AI UKRAINE™ 2021
Produced by AltexSoft

PRESTE

Automatic text summarization: is GPT3 the best and only?

Natalya Avanesova, Head of NLP

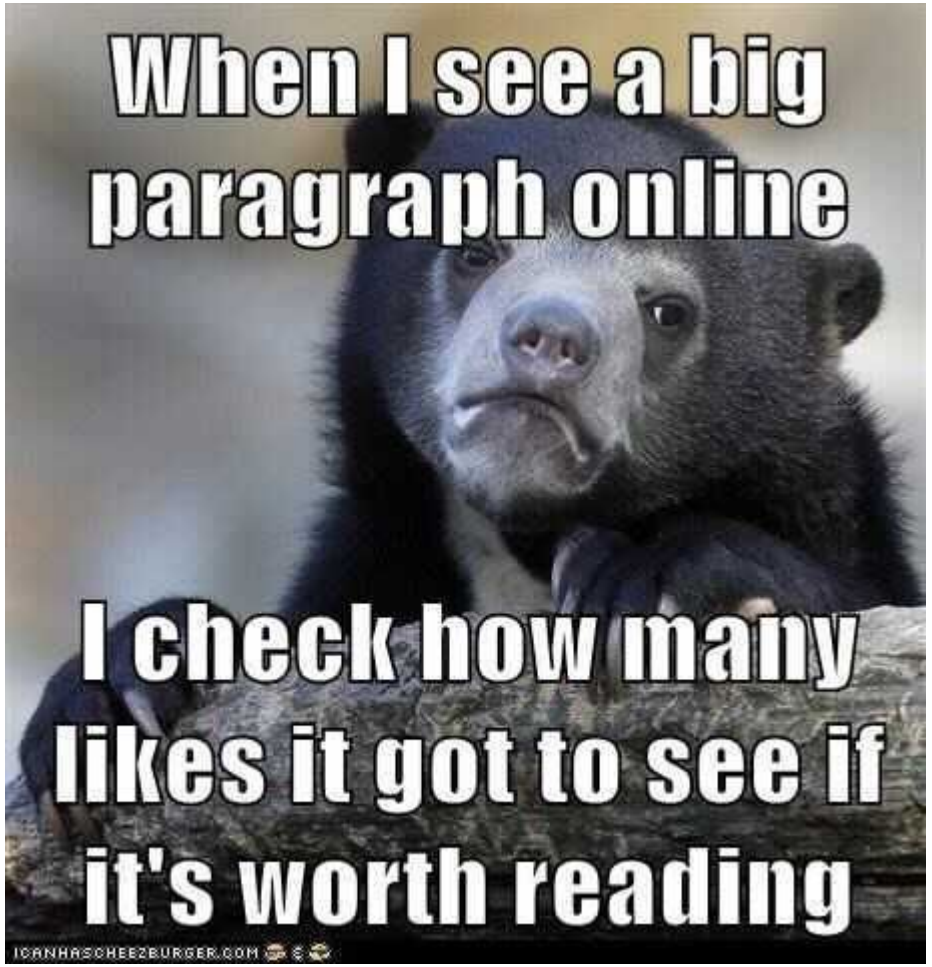
PRESTE

We carry out B2B AI projects

from design to production

SCAN ME

contact@preste.ai



https://cheezburger.com/7355291904

Every-day examples of text summaries

- headlines (from around the world)
- outlines (notes for students)
- minutes (of a meeting)
- previews (of movies)
- synopses (soap opera listings)
- reviews (of a book, CD, movie, etc.)
- digests (TV guide)
- biography (resumes, obituaries)
- abridgments (Shakespeare for children)
- bulletins (weather forecasts/stock market reports)
- sound bites (politicians on a current issue)
- histories (chronologies of salient events)

Advances in Automatic Text Summarization (The MIT Press)

