



USB

EP-4793 MINI-PROYECTO DE SOFTWARE

# T-CREo: Aplicación de Credibilidad en Twitter

Por:

Daniela Ramírez  
Jonathan Bautista  
Ka Fung  
Valeria Vera

Tutores:

Yudith Cardinale  
Irvin Dongo  
Ana Aguilera

# TABLA DE CONTENIDO

01 DESCRIPCIÓN, TRABAJOS PREVIOS Y OBJETIVOS

02 DISEÑO

03 IMPLEMENTACIÓN

04 PRUEBAS

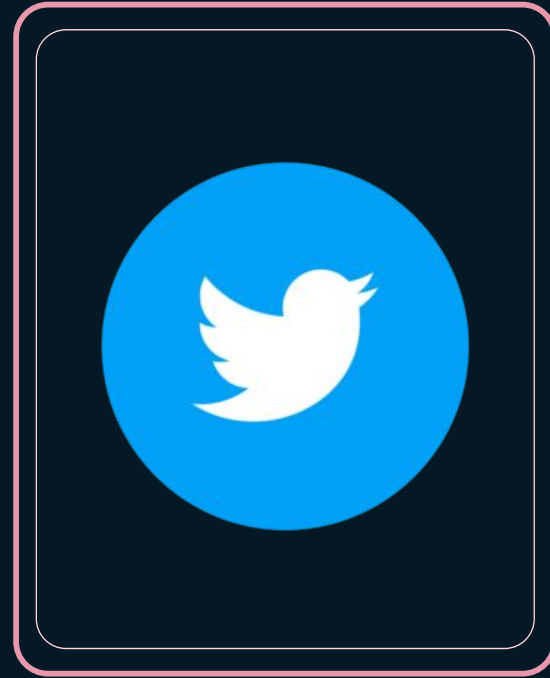
05 CONCLUSIONES Y TRABAJOS FUTUROS

# DESCRIPCIÓN

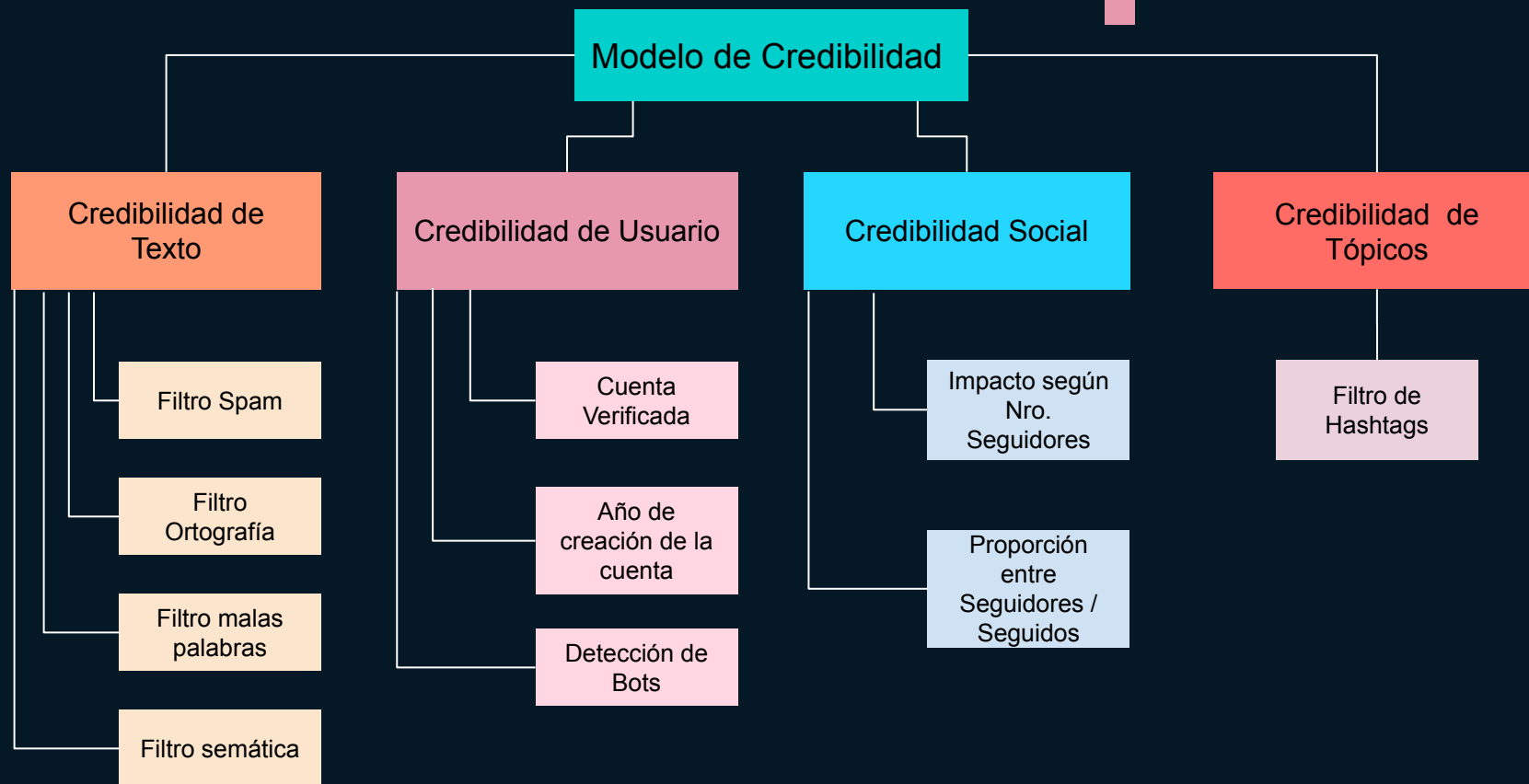
01

# T-CREo

**T-CREo** (Twitter CREDibility analysis framework) es un framework (extensión en Google Chrome) que realiza un análisis de credibilidad de tweets, de manera automática y en tiempo real.



