



**UNIVERSIDAD DE CIENCIAS
EMPRESARIALES Y SOCIALES**

**MAESTRÍA EN
INVESTIGACIÓN DE MERCADOS, MEDIOS Y OPINIÓN**

**“La relación entre el discurso de la opinión pública y el
discurso mediático, desde el diálogo entre la investigación
social y la ciencia de datos”**

Manuel Maffé

Año 2020

Tutor: Renzo Moyano

ÍNDICE GENERAL

INTRODUCCIÓN	7
CAPÍTULO 1. MARCO TEÓRICO	9
CAPÍTULO 2. METODOLOGÍA	26
<i>Problema de Investigación y Objetivos</i>	26
<i>Estrategia de Recolección de Datos</i>	28
<i>Métodos analíticos para el análisis de texto</i>	31
<i>Métodos para el análisis comparado de universos discursivos</i>	37
CAPÍTULO 3. ¿QUÉ NOS DICEN LOS MEDIOS? ANÁLISIS DE CONTENIDO SOBRE EL DISCURSO DE LOS MEDIOS DE COMUNICACIÓN	40
CAPÍTULO 4. EL PROBLEMA MÁS IMPORTANTE ANÁLISIS DE CONTENIDO SOBRE EL DISCURSO DE LA OPINIÓN PÚBLICA.....	64
CAPÍTULO 5. ¿AGENDAS RELACIONADAS? ANÁLISIS DE LA RELACIÓN ENTRE EL DISCURSO DE LA OPINIÓN PÚBLICA Y EL DISCURSO MEDIÁTICO.....	81
CAPÍTULO 6. CONCLUSIONES.....	100
BIBLIOGRAFÍA	106
ANEXO	109

ÍNDICE DE FIGURAS

Figura 1. Representación de las vertientes sobre el estudio de la Opinión Pública y Redes Sociales.....	17
Figura 2. Conformación de agenda y Objeto de Estudio.....	23
Figura 3. Organización temporal del trabajo de campo.....	29
Figura 4. Resumen de herramientas metodológicas aplicables a datos de tipo textual.....	32
Figura 5. Representación gráfica del modelo LDA.....	35

ÍNDICE DE TABLAS

Tabla 1. <i>Datos sobre consumo de medios de comunicación online seleccionados. Año 2018</i>	30
Tabla 2. <i>Distribución de publicaciones según Medio de Comunicación</i>	40
Tabla 3. <i>Coocurrencia entre series temporales correspondientes a Cantidad de Publicaciones por Medio de Comunicación</i>	42
Tabla 4. <i>Distribución de documentos según Tópico</i>	48
Tabla 5. <i>Distribución de documentos según Tópico y Medio de Comunicación</i>	57
Tabla 6. <i>Relaciones estadísticamente significativas en Tópicos y Medios de Comunicación</i>	59
Tabla 7. <i>Coeficiente Correlación de Spearman entre Medios de Comunicación</i>	62
Tabla 8. <i>Distribución de la muestra según Sexo y Rango etario</i>	64
Tabla 9. <i>Distribución de respuestas por tópico</i>	70
Tabla 10. <i>Penetración de medios de comunicación online entre los encuestados</i>	70
Tabla 11. <i>Distribución de Tópicos según Audiencias</i>	72
Tabla 12. <i>Diez palabras más utilizadas por la Opinión Pública, en relación a los Medios de Comunicación en general y cada medio de comunicación en particular (% Penetración)</i>	82
Tabla 13. <i>Diez palabras más utilizadas por los Medios de Comunicación en general, en relación a la Opinión Pública y cada medio de comunicación en particular (% Penetración)</i>	83
Tabla 14. <i>Correlación entre términos utilizados por la Opinión Pública y los Medios de Comunicación</i>	84
Tabla 15. <i>Correlación entre universos discursivos, para cada una de las audiencias principales</i>	87

ÍNDICE DE GRÁFICOS

Gráfico 1. <i>Población argentina (Total y con Conectividad a Internet) según Año</i>	11
Gráfico 2. <i>Usuarios de Redes Sociales en Argentina, según Año</i>	12
Gráfico 3. <i>Uso diario de Internet y Uso de la red social Twitter en Argentina, según Año</i>	12
Gráfico 4. <i>Distribución de publicaciones según Fecha y Medio de Comunicación</i>	41
Gráfico 5. <i>Cantidad de Apariciones por Bi-grama</i>	42
Gráfico 6. <i>Cantidad de Apariciones por Bi-grama, según Medio de Comunicación</i>	45
Gráfico 7. <i>Validación cruzada de cantidad de tópicos k sobre documentos relevados</i>	46
Gráfico 8. <i>Desagregación de tópicos en palabras con mayor contribución</i>	49
Gráfico 9. <i>Evolución de los tópicos hallados sobre la agenda mediática en el tiempo</i>	61
Gráfico 10. <i>Palabras más mencionadas para PMI</i>	65
Gráfico 11. <i>Indicador de perplejidad para PMI</i>	66
Gráfico 12. <i>Desagregación de tópicos para PMI en palabras con mayor contribución</i>	67
Gráfico 13. <i>Tamaño relativo de cada audiencia en función de su consumo de medios de comunicación online</i>	71
Gráfico 14. <i>Composición de cada audiencia principal en función de su posicionamiento político</i>	74
Gráfico 15. <i>Composición de cada audiencia principal en función de su confianza declarada en los medios de comunicación</i>	76
Gráfico 16. <i>Proporción de Usuarios con Consumo Diario de medios de comunicación online, según Audiencia</i>	77
Gráfico 17. <i>Composición de cada audiencia principal en función de su edad</i>	78

Gráfico 18. <i>Proporción de Consumidoras Mujeres, según Audiencia</i>	79
Gráfico 19. <i>Relación entre términos utilizados por los Medios de Comunicación y la Opinión Pública</i>	85
Gráfico 20. <i>Relación entre términos utilizados por cada Medio de Comunicación y la Opinión Pública</i>	86
Gráfico 21. <i>Relación entre Tamaño de la audiencia y Correlación entre discursos</i>	88
Gráfico 22. <i>Relación entre Confianza en medios de comunicación y Correlación entre discursos</i>	90
Gráfico 23. <i>Relación entre Confianza en medios de comunicación y Cantidad de Medios consumidos</i>	91
Gráfico 24. <i>Relación entre Proporción de usuarios diarios y Correlación entre discursos</i>	92
Gráfico 25. <i>Relación entre Imagen Positiva sobre Mauricio Macri y Correlación entre discursos</i>	93
Gráfico 26. <i>Relación entre Imagen Positiva sobre Cristina Fernández de Kirchner y Correlación entre discursos</i>	95
Gráfico 27. <i>Relación entre Posicionamiento Político y Correlación entre discursos</i>	96
Gráfico 28. <i>Relación entre Edad Mediana de cada audiencia y Correlación entre discursos</i>	97
Gráfico 29. <i>Relación entre Proporción de Consumidoras Mujeres en cada audiencia y Correlación entre discursos</i>	98

INTRODUCCIÓN

El presente trabajo busca contribuir a la investigación en opinión pública a partir de la incorporación de herramientas propias del ámbito de la Ciencia de Datos. En este sentido, el objetivo es realizar un aporte al desarrollo local de las ciencias sociales aplicadas, a partir del uso de herramientas analíticas provenientes de desarrollos en el mundo de la Tecnología de la Información.

Cabe señalar que este es un enfoque ya desarrollado en el mundo, generalmente bajo la terminología de *Computational Social Science*, y que encuentra incipientes desarrollos en el ámbito local, cuyos autores son, en gran parte, retomados en el presente documento. Desde el presente documento, se sostiene que este enfoque es fundamental para lograr robustecer los hallazgos de las ciencias sociales, mantener la actualidad de sus descubrimientos y, sobre todo, poder reducir la brecha entre lo que las personas dicen, sienten y hacen. En este sentido, muchas herramientas analíticas propias del mundo de la tecnología permiten integrar datos comportamentales a la información declarativa con la que, en gran parte de las ocasiones, las ciencias sociales cuentan, pudiendo desarrollar un enfoque más integral.

En este sentido, en el presente trabajo se busca vincular metodologías de tipo tradicional -en este caso, encuestas-, con herramientas propias del ámbito de la Ciencia de Datos, particularmente técnicas ligadas al procesamiento de lenguaje natural. Estas últimas herramientas citadas, por su parte, permiten realizar un análisis que tenga en cuenta un volumen mayor de datos y que, mediante un menor uso de recursos, permita llegar a resultados más asertivos y granulares, en este caso particular, respecto a la relación entre el contenido del discurso de la opinión pública y el contenido del discurso mediático. Por otro lado, los datos de carácter declarativo relevados mediante encuesta, permiten brindar riqueza al análisis sobre la convivencia y coocurrencia de estos mundos discursivos, logrando obtener resultados enmarcados en la propia perspectiva de los que opinan.

Para llevar a cabo este trabajo, en el primer capítulo se desarrolla el marco teórico del presente documento. Allí, se desarrollan los avances en materia de metodologías mixtas que aquí son retomados, a la vez que se introducen

conceptos propios de las teorizaciones sobre el comportamiento de la opinión pública frente a los medios de comunicación, y viceversa, considerando que éstos serán los “lentes” mediante los cuales serán leídos los resultados hallados. En el segundo capítulo, se desarrolla la metodología de trabajo, integrando tanto el diseño de la actual investigación, como la estrategia de relevamiento de datos y los métodos analíticos utilizados. En los capítulo tercero y cuarto, se analiza el contenido del discurso mediático y del discurso de la opinión pública, por separado, mientras que en el capítulo quinto se relacionan ambos discursos y se analiza este comportamiento bajo la óptica de la perspectiva política de los entrevistados. Por último, en el capítulo seis, se realiza una síntesis de los principales hallazgos del documento, y se desarrollan las conclusiones pertinentes en base a los objetivos planteados.

Este es un trabajo que se enmarca en un mundo profesional amplio donde, desde una multiplicidad de miradas, una cantidad innumerable de analistas se esfuerzan por darle sentido a las huellas que dejamos en entornos digitales, en forma de datos, y a lo que declaramos cuando opinamos. En este sentido, busca ser un aporte más a este conjunto de esfuerzos, en pos de mejorar, cada vez más, la calidad de nuestros hallazgos.

CAPÍTULO 1. MARCO TEÓRICO

La primera pregunta a responder mediante los antecedentes teóricos retomados en el presente trabajo es: ¿Por qué Twitter? En este sentido, cabe señalar que dicha red social constituye un marco de expresión tanto para líderes sociales - por ellos entendemos tanto líderes políticos como medios de comunicación- como para la ciudadanía en general. Por otro lado, cabe realizar un comentario respecto a su contenido, retomando las palabras de los autores Justin Grimmer y Brandon Stewart (2013), quienes sostienen que “el lenguaje es el medio para la política y el conflicto político (...) Para entender sobre qué se trata la política, necesitamos conocer sobre qué están hablando y escribiendo los actores políticos” (p. 267, traducción propia).

En este sentido, siguiendo a Ernesto Calvo (2015), en el presente trabajo se toma como base una mutación irreversible de la manera en la que se da el diálogo político, a partir de estas nuevas esferas públicas en la cual tanto la ciudadanía como los políticos y los medios de comunicación, opinan permanentemente y en constante diálogo. En palabras del autor anteriormente mencionado, estamos ante “Un foro político, social y cultural que se mantiene abierto las 24 horas del día, los 365 días del año. Un foro político irreverente que produce una cantidad de información inimaginable hasta hace tan solo una década” (Calvo, 2015, p. 11-12). Esta mutación, a su vez, impacta directamente en la cultura política de las sociedades contemporáneas, considerando que las redes sociales son ya hoy parte integral de, por ejemplo, las campañas políticas, implicando esto financiamiento, movilización de votantes, denuncia política, entre otros aspectos relevantes.

Frente a esta producción constante de información en redes sociales, cabe destacar también un eje de estabilidad que es transversal a este dinamismo: tanto los círculos de pertenencia creados en las redes sociales -esto es, las comunidades de amigos o personas que seguimos- como las actitudes y hábitos de los usuarios, son relativamente estables en el tiempo, y constituyen aquellos “lentes” mediante los cuales los usuarios consumen, y por ende interpretan, la información circundante en estos medios de expresión política. En este sentido, se presentan dos planos diferenciados de análisis, que siempre están interactuando entre sí: por un lado, el contenido fluido de las redes sociales, por

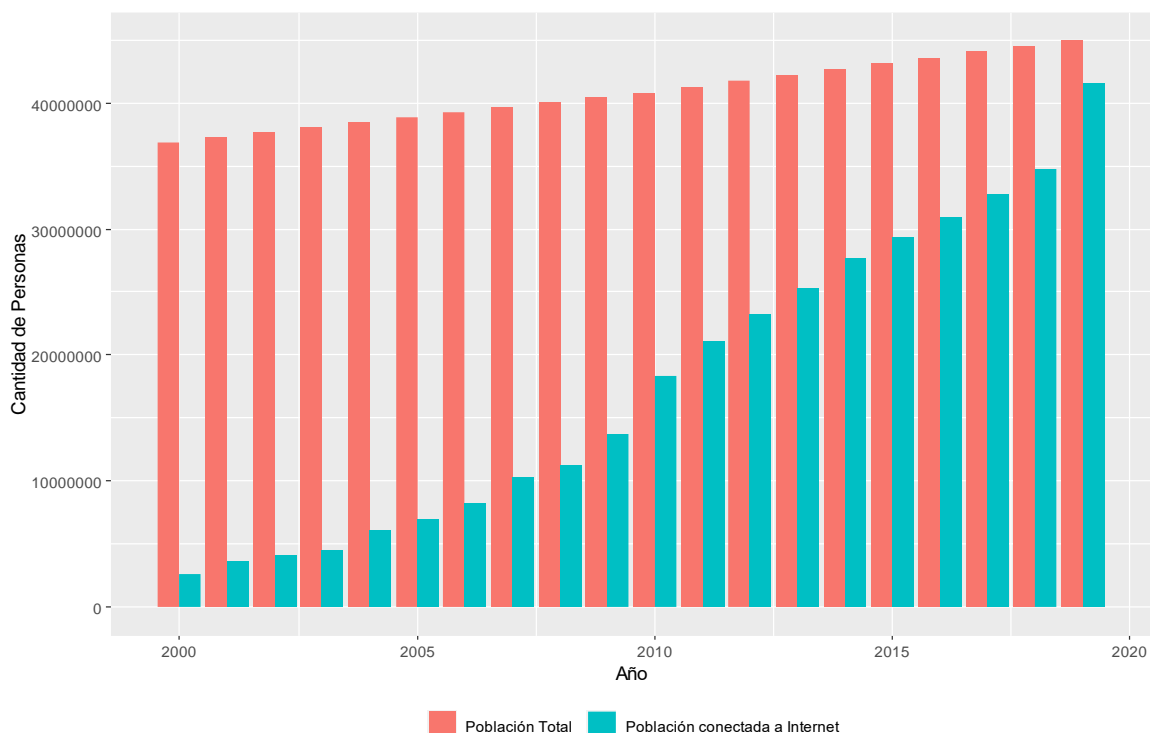
otro, las estructuras interpretativas mediante las cuales dicho contenido es consumido. En palabras de Ernesto Calvo (*íbid.*):

Contrariamente a la imagen de una red social en constante cambio y que varía de acuerdo con el tópico de discusión, la gran mayoría de las redes son estables y perduran a lo largo del tiempo. Nuestros amigos en Facebook no cambian cada mes, como tampoco lo hace el conjunto de individuos a quienes seguimos en Twitter. Nuestros ideales, gustos y preferencias mutan lentamente, así como lentamente mutan las redes sociales. Cambian las ironías, los comentarios, los contenidos, pero la estructura permanece (...). (p. 16-17).

La relevancia de dicha red social también se encuentra no solo dada por su impacto en la comunicación y cultura política, sino también por su creciente nivel de adopción, dato que también se relaciona con el nivel de conexión de la población local. A continuación, se observa la evolución anual de la población argentina junto a la evolución de la población conectada a internet:

Gráfico 1.

Población argentina (Total y con Conectividad a Internet) según Año



Fuente: Elaboración Propia a partir de Banco Mundial y We Are Social¹

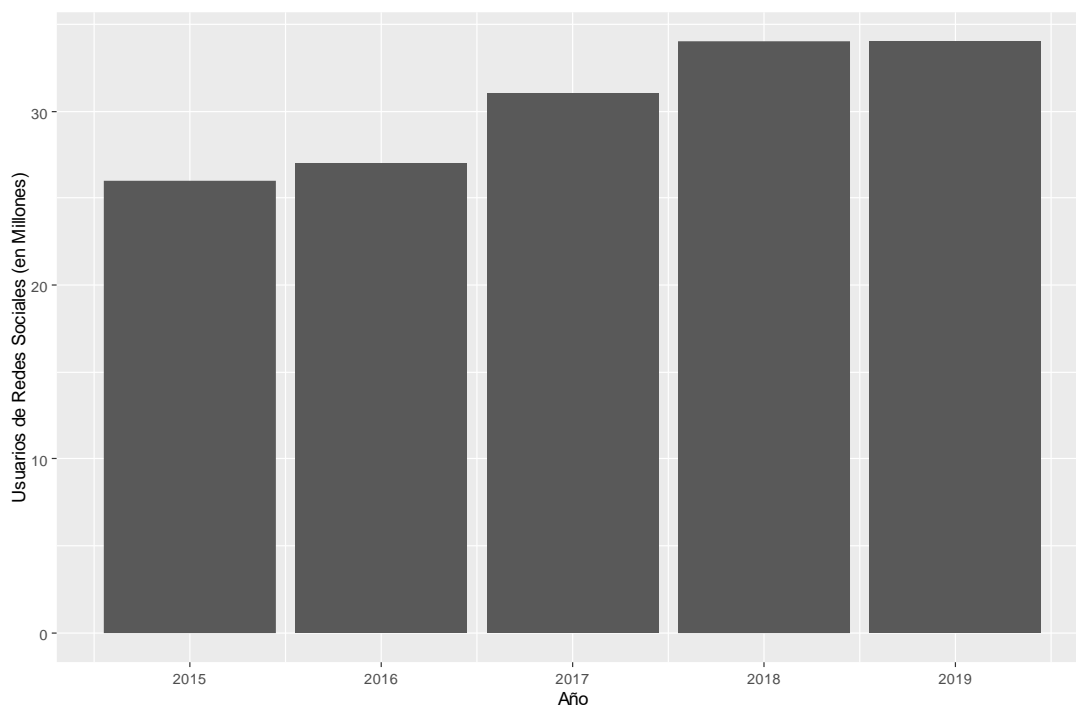
Tal como puede observarse, el crecimiento de la conectividad local es exponencial, significando cada año una penetración más significativa del servicio de internet.

En concordancia con estas condiciones que vuelven posible el uso de redes sociales para la población local, es posible corroborar cómo el uso de las redes sociales, y particularmente de Twitter, se vuelven cada vez más relevantes en los últimos años:

¹ Los datos sobre la población de Argentina por año pueden ser consultados en el sitio web del Banco Mundial (<https://data.worldbank.org/>).

Gráfico 2.

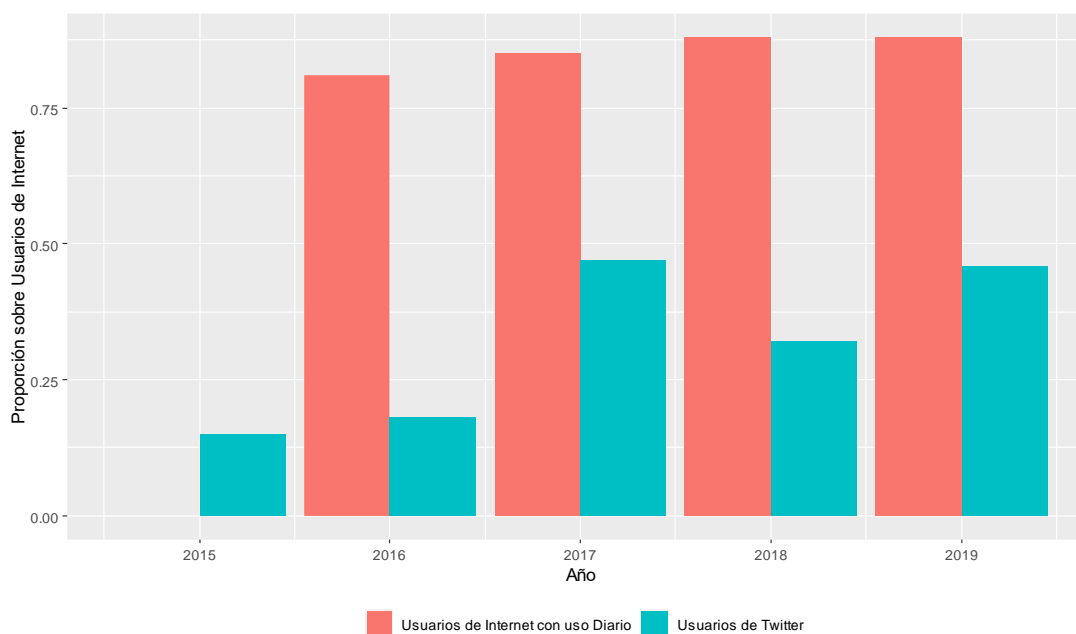
Usuarios de Redes Sociales en Argentina, según Año



Fuente: Elaboración Propia a partir de We Are Social

Gráfico 3.

Uso diario de Internet y Uso de la red social Twitter en Argentina, según Año



Fuente: Elaboración Propia a partir de We Are Social

De esta manera, es posible ver no solo la creciente importancia del uso diario de internet, sino también la creciente relevancia de la red social Twitter a partir del año 2017².

En materia de estado del arte, el presente trabajo retoma los avances en investigación de la opinión pública en base a dos vertientes diferentes: por un lado, considera las vinculaciones realizadas entre Ciencia de Datos e investigación en opinión pública, con el fin de hacer uso de las aplicaciones de nuevas herramientas analíticas a la materia; por otro lado, tiene en cuenta diferentes planteamientos existentes respecto a la relación entre la agenda de la opinión pública -esto es, los tópicos que ésta considera relevantes a la hora de pensar en problemáticas de carácter público o político- y la agenda de los medios de comunicación -es decir, los tópicos que priman en su discurso y mediante los cuales se relaciona con el resto de la sociedad, ya sea que se trate de actores individuales o colectivos-.

En relación a los estudios que toman como objeto la opinión pública a partir de la aplicación de herramientas analíticas propias de las Ciencias de Datos, resaltan dos enfoques diferenciados: aquellos estudios que, por un lado, retoman las funcionalidades analíticas propias de cada sitio web o red social para hacer uso de la información que arrojan, y a partir de allí analizar el comportamiento y la forma de pensamiento de la opinión pública; y, por otro lado, aquellos estudios que se distancian metodológicamente de los primeros, al extraer datos de primera mano desde la web, y a partir de allí realizar sus análisis. En este segundo grupo, por lo general, priman estudios de carácter experimental, mientras que en el primero se cuenta, generalmente, con estudios de carácter descriptivo o exploratorio.

En el primer caso de los dos grupos de estudios descritos, se cuenta con un grupo de investigaciones que realiza una interpretación de las formas de opinión existentes a partir de la comprensión de la producción por parte de los denominados “ciudadanos digitales”. En este sentido, se hace un especial uso

² Incluso frente a la baja en el uso notada en el año 2018, la diferencia entre los niveles promedio de uso de la red social Twitter hasta el año 2016 y después de dicho año, son notables, señalando un corte en dicho año.

de la figura del *prosumer*³, es decir aquel ciudadano que no solo consume contenido, sino que, a su vez, lo produce. En este sentido, se evidencian dos tendencias en materia de investigación en opinión pública, considerando dicha producción digital:

- a) En primer lugar, se cuenta con estudios que hacen uso de la producción de actores, tanto individuales como colectivo, en redes sociales, y que consideran dicha producción como casos testigo para comprender determinados procesos de conformación de posturas ideológicas particulares o procesos de establecimiento de agenda específicos. Dentro de esta vertiente de estudio de la opinión pública en base a la producción de contenido digital, resalta como instrumento analítico central el análisis de contenido a partir de dicha producción en redes sociales.

A modo de ejemplo, cabe destacar el trabajo académico de Natalia Domínguez (2015), quien, a partir de la producción de contenido por parte de actores políticos relevantes y medios de comunicación de carácter masivo, logra realizar un análisis crítico de la relación entre el discurso de ambos actores colectivos mencionados. Particularmente, el trabajo considera la producción de contenido digital por parte de Cristina Fernández de Kirchner en el marco de su campaña presidencial, contrastado con las publicaciones del diario La Nación, con el fin de determinar vínculos entre ambos conjuntos de publicaciones. En este sentido, en dicho trabajo se resalta el uso de las redes sociales, por parte de los actores políticos, como instrumento de campaña en base a un contenido de carácter ideológico, e incluso nota un cambio de comportamiento frente al momento específico de publicación, sea que se trate de un período de campaña o post-campaña; mientras que, por parte de los medios de comunicación, también se detecta un contenido de fuerte contenido ideológico, y sobre todo opositor a la figura política analizada, aunque, en términos metodológicos, en este punto adopta particularmente relevancia el análisis de contenido de notas periodísticas, alejándose de la producción digital.

³ Por “prosumer” se entiende al consumidor de determinado producto, servicio o plataforma, que al mismo tiempo que consume participa en la producción del mismo.

- b) En segundo lugar, se cuenta con un grupo de investigaciones que hace uso de las propias funcionalidades de las redes sociales bajo estudio, en materia de métricas y datos de uso provenientes de dichas herramientas. En este sentido, se trata de estudios de carácter descriptivo sobre aquellos tópicos que resultan relevantes para la opinión pública, en base a la información propia de dichas redes sociales.

A modo de ejemplo, cabe retomar el trabajo de Raquel Rubio García (2014), donde se buscan corroborar las hipótesis planteadas por el enfoque teórico de *agenda setting* mediante el entendimiento de los temas más relevantes para la opinión pública española considerando los tópicos más comentados por los usuarios de la red social Twitter, partiendo de definir a dicha red social como un reflejo de la agenda pública. Dicho trabajo, parte de considerar la relación entre los temas más comentados por los usuarios de Twitter, o *trending topics*, a partir de aplicaciones específicas, para luego contrastar la importancia relativa de dichos temas y la cantidad de apariciones de los mismos en noticias de los dos diarios digitales más leídos en España. En este sentido, cabe señalar que se parte de una categorización manual de dichos *trending topics* en base a un conjunto de categorías que considera, en primer lugar, los bloques característicos en los cuales suele ser dividida la información en los medios periodísticos, a los cuales se adicionan ciertas categorías consideradas como “originarias de Twitter” (Rubio García, *ibid.*: 251). Teniendo en cuenta, de esta manera, las categorías mediante las cuales los medios de comunicación organizan sus noticias, la autora halla una notable correlación entre la agenda mediática y la agenda de la opinión pública.

En el segundo grupo de estudios mencionado, se hallan trabajos que hacen uso de datos crudos de la web, con el fin de estudiar la opinión pública a partir de los mismos. En algunos casos, dicho enfoque se vincula con metodologías de tipo tradicional -por ejemplo, encuestas- para realizar estudios de carácter experimental.

Dentro de este grupo de estudios, resalta el trabajo llevado a cabo por Ernesto Calvo (*Op. Cit.*), o el correspondiente a Natalia Aruguete y Ernesto Calvo (2018). En este último estudio, los autores extraen directamente de la red social Twitter

publicaciones específicas referidos al tema del incremento de las tarifas en Argentina en el año 2016, y realizan a partir de dicho contenido un análisis en base a modelos de supervivencia⁴ con el fin de detectar comunidades de afinidad en base al contenido al cual los usuarios son expuestos. Más específicamente, se modela el tiempo que demoran los usuarios en compartir determinadas publicaciones producidas por miembros de sus propias comunidades en redes sociales, con el fin de evidenciar el tiempo de propagación característico dichas comunidades y corroborar la hipótesis de la existencia de las denominadas *burbujas informativas*⁵. Un aspecto interesante a considerar es que, en el marco de este estudio, se cuenta con un diseño metodológico mixto en el cual se contrastan los resultados obtenidos mediante el análisis de datos observacionales emergentes de las publicaciones extraídas, con datos experimentales obtenidos mediante encuesta. Por otro lado, cabe señalar que dicho experimento también toma como estímulos ciertos atributos propios de la red social evaluada, como la cantidad de seguidores de cada cuenta, el hecho de ser una cuenta verificada o no verificada, y el factor acerca que la publicación cuente con un link hacia un sitio externo. En este sentido, los autores evidencian que, en el proceso de atención selectiva evaluado, las comunidades identificadas interactúan con ideas afines, desarrollando burbujas de información separadas en base a la propagación de publicaciones correspondientes a marcos políticos diferenciados. Una cuestión a resaltar, en este punto, es que se evidencia que el tiempo de propagación de la información en la red social estudiada varía en función de la relación existente entre las características de la comunidad evaluada y el medio de comunicación que figura como autor de la publicación.

De esta manera, considerando los diferentes grupos de estudios mencionados, se identifican las siguientes vertientes metodológicas para el estudio de la opinión pública en base a la incorporación de información proveniente de redes sociales, relacionado a su vez con el tipo de estudio del cual se trate:

⁴ Los análisis de supervivencia examinan y modelan, fundamentalmente, el tiempo que tardan en producirse determinados eventos en un entorno dado.

⁵ Este concepto será desarrollado más adelante en el capítulo, siendo particularmente relevante para este trabajo.

