 The empirical material	
	4.1.1 Set A: Initial data harvesting	
	4.1.2 Set B: Available dataset	
	4.1.2.1 Algorithmic annotation	
	4.1.3 Set C: Working dataset	
	·	
	4.2 Ethical considerations 4.2.1 On data collection.	
	4.2.2 Anonymization and data aggregation	
	4.3 Fundamentals of the empirical analysis	
	4.3.1 Binary logistic regression	54
	4.3.2 The model: variables, composition and metrics	
	4.4 Data preprocessing	58
	4.6.1 Variable adjustments	58
	4.6.2 Missing values and cleaning Set D	
	4.6.4 Natural language processing	60
5	Results	61
	5.1 Descriptive statistics and trends	61

	5.1 Goodness of fit and effectiveness of the model	66
	5.2 Importance of predictors	68
6	Discussion	70
	6.1 Limitations	
	6.1.1 Internal validity	71
7	Conclusion	74
R	deferences	76
A	ppendices	85
	Appendix 1. Table of Facebook and AA variables	85
	Appendix 2. Table of terms composing the search key for Facebook posts	87
	Appendix 3. List of categories for Facebook public fan pages	88
	Appendix 4. Examples of comments and their algorithmic labeling	90
	Appendix 5. List of unique Facebook public fan pages: the news media entities	92
	Appendix 6. Timeline of the Colombian peace process 2016-2022	93
	Appendix 7. Plots of missing data: Set B vs C vs D (before any variable transformation)	94
	Appendix 8. Temporal distribution of the creation of FB posts and comments: Set B vs C vs D (from first posts on Jan. 3, 2016 to last comments on Oct. 23, 2022) (before any data transformation or cleaning)	96
	Appendix 9. Summary of covariates in clean Set D	98
	Appendix 10. Summary of the final model's binary logistic regression coefficients	. 102

1 Introduction

In contemporary contexts where sociopolitical events unfold with volatility, gaining insights into public opinion and sentiment from social media is highly valuable to mitigate the uncertainty of offline realities (Nigam et al., 2017; Palakodety et al., 2020; Wang et al., 2016).

This is especially pertinent concerning the development and sustainability of real-world peace processes – in civil war settings. In environments like this, social media harbor key pointers of how citizens perceive and eventually react to a given process (García-Perdomo et al., 2022; Nigam et al., 2017). Once harvested, processed and analyzed, these signals can enhance the understanding of the societal "pulse" and complement other traditional forms of social inquiry such as surveys and polls in this particular offline milieu (Nigam et al., 2017).

Peace processes are momentous political phenomena that, broadly, comprise a cease or reduction of violence between conflicting parties – through the resolution of the causes of the conflict or the transformation of how it is managed (Galtung & Fischer, 2013; Lederach, 2005). As such, peace processes often involve prolonged dialogues, complex negotiations and multifaceted interactions between several political stakeholders in diverse arenas (Burgess on H. Saunders, 2004). Often, these peacemaking efforts are enterprises that encounter copious and significant obstacles and threats over time. For instance, since the 1990s, substantial amounts of negotiated settlements for civil wars have failed worldwide – and many countries have regressed to violence (Newman & Richmond, 2006; Nilsson & Söderberg Kovacs, 2011; Stedman, 1997). In this context, consequently, Stedman (1997: 5) noted that peace processes particularly contribute to 'heighten uncertainty and insecurity among ordinary citizens... [regarding a renewal of war] ...because they are the ones who have the most to lose'.

During war-to-peace transitions within a country, heated debates arise at all levels and sectors of society and through different channels – for instance, reflecting the stark divisions and opposing angles regarding the expectations and desired outcomes for a given negotiated peace arrangement (Dávalos et al., 2018). By gauging the "mood" or "pulse" of the wide-ranging discussions, the different stakeholders can be informed of societal perceptions and anticipate responses throughout the development of the different phases of a negotiated peace (Nigam et al., 2017). For example, those reactions from salient agents or the general public that, might either spur or spoil the process (Stedman, 1997).

In this logic, Nigam et al. (2017), argue that the institutional character of peace agreements – for civil wars, since the Cold War – has been often inclined towards the assumption that polls and surveys integrally draft the landscape of public opinion and sentiment towards the agreement. However, as it has been evidenced, since the late 1980s the majority of peace processes for civil wars have been "fragile" and prone to last ~3.5 years before the violence reappears (Mack, 2012; Nigam et al., 2017). This, consequently, implies the existence of further (latent) elements in society - that are having an impact on the sustainability of peace processes and agreements – are presumably not being captured by the conventional forms of inquiry into public opinion (Galindo, 2017; Nigam et al., 2017). More recently, a prime example of this was observed in the context of the PA-RPA in Colombia – the case on which this thesis is concentrated. Analogies of this sort of phenomenon, evidently, can also be found in other type of events such as the 2016 Brexit referendum, the 2016 Presidential Elections in the USA, the 2016 Constitutional Referendum in Italy and the 2018 Mexican Presidential Elections, among others (Ghitis, 2021; Nigam et al., 2017; Wang et al., 2016).

Hence, given its current pervasiveness, social media becomes a potent supplementary tool for understanding the undercurrents of public opinion and sentiment which, at the same time, exert a substantial influence on real-world political dynamics beyond digital spaces (Gil de Zúñiga et al., 2014; Nau et al., 2022; Nigam et al., 2017). To understand this, the traces of online conversational data – found in web-based and mobile social media services – can be harvested, analyzed and monitored in order to detect and evaluate patterns of users' behavior

(Theocharis & Jungherr, 2021). Regarding negotiated peace settlements, hostile messages and diffusion of pro-war manifestations are key factors to be sought for in users' activities, so that offline societal tensions and threats for peace processes can be assessed more comprehensively, for instance (Palakodety et al., 2020; Tabares Higuita, 2018).

In online environments, news media entities (in their digital form) emerge as a distinct category of actors deserving significant scrutiny. Their multifaceted and influential role as the major carriers of political communication in a democracy allows them to centralize debates on critical societal matters (Dahlgren, 2009; Jamieson & Kenski, 2017). Consequently, with the affordances of social media, they wield considerable potential to shape public opinion and thus, influencing offline political participation during a peace process (Gutiérrez, 2023; Pinzón Flórez, 2020; Salazar Mahecha, 2016).

This thesis focuses on the contemporary Colombian peace process, which has been a central subject of extensive and contentious debates across social media platforms – namely Facebook and Twitter – with news media entities at the heart of these interactions (Pinzón Flórez, 2020). In 2016, the mentioned process reached a pivotal moment in the country with the signature of the PA – which formalized an end to more than five decades of war between the FARC-EP and the Colombian State. However, shortly after this event, a Plebiscite took place as the official mechanism chosen by the government to legitimize and ratify the peace accord by popular vote. With an overall low voter turnout as a backdrop, the PA was rejected by Colombian citizens by a slim margin (less than 1% of all valid votes) (Comisión de la Verdad, 2022).

Remarkably, days before the Plebiscite, the main polls and opinion studies in the country indicated a comfortable winning margin – of ~60% on average – for the vote that accepted the agreement (Galindo, 2017; Palomino, 2016). Amidst the multiple factors for its failure, Nigam et al. (2017: 354) sustained that'(...) had proaccord stakeholders used social media as a tool or as a platform to listen and understand public opinion and sentiment toward the Colombian peace agreement, the outcome of the referendum could have been different and in the longer term it

could have resulted in a more successful implementation of the peace process.'. In the end, due to the results of the Plebiscite, the original PA was renegotiated between the parties involved and its text modified – by including certain demands from the political opposition. Ultimately, this new deal (RPA) was ratified through the Senate and Congress, and began its implementation on Dec. 2016.

The RPA laid certain foundations for constructive changes and peacebuilding efforts that have benefited the Colombian society – for example, by fostering a reduction in violence, initiating processes of reconciliation, and creating opportunities for social and economic development (Comisión de la Verdad, 2022). However, despite the progresses made, its provisions are far from being fulfilled and its implementation faces a myriad of challenges – such as a broader context of ongoing violence with multiple illegal actors, insufficient funds, political opposition, security concerns of ex-combatants, structural inequities in land ownership, among others (Noticias Caracol, 2023). Yet, above all, a polarized civil society remains – rendering even more uncertain the promise of a stable and lasting peace (Borda, 2018).

The purpose of this study is to explore the debate regarding the Colombian peace process on a major social media platform such as Facebook – a central space for modern-day political communication (Ríos Hernández et al., 2018; Tufekci, 2017) and also, historically the most popular social network in the country by number of users (OOSGA, 2023). This, in order to contribute to shed light on the development of offline public opinion and sentiment since 2020 until 2022. Given the combative nature of the discussion, I try to gain a deeper understanding of how FB users expressed disagreement (broadly, towards other opinions, actors or circumstances) through comments on posts directly related to the Colombian peace process.

My particular focus lies on exploring certain features from FB posts (message-based and other associated and contextual attributes), generated by media news entities on their FB public fan pages, as predictors for the prevalence status of the type of users' disagreeing comments on posts (antagonistic or not antagonistic). This is a classification problem; where, combining non-reactive quantitative methods and automated text analyses, within a logistic regression modelling

strategy (see Methodology), I investigate how the debate is shaped by a crucial actor – as an online proxy (Nigam et al., 2017). Ultimately, to probe these digital features, I adopt a social-scientific frame set at the junction of critical peace research (within political science) and social media research (as a working area of computational social science) (Theocharis & Jungherr, 2021).

Regarding critical peace research, I make use of conceptualizations of agonistic peace based on Strömbom's (2020) advanced analytical framework that permits its empirical measurement. Also, I draw on the works within PUSHPEACE, a peace research project in the Department of Political Science at Lund University, and related literature. The viewpoint of agonistic peace suggests that peace can be achieved by creating political environments where adversaries engage in nonviolent interactions as legitimate opponents (continuing a conflict in an agonistic way), thereby altering the course of a conflict from its violent trajectories between enemies (an antagonistic interaction) (Mouffe, 2013; Strömbom, 2020; Strömbom & Bramsen, 2022). In essence, this perspective of what peace is and should be, is traversed by the conflict transformation paradigm (Lederach, 2005, 2019). Here, the focus is not in terminating a conflict, but in changing its destructive manifestations (Strömbom, 2020). As such, conflicts are innate to human interactions and can actually be a source for constructive change in society (ibid). Thus, the effectiveness of the efforts within a peace process relies on their capacity to continue a conflict agonistically between different opposing identities (at diverse societal levels) (Strömbom & Bramsen, 2022).

For my investigation, all the above provides the theoretical lenses to be applied in the analysis of social media discussions. The empirical material used for the present thesis is provided by a PUSHPEACE research initiative – AA (2022) – regarding Colombia as a study case. This entailed the collection, isolation and analysis of the extensive debate orbiting the process on FB from early 2016 until late 2022. This filtered information pool (see Set B) has ~6.2 million users' comments on posts made by context-relevant entities on their public fan pages. In essence, by operationalizing the core premises of agonistic peace, AA's initial goal was to measure the variation of constructive (agonistic) and violent (antagonistic) interactions of users over time. This analysis required a specific annotation scheme,

a group of human coders to label a sample of data according to AA's Manual and the creation of algorithms to mark all the comments in the dataset. In this thesis, I deal with a subset of ~2 million of those comments pertaining only to posts published by news media entities on their fan pages (see Set C).

1.1 Background of the study

This thesis builds upon the works of AA - a research initiative enacted by the PUSHPEACE¹ project in collaboration with the organization Analyse og Tal². Guided by agonistic peace, a central concept that will be unpacked and used throughout this study as well, AA retrospectively collected, isolated and examined the wide-ranging FB debate (comments on posts pertaining to a diversity of public fan pages) concerning the Colombian peace process. AA's data collection covers almost seven years of discussions, starting from early 2016 - i.e., before the implementation of the RPA. Targeting users' conflictive comments on posts (those expressing any sort of disagreement or opposition towards a previous comment, post or another actor), the purpose of the joint research initiative was to quantify the development of agonistic (non-violent) and antagonistic (violent) interactions and its variations over time. This, also regarding offline critical events (those with high potential impact on the overall peace process such as elections or demonstrations, for instance). As a Research Assistant³ in PUSHPEACE, I was given the opportunity to link my thesis with their research frontiers and data (see Figure 1). Deciding to do so, it is necessary to introduce here the basic elements of the analytical and empirical material available for my investigation. Respectively, in Theoretical framework and Methodology, an expanded account of agonistic peace and an in-depth description of AA's data, proceedings and results are provided.

To fully comprehend AA's work, one has to refer back to the overarching objective of PUSHPEACE – which is to enhance the understanding and promotion of

¹ PUSHPEACE: Pushing the boundaries of peace research – Reconceptualizing and measuring agonistic peace (active from 2019 - 2024). The project is funded by the Riksbankens Jubileumsfond (Bank of Sweden Tercentenary Foundation), and it is led by Professor Lisa Strömbom (Principal Investigator) in the Department of Political Science at Lund University (Sweden). The project's website can be consulted in: https://portal.research.lu.se/en/projects/pushing-the-boundaries-of-peace-researchreconceptualizing-and-me.

An analysis and research company founded in 2014, with headquarters in Denmark and Norway. It is organized as a 'democratic cooperative' and specializes in delivering insights to customers on various 'complex' topics (digital spaces, hate speech, democracy, and participation, e.g.) through diverse methods (traditional and/or digital). Its website (in English) is: https://www.andnumbers.com/.

³ From November 2022 to May 2023.

sustainable peace practices. This is pursued through the development of novel conceptualizations of peace and its empirical deployment in specific case studies (societies transitioning from enduring violent conflicts to peace or less violent contexts). Overall, the project's analytical frameworks gravitate around the notion of agonistic peace. Succinctly here, this perspective points to peace as the generation of a political space, an arena of adversarial politics, where violent trajectories of a conflict between enemies can be shifted towards non-violent interactions among legitimate adversaries (Strömbom, 2020). This entails an underlying paradigm where conflict (or discord) is understood as an inevitable dynamic in human relations yet a source for societal change (Lederach, 2005). In other words, agonistic peace works with the idea that a conflict cannot be stopped, but its character can be transformed – ideally towards non-violent expressions (Strömbom, 2020). Operating at different levels of analyses, PUSHPEACE is then devoted to examining and contrasting the outcomes of the peace processes and peacebuilding efforts in Colombia, Northern Ireland, and Israel-Palestine. Synthesizing, the project chiefly aims to settle the 'viability of these peace processes over time in terms of agonistic peace' (Strömbom, 2020: 1).

According to Strömbom & Bramsen (2022), there are two fundamental dimensions that cluster the vital aspects that determine agonistic peace: institutional dynamics (changes in institutions that allow a conflict to continue politically without violence) and relational transformation (changes in identities that permit the mutual recognition and legitimization of conflicting parts via common grounds for disagreeing). This is based on Strömbom's (2020) crafting of a new analytical framework for studying agonistic peace, which interconnects these two aspects and allows for actual measurement to provide an account of the sustainability of a given peace process and the contributions of peacebuilding efforts. Consequently, qualitative and quantitative data has been gathered through diverse methods throughout the research (interviews, surveys, ethnographic approaches, e.g.). In later stages of the project, the results of the three case studies are to be validated on larger samples of protracted (persistent over time) conflicts. In sum, it can be argued that PUSHPEACE advances knowledge of peace by clustering different threads of agonistic peace literature and making its theoretical concepts (often abstract)

applicable for the empirical analysis of war-torn societies at different levels of analysis, 'from local to national to global' (Strömbom & Bramsen, 2022: 1238).

To come full circle, under the PUSHPEACE umbrella and probing the Colombian case in greater depth, AA was instantiated as a subproject to monitor the aggregated online FB public debate tendencies regarding the contemporary post-accord context of Colombia and its descriptive association with key offline events for peace and conflict in the country. To that end, AA produced PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook (2022). This work comprised the identification of hundreds of Facebook public fan pages ran by actors who were deemed relevant in Colombia for the peace process (e.g., politicians, government organizations, news media, NGOs, thematic pages). All available comments – from 2016 until 2022, on those fan pages' posts - were collected. Subsequently, a composite search key with various terms was created and applied (to the posts) to isolate the discussions around the peace process with expressions related to it. At the end, the dimensions of the filtered database resulted in ~ 6.2 million comments on the posts of the different public fan pages. On a sample of this data, a first phase of human coding was applied – with comments as the unit of analysis – following an annotation manual that operationalized the analytical framework for agonistic peace established by PUSHPEACE. The prime objective then, was on identifying a disagreement in a comment as either antagonistic (violent) or not antagonistic (non-violent; a simplified version of agonism). Parallel to that, two algorithms were also built and defined according to that coding scheme. Then, trained on the humanannotated sample, the algorithms were implemented to label the rest of the comments of the filtered dataset – i.e., each giving a probability score and a decision for both disagreement and antagonism comments separately. Ultimately, AA's efforts were presented in the form of a final (algorithmically tagged) dataset and a presentation detailing the major quantitative findings of the research (socialized within PUSHPEACE). This dataset and the initial descriptive analyses are the point of departure of my research.

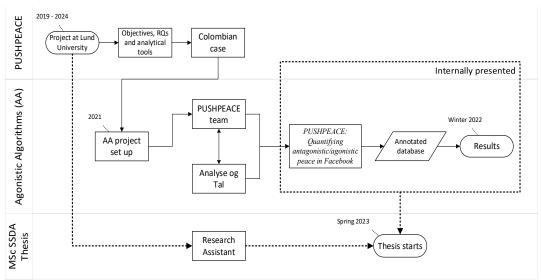


Figure 1. Flowchart describing the research background.

1.2 Research aim, objective and question

The aim of my thesis is to contribute to the understanding of how FB posts, published by news media entities on their FB public fan pages, attract users' discording comments (antagonistic and not antagonistic) through different post features. This, concerning posts made about the contemporary Colombian peace process. The study spans from Jan. 2, 2020 to Oct. 19, 2022.

With an explorative mindset, FB post features and associated attributes are interpreted as (basic) composing elements in 'the socialization of the news consumption process' within the developments of a negotiated peace settlement (Dutceac Segesten et al., 2022: 1116; Molnar, 2022; Nigam et al., 2017; Peng & Matsui, 2016). Therefore, I have the following specific research objective and research question:

(i) Explore the element or characteristic of news media FB posts that might best contribute to predict if a post will receive more antagonistic disagreements through user comments than responses without antagonism.

RQ1: Which feature of news media Facebook posts have the highest measure of variable importance in predicting the prevalence of type of disagreement (antagonistic or not) from users when discussing the Colombian peace process?

The nature of the online debate shown by AA was, non-surprisingly, conflict-ridden – since it is about a momentous period for the recurrent cycles of endemic violence and the long-sought efforts to achieve peace in Colombia (Perea Restrepo, 2006; 166). Indeed, this is in line with previous studies clearly showing that social media deliberations regarding the contemporary Colombian peace process have been a fertile ground for polarization, violent communication, and hate speech (e.g., Borda, 2018; Tabares Higuita, 2018; 160-161; Villa Gómez et al., 2023; 8).

So, by continuing scanning this contentious and divisive environment of online interactions through the optics of agonistic peace (Strömbom, 2014, 2020; Strömbom & Bramsen, 2022), this research provides macro-level insights on a

crucial parcel of the formation of public opinion regarding the Colombian transition – an operative part for ensuing offline political engagement (Bimber et al., 2012; 82; González-Bailón et al.; 1-2, 2011; Gil de Zúñiga et al., 2014; 626). With the rise of the platform-based societies (Dijck et al., 2018; 2), I ultimately contribute to understand the sway of digital news for citizens' political expressions in the midst of changes towards less violent realities.

2 Context review

The present-day armed confrontation in Colombia is a legacy from the bipartisan political violence of second half of 20th century. The conflict progressed 'initially [involving] the public armed forces and the insurgent guerilla organizations, later exacerbated by the gradual engagement of paramilitary groups, as well as by other state actors and civilian sectors' (Comisión de la Verdad, 2022: 179).

However, it is worth noting that this phenomenon can also be seen as part of an ampler historical context of recurring episodes of internal violence – even encompassing pre-Republican times (González Arana & Molinares Guerrero, 2010; Perea Restrepo, 2006). In a modern light, Colombia totals only a few decades of a "pause" from violent civil confrontations – ever since its independence processes from Spanish colony took place around 1810-1819 (ibid). Roughly, there has been a "relative break" from 1902 to 1946 (Ferry, 2012; Palacios, 2006, 2014). From there, violence has often been recycled despite the official enactment of copious amnesties, pardons and other mechanisms for multiple actors throughout time (ibid). The RPA is yet again an attempt to put forth a negotiated framework for peace, this time enacting record provisions and achieving the formal dissolution of the FARC-EP – the biggest guerilla in the country's history.

Until recently, the contemporary conflict in Colombia tallies ~9.5 million unique individuals registered as victims and ~12.5 million victimizing events perpetrated by diverse actors (with forced displacements accounting for ~78% of these cases) (RUV, 2023). Further calculations, situate the total human death toll above 450,000 (including the seminal violences from the 1940s) (Comisión de la Verdad, 2022). So forth, and more granularly, the ample register of violent circumstances resultant from war in the country (e.g., kidnapping, torture, extorsion, forced recruitment, affectations from landmines, imposed exile, sexual violence) continues to be populated namely by civilians (ibid). Altogether, this environment has been

commonly marked by elevated rates of impunity and 'grey zones' in the attribution of responsibilities (ibid: 179).

2.1 A glance at the armed conflict's roots in Colombia

The analysis of Colombia's current armed conflict – that peaked in terms of individual victims between 1995-2005 – frequently begins by tracing its origins back to La Violencia (1946-1958) (Caballero, 2016; Comisión de la Verdad, 2022; RUV, 2023). Particularly, there is ample academic agreement on the formative character of that era for the rise of the left-wing guerillas in the mid-1960s (Cairo & Ríos, 2019; Ferro & Ramón, 2002; Molano, 2000; Özmen & Medina, 2020; Roldán, 2002). Around 1950, Colombia was a diverse country of ~11 million inhabitants, ~60-70% rural – although urbanization was accelerating (Melo González, 2021; Moreno Montalvo, 2019).

La Violencia, represents an acute cycle of political violence occurring under a succession of governments (namely conservative and military) which, in turn, contributed to its persistence at both the local and national levels (Caballero, 2016). To a great extent, this was generated as a result of longstanding socioeconomic inequities and the hostile fanaticism between the two main political identities of the day – *conservador* and *liberal* (González Arana & Molinares Guerrero, 2010; Molano, 2000). Largely, the former group historically clustered the interests of the social elites, landowners and the Catholic Church; and the latter, those of reformists, great portions of urban middle classes and land-seeking peasants (Molano, 2000). Such issues of territory (land ownership and use), class and religion have been recurrent conflictual factors in Colombian society throughout time.

