



Derecha fragmentada y un enemigo compartido

Análisis textual longitudinal de la constación discursiva en las elecciones presidencial de 2025 en Chile

Tesis para optar al grado de
Magíster en Ciencia de Datos

Matías Deneken

Centro de Datos e Inteligencia Artificial - Universidad de Concepción

Profesores Guías:

Dr. Carlos Navarrete

Dr. Marcela Parada

Universidad de Concepción

Enero 2026

Agradecimientos

Esta tesina no habría sido posible sin el apoyo económico de la beca de exención de arancel por el Premio Universidad de Concepción, según consta el decreto DIRPOST-021-24.

Agradezco a mis compañeras y compañeros de generación con los cuáles compartí durante este proceso. En particular a Javiera Baeza y Florencia Pampaloni, con quién constituí un excelente equipo de trabajo desde la distancia. Sin este trabajo colaborativo la experiencia el MACI habría sido significativamente más compleja.

De la misma forma, agradezco a nivel general mi experiencia en el MACI. La pasantía realizada en la Universidad de Harvard en 2024 brindó una oportunidad y riqueza inigualable para poder conocer un ecosistema que -hasta entonces- me era ajeno. En particular, dar las gracias a mis profesores guías, Marcela Parada y Carlos Navarrete, quienes me acompañaron en este proceso formativo.

Por último, darle siempre las gracias a mi familia por estimularla a afrontar nuevos desafíos, incluso cuándo la energía es escasa.

Sin duda esta tesis combina mis intereses de la sociología, a nivel de pregrado y de magíster, y la ciencia de datos, que me ha conllevado a esta incipiente área de las *ciencias sociales computacionales*. Con sus bemoles, espero que esta tesis sirva cómo mi puntapié inicial para poder enfrentar nuevos desafíos que combine los métodos estadísticos avanzado con la revolución de la Inteligencia Artificial que hoy estamos viviendo. El tiempo nos dirá.

Presentación

Contenido de presentación/introducción general

Índice General

Agradecimientos	ii
Presentación	iii
1 Introducción	1
1.1 Justificación y relevancia del estudio	2
1.2 Preguntas de investigación	3
1.3 Objetivos de la investigación	4
1.3.1 Objetivo general	4
1.3.2 Objetivos específicos	4
2 Presentación del Problema	5
2.1 El problema: competencia política digital en un sistema fragmentado	5
2.2 Acercamiento: actores, dimensiones y niveles de análisis	6
2.2.1 Los actores: tres derechas y un adversario	6
2.2.2 Las dimensiones analíticas	7
2.2.3 Los niveles de análisis	8
2.3 Descripción de la propuesta analítica	8
2.3.1 Enfoque general	8
2.3.2 Arquitectura del análisis	9
2.3.3 Originalidad de la propuesta	10
2.4 Producto mínimo viable (MVP) y objetivos operacionales	10
2.4.1 Definición del MVP	10
2.4.2 Objetivos operacionales y entregables	11
2.4.3 Alcances y limitaciones anticipadas	11

2.5	Estrategia de validación	12
2.5.1	Validación interna	12
2.5.2	Validación externa	12
2.5.3	Criterios de éxito	13
3	Marco teórico	14
3.1	Comunicación política en entornos digitales	14
3.2	Polarización afectiva	14
3.3	Teoría del <i>framing</i>	15
3.4	Emociones políticas	15
3.5	Construcción de adversarios	16
3.6	Derechas contemporáneas	16
3.7	Reddit como espacio de deliberación política	17
3.8	Modelo analítico integrado	17
4	Diseño metodológico	18
4.1	Diseño de investigación	18
4.2	Reddit como plataforma de análisis	18
4.3	Recolección de datos mediante la API de Reddit	19
4.3.1	Procedimiento técnico	19
4.3.2	Período de recolección y fases del ciclo electoral	20
4.3.3	Estrategia de extracción incremental	20
4.3.4	Contenido extraído y filtrado temático	20
4.4	Consideraciones éticas	21
4.5	Técnicas de análisis	22
4.5.1	Análisis de sentimiento y emociones	22
4.5.2	Clasificación de marcos interpretativos (<i>framing</i>)	22
4.5.3	Análisis de redes de interacción	23
4.5.4	Análisis temporal	23
4.6	Corpus final y pipeline de procesamiento	23
5	Depuración de Datos	25

6 Exploración de Datos **26**

7 Modelado de Datos **27**

7.1 Clustering K-Means 27

7.2 Clustering DBSCAN 27

7.3 Clustering Jerárquico 27

7.4 Determinación del modelo más óptimo 27

8 Interpretación de Resultados **28**

8.1 Interpretación de componentes principales 28

8.2 Interpretación de clusters 28

8.3 Validación de los resultados 28

9 Conclusión **29**

10 Referencias **30**

Índice de Figuras

2.1	Pipeline analítico: de la recolección del corpus a los hallazgos sustantivos.	9
4.1	Flujo de recolección y procesamiento del corpus.	24

Índice de Tablas

2.1	Candidaturas incluidas en el análisis y su caracterización ideológica.	6
2.2	Niveles de análisis utilizados en la investigación.	8
2.3	Objetivos operacionales, entregables e indicadores de cumplimiento del MVP.	11
4.2	Candidaturas objetivo y patrones regex para su detección en el corpus.	21

1 Introducción

Las elecciones presidenciales chilenas de 2025 se desarrollan en un contexto de profunda reconfiguración del sistema de partidos y de transformación de las condiciones bajo las cuales se produce la comunicación política. Tras el ciclo de movilizaciones iniciado en octubre de 2019, los procesos constituyentes fallidos de 2022 y 2023, y la creciente fragmentación de las coaliciones tradicionales, el escenario electoral chileno exhibe niveles de incertidumbre, volatilidad y polarización inéditos en la historia reciente del país.

