



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
ESCUELA DE INGENIERÍA

**EVALUACIÓN DE INTERFACES DE  
VISUALIZACIÓN DE MODELO DE  
TÓPICOS DE CONSULTAS  
CIUDADANAS: UN ESTUDIO DE CASO  
DEL PROCESO CONSTITUYENTE  
PARTICIPATIVO DEL 2016 EN CHILE**

**IVANIA YOVANOVIC LÓPEZ**

Tesis para optar al grado de  
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:  
**CONSTANZA MIRANDA**

Santiago de Chile, (enero, 2021)

© 2021, Ivania Yovanovic



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
ESCUELA DE INGENIERÍA

# **EVALUACIÓN DE INTERFACES DE VISUALIZACIÓN DE MODELO DE TÓPICOS DE CONSULTAS CIUDADANAS: UN ESTUDIO DE CASO DEL PROCESO CONSTITUYENTE PARTICIPATIVO DEL 2016 EN CHILE**

**IVANIA YOVANOVIC LÓPEZ**

Tesis presentada a la Comisión integrada por los profesores:

**CONSTANZA MIRANDA**

**DENIS PARRA**

**MARÍA PAZ HERMOSILLA**

**PABLO PASTÉN**

Para completar las exigencias del grado de  
Magíster en Ciencias de la Ingeniería

Santiago de Chile, (enero, 2021)

(A Jessica, Talía, Juan Pablo, familia,  
amigas y amigos, que me apoyaron  
mucho.)

## **AGRADECIMIENTOS**

Mis más sinceros agradecimientos a Constanza Miranda e Iñaki Goñi quienes me han guiado en este camino hasta aquí, tanto personal, como académicamente.

Agradezco también a los miembros del Comité, María Paz Hermosilla, Denis Parra y Pablo Pastén, por su buena disposición y orientación en mi trabajo de investigación.

## ÍNDICE GENERAL

Pág.

|                                                                               |     |
|-------------------------------------------------------------------------------|-----|
| DEDICATORIA.....                                                              | ii  |
| AGRADECIMIENTOS .....                                                         | iii |
| INDICE DE TABLAS .....                                                        | vi  |
| INDICE DE FIGURAS .....                                                       | vii |
| RESUMEN.....                                                                  | ix  |
| ABSTRACT .....                                                                | x   |
| 1. INTRODUCCIÓN .....                                                         | 1   |
| 1.1 Caso de estudio: una Constitución para Chile.....                         | 3   |
| 2. MARCO TEÓRICO .....                                                        | 7   |
| 2.1 Participación ciudadana.....                                              | 7   |
| 2.2 Modelamiento de Tópicos .....                                             | 9   |
| 2.3 Visualización de información.....                                         | 12  |
| 2.4 Usabilidad.....                                                           | 15  |
| 3. TRABAJO RELACIONADO .....                                                  | 17  |
| 3.1 Modelamiento de tópicos en opiniones .....                                | 17  |
| 3.2 Interfaces de visualización de tópicos .....                              | 19  |
| 3.3 Trabajos vinculados a Consulta Ciudadana Proceso Constituyente 2016 ..... | 22  |
| 4. OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN .....                               | 25  |
| 5. METODOLOGÍA .....                                                          | 26  |
| 5.1 Selección y descripción de los datos .....                                | 26  |
| 5.2 Pre-procesamiento y limpieza .....                                        | 27  |
| 5.3 Transformación .....                                                      | 28  |
| 5.3.1 Transformación de [normalización].....                                  | 28  |
| 5.3.2 Transformación de [region] .....                                        | 29  |
| 5.3.3 Transformación de [comuna].....                                         | 30  |

|       |                                                               |    |
|-------|---------------------------------------------------------------|----|
| 5.4   | Modelo de tópicos .....                                       | 30 |
| 5.5   | Etiquetas de tópicos.....                                     | 33 |
| 5.6   | Estudio de usuarios.....                                      | 35 |
| 6.    | DISEÑO E IMPLEMENTACIÓN .....                                 | 37 |
| 6.1   | Situación o dominio .....                                     | 37 |
| 6.2   | Tareas y abstracción de datos.....                            | 38 |
| 6.3   | Codificación visual e idioma de interacción .....             | 39 |
| 6.4   | Interfaces de usuario.....                                    | 44 |
| 7.    | ESTUDIO DE USUARIOS .....                                     | 46 |
| 7.1   | Motivación y objetivos.....                                   | 46 |
| 7.2   | Tareas de usuario.....                                        | 46 |
| 7.2.1 | Validación de coherencia semántica del modelo .....           | 46 |
| 7.2.2 | Evaluación de usabilidad de interfaces de visualización ..... | 49 |
| 7.3   | Métricas.....                                                 | 52 |
| 7.4   | Participantes .....                                           | 53 |
| 7.5   | Diseño experimental.....                                      | 55 |
| 7.6   | Procedimiento.....                                            | 57 |
| 8.    | RESULTADOS Y ANÁLISIS .....                                   | 60 |
| 8.1   | Validación de coherencia semántica .....                      | 60 |
| 8.2   | Evaluación de usabilidad.....                                 | 63 |
| 9.    | DISCUSIÓN .....                                               | 70 |
| 10.   | CONCLUSIONES.....                                             | 74 |
|       | BIBLIOGRAFÍA.....                                             | 76 |

## INDICE DE TABLAS

|                                                                                                                                   | Pág. |
|-----------------------------------------------------------------------------------------------------------------------------------|------|
| Tabla 5-1: Extracto de algunos elementos seleccionados de la base de datos .....                                                  | 27   |
| Tabla 5-2: Palabras por tópico .....                                                                                              | 32   |
| Tabla 5-3: Etiqueta y palabras por tópico .....                                                                                   | 34   |
| Tabla 7-1: Pregunta de investigación 1 y tareas de usuario .....                                                                  | 47   |
| Tabla 7-2: Pregunta de investigación 2 y tareas de usuario .....                                                                  | 49   |
| Tabla 7-3: Cuadrado latino 2x2.....                                                                                               | 56   |
| Tabla 7-4: Preguntas por tarea y versión.....                                                                                     | 56   |
| Tabla 8-1: Resultados Precisión del modelo – Palabra intrusa .....                                                                | 61   |
| Tabla 8-2: Resultados Topic log odds - Intrusión de tópicos.....                                                                  | 63   |
| Tabla 8-3: Medias (M) y Desviación estándar (SD) de tasa de éxito, tiempo de<br>completación y satisfacción percibida (N=34)..... | 64   |

## INDICE DE FIGURAS

|                                                                                           | Pág. |
|-------------------------------------------------------------------------------------------|------|
| Figura 1-1: Temas y preguntas de la consulta.....                                         | 4    |
| Figura 1-2: Resumen de metodología del proceso .....                                      | 5    |
| Figura 1-3: Ejemplo de opinión original versus opinión normalizada.....                   | 6    |
| Figura 2-1: Niveles de participación ciudadana.....                                       | 8    |
| Figura 2-2: Representación gráfica de topic modeling .....                                | 10   |
| Figura 2-3: Interpretación matricial de topic modeling .....                              | 11   |
| Figura 2-4: Cuatro razones de por qué visualizar .....                                    | 12   |
| Figura 2-5: Preguntas fundamentales framework .....                                       | 13   |
| Figura 2-6: Diagrama framework.....                                                       | 14   |
| Figura 3-1: Sistema con mapa de árbol.....                                                | 21   |
| Figura 3-2: Sistema con mapa coroplético.....                                             | 22   |
| Figura 3-3: Plataforma de trazabilidad.....                                               | 23   |
| Figura 3-4: Plataforma Constitución Abierta.....                                          | 24   |
| Figura 5-1: Puntaje de coherencia del modelo según número de tópicos .....                | 32   |
| Figura 6-1: Mapa de árbol.....                                                            | 40   |
| Figura 6-2: Gráfico de barras apiladas .....                                              | 41   |
| Figura 6-3: Mapa coroplético.....                                                         | 42   |
| Figura 6-4: Mapa de densidad de puntos .....                                              | 43   |
| Figura 6-5: Interfaz A - Mapa de árbol .....                                              | 44   |
| Figura 6-6: Interfaz B - Barras apiladas.....                                             | 45   |
| Figura 7-1: Ejemplo Palabra Intrusa .....                                                 | 47   |
| Figura 7-2: Ejemplo Tópico Intruso.....                                                   | 48   |
| Figura 7-3: Ejemplo de tarea de categorización de tópicos .....                           | 50   |
| Figura 7-4: Ejemplo de tarea de comparación .....                                         | 51   |
| Figura 7-5: Parámetros utilizados para Power Analysis de GPower .....                     | 54   |
| Figura 7-6: Estructura cuestionario.....                                                  | 58   |
| Figura 8-1: Gráfico de respuestas correctas según dificultad - intrusión de palabras..... | 60   |

|                                                                                                                      |    |
|----------------------------------------------------------------------------------------------------------------------|----|
| Figura 8-2: Gráfico de respuestas correctas según dificultad - intrusión de tópicos .....                            | 62 |
| Figura 8-3: Medias de tasa de éxito de cada tipo de interfaz para tareas de categorización y comparación .....       | 66 |
| Figura 8-4: Tiempos medios de completación de cada tipo de interfaz para tareas de categorización y comparación..... | 67 |
| Figura 8-5: Satisfacción media percibida de cada tipo de interfaz para tareas de categorización y comparación.....   | 69 |

## RESUMEN

La participación ciudadana en la formulación de políticas requiere que participantes y políticos comprendan y puedan utilizar la información producida durante los ejercicios de consulta pública. Por ello, se han desarrollado diferentes estrategias de visualización para representar la opinión pública sistematizada. En este sentido, existe la necesidad de aportar evidencia sobre la usabilidad de estas estrategias de visualización. Entre las estrategias de visualización, el mapa de árbol es una de las más utilizadas para difundir datos y devolverlos a la población. Sin embargo, según la jerarquía de precisión de visualización hay motivos para creer que esta visualización no es la más adecuada para representar datos cuantitativos y permitir análisis en comparación con otras visualizaciones, como los gráficos de barras apiladas. El propósito de este estudio es comprender cómo los mapas de árbol y los gráficos de barras apiladas difieren en términos de efectividad, eficiencia y satisfacción en tareas de análisis como categorización y comparación. Se utilizó un diseño experimental para examinar el rendimiento de usuarios en una prueba de usabilidad basada en tareas. 34 participantes interactuaron con visualizaciones de tópicos basadas en los datos de la consulta del Proceso Constituyente del año 2016 en Chile. El ANOVA mostró que los gráficos de barras apiladas funcionan significativamente mejor para las tareas de comparación que el mapa de árbol, pero no hay diferencias significativas con respecto a las tareas de categorización. Estos hallazgos esperan determinar orientaciones de diseño para modelar y visualizar los resultados de consultas públicas futuras.

Palabras Claves: visualización de tópicos; usabilidad; modelamiento de tópicos; interacción humano-computador; consulta pública; participación ciudadana

## **ABSTRACT**

Citizen participation in policy-making requires that participants and policy-makers understand and be able to use the information produced during public consultation exercises. For this reason, different visualization strategies have been developed to represent the systematized public opinion. In this sense, there is a need to provide evidence on the usability of these visualization strategies. Among the visualization strategies, tree map is one of the most frequently used to disseminate data and to give it back to the population. However, based on the visualization precision hierarchy, there is reason to believe that this visualization is not the most suitable for representing quantitative data and allowing analysis compared to other visualizations, such as stacked bar charts. The purpose of this study is to understand how tree maps and stacked bar charts differ in terms of effectiveness, efficiency, and satisfaction in analysis tasks such as categorization and comparison. An experimental design was used to examine user performance based on a task-based usability test. 34 participants interacted with data visualizations from an open 2016 participatory constitution-making process in Chile. The ANOVA showed that stacked bar charts work significantly better for comparison tasks than the tree map, but there are no significant differences in regard to categorization tasks. These findings hope to determine design guidelines to model and visualize the results of future public consultations.

Keywords: topic visualization; usability; topic modeling; human-computer interaction; public consultation; citizen participation

## **1. INTRODUCCIÓN**

La participación ciudadana contempla el involucramiento de la ciudadanía en la toma de decisiones públicas (Baum, 2015). Arnstein (1969) formuló una escalera conceptual con ocho niveles de participación según el grado de involucramiento de la ciudadanía. Los niveles van desde la manipulación, su forma más débil, hasta el control ciudadano directo, su nivel más completo. Hoy en día, existe una tendencia entre los académicos y profesionales de las ciencias políticas a exigir una participación ciudadana más activa, directa y sustancial en la formulación de políticas (Dryzek, 2002; Floridia, 2018).

La participación pública en la formulación de políticas abarca desde la elaboración de presupuestos del gobierno local hasta la elaboración de constituciones (Bobbio, 2019). La participación ciudadana en procesos de cambio constitucional es una tendencia creciente (Pacheco i Canals, 2019). Más del 70 por ciento de las nuevas constituciones promulgadas en las últimas décadas ha involucrado a la ciudadanía (Maboudi y Nadi, 2016). Las formas de participación ciudadana van desde la elección de los redactores de la constitución hasta las consultas públicas sobre el contenido de las nuevas constituciones (Negretto, 2017).

Aunque la participación ciudadana en la elaboración de una constitución es muy deseable para lograr la legitimidad, existen desafíos persistentes en el proceso de gestión de las instituciones, especialmente en lo que respecta a la inclusión, la transparencia y la rendición de cuentas (Gluck y Brandt, 2015). Algunos países han comenzado a

incorporar nuevas tecnologías para facilitar ese proceso (Gluck y Ballou, 2014). Más allá del tema de la elaboración de constituciones, varios países han utilizado técnicas de Inteligencia Artificial (IA) para organizar y clasificar el contenido de sus consultas públicas (Noveck, 2018).

El modelamiento de tópicos es una técnica de aprendizaje de máquina que ha resultado útil para este propósito (Hsiao, Lin, Tang, Narayanan y Sarahe, 2018). Esta técnica identifica y organiza los temas subyacentes en las opiniones de forma automática (Blei, Ng y Jordan, 2003). Sin embargo, los resultados de este modelo pueden ser difíciles de entender para los ciudadanos (Cai, Sun y Sha, 2018). Es por esto que muchas consultas ciudadanas acompañan el análisis de datos con herramientas de visualización para aumentar la usabilidad de la información recopilada (Roque de Oliveira y Partidário, 2020).

Se han desarrollado diferentes interfaces de visualización para representar los modelos de tópicos (Karpovich, Smirnov, Teslya y Grigorev, 2017; Liu et al., 2019). Bajo un diseño centrado en el usuario, las interfaces han sido construidas en base a tres lineamientos (Yang, Yao y Qu, 2017). Una interfaz eficiente y útil ha de tener: (i) una visión general sobre la organización de los temas, (ii) una visión con perspectiva geográfica para datos geolocalizados y (iii) una visión particular que permita al usuario revisar el compilado de las opiniones o un extracto de ellos.

El mapa de árbol es una de las visualizaciones más utilizadas para presentar una visión general de la organización de los tópicos y permitir un análisis categórico y cuantitativo (Sukhija et al., 2019). El mayor problema con los mapas de árboles es que el uso del área de un cuadrado para codificar los datos dificulta la realización de comparaciones cuantitativas precisas, frente al uso de la longitud de una barra (Li et al., 2019). En ese sentido, otras estrategias de visualización, como los gráficos de barras apiladas, pueden ofrecer una mejor manera de codificar datos para realizar comparaciones precisas.

Este estudio busca evaluar y comparar características de usabilidad de dos tipos de interfaces de visualización de tópicos de participación ciudadana. Se compara la usabilidad del mapa de árbol y del gráfico de barras apiladas, dentro de una interfaz complementada por un mapa coroplético para información geográfica y un mapa de densidad de puntos para información puntual. Para el modelo de tópicos se utiliza como caso de estudio la Consulta Ciudadana del Proceso Constituyente 2016 en Chile. La instancia reunió a más de 200.000 ciudadanos en la discusión sobre el contenido de una eventual nueva constitución (OCDE, 2017).

### **1.1 Caso de estudio: una Constitución para Chile**

En 2016, el Gobierno de Chile convocó a un proceso constituyente para poder discutir y deliberar sobre los contenidos de una eventual nueva Constitución (Gobierno de Chile, 2015). El proceso propuesto estaba compuesto de tres

momentos continuos, estos son: fase encuentro, fase deliberación y fase soberanía (Becker Castellaro, 2018).

La **fase encuentro** o etapa participativa, se llevó a cabo entre el 23 de abril y el 6 de agosto de 2016. La metodología desarrollada por el Gobierno estableció cuatro ejes de debate: (1) valores y principios, (2) derechos, (3) deberes y responsabilidades, e (4) instituciones del Estado. Las preguntas orientadoras se pueden ver en la Figura 1-1.

|    |                             |                                                                                                         |
|----|-----------------------------|---------------------------------------------------------------------------------------------------------|
| 01 | Valores y principios        | ¿Cuáles deberían ser los principales <b>valores y principios</b> que inspiran y apoyan la Constitución? |
| 02 | Derechos                    | ¿Cuáles deberían ser los <b>derechos</b> fundamentales y universales contenidos en la Constitución?     |
| 03 | Deberes y responsabilidades | ¿Qué <b>deberes y responsabilidades</b> universales deben establecerse en la Constitución?              |
| 04 | Instituciones del Estado    | ¿Qué <b>Instituciones del Estado</b> debe incluir la Constitución y qué características deben tener?    |

Figura 1-1: Temas y preguntas de la consulta (Gobierno de Chile, 2015)

Para cada una de las preguntas, los participantes debían elegir conceptos de una lista preexistente, o sugerir nuevos. Por cada concepto seleccionado, se debía escribir un breve argumento que explicara por qué dicho concepto debía ser incluido en la Nueva Constitución. Las respuestas se consolidaron en un

documento denominado Diálogos Ciudadanos. En la Figura 1-2 se resume la metodología descrita.

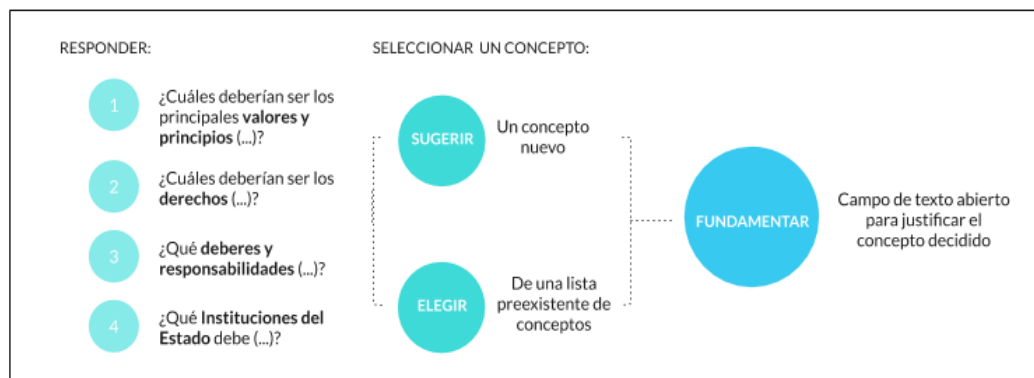


Figura 1-2: Resumen de metodología del proceso (Elaboración propia)

Según un informe de la OECD (2017), este proceso de cambio constitucional registró el porcentaje de participación ciudadana más alto en el mundo (1,3%). Además, en términos de género y edad, la distribución de la participación en el proceso fue representativa a nivel de país para estas dos dimensiones (Raveau, Couyoumdjian, Fuentes-Bravo, Rodriguez-Sickert y Candia, 2020).