As background, during 1920-1950, and under international influxes tied to the aftermath of WWI and the emerging dynamics of the Cold War, Colombia started seeing a political left gaining traction – with the rise of labor unions, socialist organizations and the creation of the Colombian Communist Party (Salazar, 2005). Their political endeavors, whether peaceful or bellicose, were often forcefully suppressed by traditionalists and conservative agents at different levels (ibid). Yet, in fact, once in power, both currents were prone to engaging in abuses towards their opponents (ibid). In the territories, the non-government war actors from either faction, often took the forms of squads of (political) brigands that spiraled violence

affecting entire populations (Acuña Rodriguez, 2014). In certain moments, liberals gradually allied with armed communist groups – finding common causes for rebellion against civil and public conservative violent siege (ibid). In that context, the *autodefensas campesinas* sprouted in various locations (e.g., municipalities of Marquetalia, Rio Chiquito, and Pato) – i.e., the prelude for FARC-EP and other guerilla entities (ibid). Ultimately, more than 50 civil war sequences took place among *liberales* and *conservadores* in La Violencia (CIDOB, 2014).

In this landscape, notably, the assassination of Jorge Eliécer Gaitán, leader of the Partido Liberal, on April 9, 1948 aggravated the violence. He sustained a widespread appeal for social justice and reform, while calling for an end to bipartisan violence. But also, he mobilized liberal masses against what was considered as oligarchy and to resist the policies of President Mariano Ospina's conservative administration (1946-1950) – mostly related to order and the economy (Caballero, 2016; Ferry, 2012). Gaitán was gunned-down by a lone political fanatic in a principal street in downtown Bogotá (official version) (Sánchez Torres, 2001). This event triggered the seismic *Bogotazo* – an extraordinary episode of civil commotion in the Colombian capital, that further contributed to the deterioration of the bipartite confrontation across the country. La Violencia became La Gran Violencia (circa 1948-1957) (Salazar, 2005). Caballero (2016) posits that, from 1947 to 1948, the total reported killings at a national level three-folded (roughly from 14,000 to 43,000 cases).

Eventually, during La Violencia Tardía (1957-1964), violence moderately decreased – in part due to an incipient political stability stirred by new bipartisan arrangements, later known as Frente Nacional (1958-1974) (ibid). However, as Archila Neira (1997) asserts, while this allowed the two political parties to alternate power in Colombia (e.g., shifting the presidency every four years since), it also deepened social divisions and sparked new hostilities due to a lack of political inclusiveness of other actors (mostly rural sectors) for years to come. Ultimately, Pécaut & González (1997) note that between 1946 and 1964, political violence left ~200,000 deaths – and substantial trails of rural populations being forcibly displaced to urban centers, reshaping the country.

As a note, Colombia's rich geographic context has been an extremely important variable for the evolution of the internal conflict – e.g. facilitating or impeding settlements, passages or hideouts (Salazar, 2005). In La Violencia, for instance, the most "active" regions were the departments of Santander and Norte de Santander in Northeastern Colombia populated by the liberal guerillas of the *llanos orientales* (concerning ~15,000 individuals in 1953) (ibid). Later on, this was replicated in the rest of the country and even, some departments were considered as "independent republics" – unconstrained by official authority (e.g., departments of Cauca, Huila, Tolima, Vichada). The dynamics of bipartisan political violence and a culture of social enmity, civil and official, traversed the country – but rural areas were consistently the most affected (Archila Neira, 1997; Caballero, 2016; González Arana & Molinares Guerrero, 2010).

This mid-20th century Colombia was the breeding ground for the foundation of several modern guerillas with liberal, socialist, communist and anti-imperialist tendencies – 'sons of La Violencia' (Ferry, 2012: 30). Since the 1960s, ~34 of such entities appeared (Observatorio de Paz y Conflicto, 2016). Until 2016, there were only three active guerillas, the oldest: ELN, EPL and FARC-EP. The latter was dissolved with the RPA.

FARC-EP was co-initiated by "the last" La Violencia liberal guerilla veteran – Manuel Marulanda Vélez⁴ or 'Tirofijo' (Sure shot) – and Jacobo Arenas (Ferry, 2012). Escaping from the siege of the Colombian army on the insurgent territory of Marquetalia, Tolima (1964), they gave form to the guerilla with ~50 armed campesinos and their families (ibid). It then reached its power climax in 1998-2002, 'controlling ~30% of the national territory with ~17,000 rural fighters in 62 fronts – along with an extense network of urban militias and spies' (ibid: 33).

Gradually, drug trafficking became an acceptable (then vital) source of income for FARC-EP's revolution — while its commanders manifested their visions of achieving political power in the country (inspired by elements from Marxism-Leninism, Castro's Cuba, and ideas from Simon Bolívar, among others). With substantial support in certain regions (mostly rural and remote), and composed

17

⁴ His actual name was Pedro Antonio Marín, he was never captured and died of old age in 2008.

mainly by rural youngsters from impoverished rural families – its sociolegal status has been traditionally disputed between belligerent and narco-terrorist typologies in Colombia (e.g., abroad, the USA considered them terrorists from 1997 until 2021) (Ferry, 2012; Hege, 2021; Valcárcel Torres, 2008).

Apart from its armed confrontation with public forces and other armed groups, it is also responsible for numerous and systematic crimes against civilians (e.g., massacres, kidnapping, extorsions, car bombs, forced recruitment of children, mortar attacks, hijacking) and for environmental affectations (Comisión de la Verdad, 2022; Ferry, 2012). In this sense, Ferry (2012: 46) asserts that, Colombia's 'poverty, unemployment and the [social] resentment against official abuses' have greatly contributed for FARC-EP's endurance.

2.2 The 2016 Colombian peace process

'Why, despite the multiple agreements and peace processes, the armed conflict has not been completely shut down and instead is recycled?' (Comisión de la Verdad 2022: 24-25). This was the all-encompassing question that steered the works of the Truth Commission in Colombia, created after the final signature and implementation of the RPA. In essence, striving (again) for peace, this accord formally concluded more than fifty years of insurgent-counterinsurgent war between the all-time largest (communist) guerilla in the country, the FARC-EP, and the Colombian State.

During all these decades of war, no conclusive military victory was ever achieved by any government nor outlaw group. So, prior to the RPA, itself a contemporary conditional amnesty, the country already had a noteworthy history with other (peace) processes entailing indults, amnesties, pardons, disarmaments, ceasefires and demobilizations – 'largely circumscribed to negotiations between the State and (armed) agents vying for political power' (Cubides-Cárdenas & Vargas Parrasi on Uprimny, 2020: 46). As context, for example, from 1820 to 1995 there were roughly 88 indults and amnesties enacted (ibid). Other perspectives contemplate that, from 1953 to 2017, more than 30 legal provisions were also created – ultimately, trying to find a way between the intricacies of peace and justice in the setting of the Colombian internal war (ibid).

As a critical point, it was not until 1982, during President Betancur's administration, that certain armed rebel groups received some formal political recognition – coinciding with a period of great growth for organizations such as FARC-EP in terms of militants and fronts (Gutiérrez Loaiza, 2012; Salazar, 2005). However, by then, 'the escalation of violence connected to drug trafficking [by large cartels] was already taking place, a factor that injected [more] dynamism and amplified the internal armed conflict in Colombia' (Comisión de la Verdad, 2022: 402). In addition, in the 1980s, paramilitarism as a counter-insurgent phenomenon in society became more pronounced (Rivera, 2007).

Thus, as a clarification, what is typically known as the "Colombian peace process", actually refers to a series of peace processes that have taken place in the country – namely, but not exclusively, government-led enterprises to terminate the war with armed insurgencies since the mid-20th century (Nigam et al., 2017). But ultimately, it can be conceived broadly as various efforts, from different social sectors, to break from settings of enduring and intense violences, which have led to the buildup of individual and collective trauma (Comisión de la Verdad, 2022; RUV, 2023).

The present-day phase of such peace process, the focus of this thesis, involves the official dissolution of FARC-EP, in an unprecedented political feat that was formalized in 2016 (see Appendix 6). That milestone claimed to target the core causes of the internal armed conflict in Colombia (even transcending particular issues between the direct stakeholders) and to set the bases for the reconstruction of the torn national social fabric. This, so as to propitiate the non-repetition of (political) violence and contribute to satisfy the demand for truth and justice of victims. Moreover, it contemplated differential foci for impending public policies regarding ethnicity, gender and environment (Dabène & Le Cour Grandmaison, 2022). Thus, it promoted issues such as the recognition of rights over territories, restitution of lands and the sustainable use of natural resources, for instance; but also, the protection of the historical memory and cultural expressions of marginalized human groups.

On Sept. 26, 2016, the then FARC-EP commander-in-chief Rodrigo Londoño and President Juan Manuel Santos⁵, signed a political peace arrangement in Cartagena de Indias – after complex approaches and negotiations that were made public since 2012, but that date back to 2010 (Mesa de Conversaciones, 2018). This peace accord, unlike others, comprised a series of unique dispositions and addressed structural issues that signaled to society an ambitious promise of a stable and lasting *territorial peace* in the country (Aya Smitmans, 2017; Comisión de la Verdad, 2022). That is, a transformative accord – for the social, political and economic – that had the victims as the center and a special regard for the most war-torn zones

⁵ Santos governed in two periods, in 2010-2014 and 2014-2018. Londoño, back then known as 'Timoleón Jiménez' or 'Timochenko', was the last "official" chief of the Estado Mayor Central of the FARC-EP since 2011.

of the country (Dabène & Le Cour Grandmaison, 2022). With the political momentum, this time there were not empty chairs at the negotiation table' (Gutiérrez, 2023: 20). This, in contrast to the image of an elusive peace of the negotiation attempt in Caguán by President Pastrana's government (1998-2002), when FARC-EP commander Manuel Marulanda Vélez did not arrive and left his chair empty at the conversations table (ibid). In consequence, in order to lastly endow the PA with popular validation and legitimacy, Santos called for a single-question Plebiscite on Oct. 2, 2016 asking: 'Do you support the final agreement to end the conflict and build a stable and lasting peace?' (El Tiempo, 2016).

Yet, as a common phenomenon expected at the start of peace processes, there was, unsurprisingly, substantial sociopolitical opposition to the negotiated peace efforts in Colombia (Santos on the Kroc Institute, in Dabène & Le Cour Grandmaison, 2022; Nigam, 2017). In the atmosphere preceding the Plebiscite, the contradictors were significantly (not solely) grouped around the figure of former President Álvaro Uribe Vélez – the once bearer of Democratic Security (García-Perdomo et al., 2022). As background, Comisión de la Verdad (2022: 193) remarks that during Uribe's two tenures (2002-2006 and 2006-2010), '[...] the greatest military and strategic effort was employed to defeat the guerrilla forces [...] while negotiations with the AUC [major paramilitary group] were taking place.'. Throughout Uribe's terms, important numbers of guerrilla and paramilitary combatants demobilized and several outlaw groups were weakened, contributing to changes in the country's perception of internal security (Verdad Abierta, 2017). However, regarding human rights, there are questions based on 'a rise of extrajudicial executions, an increase of arbitrary detentions and a paramilitarization of institutions' (Rivera, 2007: 149). Uribe's government also negated the official existence of the internal armed conflict in Colombia, contrary to what his successor's did – causing an enlargement of the political rift between former associates (Santos was once Uribe's Minister of Defense) (OACP, 2018).

The signatories of the accord and its supporters campaigned on and anticipated overcoming "all disapproval" to the process with a sound victory in the Plebiscite. Despite an important political thrust, the democratic inclusion of historically marginalized sectors, the international support, the wide-spread edges of civil

mobilization and, above all, the demand of victims for truth, justice and reparation – the pro-accord movement was defeated. Out of ~12.8 million valid votes, the 50.2% answered 'NO' in the Plebiscite (Comisión de la Verdad, 2022a; Mesa de Conversaciones, 2018; MOE, 2017). Beyond these slim margins, notably, the voter turnout was also lower than expected by the supporters of the accord, with just the 37.4% of the entire voting population taking part, out of ~35 million possible voters then (MOE, 2017). Or, rounding up, 4 out of 10 enabled voters exercised their right to vote during the Plebiscite. All this, was not foreseen in the majority of surveys made and in numerous (offline and online) opinion circles – which, apparently, discarded other visions and generated some pro-accord bubbles that were incompatible to real-world developments and thus, creating societal shock for those favoring the peace process when burst (Acosta, 2016; Basset, 2017; BBC Staff, 2016; Rizzi, 2016).

Illustratively, as one of numerous attempts to identify the main contributors for the outcome of the Plebiscite, Aya Smitmans (2017) primarily points to (manifest and latent) perceptions of the general public. Concretely, the author posited that a popular divide that encapsuled the discussions leading to the Plebiscite, was on how the FARC-EP were seen: either as a rebel group with legitimate political motives or as mere brigands. This dichotomy is key, because (some) voters might have pondered that '[...] there cannot be a PA with bandits [...] there is no pact with delinquents. There is the penal system for them [...]' (Aya Smitmans, 2017: 172).

After this rejection, the Santos' government and the FARC-EP reached another deal contemplating modifications, which was again signed on Nov. 24, 2016 in Bogotá, amidst a sociopolitical atmosphere of acute uncertainty and volatility (García-Perdomo et al., 2022). A revised 310-page agreement (the previous had 297), now included some petitions from political contradictors that still had the original accord as subject of political discussion and negotiation – the more radical opposition merely negated the (legal) existence of the agreement after the Plebiscite. Here, the signatories concluded that it kept 'the same structure and spirit' of the first deal (Cosoy, 2016). After intricate legal elaborations and political disputes on their popular legitimacy, the modified document was finally ratified by Senate and Congress in late Nov. 2016. The official implementation of the accord began on

Dec. 1, 2016 – nurturing major societal expectations and divisions. So, with this, the FARC-EP came to a formal end as an armed organization and thus, the war with the largest armed rebel group in the country's history. Also, the new impulses for sociopolitical and institutional changes, commenced their materialization.

Overall, the definitive 2016 Peace Agreement included the following six-point agenda: 1.) Integral Rural Reform, 2.) Political Participation, 3.) The End of Conflict, 4.) Solution to the Problem of Illicit Drugs, 5.) Accord on the Victims of the Conflict and, 6.) Implementation, Verification and Ratification. It also created the SIVJRNR with three pillars: the JEP, the CEV and the UBPD. Altogether, it contains 578 requirements to be met within a 15-year timeframe guided by the PMI (Dabène & Le Cour Grandmaison, 2022). Fondo Colombia en Paz is the institution in charge of administering the economic resources for its completion.

Until recently, the evaluation of the implementation of the final 2016 Peace Agreement shows contrasting outcomes. The main (formal) entities in charge of monitoring the process such as the CINEP, CELAC and The Kroc Institute note advancements. The latter even counts with the PAM, an instrument that records implementation information in "real time" (ibid). Yet, despite the progress, there is coincidence on characterizing the execution as slow and irregular (ibid). According to the PAM latest report (late 2022), around 30.8% of the accord's stipulations were completed, 19.7% were at an intermediate state, 36.7% were at a minimum state and 12.8% were not initiated at all (Quinn & Gómez Vásquez, 2023). Points 3 (end of conflict) and 6 (mechanisms of implementation) are the most accomplished; and points 1 (rural matters), 2 (political participation), 4 (drugs) and 5 (victims) are the most underdeveloped (CINEP/PPP-CERAC, 2022; Dabène & Le Cour Grandmaison, 2022; Ramírez, 2022).

Overall, as a general sketch of developments, there has been a major decrease in reported war-related victims per year: 2016-2020 saw the ~46% less, on aggregate, than in 2011-2015 (~1.5 million cases) (Indepaz, 2021; RUV, 2023). However, from 2021 until today, the accumulated numbers (~750,000) are already approaching those of the PA's first luster (~830,000) (RUV, 2023). Continuing, many former FARC-EP fighters are still actively involved in the peace process and

more than 13,000 ex-combatants reintegrated into civilian life (Dabène & Le Cour Grandmaison, 2022). From the extinct FARC-EP, a political party – first named FARC, then Comunes – was constituted. Automatically, through the accord, it was granted seats in the Senate (5) and Congress (5) until 2026 to facilitate its political competition alongside certain prerogatives - apart from the possibility of participating in further elections with due guarantees and responsibilities (ibid). Additionally, 16 special constituencies for peace are now represented in Congress, giving victims a political representation until 2030 (ibid). In this scenario, the public policy of Reconciliation, Coexistence and Non-Stigmatization has been adopted too (Quinn & Gómez Vásquez, 2023). The inclusion of the SIVJRNR in the Colombian Political Constitution and its operative arrangement, have likewise signified a historic acknowledgement of the rights of victims to truth, reparations, justice, and non-repetition – all to encourage a peaceful coexistence (Dabène & Le Cour Grandmaison, 2022; Quinn & Gómez Vásquez, 2023). In this regard, the JEP inaugurated a transitional justice system, unparalleled worldwide, compliant with the Rome Statute (i.e., the bases of the International Criminal Court); the CEV has delivered its Final Report and associated multimedia of the internal armed conflict in late 2022; and, the UBPD in its humanitarian and extrajudicial mission, (by 2023) has received ~29,000 requests for missing persons of a (calculated) universe of ~104,000 possible cases – recovering until now a dozen individuals alive and around 1000 bodies (Dabène & Le Cour Grandmaison, 2022; Giraldo, 2023; Santos in Greene, 2023).

Challenges persist in various fronts. Perhaps the most structural, concerns rural matters and their reform – in a country where ~1% of large landowners have ~40% of the land (Alviar in Dabène & Le Cour Grandmaison, 2022). The crucial rural reform, calls for ~85% of all resources needed to satisfy the RPA – and almost 80% of the budget planned for this area has gone into "solving" other public expenses between 2020-2021 (ibid). Moreover, there are major issues regarding adjudication of lands, distribution of agrarian subsidies, access to territories and financial credits for rural populations (e.g., with significant geographic and gender cleavages) (ibid). On illicit crops and mining, there is preoccupation on their effective substitution through viable long-term initiatives for campesinos due to security and financial factors, all while ensuring economic and environmental sustainability (Quinn &

Gómez Vásquez, 2023). On other chief aspects, the reintegration process is greatly threatened as former combatants struggle with a stable (legal) livelihood (Dabène & Le Cour Grandmaison, 2022). And, according to the Indepaz website (2023), from Sept. 26, 2016 until Oct. 27, 2023, security gaps and re-emerging violences have led to 1,552 social leaders and human rights defenders assassinated, 444 massacres and 399 signatories murdered. In this setting, some FARC-EP signatories have even returned to arms to continue their outlaw endeavors (e.g., Segunda Marquetalia and other fronts).

It is important to consider that, during President Duque's tenure (2018-2022), the official posture towards the RPA was more aligned with ex-President Uribe's perspective (both affiliated to the Centro Democrático party). In Duque's period, Quinn & Gómez Vásquez (2023:3) assert that 'minimal variations' of advancements of the accord's execution were recorded, in contrast to 2016-2018. Due to this lag in the agreement's application, illegal armed groups and drug traffickers found new impetus – and violence has been reconfigured (Loaiza, 2022). During this presidential mandate, ~1,000 social leaders and human rights defenders were killed – along with more than 260 signatories from the demobilized FARC-EP (Gomez-Suarez & Franz, 2020; Loaiza, 2022). Additionally, the COVID-19 pandemic arrived and the government opted for strict quarantine regimes.

All of this, contributed to the onset of unparalleled episodes social protests nationwide, jointly known as *El estallido social* (2019-2021) – kindled by the announcement of a tax reform, which was later removed. The latter are commonly framed as legitimate (and "mostly" pacific) expressions conjured by multifactorial aspects, wherein significant social layers overflowed their demands beyond institutional channels in a context of poverty, marginalization, a national health emergency, insecurity and police brutality (Comisión de la Verdad, 2022; Uprimny, 2022). This social unrest is subject to diverse theorizations (and conspirations) (e.g., part of hybrid warfare from illegal groups, consequence of the dialectics of capitalism-imperialism, infiltrated by international agents) (Bernal, 2021; Castro Riaño, 2020; Uprimny, 2022). Here, Duque's administration met the social protest with 'intermediate concessions [...] unproductive negotiations [...] but overall, severe repression' (Uprimny, 2022:135). This response – entailing (excessive)

judicializations, militarization and exception measures to reestablish order – left 'dozens of civil fatalities [mostly youngsters] and an undetermined number of missing individuals' (ibid). Armed civilians also appeared in the scene of the repression (ibid). *El Estallido social* also left a myriad of questions on the limits of legitimate protest, due to its violent edges – resulting also in the death and injury of public officers, road blockages, detriment to commerce and public goods, thefts and sieges in critical sectors of major cities (Gomez-Suarez & Franz, 2020; Semana, 2021a, 2021b; Uprimny, 2022).

By navigating this atmosphere, Gustavo Petro – a former M-19 guerilla fighter turned into a non-armed democrat – became the President (2022-2026) at his third attempt (after 2010 and 2018). Now in head of the executive, the once Senator and Mayor of Bogotá has promised full commitment to the RPA. Hence, several of the accord's dispositions have been relaunched with certain institutional and budget arrangements (CINEP/PPP-CERAC, 2022; Quinn & Gómez Vásquez, 2023). Apart from this, his administration has proposed structural reforms and an ambitious agenda for peace (Paz Total) – seeking out peace through dialogues and negotiations with a multiplicity of armed actors, even with those without political ambitions (ibid). After Petro's mandate, only one more president will ought to see the full implementation of the RPA (Quinn & Gómez Vásquez, 2023). Ultimately, a delicate situation considering its initial 15-year timeframe. Failure here, can have a massive adverse effect on the viability of other future peace initiatives (Santos in Greene, 2023; Quinn & Gómez Vásquez, 2023).

Beyond its immediate impact of preserving human lives, the RPA represented a turning yet contested point in recent Colombia. Ultimately, it changed the national political agenda – unveiling (more) the 'real problems' of the country (Dabène & Le Cour Grandmaison, 2022). In this sense, Sandra Ramírez (2022), FARC-EP delegate during the Conversation Tables, asserts that 'the peace agreement meant an advancement for the humanization of our country' (ibid). And Rodrigo Londoño (2022), posits that the agreement opened discussions that were thought to be closed (regarding victims, land ownership, quality of life of campesinos, urban segregation, minorities in politics, for instance) (ibid). In his terms, the country has become more 'politized' – meaning greater dynamism for organization and

mobilization toward social causes (ibid). Meanwhile, ex-President Santos and ex-High Commissioner for Peace Jaramillo, both praise the uniqueness of the accord and its viability amid the odds (ibid).

The negotiated peace was, at last, made within complex oppositions across all levels of society, in a context of ongoing violence (stoked by other guerillas, paramilitaries, drug cartels, FARC-EP non-signatories and dissidents, official abuses, criminal gangs, e.g.). Concurrent debates – through multiple channels and with various cues – still reflect ingrained social divisions, both historical and emergent. In this context, competing memories and divergent narratives have clashed on a myriad of issues such as victimhood, the roots of the war itself, political participation of ex-guerrilleros, forgiveness and reparations (Uribe Sarmiento & Rodríguez Cortes, 2022). This, evidently, permeates all social spheres; from high places of power to the everyday life of families and the individual. It, thus, becomes apparent that navigating political landscapes in times of transition requires a nuanced understanding of historical contexts and the intricate interplay of public perceptions and sentiments – indeed, the Colombian case is an example where the mere concession of political undertones to the "other" can (in)advertently shape the trajectory of the social conflict.