# T-CREo: Modelo de credibilidad



# Clasificación de credibilidad

- **Texto:** mide el nivel de relevancia y precisión del texto, independientemente del tema referenciado o con respecto a un tema determinado.
- **Usuario:** calcula la credibilidad de la cuenta de usuario en función de los atributos que la describen, como la fecha de creación de la cuenta, si la cuenta está verificada y si es un bot.
- **Social:** calcula la credibilidad de una publicación, relacionada o no con un tema, en base a los metadatos disponibles que describen el impacto social de la cuenta de usuario y la publicación en sí, con respecto a otros usuarios. Se calcula en función de datos como el número de seguidores y la cantidad de retweets.
- **Tópicos:** mide el nivel de similitud de tópicos entre el texto del tweet y los hashtags.

# TRABAJOS PREVIOS

## T-CREo: A Twitter Credibility Analysis Framework

Robayo, G., Cabeza, D. & Medina, S.

2020

2020

## Web Scraping versus Twitter API: A Comparison for a Credibility

Martínez, F., Quintero, Y. & Barrios, S.

## Credibility Analysis on Twitter Considering Topic Detection

Hernandez, M. & Cornejo, J.

2021

2022

## CrediBot: Applying Bot Detection for Credibility Analysis on Twitter

Quinteros, P.

# T-CREo: A Twitter Credibility Analysis Framework

- **FRONTEND**

- Extracción de datos: Twitter API y Scrapping
- Extensión en Chrome: Vista de Popup y Configuraciones

- **BACKEND**

- Credibilidad de texto  **$O(n)$**  con n cantidad de palabras:

$$\begin{aligned}\text{TextCred}(p.\text{text}) = & w_{\text{SPAM}} \times \text{isSpam}(p.\text{text}) + \\ & w_{\text{BadWords}} \times \text{bad\_words}(p.\text{text}) + \\ & w_{\text{MisspelledWords}} \times \text{misspelling}(p.\text{text})\end{aligned}$$

- Credibilidad de usuario  **$O(1)$** :

$$\text{UserCred}(p.\text{user}) = \text{Verif\_Weight}(p.\text{user}) + \text{Creation\_Weight}(p.\text{user})$$

- Credibilidad social  **$O(1)$** :

$$\text{SocialCred}(p.\text{user}) = \text{FollowersImpact}(p.\text{user}) + \text{FFProportion}(p.\text{user})$$



# Web Scraping versus Twitter API:

## A Comparison for a Credibility

### Web Scraping

Analizar etiquetas de hipertexto y recuperar información de texto sin formato incrustada en ellas

#### VENTAJAS

- Mucho más rápida, sin necesidad de aplicaciones externas.

#### DESVENTAJAS

- Se debe buscar cada campo de un tweet o un usuario.
- Programación específica de un sitio web.
- Sensible a los cambios del formato de la misma página web. Requiere de constante mantenimiento.

### Twitter API

API que facilita la interacción entre diferentes programas de software y el acceso a los servicios de Twitter

#### VENTAJAS

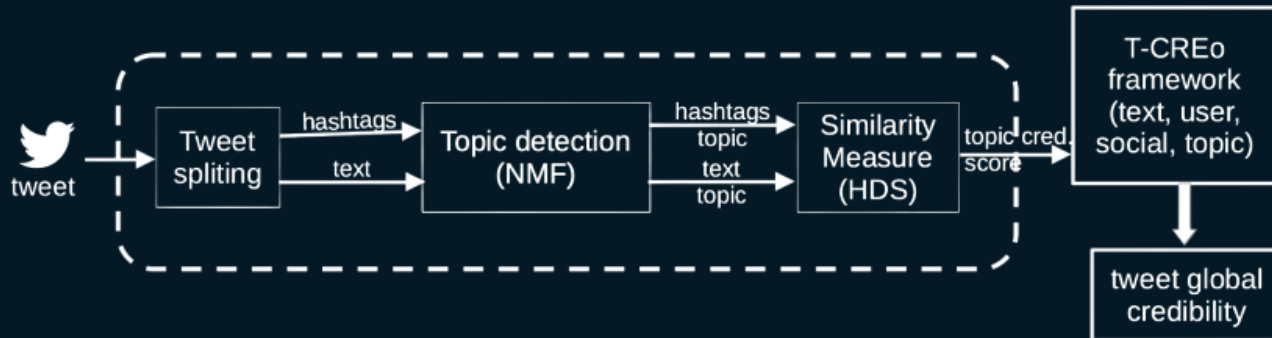
- Información más detallada y directa.
- Sólo requiere una solicitud a la API, sin necesidad de buscar cada campo de un tweet o un usuario.

#### DESVENTAJAS

- Requiere del Tweet ID y del nombre de usuario.
- Límites de volumen y tasa de solicitud.

# Credibility Analysis on Twitter Considering Topic Detection

- El tweet es preprocesado: Se separan el texto y los hashtags.
- Se identifica el tópico del texto y los tópicos de los hashtags con un modelo de *Factorización No Negativa de Matrices* (**NMF**) entrenado con 250 tópicos.
- Se calcula la similitud entre tópicos con la distancia Hellinger.