En este contexto, las plataformas digitales han adquirido un protagonismo sin precedentes como espacios de debate político, formación de opinión y articulación de identidades colectivas. Lo que ocurre en redes sociales ya no constituye un reflejo secundario de dinámicas políticas que se dirimen en otros ámbitos, sino un terreno constitutivo donde se disputan los términos del conflicto, se movilizan emociones políticas y se construyen las fronteras simbólicas que delimitan aliados y adversarios.

Una de las transformaciones más significativas del período reciente es la reconfiguración del campo de las derechas chilenas. Donde antes existía una coalición relativamente cohesionada —heredera de la transición democrática y articulada en torno a partidos como la UDI y Renovación Nacional—, hoy convergen y compiten al menos tres corrientes distinguibles: una derecha tradicional que busca renovar su apelación desde la moderación y la competencia técnica; una derecha radical-conservadora que articula posiciones morales restrictivas con retórica antielitista; y una derecha libertaria de perfil disruptivo que desafía las convenciones comunicativas establecidas. La competencia entre estas corrientes, y su eventual convergencia frente a adversarios comunes, constituye uno de los fenómenos más relevantes —y menos estudiados— de la política chilena contemporánea.

1.1 Justificación y relevancia del estudio

En la actualidad las encuestas, en general, y las de opinión, en particular, han sido parte del escrutinio público. Este tipo de medición han sido acusadas de fundar un datos sesgados y , en última instancia, serviles a las tendencias ideológicas a las que se encuentran posicionadas las que la financia o la consultora encargada.

Indepiente de esos juicios, las encuestas enfrentar problemas más allá. El desafío principal es que las tasas de respuestas cada vez son más bajas y los costos por la realización de la encuesta al individuo es cada vez más caro. La era digital brindan oportunidades nuevas dado la emergencia en del “Big Data”. Es decir, la presencia de datos masivos.

Esta investigación se justifica en función de tres órdenes de consideraciones: empíricas, teóricas y metodológicas.

Desde el punto de vista **empírico**, el ciclo electoral chileno de 2025 ofrece una oportunidad única para observar en tiempo real las dinámicas de competencia intra-derecha en un sistema político en transformación. La coexistencia de tres candidaturas de derecha con perfiles ideológicos y estrategias comunicativas diferenciadas —Evelyn Matthei, José Antonio Kast y Johannes Kaiser— permite analizar comparativamente cómo distintas corrientes construyen sus identidades, movilizan usuarios y se posicionan frente a adversarios internos y externos. La presencia de una candidatura de izquierda asociada al Partido Comunista —Jeannette Jara— añade una dimensión adicional al análisis, habilitando el estudio de los procesos de convergencia discursiva que emergen cuando actores previamente enfrentados deben articular un frente común.

Desde el punto de vista **teórico**, la investigación contribuye a debates contemporáneos sobre comunicación política digital, polarización y construcción de antagonismos. Si bien existe abundante literatura sobre el rol de las redes sociales en procesos electorales, la mayor parte de esta investigación se ha concentrado en contextos del Norte Global, particularmente Estados Unidos y Europa Occidental. Los estudios sobre América Latina, aunque crecientes, han tendido a focalizarse en casos de alta visibilidad internacional —Brasil con Bolsonaro, Argentina con Milei— sin atender suficientemente a las variaciones nacionales y a las dinámicas de competencia al interior de familias

ideológicas. Esta tesis busca contribuir a llenar ese vacío, ofreciendo un análisis detallado de las especificidades del caso chileno.

Desde el punto de vista **metodológico**, la investigación propone una aproximación que integra técnicas de ciencia social computacional con análisis interpretativo del discurso político. Esta combinación permite procesar volúmenes de datos que excederían las capacidades del análisis cualitativo tradicional, sin sacrificar la atención a las dimensiones semánticas y pragmáticas que resultan cruciales para comprender la comunicación política. La elección de Reddit como campo de observación —una plataforma subexplorada en el contexto latinoamericano— ofrece además la posibilidad de acceder a conversaciones políticas con características estructurales distintivas que enriquecen el análisis.

1.2 Preguntas de investigación

La pregunta central que orienta esta investigación es:

¿Cómo se configura y evoluciona la conversación política en Reddit durante la campaña presidencial chilena de 2025?

Esta pregunta general se desagrega en las siguientes interrogantes específicas:

1. ¿Qué patrones de encuadre (*framing*), movilización emocional y construcción de adversarios estructuran el debate político en los subreddits chilenos durante el período electoral?
2. ¿Cómo varían temporalmente las estrategias discursivas de ataque, defensa y demarcación de fronteras identitarias a lo largo de las distintas fases de la campaña?
3. ¿Qué dinámicas de competencia, diferenciación y antagonismo se observan entre las tres corrientes de derecha (tradicional, radical-conservadora y libertaria) durante la primera vuelta?
4. ¿Cómo se produce la convergencia discursiva en torno al Partido Comunista como adversario común durante la segunda vuelta, y qué tensiones persisten bajo esta convergencia aparente?

1.3 Objetivos de la investigación

1.3.1 Objetivo general

Analizar la configuración y evolución de la conversación política en Reddit durante la campaña presidencial chilena de 2025, identificando los patrones de encuadre, emoción y construcción de adversarios que estructuran el debate entre las distintas corrientes ideológicas.

1.3.2 Objetivos específicos

1. Caracterizar los marcos interpretativos (*frames*) predominantes en el discurso de cada corriente ideológica y su variación a lo largo del ciclo electoral.
2. Identificar los patrones de movilización emocional asociados a cada candidatura, distinguiendo emociones de valencia positiva, negativa y las combinaciones que configuran repertorios afectivos específicos
3. Analizar las estrategias de construcción de adversarios, atendiendo tanto a la definición de enemigos externos (inter-bloque) como a las dinámicas de diferenciación interna (intra-bloque).
4. Examinar la evolución temporal del debate, identificando puntos de inflexión, eventos catalizadores y patrones de activación discursiva asociados a momentos específicos de la campaña.

