



**4TH ARTIFICIAL INTELLIGENCE AND
LIBRARIES SYMPOSIUM: TRANSFORMING
INFORMATION ACCESS AND DISCOVERY**



Biblioteca del Congreso
Nacional de Chile / BCN

Transformación de Servicios Parlamentarios: *Evolucionando hacia un análisis oportuno de la coyuntura mediante IA y visualizaciones*

El Proyecto Diana

Francisco Cifuentes Silva
fcifuentes@bcn.cl

Abril de 2024

Algo de historia...

IFLA-LAC 2017

Modelo de enriquecimiento semántico de recursos de información

Francisco Cifuentes Silva
fcifuentes <at> bcn.cl



14 de marzo de 2017

The video shows a speaker on the left and a slide on the right. The slide is titled "¿Qué se puede identificar en un recurso de información?" and features a central "Recurso de información" icon. Surrounding it are several categories of metadata and content analysis:

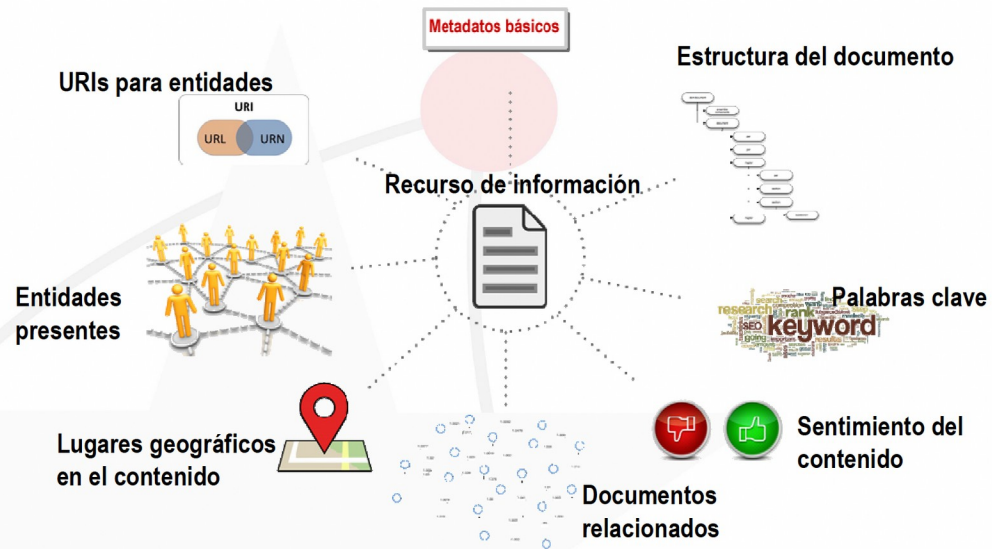
- Metadatos básicos
- Estructura del documento
- URIs para entidades
- Entidades presentes
- Lugares geográficos en el contenido
- Palabras clave / keyword
- Sentimiento del contenido
- Documentos relacionados

The video player interface at the bottom shows a progress bar at 5:33 / 21:08 and the text "Biblioteca del Congreso Nacional de Chile / BCN".

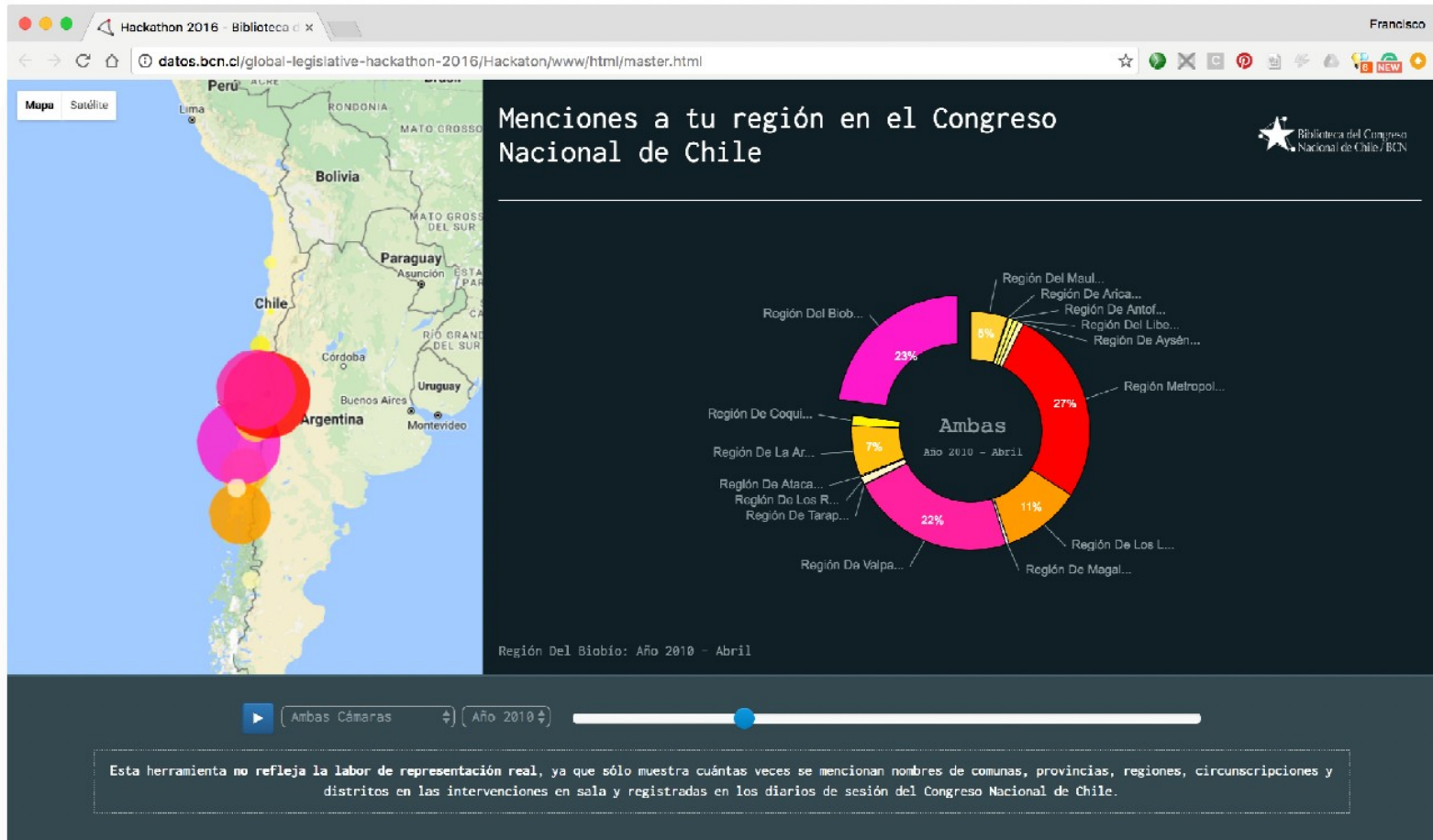
Algo de historia...