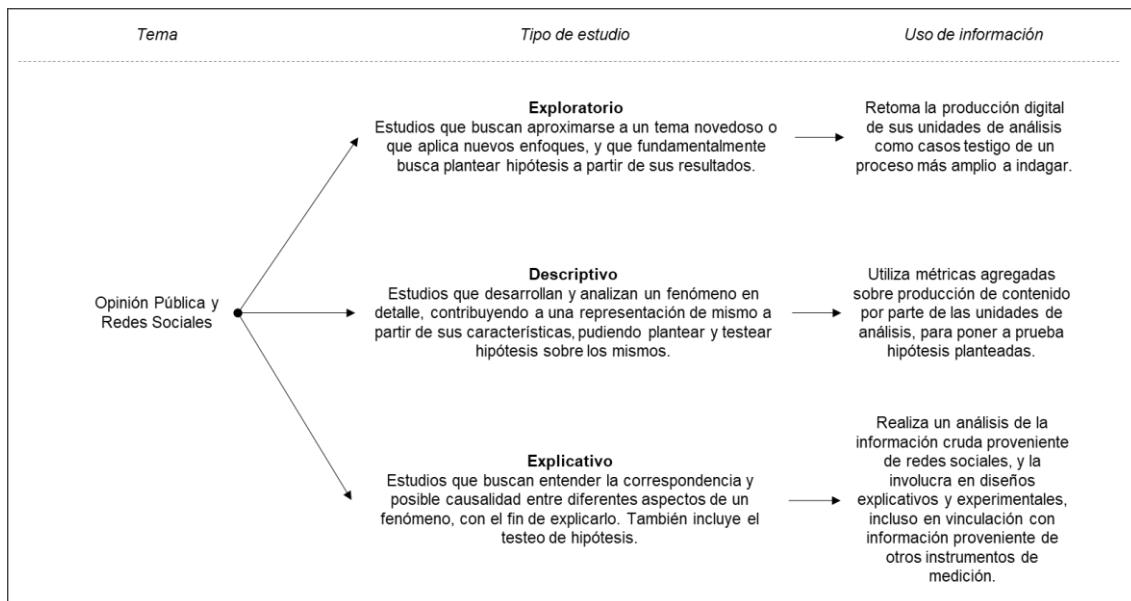


Figura 1. Representación de las vertientes sobre el estudio de la Opinión Pública y Redes Sociales. Fuente: Elaboración Propia

El presente trabajo se enmarca dentro la tercera vertiente de estudios mencionada, caracterizada por la extracción de datos crudos de las redes sociales consideradas y la vinculación con otras metodologías. Particularmente, en el presente trabajo, se desarrollarán categorías emergentes a partir del uso del lenguaje por parte de los medios de comunicación en redes sociales, a la vez del uso del lenguaje por parte de la opinión pública medida mediante encuestas, con el fin de realizar un análisis exploratorio sobre los tópicos predominantes en cada uno de dichos discursos. A su vez, sobre la unidad mínima de este vocabulario -esto es, las palabras integrantes de cada discurso-, será evaluada la relación entre ambos universos discursivos, y la forma que ésta adopta en segmentos o audiencias diferenciadas.

Ahora bien, considerando que se trata de un estudio que tiene foco en el estudio de la opinión pública y su relación con el contenido de los medios de comunicación, cabe desarrollar las herramientas teóricas que permitirán llevar a cabo la interpretación de los resultados alcanzados.

El presente estudio se enmarca en la línea de trabajo que retoma conceptos de las postulaciones sobre *agenda setting*, poniéndolos también en debate frente al

surgimiento de nuevos medios de comunicación digitales y nuevos entornos de expresión, tanto para la opinión pública como para los medios de comunicación tradicionales.

Particularmente referido a los planteamientos de *agenda setting*, si bien no existen desarrollos teóricos que establezcan los pilares de una teoría, se encuentran una diversidad de estudios que establecen, en conjunto, una guía heurística para la investigación aplicada a la opinión pública.

Los principales hallazgos de este grupo de estudios se ubican en torno a los niveles de establecimiento de agenda por parte de los medios de comunicación. En este sentido, se identifican dos niveles, el primero vinculado a las temáticas de conversación relevantes, y el segundo vinculado a atributos sobre dichas temáticas, como lo es la imagen de los candidatos (Canel, 1996; López-Escobar, 1996; López-Escobar, 1996b). Particularmente relevante para este estudio será el primer del establecimiento de agenda, dado que, si bien se comparará el consumo de medios de comunicación con la imagen de actores políticos centrales, no se hará énfasis en los atributos que desarrollan los medios de comunicación sobre dichos candidatos, sino en la coocurrencia de temas, vistos mediante palabras, entre la agenda mediática y la agenda de la opinión pública. Estas aproximaciones presentan una doble vía de entrada al problema del condicionamiento de la opinión pública por parte de los medios de comunicación, en función del actor que se esté investigando.

Al considerar el actor colectivo Opinión Pública, dichos enfoques consideran diseños de investigación tradicionales de carácter cuantitativo, basados en cuestionarios estructurados, donde el *core* de la investigación se centra en la clásica pregunta desarrollada por George Gallup sobre el Problema Más Importante (PMI)⁶ de determinada región, país, provincia o ciudad. Aquí,

⁶ La pregunta PMI consiste en una forma clásica de indagar sobre los principales problemas que, en mirada de la opinión pública, aquejan a determinado territorio nacional. Dicha pregunta, es inaugurada en investigación en opinión pública por George Gallup en la década de 1930 y relevada mensualmente en Estados Unidos por la consultora Gallup desde entonces. Originalmente en su formato "*What do you think is the most important problem facing this country today?*", es una pregunta también retomada por los estudios bajo el enfoque de Agenda Setting, implementado con el fin de poder estudiar la correlación entre la agenda mediática y la agenda de la opinión

encontramos en diversos desarrollos preguntas orientadas a segmentar a la población encuestada, basada en estímulos acerca de su posicionamiento ideológico, su confianza en la política, en los medios de comunicación y en determinadas instituciones, y su grado o intensidad en materia de consumo de medios.

Al considerar el actor Medios de Comunicación, el presente enfoque se caracteriza por llevar a cabo un análisis de contenido para diferentes formatos (prensa escrita, radio o televisión), que es categorizado en función de los objetivos de investigación planteados y el nivel de agenda sobre el que se esté trabajando.

De esta forma, y sobre la información recolectada en base a estos principios metodológicos, se acepta o rechaza la hipótesis de relación entre las temáticas y/o atributos comunicados por los medios de comunicación y las temáticas y/o atributos de mayor relevancia para la opinión pública, ya sea desde un punto de vista general o sobre sus diferentes segmentos. La aceptación o rechazo de dicha hipótesis, por su parte, se basa en el cálculo de un coeficiente *Rho de Spearman* sobre el ranking de importancia relativa elaborado sobre los diferentes estímulos para los dos actores colectivos analizados.

Los estudios basados en el enfoque teórico mencionado establecen una relación continua y en múltiples direcciones entre tres actores colectivos: la opinión pública, los medios de comunicación y los políticos. En el marco de esta interacción, se demuestra que los medios de comunicación presentan un mayor poder de influencia sobre las agendas restantes (la correspondiente a la opinión pública y a los políticos), que por su parte muestran un carácter mayormente reactivo (Aruguete, 2017a; González, G., 2017). No obstante, cabe destacar que el condicionamiento no es directo, sino que se trata de relaciones bidireccionales que presentan condicionamientos mutuos en un conglomerado de fuerzas y contrafuerzas en constante tensión.

En este sentido, cabe considerar que la aparición de las nuevas plataformas de comunicación digitales ha permitido cuestionar ciertos principios establecidos

pública. En el marco de este trabajo en particular, dicha pregunta es adapta a la forma “¿Cuál es, en tu opinión, el principal problema del país hoy en día?”

sobre la dinámica de establecimiento de agenda por parte de los medios de comunicación, ya sea por la aparición de nuevos espacios de expresión como por el cambio en los hábitos de consumo de contenido mediático. En este sentido, los principales cambios en los postulados de estas teorías refieren, sobre todo, a la forma de interacción entre las agendas. Particularmente, se remarca que, si bien la convención sobre el tiempo necesario para el establecimiento de agenda por parte de los medios de comunicación es de tres semanas, en promedio, actualmente este período de tiempo deja de ser un punto de referencia tan estricto, dada la circulación y producción de contenido de manera real. Cabe retomar, en este punto, el cambio de foco en la manera de entender la interacción desde una perspectiva técnica centrada en los medios de comunicación hacia una perspectiva social basada en el uso de dichos sistemas de comunicación y las concepciones políticas en juego, que en este marco se asocian a oportunidades de expresión gubernamental o ciudadana en una sociedad dada (Arugete, *op. Cit.*). Este será un enfoque particularmente relevante para el presente trabajo, dado que, si bien se considera el período tradicional de tiempo de establecimiento de agenda con fines metodológicos, la relación entre opinión pública y medios de comunicación es interpretada a partir del posicionamiento político de los ciudadanos.

Adicionalmente, desde la perspectiva teórica de *agenda building*, podemos señalar ciertos principios que colaboran con el establecimiento de agenda por parte de los medios de comunicación, que son producto de la interacción entre medios y el establecimiento de agenda mutuo, y contribuyen a la construcción de cierta *homogeneidad* por parte de la agenda mediática en su conjunto -esto es, sin importar particularmente qué medio de comunicación esté comunicando-. Estos principios son tres, caracterizados de la siguiente manera:

- a) el principio de *Focusing*, referido a que los medios centren la atención en los mismos temas;
- b) el principio de *Consonance*, referido a un encuadre similar por parte de estos temas por parte de los medios de comunicación;
- c) el principio de *Persistence*, que refiere a que estos temas sean cubiertos con perseverancia.

Estos tres vectores, actuando en conjunto, redundan en la homogeneización o estandarización del contenido mediático.

A su vez, cabe destacar que, siguiendo a Natalia Aruguete (2017b), los medios de comunicación son entendidos también como actores políticos con intereses y metas particulares que interactúan con otros, en el marco de una negociación por la confirmación de los sentidos públicos, configuración en la cual se vuelve particularmente relevante el “menú de temas” (Aruguete, *ibid.*) tratado. Si bien esta forma de entender los medios de comunicación es parte del *corpus* teórico de la perspectiva de *agenda building* -esto es, la cuarta fase de la teoría de *Agenda Setting*⁷-, resulta particularmente relevante para el presente trabajo el entender a la agenda mediática no tanto como un factor explicativo de la percepción de las audiencias, sino más bien como una variable resultante de un conjunto de factores institucionales, valores y hábitos que adopta una complejidad particular, de lo cual resulta que sea igualada a la agenda de las audiencias, a nivel analítico, sin imponer órdenes de causalidad. En este sentido, el planteo sobre el *establecimiento* de la agenda pasa a ser más el de la *construcción* de la misma, indicando cierta reciprocidad entre medios de comunicación, tomadores de decisión y público. Cabe retomar, de esta manera, el modelo mencionado por la autora:

El matrimonio Lang propone un modelo de cuatro pasos en el que quedan englobados los procesos de agenda setting y de agenda building. Primero, los medios destacan ciertos eventos, actividades, grupos o personalidades. Luego, los elementos de un conflicto se combinan en un marco común, subrayando o escondiendo ciertos aspectos del objeto. En tercer lugar, el tema en cuestión se asocia a símbolos secundarios, convirtiéndose en parte de un paisaje político reconocido. Finalmente, los portavoces juegan un papel

⁷ Natalia Aruguete (*Op. Cit.*) identifica cuatro etapas diferenciadas de desarrollo de la teoría de *agenda setting*, a saber: 1) se basa en el momento de confirmación del efecto cognitivo de los medios de comunicación sobre los votantes; 2) se basa en la introducción de las condiciones contingentes que mediaban la relación entre medios y públicos; 3) es la fase que abre paso a la *attribute agenda setting*, donde la influencia de los medios de comunicación pasa de ser cognitiva a actitudinal; 4) supone un giro copernicano de la perspectiva teórica, pasando del análisis de los efectos mediáticos en la agenda pública a una exploración sobre el proceso mismo de construcción de la agenda de los medios de comunicación. Aquí, se apunta a los factores que moldean las coberturas mediáticas.

importante en la promoción de los temas y los símbolos. *La primera etapa corresponde al establecimiento de la agenda. La segunda, tercera y cuarta, a la construcción de la agenda.* (Aruguete, N., *íbid.*, p. 40). Las cursivas son propias).

En este sentido, cabe señalar que el presente trabajo tiene foco en el primer paso señalado, considerando que se hará énfasis en la comunicación de eventos, actividades y personalidades por parte de los medios de comunicación, en primera instancia. De esta manera, en términos de Aruguete (*íbid.*), se toma como objeto a los “hacedores” de noticias o *agenda setters*. Ahora bien, considerando que la combinación de estos elementos se da en el marco de un lenguaje común, el análisis de la interacción entre el universo discursivo de la opinión pública y de los medios de comunicación, permitirá obtener una aproximación a la construcción de ese universo discursivo, expresado en tópicos y palabras específicas, que denominamos agenda. Por otro lado, el incorporar en el análisis características propias de los públicos como variables, permitirán entender a este último grupo no como un simple receptor, sino como un actor colectivo con intereses y valores propios, a partir de los cuales define sus patrones de consumo y elabora sus propios discursos.

Gráficamente, la definición de la agenda pública a partir de los tres actores mencionados, y el foco de este trabajo, se presentan de la siguiente manera:

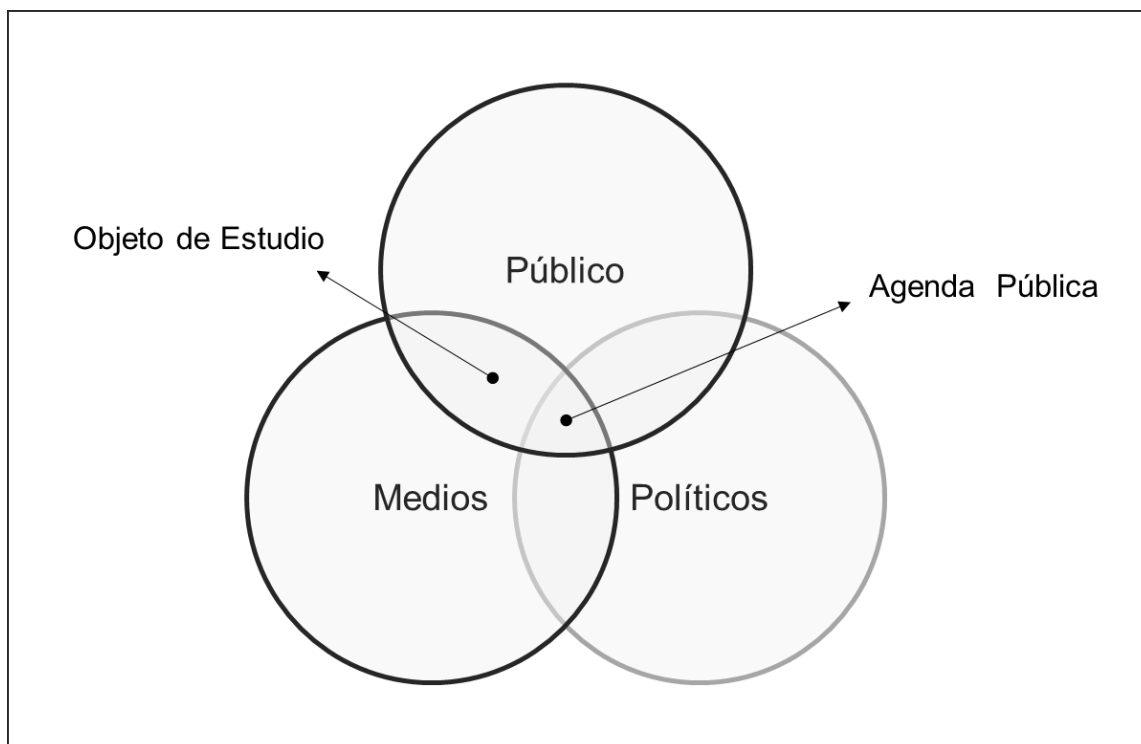


Figura 2. Conformación de agenda y Objeto de Estudio. Fuente: Elaboración Propia

Complementariamente a estos planteos, cabe destacar los hallazgos de Ernesto Calvo (*Op. Cit.*), quien resalta que, pese a su presentación como una red social horizontal, existen mecanismos jerárquicos que consolidan el posicionamiento social de determinados actores políticos, particularmente líderes políticos y medios de comunicación, basados a su vez en ciertas dinámicas de “identificación” entre los medios de comunicación masivos y las redes sociales, particularmente, mediante los cuales la información difundida en las redes sociales encuentra un alto nivel de consistencia con la transmitida en los medios de comunicación tradicionales (Aruguete, 2017a). En palabras del investigador:

La reputación importa, esa es la conclusión principal que uno debería extraer. Cuando se produce información de forma masiva, con millones de usuarios, no toda la información es creíble. La información que tiene validación autoritativa (medios, políticos, artistas, etc.) es repetida por las redes afines con mucha mayor frecuencia que la de los otros usuarios. Si bien hay millones de usuarios, la información de tan sólo algunos de ellos es retwitteada masivamente. En este último año hemos podido verificar que este fenómeno de alta concentración masiva de información se produce en

la gran mayoría de los conflictos políticos que son vehiculizados por Twitter. (Ernesto Calvo en Slimovich, A. *et al.*, 2016, p. 194).

Por último, respecto a los principios teóricos retomados en el presente trabajo, incorporamos el concepto de *agenda melding* o fusión de agendas (Calvo, E. y Aruguete, N., *op. Cit.*), desarrollado por Weaver, McCombs y Shaw (2004). Aquí, se explicita una dinámica de reducción del efecto de disonancia cognitiva por parte de las audiencias, que tiende a mimetizarse conformando comunidades que comparten valores e intereses. Este efecto, se ve maximizado gracias a las nuevas tecnologías, dado que las comunidades de pertenencia conformadas y su respectiva agenda colectiva no se limita al consumo de medios tradicionales, sino que se completa con la conectividad con otras personas y fuentes de información que se asemejan por afinidad social. Bajo estos supuestos, los individuos reciben y procesan información de acuerdo a sus creencias ideológicas previas y el grado de conectividad que mantengan con sujetos afines, intensificando los efectos de establecimiento de agenda, con sus diferentes niveles correspondientes, al interior de cada comunidad específica.

Dicha agrupación de los individuos en comunidades de intereses y valores se ve reflejado, a su vez, en el concepto de “cámara de eco” entrañado por Ernesto Calvo (Slimovich, A. *et al.*, *op. Cit.*). Dicho concepto refiere a la manera en que ideas e información son amplificadas en una comunidad dada, obteniendo como resultado que los individuos reciben una versión amplificada de información de la cual ya disponen previamente. En este sentido, mediante este mecanismo de especialización en medios de comunicación sobre los cuales ya se tiene afinidad, la opinión pública se ve expuesta a una versión amplificada de sus propios preconceptos. Cabe destacar que, si bien pensado para el análisis de medios de comunicación tradicionales, según el autor dicho concepto también aplica al análisis en el marco de la revolución tecnológica, solo que en el marco de un entramado de relaciones sociales más complejo donde los usuarios no operan pasivamente, sino que a su vez son productores activos de información.

En línea con lo planteado, al considerar esta red social como el escenario a analizar, donde se emplazan los discursos políticos anteriormente mencionados,

cabe considerar las distintas “geografías” que la definen (Calvo, *op. Cit.*). Relacionado con los conceptos de “burbuja informativa” y “cámara de eco” ya introducidos, se señala que las diferentes zonas que componen a Twitter se encuentran definidas por sus propios grupos de usuarios. En este sentido, cuando las comunidades cobran mayor intensidad y sinergia, esto es, cuando el número de conexiones es mayor, también es mayor la segregación de las narrativas a las cuales sus miembros son expuestos. En este sentido, las ideas de cada usuario sobre el mundo que lo rodea va filtrando el contenido al cual accede y genera “narrativas localmente mayoritarias” (Calvo, *íbid.*, p. 27), lo cual genera una diversidad de ideas disponibles que tienden a reconfirmar los propios preconceptos políticos, y que por ende marca una invitación a considerar la exposición a dicho contenido en vinculación con mediciones que den evidencia sobre el propio posicionamiento político de los usuarios, sus actitudes frente a las diferentes personalidades políticas y los medios de comunicación.

Considerando estos aspectos teóricos, a continuación, se desarrolla la estrategia metodológica del presente trabajo, principios mediante los cuales será abordado el problema introducido.

CAPÍTULO 2. METODOLOGÍA

En el presente apartado, será desarrollado en planteo metodológico del presente trabajo. Con dicha finalidad, se volverán explícitos el problema de investigación abordado; el objetivo general y los objetivos específicos del documento; la organización del estudio en términos de las etapas que fueron consideradas; las diferentes estrategias metodológicas utilizadas para el relevamiento y análisis de, por un lado, el discurso mediático y, por otro, el correspondiente a la opinión pública; y, por último, las herramientas analíticas utilizadas para indagar sobre la relación entre ambos universos discursivos.

Problema de Investigación y Objetivos

El *problema de investigación* planteado es comprender la relación existente entre los temas de agenda establecidos por los medios de comunicación, y los principales temas de agenda pública señalado por parte de la opinión pública, a partir de su unidad discursiva mínima: la palabra. Este problema, será abordado para el conjunto de Argentina, durante los meses de noviembre y diciembre del año 2018.

El trabajo se circunscribe en el marco del desarrollo local de la investigación en opinión pública y medios de comunicación, aunque busca complementar dicho ámbito de trabajo académico y profesional con la incorporación de desarrollos en Tecnología de la Información, particularmente en Ciencias de Datos.

Desde el presente trabajo se plantea que las publicaciones de los medios de comunicación digitales en Twitter constituyen un punto de acceso al contenido que producen, fundamentalmente porque sus titulares son replicados en su totalidad en dicha red social, y que dada la capacidad de procesamiento informático actual es posible ser extraído y analizado, pero que para evidenciar la relación entre dicho contenido y la perspectiva de la opinión pública, es válido vincular estas herramientas metodológicas con instrumentos de investigación social tradicionales que permitan comprender la jerarquización de temas realizada por la opinión pública, identificando a su vez diferentes segmentos dentro de ella, complementando la propia información brindada por la red social.

Por este motivo, se plantea la problemática sobre la existencia de correlación estadística entre los temas de conversación abordados por los principales medios de comunicación digitales, evidenciados mediante un análisis de contenido de carácter cuantitativo en función de sus publicaciones en la red social Twitter, y los temas mayormente relevantes a nivel país planteados por la población digital argentina, identificados mediante un cuestionario semiestructurado realizado de manera online, para los meses de noviembre y diciembre del año 2018. Dicha relación, como ya fue mencionado, será abordada mediante la unidad mínima evidenciada en su discurso: las palabras.

El problema de investigación anteriormente presentado encuentra su justificación en el hecho que, si bien se encuentran en curso estudios locales que hacen uso de nuevas tecnologías para acceder a contenidos mediáticos, e incluso vinculan sus resultados con metodologías de carácter tradicional -como estudios experimentales mediante encuestas-, no se hallan trabajos que hagan uso de dicho abordaje metodológico de cara a corroborar o rechazar parcialmente la relación existente entre los temas de agenda identificados por los medios de comunicación y por la opinión pública, relación ya abordada, por su parte, desde diversos puntos de vista en el ámbito de las ciencias sociales.

En este sentido, la justificación del presente trabajo es enteramente de carácter metodológico. Su motivación se halla en la inquietud por realizar una propuesta metodológica en el marco del estudio local de la opinión pública y su relación con el contenido mediático, a partir de un proceso de *web scrapping* y análisis de información textual posibilitado por los recientes desarrollos en softwares de procesamiento estadístico, pero en vinculación también con una metodología de tipo tradicional para abordar a la opinión pública, como es un cuestionario semiestructurado realizado de forma online. El principal aporte del presente trabajo, de esta forma, se halla en presentar una vinculación entre metodologías tradicionales y herramientas analíticas de reciente desarrollo, de cara a seguir fortaleciendo el estudio de la opinión pública mediante la incorporación de herramientas propias del ámbito de la ciencia de datos.

Como *objetivo general*, entonces, el presente trabajo busca indagar la forma que adopta la relación entre la agenda de los medios digitales seleccionados, cuyo

contenido es accesible a través de la red social Twitter, y la agenda de la opinión pública entre la población digital argentina, para el conjunto del territorio de Argentina, en los meses de noviembre y diciembre del año 2018.

Como *objetivos específicos*, el presente trabajo se plantea evidenciar:

- (a) Los principales temas abordados por los medios de comunicación definidos, a partir de sus noticias compartidas en Twitter;
- (b) La variación de los principales temas abordados, en función de cada medio;
- (c) Los principales medios de comunicación digitales consumidos por la población encuestada;
- (d) El Problema más importante (PMI) del país para la población digital argentina, encontrando expresión en las respuestas brindadas por parte de la muestra encuestada;
- (e) El nivel de confianza existente sobre los medios de comunicación, por parte de la muestra encuestada;
- (f) El posicionamiento político de la muestra encuestada, identificado en función de la imagen sobre los actores políticos más relevantes al momento de realizar el estudio;
- (g) La intensidad de la relación existente entre ambos universos discursivos, a partir de la coocurrencia de palabras, y su profundización en audiencias específicas.
- (h) La relación entre la intensidad que adopta la correlación entre ambos universos discursivos, y las restantes variables características de las audiencias consideradas.

Estrategia de Recolección de Datos

Para la realización del presente proyecto, se realizó un análisis en dos etapas superpuestas. En primer lugar, se llevó a cabo un proceso de *web scrapping* para extraer todos los tweets publicados por los medios de comunicación seleccionados, en un marco de tiempo determinado.

Paralelamente a este proceso, y como estrategia metodológica considerando los mecanismos de *framing*⁸ de los medios de comunicación seleccionados, fue elegido un punto dentro de la ventana de tiempo del proceso de extracción, para realizar un estudio cuantitativo que indagara sobre los problemas más relevantes para la población digital argentina⁹. Dicho marco temporal correspondiente al estudio cuantitativo, fue seleccionado considerando el consenso existente sobre los tiempos de establecimiento de tópicos de conversación en la agenda pública por parte de los medios de comunicación -el tiempo promedio de establecimiento de agenda comprobado es de tres semanas, siendo el rango establecido entre una y ocho semanas (Aruguete, 2017a).

En términos temporales, el trabajo de campo fue desarrollado de la siguiente manera:

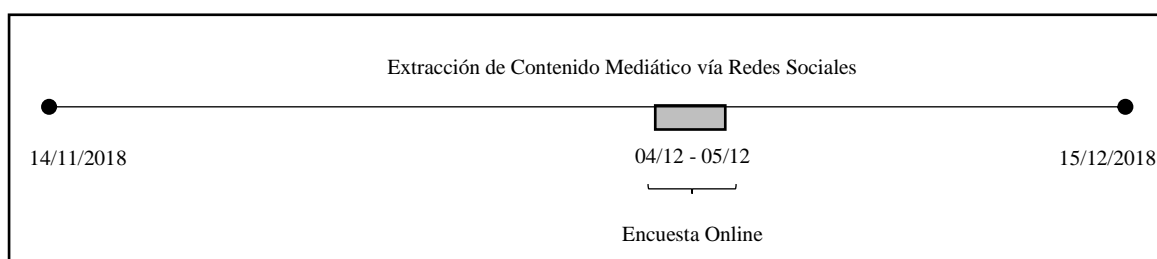


Figura 3. Organización temporal del trabajo de campo. Fuente: Elaboración Propia

En cuanto al proceso de extracción de publicaciones en redes sociales por parte de los medios de comunicación, la red social elegida fue Twitter, dado el hecho que los medios de comunicación replican los titulares publicados en sus portales online en dicho canal. Por otro lado, para la selección de medios de comunicación, fue considerado el nivel de consumo de manera online, en primer lugar, pero también el hecho de considerar un espectro ideológico amplio -en términos políticos-, con el fin de hallar representadas a una diversidad de

⁸ El concepto de *framing* o *encuadre* refiere, en ciencias sociales, a la influencia de los marcos interpretativos sobre las opiniones y decisiones de las personas. En este sentido, se considera el encuadre por parte del discurso de los medios de comunicación, por lo que se tiene en cuenta el tiempo estimado por trabajos previos, para el establecimiento de temas en agenda por parte de estos últimos.

⁹ Para un mayor desarrollo sobre la pregunta PMI, consultar el capítulo "Metodología" del presente documento.

audiencias. En este sentido, los medios de comunicación relevados fueron Infobae, Clarín, La Nación, Página 12 y Perfil, todos mediante el mismo canal de comunicación.

A continuación, se presentan datos elaborados por Reuters Institute y la Universidad de Oxford (2019), donde se señala la penetración de cada medio de comunicación seleccionado, mediante canales digitales, respecto a la población argentina para el año 2018. A su vez, se muestra la posición correspondiente en el ranking de medios de comunicación más consumidos de manera online:

Tabla 1.

Datos sobre consumo de medios de comunicación online seleccionados. Año 2018¹⁰

Medio de Comunicación	Penetración	Ranking
Infobae	42%	1
Clarín	30%	3
La Nación	22%	4
Página 12	9%	9
Perfil ¹¹	SD	SD

Fuente: Elaboración Propia en base a University of Oxford (2019)

Complementariamente, el estudio cuantitativo llevado a cabo tuvo, como objetivo general, caracterizar el problema más importante del país para la población digital mayor de 16 años, consumidora de noticias a través de canales online, para el año 2018. Dicho estudio, consideró como Universo a la población usuaria de internet consumidora de noticias a través de medios online, mayor de 16 años, NSE A-B-C-D, residentes de AMBA e Interior.

¹⁰ Se presentan datos para los medios de comunicación seleccionados. Cabe destacar que el medio de comunicación que ocupa la posición número 2 del ranking es TN, parte del Grupo Económico Clarín S.A.

¹¹ Cabe destacar que, si bien no se dispone de datos sobre penetración para el medio de comunicación Perfil, dicha marca informativa se encuentra entre las 12 marcas que representan mayor confianza para la opinión pública local, sin distinguir entre plataforma de comunicación (esto es, radio, televisión, prensa escrita, entre otras), según el estudio citado.

En cuanto al diseño muestral, basado en un nivel de confianza establecido de 95,5%, un error muestral de 5%, entendiendo que la población usuaria de internet a nivel país es de 34,79 millones (We Are Social, 2017), y sabiendo que el tamaño de la muestra se encuentra definido por

$$n = \frac{N}{1 + \frac{e^2(N-1)}{z^2 pq}}$$

Donde:

Z representa el valor Z correspondiente a un determinado nivel de confianza;

$p * q$ representa la varianza o heterogeneidad asumida de la población;

e representa el error muestral considerado; y

N representa el tamaño poblacional conocido¹²,

se define una muestra a considerar de 400 casos, seleccionados aleatoriamente dentro de la comunidad del panel online “*Oh! Panel*”.

Métodos analíticos para el análisis de texto

Posteriormente al trabajo de campo, fue realizado un análisis de los datos relevados en dos instancias. En primer lugar, fue llevado a cabo un análisis de contenido cuantitativo, tanto respecto a las publicaciones en redes sociales por parte de los medios de comunicación como respecto a las respuestas abiertas de la opinión pública, con el fin de hallar los tópicos discursivos relevantes para cada uno de dichos actores colectivos. En segunda instancia, dichos universos discursivos fueron contrastados a partir de la correlación existente entre las unidades mínimas de cada uno de ellos, es decir, sus palabras.

¹² La fórmula a la que se hace referencia es introducida en Morales Vallejo, P. (2012). Se utiliza la fórmula para la determinación del tamaño muestral sobre una población de dimensiones conocidas, y se aplica el redondeo a 400 casos (frente a una fórmula resultante en 384 casos necesarios) para asegurar la correcta performance de la muestra, una práctica ampliamente utilizada en la industria de la investigación social aplicada.