2 Presentación del Problema

2.1 El problema: competencia política digital en un sistema fragmentado

El sistema de partidos chileno atraviesa una crisis de representación sin precedentes desde el retorno a la democracia. Los indicadores son contundentes: la identificación partidaria ha caído sostenidamente desde 2010, la volatilidad electoral alcanzó máximos históricos en los ciclos 2021-2023, y la fragmentación parlamentaria dificulta la formación de coaliciones estables (PNUD, 2024; Luna, 2023). En este contexto de desalineamiento, las plataformas digitales han emergido como espacios donde se disputa la reconfiguración de las identidades políticas —particularmente en el campo de las derechas, donde coexisten proyectos en competencia que aún no han estabilizado sus fronteras.

El **problema central** que aborda esta investigación puede formularse así:

En un contexto de fragmentación partidaria y reconfiguración ideológica, ¿cómo se estructura la competencia política digital entre actores que comparten un espacio ideológico pero compiten por diferenciarse? ¿Qué estrategias discursivas despliegan? ¿Bajo qué condiciones convergen o divergen?

Este problema es relevante por tres razones:

1. **Vacío empírico:** Mientras abundan los estudios sobre polarización izquierda-derecha, la competencia *intra-bloque* —particularmente entre derechas— permanece subexplorada en América Latina.

2. **Transformación en curso:** Las derechas chilenas están en proceso de reconfiguración activa; estudiar este proceso mientras ocurre permite capturar dinámicas que análisis retrospectivos pierden.
3. **Implicancias democráticas:** Comprender cómo se construyen adversarios y fronteras identitarias en espacios digitales resulta crucial para evaluar la calidad del debate público y anticipar escenarios de polarización.

2.2 Acercamiento: actores, dimensiones y niveles de análisis

2.2.1 Los actores: tres derechas y un adversario

La investigación se focaliza en cuatro candidaturas que representan posiciones distintivas en el espectro político chileno:

Table 2.1: Candidaturas incluidas en el análisis y su caracterización ideológica.

Actor	Posición	Caracterización sintética	Base social
Evelyn Matthei	Derecha tradicional	Pragmatismo, competencia técnica, moderación discursiva	Élites empresariales, sectores medios-altos
José Antonio Kast	Derecha radical-conservadora	Conservadurismo moral, retórica antielitista, énfasis en seguridad y migración	Sectores populares conservadores, evangélicos
Johannes Kaiser	Derecha libertaria	Ultraliberalismo, provocación deliberada, códigos de subculturas digitales	Jóvenes masculinos, usuarios intensivos de internet

Actor	Posición	Caracterización sintética	Base social
Jeannette Jara	Izquierda-PC	Proyecto transformador, defensa del legado del gobierno Boric	Militancia de izquierda, sectores sindicalizados

Como se observa en la Table 2.1, el diseño compara tres derechas y un adversario común de izquierda.

Esta selección permite observar:

- **Competencia intra-derecha:** Cómo tres proyectos distintos disputan un electorado parcialmente superpuesto.
- **Convergencia anti-izquierda:** Cómo actores previamente enfrentados articulan un frente común frente al adversario ideológico.
- **Asimetrías discursivas:** Cómo diferentes corrientes movilizan repertorios retóricos y emocionales distintivos.

2.2.2 Las dimensiones analíticas

El análisis se estructura en torno a **tres dimensiones** que capturan aspectos complementarios de la comunicación política:

Dimensión 1: Encuadre (*framing*) - ¿Qué marcos interpretativos predominan en el discurso de cada corriente? - ¿Cómo definen el problema, sus causas y sus soluciones? - ¿Qué está “en juego” según cada actor?

Dimensión 2: Emoción - ¿Qué emociones movilizan las distintas candidaturas? - ¿Cuál es la proporción entre emociones positivas (entusiasmo, esperanza) y negativas (miedo, indignación)? - ¿Hacia quién se dirigen las emociones negativas?

Dimensión 3: Construcción de adversarios - ¿Quiénes son construidos como amenazas o enemigos? - ¿Qué estrategias de deslegitimación se emplean? - ¿Cómo se articulan los antagonismos inter-bloque (derecha vs. izquierda) con los intra-bloque (entre derechas)?

2.2.3 Los niveles de análisis

El diseño contempla tres niveles de observación:

Table 2.2: Niveles de análisis utilizados en la investigación.

Nivel	Unidad	Foco analítico
Macro	Corpus completo	Tendencias agregadas, evolución temporal, comparación entre subreddits
Meso	Hilos de discusión	Dinámicas conversacionales, patrones de confrontación/consenso
Micro	Posts y comentarios individuales	Estrategias retóricas específicas, ejemplos ilustrativos

La Table 2.2 resume cómo se articula el análisis en escalas macro, meso y micro.

2.3 Descripción de la propuesta analítica

2.3.1 Enfoque general

La propuesta combina **ciencia social computacional** con **análisis interpretativo del discurso político**. Esta combinación responde a una convicción metodológica: las técnicas computacionales permiten identificar patrones a escala que el análisis cualitativo tradicional no alcanza, pero requieren interpretación teóricamente informada para producir conocimiento sustantivo sobre procesos políticos.

El flujo analítico puede representarse así:

