

Suite for Tweet Enhancement STE

ANTONIO COÍN CASTRO

LUIS ANTONIO ORTEGA ANDRÉS

ÁNGEL GABRIEL RIVAS FONFRÍA

IGNACIO SALINAS SÁNCHEZ

Index

01 SCIENTIFIC PROPOSAL

- ✓ Background and state of the art
- ✓ Objectives and SWOT (DAFO)
- ✓ Methodology and work plan

02 EXPECTED RESULTS IMPACT

03 TRAINING CAPACITY

Main Researchers

Information Retrieval Group



Lara
Quijano

IP1



Iván
Cantador

IP2

SCIENTIFIC PROPOSAL

SCIENTIFIC
PROPOSAL

1

EXPECTED
RESULTS
IMPACT

2

TRAINING
CAPACITY

3

Background



Twitter



Machine
Learning

Current Status



Topic
modeling



Predicting
Popular
Content

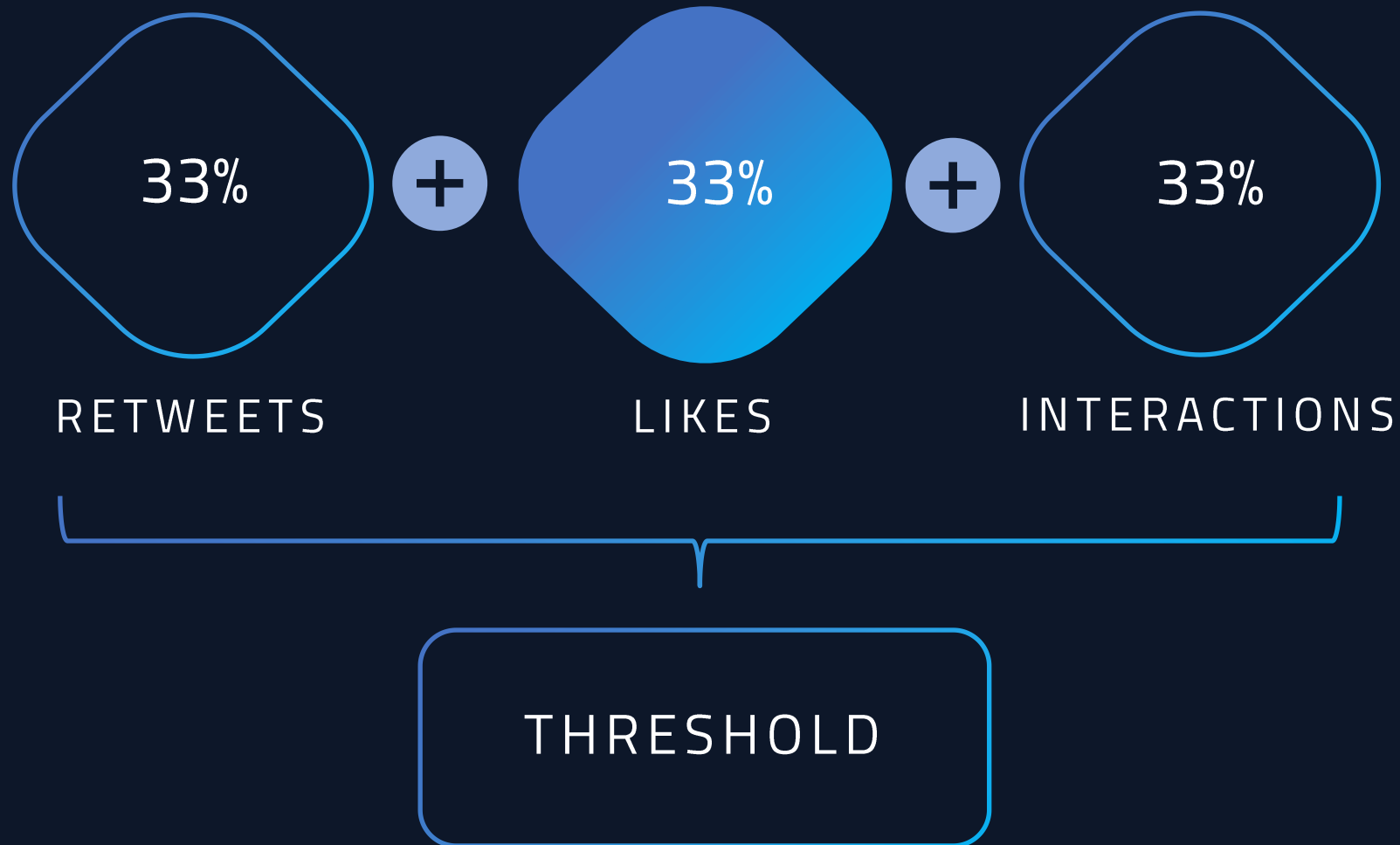


Hashtag, mention
and time
recommendation



Paraphrasing
tweets

Popularity Metric



SWOT Analysis

✓ STRENGTHS

- Machine Learning knowledge
- Comprehensive survey of the state of the art
- Working knowledge of data privacy regulations
- Accordance with Spanish R&D+I Strategy in AI

☑ OPPORTUNITIES

- Possibility of receiving project funding
- Attend international conferences
- Improve presence of public entities on social networks
- Build a collaborative community through free software

✗ WEAKNESSES

- Very stringent data privacy laws
- No formal protocols for data security standards
