Disruptions Caused by AI in the Language Industry and Education

across ‘ ‘
Language Technology
for a Globalized World.

WF **WORDFAST**

SDL

memo Q
TRANSLATOR PRO

MEMSOURCE CLOUD

Is the AI a challenge or a threat?

Disruptions or innovations in bi-and multilingual communication tasks? 2016-2022

- Rise in Translation and processing demand for multilingual content
- Machine translation (MT), Computer-assisted translation CAT)
- Terminology management
- Neural MT for professional and non-professionals

LLM- Large language models

- computational neural network-based language models with many parameters, trained on large amounts of text using machine learning algorithms

- ChatGPT and other AI-based platforms

Natural Language Processing

Natural Language Processing is the ability of a computer program to (understand?) human language as it is spoken and written. It is part of artificial intelligence.

What can we do with language data?

Summarization

Part-of-Speech-Tagging – assigns a tag to every token in a document based on its part of speech

Text categorization

Sentiment analysis – detects positive or negative sentiment from internal or external data sources. Commonly used to provide information about brands, **products or services**

Named-entity recognition – categorizes entities in text, e.g. names or locations; e.g. California (location), Mary (name)

Target groups

- Editors

- Marketing specialists

- Authors

- Academics and researchers

- Translators and interpreters

- Lawyers

- Teachers

- ....

Opportunities for all

- Visibility
- Leadership
- Networking
- New ideas and concepts
- Economization of tasks
- Labor opportunities
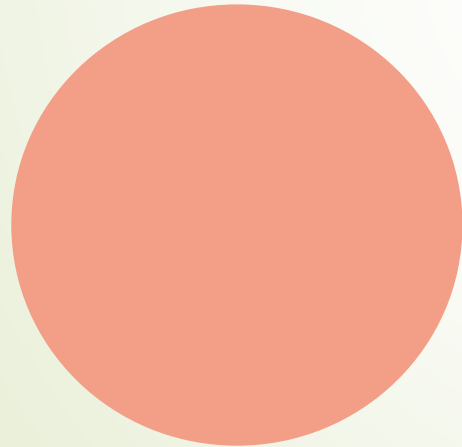- Emerging domains for research

Benefits vs Costs of AI generated models
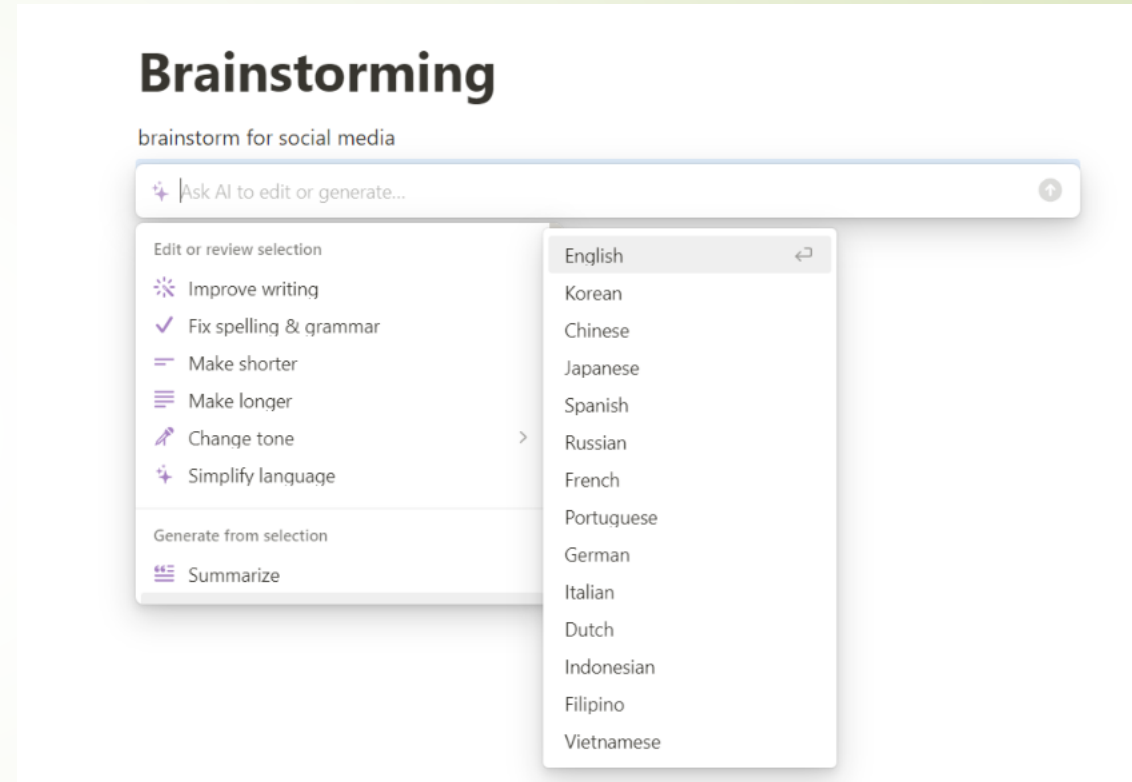**adopted by prof. Lynne Bowker**

- Adaptable
- New market needs
- Productivity boost
- Inspiration

- Hallucinations
- Data Sensitivity
- Inconsistency/Unpredictability
- Lack of creativity
- Marginalization of low-resource languages
- Potential misuse
- Rates down/remuneration issues
- Privacy
- Data dispossession
- ......
- ......

AI Tool for processing language contents

A. Writing

B. Editing

C. Translation

D. Preparation of glossaries

E. Comprehension of the read text

F. Other (research, gisting)

G. ......

CHAT GPT

- Founded in 2015.