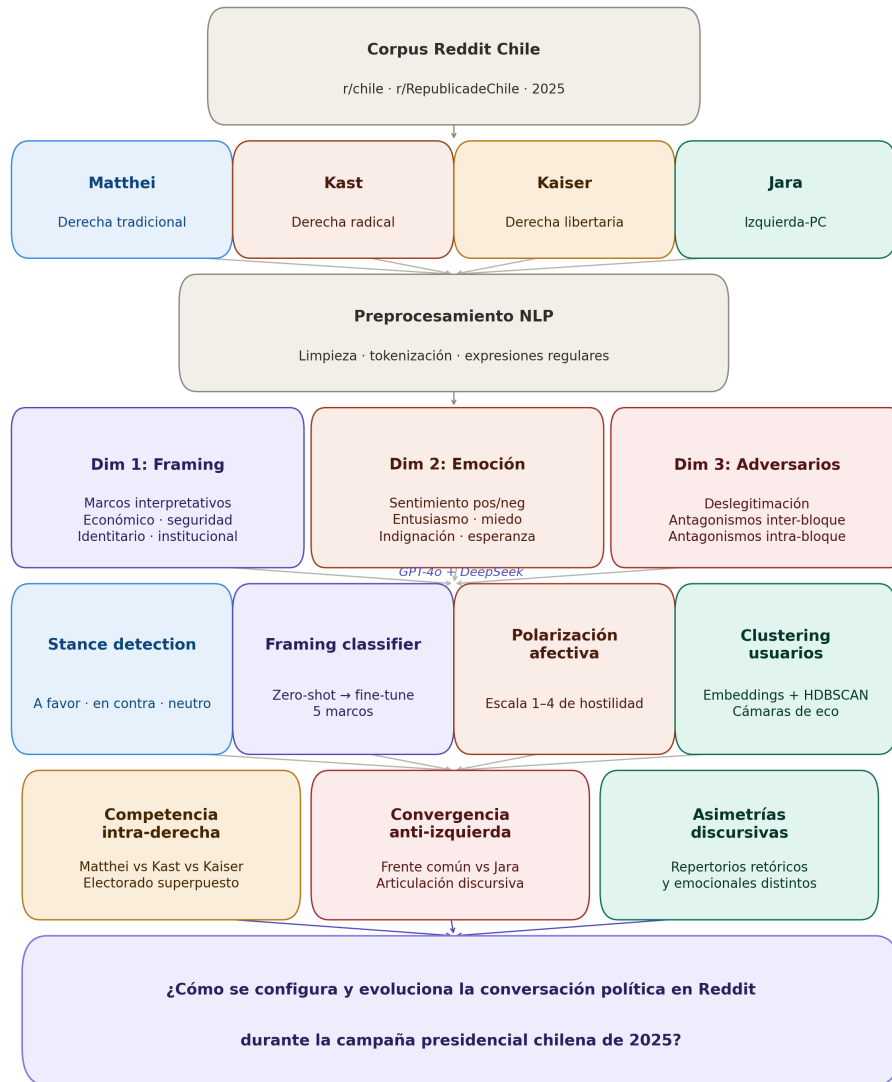


Figure 2.1: Pipeline analítico: de la recolección del corpus a los hallazgos sustantivos.

La Figure 2.1 sintetiza el encadenamiento entre corpus, procesamiento, modelamiento y hallazgos.

2.3.2 Arquitectura del análisis

Fase 1: Caracterización del corpus - Estadísticas descriptivas: volumen, distribución temporal, participación por subreddit - Mapeo de actores: frecuencia de mención de cada candidatura - Identificación de eventos: picos de actividad y su correlación con acontecimientos políticos

Fase 2: Análisis de encuadres - Modelado de tópicos para identificar temas estructurantes - Análisis de co-ocurrencias: qué conceptos aparecen asociados a cada candidatura - Comparación inter-temporal: evolución de marcos entre primera y segunda vuelta

Fase 3: Análisis emocional - Clasificación de sentimiento (positivo/negativo/neutro) - Identificación de emociones discretas cuando es posible - Construcción de “perfiles emocionales” por candidatura

Fase 4: Análisis de adversarios - Identificación de menciones cruzadas entre candidaturas - Clasificación de valencia: menciones positivas, negativas, neutras - Mapeo de redes de antagonismo

Fase 5: Síntesis e integración - Triangulación de hallazgos entre dimensiones - Construcción de narrativa analítica - Identificación de mecanismos explicativos

2.3.3 Originalidad de la propuesta

La propuesta se distingue de estudios previos en tres aspectos:

1. **Foco en competencia intra-bloque:** Mientras la mayoría de estudios examina la polarización entre izquierda y derecha, esta investigación centra la atención en las dinámicas *dentro* del campo de las derechas.
2. **Perspectiva procesual:** En lugar de fotografías estáticas, el diseño longitudinal permite observar cómo evolucionan las estrategias discursivas a lo largo del ciclo electoral.
3. **Reddit como campo:** La plataforma está subexplorada en el contexto latinoamericano, y sus características estructurales (conversaciones anidadas, votación comunitaria) ofrecen posibilidades analíticas distintivas.

2.4 Producto mínimo viable (MVP) y objetivos operacionales

2.4.1 Definición del MVP

El **producto mínimo viable** de esta investigación consiste en:

Un análisis sistemático de la conversación política en Reddit sobre las candidaturas presidenciales chilenas de 2025, que identifique y caracterice los patrones de encuadre, emoción y construcción de adversarios de cada corriente ideológica, y que documente su evolución a lo largo del ciclo electoral (agosto-diciembre 2024).

2.4.2 Objetivos operacionales y entregables

Table 2.3: Objetivos operacionales, entregables e indicadores de cumplimiento del MVP.

Objetivo	Entregable	Indicador de cumplimiento
O1: Construir corpus	Base de datos estructurada	1000 posts y 10.000 comentarios con menciones políticas
O2: Caracterizar encuadres	Modelo de tópicos interpretado	Identificación de 5 marcos distintivos por corriente
O3: Mapear emociones	Perfiles emocionales por candidatura	Clasificación de sentimiento del 100% del corpus
O4: Analizar adversarios	Red de antagonismos	Matriz de menciones cruzadas con valencia
O5: Documentar evolución	Serie temporal	Análisis por fase electoral (pre-campaña, campaña, segunda vuelta)
O6: Sintetizar hallazgos	Capítulos de resultados	Respuesta fundamentada a cada pregunta de investigación

La planificación operativa del estudio se detalla en la Table 2.3.

2.4.3 Alcances y limitaciones anticipadas

Alcances: - El análisis cubre el debate en dos comunidades de Reddit, que constituyen los principales espacios de discusión política chilena en la plataforma. - El período captura fases clave

del ciclo electoral, permitiendo observar dinámicas de cambio. - Las técnicas empleadas permiten procesar el corpus completo, no solo muestras.