- Lack of expert knowledge in probabilistic models

⚡ THREATS

- Violation of the GDPR
- Low level of support from the general public
- Long-term project → Possible competitors

Objectives

FAMILY OF EXPLICIT MODELS FOR POPULARITY PREDICTION

- Popularity prediction model based on Tweet information.
- Extend the popularity prediction model.

RECOMMENDATION SYSTEM BASED ON POPULARITY

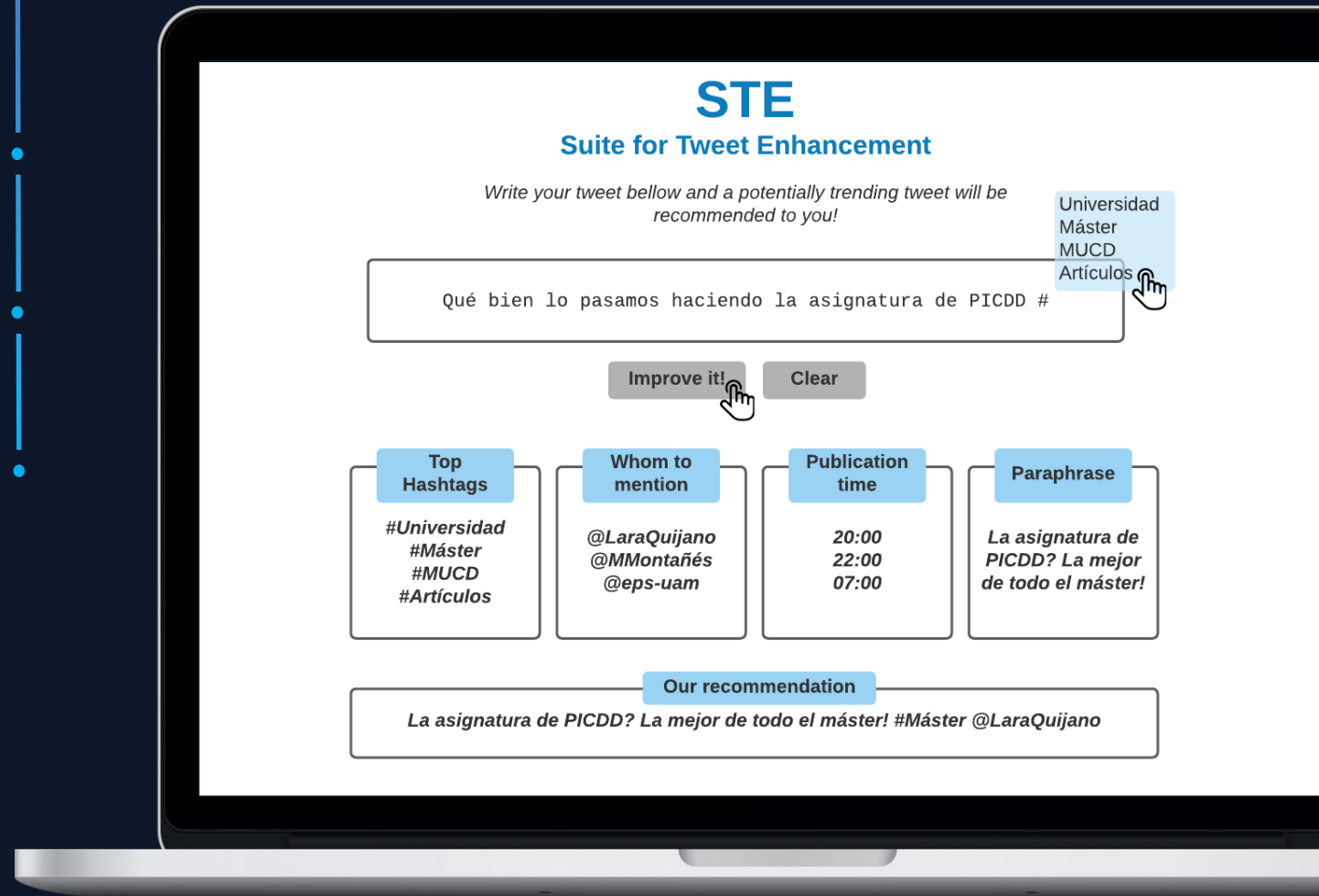
- Hashtag recommendation system.
- Publication date and time recommendation system
- Mention recommendations
- Paraphrasing the tweet
- Framework to periodically retrain the model