Los Diálogos Ciudadanos fueron procesados por el Comité de Sistematización, órgano autónomo y externo al Gobierno. Su metodología consistió en un trabajo manual de unificación sintáctica de cada opinión, sin modificar su semántica (Fierro, Fuentes, Pérez y Quezada, 2017). Los resultados se consolidaron en una base de datos. En la Figura 1-3 se observa un ejemplo.

| Opinión original:                                                                                                                                                          | Opinión sistematizada:                                                                                                                                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| que el bienestar sea mas importante.<br>eliminación de la justicia militar.<br>equidad para justicia social igualitaria.<br>eliminación de la Ley de reserva del<br>cobre. | El bienestar debe ser más importante,<br>habiendo equidad para justicia social<br>igualitaria y eliminando la ley de reserva<br>del cobre y la justicia militar |

Figura 1-3: Ejemplo de opinión original versus opinión normalizada  
(Elaboración propia)

Para poder analizar los resultados del proceso participativo, se puede acceder a las [Bases de datos](#) disponible en [datos.gob.cl](http://datos.gob.cl); o a los informes ejecutivos del [Comité de Sistematización](#) y del [Consejo Ciudadano de Observadores](#) (Becker Castellaro, 2018). Sin embargo, la participación no es fácil de interpretar teniendo una base de datos tan extensa, o por el contrario, fácil de atender casos específicos con tan solo 200 páginas.

## **2. MARCO TEÓRICO**

### **2.1 Participación ciudadana**

La participación ciudadana es un proceso amplio que contempla el involucramiento de los ciudadanos en los asuntos públicos. Estos asuntos pueden ser a nivel comunitario o nacional y en diversos ámbitos de gobierno y sociedad (PNUD, 2014). Arnstein (1969), fue una de las primeras en plantear un modelo conceptual sobre el grado de involucramiento que pueden tener los ciudadanos en estos asuntos.

Arnstein (1969) definió ocho niveles, que van desde su forma más débil, la manipulación, hasta la más completa, control ciudadano. La OECD (2015), a modo de simplificar el modelo de Arnstein, ordenó dicho modelo en tres niveles: Informar, consultar y vincular (ver Figura 2-1).

El “informar” es el nivel de menor involucramiento. Considera tener a la ciudadanía al tanto de los temas que se conversan y las decisiones que se toman. El “vincular” es el de mayor involucramiento, donde la participación de la ciudadanía puede tener una influencia importante en la toma de decisiones.

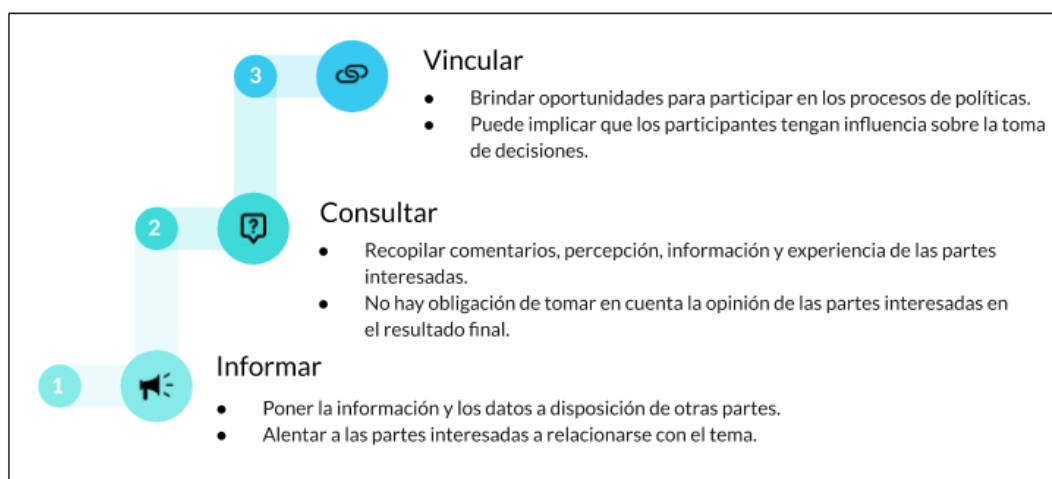


Figura 2-1: Niveles de participación ciudadana según OECD (2015), basada en Arnstein (1969)

El proceso de cambio constitucional llevado a cabo en nuestro país, según esta definición, entraría en el nivel de “consultar”. Vale decir, el interés de las autoridades de gobierno recae en recopilar comentarios y percepción de la ciudadanía, sin ser estos necesariamente vinculantes.

La participación ciudadana en procesos de reforma constitucional ha ido en aumento. De las doscientas nuevas constituciones promulgadas en las últimas cuatro décadas, más del 70 por ciento ha incorporado al menos una forma de participación ciudadana (Maboudi y Nadi, 2016).

La mayoría de estas iniciativas están comenzando a usar IA para gestionar el proceso de participación y consulta ciudadana. Se han utilizado algoritmos de

aprendizaje de máquina para sintetizar y clasificar el contenido (Hsiao, Lin, Tang, Narayanan y Sarahe, 2018). De esa manera, el uso de datos ha logrado una revisión y discusión más eficiente (Noveck, 2018).

## **2.2 Modelamiento de Tópicos**

Con el surgimiento de la ciencia de la computación, las tareas de análisis de texto se han podido automatizar. La IA es el campo científico para ello. Estas tecnologías funcionan mediante algoritmos de clasificación, análisis estadístico y distribuciones de probabilidad (Wang y Blei, 2011). Con una gran cantidad de datos de entrenamiento, la IA permite comprender distintas perspectivas que subyacen en los documentos (Radford, Narasimhan, Salimans y Sutskever, 2018).

Una de las perspectivas más comunes y útiles para entender documentos, es por sus tópicos. El proceso de aprendizaje y extracción de estos temas se denomina modelamiento de tópicos o *topic modeling* (Brynjolfsson y McAfee, 2017). Este modelo es capaz de agrupar documentos y categorizarlos en base a uno o varios tópicos. Se entiende tópico como la estructura semántica “abstracta” que caracteriza a un documento en base a sus palabras (Martin y Schuurman, 2017).

Existen dos niveles para modelar tópicos: a nivel de palabras y a nivel de documentos. Es posible agrupar palabras para crear temas y agrupar documentos según dichos temas (Ver Figura 2-2). Diversas técnicas para el modelamiento de

tópicos han surgido a lo largo de los años para lograr examinar un conjunto de documentos y descubrir cuáles podrían ser los temas y cuál es el equilibrio de temas de cada documento.

Una de las primeras aproximaciones al área que define gran parte de los conceptos a tratar es la “bolsa de palabras” o en inglés *bag-of-words*. Esta técnica asume que el orden de las palabras dentro de un documento no influye, por lo que este se ignora (Wallach, 2006). El resultado del modelo son grupos de palabras entendidos como tópicos.

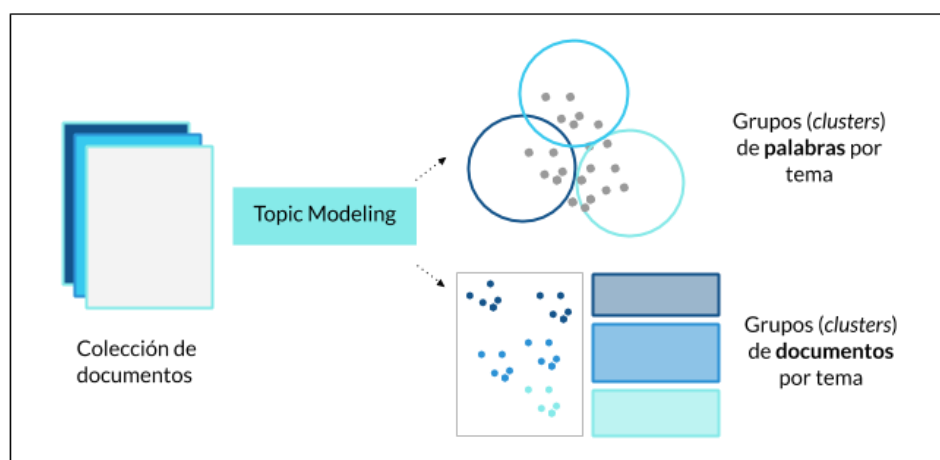


Figura 2-2: Representación gráfica de *topic modeling*

En la técnica de *tf-idf* (*term frequency-inverse document frequency*) (Salton y McGill, 1986), se elige un diccionario en base a las ocurrencias de las palabras en los documentos. Luego, por cada documento, se mantiene la cuenta de la cantidad

de ocurrencias de cada palabra del diccionario. Tras normalizar la frecuencia de las palabras, se compara con la frecuencia inversa de las palabras en todo el documento. El resultado de esta comparación es una matriz palabra-documento.

Bajo estas dos líneas de pensamiento, surge el modelo de *Latent Dirichlet Allocation (LDA)*, modelo generativo estadístico introducido por David Blei, Andrew Ng y Michael Jordan (2003). De forma sencilla, *LDA* toma una matriz de *Palabras  $\times$  Documentos ( $P \times D$ )* y genera dos matrices como resultado. La primera es la probabilidad de cada palabra de pertenecer a cada tópico o  $P(P|T)$ . La segunda es la probabilidad de cada tópico de pertenecer a cada documento o  $P(T|D)$  (Figura 2-3).

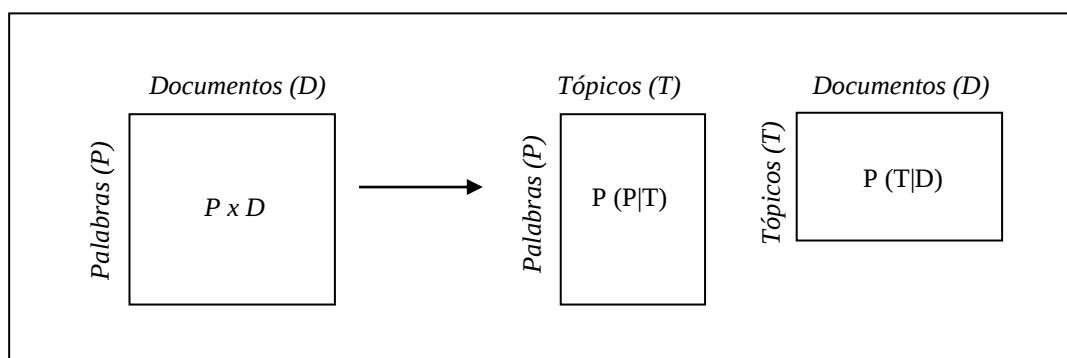


Figura 2-3: Interpretación matricial de topic modeling

El modelo generativo descrito por Blei et al., (2003) incluye la distribución de Dirichlet y otro parámetros para definir dicha probabilidad. En este contexto, tópico es entendido como un grupo de palabras que frecuentemente ocurren juntas.

## 2.3 Visualización de información

Hasta hace un tiempo, el término visualización significaba construir una imagen visual en la mente (Shorter Oxford English Dictionary, 1972). Hoy, dicha definición ha evolucionado, y es entendida como una representación gráfica de datos o conceptos (Mayr, Hynek, Salisu y Windhager, 2019). Ware (2020), indica cuatro razones de por qué es importante visualizar, las cuales se pueden observar en la Figura 2-4.

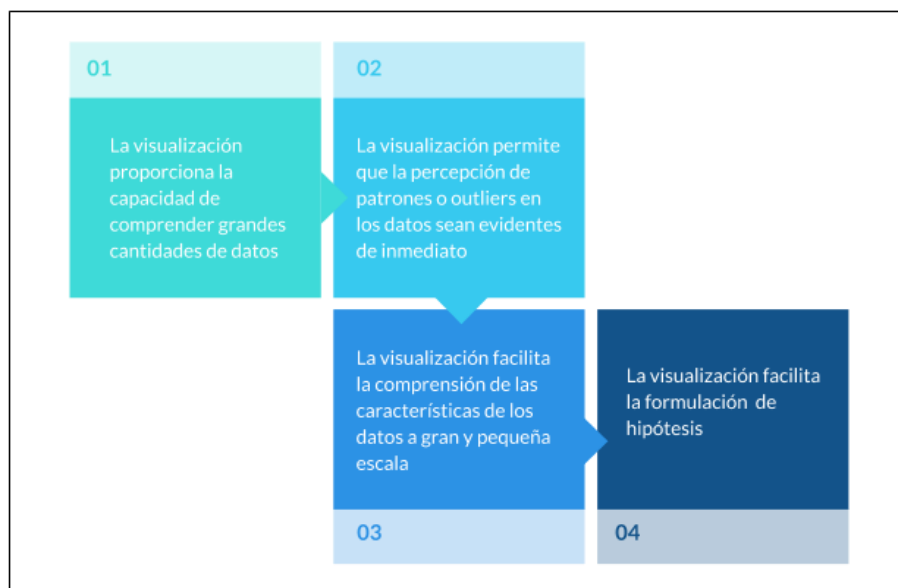


Figura 2-4: Cuatro razones de por qué visualizar según Ware (2020)

La investigación en visualización de información ha reunido investigadores de distintas áreas del conocimiento. Entre ellas, ciencia de la computación, diseño, psicología y estadísticas (Fayyad, Grinstein, y Wierse, 2002). France y Akkucuk

(2020), indican que para lograr que el diseño de estas interfaces sea efectivo, se requiere del trabajo conjunto de esas disciplinas.

Distintas metodologías han surgido para crear estas visualizaciones. Una de ellas, es la propuesta por Tamara Munzner (2009). Este método combina cuatro áreas del conocimiento: antropología/etnografía, diseño, ciencia de la computación y psicología cognitiva. Munzner propone diseñar las visualizaciones a través de tres preguntas fundamentales que se hacen de manera iterativa, y que se pueden observar en la Figura 2-5.

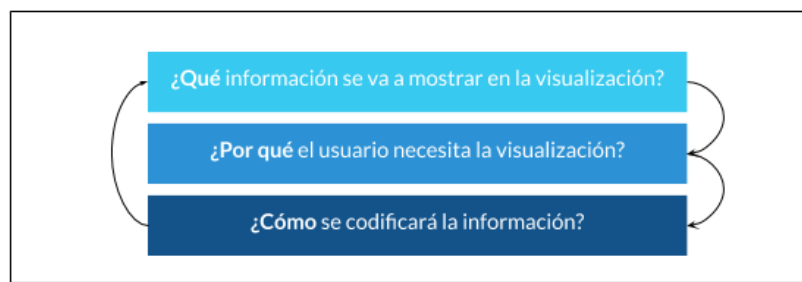


Figura 2-5: Preguntas fundamentales framework Munzner (2015)

Además, se definen cuatro niveles de diseño dependientes para poder resolver las preguntas fundamentales de manera óptima: Situación de dominio, Abstracción de tareas y datos, Codificación visual e Idioma de interacción y Algoritmo. Se muestra el *framework* en la Figura 2-6. En este, el *output* de un nivel superior, es el *input* de los niveles inferiores.

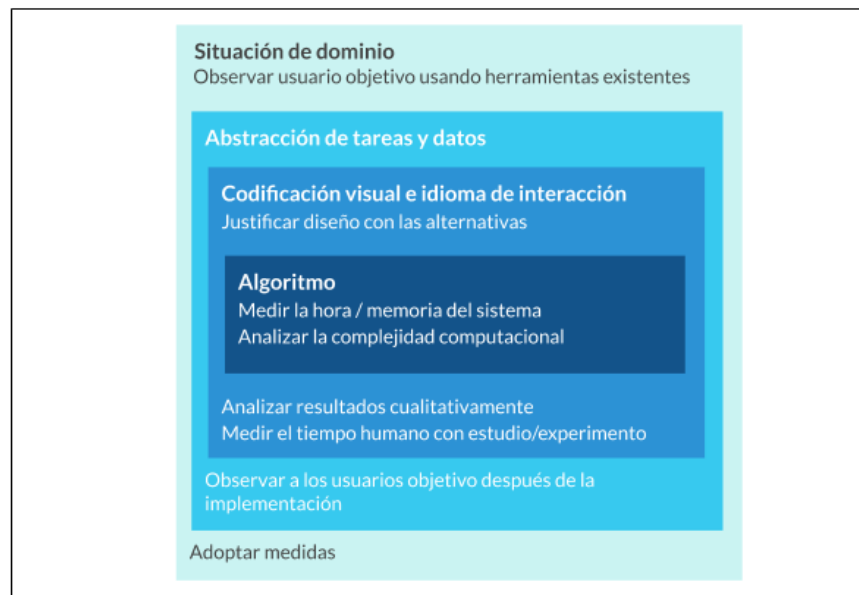


Figura 2-6: Diagrama framework Tamara (Munzner, 2015)

La descripción más detallada de las cuatro áreas es la siguiente:

- **Situación de dominio:** En este primer nivel se debe establecer el grupo de usuarios objetivo de la visualización, el dominio de su conocimiento e interés, sus preguntas y problemas, y los datos asociados.
- **Abstracción de tareas y datos:** En este segundo nivel de diseño, se responde a las preguntas *qué* y *por qué*. Se identifica qué tipo de datos se tiene y los atributos de dichos datos. Además, se define la o las acciones a realizar por el usuario (analizar, buscar o consultar) y los aspectos de interés para el usuario (todos los datos, algunos atributos, las relaciones entre los datos, etc.).
- **Codificación visual e idioma de interacción:** Este nivel responde a la pregunta *cómo*. Teniendo los conceptos de las acciones y del tipo de dato

que se tiene, se requiere elegir el idioma en que se representará la información.

- **Algoritmo:** Este último nivel, está enfocado en poder realizar todas las decisiones de diseño y código, de manera automática y óptima

## 2.4 Usabilidad

Dentro del ámbito de la interacción humano computador, la usabilidad se ha vuelto un concepto altamente utilizado a la hora de evaluar visualizaciones (Umar, Bakhat y Hassan, 2020). La usabilidad se ha hecho parte del conjunto de normas establecidas por la Organización Internacional de Normalización (ISO).

Usabilidad es entendida como el grado con que un producto puede ser usado por usuarios específicos para alcanzar metas con efectividad, eficiencia y satisfacción percibida en un contexto de uso específico (ISO 9241).

De la misma manera, cada una de esas dimensiones han sido entendidas como:

- **Efectividad:** Se refiere al rendimiento de las tareas, a la precisión con la que el usuario cumple con sus objetivos. Se suele evaluar en base a la precisión y la tasa de error o tasa de éxito (Howorko et. al, 2018).
- **Eficiencia:** Se refiere a la relación entre el nivel de efectividad y el consumo de recursos. Se utilizan por ejemplo métricas de cantidad de

clics y tiempo de completación de la tarea (Zia, Guo, Zhou, Essa y Jarc, 2019).

- **Satisfacción:** Se refiere al confort reportado por el usuario tras resolver una tarea. Se utilizan métricas de nivel de satisfacción en escalas Likert (Isaac, Abdullah, Ramayah y Mutahar, 2017).

A través de evaluaciones de usabilidad se identifican los elementos más apropiados para una interfaz. En ellas, se mide el grado de efectividad, eficiencia y satisfacción con que los usuarios pueden utilizar el sistema diseñado para resolver ciertas tareas (Manresa-Yee, Amengual, y Ponsa Asensio, 2014).

Desde el inicio de las evaluaciones de interfaces de visualización, la efectividad ha sido la principal dimensión evaluada (Szafir, Haroz, Gleicher y Franconeri, 2016). Si bien la satisfacción era importante cuando se desarrolló la norma ISO, solo recientemente se ha comenzado a considerar en las evaluaciones (Barnum, 2020). Incluso, aproximadamente la mitad de los estudios en el área no incluyen evaluaciones de ningún tipo (Elmqvist y Yi, 2015). Se necesitan estudios de usabilidad para evaluar y validar la relevancia de las interfaces de visualización (Yang et al., 2017).

### **3. TRABAJO RELACIONADO**

#### **3.1 Modelamiento de tópicos en opiniones**

La necesidad de captar la opinión del público sobre políticas públicas, campañas de marketing y preferencias de productos, ha despertado un interés creciente por parte de la comunidad científica (Poria, Cambria y Gelbukh, 2016). Lo indispensable ha sido encontrar una manera de hacer frente a los grandes números de opiniones disponibles de manera eficiente (Hagen, Uzuner, Kotfila, Harrison y Lamanna, 2015).

