A CONVERSATION
***ABOUT
ARTIFICIAL
INTELLIGENCE.

{###}//##

{+++}

Much in common. ↓

The topics and speech of democracy in its 40 years.

AN ANALYSIS OF
//PRESIDENTIAL
SPEECHES THROUGH
##ARTIFICIAL
INTELLIGENCE++

<*@/>

Co_Lab

UN
DP

Accelerator Lab
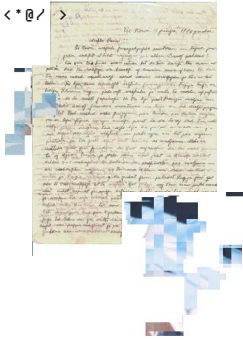**UNDP Argentina**

{###}

Index

Introduction

This year 2023 marks the 40th anniversary of the return of democracy in Argentina and the consolidation of the longest democratic period in the history of the country. During this time, Argentina has experienced crises and bonanzas, conflicts and agreements, failures, and successes; all of which were resolved within the rules of the institutional game. However, conflict and political polarization increasingly undermine the possibility for dialogue, limit the basic consensus embodied by the fusion of common values, and deepen the distance between people aligned with different political projects (Aruguete & Calvo, 2020). This deepening of political polarization is a trend that is not exclusive to Argentina, as it is present in the democracies of very diverse countries (Gu & Wang, 2022). Its causes are multiple and have been studied from different disciplines (Qureshi, 2021). One of the variables that contributes to this political polarization is the use of social media which, due to its logic of consumption and interaction. In fact, artificial intelligence (AI) algorithms are developed to reinforce the beliefs of its users, show them content related to their perspectives, and exacerbate their negative feelings (Van Bavel et al., 2021). As a result, people use and spend more time on social media because they feel emotionally compelled and affected by the content they are exposed to. In short, users' engagement is fostered as the success of social media platforms entails reaching more people to make them active user and spend more time in these platforms There is evidence of the effects that this use of AI has on the consolidation of political polarization (Tucker et al., 2018). The consequences of this can be risky for democratic governance insofar as polarization positively correlates with governmental instability (Binder, 2000), a decline in democratic quality (Casal Bértoa, 2019), conflicts and social protests (Singer, 2016), dissemination of fake news (Osmundsen, 2021), delegitimization of the media and science (Liminga, 2022), and political violence (Burton, 2023); all that, ultimately, threatens the stability of democracy (Sartori, 1976; Linz, 1978; Casal Bértoa, & Enyedi, 2019).

> → "... AI is a field that combines computer science with information to simulate human intelligence. AI systems can mimic cognitive operations, such as recognizing objects and voices, generate text and images, or make recommendations, without needing precise instructions every time they do so. This is possible because they work with models that emulate the way people think and process information. In this sense, AI systems are said to be able to 'learn'."
>
> [Moscovich, 2023] → view more

Strategies for moderating polarization are complex, because, for example, when consensus is sought to facilitate dialogue or unity, if the person or organization convening the discussion is associated with a political force or opposing partisan space, the same political polarization inherently limits this possibility. (Levy Yeyati et al., 2020).

Faced with this scenario, we raise the following question: Are there commonalities and consensuses in presidential speeches over the years that are independent of party affiliations? Can AI contribute to this enterprise?

This exercise is a renewed invitation to dialogue by UNDP, based on the identified shared priorities of Argentines, especially in this year that marks the 40th anniversary of democracy. AI was useful for this purpose because, on the one hand, it enabled us to analyze a body of speeches quickly and systematically and make different analyses. On the other hand, it helped us track signs of agreement using evidence from different party speeches. This use of AI could facilitate the reception of the results by audiences with different political preferences. In other words, although AI could display various biases, for example, derived from the information with which it was trained (Kaneko & Bollegala, 2022), on this occasion, being the issuer of this proposal of common topics and being fed with data from the entire democratic period, these biases would not be aligned with the themes of polarization of Argentina's party and/or electoral politics. Once again, UNDP aspires to perform a facilitating and articulating role between different sectors. We saw in the use of AI an opportunity to provide a legitimate input for citizenship.

UNDP as a facilitator of consensus: the history of the Argentine dialogue in 2001

During the social, economic, and political crisis of 2001, consensus and dialogue in regard to common priorities played a crucial role in reaching agreements in a context less polarized but much more conflictive in social terms. In this scenario, UNDP played a central role in convening relevant stakeholders and encouraging that conversation. Today, the context is different, but once again, UNDP can explore the shared concerns and values that can lay the foundation for a new consensus in the country, based on the core concerns and priorities of the Argentine citizens. On this occasion, 22 years later, we have additional tools to analyze large volumes of information in several ways to achieve this goal.

CHATGPT_3.5

∕PROMPT

{###}
∕∕∕
∕∕∕∕
#
##

{###}∕∕#

{###}∕∕#

In order to accomplish our purpose, we developed a strategy that employs natural language processing (NLP) tools, a branch of AI that enables computers to understand, generate, and manipulate human language. Specifically, we trained an AI model using ChatGPT 3.5 to analyze presidential inaugural addresses and speeches at the opening of legislative sessions. In total, we examined 52 speeches delivered from the return of democracy to the present, in three instances of analysis. The first instance involved a descriptive analysis of the speeches themselves, which entailed the evaluation of their extension and the frequency of occurrences of certain key concepts. The second part focused on examining whether the speeches addressed 30 policy topics and how they did it. In this instance, we used AI to analyze several aspects, such as the relationships between topics, the frequency and time of occurrences of certain concepts and words, and the way the presidents approach each topic, among others. In the third instance, the emphasis was on creating a unified speech based on the processing of all the other complete speeches. These presidential messages reflect the state of the nation at that time, the vision each president has for the country, the topics that were considered a priority, and how these topics were to be addressed (Bercholc & Bercholc, 2013). The generated speech focuses on main themes emerging in the presidential messages such as economy and economic stability; education, public health, and primary health care; science, technology, and culture; foreign policy and foreign affairs; self-esteem of the Argentine people and institutional reconstruction; democracy and social justice; integration in the region and the world; and improvement objectives.

The speech of democracy in its 40 years

We analyzed 52 inaugural addresses and speeches at the opening of legislative sessions of democratically elected presidents who have delivered at least one of these speeches from 1983 to the present.

We were also able to explore the possibilities and limitations of this new AI tool, particularly for political speech analysis exercises. Specifically, we used machine learning tools, from the Transformers architecture, which enables the use of texts as data and is included in algorithms such as ChatGPT. Models based on this architecture have proven to be effective in understanding and classifying text with high precision, as well as in generating text with a fluency and naturalness that emulates human language (Tunstall, von Werra & Wolf, 2022; Liu et al., 2019; He et al., 2020). These models have been used in the analysis of legislative speeches in Latin America (German, Micozzi, & Vallejo Vera, 2022), presidential debates in the United States (Bonikowski, Luo, & Stuhler, 2022), and policy in England (Abercrombie et al., 2019). However, to the best of our knowledge, this type of systematic speech analysis carried out to analyze policy topics, search for commonalities, and generate a unified text using ChatGPT is unique. In this sense, this project is presented as a first experiment to assess the potential of these tools. This document is a summary of the manuscript El discurso de la democracia argentina en sus 40 años [The Speech of Argentine Democracy in its 40 Years] (Vallejo Vera, 2023) that contains a more extensive analysis of the methodological approach, as well as additional evidence on analyzed topics, and is available to the reader.

The rest of this document will first briefly describe the speeches and explain the strategy we carried out for their analysis. The second part will focus on the topics and analyze them in depth. The third part will describe the steps for the generation of the unified speech and present the results of this exercise. Finally, there will be a discussion of the results as a conclusion.

Presidential speeches

{+++}

{###}

###

{###}

{###}

LENGTH

+++PROMPT

CONTENT

Characteri- zation of the presidents' speeches

First, we wanted to know how the speeches of the presidents looked like. Once we processed them, we analyzed their length and content, showing variations over time and per president.

{###}
#
##

{###}
//##

{###}

Step by step. The processing of speeches

The analyzed evidence emerged from different processing and analysis strategies of the 52 presidential speeches delivered from the return of democracy to the present day, that is, from 1983 to 2023. From these texts, as can be seen in Table 1, we first preprocessed data, eliminating fragments and special characters that do not belong to the specific content of each presidential speech.

DELETED TEXT	REASON
Prefaces, foot notes, page numbers.	Page format text.
Words from non-relevant political stakeholders [e.g., the President of Congress], official greetings, and meeting adjournments.	The majority of the documents in which the speeches were included are written records of the plenary sessions of the Argentine Congress. During the sessions, the President of the Congress, the Secretary, and, on occasion, the public, also participate. All of this information is deleted. We keep the speech in its entirety, including the opening and closing remarks.
Special characters and shorthand errors.	Flaws in the text.

Table 1: Pre-processing of speeches

In the same order, Table 2 presents the main analysis and processing of the corpus of texts, which included editing the content of each speech by removing stopwords and other words that were not considered to prevent biases in the analyses.

DELETED TEXT	REASON
Stopwords ['empty' words, e.g., articles, personal pronouns, etc.].	These are words that are used to give grammatical meaning to a text, but do not have any substantial meaning.
The following words: "applause", "p.", "march", "country", "year", "years", "if", "to be", "argentine", "gallery", "hundreds", "millions", "argentines", "hundred", "national", "today", "every", "everyone", "everything", "whole", "government", "session", "legislative", "thousand", "honorable", "sr.", and "president".	These are recurring words in presidential speeches but they are not related to the topics analyzed.
Special characters and shorthand errors.	Flaws in the text.

"Table 2: Processing of speeches"

As for their length, President Cristina Fernández delivered an average of 17,360 words per speech, while President Rodríguez Saá delivered a single speech of 2,487 words. Specifically, Figure 1 shows the distribution of the number of words per speech over time and per presidency.
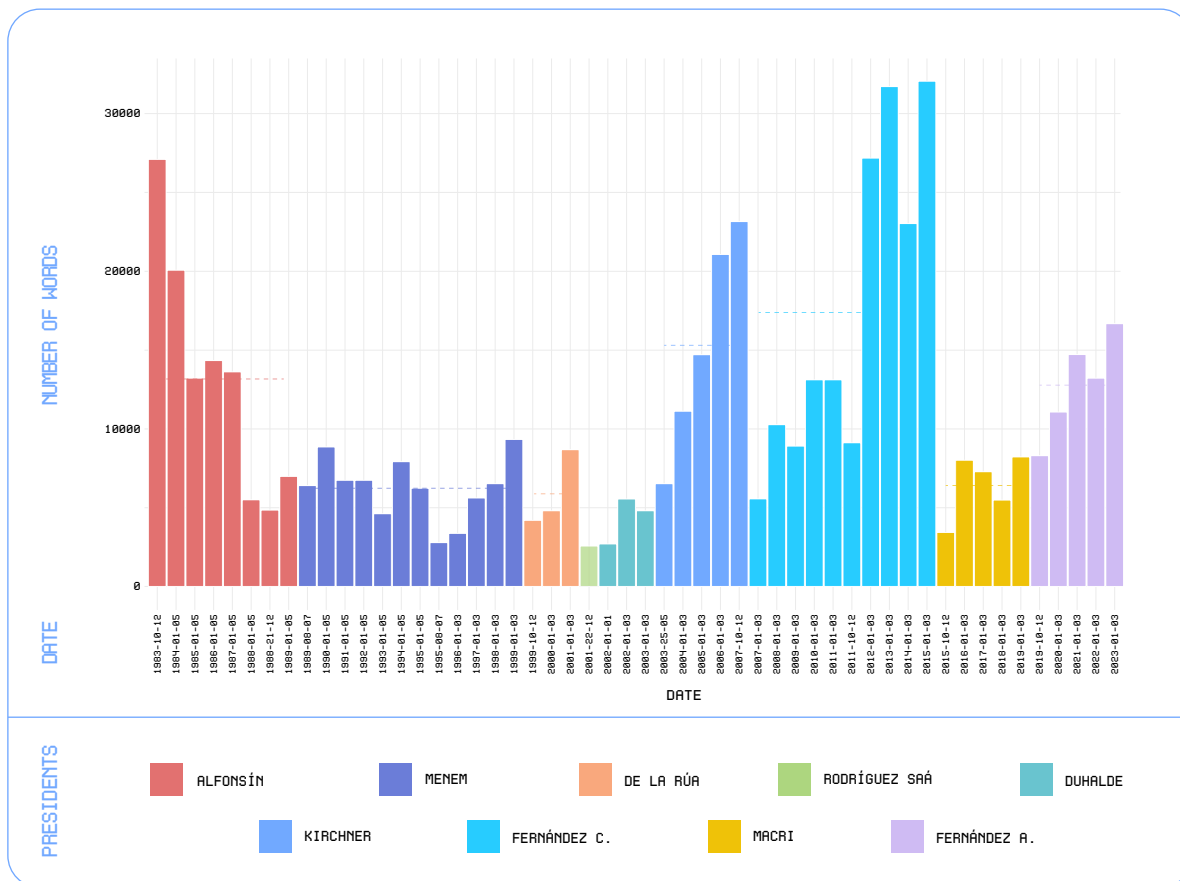

