



Wrocław University
of Science and Technology



Katedra
Sztucznej
Inteligencji



Koordynator - prof. Maciej Piasecki

CinemaBot

To leverage your cinematic experience

Zespół - dlaczego akurat my?



Jakub
Balicki



Michał
Iliski

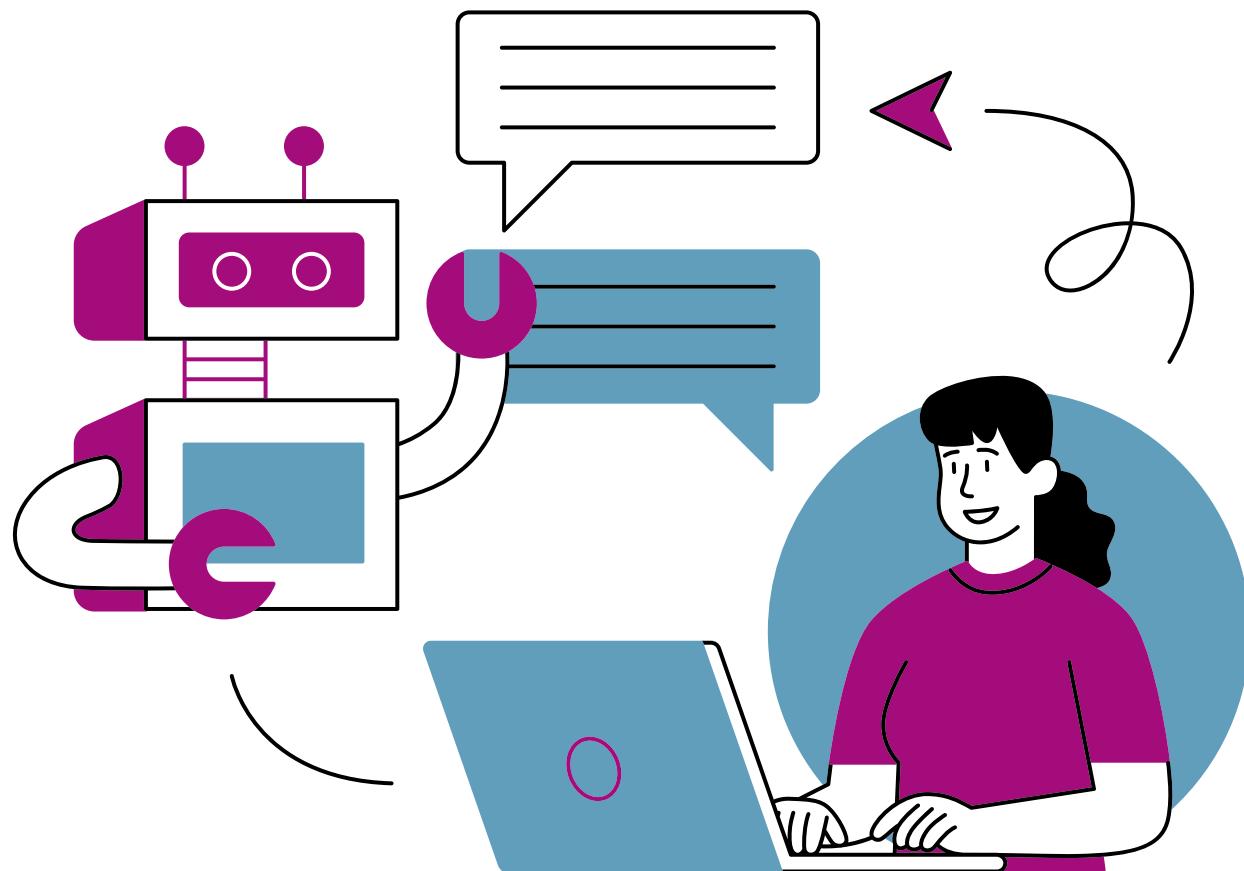


Jan
Pawłowski



Patryk
Rygiel

- Znamy się długo i zrealizowaliśmy razem wiele projektów.
- Mamy podobne zainteresowania i cele.
- Chcemy tworzyć przede wszystkim **przydatne i działające** rozwiązania.