Las técnicas de procesamiento del lenguaje natural han resultado útiles para abordar este desafío. En específico, los modelos de tópicos se han utilizado para la categorización y agrupamiento de opiniones, junto con el análisis de sentimientos (Srivastava y Sahami, 2009). Sus principales fuentes de estudio han sido valoraciones de clientes, comentarios en redes sociales y participaciones ciudadanas, todas emitidas a través de la web.

En el ámbito de valoraciones de clientes, Calheiros, Moro y Rita (2017), realizaron una clasificación de sentimientos utilizando varias fuentes diferentes de comentarios de clientes de un hotel a través de un enfoque de minería de texto. El algoritmo de LDA se aplicó para reunir temas relevantes que vislumbraran las

principales oportunidades de mejora del hotel. A través de este estudio, se pudo validar el uso de distintas opiniones en torno a preguntas en común.

En redes sociales, los *tweets* han sido ampliamente analizados. Ghosh y Guha (2013), hicieron un análisis para revelar opiniones en torno a temas de salud pública. Para el análisis se utilizó el modelo LDA. Del modelo se extrajeron temas coherentes y significativos que pudieron ser validados. Por lo tanto fue posible concluir que, técnicas de automatización como el LDA proporcionan un gran potencial a la hora de analizar un conjunto de opiniones de texto corto.

En el caso de participación ciudadana, Hagen et al. (2015), utilizaron un enfoque de modelado de tópicos para descubrir temas emergentes dentro de una encuesta pública realizada en USA. Aplicaron el algoritmo de LDA a una colección de 2,850 textos y produjeron 30 temas emergentes. Finalmente, sus resultados demostraron que el modelamiento de tópicos tiene el potencial de permitir la interpretación de grandes cantidades de sugerencias de políticas generadas por los ciudadanos.

Se han logrado varios avances en el modelamiento de la opinión pública. Sin embargo, los estudios recientes tienden a centrarse principalmente en el análisis de opinión de redes sociales (Parra et al., 2016; Steinskog, Therkelsen y Gambäck, 2017; Yu, Xu, Wang y Ni, 2019; Kant, Weisser y Säfken, 2020). El modelamiento

de opiniones de instancias institucionales de participación ciudadana sigue sin prestar suficiente atención (Mukherjee, 2019). De esa manera, sería más manejable abordar el creciente interés y demanda hacia una participación ciudadana más activa, directa y sustancial en la formulación de políticas (Dryzek, 2002; Floridia, 2018).

### **3.2 Interfaces de visualización de tópicos**

Los resultados del LDA pueden ser poco intuitivos dada su notación matemática (Cai, Sun y Sha, 2018). Afortunadamente, han surgido nuevas áreas de investigación de la intersección entre las humanidades y la informática (Uboldi y Caviglia, 2015). Desde el diseño, la antropología y la psicología, se han establecido metodologías de trabajo para construir representaciones visuales de estas expresiones matemáticas. Hoy se conocen como visualizaciones de tópicos.

En los últimos años se han desarrollado varias interfaces de visualización de tópicos (Gardner et al., 2010; Chaney y Blei, 2012; Snyder et al., 2013) Uno de los objetivos principales de este tipo de visualizaciones es permitir a los usuarios explorar el conjunto de documentos e identificar los tópicos (Ganesan et al., 2015).

En base a un proceso de diseño centrado en el usuario, hay tres aspectos fundamentales a la hora de construir sistemas de visualizaciones de tópicos (Yang et al., 2017). Se debe presentar una visión general sobre la organización de los

temas encontrados; una visión exploratoria de los temas; y una visión particular que permita al usuario revisar el compilado de los documentos o un extracto de ellos.

Para entregar una visión general, se utilizan comúnmente visualizaciones interactivas que permitan mostrar tendencias y clusterización entre los tópicos. Dentro de los recientemente estudiados se encuentran gráficos de barras apiladas (Sievert y Shirley, 2014) y mapas de árbol (Ganesan et al., 2015).

Los gráficos de barras apiladas, son aquellos en que las barras se dividen en variables nominales. *Theme River* (Havre, Hetzler, Whitney y Nowell, 2002) es una línea de tiempo que indica el flujo de temas de documentos. El *Name Voyager* visualiza un gráfico de la popularidad de los nombres de bebés durante el siglo pasado (Wattenberg, 2005).

Los mapas de árbol (Figura 3-1) utilizan rectángulos de diferentes tamaños y colores para transmitir valores numéricos y nominales. *LDAExplore* (Ganesan et al., 2015) visualiza los temas de un modelo que utiliza resúmenes de artículos de visualización de información, donde cada resumen se considera un documento.

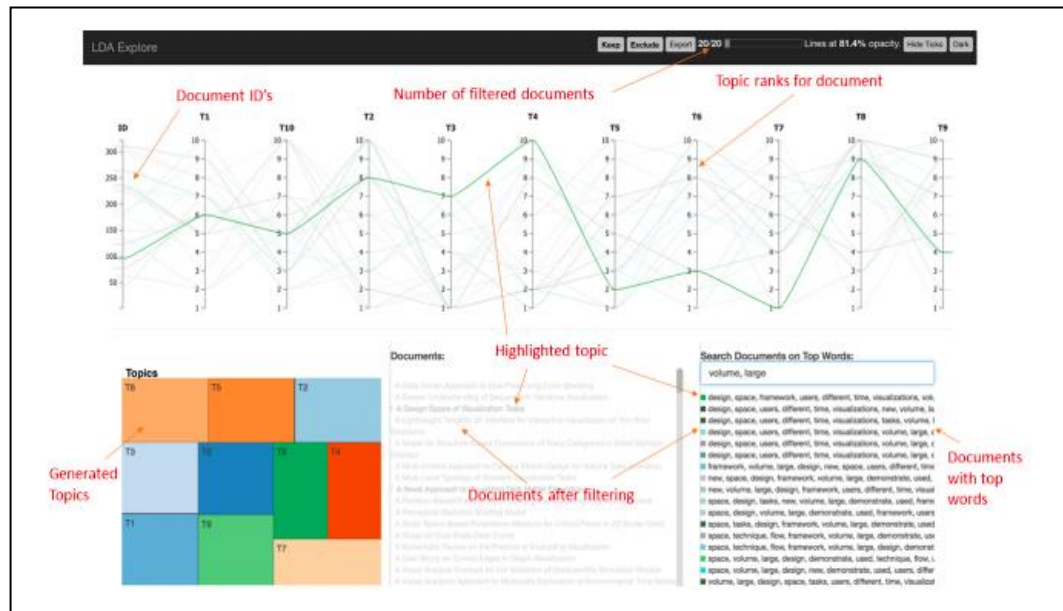


Figura 3-1: Sistema con mapa de árbol (Ganesan et al., 2015)

Para entregar una visión exploratoria, los datos geolocalizados se han mostrado a través de visualizaciones de tendencia espacial como los mapas coropléticos (Koylu, 2019). Ver Figura 3-2. Estos mapas muestran áreas geográficas coloreadas según una variable numérica. Esto proporciona una forma de visualizar patrones sobre un área geográfica determinada.

Para entregar una visión particular, los mapas de densidad de puntos han sido una forma de detectar patrones y distribución espaciales mientras se puede leer cada documento a través de una ventana emergente (Cai, Sun y Sha, 2018). De esa manera se pueden localizar opiniones de cierta región, o simplemente explorar opiniones.

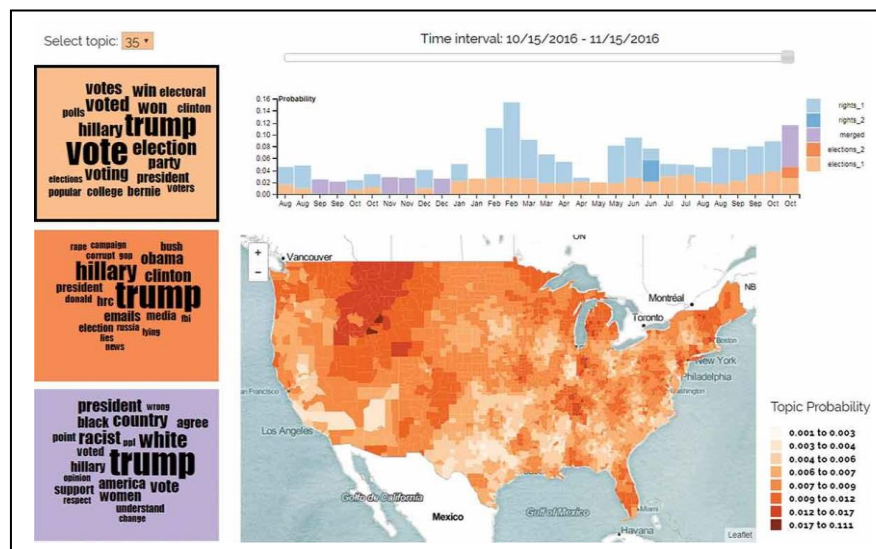


Figura 3-2: Sistema con mapa coroplético (Koylu, 2019)

Existen diferentes alternativas de visualización para codificar la misma información. Es importante que estas alternativas cumplan con los criterios de expresividad (Mackinlay, 1986). Dados dos lenguajes gráficos que expresan alguna información, la pregunta obvia es qué lenguaje involucra un diseño que especifica la presentación más efectiva. A diferencia de la expresividad, que sólo depende de la sintaxis y semántica del lenguaje gráfico, la efectividad también depende de las capacidades del perceptor (Qu y Hullman, 2016).

### 3.3 Trabajos vinculados a Consulta Proceso Constituyente 2016

#### • Plataforma de trazabilidad

Plataforma orientada a visualizar los contenidos del Borrador Constitucional propuesto por el Gobierno de Michelle Bachelet (Boletín N° 11.617-07 del 6 de

marzo del 2018), y su relación con las Bases Ciudadanas. La plataforma de trazabilidad está formada por dos componentes principales, un panel con el Borrador Constitucional y un índice de conceptos.

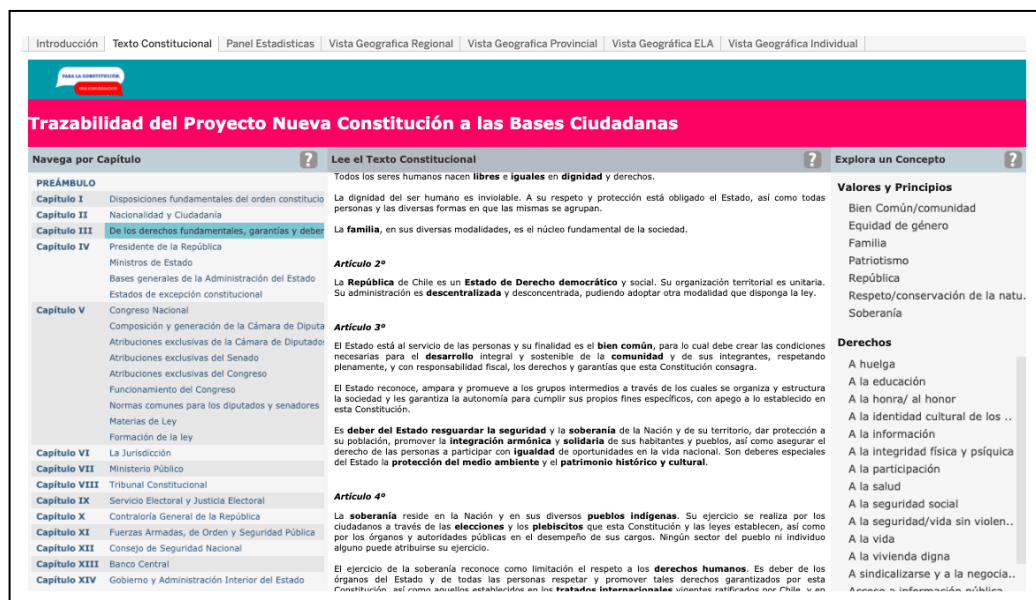


Figura 3-3: Plataforma de trazabilidad (Una Constitución para Chile, 2018)

El usuario puede revisar el texto y ver resaltados los conceptos mencionados en la consulta (Figura 3-3). Al seleccionar un concepto, se puede acceder a información sobre la distribución geográfica de las menciones de dicho concepto, o información estadística del mismo.

### • Constitución Abierta

Plataforma orientada a presentar en formato abierto, las opiniones de la consulta ciudadana desde una perspectiva geográfica. Constitución Abierta está formada

por tres componentes principales, un buscador por conceptos, un mapa para explorar y una tabla descriptiva (Figura 3-4).

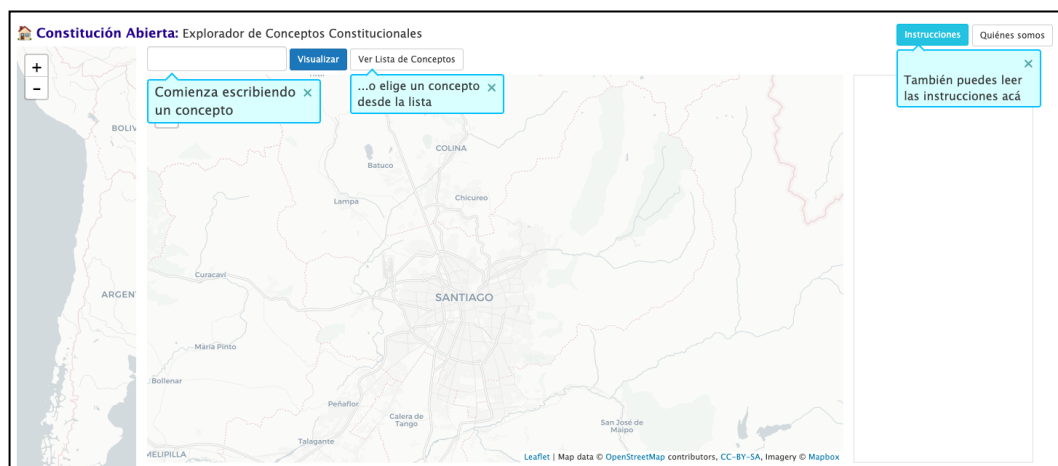


Figura 3-4: Plataforma Constitución Abierta (IMFD, 2018)

El usuario puede escribir un concepto en el buscador y automáticamente se visualiza en el mapa la distribución de menciones por comuna. Al interactuar con alguna comuna, se despliega una tabla que detalla cada una de las instancias de participación en que fue mencionado dicho concepto y otros datos estadísticos.

Cabe mencionar que esta plataforma fue construida por miembros del Instituto Milenio de Investigación Fundamentos de los Datos basándose en investigaciones previas de Fierro et al. (2017).

#### **4. OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN**

El objetivo general de este trabajo es comparar la usabilidad de dos interfaces de visualización que utilizan el modelo de tópicos de la Consulta Ciudadana del Proceso Constituyente 2016. Para lo anterior, se definen los siguientes objetivos específicos:

**O1:** Implementar, entrenar y validar un algoritmo de modelación de tópicos.

**O2:** Desarrollar, evaluar y comparar dos interfaces de visualización de tópicos.

En este trabajo se implementó y entrenó el algoritmo de LDA para realizar un modelo de tópicos. Luego se hizo una validación de coherencia semántica con un grupo de usuarios. Posteriormente se desarrollaron dos interfaces de visualización de tópicos. La primera incluyó un gráfico de barras apiladas, y la segunda un mapa de árbol. Finalmente, se midió la efectividad a través de la tasa de éxito, la eficiencia a través del tiempo de completación y satisfacción percibida a través del nivel de satisfacción en tareas de categorización y comparación con usuarios.

A través de este trabajo se busca responder a las siguientes preguntas de investigación:

**P1:** ¿El modelo de tópicos LDA puede lograr una categorización coherente de las opiniones de la consulta ciudadana 2016?

**P2:** ¿Cuál es la mejor interfaz de visualización para analizar el modelo de tópicos de las opiniones de la consulta ciudadana 2016?

**P2.1:** ¿Qué interfaz es mejor para realizar análisis de categorización entre los tópicos?

**P2.2:** ¿Qué interfaz es mejor para realizar análisis de comparación entre los tópicos?

## 5. METODOLOGÍA

El primer paso en el desarrollo de este trabajo fue el procesamiento y análisis de los datos. La metodología escogida para ello fue la metodología KDD - *Knowledge Discovery in Databases* (Fayyad, Piatetsky-Shapiro y Smyth, 1996). KDD consta de un proceso iterativo de cinco etapas (Debusse, de la Iglesia, Howard, Rayward-Smith, 2001). Las etapas son: selección de los datos, pre-procesamiento y limpieza, transformación, minería de datos y evaluación.

### 5.1 Selección y descripción de los datos

La base de datos seleccionada corresponde a las Bases Ciudadanas, documento de opiniones normalizadas por el Comité de Sistematización. Esta cuenta con un total de 205,357 opiniones organizadas en los cuatro temas mencionados en el capítulo 1. El corpus tiene un total de 4.653.518 palabras, lo que da un promedio de 23 palabras por opinión ( $\sigma = 13$ ).

Del total de 21 atributos que posee la base de datos, se seleccionaron los siguientes para trabajar el modelo y las visualizaciones. En la Tabla 5-1 se observa un extracto de ellos:

- **id**: Identificador único de cada entrada en la base de datos.

- **región:** Número de la región de Chile en la cual se llevó a cabo el encuentro registrado. Puede tomar los valores 1 a la 15. Su codificación se corresponde con la oficial.
- **comuna:** Número de la comuna de la región en la cual se llevó a cabo el encuentro registrado. Puede tomar los valores 1 al 346. Su codificación es arbitraria.
- **Normalización:** Fundamento normalizado por el Comité de Sistematización.

Tabla 5-1: Extracto de algunos elementos seleccionados de la base de datos

| <b>id</b> | <b>región</b> | <b>comuna</b> | <b>normalización</b>                                                                      |
|-----------|---------------|---------------|-------------------------------------------------------------------------------------------|
| 45349     | 13            | 307           | Debe haber igualdad entre hombres y mujeres.                                              |
| 45235     | 10            | 248           | Hombres y mujeres deben tener los mismos derechos y deberes para una sociedad equitativa. |
| 44861     | 9             | 170           | Se debe garantizar la igualdad de género, de derechos y oportunidades.                    |

## 5.2 Pre-procesamiento y limpieza

Una vez seleccionados los datos a utilizar, se prosiguió a revisar la relevancia de los datos. Se removieron aquellas entradas cuya información no permite el análisis de tópicos. Vale decir, aquellos registros cuyos fundamentos estuviesen vacíos,

con una notación de “blanco” o con un asterisco, y opiniones recopiladas desde el extranjero.

Tras la limpieza, se redujo el tamaño del set de datos a 185,458 entradas, lo que implica una reducción del 9,8% sobre el set de datos original. A pesar de la reducción en la cantidad de los datos, los datos restantes aún hacen factible el modelo.

### **5.3 Transformación**

Para poder trabajar los datos de forma eficiente, se debieron transformar tres de los atributos del set de datos original, que son la normalización, la región y la comuna donde se registró la participación.

#### **5.3.1 Transformación de [normalización]**

Se utilizaron las librerías de Python, NLTK Toolkit y Gensim (Bird y Loper, 2004). El primer paso fue homogeneizar el texto al eliminar mayúsculas, acentos y puntuación. A esta acción se le conoce como tokenización.

El segundo paso, fue remover las palabras que no aportan valor al significado del texto. Tales como los artículos, pronombres y preposiciones.

A este tipo de palabras se les conoce como *stopwords*. Además, se eliminaron palabras de menos de 3 letras.

Finalmente, se realizó una lematización. Vale decir, reemplazar las palabras por su forma más aceptada en el diccionario. Donde las palabras en tercera persona cambian a primera persona y los verbos en pasado y futuro cambian a presente.

De esta transformación se generó el atributo *preprocesado*, que contiene la normalización de la opinión limpia para poder ser procesada.

### **5.3.2 Transformación de [region]**

Las regiones de Chile vienen codificadas de la 1 a la 15. Sin embargo, un par de años después de la consulta, se creó una nueva región. La región número 8, Región del Biobío, se dividió y dio origen a la Región del Ñuble, la número 16. Las opiniones correspondientes fueron reasignadas a la región 16. Por último, se transformó la codificación de cada región por su nombre. Para todo lo anterior se utilizó la librería *pandas* de Python. De la transformación se originó el atributo *nombre región*.