IFLA-LAC 2017

¿Qué se puede identificar en un recurso de información?



Pruebas de concepto (2017)



Pensar en las posibilidades...

Primer objetivo estratégico BCN

Fortalecer asesoría parlamentaria

BCN cuenta con un archivo importante:

- Prensa nacional
- Historia política
- Legislación
- Historia legislativa

Qué se propone

Proyecto Diana

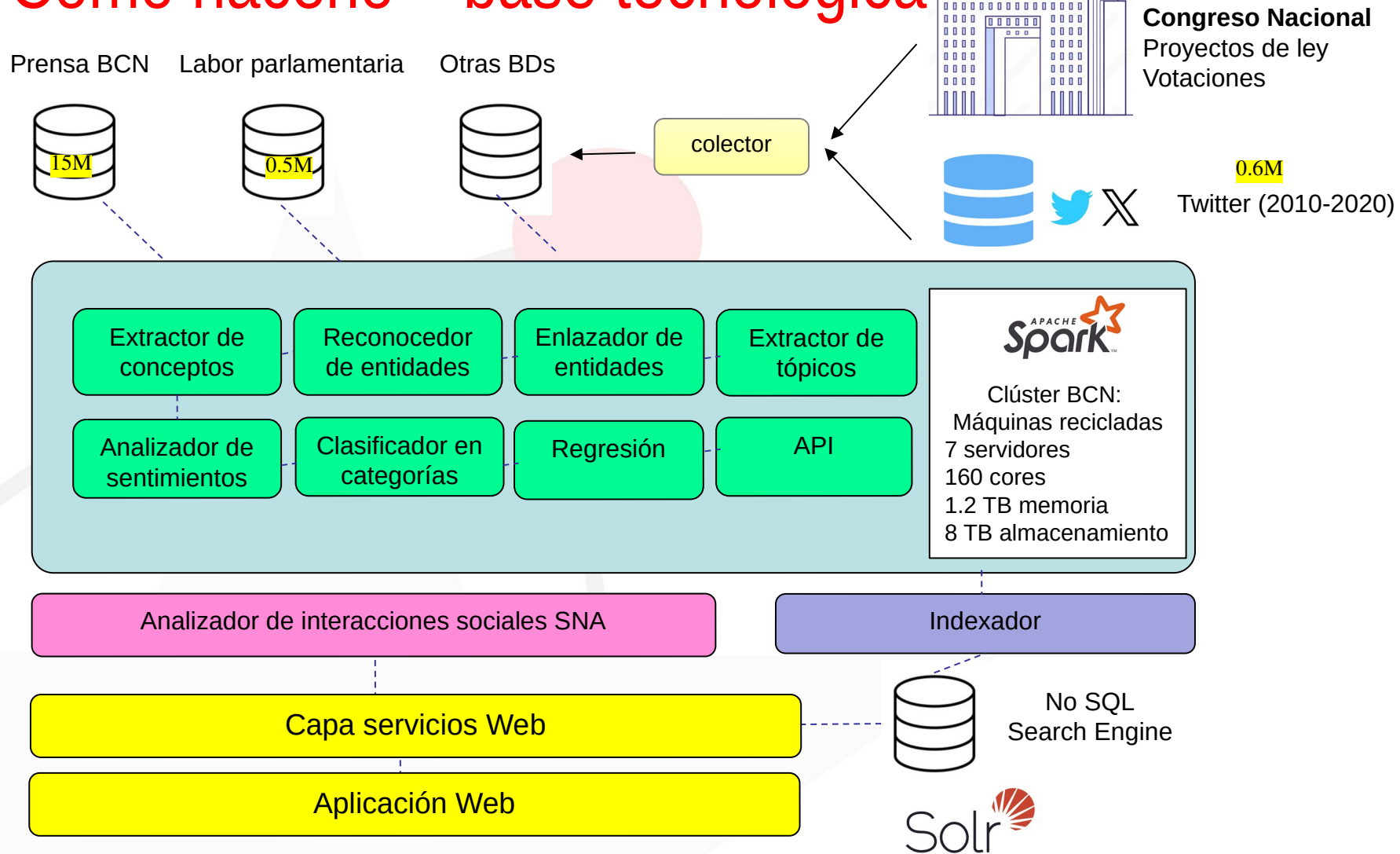


Objetivo del Proyecto Diana



Construir un sistema inteligente que permita **identificar y entregar información de interés** a la comunidad parlamentaria de **forma proactiva, bajo oferta** y de forma **personalizada**