Figure 1: Number of words mentioned in each presidential speech. The texts include inaugural addresses and speeches at the opening of legislative sessions and extraordinary sessions. The average number of words per speech of each president is shown with a dotted line.

next page
fig. 2

The 10 most mentioned words in all of the presidential speeches (Fig. 2), excluding the word **Argentina**, are **social**, the most recurrent, followed by **policy** and **labor**. Presidents who ruled for more years, for example, Carlos Menem and Cristina Fernández, have more speeches and this influences the weight of their speeches in the analyses of the number of words mentioned. Insofar as the analysis does not seek to be representative but exhaustive of this body of evidence, the frequency of word occurrence shapes the general vision that characterizes the forty years of Argentine democracy, regardless the fact that some presidents gave more speeches than others.
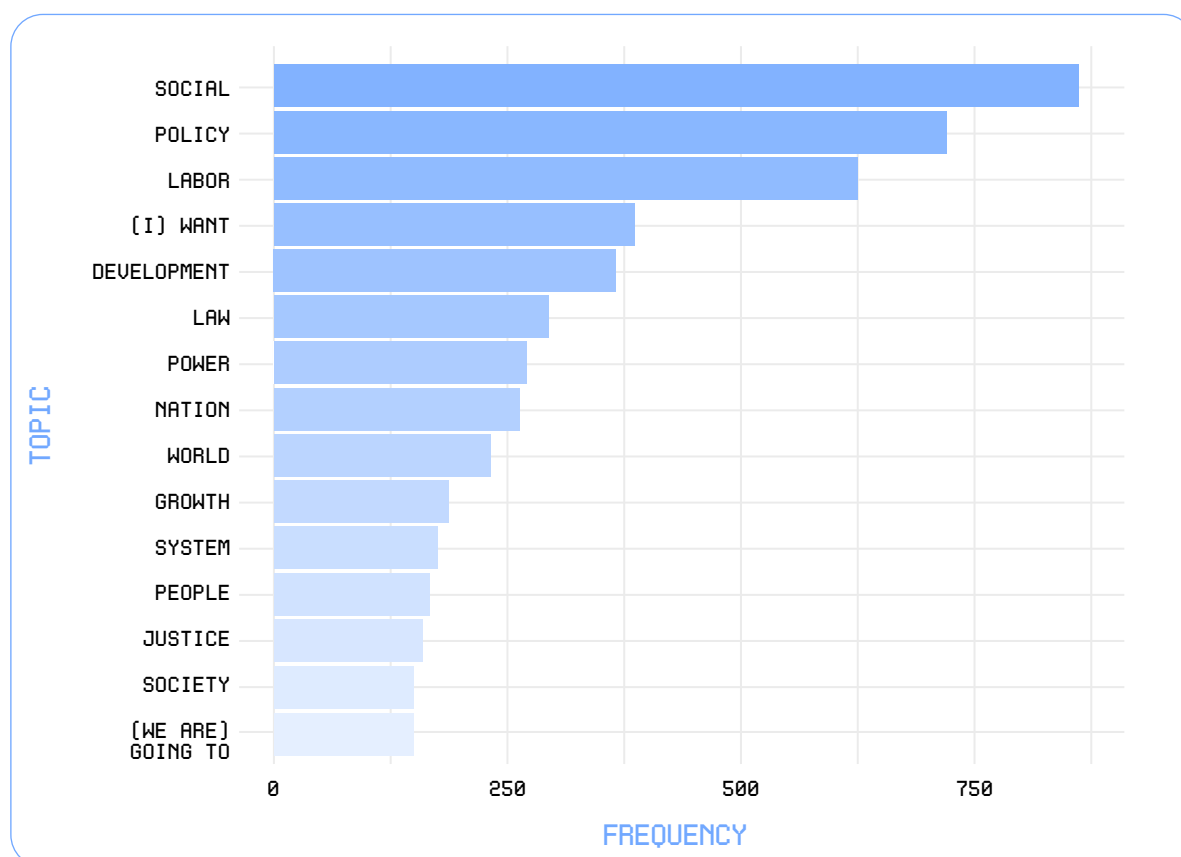
Figure 2: The 15 most mentioned words in the presidential speeches of Argentine presidents from 1983 to 2023. For this figure, all the speeches were concatenated, and the words extracted.

PRESIDENT	START AND END OF PRESIDENCY	TERM OF OFFICE	NUMBER OF SPEECHES INCLUDED
Raúl Alfonsín	December 10, 1983 – July 8, 1989	5 years and 7 months	8
Carlos Menem	July 8, 1989 – December 10, 1999	10 years and 5 months	12
Fernando De la Rúa	December 10, 1999 – December 21, 2001	2 years	3
Adolfo Rodríguez Saá	December 23, 2001 – December 30, 2001	7 days	1
Eduardo Duhalde	January 2, 2002 – May 25, 2003	1 year and 4 months	3
Néstor Kirchner	May 25, 2003 – December 10, 2007	4 years and 7 months	5
Cristina Fernández	December 10, 2007 – December 9, 2015	8 years	10
Mauricio Macri	December 10, 2015 – December 10, 2019	4 years	5
Alberto Fernández	December 10, 2019 – In progress as of the publication of this report	4 years	5
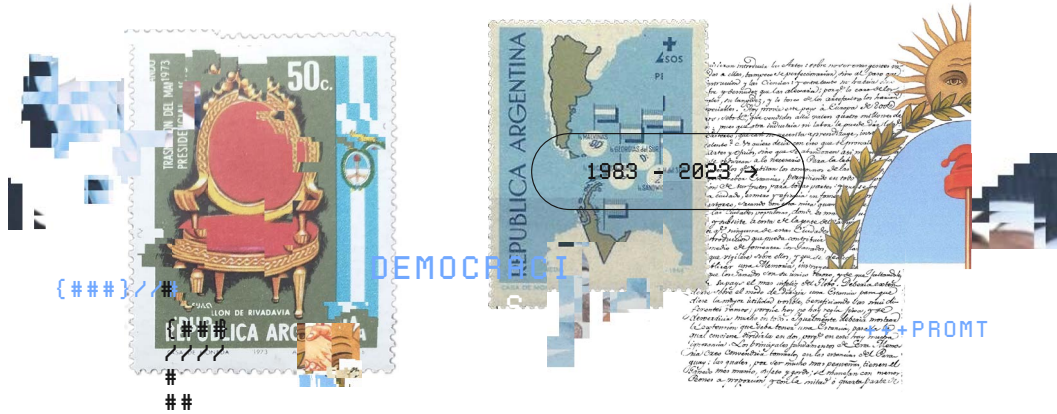
At a more specific level, the words most mentioned by the presidents in their speeches reveal the priorities they addressed in that context of the country (see Figures 3 and 4). For example, emphasizing the democratization process was central to the first speeches in the early 1980s, while more recent speeches, no longer overshadowed by the dictatorship, shifted their focus to other topics that were considered relevant at that time. For example, from Duhalde's presidency onwards, one of the most mentioned words was **labor**, marking a change in the main themes.
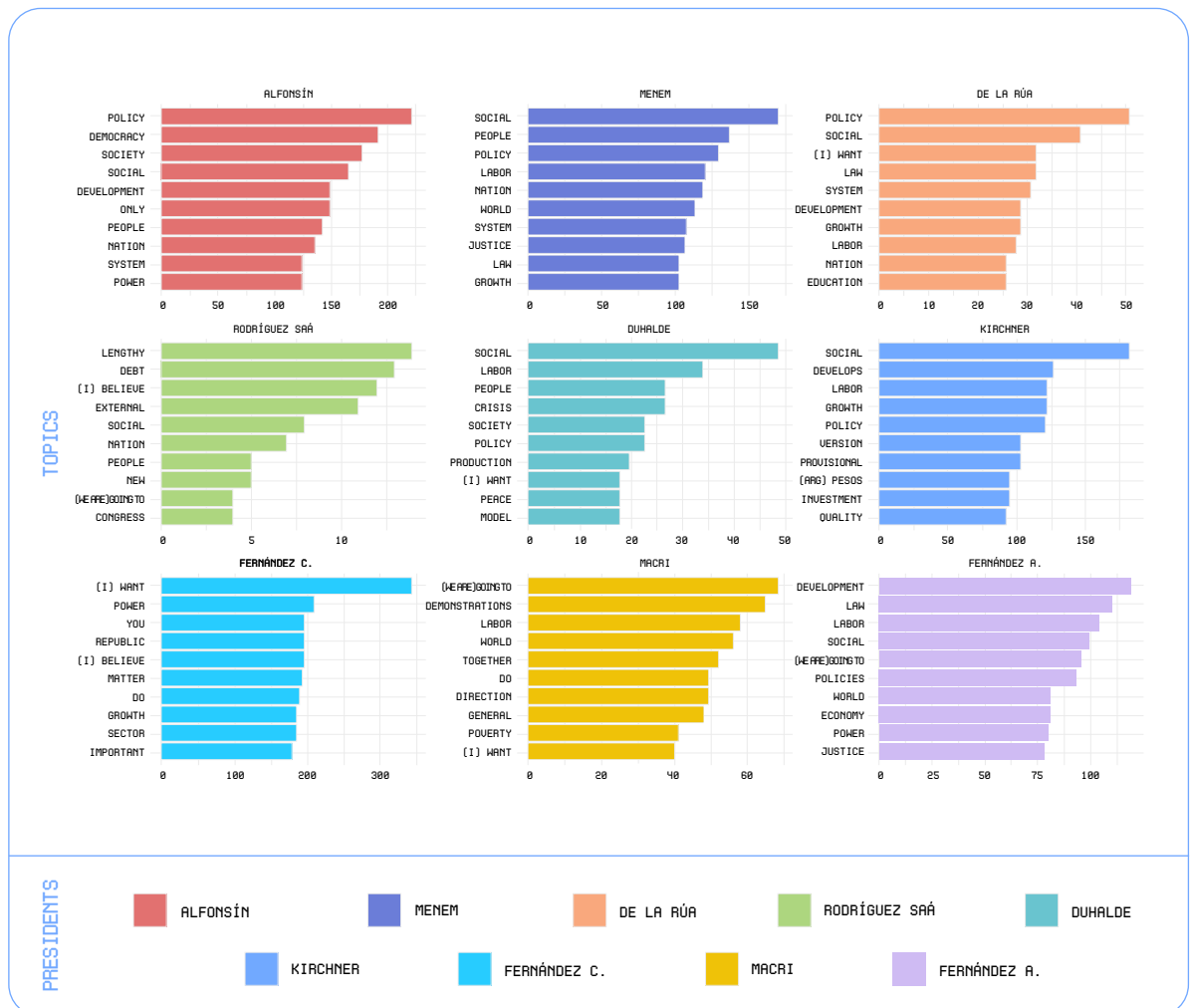
Figure 3: The ten most mentioned words in the speeches of Argentine presidents from 1983 to 2023, per president.