### 5.3.3 Transformación de [comuna]

Las comunas de Chile vienen codificadas de la 1 a la 346. A excepción de la comuna número 1000 que hace referencia a cualquier otra localidad fuera de Chile (participación de chilenos en el extranjero). La codificación la llevó a cabo el Comité de Sistematización sin ninguna notación estándar. Para poder descifrar la codificación, los datos sistematizados se cruzaron a través del id, con los datos sin sistematizar, originando el atributo *nombre comuna*. Por otro lado, a través de la API de Google Maps, se obtuvieron las coordenadas de cada comuna. De la transformación se originó el atributo *latitud* y *longitud*.

## 5.4 Modelo de tópicos

Para obtener el modelo, se utilizó la librería de Python, Gensim (Rehurek y Sojka, 2010), que es una librería especializada para realizar modelamiento de tópicos. Gensim está diseñada para extraer automáticamente la semántica de los documentos de manera eficiente (Araya, 2018).

Lo primero fue crear un diccionario para cada documento. Este diccionario indica cuántas palabras contiene el documento y cuántas veces aparecen esas palabras en ese documento. Gensim genera una matriz de (palabra\_id, palabra\_frecuencia). Por ejemplo, (0, 1) en la salida, implica que la palabra 0 aparece una vez.

Del diccionario se filtraron aquellas palabras que aparecen en menos de 15 documentos y se conservaron las 100,000 palabras más frecuentes. Luego, se calculó la frecuencia de término – frecuencia inversa de documento o *term frequency - inverse document frequency* (TF-IDF). Esta medida numérica refleja cuán relevante es una palabra para un documento en una colección.

LDA es una técnica no supervisada, lo que significa que antes de ejecutar el modelo no sabemos cuántos tópicos se tiene. La coherencia temática es una de las principales técnicas utilizadas para estimar el número de tópicos (Mimno, 2011).

Para encontrar el número óptimo, se construyeron muchos modelos LDA con diferentes valores del número de temas ( $k$ ) y se eligió el de mayor valor de coherencia ( $c_v$ ). La elección de una " $k$ " que marca el final de un rápido crecimiento de la coherencia del tópico, suele ofrecer temas significativos e interpretables.

La Figura 5-1 muestra que el puntaje de coherencia aumenta con la cantidad de tópicos, con una disminución entre 15 y 25.  $K=30$  tiene alto puntaje de coherencia, pero probablemente tenga palabras clave repetidas entre los tópicos. Es por ello, que el " $k$ " escogido es de 15 tópicos, cuyo puntaje de coherencia es de 0,533.

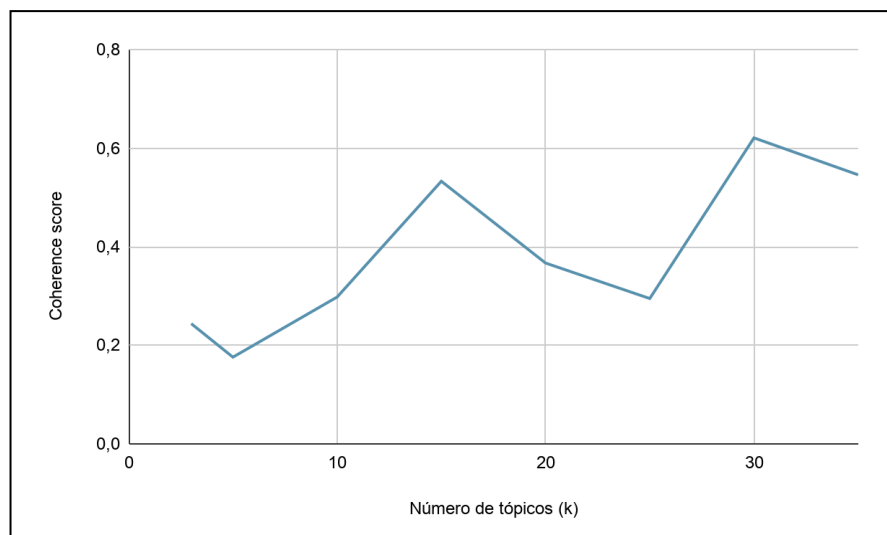


Figura 5-1: Puntaje de coherencia del modelo según número de tópicos (Elaboración propia)

El tiempo de entrenamiento promedio fue de 26 minutos. El *output* del modelo, fue la probabilidad de cada documento de pertenecer a cada tópico, y la probabilidad de cada palabra de pertenecer a cada documento. La Tabla 5-2 indica los 15 tópicos seleccionados por el modelo, compuestos por sus 6 palabras más probables. El número entre paréntesis de cada palabra, indica la probabilidad de la palabra de pertenecer al tópico en cuestión.

Tabla 5-2: Palabras por tópico

|                          |                       |                       |                      |                        |                         |
|--------------------------|-----------------------|-----------------------|----------------------|------------------------|-------------------------|
| participación<br>(0,339) | democracia<br>(0,164) | decisiones<br>(0,152) | consultas<br>(0,137) | vinculantes<br>(0,119) | plebiscito<br>(0,117)   |
| naturaleza<br>(0,148)    | medio<br>(0,087)      | ambiente<br>(0,085)   | cuidar<br>(0,067)    | recursos<br>(0,043)    | conservación<br>(0,043) |
| nueva                    | constitución          | pueblo                | defensor             | asamblea               | constituyente           |

|                          |                       |                       |                           |                           |                              |
|--------------------------|-----------------------|-----------------------|---------------------------|---------------------------|------------------------------|
| (0,095)                  | (0,082)               | (0,046)               | (0,042)                   | (0,031)                   | (0,029)                      |
| congreso<br>(0,084)      | unicameral<br>(0,079) | funciones<br>(0,075)  | parlamentarios<br>(0,062) | representativo<br>(0,042) | electo<br>(0,029)            |
| educación<br>(0,183)     | salud<br>(0,168)      | derecho<br>(0,146)    | acceso<br>(0,065)         | gratuita<br>(0,057)       | digna<br>(0,022)             |
| gobierno<br>(0,071)      | regional<br>(0,062)   | autonomía<br>(0,023)  | local<br>(0,022)          | mayor<br>(0,018)          | descentralización<br>(0,015) |
| leyes<br>(0,063)         | igualdad<br>(0,045)   | justicia<br>(0,043)   | normas<br>(0,024)         | todos<br>(0,022)          | cumplimiento<br>(0,012)      |
| convivencia<br>(0,042)   | vida<br>(0,030)       | bien<br>(0,028)       | común<br>(0,022)          | fundamental<br>(0,017)    | respeto<br>(0,017)           |
| todos<br>(0,222)         | derechos<br>(0,120)   | igualdad<br>(0,090)   | condición<br>(0,049)      | personas<br>(0,029)       | discriminación<br>(0,023)    |
| poder<br>(0,032)         | judicial<br>(0,032)   | ejecutivo<br>(0,023)  | legislativo<br>(0,023)    | funciones<br>(0,022)      | independientes<br>(0,021)    |
| patrimonio<br>(0,149)    | cultural<br>(0,094)   | proteger<br>(0,072)   | nuestra<br>(0,038)        | identidad<br>(0,028)      | pueblos<br>(0,021)           |
| derechos<br>(0,042)      | humanos<br>(0,032)    | protección<br>(0,029) | promoción<br>(0,026)      | vida<br>(0,025)           | esencial<br>(0,024)          |
| economía<br>(0,052)      | libre<br>(0,051)      | mercado<br>(0,021)    | promover<br>(0,018)       | recursos<br>(0,018)       | propios<br>(0,014)           |
| transparencia<br>(0,082) | cargos<br>(0,078)     | públicos<br>(0,042)   | control<br>(0,031)        | atribuciones<br>(0,023)   | fiscalización<br>(0,023)     |
| igualdad<br>(0,094)      | derechos<br>(0,089)   | garantizar<br>(0,074) | equidad<br>(0,059)        | brechas<br>(0,033)        | resolver<br>(0,027)          |

## 5.5 Etiquetas de tópicos

La etiqueta de un tópico corresponde a una palabra, o secuencia de palabras, que es semánticamente significativa y cubre el significado latente de un tópico (Mei et al. 2007). Una forma estándar de etiquetar un tópico es representarlo a través de

sus palabras principales. Sin embargo, buenas etiquetas a menudo se expresan mejor con palabras que no necesariamente están entre las principales del tópico (Lau et al., 2011). Las etiquetas para este modelo, se muestran en la Tabla 5-3.

Tabla 5-3: Etiqueta y palabras por tópico

| <b>Etiqueta</b>        | <b>Palabras</b>                                                           |
|------------------------|---------------------------------------------------------------------------|
| 1 Participación        | participación, democracia, decisiones, consultas, vinculantes, plebiscito |
| 2 Medio ambiente       | naturaleza, medio, ambiente, cuidar, recursos, conservación               |
| 3 Nueva constitución   | nueva, constitución, pueblo, defensor, asamblea, constituyente            |
| 4 Congreso             | congreso, parlamento, unicameral, representativo, democrático, electo     |
| 5 Educación y salud    | educación, salud, derecho, acceso, gratuita, digna                        |
| 6 Descentralización    | gobierno, regional, autonomía, local, mayor, descentralización            |
| 7 Igualdad ante la ley | leyes, igualdad, justicia, normas, todos, cumplimiento                    |
| 8 Bien común           | convivencia, vida, bien, común, fundamental, respeto                      |
| 9 Igualdad de derechos | todos, derecho, igualdad, condición, personas, discriminación             |
| 10 Poderes del Estado  | poder, judicial, ejecutivo, legislativo, funciones, independientes        |
| 11 Patrimonio cultural | patrimonio, cultural, pueblos, proteger, nuestra, identidad               |
| 12 Derechos humanos    | derechos, humanos, protección, promoción, vida, esencial                  |
| 13 Economía local      | economía, libre, mercado, promover, recursos, propios                     |
| 14 Transparencia       | transparencia, cargos, públicos, control, atribuciones,                   |

| fiscalización         |                                                            |
|-----------------------|------------------------------------------------------------|
| 15 Igualdad y equidad | igualdad, derechos, garantizar, equidad, brechas, resolver |

## 5.6 Estudio de usuarios

Para poder cumplir con los objetivos y responder a las preguntas de investigación planteadas en el Capítulo 4, se realizó un estudio de usuario.

Para la metodología de estudio de usuario (Moreno, 2008) se consideraron 6 aspectos:

- **Definir muestreo y usuarios:** Muestreo no probabilístico. Se decidió qué y cuántos usuarios forman parte del estudio a través de criterios de inclusión y exclusión y softwares estadísticos respectivamente.
- **Escoger método de recogida de datos:** Se empleó un método directo en la recogida de datos a través de un cuestionario online.
- **Elaborar cuestionario:** Se seleccionaron temas, preguntas y tipo de respuestas. Se usó el diseño *Within-Subjects Design*, aplicando balanceo de preguntas para evitar un sesgo correspondiente al orden de las preguntas.
- **Desarrollar protocolos:** Si bien la investigación no se sometió al proceso de aprobación del Comité de Ética UC, se siguió el protocolo estipulado por Ética UC para llevar a cabo el estudio. Se redactó una invitación a participar y se creó un consentimiento informado (Ver Anexo A).

- **Validar cuestionario:** Con un grupo pequeño de usuarios se comprobó si la estructura del mismo era adecuada y si las preguntas estaban redactadas de forma clara y precisa.
- **Enviar y recoger cuestionario:** Se agendaron sesiones de 30 minutos vía zoom para responder el cuestionario online con asistencia remota.
- **Analizar los datos:** Una vez recibidos los cuestionarios se hizo preciso normalizar, tabular y analizar los datos para evaluar las hipótesis planteadas.

## **6. DISEÑO E IMPLEMENTACIÓN**

Para las decisiones de diseño se siguió el *Framework* de Tamara Munzner explicado en el Capítulo 2.

### **6.1 Situación o dominio**

Los datos seleccionados, provienen de la consulta ciudadana del Proceso Constituyente 2016. Estos datos registran opiniones geolocalizadas y, posteriormente, categorizadas por el modelo de tópicos del Capítulo 5.

El grupo de usuarios objetivo corresponde a autoridades públicas. Su interés radica en analizar las respuestas de la consulta pública, a nivel general y particular, de manera eficiente, efectiva y satisfactoria. Estos usuarios no poseen, necesariamente, experiencia en el área de visualización de datos, ni de modelación de tópicos.

En base a lo anterior, el usuario objetivo, tiene específicamente el interés de (1) identificar los temas más discutidos en la consulta, (2) descubrir tendencias o patrones de los temas discutidos de manera espacial, y (3) revisar opiniones referentes a los temas más discutidos en detalle.

Actualmente, los usuarios pueden acceder a esta información a nivel general, a través del informe de Sistematización del Proceso Constituyente. El informe entrega una visión panorámica de los temas más tratados mediante descripciones, tablas y códigos de colores. En cuanto a nivel particular, es posible acceder cada a una de las opiniones en la base de datos del Portal de Datos Abiertos ([datos.gob.cl](http://datos.gob.cl)).

## 6.2 Tareas y abstracción de datos

En base a las necesidades del usuario, se espera que la visualización permita:

1. Identificar los temas más discutidos en la consulta: **Analizar y consumir datos** para descubrir la distribución y organización de los tópicos. Cuáles son los tópicos, cual es más importante que otro, y cómo se relacionan entre sí.
2. Descubrir tendencias o patrones de los temas discutidos de manera espacial: **Analizar y consumir datos** para conocer la distribución geográfica de los temas más discutidos, comparar la importancia de cada tema entre regiones y hallar aspectos en común o *outliers*. Además de **buscar** la importancia de un tópico en alguna región en particular.
3. Para revisar opiniones referentes a los temas más discutidos en detalle: **Explorar** por el mapa para navegar por las opiniones, o **buscar** por tema más discutido o región.

Los datos se encuentran actualmente en formato relacional. Cada opinión se encuentra tabulada por sus atributos: fundamento, región, comuna, latitud, longitud, y probabilidad de pertenecer a cada tópico. Cada tópico se encuentra tabulado por sus palabras más probables, y el peso de cada una.

### **6.3 Codificación visual e idioma de interacción**

Con el fin de atender a las tareas descritas con anterioridad, se definirá la codificación visual y la interacción para las visualizaciones:

#### **• Codificación visual**

1. Identificar los temas más discutidos en la consulta:

Se han propuesto dos alternativas de visualización que cumplen con criterios de expresividad según la literatura (Mackinlay, 1986). Vale decir, ambas presentan a través de lenguaje gráfico toda y solamente la misma información.

La información a representar en ambas alternativas fue: mostrar cómo una categoría más grande se divide en categorías más pequeñas (tópicos) y cuál es la relación de cada parte sobre la cantidad total (distribución de categorías y tópicos).

Dados dos lenguajes gráficos que expresan una misma información, la pregunta obvia es qué lenguaje involucra un diseño que especifica la presentación con

mejor usabilidad. A diferencia de la expresividad, que sólo depende de la sintaxis y semántica del lenguaje gráfico, la usabilidad también depende del usuario.

### **Alternativa 1: Mapa de árbol o *Tree map***

La relación de cada tópico y cada categoría sobre la cantidad total, se representa a través del área o tamaño de cada cuadrado, donde a mayor área se tiene una mayor importancia. Por otro lado, para codificar cada categoría se asignaron 5 colores de diferentes tonos y se asignó una posición contigua para indicar relación (ver Figura 6-1).

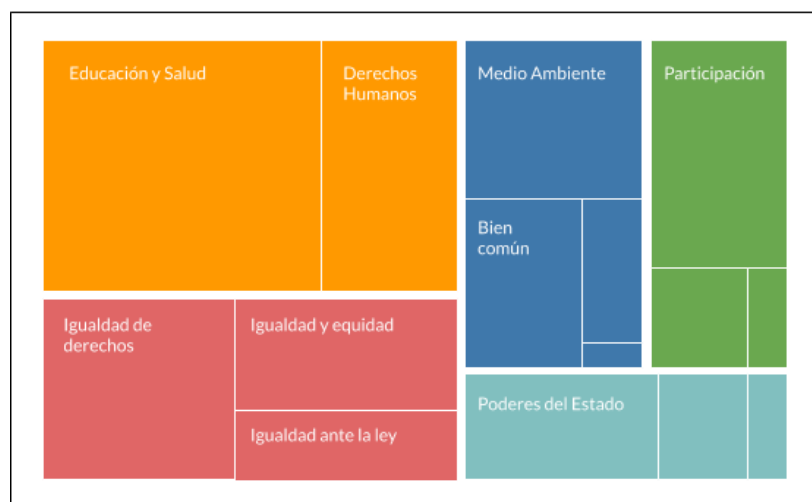


Figura 6-1: Mapa de árbol (Elaboración propia)

### **Alternativa 2: Gráfico de barras apiladas o *Stacked Bar Chart***

La relación de cada tópico y cada categoría sobre la cantidad total, también se representa a través del tamaño, sin embargo, además se codifica por su posición.

Por otro lado, para codificar cada categoría se asignaron los 5 mismos colores y se asignó una posición apilada para indicar relación (ver Figura 6-2).

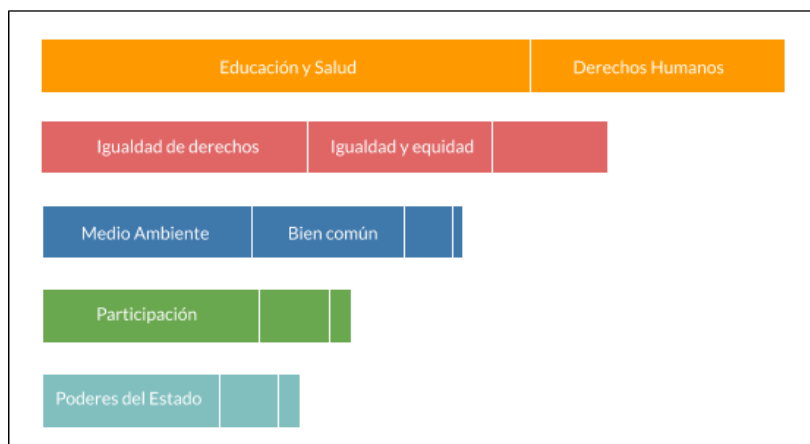


Figura 6-2: Gráfico de barras apiladas (Elaboración propia)

## 2. Descubrir tendencias o patrones de los temas discutidos de manera espacial:

Se escogió un mapa coroplético o *choropleth map*. Este mapa muestra las regiones del país con distinta saturación de color en relación a una variable de los datos, en este caso la importancia de un tópico. Lo anterior, proporciona una forma de visualizar la importancia de cada tópico sobre un área geográfica, que puede mostrar variaciones o patrones en la ubicación visualizada (ver Figura 6-3).