Para el análisis de contenido, en ambos casos y dado el volumen de información relevado para ambas audiencias, fueron utilizadas herramientas propias del procesamiento de lenguaje natural en el ámbito de la ciencia de datos. En este sentido, en el marco del análisis cuantitativo de texto en ciencias sociales, se identifican una serie de metodologías analíticas diferenciadas, en referencia a dicho tipo de datos, que son identificadas exhaustivamente en Grimmer y Stewart (*Op. Cit.*), a saber:

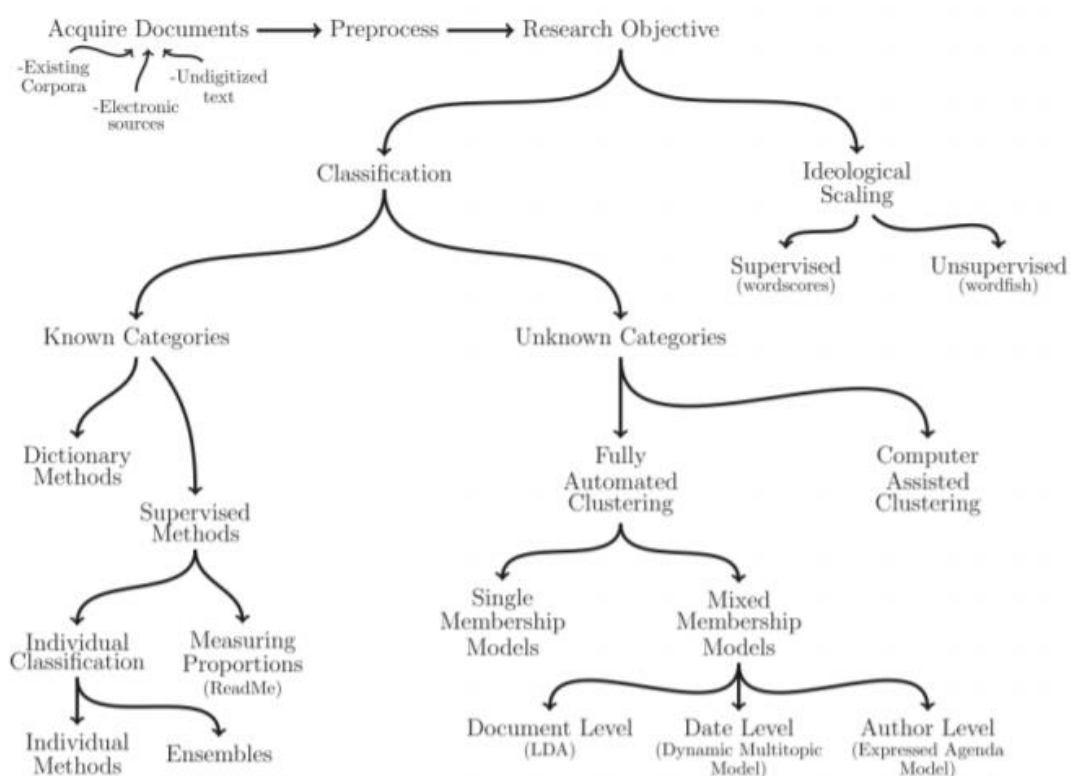


Figura 4. Resumen de herramientas metodológicas aplicables a datos de tipo textual. Fuente: Grimmer y Stewart (2013)

Partiendo del trabajo anteriormente citado, cabe considerar sus advertencias metodológicas a la hora de utilizar este tipo de herramientas. Siguiendo a los autores, los métodos anteriormente nombrados deben ser considerados como herramientas para *amplificar* y *aumentar* un análisis cuidadoso, pero no métodos cuyos resultados puede considerarse como esencialmente correctos. De hecho, los autores afirman que los métodos automatizados aplicados al lenguaje natural son esencialmente *modelos incorrectos de lenguaje*, por lo que la validación por

parte del analista se vuelve un aspecto prioritario, considerando que la correcta performance de estos modelos sobre un set de datos específico no puede ser garantizado de antemano.

Cabe resaltar, de esta manera, los cuatro principios mencionados por los autores para el análisis cuantitativo de texto:

- (a) Todos los modelos cuantitativos sobre el lenguaje son incorrectos, pero algunos de ellos son útiles;
- (b) Los modelos cuantitativos para textos amplifican los recursos y aumenta la mirada de los analistas;
- (c) No hay un método globalmente correcto para el análisis automatizado de texto;
- (d) Validar, validar, validar.¹³

En este sentido, la complejidad propia del lenguaje implica que todos los métodos necesariamente fallan en proveer un modelo completamente asertivo sobre el lenguaje que tratan, aunque brindan herramientas útiles para realizar inferencias productivas sobre determinado set de datos. Así, la principal utilidad de estos métodos resulta en asignar documentos a determinadas características -en algunos casos emergentes y en otros preestablecidas-, muchas veces permitiendo descubrir nuevas categorías y formas de ordenar los documentos bajo análisis.

Teniendo en cuenta estas consideraciones, cabe destacar la necesidad de validar los resultados obtenidos mediante el modelo cuantitativo seleccionado, considerando que la ganancia obtenida por parte de los analistas en términos de tiempo y recursos necesarios para procesar el conjunto de documentos en consideración, puede verse mermado por la falta de validez de dichos resultados.

La validación de los tópicos o categorías obtenidas, por su parte, se determina fundamentalmente a partir del *qué* de cada una de ellas, es decir de su contenido en materia del texto que contienen. Generalmente, la forma de validar los tópicos obtenidos es extraer sobre cada categoría los diez o quince textos más representativos de cada uno, y evaluar su validez semántica -esto es, que el analista encuentre tanto coherencia interna entre los documentos que componen

¹³ Cfr. Grimmer *et al.*, *Op Cit.*, p. 269. Traducción propia.

un tema, como también cierta heterogeneidad respecto a otros documentos externos, que permita establecer fronteras válidas entre categorías-.

Siguiendo el planteo metodológico de los autores, el presente trabajo se basa en la aplicación de algoritmos ligados al *topic modeling*, partiendo de un objetivo de investigación basado en la clasificación de texto en función de categorías desconocidas. Particularmente, al tratarse de una clusterización completamente automática a nivel de documento (esto es, a nivel de cada tweet o respuesta proveniente de la encuesta), se implementa el algoritmo más común en este tipo de usos: el Latent Dirichlet Allocation o LDA (Barberá, P., 2018).

Los algoritmos de este tipo son particularmente útiles a la hora de descubrir los principales temas o tópicos en el marco de un *corpus* de texto no estructurado. En este sentido, son utilizados para organizar una colección de textos en el marco de categorías emergentes de los propios datos, dependiendo del investigador únicamente en la decisión acerca de la cantidad de tópicos a hallar en el texto. Particularmente, el algoritmo LDA provee un modelo generativo que describe cómo fueron creados los documentos de un set de datos determinados, en función de sus unidades mínimas: las palabras. En este sentido, cabe destacar que: a) cada tópico es una distribución de probabilidad específica sobre un vocabulario definido fijo, mientras que, b) cada documento es una colección de palabras, generado a partir de una distribución de tipo multinomial para cada uno de los tópicos definidos, lo cual lo define como una mezcla particular de tópicos latentes.

De esta manera, se define un modelo probabilístico en base a tres instancias:

1. Se define $\theta_i \sim \text{Dirichlet}(\alpha)$;
2. Se define $\beta_k \sim \text{Dirichlet}(\delta)$;
3. Para cada palabra en cada documento i :
 - a. Se define un tópico $z_m \sim \text{Multinomial}(\theta_i)$;
 - b. Se define una palabra $w_{im} \sim \text{Multinomial}(\beta_{i,k=z_m})$.

Donde:

α representa el parámetro de Dirichlet previo a la distribución de tópicos sobre documentos;

θ_i representa la distribución de probabilidad sobre los diferentes tópicos, para cada documento i ;

δ representa el parámetro de Dirichlet previo a la distribución de palabras sobre los tópicos;

β_k representa la distribución de probabilidad sobre el vocabulario utilizado en un cuerpo de documentos, para cada tópico k -o “distribución tema-palabra”-;

z_m representa la asignación de un tópico latente a cada documento;

w_{im} representa la asignación de un tópico latente a cada palabra utilizada en los diferentes documentos.

Gráficamente, dicho modelo se expresa de la siguiente manera:

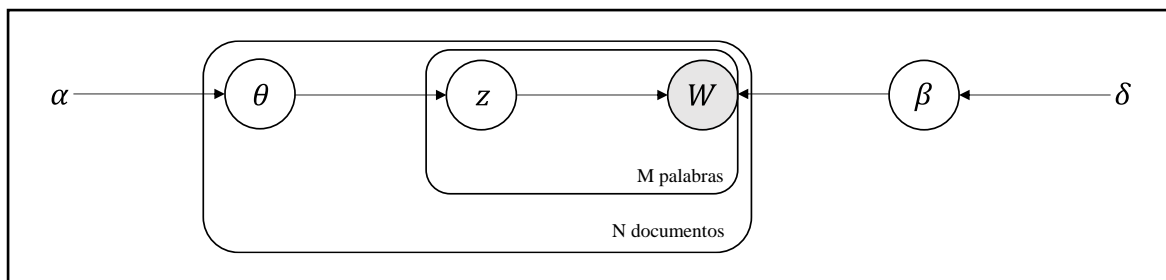


Figura 5. Representación gráfica del modelo LDA. Fuente: Elaboración Propia

De donde se infiere que β refiere a probabilidad de pertenencia a cada tópico k por parte de cada palabra m ; θ refiere a la probabilidad de pertenencia a cada tópico k por parte de cada documento i ; y Z refiere a la asignación de un tópico k para cada palabra w .

En el marco de la reducción de complejidad que implica moverse de textos a categorías, cada unidad de análisis es un *texto* o *documento*, que puede ser un tweet, una publicación en otra red social, una respuesta abierta en una encuesta, u otro elemento de texto; mientras que la población de textos es denominada *corpus*.

A fines del análisis, los modelos cuantitativos de texto eliminan el orden existente en las palabras en el marco de los documentos que las integran, y los tratan como *bolsas de palabras*. Cabe destacar, en este sentido, que las palabras únicas o *unigrams* son consideradas suficientes a la hora de evidenciar el sentido de una pieza de este tipo de texto en estos análisis. Posteriormente a descartar el orden en las palabras, se procede a limpiar el texto que será objeto de análisis, descartando elementos de puntuación, letras mayúsculas, palabras comunes en cada lenguaje -esto es, las denominadas “stopwords” o palabras que no brindan significado específico por sí mismas, como “de”, “la”, “para”, entre otras-, y palabras particularmente infrecuentes -por ejemplo, aquellas que aparecen en menos del 1% de los textos que componen determinado *corpus* -.

El resultado de pre-procesar el texto de esta manera, es que cada documento i ($i = 1, \dots, N$) es representado en un vector que cuenta la cantidad de ocurrencias de cada palabra única M , siendo este vector $W_i = (W_{i1}, W_{i2}, \dots, W_{iM})$. Cada W_{iM} cuenta el número de ocurrencia de las m palabras en cada documento i . Por último, la unificación de esta colección de vectores en una matriz resulta es una *document-term matrix*, utilizada para el análisis. Si bien este tratamiento del texto lleva a una particular reducción de la información contenida en los documentos, se encuentra demostrado que esta representación simple de los textos es suficiente para inferir propiedades sustantivamente interesantes sobre los mismos (Grimmer *et al.*, *Op. Cit.*).

En el marco de este análisis, queda bajo criterio del investigador elegir una cantidad de tópicos k a la hora de realizar el análisis de texto en base a dicha metodología. En este sentido, Barberá, P. (2017) introduce la posibilidad de determinar dicha cantidad a partir de una técnica de validación cruzada donde se estima una medida de *perplejidad*, que indica cuán bien una distribución de probabilidad predice una muestra. Cuando es utilizada para comprar modelos de probabilidad, un menor valor del indicador indica que la distribución de probabilidad permite predecir correctamente la muestra. En este trabajo en particular, se elige una cantidad de tópicos k que minimice sustantivamente dicha medida en relación a las restantes cantidades de grupos evaluadas. Gráficamente, se observará un punto de inflexión en la curva de valores de dicho indicador, sobre una cantidad de tópicos dada.

Métodos para el análisis comparado de universos discursivos

Por último, resta considerar la manera de comparar ambos universos discursivos. Considerando la unidad mínima de dichos discursos, la comparación será realizada a partir de la coocurrencia de palabras entre ambos¹⁴. En este sentido, el análisis se basa en identificar la frecuencia de cada palabra para cada actor colectivo, sobre el conjunto de palabras coincidentes entre ambos universos discursivos, con el fin de evaluar su correlación. Así, sobre una matriz genérica A conformada por las m palabras coincidentes como unidades de análisis, cada n universo discursivo como variable, y su frecuencia relativa a como valores, se calcula el coeficiente de correlación de Pearson correspondiente. A saber, dicha matriz adquiere la forma:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Partiendo de la consideración que $n = 2$ en todos los casos, es decir que siempre se realizarán comparaciones de a pares, y sabiendo que siempre habrá un vector op referido a la frecuencia relativa de cada palabra coincidente para el segmento de la opinión pública seleccionado para la comparación, y otro vector mc correspondiente a los valores del segmento de medios de comunicación seleccionado. Sobre dicha matriz genérica, se estima cada coeficiente de correlación de Pearson correspondiente, a saber:

¹⁴ Cabe destacar, en este punto, que un enfoque posible, dado el marco teórico considerado y los usuales análisis en la investigación en opinión pública, hubiera sido el corroborar la correlación entre la importancia relativa de los diferentes temas que componen cada universo discursivo. Ahora bien, tal como se verá a continuación, los temas discursivos predominantes en cada actor colectivo no son coincidentes, por lo que el análisis debe buscar un nivel mayor de granularidad para comprender acabadamente la relación entre dichos universos discursivos. Esto se encuentra justificado, por su parte, en el hecho que los medios de comunicación resaltan hechos o personalidades, por lo general, mientras que la opinión pública tiende a sostener un discurso mayormente asentado y de carácter general, tal como se desarrolló en el capítulo previo.

$$\rho_{op,mc} = \frac{\sigma_{opmc}}{\sigma_{op}\sigma_{mc}}$$

Donde *op* refiere al vector de la frecuencia relativa para la opinión pública sobre las palabras coincidentes, mientras que *mc* refiere a un vector de las mismas características, correspondiente a los medios de comunicación seleccionados para la comparación.

En base a dicho indicador, será posible evidenciar la correlación existente entre los universos discursivos comparados. Por otro lado, a partir de la representación gráfica de la coocurrencia de dichos discursos en base a sus unidades mínimas, será posible evidenciar las palabras que mayormente representan dicha relación, en caso de ser requerido.

Resta considerar, por otro lado, la manera en que dicha correlación será evaluada. En este sentido, cabe destacar que se definen audiencias en función de los medios de comunicación online que declaran consumir los entrevistados, para luego profundizar en la correlación específica que mantengan los universos discursivos hacia el interior de dichos grupos, y la relación de este indicador con otras variables relevadas que expresen su posicionamiento político, tanto respecto a los medios de comunicación como a actores políticos centrales.

Particularmente, será evaluada la relación entre el coeficiente de correlación obtenido y la respuesta de los entrevistados a cinco preguntas, a saber:

- a. “En una escala del 1 al 10, siendo 1 nada de confianza y 10 total confianza, te pedimos que indiques cuánta confianza te generan los medios de comunicación online en general”;
- b. “¿Con qué frecuencia consumís estos medios online?”;
- c. “En una escala del 1 al 10, siendo 1 totalmente negativa y 10 totalmente positiva, te pedimos que nos indiques la imagen que tenés de Mauricio Macri”;
- d. “En una escala del 1 al 10, siendo 1 totalmente negativa y 10 totalmente positiva, te pedimos que nos indiques la imagen que tenés de Cristina Fernández de Kirchner”;

e. “En una escala del 1 al 10, siendo 1 izquierda y 10 derecha, te pedimos que indiques tu posicionamiento político”.

Sobre la primera, la segunda y la tercera pregunta, será definida la proporción de confianza o imagen positiva, para cada uno de los segmentos, considerando la razón entre la cantidad de entrevistados que hayan puntuado entre 7 y 10 estas preguntas, y el total de entrevistados que componga la audiencia. Considerando estos tres indicadores, junto con la proporción de usuarios diarios derivada de la segunda pregunta y la mediana del posicionamiento político indicado, será llevado a cabo un análisis gráfico de la relación entre los mismos y el nivel de correlación hallado entre discursos. A su vez, este análisis será complementado con datos sociodemográficos de cada audiencia, como la edad mediana de cada grupo o la proporción de mujeres hacia el interior de cada uno de ellos, entendiendo a este último indicador como expresión de la distribución según género de cada segmento.

CAPÍTULO 3. ¿QUÉ NOS DICEN LOS MEDIOS? ANÁLISIS DE CONTENIDO SOBRE EL DISCURSO DE LOS MEDIOS DE COMUNICACIÓN

En el presente capítulo, se realiza el análisis de contenido sobre las publicaciones de cuentas oficiales extraídas sobre los medios de comunicación seleccionados, durante el día 14 de noviembre de 2018 y el día 15 de diciembre de 2018. Particularmente, se analizan 21.486 publicaciones distribuidas de la siguiente manera, en función de los medios de comunicación:

Tabla 2.

Distribución de publicaciones según Medio de Comunicación

Cuenta Oficial	Cantidad de Publicaciones	Porcentaje
infobae	10.417	48%
clarincom	3.485	16%
perfilcom	3.096	14%
LANACION	2.781	13%
pagina12	1.707	8%
Total	21.486	100%

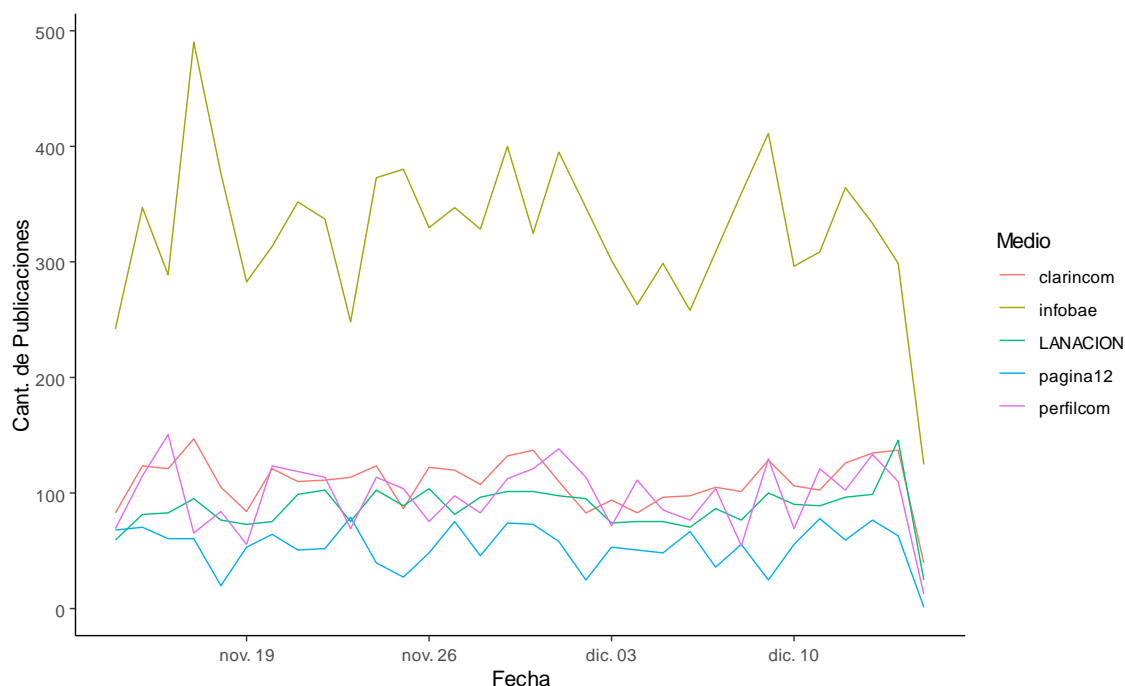
Fuente: Elaboración Propia

Se evidencia cómo Infobae, por ser un medio nativo digital, representa prácticamente la mitad del cuerpo de publicaciones extraídas, mientras que en los restantes medios se observa un comportamiento similar en términos de volumen.

Al observar el comportamiento de los medios de comunicación seleccionados en el tiempo, es posible observar la actividad diferencial de Infobae por sobre los restantes medios. No obstante, se evidencia un comportamiento relativamente similar por parte de todos los medios, en cuanto a la cantidad de publicaciones por fecha, mostrando reacciones similares según el día de publicación.

Gráfico 4.

Distribución de publicaciones según Fecha y Medio de Comunicación



Fuente: Elaboración Propia

Como puede observarse, Página 12 es el medio de comunicación que comporta relativamente diferente en el grupo de cuentas oficiales seleccionado, marcando ciertas contratendencias en cuanto a los días de publicación que son pico de actividad en canales digitales. En este sentido, se evidencia su comportamiento diferencial al analizar la coocurrencia entre estas series temporales¹⁵: mientras todos los medios de comunicación mantienen entre sí una correlación de intensidad media o alta -hecho particularmente relevante entre medios tradicionales e históricos como La Nación y Clarín-, Página 12 solo muestra una correlación de intensidad significativa con Clarín, diferenciándose de los restantes.

¹⁵ La coocurrencia entre series temporales es medida a partir del coeficiente de correlación R de Pearson que registran los diferentes medios de comunicación entre sí, en función de los días de publicación y la cantidad de publicaciones por registro.

Tabla 3.

Coocurrencia entre series temporales correspondientes a Cantidad de Publicaciones por Medio de Comunicación

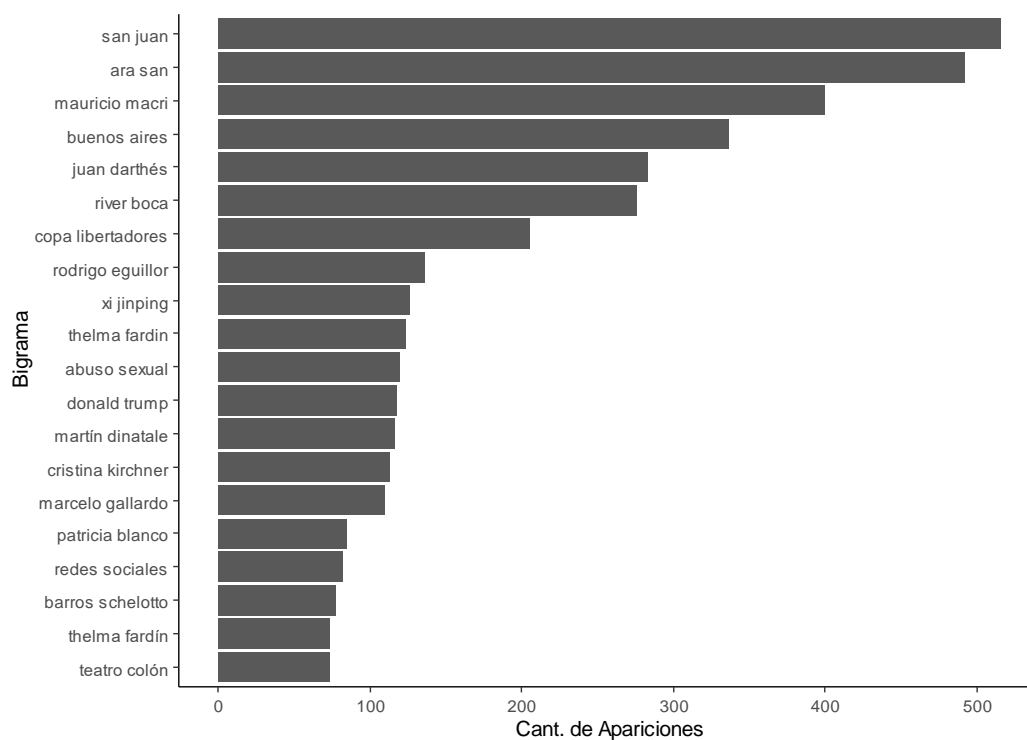
	clarincom	infobae	LANACION	pagina12	perfilcom
clarincom	1,000	0,666	0,736	0,535	0,536
infobae	0,666	1,000	0,593	0,085	0,421
LANACION	0,736	0,593	1,000	0,260	0,580
pagina12	0,535	0,085	0,260	1,000	0,280
perfilcom	0,536	0,421	0,580	0,280	1,000

Fuente: Elaboración Propia

En cuanto al contenido publicado por los medios de comunicación seleccionados, es posible considerar los *bi-gramas* -esto es, las combinaciones de dos palabras- con mayor ocurrencia en el corpus de publicaciones elaborado con el fin de hallar los términos mayormente relevantes en su discurso.

Gráfico 5.

Cantidad de Apariciones por Bi-grama



Fuente: Elaboración Propia

Como puede observarse, el concepto de San Juan y ARA San son los que presentan una mayor cantidad de apariciones, relacionados con el incidente ocurrido hacia finales de 2017. Particularmente, en el corpus de noticias relevado, el tema cumple un rol relevante dado el cumplimiento de un año luego de la desaparición del submarino y la demanda de respuestas por parte de los familiares de la tripulación a las autoridades gubernamentales.

Por otro lado, se evidencia cierta preponderancia de la figura de Mauricio Macri y Buenos Aires, hecho que se relaciona con la cumbre del G20 realizada en dicha ciudad entre el 30 de noviembre y el 1 de diciembre. Entre los hechos relevantes se destacan no solo las intervenciones del entonces Presidente y las referencias a la Ciudad de Buenos Aires como punto del evento o por cuestiones referidas a la organización vial, sino también su intercambio con diferentes líderes o la llegada de algunos de ellos al país, entre los que se destacan las figuras de Xi Jinping y Donald Trump.

En tercer lugar, es posible evidenciar la relevancia de los temas de género en los medios de comunicación seleccionados. Particularmente, resaltan aquellos hechos con alto impacto mediático como abusos y acosos sexuales protagonizados por actores reconocidos como Juan Darthés, y vinculado a este hecho su denunciante, Thelma Fardín. También se vincula a este grupo el nombre de Rodrigo Eguillor, el hijo de una fiscal, denunciado en diversas ocasiones por abuso, quien hoy se encuentra condenado a prisión.

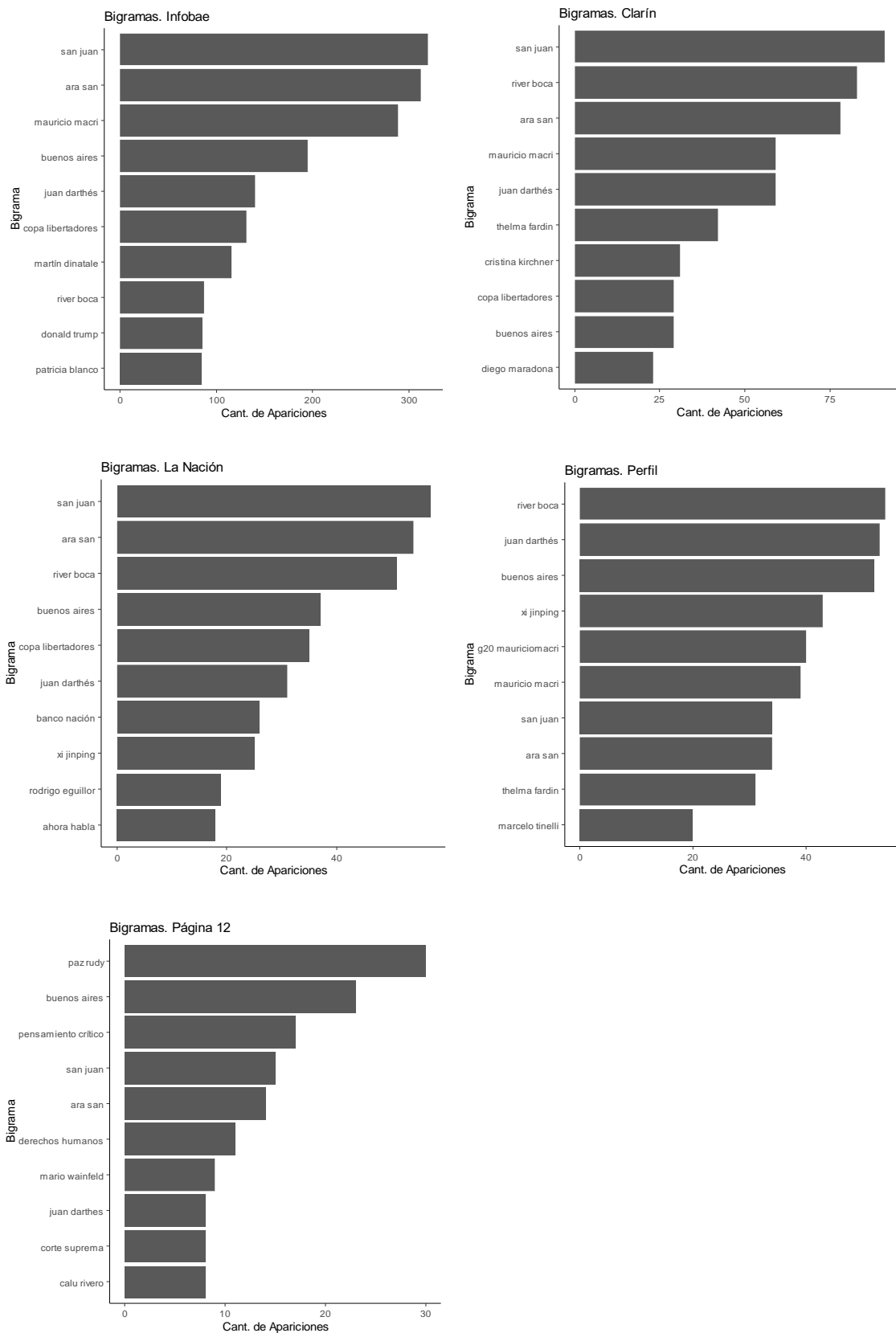
Por último, se evidencia la importancia de los términos relacionados con el fútbol, donde el hecho más relevante del período considerado se asocia al incidente ocurrido previo a la final de la Copa Libertadores que iban a protagonizar River y Boca. En este sentido, resalta la coocurrencia de estos términos en el corpus relevado.

A continuación, se muestran los bi-gramas mayormente utilizados por cada medio de comunicación. Como puede observarse, el cumplimiento del aniversario del incidente del ARA San Juan ha sido el tema mayormente mencionado por parte de Infobae, Clarín y La Nación, aunque también ocuparon un rol relevante en las publicaciones de Perfil y Página 12. Para Perfil, por su parte, los términos más utilizados se asocian a la final de la Copa Libertadores protagonizada por River y Boca, mientras que en las publicaciones de Página 12 resalta su recurso

del humor para comunicar, apoyándose en la figura de Paz Rudy, además del surgimiento de nuevos tópicos, tal como Derechos Humanos.

Gráfico 6.