Cómo hacerlo – base tecnológica



Casos de uso IA y visualizaciones

4 herramientas desarrolladas con IA dentro de Diana:

- I) Extractor automático de conceptos clave
- II) Clasificador de análisis de sentimientos
- III) Predictor de aprobación de la ley
- IV) Análisis de asociaciones usando SNA

I.- Extractor automático de conceptos clave

Cómo hacerlo:

1) Combinando:

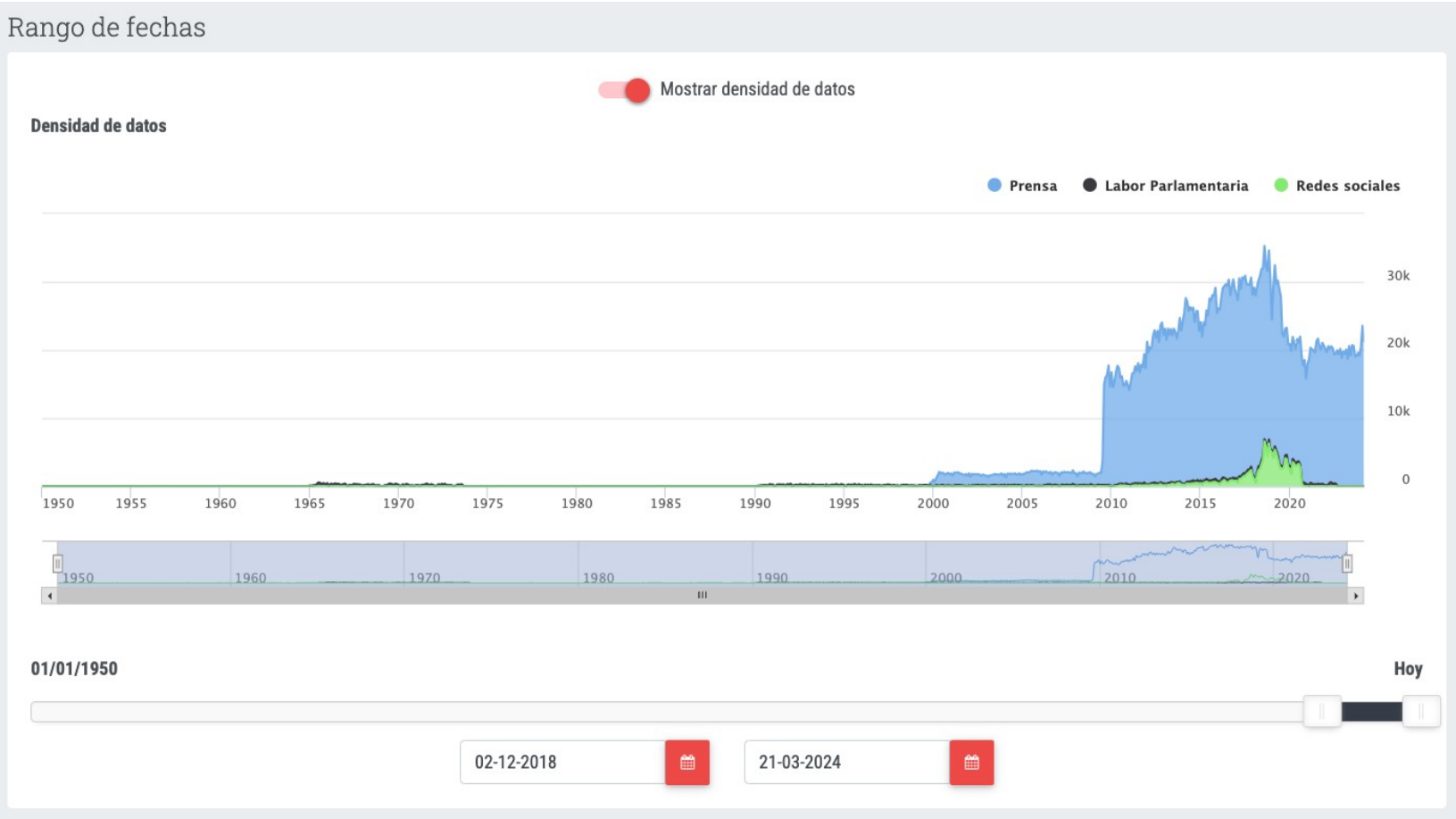
- Utilización de Named Entity Recognizer Open Source (Stanford NER + Spacy)
- Aplicación de algoritmo TF-IDF sobre n-gramas de palabras (1-4)
- Aplicación de algoritmo de Análisis Semántico Latente (LDA)