When analyzing the frequency with which certain words are mentioned over time, it is observed that President Macri focused more on structuring his speeches around the topic of **poverty**, while President Alberto Fernández centered his presidential messages around the term **development**.
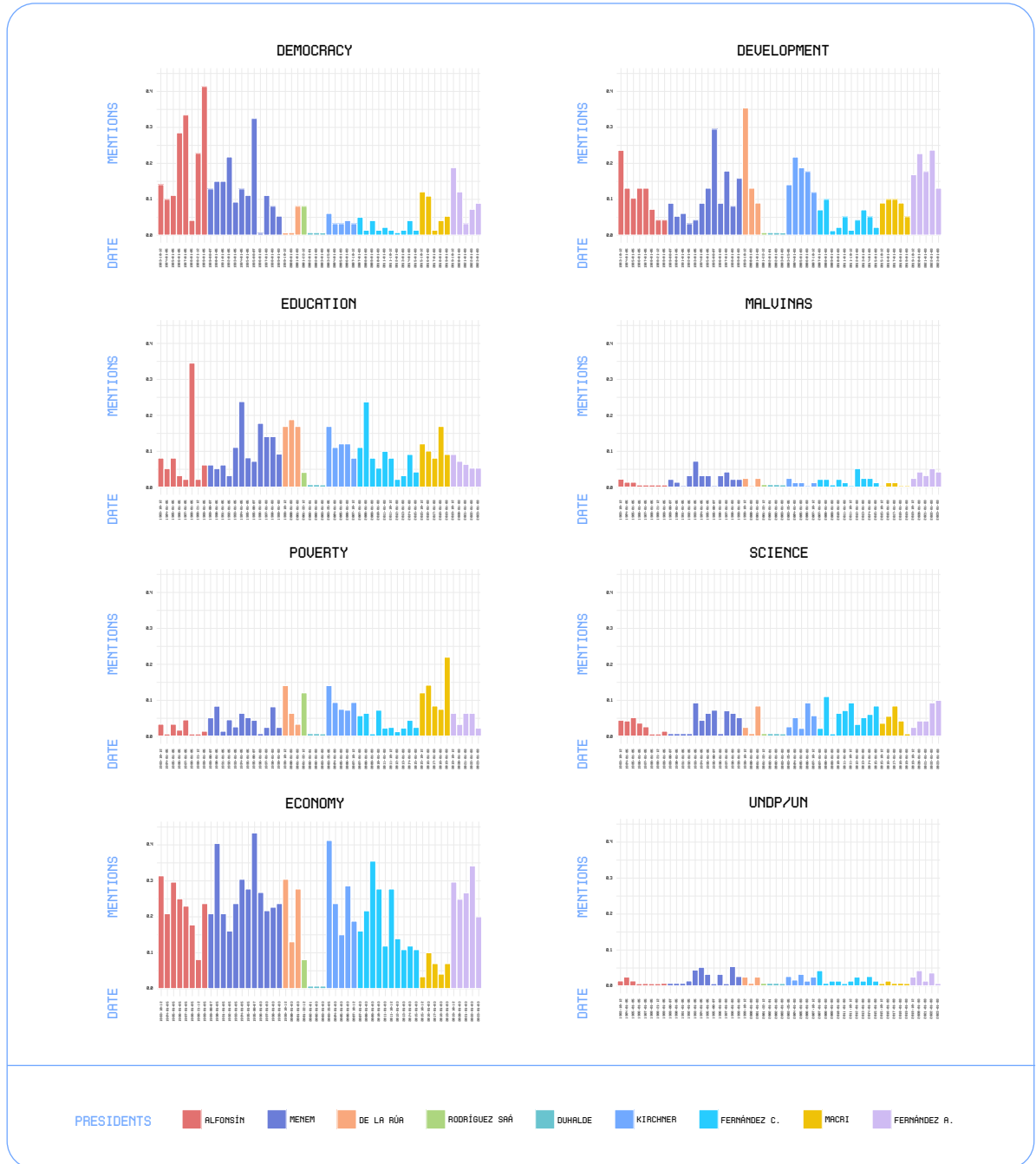
Figure 4: Percentage of mentions of topics of interest over time. The percentage is estimated by counting the number of terms mentioned divided by the total number of words in each speech.

In short, the variation in the main themes of the speeches of each president throughout his or her term reflects the importance and influence of election cycles, inner political dynamics, and the audience they addressed through their presidential addresses. In other words, these variations were aligned with each president's specific vision of the country, their policies, and the state of affairs, rather than party differences.

Speech topics

The analysis of the democracy speeches topics in its 40 years showed that the most mentioned word is **development**, followed by **macro-economics**, and **democracy**. In this sense, Figure 5 shows the ranking of topics according to the number of mentions. When conducting this analysis, the fact that the connotation of or focus on the topics may vary over time must be considered. For example, **development** refers to both human and economic development, while **democracy** may have a different connotation in the speeches pronounced immediately after the dictatorship compared to the more recent presidential messages where, for example, democratization of access to certain rights is mentioned. Furthermore, analyzing what is missing in the speeches is crucial. Topics like **gender, environment, sports, foreign affairs, migration**, and **media** received scant attention, revealing significant gaps in the discourses.
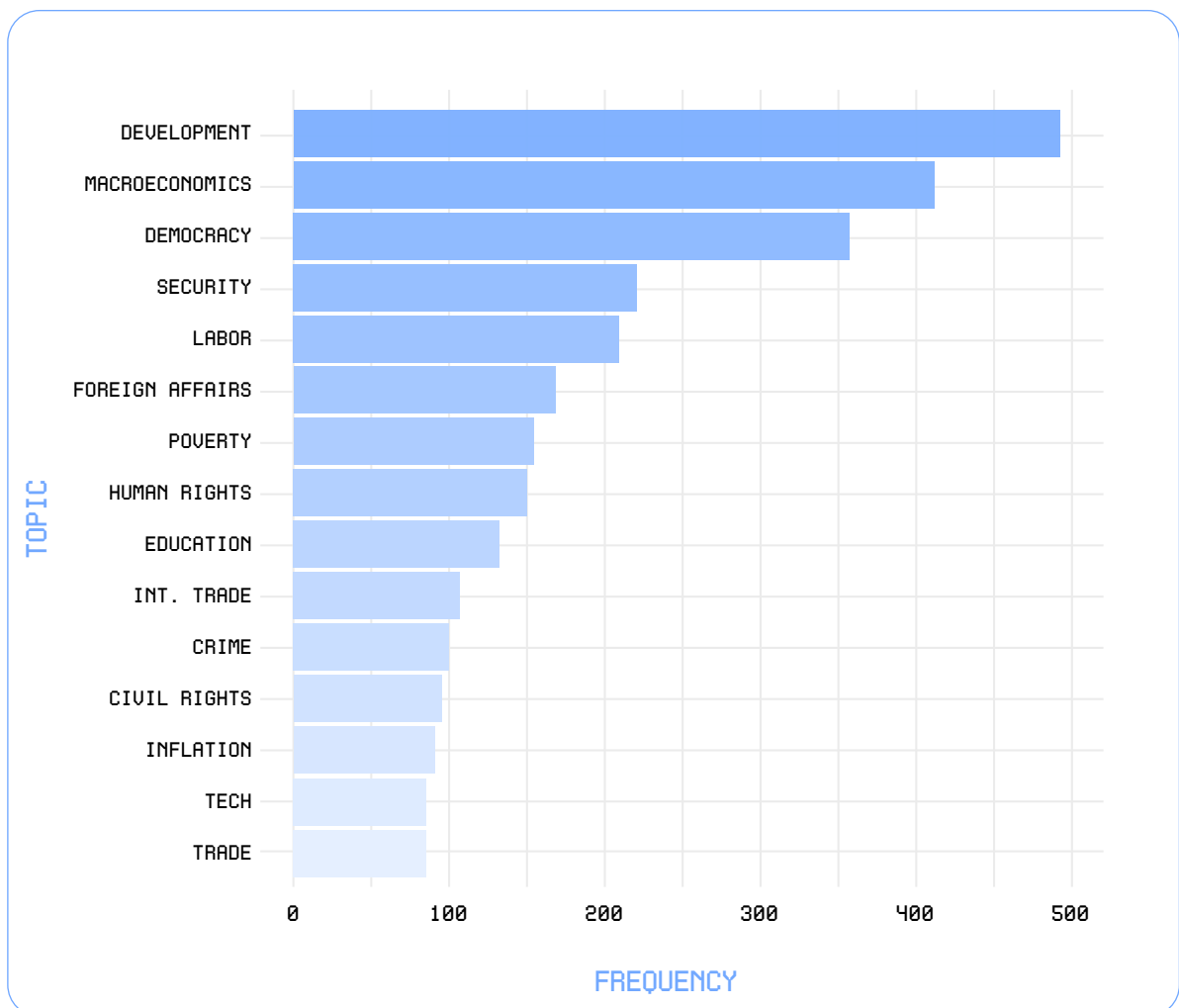
Figure 5: The 15 most mentioned topics in the presidential speeches from 1983 to 2023.