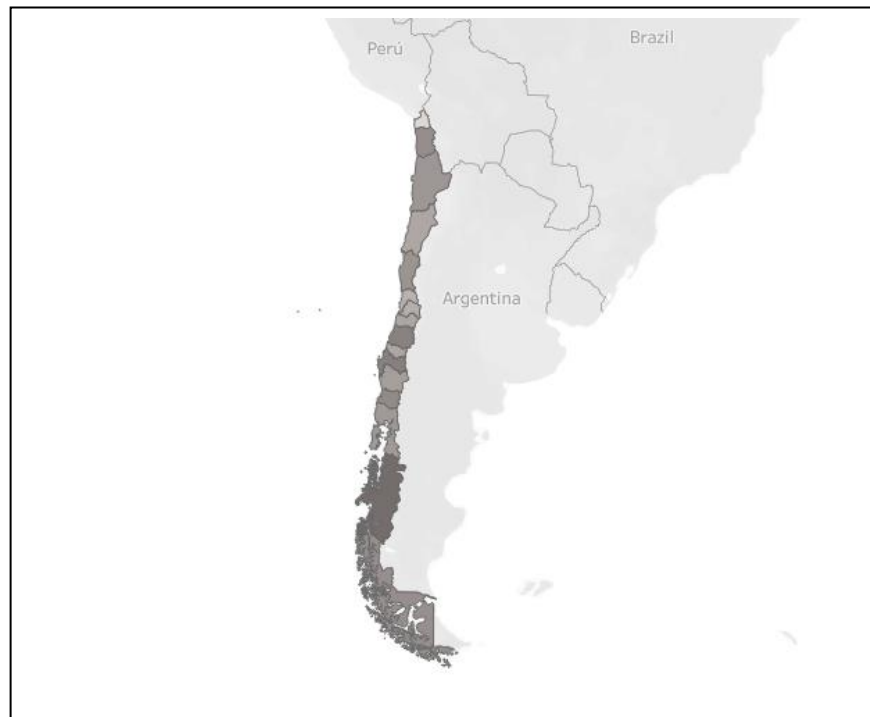


Figura 6-3: Mapa coroplético (Elaboración propia)

3. Revisar opiniones referentes a los temas más discutidos en detalle:

Se escogió un mapa de densidad de puntos. Este mapa es una forma de detectar patrones espaciales o la distribución de datos sobre una región geográfica, colocando puntos de igual tamaño sobre el mapa. El color de cada punto, representa el tópico con mayor probabilidad de esa opinión. Por temas de simplicidad, se escogieron aquellas opiniones cuyo tópico más probable es mayor o igual a 0,75. De este modo, las opiniones se condicen significativamente con el tópico en cuestión. Ver Figura 6-4.

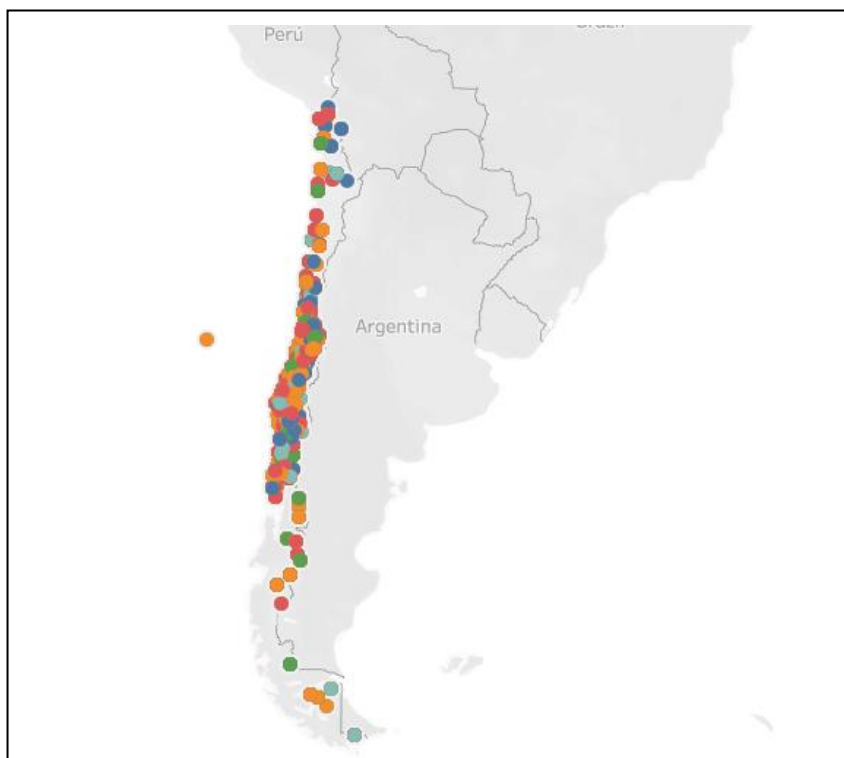


Figura 6-4: Mapa de densidad de puntos (Elaboración propia)

#### • Idioma de interacción

Las tres visualizaciones se presentan en una interfaz de manera yuxtapuesta e interconectadas. Se pueden leer los nombres de los tópicos al desplazarse sobre ellos. Al seleccionar un tópico ambos mapas se filtran según ese tópico. Al seleccionar una región, se muestran solo los datos de esa región. Al pasar el cursor sobre un punto del mapa, una ventana emergente muestra el fundamento normalizado, su comuna y región.

## 6.4 Interfaces de usuario

La interfaz representa información general y particular de las opiniones y su categorización. Permite al usuario (a) identificar la organización y distribución de los tópicos de opiniones, (b) reconocer la distribución espacial de cada tópico por zona geográfica y (c) explorar las opiniones en detalle con referencia geográfica.

Con lo anterior ya definido, se realizó la implementación del sistema descrito. La herramienta utilizada fue Tableau. En la Figura 6-5 se observa la interfaz A conformada por (a) mapa de árbol, (b) mapa coroplético y (c) mapa de puntos.

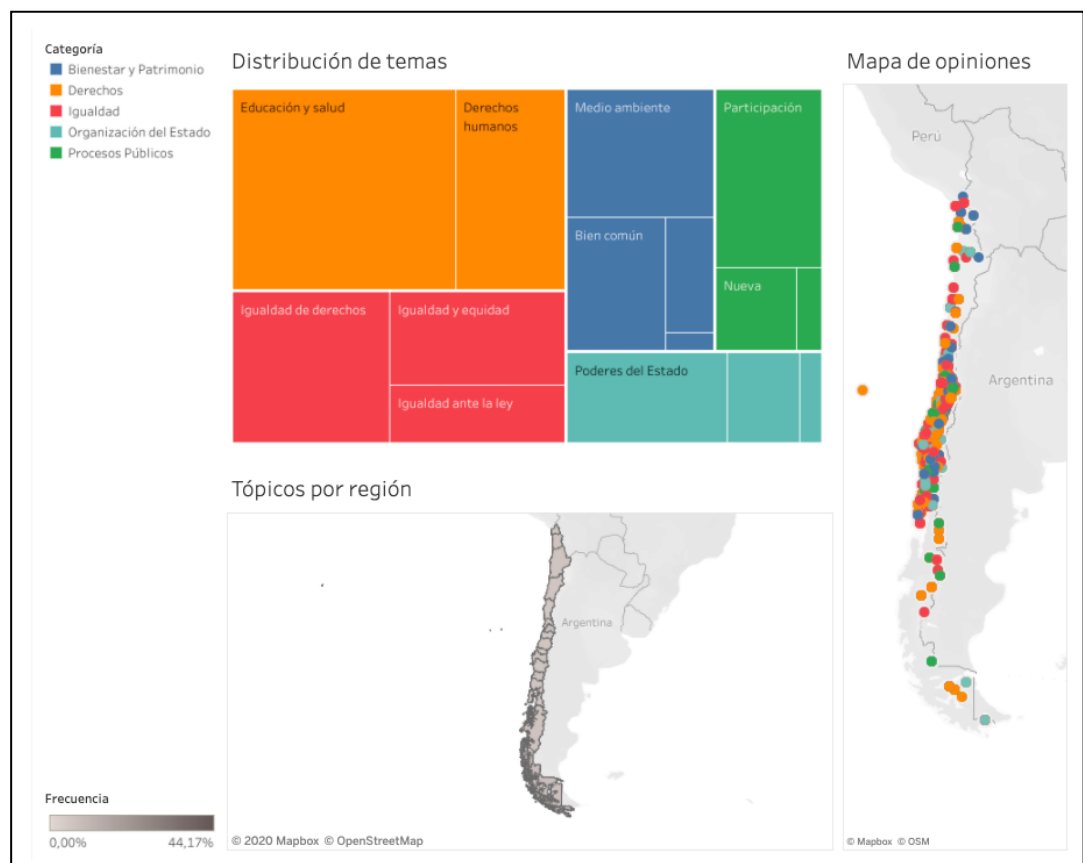


Figura 6-5: Interfaz A - Mapa de árbol

En la Figura 6-6 se observa la interfaz B conformada por (a) barras apiladas, (b) mapa coroplético y (c) mapa de puntos.

La distribución de la interfaz se pensó de manera que el usuario primero interactúe con la vista general, intermedia y luego la particular. Respetando esa distribución, se asignó arriba a la izquierda el mapa de árbol o las barras apiladas, abajo a la izquierda el mapa coroplético y a la derecha el mapa de densidad de puntos. A la izquierda se observan las leyendas de colores.

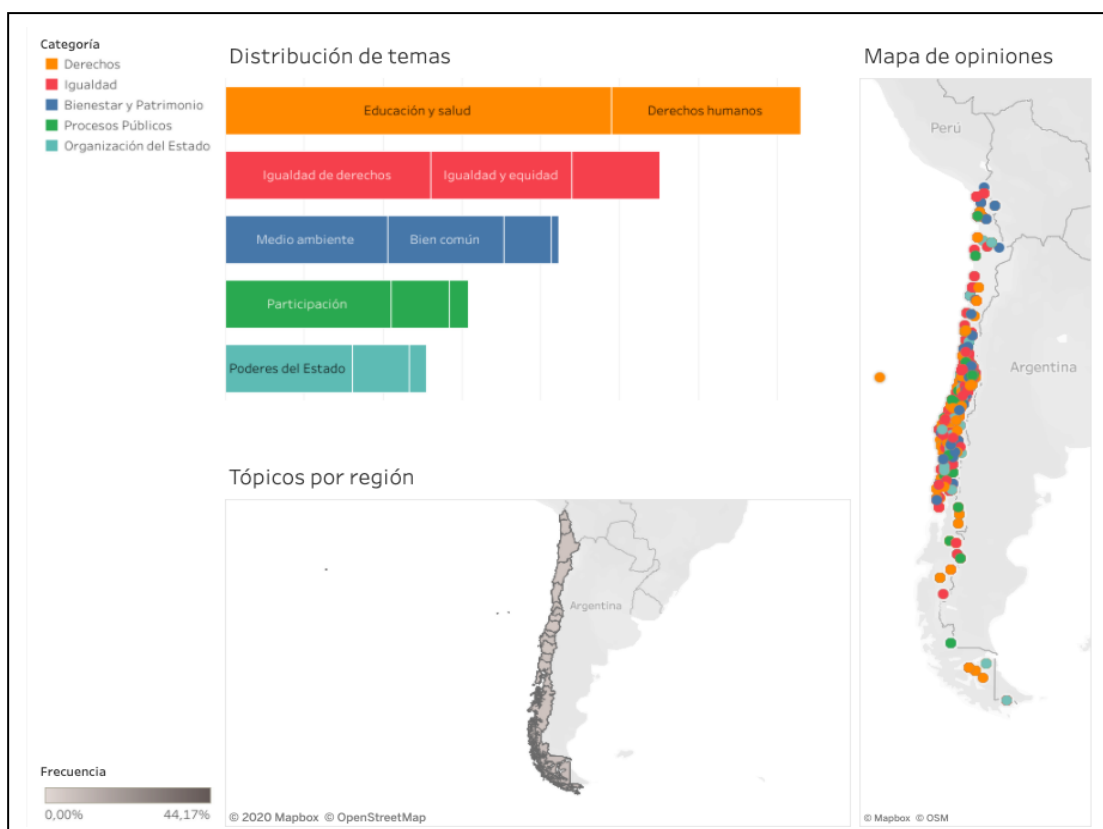


Figura 6-6: Interfaz B - Barras apiladas

## **7. ESTUDIO DE USUARIOS**

### **7.1 Motivación y objetivos**

Este estudio tuvo dos objetivos principales, el primero fue validar la coherencia semántica del modelo implementado. El segundo, fue evaluar la usabilidad de las interfaces con barras apiladas y con mapa de árbol.

En la evaluación de usabilidad se exploraron los efectos de dos factores, tipo de interfaz y tipo de tarea. Ambos, sobre la tasa de éxito de la tarea, el tiempo de completación, y la satisfacción percibida de las tareas.

Se tienen dos tipos de interfaces, la interfaz A con mapa de árbol y la interfaz B con barras apiladas, ambas explicadas en el Capítulo 6. Se tienen dos tipos de tareas, tareas de categorización, y tareas de importancia relativa.

### **7.2 Tareas de usuario**

#### **7.2.1 Validación de coherencia semántica del modelo**

La validación del modelo a través de este estudio de usuarios, busca responder a la primera pregunta de investigación planteada. Para evaluarla, los participantes debieron resolver dos tipos de tareas. Ver tabla 7-1.

Tabla 7-1: Pregunta de investigación 1 y tareas de usuario

| Pregunta                                                                                                       | Tareas de usuario                                   |
|----------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|
| ¿El modelo de tópicos LDA puede lograr una categorización coherente de las opiniones de la consulta ciudadana? | 1. Intrusión de palabras<br>2. Intrusión de tópicos |

La **tarea de intrusión de palabras** busca que el participante seleccione de un listado de palabras, aquella que no encaja con el resto. Vale decir, la palabra intrusa. El listado de palabras incluye las 5 palabras principales del tópico más una palabra con menor probabilidad de pertenecer a ese tópico. Ver Figura 7-1.

Selecciona la palabra que **no** encaja con las demás

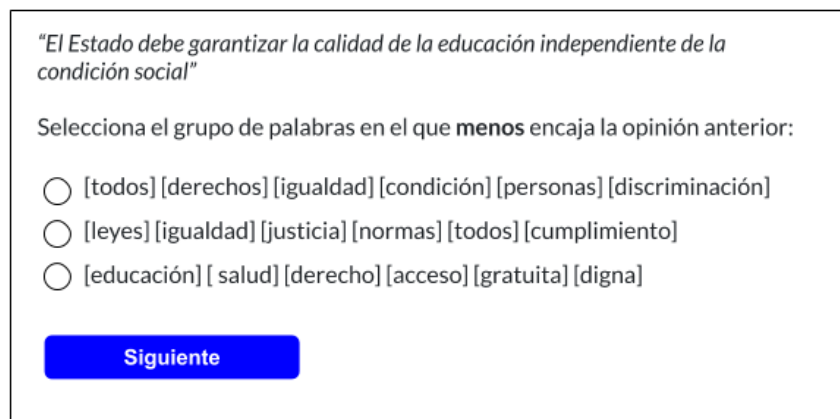
- ☐ Plebiscitos
- ☐ Consultas
- ☐ Naturaleza
- ☐ Decisiones
- ☐ Participación
- ☐ Democracia

**Siguiente**

Figura 7-1: Ejemplo Palabra Intrusa

Para la tarea de palabra intrusa se hicieron tres preguntas, con tres tópicos aleatorios. Una pregunta de dificultad baja, cuya palabra intrusa se encontraba en el percentil 75. Una de dificultad media, cuya palabra intrusa se encontraba en el percentil 50. Una de dificultad alta, cuya palabra intrusa se encontraba en el percentil 25.

La **tarea de intrusión de tópicos** busca que el participante seleccione el tópico que menos relación tenga con cierta opinión dada. Vale decir, el tópico intruso. Se presenta una opinión y dos alternativas de tópicos intrusos. El tópico intruso es parte de los tópicos menos probables de dicha opinión. Ver Figura 7-2.



*"El Estado debe garantizar la calidad de la educación independiente de la condición social"*

Selecciona el grupo de palabras en el que **menos** encaja la opinión anterior:

- ☐ [todos] [derechos] [igualdad] [condición] [personas] [discriminación]
- ☐ [leyes] [igualdad] [justicia] [normas] [todos] [cumplimiento]
- ☐ [educación] [salud] [derecho] [acceso] [gratuita] [digna]

**Siguiente**

Figura 7-2: Ejemplo Tópico Intruso

Para la tarea de tópico intruso se hicieron tres preguntas, con tres opiniones aleatorias. Una pregunta de dificultad baja, cuyo tópico intruso tenía un 2%

de probabilidad de relacionarse con la opinión. Una de dificultad media, cuya probabilidad era de aproximadamente 6%. Una de dificultad alta, cuya probabilidad era aproximadamente del 10%.

### 7.2.2 Evaluación de usabilidad de interfaces de visualización

La evaluación de herramientas de analítica visual, o interfaces, busca responder a las siguientes preguntas de investigación a través de las tareas señaladas en la Tabla 7-2.

Tabla 7-2: Pregunta de investigación 2 y tareas de usuario

| Pregunta                                                                                | Tareas de usuario                                              |
|-----------------------------------------------------------------------------------------|----------------------------------------------------------------|
| P2.1 ¿Qué interfaz es mejor para realizar análisis de categorización entre los tópicos? | 1. Categorización de tópicos<br>2. Categorización de opiniones |
| P2.2 ¿Qué interfaz es mejor para realizar análisis de comparación entre los tópicos     | 1. Comparación general<br>2. Comparación entre regiones        |

La **tarea de categorización** tiene dos variantes, una tarea de categorización de tópicos, y otra de categorización de opiniones. La primera está enfocada en el uso de la visualización de mapa de árbol o barras apiladas (Figura 7-3). Se espera que el participante identifique la correcta categorización de los tópicos.

¿Qué tópico forma parte de la misma categoría que **Derechos Humanos**?

☐ Igualdad de derechos

☐ Educación y salud

**Siguiete**

Figura 7-3: Ejemplo de tarea de categorización de tópicos

La segunda está enfocada en el uso complementario entre mapa de árbol o barras apiladas, y mapa de densidad de puntos. Se espera que el participante identifique correctamente a qué categoría corresponde cierta opinión.

La **tarea de comparación** tiene dos variantes, una de comparación entre tópicos a nivel general, y otra de comparación entre tópicos a nivel regional. La primera está enfocada en la visualización de mapa de árbol o barras

apiladas (Figura 7-4). Se espera que el participante sea capaz de distinguir qué tópico fue más predominante que otro.

Educación y Salud      Derechos Humanos

Igualdad de derechos      Igualdad y equidad

Medio Ambiente      Bien común

Participación

Poderes del Estado

¿Cuál es el tópico **más** mencionado entre "Nueva Constitución" y "Patrimonio Cultural"?

☐ Nueva Constitución

☐ Patrimonio Cultural

**Siguiete**

Figura 7-4: Ejemplo de tarea de comparación

La segunda está enfocada en el uso complementario entre mapa de árbol o barras apiladas, y mapa coroplético. Se espera que el participante logre distinguir qué tópico es más predominante que otro en una región en particular.

### 7.3 Métricas

Las métricas consideradas para realizar la evaluación de usabilidad en este estudio incluyen medidas de desempeño (tasa de éxito de la tarea, tiempo de completación de la tarea), así como medidas auto reportadas (satisfacción percibida) (Sørum, Andersen y Vatrapu, 2012).

- **Tasa de éxito de la tarea**

Corresponde a una medida de efectividad. Se asignó un 1 cada vez que un participante logró completar una tarea correctamente y se asignó un 0 por responder incorrectamente o no completar la tarea. Se calculó automáticamente tras responder a una tarea.

- **Tiempo de completación de la tarea**

Corresponde a una medida de eficiencia. Se consideró el tiempo transcurrido entre que el usuario ingresa a la tarea y pasa a la siguiente. Se expresa en segundos. Se tomó automáticamente tras responder una tarea.

- **Nivel de Satisfacción percibida durante la tarea**

Corresponde a una medida de satisfacción. A través de tres preguntas, se calificó la facilidad de una tarea, la cantidad de tiempo que la tarea tardó en completarse y el nivel de apoyo recibido durante todo el proceso. Se recopiló utilizando las siguientes preguntas del cuestionario After Scenario (Lewis, 1995).

1. En general, estoy satisfecho con la facilidad de completar las tareas en esta interfaz.
2. En general, estoy satisfecho con la cantidad de tiempo que tomó completar las tareas en esta interfaz.
3. En general, estoy satisfecho con la información entregada para completar las tareas.

Cada una se evaluó con una escala Likert de calificación de 7 puntos (1 = Totalmente en desacuerdo 7 = Totalmente de acuerdo).

#### 7.4 Participantes

Para determinar el número de participantes, o tamaño de la muestra, se utilizó el *software* GPower a través de un *Power Analysis* (Mayr et al., 2007). Se consideró un test ANOVA. Este test compara las medias de dos mediciones para determinar si es estadísticamente significativa la diferencia entre las medias (*F-test/ ANOVA: Repeated measures, within factors*).

