



05/13/2025

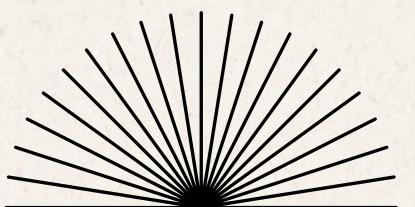


# DEPLOYMENT PLAN

**Project: News article popularity analysis**

**PRESENTED BY:**

Márton Salamon, Olivér Karsai



# Agenda

03	<b>Project Objectives and Scope</b>
04	<b>Data Protection</b>
05	<b>Technical Deployment Requirement</b>
06	<b>Personal Deployment Requirements</b>
07	<b>Deployment Timeline</b>
08	<b>Data Governance</b>
09	<b>Prototype</b>
10	<b>Post-Deployment</b>

# Project Purpose and Metrics

- Improve the efficiency of newspaper agency (Telex) by analyzing the popularity of historical news article data from their website
- Improved efficiency expected to lower costs and increase revenue

## Project Performance Metrics

**Cost Savings: \$52,000**

Sped-up decision making

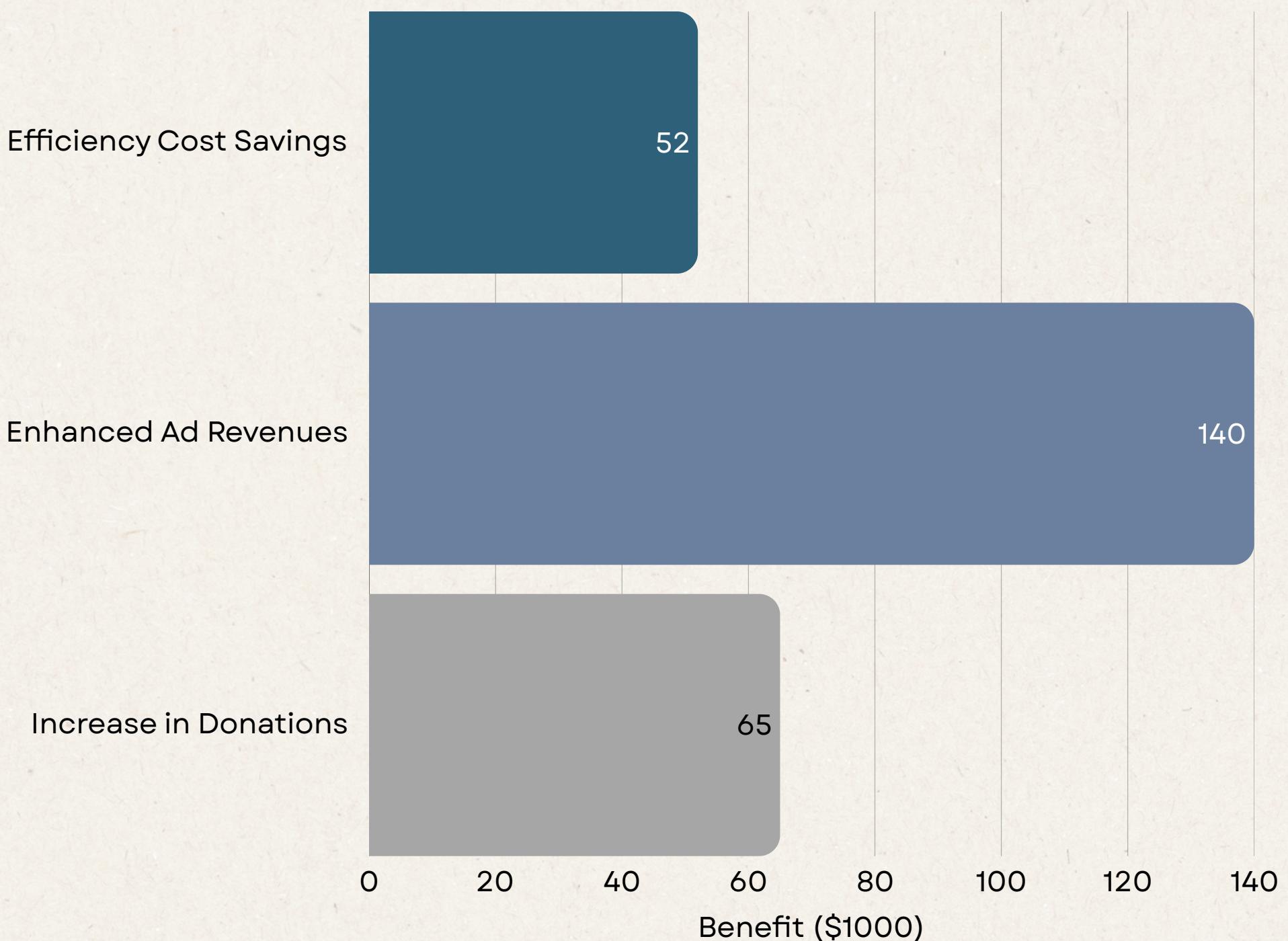
**Increased Ad Revenues: \$140,000**

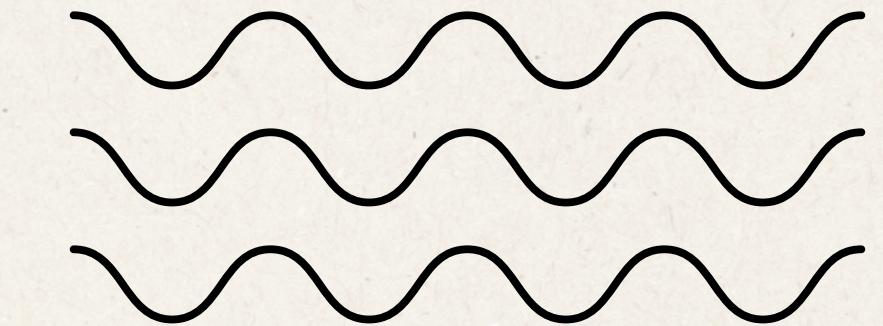
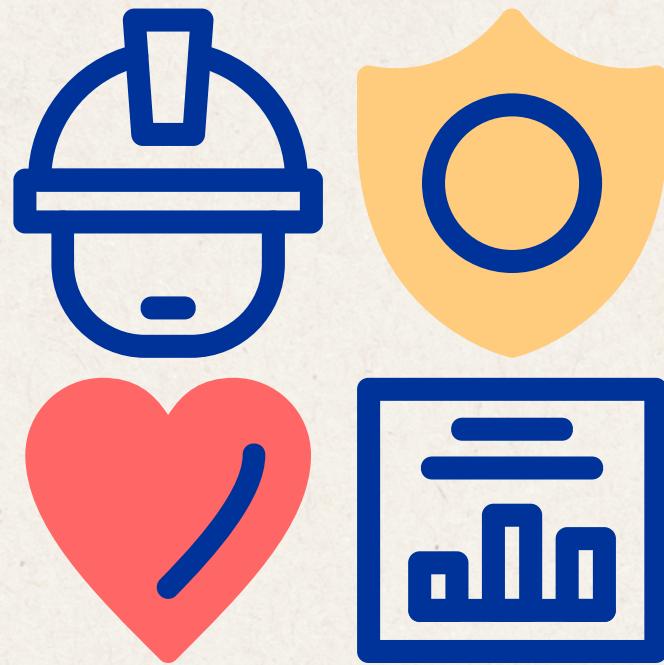
Popular articles generate more clicks