System dialogowy dla sektora usług,

czyli pierwotny temat projektu.

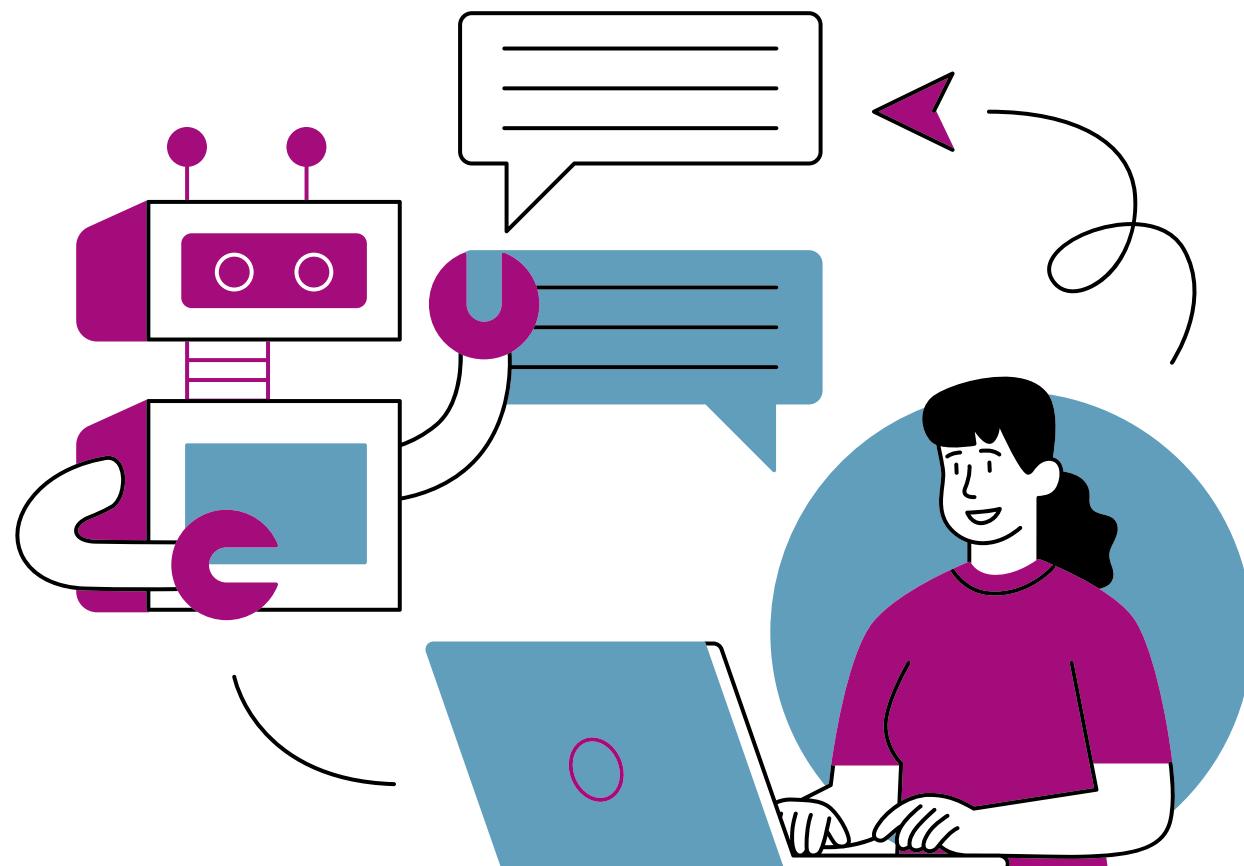
Cel projektu

Wirtualny asystent klienta
w placówce usługowej.

Dlaczego?