2) Utilización de facetas sobre motor de búsqueda SolR

3) Visualización de datos



Datos sobre los que se están consultando



Plataforma de búsqueda

Texto de búsqueda



Buscar por...
incendio

Búsqueda específica

Rango de fechas

Mostrar densidad de datos

01/01/1950 Hoy

18-03-2014  25-03-2024 

Filtros

Institución


- Senado
- Cámara de diputados
- Congreso nacional
- Sin institución

Datos




Redes Sociales


- Twitter
- Facebook
- Instagram
- Página personal

Labor Parlamentaria

- Sala de sesiones 

Prensa

- Prensa escrita 
- Prensa online 
- Revistas 

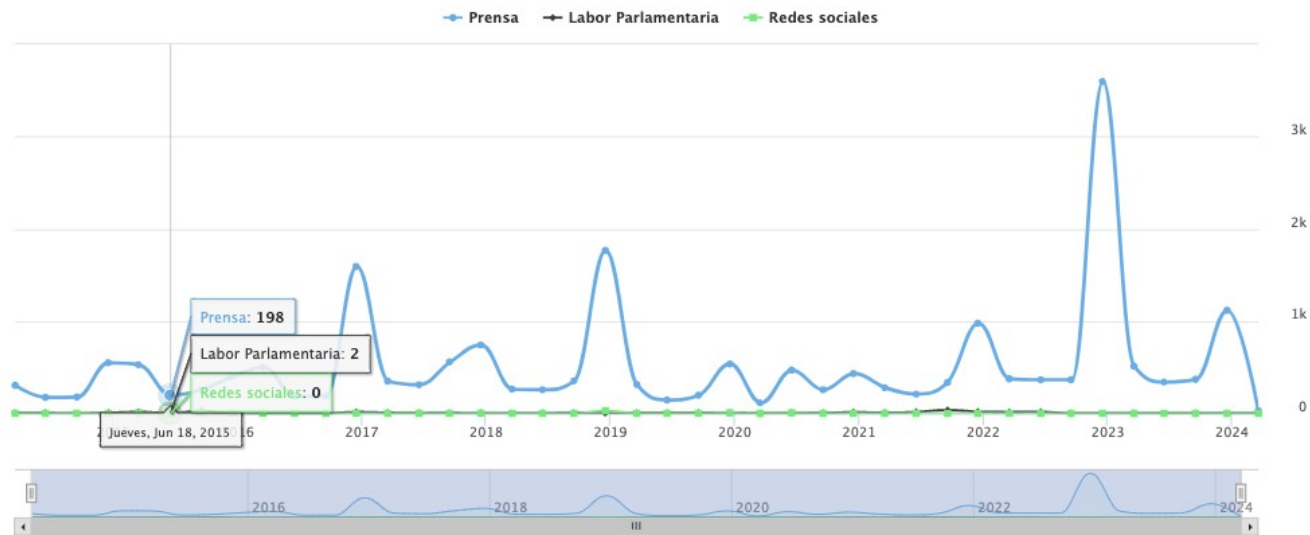
FILTRAR **LIMPIAR DATOS** 

Búsqueda por palabra
"incendio" los últimos 10 años

Describiendo concepto: Araucania, relacionado a la búsqueda incendio



Aparición del concepto Araucania por tipo de medio



Precisión Gráfica:

