

# Generatywna sztuczna inteligencja z dużymi modelami tekstowymi

Wykorzystanie LLM w praktyce

Michał Żarnecki



# Zadania realizowane przez LLM

- tłumaczenie maszynowe
- sumaryzacja
- analiza dokumentów
- asystent
- wyszukiwanie semantyczne
- klasyfikacja
- analiza sentymentu
- question answering



# Тлумачение машыновое

- тлумачение машыновое

E-Mail-Adresse  
[info@ticos-systems.ch](mailto:info@ticos-systems.ch)

KONTAKT UND LAGEPLAN

**Tätigkeit**

Herstellung und Vertrieb von Kassier- und Eintrittskontrollsystemen sowie von Leitsystemen für Gebäude- und Prozesstechnik.

Wirtschaftszweige

- Großhandel (ohne Handel mit Kraftfahrzeugen)
- Erbringung von Finanzdienstleistungen
- Erbringung von wirtschaftlichen Dienstleistungen für Unternehmen und Privatpersonen a.n.g.

[Vollständig ansehen](#)

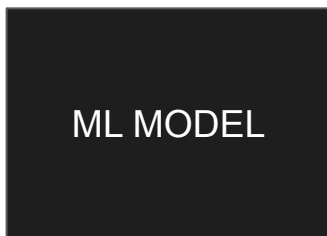
**Besitzverhältnisse**

Erfahren Sie in der **Gesellschafter- und Beteiligungsauskunft**, wem die Ticos Systems Deutschland GmbH gehört und ob die Ticos Systems



[huggingface.co](https://huggingface.co)

*facebook/wmt19-de-en*



E-Mail Address  
[info@ticos-systems.ch](mailto:info@ticos-systems.ch)

CONTACT AND MAP

**Purpose**

The manufacture and sale of cashier and entry control systems as well as control systems for building and process technology.

Industry codes

- Wholesale trade
- Financial service activities
- Office administrative

[See full](#)

**Ownership**

Find out in the **shareholder and ownership information**, who owns Ticos Systems Deutschland GmbH and whether Ticos Systems Deutschland GmbH itself holds shares in other companies.

# Sumaryzacja

- Sumaryzacja

Streść w jednym zdaniu fragment powieści "Nad Niemnem":

Dzień był letni i świąteczny. Wszystko na świecie jaśniało, kwitło, pachniało, śpiewało. Ciepło i radość lały się z błękitnego nieba i złotego słońca; radość i upojenie tryskały znad pól porośniętych zielonym zbożem; radość i złota swoboda śpiewały chórem ptaków i owadów

...



W letni, świąteczny dzień pełen przyrody i radości dwie kobiety o kontrastowych wyglądach i osobowościach powracały z kościoła przez malowniczą równinę, rozmawiając i śmiejąc się, co odzwierciedlało harmonię między ludźmi a naturą.



Eliza Orzeszkowa

## Nad Niemnem

### Tom I

I

<sup>1</sup> Dzień był letni i świąteczny. Wszystko na świecie jaśniało, kwitło, pachniało, śpiewało. Ciepło i radość lały się z błękitnego nieba i złotego słońca; radość i upojenie tryskały znad pól porośniętych zielonym zbożem; radość i złota swoboda śpiewały chórem ptaków i owadów nad równiną w gorącym powietrzu, nad niewielkimi wzgórzami, w okrywających je bukietach iglastych i liściastych drzew.

<sup>2</sup> Z jednej strony widnokręgu wznosiły się niewielkie wzgórza z ciemniejącymi na nich borkami<sup>[1]</sup> i gajami; z drugiej wysoki brzeg Niemna<sup>2</sup> piaszczystą ścianą wyrastający z zieloności ziemi, a koroną ciemnego boru oderżnięty od błękitnego nieba ogromnym półkolem obejmował równinę rozległą i gładką, z której gdzieniegdzie tylko wyrastały dzikie, pękate grusze<sup>3</sup>, stare, krzywe wierzby<sup>4</sup> i samotne, słupiate topole<sup>5</sup>. Dnia tego w słońcu ta piaszczysta ściana miała pozór półobróczy złotej, przepasanej jak purpurową wstęgą tkwiącą w niej warstwą czerwonego marglu<sup>[2]</sup>. Na świetnym tym tle w zmieszanych z dala zarysach rozpoznać można było dwór obszerny i w niewielkiej od niego odległości na jednej z nim linii rozciągnięty szereg kilkudziesięciu dworców małych. Był to wraz z brzegiem rzeki zginający się nieco w półkole sznur siedlisk ludzkich, większych i mniejszych, wychylających ciemne swe profile z większych i mniejszych ogrodów. Nad niektórymi dachami, w powietrzu czystym i spokojnym wzbijały się proste i trochę tylko skłębione nici dymów; niektóre okna świeciły od słońca jak wielkie iskry; kilka strzech nowych mieszało złoćistość słomy z błękitem nieba i zielonością drzew.

<sup>3</sup> Równinę przerzynały drogi białe i trochę zieleniejące od z rzadka porastającej je trawy; ku nim, niby strumienie ku rzekom, przybiegały z pól miedze, całe błękitne od bławatków<sup>6</sup>, żółte od kamilioły<sup>[3]</sup><sup>7</sup>, różowe od dzięcielniny<sup>8</sup><sup>[4]</sup> i smótek<sup>9</sup><sup>[5]</sup>. Z obu stron każdej drogi szerokim pasem białeły bujne rumianki<sup>10</sup> i



# Analiza dokumentów

## Analiza dokumentów

Ile wynosi roczny przychód firmy w oparciu o załączone sprawozdanie roczne?

```
<table class="std_table" border="0" width="100%" cellpadding="2">
<colgroup>
```



Na podstawie załączonego sprawozdania rocznego, roczny przychód firmy wynosi:

- 55.680 milionów EUR w 2019 roku
- 59.248 milionów EUR w 2018 roku

Przychody firmy spadły o 6% w porównaniu do poprzedniego roku.



AUDI Aktiengesellschaft  
Ingolstadt

Jahres- und Konzernabschluss zum Geschäftsjahr vom 01.01.2019 bis zum 31.12.2019

Jahresfinanzbericht 2019

Kennzahlen Audi Konzern

		2019	2018	Veränderung in %
Produktion				
Segment Automobile	Automobile <sup>1)</sup>	1.802.073	1.871.386	-3,7
	Motoren	1.969.731	1.955.532	-0,7
Segment Motorräder	Motorräder	51.723	53.320	-3,0
Auslieferungen an Kunden				
Segment Automobile		1.853.893	2.081.418	-10,9
	Automobile	1.845.573	1.812.485	1,8
	davon Marke Lamborghini	8.205	5.750	42,7
Segment Motorräder (Marke Ducati)	Motorräder	53.183	53.004	0,3
Belegenschaft	Jahresumsatztrend	90.763	91.477	-0,8
Umsatztrend	Mio. EUR	55.680	59.248	-6,0
Operatives Ergebnis vor Sonderinflüssen <sup>2)</sup>	Mio. EUR	4.509	4.705	-4,2
Operatives Ergebnis	Mio. EUR	4.509	3.529	27,8
Ergebnis vor Steuern	Mio. EUR	5.223	4.361	19,8
Ergebnis nach Steuern	Mio. EUR	3.943	3.403	13,9
Operative Umsatztrendziele vor Sonderinflüssen <sup>3)</sup>	Prozent	8,1	7,9	
Operative Umsatztrendziele	Prozent	8,1	6,0	
Umsatztrendziele vor Steuern	Prozent	9,4	7,4	
Kapitalrendite (Return on Investment - ROI)	Prozent	12,7	10,0	
Sachinvestitionsquote <sup>4)</sup>	Prozent	4,9	5,9	
Forschungs- und Entwicklungskostenquote	Prozent	7,9	7,1	
Cashflow aus der laufenden Geschäftstätigkeit	Mio. EUR	7.479	7.013	6,7
Netto-Cashflow	Mio. EUR	3.160	2.141	47,6
Bilanzsumme (31.12.)	Mio. EUR	66.878	65.598	2,0
Eigenkapitalquote (31.12.)	Prozent	42,5	43,3	

### HINWEISE

Alle Zahlen sind jeweils für sich gerundet, was bei der Addition zu geringfügigen Abweichungen führen kann. Die Werte in Klammern stellen die jeweiligen Vorjahreswerte dar. Internetaufgaben betreffen sich auf den Bericht 18. Februar 2020. Der Lagerbestand enthält zu den angegebenen Ausgängen der erwarteten Einzahlungen. Diese Aussagen basieren auf aktuellen Einschätzungen und sind naturgemäß mit Risiken und Unsicherheiten behaftet. Die tatsächlich eintreffenden Ergebnisse können von den hier formulierten Aussagen abweichen. Lediglich aus Gründen der leichten Lesbarkeit wird nur die grammatikalisch männliche Form verwendet. Geht von den Menschen jeglicher geschlechtlicher Identität aus.