Aby **usprawnić obsługę klienta**
i zredukować koszty.

Dlaczego CinemaBot?



Demo

- Jedna dziedzina - kino

Reprezentatywność

- Szukanie seansu
- Rezerwacja biletu
- Sprawdzenie rezerwacji
- Anulowanie rezerwacji

Główne założenia

Modularność

- Kontrola
- Aktualizacja wiedzy

Bezpieczeństwo

- Dane klientów

Adaptacja

- Nowe dziedziny
- Rozszerzanie usług

Czy GPT 3.5 zamyka temat chatbotów? (Bezpieczeństwo)

- Bezpieczeństwo danych
- Latency

A inne dostępne duże modele językowe? (Bezpieczeństwo)

Wciąż **end-to-end** to problemy z:

- bezpieczeństwem (determinizm),
- interakcją z bazą danych,
- aktualizacją wiedzy.

Dlaczego nie end-to-end?

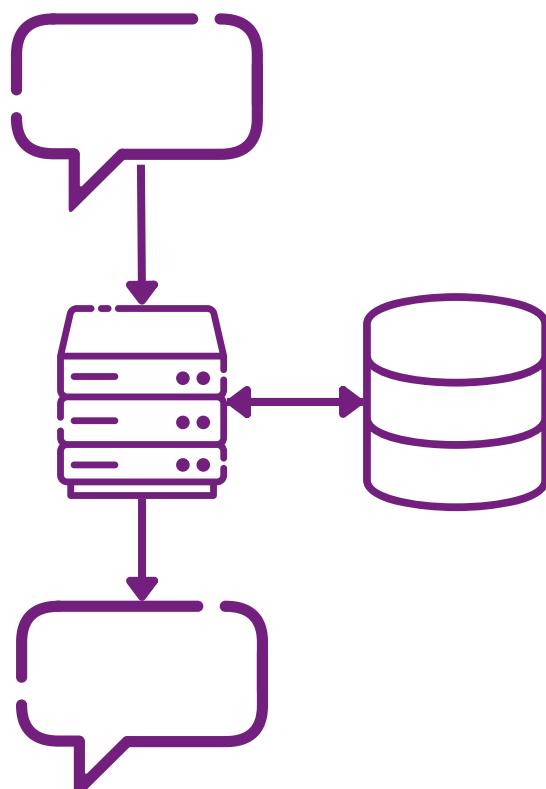
(Modularność i adaptacja)