3 Theoretical framework

In the fields of international relations and peace studies, the concept of agonism has offered theoretical possibilities for the advancement of diverse analytical models – at different levels – that allow to re-interpret what peace is and should be, as argued by Shinko (2008). This study adopts the notion of peace within the agonistic context, as understood by PUSHPEACE researchers and empirically probed by AA in social media.

As introduced earlier, affirming the plural and contested nature of peace, PUSHPEACE investigates whether and how a peace process is contingent on its potential to transform a deep-rooted violent conflict between collective identities, rather than extinguishing dispute from human interactions altogether (Strömbom, 2020; Strömbom & Bramsen, 2022). Its initial premise is that the success of a peace process, and subsequent peacebuilding efforts, is then embedded in continuing a conflict in an agonistic way – discarding conventional conflict resolution logics that are consensus-oriented or based on conflict termination (Strömbom, 2020). Thus, it questions liberal and neoliberal peace models and standardized top-down approaches (ibid). When practiced, in recent decades worldwide, such perspectives have yielded less than "optimal" outcomes in terms of the longevity of peace agreements - before recurring into civil violent conflicts (ibid). In this logic, a dimension of agonistic peace – relational transformation – was examined by AA in FB users' discussions, having the Colombian peace process as a case study. This is a way to analyze, in practice, a relevant (digital) forum where social identities are met, renegotiated and, possibly recognized and reconciliated within a conflict (Lehti, 2016; Strömbom & Bramsen, 2022).

Given the dual nature of social media for research – both an object (e.g., how content is consumed or generated by users) and an instrument (e.g., as a data bank) (Zeller on Welker and Kloß, 2022: 363-364) – AA's harvesting and analysis of FB data focused on the empirical study of 'behavioural and communicative patterns'.

It entailed identifying the action and participation of users, a fundamental area in social media research, as antagonistic or not antagonistic exchanges regarding the peace process in Colombia (Nau et al., 2022). Acknowledging this duality of social media during a peace process as a basis, I intend to focus on how news media messaging on FB affects the way in which users express their (discording) expressions – ultimately, using the lenses of agonistic peace.

3.1 Social media in a war-to-peace transition

In connection to a crucial political juncture like the conclusion of a civil war via negotiated peace settlements, Nigam et al. (2017) posit that social media wield a substantial influence – at diverse social levels – over the decision-making surrounding political participation over a peace agreement. Ultimately, shaping the (offline) evolution of mobilization, endorsement, and dissent concerning the given peace process (Cárdenas, 2021; Nigam et al., 2017). The essential mechanisms by which this occurs, are fundamentally facilitated by technologic affordances (Nigam et al., 2017; Palakodety et al., 2020; Theocharis & Jungherr, 2017, 2021). They ultimately become mediating tools for political information (Cárdenas, 2021). In turn, insights drawn from the analysis of the register of social media interactions, as signals or proxies, can be productive to offline realities concerning the peace process (Nigam et al., 2017).

The rise of digital technology has reconfigured social systems – bringing about variations in their compositions and operations by enhancing human capabilities at an accessible extent (Theocharis & Jungherr, 2017, 2021). The profuse emergence and popular usage of technologies such as online social networks and social networking sites, have consequently constituted a 'paradigmatic shift' for media and communications (Nau et al., 2022: 14-15). Through multiple affordances, these technologies now provide unparalleled opportunities in accessibility, reach and immediate diffusion of information, – along with advanced interactivity for users' discussions (Quan-Haase & Sloan, 2022).

The pervasive and disputed term *social media* lodges an important portion of these digital communication technologies nowadays (Nau et al., 2022). It does so under a guise of a 'digital Swiss army knife' – i.e., a continuous amalgam of features for connections-communications (ibid: 13-14, 17, 22-23). Due to its perceived nebulous nature by various authors, Quan-Haase & Sloan (2022; 15) propose an integrating definition following a review of (recent) literature – which is the one I adopt: 'Social media are web-based and mobile services that allow individuals, communities, and organizations to collaborate, connect, interact, and build community by enabling them to create, co-create, modify, share, and engage with content (user- or bot-generated)'. Such spheres, belong to a hybridized media ecosystem – where traditional modes of information dissemination and digital functionalities coexist and amalgamate (Chadwick, 2017; Nielsen & Schrøder, 2014; Thurman et al., 2019).

Focusing on FB, it mainly blends two typologies of social media in a 'modular' way: *social networking* and *messaging* (Nau et al., 2022: 15-16). On the one hand, it is a web-based service that allows users to establish public or semi-public profiles in a defined setting; through this, they can generate a list of other users they share a reciprocal connection with and, survey this listing and those directories created by others (Boyd & Ellison in ibid). On the other hand, FB Messenger facilitates the exchange of text and voice messages between users, along with making calls possible (audio and video supported), group chats, and the sharing of a variety of other media (Nau et al. on Gangneux, 2022).

About political communication, the digital revolutions have created unprecedented avenues for social and political interplay – concurrently leaving massive records of these virtual interactions (Theocharis & Jungherr, 2021). FB, as a source and gateway of political information, and has changed its 'production, distribution and consumption dynamics' (Nielsen & Schrøder, 2014; Theocharis & Jungherr, 2021: 1). Here, the flow of news information has transformed into a more dynamic and interactive cycle, interconnecting audiences who actively engage in creating and sharing news content in real-time (Chadwick, 2017; Theocharis & Jungherr, 2021; Wilding et al., 2018). These shifts have brought about profound and extensive consequences, including a disrupted news media landscape with a multiplicity of

information diets; where challenges such as the substitution of political news with entertainment and the spread of dis- and misinformation are frequent (Theocharis & Jungherr, 2021: 1; Wilding et al., 2018). Overall, these are 'new, noisy and deeply confusing information environments' (Theocharis & Jungherr, 2021: 2).

Surrounding the actual development of a peace process, the digital traces of (public) social media interactions can be, for instance, collected and examined regularly to identify and assess patterns of user-generated behaviors – where, hostility and diffusion of pro-war content are of particular interest to contribute to assess the viability or sustainability of a given process (Nigam et al., 2017; Palakodety et al., 2020; Wilson et al., 2012). Amidst these intricate communication courses, online violent dynamics may precipitate further fabrication and polarization of divergent social narratives – thus, possibly rendering even more challenging offline pursuits such as the search for reconciliation and the promotion of accountability in a warto-peace transition (Gutiérrez, 2023; Nigam et al., 2017; Palakodety et al., 2020; Tabares Higuita, 2018).

In online ecosystems, legacy and emergent news media entities still 'constitute one of the terrains where the construction of political subjectivity takes place', in Mouffe's (2013: 134) terms (Chadwick, 2017; Cook, 2012; Dahlgren, 2009; Gil de Zúñiga et al., 2014; Nielsen & Schrøder, 2014). Extant news media or journalistic agents have adapted to these digital complexities and continue to play a salient role in shaping public opinion and sentiment as central hubs for wide-ranging discussions on main political events (Chadwick, 2017; Dahlgren, 2009; Theocharis & Jungherr, 2021).

Regarding the PA-RPA, political communications in Colombia had already been witnessing important changes with long-term effects (Borda, 2018; Gutiérrez, 2023). As a brief background, in 2016 internet was growing steadily in Colombia to the point that, by the last trimester, its penetration reached the 32.5% of the national population with ~16 million subscribers – an increase of ~6 percentage points with respect to the same period in 2015 (Gutiérrez, 2023; Ministerio TIC, n.d.). Along this trend, social media surged to the extent that it became – by far – the chief motive to use internet by the time of the Plebiscite (Salazar Mahecha,

2016). Evidently, on the one hand, this facilitated that the stakeholders of the PA-RPA and the general public stayed informed-updated on relevant matters and engaged in more plural and inclusive dialogues about the momentous event (Nigam et al., 2017). On the other hand, however, it gave entrance to a spate of online 'misrepresentations, lies and exaggerations' about the sociopolitical conflict (Salazar Mahecha, 2016: 51).

In the Colombian landscape of social media, FB and Twitter have been the main grounds for this kind of communicative processes – often utilized as "trenches" of overflowing expressions during the developments surrounding the PA-RPA (Misión de Observación Electoral (MOE), 2017; teleSUR, 2016). As Uribe Sarmiento & Rodríguez Cortes (2022) remind us, the war was not only fought in the battlefields. The expansion of the armed conflict beyond its immediate settings in Colombia (often remote and rural) found fertile ground in social media platforms – becoming an effective realm for the transmission of past and new imageries, representations and symbolisms (ibid). Particularly, those related to a "sworn enemy" – who ought to be crushed (ibid). Indeed, in certain passages of the peace process, social media usage in Colombia resulted in exemplary highpoints of political proselytism, echo chambers, misinformation and polarization with great impact for offline realities (Borda, 2018; Cárdenas, 2021; Ríos Hernández et al., 2018; Tabares Higuita, 2018).

As a simple illustration, in the Final Report of Comisión de la Verdad (2022: 63), there is a brief note of an interview (Oct. 2016) in La República newspaper with Juan Carlos Vélez Uribe – the ex-manager of the NO campaign for the Plebiscite pertaining to the Centro Democrático. Regarding how the electoral strategy was conducted, especially in social media he said: "We appealed to outrage [instead of explaining the disadvantages of the PA], we wanted people to get out and vote irately [...] A councilman sent me an image of Santos and 'Timochenko' with a message questioning why money was being given to the guerrillas when the country was in a mess. I posted it on Facebook, and it reached six million people." Overall, although a mere glimpse to a campaign strategy, the NO was backed and financed by Colombian moguls such as the 'Ardila Lülle group, Grupo Bolívar, Grupo Uribe, Colombiana de Comercio and Codiscos' (El Espectador, 2016). Some of

them owners of highly relevant news media structures too. So, a substantial part of the NO message was carefully tailored and amplified, through many challenges, to accentuate class and geographic cleavages that generated and deepened hostile discord (ibid). After his statements, Vélez Uribe "quit" Centro Democrático – but the strategy was already open to public scrutiny (ibid).

3.1.1 News stories framing: insights from war and peace journalism

Information offered by multiple journalists and news media agents substantially allows citizens to better participate in a 'political life', according to their preferences (Dahlgren, 2009; D'Angelo, 2019; Harrison & Pukallus, 2023: 121). Yet, concerning sociopolitical communications, D'Angelo (2019: 3) posits that 'getting informed [through several paths and signs] [normally] engenders disagreement among individuals and political factions'. Amidst violent contexts, the way conflicts are portrayed and discussed through news stories can influence the discrepant sociopolitical inclinations to move either towards (further) hostilities or to non-violent dynamics (García-Perdomo et al., 2022). This is greatly based on the normative 'political and civic-developmental role' of news media activities (crucial in a democracy) (Harrison & Pukallus, 2023: 121). In this logic, war and peace journalism scholarship assesses distinct frames and framing effects when informing about war (and post-war) scenarios (García-Perdomo et al., 2022).

As background, news framing analysis has been developing since the 1970s as an 'integrative concept' by journalism researchers (D'Angelo, 2019: 1). Acknowledging the difficulty in having a cohesive paradigm for how human communication frames are built and set, along with their possible consequences at all societal levels, – social scientific endeavors are guided here by multiple domain-specific demands met with various theories and methodologies (Dahlgren, 2009; D'Angelo, 2019). Also, importantly, the provision, socialization and digestion of information highly varies across contexts and 'social environments' (Harrison & Pukallus, 2023: 122). Albeit the theoretical pluralism of frames and framing, there are scholar efforts to consolidate a basic model with four core types of frames; which are interconnected in many ways in news stories, impacting how information is produced, transmitted, and decoded: 'audience, journalist, news and issue' (D'Angelo, 2019: 8).

So, audience frames comprise adapting news media to fit diverse audience sensitivities and expectations; it thus, entails recognizing that individuals, through certain mechanisms, actively process and interpret information they come across or access and select based on their unique knowledge and experiences – to later affect

their sociopolitical decision-making and manifestations (ibid). Journalist frames (and newsroom frames) incorporate the judgements, biases, and professional conventions of those producing and socializing news (ibid). News frames are thematic and narrative constructs used to persuade an audience by conveying specific kinds of information and emphasizing some of its characteristics (e.g., the overall angle for issues) (ibid). Finally, issue frames deal with the deliberate placement and depiction of concrete subjects or problems within news media coverage (ibid). Together, these elements create assemblages that orient human communication, via verbal or and non-verbal signals that 'define a situation, describe its attributes and interpret its structures and rules' – thus, swaying public opinion and sentiment (Cappella & Jamieson, 1997; D'Angelo, 2019: 1; Gitlin, 2003).

With this considered and back to human communication in violent settings; on the one hand, there is war journalism. This style broadly focuses on and sensationalizes the perceptible outcomes of violence (e.g. with dramatic language and imagery), uses military dialectics without providing a rich background of the conflict and, frequently amplifies the narratives of dominant factions involved (García-Perdomo et al. on Galtung and Ruge, 2022). This news reporting strategy tends to exacerbate the audience's tensions and perpetuate dehumanizing narratives between antagonized parts – this is especially delicate in the formation of public opinion in the onset of a negotiated peace settlement following a civil war (García-Perdomo et al., 2022; Harrison & Pukallus on Paris, 2023; Nigam et al., 2017). Conversely, peace journalism is characterized by offering in-depth coverage, contextual analysis of the conflict and the inclusion of marginalized expressions – even exploring potential resolutions and pathways for conflict transformation (García-Perdomo et al., 2022; Harrison & Pukallus, 2023). In this sense, this peace style creates and promotes possibilities for considering non-violence (García-Perdomo et al., 2022).

Regarding news stories in social media here, predicting and understanding their framing effects on the behavior of the online audience is highly challenging and variable – due to the multiple confounders in the relationships within a hybrid media ecosystem in constant change (Chadwick, 2017; García-Perdomo et al., 2022; Valenzuela et al., 2017). Previous (experimental and non-experimental) research

has focused on different of social media news stories attributes and mechanisms (e.g. reach, values or positions messaged, types of associated multimedia) evaluated through various outcome metrics (e.g. shares or recommendations, length of users' conversations, type of users' comments and reactions) (Banks et al., 2021; García-Perdomo et al., 2022; Valenzuela et al., 2017). Overall, although the results of some studies 'do not correlate' (García-Perdomo et al., 2022: 7), there is amounting empirical evidence supporting the effective sway of frames on users' reactions (e.g., ultimately showing shifts or reinforcement of beliefs towards a sociopolitical issue) (Valenzuela et al., 2017). Interestingly, regarding the Colombian case on FB, García-Perdomo et al. (2022: 15) concluded a prevalence of war frames in news stories (in n= 841 articles) by local journalistic agents (contrary to foreign media) when informing about the peace process (June - Oct. 2016) – particularly, there was an inclination to "give voice" to dominant political factions and a 'two-party conflict reporting'. Additionally, the authors posit that social media audiences were also prone to discuss the negotiated peace settlement in conflictive and not in negotiation terms (ibid). Finally, although acknowledging the lack of causal robustness of these results, the authors argue that these findings might have contributed to the results of the Plebiscite and the (later) lagged implementation of the RPA (ibid).

3.1.2 The lenses of agonistic peace

The term agonism, a derivative in modern English from the Greek agon, namely points to a competition, contest or struggle (Daqing, 2010; Howatson, 2011). Its etymological understanding is multidimensional and frequently suggests a coming together of individuals engaging in various types of disputations – whether they are casual or formal, physical or intellectual, as argued by W. Daqing (2010). A primary idea that transpires from this, is the interplay and differentiation between opposing forces vying for a specific objective (e.g., a prize, honor, privileges, resources) – often with a conflicting dyad as a starting point of departure. As a historical note, some scholars suggest that while competitiveness is a widespread trait in human behavior, within the Classical Greek context, this was experienced as a pervasive system and tenet that permeated the daily lives of citizens across different citystates – as a result, the agon is seen as a culturally distinctive and influential element (Barker, 2009; Daqing, 2010). According to Barker (2009), the agonal or agonistic spirit of the ancient Greeks is well represented in the profuse historical accounts of competitive festivals, sport events and debates, for instance - which centralized their public life and impacted the world thereafter.

As a concept that re-emerged in Western political thought – notably through the works of Burckhardt, Nietzsche, and later championed by Schmitt in the 1920s – agonism has been treated in diverse strands of scholarly literature (Daqing, 2010; Scudder & White, 2023). These various threads offer a multifaceted perspective on how socio-political conflicts are understood and addressed in different contexts – as seen in different accounts of agonistic politics in Arendt, Connolly and Foucault, for instance. In such a context, Shinko (2008) – a key reference within the works of PUSHPEACE – asserts that peace, under an agonistic paradigm, can be re-imagined as a dynamic ground of open-ended political competition; in contrast to the static vision often associated with the liberal tradition (absence of war). In this regard, in order to go beyond conventional interpretations that ultimately reinforce notions of elusiveness and unattainability of peace – it is essential to re-emphasize its political dimension and strengthen the scholarly problematization of the issue of violence in human relationships (ibid).

Before delving further into the development of agonistic peace, it is essential to acknowledge an underlying pattern that emerges when characterizing sociopolitical conflicts, based on the perspectives mentioned before. It is the view that conflict is a non-linear and cyclic process – which is both, innate to human interactions and necessary for the shaping of societal identities (Lederach, 2019; Mouffe, 2013; Shinko, 2008; Strömbom, 2014, 2020). Conflicts, under this vision, are conceived as the very fabric of how a society is articulated and evolves (Körppen & Ropers, 2011; Lederach, 1996, 2005; Shinko, 2008). In this sense, human conflict has exhibited two basic forms throughout time: one marked by hostility, destruction, and violence, and the other characterized by constructive, non-violent interactions (Strömbom & Bramsen on Buckley-Zistel and Kriesberg, 2022). Agonistic peace, then, pivots on authors such as Lederach (1996), who contend that generating constructive changes is yet feasible in these social assemblages. This ultimately involves shifting 'the flow of human interaction in social conflict from cycles of destructive relational violence toward cycles of relational dignity and respectful engagement' (Lederach, 2005: 42). In sum, the realization that conflicts can be transformed - not just terminated, transcended or be given an absolute stop – is key to understand the ensuing premises of agonistic peace in the context of PUSHPEACE (Strömbom, 2020).

Indeed, within this scope, the relational aspect of peace is a crucial element. With a keen focus on the complexity of social connections, this idea has been notably enriched and influenced here by the insights of scholars such as Connolly (1993), Lederach (2005), Galtung & Fischer (2013), Richmond (2020), and Mac Ginty (2021). This comprehension of peace underscores a necessity to refocus analytical efforts toward power dynamics and structures in relationships – as opposed to seeing moral prescriptions upon which the foundations of peace might rely (Strömbom & Bramsen, 2022). As a consequence, the agonistic face of peace allows to identify and question out any relational configuration that might be impeding a continuous socio-political contestation and, possibly be contributing to a violent conflict (e.g., historical contexts, economic variables, political undercurrents) (Shinko, 2008). As a consequence, there are scholar avowals to 'transform unequal relationships' of power (Strömbom & Bramsen, 2022; 1237-1238) and to caution against any conception of peace that promotes a re-instalment

of 'hegemonic structures of domination, exclusion, and marginalization' (Shinko, 2008; 488). Therefore, there is a particular attention on spotting diverse manifestations of political segregation (Mouffe, 2005, 2013; Shinko, 2008). Within this framing, agonistic peace can be also contemplated as an way to 'problematize, resist and possibly alter' those operating mechanisms – from the local to the international spheres (Shinko, 2008: 476). After considering such factors, Shinko (2008) – drawing significant inspiration from Foucault – posits that the process of generating a moment for the agonistic struggle can commence.

Following this logic, the agonistic approach suggests conceiving peace a space-time of dynamic and continuous political contestation between conflicting agents (Mouffe, 2013; Shinko, 2008; Strömbom & Bramsen, 2022). Here, for instance, different social parts can engage in adversarial disputes to express their judgments, advocate for their causes, or compete for political power in a non-violent manner. The notion that the competition remains unconclusive holds philosophical significance, as it implies that all positions and determinations taken by the involved parties are inherently transient (Shinko, 2008). Consequently, the disputing individuals – from their subjectivity – must acknowledge and remain mindful of the forthcoming developments and the evolving dynamics of the relationship that they share with the counterpart (ibid). Thus, as fundamental pillars for the creation and maintenance of agonistic peace, the competing agents must recognize and respect each other – adhering to certain norms to see this through (e.g., democratic ideals and the rule of law) (Scudder & White, 2023). With these parameters, theoretically, differences can be addressed non-violently without entailing its eradication (Shinko, 2008; Strömbom & Bramsen, 2022). However, the practical implementation of the agonistic approach (for peace) depends on a contextual appropriateness - thus, highly challenged in settings with severe violence and unequitable distribution of power and economic resources (Aggestam et al., 2015; Scudder & White, 2023; Strömbom & Bramsen, 2022). Moreover, when viable (even at the conflict's periphery), agonistic peacebuilding should balance fluctuating sociopolitical diversity with the respect between the plural expressions without standardized approaches and platforms (ibid).

Synthesizing, the emphasis of agonistic peace is on thoroughly changing the interplay between social agents that have each other as targeted foes to be obliterated, onto a respectful yet disputed rapport between rivals (Strömbom on Lehti, 2020). That is, a relationship in which opposing parts no longer regard each other as enemies to be annihilated because of their differences – or, in other words, shifting an antagonistic way of relating as argued by Strömbom & Bramsen (2022). Instead, the agents become 'adversaries who recognize the legitimacy of the demands of their opponent' without resorting to violence – constituting an agonistic interaction, in Mouffe's terms (2013: 130). Any effort to stop a conflict – by seeking win-win solutions, harmony, consensus, and mutual understanding between opponents as ultimate goals - might, for instance, curtail the potential for the legitimate contestation against an existing status quo (Strömbom & Bramsen, referring to Mouffe, 2022). Furthermore, such attempts can increase the likelihood of a conflict relapsing into violence (ibid). In other words, if a conflict is not given the opportunity to continue and develop – with channels of expression as relief valves as proposed by agonistic peace – it tends to antagonism (ibid).

According to Strömbom (2020; 9, 16), this long-sought shift towards non-violence encompasses all societal levels and domains, and should be adapted to build an inclusive society (with significant representation of agents and attendance to their demands) while respecting institutions as limits for the agency of each part. In general, as noted by Strömbom & Bramsen (2022; 1239), the channeling and inhibition of social violent vectors into constructive expressions while sustaining the conflict is what is understood as the creation of agonistic peace. The means by which this is realized is through the delegitimization of violence, politicizing the conflict, and providing opportunities for a conflict to progress agonistically – elements covered in the above-mentioned themes of institutional dynamics and relational transformation (ibid).