AGGREGATION OF RECOMMENDER SYSTEMS IN COMMON USER INTERFACE

- Unify the user interface and the recommendation system

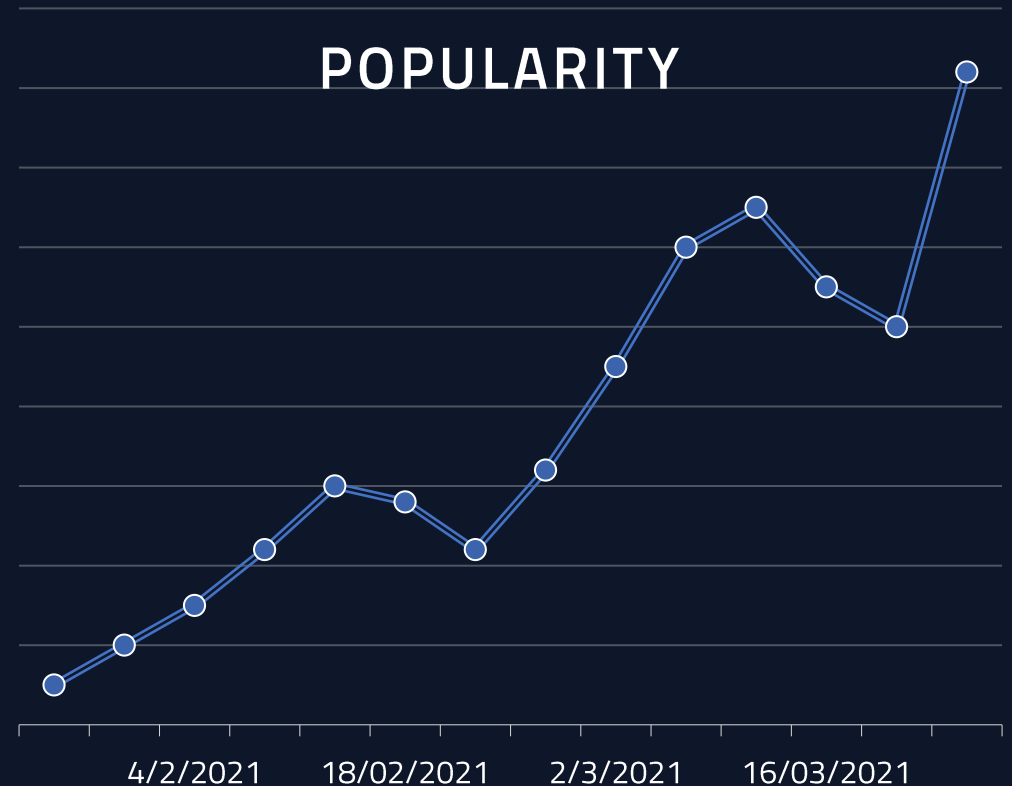
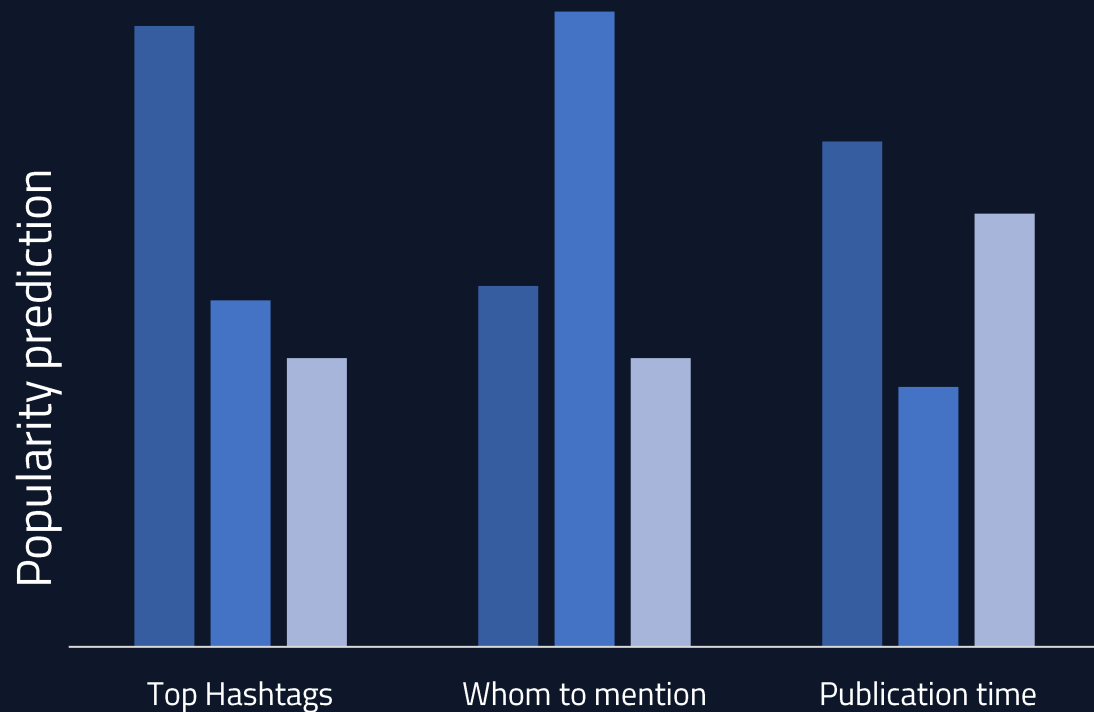
- User interface with predictive analysis results