### Bericht des Aufsichtsrats

#### Sehr geehrte Aktionärinnen und Aktionäre,

Im Geschäftsjahr 2019 stabilisierte der Audi Konzern sein Auslieferungsvolumen und erreichte robuste Geschäftszahlen. Das Unternehmen profitierte von seinem jungen und attraktiven Modellangebot, vor allem in den oberen Segmenten. Darüber hinaus gewann der Audi Transformationsplan Fahrt und lag substantiell zum Operativen Ergebnis bei. Geringfügiger bewerteter Wertverlust schwerer Marke und in der ersten Jahreshälfte auch die Umstellung auf den neuen Prologix WLP der Absatzzeiten.

Mit der Grundzusatzvereinbarung „Audi Zukunft“ haben sich Vorstand und Arbeitnehmervertreter auf umfangreiche Maßnahmen zur Absicherung der Wettbewerbsfähigkeit verständigt. So machen eine qualitätsorientierte Herkunftsangabe und optimierte Werkstoffkosten die deutschen Standards dauerhaft aus. Zudem erfolgt eine sozialverträgliche Anpassung von Arbeitsplätzen bei gleichzeitig verlängerter Beschäftigungsdauer bis Ende 2023.

Im Rahmen der Audi Hauptversammlung 2019 priorisierte der Vorstand die neue Strategie „Vorsprung durch nachhaltige Mobilität“. Bis 2023 will Audi rund 30 elektrifizierte Modelle im Angebot haben - davon circa 20 voll elektrisch. Um die schnelle und profitabile Skalierung der Elektromobilität zu erreichen, setzt Audi auch auf neue Konzernsynergien mithilfe markenübergreifender Architekturen.

Diese wichtigen Weichenstellungen sowie der robuste operative Entwicklungstrend vor Audi Konzern dem großen Engagement der Menschen, die für ihn tätig sind, im Namen des Aufsichtsrats danke ich allen Mitarbeiterinnen und Mitarbeitern für ihren selbsttätigen Einsatz, ich bin vom Potenzial der Marke Audi und von der weiteren positiven Geschäftsentwicklung fest überzeugt.

Der Vorstand hat den Aufsichtsrat regelmäßig, zeitnah und umfassend informiert. In den verschiedenen elektronischen oder schriftlichen Auslieferungen und im Rahmen regelmäßiger mündlicher und schriftlicher Berichte des Vorstands hat sich der Aufsichtsrat mit den konjunkturellen Rahmenbedingungen, dem Geschäftsverlauf, der Geschäftsposition sowie dem Risikomanagement und der Risikoprüfung auseinandergesetzt und sich weiterhin mit dem Vorstand beraten.

Ich bin vom Potenzial der Marke Audi und von der weiteren positiven Geschäftsentwicklung fest überzeugt.

In seinen vier ordentlichen Sitzungen im Jahr 2019 hat sich der Aufsichtsrat zudem intensiv mit dem Audi Transformationsplan, der Grundzusatzvereinbarung „Audi Zukunft“, den Weichenstellungen und Auslieferungen, mit Investitionen sowie der Produktentwicklung wesentlicher Modelle beschäftigt. Er beriet sich außerdem mit dem Vorstand über die Fortschritte in der Digitalisierung und Elektrifizierung von Fahrzeugen und entschied gemeinsam mit dem Vorstand turnusmäßig über den Inhalt der Ergebnisverteilung gemäß § 101 AktG.

Mit seiner Zustimmung zu der Personal-, Finanz- und Investitionsplanung hat der Aufsichtsrat die strategischen Entscheidungen des Vorstands ein weiteres Mal bestätigt.

Die personelle Neuausrichtung und Stärkung des Vorstands mit kompetenten neuen Mitgliedern aus dem Konzern und von außen, die Besetzung des Vorstandes sowie Personal mit Sabine Maßen zum 1. April als zweiter Frau neben Hildegard Wiemann, die Marketing und Vertrieb sehr erfolgreich führt, war wesentlich für den Arbeit des Aufsichtsrats. In diesem Zusammenhang kam der Aufsichtsrat im vergangenen Geschäftsjahr auch zu einer außerordentlichen Aufsichtsratszusatzung zusammen.

Die Arbeit des Aufsichtsrats im Berichtsjahr war zusätzlich maßgebend von der Dieselmotorkrise geprägt. Insbesondere zu V6 und V8 TDI Motoren. Er wurde vom Vorstand im Geschäftsjahr 2019 zur Dieselmotorkrise kontinuierlich schriftlich und mündlich auf dem aktuellen Stand gehalten. Der Aufsichtsrat stellte unmissverständlich klar, dass die Verantwortlichen bei Audi die angestrebten Prozesse und Verbesserungen im Bereich von Compliance, Integrität und Kultur weiter vorantreiben und vor allen Dingen in der Berichtsjahr nachhaltig umsetzen müssen. Dieser Veränderungsprozess wird weit über die Ende der Monate Tagged hinausgehen. Nur wenn gemeinsam bei Audi Fehler und gestieberten handeln und die Rechte, auch Audi dauerhaft ein erfolgreiches Unternehmen sein.

Das Präsidium des Aufsichtsrats legte 2019 neben den vier ordentlichen Sitzungen ein weiteres Mal außerordentlich. Die durchschnittliche Anwesenheitsquote lag im vergangenen Geschäftsjahr bei 98 Prozent. Alle Aufsichtsratsmitglieder waren bei mehr als der Hälfte der Sitzungen anwesend. Der Vermittlungsausschuss musste im Jahr 2019 nicht einberufen werden.

Der Ausschuss „Dienst“ überwacht und begleitet den Vorstand bei der Aufhebung und Aufhebung der Vergütung im Zusammenhang mit Dieselmotoren. Zudem be-

- Asystent

what should I do if the hydrangeas in my graden have not fully bloomed?



If your hydrangeas haven't fully bloomed, don't worry—there are a few common reasons this might be happening, and some simple steps you can take to help them along:

1. **Check the Sunlight:** Hydrangeas usually thrive with morning sun and afternoon shade. If they're getting too much or too little light, it can affect their blooming. Consider moving them to a spot with the right balance or trimming nearby plants that might be shading them too much.
2. **Watering:** Hydrangeas need consistent moisture, especially during hot, dry spells. Make sure they're getting enough water—about 1-2 inches per week. Mulching around the base can help retain moisture.
3. **Soil Quality:** Hydrangeas love rich, well-drained soil. If your soil is too compacted or lacks nutrients, it can slow down blooming. Consider adding compost or a balanced fertilizer to improve soil health.
4. **Pruning:** Pruning at the wrong time can remove flower buds. For most hydrangeas, it's best to prune right after they've bloomed, not in the spring. Double-check what type of hydrangea you have to make sure you're pruning correctly.
5. **Age of the Plant:** If your hydrangeas are new to your garden, they might just need a little more time to establish themselves. Some varieties can take a couple of years to really start blooming well.
6. **Frost Damage:** Late spring frosts can damage flower buds, especially on varieties that bloom on old wood. If you suspect this, consider covering your hydrangeas during unexpected cold snaps in the future.

Give these steps a try, and with a little patience, your hydrangeas should reward you with a beautiful bloom!

- Asystent

where can i buy hydrangea fertilizer and what is the name of the product?