With this data, it is also possible to analyze the mention frequency of each topic over time (Figure 6). The focus on **democracy**, for example, was most noticeable during the first few years after the end of the dictatorship. A contrary trend, although to a lesser extent, is seen in topics such as **technology**. In the 1980's, it was often not addressed while, in the last decade, technology always had a place in presidential speeches. Conversely, other topics have been constantly present over the years, such as **development,** which has been mentioned in 51 of the 52 speeches. Similarly, **education, health,** and **macroeconomics** have maintained a relatively constant presence, with periods of increased focus on their importance. On the contrary, there are topics such as **gender**, only mentioned in 7 of the 52 speeches, which indicate the lack of a gender-based approach in presidential speeches.
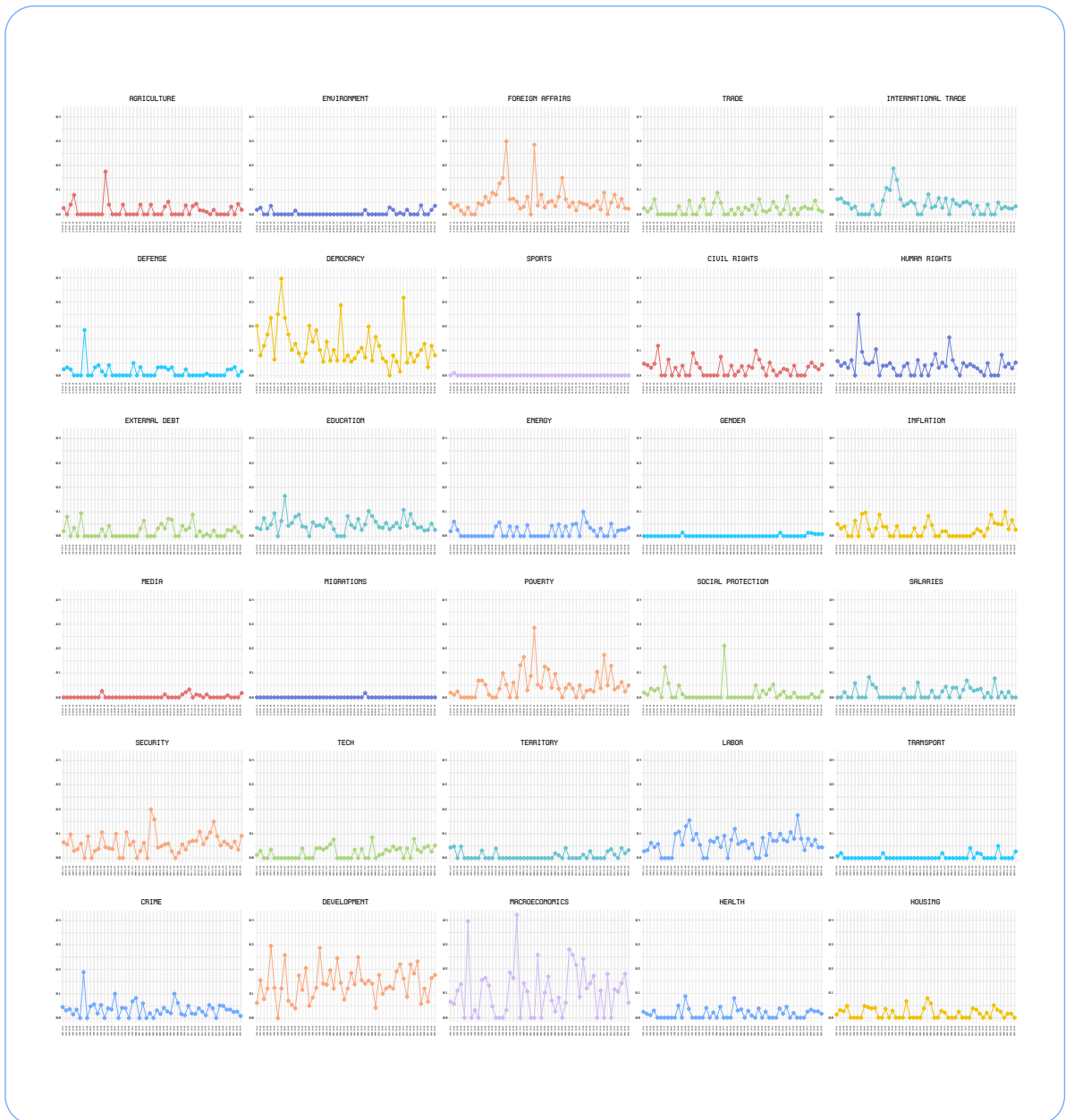
Figure 6: Percentage of mentions per topic in each speech through the years.

Steps to extract the topics of the speeches

The first step was to choose and define the topics we wanted to identify. For this purpose, we developed a codebook with 30 topics, each of which is accompanied by a definition based on related words. For example, if we are interested in identifying the parts of the speech that cover poverty topics, we can define it as follows: "The speech covers topics related to poverty and includes: poverty threshold, extreme poverty, indigence, Unsatisfied Basic Needs (UBN), inequality, income distribution, vulnerability and social conditions, and multidimensional poverty" (the complete list of topics with their keywords is available to the reader in the Annex).
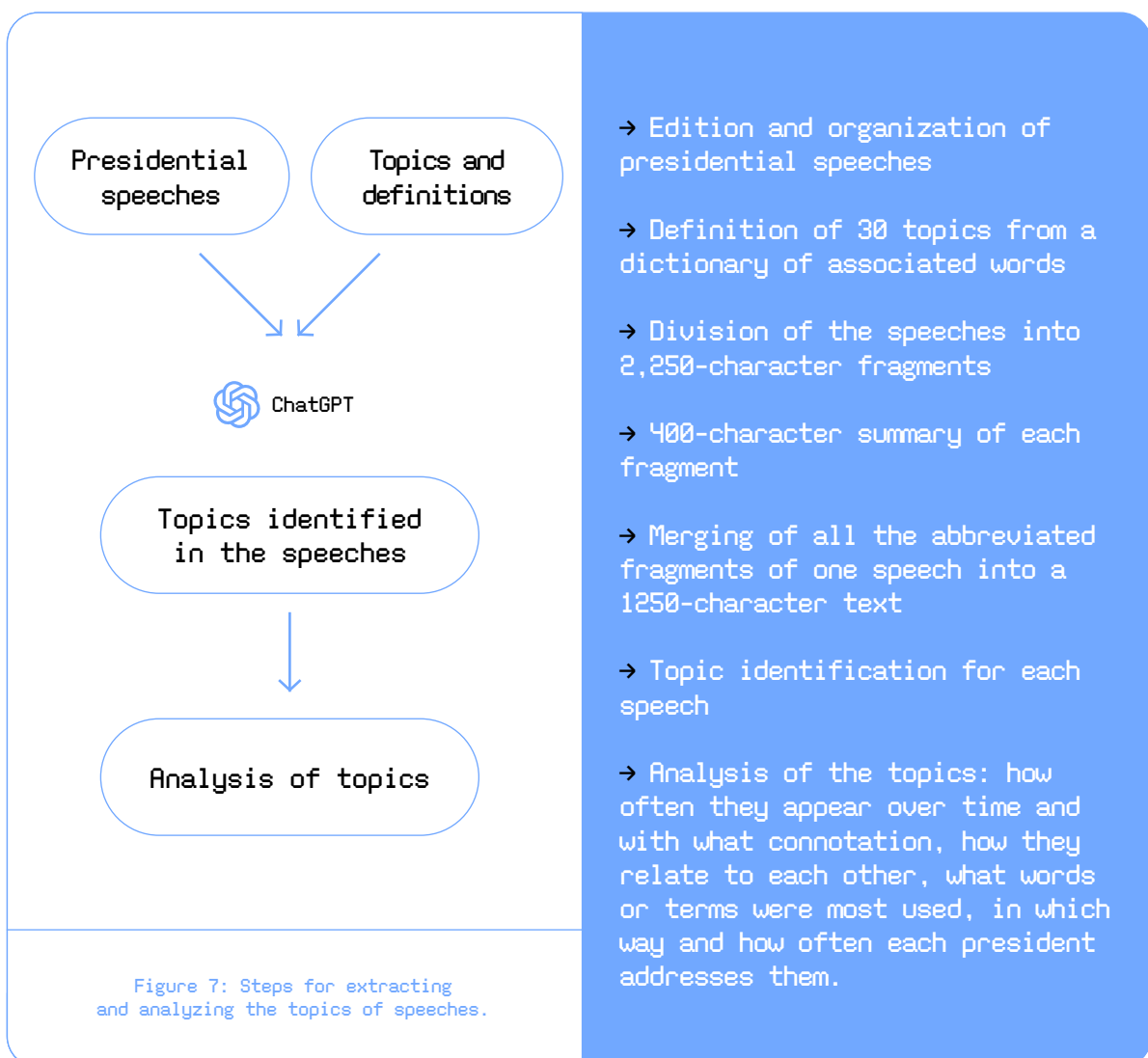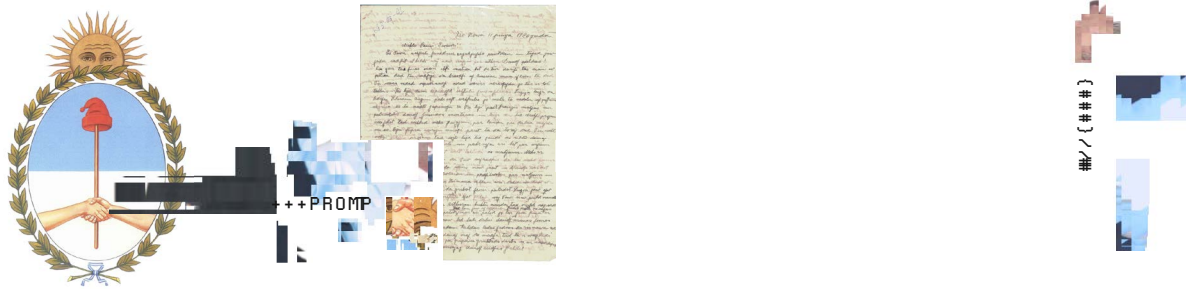
Once we developed a codebook per each topic, we proceeded to divide the speeches into fragments of 2,250 characters (chunks) which enabled us to stick to the limit of 4,000 characters that ChatGPT has between input (the question) and output (the answer). Each of these chunks is then sent to ChatGPT with a prompt asking if that fragment addresses a particular topic (along with the definition of the topic) and, if so, ChatGPT is asked to copy exactly the specific part in which it appears. This action is repeated for each chunk and each of the topics. In this way, we obtain the topics addressed in all the fragments of each speech.


```
┌──────────────────────────────────┐──────────────────────────────────┐
│                                  │                                  │
│  ╭───────────╮  ╭───────────╮    │  → Edition and organization of   │
│  │Presidential│  │  Topics and│   │    presidential speeches         │
│  │  speeches  │  │ definitions│   │                                  │
│  ╰───────────╯  ╰───────────╯    │  → Definition of 30 topics from a│
│                                  │    dictionary of associated words│
│             ChatGPT              │                                  │
│                                  │  → Division of the speeches into │
│  ╭───────────────────────╮       │    2,250-character fragments     │
│  │    Topics identified   │       │                                  │
│  │    in the speeches     │       │  → 400-character summary of each │
│  ╰───────────────────────╯       │    fragment                      │
│                                  │                                  │
│  ╭───────────────────────╮       │  → Merging of all the abbreviated│
│  │    Analysis of topics  │       │    fragments of one speech into a│
│  ╰───────────────────────╯       │    1250-character text           │
│                                  │                                  │
│                                  │  → Topic identification for each │
│                                  │    speech                        │
│                                  │                                  │
│                                  │  → Analysis of the topics: how   │
│  Figure 7: Steps for extracting  │    often they appear over time and│
│  and analyzing the topics of     │    with what connotation, how they│
│  speeches.                       │    relate to each other, what words│
│                                  │    or terms were most used, in which│
│                                  │    way and how often each president│
│                                  │    addresses them.               │
└──────────────────────────────────┘──────────────────────────────────┘
```

Figure 7: Steps for extracting and analyzing the topics of speeches.