4 Methodology

At the intersection of computational social science and traditional social science, this research works with social media data – collected from FB by AA using unobtrusive methods. On the one side, this framing entails conceiving society's digital traces as valuable and practical data sources (Ledford on Keuschnigg and Strohmaier, 2020; 329; Samuel & Buchanan, 2020; 3). On the other side, more importantly, it implies that these imprints are bound to social-scientific theoretical parameters to make sense of them and inform 'disciplinary-specific research question' (Nau et al., 2022; 18).

While different facets of the Colombian peace process on social media have undergone extensive consideration and examination, its multiple dimensions and connections to offline phenomena are yet to be fully exhausted (Botero, 2017; Cárdenas, 2021; García-Perdomo et al., 2022; Nigam et al., 2017; Ríos Hernández et al., 2018; Ruiz et al., 2022; Tabares Higuita, 2018). The perused research landscape involves different social media platforms, diverse focalized time frames, various types of data and assorted methodologies utilized, for instance. Regarding FB-based studies, their numbers and temporal scopes call for expansion, since they tend to gravitate towards the 2016 Plebiscite understandably (Cárdenas, 2021; Nigam et al., 2017; Ríos Hernández et al., 2018; Tabares Higuita, 2018). Notably, in general terms, the accessibility of social media datasets, is a major entry barrier for this sort of research (Jungherr, 2017; Theocharis & Jungherr, 2021). Hence, the value of using AA data.

In this context, I focus on news media FB posts as information and debate hubs for the peace process, where users' reactions (i.e., disagreements via textual comments, assigned as antagonistic or not) can be shaped regarding the Colombian peace process (2016-2022). I frame this phenomenon as a prediction problem – a classification task – and I study it with an exploratory supervised machine learning 'modeling mindset', following a logistic regression strategy (Molnar, 2022, 2023).

Here, the model itself is algorithmic (a function) and not made up of statistical assumptions-hypotheses and its 'performance is valued over internal validity' (i.e., by assessing the accuracy of the model on novel data and other precision scores) (ibid: 49). Based on AA's work in a 'novel research environment' for the empirical assessment of agonistic peace, the inputs for my model entail a combination of textual and numerical features (Molnar, 2022, 2023; Schroeder, 2020; Theocharis & Jungherr, 2021: 8). Thus, within my modeling strategy, feature engineering for certain input variables includes automated content analysis also. My aim is to determine which of those elements is the most relevant concerning the enhancement of the model's predictive performance (R. D. Peng & Matsui, 2016).

Given the inherent controversial context of a negotiated peace settlement, this is intended to accumulate a body of evidence of FB signals and look for elements in news media's FB posts that contribute to "tilt" the way in which users disagree (e.g., with ideas, people, processes) – in either violent (antagonistic) or non-violent (not antagonistic) manners. Important aspects for comprehending and promoting real-world peacebuilding – through the lenses of agonistic peace. The possible underlying factors, latent trends or patterns in the interactions, are subject of further developments or tests in future research puzzles (e.g., validation across social media platforms and with survey data, model performance comparisons), in terms of the theoretical framework used.

Theocharis & Jungherr's (2021: 8) overarching premise asserting that 'grand-theorizing is not viable in the field of social science due to the highly contextualized and dependent nature of its findings', elementally sustains the pertinence of my explorative approach. Apart, from the narrowed (yet justified) focus, I acknowledge using a specific section of social media data from a single platform, as a signal archive for offline phenomena. In this sense, any social media data point holds symbolic value (see Limitations) (Donnay, 2017; Molnar, 2022; Schroeder, 2020). Here, for instance, one observation can encapsulate both explicit and implicit representations of users' behavior – demanding nuanced understanding (Schroeder, 2020; Theocharis & Jungherr, 2017, 2021). Thus, linking (online) signals with concrete (offline) behaviors presents a great challenge – due to the dynamic nature

and diversity of social cues, each embedded within different social media platforms and shaped by platform-specific functionalities (ibid).

Ultimately, albeit the restrictions and without the pretention to generalize the findings, my study follows the logic that social media insights can substantially benefit the decision-making of different stakeholders during a negotiated peace settlement – eventually, influencing its sustainability. This, by monitoring and providing valuable and complementary information (apart from polls and other social inquiries) about public opinion and sentiment during transitions towards less violent realities – guided by the notion of agonistic peace (Cárdenas, 2021; Nigam et al., 2017).

4.1 The empirical material

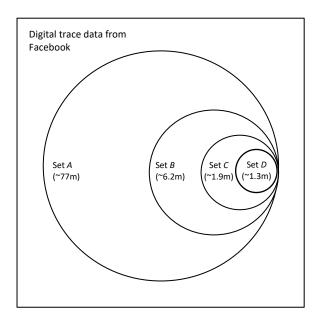


Figure 2. Diagram of the empirical material divided into four primary sets of information containing the comments (in millions) of Facebook users across posts published by the different public fan pages.

4.1.1 Set A: Initial data harvesting

AA's final research product was *PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook* (2022) – which comprised, among others, social media datasets and documentation regarding findings and procedures. According to its internal socialization, the AA research initiative started with the identification and curation of ~600 Facebook public fan pages. This selection targeted pages (assumed to be) managed by individual and collective agents who were deemed relevant to the offline Colombian peace process context, in broad terms. From those digital spaces, the AA team retrospectively gathered all accessible comments of (anonymized) users made across all page posts created between 2016 and 2022. The data collection was made in two moments: one from April/May 2021, and the second in June 2022. This concluded in ~77 million observations gathered along the variables allowed by Facebook's API when accessed by AA.

4.1.2 Set B: Available dataset

The empirical material that I was given access to was a narrowed-down version of that first data scraping. The dimensions of this dataset are of N= 6,217,154 rows of comments along with 38 variables (see Appendix 1). This was the result of the elaboration and application of a composite search key made up of 62 common terms in Spanish by AA – based on rules established by Analyse og Tal – to isolate the posts directly related to the Colombian peace process (see Appendix 2). The terms or words, were repeatedly tested and refined to avoid unrelated posts and thus, irrelevant comments as much as possible.

Those comments were nested in 72,441 unique posts made by 565 different public pages. At the same time, those pages are classified within 82 categories, clustering the different actors, institutions and sectors running them (such as civil organizations, NGOs, politicians, news media and government-related) (see Appendix 3). The pages have a total of 336 unique geolocations. Regarding the concrete time span of this dataset (with format %Y-%m-%d %H:%M:%S in UTC time zone), the earliest post was created on 2016-01-03 23:01:49 and the latest on 2022-06-24 13:47:14. For comments, the earliest was published on 2016-01-04 00:28:11 and the latest on 2022-10-23 06:57:17. Overall, in this dataset there were 11,686,900 missing values (NAs) in total. These cases were found in the following 8 variables: "score 0 disagreement", "score 1 disagreement", "score 0 antagonism", "score 1 antagonism", "disagreement", "antagonism", "comment_count", "page like count" (see Appendix 7).

4.1.2.1 Algorithmic annotation

Notably, all the comments in this dataset have been labeled by two different algorithms created by Analyse og Tal – one for detecting *disagreement* and the other for *antagonism*. Every comment received a probability score from 0 to 1, with a true-or-false interpretation logic for both categories. Subsequently, each comment was given a conclusive classification tag as either *disagreement* or *not disagreement*

and *antagonism* or *not antagonism*, respectively (see Appendix 4). The label *agonism* was ultimately removed by AA following PUSHPEACE.

Breaking this down, initially, the presence of a *disagreement* in a comment was sought for and detected by a first trained algorithm. This meant effectively identifying if a comment contained an expression of textual dissent – which could be either directed towards the same post, previous comments or statements, other users, actors or policies in general, etc. This was embedded, and thus examined, within the context of the online discussion on the Colombian peace and conflict processes – although, evidently, there were cases in which comments deviated from this (e.g., users commenting on other unrelated issues or ads). Yet, for AA's research aims, the only comments that were considered not relevant for further study were the ones labeled *not disagreement*. Secondly, once a *disagreement* was distinguished, each comment was labeled either as *antagonism* or *not antagonism*. The former category implied detecting any expression of violence, enmity, hate or misrecognition with online or offline referrals; and the latter, the absence of any kind hostile communication while manifesting *disagreement*.

On the Methodological Notes (2022) by M. Morales, within AA's research documentation, is registered that the algorithmic marking was ultimately trained on the annotations of three hired human coders on a sample of Facebook data – based on AA's Annotation Manual. These individuals were all Spanish-speaking females, with an educational background in the social sciences and from ages ranging from 19 to 21. Various meetings and coding workshops were organized between AA members and them since Aug. 2021 – where instruction, practice and discussions took place. Even, a WhatsApp group was made as a support platform to discuss and solve issues (e.g., questions on hard-to-classify comments) as the coding process went on. During the process, these matters were documented and kept track of to better inform the overall research (e.g., files on "tricky" comments). AA presentation (2022) showed that in five months, the annotators coded ~24,000 comments. This was then "fed" to train and test the classification algorithms.

Overall, the human coding was an iterative process in which key issues were addressed – such as the coders' comprehension of all research concepts once they

had gone over comprehensive parts of the data to be labeled. In this sense, a first stage of human annotation was finished in Dec. 2021, and a concluding session was held for final reflections. It was the case that a re-annotation phase began shortly after due to significant discrepancies between coders regarding the concept of *agonism* – a third and final category initially planned to be measured by AA (very scarce in the human-annotated comments, ~3%). After reviewing a sample of recoded *agonistic* comments, PUSHPEACE Principal Investigator and the rest of the team decided to have the category of *not antagonistic* as a simplified version of *agonism*. Here, some of the characteristics identified in the re-labeled subset of comments with *agonism*, were considered insufficient to be understood in the conceptual level of *thin recognition* under the *identity change* aspect of *agonism*, for instance (Strömbom, 2020). In sum, the final model used by human coders – who provided the training datasets for the algorithms to be applied on the filtered AA dataset – is reflected below.

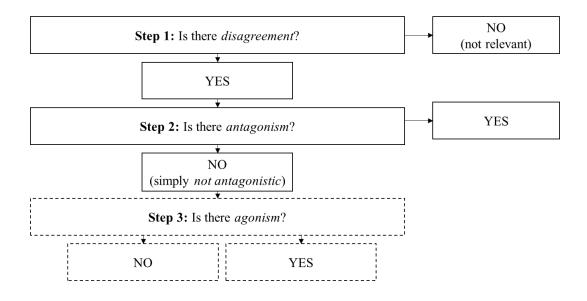


Figure 3. The complete Annotation Model. Adapted from the Annotation Manual within the files of *PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook (2022)*. The dashed contours indicate the elements that were not applied in the end by AA.

In this study, the information recorded by AA is regarded as a valuable window into what Theocharis & Jungherr (2021: 6-7) referred to as 'probably the most vibrant political information environments of our time [denoting social media platforms]'.

This sort of digital trace data coming from those atmospheres offers a thorough description of how users interact with and through digital services, revealing contextual information that was inaccessible in former times (Nau et al., 2022). The availability of the AA dataset – spanning nearly seven years – allows to explore the dynamics of the contemporary Colombian peace process comprehensively, regarding the evolving discourses surrounding it on Facebook – the most widely used social network in the country. All this, in recent times where acquiring data through social media APIs for academic research has become increasingly challenging and resource-intensive (Breuer et al., 2022). Moreover, using this data implies an evident advantage of working with a novel implementation of the analytical framework for agonistic peace, as understood by PUSHPEACE.

4.1.3 Set C: Working dataset

According to the objectives of this study, from the 82 categories of Facebook public fan pages contained in the AA dataset, I solely focus on news media entities. These agents are widely recognized for informing and engaging the public through the presentation of current events, inquiries, analyses, and commentaries on local and global matters; distributing information across different channels (e.g., print, broadcast, online, and social media) (Cook, 2012). Following AA's metadata in "category" (see Appendix 3), all cases that fell under the conceptual umbrella of news media entities were identified and isolated – in total it concerned 19 types of actors running a FB public fan page ⁶.

Those items (along associated data) were retrieved from Set B, composing the final working dataset Set C with N= 1,966,687 rows (comments) and 38 variables. Here, those users' comments were made across 17,672 different posts on 64 unique public fan pages (see Appendix 5). In terms of geolocation, there are 22 unique page

⁶ These were (in their original names): "Newspaper", "Tv-kanal", "Medier/nyheder", "TV Network", "Media/News Company", "Nyheds- og mediewebsite", "Tidsskrift", "Radiostation", "Avis", "News & media website", "Radiostation", "Medie-/nyhetsbedrift", "Udsendelses- og medieproduktionsselskab", "Media/news company", "Kringkastings- og medieproduksjonsselskap", "Medier", "News & Media Website", "TV-kanal" and "Radiokanal".

pointers. Concerning time ranges, the first post was made on 2016-08-17 02:39:28 and the last on 2022-06-24 13:47:14. On the other hand, the first comment was created on 2016-08-17 04:03:29 and the last on 2022-10-19 00:26:13. Regarding NAs, there are 3,335,919 cases in total, distributed along 4 variables: "comment count" (1,519,731),"score 0 antagonism" (605,396),"score 1 antagonism" (605,396), "antagonism" (605,396) (see Appendix 7). The last three variables have NAs simply because several users' comments were not disagreements – and were left blank without a respective score and label. They do not actually represent missing information, but rather "procedures" of the algorithmic labeling. Regarding "comment count", it is not clear why this the total count of comments on a post was not computed initially in ~70% (4,375,108) of all observations' metadata in Set B. In Set C, this represented ~77% of the total rows; and, in terms of unique FB posts, 11,657 out of 17,672 had NAs in this sense. At page level, this meant 62 of 64 had incomplete information. However, this was later solved by building simple functions using "id comment" to group all unique comments associated to each unique "id post". Thus, without replacing or removing NAs, there was another way to proceed with the analysis.

Overall, Set C harbors \sim 2.5% of the observations (users' comments) in Set A, and \sim 31.6% of Set B. On its own, Set B is \sim 8% of Set A. There, in other words, the keywords used by AA to isolate the posts concerning the Colombian peace process reduced the initial information by \sim 92%.

4.1.3 Set D: Only disagreements

Regarding the algorithmic classification of comments in Set C N=1,966,687 comments, 43.18% were antagonistic disagreements (849,391); 26.02% were disagreements without antagonism (511,900); and, 30.78% were not disagreements (605,396). Very similar proportions to what AA found in Set B.

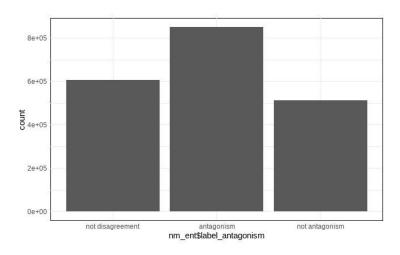


Figure 4. Summary plot of target variable "label_antagonism" in Set C with the frequency of its three levels (without removal of NAs). In Set D, user comments showing "not disagreement" are filtered out and hence, this level is dropped.

Since I seek to uncover key predictors from FB posts that attract and influence antagonism and not antagonism from users, I remove all comments without disagreements (605,396) to create Set D. This resulted in N = 1,361,291 rows with 38 variables, and now $\sim 62.39\%$ (849,391) were antagonistic comments and $\sim 37.60\%$ not antagonistic (511,900).

Here, there are 16,146 different posts enabled by 64 unique public fan pages. In Set D, the first post was made on 2017-01-10 16:34:07 and the last on 2022-06-24 13:47:14. Concerning comments, the first was created on 2017-01-10 18:08:03 and the last on 2022-10-19 00:26:13. Finally, once the comments without disagreements are dropped, the only NA that remains (before any transformation and cleaning) is "comment_count" which is not problematic as exposed earlier. In total, Set D is ~69% of set C.

4.2 Ethical considerations

4.2.1 On data collection

As a Research Assistant in PUSHPEACE, during Nov. 2022 – May 2023, I was given the opportunity to work with the AA dataset to study the Colombian case in greater depth for this student project. I had no involvement in the collection of the data used for this thesis. That is, I did not have any participation in the data retrieval, parsing, storing and in the original queries made - in Zeller's (2022) terms. Consequently, my understanding on issues such as the precise utilization of the API for extracting information from Facebook and certain steps of data processing, is limited (the programming script was restricted, for instance). This is also the case for details on the first phases of human annotation (e.g., limited knowledge on the different intercoder reliability scores) and on the details of the ethical registration and approval of AA's initiative. Nevertheless, within the scope of this study, my efforts have been directed towards consolidating and presenting relevant elements concerning the empirical data and AA's activities to the best extent possible. I base my account on the files and documentation related to AA in the private archives of PUSHPEACE (in Box Drive) and on the interactions I had with project members. I ultimately accessed the data through a protected cloud server provided by Analyse og Tal in a comma-separated value (CSV) file.

In this thesis, I acknowledge the previous work with this dataset made in *PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook (2022)*, while disclosing only some of its preliminary findings here. The following analyses made in this investigation are the result of my own interpretation and management of the mentioned data source and its related documentation, elaborated by AA members; and, by no means represent the views within PUSHPEACE. In my approach here, I try to accentuate aspects of transparency and replicability.

4.2.2 Anonymization and data aggregation

As Donnay (2017) and Theocharis & Jungherr (2017, 2021) note, extensive datasets – such as the one provided by AA – also introduce significant challenges and concerns in the process of social scientific inquiry. In essence, there are increasing questions on the informed consent of the individuals whose information is being used and, on the privacy and ownership of said data (Floridi & Taddeo, 2016; Samuel & Buchanan, 2020; Theocharis & Jungherr, 2017, 2021). The complexity increases when it comes to researching sociopolitical violent (on- and offline) contexts; – where, for example, the revelation of individual-level information on political stances or views can be an overt exposure to danger for those encompassed in an investigation (Donnay, 2017).

Acknowledging the ethical responsibilities, this study is grounded in the core tenets of data protection for individuals and groups (Donnay, 2017; Floridi & Taddeo, 2016). Consequently, it primarily seeks to prevent unintended consequences arising from the use of social media data as an approximation for offline public opinion and sentiment (Barrett-Maitland & Lynch, 2020) – although, I was not responsible for the initial data gathering. Therefore, my analyses are always reported at an aggregate level, where it is not possible to draw attention to any particular individual-group nor infer personal-group identification (Donnay, 2017; Floridi & Taddeo, 2016). Furthermore, it is worth noting that all users (represented by their Facebook profile names) were already anonymized by AA members in the available dataset (Set B). Overall, from this data collection, sensible variables such as those related to unique comment identifiers and unique post identifiers (see Appendix 1) - which could be used to track the actual location of specific comments within Facebook posts and thus, identify the authors and their context – are only used to cluster elements and operate aggregate-level metrics such as users' engagement (count of comments and reactions on posts) and type of attention (rate of antagonism on posts).

To recapitulate and draw a parallel to my study regarding the ethical approach, AA's research initiative had anonymous user comments – across posts of Facebook pages

– as the primary units of analysis. These comments underwent algorithmic labeling and then were examined over time to identify online patterns of antagonism and agonism related to the Colombian peace process (2016-2022). In this thesis, the units of analysis are the posts enacted by news media entities. Similarly, the annotated comments from users are considered on aggregate to examine the rates of engagement levels and the types of exchanges regarding the content of posts through time. Thus, the focus is solely on the aggregated-level measures and not in individual-level statements or opinions.

4.3 Fundamentals of the empirical analysis

Unit of analysis: A unique FB post published by an agent considered important to the Colombian context on its FB public fan page (see Set B). Each post primarily disseminates content concerning the Colombian peace process, as understood by AA. Hence, it enables user reactions – signaling different perceptions regarding the controversial war-to-peace transition. Having such FB user comments on aggregate (by amount and type), allows for a macro examination of the elements from posts that attract and shape user expressions.

Population: Universe of all unique FB posts derived from Set B.

Sample: The criteria is to encompass the yearly periods with the greatest production of FB user comments on the available FB posts published by news media entities (see Set D, Appendix 5 and Missing values and cleaning Set D).

Representativeness: Achieving representativeness initially becomes a consideration of the AA dataset's intrinsic characteristics (see The empirical material). The heterogeneity of users and pages, the range of all interactions, their circumstantial importance and their distribution through time, for example, basically depend on the selection of the agents running FB public fan pages that publish the FB posts and the isolation of the overall social media publications through keywords (see Limitations and Appendix 2).

4.3.1 Binary logistic regression

A binary logistic regression is used to determine the direct probabilities of a dichotomous categorization problem, between 0 and 1, through a logistic function (fitted curved line or hyperplane to the data, depending on the number of explanatory variables) (Harrell, Jr., 2015; Molnar, 2022, 2023). As an extension of linear regression (from generalized linear models), it offers the possibility of a 'meaningful threshold' to differentiate two final (categorical) classes (Molnar, 2023). Here, the aim is to predict the prevalence of the type of user disagreement on a FB post published by a news media entity (outcome variable or target), as either prevalent antagonism (Y= 1) or antagonism not prevalent (Y= 0). Ultimately, this is done by exploring the interactions of certain extracted features of FB posts – which, in turn, can contribute to influence the type of user textual disagreement to varying extents.

Developed by Cox, Walker and Duncan (1950-1960) as a statistical model, binary logistic regression is framed here as a model class within the families of supervised machine learning (Harrell, Jr., 2015; Molnar, 2022, 2023). According to Molnar (2022: 62), contrary to a (formal) statistical model where certain assumptions guide the modeling process of 'nature's unknown true mechanisms' for a given phenomenon, a supervised learning approach 'mimics the outputs of nature' without necessarily offering a 'reasonable substitute for nature's mechanism'. However, both (modeling mindsets) can lead to either helpful or unproductive models (ibid). Regarding my thesis, this offers a practical and exploratory strategy for the framework of agonistic peace on social media as posited in my aim and objective.