**Increase in Donation Revenue: \$65,000**

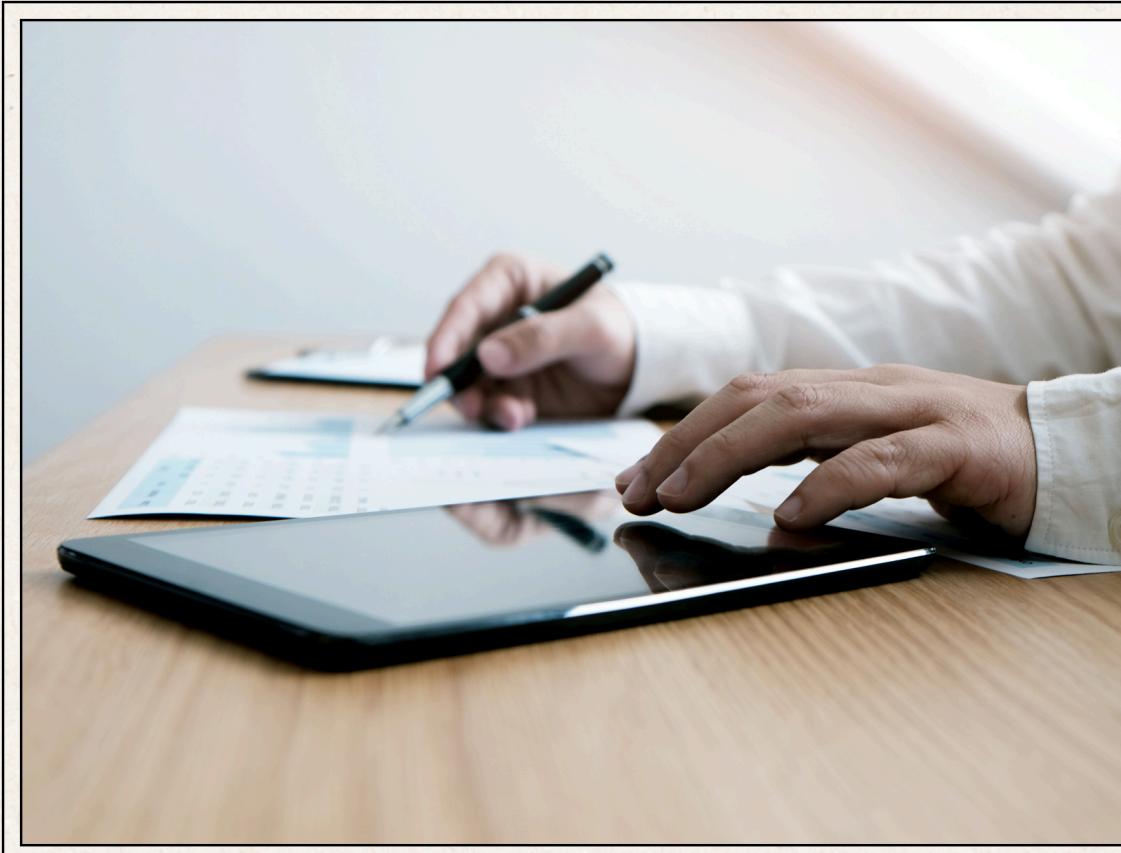
Popular articles receive more donations

**Total Benefit: \$257,000**





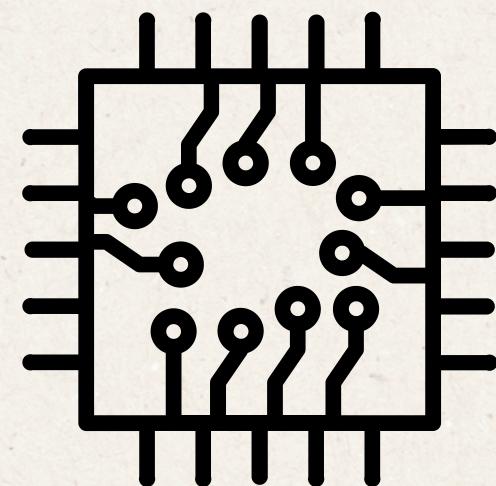
# Processed Data and Data Protection



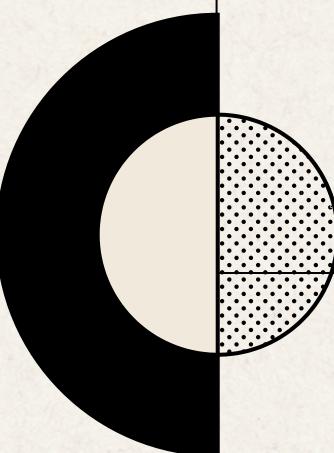
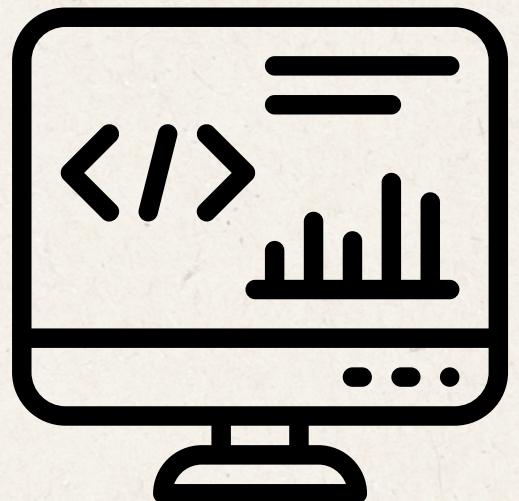
- Used news article data in the model
  - Article text
  - Author names
  - Date of the article
  - URL
  - Article categorization
  - Social media shares
- All these data are internal, owned by the company

# Technical Deployment Requirements

## HARDWARE



## SOFTWARE



	<b>On-premise or cloud?</b>	On-premise solution
<b>Deployment platform</b>		<ul style="list-style-type: none"> <li>• Server with dockerization environment</li> <li>• Streamlit: free open-source data science web app</li> </ul>
<b>Server specification</b>		High performance needs for running model (RAM heavy)