Limitaciones: - Reddit no es representativo del electorado chileno; los hallazgos refieren al debate en esta plataforma específica, no a la opinión pública general. - Los usuarios de Reddit tienden a ser más jóvenes, más masculinos y más politizados que la población general. - El análisis de sentimiento automatizado tiene márgenes de error, particularmente con ironía, sarcasmo y jerga local. - La atribución ideológica de usuarios es inferida a partir de su comportamiento discursivo, no de autodeclaraciones.

2.5 Estrategia de validación

2.5.1 Validación interna

Triangulación metodológica: Los hallazgos de cada técnica (modelado de tópicos, análisis de sentimiento, análisis de redes) son contrastados entre sí. Un patrón se considera robusto cuando emerge consistentemente de múltiples aproximaciones.

Verificación cualitativa: Los patrones identificados computacionalmente son verificados mediante lectura de casos. Cuando el análisis automatizado clasifica un comentario como “negativo hacia Matthei”, se revisan ejemplos para confirmar que la clasificación captura el fenómeno de interés.

Sensibilidad paramétrica: Las técnicas con parámetros ajustables (número de tópicos, umbrales de clasificación) son ejecutadas con múltiples configuraciones para verificar que los hallazgos no dependan de decisiones arbitrarias.

2.5.2 Validación externa

Triangulación con fuentes externas: Los patrones temporales identificados en Reddit son contrastados con el calendario de eventos políticos documentado en medios de comunicación. La correlación entre picos de actividad y eventos específicos (debates, declaraciones, encuestas) proporciona validación externa.

Consistencia con literatura: Los hallazgos son discutidos a la luz de estudios previos sobre comunicación política digital, derechas latinoamericanas y polarización. La convergencia con patrones documentados en otros contextos refuerza la validez; las divergencias son examinadas como posibles especificidades del caso chileno.

2.5.3 Criterios de éxito

La investigación se considerará exitosa si:

1. **Responde las preguntas:** Cada pregunta de investigación recibe una respuesta fundamentada empíricamente.
2. **Identifica patrones no triviales:** Los hallazgos van más allá de lo que podría inferirse sin análisis sistemático.
3. **Documenta variación:** El análisis captura diferencias significativas entre corrientes ideológicas, entre subreddits y entre fases electorales.
4. **Genera interpretaciones plausibles:** Los patrones identificados son interpretados mediante mecanismos teóricamente informados.
5. **Reconoce limitaciones:** El análisis explicita qué puede y qué no puede concluirse a partir de los datos disponibles.

3 Marco teórico

3.1 Comunicación política en entornos digitales

Las plataformas digitales no amplían cuantitativamente la esfera pública: la reconfiguran cualitativamente. boyd (2010) identificó cuatro propiedades estructurales —persistencia, replicabilidad, escalabilidad y buscabilidad— que distinguen los entornos digitales de espacios comunicativos previos. A estas se suman la **mediación algorítmica** (que privilegia contenidos emocionalmente activadores), el **colapso de contextos** (convergencia de audiencias antes separadas) y las **métricas de rendimiento** (retroalimentación inmediata que orienta la conducta comunicativa).

Más allá de sus propiedades técnicas, las plataformas deben entenderse como infraestructuras de poder privadas que ejercen moderación, optimizan el *engagement* y consolidan efectos de red con consecuencias desproporcionadas para la calidad del debate democrático (Gillespie 2018; Dijck 2018).

3.2 Polarización afectiva

La **polarización afectiva** designa el distanciamiento emocional entre grupos políticos: no solo divergen en posiciones, sino que se desagradan y se atribuyen mutuamente malas intenciones (Iyengar 2019). Es analíticamente distinta de la polarización ideológica: ambas pueden variar de manera independiente, y la evidencia comparada sugiere que la afectiva ha crecido más rápidamente en democracias contemporáneas (Boxell 2022).

Sus principales mecanismos en entornos digitales incluyen la exposición selectiva amplificada por algoritmos, la mayor visibilidad de voces extremas, la reducción de costos de expresión hostil y las

dinámicas de señalización grupal recompensadas con *likes* (Brady 2017; Settle 2018). Las consecuencias para la democracia son significativas: erosión de la confianza institucional, resistencia al compromiso y tolerancia a la transgresión normativa cuando proviene del propio bando (Graham 2020).

3.3 Teoría del *framing*

El encuadre consiste en seleccionar aspectos de la realidad y hacerlos prominentes para promover una definición del problema, una interpretación causal y una evaluación moral (Entman 1993). La competencia política es, en parte significativa, competencia entre marcos alternativos: quien establece el marco dominante condiciona las opciones que parecen razonables.

Se distinguen **marcos genéricos** (conflicto, interés humano, consecuencias económicas, atribución de responsabilidad, moralidad) y **marcos específicos** propios de cada controversia (Semetko 2000). Para campañas electorales resulta útil añadir la distinción entre marcos diagnósticos (¿cuál es el problema?), pronósticos (¿qué debe hacerse?) y motivacionales (¿qué está en juego?).

En entornos digitales, el *framing* es observable como proceso: los marcos son propuestos, contestados y reformulados en tiempo real. La estructura anidada de Reddit —con sistemas de votación y conversaciones persistentes— es particularmente adecuada para este análisis longitudinal (Meraz 2013).

3.4 Emociones políticas

Las emociones no son ruido que distorsiona el juicio racional: son componentes constitutivos de la cognición y la acción política (Marcus 2000). La teoría de la inteligencia afectiva distingue un sistema de disposiciones (activado en entornos familiares, refuerza hábitos mediante entusiasmo o aversión) y un sistema de vigilancia (activado ante amenaza o incertidumbre, genera ansiedad que puede interrumpir el procesamiento automático y abrir espacio a la persuasión).

Para el análisis empírico, el concepto de **repertorio afectivo** —la configuración característica de emociones movilizadas por un actor, su combinación y los objetos hacia los que se dirigen— resulta

más útil que el registro de emociones aisladas (Jasper 2011). Permite preguntar: ¿qué proporción del contenido emocional se dirige al endogrupo versus al exogrupo? ¿Cómo articula el actor emociones positivas y negativas?

