

# Smart New York Times chatbot

by Anatole Lapellerie, DIA 4, ESILV A4

This chatbot has been developed by Anatole Lapellerie for a Chatbot and recommendation systems course at ESILV. It contains a Facebook Messenger chatbot side using The New York Times APIs and a recommendation item-based system side using user's preferences on the requested articles.

GitHub: <https://github.com/Anatole-Lapellerie/ESILV-S8-Chatbot-and-recommendation-systems>



## SUMMARY

Subject .....	1
How to use the chatbot .....	2
Connection .....	2
Conversation.....	2
How the recommender engine works .....	3
Principle.....	3
Data saving .....	3
Demonstration.....	4
Example of conversation with the chatbot .....	4
Example ratings and scores .....	5

## SUBJECT

This Smart New York Times chatbot aims to inform the user of the latest news, by showing articles published on The New York Times website<sup>1</sup>. It uses APIs provided by The New York Times Developer Network<sup>2</sup>. These APIs are the following: Article Search API, Most Popular API and Top Stories API. For the future, some other APIs could be used as Semantic API or Times Tags API.

The user asks for articles or list of articles and can give a feedback to the bot to update its recommender engine dataset.

---

<sup>1</sup> The New York Times website: <https://www.nytimes.com/>.

<sup>2</sup> The New York Times Developer Network website: <https://developer.nytimes.com/>.

# HOW TO USE THE CHATBOT

## Connection

Please read the README.md file for a more complete installation guide.

After having created a Node.js project and installed the required modules, run the app, run Ngrok on the port 3000, set up all the connections with the Facebook Developer application and begin the conversation with the Facebook page on Messenger.

It understands only English language.

## Conversation

The chatbot can answer to different situations. Here are the different patterns made:

- **Greetings:** *Hello → Hi! What do you want to know today?* The bot proposes example of answers that can use the user by clicking on buttons.
- **End:** *Thank you, bye → Thanks, see you later!*
- **Asking for the important headlines:** *What is the news of the day? → Here are the headlines...* The bot proposes a list of titles of the most important articles of the day.
- **Asking for the popular articles:** *What are the most shared articles of the last 7 days? → Here are the most shared articles...* The bot proposes a list of titles of the most shared, emailed or viewed articles of the last 1, 3 or 7 days, following the user specifications.
- **Asking for headlines in a specific category:** *What happened in sports today? → Here are the headlines in sports...* The bot proposes a list of titles of the most important articles of the day in the specified category, following the New York Times desk categories of the website.
- **Asking for a specific article:** *Tell me more about Pfizer vaccine efficiency → I've maybe something for you...* The bot presents the desk category, the title, the author and the resume of the article. It also puts the link to the entire article on the website. After this message, the bot proposes two feedback answers (a positive and a negative) in buttons. According to the choice of the user, its preferences are updated for the recommendation system.
- **Asking for a customized recap:** *Give me a customized recap of today. → Here is your personal recap of today.* The bot selects 3 articles from the 3 desk categories that have the best recommendation score<sup>3</sup>.

---

<sup>3</sup> Because of a GET problem in queries with the New York Times Article Search API when adding desk category filter, the bot is yet not able to really show the 3 selected articles. However, it details what are the 3 recommended desk categories.

## HOW THE RECOMMENDER ENGINE WORKS

### Principle

The implemented recommendation system is an item-based engine using desk categories of the website as items.

Each user, identified by a Facebook Messenger Id, is associated to all newspaper desk categories by evolving scores, that changes following what does the user asks for. For example, if he demands to see the daily recap of articles concerning politics, the politics desk category's score will be revised upwards. Likewise, after asking for an article, if the user responds positively to the bot's feedback, the linked desk category's score will be increased. Responding "yes" to a feedback concerning an article dealing with Joel Embiid's last basketball game will improve the sports score.

Then, scores of each user are used to find similarities between desk categories. It allows the system to know what other desk categories may interest the user. For example, if every user that likes automobiles are also liking trips, the system will suppose that a new user that likes automobiles will also appreciate travel articles.

Finally, all these data are used for the customized short recap that can ask a user to the chatbot.

A long-term goal could be to use the New York Times keywords API and keywords data on articles metadata to generalize the recommender system to keywords, that are more precise than desk categories.

### Data saving

All the data concerning rates and scores is saved on the recommendation repository as csv. You can find examples on the GitHub repository<sup>4</sup>.

As the engine saves ratings and scores in csv, you can find examples of saves in the folder recommendation.

---

<sup>4</sup> These data are not reused yet because of problems faced with *creadReadStream* function.

# DEMONSTRATION

## Example of conversation with the chatbot

Here is an example of conversation with the chatbot. You can also watch the video of the conversation on YouTube: <https://youtu.be/GS9iFAWCbgU>.

**Smart News Bot**  
0 personnes aiment ça  
Société de médias/d'actualités  
18:08

Hi

Hi! What do you want to know today?

Data provided by The New York Times

What are the news?