Following Molnar's (2022) notations and comments, the formula for the probability for Y=1 is,

$$\pi = P(Y = 1|X) = \frac{1}{1 + exp(-(\beta_0 + \beta_0 X_1 + ... + \beta_n X_n))},$$

where π represents the probability of an antagonistic comment occurring based on predictors (numerical or categorical covariates) $X_1...X_n$, and their corresponding coefficients or parameters $\beta_0, \beta_1...\beta_n$ (ibid). Yet, in understanding these coefficients in relation to (actual) probabilities, the logarithm of odds (logits or log odds) is employed to have an anticipated number of Y = 1 for every Y = 0, as follows (ibid),

Odds =
$$\frac{\pi}{1-\pi}$$
 = $exp(\beta_0 + \beta_0 X_1 + ... + \beta_n X_n)$,

Log odds =
$$\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_0 X_1 + \dots + \beta_n X_n$$
,

In this way, the odds (which could be derived from counts too) of having a FB post with prevalent antagonism and those of having a post where antagonism is not prevalent, can be compared in a symmetrical scale (Molnar, 2023). Finally, the odds ratio (OR) and the log odds ratio or log(OR) are calculated to compare the odds of prevalent antagonism happening with those occurrences where antagonism was not prevalent; also with a symmetrical interpretation log scale providing the effect size (strength) and direction in the association between the covariates and target variable (Bland & Altman, 2000; Molnar, 2023),

$$Odds ratio = \frac{odds_1}{odds_2},$$

$$Log odds ratio = \frac{\ln(Odds_1)}{\ln(Odds_2)}$$

According to Molnar (2022, 2023), with these expressions, the interpretation of the binary logistic regression is that when a continuous predictor X_n increases by one unit, there is a 'multiplicative change' in the OR by a factor $exp(\beta_n)$ (ibid). For a log(OR), this change represents an increment 'by the value of the corresponding weight' (ibid). When the covariate is categorical, the interpretation entails comparing several categories to a preselected reference category (ibid). In order to determine the statistical significance and the confidence intervals of log(OR), the Wald test is used, calculating the standard deviations of log(OR) from 0 (ibid).

To assess the performance of my binary regression model more robustly, I adopt a principled approach and partition Set D (ibid). Leveraging the initial_split() function from the rsample package, I (randomly) allocate 75% of Set D for model training and the remaining 25% for testing – keeping a distinct section to examine the model's predicted performance on new information.

4.3.2 The model: variables, composition and metrics

For a predictive model for the predominance of the type FB user of disagreement, inspired by Nigam et al.'s analysis of Twitter data for the Colombian pre-2016 Plebiscite context in *Harvesting Social Signals to Inform Peace Processes Implementation and Monitoring* (2017) – I use a customized version of their model of social signals concerning public sentiment on the Colombian peace process (ibid: 343-344, 353). I use a selection of FB post features parting from their proposal of categories of predictive features to be extracted from Tweets to analyze public opinion: here, message-based (to capture structure, style and sentiment); several post attributes (as insights on time and engagement) and page-related elements (providing further context on publisher popularity and location). Ultimately, the idea is to explore which of those variables is the most dominant for estimating antagonistic user comments.

Category Variables		Description	Level of measurement				
Dependent or outcome (Y)	Prevalence of antagonism	From the aggregate count of antagonistic user comments pertaining to a specific post, a percentage is of this type of disagreement is calculated of all comments on the post. There is prevalence of antagonism if its share is > 50%, and is not prevalent if =< 50%.	Two categorical outcomes: "prevalent" (1) or "not prevalent" (0)				
Predictive feature extraction from FB posts							
X _(i) Message- based	 Tokens Verbs Adverbs Nouns Proper nouns Adposition 	No. of words in a post message No. of verbs in a post message No. of adverbs in a post message No. of nouns in a post message No. of proper nouns in a post message No. of adpositions in a post message	Numerical Numerical Numerical Numerical Numerical Numerical				

	7	Determiners	No of determiners in a past massage	Numerical	
	7.		No. of determiners in a post message		
8. Pronouns		Pronouns	No. of pronouns in a post message	Numerical	
9. Adjectives		Adjectives	No. of adjectives in a post message	Numerical	
10. War concept		War concept	Presence of an AA term related to war	Factor: TRUE, FALSE	
	11.	Sentiment	Prevalence of no. positive or negative words	Factor: Mostly negative,	
				Mostly positive, Equal	
	1.	"year_post"	Year of creation of post	Factor: 2020, 2021, 2022	
				Numerical	
	2.	"total_comments"	Aggregate count of comments on a post	Numerical	
X _(ii)	3.	"total_like"	Aggregate count of likes on a post	Numerical	
Post	4.	"total_shares"	Aggregate count of shares on a post	Numerical	
attributes 5. "total		"total_love"	Aggregate count of love on a post	Numerical	
	6.	"total_angry"	Aggregate count of angry on a post	Numerical	
	7.	"total_sad"	Aggregate count of sad on a post	Numerical	
	8.	"total_haha"	Aggregate count of haha on a post	Numerical	
	9.	"total_wow"	Aggregate count of wow on a post		
X _(iii)	1.	"page_like_count"	Count of likes each page has (popularity)	Numerical	
Page-	2.	"location_json"	Location reported of a page	Factor: Bogotá, Rest	
related					

4.4 Data preprocessing

All the computer programming necessary for data preparation, analysis and presentation was done using the R software v. 4.3.1 with the RStudio integrated development environment (IDE) v. 2023.9.1.494. Here, the steps are based on the context of the (iterative) epicycles of data analysis, proposed in *The Art of Data Science* (2016) by Peng & Matsui. The (commented) coding script can be found in https://github.com/FelipeVillota/t ssda.

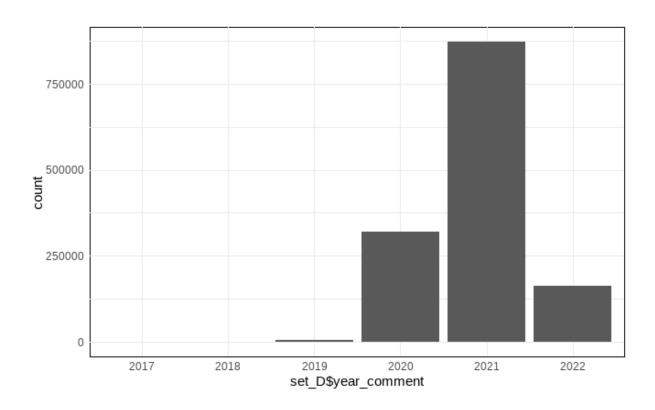
4.6.1 Variable adjustments

I assigned the correct data types in R for the target or outcome variable "label_antagonism" and the different categorical and numerical predictors. The adjustments occurred at different phases (i.e., on the distinct Sets). In general aspects, no original variables were relabeled. Yet, some factors required manual recoding of their levels due to an original inconsistent formatting. This was the case of location_json, in which I extracted the cities as levels when available (some were identified using postal codes and further geographic coordinates, others were left as unknown). Ultimately, the final values or categories (those of the clean Set D) are seen in Appendix 1. Other categorical and numerical variables of interest were simply created as aggregate measures or counts from FB posts associated attributes to become predictors (see The model).

4.6.2 Missing values and cleaning Set D

After checking for missing values in all sets, once Set D was created the variable "comment_count" was dropped (no observations affected). After factor recoding, there were blank spaces in the levels of "location_json" (208,437) and "status_type" (1,057,100), but were not removed since they represented significant information. The former was finally used as a two-level factor harboring cities: "Bogotá" (840,498) and "Rest" (520,793, including various local and international cities along with unreported places). The latter, "status type", was not used in the model.

However, according to my sample criteria, I only keep comments made on 2020, 2021 and 2022, the rest are removed (4,496) from Set D (see Figure 5). Right after this, all unique posts with blank spaces in "post_message" (153) were dropped, further eliminating 6,820 observations. Finally, with this selection some of the comments were made on posts published in 2019 – which initially is not problematic, but since they were 11 comments made in 9 posts (< 0.001% of the total observations in the remaining Set D), those comments were removed too. The clean Set D ended up with N= 1,349,964 comments across 15,509 unique posts by 63 different pages.



	2016	2017	2018	2019	2020	2021	2022	Total
Comments	0	2	138	4,356	321,063	872,363	163,369	1,361,291
Posts	0	1	11	472	3,484	9,244	2,934	16,146

Figure 5. Cumulative of user comments and unique FB posts made per year in Set D (yet to be cleaned).

4.6.4 Natural language processing

In extracting message-based features for the predictive model, I used the R package udpipe v.0.8.11 (Straka & Straková, 2017; Wijffels et al., 2023). So, I approached the data within "post_message" as a sequence of words in Spanish (Silge & Robinson, 2017) – to manipulate them through udpipe capabilities (by downloading a pre-trained model in the concerning language). Primarily, I established cleaning parameters within the text processing pipeline – including lowercase conversion, removal of punctuation, numbers, special characters or symbols, and extra spaces (Benoit, 2020; Silge & Robinson, 2017). No stop words were removed, since my aim was to capture a comprehensive range of linguistic features from each "post_message". Then, I parsed the textual data into single-word tokens or unigrams. With this, I ran a linguistic analysis by producing annotations for tokens and part-of-speech (POS) tags – i.e., low-level natural language processing tasks (Nadkarni et al., 2011). Then, I calculated for each unique "id_post" and "post_message" the total counts of resulting elements like verbs, adverbs, nouns, proper nouns, among others.

Taking advantage of this process, I created the "war_concept" variable; aiming to identify specific textual elements that could possibly emphasize conflictive social relations and violence or war within the dataset. So, with this logic, a simple selection of terms was extracted from AA's composite search key (see Appendix 2), based on the literature review made on war and peace journalism. Each sentence in clean Set D was examined against this selection using an R function that scanned for matches, determining whether or not any of these terms were present in a post message. Finally, a binary categorization was applied, denoting the presence (TRUE) or absence (FALSE) of these concepts in each sentence. In other aspects, the "Sentiment" variable was derived by evaluating the emotional tone of each post message (as a clean word string or sentence) and categorizing them as predominantly positive, negative, or neutral (referring to which type of words in a post message were mostly used). The syuzhet package was utilized to extract the sentiment scores from the Spanish-language text in the dataset (Isasi, 2021; Jockers, 2023).

5 Results

5.1 Descriptive statistics and trends

After an extensive exploratory data analysis, for Sets B, C and D (before and after cleaning) there were extremely high spikes – daily creation volumes of both, unique posts and comments – concentrated in 2021 (see Appendix 8 and Figure 7) Suggesting that the composite search key, and not the selection of news media entities as a particular category of FB public fan pages nor isolating user disagreements, substantially affected the balance of the timeframe – ultimately, it was skewed towards the second waves of the *Estallido social* of 2021 (Appendix 6).

From the extant information in the clean Set D, there were 1,349,964 FB user comments expressing disagreements across 15,509 different FB posts from Jan. 2, 2020 (14:09:45 UTC) until Oct. 19, 2022 (00:26:13 UTC) (a span of 1022 days, including the last). On the days that actually had comments (935), the summary statistics are as follows: Min.= 1 (on various days), 1st Qu.= 92, Median= 240, Mean= 1,444, 3rd = 841, Max.= 66,227 (May 5, 2021). Regarding the type of user disagreement, overall, there were N = 842,852 (~62%) antagonistic comments and N = 507,112 (~37%) not antagonistic disagreements. In this sense, their respective daily maximums were 40,415 antagonistic comments and 25, 812 not antagonistic comments – both occurring on May 5, 2021. And, May 2021 was the moment with more user disagreements made (537,933 cases or ~40% of all comments in clean Set D) and where the majority of antagonism was concentrated (319,922 cases or ~38% of all antagonistic comments in clean Set D).

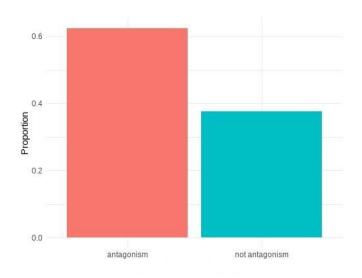


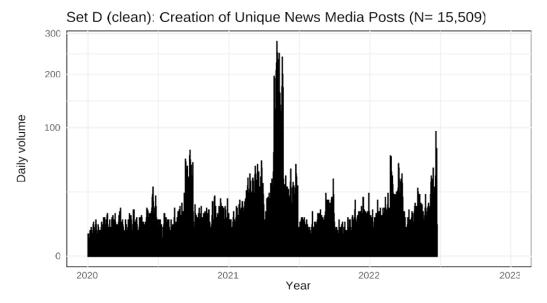
Figure 6. Final proportions of type of user disagreement in clean Set D N=1,349,964 user comments.

The context review and AA's previous results (internally presented on Dec. 15, 2022) already evidenced a conflictive environment (with prevalent hostility) in social media concerning the Colombian peace process in similar proportions. Precisely, in Set B, from Jan. 3, 2016 to June 24, 2022, out of ~6.2 million user comments, ~4 million (~64%) contained *disagreements*. Of this portion, ~65% were antagonistic (~2.6 mill.) and ~35% were not antagonistic (~1.4 mill.).

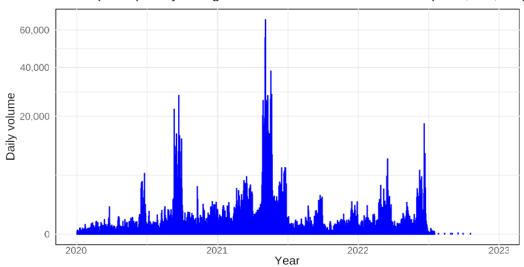
Concerning my target variable of the share of antagonism on posts ("prevalent" if the percentage of antagonistic comments >50% of total comments, if not, "not prevalent"), from the 15,509 unique posts, ~60% (9,254) were posts where antagonism in comments was prevalent; and ~40% (6,255) where it was not. Its summary is (in percentages), Min.= 0.00 (on 1,534 posts), 1st Qu.= 41.18, Median= 58.33, Mean= 55.48, 3rd = 72.73, Max.= 100.00 (on 1,596 posts). Regarding some of the model's numerical predictive covariates pertaining to clean Set D, there are noticeable differences when grouped and compared between posts with "prevalent" and "non-prevalent" antagonistic disagreements. For instance, when examining the total comments in a post (a metric of conversational engagement), it displays a Mean= 35 (SD= 125) for non-prevalent cases and a Mean= 122 (SD= 531) for prevalent ones; suggesting that, on average, longer user conversations on posts favor the pervasiveness of antagonistic comments. In terms of the popularity of the publisher of posts, the most liked page was Noticias Caracol (6,395,004 likes) and

the least Comité Nacional del Paro (1,748 likes). However, the latter is not (actually) a news media entity. Initially, it was a civilian cluster of groups and associations that were involved in the Estallido social of 2021 – mainly, integrated by Colombian labor unions (Sánchez, 2021). Yet, this label was kept to preserve AA's original system, and because it was the only "peculiar" case in the group of news media entities (see Appendix 5). On average, the more likes a page has, it is prone to receive more antagonistic comments across its posts.

As further illustrations, to shed light on the distribution nuances of categorical variables within both, "prevalent" and "not prevalent" categories – accentuating distinctive trends across these segments: within the geographical context of the publisher of posts, the majority of occurrences in the "not prevalent" category belong to the "Rest" classification (~62%) while ~38% associate with "Bogotá". In the "prevalent" segment, ~46% is attributed to "Rest" and ~54% to "Bogotá." So, on average, Bogotá-based news media entities tend to have a bigger share of antagonistic comments across their posts. The variable "war concept" shows ~82% as "FALSE" and ~18% as "TRUE" within the "not prevalent" domain; and, in the "prevalent" classification, "FALSE" accounts for ~76%, and "TRUE" constitutes ~24% of the cases, indicating that a war concept in a post message does not necessarily spur the dominance of antagonism in user responses. Lastly, examining the "Sentiment" variable, in the "not prevalent" category, ~23% align with "Equal," ~34% lean towards "More negative words," and the remaining ~43% express "More positive words." Correspondingly, the "prevalent" subset records ~24% under "Equal," ~33% associated with "More negative words," and ~43% embracing "More positive words". For more details on all covariates, see Appendix 9.



Set D (clean): Only Disagreements on News Media Posts (N= 1,349,964)



Set D (clean): By Type of Disagreement on News Media Posts (N= 1,349,964)

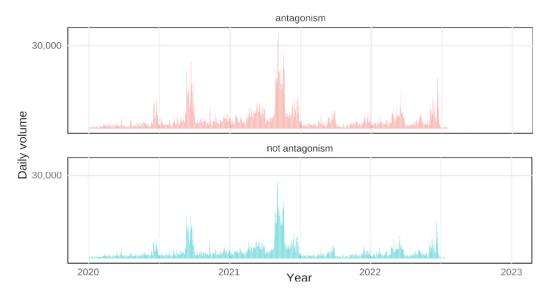


Figure 7. Histograms of clean Set D.

#	Page name	Page like	Unique posts	Discording	Antagonism
	rage name	count	published	comments	received (%)
1	Diario La Libertad	19,558	52	425	72.23
2	TeleRoble	13,790	68	1,467	71.98
3	Periódico Digital Erbol	354,250	478	55,680	70.31
4	Revista Semana	2,519,475	352	134,767	68.08
5	Colombia2020	76,484	1,830	35,673	67.28
59	Mi Diario Valledupar	3,050	2	3	33.33
60	Agencia de Noticias – Universidad	67,581	7	12	33.33
	Nacional de Colombia	07,301	,	12	33.33
61	RTV Televisión	7,619	20	107	31.77
62	La Nota Económica	3,840	2	14	14.28
63	Hacemos Memoria	13,188	18	29	10.34
					I

Table 1. Share of antagonism at page-level: "head and tail" of all the unique FB pages ran by news media entities in clean Set D (63); ranked by the share of antagonistic comments received from users in total; in the context of page popularity, activity and attention received on aggregate (from Jan. 2, 2020 to Oct. 19, 2022).



#	Attachment title of unique FB post	Page name	Creation time (UTC)	Antagonistic comments
1	Iván Duque: ""Diseñamos una estrategia contra el	El Espectador	2021-05-05 18:38:42	14,748
	vandalismo a nivel nacional""			
2	Paro Nacional: ""La ciudadanía debe agradecer	El Espectador	2021-05-04 22:52:31	14,634
	la labor de la Policía"", Iván Duque			
3	El senador Gustavo Petro se unió a las	Noticias Caracol	2021-05-19 18:32:32	9,066
	movilizaciones de este 19 de mayo			
4	El vandalismo en las manifestaciones del 21S en	Revista Semana	2020-09-22 18:59:50	8,951
	Colombia Vicky en Semana			
				I

Table 2. Top 4 unique FB posts (and their original attachment title) with the most cases of an antagonistic comments from users

5.1 Goodness of fit and effectiveness of the model

The clean Set D of N= 15,509 unique FB posts was split into a training set (randomly) containing a 75% (11,631 posts), and a testing set with the 25% (3,878 posts). The final binary logistic regression equation and coefficients containing 22 predictors, applied on the training set is (in R) is: Y (prevalence) = -0.218 + (0.009)*count token + (0.011)*count VERB + (-0.019)*count ADV 0.038)*count NOUN + (-0.056) *count PROPN + (0.000)*count ADP (0.008)*count DET + (-0.041)*count PRON + (0.004)*count ADJ(0.189)*year post2022 (0.003)*year post2021 ++(0.003)*total comments in post + (0.000)*total like + (0.000)*total shares + (0.000)*total angry (0.000)*total sad (-0.001)*total love (0.002)*total haha + (-0.002)*total wow + (0.000)*page like count (0.517)*location jsonBogotá +(0.437)*war conceptTRUE (0.005)*SentimentMore negative words $\} + (0.061)$ *SentimentMore positive words

Regarding the categorical covariates, [2020] is the reference level for the year of FB post, [Rest] for the predictor concerning the location of the FB fan page enabling the post, [False] regarding concept the inclusion of a war concept on a post message and [Equal] for the number of negative and positive words on a post (see the estimates in Appendix 10).

#	Model	Log likelihood ratio test (IIh)	-2 Log likelihood (-2llh)	AIC	ВІС
1	Null	-7,844	15,688	15,691	15,698
2	Message-based	-7,765	15,531	15,557	15,652
3	Attributes	-7,483	14,967	14,989	15,070
4	Page-related	-7,662	15,325	15,331	15,353
5	All predictors	-7,321	14,642	14,692	14,876

Table 3. Comparison of all models implemented, alongside their performance indices (rounding up)

As a summary, the null model was the baseline model (0 predictors). This serves as the reference point to predicting the prevalence status of antagonism on FB posts based on overall characteristics of the training set (Peng et al., 2002). Lower values of -2 Log Likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) suggest a better model fit in general, since they signal a balance between the quantity of parameters (covariates) and the goodness of fit to the data (Cavanaugh & Neath, 2019). So, it is expected to have highest values in those indexes here and a McFadden $R^2 = 0$. The latter being a commonly used test for assessing logistic regression models, indicating in this case an inadequate (null) fit for explaining the variance of prevalence status (Ehrman & Kline, 2022; Peng et al., 2002). Evidently, there is need for entering predictive variables into the baseline reference; so, I introduced exploratively, the following predictive input categories: the message-based model, which incorporated predictors (11) based on a FB post message content. The attributes model, which was made focused on extracted post attributes regarding engagement and a basic temporal context (9). Later, the pagerelated model explored predictors related to the FB public fan page "behind" the FB posts (2). And, the last model, utilized all available predictors (22).

Among the models tested, the comparative analysis showed that the "All Predictors" model demonstrated significantly the best fit from the other models (Chi^2 = 1,046.812, df = 24, p < 0.001; AIC= 14,692; llh= -7,321; -2 llh = 14,642). Meaning that the final model effectively captured more portion of the variance inherent in the training data through all the predictors than other models (Peng et al., 2002). However, the final model itself is not a very good fit for the data, because -2llh remains a large value (when it is ideal to be closer to 0, a perfect fit), and the McFadden $R^2 \sim 7\%$ (and R^2 Tjur $\sim 8.6\%$), suggesting a very small proportion of variance explained by the model (ibid).

In terms of prediction accuracy, it is important to restate that the database (entire clean Set D) revealed a \sim 60% of FB posts with prevalent antagonism in user comments, while in a \sim 40% there was prevalence of disagreements without antagonism (analogous proportions are reflected in the training and test sets). The overall prediction accuracy of the final model on the testing set (new data) was of 64% (2,489 correctly classified cases out of 3,878 posts in the testing set). And, separating this predictive performance

by the levels of the target variable: for the case of post with prevalent antagonism the accuracy was of 80% (1,853 correct cases out of 2,317 posts with prevalent antagonism) and for the case of prevalence of non-antagonistic disagreements was of 40.7% (636 correct cases out of 1,561 posts with no prevalence of antagonism). The overall area under the curve (AUC) measure (ranging from 0 to 1) for the final model is 0.68, meaning a moderate ability in distinguishing between the two classes.

5.2 Importance of predictors

The importance of a predictor can be determined by comparing its unique contribution to the variance when considered alone versus in combination with other predictors (seen in the coefficient, i.e. the shift in the log odds when the covariate increases by 1 unit) (Filho, 2023; RITHP, 2023). Key predictors demonstrating statistical significance and confidence intervals different from zero include: "count token", "count NOUN", "count PROPN", "count PRON", "year post [2022]", "total comments in post", "total love", "total sad", "total haha", "total wow", "page like", "location_json [Bogotá]", and "war concept [TRUE]". In this context, "location_json [Bogotá]" is the most important predictor.