Se obtuvo un tamaño de la muestra a estudiar de 34 personas a través de un análisis *a priori* y los siguientes parámetros elegidos (Figura 7-5),

- Nivel de significancia ( $\alpha$ ) = .05
- Poder estadístico ( $1 - \beta$ ) = .8

- Tamaño del efecto ( $d$ ) = .25- efecto pequeño (Cohen, 1988)

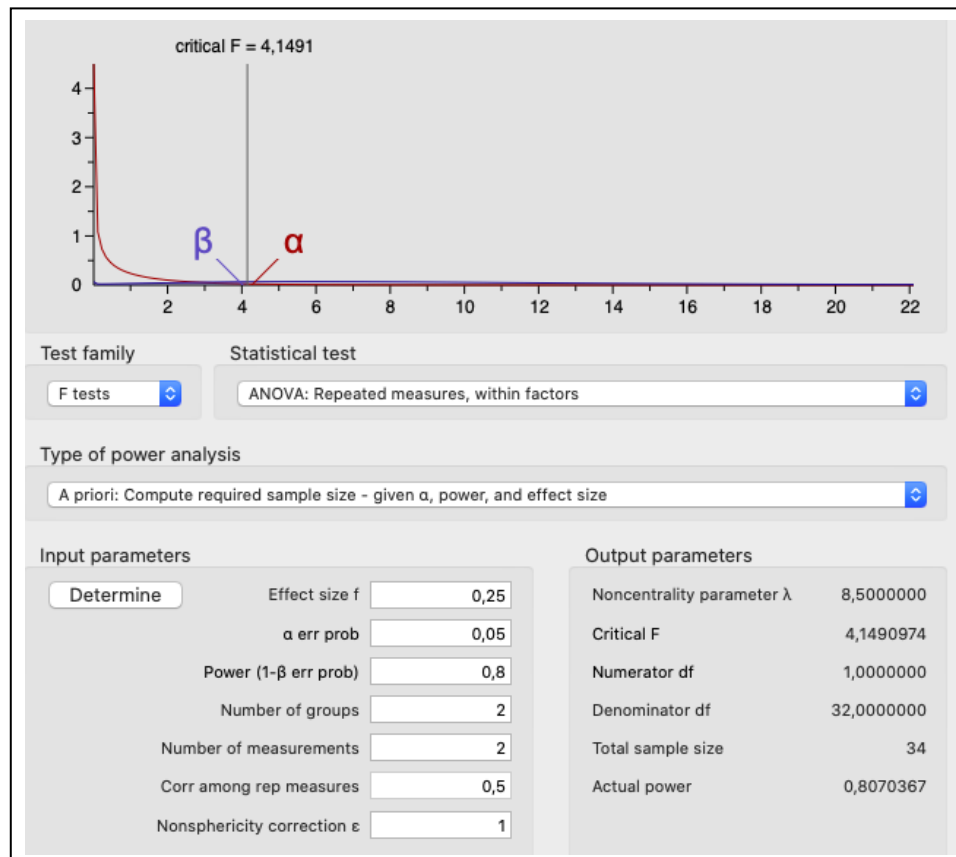


Figura 7-5: Parámetros utilizados para Power Analysis de GPower

La edad de los participantes reclutados se encontraba entre 25 y 30 años. Se limitó a este rango de edad debido a su ventaja en competencias digitales y para evitar efectos asociados con la edad (Hanson, 2011). La edad promedio de los participantes fue de 27 años. En cuanto al sexo de los participantes, 19 (56%) eran mujeres y 15 (44%) eran hombres. Se consideró como criterio de exclusión tener

un manejo o experiencia con visualizaciones de datos, mayor a 5 en una escala de 1 a 10.

Todos los participantes utilizaron el mismo equipo y realizaron las tareas asignadas en un tiempo de 30 minutos. Ninguno de los participantes recibió compensación económica. Se pidió a los participantes que firmaran un formulario de consentimiento y se les informó que podían abandonar el estudio en cualquier momento.

### **7.5 Diseño experimental**

Se utilizó un diseño intra-sujeto de 4 tipos de tareas x 2 tipos de interfaz. Cada participante utilizó los dos tipos de interfaz explicados en el Capítulo 6. Combinando cada interfaz con cada tipo de pregunta, se originan 8 preguntas a responder por usuario.

Se realizaron dos versiones para contrarrestar la fatiga cognitiva y para controlar los efectos de aprendizaje asociados (MacKenzie, 2013). La Tabla 7-3 muestra el Cuadrado Latino de 2x2 utilizado para determinar el orden de las interfaces utilizadas por los participantes en el estudio.

Tabla 7-3: Cuadrado latino 2x2

| <b>Versión</b> | <b>Interfaz 1</b>                                                          | <b>Interfaz 2</b>                                                          |
|----------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| 1              | A: mapa de árbol + mapa coroplético + mapa densidad de puntos              | B: gráfico de barras apiladas + mapa coroplético + mapa densidad de puntos |
| 2              | B: gráfico de barras apiladas + mapa coroplético + mapa densidad de puntos | A: mapa de árbol + mapa coroplético + mapa densidad de puntos              |

A cada participante se le asignó la versión 1 o 2, según el orden de su disponibilidad. Es decir, al primer participante se le asignó la versión 1, al segundo la versión 2 y así sucesivamente. Las preguntas de cada versión se detallan en la Tabla 7-4:

Tabla 7-4: Preguntas por tarea y versión

| <b>Id</b> | <b>Tipo de tarea</b>        | <b>Versión 1</b>                                                                                                                                                                                                                                                                                       | <b>Versión 2</b>                                                                                                                                                                                    |
|-----------|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1         | Categorización de tópicos   | ¿Qué tópico forma parte de la misma categoría que “Bien común”?                                                                                                                                                                                                                                        | ¿Qué tópico forma parte de la misma categoría que “Derechos humanos”?                                                                                                                               |
| 2         | Categorización de opiniones | ¿A qué categoría pertenece la siguiente opinión de Juan Fernández? “Las relaciones cívicas propositivas se consideran esenciales para la protección de los derechos humanos en nuestra sociedad permanentemente, para así lograr la no discriminación por opiniones divergentes, políticas, de género” | ¿A qué tópico podría pertenecer la siguiente opinión de Cabo de Hornos? “Creación de leyes y normativas regionales debe posibilitar la administración de recursos propios en forma descentralizada” |
| 3         | Comparación general         | ¿Cuál es el tópico más frecuente entre “Poderes del Estado” y “Medio Ambiente”?                                                                                                                                                                                                                        | ¿Cuál es el tópico más frecuente entre “Nueva Constitución” y “Patrimonio Cultural”?                                                                                                                |

|   |                             |                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                                                        |
|---|-----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4 | Comparación por región      | En la región de Antofagasta, ¿cuál es el tópico más frecuente entre "Igualdad ante la Ley" y "Participación"?                                                                                       | En la región de Los Lagos, ¿cuál es el tópico más frecuente entre "Congreso" y "Transparencia"?                                                                                                                                                                                                        |
| 5 | Categorización de tópicos   | ¿Qué tópico forma parte de la misma categoría que “Derechos humanos”?                                                                                                                               | ¿Qué tópico forma parte de la misma categoría que “Bien común”?                                                                                                                                                                                                                                        |
| 6 | Categorización de opiniones | ¿A qué tópico podría pertenecer la siguiente opinión de Cabo de Hornos? "Creación de leyes y normativas regionales debe posibilitar la administración de recursos propios en forma descentralizada" | ¿A qué categoría pertenece la siguiente opinión de Juan Fernández? “Las relaciones cívicas propositivas se consideran esenciales para la protección de los derechos humanos en nuestra sociedad permanentemente, para así lograr la no discriminación por opiniones divergentes, políticas, de género” |
| 7 | Comparación general         | ¿Cuál es el tópico más frecuente entre “Nueva Constitución” y “Patrimonio Cultural”?                                                                                                                | ¿Cuál es el tópico más frecuente entre “Poderes del Estado” y “Medio Ambiente”?                                                                                                                                                                                                                        |
| 8 | Comparación por región      | En la región de Los Lagos, ¿cuál es el tópico más frecuente entre "Congreso" y "Transparencia"?                                                                                                     | En la región de Antofagasta, ¿cuál es el tópico más frecuente entre "Igualdad ante la Ley" y "Participación"?                                                                                                                                                                                          |

## 7.6 Procedimiento

El estudio se realizó a través de un cuestionario online, alojado en una aplicación web node.js que se desarrolló especialmente para ello. Se agendó con cada participante, una sesión de 30 minutos con asistencia remota vía Zoom. El participante accedió al cuestionario a través de un sitio web.

El cuestionario estaba formado por tres partes como indica la Figura 7-6. La primera parte correspondió al consentimiento informado y toma de datos de los usuarios. La segunda, a la evaluación de coherencia semántica del modelo. La tercera, a la evaluación de usabilidad de las herramientas de analítica visual.

- Parte 0 - Introducción

Explicación general del contexto y objetivo del estudio. Además, de toma de datos de los usuarios para el análisis posterior (género, edad, ocupación y experiencia con visualizaciones).

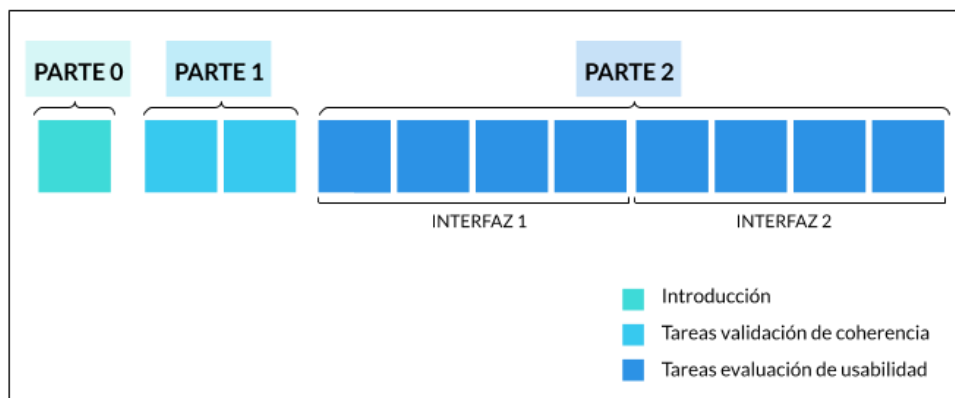


Figura 7-6: Estructura cuestionario

- Parte 1 - Evaluación de coherencia semántica del modelo

Se detalla la explicación de las tareas a resolver ejemplificando el tipo de pregunta a responder a través de un ejercicio de prueba. Se realizan tres preguntas de selección múltiple.

- Parte 2 - Evaluación de usabilidad de herramientas de analítica visual

Se detalla la explicación de las tareas a resolver ejemplificando el tipo de pregunta a responder a través de un ejercicio de prueba. Se realizan 8 preguntas de selección múltiple (4 por cada interfaz). Después de cada interfaz se debe responder el cuestionario de satisfacción percibida.

La navegación seguía un orden establecido y se explicitó que no había tiempo límite para responder. Cada sección recogía de manera automática por cada usuario, las respuestas y el tiempo, en un archivo .csv.

## 8. RESULTADOS Y ANÁLISIS

### 8.1 Validación de coherencia semántica

Para evaluar la coherencia semántica, se siguió la metodología de intrusión de palabras. Se solicitó a los participantes identificar la palabra intrusa; si el tópico era coherente, debió ser fácil detectar la palabra fuera de lugar. Si no lo era, es probable que los participantes adivinaran y eligieran incorrectamente.

Del total de 34 participantes, 30 participantes acertaron en la pregunta de baja dificultad; 24 participantes acertaron en la de dificultad media; y 18 acertaron en la de dificultad alta. En la Figura 8-1 se muestran los resultados.

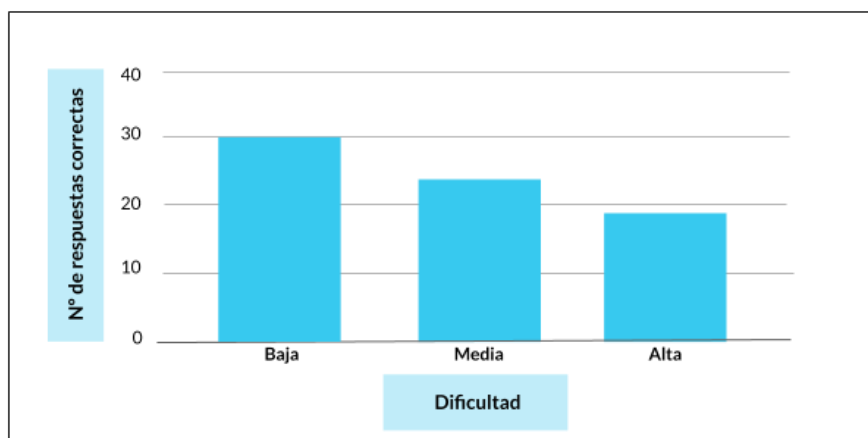


Figura 8-1: Gráfico de respuestas correctas según dificultad - intrusión de palabras

Chang et al. (2009) propusieron la precisión del modelo como una forma de evaluar la tarea de intrusión de palabras, mediante la cual simplemente calculan la

proporción de participantes que seleccionaron correctamente la palabra intrusa para un tópico determinado. En la Tabla 8-1 se resumen los resultados.

Tabla 8-1: Resultados Precisión del modelo – Palabra intrusa

|                  |      |
|------------------|------|
| Dificultad baja  | 0.89 |
| Dificultad media | 0.71 |
| Dificultad alta  | 0.53 |
| Promedio         | 0.71 |

Para precisión del modelo, los valores posibles están entre 0 y 1, donde el mejor resultado posible es 1. De lo anterior se desprende que se tiene mejor desempeño en preguntas de menor dificultad (0.89) que en aquellas de mayor dificultad (0.53). La media de la precisión del modelo cumple con criterios suficientes de calidad ( $> 0.6$ ) según Bhatia, Lau y Baldwin (2018).

En la segunda tarea se solicitó a los participantes identificar el tópico intruso; si el tópico era coherente con la opinión, debió ser fácil detectar el tópico fuera de lugar. Si no lo era, es probable que los participantes adivinaran y eligieran incorrectamente.

Del total de 34 participantes, en la pregunta 1, 29 participantes acertaron en la pregunta de baja dificultad; 25 participantes acertaron en la de dificultad media; y

12 participantes en la de dificultad alta. En la Figura 8-2 se muestran los resultados.

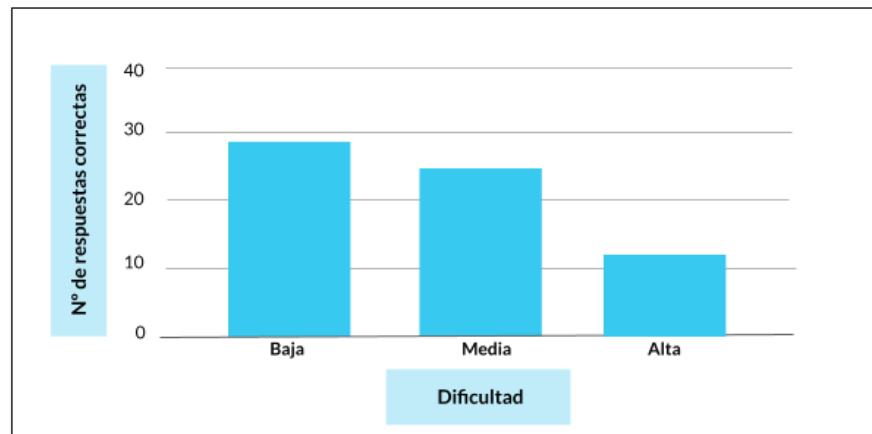


Figura 8-2: Gráfico de respuestas correctas según dificultad - intrusión de tópicos

Chang et al. (2009) propuso topic log odds ("TLO") como un medio para evaluar la tarea de intrusión de tópicos. Los autores definieron topic log odds como el logaritmo del ratio entre la proporción del tópico intruso y la proporción del no intruso. En la Tabla 8-2 se resumen los resultados.

Para topic log odds, el límite superior es 0. Esto se logra cuando los participantes escogen un intruso con proporción no superior al del intruso real. En este caso vemos que, a mayor dificultad, mejor topic log odds.

Tabla 8-2: Resultados Topic log odds - Intrusión de tópicos

|                  |       |
|------------------|-------|
| Dificultad baja  | -1.06 |
| Dificultad media | -0.46 |
| Dificultad alta  | -0.06 |
| Promedio         | -0.53 |

## 8.2 Evaluación de usabilidad

Este estudio utilizó un diseño intra-sujetos para explorar los efectos de dos factores (tipo de tarea y tipo de interfaz) en la tasa de éxito, el tiempo de completación, y la satisfacción percibida. Dada la configuración del experimento, se midieron las diferencias de media utilizando análisis de varianza de medidas repetidas de dos vías.

No hay valores atípicos significativos en ninguna celda del diseño. La variable dependiente cumple con el supuesto de normalidad según el test de Shapiro-Wilk. La prueba de Mauchly mostró que los datos cumplían con el supuesto de esfericidad. Cuando los valores fueron significativos, se realizaron comparaciones post hoc con ajuste de Bonferroni. Los datos experimentales se analizaron con R Studio.

La tabla 8-3 proporciona las estadísticas descriptivas de los resultados. Muestra las medias y las desviaciones estándar de las medidas de rendimiento en cuatro tipos de tareas (categorización de tópicos, categorización de opiniones, comparación general y comparación entre regiones) y dos tipos de interfaces (A: con mapa de árbol y B: con barras apiladas).

Tabla 8-3: Medias (M) y Desviación estándar (SD) de tasa de éxito, tiempo de completación y satisfacción percibida (N=34)

| Tipo de tarea               | Tipo de interfaz | Tasa de éxito |           | Tiempo completación |           | Satisfacción percibida |           |
|-----------------------------|------------------|---------------|-----------|---------------------|-----------|------------------------|-----------|
|                             |                  | <i>M</i>      | <i>SD</i> | <i>M</i>            | <i>SD</i> | <i>M</i>               | <i>SD</i> |
| Categorización de tópicos   | Mapa de árbol    | 1.00          | 0.00      | 13.43               | 6.10      | 6.58                   | 0.53      |
|                             | Barras apiladas  | 0.97          | 0.17      | 14.65               | 4.76      | 6.22                   | 0.56      |
| Categorización de opiniones | Mapa de árbol    | 0.79          | 0.41      | 33.03               | 10.72     | 6.29                   | 0.98      |
|                             | Barras apiladas  | 0.71          | 0.46      | 31.73               | 11.36     | 6.13                   | 0.95      |
| Comparación general         | Mapa de árbol    | 0.38          | 0.49      | 38.29               | 8.90      | 3.92                   | 1.16      |
|                             | Barras apiladas  | 0.88          | 0.33      | 18.85               | 3.92      | 5.25                   | 0.58      |
| Comparación por región      | Mapa de árbol    | 0.44          | 0.50      | 43.13               | 7.24      | 4.12                   | 0.67      |
|                             | Barras apiladas  | 0.85          | 0.36      | 22.25               | 5.68      | 5.53                   | 0.88      |

La tasa de éxito indica el número de respuestas correctas sobre el total de respuestas. El tiempo de completación indica el tiempo (en segundos) empleado por los participantes para completar una tarea específica usando una interfaz en particular. La satisfacción percibida indica el nivel de comodidad que sintieron los

participantes al completar una tarea en particular usando una interfaz específica (1 a 7).

### **8.2.1 Efecto de tipo de tarea y tipo de interfaz en tasa de éxito**

El ANOVA de dos vías de medidas repetidas, mostró un efecto principal del tipo de tarea,  $F(3,58) = 41.935$  ( $p < .001$ ), tipo de interfaz,  $F(1,58) = 14.642$  ( $p < .001$ ) y una interacción entre el tipo de tarea y tipo de interfaz,  $F(3,58) = 6.922$  ( $p < .001$ ), en la tasa de éxito.

Las tasas medias de éxito del gráfico de barras apiladas fueron las más altas en las tareas de comparación ( $M = 0.88$ ,  $SD = 0.33$  y  $M = 0.85$ ,  $SD = 0.36$ ) lo que indica que los participantes se desempeñaron mejor en las tareas de comparación general y según región, cuando usaron el gráfico de barras apiladas (ver Figura 8-3).