Trimestral

Sentimiento asociado al concepto



Términos co-ocurrentes





NUBE

MEDIOS

LINEA TEMPORAL

EVOLUCIÓN CONCEPTOS

RESULTADOS

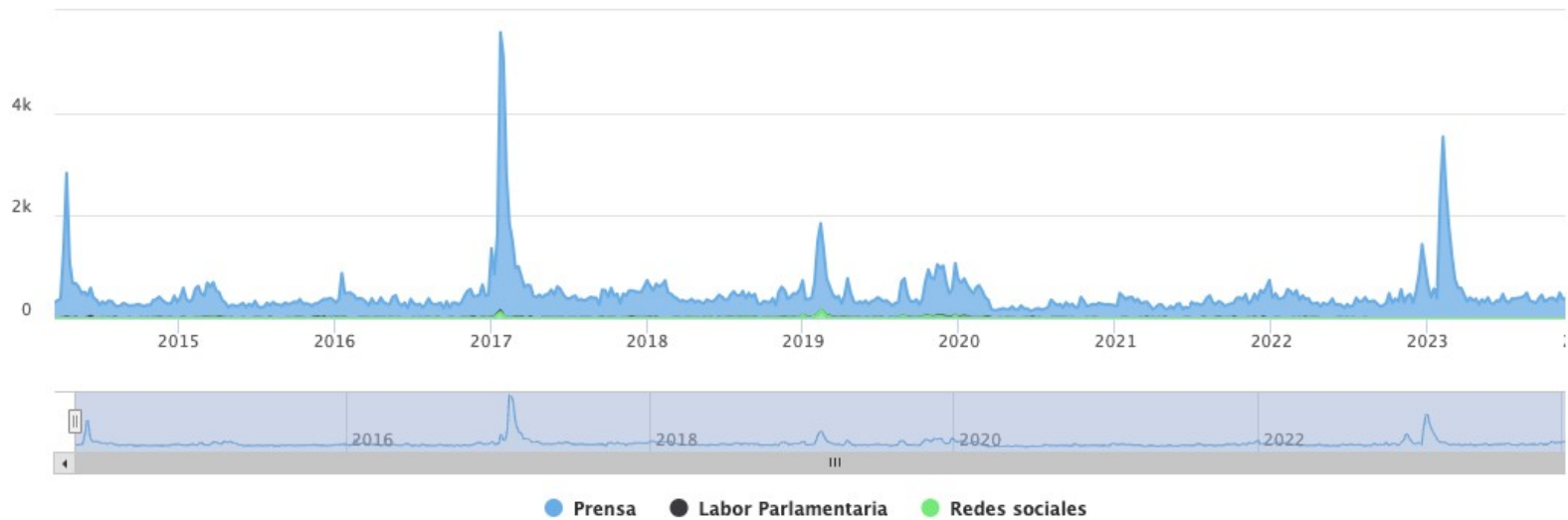
SENTIMIENTO

GEOREFERENCIA

CLUSTERIZAR

PARLAMENTARIOS

Total de documentos en el periodo



Precisión Gráfica

Semanal

NUBE

MEDIOS

LINEA TEMPORAL

EVOLUCIÓN CONCEPTOS

RESULTADOS

SENTIMIENTO

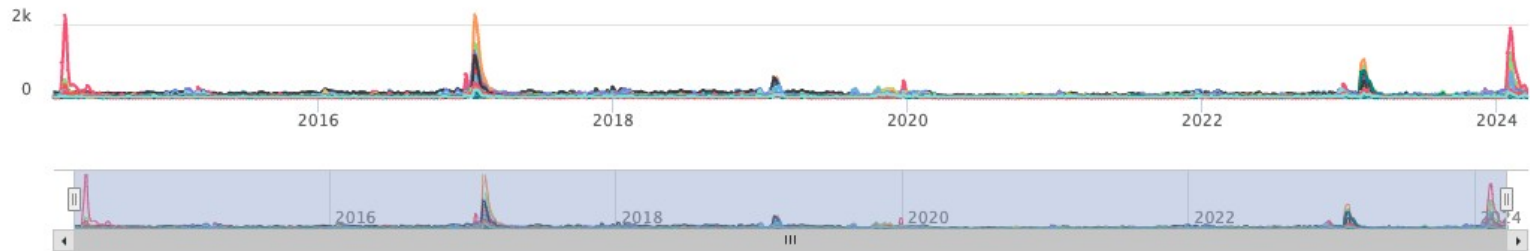
GEOREFERENCIA

CLUSTERIZAR



PARLAMENTARIOS

Evolución de conceptos co-ocurrentes más relevantes asociados a incendio



- Bomberos
- Conaf
- Gobierno
- Cuerpo de Bomberos
- Santiago
- Sebastian Pinera
- hectareas
- damnificados
- incendio
- Valparaiso
- emergencia
- fuego
- Biobio
- Maule
- Ministerio
- PDI
- Chile
- Carabineros
- forestales
- incendios
- propagacion
- Corporacion Nacional Forestal
- voluntarios
- incendios forestales
- sinistro
- Araucania
- cuerpo
- Onemi
- Ministerio Publico
- compania

Precisión Gráfica

Semanal

Cantidad de Conceptos

30

10

20

30

40

50

60

70

Opciones

Series:

Ocultar Mostrar

Gráfico:

Líneas Áreas

II.- Clasificador de análisis de sentimientos

- Interfaz de entrenamiento para usuarios
- Permite entender tendencias generales
- Diseño de características (50 aproximadamente), algunas de ellas:

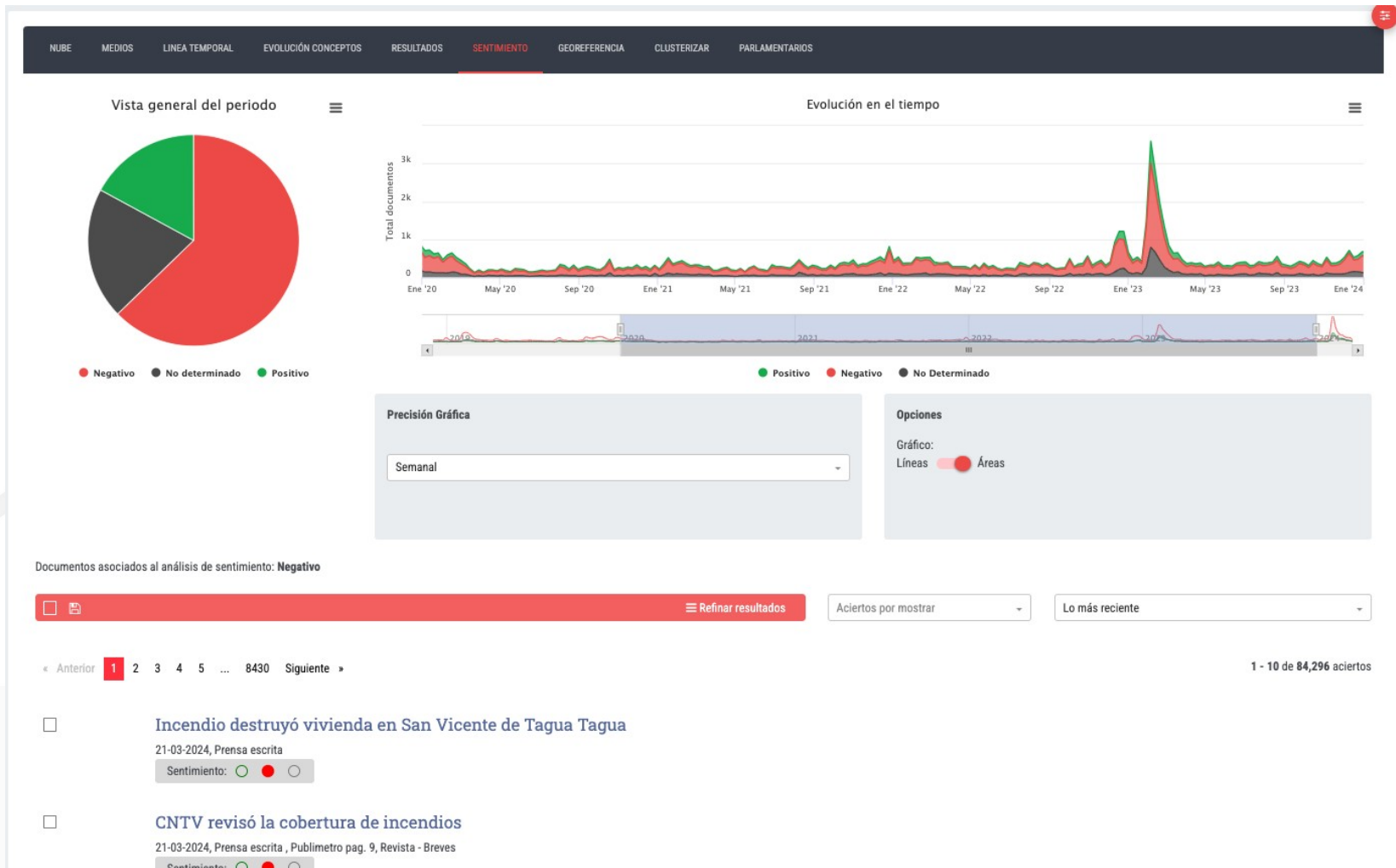
- CantidadExclamacion
- TotalEntidades
- Largo del texto
- CantidadSegundaPersona
- SumaPuntajePolaridadPalabras
- CantidadSi
- CantidadNo
- CantidadPalabrasPositivas
- CantidadMayusculas
- CantidadEmoticonTriste
- CantidadPregunta
- CantidadPalabras
- CantidadEmoticonFeliz
- CantidadTerceraPersona
- PuntajePalabrasPositivas
- CantidadPrimeraPersona
- PuntajePalabrasNegativas
- CantidadPalabrasNegativas