- Create embedded web browser add-on for Twitter website



Dashboard

Unified user-friendly interface for efficient tweet tuning



Applications



**Public
Institutions**

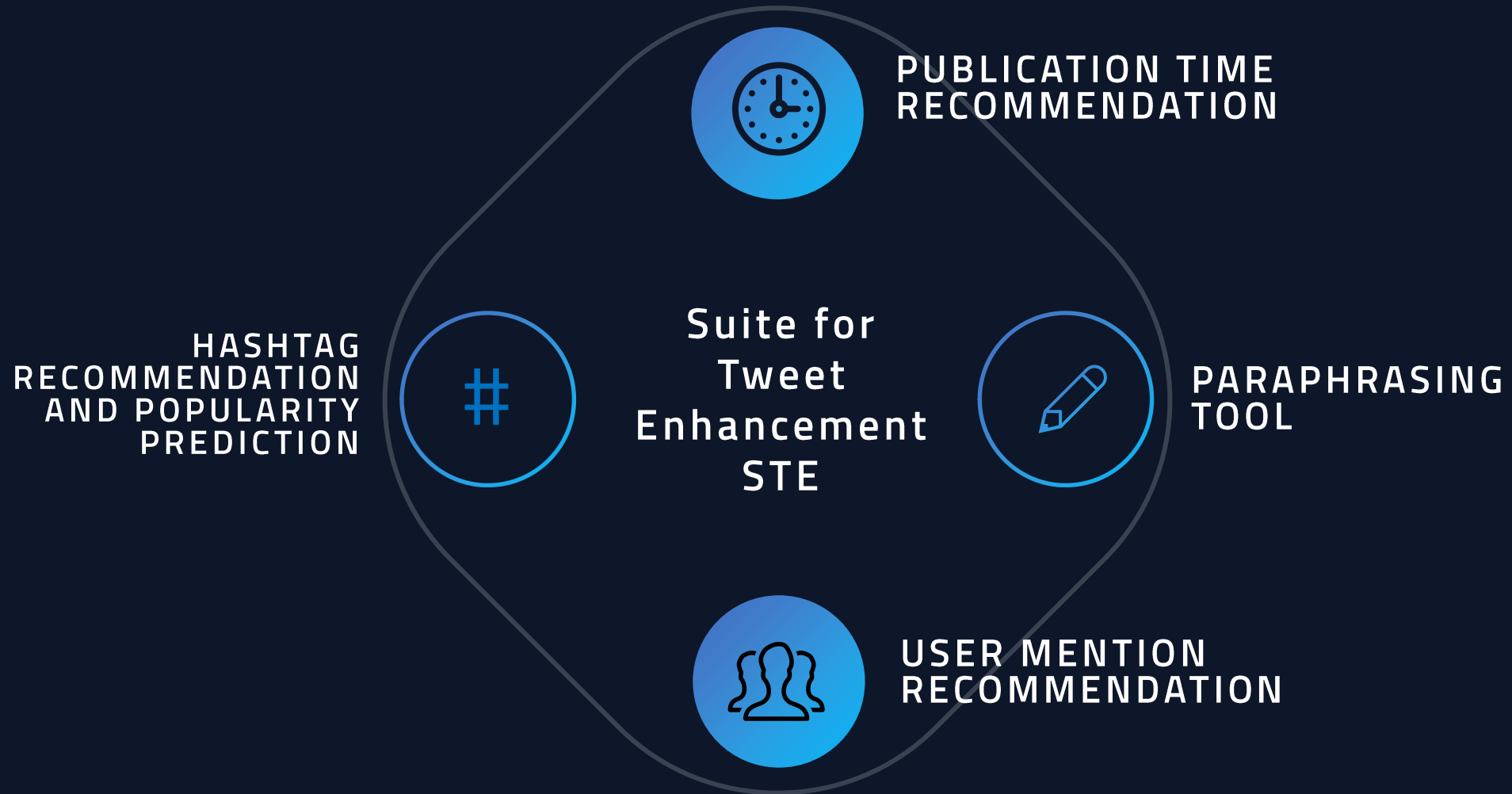


**General
Users**



**Private
Businesses**

Methodology



Schedule

YEAR 1

Popularity
Prediction Models

State of the art

Hashtags,
publication date and
twitter mentions

Tweet content

YEAR 2

A recommendation system

YEAR 3

User interface

State of the art

Twitter
recommendation
system

Paraphrasing tweet

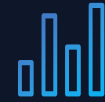
Predictive analysis
Interface

Recommendation
Interface

YEAR 4

Applications

Madrid cultural activities
publications impact



Local newspaper
Subscribers



Budget

	COST (€)	DURATION	QUANTITY	TOTAL COST
Programmer	2000 €/MONTH	4*12 MONTHS (14 PAYMENTS)	1	114000 €
Application Testers	100 €/MONTH	DEPENDING ON REQUIREMENTS	DEPENDING ON REQUIREMENTS	MAX 2500 €
Material (computers, servers, etc.)	-	-	4 COMPUTERS, 5 SERVERS, 4 MONITORS	10800 €
Congresses, dissemination and training	-	THE ENTIRE LIFETIME OF THE PROJECT	-	MAX 20000 €
TOTAL				147,000.00 €