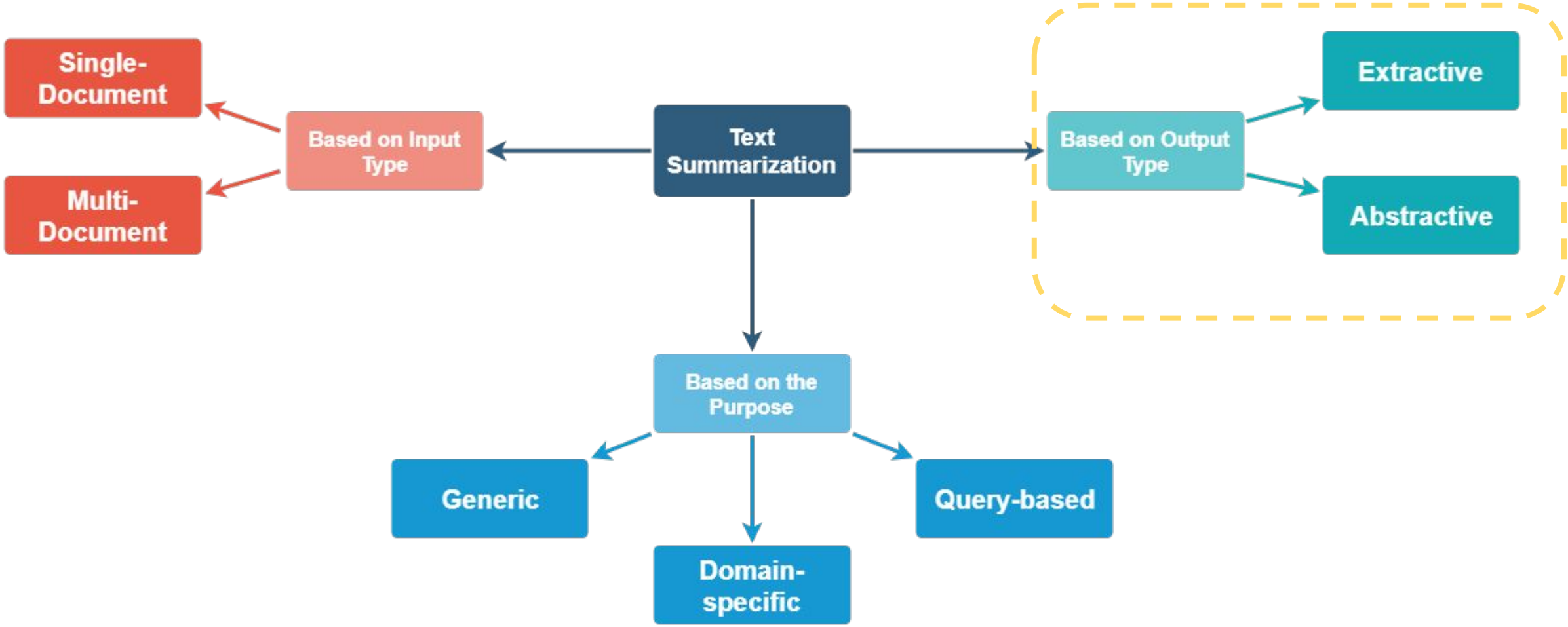
https://www.blendspace.com/

Long story short... What is the need for text summarization?

- Summaries reduce reading time.
- When researching documents, summaries make the selection process easier.
- Summarization improves the effectiveness of indexing.
- Summarization algorithms are less biased than human summarizers.
- Personalized summaries are useful in question-answering systems as they provide personalized information.
- Summarization tools reduce the need to find expertise outside the company

PRESTE

Approaches for text summarization



Types of text summarization. Source: Chauhan 2018.

PRESTE

(a) Extractive Summarization

Source Text: Peter and Elizabeth took a taxi to attend the night party in the city.

While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Peter and Elizabeth attend party city. Elizabeth rushed hospital.