Se implementó clasificador Ridge
Usando scikit-learn

Con un 85% de precisión media
estimada en la clasificación



II.- Clasificador de análisis de sentimientos



II.- Clasificador de análisis de sentimientos

NUBE MEDIOS LINEA TEMPORAL EVOLUCIÓN CONCEPTOS RESULTADOS SENTIMIENTO GEOREFERENCIA CLUSTERIZAR PARLAMENTARIOS

Búsqueda por palabra "viejito pascuero"

Vista general del periodo



● Positivo ● No determinado ● Negativo



Precisión Gráfica: Semanal

Opciones: Gráfico: Líneas Áreas

Documentos asociados al análisis de sentimiento: **Positivo**

Refinar resultados Aciertos por mostrar Ordenar por

« Anterior 1 2 3 4 5 ... 247 Siguiente » 1 - 10 de 2,466 aciertos

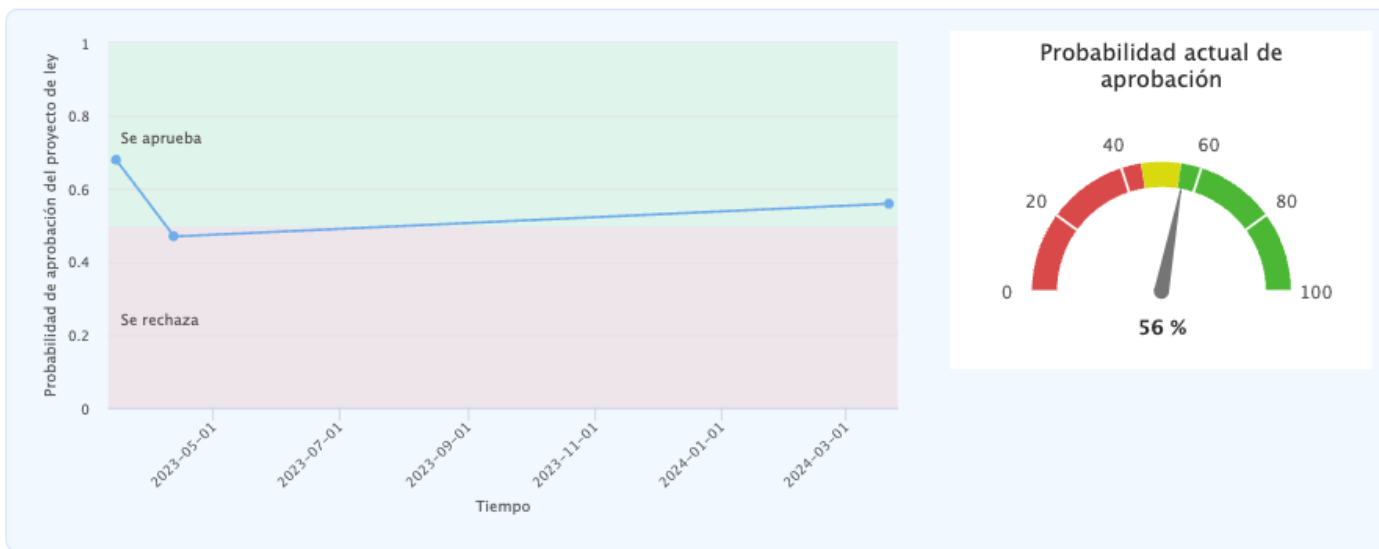
- Celebración de Navidad en Servicio de Salud Magallanes**
25-12-2022, Prensa escrita, El Pinguino pag. 10, Sociales
Sentimiento: ● ○ ○
- ¿Dónde viene el Viejito Pascuero? Sigue en TIEMPO REAL su recorrido entregando regalos**
24-12-2023, Prensa online, 24horas.cl, Portada
Sentimiento: ● ○ ○

III.- Predictor de aprobación de la ley

Moderniza la gestión de conservadores, notarios y archiveros judiciales (id:20858)

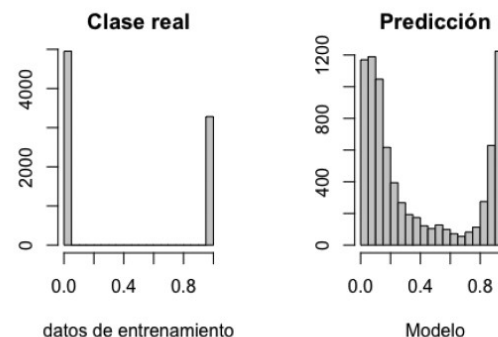
Nº boletín:	Ver tramitación boletín nº 15291-07
Iniciativa:	Moción Parlamentaria
Fecha de presentación:	2022-08-17
Estado:	En tramitación
Organismo fuente:	Cámara de Diputados