Cantidad de Apariciones por Bi-grama, según Medio de Comunicación



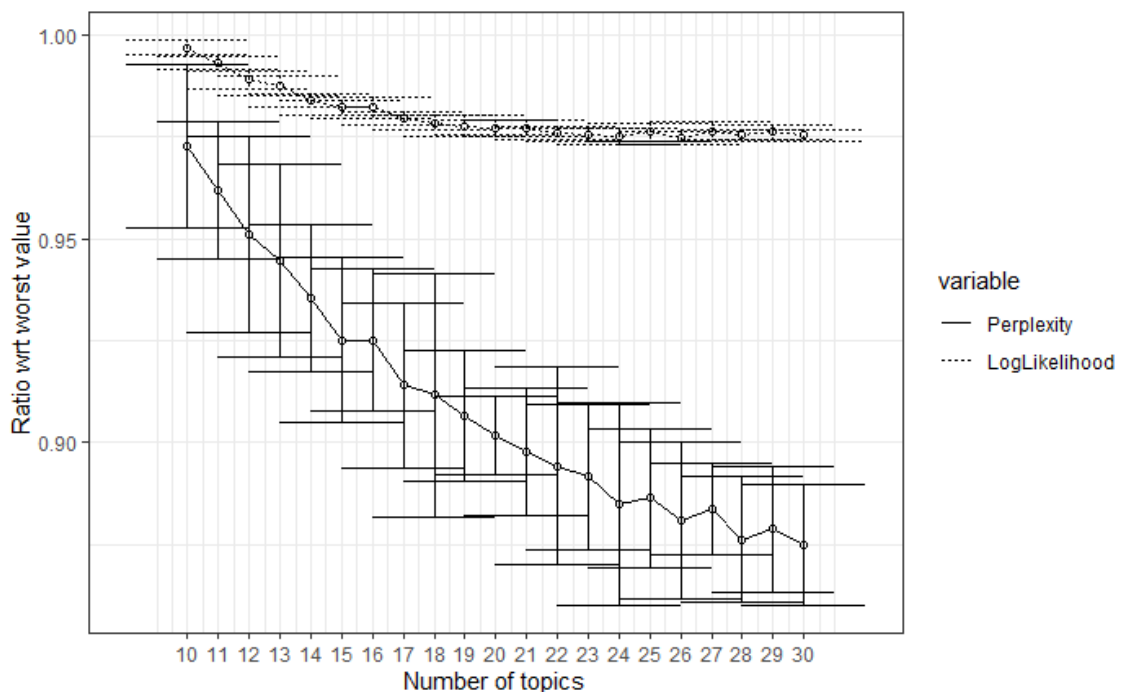
Fuente: Elaboración Propia

Ahora bien, una vez descritos los términos mayormente frecuentes en las publicaciones de los medios de comunicación seleccionados, es posible preguntarse sobre qué temas o tópicos se basa el discurso de estos medios de comunicación. Si bien se observa que existen términos frecuentes en las publicaciones extraídas, e incluso se evidencia una coincidencia respecto a su uso por parte de los medios de comunicación elegidos, dichos términos no abarcan la totalidad de la comunicación proveniente de dichas fuentes en el período de tiempo considerado.

Con este fin, se utilizan las técnicas de *Text Mining* anteriormente introducidas, con el fin de explorar la totalidad del *corpus* extraído. En este sentido, los tópicos correspondientes al discurso analizado son inferidos a partir de la implementación del algoritmo LDA, mientras que la cantidad de tópicos es definida en función del indicador de perplejidad estimado.

Gráfico 7.

Validación cruzada de cantidad de tópicos k sobre documentos relevados



Fuente: Elaboración Propia

Como puede observarse, existen una diversidad de puntos de inflexión en el gráfico, que minimizan la medida de perplejidad e indican distribuciones de probabilidad pasibles de explicar la muestra de documentos extraídos. Sobre la diversidad de modelos posibles, en el presente trabajo se considera un modelo descriptivo basado en 16 tópicos, considerando que establece un punto de inflexión significativo, a la vez que mantiene la parsimonia del modelo para fines interpretativos.

Considerados estos 16 tópicos, se evalúa la dimensión relativa de cada uno de ellos, una vez que se ha seleccionado el tópico de mayor probabilidad correspondiente a cada documento. Como puede observarse, el tópico número 7 representa aquel con mayor volumen en el corpus de documentos relevados, abarcando 2.285 de ellos (11% del total). A éste, le sigue el tópico número 2, representando el 8% del total, y los tópicos 1, 3, 4, y 5, los cuales representan el 7% del total de publicaciones.

Tabla 4.

Distribución de documentos según Tópico

Tópico	Cantidad de Publicaciones	Porcentaje
1	1.469	7%
2	1.723	8%
3	1.583	7%
4	1.473	7%
5	1.392	7%
6	1.465	7%
7	2.285	11%
8	1.302	6%
9	1.156	5%
10	1.043	5%
11	1.352	6%
12	1.036	5%
13	1.046	5%
14	1.043	5%
15	1.086	5%
16	954	4%

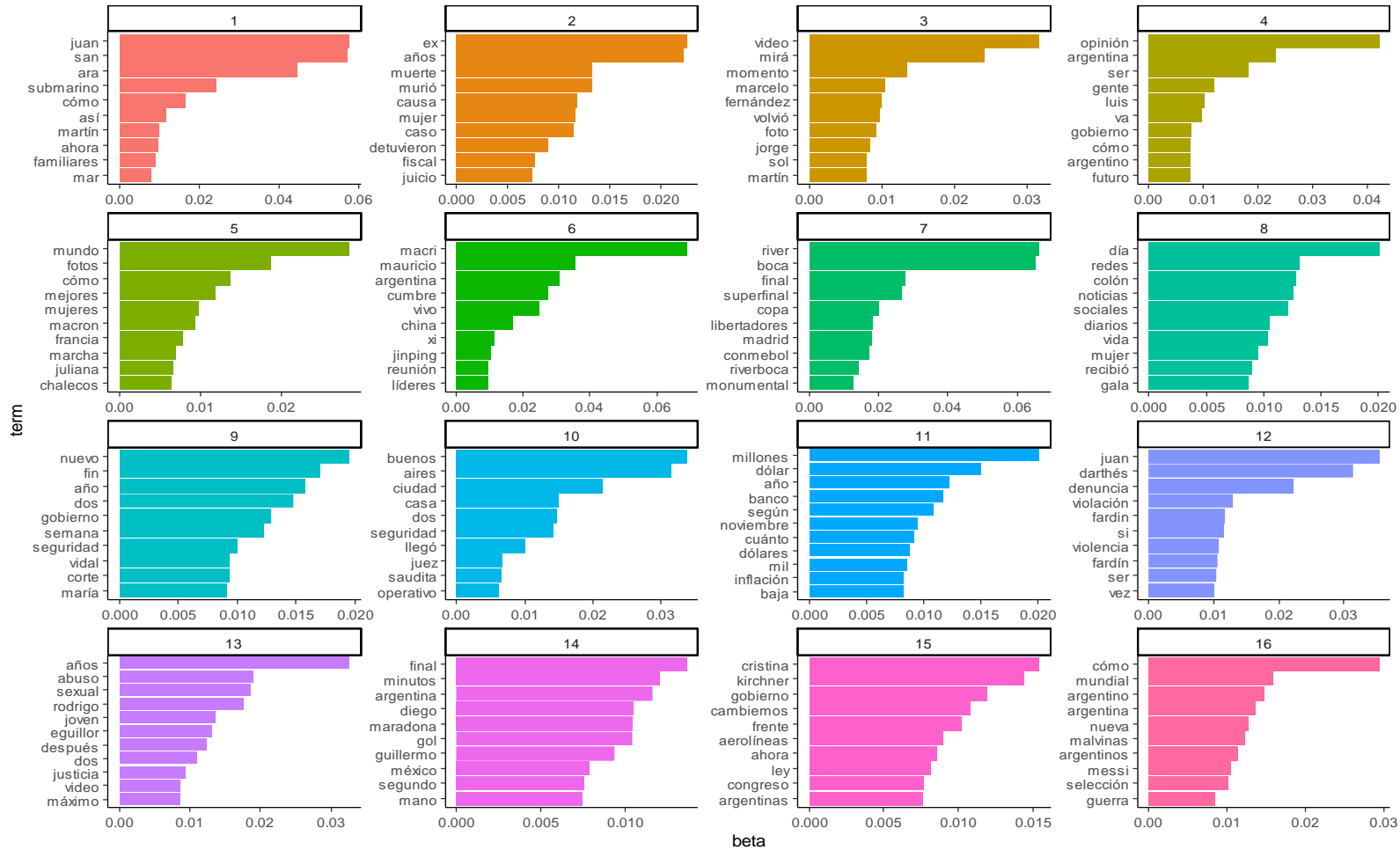
Fuente: Elaboración Propia

A continuación, se presentan gráficamente los tópicos extraídos, donde es posible observar los términos que mayormente contribuyen a cada uno de ellos. En este sentido, es posible comprender la composición de cada tema a partir de la contribución relativa de cada palabra frente a estos, en función de su coeficiente β_i estimado.

Gráfico 8.

Desagregación de tópicos en palabras con mayor contribución

Fuente: Elaboración Propia



Sobre el análisis previo, se define cada uno de los tópicos en función de su contenido, a la vez que se presentan publicaciones que expresan las características enunciadas:

1. *ARA San Juan*: representa aquellas publicaciones que tratan el tema del incidente ocurrido con el submarino ARA San Juan. Ejemplos:
 - *“Tras el hallazgo del ARA San Juan el ministro de Defensa dijo que no tienen los medios tecnológicos para rescatar el buque y que ahora resta determinar las causas de lo ocurrido y ‘encontrar a los responsables si es que los hubo’”,* Página 12, 18/11/2018
 - *“Así se navega a bordo del Seabed Constructor el buque que encontró el ARA San Juan Cómo es por dentro”,* Infobae, 17/11/2018
 - *“SOMOS COMO UNA FAMILIA Recuerdo de cuando el ARA San Juan volvía al agua luego de una larga etapa de reparación El comandante de ese entonces daba un mensaje de agradecimiento a la tripulación Imágenes que resumen el espíritu de camaradería”,* Clarín, 17/11/2018
2. *Causas de asesinatos y femicidios*: representa aquellas publicaciones que refieren a asesinatos o femicidios, relacionando muertes, causas y juicios con mujeres. Ejemplos:
 - *“Habla el femicida Fernando Farré desde la cárcel Me quieren matar porque desarmé el complot del ex saxofonista de los Cadillac para asesinar al fiscal Por Fernando Soriano”,* Infobae, 05/12/2018
 - *“Desde la cárcel habla el ex Fabulosos Cadillacs acusado por Farré de planear la muerte de un fiscal El femicida miente y vive empastillado Por Federico Fahsbender”,* Infobae, 09/12/2018
 - *“Crimen por encargo el fiscal cree que Macarrón contrató un sicario para matar a su esposa Nora Dalmasso Por Rodolfo Palacios”,* Infobae, 21/11/2018
3. *Espectáculos*: refiere a publicaciones del mundo del espectáculo haciendo mención de celebridades relevantes como Flor Vigna, Laurita Fernández, Sol Pérez, entre otros. Ejemplos:
 - *“En medio del escándalo entre Flor Vigna y Laurita Fernández Gimena Accardi publicó un video besando a Nicolás Cabré”,* Infobae, 14/11/2018

- *“FUERTES DECLARACIONES Nacha Guevara hizo tremendos comentarios sobre los rasgos físicos de los hijos de Marley Luciana Salazar y Flavio Mendoza Mirá el video”, Perfil, 29/11/2018*
 - *“INCORRECTA Nora Cárpena fundamentó las razones de su licencia del programa de Moria Casán tras la llegada de Nacha Guevara a la conducción Mirá sus declaraciones”, Perfil, 23/11/2018*
4. *Política argentina*: refiere a publicaciones sobre política argentina (interna, haciendo referencia a acuerdos políticos u opiniones de políticos influyentes, y externa, con publicaciones sobre la relación con Brasil y Bolsonaro o el FMI). Ejemplos:
- *“Massa Pichetto Urtubey y Schiaretti sumaron el apoyo de siete gobernadores peronistas de cara a Por Joaquín Mugica Díaz”, Infobae, 21/11/2018*
 - *“Sergio Massa y Gustavo Bordet se reunieron en Entre Ríos para avanzar en la construcción de una alternativa electoral”, Infobae, 15/11/2018*
 - *“Sergio Massa se reunió con el FMI Planteamos que el camino es crecer y generar empleo no ajustar la economía”, Infobae, 18/11/2018*
5. *Noticias internacionales*: mayormente representado por la palabra mundo, representa noticias ligadas a conflictos internacionales como conflictos en Irán o Francia, pero también a reconocimientos de argentinos en el exteriormente, como el correspondiente a Mauro Colagreco (chef argentino). Ejemplos:
- *“Irán ahorcó al sultán de las monedas por contribuir a las alzas de precios acumulando oro aprovechándose de los problemas financieros del país También colgaron a un asistente”, Perfil, 16/11/2018*
 - *“Emmanuel Macron sobre los desmanes en París Los culpables de esta violencia no quieren ninguna reforma sólo quieren el caos”, Infobae, 01/12/2018*
 - *“Mauro Colagreco la mente brillante de la gastronomía argentina que se alza en el podio de los mejores chefs del mundo Por Silvina Reusmann”, Infobae, 17/11/2018*
6. *Cumbre G20*: refiere a las publicaciones que comentan la cumbre de G20 ocurrida a fines de 2018, con una fuerte incidencia de Mauricio Macri, pero también con referencia a los líderes mundiales participantes. Ejemplos:

- *“El gobierno se mostró exultante por el resultado de la cumbre y destacó el rol de mediador que jugó Mauricio Macri Por Martín Dinatale”*, Infobae, 02/12/2018
 - *“La batuta de Barenboim un caballo de polo y dos bicicletas chinas los regalos que intercambié Mauricio Macri con Xi Jinping”*, Infobae, 02/12/2018
 - *“AHORA Habla Xi Jinping El presidente Macri y yo hemos tenido una reunión muy productiva en donde hablamos de los temas con miras de desarrollo a largo plazo y hemos acordado planificar abrir una nueva era de la relación comercial en el proceso de apertura”*, La Nación, 02/12/2018
7. *Conflicto superfinal Copa Libertadores*: refiere a los incidentes ocurridos previamente a la final de la Copa Libertadores, hechos que resultaron en la suspensión del partido, además de los hechos que le sucedieron y los resultados del partido jugado en Madrid. Ejemplos:
- *“REITERAMOS Autorizaron a Rafael Di Zeo a viajar a España para la final de la Copa Libertadores entre River y Boca”*, Infobae, 05/12/2018
 - *“Marcelo Gallardo probó variantes Juanfer Quintero Rodrigo Mora y los cinco defensores Cuál de estos tres jugadores debería salir a la cancha desde el comienzo”*, Infobae, 24/11/2018
 - *“AHORA Alejandro Domínguez presidente de Conmebol sobre la final entre Boca y River Se jugará en sede a definirse”*, Clarín, 27/11/2018
8. *Gala Teatro Colón*: refiere a las publicaciones sobre la Gala que presenciaron los líderes globales en el Teatro Colón durante el G20, incluyendo los artistas que actuaron allí, las reacciones de los líderes y las repercusiones. Ejemplos:
- *“La gala del Teatro Colón presidentes invitados especiales y la intimidad de la gran noche”*, Infobae, 30/11/2018
 - *“Julio Bocca volvió a bailar en el Teatro Colón en la gala multidisciplinaria Argentum bajo la dirección de Ricky Pashkus”*, Infobae, 30/11/2018
 - *“Danza música de todo el país y sorpresas cómo será el show multidisciplinario que verán los líderes en el Teatro Colón”*, Infobae, 28/11/2018
9. *Medidas locales de gobierno*: refiere a publicaciones que relacionan acciones de Gobierno y el Impuesto a las ganancias. Particularmente, se debate en

torno a un bono pagado a fin de año (con fuerte relación con María Eugenia Vidal) y su alcance por dicho impuesto, a la vez que el debate en torno a si los nuevos jueces serán alcanzados por el impuesto mencionado. Por otro lado, se hace referencia al debate en torno a temas como la inseguridad y las herramientas de las fuerzas de seguridad para actuar. También se hace referencia a fallos clave por parte de la Corte. Ejemplos:

- *“María Eugenia Vidal adelantará el pago del aguinaldo a los estatales y otorgará un bono para los jubilados que menos ganan”*, Infobae, 23/11/2018
- *“Germán Garavano sobre el nuevo Código Procesal Hubo resistencia de jueces por una visión mezquina del poder Por Martín Angulo”*, Infobae, 08/12/2018
- *“El bono divide a Macri Vidal y Larreta El jefe de Gobierno descarta pagarlo la gobernadora aún no lo definió”*, Página 12, 14/11/2018

10. *Hechos de inseguridad y conflicto*: refiere a protestas como la realizada por el cierre del Buenos Aires Design, a la vez que se relaciona con hechos de inseguridad como el ataque con explosivos al juez Bonadío. A su vez, este último hecho se relaciona con el explosivo utilizado en Recoleta por anarquistas. Ejemplos:

- *“BUENOS AIRES ALERTA Tras la explosión en el Cementerio de Recoleta y la amenaza al Juez Bonadio encuentran artefacto sospechoso en la Galería Obelisco Norte y reportan amenaza de bomba al Departamento Central de la Policía en el bajo porteño”*, Perfil, 15/11/2018
- *“URGENTE Encuentran objeto sospechoso en Galería Obelisco Norte Se trata de un artefacto montado sobre la carcasa de un TV aprox x La división explosivos se encuentra trabajando en el lugar”*, Perfil, 15/11/2018
- *“En el día de ayer hubo dos ataques con explosivos Uno fue en la casa del juez Claudio Bonadio y otro en el cementerio de la Recoleta Por los dos ataques hay detenidos”*, La Nación, 15/11/2018

11. *Economía*: refiere a publicaciones relacionadas al precio del dólar, el aumento de precios (por ejemplo, en combustibles) y su impacto en la inflación, y las medidas del BCRA en relación a la definición de tasas de interés. Ejemplos:

- *“Por la inflación más ahorristas se vuelcan a los plazos fijos UVA cuáles son sus ventajas y riesgos Por Ximena Casas”, Infobae, 17/11/2018*
- *“Yerba mate una industria de millones de ventas pero desafiada por las tasas de interés y la inflación Por Francisco Bueno”, Infobae, 24/11/2018*
- *“Por tercer mes consecutivo cayeron las ventas en los supermercados y grandes centros de compra Por Daniel Sticco”, Infobae, 26/11/2018*

12. *Casos de abuso y violencia de género: vinculado sobre todo a la denuncia de Juan Darthés por parte de Thelma Fardín, refiere a las publicaciones que toman el tema de la violencia de género, el machismo y el acoso callejero. Ejemplos:*

- *“Florencia Ortiz ex compañera de Juan Darthés y Thelma Fardín en Patito Feo Él era un tipo callado tenía un lado oscuro Por Nieves Otero”, Infobae, 12/12/2018*
- *“SE FILTRÓ UN MENSAJE DE JUAN DAHRTÉS El actor le envió un escrito vía Whatsapp a sus amigos tras ser denunciado por violación En minutos hará su descargo en Más que noticias con Mauro Viale”, Clarín, 13/12/2018*
- *“El mensaje de Laura Esquivel a Thelma Fardín tras la denuncia a Juan Darthés por violación Sos inmensamente valiente”, Infobae, 12/12/2018*

13. *Causas de abuso y femicidios con fuerte tratamiento judicial: también refiere a abuso y asesinato de mujeres, pero en relación a causas por fuerte impacto en la justicia. Partiendo de la causa de Rodrigo Eguillor denunciado por abuso (hecho con mayor impacto en las publicaciones), relaciona notas referidas al asesinato por parte de Barreda, entre otras. Ejemplos:*

- *“Rodrigo Eguillor fue denunciado por una joven de años Crece el escándalo del hijo de una fiscal acusado de abuso sexual Por Milton Del Moral”, Infobae, 04/12/2018*
- *“La historia del hombre al que denunciaron por abuso sexual y la Justicia no quiso detener mató a toda su familia y después se suicidó Por Bárbara García Crespo”, Infobae, 09/12/2018*
- *“IMÁGENES SENSIBLES El momento en que Rodrigo Eguillor el joven acusado de abuso forcejea con su denunciante en el balcón”, Infobae, 04/12/2018*

14. *Deporte*: hace referencia a noticias deportivas como resultados de partidos de fútbol (local e internacional) y tenis. Ejemplos:

- *“Comienza la definición por penales En los minutos igualaron a con goles de Zampedri y Faravelli respectivamente Expectativa en Mendoza por conocer al campeón”*, Infobae, 07/12/2018
- *“Comenzó el complemento en el Minella están empatando con goles de Pity Martínez y L Faravelli respectivamente En el Lobo se fue expulsado Bonifacio”*, Infobae, 29/11/2018
- *“Gerardo Martino salió campeón en la MLS con Atlanta United y ya piensa en México”*, La Nación, 09/12/2018

15. *Internas políticas locales*: refiere a las acciones políticas de las fuerzas locales, como la construcción de un bloque amplio peronista, y noticias sobre sus protagonistas y causas relacionadas como los hechos de corrupción relacionados con Cristina Fernández de Kirchner. Por otro lado, refleja internas dentro del bloque Cambiemos, particularmente relacionado a la salida de Monzó. Por último, en relación a la política local y el debate desde diferentes perspectivas políticas, surge como protagonista el conflicto gremial en Aerolíneas Argentinas y el intento del gobierno por implementar la conciliación obligatoria. Ejemplos:

- *“CONFLICTO Aerolíneas Argentinas anunció sanciones a empleados por la retención de tareas que paralizaron más de vuelos el de noviembre Los sindicatos anunciaron nuevas asambleas que complicarían las operaciones”*, Clarín, 20/11/2018
- *“Carolina Stanley Hay sectores que buscan exacerbar la angustia y creen en la lógica de cuanto peor mejor”*, Infobae, 08/12/2018
- *“Crisis en Cambiemos reproches cruzados entre el PRO y la UCR por la derrota en el Consejo de la Magistratura Por Silvia Mercado”*, Infobae, 20/11/2018

16. *Deporte nacional y seleccionados nacionales*: hace referencia a las publicaciones que tratan el tema de los diferentes seleccionados locales, debatiendo particularmente el rol de Messi en el seleccionado de fútbol masculino.

Ejemplos:

- *“INFOGRAFÍAS A años de su debut es Lionel Messi el mejor futbolista de la historia cinco infografías para que usted decida”*, Infobae, 16/11/2018
- *“Chiquito Romero le contestó a Mauro Icardi y abrió otra grieta en la interna del vestuario de la Selección”*, Infobae, 22/11/2018
- *“La Selección femenina jugará el Mundial de Francia crónica de una batalla que ganaron las jugadoras de principio a fin”*, Infobae, 14/11/2018

En función de los temas desarrollados, al contrastar la proporción que ocupa cada tópico respecto al total de publicaciones de cada medio de comunicación, no se observa una dependencia en términos estadísticos¹⁶, por lo que puede afirmarse que los temas sobre los cuales se desarrollen publicaciones no depende del medio de comunicación que esté comunicando dicha noticia. Esto resulta relevante, particularmente, a la hora de considerar el principio de *Focusing* introducido al comienzo del presente trabajo, donde fue explicado que, para lograr establecer ciertos temas de agenda, los medios de comunicación tienden a centrar su discurso sobre los mismos tópicos, en conjunto.

¹⁶ Los resultados del Test de Chi Cuadrado para la tabulación cruzada presentada arroja un P-valor igual a 1, correspondiente a un valor de χ^2 de 0 para 60 grados de libertad.

Tabla 5.

Distribución de documentos según Tópico y Medio de Comunicación

Tópico	clarincom	infobae	LANACION	pagina12	perfilcom
ARA San Juan	8%	7%	7%	7%	6%
Causas de Asesinatos y Femicidios	9%	8%	8%	10%	7%
Espectáculos	9%	6%	6%	4%	14%
Política Argentina	5%	7%	6%	9%	7%
Noticias Internacionales	6%	6%	7%	7%	6%
Cumbre G20	5%	7%	7%	6%	8%
Conflicto superfinal Copa Libertadores	12%	12%	13%	4%	8%
Gala Teatro Colón	6%	7%	6%	5%	5%
Medidas locales de Gobierno	5%	5%	5%	8%	5%
Hechos de Inseguridad y Conflicto	4%	5%	5%	5%	6%
Economía	5%	7%	7%	8%	6%
Casos de abuso y violencia de género	6%	4%	6%	6%	5%
Causas de abuso y femicidios con tratamiento judicial	6%	5%	5%	4%	4%
Deporte	5%	5%	5%	3%	4%
Internas Políticas Locales	4%	5%	5%	9%	5%
Deporte Nacional y Seleccionados	4%	5%	4%	4%	4%
Total	100%	100%	100%	100%	100%

Fuente: Elaboración Propia

Como puede observarse, el tema referido al conflicto de la final de la Copa Libertadores protagonizado por River y Boca es el tema que mayor espacio

ocupa en la agenda mediática en la red social Twitter, en la mayoría de los medios de comunicación seleccionados. Las excepciones, en este caso, son Página 12 y Perfil: mientras que para el primero el tema mayormente relevante en agenda son causas de asesinatos y femicidios, para el segundo el tema más relevante son aquellos hechos relacionados al mundo del espectáculo.

En términos generales, resalta la manera en que los medios de comunicación aquí presentados construyen una agenda, en redes sociales, basada en temas de coyuntura, como el espectáculo, deportes o conflictos específicos, por sobre temas estructurales como la pobreza o la economía.

Con el fin de evaluar la distribución de tópicos entre los diferentes medios de comunicación considerados, y contrastar la falta de relación estadística en términos generales evidenciada previamente con la particular relación entre medios hacia el interior de cada tópico, es posible realizar un test Z entre proporciones, de a pares, indicando las relaciones estadísticamente significativas que tienen lugar en cada caso. Dichas relaciones, se expresan a continuación¹⁷:

¹⁷ Se indican las relaciones estadísticamente significativas con un P-valor < 0,05, según un Test Z para proporciones bilateral o de dos colas. Para cada par de categorías, se indica la clave de la columna de mayor porcentaje. Una referencia a esta prueba estadística puede ser encontrada en Manuel García Ferrando (1989, p.292) o en Humbert M. Blalock (1960, p. 176).

Tabla 6.

Relaciones estadísticamente significativas en Tópicos y Medios de Comunicación

Tópico	clarincom	infobae	LANACION	pagina12	perfilcom
	(A)	(B)	(C)	(D)	(E)
ARA San Juan					
Causas de Asesinatos y Femicidios	E			E	
Espectáculos	B C D		D		A B C D
Política Argentina		A C		A C	A
Noticias Internacionales					
Cumbre G20		A	A		A
Conflicto superfinal Copa Libertadores	D E	D E	D E		D
Gala Teatro Colón	E	E			
Medidas locales de Gobierno				A B C E	
Hechos de Inseguridad y Conflicto					
Economía		A	A	A E	
Casos de abuso y violencia de género	B		B	B	
Causas de abuso y femicidios con tratamiento judicial	E				
Deporte	D	D E	D		
Internas Políticas Locales		A	A	A B C E	A
Deporte Nacional y Seleccionados					

Fuente: Elaboración Propia

Tal como puede observarse, hay ciertas relaciones entre tópicos y medios que describen las particularidades de estos últimos. A continuación, se realiza una descripción del comportamiento comunicacional de cada medio, en relación a los tópicos de comunicación de los restantes:

- a. Clarín se distingue por comunicar acerca de causas de asesinatos y femicidios, la gala en el Teatro Colón, y las causas de abuso con fuerte

- impacto judicial frente a Perfil, por comunicar causas de abuso y violencia de género frente a Infobae, deporte y conflictos asociados al mismo frente a Página 12, y espectáculos frente a Infobae, La Nación y Página 12;
- b. Infobae se distingue de Clarín dado el volumen que representa la política argentina, la Cumbre G20, la economía y las internas políticas locales en su comunicación, a la vez que se distingue de Página 12 y Perfil al comunicar deporte y conflictos asociados al deporte;
 - c. La Nación se diferencia principalmente por el lugar que ocupa en su agenda en redes sociales la Cumbre G20, la economía y las internas políticas locales frente a Clarín, los espectáculos, el deporte y conflictos derivados del deporte frente a Página 12, y los casos de abuso y violencia de género frente a Infobae;
 - d. Página 12 se distingue del conjunto de medios considerados, por el lugar particular que ocupan las medidas de gobierno local y las internas políticas locales en su agenda, además del lugar que le brinda a la política argentina frente a Clarín y La Nación, la economía frente a Clarín y Perfil, los femicidios frente a Perfil, y las causas de abuso y violencia de género frente a Infobae;
 - e. Por último, Perfil se distingue del resto de los medios de comunicación considerados, por el lugar que ocupan los espectáculos en su agenda. A su vez, se distingue de Clarín por el lugar que ocupa la política argentina, la Cumbre G20 y las internas políticas locales, a la vez que se distingue de Página 12 por el lugar que le brinda al conflicto derivado de la Superfinal de la Copa Libertadores.

En base a estas caracterizaciones, si bien se observa que todos los tópicos identificados son relevantes para el conjunto de medios de comunicación seleccionados, puede distinguirse a Página 12 como el medio de comunicación que mayormente se diferencia del resto, fundamentalmente, por el lugar que ocupa en su agenda las medidas locales de gobierno y las internas políticas locales, a la vez que por la manera que en los restantes medios de comunicación se diferencian de éste, en base a diferentes tópicos.