# CrediBot: Applying Bot Detection for Credibility Analysis on Twitter

- Dado un usuario de Twitter, extrae la cantidad de
  - Del usuario: Seguidores, siguiendo, estados, favoritos, listas.
  - De tweets: Texto, lenguaje, favoritos, retweets.
- Se identifica si el usuario es un bot con un modelo de *Random Forest*, entrenado con un dataset de 500.000 cuentas.
- Si el usuario trata de un bot, se penaliza su puntuación obtenido en la credibilidad del usuario:

$$UserCred'(p.user) = \begin{cases} UserCred(p.user) & \text{Si } p.user \text{ no es bot,} \\ 0,85 \times UserCred(p.user) & \text{Si } p.user \text{ es bot y } UserCred(p.user) \text{ mayor que 50,} \\ 0,75 \times UserCred(p.user) & \text{Si } p.user \text{ es bot y } UserCred(p.user) \text{ entre 35 y 50,} \\ 0 & \text{De lo contrario,} \end{cases}$$

# CrediBot: Applying Bot Detection for Credibility Analysis on Twitter

## Análisis semántico

- Dado un tweet, se realiza un análisis de qué tan similares y qué tan relacionadas están las palabras entre sí.
  - Tokenización: Se segmenta el texto en palabras y se realiza una “agregación” de palabras acortadas o abreviadas.
  - Revisión léxica: Cada palabra es identificada como sustantivo, adjetivo, verbo, adverbios. Se genera un árbol de análisis.
  - Revisión semántico:
    - Si tiene entidades (lugar, organización, persona), consulta a DBpedia y Wordnet para comparar coherencia entre entidades.
    - Si no, compara la coherencia entre palabras con Wordnet.
  - Cálculo del promedio de similitud.

# OBJETIVOS

01

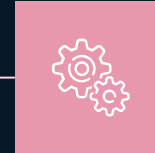
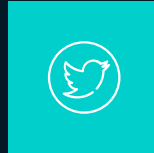
# PLANTEAMIENTO DEL PROBLEMA

- El empleo de *scraping* no resulta mantenible dado que varía según las actualizaciones de la aplicación Twitter.
- Desde Enero 2023, Chrome Extension ha modificado el empleo y creación de extensiones, al cambiar de Manifest V2 a Manifest V3.
- Desde Abril 2023, Twitter API ha cambiado de versión de v1.1 a v2.0, como también su uso gratuito a una suscripción mensual.



# OBJETIVOS

Eliminar métodos de acceso de información de Twitter secundarios como scrapping.



Refactorizar el código y brindar modularidad.

Incorporar los módulos de detección de tópicos, detección de bots y análisis semántico.



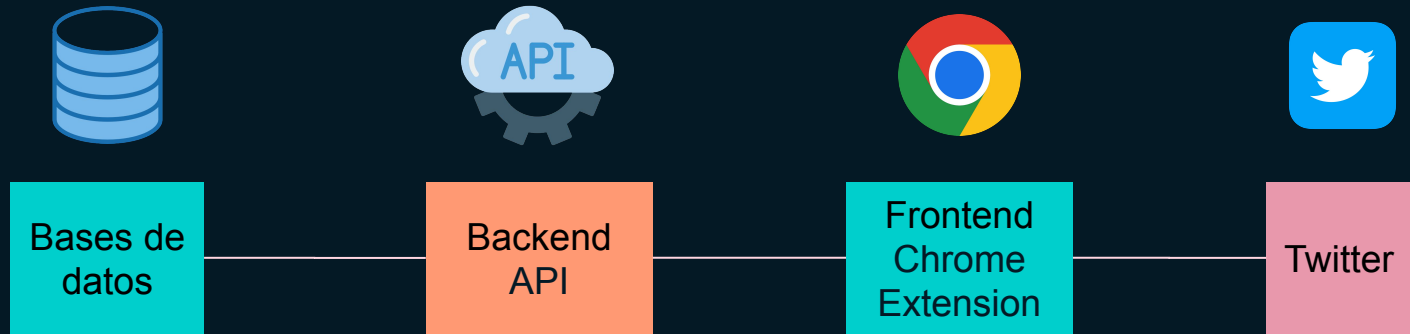
Actualizar el Plugin de Google Chrome con las funcionalidades de T-CREo.