<b>Analysis tools</b>		<ul style="list-style-type: none"> <li>• spaCy: free open-source NLP library</li> <li>• Gensim: free open-source document vectorization library</li> </ul>

## 01 Users of the System

Article authors

## 02 Stakeholders

- Content production management
- HR
- Sales

## 03 Change Management

- Convince authors of added value and efficiency improvement
- Hold trainings
- Give incentives, reward early adopters

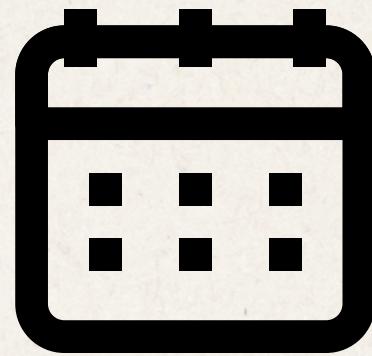


# Personnel Deployment Requirements

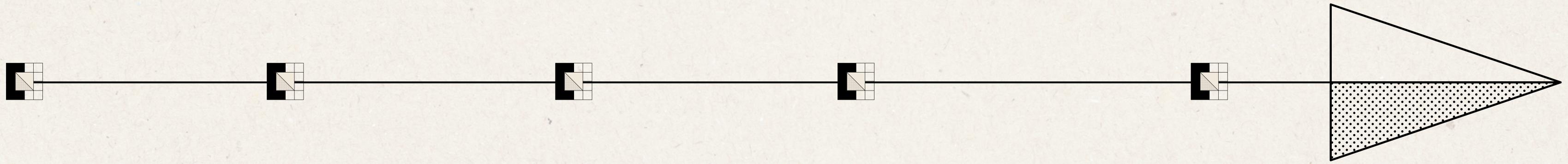
The adoption of the new application by the article writers to optimize their writing style is an important barrier to overcome

# Deployment Timeline

Development and initial deployment timeline of the news article analytical application



07/11



## **Define Phase**

Development of the prototype

## **Measurement Phase**

Unit testing, Model validation, and Performance testing

## **Analysis Phase**

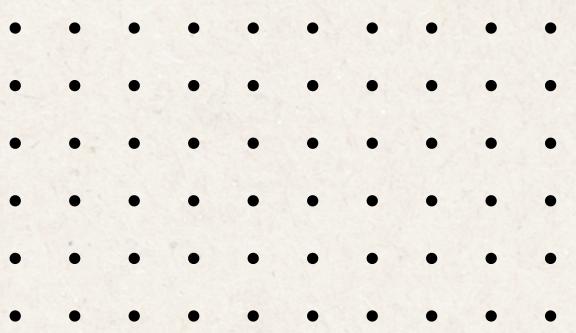
Analyzing testing results and making changes accordingly