Predicción de aprobación



III.- Predictor de aprobación de la ley

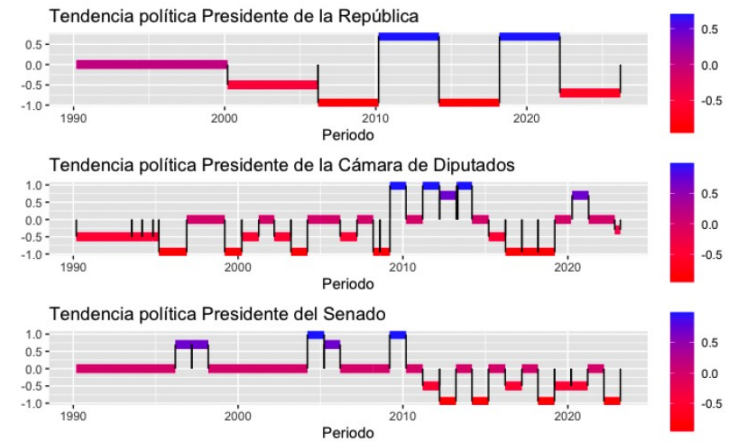
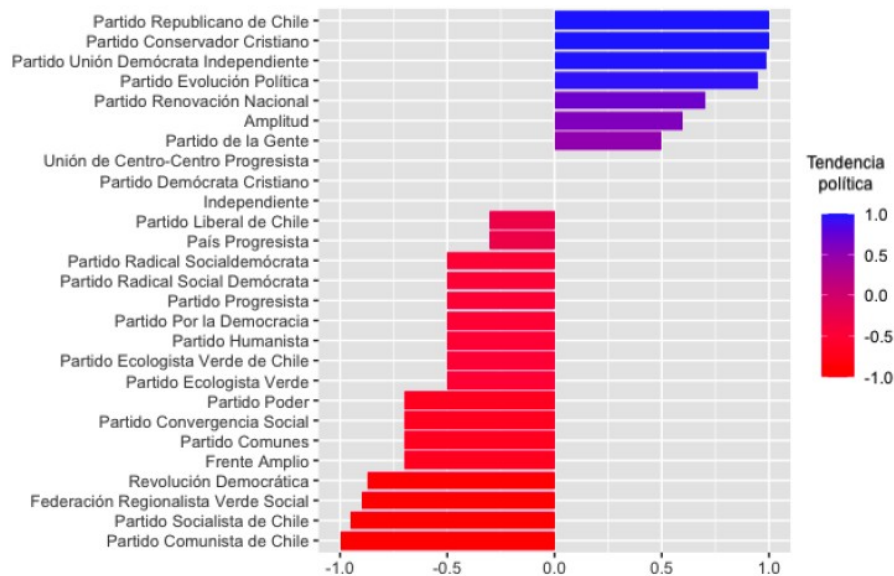
- Se implementó regresión logística en R
- Se entrenó con los datos de proyectos de ley aprobados y rechazados/archivados
- Se definieron características en base a datos de la tramitación y en base al índice de tendencia política del momento
- Modelo alcanza un 86% de precisión



Intercepto	***
id_iniciativa	***
total_autores_carga_cero	***
promedio_carga	***
promedio_carga_neg	***
mes_ultima_tramitacion_sil	***
anio_ultima_tramitacion_sil_hay_elecciones	**
total_dias_en_tramitacion	***
carga_politica_ejecutivo_al_ingreso	***
carga_politica_presidencia_senado_en_ultima_tramitacion_sil	**

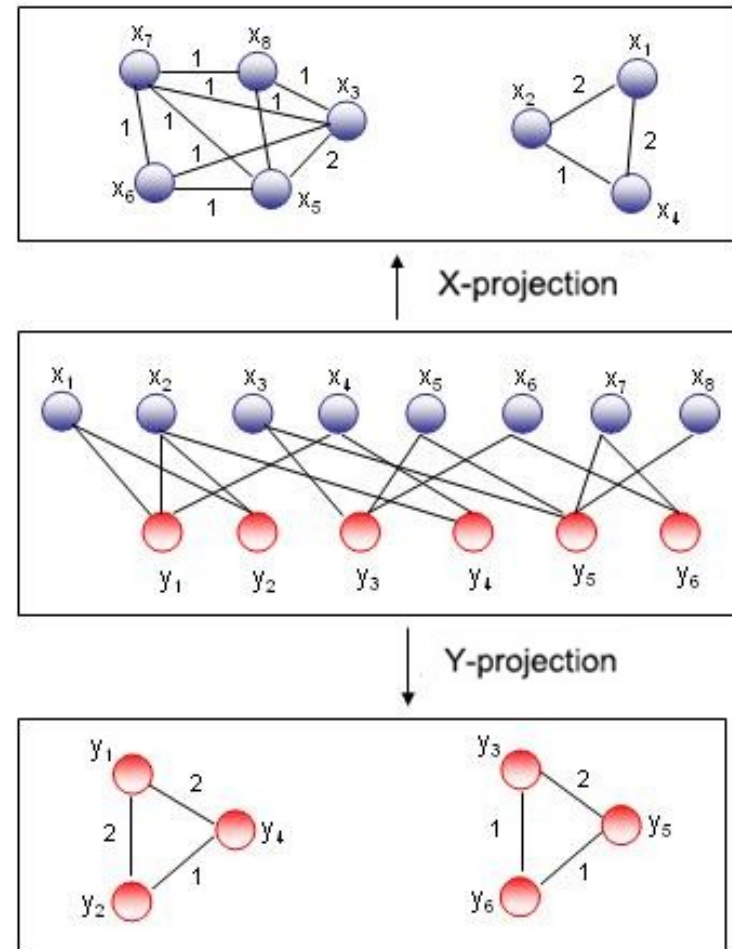
III.- Predictor de aprobación de la ley

Índice de tendencia política por partido



IV.- Análisis de asociaciones usando SNA

- Aplicación de técnica de proyección de grafos bipartitos persona – documento
- Combinación con búsqueda por texto + visualización