ARQUITECTURA

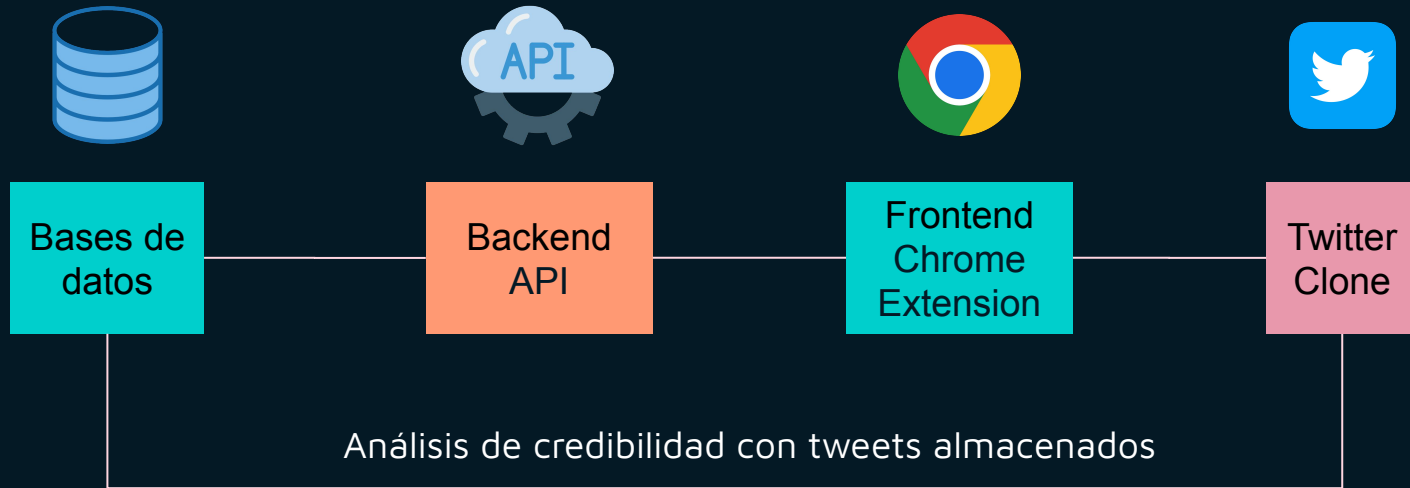
02



# ARQUITECTURA PREVIA (Con Twitter API)



# ARQUITECTURA ACTUAL (Sin Twitter API)



IMPLEMENTACIÓN

03

# Familiarización

- Se estudió los trabajos previos.
- Se configuró el ambiente de desarrollo en un repositorio privado en GitHub.
- Se realizó un estudio de las herramientas empleadas y a emplear en el proyecto.
- Se decidió refactorizar el proyecto desde cero, dado:
  - Limitaciones con Twitter API desde Abril 2023.
  - Limitaciones con Chrome Manifest V2 desde Enero 2023.
  - Librerías y frameworks obsoletos, difíciles de actualizar o incompatibles.

# Twitter Clone

The background is a dark navy blue. It is decorated with various geometric elements: small squares in solid colors (pink, teal, orange) and as thin white outlines, and thin white vertical lines of varying lengths scattered across the canvas.

# HERRAMIENTAS: Twitter Clone

ANTERIOR

ACTUAL

DESCRIPCIÓN

Lenguaje

-

TypeScript

Se escogió un lenguaje tipado.  
Más escalable.

Frameworks

-

React JS

Implementación con mayor  
rapidez y organizado en  
componentes. Uso de  
herramientas como Tailwind.