1. Split text into sentences.
2. Parse through each sentence, identify which ones are more important than others based on certain parameters and assign them a score.
3. Pick out the sentences that rank above others.
4. Add selected sentences to the summary

(b) Abstractive Summarization

Source Text: Peter and Elizabeth took a taxi to attend the night party in the city.

While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Elizabeth was hospitalized after attending a party with Peter.

1. Read the text.
2. Analyze the underlying meaning of the text and the sentences.
3. Pick out the important topics and create new sentences (may or may not use vocabulary from the original text).
4. Add new sentences to the summary

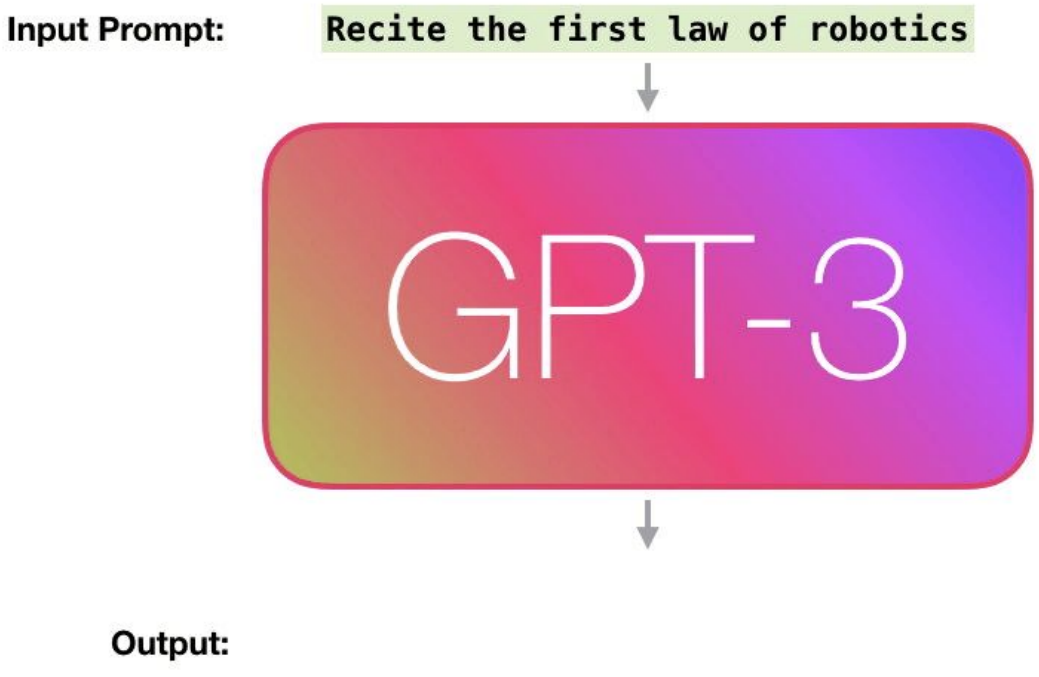
Illustrating extractive vs abstractive summarization. Source: Adapted from Opidi 2019

PRESTE

GPT3* by OpenAI

**Generative Pre-Trained Transformer -3*

- uses 175 billion parameters
- was trained with almost all available data from the Internet
- can excel in task-agnostic performance without fine tuning
- uses prompt programming
- is not free
- not available for everyone
- can be used only through API



http://jalammar.github.io/how-gpt3-works-visualizations-animations/

PRESTE

PEGASUS* by Google

**Pre-training with Extracted Gap-sentences for Abstractive Summarization*

- designed for abstractive summarization
- achieves state-of-the-art results on 12 diverse summarization datasets
- CNN/DailyMail dataset used for training
- can be fine-tuned the on your dataset
- open-sourced
- can used locally