IV.- Análisis de asociaciones usando SNA

Búsqueda por palabra “AFP” sobre mensajes twitter

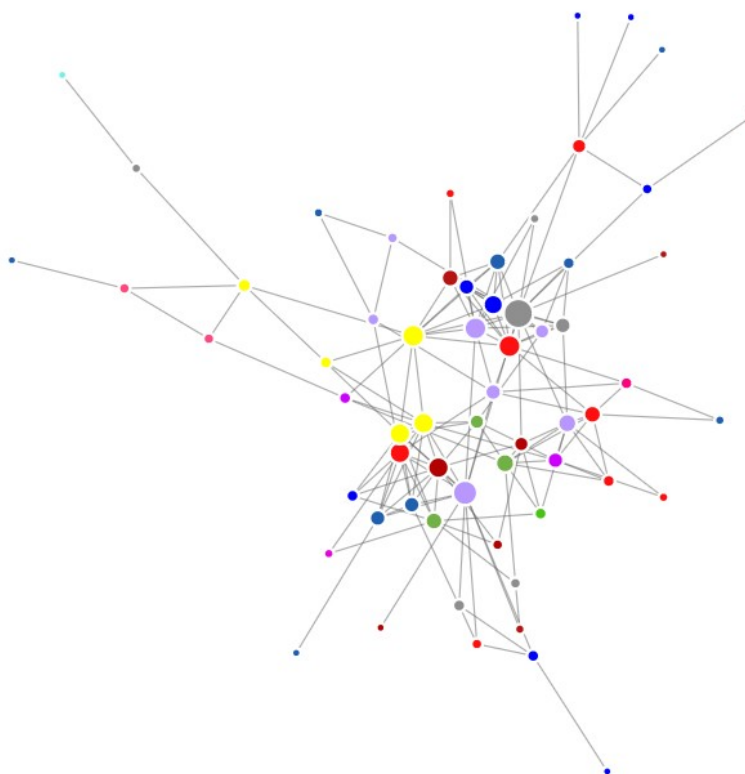
Buscar una persona 

Ximena Cecilia Rincón González	18 conexiones
Raúl Soto Mardones	14 conexiones
Yasna Provoste Campillay	12 conexiones
Alfonso De Urresti Longton	12 conexiones
Juan Ignacio Latorre Riveros	12 conexiones
Maite Orsini Pascal	11 conexiones
Gael Yeomans Araya	11 conexiones
Gastón Saavedra Chandía	11 conexiones
Karol Cariola Oliva	11 conexiones
David Sandoval Plaza	10 conexiones

10 de 62

Filtro por género

Red de co-ocurrencia



IV.- Análisis de asociaciones usando SNA

Búsqueda por palabra “AFP” sobre mociones y proy. acuerdo

Buscar una persona 

Carlos Bianchi Chelech **25 conexiones**

Ricardo Lagos Weber **21 conexiones**

José Miguel Insulza Salinas **20 conexiones**

Pedro Araya Guerrero **20 conexiones**

Francisco Huenchumilla Jaramillo **20 conexiones**

Francisco Chahuán Chahuán **20 conexiones**

Alfonso De Urresti Longton **19 conexiones**


Jorge Soria Quiroga **19 conexiones**

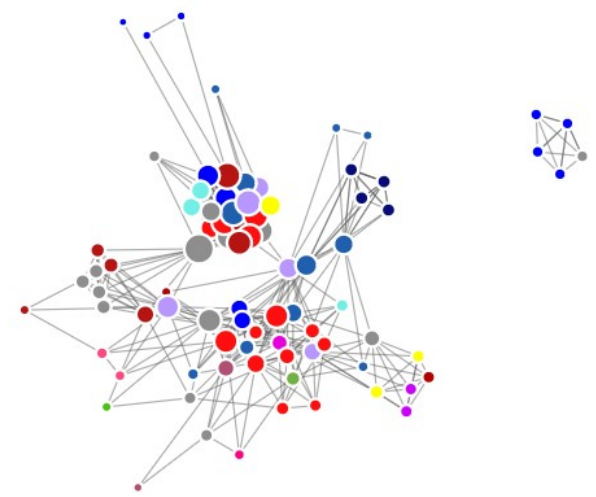
Ximena Ordenes Neira **19 conexiones**

Álvaro Elizalde Soto **19 conexiones**

10 de 86

Filtro por género

Red de co-ocurrencia 



Conclusiones

1. Los datos son el **nuevo petróleo**
2. No todo es Chat GPT (o LLM's)
3. Existen técnicas de IA para una infinidad de problemas entregando nuevas posibilidades de solución
4. Eficiencia en análisis de información para el parlamento

A stylized, minimalist figure composed of light gray lines and shapes. The figure has a large, solid red circle for a head. Its body is a vertical line, and its arms are represented by curved lines extending from the shoulders. The figure is positioned behind the main text.

Muchas gracias