La **indignación moral** merece atención especial: combina valencia negativa con alta activación y orientación a la acción, se difunde más que otras emociones y genera bajo costo de expresión con recompensas sociales inmediatas (Crockett 2017).

3.5 Construcción de adversarios

Las identidades colectivas no son esencias preexistentes: emergen a través de la diferenciación. Todo “nosotros” se constituye en relación con un “ellos” (Laclau 1985). En contextos electorales, analizar cómo se construyen los adversarios —qué atributos se les asignan, qué estrategias se emplean para deslegitimarlos— es una vía privilegiada para comprender la configuración del espacio político.

Mouffe (2005) distingue **antagonismo** (relación entre enemigos sin espacio común, suma cero) de **agonismo** (relación entre adversarios que comparten el marco democrático). Esta distinción permite preguntar si los actores construyen a sus oponentes como participantes legítimos o como amenazas existenciales.

Las estrategias de deslegitimación incluyen: asociación negativa, atribución de intenciones ocultas, ridiculización, esencialización y construcción de amenaza. Para esta investigación resulta además crucial distinguir entre **fronteras inter-bloque** (izquierda vs. derecha) y **fronteras intra-bloque** (entre corrientes de derecha), que generan tensiones estratégicas entre diferenciación y convergencia.

3.6 Derechas contemporáneas

La tipología de Mudde (2007) distingue **extrema derecha** (rechaza principios democráticos) de **derecha radical** (los acepta nominalmente pero los tensiona mediante nativismo, autoritarismo y populismo). A estas se suman la **derecha tradicional** (privilegia estabilidad y continuidad institucional) y la **derecha libertaria** (ultraliberalismo, crítica al establishment, códigos digitales).

Estas corrientes difieren no solo ideológicamente sino en sus estrategias comunicativas: la derecha tradicional emplea registro institucional y moderado; la radical populista moviliza agravios y retórica de confrontación; la libertaria combina argumentos económicos con estilos provocadores y referencias a subculturas de internet (Norris 2019).

La coexistencia de las tres derechas en el caso chileno genera dinámicas de competencia intra-bloque —disputa por electorados superpuestos, estrategias de diferenciación, presiones de convergencia frente al adversario común— que constituyen un laboratorio analítico de particular valor.

3.7 Reddit como espacio de deliberación política

Reddit se estructura en torno a comunidades temáticas (*subreddits*) antes que en redes de relaciones personales, lo que privilegia el debate sustantivo sobre la identidad individual. Sus comentarios anidados permiten seguir el desarrollo de argumentos y contraargumentos; el sistema de votación ofrece un indicador de resonancia comunitaria; y la persistencia habilita el análisis longitudinal (Proferes 2021).

En el contexto chileno, **r/chile** (generalista, tendencia centro-izquierda) y **r/RepublicadeChile** (surgida como escisión, tolerancia de discursos más confrontacionales) constituyen un caso comparativo que permite examinar cómo una misma controversia es procesada de manera diferenciada según la cultura discursiva de la comunidad.

3.8 Modelo analítico integrado

El modelo integrado sigue una secuencia analítica:

1. **Contexto:** fragmentación partidaria, reconfiguración de derechas y plataformización.
2. **Actores:** Matthei, Kast, Kaiser y Jara.
3. **Dimensiones:** framing, emoción y construcción de adversarios.
4. **Dinámicas:** competencia intra-derecha, convergencia anti-izquierda y evolución temporal.
5. **Resultado:** identificación de patrones discursivos por corriente ideológica.

4 Diseño metodológico

4.1 Diseño de investigación

Esta investigación adopta un diseño **longitudinal observacional** basado en datos digitales generados por usuarios (*user-generated content*). El carácter longitudinal permite capturar la evolución del debate político a lo largo del ciclo electoral, identificando tendencias, puntos de inflexión y dinámicas de cambio. El carácter observacional preserva las propiedades naturales de la interacción en las plataformas estudiadas, sin intervención experimental.

Metodológicamente, la investigación se inscribe en la **ciencia social computacional** (Lazer 2009; Salganik 2019): integra técnicas de procesamiento de datos a gran escala con marcos teóricos y preguntas sustantivas de las ciencias sociales. El diseño combina componentes **descriptivos** — caracterizar la conversación política en Reddit durante el período estudiado— con componentes **explicativos**, orientados a identificar los factores y mecanismos que dan cuenta de los patrones observados.

4.2 Reddit como plataforma de análisis

Reddit es una plataforma de agregación de contenido y discusión organizada en comunidades temáticas denominadas *subreddits*, cada una con sus propias normas, moderadores y cultura discursiva. A diferencia de plataformas organizadas en torno a redes de seguidores o amistades, Reddit estructura la participación en torno a **intereses y temas compartidos**, lo que favorece conversaciones más extensas y argumentadas que en otros entornos digitales.

Su arquitectura de **comentarios anidados** (*threaded discussions*) permite seguir el desarrollo de argumentos y contraargumentos en múltiples niveles, haciendo visible la estructura deliberativa del debate. El sistema de votación (*upvotes/downvotes*) ofrece un indicador de resonancia comunitaria, y la persistencia de los hilos habilita el análisis longitudinal de conversaciones en desarrollo.

Los usuarios operan bajo seudónimos no vinculados a su identidad real, lo que puede facilitar la expresión de posiciones polémicas. Las comunidades son autogobernadas: cada subreddit define sus propias reglas de participación, generando culturas discursivas distintivas que hacen de la comparación entre comunidades un ejercicio analíticamente valioso.

En el contexto chileno, dos subreddits concentran la discusión política nacional:

- **r/chile**: Comunidad generalista sobre temas chilenos, con mayor volumen de usuarios y diversidad temática. Ha tendido históricamente hacia posiciones de centro-izquierda, aunque con presencia de voces diversas.
- **r/RepublicadeChile**: Comunidad surgida como escisión de r/chile, con políticas de moderación más permisivas y una composición ideológica tendiente a la derecha. Tolera discursos más confrontacionales y concentra usuarios críticos de la moderación de r/chile.