https://1.bp.blogspot.com/-TSor4o51jGI/Xt50lkj6bII/AAAAAAAAGDs/

PRESTE

T5* by Google

**Text-to-Text Transfer Transformer*

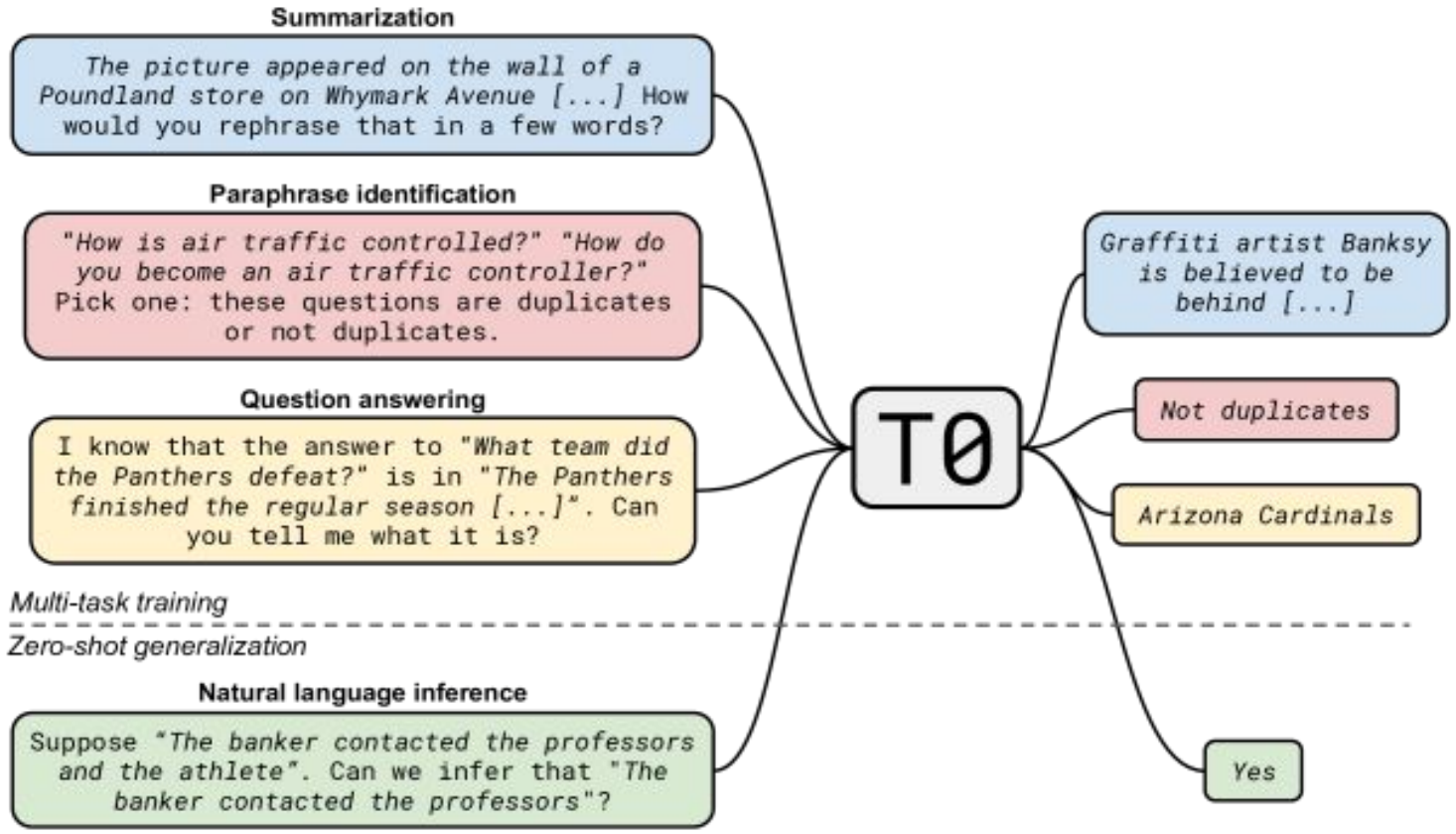
- is pre-trained on several unsupervised and supervised objectives
- is trained on the C4 dataset
- the model is conditioned to generate the correct output based on a textual prompt
- can be fine-tuned
- open-sourced
- can be used locally