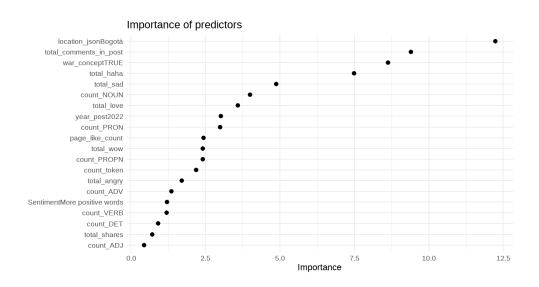


Figure 8. Variable importance plot (normalized scores from vip() function).

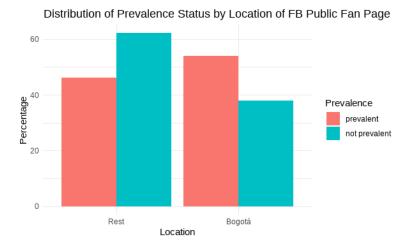


Figure 9. Plot displaying the distribution (in percentages) of prevalence status across the most important predictor.

6 Discussion

Succinctly, the small explanatory power of the model reflects the substantial uncertainty associated with individual and aggregate user behaviors in the expansive realm of social media (in this case in determining mechanisms that explain the prevalence status of user antagonism on FB posts published by news media entities) (Nau et al., 2022). The results call for evident further (model) refinement. Precisely, they underscore the necessity to revisit variable selection and their constitution (due to the model underfitting the data), scrutinize possible collinearity issues, and include additional contextual predictors, e.g. (Peng et al., 2002). Despite these aspects, the model showed a moderate ability in differentiating between posts without prevalent antagonism and posts with prevalent antagonism. And, in the end, it performed with best accuracy in identifying posts with prevalent antagonism (with a substantial difference in comparison with the other target level). So, there is a manifest margin for improvement regarding the imbalance in estimating non-antagonistic disagreements.

Interestingly, in this predictive context, the (statistically significant) covariate related to the location of the publisher of the posts, was the most important feature in the model. In the final estimates, holding other predictors constant, posts associated with Bogotá are linked with a 0.517 increase in the log odds of prevalence of user antagonism compared to the reference level of location (rest of cities in Colombia and abroad). This suggests a stronger association with the target variable and a higher likelihood of encountering prevalent user antagonism on posts enabled by news media entities based in Bogotá, compared to those based in the rest of the locations.

6.1 Limitations

While it is difficult to account for all limitations, I focus on the interpretation of internal validity and frame it according to the concerns brought up by the convergence of computational and conventional approaches to social science and subsequent academic controversy – particularly, when dealing with the use of social media data for research (Schroeder, 2020; Sen et al., 2021; Stier et al., 2020; Theocharis & Jungherr, 2017, 2021).

As a backdrop, Ledford (2020: 329) has characterized two groups regarding this wide-ranging debate, 'computational-oriented investigators and traditionalists'. In this sense, critics of computational social science contend that the counterpart's corpus of work is frequently descriptive (looking for patterns-trends and not for causal phenomena) (Ledford, 2020; Theocharis & Jungherr, 2017, 2021). Additionally, that its researchers are not thoroughly acquainted with social-science theories (explanations or models on human behavior) and previous body of knowledge (ibid). Moreover, there are acute claims made against the representativeness of the data, its hygiene (how it is gathered and manipulated), and the generalizability of the studies (which significantly vary in scale, topics, platforms, e.g.) (ibid). Nonetheless, on their part, computational researchers assert the 'cloudiness and low testability' of several social-science theories (poorly defined) and highlight the lack of (computer-technical) training of some traditional social-science practitioners engaging with computational approaches (Leford: 330).

6.1.1 Internal validity

Attending to some considerations on the complexity of the empirical material used and the results of the analysis, it is worth restating that the criteria structured by AA to isolate the discussion on the Colombian peace process, while efficient, constrained the breadth and diversity of information available for analysis. Considering the multifaceted nature of the [peace] process — and without discounting the relevance of the chosen search key—incorporating complementary elements like hashtags or broader keyword variations could have enriched the

dataset, potentially offering a more comprehensive understanding of the dynamics of the 2021 *Estallido Social* and beyond (e.g., capturing more user activity since 2016 and other events) (for example, see Nigam et al., 2017).

With this noted, the data retrospectively extracted by AA from FB, does not necessarily represent a comprehensive cross-section of the Colombian population – despite being the most popular social networking site in the country (by number of users) (OOSGA, 2023). It does not even guarantee that all users are Colombian citizens or residents – not even human (an undetermined percentage might be bots) (Samuel & Buchanan, 2020). The sample of users (commenting on posts enacted by news media entities) in AA's dataset might also be biased in terms of certain demographic traits such as such as age, gender, income, education level, digital literacy, occupation, ethnicity, marital status, or geographic location. Hence, the composition and distribution of different groups might lead to skewed or incomplete findings regarding public opinion and sentiment of the overall (national) population – i.e. through 'under-representation or over-representation of societal expressions' (Salem, 2017: 12). Furthermore, regarding their activity, Facebook users might have published inconsistent personal information or could be emitting just "noise" through their comments on posts (unrelated or repeating expressions). In this context, it is worth noting that social media accounts also fluctuate over time – as some are closed and some are created constantly anew, and other individuals might have multiple accounts at the same time or multiple individuals can have one account, e.g. (Sen et al., 2021; Stier et al., 2020). Within this social media realm, information related to geolocation (can be made public or private, or made visible with "second" intentions too). Finally, regarding the management of the FB public fan pages, it is not clear if there are (undisclosed) mechanisms that act as an "audience selector" and to what extent it functions, if applied.

The debate of using social media data for social science investigations has continued in recent years – establishing key questions, for instance, on the treatment of social media observations, the necessity or not of human participant consent, the ownership of the data and, the actual operations of social media platforms and the gratifications user get from their usage (Barrett-Maitland & Lynch, 2020; Jungherr, 2017; Nau et al., 2022; Samuel & Buchanan, 2020; Theocharis & Jungherr, 2021;

Zeller, 2022). Despite these inquiries that mold the meaning of social media research (and online-offline connections), the scholar avowal (in the emergent field of computational social science) is to assure transparency and enhance replicability in a relevant study context (Jungherr, 2017; Nigam et al., 2017; Palakodety et al., 2020). Following this, external data validation of the research should be pursued (e.g., with cross-platform studies and comparisons with traditional surveys and polls) (ibid).

7 Conclusion

As final a reflection, the state of public opinion and sentiment during the period comprehended in the study (2020-2022), was traversed and constituted by a widespread unrest of significant edges of Colombian civil society (Bernal, 2021; Gaviria, 2023; Uprimny, 2022). As a pertinent reading, in his book *La explosión controlada* (2023: 26), A. Gaviria asserts that a prevalent social indignation was translated into a denominated an 'unstoppable demand' (for social, political and economic change). Within the scope of this study, in Facebook, some of those (controversial) expressions – framed within the Colombian peace process – frequently took an antagonistic form. Also, the heightened prominence of Bogotá as a predictive contributor in the model – as the location of news media post publishers – underscores a relationship between critical news propagation and a possible center-periphery communication style or connotation in covering the discussions about the peace process, within the virtual sphere.

In general terms, this (public sentiment) might be understood by an abrupt interruption – caused by the COVID-19 pandemic and its effects – to the dominant "optimistic" narrative and the factual socioeconomic growth Colombia had experienced during the last two decades (ibid). The author posits that this progress, was then revealed as frail and relative – specially, concerning the development and living of conditions of middle urban classes and the incompleteness of the Colombian welfare state (ibid). The health emergency caused higher rates of impoverishment (the country had a strict regime on quarantines), which in turn, caused more exposures and (death) risk to the pandemic (e.g., people looked more for unprotected and informal jobs for a livelihood) – ultimately, a cycle that detonated unprecedented episodes of civil unrest (mobilizations, protests, vandalic acts met with official repression and partial accords) particularly acute in 2021 (Gaviria, 2023; Uprimny, 2022).

In this sense, Uprimny (2022) sketched out the multifactorial nature of the causes of the Estallido. Based on J. Elster, Gaviria (2023) schematizes this as a superposition of structural, mid-term and supervening elements (e.g., a mixture of long-dated inequities,

the Santos-Uribe political divide, the predominance of hostility as the style for political communication along with new channels, low trust in traditional institutions, the perception of corruption of elites and political agents, the "slow" implementation of the RPA, the pandemic and a proposed tax reform) that gave way to the phenomenon. However, the complexity of this, both authors ultimately emphasize the centrality of the fluctuations of public opinion in the unfolding of sociopolitical events (Gaviria, 2023; Nigam et al., 2017; Uprimny, 2022). And that, when the State's response to social demands is made with abuse (e.g. with violence or without serious commitment), the public opinion only gets more exacerbated and hostile (ibid).

According to Gaviria (2023: 27), the waves of the Estallido social (2019-2021) occurred under Iván Duque's (socially perceived) 'illegitimate and disconnected' tenure. And, against this backdrop, Gustavo Petro claimed victory in the May-June 2022 Presidential elections by a slender margin, championing a message that challenged traditional elites, social groups and institutions (ibid). Therefore, both phenomena – Petro's presidency and the Estallido – share the same seminal (and volatile) sentiment of popular indignation and expectation, which should be kept track of concerning the development of the RPA (ibid).

References

- Acosta, L. J. (2016, September 9). El "Sí" lograría amplio respaldo en plebiscito por la paz en Colombia: Encuesta [Review of *El "Sí" lograría amplio respaldo en plebiscito por la paz en Colombia: Encuesta*, by G. Donoso]. *Reuters*. https://www.reuters.com/article/colombia-paz-plebiscito-idLTAKCN11F1TO
- Acuña Rodriguez, O. Y. (2014). Bandolerismo politico en Boyaca (Colombia), 1930-1953/Political brigandage in Boyaca (Colombia), 1930-1953. *Antropologia y Sociologia: Virajes*, 16(2), 229–254.
- Aggestam, K., Cristiano, F., & Strömbom, L. (2015). Towards agonistic peacebuilding? Exploring the antagonism–agonism nexus in the Middle East Peace process. *Third World Quarterly*, 36(9), 1736–1753. https://doi.org/10.1080/01436597.2015.1044961
- Archila Neira, M. (1997). *El frente nacional: Una historia de enemistad social*. https://repositorio.unal.edu.co/handle/unal/30363
- Aya Smitmans, M. T. (2017). El Proceso de Paz en Colombia: Dos pasos adelante, un paso atrás. *Estudios Internacionales (Santiago)*, 49(187), 163–179. https://doi.org/10.5354/0719-3769.2017.47034
- Banks, A., Calvo, E., Karol, D., & Telhami, S. (2021). #PolarizedFeeds: Three Experiments on Polarization, Framing, and Social Media. *The International Journal of Press/Politics*, 26(3), 609–634. https://doi.org/10.1177/1940161220940964
- Barcelona Centre for International Affairs CIDOB. (2014). *Conflicto en Colombia:*Antecedentes históricos y actores. CIDOB.

 https://www.cidob.org/publicaciones/documentacion/dossiers/dossier_proceso_de_paz
 _en_colombia/dossier_proceso_de_paz_en_colombia/conflicto_en_colombia_anteced
 entes historicos y actores
- Barrett-Maitland, N., & Lynch, J. (2020). Social Media, Ethics and the Privacy Paradox. In *Security and Privacy From a Legal, Ethical, and Technical Perspective*. IntechOpen. https://doi.org/10.5772/intechopen.90906
- Basset, Y. (2017). Claves del rechazo del plebiscito para la paz en Colombia. *Estudios Políticos (Medellín)*, 52. https://doi.org/10.17533/udea.espo.n52a12
- BBC Staff. (2016). Colombia: Ganó el "No" en el plebiscito por los acuerdos de paz con las FARC. *BBC News Mundo*. https://www.bbc.com/mundo/noticias-america-latina-37537187
- Benoit, K. (2020). Text as Data: An Overview. In *The SAGE Handbook of Research Methods in Political Science and International Relations* (pp. 461–497). SAGE Publications Ltd. https://doi.org/10.4135/9781526486387.n29
- Bernal, J. C. U. (2021). Teorización de la conflictividad híbrida en el estallido social colombiano. *Revista de Las Fuerzas Armadas*, 258, 63–74.
- Bland, J. M., & Altman, D. G. (2000). The odds ratio. *BMJ*: *British Medical Journal*, 320(7247), 1468.
- Borda, S. (2018, May). *Presidenciales en Colombia: ¿polarización o deterioro de la conversación política?* Nueva Sociedad | Democracia y Política En América Latina. https://www.nuso.org/articulo/presidenciales-en-colombia-polarizacion-o-deterioro-de-la-conversacion-politica/

- Botero, S. (2017). El plebiscito y los desafíos políticos de consolidar la paz negociada en Colombia. *Revista de Ciencia Política (Santiago)*, *37*(2), 369–388. https://doi.org/10.4067/s0718-090x2017000200369
- Breuer, J., Kmetty, Z., Haim, M., & Stier, S. (2022). User-centric approaches for collecting Facebook data in the 'post-API age': Experiences from two studies and recommendations for future research. *Information, Communication & Society*, 1–20. https://doi.org/10.1080/1369118X.2022.2097015
- Caballero, A. (2016). *Historia de Colombia y sus oligarquías (1498—2017)* (M. A. Pautassi, T. Méndez, C. Cuervo, & N. Pernett, Eds.). Ministerio de Cultura Biblioteca Nacional de Colombia. https://bibliotecanacional.gov.co/es-co/proyectos-digitales/historia-de-colombia/libro/index.html
- Cairo, H., & Ríos, J. (2019). Las élites políticas y la paz territorial en Colombia: Un análisis de discurso en torno al Acuerdo de Paz. *Revista Española de Ciencia Política*, 50, 91–113. https://doi.org/10.21308/recp.50.04
- Cappella, J. N., & Jamieson, K. H. (1997). Spiral of Cynicism: The Press and the Public Good. Oxford University Press.
- Cárdenas, J. D. (2021). El plebiscito del 2016 en Colombia: Framing, resonancia cultural y comunicación política en los contenidos visuales de Facebook. *Ciencia Política*, 16(31), 235–273. https://doi.org/10.15446/cp.v16n31.88278
- Castro Riaño, L. C. (2020). La protesta social en América Latina: Una aproximación a su fisonomía a propósito de los estallidos sociales de 2019. *RUMBOS TS Un Espacio Crítico para la Reflexión en Ciencias Sociales*, 23, 159–184. https://doi.org/10.51188/rrts.num23.418
- Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Computational Statistics*, 11(3), e1460. https://doi.org/10.1002/wics.1460
- Chadwick, A. (2017). *The Hybrid Media System: Politics and Power*. Oxford University Press.
- CINEP/PPP-CERAC. (2022). Duodécimo Informe de verificación de la implementación del Acuerdo Final de Paz en Colombia. Centro de Investigación y Educación Popular / Programa por la Paz | Centro de Recursos para el Análisis de Conflictos. https://www.cinep.org.co/informes-generales/
- Comisión de la Verdad. (2022). Hallazgos y Recomendaciones de la Comisión de la Verdad de Colombia. In *Informe Final de la Comisión para el Esclarecimiento de la Verdad, la Convivencia y la No Repetición* (1st ed., 1–24, p. 792). Comisión de la Verdad, Sistema Integral para la Paz. https://www.comisiondelaverdad.co/sites/default/files/descargables/2022-08/FINAL%20CEV HALLAZGOS DIGITAL 2022.pdf
- Cook, T. E. (2012). Governing With the News, Second Edition: The News Media as a Political Institution. University of Chicago Press.
- Cosoy, N. (2016). ¿En qué se diferencia el nuevo acuerdo de paz entre el gobierno de Colombia y las FARC del que fue rechazado en el plebiscito? *BBC News Mundo*. https://www.bbc.com/mundo/noticias-america-latina-37965382
- Cubides-Cárdenas, J., & Vargas Parrasi, L. G. (2020). Amnistías e indultos en Colombia, 1953-2017. Su jurisprudencia a la luz de las sentencias de la Corte Interamericana de Derechos Humanos (pp. 43–81). https://doi.org/10.21830/9789585287860.02
- Dabène, O., & Le Cour Grandmaison, R. (2022, January 6). Colombia: 5 años después de los Acuerdos de Paz. *El Grand Continent*. https://legrandcontinent.eu/es/2022/01/06/colombia-5-anos-despues-de-los-acuerdos-de-paz/

- Dahlgren, P. (2009). *Media and Political Engagement* (W. L. Bennett & R. M. Entman, Eds.). Cambridge University Press.
- D'Angelo, P. (2019). Framing Theory and Journalism. In T. P. Vos & F. Hanusch (Eds.), *The International Encyclopedia of Journalism Studies* (pp. 1–10). Wiley.
- Dávalos, E., Morales, L. F., Holmes, J. S., & Dávalos, L. M. (2018). Opposition Support and the Experience of Violence Explain Colombian Peace Referendum Results. *Journal of Politics in Latin America*, 10(2), 99–122. https://doi.org/10.1177/1866802X1801000204
- Donnay, K. (2017). Big Data for Monitoring Political Instability. *International Development Policy* | *Revue Internationale de Politique de Développement*, 8.1, Article 8.1. https://doi.org/10.4000/poldev.2468
- Dutceac Segesten, A., Bossetta, M., Holmberg, N., & Niehorster, D. (2022). The cueing power of comments on social media: How disagreement in Facebook comments affects user engagement with news. *Information, Communication & Society*, 25(8), 1115–1134. https://doi.org/10.1080/1369118X.2020.1850836
- Ehrman, R. R., & Kline, J. A. (2022). Primer on Logistic Regression for Emergency Care Researchers. *The Journal of Emergency Medicine*, *63*(5), 683–691. https://doi.org/10.1016/j.jemermed.2022.09.032
- El Espectador. (2016). *La cuestionable estrategia de campaña del No* [Text]. ELESPECTADOR.COM. https://www.elespectador.com/politica/la-cuestionable-estrategia-de-campana-del-no-article-658862/
- Ferro, J. G., & Ramón, G. U. (2002). El orden de la guerra: Las FARC-EP, entre la organización y la política. Pontificia Universidad Javeriana.
- Ferry, S. (2012). Violentología: Un manual del conflicto colombiano (Primera edición). Icono.
- Filho, M. (2023, March 30). *How To Get Feature Importance In Logistic Regression*. https://forecastegy.com/posts/feature-importance-in-logistic-regression/
- Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360
- Galindo, J. (2017, October 6). ¿Por qué fallan tanto las encuestas en Colombia? *El País*. https://elpais.com/internacional/2017/10/06/colombia/1507261563_998951.html
- Galtung, J., & Fischer, D. (2013). *Johan Galtung: Pioneer of Peace Research* (Vol. 5). Springer. https://doi.org/10.1007/978-3-642-32481-9
- García-Perdomo, V., Harlow, S., & Brown, D. K. (2022). Framing the Colombian Peace Process: Between Peace and War Journalism. *Journalism Practice*, 1–24. https://doi.org/10.1080/17512786.2022.2062428
- Ghitis, F. (2021, October 28). A Mexican presidential candidate is getting an unexpected boost from Trump—And Putin. *Washington Post*. https://www.washingtonpost.com/news/democracy-post/wp/2018/01/11/a-mexican-presidential-candidate-is-getting-an-unexpected-boost-from-trump-and-putin/
- Gil de Zúñiga, H., Molyneux, L., & Zheng, P. (2014). Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships. *Journal of Communication*, 64(4), 612–634. https://doi.org/10.1111/jcom.12103
- Giraldo, K. (2023, August 29). *UBPD: La esperanza para miles de familias de Colombia*. https://www.radionacional.co/actualidad/ubpd-la-esperanza-para-miles-de-familias-que-mantienen-viva-la-memoria
- Gitlin, T. (2003). The Whole World Is Watching: Mass Media in the Making and Unmaking of the New Left. University of California Press.