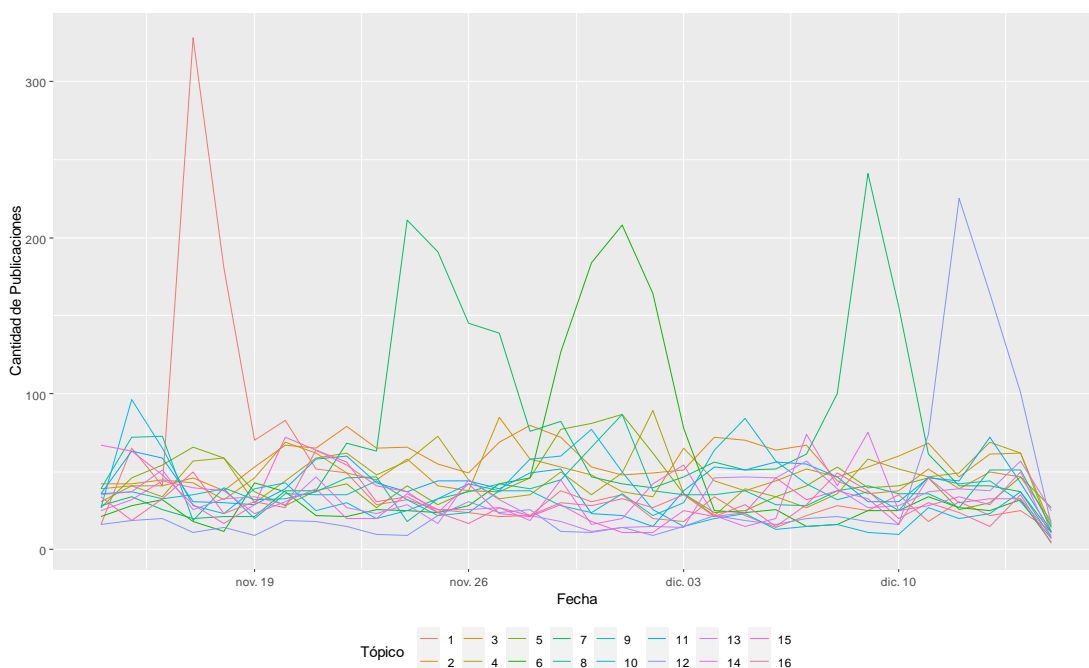
Una manera de corroborar el principio de *Persistence* para el establecimiento de agenda, es evaluar la relevancia dada a los diferentes tópicos hallados a lo largo

del tiempo, por parte de los medios de comunicación seleccionados. En este sentido, si bien se evidencia cierta consistencia temporal para todos los tópicos hallados -esto es, cierta perdurabilidad en el tiempo de los temas latentes a los eventos resaltados, o incluso de los mismos eventos-, hecho que facilita la construcción de agenda por parte de los medios de comunicación, también se identifican ciertos picos en relación a eventos particulares, en el período de tiempo seleccionado.

Particularmente, se nota la relevancia del tópico “ARA San Juan” hacia mediados del mes de noviembre; “Cumbre G20” hacia los primeros días del mes de diciembre; y “Casos de abuso y violencia de género” hacia mediados de diciembre. Cabe destacar que el tema “Conflicto Superfinal Copa Libertadores” encuentra dos picos temporales de relevancia, uno hacia fines del mes de noviembre y otro hacia la segunda semana del mes de diciembre. Gráficamente:

Gráfico 9.

Evolución de los tópicos hallados sobre la agenda mediática en el tiempo



Fuente: Elaboración Propia

Con el fin de contrastar la coocurrencia en términos de su actividad en Twitter por parte de los medios de comunicación seleccionados, complementariamente

a los test estadísticos realizados previamente, es posible calcular una correlación de Spearman en función del lugar relativo que ocupa cada tema en la agenda de cada medio de comunicación. En este sentido, se presentan a continuación los resultados en base a un ranking elaborado sobre la proporción de cada tema sobre el total de publicaciones por medio:

Tabla 7.

Coefficiente Correlación de Spearman entre Medios de Comunicación

	Clarín	Infobae	La Nación	Página 12	Perfil
Clarín	1	0,4977909	0,5886683	-0,06769685	0,4270987
Infobae	0,49779087	1	0,8285506	0,28403245	0,6863034
La Nación	0,5886683	0,8285506	1	0,39411765	0,7682121
Página 12	-0,06769685	0,2840325	0,3941176	1	0,3311259
Perfil	0,42709867	0,6863034	0,7682121	0,33112592	1

Fuente: Elaboración Propia

Como puede observarse, la agenda de Infobae encuentra alta correlación con la agenda de La Nación y Perfil, mientras que estos dos últimos también comparten agenda en una intensidad considerable. Por otro lado, la relación entre las agendas de los medios restantes es de intensidad media, considerando a modo de ejemplo la relación entre La Nación y Clarín, o entre Clarín y Perfil. Por último, Página 12 es el medio de comunicación que menor relación presenta con los restantes, diferenciación que podía verse en el análisis de coocurrencia temporal previamente realizado. En este sentido, el medio de comunicación con el que mayormente comparte agenda es con La Nación, seguido de Perfil, aunque la intensidad con la que lo hace no es elevada.

En base a lo anteriormente dicho, destaca la estrecha relación que existe entre las agendas mediáticas, en la red social Twitter, de los medios de comunicación La Nación e Infobae, Perfil y La Nación, y Perfil e Infobae. De carácter levemente más débil, aunque no despreciable, resalta la coocurrencia discursiva entre Clarín y La Nación. Por otro lado, resalta la autonomía relativa que se evidencia para Página 12, en términos de los tópicos que priman en el discurso analizado.

Complementariamente a estos resultados, cabe señalar la elevada coocurrencia, en términos temporales, que se había identificado previamente entre Clarín y La Nación.

En este sentido, la coocurrencia temporal, junto con su mediana coocurrencia en términos de temas tratados, habla de una complementariedad entre ambos discursos, sobre una base común. Sobre estos resultados, cabe proceder a analizar el discurso de la opinión pública, para luego profundizar en la relación entre ambas agendas.

CAPÍTULO 4. EL PROBLEMA MÁS IMPORTANTE ANÁLISIS DE CONTENIDO SOBRE EL DISCURSO DE LA OPINIÓN PÚBLICA

En el presente capítulo se realiza un análisis de las respuestas obtenidas mediante encuesta a la clásica pregunta sobre el Problema Más Importante (PMI)¹⁸.

Metodológicamente, se consideran 410 respuestas recolectadas mediante una encuesta online realizada por panel, en el punto medio del período de extracción de contenido considerado para el análisis de los medios de comunicación anteriormente presentado. Particularmente, la muestra se compone de 410 casos de entre 18 y 70 años de edad, habitantes de Argentina, de todos los niveles socioeconómicos, usuarios de internet. En base a la población encuestada, se obtiene la siguiente distribución en base al sexo y rangos de edad elaborados en base a la edad promedio de la muestra (45 años) +- 1 desvío estándar (11 años):

Tabla 8.

Distribución de la muestra según Sexo y Rango etario

Sexo	Rango Etario					Total general
	18-21	22-33	34-45	46-57	58-70	
<i>Femenino</i>	1,95%	10,98%	21,95%	19,27%	8,05%	62,20%
<i>Masculino</i>	0,24%	3,90%	14,15%	13,41%	6,10%	37,80%
Total general	2,20%	14,88%	36,10%	32,68%	14,15%	100,00%

Fuente: Elaboración Propia

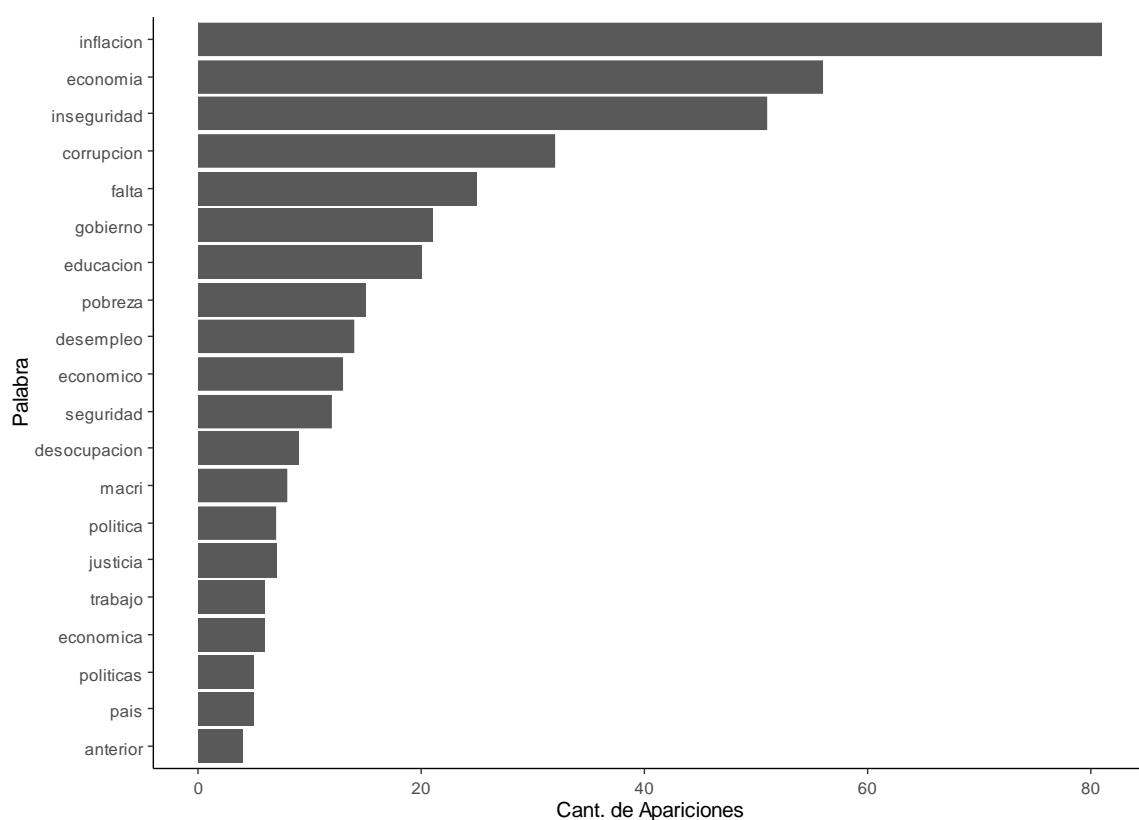
Como una primera aproximación al problema, es posible realizar un conteo de las palabras mayormente mencionadas por los encuestados a la hora de responder a dicha pregunta. En este sentido, cabe destacar que entre las respuestas a la pregunta mencionada se detectan 348 palabras diferenciadas,

¹⁸ Para más información sobre la pregunta PMI, consultar el capítulo “Marco Teórico” del presente trabajo.

sin considerar las denominadas “stopwords”¹⁹, que en conjunto representan un total de 818 menciones. A continuación, en la presentación gráfica de las veinte palabras más relevantes, se observa cómo los primeros dos problemas, en función de sus menciones, tienden a ser económicos, particularmente referidos a la problemática de la inflación y a la economía en general, seguidos de la inseguridad y la corrupción.

Gráfico 10.

Palabras más mencionadas para PMI



Fuente: Elaboración Propia

Ahora bien, considerando que las respuestas dadas a la pregunta mencionada contienen un conjunto de términos conceptualmente relacionados para aquel

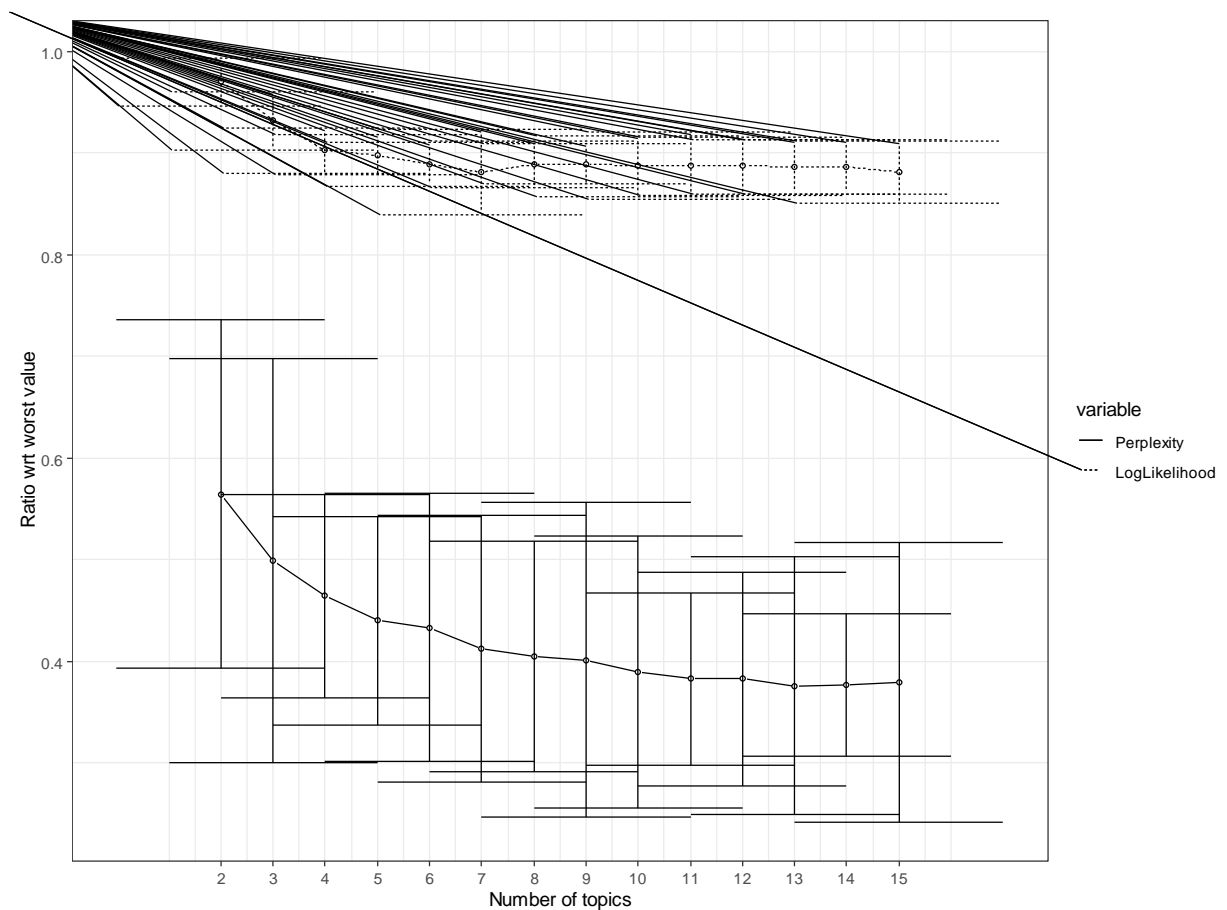
¹⁹ En el marco del análisis de texto y lenguaje natural, como ya fue mencionado en el capítulo “Metodología”, se conoce a las “stopword” como aquellas palabras que cumplen una función de apoyo en la elaboración de una oración o frase con sentido dentro de un discurso dado, pero que no tienen significado en sí mismas. Por este motivo, dichas palabras suelen ser descartadas de los análisis. Ejemplos de éstas son *a, así, cada, el, en, la, los, para, por*, entre otras.

que responde, es posible realizar un análisis de contenido en base a los tópicos comprendidos por estas respuestas, de la misma manera en la que fue realizada la exploración del discurso mediático.

Tal como puede observarse en el gráfico que sigue, el indicador de perplejidad estimado indica una ganancia significativa de poder explicativo al considerar una distribución en base a seis tópicos, por lo que ésta será la cantidad de temas considerados para el análisis.

Grafico 11.

Indicador de perplejidad para PMI



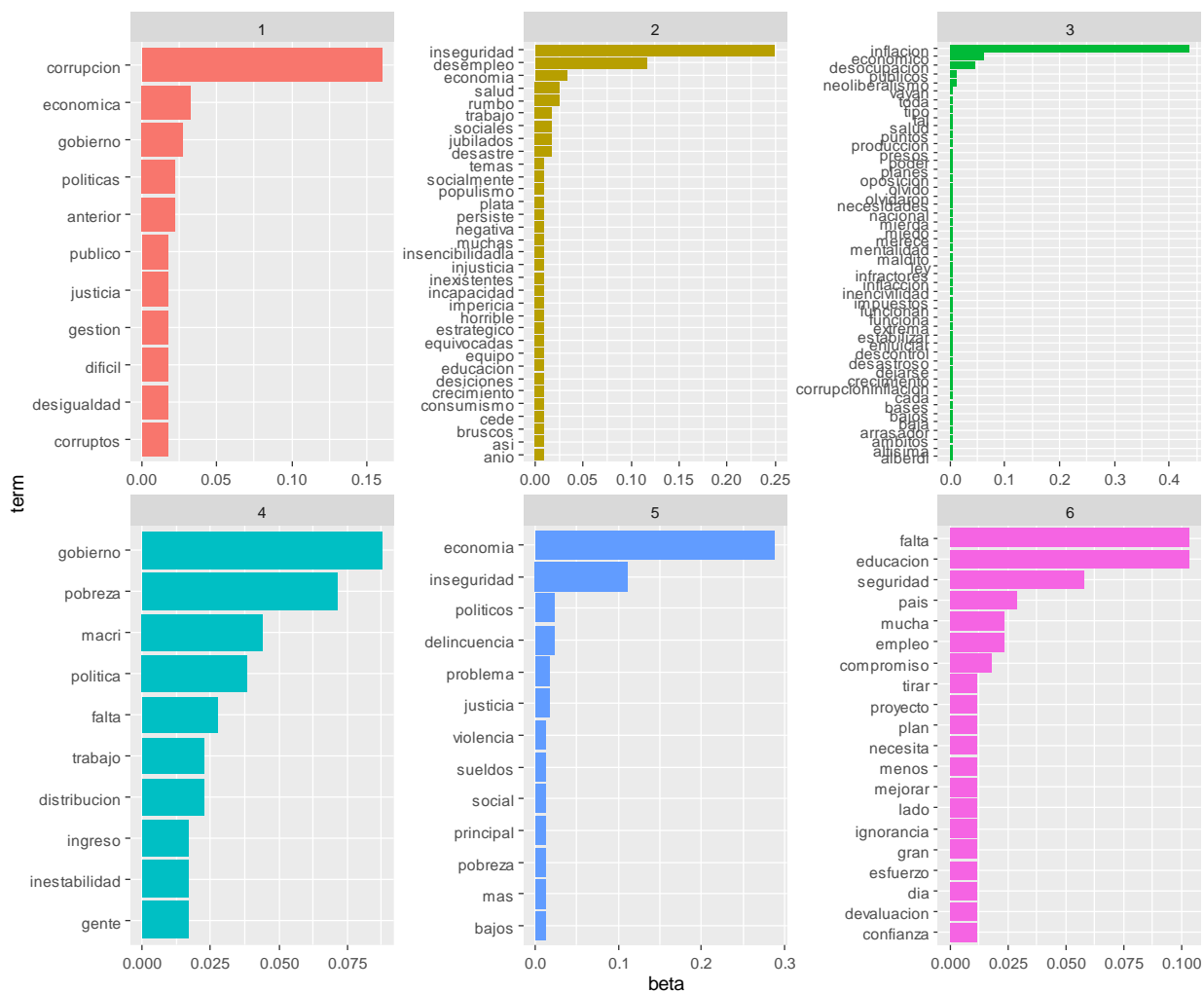
Fuente: Elaboración Propia

Una vez considerada dicha cantidad de temas, es posible evidenciar su composición, tanto a partir de aquellas palabras que mayor contribuyen a la composición de dicho tópico como mediante la observación de respuestas características para cada uno de ellos.

A continuación, se muestran los términos que mayormente distinguen a cada tópico en base a su coeficiente β_i estimado:

Grafico 12.

Desagregación de tópicos para PMI en palabras con mayor contribución



Fuente: Elaboración Propia

Como puede observarse, se hallan seis tópicos diferenciados detallados a continuación:

1. **Corrupción:** representa aquellas respuestas que consideren que el problema más importante del país se centra en la corrupción política, fuertemente ligada al “kirchnerismo”, y la falta de velocidad de acción por parte de la justicia. Marginalmente, se nota la aparición de temas económicos. Ejemplos:

- *“Las incesantes subas en muchísimos rubros. La falta de credibilidad en la justicia contra la terrible corrupción del gobierno anterior”*, respuesta número 168, masculino, 41 años.
 - *“La Justicia lenta para enjuiciar a los corruptos, delincuentes y asesinos”*, respuesta número 21, masculino, 55 años.
 - *“Corrupción gobernante, negocios particulares”*, respuesta número 235, masculino, 53 años.
2. *Inseguridad ciudadana y económica*: representa aquellas respuestas donde los principales problemas del país son la inseguridad, la economía y el desempleo. Su relación se debe a que, en diversas ocasiones, la inseguridad se traduce en inseguridad económica, haciendo referencia también a la situación de los jubilados. Ejemplos:
- *“La inseguridad, desempleo”*, respuesta número 20, masculino, 63 años.
 - *“Economía, jubilados, salud, educación, inseguridad”*, respuesta número 336, femenino, 50 años.
 - *“Economía e inseguridad”*, respuesta número 204, femenino, 22 años.
3. *Economía e inflación*: refiere a aquellos casos donde el principal problema del país refiere a la economía, pero particularmente a un aspecto puntual de este fenómeno: la inflación. Ejemplos:
- *“Inflación altísima”*, respuesta número 91, masculino, 36 años.
 - *“Neoliberalismo económico”*, respuesta número 153, femenino, 49 años.
 - *“La inflación”*, respuesta número 8, femenino, 54 años (entre otros).
4. *Gestión política y pobreza*: refiere a aquellas respuestas donde el principal problema está dado por la gestión política y sus consecuencias, particularmente el incremento de la pobreza en el país. Ejemplos:
- *“Que es increíble que no se den cuenta en el gobierno la cantidad de gente que está creciendo y llevando a la pobreza”*, respuesta número 185, femenino, 26 años.
 - *“El gobierno y su insensibilidad. La falta de trabajo trae hambre y secuelas terribles”*, respuesta número 273, femenino, 53 años.
 - *“El presidente Mauricio Macri”*, respuesta número 260, femenino, 19 años.
5. *Economía y política económica*: refiere a la economía en general como principal problema del país, pero a diferencia del tópico 3, incorpora en este

caso a la política como parte del problema económico y, en consecuencia, a la política económica de la gestión de turno. Ejemplos:

- *“Hay más de uno: economía, sueldos bajos, altos impuestos y servicios, corrupción”*, respuesta número 160, femenino, 56 años.
- *“El principal problema es la economía”*, respuesta número 61, femenino, 43 años.
- *“Economía, injusticia, inequidad, inseguridad”*, respuesta número 302, femenino, 41 años.

6. *Desarrollo*: refiere a diversos aspectos que componen una visión a largo plazo del país en forma de un proyecto de desarrollo. Entre dichos factores, se encuentra la educación, el respecto de ciertas normas, el consumo de drogas, la falta de respeto o pluralidad, entre otros. Ejemplos:

- *“Ignorancia en los tres poderes que hacen que no se tomen las iniciativas para mejorar el país: falta de estudios universitarios, falta de compañerismo, mucha ignorancia en economía”*, respuesta número 350, femenino, 26 años.
- *“Falta de proyecto de país”*, respuesta número 125, femenino, 49 años.
- *“Falta de compromiso y un equipo adecuado de trabajo”*, respuesta número 141, femenino, 23 años.

Al considerar la distribución de las respuestas obtenidas en función de los tópicos asignados a cada una de ellas, se evidencia que, sobre el total de 402 respuestas categorizadas²⁰, el tópico más popular es el relacionado a la problemática de la inflación, seguido por la corrupción y la política económica.

²⁰ Al momento de realizar el análisis, es posible que algunas respuestas sean descartadas, tanto porque no tienen contenido como porque el contenido de la misma no es asignable a un conjunto de palabras (por ejemplo, en el caso que sean completadas por un signo de puntuación como un punto -“.”-). En este caso, las respuestas de los encuestados consideradas fueron 402, sobre un total de 410 respuestas relevadas (98% de los casos).

Tabla 9.

Distribución de respuestas por tópico

Tópico	Cantidad	Porcentaje
Corrupción	80	19,90%
Inseguridad ciudadana y económica	58	14,43%
Economía e inflación	98	24,38%
Gestión pública y pobreza	56	13,93%
Economía y política económica	74	18,41%
Desarrollo	36	8,96%
Total	402	100%

Fuente: Elaboración Propia

Una vez hallados estos tópicos, cabe considerar la distribución de cada uno de ellos en función del consumo declarado de medios de comunicación, por parte de los entrevistados. En este sentido, se presenta a continuación la penetración de cada medio de comunicación en relación a la muestra encuestada:

Tabla 10.

Penetración de medios de comunicación online entre los encuestados

Medio de Comunicación	n	%
Medios de comunicación online	372	90,73%
Infobae	253	61,71%
Clarín	221	53,90%
La Nación	190	46,34%
Página 12	83	20,24%
Perfil	52	12,68%

Fuente: Elaboración Propia

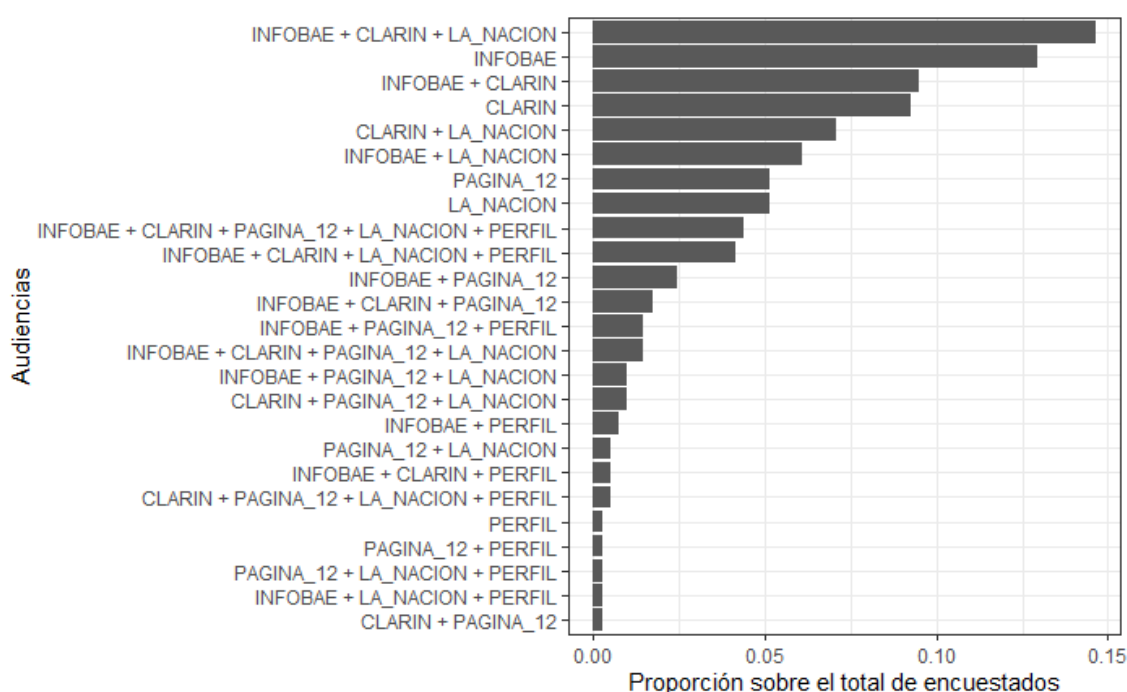
Como puede observarse, el consumo de medios de comunicación de manera online representa a la mayor parte de la población encuestada, mientras que se confirma que la importancia relativa de cada medio de comunicación en la muestra replica el lugar que ocupa cada uno de ellos a nivel nacional, respecto a su consumo por canales digitales²¹.

²¹ Cfr. Capítulo “Metodología” del presente trabajo.

Considerando este marco inicial, es posible indagar acerca del consumo específico de cada medio de comunicación y su solapamiento mutuo, por parte de la audiencia encuestada. En este sentido, a continuación se presentan, de forma gráfica, los diferentes patrones de consumo existentes en la muestra respecto a las noticias en formato digital:

Gráfico 13.

Tamaño relativo de cada audiencia en función de su consumo de medios de comunicación online



Fuente: Elaboración Propia

Tal como queda expresado, las audiencias exclusivas ocupan un rol relevante dentro de la distribución, particularmente en referencia a Infobae y Clarín. Por otro lado, resalta la manera en que Infobae, Clarín y La Nación ocupan, en base a diferentes modalidades de consumo, los primeros lugares en la definición de audiencias. Por último, cabe destacar la relevancia de Página 12 como medio de comunicación que cuenta con una audiencia exclusiva tal que, igualando a La Nación en importancia relativa, quiebra el patrón de consumo anteriormente mencionado.

Considerando las audiencias principales sobre conjunto de audiencias, cabe corroborar la manera en que se distribuyen los tópicos de agenda explorados²².

Tabla 11.
Distribución de Tópicos según Audiencias

Audiencia	Tópico					
	Corrupción	Inseguridad Ciudadana y Económica	Economía e Inflación	Gestión Pública y Pobreza	Economía y política económica	Desarrollo
Infobae + Clarín + La Nación	10,00%	8,20%	18,18%	10,34%	15,07%	29,03%
Infobae	16,25%	14,75%	17,17%	6,90%	6,85%	9,68%
Infobae + Clarín	7,50%	16,39%	8,08%	8,62%	9,59%	9,68%
Clarín	6,25%	16,39%	10,10%	5,17%	10,96%	6,45%
Clarín + La Nación	3,75%	11,48%	8,08%	5,17%	5,48%	12,90%
Infobae + La Nación	7,50%	4,92%	7,07%	6,90%	4,11%	6,45%
Página 12	2,50%	0,00%	3,03%	18,97%	4,11%	6,45%
La Nación	6,25%	6,56%	5,05%	1,72%	6,85%	0,00%
Total general	100%	100%	100%	100%	100%	100%

Fuente: Elaboración Propia

Como puede observarse, la gran mayoría de las menciones sobre el tópico Desarrollo, Economía y política económica y Economía e inflación, se concentran en consumidores de Infobae + Clarín + La Nación, a diferencia de los restantes temas. Por otro lado, Corrupción tiende a concentrarse en la audiencia exclusiva de Infobae, mientras que Inseguridad ciudadana y económica tiende a

²² Se presentan datos para las primeras ocho audiencias, dado que las restantes no cuentan con un volumen tal como para calcular el coeficiente de correlación de Pearson correspondiente para desarrollar la relación entre universos discursivos. Estas son las ocho audiencias que serán consideradas en el resto del documento, representando en su conjunto al 70% de los entrevistados.

concentrarse en la audiencia exclusiva de Clarín y la audiencia de Infobae + Clarín.

En contraste a lo desarrollado previamente, cabe resaltar la particularidad de la audiencia exclusiva de Página 12, sobre la cual tiende a concentrarse significativamente el tópico Gestión Pública y Pobreza, problemática secundaria en el resto de los grupos considerados. Estos resultados, por su parte, explican la baja penetración del tópico aquí mencionado, considerando el pequeño tamaño relativo de la audiencia específica donde dicho tema adquiere mayor relevancia.