Dado el efecto de interacción entre el tipo de interfaz y tipo de tarea en la tasa de éxito, se hicieron comparaciones post hoc. Estas comparaciones confirmaron una diferencia significativa en la tasa de éxito de ambas interfaces para las tareas de comparación general ( $p < .001$ ) y para las tareas de comparación por región ( $p < .001$ ).

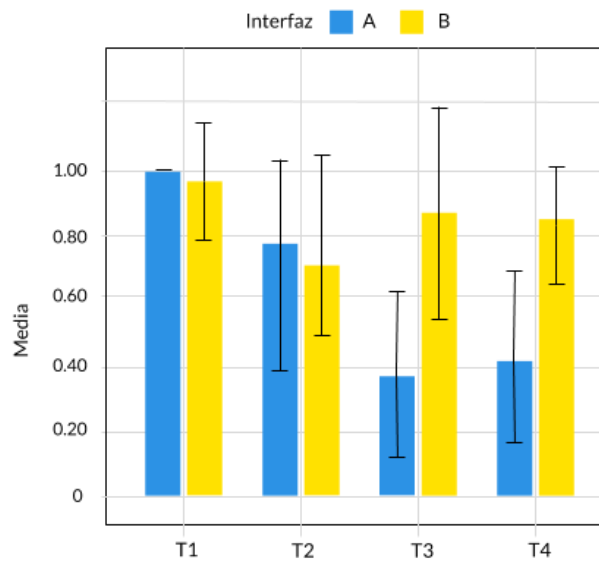


Figura 8-3: Medias de tasa de éxito de cada tipo de interfaz para tareas de categorización y comparación

No se encontraron diferencias significativas de la tasa de éxito entre los tipos de interfaz para las tareas de categorización. Este resultado sugiere que, para las tareas de categorización de tópicos y opiniones, ambas interfaces ofrecieron un rendimiento similar con respecto a la tasa de éxito.

### 8.2.2 Efecto de tipo de tarea y tipo de interfaz en tiempo de completación

El ANOVA de dos vías de medidas repetidas, mostró un efecto principal del tipo de tarea,  $F(3,58) = 7.421$  ( $p < .001$ ), tipo de interfaz,  $F(1,58) = 13.773$  ( $p < .001$ ) y una interacción entre tipo de tarea y el tipo de interfaz,  $F(3,58) = 6.340$  ( $p < .001$ ), en el tiempo de completación de la tarea.

Los tiempos medios de completación del gráfico de barras apiladas fueron los más bajos en las tareas de comparación ( $M = 18.85$ ,  $SD = 3.92$  y  $M = 22.25$ ,  $SD = 5.68$ ) lo que indica que los participantes se desempeñaron mejor en las tareas de comparación general y según región, cuando usaron el gráfico de barras apiladas (ver Figura 8-4).

Dado el efecto de interacción entre el tipo de interfaz y tipo de tarea en el tiempo de completación, se hicieron comparaciones post hoc. Estas comparaciones confirmaron una diferencia significativa en el tiempo de completación de ambas interfaces para las tareas de comparación general ( $p = .002$ ) y para las tareas de comparación por región ( $p = .001$ ).

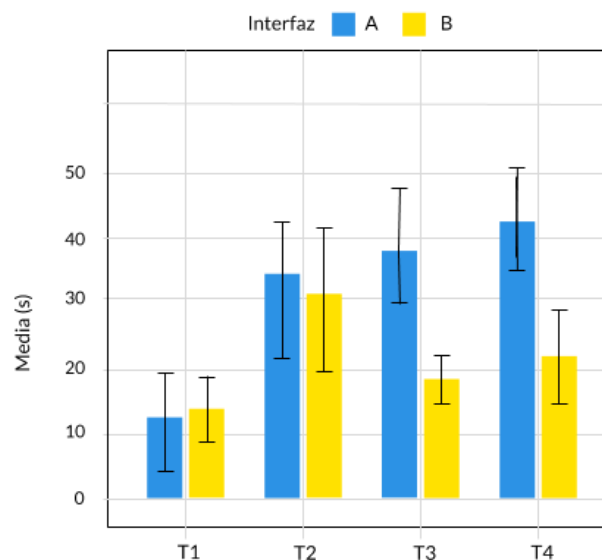


Figura 8-4: Tiempos medios de completación de cada tipo de interfaz para tareas de categorización y comparación

No se encontraron diferencias significativas de la tasa de completación entre los tipos de interfaz para las tareas de categorización. Este resultado sugiere que, para las tareas de categorización de tópicos y opiniones, ambas interfaces ofrecieron un rendimiento similar con respecto al tiempo de completación.

### **8.2.3 Efecto de tipo de tarea y tipo de interfaz en satisfacción percibida**

El ANOVA de dos vías de medidas repetidas, mostró un efecto principal del tipo de tarea,  $F(3,58) = 4.026$  ( $p = .011$ ) y tipo de interfaz,  $F(1,58) = 12.001$  ( $p = .001$ ) y una interacción entre el tipo de tarea y el tipo de interfaz,  $F(3,58) = 5.107$  ( $p = .003$ ), en la satisfacción percibida de la tarea.

Para las tareas de comparación, la satisfacción media percibida tras utilizar el gráfico de barras apiladas fueron las más altas ( $M = 5.25$ ,  $SD = 0.58$  y  $M = 5.53$ ,  $SD = 0.88$ ). Este resultado sugiere que los participantes percibieron que el gráfico de barras apiladas es el más fácil de usar para completar comparaciones generales y por región (ver Figura 8-5).

Dado el efecto de interacción entre el tipo de interfaz y tipo de tarea en la satisfacción percibida, se hicieron comparaciones post hoc. Estas comparaciones confirmaron una diferencia significativa en el nivel de

satisfacción de ambas interfaces para las tareas de comparación general ( $p = .024$ ) y para las tareas de comparación por región ( $p = .029$ ).

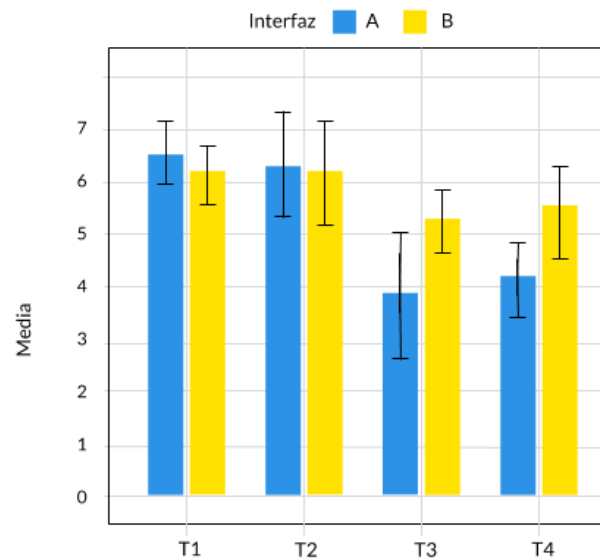


Figura 8-5: Satisfacción media percibida de cada tipo de interfaz para tareas de categorización y comparación

Si bien la satisfacción media percibida en las tareas de categorización es mayor en el mapa de árbol, no se encontraron diferencias significativas del nivel de satisfacción percibida entre los tipos de interfaz para las tareas de categorización. Este resultado sugiere que, para las tareas de categorización de tópicos y opiniones, ambas interfaces ofrecieron un confort similar.

## 9. DISCUSIÓN

Teniendo en cuenta la necesidad de transparencia, rendición de cuentas y comprensión del público, visualizar los datos de participación pública sin evidencia de usabilidad no es suficiente. Por un lado, estudios recientes han demostrado la eficacia de modelar y visualizar la opinión pública, pero principalmente en entornos informales y redes sociales (Steinskog, Therkelsen & y Gambäck, 2017; Yu, Xu, Wang & Ni, 2019; Kant, Weisser & Säfken, 2020). Por otro lado, la evaluación de las estrategias de visualización a menudo ha carecido de una operacionalización más completa de la usabilidad y se ha centrado principalmente en medir la efectividad.

En este trabajo se estudió la coherencia semántica del modelo de tópicos LDA que utilizó los datos de la consulta ciudadana del proceso constituyente 2016 en Chile. Se calculó la coherencia a través de las métricas de precisión del modelo y topic log odds obtenidas de tareas de intrusión de palabras e intrusión de tópicos con usuarios. Ambas métricas resultaron favorables para el modelo en cuestión.

Por otro lado, se estudió la usabilidad de dos interfaces de visualización diferentes utilizando datos del modelo de tópicos validado. Se buscó comparar la usabilidad del mapa de árbol, que es la técnica más utilizada, y el gráfico de barras apiladas, que es conceptualmente más apto para el análisis cuantitativo. Se realizó un ANOVA de medidas repetidas de dos vías para determinar si había diferencias significativas entre las dos interfaces en términos de tasa de éxito, tiempo de completación y satisfacción

percibida cuando se utilizaban para dos tareas cognitivas diferentes (categorización de datos y comparación). Este enfoque se utilizó para determinar si la usabilidad de las dos interfaces era similar o diferente cuando se usaba para resolver diferentes tareas.

Los resultados del ANOVA mostraron que existe un efecto de interacción significativo entre el tipo de tarea y la interfaz en las tres dimensiones de usabilidad. Esto sugiere que las interfaces producen resultados diferentes, pero solo en determinadas tareas. Para saber qué interfaz produce mejores resultados en qué tarea, se utilizó una comparación por pares. Este análisis mostró que el gráfico de barras apiladas produce mejores resultados en todas las dimensiones de usabilidad, pero solo al resolver las tareas 3 y 4, que son tareas de comparación. Las tareas 1 y 2 (tareas de categorización) no mostraron diferencias significativas.

Los hallazgos apoyan la hipótesis de que el gráfico de barras apiladas es una mejor alternativa a la visualización del mapa de árbol para tareas cognitivas que involucran comparación. Esto puede explicarse por el hecho de que es más difícil para el ojo humano hacer comparaciones precisas cuando los datos están codificados a través del área de una figura (Few, 2009), como lo hace el mapa de árbol. Por otro lado, la ventaja del gráfico de barras apiladas puede estar respaldada por la jerarquía de precisión de visualización (Mackinlay, 1986), donde el uso de la posición y la longitud para comunicar información es la forma más precisa de mostrar datos cuantitativos.

En este sentido, estos hallazgos apoyan la idea de que la usabilidad de las estrategias de visualización es específica de la tarea y, por lo tanto, la evaluación de visualizaciones de participación ciudadana debe dirigirse a operaciones cognitivas específicas. Se necesita más debate e investigación para determinar qué tareas cognitivas se pueden utilizar como estándares de éxito para los datos de participación pública.

En resumen, la contribución de esta investigación es triple. En primer lugar, validamos el modelo LDA para textos de opinión cortos y en español a través de su coherencia semántica. En segundo lugar, proporcionamos evidencia de la usabilidad de dos técnicas de visualización utilizando datos de un ejercicio de participación ciudadana masiva. Nuestros resultados mostraron que la visualización del gráfico de barras apiladas produce mejores resultados para las tareas de comparación y que las tareas de categorización no muestran diferencias significativas. Según estos resultados, el gráfico de barras apiladas es más apropiado para una visualización que busca categorizar y a la vez comparar datos cuantitativos. En segundo lugar, nuestra investigación indica que el rendimiento y la usabilidad de las estrategias de visualización son específicas de la tarea y, por lo tanto, no se debe realizar una evaluación de usabilidad genérica. A su vez, esta conclusión se suma a la complejidad de devolver información al público, ya que las estrategias de visualización tienen que dar cuenta de las operaciones cognitivas esperadas de sus usuarios finales.

Esta investigación se enfocó en validar un modelo reconocido, como lo es el LDA, en datos de opinión pertenecientes a un ejercicio de participación ciudadana masivo. Sin embargo, no se consideraron otros modelos para poder elegir el mejor entre ellos. En la literatura se han desarrollado otros modelos que podrían ser útiles pero que no han sido validados con este tipo de datos aún. Sería beneficioso para el campo del procesamiento de lenguaje natural y de las políticas públicas validar nuevos modelos más precisos.

Este estudio se centró en el enfoque cuantitativo para evaluar la usabilidad de las interfaces. Sin embargo, los datos experimentales no capturaron la experiencia subjetiva del usuario al realizar las tareas. Sería útil poder triangular los resultados experimentales con información cualitativa para comprender mejor cómo el público representa el manejo de la información y cómo se relaciona el tema de la información con la tarea cognitiva.

Por otro lado, los participantes tenían cierto nivel de conocimientos informáticos, ya que todos eran profesionales recién graduados. El conocimiento previo y la edad pueden influir en los resultados del experimento. Por lo tanto, es posible que los hallazgos del estudio no se puedan generalizar fácilmente a los usuarios que no utilizan computadoras y a los usuarios mayores. La investigación futura debe abordar cómo las características de los participantes y cómo el tema de la información impacta la usabilidad de las interfaces visuales.

## **10. CONCLUSIONES**

Este estudio evaluó la usabilidad de dos interfaces de visualización de tópicos: una que incluyó gráfico de barras apiladas y otra que incluyó mapa de árbol, ambas para respaldar tareas de categorización y comparación de importancia relativa entre los tópicos del modelo entrenado. Se utilizó la tasa de éxito, el tiempo de completación, y la satisfacción percibida como medidas de usabilidad.

Los resultados experimentales mostraron que el gráfico de barras apiladas permitió a los participantes completar las comparaciones en la menor cantidad de tiempo, con hasta incluso 21 segundos de diferencia, con un 50% de mayor precisión y con 1,4 puntos más de satisfacción que el mapa de árbol.

Se atribuye una mejor usabilidad del gráfico de barras apiladas a dos factores. En primer lugar, la representación visual de los datos tenía un significado coherente con respecto a la tarea del usuario en cuestión. En segundo lugar, esta coherencia redujo el esfuerzo de evaluar con precisión y rapidez.

Destacamos cómo estos hallazgos podrían contribuir al diseño de herramientas de visualización de tópicos útiles. Resumir, categorizar y visualizar datos de consultas ciudadanas permite una mayor visibilidad de la opinión pública. El promover estas prácticas puede incentivar la participación ciudadana y su inclusión en la toma de

decisiones del Estado. Se podría lograr con ello, aumentar la confianza y representatividad de la ciudadanía.

Hay varias direcciones para dar seguimiento a esta investigación y abordar sus limitaciones actuales. Por ejemplo, la usabilidad de visualizaciones de tópicos se puede evaluar en un estudio longitudinal donde autoridades y representantes de gobierno utilicen las visualizaciones para respaldar sus tareas diarias en el mundo real. Asimismo el uso de visualizaciones se puede medir y comparar durante un período prolongado para comprender la usabilidad de las visualizaciones.

Dado que las experiencias y los gustos del usuario afectan la usabilidad de un objeto, un estudio que combine métodos cualitativos y cuantitativos ayudaría a obtener información más amplia y profunda sobre el uso de visualizaciones de tópicos. Los datos de entrevistas y por ejemplo, pueden arrojar mejores conocimientos sobre la razón por la cual un tipo específico de visualización funciona mejor que otros.

## BIBLIOGRAFÍA

- Abran, A., Khelifi, A., Suryn, W. y Seffah, A. (2003). “Usability Meanings and Interpretations in ISO Standards”. *Software Quality Journal*, 11(4), 325-338
- Allen, J. (1995). *Natural language understanding*. Pearson.
- Araya, F. B. (2018). Metodología para la construcción automática de un corpus de dominio específico.
- Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of planners*, 35(4), 216-224.
- Barnum, C. M. (2020). *Usability testing essentials: ready, set... test!* Morgan Kaufmann
- Baum, H. S. (2015). *Citizen participation*. In: N.J. Smelser and P.B. Baltes, eds. *International encyclopedia of the social and behavioral sciences*. New York: Elsevier, 1840–1846
- Bhatia, S., Lau, J. H., y Baldwin, T. (2018). Topic intrusion for automatic topic model evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 844-849).
- Becker Castellaro, S. (2018). Análisis deliberativo a la participación en procesos constituyentes: El caso chileno (2015-2018).
- Bisarya, S. (2015). Presentación sobre cambio constitucional. *Seminario Internacional: Cambio Constitucional en Democracia*. Santiago: SEGPRES; Banco Interamericano de Desarrollo; IDEA Internacional & PNUD.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng, A. Y., y Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3, 993–1022.
- Bobbio, L. (2019). Designing effective public participation. *Policy and Society*, 38(1), 41–57. <https://doi.org/10.1080/14494035.2018.1511193>
- Brynjolfsson, E., y McAfee, A. (2017). What’s driving the machine learning explosion? *Harvard Business Review*, 1704(4), 201-213.

Cai, G., Sun, F., y Sha, Y. (2018). Interactive Visualization for Topic Model Curation. In *IUI Workshops*.

Calheiros, A. C., Moro, S., y Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. *Journal of Hospitality Marketing & Management*, 26(7), 675-693.

Chaney, A. J. B., y Blei, D. M. (2012). Visualizing topic models. In *Sixth International AAAI Conference on Weblogs and Social Media-ICWSM*, pp. 419-422, 2012.

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., y Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems* (pp. 288-296).

Chen, M., Mao, S., y Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>

Cho, H. W. (2019). Topic Modeling. *Osong public health and research perspectives*, 10(3), 115.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Hillsdale, NJ: LEA.

Chuang, J., Manning, C. D., y Heer, J. (2012). Termite: Visualization techniques for assessing textual topic models. In *Proceedings of the international working conference on advanced visual interfaces* (pp. 74-77).

Colin, W. (2004). *Information visualization: perception for design*. San Francisco, CA: Morgan Kaufmann.

Debusse J., de la Iglesia B., Howard C., y Rayward-Smith V. (2001) Building the KDD Roadmap. In: Roy R. (eds) *Industrial Knowledge Management*. Springer, London

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., y Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407. doi: 10.1002/(SICI)1097-4571(199009)41:6h391::AID-ASII3.0.CO;2-9

Dryzek, J. S. (2002). *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford:Oxford University Press.

Elmqvist, N., y Yi, J. S. (2015). Patterns for visualization evaluation. *Information Visualization*, 14(3), 250-269.

Fayyad, U., Grinstein, G. G., y Wierse, A. (Eds.). (2002). *Information visualization in data mining and knowledge discovery*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Few, S. (2009). *Now you see it: simple visualization techniques for quantitative analysis* (No. Sirsi) i9780970601988).

Fierro, C., Fuentes, C., Pérez, J., y Quezada, M. (2017). 200k+ crowdsourced political arguments for a new chilean constitution. In *Proceedings of the 4th Workshop on Argument Mining* (pp. 1-10).

Floridia, A. (2018). *The Origins of the Deliberative Turn*. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, y M. Warren (Eds.), *The Oxford Handbook of Deliberative Democracy* (pp. 34–54). Oxford: Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780198747369.013.25>

France, S. L., y Akkucuk, U. (2020). A review, framework, and R toolkit for exploring, evaluating, and comparing visualization methods. *The Visual Computer*, 1-19.

Ganesan, A., Brantley, K., Pan, S., y Chen, J. (2015). Ldaexplore: Visualizing topic models generated using latent dirichlet allocation. *arXiv preprint arXiv:1507.06593*.

Ghosh, D., y Guha, R. (2013). What are we ‘tweeting’ about obesity? Mapping tweets with topic modeling and Geographic Information System. *Cartography and geographic information science*, 40(2), 90-102.

Ginsburg, T., Elkins, Z., y Blount, J. (2009). Does the Process of Constitution-Making Matter? *Annual Review of Law and Social Science*, 5, 201-223.

Gluck, J., y Ballou, B. (2014). *New Technologies in Constitution Making*. Washington, D.C.

Gluck, J., y Brandt, M. (2015). *Participatory and inclusive constitution-making*. Washington, D.C.