## **Go-Live - Initial Deployment Phase**

Making available the first version

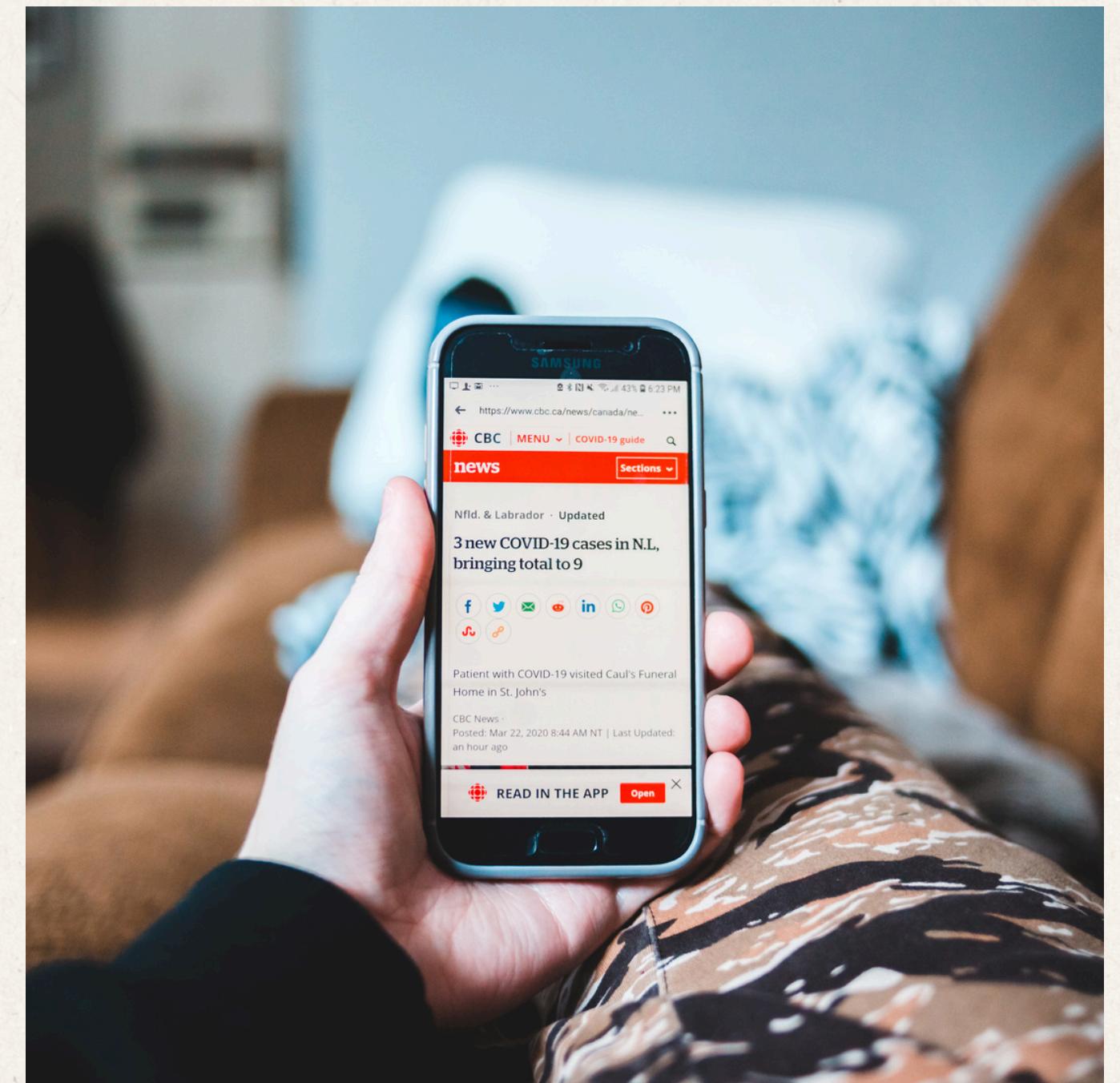
## **Improvement Phase**

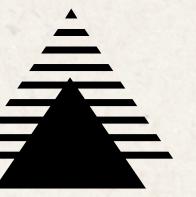
Post-deployment evaluation and improvement