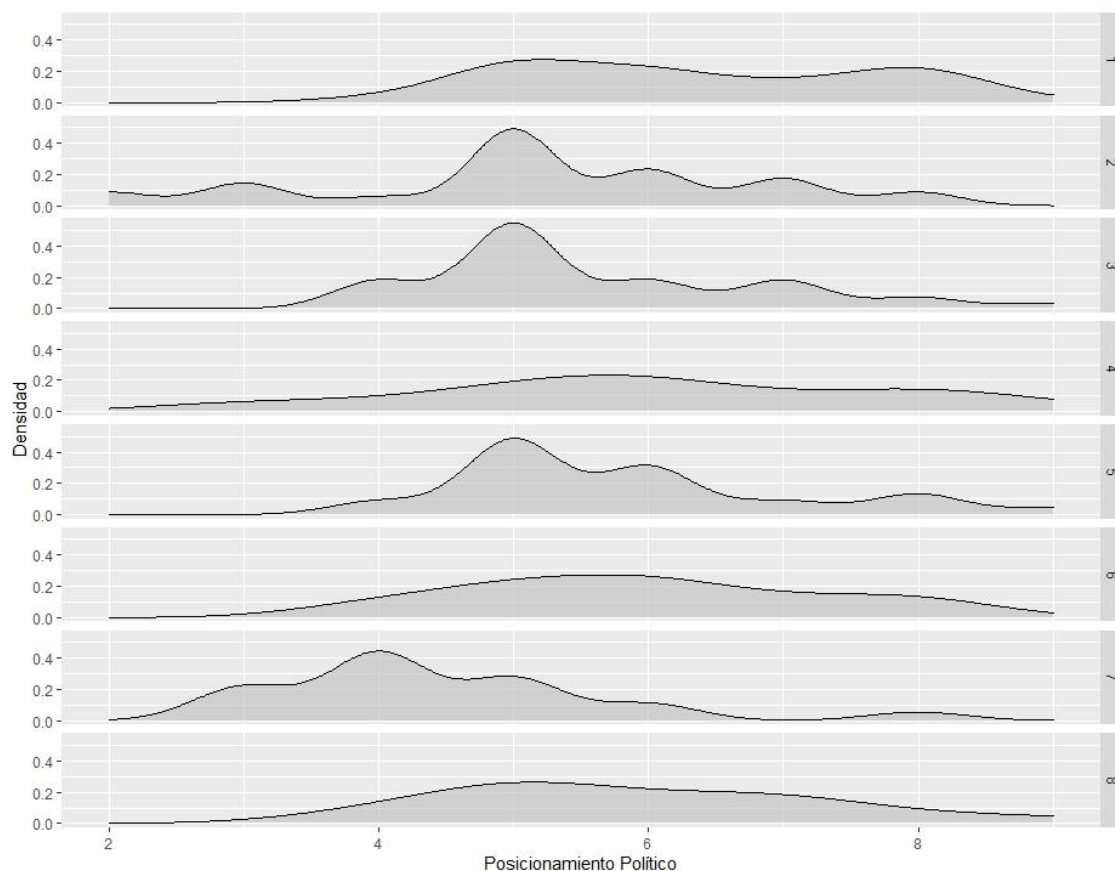
Sobre la base de las diferencias notadas entre la relación de los diferentes tópicos hallados y las audiencias identificadas, cabe introducir también la diferenciación de las audiencias en otros aspectos centrales que hacen a sus características. Particularmente, cabe destacar su posicionamiento político declarado, su confianza en los medios de comunicación y su frecuencia de consumo de los mismos, su edad y su distribución en materia de género.

A continuación, se expresan gráficamente las características de cada grupo en base a su posicionamiento político, entendiendo éste como su autodefinición, por parte de los entrevistados, en el marco de una escala izquierda-derecha²³.

²³ En el marco del cuestionario implementado, el dato presentado corresponde a la pregunta “En una escala del 1 al 10, siendo 1 izquierda y 10 derecha, te pedimos que indiques tu posicionamiento político”

Gráfico 14.

Composición de cada audiencia principal en función de su posicionamiento político



Fuente: Elaboración Propia

Tal como puede observarse, si bien todas las audiencias tienden a concentrarse en una posición política ligada al centro, se ve cómo la audiencia número 7, correspondiente a los consumidores exclusivos de Página 12, tienden a concentrarse en torno a un posicionamiento político de izquierda. Por otro lado, es posible notar la manera en que las distribuciones de las audiencias de La Nación -esto es, la audiencia número 8-, Clarín -la audiencia número 4-, y la correspondiente a Infobae + Clarín + La Nación -esto es, la audiencia número 1-, tienden a hallar una concentración relativa levemente mayor que las restantes, en torno a las posiciones de centro-derecha.

Complementariamente a estos resultados, cabe destacar que la audiencia exclusiva de Página 12 es la que presenta menor dispersión en torno al posicionamiento ideológico promedio del grupo, adoptando ésta una variación

considerablemente menor que las audiencias exclusivas de Infobae, Clarín o La Nación. Cabe resaltar, a su vez, que la audiencia exclusiva de Infobae es la audiencia de mayor dispersión en torno al posicionamiento político promedio del grupo, incluso considerado las audiencias mixtas²⁴. En este sentido, se evidencia una diferencia en la posición político-ideológica de los diferentes grupos, que será retomada más adelante en el presente trabajo, expresada ésta también en la imagen de estos grupos sobre los principales actores políticos del momento.

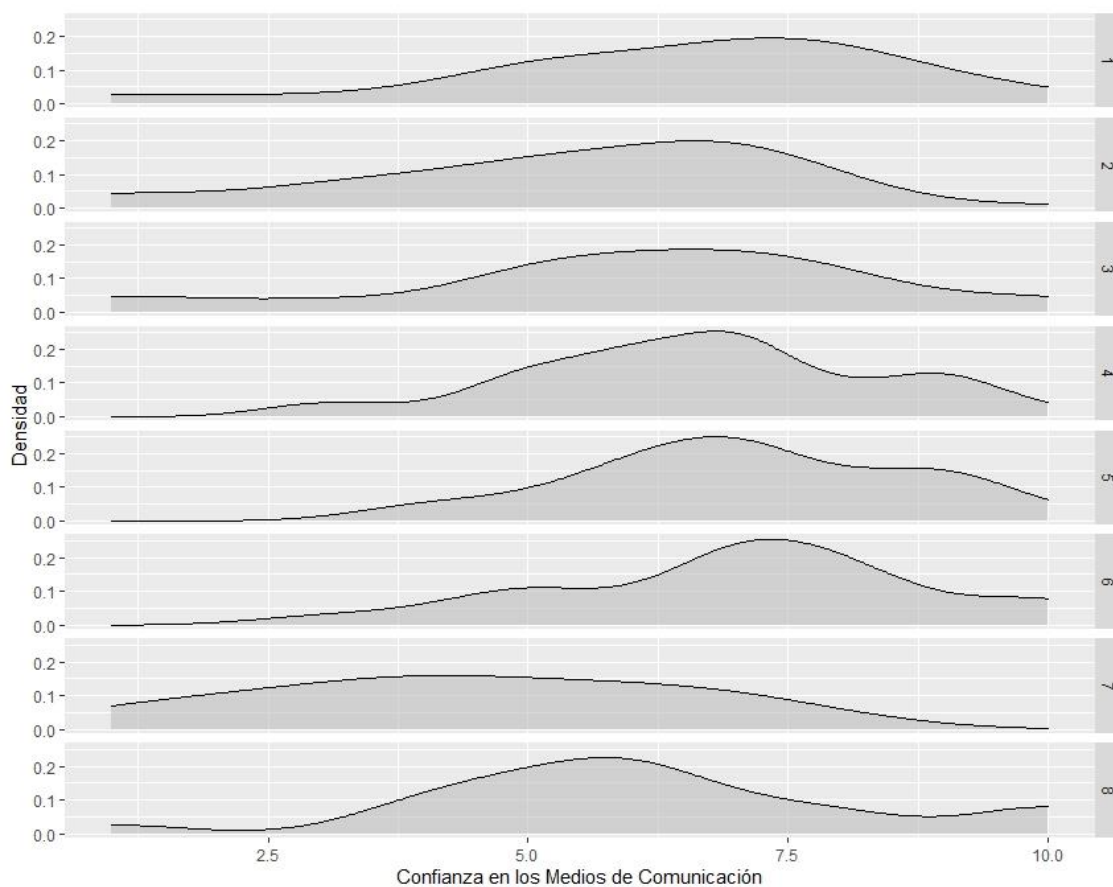
Considerando un aspecto complementario del posicionamiento político por parte de los encuestados, es posible definir cada audiencia identificada en función de su confianza declara en los medios de comunicación²⁵. En este sentido, se observa la siguiente distribución sobre la variable indicada, para cada segmento:

²⁴ Los estadísticos descriptivos correspondientes al posicionamiento político de cada audiencia, al igual que los correspondientes a la confianza y el consumo de medios de comunicación, pueden ser encontrados en el Anexo del presente trabajo.

²⁵ En el marco del cuestionario implementado, el dato presentado corresponde a la pregunta “En una escala del 1 al 10, siendo 1 nada de confianza y 10 total confianza, te pedimos que indiques cuánta confianza te generan los medios de comunicación online en general.

Gráfico 15.

Composición de cada audiencia principal en función de su confianza declarada en los medios de comunicación



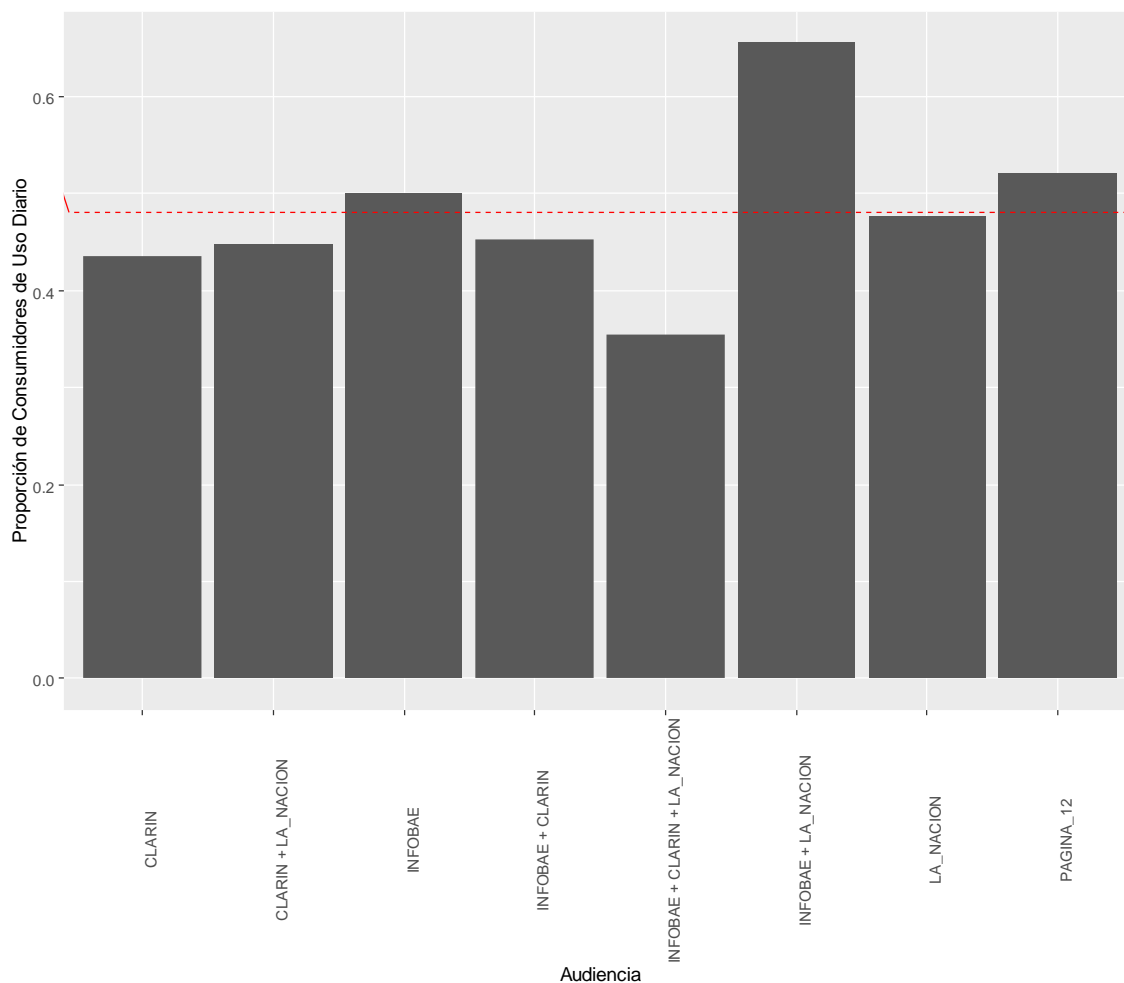
Fuente: Elaboración Propia

Se evidencia, en este marco, una considerable confianza para con los medios de comunicación, por parte de todo el conjunto de audiencias. Resalta, no obstante, la audiencia exclusiva de Página 12 por su considerable falta de confianza en los mismos, a la vez que la concentración de los lectores exclusivos de La Nación, en torno a los valores medios de la escala considerada.

Complementariamente, es posible describir a las audiencias en función de la frecuencia con la cual consumen los medios que declaran leer. En este sentido, se presenta a continuación la proporción de usuarios habituales de dichos medios online, es decir, aquellos que los consumen a diario.

Gráfico 16.

Proporción de Usuarios con Consumo Diario de medios de comunicación online, según Audiencia



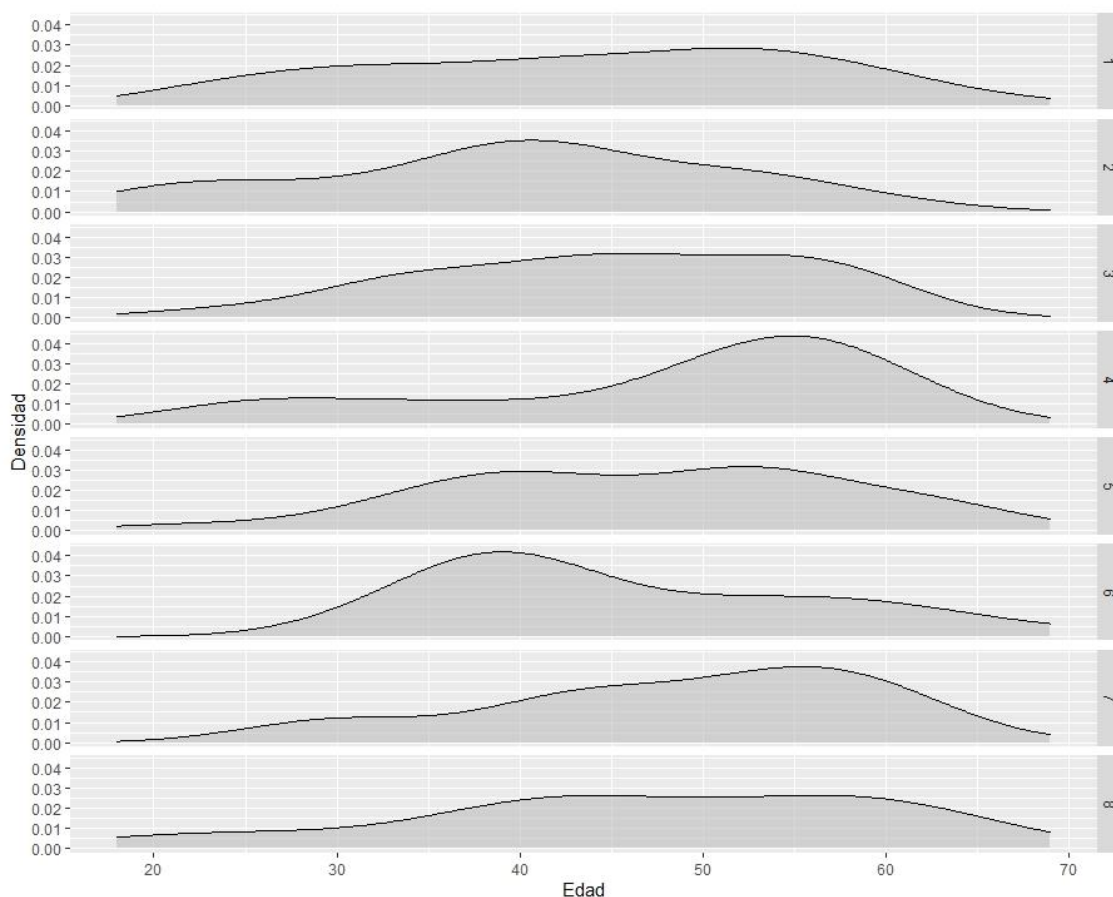
Fuente: Elaboración Propia

Tal como puede observarse, si bien el consumo diario de los medios de comunicación online declarados es elevado en todo el conjunto de audiencias consideradas, cabe señalar la importancia de esta dinámica de consumo en la audiencia de Infobae + La Nación. Dicho segmento, con una representación de usuarios diarios de un 65% sobre el total de dicha audiencia, presenta un valor considerablemente superior al 48% de usuarios diarios, estadístico que representa la proporción de usuarios de consumo diario promedio, entre las diferentes audiencias consideradas (este valor es mostrado, en el gráfico, mediante la línea roja punteada que atraviesa las diferentes audiencias).

En cuanto a las características sociodemográficas de las audiencias consideradas, cabe destacar, en primer lugar, la diferencia de estos grupos en torno a la edad de sus integrantes. Gráficamente:

Gráfico 17.

Composición de cada audiencia principal en función de su edad



Fuente: Elaboración Propia

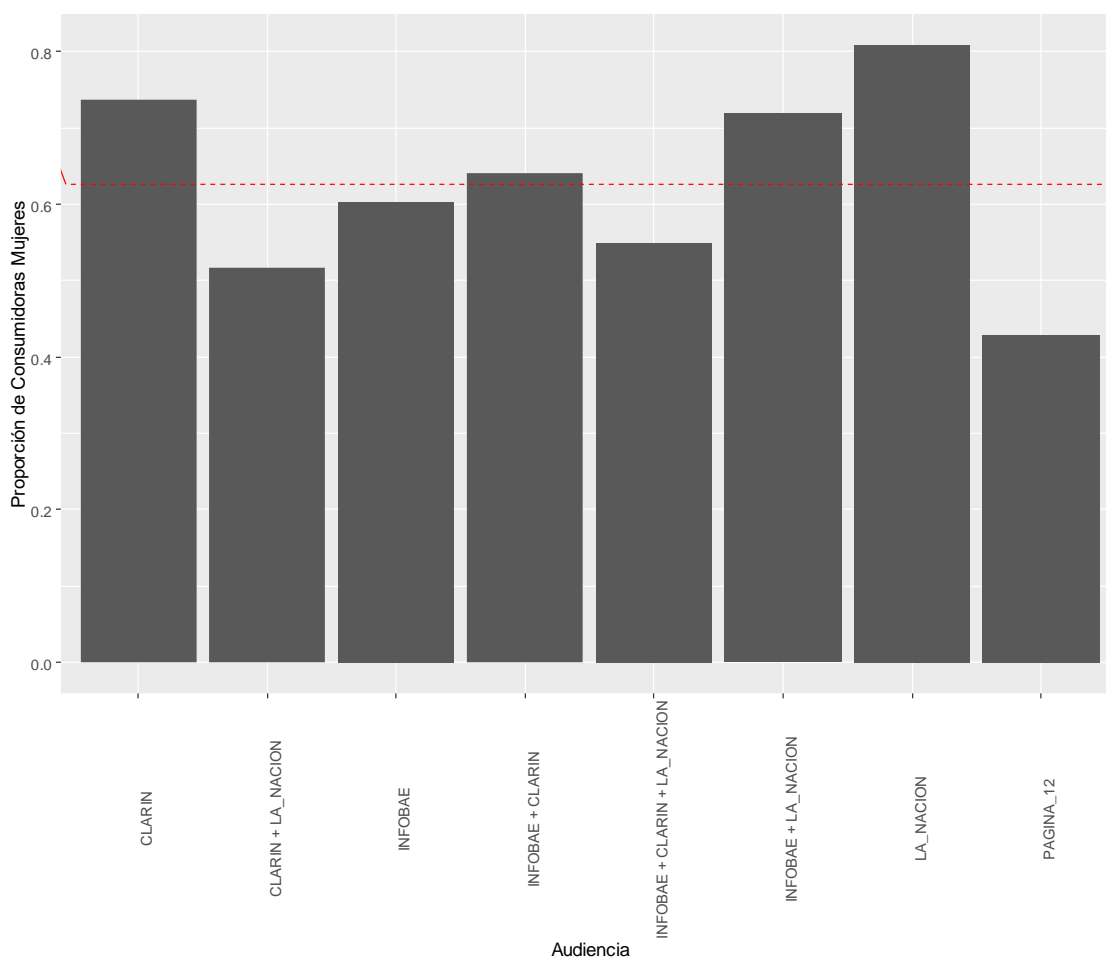
Como puede observarse, no se encuentran diferencias significativas en torno a la edad de las audiencias, con excepción de las audiencias de Infobae + Clarín + La Nación, y la audiencia exclusiva de Infobae -esta última, particularmente-, que tienden a mostrar una mayor concentración de lectores de edad más joven. Por otro lado, pueden distinguirse las audiencias exclusivas de Clarín y Página 12, por la edad mayor de sus consumidores online. A su vez, cabe señalar que,

en materia de dispersión en base a la edad de cada uno de estos grupos, no se observan diferencias significativas²⁶.

Por último, con el fin de terminar de describir estas audiencias en términos demográficos, se presenta a continuación la proporción de usuarios mujeres, correspondiente a cada una de ellas:

Gráfico 18.

Proporción de Consumidoras Mujeres, según Audiencia



Fuente: Elaboración Propia

Así, se evidencia que las audiencias donde el segmento femenino es más fuerte son las audiencias exclusivas de La Nación y Clarín, junto a las audiencias mixtas de consumidores de Infobae + La Nación y de Infobae + Clarín. Cabe

²⁶ Los estadísticos descriptivos correspondientes a las variables demográficas de las audiencias identificadas, pueden encontrarse en el Anexo del presente trabajo.

señalar que, si bien todas las audiencias consideradas se muestran equilibradas en cuanto a su composición en materia de género, los segmentos anteriormente mencionados se encuentran por sobre la media de dicho estadístico para el conjunto de audiencias considerado. Por otro lado, resta destacar la particularidad de la audiencia exclusiva de Página 12, la cual es la que menos proporción de mujeres presenta entre sus lectores.

En base a la evidencia proveída por los análisis particularizados del discurso de los medios de comunicación seleccionados en la red social Twitter, por un lado, y de la opinión pública usuaria de internet, por el otro, es posible evaluar el grado de coincidencia -o discrepancia- entre ambos universos discursivos. Dicho análisis, será realizado en el capítulo siguiente.

Este análisis, por su parte, será complementado por las particularidades que presenta cada audiencia, tanto en términos de su posicionamiento político como respecto a los medios de comunicación, a sus características sociodemográficas. De esta manera, la relación no lineal expresada a continuación, entre las características de las audiencias, sus patrones de consumo de medios de comunicación, su posicionamiento político correspondiente, y la correlación evidenciada entre su discurso y el discurso mediático, permite retomar y confirmar la posición planteada en el marco teórico del presente trabajo, donde se señaló que no existe un orden de causalidad entre las agendas analizadas, sino que se ubican en el mismo nivel en términos analíticos, y ambas son resultantes de un conjunto complejo de factores.

CAPÍTULO 5. ¿AGENDAS RELACIONADAS? ANÁLISIS DE LA RELACIÓN ENTRE EL DISCURSO DE LA OPINIÓN PÚBLICA Y EL DISCURSO MEDIÁTICO

“En Twitter, la gran mayoría de los que me rodean están de acuerdo con lo que yo pienso” (Calvo, E., 2015, p. 29)

Como pudo evidenciarse en los capítulos previos, el discurso de los medios de comunicación, particularmente aquel discurso correspondiente a los medios de comunicación seleccionados en el marco de la red social Twitter, no coincide explícitamente con el discurso de la opinión pública. Esto se expresa tanto en los tópicos resultantes del análisis de cada uno de los tipos de actores -esto es, medios de comunicación y opinión pública-, como en la poca capacidad explicativa que encuentra el modelo resultante del corpus proveniente de la encuesta respecto al discurso de los medios de comunicación²⁷. Si bien esto podría rechazar, a priori, la hipótesis subyacente al presente trabajo, es en realidad una expresión del momento de construcción de agenda considerado por parte de los medios de comunicación en el presente trabajo²⁸, en contraposición a la sedimentación propia de la agenda pública en grandes tópicos generales, expresado en el discurso de la opinión pública.

Dada dicha falta de relación entre las agrupaciones en tópicos de cada discurso, una vía alternativa para analizar la relación entre estos universos simbólicos es partir desde un punto de vista desagregado, es decir considerando los términos que cada uno de ellos emplea, individualmente. Esto permitirá construir un nexo entre ambos discursos, a partir de la unidad mínima de cada uno de ellos. En este sentido, se presentan a continuación los diez términos mayormente

²⁷ Para llegar a esta conclusión, fueron implementadas técnicas analíticas que no proveyeron resultados significativos. Particularmente, se implementó el modelo LDA resultante del análisis exploratorio en base a tópicos realizado en el corpus derivado de la encuesta, sobre el texto extraído en formato de tweet, representante del discurso mediático. En este caso, los textos identificados para cada tópico no presentaban coherencia semántica con la definición de cada uno de estos temas en base a la encuesta.

²⁸ Para una explicación más desarrollada de los diferentes momentos de establecimiento de agenda, ver el capítulo “Marco Teórico” del presente trabajo.

utilizados por cada uno de los actores analizados -opinión pública, por un lado, y medios de comunicación, por el otro-, acompañado de su frecuencia en el discurso de los actores restantes.

Tabla 12.

Diez palabras más utilizadas por la Opinión Pública, en relación a los Medios de Comunicación en general y cada medio de comunicación en particular (% Penetración)

Palabra	Opinión Pública	Medios de Comunicación	Clarín	Infobae	La Nación	Página 12	Perfil
inflación	9,54%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
economía	6,85%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
inseguridad	6,11%	0,01%	0,01%	0,00%	0,00%	0,00%	0,01%
corrupción	3,79%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
falta	3,06%	0,03%	0,04%	0,02%	0,04%	0,06%	0,05%
gobierno	2,57%	0,28%	0,19%	0,30%	0,28%	0,43%	0,24%
educación	2,32%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
pobreza	1,83%	0,02%	0,00%	0,02%	0,05%	0,03%	0,01%
desempleo	1,71%	0,00%	0,00%	0,00%	0,00%	0,01%	0,00%
económico	1,59%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%

Fuente: Elaboración Propia

Tabla 13.

Diez palabras más utilizadas por los Medios de Comunicación en general, en relación a la Opinión Pública y cada medio de comunicación en particular (% Penetración)

Palabra	Medios de Comunicación	Opinión Pública	Clarín	Infobae	La Nación	Página 12	Perfil
juan	0,57%	0,00%	0,72%	0,64%	0,52%	0,28%	0,36%
river	0,54%	0,00%	0,57%	0,71%	0,69%	0,11%	0,10%
boca	0,53%	0,00%	0,57%	0,70%	0,66%	0,10%	0,07%
argentina	0,51%	0,00%	0,42%	0,61%	0,55%	0,26%	0,37%
macri	0,47%	0,98%	0,35%	0,54%	0,52%	0,55%	0,33%
años	0,41%	0,00%	0,52%	0,42%	0,38%	0,32%	0,33%
tras	0,36%	0,00%	0,35%	0,33%	0,32%	0,15%	0,54%
san	0,35%	0,00%	0,45%	0,42%	0,38%	0,18%	0,14%
final	0,31%	0,00%	0,38%	0,23%	0,52%	0,15%	0,39%
video	0,29%	0,00%	0,33%	0,21%	0,15%	0,15%	0,64%

Fuente: Elaboración Propia

Como puede observarse, entre los términos mayormente utilizados por la opinión pública resalta el concepto de “gobierno” como un punto de contacto con los medios de comunicación, mientras que, partiendo del discurso mediático, se encuentra un punto de contacto en el concepto de “macri”. En este marco, se plantea la necesidad de partir del universo de palabras coincidentes entre ambos discursos, con el fin de evidenciar si la importancia relativa que cada uno de estos discursos le brinda a cada concepto, se ve reflejado en el discurso de la audiencia opuesta. Si bien la cantidad de conceptos compartidos es considerablemente baja respecto al total de términos empleados -245 términos sobre un total de 21.475 entre ambos grupos, es decir un 1,14%-, esto se debe sobre todo al volumen de contenido por parte de los medios de comunicación, frente a la capacidad de la opinión pública de aglomerar esta serie de eventos en grandes temas, evocados de manera sintética y general.

Al evaluar dicha la relación entre ambos discursos sobre este universo discursivo coincidente mediante el cálculo del coeficiente r de Pearson, se corrobora una relación de baja intensidad en todos los casos, únicamente tendiendo a una relación de media intensidad en el caso de Página 12. En este sentido, es

destacable que la relación más alta se registre respecto al medio de comunicación explícitamente opositor al gobierno en gestión en el momento del análisis, del cual, además, se ha demostrado que es el medio con mayor independencia relativa respecto a los tópicos que trata en su discurso en redes sociales²⁹.

Si bien de baja intensidad, cabe destacar que dichas relaciones son de carácter significativo, con un nivel de significación de 0,05 en todos los casos, lo cual permite demostrar la relevancia de las relaciones halladas, pese a dicha intensidad.

Tabla 14.

Correlación entre términos utilizados por la Opinión Pública y los Medios de Comunicación

	Medios de Comunicación (total)	Infobae	Clarín	La Nación	Perfil	Página 12
<i>Coef. R de Pearson</i>	0,215	0,201	0,185	0,207	0,211	0,396
<i>Obs.</i>	245	199	185	180	189	157
Opinión Pública <i>t</i>	3,4306	2,8865	2,5466	2,8302	2,9494	5,3747
<i>df</i>	243	197	183	178	187	155
<i>p-value</i>	0,0007	0,0043	0,0117	0,0052	0,0036	0,0000

Fuente: Elaboración Propia

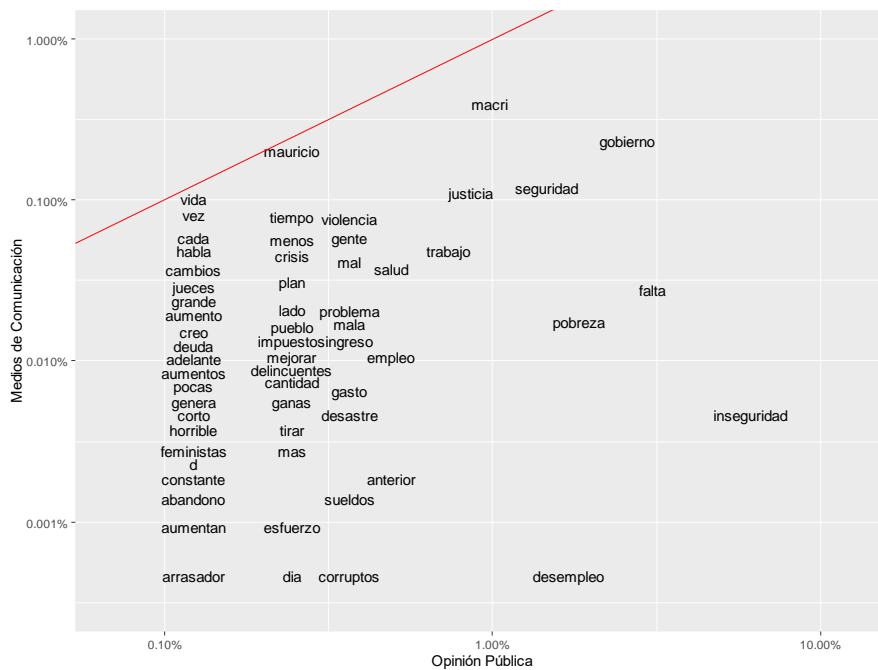
A continuación, es presentada gráficamente la relación entre los conceptos utilizados por los medios de comunicación en conjunto y la opinión pública. Como puede evidenciarse, las relaciones de mayor intensidad de hallan en torno a conceptos asociados a la gestión actual de gobierno como “Mauricio” o “macri”, y relacionado a esto el concepto “gobierno”, aunque también se evidencia una relación no menor en torno a problemas estructurales como “seguridad”, “justicia”, “pobreza”, “salud” o “trabajo”. Cabe destacar que, al ser esperable la

²⁹ Cfr. Capítulo “Análisis de Contenido. Medios de Comunicación”.

baja frecuencia relativa de cada término dada la diversidad de conceptos que integran un discurso dado, lo que vale considerar es la relación entre dichas frecuencias relativas respectivas a cada universo discursivo, antes que la penetración específica de cada término.