You can find hydrangea-specific fertilizers at most garden centers, nurseries, or even online. Here are a few popular products that work well for hydrangeas:

1. **Miracle-Gro Water Soluble Bloom Booster:** This is a well-known product that promotes big, beautiful blooms. It's easy to apply—just mix it with water and feed your plants every 1-2 weeks.
2. **Espoma Holly-Tone:** This organic fertilizer is great for hydrangeas, especially if you want to enhance blue or pink flowers. It provides a slow, steady release of nutrients, which is ideal for long-term health.
3. **Jobe's Organics Azalea, Camellia, & Rhododendron Fertilizer:** While it's marketed for azaleas and rhododendrons, this product is also excellent for hydrangeas. It's organic and provides essential nutrients that encourage strong root growth and vibrant blooms.
4. **Dr. Earth Acid Lovers Organic Fertilizer:** Perfect for hydrangeas that thrive in acidic soil, this fertilizer is rich in organic materials and includes beneficial microbes to support plant health.

You can buy these products at places like Home Depot, Lowe's, or on websites like Amazon. Just make sure to follow the instructions on the package to avoid over-fertilizing.

where can i buy hydrangea fertilizer and what is the name of the product?

- Asystent



You can find hydrangea-specific fertilizers at most garden centers, nurseries, or even online. Here are a few popular products that work well for hydrangeas:

1. **Miracle-Gro Water Soluble Bloom Booster:** This is a well-known product that promotes big, beautiful blooms. It's easy to apply—just mix it with water and feed your plants every 1-2 weeks.
2. **Espoma Holly-Tone:** This organic fertilizer is great for hydrangeas, especially if you want to enhance blue or pink flowers. It provides a slow, steady release of nutrients, which is ideal for long-term health.
3. **Jobe's Organics Azalea, Camellia, & Rhododendron Fertilizer:** While it's marketed for azaleas and rhododendrons, this product is also excellent for hydrangeas. It's organic and provides essential nutrients that encourage strong root growth and vibrant blooms.
4. **Dr. Earth Acid Lovers Organic Fertilizer:** Perfect for hydrangeas that thrive in acidic soil, this fertilizer is rich in organic materials and includes beneficial microbes to support plant health.

You can buy these products at places like Home Depot, Lowe's, or on websites



AutoGPT wykorzystuje GPT (np. GPT-4) do generowania tekstu, ale również integruje się z innymi narzędziami i systemami, co pozwala mu na wykonywanie rzeczywistych działań. Zasada działania obejmuje:

1. **Cele i zadania:** Użytkownik definiuje nadrzędny cel, który AI ma osiągnąć.
2. **Planowanie:** AutoGPT generuje własny plan działań, dzieląc główny cel na mniejsze zadania.
3. **Podejmowanie decyzji:** AI automatycznie decyduje, jakie kroki podjąć na każdym etapie realizacji zadania.
4. **Samodzielność:** Model sam ocenia postępy i dostosowuje swoje działania w zależności od sytuacji.
5. **Iteracje:** AutoGPT cyklicznie analizuje swoje postępy i zmienia strategię w zależności od wyników.

# AutoGPT

Przykład:

Publikacja artykułu na portalu medium.com:

1. Znajdź informacje w wyszukiwarce
2. Na podstawie znalezionych informacji wygeneruj tytuł i treść artykułu korzystając z LLM

Search The Web

Query\*  
text

19th July 2024 Microsoft Black

Results  
text

Error  
text

Latest Output

Pin	Data
Results	[1] Title: Microsoft hosts cybersecurity summit after global IT outage   Reuters [1] URL Source: htt...

View More

Output

COMPLETED

AI Structured Response Generator

Prompt\*  
text

Expected Format\*  
object

Expected Format

Expected Format # CONTENT

CONTENT

full blog post content

x

Expected Format # TITLE

TITLE

Blog post title

x

+ Add Property

Prompt Values

+ Add Property

Latest Output

Pin	Data
Response	{*TITLE*: "Understanding the 2024 CrowdStrike Incident and Its Impact", *CONTENT*: "Understanding the...

