

The background features a dark blue grid. A magnifying glass with a silver rim and a dark handle is positioned in the upper right. Two jagged lines, one dark blue and one dark red, trend across the image. The blue line starts high on the left and trends downwards towards the bottom right. The red line starts lower on the left and trends upwards towards the top right. The magnifying glass is focused on the red line's upward trend.

# Natural Language Processing for Trend Forecasting

Heládio Sampaio Lopes

Computer Engineering (ITA 2020)

# CONTENT



1. INTRODUCTION
2. LITERATURE TO REVIEW
3. RELATED WORKS
4. MATERIALS AND METHODS
5. ROADMAP



## 1. INTRODUCTION

## 2. LITERATURE TO REVIEW

## 3. RELATED WORKS

## 4. MATERIALS AND METHODS

## 5. ROADMAP

# INTRODUCTION



Over the years, more and more knowledge is generated and we humans are not able to process such an amount of information. Natural language processing emerges as a technology capable of assisting us in this hard task.

# CONTENT



1. INTRODUCTION

2. LITERATURE TO REVIEW

3. RELATED WORKS

4. MATERIALS AND METHODS

5. ROADMAP

# LITERATURE TO REVIEW



# CONTENT



1. INTRODUCTION

2. LITERATURE TO REVIEW

3. RELATED WORKS

4. MATERIALS AND METHODS

5. ROADMAP

## RELATED WORKS





# CONTENT



1. INTRODUCTION

2. LITERATURE TO REVIEW

3. RELATED WORKS

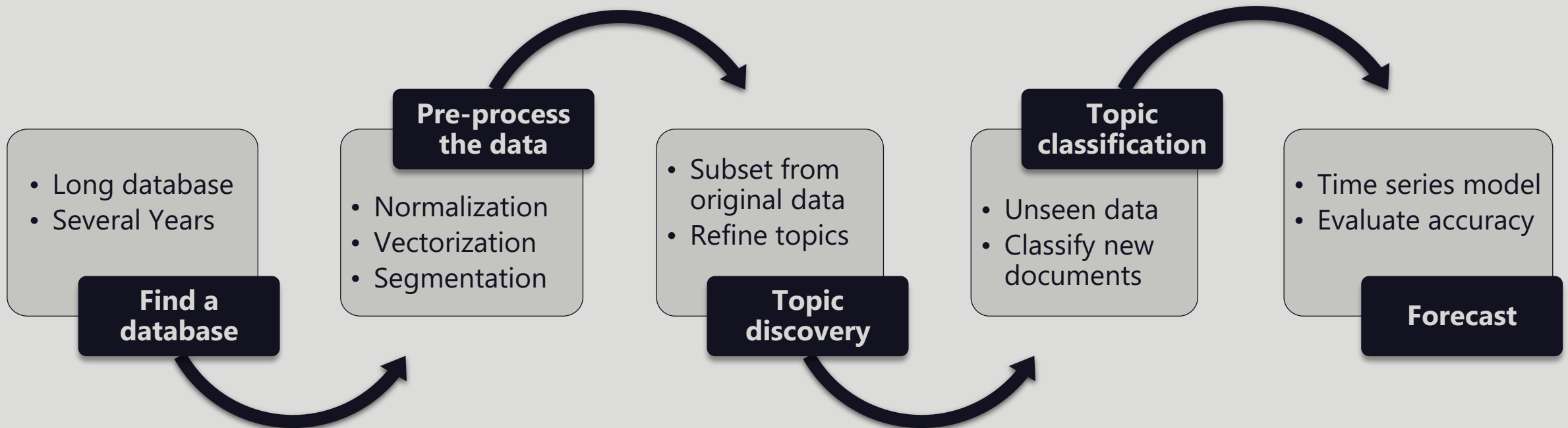
4. MATERIALS AND METHODS

5. ROADMAP

# MATERIALS AND METHODS | Objectives



As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



# MATERIALS AND METHODS | Database



As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



Wikipedia Daily News



Newspapers Articles



Academic Papers



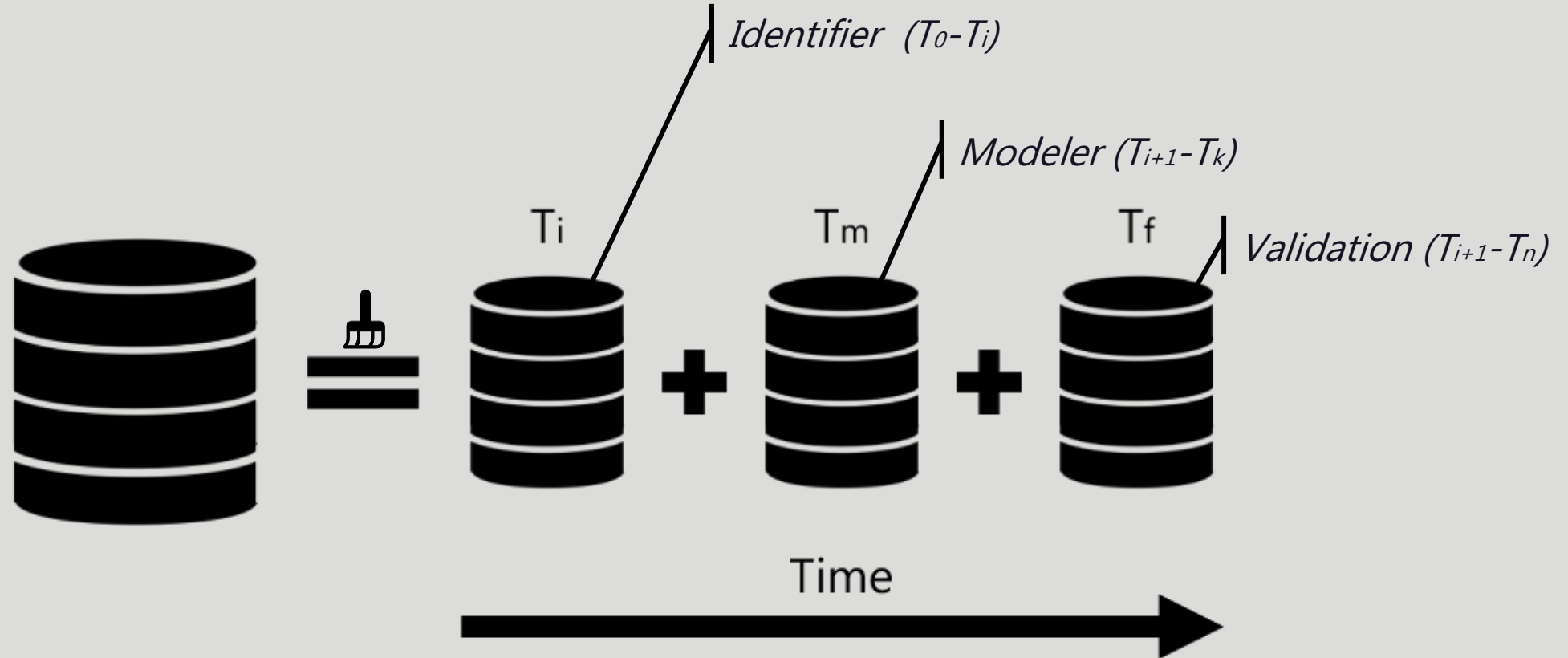
Social Media - Reddit



# MATERIALS AND METHODS | Pre-processing the Data



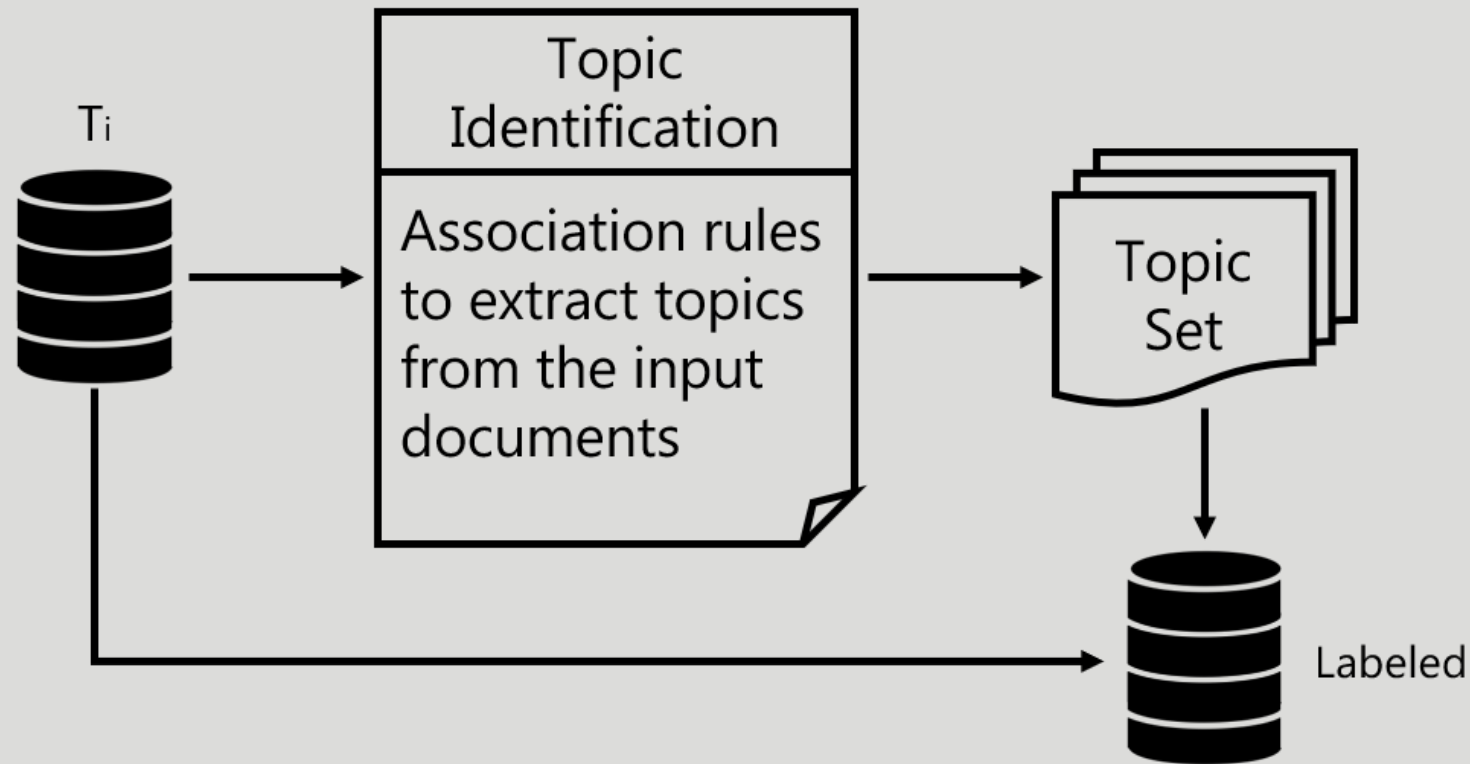
As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