Gráfico 19.

Relación entre términos utilizados por los Medios de Comunicación y la Opinión Pública

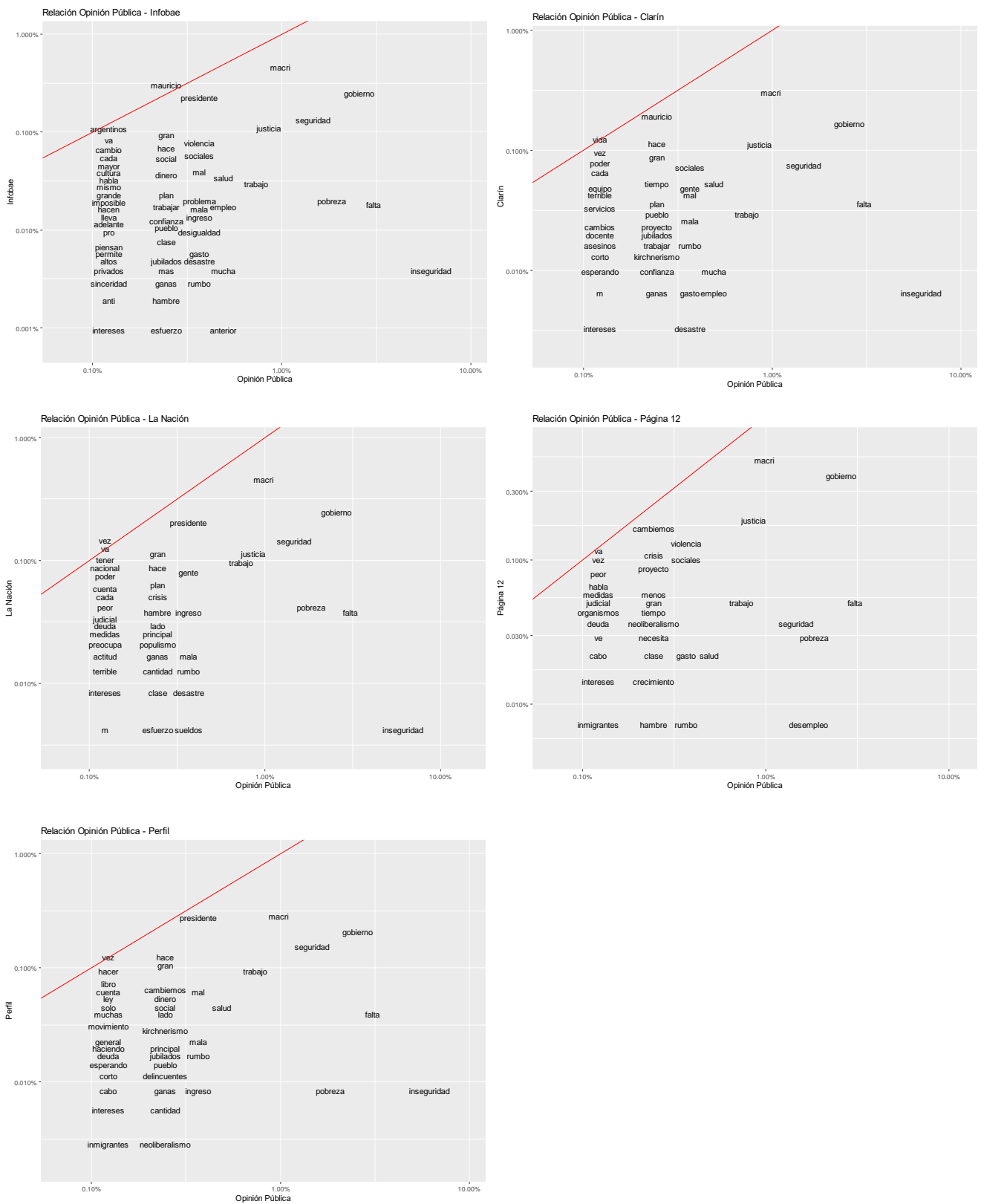


Fuente: Elaboración Propia

Complementariamente a lo anteriormente dicho, es posible evidenciar la relación entre los términos utilizados por la Opinión Pública y cada Medio de Comunicación, particularmente. Como puede observarse a continuación, en todos los casos la tendencia de la relación adopta características similares, aunque difiriendo levemente en intensidad, mientras que los conceptos relacionados encuentran elevada similitud con lo observado en el caso del análisis agregado. Particularmente, resalta la relevancia que adopta el término “trabajo” en el caso de los medios de comunicación La Nación y Perfil, a la vez que la pérdida de relevancia de la palabra “justicia” en relación a este último medio.

Gráfico 20.

Relación entre términos utilizados por cada Medio de Comunicación y la Opinión Pública



Fuente: Elaboración Propia

Una vez expuesta, en términos generales, la relación entre el contenido compartido en redes sociales por parte de los medios de comunicación seleccionados y la descripción del problema más importante según la opinión pública, es posible profundizar dicho análisis a partir de la exploración de la relación entre las expresiones de cada medio de comunicación en particular y su audiencia específica. En este sentido, se replicará el análisis de correlación realizado, esta vez a partir de una segmentación en audiencias en función de los medios de comunicación que declaren consumir los encuestados, realizada en el capítulo previo.

Considerando las ocho audiencias principales, se presenta a continuación la correlación calculada entre su discurso y el discurso mediático que consume.

Tabla 15.

Correlación entre universos discursivos, para cada una de las audiencias principales

Audiencia	Coef. R de Pearson	Obs.³⁰	t	df	p-value
INFOBAE + CLARIN + LA NACION	0,062	70	0,512	68	0,610
INFOBAE	0,155	152	1,922	150	0,056
INFOBAE + CLARIN	0,079	41	0,494	39	0,624
CLARIN	0,191	72	1,626	70	0,108
CLARIN + LA NACION	-0,142	25	-0,690	23	0,497
INFOBAE + LA NACION	0,420	34	2,619	32	0,013
PÁGINA 12	0,646	66	6,766	64	0,000
LA NACION	0,211	72	1,803	70	0,076

Fuente: Elaboración Propia

En este sentido, cabe destacar cómo la relación mayormente significativa en términos de su intensidad, se da entre Página 12 y su audiencia exclusiva. Por otro lado, cabe señalar la considerable correlación entre la audiencia que

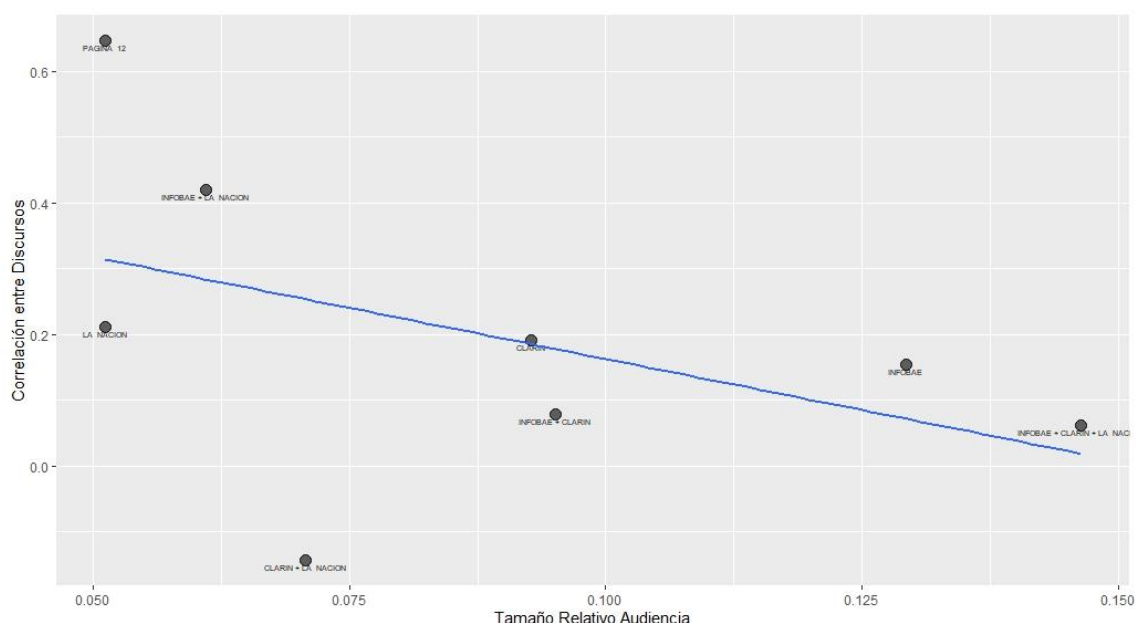
³⁰ “Observaciones” refiere, en este caso, a la cantidad de términos en común entre ambos universos discursivos, sobre los cuales se evalúa la intensidad de dicha relación.

consume Infobae y La Nación, respecto al discurso de dichos medios de comunicación³¹.

Complementariamente a esto, es posible indagar si la lealtad de los integrantes de cada audiencia al contenido de los medios de comunicación, se relaciona con el tamaño relativo de cada una de ellas. En este sentido, se presenta la comparación entre el tamaño relativo de cada audiencia y la correlación entre el contenido de su discurso y el de los medios de comunicación que la definen.

Gráfico 21.

*Relación entre Tamaño de la audiencia y Correlación entre discursos*³²



Fuente: Elaboración Propia

Tal como puede observarse, existe una relación negativa entre estas dos variables: a medida que la audiencia representa una mayor proporción entre la población, menos se correlaciona su discurso con el de los medios de comunicación online. Por otro lado, se evidencia el rol protagónico que adopta

³¹ Si bien gran parte de las relaciones evaluadas no presenta carácter significativo en términos estadísticos, esto puede estar relacionado con la cantidad de observaciones - es decir, de términos- que componen al universo sobre el cual dichas relaciones son evaluadas. En este sentido, cabe considerar estas magnitudes en función del ordenamiento que establecen entre las diferentes audiencias, y su relación con otras variables.

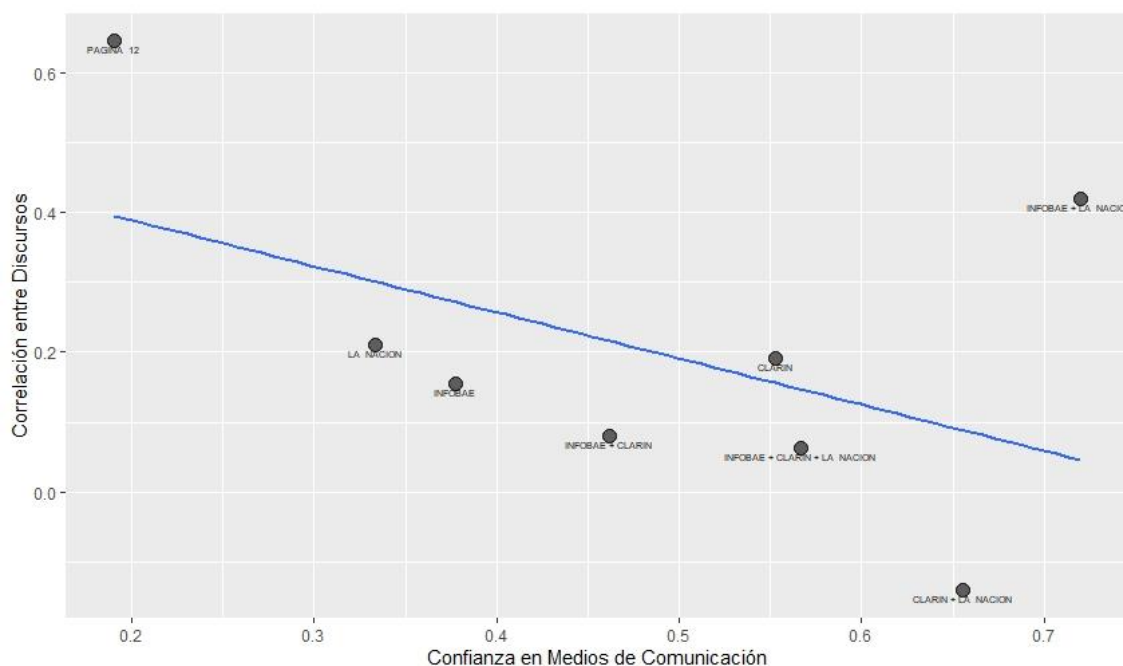
³² Cada nombre en el gráfico representa los medios de comunicación consumidos de manera online por cada audiencia.

Página 12 como medio comunicación de nicho, a la hora de definir esta relación: representa la audiencia de menor tamaño relativo junto a La Nación, entre las audiencias consideradas, pero resalta la alta fidelidad de su audiencia a su contenido. Complementariamente a este planteo, cabe destacar cómo el resto de las audiencias exclusivas -esto es, Clarín e Infobae-, mantienen un similar nivel de correlación entre los universos discursivos considerados, respecto a La Nación, incluso teniendo en cuenta que su tamaño relativo es mayor.

En este sentido, cabe considerar si las diferencias notadas entre las diferentes audiencias y su relación con el contenido mediático consumido encuentra relación con particularidades ideológicas propias de estos grupos. Específicamente, indagaremos en primer lugar sobre la relación entre la correlación de los universos discursivos y la confianza declarada sobre los medios de comunicación, mientras que por otro lado consideraremos la relación existente entre dicha correlación y la imagen declarada de estas audiencias sobre los dos representantes políticos más influyentes del momento: Mauricio Macri, en representación de la postura oficialista de gobierno, y Cristina Fernández de Kirchner, como representante principal de la oposición política al gobierno.

Gráfico 22.

*Relación entre Confianza en medios de comunicación y Correlación entre discursos*³³



Fuente: Elaboración Propia

Cabe notar, a partir de lo observado en el gráfico previo, la manera en que la correlación entre los universos discursivos considerados cae a medida que incrementa la confianza declarada respecto a los medios de comunicación, por parte de cada audiencia.

Si bien, a priori, puede resultar en una observación contradictoria, a partir de notar que, en el cuadrante de mayor confianza y menor correlación, también predominan aquellas audiencias que consumen una diversidad de medios de comunicación de manera online, es posible inferir una cierta amplitud en la mirada respecto a la problemática del país por parte de estas audiencias, producto, en parte, del consumo de una diversidad de voces mediáticas. Por otro lado, cabe destacar el efecto que adopta el hecho de consumir una diversidad de medios de comunicación sobre la confianza en dichos medios, notando que

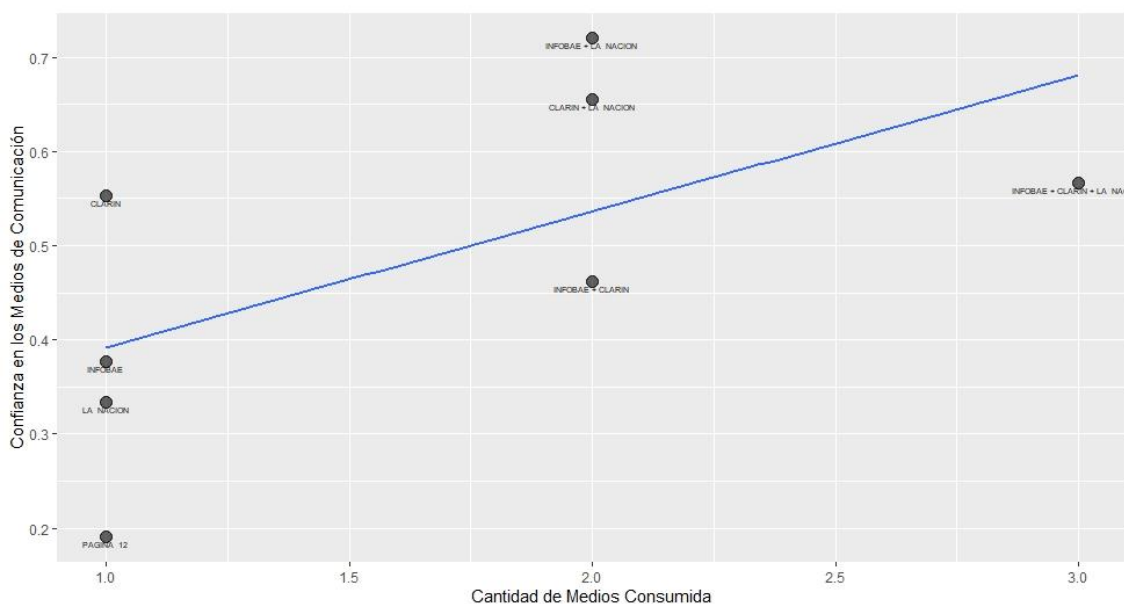
³³ Cada nombre en el gráfico representa los medios de comunicación consumidos de manera online por cada audiencia. Se define como “confianza en medios de comunicación” a aquella proporción de entrevistados que indicó su nivel de confianza en los medios de comunicación, en el marco de una escala de 1 a 10, con un puntaje de entre 7 y 10 puntos.

la mayor confianza se deposita entre aquellos que, con excepción de la audiencia exclusiva de Clarín, consumen más de una voz mediática.

Con el fin de corroborar este planteo, a continuación, se muestra gráficamente la relación positiva entre la cantidad de medios de comunicación consumida y la confianza en los medios, en base a las audiencias identificadas:

Gráfico 23.

Relación entre Confianza en medios de comunicación y Cantidad de Medios consumidos

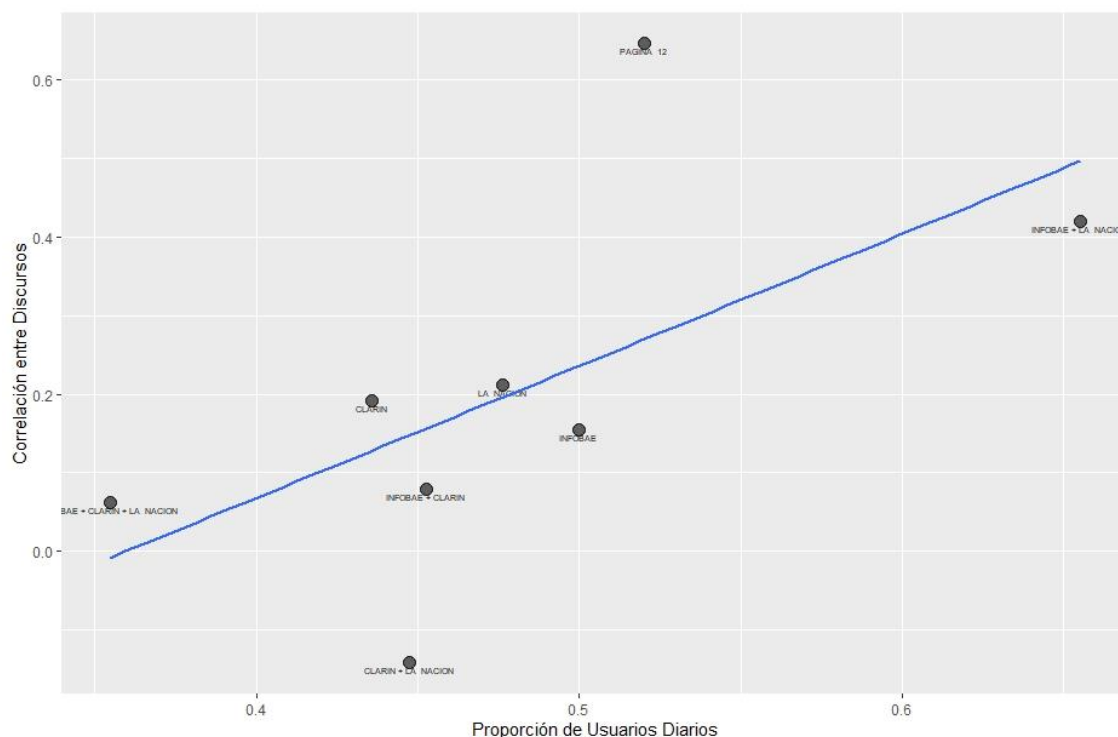


Fuente: Elaboración Propia

Teniendo en cuenta, a la vez que la confianza en los medios de comunicación y la cantidad de medios consumida por cada segmento, la dinámica de consumo de dichos medios -esto es, la frecuencia con la que lo hacen-, se evidencia una relación positiva entre la correlación existente entre los universos discursivos considerados y la proporción de usuarios diarios en cada audiencia. En este sentido, a medida que se observa una mayor cantidad de usuarios habituales a los medios de comunicación online seleccionados, se evidencia una mayor correlación entre el discurso de estos medios y el de sus audiencias. Gráficamente, dicha relación se expresa de la siguiente manera:

Gráfico 24.

Relación entre Proporción de usuarios diarios y Correlación entre discursos



Fuente: Elaboración Propia

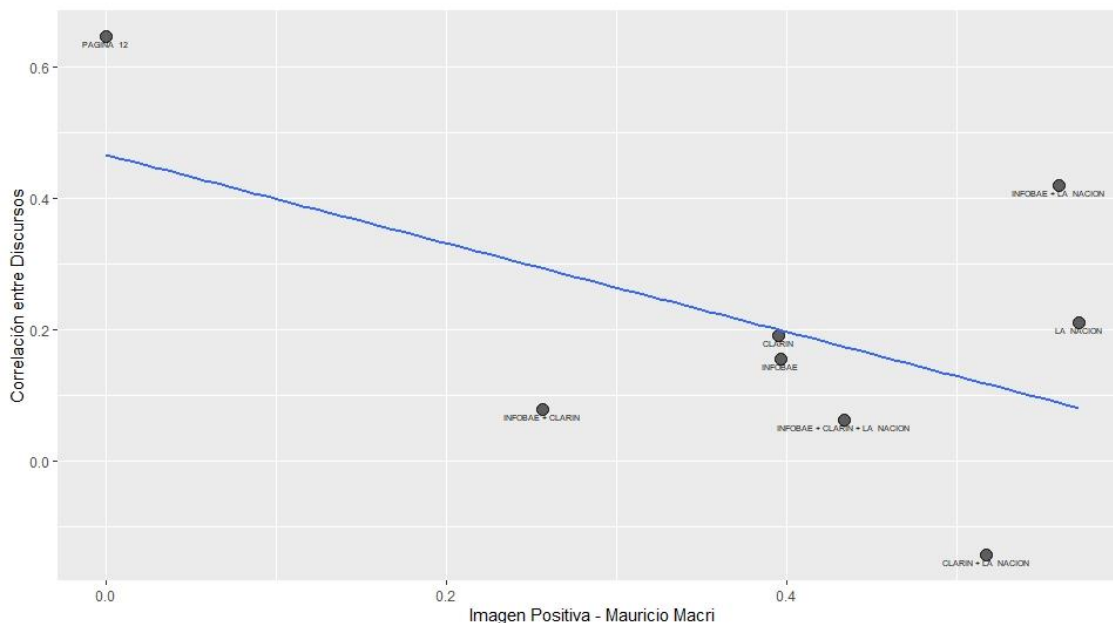
Adicionalmente a lo planteado, resta considerar la particular postura de la audiencia exclusiva de Página 12. Tal como puede observarse, este conjunto de encuestados presenta el menor nivel de confianza en los medios de comunicación, pero a su vez representa el mayor nivel de correlación encontrado entre su discurso y el discurso mediático del canal informativo que consume. Si bien se ha demostrado que el consumir un solo medio de comunicación se relaciona con la menor confianza depositada en ellos, cabe contrastar la correlación entre su discurso y el de su audiencia, con la posición ideológico-política de esta última, con el fin de indagar si esta fuerte coincidencia con el discurso mediático acompañado de un notable distanciamiento político con la figura de los medios de comunicación, se debe en realidad a su propio posicionamiento político.

Para realizar dicha exploración, se presenta a continuación la relación entre la correlación de los universos discursivos y la imagen positiva de los dos principales políticos al momento de realizar el estudio, para cada una de las audiencias. Adicionalmente, se presenta la relación entre el primer indicador y la

autodefinición de los entrevistados de cada audiencia, en base al binomio político izquierda-derecha.

Gráfico 25.

*Relación entre Imagen Positiva sobre Mauricio Macri y Correlación entre discursos*³⁴



Fuente: Elaboración Propia

Tal como puede observarse, Mauricio Macri presenta una alta imagen positiva entre ciertas audiencias mixtas que incluyen a Clarín, Infobae y La Nación, y resalta particularmente su imagen positiva entre la audiencia exclusiva de este último medio de comunicación.

Por otro lado, si bien su imagen se debilita levemente entre las audiencias exclusivas de Clarín e Infobae, cabe señalar que los consumidores de dichos medios mantienen aún una imagen positiva notable sobre este actor político. Es de destacar, en este sentido, la manera en que dicha imagen se debilita a la hora de considerar el consumo de estos medios de comunicación de manera conjunta,

³⁴ Cada nombre en el gráfico representa los medios de comunicación consumidos de manera online por cada audiencia. Se define como “imagen positiva” a aquella proporción de entrevistados que indicó su imagen sobre Mauricio Macri, en el marco de una escala de 1 a 10, con un puntaje de entre 7 y 10 puntos.

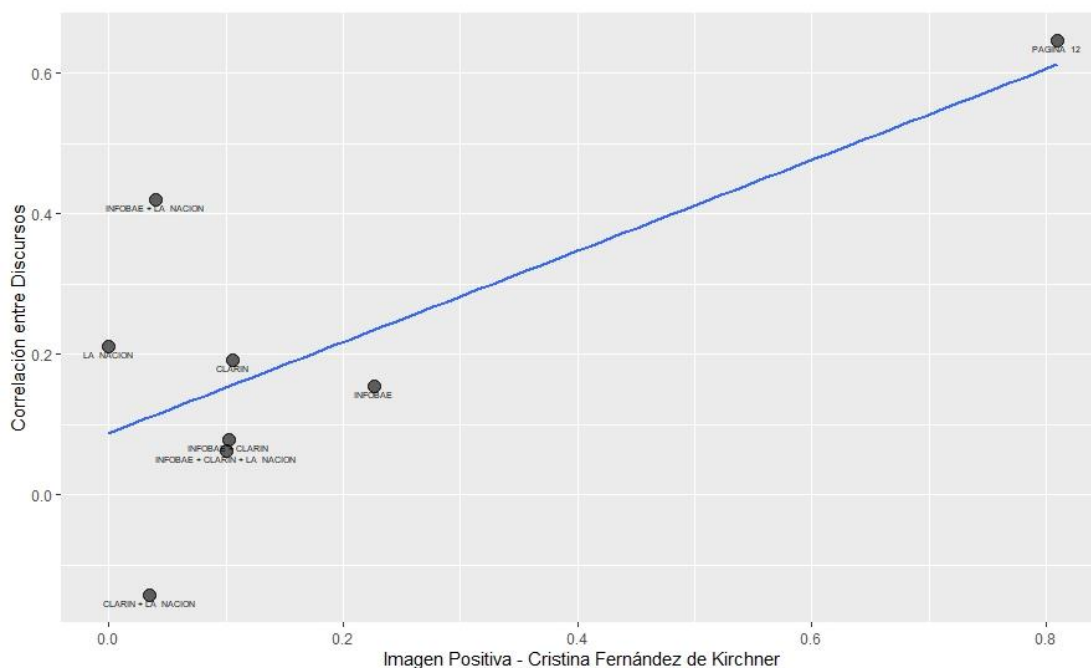
pudiendo inferir que este contenido mediático no refuerza, conjuntamente, dicho posicionamiento político.

Por último, resta considerar a la audiencia exclusiva de Página 12. En relación a lo considerado previamente, el posicionamiento político de dicha audiencia encuentra amplia concordancia con la postura adoptada frente a los medios de comunicación, siendo éstas dos dimensiones donde este grupo logra diferenciarse ampliamente de las restantes audiencias. Complementariamente a esto, cabe destacar la manera en que se da dicha diferenciación, donde no solo se evidencia un distanciamiento frente a la figura de Mauricio Macri, sino que se reafirma el distanciamiento frente a los restantes medios de comunicación, en función de la evidente relación negativa entre la valoración positiva de dicho actor político y la correlación entre el discurso propio y el del contenido mediático consumido.

Con el fin de notar la posición opositora a la recientemente descrita, a continuación, se presenta la relación entre el principal indicador evaluado y la imagen positiva de Cristina Fernández de Kirchner:

Gráfico 26.