Hagen, L., Uzuner, Ö., Kotfila, C., Harrison, T. M., y Lamanna, D. (2015). Understanding citizens' direct policy suggestions to the federal government: a natural language processing and topic modeling approach. In *2015 48th Hawaii International Conference on System Sciences* (pp. 2134-2143). IEEE.

Hall, M., Clough, P., y Stevenson, M. (2012, September). Evaluating the use of clustering for automatically organising digital library collections. In *International Conference on Theory and Practice of Digital Libraries* (pp. 323-334). Springer, Berlin, Heidelberg.

Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). Themeriver: Visualizing thematic changes in large document collections. *IEEE transactions on visualization and computer graphics*, 8(1), 9-20.

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., y Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction*. ACM.

Hofmann, T. (1999). Probabilistic latent semantic analysis. En Proceedings of the fifteenth conference on uncertainty in artificial intelligence (pp. 289–296). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. Descargado de <http://dl.acm.org/citation.cfm?id=2073796.2073829>

Hsiao, Y. T., Lin, S. Y., Tang, A., Narayanan, D., y Sarahe, C. (2018). vTaiwan: An empirical study of the open consultation process in Taiwan.

Isaac, O., Abdullah, Z., Ramayah, T., y Mutahar, A. M. (2017). Internet usage, user satisfaction, task-technology fit, and performance impact among public sector employees in Yemen. *The International Journal of Information and Learning Technology*.

ISO. (1998). *ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (vdts) – part 11: guidance on usability*.

Kairam, S., Riche, N. H., Drucker, S., Fernandez, R., y Heer, J. (2015). Refinery: Visual exploration of large, heterogeneous networks through associative browsing. In *Computer graphics forum* (Vol. 34, No. 3, pp. 301-310).

Kant, G., Weisser, C., y Säfken, B. (2020). TTLocVis: A Twitter Topic Location Visualization Package. *Journal of Open Source Software*, 5(54), 2507.

Kapadia, S. (2019). Evaluate Topic Models: Latent Dirichlet Allocation (LDA). 21 Abril 2020, de Towards Data Science Sitio web: <https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0>

Karpovich, S., Smirnov, A., Teslya, N., y Grigorev, A. (2017, April). Topic model visualization with ipython. In *2017 20th Conference of Open Innovations Association (FRUCT)* (pp. 131-137). IEEE.

Koylu, C. (2019). Modeling and visualizing semantic and spatio-temporal evolution of topics in interpersonal communication on Twitter. *International Journal of Geographical Information Science*, 33(4), 805-832.

- Lau, J. H., Grieser, K., Newman, D., y Baldwin, T. (2011, June). Automatic labelling of topic models. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies* (pp. 1536-1545).
- Lau, J. H., Newman, D., y Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 530-539).
- Landauer, T. K., Foltz, P. W., y Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3), 259-284.
- Lee, B., Riche, N. H., Karlson, A. K., y Carpendale, S. (2010). Sparkclouds: Visualizing trends in tag clouds. *IEEE transactions on visualization and computer graphics*, 16(6), 1182-1189.
- Lewis, J. R. (1995) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction*, 7:1, 57-78. Abstract | About quest.cgi
- Li, G., Zhang, Y., Dong, Y., Liang, J., Zhang, J., Wang, J., y Yuan, X. (2019). BarcodeTree: Scalable comparison of multiple hierarchies. *IEEE transactions on visualization and computer graphics*, 26(1), 1022-1032.
- Liu, Q., Chen, Q., Shen, J., Wu, H., Sun, Y., y Ming, W. K. (2019). Data analysis and visualization of newspaper articles on thirdhand smoke: a topic modeling approach. *JMIR medical informatics*, 7(1), e12414.
- Ma, B., Zhang, N., Liu, G., Li, L., y Yuan, H. (2016). Semantic search for public opinions on urban affairs: A probabilistic topic modeling-based approach. *Information Processing & Management*, 52(3), 430-445.
- Maboudi, T., y Nadi, G. P. (2016). Crowdsourcing the Egyptian constitution: social media, elites, and the populace. *Political Research Quarterly*, 69(4), 716-731.
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *Acm Transactions On Graphics (Tog)*, 5(2), 110-141.
- Manning, C. D., Raghavan, P., y Schütze, H. (2008). Scoring, term weighting and the vector space model. *Introduction to information retrieval*, 100, 2-4.
- Manresa-Yee, C., Amengual, E., y Ponsa Asensio, P. (2014). La usabilidad de las interfaces basadas en visión. *FAZ*, (7), 12-31.

- Maña, M. J., De Buenaga, M., y Gómez, J. M. (1998). Diseño y evaluación de un generador de resúmenes de texto con modelado de usuario en un entorno de recuperación de información. *Procesamiento del Lenguaje Natural*, 23.
- Martin, M. E., y Schuurman, N. (2017). Area-based topic modeling and visualization of social media for qualitative GIS. *Annals of the American Association of Geographers*, 107(5), 1028-1039.
- Mayr, E., Hynek, N., Salisu, S., y Windhager, F. (2019). Trust in information visualization. In *TrustVis workshop*. The Eurographics Association, Porto.
- Mayr, S., Erdfelder, E., Buchner, A., y Faul, F. (2007). A short tutorial of GPower. *Tutorials in quantitative methods for psychology*, 3(2), 51-59.
- Mei, Q., Shen, X., y Zhai, C. (2007, August). Automatic labeling of multinomial topic models. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 490-499).
- Mimno, D., Wallach, H., Talley, E., Leenders, M., y McCallum, A. (2011, July). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 262-272).
- Moreno, C. M. (2008). Metodología de investigación en estudios de usuarios. *Revista general de información y documentación*, 17(2), 129-149.
- Morstatter, F., y Liu, H. (2017). In search of coherence and consensus: measuring the interpretability of statistical topics. *The Journal of Machine Learning Research*, 18(1), 6177-6208.
- Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 921-928. Descargado de <http://dx.doi.org/10.1109/TVCG.2009.111> doi: 10.1109/TVCG.2009.111
- Munzner, T. (2015). Visualization analysis and design. CRC Press.
- Mukherjee, S. (2019). How IT allows E-Participation in Policy-Making Process. *arXiv preprint arXiv:1903.00831*.
- Negretto, G. L. (2017). *Constitution Making in Democratic Constitutional Orders: The Challenge of Citizen Participation. Let the people rule*.
- Newman, D., Baldwin, T., Cavedon, L., Huang, E., Karimi, S., Martínez, D. y Zobel, J. (2010). Visualizing search results and document collections using topic maps. *Web Semantics: Science, Services and Agents on the World Wide Web*, 8(2-3), 169-175.

- Nilsson, N. J. (1996). Introduction to Machine Learning (An Early Draft to a proposed textbook). Descargado de <https://ai.stanford.edu/~nilsson/mlbook.html>
- Nikolenko, S. I., Koltcov, S., y Koltsova, O. (2017). Topic modelling for qualitative studies. *Journal of Information Science*, 43(1), 88-102.
- Noveck, B. S. (2018). Crowdlaw: Collective intelligence and lawmaking. *Analyse & Kritik*, 40(2), 359-380.
- OECD (2015), “Policy Shaping and Policy Making: The Governance of Inclusive Growth”, background report to the Public Governance Ministerial Meeting, 28 October 2015, OECD, Paris, [www.oecd.org/governance/ministerial/the-governance-of-inclusive-growth.pdf](http://www.oecd.org/governance/ministerial/the-governance-of-inclusive-growth.pdf).
- OECD (2017). Public Governance Reviews. Chile scan report on the citizen participation in the constitutional process. Technical report.
- Pacheco i Canals, J. (2019). A normative model for deliberative constitution-building. *Revista Catalana de Dret Públic*, 58, 166–181. <https://doi.org/http://dx.doi.org/10.2436/rcdp.i58.2019.3225>
- Parra, D., Trattner, C., Gómez, D., Hurtado, M., Wen, X., y Lin, Y. R. (2016). Twitter in academic events: a study of temporal usage, communication, sentimental and topical patterns in 16 computer science conferences. *Computer Communications*, 73, 301-314.
- Paul, M. J., y Girju, R. (2010, July). A Two-Dimensional Topic-Aspect Model for Discovering Multi-Faceted Topics. In *AAAI* (Vol. 51, p. 61801).
- Poria, S., Cambria, E., y Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108, 42-49.
- Qu, Z., y Hullman, J. (2016, October). Evaluating visualization sets: Trade-offs between local effectiveness and global consistency. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (pp. 44-52).
- Raveau, M. P., Couyoumdjian, J. P., Fuentes-Bravo, C., Rodriguez-Sickert, C., y Candia, C. (2020). Citizens at the forefront of the constitutional debate: Participation determinants and emergent content in Chile. *arXiv preprint arXiv:2006.00140*.
- Radford, A., Narasimhan, K., Salimans, T., y Sutskever, I. (2018). Improving language understanding by generative pre-training. [www.s3-us-west-2.amazonaws.com/openai-assets/researchcovers/languageunsupervised/language understanding paper.pdf](http://www.s3-us-west-2.amazonaws.com/openai-assets/researchcovers/languageunsupervised/language%20understanding%20paper.pdf).

- Reisinger, J., Waters, A., Silverthorn, B., y Mooney, R. J. (2010). Spherical topic models. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 903-910).
- Repo, P., Anttonen, M., Mykkänen, J., y Lammi, M. (2018). Lack of congruence between European citizen perspectives and policies on circular economy. *European Journal of Sustainable Development*, 7(1), 249-264.
- Roque de Oliveira, A., y Partidário, M. (2020). You see what I mean? – A review of visual tools for inclusive public participation in *EIA decision-making processes*. *Environmental Impact Assessment Review*, 83, 106413. <https://doi.org/10.1016/j.eiar.2020.106413>
- Saavedra, C. F. (2016). El proceso constituyente: La vía chilena hacia el cambio constitucional. *Anuario de Derecho Público*, (1), 107-130.
- Saket, B., Endert, A., y Demiralp, Ç. (2018). Task-based effectiveness of basic visualizations. *IEEE transactions on visualization and computer graphics*, 25(7), 2505-2512.
- Salton, G., y McGill, M. J. (1986). *Introduction to modern information retrieval*. New York, NY, USA: McGraw-Hill, Inc.
- Sampieri, R. H., Collado, C. F., Lucio, P. B., Valencia, S. M., y Torres, C. P. M. (1998). *Metodología de la investigación* (Vol. 6). México, DF: McGraw-hill.
- Sathi, V. R., y Ramanujapura, J. S. (2016). A Quality Criteria Based Evaluation of Topic Models.
- Seitsonen, O. (2017). Crowdsourcing cultural heritage: public participation and conflict legacy in Finland. *Journal of Community Archaeology & Heritage*, 4(2), 115-130.
- Sievert, C., y Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces* (pp. 63-70).
- Smith, A., Hawes, T., y Myers, M. (2014). Hierarchie: Visualization for hierarchical topic models. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (pp. 71-78).
- Sørum, H., Andersen, K. N., y Vatrapu, R. (2012). Public websites and human–computer interaction: an empirical study of measurement of website quality and user satisfaction. *Behaviour & Information Technology*, 31(7), 697-706.

Srivastava, A. N., y Sahami, M. (2009). Text mining: Classification, clustering, and applications. *Boca Raton, FL: CRC Press.*

Steinskog, A., Therkelsen, J., y Gambäck, B. (2017, May). Twitter topic modeling by tweet aggregation. In *Proceedings of the 21st nordic conference on computational linguistics* (pp. 77-86).

Steyvers, M., y Griffiths, T. (2007). Probabilistic topic models. *Handbook of latent semantic analysis*, 427(7), 424-440.

Sukhija, N., Tatineni, M., Brown, N., Van Moer, M., Rodriguez, P., y Callicott, S. (2016, July). Topic modeling and visualization for big data in social sciences. In *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)* (pp. 1198-1205). IEEE.

Szafir, D. A., Haroz, S., Gleicher, M., y Franconeri, S. (2016). Four types of ensemble coding in data visualizations. *Journal of vision*, 16(5), 11-11.

Torres Ponjuán, D. (2010). *La Visualización de la Información en el entorno de la Ciencia de la Información*. Granada: Universidad de Granada.

Tvinnereim, E., y Fløttum, K. (2015). Explaining topic prevalence in answers to open-ended survey questions about climate change. *Nature Climate Change*, 5(8), 744.

Uboldi, G., & Caviglia, G. (2015). Information visualizations and interfaces in the humanities. In *New Challenges for Data Design* (pp. 207-218). Springer, London.

Umar, M. M., Bakhat, M. U., y Hassan, M. (2020). Mapping HCI Principles to Evaluate the Usability of Learning Applications for CCI User. *International Journal of Computer Science and Telecommunications*, 11(2).

Wallach, H. M. (2006). Topic modeling: beyond bag-of-words. En *Proceedings of the 23rd international conference on machine learning* (pp. 977–984).

Wang, C., y Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 448-456).

Ward, M., Grinstein, G., y Keim, D. (2010). *Interactive data visualization: Foundations, techniques, and applications*. Natick, MA, USA: A. K. Peters, Ltd.

Ware, C. (2020). *Information visualization: perception for design*. Morgan Kaufmann.

Wattenberg, M. (2005, October). Baby names, visualization, and social data analysis. In IEEE Symposium on Information Visualization, 2005. *INFOVIS 2005*. (pp. 1-7). IEEE.

Yang, Y., Yao, Q., y Qu, H. (2017). VISTopic: A visual analytics system for making sense of large document collections using hierarchical topic modeling. *Visual Informatics*, 1(1), 40-47.

Yao, L., Mimno, D., y McCallum, A. (2009). Efficient methods for topic model inference on streaming document collections. En *Proceedings of the 15th acm sigkdd international conference on knowledge discovery and data mining* (pp. 937–946).

Yu, D., Xu, D., Wang, D., y Ni, Z. (2019). Hierarchical topic modeling of Twitter data for online analytical processing. *IEEE Access*, 7, 12373-12385.

Zia, A., Guo, L., Zhou, L., Essa, I., y Jarc, A. (2019). Novel evaluation of surgical activity recognition models using task-based efficiency metrics. *International journal of computer assisted radiology and surgery*, 14(12), 2155-2163.

## Anexo A. Información estudio de usuario

### A.1 Consentimiento informado

### Consentimiento Informado

Por favor lea esta información cuidadosamente antes de decidir su participación en el estudio:

Usted ha sido invitado a participar de un estudio sobre usabilidad de herramientas de analítica visual para topic modeling de participación ciudadana.

El estudio es representado por la alumna de Magíster de Ciencias de la Ingeniería, Ivania Yovanovic, en supervisión de la profesora Constanza Miranda, de Diseño en Ingeniería (Dilab), de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile.

**¿De qué trata el estudio?**

El estudio consiste en evaluar la coherencia semántica del modelo de tópicos implementado y en investigar la eficiencia, efectividad y satisfacción de dos herramientas de analítica visual, con datos de participación ciudadana.

**¿Cómo participar?**

Para concretar su participación, le pedimos que conteste un cuestionario, que le tomará entre 20 y 30 minutos. Su participación en este estudio es voluntaria. No se prevén eventuales riesgos por completar el cuestionario, sin embargo usted es libre de suspender su participación en cualquier momento.

**¿De qué manera se utilizan las respuestas?**

Toda información que usted nos entregue será tratada de manera confidencial. Cualquier información que se haga pública será de manera agregada y sin posibilidad que sea posible identificarlo personalmente. La confidencialidad de sus respuestas será protegida en todo momento. Los datos sólo serán utilizados para la presente investigación.

**¿Qué debo hacer si necesito más información?**

Si lo desea, puede enviar un correo electrónico a la investigadora responsable, (iyovanovic@uc.cl), para recibir una copia de las publicaciones que se escriban basadas en esta investigación.

Haciendo clic en la casilla, usted confirma que ha leído y comprendido esta descripción y está de acuerdo con participar del estudio.

☐ Acepto el consentimiento informado

Volver al menúSiguiente

### A.2 Pre encuesta

### Pre-encuesta

Edad

Ocupación

Género

AtrásEnviar

## A.3 Contexto del estudio

### Contexto

En 2016, el Gobierno de Chile inició un nuevo proceso de cambio constitucional. En esta línea, se definió un mecanismo abierto a la ciudadanía. Se esperaba que ciudadanos participaran y pudieran incidir en la discusión constitucional.

Tras la consulta ciudadana, se recopilaron más de 200.000 opiniones de ciudadanos chilenos. Esta es la base de datos más extensa de participación ciudadana en reformas constitucionales en el mundo.

Para poder conocer la opinión de la ciudadanía, es decir, sistematizar aquellas 200.000 opiniones, la capacidad humana no es suficiente. Es por ello que en este estudio se recurrió a un modelo computacional para poder analizar todas aquellas opiniones de manera más eficiente. Sin embargo, la coherencia con que un computador opera debe ser verificada.

Por lo tanto, surge la siguiente pregunta: ¿Puede el modelo, categorizar las opiniones de manera coherente para un humano?

[Volver al menú](#)[Siguiendo](#)

### Contexto

Para responder esta pregunta, debemos entender la manera en que el modelo *categoriza* estas opiniones. En primer lugar, hay que saber que el computador recibe las 200.000 opiniones y las categoriza según su contenido, es decir, según sus palabras.

El modelo agrupa palabras según el contexto. Cada uno de esos grupos de palabras corresponde a un tópico. Por ejemplo, si procesáramos un libro de historia, algunos de los tópicos podrían ser:

Tópico 1: [Mapuche] [Aymara] [Alacalufe] [Atacameño] [Diaguita] [Ona]

Tópico 2: [Alessandri] [Frei] [O'Higgins] [Allende] [Montt] [Pinochet]

Tópico N: [Salitre] [Chile] [Tratado] [Perú] [Bolivia] [Esmeralda]

Lo que buscamos a través de este estudio, es verificar si estos grupos de palabras, fueron agrupados **coherentemente**. Para ello, deberás responder dos tipos de ejercicios.

[Atrás](#)[Siguiendo](#)

## A.4 Tutorial tareas Parte I

### Tutorial

#### La palabra intrusa

El primer ejercicio se trata de descubrir la palabra intrusa. Cada tema se presentará como una lista de palabras, donde existe una "palabra intrusa". Tu misión será identificar aquella palabra intrusa que **no** encaja con las demás.

Deberás responder tres preguntas de este tipo. Donde cada pregunta es sobre un tema diferente. El orden de las palabras es aleatorio. Por ende, la palabra intrusa puede estar en cualquier posición.

Recordando el ejemplo del libro de historia, aquí hay un nuevo listado de palabras. ¿Puedes descubrir la palabra intrusa?

Tópico: [Colón] [Descubrimiento] [Pinta] [Niña] [Magallanes] [Atlántico] [América]

Al comienzo de cada ejercicio habrá una pregunta de ejemplo, con palabras y tópicos de la consulta.

[Atrás](#)[Siguiendo](#)

# Tutorial

## El tópico intruso

El segundo ejercicio se trata de descubrir el tópico intruso. En este caso se presentará un breve texto y se plantearán dos tópicos que resumen ese texto, donde uno de ellos es un "tópico intruso". Tu misión será identificar aquel conjunto de palabras, o tópico que **menos** encaja con el texto presentado.

Deberás responder tres preguntas de este tipo. Donde cada pregunta es sobre un texto con tópicos diferentes. El orden de los tópicos es aleatorio. Por ende, el tópico intruso puede estar en cualquier posición.

Recordando el ejemplo del libro de historia, aquí se presenta un breve texto y dos tópicos asociados.

Texto: *La Primera Guerra Mundial, anteriormente llamada la Gran Guerra, fue una confrontación bélica centrada en Europa que empezó el 28 de julio de 1914 y finalizó el 11 de noviembre de 1918, cuando Alemania aceptó las condiciones del armisticio.*

Tópico 1: [Guerra] [Confrontación] [armas] [Alemania] [Bélica] [Europa]

Tópico 2: [Alemania] [Imperio] [Británico] [Stalingrado] [Aliados] [Italia]

¿Puedes descubrir el tópico intruso?

Al comienzo de cada ejercicio habrá una pregunta de ejemplo, con opiniones y tópicos de la consulta.

Atrás

Siguiente