Additionally, the 10 most mentioned topics per president (Figure 8) indicate, to some extent, the priority that each president gave to these topics, as well as the context in which he or she assumed the presidency and governed. For example, Rodríguez Saá mostly mentioned **development** and **social protest**, which is to be expected given the context of financial, political, and social crisis in our country at that time. On the contrary, in the speeches of Macri and A. Fernández, particular attention was paid to the topics of **labor, poverty,** and **social security,** although the presidents addressed them differently.
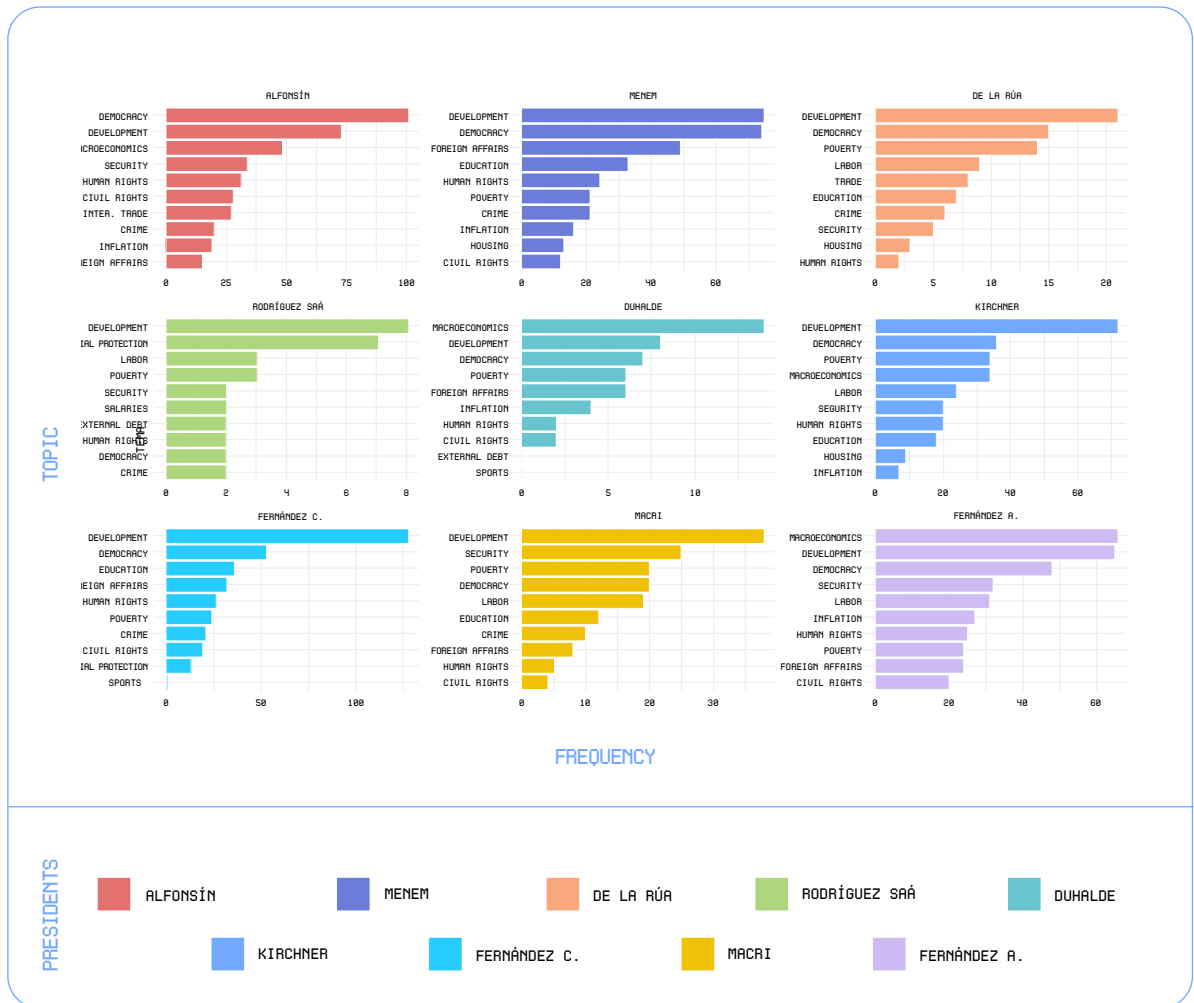
Figure 8: The 10 most mentioned topics in presidential speeches per president.

Similarly, we can see to what extent the presidents combine the different topics through the correlation matrix (Figure 9). Throughout the period analyzed, **development** has frequently been associated with **technology, education,** and **health** while **democracy** has been linked to **civil rights** and **human rights.** It should be clarified that these correlations include all topic associations over time, but each president may have made his or her own connections between the topics, and that some associations may even be a product of chance because of the way presidential speeches are traditionally structured. The detail of how the presidents addressed each topic is available in the **infographic** on the project's microsite.
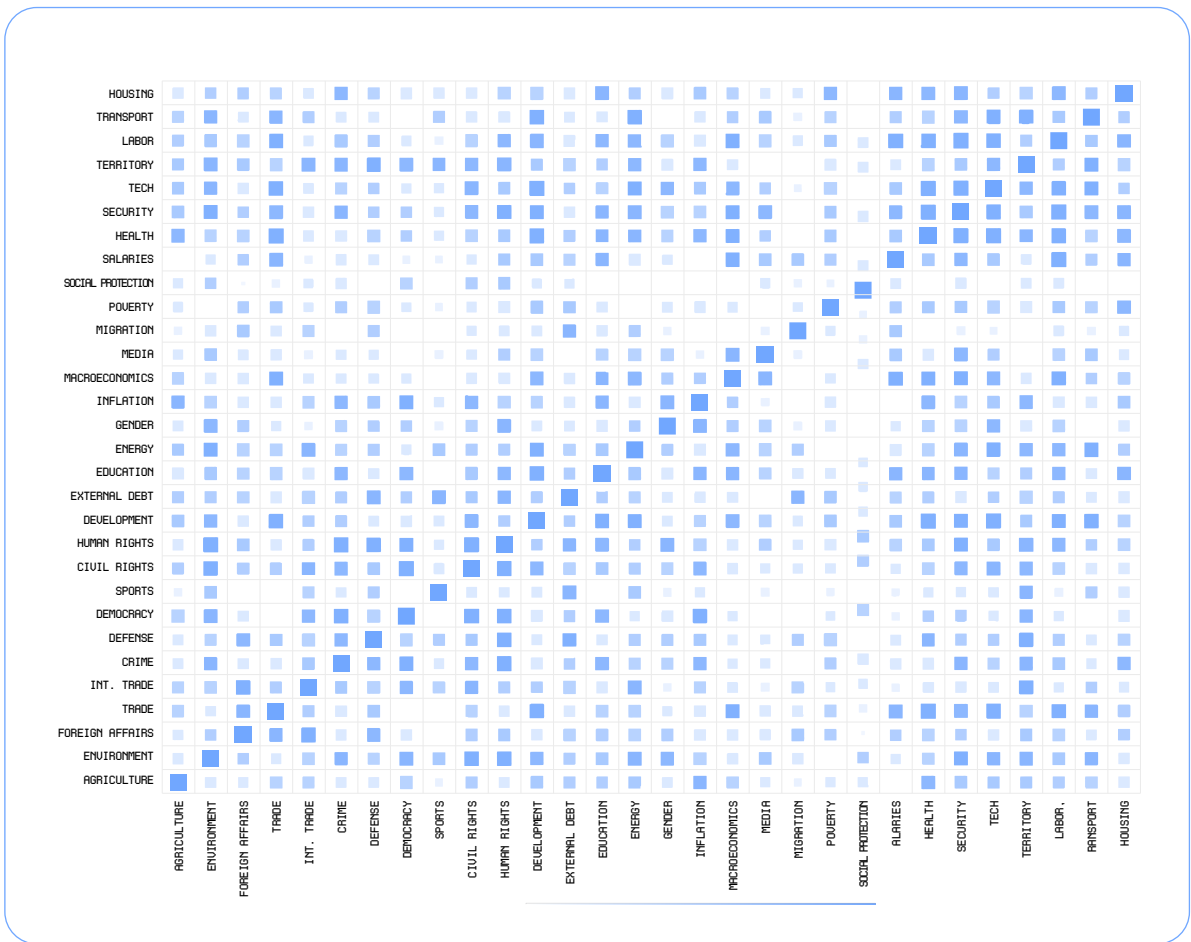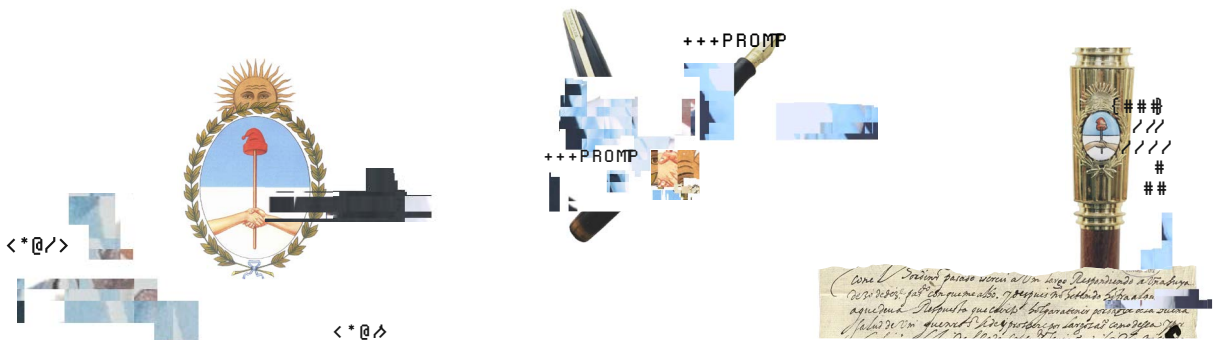

Figure 9: Correlation matrix of topics mentioned in the speeches

The most mentioned topics in the spee-ches in these 40 years of democracy were development, macroeconomics, and demo-cracy. However, their connotation or focus may vary over time and for each presi-dent. To learn what the 10 most mentioned topics were and how they relate to each other using an interactive tool, visit the infographic.

www.inteligenciaargentina.org →

In short, analyzing the topics and words of each speech gives us a deeper understanding of each president's vision, priorities, and concerns, as well as the context in which he or she governed. Therefore, it is essential to consider the temporal dimension when analyzing speeches and not only address them in isolation. Furthermore, it is important to stress that relevance is not limited to what is mentioned; it is also interesting to pay attention to topics that are not discussed (or those that are rarely mentioned) and question why they are not considered, as well as to identify the presidents who did discuss them.
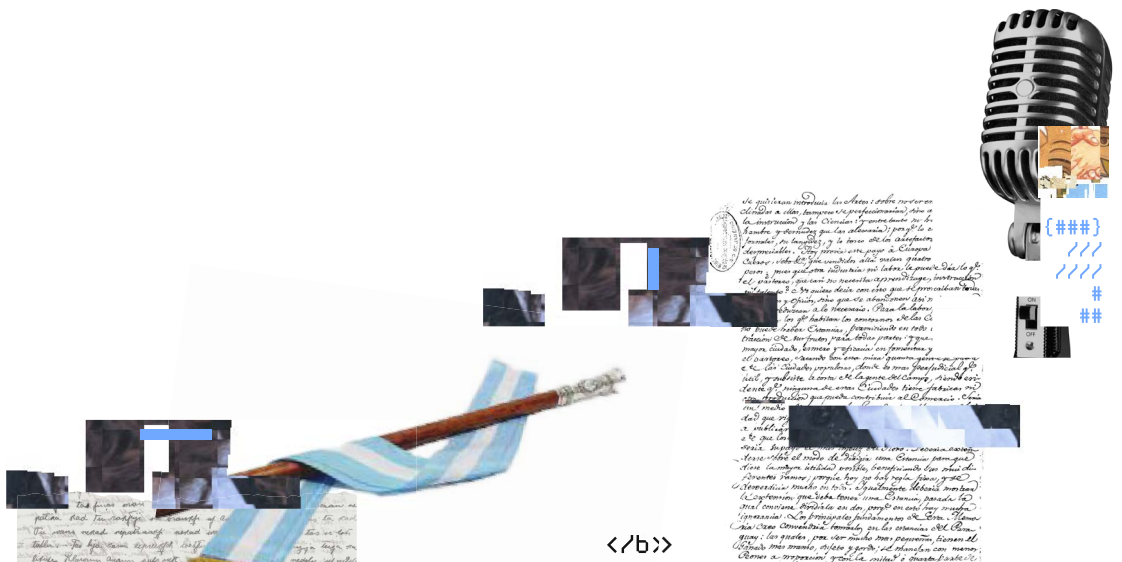
The unified speech of Argentine democracy in its 40 years

The evidence used to generate the unified speech was the same as the one used for the analysis of the topics. After editing and removing words, fragments, and special characters from the speeches, we proceeded to summarize them. It should be mentioned that most AI tools have limitations in terms of extension when processing texts. Therefore, this forced us to develop a strategy to, on the one hand, work with the texts in chunks and, on the other, keep their specificity and literality. In our third phase, the main themes of each presidential text were extracted through ChatGPT. Finally, the fourth and final step unified the speeches in ChatGPT using the main themes.