EXPECTED RESULTS IMPACT

SCIENTIFIC
PROPOSAL

1

EXPECTED
RESULTS
IMPACT

2

TRAINING
CAPACITY

3

Scientific and technological impact

- ✓ Twitter Popularity Prediction
- ✓ Recommender System
- ✓ Other Social Media

Dissemination and internationalization

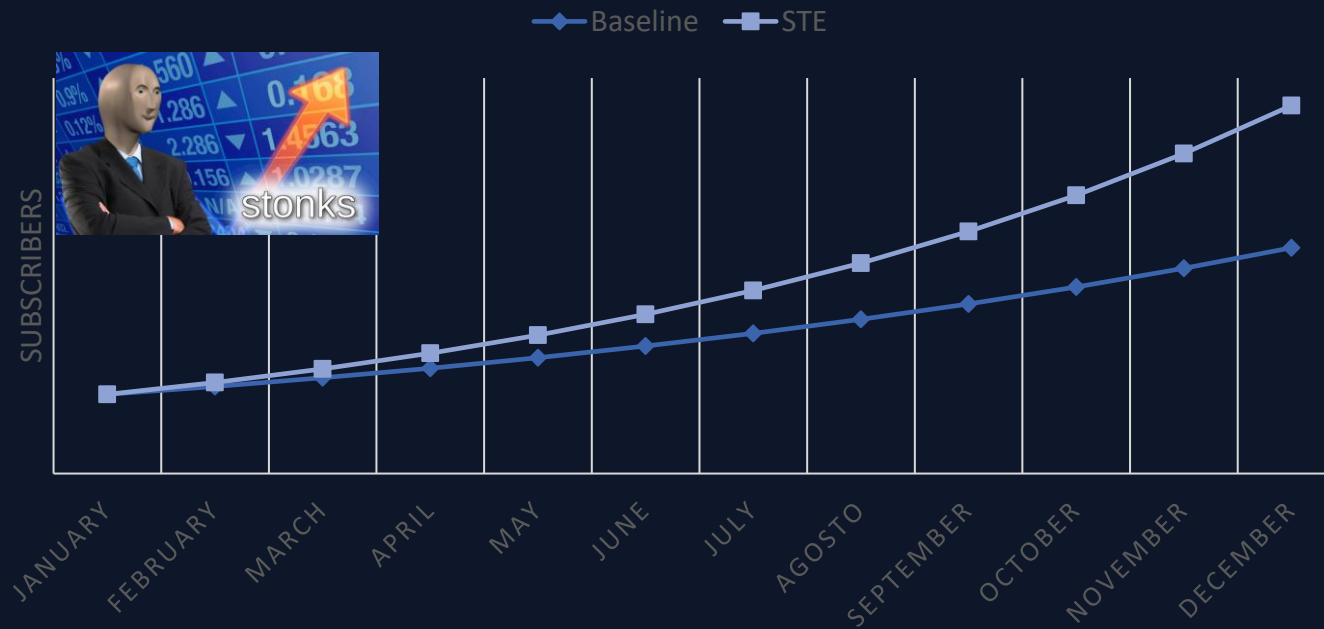


Knowledge transfer



Social and economic impact

Subscribers prediction



TRAINING CAPACITY

SCIENTIFIC
PROPOSAL

1

EXPECTED
RESULTS
IMPACT

2

TRAINING
CAPACITY

3

Training plan