BD

-

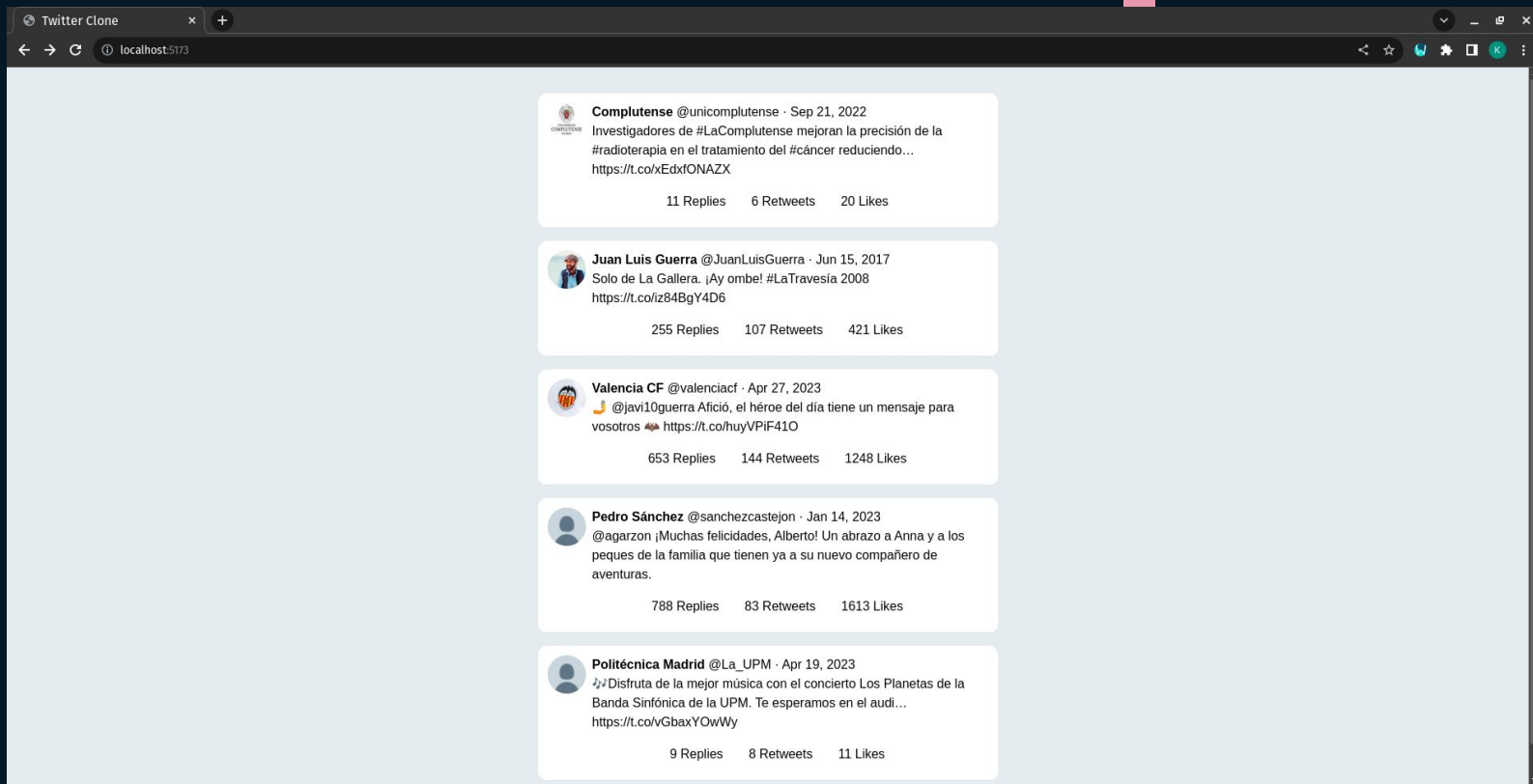
MongoDB

Complemento adecuado para  
TypeScript

# Implementación del Twitter Clone

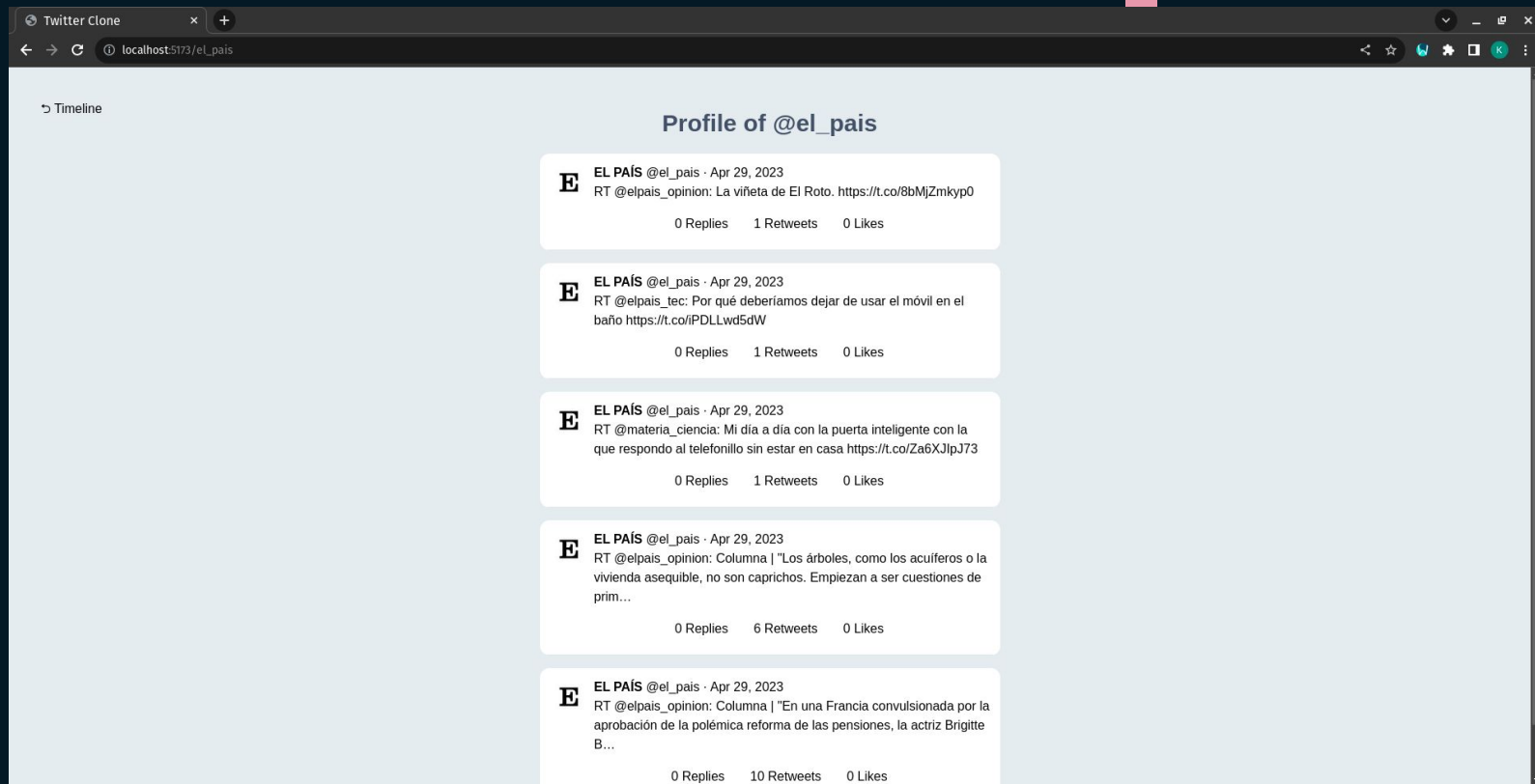
- Se conectó con la base de datos que contiene tweets recolectados.
- Se crearon las vistas de:
  - Timeline.
  - Usuario de Twitter.
- Dichas vistas se implementaron manteniendo las mismas etiquetas que posee la página de Twitter, para facilitar el scraping en caso de que se quiera migrar con Twitter API.

# Twitter Clone Timeline





# Twitter Clone Perfil



# Extensión T-CREo

The background features a dark navy blue field with a sparse, abstract pattern of geometric elements. These include thin, light-colored vertical lines of varying lengths and small squares in three colors: teal, pink, and orange. Some squares are solid, while others are hollow outlines. The elements are scattered across the frame, creating a modern, minimalist aesthetic.

# HERRAMIENTAS: Back-end Extensión

ANTERIOR      ACTUAL      DESCRIPCIÓN

Lenguaje

TypeScript

TypeScript

Se mantuvo el uso de TypeScript.