Steps to producing a unified speech

With speeches ready for analysis, we use the OpenAI API, that is, the version that enables us to program each of the steps that will produce the unified speech. This includes sending the speeches to the software and programming the prompts for the algorithm to process these texts, generate a new one, and identify the topics of each speech. However, it was necessary to adapt the inputs, since ChatGPT can only read and produce 4,000 characters at a time —including prompts—. In addition, ChatGPT cannot store information from previous iterations.

To overcome these problems, each of the speeches was summarized, dividing them first into 2,000-character sections when the text exceeded 4,000. ChatGPT was then asked to summarize each of these sections, with a limit of 400 characters per summary. Based on these results, ChatGPT followed the order to merge these summaries into a single 1,250-character text, and thus we obtained the abbreviated version of each presidential speech.

In the third stage, to obtain the main themes of each presidential message, ChatGPT was reminded that the summaries were of his authorship and given the following instruction: "What are the main themes of the speech?", with a response limit of 200 characters.

Finally, based on the main themes previously obtained, the final step of our work was to generate a sole unified speech. However, due to the character limitation, it is only possible to include and process the main themes of 10 speeches at the same time. Therefore, based on the main themes of 10 speeches chosen at random (up to a total of 52), ChatGPT created different partial speeches. Finally, the software took the 5 generated partial speeches and delivered the final unified
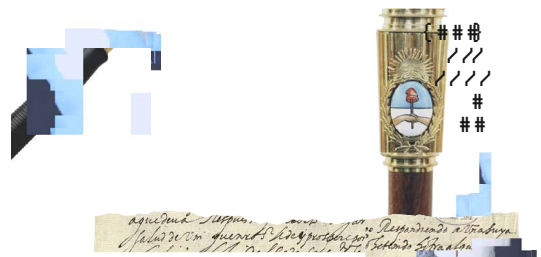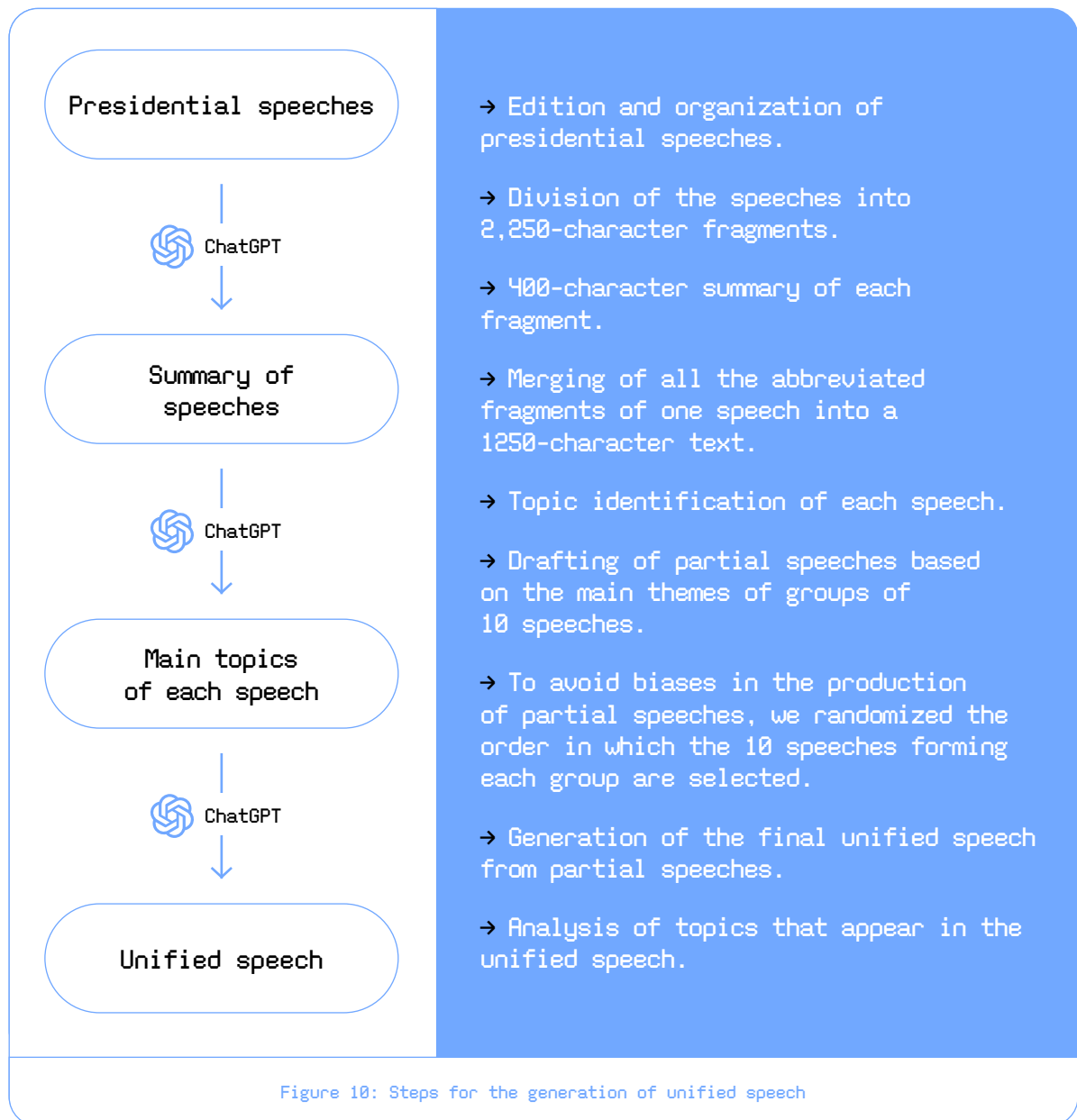
#//[###}

speech. Lastly, as a complementary analysis, we used ChatGPT to identify the main themes that appeared in the unified speech and compare them with the aforementioned topics used in the code-books as explained in the previous section.

In different tests, we noticed that, if we followed the chronologi-cal order of the inputs to feed the software and obtain the unified speech this biased the different results in two ways. On the one hand, it always seemed to reflect topics of a certain context, for example, it mentioned the pandemic or some very specific event of the state of affairs of a given speech. On the other hand, as previously mentioned, there are two presidents with a greater number of speeches due to their lengthy terms of office, which seemed to have more weight if we feed the software in chronological order. Unlike the previous section, where the difference in the number of speeches did not introduce biases to assess the relevance of certain topics in the texts throu-ghout the 40 years of democracy, this exercise of the unified speech tries to capture both the most mentioned main themes and those that, although they were not higher in the frequency of occurrence, indicate a diversity of concerns worth to be found. Consequently, we tried different procedures to **capture that intrinsic diversity, not necessarily aligned with how many times a theme appears, and this is a singular contribution of AI analysis.** Thus, also to obtain a timeless and balanced result among the different presidents. The solution was to randomize the order of the speeches with which the main themes are fed for ChatGPT to prevent biases in the production of partial texts.

{+++}

Presidential speeches

→ Edition and organization of presidential speeches.

→ Division of the speeches into 2,250-character fragments.

ChatGPT

→ 400-character summary of each fragment.

Summary of speeches

→ Merging of all the abbreviated fragments of one speech into a 1250-character text.

ChatGPT

→ Topic identification of each speech.

→ Drafting of partial speeches based on the main themes of groups of 10 speeches.

Main topics of each speech

→ To avoid biases in the production of partial speeches, we randomized the order in which the 10 speeches forming each group are selected.

ChatGPT

→ Generation of the final unified speech from partial speeches.

Unified speech

→ Analysis of topics that appear in the unified speech.

Figure 10: Steps for the generation of unified speech

Variations and robustness tests

+++PROMP

The results obtained by ChatGPT are not deterministic, but probabilistic. This implies that the results are slightly different with each iteration of the process. Therefore, several repetitions of the steps described above were performed and different versions of unified speeches were compared. In this analysis, we found that the different texts maintained a structure and focused on similar main themes. Furthermore, the results generated were consistent given the different specifications of the parameters configuring the ChatGPT text generation feature, such as temperature (how predictable or consistent text generation is to similar entries), frequency penalty (how much repetition it supports), and different prompts about the specificity of the requested response.

The speech ↓

Citizens of the Argentine Republic,

I address you today to discuss the issues that are fundamental to the development and progress of our country. All the presidents who preceded me have spoken about the need to improve the country's economy, the importance of maintaining economic stability and controlling inflation, and the need for a State that is present and promotes a sustainable development and a just society. These are topics that remain relevant today.

In addition, we must remember the importance of education, public health, and primary health care. Science, technology, and culture are crucial to the development of the country, and we must continue to invest in them.

Argentina's foreign policy is also an important issue. We must commit ourselves to multilateralism and the promotion of human rights and work together with other countries to achieve our goals.

We must also recall the importance of the self-esteem of the Argentine people and the need for institutional reconstruction to achieve sustained economic growth with equity and social inclusion.

Democracy and social justice are essential values that we must defend and promote. We must work to eradicate poverty and social exclusion and ensure access to education, health, security, and justice for all citizens.

It is important that we work to integrate the country into the region and the world, encouraging trade and international cooperation. We must adapt to changes and not be afraid of the financial, productive, and monetary globalization of the world.

In short, we must work together to improve the country's economy, foster equality, and social justice; invest in education, health, and culture; and commit ourselves to multilateralism and the promotion of human rights.

If we work together, we can build a better and more prosperous Argentina for all. Thank you!

What topics does the unified speech include?

The last exercise was a meta-analysis in which the AI analyzed its own production. We asked ChatGPT 4.0 to identify the main themes it could recognize from the unified speech it had generated previously. In particular, AI identified that this speech addresses topics such as **economics** and **economic stability; education, public health,** and **primary health care; science, technology,** and **culture; foreign policy** and **foreign affairs;** the **self-esteem of the Argentine people** and **institutional reconstruction; democracy** and **social justice; integration in the region and the world;** and, finally, **improvement objectives,** referring to goals to be achieved based on joint effort.

It should be remembered that, throughout this study, we worked in two different ways to identify the topics. First, 30 topics were identified and defined in all the speeches based on a glossary of associated words. This helped analyze the type of topic, the frequency of occurrences, its relationship with other topics, the similarities between presidents when mentioning them, etc. Second, to process the speeches and produce the unified speech, ChatGPT was asked to identify the main themes present in the presidential speeches and to use them to generate the speech in question.

But... how do the topics arising from the unified speech relate to the 30 predefined topics? The first observation is that ChatGPT was not instructed to use this list of topics, that is, the unified speech does not go through this filter. The topics that appeared were some of those mentioned by the presidents that the AI itself prioritized.

When we compare the list of topics that AI selected with the one that we had initially considered a priority, we observe the following:

→ Some topics in the unified speech match the predefined list of 30 topics. For example, topics such as **education, health, technology,** and **democracy** appear in the speech and had been defined in this way in the list of topics. Others such as **economics** and **economic stability** could fall under the **macroeconomics** topic of the list, while **foreign policy, foreign affairs**, and **integration in the region and the world** can be encompassed within **foreign affairs** and **international trade**.