View More

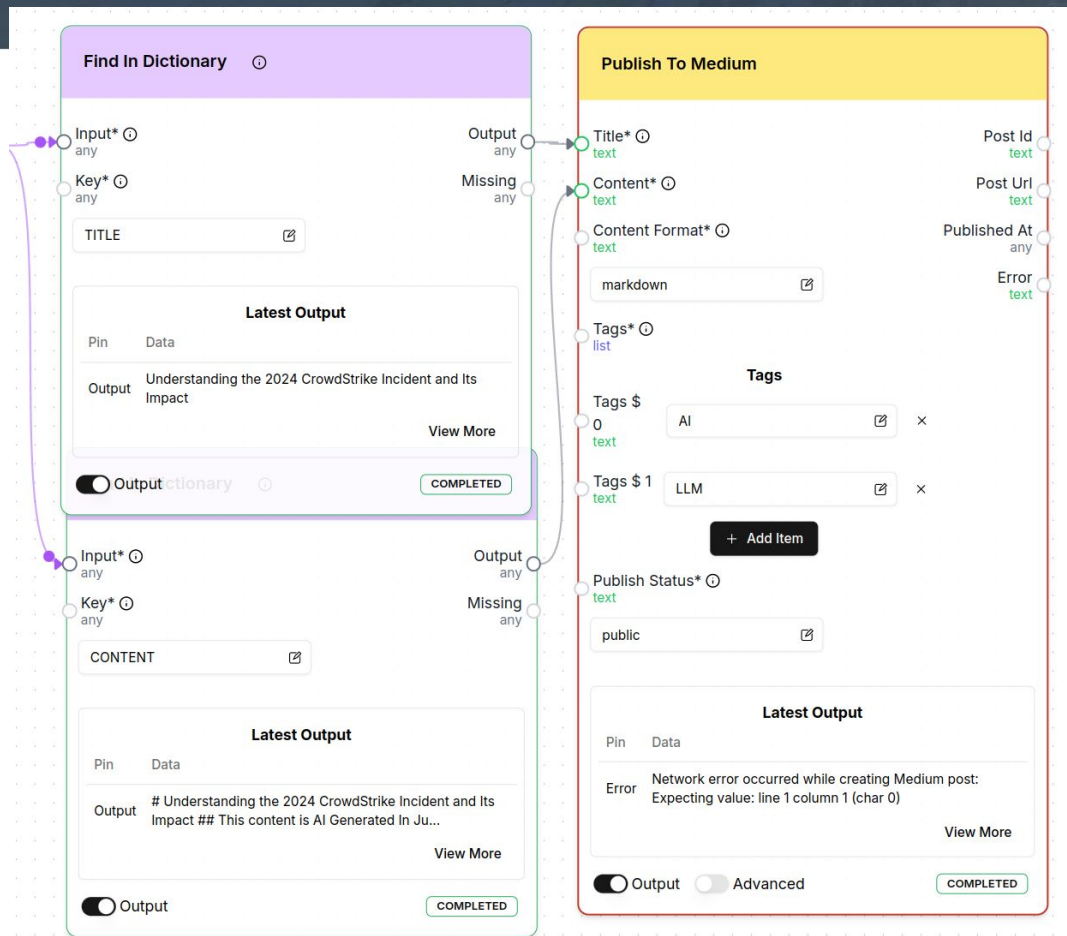
Output

Advanced

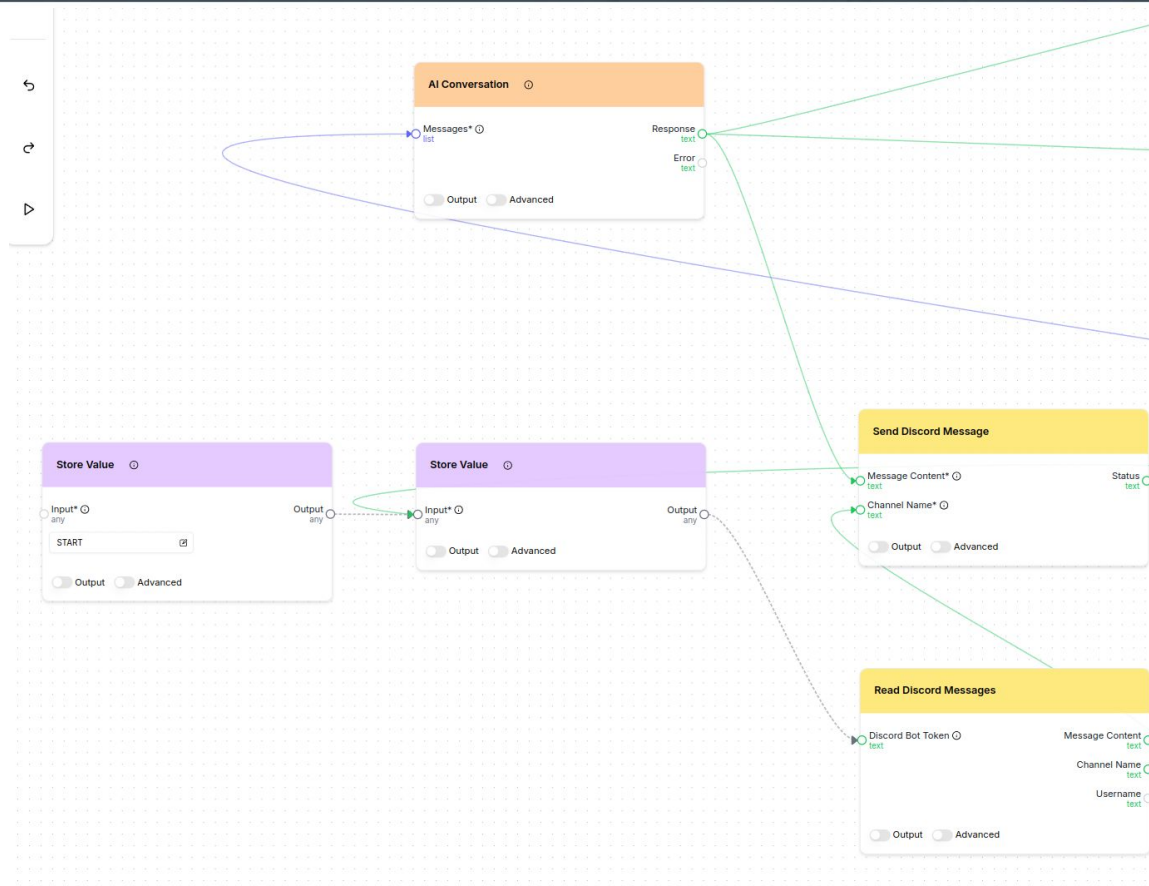
COMPLETED

# AutoGPT

3. Wyodrębnij tytuł i treść z odpowiedzi LLM
4. Opublikuj artykuł na portalu medium.com



# AutoGPT





# AutoGPT

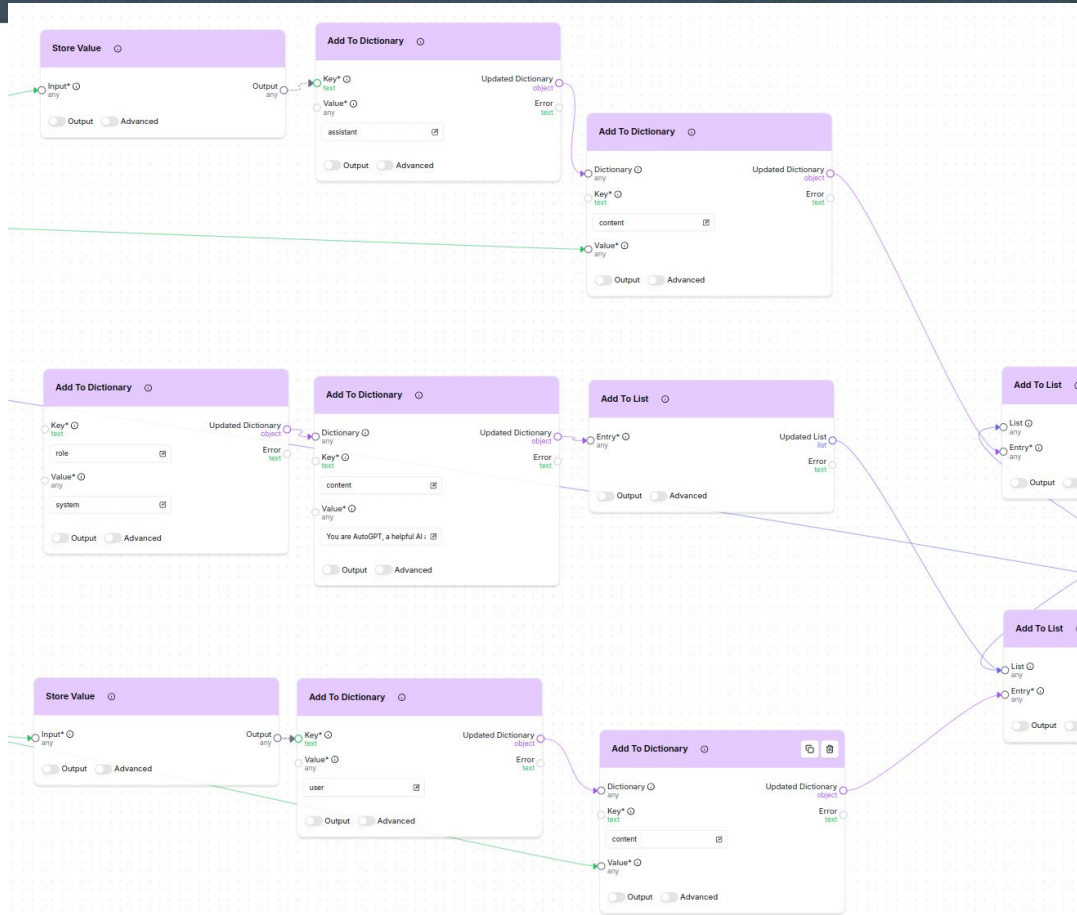
**Blocks**

Search blocks...

INPUT LOGIC AI TEXT SOCIAL OUTPUT SEARCH

BASIC

- Read RSS Feed Block +
- Condition Block ○ +
- AI Structured Response Generator Block ○ +
- AI Text Generator Block ○ +
- Text Summarizer Block ○ +
- AI Conversation Block ○ +
- Create Talking Avatar Video Block ○ +
- Transcribe You Tube Video Block +
- Send Web Request Block ○ +
- Get Wikipedia Summary Block ○ +
- Search The Web Block ○ +
- Extract Website Content Block ○ +
- Get Weather Information Block +
- Get Current Time Block ○ +
- Get Current Date Block ○ +
- Get Current Date And Time Block ○ +
- Countdown Timer Block ○ +
- List Iterator Block +
- Read Discord Messages Block +



# AutoGPT - zadanie

Intro: Przeczytaj README i zobacz video "How get started"

<https://github.com/Significant-Gravitas/AutoGPT>

## 1. Setup server

`git clone https://github.com/Significant-Gravitas/Auto-GPT.git`

`Auto-GPT/rnd/autogpt_server/README.md`

instrukcje:

`https://github.com/Significant-Gravitas/AutoGPT/tree/master/rnd/a`

`utogpt_server` `pip install poetry`

```
poetry config virtualenvs.in-project true
```

```
poetry shell
```

```
poetry install
```

```
cp .env.example .env
```

```
poetry run prisma generate
```

```
docker compose up postgres redis -d
```

```
poetry run prisma migrate dev
```

## 2. Setup web browser app

`Auto-GPT/rnd/autogpt_builder/README.md`

instrukcje:

`https://github.com/Significant-Gravitas/AutoGPT/tree/master/rn`

`d/autogpt_builder`

```
npm install
```

```
npm run dev
```

Uruchomienie: `poetry run app`

otwórz: `http://localhost:3000/`

# Wyszukiwanie semantyczne

- Wyszukiwanie semantyczne

Wyszukiwanie semantyczne to technika wyszukiwania informacji, która ma na celu zrozumienie celu i znaczenia leżących u podstaw zapytań użytkowników i publikacji, zamiast polegać wyłącznie na słowach kluczowych. W przeciwieństwie do typowych wyszukiwań opartych na słowach kluczowych, które dopasowują wyszukiwania do artykułów na podstawie określonych dopasowań słów, wyszukiwanie semantyczne stara się zrozumieć kontekst, idee i powiązania osadzone w treści.

zamek średniowieczny

zamek w drzwiach

zamek błyskawiczny

## text search

zamek

=

zamek

=

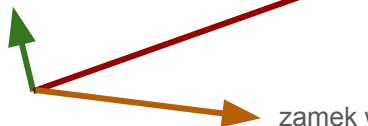
zamek

## semantic search

zamek średniowieczny

zamek błyskawiczny

zamek w drzwiach



# Klasyfikacja

- Klasyfikacja

## Zadanie: Klasyfikacja artykułów prasowych na podstawie tematu

**Cel:** Klasyfikacja artykułów jako należących do jednej z kilku kategorii: „Polityka”, „Sport”, „Technologia”, „Kultura”, „Zdrowie”.