Frameworks

Express

Express

Se mantuvo el uso de Express por su flexibilidad.

BD

-

MongoDB

Complemento adecuado para Express

# Implementación del Back-end

- Se conectó con la base de datos proporcionada.
- Se instalaron nuevas dependencias:
  - *bad-words* para la detección de malas palabras.
  - *dictionary-en, dictionary-es, dictionary-fr* para la ortografía.
  - *npell* para la creación de verificadores de ortografía
- Se implementó nuevamente las funciones que realizan el análisis de:
  - Credibilidad del Texto
  - Credibilidad Social
  - Credibilidad del Usuario

# Módulo de Tópicos

## Integración en T-CREo

- Configuración de un entorno virtual de *Python*.
- Instalación de dependencias de *Python*.
- Modificación de funciones del cálculo de credibilidad de tópicos:
  - Verificación si el tweet posee hashtags.
  - Verificación si los hashtags están presentes en el modelo.
- Integración de *Python* a *Typescript*:
  - Uso de *node:child\_process* para generar subprocessos.
- Implementación de la función de la credibilidad de tópicos tanto en el front-end como el back-end de la extensión T-CREo.

# Módulo de Bots

## Integración en T-CREo

- Instalación y actualización de dependencias de *Python*.
- Modificación de funciones del cálculo de credibilidad de bots:
  - Eliminación del uso de Twitter API en *Python* (*Tweepy*).
  - Cambios menores en funciones de retorno de la predicción.
- Integración de Python a Typescript:
  - Uso de *node:child\_process* para generar subprocessos.
- Incorporación de la identificación de bot en el cálculo de credibilidad del usuario.

# HERRAMIENTAS: Front-end Extensión

ANTERIOR      ACTUAL      DESCRIPCIÓN

Lenguaje

JavaScript

TypeScript

Se escogió un lenguaje tipado.  
Más escalable.

Frameworks

-

React JS

Implementación con mayor  
rapidez y organizado en  
componentes. Uso de  
herramientas como Tailwind.

Chrome

V2

V3

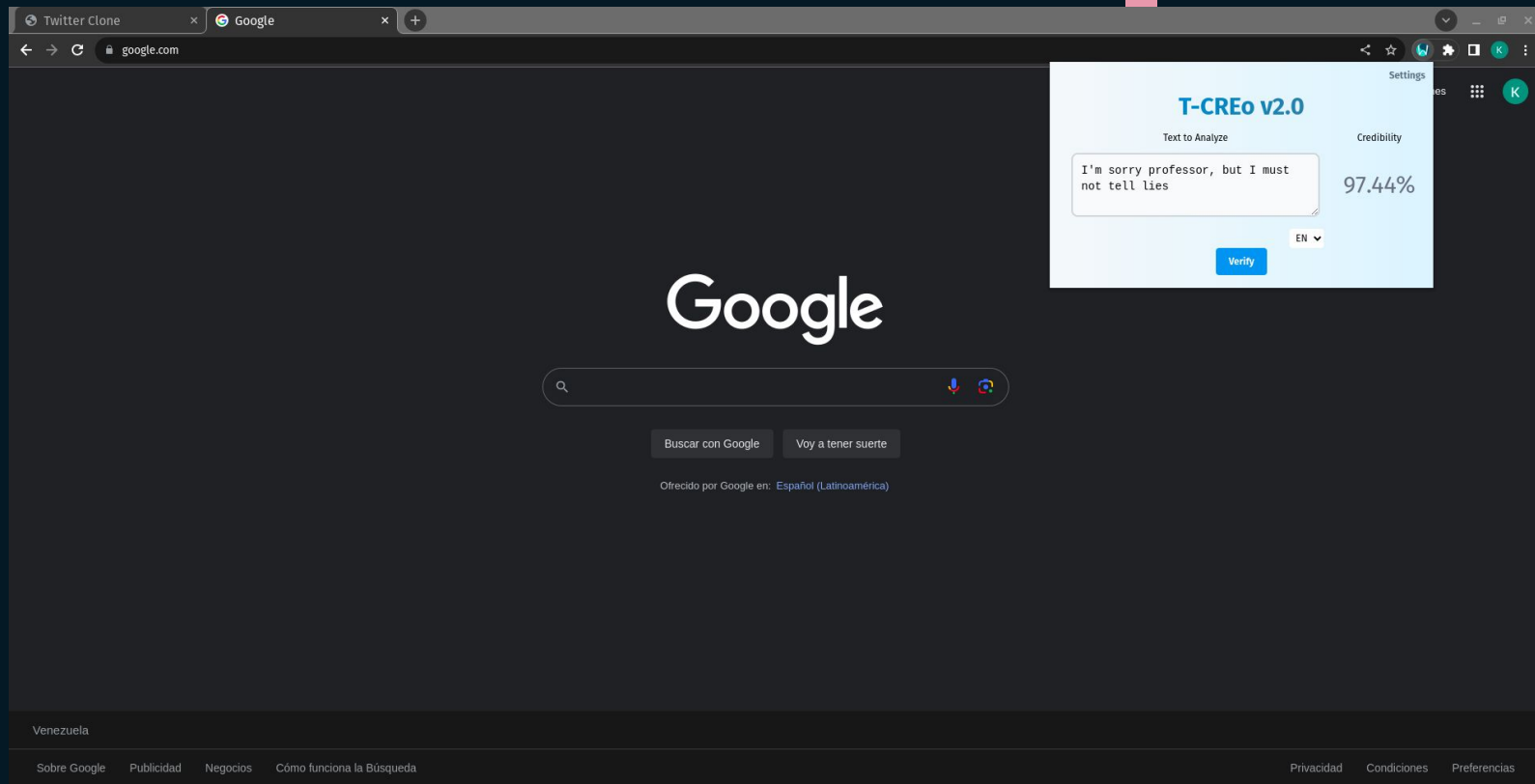
V2 se encuentra obsoleto.