# MATERIALS AND METHODS | Topic Identification



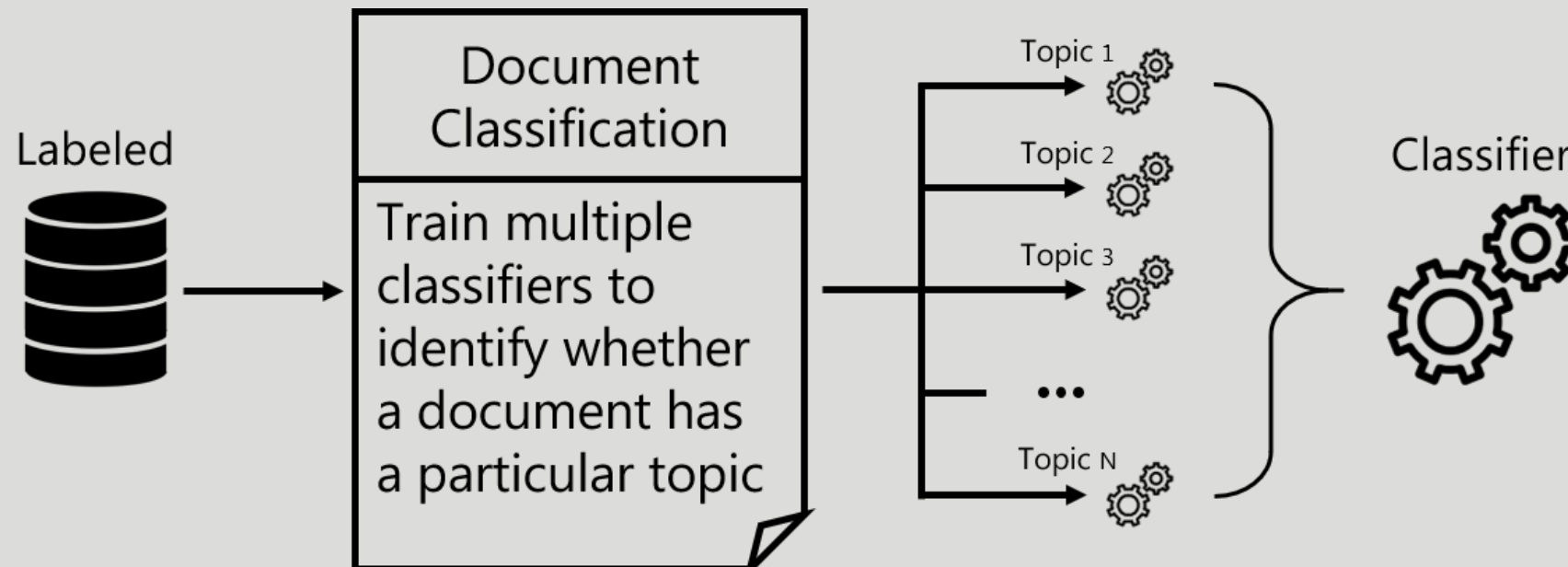
As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