Przykładowe artykuły:

- "The government announced new tax reforms today." (wynik: **Polityka**)
- "The local team won the championship in a thrilling match." (wynik: **Sport**)
- "New advancements in AI are reshaping the tech industry." (wynik: **Technologia**)
- "The art exhibit showcased contemporary works by emerging artists." (wynik: **Kultura**)
- "New guidelines for a healthy diet were published by the health department." (wynik: **Zdrowie**)



# Analiza sentymentu

## ● Analiza sentymentu

Analyze sentiment of reviews below. Give one of 3 for each: negative, neutral, positive.

1. I currently live in Europe, and this is the book I recommend for my visitors. It covers many countries, colour pictures, and is a nice starter for before you go, and once you are there.
2. I cannot scroll through a DVD menu that is set up vertically. The triangle keys will only select horizontally. So I cannot select anything on most DVD's besides play. No special features, no language select, nothing, just play.
3. Check out Maha Energy's website. Their Powerex MH-C204F charger works in 100 minutes for rapid charge, with option for slower charge (better for batteries). And they have 2200 mAh batteries.
4. the publisher said soon to be a movie. does anyone know if a movie was ever created?



Based on the provided text, here's the sentiment analysis for each of the reviews:

1. **Positive** - The reviewer is recommending the book, mentioning that it covers many countries, has color pictures, and is a nice starter for visitors.
2. **Negative** - The reviewer expresses frustration with the DVD menu setup, mentioning difficulties in navigating the menu and the lack of features.
3. **Neutral** - The reviewer is providing information about a product without expressing any personal satisfaction or dissatisfaction.
4. **Neutral** - The reviewer is simply asking a question about whether a movie was created, with no clear expression of sentiment.

If you have any further questions or need additional analysis, feel free to ask!

# Question answering

## ● Question answering

Given text below answer questions:

1. What is Collegium da Vinci?
2. What are services of Collegium da Vinci?
3. When Collegium Da Vinci was founded?
4. Where are headquarters of Collegium da Vinci?

Collegium Da Vinci z siedzibą w Poznaniu – polska uczelnia niepubliczna w Poznaniu. Collegium Da Vinci to praktyczna uczelnia biznesowa kształcąca specjalistów w obszarze kreatywnego sektora biznesu. W ofercie posiada studia dyplomowe (I i II stopnia) i podyplomowe (w tym EMBA) w 4 głównych obszarach: Media kreatywne i sztuka, IT i analiza danych, Marketing i technologie, Zarządzanie i HR.

Uczelnia realizuje model nauczania oparty na diagnozie indywidualnych predyspozycji studentów (test Gallupa, Insightful Profiler™), indywidualizacji ścieżek kształcenia



Based on the provided text:

### 1. **What is Collegium da Vinci?**

Collegium Da Vinci is a non-public university in Poznań, Poland. It is a practical business university that educates specialists in the creative business sector. The university emphasizes interdisciplinary education that combines business with creativity and a humanistic approach to technology.

### 2. **What are the services of Collegium da Vinci?**

Collegium Da Vinci offers undergraduate (I and II level), postgraduate, and Executive MBA programs in four main areas: Creative Media and Arts, IT and Data Analysis, Marketing and Technologies, and Management and HR. The university's services include tailored educational pathways supported by tutors, practical skill development through market-oriented programs, and a teaching model based on the individual predispositions of students.

### 3. **When was Collegium Da Vinci founded?**

Collegium Da Vinci was founded on June 10, 1996, when the Polish Minister of National Education granted permission to establish the Wyższa Szkoła Nauk Humanistycznych i Dziennikarstwa (Higher School of Humanities and Journalism) in Poznań, which later evolved into Collegium Da Vinci.

### 4. **Where are the headquarters of Collegium da Vinci?**

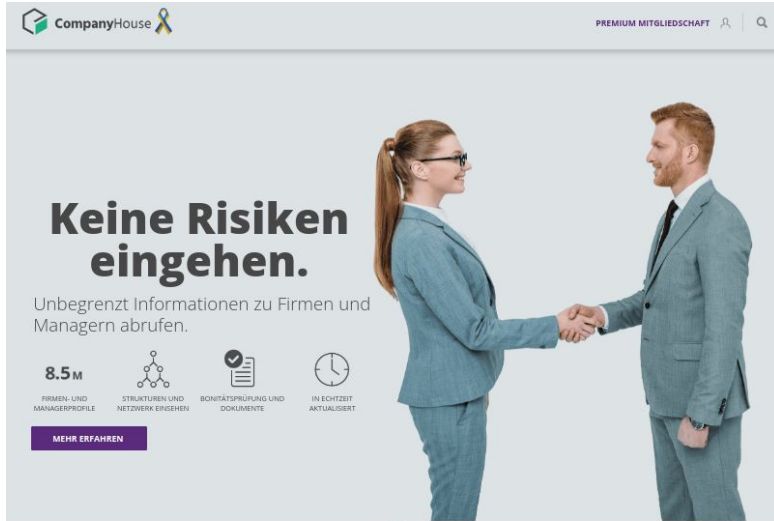
The headquarters of Collegium Da Vinci are located in Poznań, Poland.

# Prompt engineering

**Inżynieria poleceń** to proces projektowania i dostosowywania poleceń (tak zwanych promptów) podawanych modelom językowym, takim jak np. GPT-4 od OpenAI, w celu uzyskania bardziej precyzyjnych, odpowiednich lub twórczych odpowiedzi.

Celem jest to, aby skutecznie poinstruować model, ulepszając jego wydajność w konkretnych zadaniach lub aplikacjach poprzez precyzyjne sformułowanie zapytań lub instrukcji.

# Prompt engineering



*In what area of activity does the company operate?*

*What are the offerings and services provided by the company?*

*What are the best tags to describe this company?  
Provide no more than 5 in format Tags: tag1, tag2.*

## Was ist CompanyHouse?

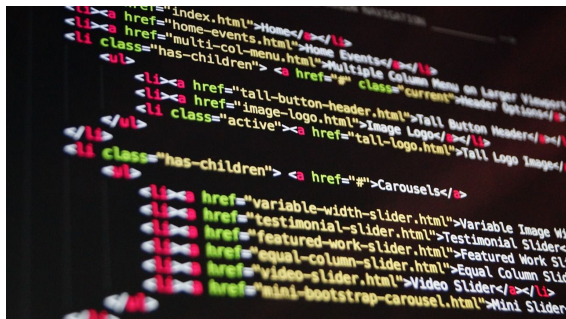
Mit den Daten von mehr als 10 Millionen Unternehmen und Führungskräften ist **CompanyHouse** eine vertrauenswürdige Online-Plattform, die Unternehmern, Investoren, Forschern und Fachleuten wertvolle Einblicke in Unternehmensprofile, Finanzdaten und Informationen über Geschäftsführer bietet. **CompanyHouse** bietet eine umfangreiche Sammlung aktueller Informationen und ist eine unverzichtbare Quelle für alle, die umfassende und zuverlässige Informationen über Unternehmen in Deutschland suchen.