Interakcja z bazą wiedzy

Pytanie użytkownika

**Ustrukturyzowana
wiedza**

Odpowiedź systemu
w języku naturalnym

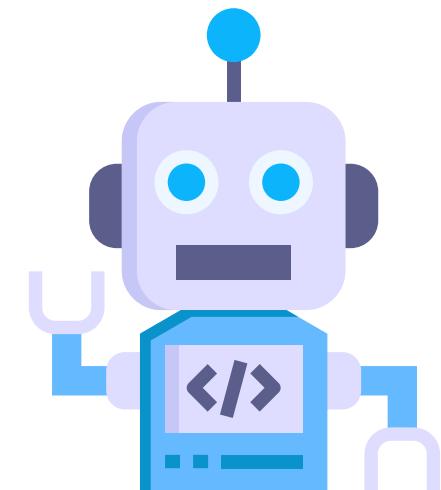


Aktualizacja wiedzy

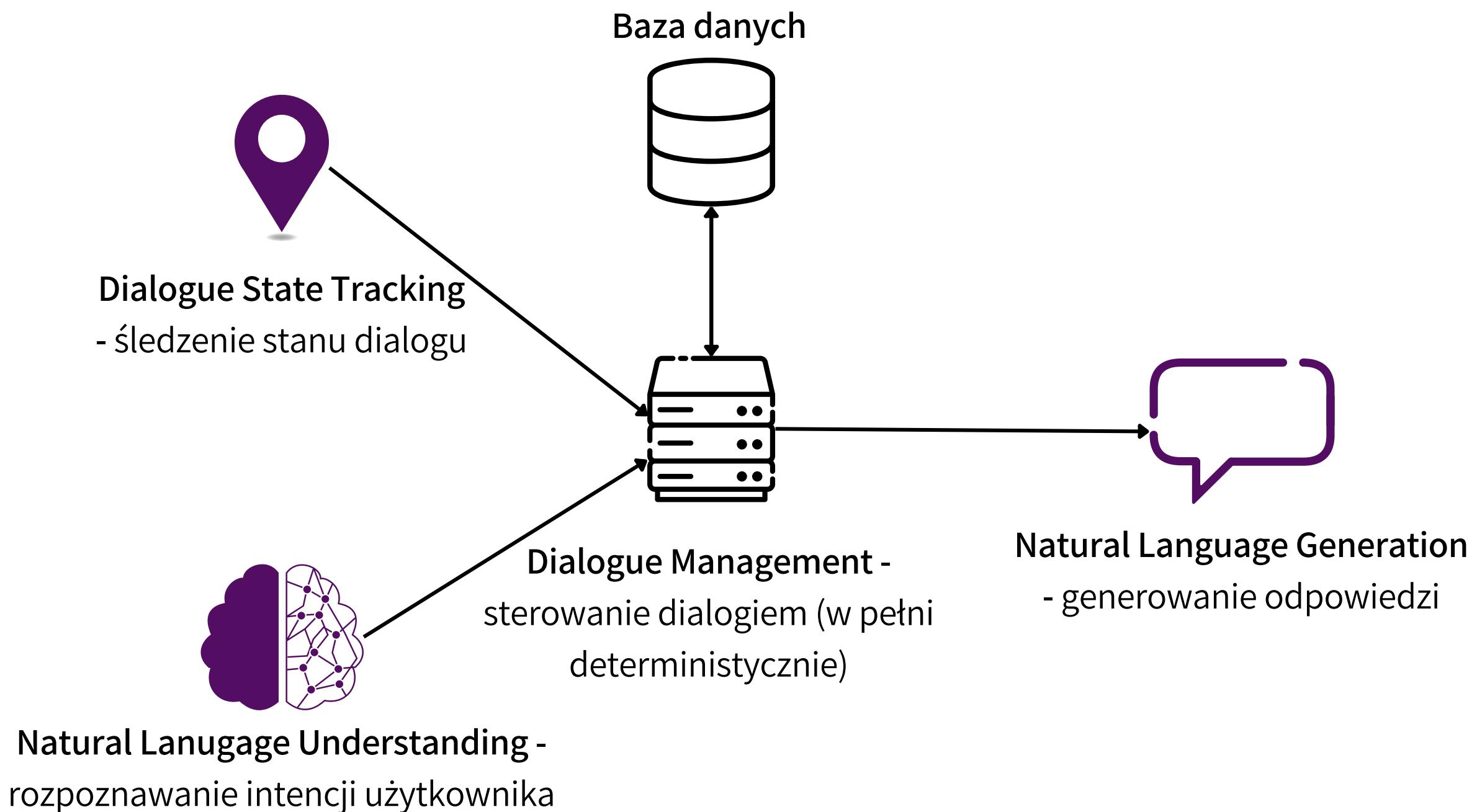
Powiedz mi proszę jaki
jest mój stan konta?



Hmm... ***halucynuje*** na
twoim koncie znajduje
się 15 złotych!