- Gomez-Suarez, A., & Franz, T. (2020, September 20). *Protestas legítimas o vandalismo:* ¿Qué ocurre en Colombia? The Conversation. http://theconversation.com/protestas-legitimas-o-vandalismo-que-ocurre-en-colombia-146361
- González Arana, R., & Molinares Guerrero, I. (2010). LA VIOLENCIA EN COLOMBIA: UNA MIRADA PARTICULAR PARA SU COMPRENSIÓN. DE CÓMO PERCIBIMOS LA VIOLENCIA SOCIAL A GRAN ESCALA Y HACEMOS INVISIBLE LA VIOLENCIA NO MEDIÁTICA. *Investigación y Desarrollo*, 18(2), 346–369.
- Greene, T. (2023, November 17). *Juan Manuel Santos: I copied aspects of Northern Ireland peace agreement*. Al Jazeera. https://www.aljazeera.com/news/2023/11/17/colombias-santos-i-copied-aspects-of-northern-ireland-peace-agreement
- Gutiérrez, J. D. (2023). Imaginarios de guerra y paz en los memes de las campañas del sí y el no durante el plebiscito de 2016 en Colombia. *Maguaré*, *37*(1), Article 1. https://doi.org/10.15446/mag.v37n1.107563
- Gutiérrez Loaiza, A. (2012). Negociaciones de paz en Colombia, 1982-2009. Un estado del arte. *Estudios Políticos*, 40, 175–200.
- Harrell, Jr., F. E. (2015). Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer International Publishing. https://doi.org/10.1007/978-3-319-19425-7
- Harrison, J., & Pukallus, S. (2023). The civil norm building role of news journalism in post-civil war settings. *Journalism*, 24(1), 120–138. https://doi.org/10.1177/14648849211072947
- Hege, S. (2021, December 16). *Colombia: U.S. Aids Peace by Lifting 'Terrorist' Label From FARC*. United States Institute of Peace. https://www.usip.org/publications/2021/12/colombia-us-aids-peace-lifting-terrorist-label-farc
- Indepaz. (2021, November 24). 5 años del acuerdo de paz Balance en cifras de la violencia en los territorios. https://indepaz.org.co/5-anos-del-acuerdo-de-paz-balance-en-cifras-de-la-violencia-en-los-territorios/
- Isasi, J. (2021). Sentiment Analysis with "syuzhet" using R. *Programming Historian*. https://programminghistorian.org/en/lessons/sentiment-analysis-syuzhet
- Jamieson, K. H., & Kenski, K. (2017). Political Communication: Then, Now, and Beyond. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford Handbook of Political Communication* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199793471.013.77
- Jockers, M. (2023). *Introduction to the Syuzhet Package* [Computer software]. https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html
- Jungherr, A. (2017). Normalizing Digital Trace Data.
- Körppen, D., & Ropers, N. (2011). Addressing the Complex Dynamics of Conflict Transformation. In *The Non-Linearity of Peace Processes: Theory and Practice of Systemic Conflict Transformation*. Verlag Barbara Budrich.
- Lederach, J. P. (1996). *Preparing For Peace: Conflict Transformation Across Cultures*. Syracuse University Press.
- Lederach, J. P. (2005). *The moral imagination: The art and soul of building peace*. Oxford University Press.
- Lederach, J. P. (2019). Conflict Transformation. *Beyond Intractability*. https://www.kennethacha.com/wp-content/uploads/2020/01/Conflict-Transformation_-Beyond-Intractability.pdf
- Ledford, H. (2020). How Facebook, Twitter and other data troves are revolutionizing social science. *Nature*, *582*(7812), 328–330. https://doi.org/10.1038/d41586-020-01747-1

- Lehti, M. (2016). From Antagonism to Agonistic Peace: Rethinking Identities and Dialogic Transformation.
- Loaiza, M. V. (2022, August 6). La Colombia que deja Iván Duque: Desafíos, fracasos y logros. *CNN*. https://cnnespanol.cnn.com/2022/08/06/colombia-deja-ivan-duque-violencia-desafíos-logros-orix/
- Mack, A. (2012, August 10). Even Failed Peace Agreements Save Lives. *Political Violence at a Glance*. https://politicalviolenceataglance.org/2012/08/10/even-failed-peace-agreements-save-lives/
- Melo González, J. O. (2021). Ciudad y campo en Colombia hasta comienzos del siglo XX. De la utopía urbana a la ruralización, y a la urbanización acelerada. *CEPAL* | *Naciones Unidas*.
- Ministerio TIC. (n.d.). *Total de suscriptores de Internet*. Retrieved July 24, 2023, from https://colombiatic.mintic.gov.co/estadisticas/stats.php?&pres=content&jer=1&cod=&id=34#TTC
- Misión de Observación Electoral (MOE). (2017). Plebiscito de Refrendación de los Acuerdos de Paz. Informe de Observación.
- Molano, A. (2000). The Evolution Of The Farc: A Guerrilla Group's Long History. *NACLA Report on the Americas*, *34*(2), 23–31. https://doi.org/10.1080/10714839.2000.11722627
- Molnar, C. (2022). *Modeling Mindsets: The Many Cultures of Learning from Data*. Christoph Molnar c/o Mucbook Clubhouse.
- Molnar, C. (2023). *Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/
- Moreno Montalvo, G. (2019). *Breve historia de Colombia*. Diario La República. https://www.larepublica.co/analisis/gustavo-moreno-montalvo-2565659/breve-historia-de-colombia-2894731
- Mouffe, C. (2005). The Return of the Political. Verso.
- Mouffe, C. (2013). Agonistics: Thinking The World Politically. Verso Books.
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464
- Nau, C., Quan-Haase, A., & McCay-Peet, L. (2022). Defining Social Media and Asking Social Media Research Questions: How Well Does the Swiss Army Knife Metaphor Apply? In *The SAGE Handbook of Social Media Research Methods* (pp. 13–26). SAGE Publications Ltd. https://doi.org/10.4135/9781529782943
- Newman, E., & Richmond, O. (2006). Peace building and spoilers. *Conflict, Security & Development*, 6(1), 101–110. https://doi.org/10.1080/14678800600590728
- Nielsen, R. K., & Schrøder, K. C. (2014). The Relative Importance of Social Media for Accessing, Finding, and Engaging with News: An eight-country cross-media comparison. *Digital Journalism*, *2*(4), 472–489. https://doi.org/10.1080/21670811.2013.872420
- Nigam, A., Dambanemuya, H. K., Joshi, M., & Chawla, N. V. (2017). Harvesting Social Signals to Inform Peace Processes Implementation and Monitoring. *Big Data*, *5*(4), 337–355. https://doi.org/10.1089/big.2017.0055
- Nilsson, D., & Söderberg Kovacs, M. (2011). Revisiting an Elusive Concept: A Review of the Debate on Spoilers in Peace Processes 1. *International Studies Review*, 13(4), 606–626. https://doi.org/10.1111/j.1468-2486.2011.01080.x
- Noticias Caracol (Director). (2023, April 11). "Quiero aplicarlo, pero vale \$150 billones": Petro advierte falta de recursos para acuerdo de paz. https://www.youtube.com/watch?v=qD8gXuHBCC0

- Observatorio de Paz y Conflicto. (2016). Organizaciones Guerrilleras en Colombia desde la década de los sesenta. Universidad Nacional de Colombia | Sede Bogotá.
- Oficina del Alto Comisionado para la Paz | Presidencia de la República. (2018). La fase exploratoria y el camino hacia el Acuerdo General. Del 7 de agosto de 2010 al 17 de octubre de 2012. In *Biblioteca del Proceso de Paz con las FARC-EP: Vol. I* (Presidencia de la República).
- OOSGA. (2023). Social Media in Colombia—2023 Stats & Platform Trends. https://oosga.com/social-media/col/
- Özmen, Y. P., & Medina, N. (2020). Conflict and Violence in Colombia: The Breaking Points in the Peace Process with the FARC. *Turkish Journal of Political Science*, *3*(2), 113–135.
- Palacios, M. (2006). *Between legitimacy and violence: A history of Colombia, 1875–2002*. Duke University Press. https://scholar.google.com/citations?view_op=view_citation&hl=es&user=eIpSglkAA AAJ&citation for view=eIpSglkAAAAJ;qjMakFHDy7sC
- Palacios, M. (2014). VIOLENCIA POLÍTICA EN COLOMBIA 1958-2010. Revista de Economía Institucional, 16(30), 357–364.
- Palakodety, S., KhudaBukhsh, A. R., & Carbonell, J. G. (2020). *Hope Speech Detection: A Computational Analysis of the Voice of Peace* (arXiv:1909.12940). arXiv. http://arxiv.org/abs/1909.12940
- Pécaut, D., & González, L. (1997). Presente, pasado y futuro de la violencia en Colombia. *Desarrollo Económico*, 36(144), 891–930. https://doi.org/10.2307/3467131
- Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, 96(1), 3–14.
- Peng, R. D., & Matsui, E. (2016). The Art of Data Science. Leanpub, 162.
- Pinzón Flórez, A. (2020). Influencia de los medios de comunicación en el plebiscito por la paz de 2016 en Colombia. *Inciso*, 22(2), 293–315.
- Quan-Haase, A., & Sloan, L. (2022). Introduction. In *The SAGE Handbook of Social Media Research Methods* (1–0). https://doi.org/10.4135/9781529782943
- Quinn, L., & Gómez Vásquez, M. (2023). *Implementation of the Colombian Peace Accord Reaches its Sixth Year* (16). Kroc Institute for International Peace Studies, Keough School for Global Affairs, University of Notre Dame. https://doi.org/10.7274/6395w666821
- Ramírez, Y. (2022). *Tan solo el 30 % de los acuerdos de paz se ha completado: Instituto Kroc*. https://www.radionacional.co/actualidad/paz/acuerdos-de-paz-solo-se-hacompletado-el-30-instituto-kroc
- Registro Único de Víctimas (RUV). (2023, July 10). Registro Único de Víctimas (RUV). Unidad para las Víctimas. https://www.unidadvictimas.gov.co/es/registro-unico-devictimas-ruv/37394
- Ríos Hernández, I. N., Pérez-Marín, M., & Morillo Puente, S. (2018). La agenda setting en la red social Facebook: Campaña del plebiscito por la paz en Colombia. *Kepes*, *15*(17), 93–117. https://doi.org/10.17151/kepes.2018.15.17.5
- RITHP. (2023, January 1). Logistic Regression for Feature Selection: Selecting the Right Features for Your Model. *Medium*. https://medium.com/@rithpansanga/logistic-regression-for-feature-selection-selecting-the-right-features-for-your-model-410ca093c5e0
- Rivera, E. de J. V. (2007). Historia del paramilitarismo en Colombia. *História (São Paulo)*, *26*, 134–153. https://doi.org/10.1590/S0101-90742007000100012
- Rizzi, A. (2016, October 3). Muchos 'likes', pocos votos. *El País*. https://elpais.com/internacional/2016/10/03/actualidad/1475485483_032686.html

- Roldán, M. (2002). Blood and Fire: La Violencia in Antioquia, Colombia, 1946-1953. Duke University Press.
- Ruiz, J. I., Castro-Abril, P., López-López, W., Páez, D., Méndez, L., Castro-Molinares, S., Yadira-Cepeda, Z., Caicedo-Bucheli, M. A., Amaris, M. del C., Moncayo, J. E., Camelo-Mendoza, R., Orduz-Gualdron, F. S., Beltrán-Espitia, M., Mongui, Z. L., Domínguez, E., Alejo-Riveros, A., Pérez-Cervantes, L. E., Castillo, C. O., Alvarado-Pinzón, L., ... Arizabaleta, M. P. (2022). La Comisión de la Verdad en Colombia: Conocimiento, percepción, eficacia y emociones asociadas. Revista de Psicología (PUCP), 40(1), 119–154. https://doi.org/10.18800/psico.202201.005
- Salazar, G. (2005). *Mirada crítica al conflicto armado colombiano*. Memorias del Seminario-Taller Internacional de Contextualización sobre Desplazamiento Forzado y Refugio en Zonas de Frontera., Bogotá: Defensoría del Pueblo. http://www.acnur.org/t3/fileadmin/scripts/doc.php
- Salazar Mahecha, S. (2016). La conversación sobre el Plebiscito de Refrendación en las Redes Sociales. In *Medios de Comunicación y Plebiscito de Refrendación de los Acuerdos de Paz* (pp. 37–55). Misión de Observación Electoral (MOE).
- Salem, F. (2017). Social media and the internet of things. Towards data-driven policymaking in the Arab world: Potential, limits and concerns. *Arab Social Media Report*, 7. https://ora.ox.ac.uk/objects/uuid:fab5afd1-30b5-479c-8231-fbf6bf780017
- Samuel, G., & Buchanan, E. (2020). Guest Editorial: Ethical Issues in Social Media Research. *Journal of Empirical Research on Human Research Ethics*, 15(1–2), 3–11. https://doi.org/10.1177/1556264619901215
- Sánchez Torres, C. A. (2001). La decisión judicial sobre la muerte de Gaitán. *Estudios Socio-Jurídicos*, *3*(1), 60–92.
- Schroeder, R. (2020). Big data and cumulation in the social sciences. *Information, Communication & Society*, 23(11), 1593–1607. https://doi.org/10.1080/1369118X.2019.1594334
- Scudder, M. F., & White, S. K. (2023). The Agonistic Face. In M. F. Scudder & S. K. White (Eds.), *The Two Faces of Democracy: Decentering Agonism and Deliberation* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780197623886.003.0004
- Semana. (2021a, April 29). Paro nacional 2021: Millonarias pérdidas para el comercio tras actos de vandalismo en las principales ciudades de Colombia. Semana.com Últimas Noticias de Colombia y el Mundo.

 https://www.semana.com/economia/empresas/articulo/paro-nacional-2021-millonarias-perdidas-para-el-comercio-tras-actos-de-vandalismo-en-las-principales-ciudades-de-colombia/202113/
- Semana. (2021b, May 15). *Manifestantes humillan a la Policia en medio de las protestas*. Semana.com Últimas Noticias de Colombia y el Mundo. https://www.semana.com/nacion/articulo/manifestantes-humillan-a-la-policia-en-medio-de-las-protestas/202129/
- Sen, I., Floeck, F., Weller, K., Weiss, B., & Wagner, C. (2021). *TED-On: A Total Error Framework for Digital Traces of Human Behavior on Online Platforms* (arXiv:1907.08228). arXiv. https://doi.org/10.48550/arXiv.1907.08228
- Shinko, R. E. (2008). Agonistic Peace: A Postmodern Reading. *Millennium*, *36*(3), 473–491. https://doi.org/10.1177/03058298080360030501
- Silge, J., & Robinson, D. (2017). Text Mining with R: A Tidy Approach (1st edition). O'Reilly Media
- Stedman, S. J. (1997). Spoiler Problems in Peace Processes. *International Security*, 22(2), 5–53.

- Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field. *Social Science Computer Review*, *38*(5), 503–516. https://doi.org/10.1177/0894439319843669
- Straka, M., & Straková, J. (2017). Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. In J. Hajič & D. Zeman (Eds.), *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies* (pp. 88–99). Association for Computational Linguistics. https://doi.org/10.18653/v1/K17-3009
- Strömbom, L. (2014). Thick recognition: Advancing theory on identity change in intractable conflicts. *European Journal of International Relations*, 20(1), 168–191. https://doi.org/10.1177/1354066112439217
- Strömbom, L. (2020). Exploring analytical avenues for agonistic peace. *Journal of International Relations and Development*, 23(4), 1–23. https://doi.org/10.1057/s41268-019-00176-6
- Strömbom, L., & Bramsen, I. (2022). Agonistic peace: Advancing knowledge on institutional dynamics and relational transformation. *Third World Quarterly*, *43*(6), 1237–1250. https://doi.org/10.1080/01436597.2022.2054797
- Tabares Higuita, L. X. (2018). Análisis del discurso violento y de odio en dos grupos de Facebook contra la candidatura de Rodrigo Londoño 'Timochenko' a la presidencia de Colombia. *index.comunicación*, 8(3), 157–184.
- teleSUR. (2016, October 2). *Plebiscito consigue más de 2 millones de menciones en redes*. teleSURtv.net. https://www.telesurtv.net/news/Plebiscito-consigue-mas-de-2-millones-de-menciones-en-redes-20161002-0023.html
- Theocharis, Y., & Jungherr, A. (2017). The empiricist's challenge: Asking meaningful questions in political science in the age of big data. *Journal of Information Technology* & *Politics*, 14(2), 97–109. https://doi.org/10.1080/19331681.2017.1312187
- Theocharis, Y., & Jungherr, A. (2021). Computational Social Science and the Study of Political Communication. *Political Communication*, *38*(1–2), 1–22. https://doi.org/10.1080/10584609.2020.1833121
- Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My Friends, Editors, Algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, 7(4), 447–469. https://doi.org/10.1080/21670811.2018.1493936
- Tufekci, Z. (2017). Twitter and Tear Gas: The Power and Fragility of Networked Protest. Yale University Press.
- Uprimny, R. (2022). El «estallido social» colombiano: Reflexiones sobre protesta y derechos humanos en democracias débiles. *Deusto Journal of Human Rights*, *10*, Article 10. https://doi.org/10.18543/djhr.2625
- Uribe Sarmiento, J. J., & Rodríguez Cortes, A. B. (2022). Tensiones del acuerdo de paz: Entre la estatalidad y la cotidianidad del hogar. *Ciudad Paz-ando*, *15*(1), Article 1. https://doi.org/10.14483/242278X.18826
- Valcárcel Torres, J. M. V. (2008). BELIGERANCIA, TERRORISMO Y CONFLICTO ARMADO: NO ES UN JUEGO DE PALABRAS.
- Valenzuela, S., Piña, M., & Ramírez, J. (2017). Behavioral Effects of Framing on Social Media Users: How Conflict, Economic, Human Interest, and Morality Frames Drive News Sharing: Framing Effects on News Sharing. *Journal of Communication*, 67(5), 803–826. https://doi.org/10.1111/jcom.12325
- Verdad Abierta. (2017). La desmovilización: El proceso de paz (2003-2006). *VerdadAbierta.com*. https://verdadabierta.com/periodo4/
- Wang, Y., Li, Y., & Luo, J. (2016). Deciphering the 2016 U.S. Presidential Campaign in the Twitter Sphere: A Comparison of the Trumpists and Clintonists. *Proceedings of the*

- *International AAAI Conference on Web and Social Media*, 10(1), Article 1. https://doi.org/10.1609/icwsm.v10i1.14783
- Wijffels, J., BNOSAC, Linguistics, I. of F. and A., Physics, F. of M. and, Prague, C. U. in, Republic, C., Straka, M., & Straková, J. (2023). *udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the "UDPipe" "NLP" Toolkit* (0.8.11) [Computer software]. https://cran.r-project.org/web/packages/udpipe/index.html
- Wilding, D., Fray, P., Molitorisz, S., & McKewon, E. (2018). *The Impact of Digital Platforms on News and Journalistic Content*. University of Technology Sydney, NSW. https://www.accc.gov.au/system/files/ACCC+commissioned+report+-+The+impact+of+digital+platforms+on+news+and+journalistic+content,+Centre+for+Media+Transition+(2).pdf
- Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, 7(3), 203–220. https://doi.org/10.1177/1745691612442904
- Zeller, F. (2022). The Good, the Bad, and the How-to of Analysing Social Media Data. In *The SAGE Handbook of Social Media Research Methods* (1–0, pp. 363–381). SAGE Publications Ltd. https://doi.org/10.4135/9781529782943

Appendices

Appendix 1. Table of Facebook and AA variables

Here, there are all variables and relevant information (this includes the fields allowed by Facebook's API and the variables created by AA for the algorithmic annotation of comments on disagreement and antagonism). Transformations and levels of measurement are final and correspond to Set D.

No.	Label of variable	Item description	Initial Data type	Transformation in R	Values, ranges and levels
1	"X"	Generic identifier of a comment	Integer	-	ℤ⁺₀
2	"comment_message"	Content of a comment	Character	-	string
3	"created_time_comment"	Date of publication of a comment	Character	"POSIXct", "POSIXt" with format %Y-%m-%d %H:%M:%S" (UTC)	from "2016-01-04 00:28:11" to "2022-10-23 06:57:17 "
4	"love_comment"	Reaction (love) on a comment	Numeric	-	ℤ⁺₀
5	"like_comment"	Reaction (like) on a comment	Numeric	-	ℤ⁺₀
6	"angry_comment"	Reaction (anger) on a comment	Numeric	-	ℤ⁺₀
7	"sad_comment"	Reaction (sadness) on a comment	Numeric	-	\mathbb{Z}_0^+
8	"haha_comment"	Reaction (laughter) on a comment	Numeric	-	ℤ⁺₀
9	"wow_comment"	Reaction (surprise) on a comment	Numeric	-	ℤ⁺₀
10	"label_disagreement"	Identification of disagreement	Character	Factor	Two levels: "disagreement" or "not disagreement"
11	"label_antagonism"	Identification of antagonism	Character	Factor	Two levels= "antagonism", "not antagonism"
12	"score_0_disagreement"	Algorithmic score for "not disagreement"	Numeric	-	$(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$
13	"score_1_disagreement"	Algorithmic score for "disagreement"	Numeric	-	$(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$
14	"score_0_antagonism"	Algorithmic score for "not antagonism"	Numeric	-	$(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$
15	"score_1_antagonism"	Algorithmic score for "antagonism"	Numeric	-	$(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$
16	"disagreement"	Algorithmic decision on disagreement	Numeric	Factor	Two levels: "0" = "Disagreement", "1" = "Not Disagreement"
17	"antagonism"	Algorithmic decision on antagonism	Numeric	Factor	Two levels: "0" = "Antagonism", "1" = "Not Antagonism"

18	"post message"	Content of a post	Character	-	text string
19	"status_type"	Type of post	Character	Factor	Six categories: "added_video", "shared_story", "added_photos", "mobile_status_update", "wall_post", ""
20	"attachment_title"	Title of attachment of a post	Character	-	string
21	"attachment_description"	Description of attachment of a post	Character	-	string
22	"created_time_post"	Date of publication of a post	Character	"POSIXct", "POSIXt" with format %Y-%m-%d %H:%M:%S" (UTC)	from "2016-01-03 23:01:49" to "2022-06-24 13:47:14"
23	"comment_count"	Number of comments on a post	Numeric	-	ℤ⁺₀
24	"love_post"	Reaction (love) on a post	Numeric	-	ℤ⁺₀
25	"like_post"	Reaction (like) on a post	Numeric	-	ℤ+₀
26	"angry_post"	Reaction (anger) on a post	Numeric	-	ℤ+₀
27	"sad_post"	Reaction (sadness) on a post	Numeric	-	ℤ⁺₀
28	"haha_post"	Reaction (laughter) on a post	Numeric	-	ℤ+₀
29	"wow_post"	Reaction (surprise) on a post	Numeric	-	ℤ⁺₀
30	"shares"	Number of unique shares of a post	Numeric	-	ℤ⁺₀
31	"page_name"	Title of a page	Character	Factor	64 levels: see Appendix 5
32	"category"	Type of actor running the page	Character	Factor	19 levels: see 4.1.3 Set C
33	"page_like_count"	Number of likes of a page	Numeric	-	ℤ+₀
34	"website"	Link to a website of a page	Character	-	string
35	"location_json"	Geolocation of a page	Character	Factor	2 levels: "Bogotá", "Rest"
36	"id_comment"	Unique identifier of a comment	Character	-	string
37	"id_post"	Unique identifier of a post	Character	-	string
38	"id_page"	Unique identifier of a page	Numeric	Character	String

Here, there are the 62 terms in Spanish that make the composite search key*. This query was applied to Facebook posts in order to isolate the discussions on the Colombian peace process in AA's first Facebook data collection of ~77 million comments (from 2016 to 2022). Once applied, a subsequent dataset of ~6.2 million comments resulted.

1	comision do la vordad	22	lídar AND naz	43	protesta
	_comision_de_la_verdad_		_líder AND _paz_		
2	_comisión_de_la_verdad_	23	_lider AND _social	44	_restitucion_de_tierras_
3	_comite_del_paro_	24	_líder AND _social	45	_restitución_de_tierras_
4	_comité_del_paro_	25	_manifest AND _bogota_	46	_terroris
5	_conflictividad_	26	_manifest AND _bogotá_	47	acuerdodepaz_
6	_conflicto_ AND plebiscito_	27	_memoria AND _conflicto_	48	desmoviliza
7	_conflicto_armado_	28	_memoria AND _paz_	49	disidencia
8	_construccion_de_paz_	29	_museo_de_la_memoria_	50	eln
9	_construcción_de_paz_	30	_paro_	51	excombatiente
10	_cultivo AND _ilicito	31	_paz_ AND _acuerdo_	52	exguerriller
11	_cultivo AND _ilícito	32	_paz_ AND _dialogo_	53	farc
12	_disturbio	33	_paz_ AND _implementacion_	54	jep
13	_gran_conversacion_nacional_	34	_paz_ AND _implementación_	55	mina AND _antipersonal
14	_gran_conversación_nacional_	35	_paz_ AND _justicia_	56	paramilitar
15	_grupo AND _armado	36	_paz_ AND _negociacion_	57	pazconlegalidad_
16	_guerra_	37	_paz_ AND _negociación_	58	pdet_
17	_lider AND _asesin	38	_paz_ AND _proceso_	59	posconflicto_
18	_líder AND _asesin	39	_paz_ AND _reconciliacion_	60	reconciliación_
19	_lider AND _masacr	40	_paz_ AND _reconciliación_	61	reincorpora
20	_líder AND _masacr	41	_paz_ AND legalidad_	62	reintegra
21	_lider AND _paz_	42	_paz_ AND plebiscito_		

* In the file Search_Key_09.06.21 elaborated by the AA team within PUSHPEACE: Quantifying antagonistic/agonistic peace in Facebook (2022), the rules were: '(...) terms only in small caption; only one search term per cell (in Excel); use the shortest variation of a word (i.e., some stems in this case) or multiple versions of the word to include different conjugations and/or tenses; check if the search term can be found as part of other words and with a different/wrong meaning. As an example, _eln_ indicates a space before and after the search term. If a search term consists of more than one word, separate by the following rules: brutalidad_policial indicates that the words must be in that exact order; and the spelling brutalidad AND policial indicates that the words must be used in the same publication (Facebook post) but not necessarily in that order (...)'. It should be noted that for some words, the AA team also adjusted for accents (e.g., in restitución, comisión, negociación, etc.).