- A language model that works on 570 GB text datasets.

- The Largest neural network at launching (175 billion parameters).

- It is able to predict the next token given a sequence of tokens.

- It is able to learn.

- GPT 3.5 data is only until 2021.

- Images on the internet are not accurate ("blurry jpeg").

- Artificial intelligence can "mislead" and give nonsensical answers.

- Confidential information.

- Inaccurate machine translation than other tools?

- Less accurate in less spoken languages - (see Hendy et al. 2023).

- Comparable for EN > EC general, less accurate for EN > DE + domain specific - (Savenkov 2023).

Chatbots and Platforms

Chatbots:

Bing

Google Bard

Platforms:

GPT 3,5

ChatGPT Plus

Phrase

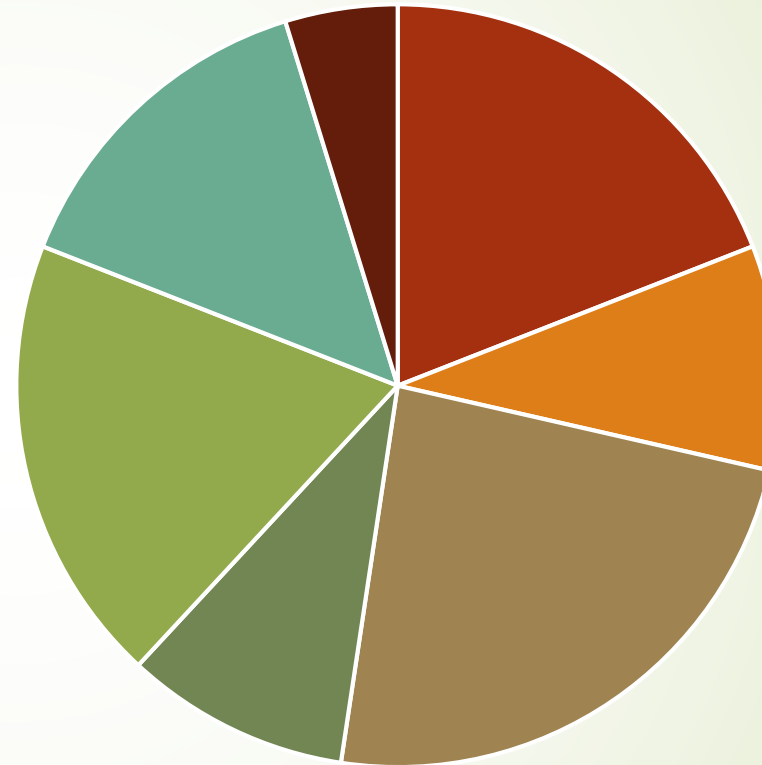Grammarly

Gaby-T

Notion

CoTranslatorAI

Other:

DeepL

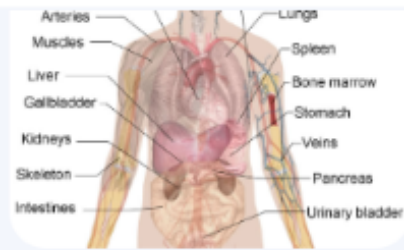Memrise (for language learning)

....

■ Writing ■ Editing ■ Translation ■ Glossaries ■ Text understanding ■ Others ■ No experience

What else is AI good for?

AI does not replace professional translation!

- Brainstorming,

- Expository writing,

- Correct spelling and grammar,

- Looking up vocabulary and terms

- Doing research to understand the theme (Scispace, Copilot,Textoptimizer)

- Searching equivalents, synonyms, antonyms, etc.,

- Condensing sentences,

- Breaking up run-on sentences,

- Reducing repetitions,

- Replacing words,

- Paraphrasing (ParaphraseTool, Simplif.ai, Genei)

- Formatting the text, such as headings,

- Changing the tone or style of the language content,

- Creating mono-, bi- or multilingual terminology glossaries

- Aligning documents

- Conversion of measure units

- Transcriptions (otter.ai, Sonix. ai, Whisper. ai)

Sure, here is a table describing the parts of the human body in the picture, how they work, and their Slovak equivalents:

English	Slovak	Function
Brain	Мозг	The control center of the body, responsible for thinking, feeling, and movement.
Heart	Srdce	A muscular organ that pumps blood throughout the body.
Lungs	Pľúca	Organs that allow the body to breathe and take in oxygen.

Be inventive, do a lot of free brainstorming

- Enter a command for the AI in a common language (in most languages)
- Don't be afraid to use creative prompts, such as "10 unconventional ideas".
- Select expressions from information gathered through brainstorming
- Create text from your own ideas
- Create a longer/shorter...
- Write a text based on the assignment
- Write a text from an outline point
- Write a text from a series of outline points
- Translate the text into another language

LLM Limitations

- variations in training data
- harmful effects in biased responses
- training data blind spots hinder the model
- the scope of the datasets is limited
- ChatGPT has limited knowledge of events before 2021
- ChatGPT models have inaccurate responses
- it produces incorrect answers without realising that the answer is not based on reality
- e.g. the model can generate a price quote based only on the training data

AI Prospectives

- Proactive vs. Reactive approach (critical, decision-making);
- Mobilizing associations (guidelines, policies);
- Diversifying skillsets (consulting);
- Definining use cases (where and why);
- Developing creativity (broadening contexts)
- Define legal frameworks
- Define the educational goals (professionals, experts, clients)

Questions?

Thank you for your attention

jozef.stefcik@euba.sk