PRESTE

T0*

**T Zero*

- shows zero-shot task generalization on English natural language prompts, outperforming GPT-3 on many tasks, while being 16x smaller
- T0* models are based on T5
- open-sourced
- numerous English supervised datasets converted into prompts for training



PRESTE

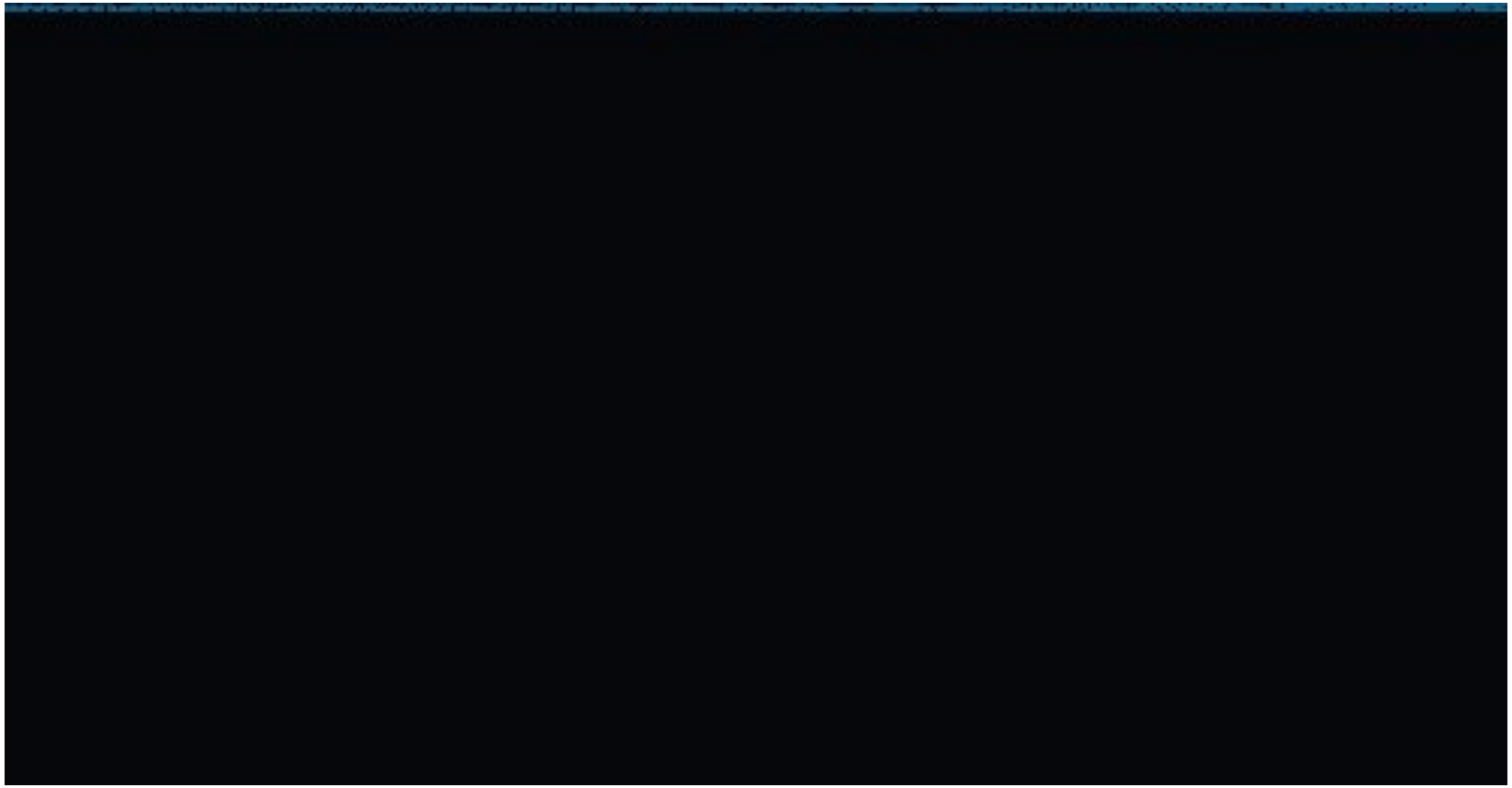
Metrics used for summarization results evaluation