*Relación entre Imagen Positiva sobre Cristina Fernández de Kirchner y Correlación entre discursos*³⁵



Fuente: Elaboración Propia

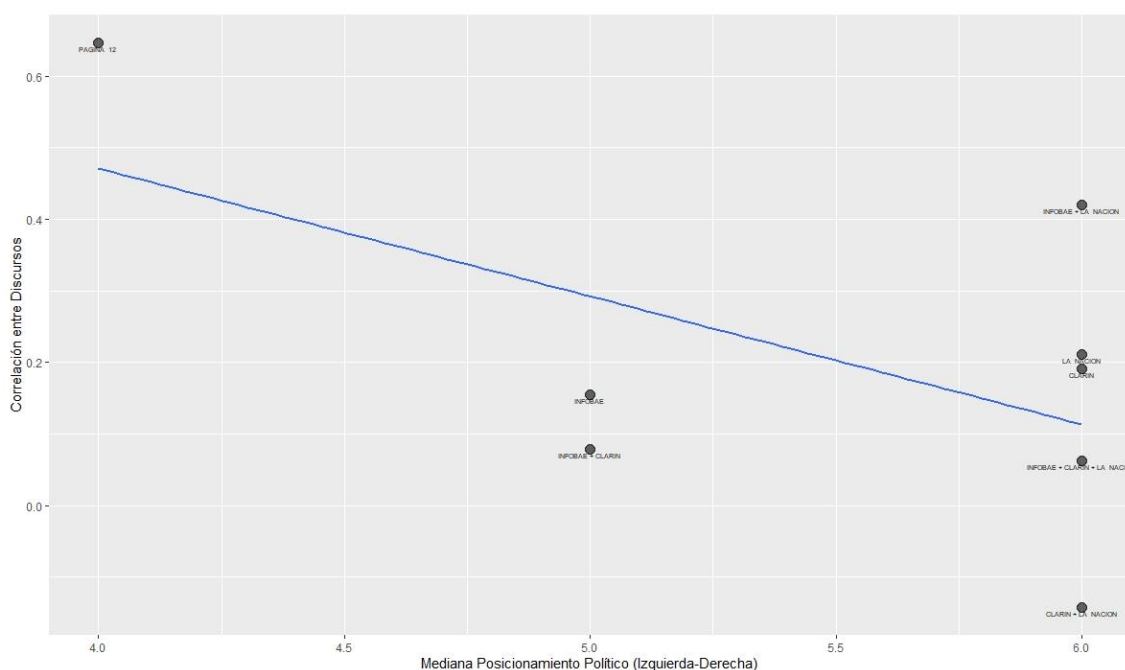
Como resulta evidente, la relación positiva existente entre la valoración de Cristina Fernández de Kirchner y la correlación entre los universos discursivos, reafirma el análisis anteriormente introducido. En este sentido, se evidencia la manera en que la audiencia exclusiva de Página 12 mantiene su distanciamiento respecto a los demás segmentos, reafirmando su postura en contraposición a la postura política extrema opuesta, a saber, la audiencia exclusiva de La Nación, quien mantiene tanto la menor imagen positiva de Cristina Fernández, como la mayor imagen positiva de Mauricio Macri. Por último, cabe destacar la manera en que, tanto la confianza en los medios de comunicación como la correlación entre el discurso propio y el contenido mediático consumido, se ven impregnados por el posicionamiento político declarado.

³⁵ Cada nombre en el gráfico representa los medios de comunicación consumidos de manera online por cada audiencia. Se define como “imagen positiva” a aquella proporción de entrevistados que indicó su imagen sobre Cristina Fernández, en el marco de una escala de 1 a 10, con un puntaje de entre 7 y 10 puntos.

Con el fin de terminar de describir este posicionamiento por parte de las audiencias, cabe introducir la relación existente entre el posicionamiento político por parte de las audiencias, y la correlación hallada entre universos discursivos.

Gráfico 27.

*Relación entre Posicionamiento Político y Correlación entre discursos*³⁶



Fuente: Elaboración Propia

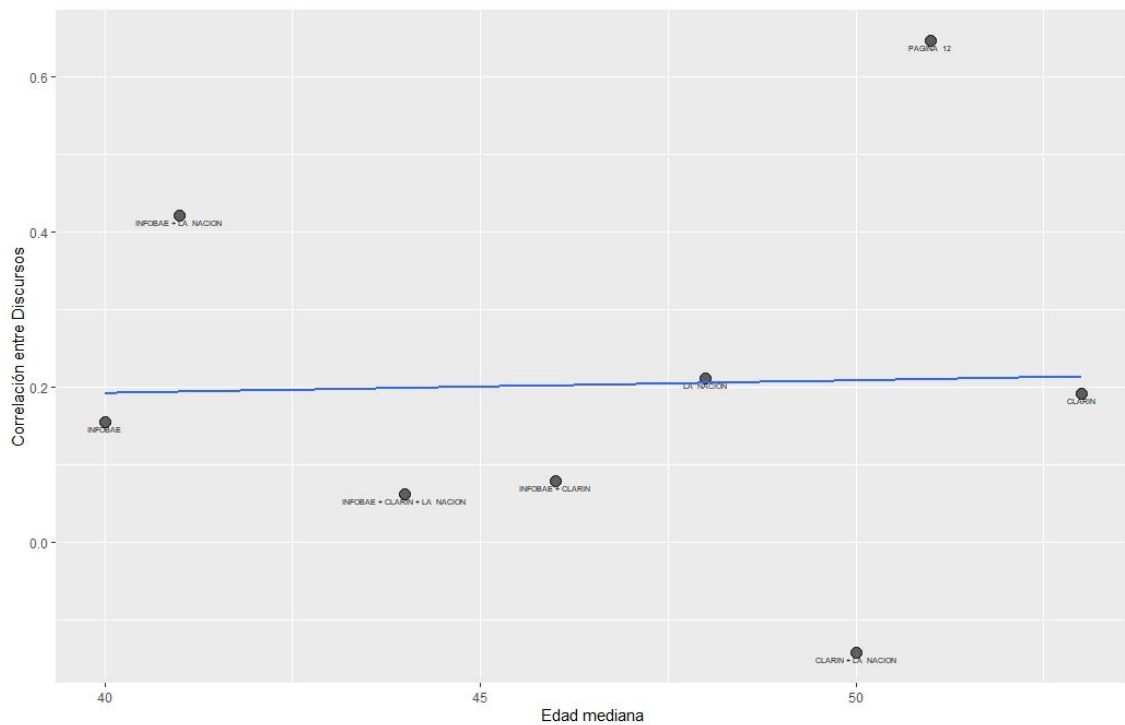
De esta manera, se confirman las observaciones realizadas previamente: la audiencia exclusiva de Página 12 logra diferenciarse notablemente de las restantes, no solo por su alta correlación con el discurso mediático que consume, sino también por sus diferencias ideológicas respecto a los demás segmentos considerados. Esto, por su parte, se corrobora no solo partiendo de la imagen de los dos principales actores políticos del momento, sino también considerando la autodefinición de los entrevistados en materia de posicionamiento político.

³⁶ En este caso, se entiende “Posicionamiento Político” como la mediana correspondiente a cada audiencia, en el marco de la autodefinición política de cada grupo sobre la escala de posicionamiento ideológico señalada en el capítulo previo. Este, junto con otros estadísticos descriptivos, pueden ser encontrados en el Anexo del presente trabajo.

Como última instancia a la hora de comprender el comportamiento y composición de estas audiencias, y la forma y magnitud que adopta su relación con el contenido mediático que consume, resta considerar la manera en que dicho grado de correlación se asocia con variables demográficas. En este sentido, se presenta a continuación el análisis gráfico de dicha asociación con la edad mediana de cada grupo, y su proporción de consumidoras mujeres:

Gráfico 28.

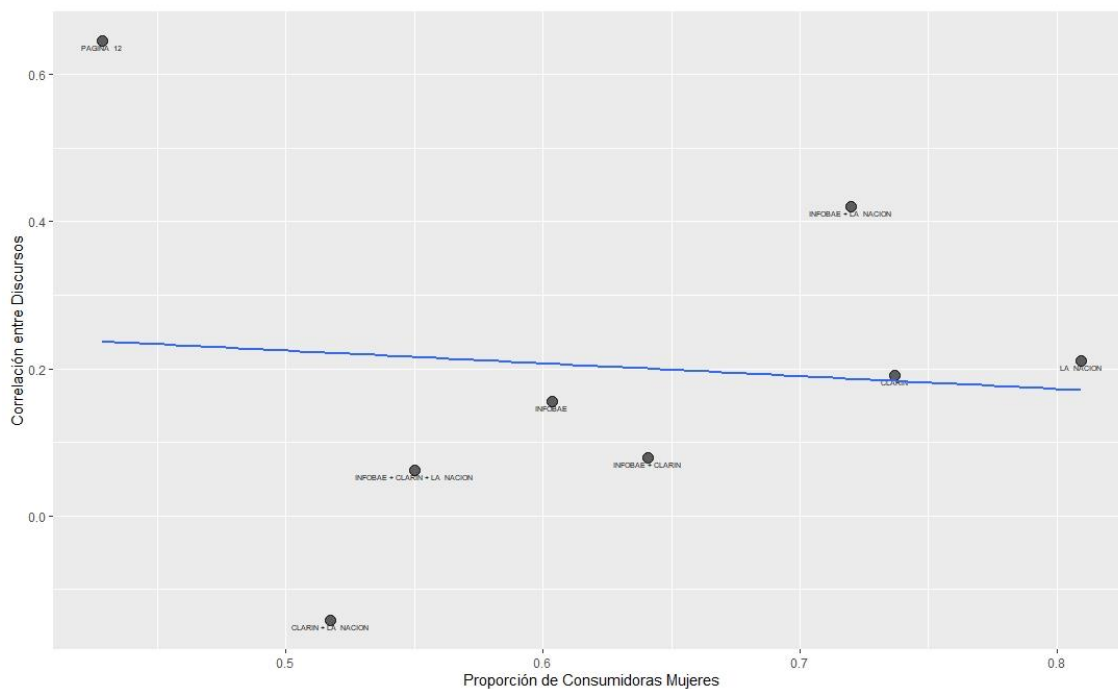
Relación entre Edad Mediana de cada audiencia y Correlación entre discursos



Fuente: Elaboración Propia

Gráfico 29.

Relación entre Proporción de Consumidoras Mujeres en cada audiencia y Correlación entre discursos



Fuente: Elaboración Propia

Tal como puede observarse, no se evidencia una relación significativa entre la correlación del discurso de cada audiencia y el discurso de los medios de comunicación que consume, y las características demográficas consideradas de dichas audiencias. En este sentido, no solo resalta la poca diferenciación de los grupos en estos aspectos -hecho observado en el capítulo previo-, sino también su poca influencia respecto al comportamiento observado de dichas audiencias en relación al contenido mediático que consumen.

Sobre los análisis considerados, en términos generales, se desprende entonces que:

- a. Existe una relación negativa entre la correlación del discurso de cada audiencia y el discurso mediático, y el tamaño relativo de cada una de ellas;
- b. Dicha relación negativa se replica frente a la confianza declarada en los medios de comunicación, donde resalta el hecho que el cuadrante de mayor

- confianza declarada y menos correlación entre universos discursivos se encuentra integrado prácticamente en su totalidad por audiencias mixtas;
- c. Un factor influyente en la confianza en los medios de comunicación resulta ser la cantidad de medios consumida, considerando la relación positiva encontrada entre la cantidad de medios que se consumen y la confianza en los mismos;
 - d. La correlación entre los universos discursivos y el posicionamiento político declarado muestran tener una alta relación, considerando que se trata de una relación donde la audiencia exclusiva de Página 12 tiene a definir la tendencia mediante un fuerte distanciamiento de las restantes audiencias, las cuales, por su parte, tienden a concentrarse sobre la mirada opuesta de dicha audiencia exclusiva. En este sentido, se observan diferencias notables no solo respecto al posicionamiento ideológico declarado por parte de los entrevistados, sino también respecto a su imagen sobre los principales líderes políticos del momento;
 - e. La correlación entre universos discursivos no encuentra relación con características demográficas de los grupos considerados, como lo son su edad mediana y su distribución en función del género de sus integrantes. Esto, por su parte, reconfirma la poca capacidad de diferenciación de dichas variables respecto a las audiencias consideradas, ya vista en el capítulo previo.

Sobre este conjunto de observaciones generales es posible, entonces, trazar las conclusiones del presente trabajo, en relación a los planteos teóricos desarrollados en un comienzo.

CAPÍTULO 6. CONCLUSIONES

A lo largo del presente trabajo, fue estudiado de manera descriptiva el discurso por parte de los medios de comunicación, por un lado, y el discurso de la opinión pública en materia del problema más importante del país.

Tal como pudo corroborarse, los tópicos en los cuales fue desagregado cada uno de dichos discursos no son coincidentes a priori, y a la vez el modelo de lenguaje correspondiente a los tópicos elaborados sobre el discurso de la opinión pública, no constituye un buen clasificador del discurso mediático. Frente a este desafío, cabe retomar algunas distinciones realizadas en el apartado teórico del presente documento.

Como fue mencionado, el trabajo retoma, desde lo interpretativo, los principales conceptos del corpus teórico que componen a los postulados de *agenda setting*, considerando aquí los diferentes momentos de su desarrollo ya enunciados. Dentro de este conjunto de planteos, se establecen cuatro instancias de establecimiento de agenda, por parte de los medios de comunicación, los cuales vale la pena rever. Estos son:

1. En primer lugar, los medios de comunicación resaltan eventos, actividades o personalidades;
2. Luego, un marco común es construido, sobre ciertos aspectos específicos de dicho objeto;
3. Este tema, por su parte, es asociado a símbolos secundarios, ubicándose así en un paisaje político reconocido;
4. Por último, se da la promoción de temas y símbolos por parte de portavoces reconocidos.

Considerando estos pasos diferenciados, cabe mencionar que el relevamiento y análisis de contenido respecto a los medios de comunicación, se da en la primera instancia, teniendo en cuenta eventos y personalidades destacadas por los mismos. Por otro lado, en función del relevamiento sobre la opinión por parte de las audiencias de dichos medios de comunicación, notamos la aparición de estos “paisajes políticos reconocidos”, expresados en los grandes tópicos generales mencionados como problemas, mayormente de carácter estructural, ligados al territorio nacional.

En este sentido, y al evidenciar estos planos diferenciados de expresión entre las poblaciones estudiadas, fue planteado, desde lo metodológico, abordar la relación de dichos discursos sobre la base de su unidad mínima: las palabras que los constituyen. Aquí, vale la pena resaltar los trabajos metodológicos de Grimmer *et al.* (*Op. Cit.*), quienes logran evidenciar que la palabra integrante de un documento -sea éste una frase, una oración o un *tweet*-, logra capturar la mayor parte del sentido que encierra dicho marco.

Partiendo de dichas consideraciones metodológicas, fue posible establecer el sentido e intensidad de la relación entre ambos universos discursivos analizados. Por otro lado, a partir de las diferentes variables relevadas de manera complementaria al problema más importante mediante la metodología de encuesta, fue posible realizar un análisis de mayor granularidad sobre dicha relación.

En primer lugar, fue posible evidenciar aquellas audiencias frente a las cuales nos encontrábamos, a partir del consumo declarado de medios de comunicación de manera online. Sobre estas audiencias, fue posible corroborar la manera en que la correlación inicialmente vista entre ambos universos, se amplificaba o relajaba. En segunda instancia, gracias a variables de tipo declarativas relevadas mediante el instrumento de medición considerado, también fue posible entender en profundidad la relación entre la correlación discursiva hallada y las características que componen a cada audiencia. Vale la pena, en este punto, realizar un repaso de los principales hallazgos del trabajo:

- a) Considerando únicamente los medios de comunicación, fueron hallados 16 tópicos que lograban sintetizar el contenido de su discurso. Si bien en algunos casos relacionados, los temas mostraban particular validez semántica, y por ende una diferenciación notable en relación al resto;
- b) En función de estos tópicos hallados, fue evidenciada su consistencia en el tiempo y la falta de relación entre el tópico en cuestión y el medio de comunicación seleccionado -esto último, en términos generales, aunque notamos que, frente a algunos temas específicos, los medios de comunicación se diferenciaban entre sí, parcialmente-, observaciones que permitieron corroborar los principios de *Focusing* y *Persistence* introducidos en el apartado teórico, considerando a su vez la coocurrencia temporal

hallada en base al análisis del comportamiento de publicación de dichos medios de comunicación;

- c) En materia de la opinión pública relevada respecto al problema más importante, fue evidenciado un discurso mayormente ligado a problemas estructurales en relación al discurso de los medios de comunicación, sobre el cual se pudieron definir 6 tópicos con elevada validez semántica. Cabe destacar, en este sentido, la menor segmentación de las problemáticas por parte de la opinión pública, que tiende a tener un discurso considerablemente más laxo y general que el de los medios;
- d) Sobre el consumo declarado de medios de comunicación online, fueron desarrolladas audiencias diferenciadas entre los entrevistados, sobre las cuales fueron consideradas únicamente aquellas que presentaban un volumen de respuestas tal como para poder realizar los cálculos estadísticos pertinentes. Sobre estas audiencias definidas, fueron analizadas las diferencias notadas respecto al tópico de conversación que presentaba mayor predominancia, a la vez que aquellas diferencias respecto a otras variables relevadas, como su posicionamiento político, su confianza en los medios de comunicación, su dinámica de consumo de dichos medios, su género y su edad;
- e) A partir de considerar la unidad mínima de dichos discursos como unidad de análisis, y teniendo en cuenta la información anteriormente introducida como meta-datos que brindaban una profundidad y riqueza diferencial a las relaciones encontradas, fue analizada la correlación entre ambos universos discursivos. En este sentido, cabe mencionar que, entre los medios de comunicación en general y la opinión pública, fue hallada una relación débil pero significativa que, si bien no encontraba grandes diferencias al considerar el discurso de la opinión pública frente al de cada medio de comunicación en particular, adoptaba una magnitud particularmente relevante entre los conceptos coincidentes en el discurso de la opinión pública y el medio de comunicación Página 12;
- f) En base a estos primeros resultados, fue considerada cada una de las audiencias halladas, y la correlación entre su discurso y el de los medios de comunicación que consume fue vista a través de los “lentes” que brindaban los meta-datos relevados mediante encuesta. En este sentido, pudo

identificarse la particular diferenciación de la audiencia exclusiva del medio de comunicación anteriormente mencionado -Página 12-, no solo en materia de la consonancia de su discurso con el discurso mediático, sino también en materia de posicionamiento ideológico en términos partidarios, como también respecto a los medios de comunicación. De esta manera, se evidencia la particular importancia de la recolección de meta-datos mediante otros instrumentos de medición complementarios, por ejemplo, a la técnica de *web scrapping* utilizado;

- g) Sobre la base de las relaciones encontradas entre el comportamiento de las audiencias y la correlación discursiva, se desprenden algunas relaciones generales. A saber:
- a. La correlación entre universos discursivos se relaciona negativamente con el tamaño relativo de las audiencias. Esto es, cuando una audiencia adopta un carácter “de nicho”, la concordancia entre su discurso y el discurso mediático es mayor;
 - b. Dicha correlación se relaciona negativamente con la confianza en los medios de comunicación, hallando que cuanto mayor confianza se declara en éstos, menos se condice el discurso de dicha audiencia con el mediático;
 - c. Esto, a su vez, se relaciona con la cantidad de medios de comunicación consumida, siendo que a medida que incrementa la cantidad de medios que consume cada audiencia, mayor es la confianza que depositan en ellos;
 - d. La correlación entre universos discursivos encuentra una consistente relación con el posicionamiento político de cada audiencia. En este sentido, se identifican tres patrones: la correlación mantiene una relación negativa tanto con la imagen positiva sobre Mauricio Macri como con el posicionamiento ideológico hacia una postura de derecha, mientras que mantiene una relación positiva con la imagen positiva de Cristina Fernández de Kirchner;
 - e. Por último, no se halla una relación considerable entre dicha correlación y características sociodemográficas de las diferentes audiencias, como son la edad mediana o la distribución por género.

En este punto, cabe realizar un comentario final acerca de la vinculación metodológica que tuvo lugar en el presente trabajo, y que responde al objetivo del presente documento.

Considerando las diferentes audiencias conformadas, la distribución de cada tópico hallado al interior de cada una, y su comportamiento respecto al contenido compartido por los medios de comunicación en función de características relevadas mediante encuesta, fue posible notar la aparición del efecto de “cámara de eco” o la conformación de “burbujas informativas” mencionadas en el apartado teórico.

En este sentido, la evidencia sobre la conformación de dichos grupos y sobre la manera en que sus características y opiniones tiñen la relación que mantienen con el contenido mediático, únicamente era posible de observar mediante una metodología tradicional como la encuesta. La encuesta, en este caso, permitió unir meta-datos propios de una dimensión “declarativa” por parte de los usuarios de los medios de comunicación, con los hallazgos propios de herramientas analíticas provenientes del ámbito de la Ciencia de Datos; dichas herramientas, por su parte, permitieron realizar un análisis asertivo de un gran volumen de datos, logrando obtener tanto una mirada general, como medidas estadísticas particulares para cada uno de los grupos considerados.

De esta manera, además de la contribución particular del presente trabajo a la tradición local de estudios sobre la relación entre los medios de comunicación y la opinión pública, se destaca también su contribución en términos del dispositivo metodológico utilizado: teniendo en cuenta que el principal objetivo de este trabajo se basaba en vincular técnicas de recolección de datos y métodos analíticos propios del universo tradicional de la investigación social aplicada a la opinión pública con técnicas de recolección y análisis característicos del universo de las Ciencias de Datos, se considera que el principal aporte del presente documento se expresa en mostrar la potencialidad de dicha articulación metodológica, pasible de ser implementada en diversos ámbitos de la investigación social aplicada.

Sobre estos hallazgos, entonces, se plantea la necesidad de seguir fortaleciendo los vínculos metodológicos y analíticos aquí presentados, con el fin de poder

seguir evidenciando los intrincados vínculos entre lo que las personas dicen, lo que las personas sienten, y su comportamiento, particularmente observable en plataformas digitales.

BIBLIOGRAFÍA

- Aruguete, N. (2017a). The agenda setting hypothesis in the new media environment. *Revista Comunicación y Sociedad*, Núm. 28, pp. 35-58.
- Aruguete, N. (2017b). Agenda building. Revisión de la literatura sobre el proceso de construcción de agenda mediática. *Revista Signo y Pensamiento* 70, Vol. XXXVI, pp. 36-52.
- Barberá, P. (2017). Training Session: Topics Models. Networked Democracy Lab, University of Southern California. Recuperado de: <https://github.com/netdem-USC/training>
- Barberá, P. (2018). Quantitative Text Analysis. Applications to Social Media Research. London School of Economics. Recuperado de: <http://pablobarbera.com/text-analysis-vienna>
- Barberá, P. (2019). Computational Social Science. School of International Relations, University of Southern California. Recuperado de: <http://pablobarbera.com/POIR613/index.html>
- Benoit, K (2018). Quantitative text analysis: overview and fundamentals. Recuperado de: <https://kenbenoit.net/quantitative-text-analysis-tcd-2018>
- Blalock, H.M. (1960). *Social Statistics*. Londres: McGraw-Hill Book Company.
- Calvo, E. (2015). *Anatomía política de Twitter en Argentina*. Ciudad Autónoma de Buenos Aires: Capital Intelectual.
- Calvo, E. y Arguete, N. (2018). #Tarifazo. Medios tradicionales y fusión de agendas en redes sociales. *Revista Inmediaciones de la Comunicación*, Vol. 13, pp. 189-213.
- Canel, M., Llamas, J. y Rey, F. (1996). El primer nivel del efecto agenda setting en la información local: los “problemas más importantes” de la ciudad de Pamplona. Recuperado de: <https://scholar.google.com.ar/citations?user=FurzDEwAAAAJ&hl=es#>
- Casermeyro De Pereson, A. (2003). *Los medios en las elecciones: la agenda setting en la Ciudad de Buenos Aires*. Buenos Aires: Educa.
- Dominguez, N. (2015). Política 2.0: La comunicación política mediada por Twitter. (Tesis de Doctorado). Recuperado de: <http://sedici.unlp.edu.ar/handle/10915/46845>
- García Ferrando, M. (1989). *Socioestadística. Introducción a la estadística en sociología*. Madrid: Alianza Editorial.
- González, G. F. (Coord.) (2017). *Comunicación política: periodistas, políticos y la opinión pública: definiciones, conceptos e investigación de campo*. La Plata:

Universidad Nacional de La Plata. Facultad de Periodismo y Comunicación Social.

Grimmer, J. y Stewart, B. (2013). Text as Data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* (Summer 2013) 21 (3): pp. 267-297.

López-Escobar, E., Llamas, J. y Rey, F. (1996). La agenda entre los medios: primer y segundo nivel. Recuperado de: <https://scholar.google.com.ar/citations?user=FurzDEwAAAAJ&hl=es#>

López-Escobar, E., McCombs, M. y Rey, F. (1996b). La imagen de los candidatos: el segundo nivel de la agenda-setting. Recuperado de: <https://scholar.google.com.ar/citations?user=FurzDEwAAAAJ&hl=es#>

Morales Vallejo, P. (2012). Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos? Madrid: Facultad de Humanidades, Universidad Pontificia Comillas.

Silge, J. y Robinson, D. (2020). Text Mining with R. Estados Unidos: O'Reilly. Recuperado de: <https://www.tidytextmining.com/>

Slimovich, A. y Fraticelli, D. (2016). Anatomía política de Twitter. Entrevista a Ernesto Calvo. Buenos Aires: *Revista Letra. Imagen. Sonido (L.I.S.)*. Año VIII, Núm 15., Primer Semestre 2016, pp. 191-198.

Rubio García, R. (2014). Twitter y la teoría de Agenda Setting: mensajes de la opinión pública digital. *Revista Estudios sobre el Mensaje Periodístico*, Vol. 20, pp. 249-264.

UNIVERSITY OF OXFORD, REUTERS INSTITUTE (2019). Digital News Report 2019. Recuperado de: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-06/DNR_2019_FINAL_0.pdf

WE ARE SOCIAL (2015). Digital in 2015: Argentina. Recuperado de: <https://datareportal.com/reports/digital-2015-argentina>

WE ARE SOCIAL (2016). Digital in 2016: Argentina. Recuperado de: <https://datareportal.com/reports/digital-2016-argentina>

WE ARE SOCIAL (2017). Digital in 2017: Argentina. Recuperado de: <https://datareportal.com/reports/digital-2017-argentina>

WE ARE SOCIAL (2018). Digital in 2018: Argentina. Recuperado de: <https://datareportal.com/reports/digital-2018-argentina>

WE ARE SOCIAL (2019). Digital in 2019: Argentina. Recuperado de: <https://datareportal.com/reports/digital-2019-argentina>

ZHANG, Z. (2018). Text Mining for Social and Behavioral Research Using R. A case study on Teaching Evaluation. Recuperado de: <https://books.psychstat.org/textmining/index.html>

ANEXO

Tabla 16.

Datos estadísticos correspondiente a cada audiencia definida.

Audiencia	Tamaño relativo	R de Pearson	Obs	t	df	p-value
INFOBAE + CLARIN + LA NACION	0,146	0,062	70	0,51187	68	0,6104
INFOBAE	0,129	0,155	152	1,9221	150	0,05649
INFOBAE + CLARIN	0,095	0,079	41	0,49437	39	0,6238
CLARIN	0,093	0,191	72	1,6263	70	0,1084
CLARIN + LA_NACION	0,071	-0,142	25	- 0,68959	23	0,4974
INFOBAE + LA_NACION	0,061	0,42	34	2,6185	32	0,01338
PAGINA_12	0,051	0,646	66	6,7662	64	4,776E-09
LA_NACION	0,051	0,211	72	1,803	70	0,07569
INFOBAE + CLARIN + PAGINA 12 + LA_NACION + PERFIL	0,044	SD	SD	SD	SD	SD
INFOBAE + CLARIN + LA NACION + PERFIL	0,041	-0,1211394	11	- 0,36611	9	0,7227
INFOBAE + PAGINA 12	0,024	0,4088475	12	1,4167	10	0,187
INFOBAE + CLARIN + PAGINA 12	0,017	SD	SD	SD	SD	SD
INFOBAE + PAGINA 12 + PERFIL	0,015	0,9822573	7	11,712	5	7,977E-05
INFOBAE + CLARIN + PAGINA 12 + LA NACION	0,015	SD	SD	SD	SD	SD
INFOBAE + PAGINA 12 + LA NACION	0,010	SD	SD	SD	SD	SD
CLARIN + PAGINA 12 + LA NACION	0,010	SD	SD	SD	SD	SD
INFOBAE + PERFIL	0,007	SD	SD	SD	SD	SD
PAGINA 12 + LA_NACION	0,005	SD	SD	SD	SD	SD
INFOBAE + CLARIN + PERFIL	0,005	SD	SD	SD	SD	SD
CLARIN + PAGINA 12 + LA NACION + PERFIL	0,005	SD	SD	SD	SD	SD
PERFIL	0,002	SD	SD	SD	SD	SD
CLARIN + PAGINA 12	0,002	SD	SD	SD	SD	SD
PAGINA 12 + PERFIL	0,002	SD	SD	SD	SD	SD
INFOBAE + LA NACION + PERFIL	0,002	SD	SD	SD	SD	SD
PAGINA 12 + LA NACION + PERFIL	0,002	SD	SD	SD	SD	SD

Fuente: Elaboración Propia

Tabla 17.

Datos sobre posicionamiento político, correspondiente a cada audiencia definida sobre las ocho audiencias principales.

Audiencia	Promedio de Imagen Macri	Proporción Aceptación Macri	Promedio de Imagen CFK	Proporción Aceptación CFK	Mediana del Posicionamiento Político	Promedio del Posicionamiento Político	Desvío Estándar Posicionamiento Político
INFOBAE + CLARIN + LA_NACION	5,695	0,433	2,431	0,1	6	6,5	1,722
INFOBAE	4,520	0,396	3,547	0,226	5	5,32	2,368
INFOBAE + CLARIN	4,816	0,256	2,692	0,102	5	6,05	2,064
CLARIN	5,371	0,394	2,568	0,105	6	6,10	2,252
CLARIN + LA NACION	6,107	0,517	1,931	0,034	6	6,27	1,771
INFOBAE + LA NACION	6,080	0,56	2,200	0,04	6	6,28	1,671
PAGINA 12	1,000	0	8,300	0,809	4	4,28	1,419
LA_NACION	6,524	0,571	1,900	0	6	6,71	2,053

Fuente: Elaboración Propia

Tabla 18.

Datos sobre consumo e imagen de medios, correspondiente a cada audiencia definida sobre las ocho audiencias principales.

Audiencia	Promedio de Confianza en los Medios	Proporción Confianza en los Medios	Proporción de Consumo Diario
INFOBAE + CLARIN + LA NACION	6,45	56,67%	35,48%
INFOBAE	5,43	37,74%	50,00%
INFOBAE + CLARIN	6,03	46,15%	45,28%
CLARIN	6,66	55,26%	43,59%
CLARIN + LA NACION	7,07	65,52%	44,74%
INFOBAE + LA NACION	7,04	72,00%	65,52%
PAGINA 12	4,43	19,05%	52,00%
LA NACION	6,14	33,33%	47,62%

Fuente: Elaboración Propia

Tabla 19.

Datos demográficos, correspondiente a cada audiencia definida sobre las ocho audiencias principales.

Audiencia	Edad Mediana	Edad Promedio	Desvío Estándar Edad	Proporción Mujeres
INFOBAE + CLARIN + LA NACION	44	43,867	12,07107	55,00%
INFOBAE	40	39,811	11,225069	60,38%
INFOBAE + CLARIN	46	45,128	10,039867	64,10%
CLARIN	53	47,921	11,627976	73,68%
CLARIN + LA_NACION	50	46,966	10,675348	51,72%
INFOBAE + LA NACION	41	45,760	10,304853	72,00%
PAGINA_12	51	48,667	10,135745	42,86%
LA_NACION	48	46,667	12,827055	80,95%

Fuente: Elaboración Propia