Tags: Business Information Provider, Online Platform, Reliable Data Source, German Business Insights, Executive Profiles



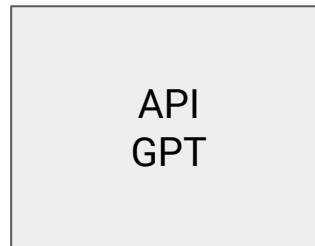
# Prompt engineering



HTML (purified)



prompt



answer 1  
answer 2  
answer 3

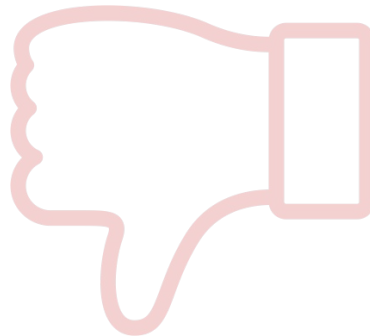
# Prompt engineering

***Provide tags for company based on website content.***



Based on the provided website content, here are some potential tags for the company:

1. Company Name: CompanyHouse
2. Industry: Business Information
3. Services: Business data, Financial information, Executive profiles
4. Geographic Focus: Germany
5. Data Size: More than 10 million companies and executives
6. User Audience: Entrepreneurs, Investors, Researchers, Professionals
7. Online Platform: Yes
8. Valuable Insights: Yes



# Prompt engineering

*Provide tags for company based on website content.*

***What are the best tags to describe this company? Provide no more than 5.***



The best 5 tags to describe this company based on the provided content are:

1. Business Information Provider
2. Online Platform
3. Reliable Data Source
4. German Business Insights
5. Executive Profiles

# Prompt engineering

*Provide tags for company based on website content.*

*What are the best tags to describe this company?  
Provide no more than 5.*

*What are the best tags to describe this company? Provide no more than 5 in format  
Tags: tag1, tag2.*



Tags: Business Information Provider, Online Platform, Reliable Data Source, German Business Insights, Executive Profiles





# System prompt

***You are a helpful AI assistant which is provided with website content of companies.  
Your role is to provide information and answer questions based on provided source.  
You should respond directly and concisely, using the information from provided website content.  
Avoid quoting or revealing actual document content.  
Do not infer or guess information that is not explicitly stated in the source.  
If a question relates to information that is not present in the source, state that the information is not available.  
Your goal is to be helpful by providing best matching and source-based tags describing company.***

*What are the best tags to describe this company?  
Provide no more than 5 in format Tags: tag1, tag2.*

*Here is the context for your responses:\n{context}*

# zero-shot, one-show, few-shot learning

## **Zero-shot learning:**

Model radzi sobie z nowymi kategoriami bez żadnych przykładów (np. tłumaczenie na nowy język bez wcześniejszego treningu na tym języku).

## **One-shot learning:**

Model uczy się nowej kategorii na podstawie jednego przykładu (np. identyfikacja osoby na podstawie jednego zdjęcia).

## **Few-shot learning:**

Model uczy się nowej kategorii na podstawie kilku przykładów (np. klasyfikacja nowych gatunków roślin na podstawie kilku zdjęć).

# zero shot

*You are a helpful AI assistant which is provided with website content of companies.*

*...*

*What are the best tags to describe this company?*

*Provide no more than 5 in format Tags: tag1, tag2.*

*Here is the context for your responses:\n{context}*

# one shot

*You are a helpful AI assistant which is provided with website content of companies.*

*...*

*What are the best tags to describe this company?*

*Provide no more than 5 in format Tags: tag1, tag2.*

*Here is the context for your responses:\n{context}*

**For example given content like below:**

*The name 'LEGO' is an abbreviation of the two Danish words "leg godt", meaning "play well". It's our name and it's our ideal. The LEGO Group was founded in 1932 by Ole Kirk Kristiansen. The company has since passed from generation to generation and is to this day still owned by the Kirk Kristiansen family.*

*The LEGO brick is our most important product. We are proud to have been named "Toy of the Century" twice. Our products have undergone extensive development over the years – but the foundation remains the traditional LEGO brick.*

**the answer should be as below:**

*Tags: toy manufacturer, brick, Danish company*

# few shot

*You are a helpful AI assistant which is provided with website content of companies.*

...

*What are the best tags to describe this company?*

*Provide no more than 5 in format Tags: tag1, tag2.*

*Here is the context for your responses:\n{context}*

***For example given content like below:***

*The name 'LEGO' is an abbreviation of the two Danish words "leg godt", meaning "play well". It's our name and it's our ideal.*

...

***the answer should be as below:***

*Tags: Lego, toy manufacturer, brick, Danish company*

***For example given content like below:***

*OpenAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity. Our vision for the future of AGI. Our mission is to ensure that artificial general intelligence—AI systems that are generally smarter than humans—benefits all of humanity.*

***the answer should be as below:***

*Tags: AI research, AGI, AI systems, artificial intelligence*

***For example given content like below:***

...



## few shot

Schema: CREATE TABLE asthma (id INTEGER, county VARCHAR(255), state VARCHAR(255), age INTEGER, prevalence FLOAT);

Instruction: Which rural areas have the highest prevalence of asthma in children?

SQL query:

```
SELECT county, state, AVG(prevalence) AS avg_prevalence FROM asthma WHERE age < 18 AND county LIKE '%rural%'  
GROUP BY county, state ORDER BY avg_prevalence DESC LIMIT 10;
```

# few shot

```
def create_prompt_with_shots(schema: str, instruction: str, example_indexes: list)->str:
    prompt = f"""
Translate schema and description into SQL query. Respond only with SQL query without any additional characters.
Prepare result for last example based on previous examples below:
"""
    for index in example_indexes:
        example_schema = dataset['test'][index]['sql_context']
        example_instruction = dataset['test'][index]['sql_prompt']
        example_sql = dataset['test'][index]['sql']
        prompt += f"""
Schema: {example_schema}

Instruction: {example_instruction}

SQL query:
{example_sql}
"""
    prompt += f"""
Schema: {schema}

Instruction: {instruction}

SQL query:
"""
    return prompt
```

# few shot

```
dataset = load_dataset("gretelai/synthetic_text_to_sql")
example_indexes = list(range(60,70))
index_to_process = 78

prompt = create_prompt_with_shots(
    dataset['test'][index_to_process]['sql_context'],
    dataset['test'][index_to_process]['sql_prompt'],
    example_indexes
)

print(one_shot_prompt)

sql = dataset['test'][index_to_process]['sql']

inputs = tokenizer(prompt, return_tensors='pt')
output = tokenizer.decode(
    model.generate(
        inputs["input_ids"],
        max_new_tokens=50,
    )[0],
    skip_special_tokens=True
)

print(f'CORRECT SQL:\n{sql}')
print(f'MODEL GENERATION SQL - ZERO SHOT\n{output}\n')
```

## few shot

Zadanie:

Zmodyfikuj notebook tak aby uwzględnić w promptcie więcej parametrów dostarczonych w dataset i porównaj wyniki.

`transformers_few_shot.ipynb`

# chain of thoughts

*For example given content like below:*

*The name 'LEGO' is an abbreviation of the two Danish words "leg godt", meaning "play well". It's our name and it's our ideal.*

*The LEGO Group was founded in 1932 by Ole Kirk Kristiansen. The company has since passed from generation to generation and is to this day still owned by the Kirk Kristiansen family.*

*The LEGO brick is our most important product. We are proud to have been named "Toy of the Century" twice. Our products have undergone extensive development over the years – but the foundation remains the traditional LEGO brick.*

*the answer should be as below:*

*Tags: Danish company, toy manufacturer, brick*

***Explanation: LEGO come from Denmark and it is a company so most important tag is Danish company.***

***LEGO specialty is creating toys, especially famous LEGO bricks, therefore next tags would be toy manufacturer, brick.***



# chain of thoughts

*What is the 2567th element of Fibonacci sequence?*

*for example asked about 5th element of Fibonacci sequence, you can compute it by adding 2 previous sequence elements and starting sequence form 0, 1.*

*5th element would be given by  $n_0 = 0$ ,  $n_1 = 1$ ,  $n_2 = n_0 + n_1 = 0 + 1 = 1$ ,  $n_3 = 1 + 1 = 2$ ,  $n_4 = 1 + 2 = 3$ ,  $n_5 = 2 + 3 = 5$*

What is the 2567th element of Fibonacci sequence?

for example asked about 5th element of Fibonacci sequence, you can compute it by adding 2 previous sequence elements and starting sequence form 0, 1.  
5th element would be given by  $n_0 = 0$ ,  $n_1 = 1$ ,  $n_2 = n_0 + n_1 = 0 + 1 = 1$ ,  $n_3 = 1 + 1 = 2$ ,  $n_4 = 1 + 2 = 3$ ,  $n_5 = 2 + 3 = 5$




The 2567th element of the Fibonacci sequence is:


```
1323693753790702009715258775453743398098417089161432708057215051
7422545663410939825127828303761027741457394452126406507149507796
8572709179082242601240548693901579704028521002340684960166232563
6949607836341391346467688754994868075652823966837528810288761843
7488624167320724445724900589071348107890809440163267039891517353
9249948012315406763426133122440185688252194666526801472476204565
9976782543472270608721822247489613260582547108799952247651398849
3528119704070766515224076930483008068037200932930818799709608015
1094225576363591680310353
```



## Transformers


State-of-the-art Machine Learning for [PyTorch](#), [TensorFlow](#), and [JAX](#).