→ There are new topics emerging in the unified speech that had not been defined in the list of topics, but mentioned by the presidents in their speeches, which is the case of the s**elf-esteem of the Argentine people, institutional reconstruction, social justice, and improvement objectives.**

→The topics in the unified speech are not necessarily the most mentioned by the presidents in their speeches, in other words, the AI's criterion of inclusion is not directly related to the

number of mentions. For example, the unified speech covers topics such as **public health, primary health care, self-esteem of the Argentine people,** and **institutional reconstruction**, even though these main themes have not been frequently mentioned by presidents and excludes topics such as **security** and **crime** that were most frequently mentioned.

In short, we can analyze the topics of the unified speech in two ways: through the predefined topics that we used to analyze the topics of the presidential speeches or through the main themes addressed by the presidents in their speeches that were the basis for producing the unified speech. As mentioned, the list includes only 30 topics and thus some themes of the unified speech may be left out of this list. Moreover, the topics of the unified speech do not necessarily relate to the importance or frequency with which the presidents mentioned them, since the decision to include or not include certain topics in the unified speech was made by the AI itself.

+++PROMP

{###}

{###}

<*@/>

What this exercise taught us: discussion and closure

In this analytical exercise, natural language models were used to process the speeches of the Argentine presidents during the most extensive democratic period in the history of the country, spanning from 1983 to 2023, to identify the shared priorities of the Argentine citizens and engage in dialogue to facilitate consensus and agreements based on these priorities. To achieve this goal, we carried out an exercise consisting of three parts: the analysis of the speeches, the exploration of the topics, and the generation of a unified speech. The results showed that the association between certain topics is frequent while the mention of certain topics would seem to depend, to a greater extent, on a temporal factor rather than on a partisan one, since this association is strongly linked to the socio-political context in which each president governed.

Additionally, the unified speech emerges as a construction of language that differs greatly from the sum of its parts, since it does not necessarily reflect the most mentioned topics of the analyzed speeches according to the glossary of topics. Although each generated speech was different, the strategy of randomizing the order of the analyzed speeches, with which the unified speech was created, resulted in the generated speech being very realistic and the general issues raised being perceived as timeless, that is, not too circumscribed to a certain period.

In the technical field, the results obtained are presented as an innovative application of ChatGPT in the analysis automation and the creation of text of a political nature. This exercise gives us the opportunity to understand the virtues and limitations of this tool, which in turn enables us to determine new possibilities of use. Furthermore, this exercise shows that the creation of a unified speech implies more than just the merging of different text fragments, as it requires identifying the common topics, combining them, and creating a voice that maintains the style of each president. In fact, the results show that ChatGPT can emulate human discourse in a reasonable way, which opens up possibilities for future projects that combine text analysis and creation.

In this context, our findings are particularly relevant to democratic governance. These have two comparative advantages that can turn this input into a tool with potential legitimacy for different audiences. First, AI is a tool free of political or partisan biases in text generation, as these texts are not associated with an issuer of a given political party, and have legitimacy, allowing the consensus found in this exercise to open a public conversation on common ground and to be a resource to promote democratic dialogue. Second, AI helped us process large volumes of text and find different indicators that characterize presidential speeches of different eras and parties. As a result, cross-cutting themes of consensus transcending party affiliations were identified.

In turn, thanks to this analysis, we were able to understand certain opportunities and limitations of AI tools applied to the political context. As for the opportunities, we can say that it enables the rapid and systematic analysis of large volumes of texts and topics, to see how they relate to each other, the frequency with which they are mentioned, and the connotation given to them in each period. As regards limitations, we found that ChatGPT tends to resort to typical contributions and faces the potential presence of biases in the results, due to the nature of the data with which it is trained. In addition, this tool cannot be modified and audited (it is not open source) and has high access costs (both monetary and technical). However, it should be noted that these are constantly evolving tools and thus the various restrictions could disappear or be reduced soon.

To sum up, this exercise helped us analyze the topics of the presidential speeches of Argentine democracy on its 40th anniversary, as well as generate a speech that reflects the values shared by the different presidents of this period. Today, this was done through AI, in the past, it was done through the organization of dialogue tables, but there is no doubt that, with different tools, UNDP aspires to contribute to consensus-building and democratic dialogue, which is essential for social cohesion and inclusion through development policies, and this is a new effort in this regard.

<*@/>

{+++}

DEMOCRACIA
40 AÑOS↑

{###}
////
#
##
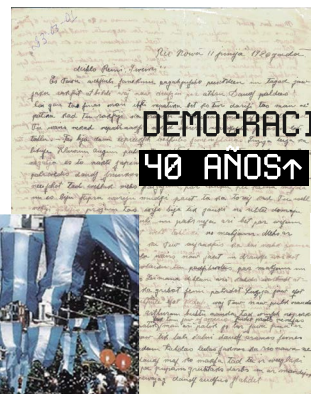
References

→ Abercrombie, G., Nanni, F., Batista-Navarro, R., & Ponzetto, S. P. (2019). Policy preference detection in parliamentary debate motions. *Association for Computational Linguistics*.

→ Alemán, E., Micozzi, J. P., & Vallejo Vera, S. (2023). Congressional Committees, Electoral Connections, and Legislative Speech. *Political Research Quarterly*, 76(2), 994-1011.

→ Aruguete, N., & Calvo, E. (2020). s Coronavirus en Argentina: Polarización partidaria, encuadres mediáticos y temor al riesgo. *Revista Saap*, 14(2), 280-310.

→ Bercholc, J. O., & Bercholc, D. (2013). El concepto de calidad institucional en los discursos presidenciales en la Argentina. *Revista Interdisciplinar do Direito-Faculdade de Direito de Valença*, 10(1).

→ Binder, S. (2000). Going Nowhere: A Grid-locked Congress. *The Brookings Review*, 18 (1), 16-19.

→ Bonikowski, B., Luo, Y., & Stuhler, O. (2022). Politics as Usual? Measuring Populism, Nationalism, and Authoritarianism in US Presidential Campaigns (1952–2020) with Neural Language Models. *Sociological Methods & Research*, 51(4), 1721-1787.

→ Burton, J. (2023). Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. *Technology in Society*, 75, 102262. https://doi.org/10.1016/j.techsoc.2023.102262

→ Casal Bértoa, F. (2019). Causas y consecuencias de la polarización: ¿Qué es lo que sabemos? *Cuadernos de Pensamiento Político*, 64, 5–16. https://www.jstor.org/stable/26822524

→ Casal Bértoa, F. & Enyedi, Z. (2019). Party System Closure: Alliances and Innovations between 1848 and 2018. Manuscrito para la editorial Oxford University Press.

→ Gu, Y., & Wang, Z. (2022). Income Inequality and Global Political Polarization: The Economic Origin of Political Polarization in the World. *Journal of Chinese political science*, 27(2), 375–398. https://doi.org/10.1007/s11366-021-09772-1

→ He, P., Liu, X., Gao, J., & Chen, W. (2020). Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv*:2006.03654.

→ Kaneko, M., & Bollegala, D. (2022). Unmasking the mask–evaluating social biases in masked language models. *In Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, No. 11, pp. 11954-11962).

→ Moscovich, L. (1 de septiembre de 2023). IA: Inteligencia Argentina. Una conversación sobre la inteligencia artificial: Episodio I. Todos pueden cocinar ¡y usar IA! PNUD Argentina. Disponible en https://www.undp.org/es/argentina/blog/ia-inteligencia-argentina-una-conversacion-sobre-la-inteligencia-artificial

→ Liminga, A. (2022). Politicians as communicators of delegitimizing criticism towards epistemic authorities: A study of political hostility towards news media and science in Sweden (Dissertation). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-481116

→ Levy Yeyati, E., Moscovich, L., & Abuin, C. (2020). Leader over policy? The scope of elite influence on policy preferences. Political Communication, 37(3), 398-422.

→ Linz, J. J. (1978). The Breakdown of Democratic Regimes. *John Hopkins University Press*

→ Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv*:1907.11692.

→ Odmundsen, M., Bor, A., Vahlstrup, P., Bechmann, A., & Petersen, M. (2021). Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. American Political Science Review, 115(3), 999-1015. doi:10.1017/S0003055421000290

→ Qureshi, I., Bhatt, B., Gupta, S. & Tiwari, A. (2021). Causes, Symptoms and Consequences of Social Media Induced Polarization (SMIP).

→ Sartori, G. (1976). Parties and Party Systems: A framework for analysis. *Cambridge University Press*

→ Singer, M. (2016). Elite Polarization and the Electoral Impact of Left-right Placements: Evidence from Latin America, 1995-2009. *Latin American Research Review*, 51(2), 174-194. doi:10.1353/lar.2016.0022

→ Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... & Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature*.

→ Tunstall, L., von Werra, L. & Wolf, T. (2022). Natural Language Processing with Transformers. O'Reilly Media, Inc.

→ Van Bavel, J. J., Rathje, S., Harris, E., Robertson, C., & Sternisko, A. (2021). How social media shapes polarization. *Trends in cognitive sciences*, 25(11), 913–916.

Annex:

List of topics with their keywords

1. **Macroeconomics:** We divided the topic of macroeconomics into four subtopics of particular interest:

    a. **External debt:** The speech covers topics related to external debt, including:

        i. International financial institutions (e.g., International Monetary Fund, World Bank, Paris Club).

        ii. Concepts related to debt repayment, including conditionalities, shock policies, austerity measures, stabilization plans, debt restructuring, taxes.

        iii. Country risk, debt default.

        iv. Bondholders, vulture funds.

        v. Agricultural export sector, entry and exit of dollars.

    b. **Poverty:** The speech covers topics related to poverty, including:

        i. Poverty threshold, extreme poverty, indigence.

        ii. Unsatisfied Basic Needs (UBN).

        iii. Inequality, income distribution.

        iv. Vulnerability, social conditions.

        v. Multidimensional poverty.

    c. **Inflation:** The speech covers topics related to inflation, including:

        i. Hyperinflation.

        ii. Currency devaluation.

        iii. Price control, price agreements, price makers.

        iv. Monetary policy, the central bank's reserves.

    d. **Macroeconomics (general):** The speech covers topics related to domestic macroeconomic policy, including:

        i. Interest rates, cost of living, prices.

        ii. Unemployment rates, underemployment, impact of unemployment.

        iii. Industry regulation, industry subsidies, industry profits and taxes.

2. **Civil rights:** We divided the topic of civil rights into three subtopics of particular interest:

    a. **Democracy:** The speech covers topics related to democracy, including:

        i. Electoral integrity, elections, electoral transparency, electoral fraud.

        ii. Clientelism, patronage, and electoral favors.

        iii. Right to vote.

        iv. Democracy, democratic quality.

        v. Military dictatorship and authoritarianism.

        vi. Electoral fraud.

        vii. Electoral reforms, re-election, gender quota, and Quota Act.

        viii. Campaign and Political Parties Financing Act.

   b. **Human rights:** The speech covers human rights topics, including:

     i. Human rights.

     ii. Crimes against humanity.

     iii. The disappeared.

     iv. International convention on human rights, crimes against humanity, international courts.

     v. Military uprising and military insurrectionist popularly known Carapintadas [Painted Faces].

     vi. Pardon, Pardon Laws.

     vii. State terrorism, subversion, political violence.

   c. **Civil rights (general):** The speech covers topics related to civil rights and rights of minorities and different populations:

     i. (Laws against) Discrimination against ethnic or racial minorities.

     ii. (Laws against) Discrimination based on gender or sexual orientation, sexual diversity, gender identity, gender expression, and body diversity or discrimination against transgender people, trans people, and transvestites.

     iii. (Laws against) Age discrimination.

     iv. (Laws against) Disability discrimination.

     v. Sizes Act.

     vi. Freedom of expression.

     vii. Right to privacy.

     viii. Divorce.