Esta dualidad constituye un caso comparativo de valor: permite examinar cómo una misma controversia política es procesada de manera diferenciada según la cultura discursiva de la comunidad.

4.3 Recolección de datos mediante la API de Reddit

4.3.1 Procedimiento técnico

Los datos fueron extraídos mediante la API pública de Reddit utilizando la librería PRAW (*Python Reddit API Wrapper*). Esta aproximación —distinta del *web scraping* no autorizado— accede de forma estructurada a contenido público respetando los límites de uso establecidos por la plataforma y garantizando la trazabilidad y reproducibilidad del proceso.

4.3.2 Período de recolección y fases del ciclo electoral

La recolección abarcó **cinco meses**, del **1 de agosto al 31 de diciembre de 2025**, cubriendo las siguientes fases del ciclo electoral:

Período	Fase electoral
Agosto – octubre 2025	Emergencia de candidaturas, posicionamiento inicial, primeras definiciones programáticas
Noviembre – diciembre 2025	Escalamiento de campaña, consolidación electoral, estructuración del debate público

4.3.3 Estrategia de extracción incremental

El proceso fue **incremental y acumulativo**, ejecutado en múltiples iteraciones que: (1) recuperaban los posts más recientes de cada subreddit (250–1.000 posts por ejecución); (2) descargaban el árbol completo de comentarios mediante `replace_more(limit=None)` para capturar todos los niveles de anidación; (3) integraban nuevos datos con extracciones previas en archivos Parquet; y (4) eliminaban duplicados mediante el identificador único nativo de Reddit (`post_id`), que garantiza unicidad por diseño de la plataforma.

4.3.4 Contenido extraído y filtrado temático

Para cada publicación dentro del rango temporal se extrajeron dos niveles:

Posts: título, cuerpo (*selftext*), autor, fecha, puntuación, número de comentarios y flair temático.

Comentarios: cuerpo del comentario, autor, fecha, puntuación e identificador del comentario al que responde.

El filtrado temático se realizó mediante **expresiones regulares** que detectan menciones de las cuatro candidaturas de interés, manejando variaciones ortográficas (acentuación, nombres completos vs. apellidos, apodos y errores tipográficos frecuentes):

Table 4.2: Candidaturas objetivo y patrones regex para su detección en el corpus.

Candidatura	Corriente ideológica	Patrones de búsqueda
Evelyn Matthei	Derecha tradicional	<code>matthei, matei, evelyn</code>
José Antonio Kast	Derecha radical-conservadora	<code>kast, jose antonio kast, jak</code>
Johannes Kaiser	Derecha libertaria	<code>kaiser, johannes, jkaiser</code>
Jeannette Jara	Izquierda-PC	<code>jara, jeannette, jeanette</code>

Esta selección permite analizar tanto la **competencia intra-derecha** entre tres corrientes como la **convergencia discursiva** frente al adversario ideológico común (ver Table 4.2).

4.4 Consideraciones éticas

La investigación se rige por principios establecidos para la investigación con datos digitales (Salganik 2019; Zimmer 2017):

Datos públicos: Todo el contenido recolectado proviene de espacios públicos de Reddit. No se accedió a mensajes privados ni a información restringida.

Privacidad: El análisis no reporta identificadores individuales ni busca vincular cuentas con identidades reales. Los ejemplos citados son anonimizados cuando resulta pertinente.

No intervención: La investigación no actúa sobre las comunidades estudiadas ni expone a sus miembros a riesgos adicionales; el foco es el análisis de dinámicas colectivas.

Reproducibilidad: Los scripts de recolección y procesamiento están documentados para permitir la replicación del estudio respetando los términos de uso de la API.

4.5 Técnicas de análisis

El análisis combina la capacidad de procesamiento de modelos de lenguaje de gran escala con validación humana sistemática. Se utilizaron dos APIs de forma complementaria: **GPT-4o** (OpenAI) y **DeepSeek**, seleccionados por su capacidad de comprensión del español rioplatense y chileno, su desempeño en tareas de clasificación de texto político y la posibilidad de controlar los prompts con precisión. Ningún modelo reemplaza el juicio analítico: en todos los casos, las clasificaciones automatizadas son revisadas y validadas manualmente sobre muestras representativas antes de escalar al corpus completo.

4.5.1 Análisis de sentimiento y emociones

Cada fragmento del corpus fue procesado mediante prompts estructurados que solicitaban a los modelos clasificar: **polaridad** (positiva, negativa, neutra), **intensidad emocional** (baja, media, alta) y **emoción predominante** (alegría, miedo, ira, desprecio, indignación, esperanza). Los prompts incluían definiciones operacionales de cada categoría y ejemplos de calibración extraídos del propio corpus de Reddit chileno.

La clasificación automática fue validada mediante codificación manual de una muestra aleatoria estratificada por candidatura y subreddit, calculando coeficientes de acuerdo intercalificador (Cohen's κ) para establecer umbrales de confiabilidad aceptables antes de aplicar el esquema al corpus completo.

4.5.2 Clasificación de marcos interpretativos (*framing*)

Para identificar los marcos interpretativos predominantes, se diseñaron prompts que solicitaban a los modelos asignar cada fragmento a uno o más de los marcos definidos teóricamente: económico, seguridad, identitario, institucional y medioambiental. Los prompts especificaban indicadores lingüísticos de cada marco y pedían justificación de la clasificación, lo que permitió refinar iterativamente las definiciones operacionales.

Al igual que en el análisis de sentimiento, las clasificaciones fueron validadas manualmente antes de escalar. Los casos de desacuerdo entre GPT-4o y DeepSeek fueron marcados para revisión humana prioritaria, convirtiendo la divergencia entre modelos en un indicador de ambigüedad analítica.