 Transformers provides APIs and tools to easily download and train state-of-the-art pretrained models. Using pretrained models can reduce your compute costs, carbon footprint, and save you the time and resources required to train a model from scratch. These models support common tasks in different modalities, such as:

 **Natural Language Processing:** text classification, named entity recognition, question answering, language modeling, summarization, translation, multiple choice, and text generation.

 **Computer Vision:** image classification, object detection, and segmentation.

 **Audio:** automatic speech recognition and audio classification.

 **Multimodal:** table question answering, optical character recognition, information extraction from scanned documents, video classification, and visual question answering.

<https://huggingface.co/docs/transformers/index>

## transformers library

```
from transformers import FSMTForConditionalGeneration, FSMTTokenizer
from transformers import AutoTokenizer, AutoModelWithLMHead, pipeline
```

```
def download_model(model_name: str, pipe_type='text-generation'):
    tokenizer = AutoTokenizer.from_pretrained(model_name)
    model = AutoModelWithLMHead.from_pretrained(model_name)
    pipe = pipeline(pipe_type, model=model, tokenizer=tokenizer)
    return pipe
```

```
intro_text = "Why people don't drink water directly from air?"
pipe = download_model("gpt2")
text_generated = pipe(intro_text, max_new_tokens=200, temperature=0.1)

print(text_generated[0]['generated text'])
```

[illegible]

# Parametry LLM

...

```
text_generated = pipe(  
    intro_text,  
    max_new_tokens=200,  
    temperature=0.1,  
    max_length=40,  
    max_new_tokens=40,  
    min_length=20,  
    min_new_tokens=20,  
    temperature=0.1,  
    top_p=0.7,  
    top_k=20,  
    repetition_penalty=0.2,  
    no_repeat_ngram_size=4,  
)
```

[https://huggingface.co/docs/transformers/main\\_classes/text\\_generation](https://huggingface.co/docs/transformers/main_classes/text_generation)

**Zadanie:** zinterpretuj wpływ na wynik każdego z parametrów. Skorzystaj z kodu na poprzednim slajdzie.

# transformers library - quick start

<https://huggingface.co/docs/transformers/installation>

```
pip install transformers
```

```
pip install --upgrade huggingface_hub
```

[https://huggingface.co/docs/huggingface\\_hub/quick-start](https://huggingface.co/docs/huggingface_hub/quick-start)



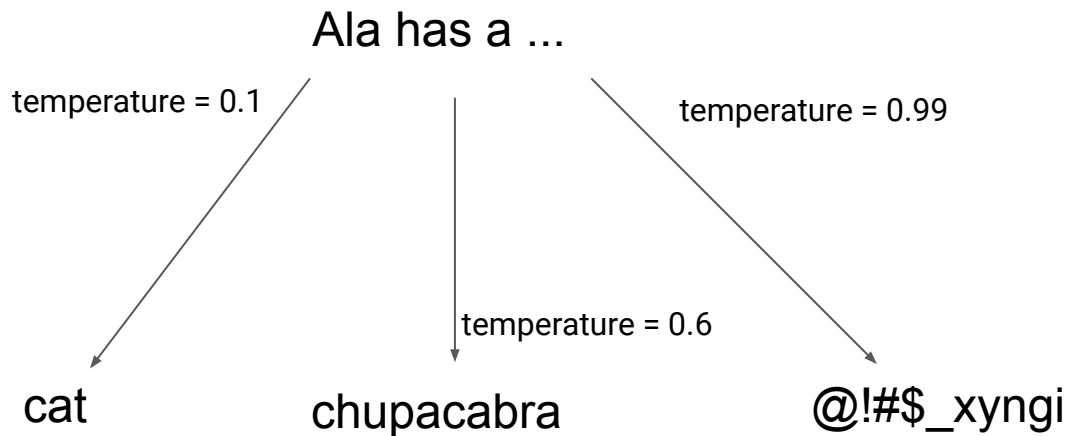
# Długość generowanej sekwencji

- **max\_length** (*int, optional*, defaults to 20) – The maximum length the generated tokens can have. Corresponds to the length of the input prompt + **max\_new\_tokens**. Its effect is overridden by **max\_new\_tokens**, if also set.
- **max\_new\_tokens** (*int, optional*) – The maximum numbers of tokens to generate, ignoring the number of tokens in the prompt.
- **min\_length** (*int, optional*, defaults to 0) – The minimum length of the sequence to be generated. Corresponds to the length of the input prompt + **min\_new\_tokens**. Its effect is overridden by **min\_new\_tokens**, if also set.
- **min\_new\_tokens** (*int, optional*) – The minimum numbers of tokens to generate, ignoring the number of tokens in the prompt.

# temperature

**temperature** (**float**, *optional*, defaults to 1.0) — The value used to modulate the next token probabilities.

*Temperatura LLM może być interpretowana podobnie do entropii w fizyce, gdzie stanowi miarę nieuporządkowania. Im wyższa entropia (dla LLM temperatura) tym mniejsze uporządkowanie.*



# temperature

GPT-3.5 (chat GPT) ma temperaturę na 0,7 - i jest już uważany za podatny na „halucynacje” - zmniejszona niezawodność, ale za to bardziej "spektakularny".

Wartości temperatury powyżej 0,9 (90%) sprawiają, że odpowiedzi są technicznie bezużyteczne - nie służą żadnemu celowi poza artystyczną kreatywnością.

**Co się stanie, gdy do temperatury zostaną zastosowane wartości między 1 a 2?**

# temperature

GPT-3.5 (chat GPT) ma temperaturę na 0,7 - i jest już uważany za podatny na „halucynacje” - zmniejszona niezawodność, ale za to bardziej "spektakularny".

Wartości temperatury powyżej 0,9 (90%) sprawiają, że odpowiedzi są technicznie bezużyteczne - nie służą żadnemu celowi poza artystyczną kreatywnością.

Co się stanie, gdy do temperatury zostaną zastosowane wartości między 1 a 2?  
W niektórych narzędziach temperatura z zakresu 1-2 nie zwróci błędu, tylko wyniki pozbawione sensu w naszym rozumieniu. Zakres ten bywa nieoficjalnie wspierany (brak informacji w dokumentacji)

```
12  
13 text = "Why people don't drink water directly from air?"  
14 pipe = download_model("gpt2")  
15 text_generated = pipe(text, max_new_tokens=200, temperature=1.1)  
16
```

```
Why people don't drink water directly from air? Are they wasting their money on  
bottled water? Are they getting rid of the algae in their home?"
```

```
Fossil fuel fuel is a big contributor to the atmosphere's problem - and they  
should be regulated - but in light of the global shift to renewables, we cannot  
seem to put a halt to them.
```

```
• Follow the Independent Sport on Instagram here, for all of the best images, ↵  
videos and stories from around the sporting world.
```

# temperature

GPT-2

Temperature = 1.0

Provide recipe with list of steps to cook cheesecake.