# Implementación: Front-end de T-CREo

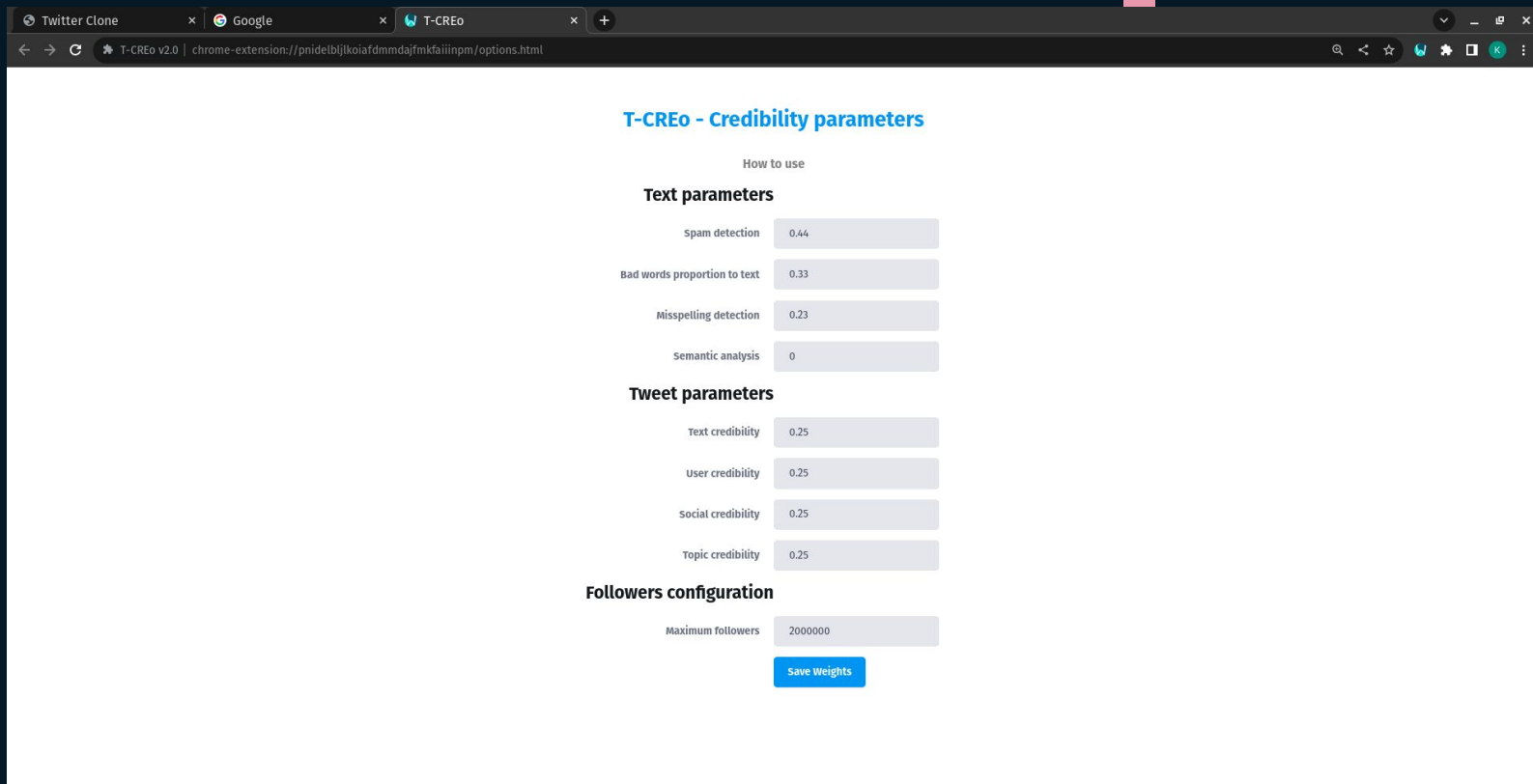
- Se reestructuró el front-end del trabajo previo para que sea adecuado con el framework de React.
- Se crearon las vistas de:
  - Popup: Vista principal de la extensión.
  - Options: Vista para la configuración de parámetros.
- Se conectó con Chrome Extension con Manifest v3, lo cual requirió actualizar la configuración y reestructurar el código.
- Se conectó el Front-end con el Back-end.



# FASE II: Popup de T-CREo



# FASE II: Configuración de T-CREo



The screenshot shows a web browser with three tabs: 'Twitter Clone', 'Google', and 'T-CREo'. The address bar displays 'chrome-extension://pnidelbljlkoiafdmdmdajfmkfalliinp/ options.html'. The page content is titled 'T-CREo - Credibility parameters' and includes a 'How to use' section. Below this, there are three main configuration sections: 'Text parameters', 'Tweet parameters', and 'Followers configuration'. Each section contains several input fields with numerical values. A 'Save Weights' button is located at the bottom right of the configuration area.