Metrics	Description	Explanations	References
Tokens Comparison	in a new text created from both input texts unique and common tokens are highlighted with different colours	• highlighted with light green = unique tokens for full text • highlighted with pink = unique tokens for summarized text • not highlighted = common tokens for both texts	good abstractive summary needs to contain more unique (pink coloured) and less common with full text tokens
Cosine Similarity	is the cosine of the angle between two vectors, created from two input texts. The cosine shows semantica similarity between them	• 0 = no similarity • 1 = identical	correct cut-off value always depends on the situation; in our case **0.5** might be a good starting point

"best" summary requires more pink coloured text + cosine similarity >= 0.5

PRESTE

Let Transformers summarize for you...



Summarize: Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines, and ensure every country receives them and can roll them out to protect their people, starting with the most vulnerable.

GOOGLE/PEGASUS-XSUM: SUMMARY

The World Health Organization (WHO) is working with partners to develop, manufacture and deploy safe and effective vaccines.

TOLOKA/T5-LARGE-FOR-TEXT-AGGREGATION: SUMMARY

equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines in clinical and preclinical development, generally updated twice a week.

BIGSCIENCE/T0PP: GENERATED TEXT

As the world's largest outbreak of seasonal influenza draws to a close, it is heartening to see so many vaccines in development to protect against the new virus.

GPT3 SUMMARIZATION: OUTPUT

The world is mobilizing to develop and deploy vaccines against the COVID-19 pandemic. Vaccines are a game-changing tool, but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds.

GPT3 result

Summaries Evaluation

full text

Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not

summarized text

The world is mobilizing to develop and deploy vaccines against the COVID-19 pandemic. Vaccines are a game-changing tool, but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds.

Clear

Submit

differences between texts: Output 0.04s

\+ -

EquThe world itas moble access to safe and effective vaccines is critical to endizing to develop and deploy vaccines against the COVID-19 pandemic. so Vaccint is huges are a game-changing tool, but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, phy encouraging to sicee so manlly vaccines proving and goisng intao developmentcing and avoiding crowds. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines, and ensure every country receives them and can roll them out to protect their people, starting with the most vulnerable.