4.5.3 Análisis de redes de interacción

A partir de la estructura de respuestas anidadas de Reddit, se construyeron **redes de réplica** donde los nodos representan usuarios y las aristas representan relaciones de respuesta (comentario a comentario). El análisis de estas redes permite identificar subcomunidades discursivas, patrones de confrontación o cooperación entre usuarios de distintas orientaciones ideológicas, y la presencia de estructuras de cámara de eco en cada subreddit.

4.5.4 Análisis temporal

El diseño longitudinal habilita el análisis de **series temporales** de los indicadores de interés: volumen de menciones, tono emocional promedio y distribución de marcos temáticos semana a semana. Se identifican tendencias, ciclos y puntos de quiebre, así como correlaciones entre eventos externos (debates, declaraciones, controversias) y cambios en los patrones discursivos.

4.6 Corpus final y pipeline de procesamiento

Al cierre de la recolección (31 de diciembre de 2025), el corpus contiene todos los posts y sus árboles completos de comentarios publicados en r/chile y r/RepublicadeChile durante el período que mencionan al menos una de las cuatro candidaturas de interés, junto con sus metadatos asociados. La estructura del corpus permite análisis longitudinales, de red e interaccionales a lo largo del ciclo electoral completo.

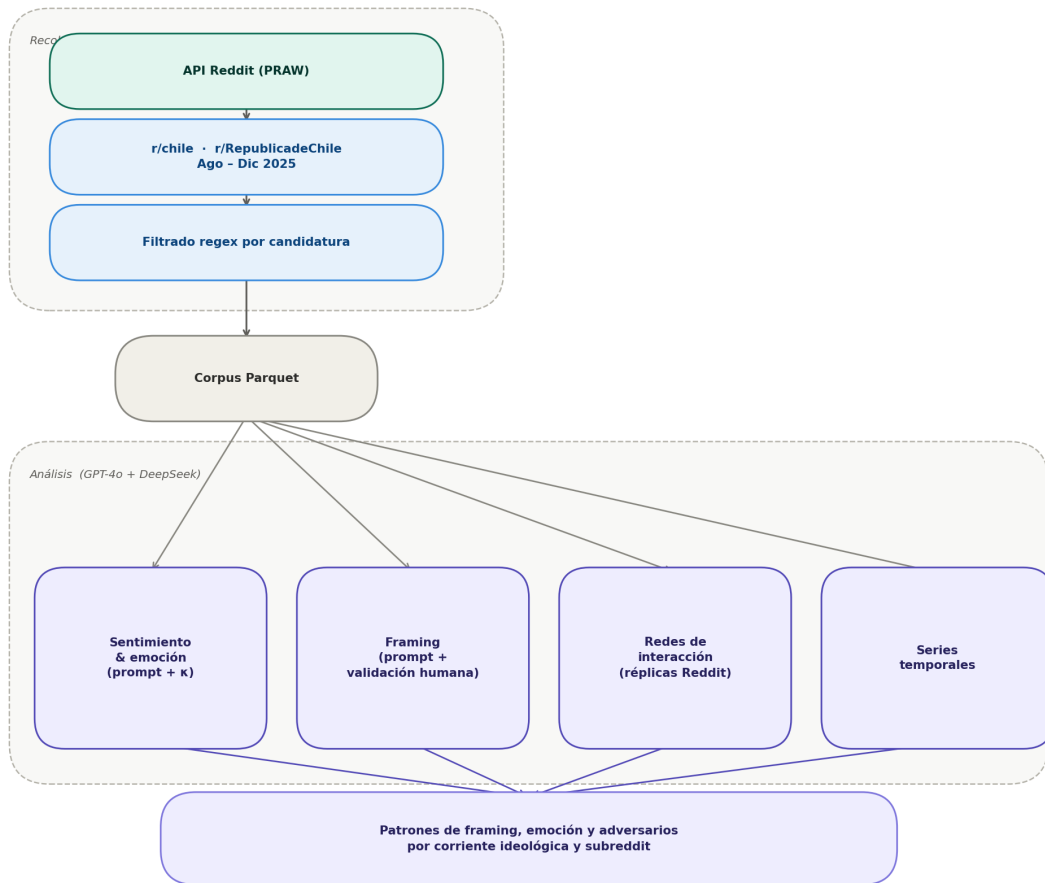


Figure 4.1: Flujo de recolección y procesamiento del corpus.

5 Depuración de Datos

6 Exploración de Datos

7 Modelado de Datos

[Contenido de modelado de datos]

7.1 Clustering K-Means

[Contenido sobre clustering K-Means]

7.2 Clustering DBSCAN

[Contenido sobre clustering DBSCAN]

7.3 Clustering Jerárquico

[Contenido sobre clustering jerárquico]

7.4 Determinación del modelo más óptimo

[Contenido sobre determinación del modelo más óptimo]

8 Interpretación de Resultados

[Contenido de interpretación de resultados]

8.1 Interpretación de componentes principales

[Contenido sobre interpretación de componentes principales]

8.2 Interpretación de clusters

[Contenido sobre interpretación de clusters]

8.3 Validación de los resultados

[Contenido sobre validación de resultados]

9 Conclusión

[Contenido de la conclusión]

10 Referencias

Boxell. 2022. "Boxell2022."
Brady. 2017. "Brady2017."
Crockett. 2017. "Crockett2017."
Dijck, van. 2018. "vanDijck2018."
Entman. 1993. "Entman1993."
Gillespie. 2018. "Gillespie2018."
Graham. 2020. "Graham2020."
Iyengar. 2019. "Iyengar2019."
Jasper. 2011. "Jasper2011."
Laclau. 1985. "Laclau1985."
Lazer. 2009. "Lazer2009."
Marcus. 2000. "Marcus2000."
Meraz. 2013. "Meraz2013."
Norris. 2019. "Norris2019."
Proferes. 2021. "Proferes2021."
Salganik. 2019. "Salganik2019."
Semetko. 2000. "Semetko2000."
Settle. 2018. "Settle2018."
Zimmer. 2017. "Zimmer2017."