## T-CREo - Credibility parameters

How to use

### Text parameters

Spam detection	0.44
Bad words proportion to text	0.33
Misspelling detection	0.23
Semantic analysis	0

### Tweet parameters

Text credibility	0.25
User credibility	0.25
Social credibility	0.25
Topic credibility	0.25

### Followers configuration

Maximum followers	2000000
-------------------	---------

[Save Weights](#)

# Twitter Clone + T-CREo

The screenshot displays a web browser window titled "Twitter Clone" at the URL "localhost:5173". The main content area shows a feed of tweets. Each tweet includes a user profile picture, name, handle, date, text, and engagement metrics (Replies, Retweets, Likes). A green "T-CREo Credibility" score is displayed below each tweet. A modal window titled "T-CREo v2.0" is overlaid on the right, showing a "Text to Analyze" input field with the text "I'm sorry professor, but I must not tell lies", a "Credibility" output of "0%", a "Verify" button, and a message "You are currently on a Twitter page" with a "Verify Page Tweets" button. The browser's address bar and various icons are visible at the top.

Twitter Clone

localhost:5173

**Rafa Nadal** @RafaelNadal · Mar 27, 2020  
RT @MallorcaChamps: Con orgullo nos unimos 🌱 a la genial iniciativa de @RafaelNadal y @paugasol con #CruzRojaResponde a la que está respond...

0 Replies 115 Retweets 0 Likes

T-CREo Credibility: 82.81%

**Politécnica Madrid** @La\_UPM · Dec 15, 2022  
Sobre #tecnociudades y artefactos urbanos evolutivos ahondarán expertos de la #UPM, como Fernando Vela, José Maria...  
<https://t.co/U6oQvAn6vX>

0 Replies 0 Retweets 0 Likes

T-CREo Credibility: 70.32%

**Rafa Nadal** @RafaelNadal · Nov 10, 2017  
RT @Kia\_Motors: We believe in the power to surprise. See how #Kia's special visit to Keon Park Tennis Club made everyone smile - the #AusOp...

0 Replies 56 Retweets 0 Likes

T-CREo Credibility: 78.4%

**Politécnica Madrid** @La\_UPM · Nov 24, 2022  
Mañana comienza #Climathon Madrid 2022 con el objetivo de buscar soluciones a la #emergenciadclimática. Porque todos...  
<https://t.co/MgtJEFgxZu>

0 Replies 1 Retweets 1 Likes

T-CREo Credibility: 70.73%

**UC3M** @uc3m · May 11, 2022

**T-CREo v2.0**

Text to Analyze

I'm sorry professor, but I must not tell lies

Credibility

0%

EN

Verify

You are currently on a Twitter page

Verify Page Tweets



PRUEBAS

04

# Pruebas de limpieza de texto

Prueba	Estado	Tiempo ejec. BD cloud	Tiempo ejec. BD local
Pruebas de remoción de hashtags	Pasó	73 ms	104 ms
Prueba de remoción de menciones	Pasó	17 ms	21 ms
Prueba de remoción de emojis	Pasó	37 ms	43 ms
Prueba de remoción de puntuación	Pasó	68 ms	25 ms
Prueba de remoción de espaciado extra	Pasó	39 ms	15 ms
Prueba de remoción de enlaces	Pasó	30 ms	15 ms

# Pruebas de credibilidad de texto

Prueba	Estado	Tiempo ejec. BD cloud	Tiempo ejec. BD local
Textos con malas palabras	Pasó	6371 ms	248 ms
Textos con spam	Pasó	132 ms	70 ms
Textos con errores ortográficos	Pasó	245 ms	111 ms
Textos con mezcla de pruebas mencionadas	Pasó	194 ms	102 ms

# Pruebas de manejos de errores

Prueba	Estado	Tiempo ejec. BD cloud	Tiempo ejec. BD local
Retornar Error 400	Pasó	20 ms	15 ms
Retornar Error 401	Pasó	31 ms	3 ms
Retornar Error 403	Pasó	31 ms	2 ms
Retornar Error 404	Pasó	31 ms	3 ms
Retornar Error 500	Pasó	33 ms	3 ms

# Pruebas de validación de credibilidad de texto

Prueba	Estado	Tiempo ejec. BD cloud	Tiempo ejec. BD local
Pruebas de validación en textos simples	Pasó	6898 ms	428 ms
Pruebas de validación en tweets	Pasó	15536 ms	10628 ms



# Prueba de estatus del endpoint

Prueba	Estado	Tiempo ejec. BD cloud	Tiempo ejec. BD local
/health endpoint retorna status 200	Pasó	11720 ms	154 ms

# Comparación de los tiempos de ejecución con BD local y en cloud

N. de tweets analizados	Tiempo ejec. nube	Tiempo ejec. local
5	15679 ms	41881 ms
10	18724 ms	78548 ms
20	94255 ms	147700 ms
30	48222 ms	210320 ms
40	47787 ms	309684 ms
50	70222 ms	353782 ms

N. de tweets analizados	Tiempo ejec. nube	Tiempo ejec. local
60	82833 ms	418230 ms
70	56231 ms	392061 ms
80	72855 ms	434683 ms
90	81209 ms	475613 ms
100	106087 ms	589190 ms
AVG	63059,45	313,790.18

# CONCLUSIONES Y TRABAJOS FUTUROS

05

# CONCLUSIONES Y TRABAJOS FUTUROS

- Se presenta una nueva versión del framework capaz de crear análisis de credibilidad más concisos gracias a la detección de tópicos y detección de bots.
- Considerar nuevos parámetros que puedan surgir con los constantes cambios en Twitter y puedan mejorar el framework.
- Continuar con T-CREo en redes similares a Twitter, tales como Facebook Threads y Mastodon.



GRACIAS POR  
SU ATENCIÓN