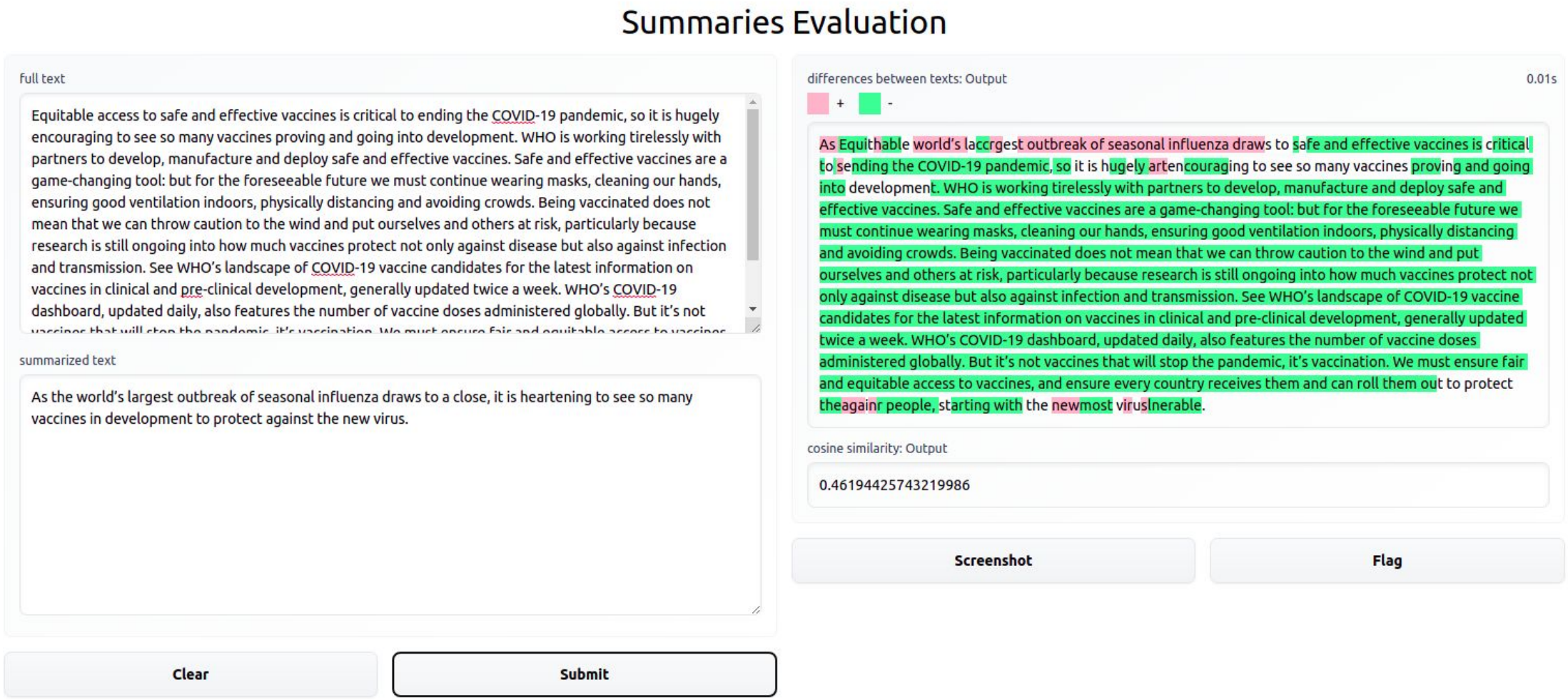
cosine similarity: Output

0.6064343625611002

Screenshot

Flag

T0pp result


Summaries Evaluation
full text
Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not
summarized text
As the world's largest outbreak of seasonal influenza draws to a close, it is heartening to see so many vaccines in development to protect against the new virus.
Clear
Submit
differences between texts: Output
0.01s
+
-
As Equithable world's laccrgest outbreak of seasonal influenza draws to safe and effective vaccines is critical to sending the COVID-19 pandemic, so it is hugely artencouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines, and ensure every country receives them and can roll them out to protect theagainr people, starting with the newmost virusInerable.
cosine similarity: Output
0.46194425743219986
Screenshot
Flag

Pegasus result

Summaries Evaluation

full text

Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not

summarized text

The World Health Organization (WHO) is working with partners to develop, manufacture and deploy safe and effective vaccines.

Clear | Submit

differences between texts: Output 0.01s

\+ -

TEquitable access to safe and effective vaccines is critical to ending the CWorld Health OrgVID-19 pandemic,zation (so it is hugely encouraging to see so many vaccines proving and going into development. WHO) is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines, and ensure every country receives them and can roll them out to protect their people, starting with the most vulnerable.

cosine similarity: Output

0.546358364708153

Screenshot | Flag

T5 result

Summaries Evaluation

full text

Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines

summarized text

equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines in clinical and preclinical development, generally updated twice a week.