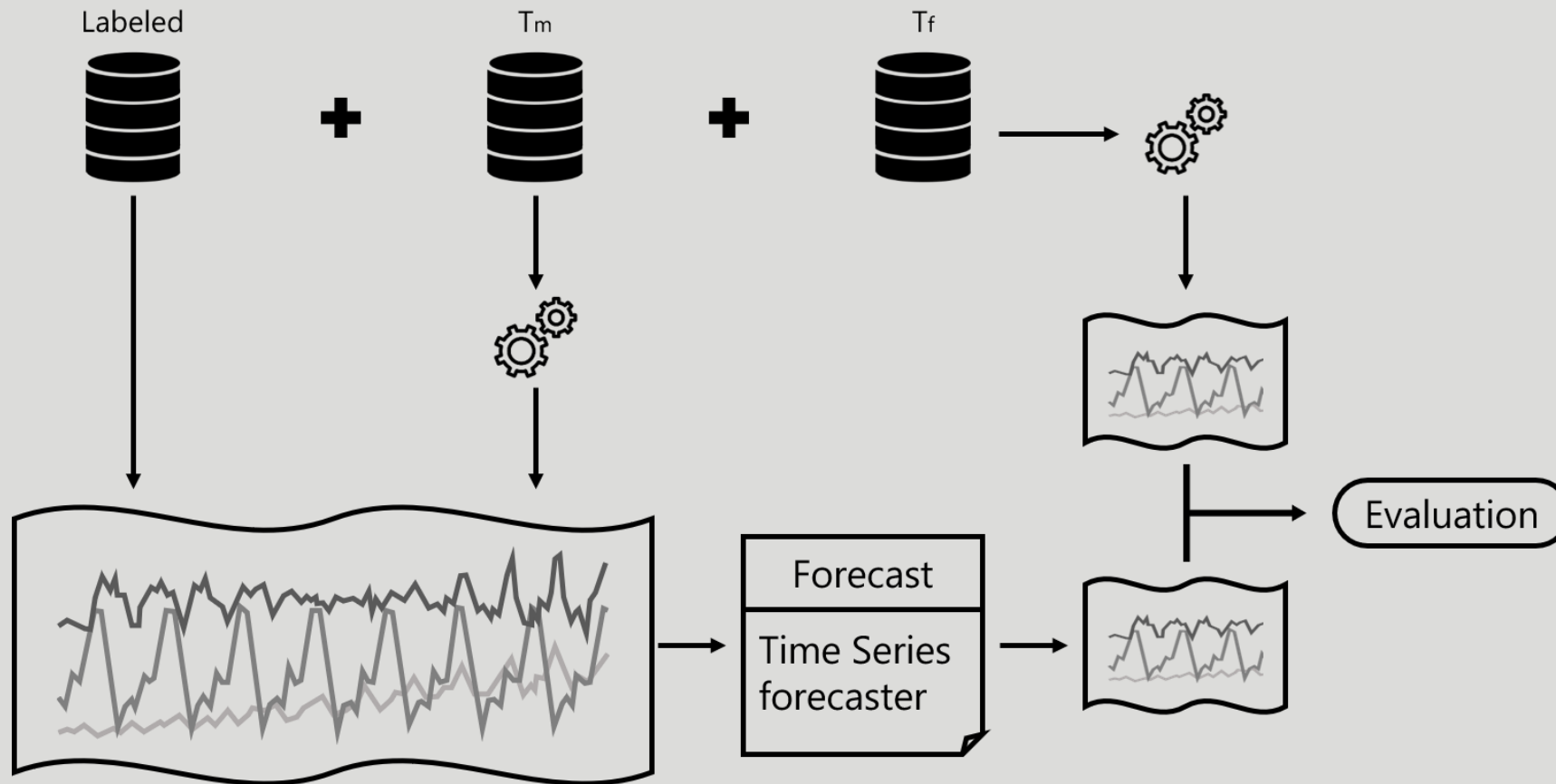
# MATERIALS AND METHODS | Document Classification



As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



As discussed earlier, we want to build models capable of make predictions regarding the evolution of discovered topics in a set of documents and identify the discovered topics in real time.



# CONTENT



1. INTRODUCTION
2. LITERATURE TO REVIEW
3. RELATED WORKS
4. MATERIALS AND METHODS
5. ROADMAP



# ROADMAP



In view of the problem's complexity, we can elaborate a schedule with the proposed tasks in the previously. The table above show the tasks over the remains months until the end of this work.

Sprint	Start Date	End Date	Duration	Task
#1	August 3	August 16	14 days	- Choose a database - Pre process the database
#2	August 17	September 6	21 days	- Topic Identification
#3	September 7	September 27	21 days	- Document Classification
#4	September 28	October 18	21 days	- Time Series Forecast
#5	October 19	November 8	21 days	- Test and fix bugs

**Obrigado!**