# Data Governance

- **Access Controls:** authors are only authorized to access aggregated data, not on author level
- **Data Quality:** stop word dictionaries and lemmatization are used to filter out non-sensical data
- **Backup and Recovery Plan:** the model is executed monthly, with the output data first backed-up, then used in the application





# Prototype

Streamlit application

09/11

## Telex Article Topic Modeling

Select Date Range

2025/01/01 – 2025/01/31

Select All Tags

Number of articles selected: 1020

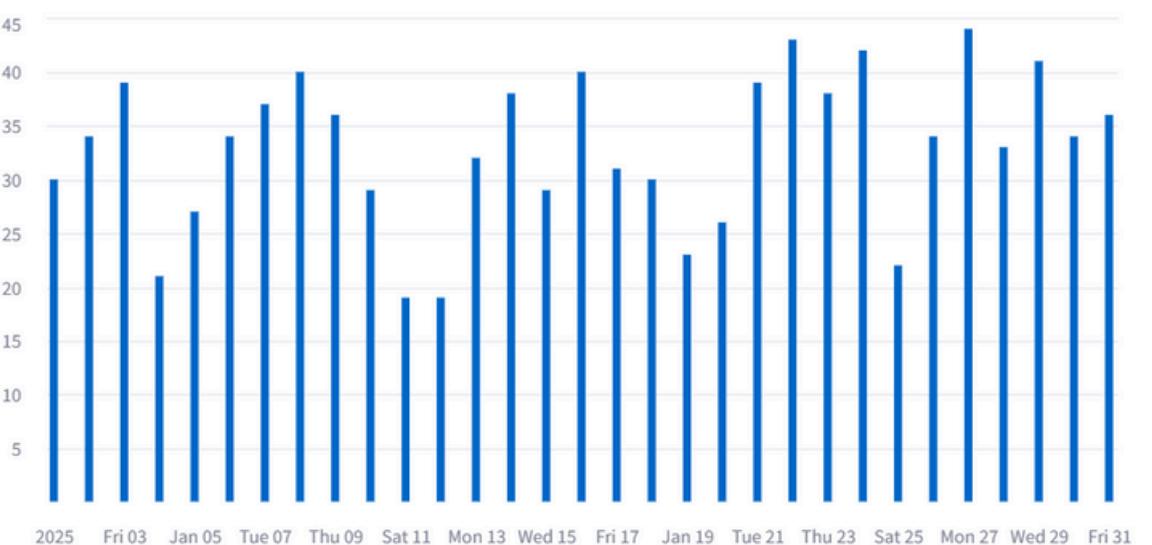
Run Topic Modeling

Modeling completed!