3. **Health:** The speech covers topics related to public health and obras sociales [health insurance plans]:

   a. Public health system, obras sociales system.

   b. Health coverage (public and obras sociales).

   c. Medicines, clinics, pharmacies, and pharmaceutical companies.

   d. Hospitals, ambulances, laboratories, primary health care.

   e. Doctors, nurses.

   f. Health and disease prevention programs.

   g. Mental health.

   h. Use and regulation of tobacco and alcohol, addictions.

   i. Research and development of medicines.

   j. Access to health, quality of care.

   k. New technologies and health.

   l. Pandemic, endemic.

   m. Sexual and reproductive health, application of voluntary termination of pregnancy, contraceptives, abortion, emergency contraceptives.

   n. Menstrual health, menstrual hygiene management.

   o. Sexually transmitted infections, HIV.

4. **Agriculture:** The speech covers topics related to agriculture and commodities:

   a. Regulation on the import and export of agricultural products.

   b. Subsidies, withholding taxes, agricultural investment and taxes.

   c. Export quotas.

    d.   Food regulation (general).

    e.   (Control of) pests.

    f.   Fishing and aquaculture.

    g.   Arable land.

    h.   Irrigation systems.

    i.   Productive/agricultural frontier.

    j.   Family farming, subsistence, landless farmers.

    k.   No-till farming.

    l.   The countryside, elite, power sectors, soybean, soybean producers.

    m.   Transgenics, biotechnology.

    n.   Child labor in rural areas.

5.   **Labor:** We divided the labor topic into two subtopics of particular interest:

    a.   **Labor (general):** The speech covers topics related to labor laws, workers, pensions:

        i.   Trade unions.

        ii.   Topics related to teachers' unions, police unions, doctors' unions, etc., should also be categorized as labor.

        iii.   Pensions, contributions.

        iv.   Labor Act, labor market flexibility, precarious work.

        v.   Public sector workers.

        vi.   Domestic work.

        vii.   Labor informality, unregistered employment.

        viii.   Labor market.

        ix.   Employment quotas, trans employment quota.

        x.   Child labor (general).

        xi.   Illegal forced labor.

    b.   **Salaries:** The speech covers topics related to salaries:

        i.   Regulation on employment and salaries.

        ii.   Negotiations between workers and employers, collective bargaining.

        iii.   Minimum wage, wage gaps (related to gender, education, migration, etc.).

6.   **Social protest:** The speech covers issues related to protest and social conflict:

    a.   Strikes and labor union protests.

    b.   Repression in strikes and labor union protests.

    c.   Pickets.

    d.   Social mobilization.

7.   **Education:** It refers to education, the Education Act, or institutions related to education:

    a.   Laws, regulations, and funding of schools, institutions, and universities (public and private).

    b.   Access to education.

    c.   Literacy programs, bilingual education, special education.

    d.   Vocational training programs.

    e.   Deaths and recognition of key figures in education, culture, or science.

    f.   Construction of schools, institutions, and universities.

g.   Teacher training and teacher evaluation.

h.   Performance, student results, and education quality.

8.   **Environment:** It refers to topics about the condition, preservation, and regulation of the environment.

a.   Drinking water, water pollution, water reserves, etc.

b.   Waste management (e.g., chemicals), urban solid waste.

c.   Air pollution, carbon emissions.

d.   Recycling.

e.   Preservation of native forests (Forest Act), jungles, natural resources, animals, water, glaciers, etc.

f.   Preservation of the environment in general.

g.   Climate change (adaptation and mitigation), climate crisis, climate action.

h.   Natural disasters and risk management.

9.   **Energy:** It refers to energy topics, including oil-related topics:

a.   Investment and regulation of nuclear energy.

b.   (Regulation of) Electricity, including electricity generation.

c.   (Regulation of) Oil, coal, and natural gas, including their extraction and transportation.

d.   Unconventional reservoirs, Vaca Muerta ("Dead Cow", shale gas deposit and shale oil reserve)

e.   Renewable energies or new sources of energy.

f.   Energy transition, lithium, hydrogen.

g.   Energy infrastructure.

10. **Migration:** It refers to migration-related topics:

a.   Migrants to other countries, migrants from other countries, and internal migration.

b.   Refugees.

c.   Citizenship and naturalization.

d.   Emigration.

11. **Indigenous peoples:** The speech covers topics related to indigenous peoples.

12. **Transport:** It refers to public transport, roads, etc.:

a.   Construction, maintenance, and financing of public transport and mass transit (buses, trains, trams, etc.).

b.   Construction and maintenance of roads and tracks.

c.   Regulation and investment in maritime and air transport.

d.   Electric transportation in cities.

e.   Transport subsidies.

13. **Law Enforcement and Crime:** It refers to laws, the legal system, and crimes (excludes laws and regulations on public institutions and the branches of government):

   a.   (Laws on) white collar crimes (e.g., corruption, organized crime, fraud, cybercrimes, money laundering).

   b.   (Laws on) illegal drug crimes (e.g., drug trafficking, mafias, drug-related violence).

   c.   Supreme Court and court management.

   d.   Prison management.

   e.   Reform of the Civil Code.

   f.   Penal reform. Impunity.

   g.   Topics related to family and child violence.

   h.   Topics related to the regulation and financing of police forces or national security.

   i.   Topics related to police violence (e.g., trigger-happy cases).

   j.   Internal Security Act.

14. **Social security and social protection:** It refers to topics generally related to social issues and inclusion:

   i.   Assistance for low-income families (e.g., conditional cash transfer programs, workfare programs).

   ii.   Assistance for the elderly, people with disabilities, children. Disability/PAMI [Comprehensive Medical Attention Program].

   iii.   Social security institutions (e.g. ANSES [National Social Security Administration]).

   iv.   Conditional cash transfers, social handouts, social programs, pension reform, privatization, AFJP [Retirement and Pension Fund Administrators], universal basic income, retirements and pensions, domestic work, private-home domestic workers.

   v.   *Jefes o Jefas del Hogar* [Heads of Household] Program, Argentina Trabaja [Argentina Works] program, Caja PAN (National Food Plan).

15. **Housing:** It refers to housing and urban development:

   a.   Regulations and laws for tenants, landlords, and properties.

   b.   Rural and urban development (e.g. building permits).

   c.   Popular, social, or public housing.

   d.   (Decrease in) homeless populations.

   e.   Sustainable housing.

   f.   Rental Act, mortgage loans.

   g.   Socio-urban integration.

   h.   Home ownership.

16. **Trade (domestic):** It refers to topics affecting trade and domestic markets:

   a.   (Regulation of) financial institutions and the banking system.

   b.   (Regulation of) electronic commerce and fintech.

   c.   (Regulation of) insurers.

   d.   Bankruptcy.

   e.   Small and medium-sized enterprises.

    f. Patents, intellectual property rights, royalty, and copyright.

    g. Gambling regulation.

17. **Sports:** It refers to sporting events, sports personalities, or their regulation.

    a. Olympic Games.

    b. Football World Cup.

18. **Defense:** It refers to national security (i.e., national army):

    a. Financing and regulation of the national army.

    b. National Defense Act.

    c. Military intelligence, espionage.

    d. War and armed conflicts.

    e. Nuclear weapons.

    f. Military assistance, military reserves, military service.

    g. Military courts (martial court) and military abuse.

    h. Malvinas and veterans.

    i. Border areas, border controls, and boundary demarcation.

19. **Technology:** It refers to technology, science, and communications:

    a. Regulation and investment in telecommunications.

    b. (Investment in) Science and technology.

    c. Technology industry, Internet regulation, and cybersecurity.

    d. Research and development in computing science.

    e. Basic science, hard science, scientific careers, PhD scholarships, software, CONICET [National Scientific and Technical Research Council].

    f. Scientific migration and brain drain.

    g. Satellites.

    h. Access to space.

    i. Development of telecommunications infrastructure.

    j. Development of digital infrastructure/Internet.

    k. Digital inclusion, Internet access, digital training, and digital literacy.

20. **Media:** The speech covers topics related to media:

    a. Media Act.

    b. Radio and television transmission regulation.

    c. Monopoly, media concentration, media corporations.

21. **International trade:** The speech covers topics related to:

    a. Free trade agreements, investment treaties.

    b. International and regional trade treaties: Pacific Alliance, MERCOSUR, etc.

    c. Export regulation, promotion, and control.

    d. Foreign investment (private and public).

    e. Duties, imports, and customs.

    f. Balance of trade.

    g. Exchange rates.

    h. International or multilateral financing: IMF, WB, IDB, CAF, etc.

22. **Foreign affairs:** The speech covers topics related to international issues and aid:

    a. Regulation of maritime exploitation and international waters.

    b. Foreign policy in general, international alignments.

    c. International agreements and treaties, compliance with MEAs.

    d. International and regional treaties: MERCOSUR, NAFTA, CELAC, UNASUR, ALBA, UN, etc.

    e. International organizations (e.g., United Nations, OAS, UNESCO, Inter-American Court of Justice, etc.).

    f. Deaths and recognition of international political figures.

23. **Government operations:** It refers to the following topics:

    a. Relations between government institutions.

    b. Bureaucracy and bureaucrats.

    c. Postal service (i. e., national post) and stamps.

    d. Nominations for public office (e.g., judges, comptrollers, superintendents, etc.).

    e. Relations between branches of government (i.e., legislative, executive, judicial).

    f. Regulation of political campaigns, campaign financing, polling places.

    g. Censuses and statistics: INDEC [National Institute of Statistics and Censuses], transparency.

    h. Privatizations, federalism, provinces, municipalities, nationalizations, decentralization, centralization.

    i. Modernization of the state.

    j. Tax Sharing Regime (*co-participation*).

24. **Territorial demarcations:** It refers to the creation of new territorial demarcations and public lands:

    a. Creation of new local governments, municipalities, provinces, cities, towns, administrative areas, etc.

    b. Creation and naming of national parks.

25. **Conjuncture:** It refers to policies responding to immediate circumstances:

    a. Mediatized political events.

    b. Political accusations.

    c. Deaths and recognition of political figures.

    d. Scandals and impeachments.

    e. Corruption of political figures.

    f. Internal politics (refers to conflicts).

    g. Alliances or political unity.

26. **Development:** The speech covers topics related to development such as:
    a. Economic development.
    b. Human development.
    c. Sustainable development goals (SDGs), 2030 agenda, millennium goals.

27. **Gender:** The speech covers topics related to gender such as:
    a. Feminism.
    b. Family models, heads of household, economic autonomy of women.
    c. Gender-based violence, political violence against women, digital gender-based violence/digital violence against women.
    d. Topics of sexual diversity, transgender people, LGBTIQ+.
    e. Care, Comprehensive Care System.
    f. Masculinities.
    g. Transgender and non-binary childhood and adolescence.
    h. Gender equality, gender gaps, gender inequalities.
    i. Gaps and access to the labor market, land, salaries, health, political spaces, etc.

Accelerator lab <co_lab>
UNDP Argentina
IA Argentine Intelligence

TEAM   UNDP ARGENTINA

CLAUDIO TOMASI
Resident Representative at UNDP Argentina

VALENTÍN GONZÁLEZ LEÓN
Assistant Representative at UNDP Argentina

ACCELERATOR LAB (CO_LAB)

LORENA MOSCOVICH
Head of Experimentation

MARÍA VERÓNICA MORENO
Head of Solutions Mapping

ABOUT THIS REPORT