Conferences



Internships



**International
Collaborations**

Conclusions

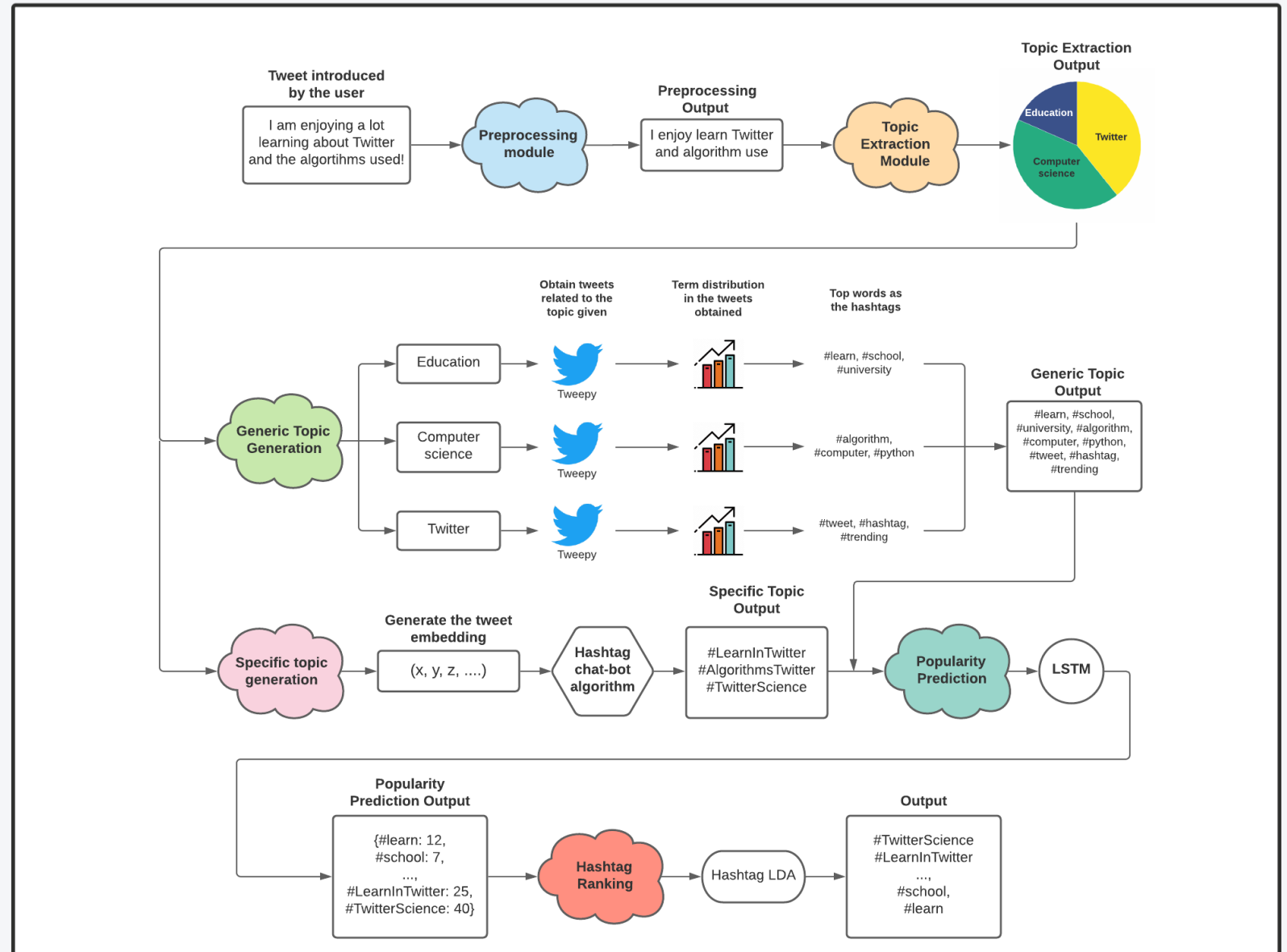
- 01 Full Recommendation System
- 02 Knowledge Transfer
- 03 Training Opportunities



"That's all Folks!"

Questions?

Workflow



Data collection and preprocessing

- ✓ Twitter API + Tweepy
- ✓ Common datasets: SEISMIC
- ✓ NLP Toolkit

Validation of the system



**HYBRID
METRIC**

**Popularity-
relevance
trade-off**



**TEMPORAL
METRIC**

**Measure
impact over
time**