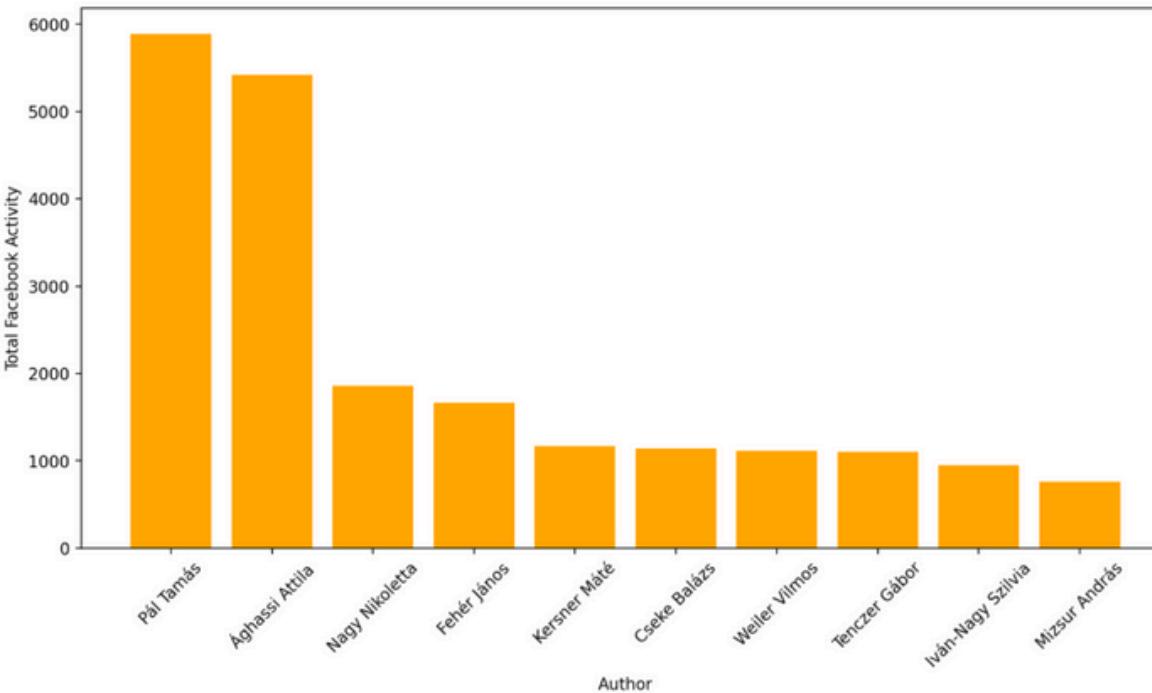
Download KÜLFÖLD



Number of Articles by Date



Top 10 Authors by Total Facebook Activity



### Features

Article category and Date filtering

Topic modeling determining the most used key words

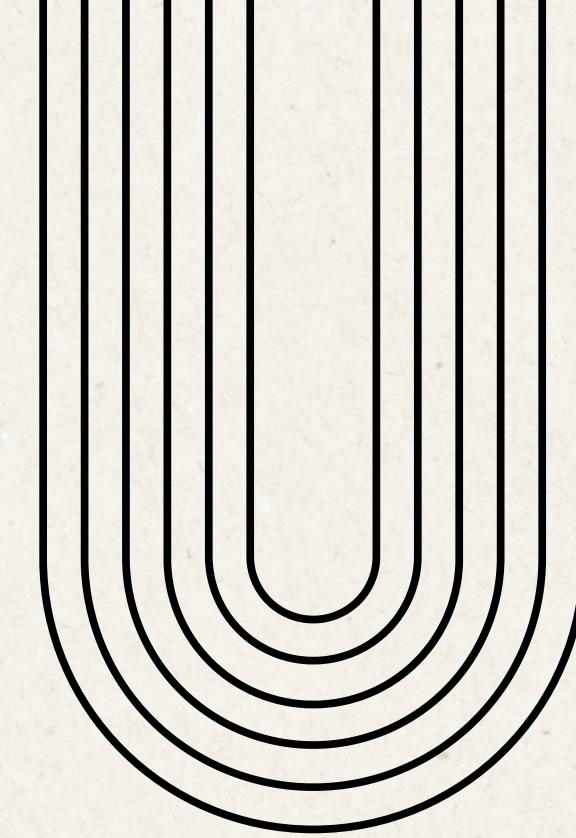
Displaying social media activity per topic

Charts for distribution of the number of articles

Author ranking based on social media activity

Download analysis results

# Post-Deployment Plan



Potential improvements to the prototype and the initial release version



## Performance tuning

Improve modeling speed for larger article categories



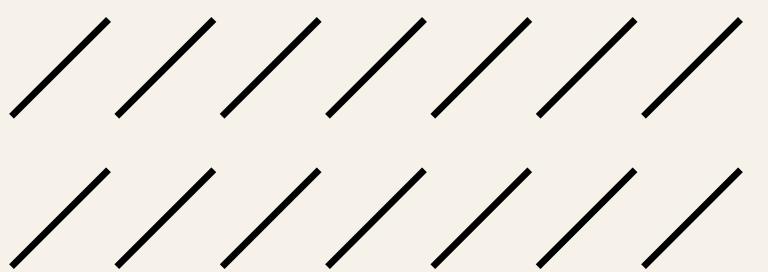
## Plan for scalability

Incorporate larger datasets (such as yearly data)



## Login feature

With documentation established, users will need to login to access website



# Thank you for your attention!

## PROJECT DETAILS

**Name** News article popularity analysis

**Owners** Márton Salamon, Olivér Karsai