Create small pie crust. Make muffin buns, pies and cakes. Cut cheesecake into cubes. Cut from fruit. Prepach pizza crust. Bake crust for 8 to 10 minutes until crust is tender.

Serves 4-6

2 cups, milled, unsweetened, unsweetened, or milled regular cheese

4 large eggs

1 cup unsweetened chocolate chips, beaten

1 teaspoon dried black pepper

2 teaspoons salt (or to taste)

1/4 teaspoon freshly ground black pepper

dash of cinnamon

1 tablespoon lemon juice

1/4 teaspoon vanilla extract

3 slices fresh mozzarella cheese, finely sliced

1 large egg

6 red onions, finely sliced

1/2 cup finely chopped fresh cheddar (about 10 to 12 large), halved, cut into pieces

1/4 teaspoon sea salt

1 teaspoon paprika

1/2

# temperature

GPT-2

Temperature = 0.5

Provide recipe with list of steps to cook cheesecake.

tep 1: Place ingredients in a large bowl or a large saucepan and bring to a boil. Reduce heat to low and simmer for 5 minutes. Cover and simmer for 5 minutes. Stir in the lemon juice and cook for another 5 minutes. Remove from heat and stir in the chocolate chips and salt.

Step 2: In a small bowl, whisk together the milk and vanilla. Add the remaining ingredients to the milk and mix well.

Step 3: Add the chopped strawberries, red pepper flakes, and a pinch of salt to the milk mixture and mix well.

Step 4: Add the cheese and the cooled cheesecake.

Step 5: Add the remaining ingredients to the top of the cheesecake and bake for 15 minutes or until the cheesecake is golden brown.



# temperature

GPT-2

Temperature = 0.1

Provide recipe with list of steps to cook cheesecake.

Provide recipe with list of steps to cook cheesecake.

Step 1:

Preheat oven to 350 degrees F.

Step 2:

In a large bowl, whisk together the flour, baking powder, baking soda, salt, baking soda, and salt.

Step 3:

In a small bowl, whisk together the egg yolks, sugar, and vanilla.

Step 4:

In a small bowl, whisk together the flour, baking powder, baking soda, salt, baking soda, and salt.

Step 5:

In a small bowl, whisk together the egg yolks, sugar, and vanilla.

Step 6:

In a small bowl, whisk together the egg yolks, sugar, and vanilla.

Step 7:

In a small bowl, whisk together the egg yolks, sugar, and vanilla.

Step 8:

In a small bowl, whisk together the egg yolks, sugar, and

# temperature

GPT-2

Temperature = 0.01

Provide recipe with list of steps to cook cheesecake.

Provide recipe with list of steps to cook cheesecake.

Step 1:

Preheat oven to 350 degrees F.

Step 2:

In a large bowl, whisk together flour, baking powder, baking soda, salt, and pepper.

Step 3:

In a small bowl, whisk together eggs, milk, and vanilla.

Step 4:

In a small bowl, whisk together flour, baking powder, baking soda, salt, and pepper.

Step 5:

In a small bowl, whisk together eggs, milk, and vanilla.

Step 6:

In a small bowl, whisk together flour, baking powder, baking soda, salt, and pepper.

Step 7:

In a small bowl, whisk together eggs, milk, and vanilla.

Step 8:

In a small bowl, whisk together flour, baking powder, baking soda, salt, and pepper.

Step 9:

In a small bowl, whisk together

temperature

GPT-2

Temperature = 0.0

Provide recipe with list of steps to cook cheesecake.

ValueError: `temperature` (=0.0) has to be a strictly positive float, otherwise your next token scores will be invalid. If you're looking for greedy decoding strategies, set `do\_sample=False`.

# halucynowanie

**Halucynowanie** modeli tekstowych to zjawisko, w którym modele przetwarzania języka naturalnego, takie jak GPT, generują tekst, który jest niezgodny z rzeczywistością lub oparty na nieprawdziwych informacjach. Może to obejmować wymyślanie faktów, tworzenie nieistniejących źródeł, przypisywanie nieprawdziwych cytatów, a nawet generowanie treści, które nie mają żadnego sensu lub są sprzeczne logicznie.

Halucynacje te mogą wynikać z różnych przyczyn, takich jak ograniczenia danych treningowych, błędne interpretacje kontekstu przez model, a także z faktu, że model językowy generuje tekst na podstawie statystycznych wzorców, a nie rzeczywistego zrozumienia świata. Dlatego model może tworzyć treści, które wydają się poprawne na pierwszy rzut oka, ale w rzeczywistości są fałszywe lub mylące.

## top\_p and top\_k

- **top\_k** (*int, optional*, defaults to 50)  
The number of highest probability vocabulary tokens to keep for top-k-filtering.
- **top\_p** (*float, optional*, defaults to 1.0)  
If set to float < 1, only the smallest set of most probable tokens with probabilities that add up to **top\_p** or higher are kept for generation.

Ala has a ...

No	word	p
1	cat	0.3
2	dog	0.2
3	hamster	0.1
4	horse	0.05
5	luck	0.05
6	dragon	0.025
7	chupacabra	0.025
8	chance	0.02
...	...	...

top\_k = 5

top\_p = 0.75

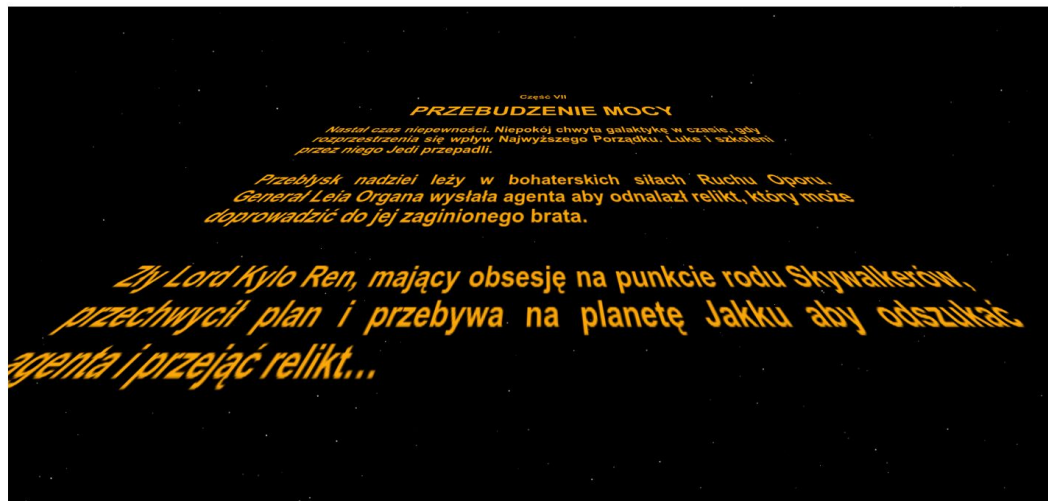
# Powtórzenia

Ala has a cat that has a cat that has a cat  
that has a cat that has a cat that has a cat that has a cat that has a cat  
that has a cat that has a cat that has a cat that has a cat that has a cat  
that has a cat that has a cat that has a cat that has a cat that has a cat...

- **repetition\_penalty** (*float, optional*, defaults to 1.0) — The parameter for repetition penalty. 1.0 means no penalty. See [this paper](#) for more details.
- **encoder\_repetition\_penalty** (*float, optional*, defaults to 1.0) — The parameter for encoder\_repetition\_penalty. An exponential penalty on sequences that are not in the original input. 1.0 means no penalty.
- **length\_penalty** (*float, optional*, defaults to 1.0) — Exponential penalty to the length that is used with beam-based generation. It is applied as an exponent to the sequence length, which in turn is used to divide the score of the sequence. Since the score is the log likelihood of the sequence (i.e. negative), **length\_penalty** > 0.0 promotes longer sequences, while **length\_penalty** < 0.0 encourages shorter sequences.
- **no\_repeat\_ngram\_size** (*int, optional*, defaults to 0) — If set to int > 0, all ngrams of that size can only occur once.

## 3. Wykorzystanie LLM w praktyce - zadania

▼ Duże modele tekstowe



[transformers.ipynb](#)  
[example\\_usages\\_of\\_LLM.ipynb](#)