The highlighted terms were the ones considered to compose the "war_concept" variable.

List of the original 82 categories of the Facebook public fan pages targeted by AA.

No.	Page category (some in Danish/Norwegian)	Page category (translated to English)
1	Avis	Newspaper
2	Cause	Cause
3	Charity Organization	Charity Organization
4	College & University	College & University
5	Community	Community
6	Community Organization	Community Organization
7	Educational Research Center	Educational Research Center
8	Embedsmand	Official
9	Fællesskab	Community
10	Government building	Government building
11	Government Official	Government Official
12	Government official	Government official
13	Government organisation	Government organization
14	Government Organization	Government Organization
15	Ideell organisasjon	Non-profit organization
16	Ikke-statlig organisasjon (NGO)	Non-governmental organization (NGO)
17	Interesse	Interest
18	Interest	Interest
19	Kringkastings- og medieproduksjonsselskap	Broadcasting and media production company
20	Local Business	Local Business
21	Local business	Local business
22	Lokal bedrift	Local Holding
23	Lokal virksomhed	Local Business
24	Mad og drikkevarer	Food and Drinks
25	Mærkesag	Brand Case
26	Media/News Company	Media/News Company
27	Media/news company	Media/news company
28	Medicinsk forskningscenter	Medical Research Center
29	Medie-/nyhetsbedrift	Media/news company
30	Medier	Media
31	Medier/nyheder	Media/News
32	Museum	Museum
33	News & media website	News & media website
34	News & Media Website	News & Media Website
35	Newspaper	Newspaper
36	NGO (ikke-statslig organisation)	NGO (non-governmental organization)
37	Non-governmental organisation (NGO)	Non-governmental organization (NGO)
38	Non-Governmental Organization (NGO)	Non-Governmental Organization (NGO)
39	Nonprofit Organization	Nonprofit Organization
40	Nonprofitorganisation	Nonprofit organization
41	Nyheds- og mediewebsite	News and media website
42	Offentlig funksjonær	Public servant

43	Offentlig person	Public person
44	Offentlig service	Public Service
45	Offentlige, kommunale og statlige tjenester	Public, municipal and state services
46	Organisation	Organization
47	Personlig blog	Personal blog
48	Political Candidate	Political Candidate
49	Political organisation	Political organization
50	Political Organization	Political Organization
51	Political party	Political party
52	Political Party	Political Party
53	Politician	Politician
54	Politiker	Politician
55	Politisk organisation	Political organization
56	Politisk parti	Political Party
57	Public & Government Service	Public & Government Service
58	Public Figure	Public Figure
59	Public figure	Public figure
60	Rådhus	Town Hall
61	Radio station	Radio station
62	Radiokanal	Radio Channel
63	Radiostation	Radio station
64	Regeringswebsite	Government website
65	Religiøs organisation	Religious organization
66	Samfundsorganisation	Community organization
67	Samfunn	Society
68	Samfunnshus	Community center
69	Sociale tjenester	Social services
70	Statlig organisasjon	Government Organization
71	Statslig bygning	Government building
72	Statsstyret organisation	Government-owned organization
73	Tidsskrift	Magazine
74	TV Network	TV Network
75	Tv-kanal	TV channel
76	TV-kanal	TV Channel
77	Uddannelse	Education
78	Uddannelseswebsite	Education website
79	Udsendelses- og	Broadcasting and media production
	medieproduktionsselskab	company
80	Universitet	University
81	Universitet og høgskole	University and college
82	University	University

Appendix 4. Examples of comments and their algorithmic labeling

A brief vignette of N=9 random comment messages from the available AA dataset, in the context of their algorithmic labeling (i.e., with due possible classification tags and their scores).

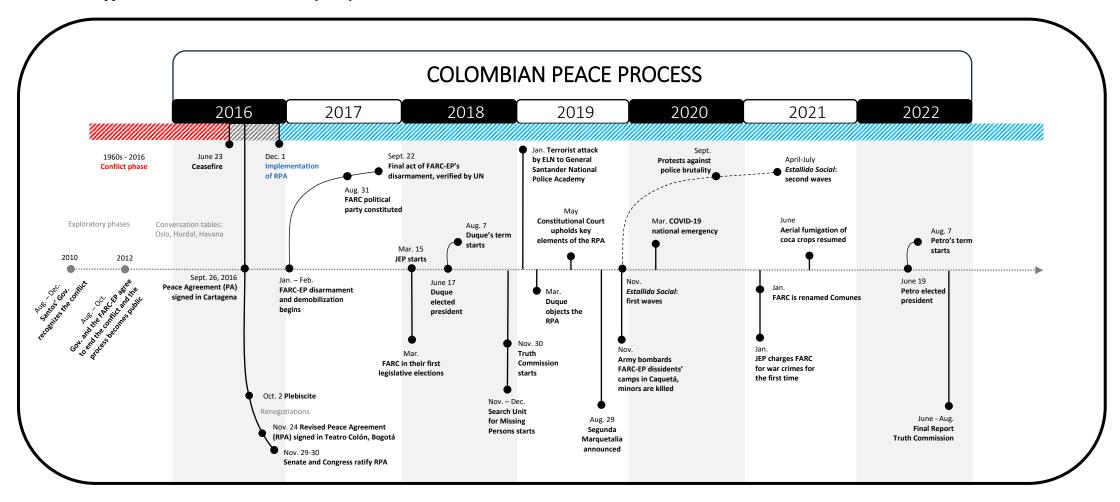
			:	Scores (proba	bilities 0 to 1)
Final algorit	thmic tags]	Algorithm disagre		_	2. Is there onism?
1 st step	2 nd step	Comment message	true	false	true	false
		Example 1. Excelente, propiciar sinergia insterinstitucional trae grandes beneficios para la comunidad. Felicitaciones señor Alcalde por esta importante gestion 🖎 🖎	0.02023431	0.9797657	NA	NA
not disagreement	NA	Example 2. Justicia, justicia, justicia, para Lucas y todos los jóvenes caídos!!!	0.2547308	0.7452692	NA	NA
		Example 3. Hay alguna manera de apoyar a la Comisión con trabajo voluntario, a mi particularmente me interesaría mucho!	0.02344085	0.9765592	NA	NA
	not antagonism	Example 4. [Name] hablando de ojos, aquí no hacemos nada sacandonos los ojos, mejor miremos que granito de arena aportamos a un país tan golpeado como es el nuestro	0.8758861	0.124114	0.09061995	0.90938
		Example 5. Les quedó grande la silla a los Alcaldes del área metropolitana, problema de años que tiene su punto de ebullición y aún sin una respuesta de fondo.	0.9296343	0.07036567	0.0412791	0.9587209
disagreement		Example 6. [Name] Se ve que Ud, no tiene hijos y nietos, yo sí y voy a las marchas por ellos, se merecen un mejor País, con oportunidades de estudio y trabajo dignos, recuerde que somos el segundo País con la educación más deficiente después de Etiopía, el empleo para graduados de U. con salario mínimo. Porque cree se van del País ?	0.6097182	0.3902818	0.1809351	0.8190649
	antagonism	Example 7. Con razón quieren acabar con la corte, es el único tropiezo que les queda para imponer la dictadura. Necesitan una corte de bolsillo	0.9735976	0.02640231	0.7779551	0.2220449
		Example 8. Loka decrépitaeste senador es un imbécil promoviendo el.odio y la diferencia entre el pueblo El.cobrando millones .son doble moral los senadores .como esteimbécil	0.9927249	0.00727502	0.99515	0.004850043

personas de bien que hagan algo por el pais			Example 9. Toca es fumigar ese congreso para que quede limpio de cucarachas y ratas ypuedan entrar personas de bien que hagan algo por el pais	0.7989146	0.2010854	0.532207	0.467793
---	--	--	---	-----------	-----------	----------	----------

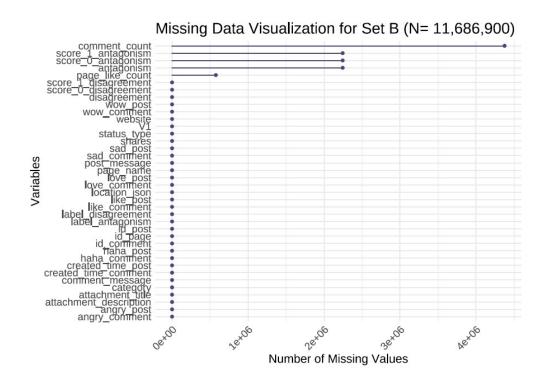
Appendix 5. List of unique Facebook public fan pages: the news media entities

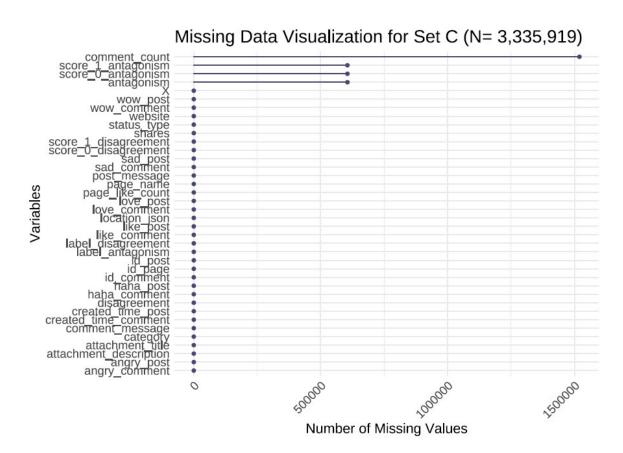
1	Actualidad Zuliana Noticias	33	La Maga
_	Agencia de Noticias - Universidad Nacional de	33	La Mega
2	Colombia	34	La Nota Económica
3	Antena 2	35	La Opinión Cúcuta
4	Blu Radio	36	La Patria Manizales
5	CABLENOTICIAS	37	Mi Diario Valledupar
6	Canal Capital	38	Noticias Caracol
7	Canal CNC Cartagena	39	Noticias Popayán y Cauca
8	Canal Teleantioquia	40	Noticias RCN
9	Canal Telemedellín	41	Noticias YA FM
10	Colombia Viral	42	Notidespeguecucuta
11	Colombia 2020	43	Periódico Digital Erbol
12	Comité Nacional de Paro	44	Periódico El Nuevo Liberal
13	Cromos	45	Periódico La Nación
14	Diario del Cesar	46	Periódico Umarí - Amazonas
15	Diario del Huila	47	Portafolio
16	Diario gratuito ADN Colombia	48	Proclama Cauca y Valle
17	Diario La Libertad	49	Radio Bolivariana
18	Diario La República	50	Radio Nacional de Colombia
19	Diario Occidente	51	RCN Radio
20	El Colombiano	52	Red Emergencias Valledupar Oficial
21	El Diario Pereira	53	RED+ Noticias
22	El Espectador	54	Revista Dinero
23	El Heraldo Barranquilla	55	Revista Gatopardo
24	El Nuevo Día - Colombia	56	Revista Semana
25	El País Cali	57	RTV Television
26	El Tabloide	58	Semanario El Extra de San Andrés
27	El Tiempo	59	Tele San Jacinto
28	El Universal	60	Telepacífico
29	Hacemos Memoria	61	TeleRoble
30	La 10 Diario Deportivo	62	Tolima Stereo 92.3 FM
31	La Crónica del Quindío	63	TSM Noticias
32	La Cuarta Vía	64	Univalle Estéreo

Appendix 6. Timeline of the Colombian peace process 2016-2022

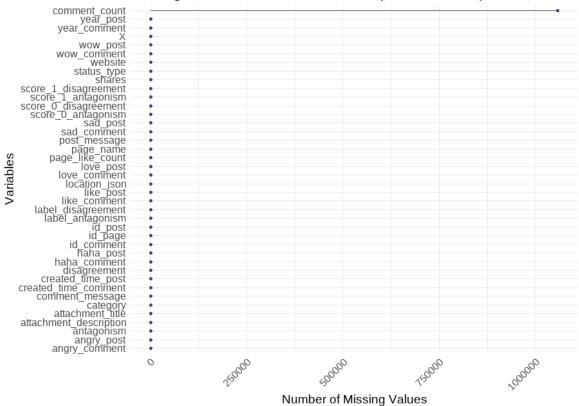


Appendix 7. Plots of missing data: Set B vs C vs D (before any variable transformation)

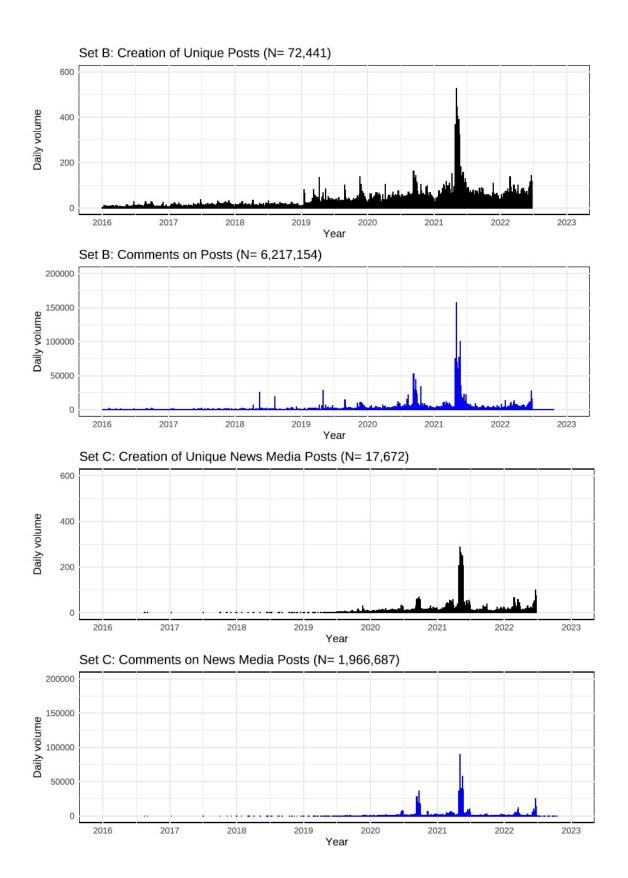


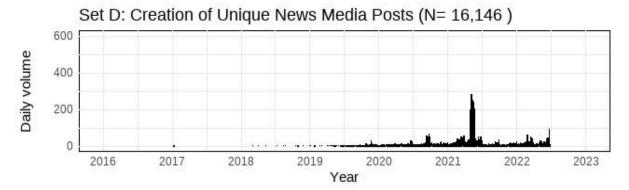




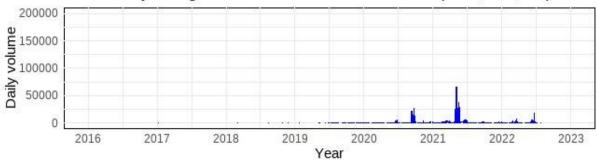


Appendix 8. Temporal distribution of the creation of FB posts and comments: Set B vs C vs D (from first posts on Jan. 3, 2016 to last comments on Oct. 23, 2022) (before any data transformation or cleaning)





Set D: Only Disagreements on News Media Posts (N= 1,361,291)



Appendix 9. Summary of covariates in clean Set D

Variable	not prevalent, N = 6,255 ¹	prevalent, N = 9,254 ¹
count_token		
Median (IQR)	33 (22 - 50)	32 (22 - 44)
Mean (SD)	62 (135)	50 (93)
Range	1 - 5,161	1 - 2,845
count_VERB		
Median (IQR)	4 (2 - 6)	3 (2 - 5)
Mean (SD)	7 (16)	6 (11)
Range	0 - 802	0 - 345
count_ADV		
Median (IQR)	1 (0 - 2)	1 (0 - 2)
Mean (SD)	2 (5)	1 (4)
Range	0 - 219	0 - 143
count_NOUN		
Median (IQR)	9 (6 - 14)	9 (6 - 12)
Mean (SD)	16 (33)	13 (23)
Range	0 - 1,053	0 - 636
count_PROPN		
Median (IQR)	0 (0 - 1)	0 (0 - 0)
Mean (SD)	0 (1)	0 (1)
Range	0 - 31	0 - 18
count_ADP		
Median (IQR)	6 (4 - 10)	6 (4 - 9)
Mean (SD)	12 (24)	10 (17)
Range	0 - 820	0 - 468

Variable	not prevalent, N = 6,255 ¹	prevalent , N = 9,254 ¹
count_DET		
Median (IQR)	6 (3 - 9)	6 (4 - 8)
Mean (SD)	10 (21)	8 (15)
Range	0 - 748	0 - 456
count_PRON		
Median (IQR)	1 (0 - 2)	1 (0 - 2)
Mean (SD)	2 (7)	2 (5)
Range	0 - 412	0 - 166
count_ADJ		
Median (IQR)	3 (1 - 4)	2 (1 - 4)
Mean (SD)	5 (11)	4 (8)
Range	0 - 274	0 - 201
year_post		
2020	1,366 / 6,255 (21.839%)	2,091 / 9,254 (22.596%)
2021	3,717 / 6,255 (59.424%)	5,409 / 9,254 (58.450%)
2022	1,172 / 6,255 (18.737%)	1,754 / 9,254 (18.954%)
total_comments_in_post		
Median (IQR)	6 (2 - 23)	22 (5 - 88)
Mean (SD)	35 (125)	122 (531)
Range	1 - 3,304	1 - 22,512
total_like		
Median (IQR)	62 (23 - 173)	89 (31 - 260)
Mean (SD)	276 (1,278)	403 (1,682)
Range	0 - 58,525	0 - 76,036

Variable	not prevalent, N = 6,255 ¹	prevalent, N = 9,254 ¹
total_shares		
Median (IQR)	18 (5 - 62)	26 (7 - 99)
Mean (SD)	126 (715)	227 (1,595)
Range	0 - 31,752	0 - 84,518
total_love		
Median (IQR)	2 (0 - 13)	3 (0 - 16)
Mean (SD)	46 (327)	58 (401)
Range	0 - 17,264	0 - 16,816
total_angry		
Median (IQR)	1 (0 - 8)	6 (1 - 33)
Mean (SD)	22 (170)	93 (847)
Range	0 - 5,784	0 - 58,613
total_sad		
Median (IQR)	1 (0 - 7)	2 (0 - 10)
Mean (SD)	37 (337)	41 (276)
Range	0 - 18,004	0 - 14,749
total_haha		
Median (IQR)	1 (0 - 7)	7 (1 - 44)
Mean (SD)	21 (131)	128 (772)
Range	0 - 5,009	0 - 28,796
total_wow		
Median (IQR)	1 (0 - 5)	2 (0 - 8)
Mean (SD)	9 (32)	13 (48)
Range	0 - 857	0 - 2,180

Variable	not prevalent, $N = 6,255^{1}$	prevalent , N = 9,254 ¹
----------	--------------------------------	---

page_like_count		
Median (IQR)	219,732 (76,484 - 792,616)	486,928 (102,784 - 2,519,475)
Mean (SD)	920,407 (1,543,853)	1,371,922 (1,844,822)
Range	1,748 - 6,395,004	1,748 - 6,395,004
location_json		
Rest	3,886 / 6,255 (62.126%)	4,265 / 9,254 (46.088%)
Bogotá	2,369 / 6,255 (37.874%)	4,989 / 9,254 (53.912%)
war_concept		
FALSE	5,130 / 6,255 (82.014%)	7,057 / 9,254 (76.259%)
TRUE	1,125 / 6,255 (17.986%)	2,197 / 9,254 (23.741%)
Sentiment		
Equal	1,464 / 6,255 (23.405%)	2,225 / 9,254 (24.044%)
More negative words	2,099 / 6,255 (33.557%)	3,061 / 9,254 (33.078%)
More positive words	2,692 / 6,255 (43.038%)	3,968 / 9,254 (42.879%)

Appendix 10. Summary of the final model's binary logistic regression coefficients

	prevalence					
Predictors	Odds Ratios s	td. Beta	CI	standardized CI	I Statistic	р
(Intercept)	0.80 ***	1.26	0.71 - 0.91	1.11 - 1.43	-3.39	0.001
count token	1.01 *	2.81	1.00 - 1.02	1.07 - 7.08	2.19	0.029
count VERB	1.01	1.16	0.99 - 1.03	0.91 - 1.50	1.20	0.228
count ADV	0.98	0.92	0.96 – 1.01	0.82 - 1.04	-1.36	0.173
count NOUN	0.96 ***	0.34	0.94 - 0.98	0.20 - 0.58	4.00	<0.001
count PROPN	0.95 *	0.94	0.90 - 0.99	0.89 - 0.99	-2.41	0.016
count ADP	1.00	1.00	0.98 - 1.02	0.72 - 1.42	0.00	1.000
count DET	1.01	1.15	0.99 - 1.03	0.86 - 1.56	0.91	0.362
count PRON	0.96 **	0.78	0.93 - 0.99	0.66 - 0.92	-3.00	0.003
count ADJ	1.00	1.04	0.99 - 1.02	0.89 - 1.21	0.45	0.653
year post [2021]	1.00	1.00	0.91 - 1.11	0.91 - 1.11	0.06	0.953
year post [2022]	1.21 **	1.21	1.07 – 1.37	1.07 - 1.37	3.02	0.003
total comments in post	1.00 ***	4.44	1.00 - 1.00	3.27 - 6.10	9.40	<0.001
total like	1.00	1.01	1.00 - 1.00	0.87 - 1.17	0.10	0.920
total shares	1.00	1.04	1.00 - 1.00	0.95 - 1.16	0.72	0.471
total love	1.00 ***	0.78	1.00 - 1.00	0.68 - 0.89	-3.59	<0.001
total angry	1.00	1.32	1.00 - 1.00	0.99 - 1.86	1.71	0.087
total sad	1.00 ***	0.86	1.00 - 1.00	0.81 - 0.91	4.88	<0.001
total haha	1.00 ***	4.00	1.00 - 1.00	2.81 - 5.81	7.50	<0.001
total wow	1.00 *	0.92	1.00 - 1.00	0.86 - 0.98	-2.42	0.016
page like count	1.00 *	1.06	1.00 - 1.00	1.01 - 1.11	2.44	0.015
location json [Bogotá]	1.68 ***	1.68	1.54 - 1.82	1.54 - 1.82	12.23	<0.001
war concept [TRUE]	1.55 ***	1.55	1.40 - 1.71	1.40 - 1.71	8.62	<0.001
Sentiment [More negative words]	1.00	1.00	0.91 – 1.11	0.91 – 1.11	0.09	0.927
Sentiment [More positive words]	1.06	1.06	0.96 – 1.17	7 0.96 – 1.17	1.21	0.226
Observations	11631					
R ² Tjur	0.086					
Deviance	14641.900					
AIC	14691.900					
log-Likelihood	-7320.950					

*p<0.05 **p<0.01 ***p<0.001