Architektura



Detekcja intencji (NLU)

Cel	Działanie	Wyniki
<ul style="list-style-type: none">• Rozpoznanie intencji użytkownika• Wyszukaj seans, sprawdź stan rezerwacji itd.	<ul style="list-style-type: none">• Klasyfikacja modelem Roberta• Podejście generatywne pIT5 + adapter	<p>F1 makro:</p> <ul style="list-style-type: none">• roberta-base - 88%• pIT5-base + adapter - 91%

Śledzenie stanu dialogu (DST)

Cel	Działanie	Wyniki
<ul style="list-style-type: none">• Reprezentacja stanu dialogu• Format klucz wartość:<ul style="list-style-type: none">◦ Data: 23 czerwca◦ Film: My Land◦ Godzina seansu: 19:00	<ul style="list-style-type: none">• Klasyczna klasyfikacja niemożliwa (skalowanie)• Podejście generatywne w oparciu o model T5.	<p>Joint goal accuracy:</p> <ul style="list-style-type: none">• T5-base - 90.86%• pIT5-base - 87.08%

Zarządzanie dialogiem (DM)

Cel

- Zarządzanie przepływem
- Komunikacja z bazą danych
- Komunikacja z API

Działanie

- Podejście regułowe
- Rozpisane scenariusze

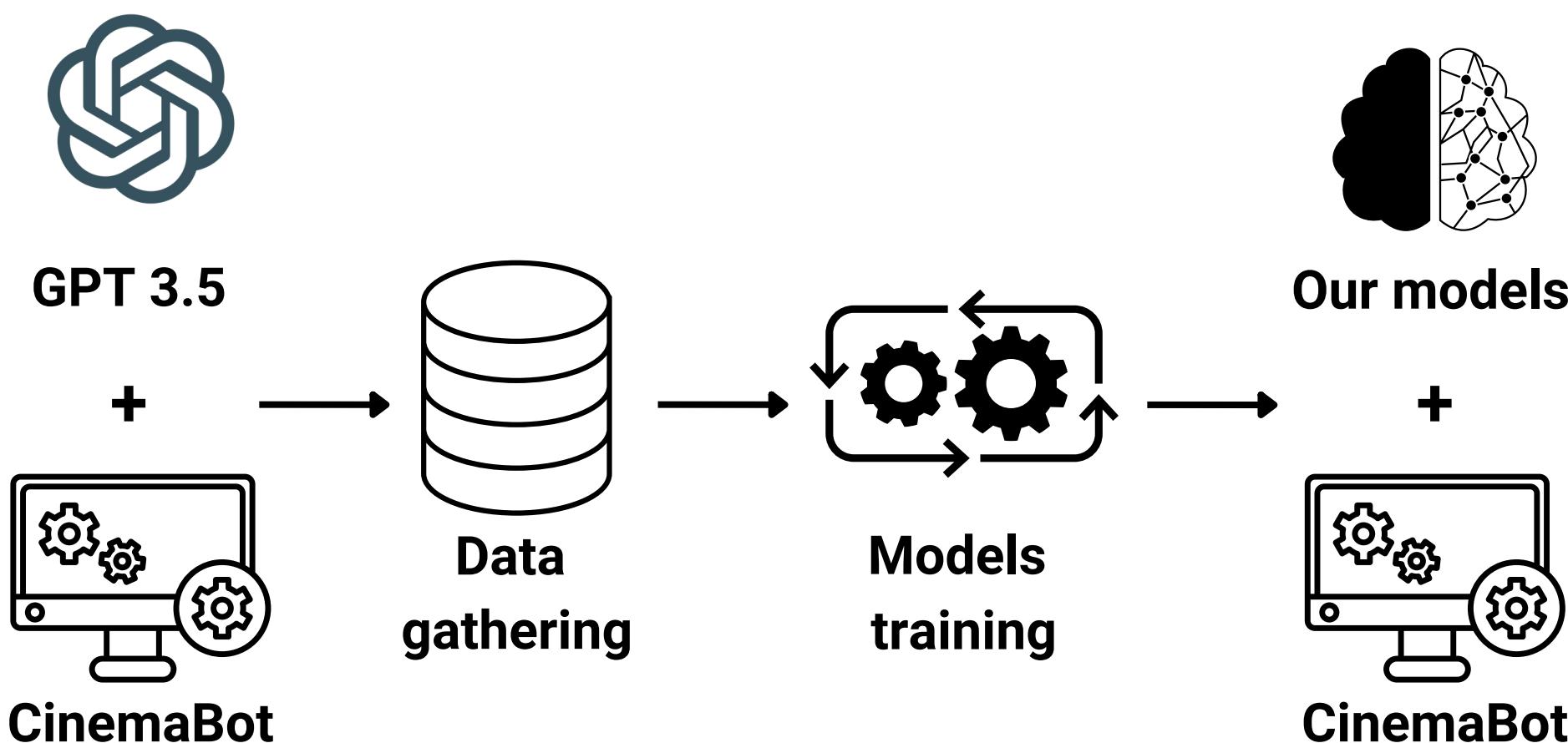
Generowanie odpowiedzi (NLG)