Here are the main news of the day:

- Biden Details \$2 Trillion Plan to Rebuild Infrastructure and Reshape the Economy
- What's in Biden's Infrastructure Plan?
- 'It Wasn't Right': Young Witnesses Offer Emotional Testimony in Chauvin Trial
- The New Normal in N.Y.: High Virus Rates and a Steady Stream of Cases
- How Food Banks Succeeded and What They Need Now
- 'Mom Is Really Different': Nursing Homes Reopen to Joy and Grief

Tell me if you want to know more about an article.

Give me my customized recap.

The best articles I can show you are dealing with politics, travel and us.

What happened in sports?

Here are the main news of the day in the category sports:

- Hernandez and Strawberry. Reyes and Wright. Lindor?
- Gonzaga Awaits U.C.L.A. in the Final Four After Bruins Stun Michigan
- N.C.A.A. Women's Final Four: A First-Timer Crashes a Group of Powers
- With 17 Games, the N.F.L. Evolves for Streaming Generation
- Soccer Samples Streetwear and Loves the Fit
- South Carolina and Stanford Round Out Women's Final Four

Tell me if you want to know more about an article.

Tell me more about Hernandez and Strawberry

I have maybe something for you... Are you satisfied with this article?

HERNANDEZ AND STRAWBERRY. REYES AND WRIGHT. LINDOR? By David Waldstein

Francisco Lindor hasn't committed to the Mets beyond this season, but he has the talent and track record to match up with the best all-around players in franchise history.

Read more:  
<https://www.nytimes.com/2021/03/31/sports/baseball/francisco-lindor-mets.html>

Yes, I like this

Glad to hear it 🥳

What are the most shared articles of the last 7 days?

Here are the more shared articles of the last 7 days:

- Matt Gaetz Is Said to Face Justice Dept. Inquiry Over Sex With an Underage Girl
- The Curious Case of the Cinnamon Toast Crunch Box
- 'A Very Big Problem.' Giant Ship in the Suez Remains Stuck.
- Mother of 6 Fatally Shot in Road Rage Episode, Police Say
- Alabama and Georgia Dig Out From Tornadoes That Left at Least 6 Dead
- Nike Sues Over Unauthorized 'Satan Shoes'

Tell me if you want to know more about an article.

Tell me more about Emmanuel Macron

I have maybe something for you... Are you satisfied with this article?

A THIRD NATIONAL LOCKDOWN SEEMS LIKELY IN FRANCE AS HOSPITALS ARE OVERWHELMED. By Constant Méheut

President Emmanuel Macron will address the country tonight and is expected to announce far-reaching new restrictions, a last resort that he has long tried to avoid.

Read more:  
<https://www.nytimes.com/2021/03/31/world/covid-france-lockdown.html>

Yes, I like this

Glad to hear it 🥳

But I don't like lockdowns

Sorry, I didn't understand what you asked... 😞

Give me my customized recap...

The best articles I can show you are dealing with travel, us and world.

Ok, bye!

Thanks, see you later! 🙌

Data provided by The New York Times

## Example ratings and scores

Here is an example of a save of ratings table in CSV:

```
user_id,arts,automobiles,books,business,fashion,food,health,home,insider,ma
gazine,movies,nyregion,obituaries,opinion,politics,realstate,science,sport
s,sundayreview,technology,theater,travel,upshot,us,world
00001,1,4,1,4,1,2,2,2,3,1,3,4,1,4,5,5,2,5,2,3,1,4,1,5,3
00002,4,1,4,2,3,5,3,3,2,5,4,2,2,2,2,1,1,1,4,2,4,3,2,3,2
00003,2,1,1,3,2,3,2,1,1,3,3,1,1,3,4,1,4,4,2,3,1,5,3,3,5
XXXXX,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,1,1,1,1,1,1,2
```

And here is an example of save of scores table in CSV calculated from the first table:

```
user_id,arts,automobiles,books,business,fashion,food,health,home,insider,ma
gazine,movies,nyregion,obituaries,opinion,politics,realstate,science,sport
s,sundayreview,technology,theater,travel,upshot,us,world
00001,4.60,4.98,4.29,5.56,4.40,5.46,4.77,4.58,4.77,5.13,5.57,5.11,3.93,5.56
,6.09,5.24,4.95,6.42,4.97,5.19,4.29,6.17,4.54,6.02,6.04
00002,4.98,4.19,4.78,5.07,4.63,5.63,4.73,4.53,4.40,5.41,5.54,4.62,4.14,5.07
,5.56,4.45,4.65,5.76,5.07,4.88,4.78,5.94,4.49,5.65,5.55
00003,4.60,4.31,4.18,5.24,4.33,5.32,4.54,4.25,4.31,5.09,5.34,4.53,3.79,5.24
,5.72,4.55,4.86,5.99,4.81,4.90,4.18,5.97,4.51,5.63,5.76
XXXXX,3.09,2.90,2.83,3.67,2.84,3.86,3.13,2.88,2.89,3.61,3.86,3.13,2.36,3.67
,4.16,3.16,3.16,4.42,3.37,3.37,2.83,4.40,2.89,4.15,4.16
```