Clear | Submit

differences between texts: Output 0.25s

\+ -

Eequitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic, so it is hugely encouraging to see so many vaccines proving and going into development. WHO is working tirelessly with partners to develop, manufacture and deploy safe and effective vaccines. Safe and effective vaccines are a game-changing tool: but for the foreseeable future we must continue wearing masks, cleaning our hands, ensuring good ventilation indoors, physically distancing and avoiding crowds. Being vaccinated does not mean that we can throw caution to the wind and put ourselves and others at risk, particularly because research is still ongoing into how much vaccines protect not only against disease but also against infection and transmission. See WHO's landscape of COVID-19 vaccine candidates for the latest information on vaccines in clinical and pre-clinical development, generally updated twice a week. WHO's COVID-19 dashboard, updated daily, also features the number of vaccine doses administered globally. But it's not vaccines that will stop the pandemic, it's vaccination. We must ensure fair and equitable access to vaccines, and ensure every country receives them and can roll them out to protect their people, starting with the most vulnerable.

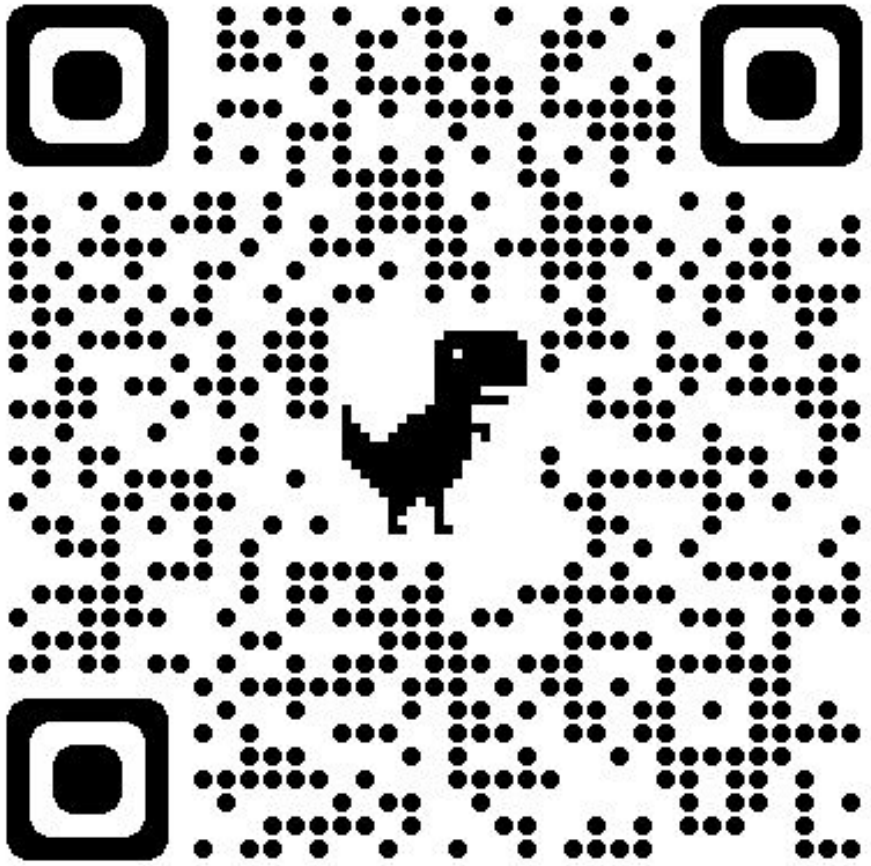
cosine similarity: Output

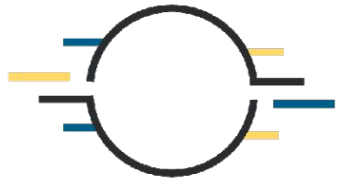
0.6922935180190629

Screenshot | Flag

PRESTE

Let T0 summarise for you...





PRESTE

THANK YOU!