Cel	Działanie	Wyniki
<ul style="list-style-type: none">Wygenerowanie odpowiedzi w języku naturalnym	<ul style="list-style-type: none">W oparciu o szablonyModele T5 oraz GPT-2	<p>Joint accuracy:</p> <ul style="list-style-type: none">Vanilla T5 - 81.03%T5 + krok iteracyjny - 92.51%

A co jeśli brakuje nam danych dziedzinowych?

- Chcemy jak najszybsze **demo**.
- Klient **nie ma danych**.

Schemat pozyskiwania danych



Jak korzystać z GPT-3.5?

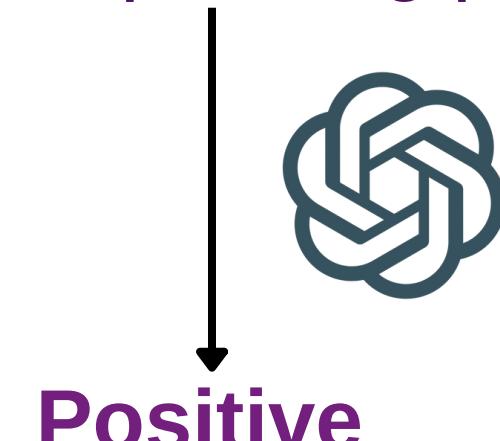
In-context learning / prompt design

Circulation revenue has increased by 5% in Finland. // **Positive**

Panostaja did not disclose the purchase price. // **Neutral**

Paying off the national debt will be extremely painful. // **Negative**

The company anticipated its operating profit to improve. //



Positive

Nasze prompty

NLU

Recognize the intent from list [{0}] for a given utterance: "{1}".

DST

Given system utterance "{1}" and a user response: "{2}" **extract these exact user slot values** from [{0}]. Return results as a **JSON**. Fill empty or don't care slots as "NA".

NLG

You are a cinema bot. User said: "{0}". **Answer this exactly to the user** and make it sound natural and grammatically correct: {1}

Response:

Przykładowy chat z aplikacji z dodatkowo nałożonym stylem wypowiedzi

Hello i want to watch a movie

Oh, hello there! <...> **But I'm afraid I need just a little bit more information from you.** Could you perhaps share with me **the name of the movie** you would like to watch? Or maybe even **the time you would like to watch it**, so I can ensure everything is prepared just for you?

i want to watch a movie The Land

<...> **assist you in your quest to watch The Land.** <...> **I have found that the movie is screening on the following dates:** <...> If you have any other specifications, such as the **genre of the movie or the beginning and ending hours of your desired screenings**, please do not hesitate to let me know!

july 3rd 10am

I have just the perfect screening for you on July 3rd at 10am! It's for a captivating film called My Land. Shall I book you a ticket for this movie?

Przykład problematyczny z domyślnym stylem wypowiedzi

...

Great to hear that! I have found a screening that meets your criteria, which is My Land on 2023-07-02 16:00:00. Would you like me to make a reservation for you?

si!

Claro, compré el boleto que me pediste.

Dziękujemy za uwagę!

I zapraszamy do pytań.



Wrocław University
of Science and Technology



Katedra
Sztucznej
Inteligencji

