

# **Krosna XXI wieku**

**(czyli GPT zabierze nam pracujących)**

Wystąpienie na Konferencji Technologie Jutra  
w Zespole Szkół nr 2 w Wieluniu



# Plan

- Teraźniejszość-, czyli modele językowe
- Teraźniejszość+, czyli Microsoft Copilot
- Nowe czy nie? Retrospekcja
- Co robić, jak żyć?
- (Dodatek) Przyszłość+, czyli generalna (uniwersalna) SI



# Modele językowe i GPT

- Czym jest GPT?
  - Modelem językowym (tylko!)
  - Modelem językowym (przełomowym ze względu na łatwość użycia)
- Czym jest model językowy?
  - Modelem matematycznym, który próbuje odwzorować znaczenie elementów języka (słów) i związki między nimi
  - KRÓL – MĘŻCZYZNA + KOBIETA = KRÓLOWA



# Modele językowe i GPT

- Czy znamy modele językowe inne niż GPT?
- Oczywiście:
  - Podpowiedzi tekstu w telefonach
  - Autokorekta w edytorze
  - Podpowiedzi w środowiskach programistycznych
  - Wyszukiwarki i translatory
    - (które już od dawna działają na wysokim poziomie)



# Modele językowe i GPT

- Na czym polega rewolucyjność GPT?
  - Bardzo dokładne odwzorowanie języka naturalnego i to nie tylko angielskiego
- Jak zostało to osiągnięte?
  - Dzięki poświęceniu ogromnych zasobów obliczeniowych (a także ludzkiej pracy) na wytworzenie modelu metodami uczenia maszynowego, na ogromnym zasobie danych
  - Dla porównania, wiele poprzednich modeli językowych polegało na odwzorowaniu zasad języka w formie dedykowanego algorytmu



## Modele językowe i GPT

- W uproszczeniu można powiedzieć, że wielkie modele językowe są przede wszystkim gigantyczna pamięcią słów oraz ich powiązań
- Odpowiednio często powtarzające się w danych uczących sekwencje zostają przełożone na liczbowa reprezentację powiązań, zaś dzięki olbrzymiej skali pozwala to uzyskać spektakularne wyniki, nawet pseudotwórcze (poezja!)
- Stanowi to pośrednią ilustrację powtarzalności konstrukcji języka, zaś anegdotyczny dowód przeprowadziłem nieświadomie na sobie samym!



# Modele językowe i GPT

- Czego GPT nie robi (czym nie jest)?
  - Nie wytwarza nowej wiedzy
  - Nie „rozumie” i nie ma „poglądów” – choć wytwarza złudne wrażenie
    - Sam zresztą to „przyznaje”
    - Nie rozwiązuje problemów np. matematycznych (pozornie tak – ale poprzez zacytowanie rozwiązania istniejącego wcześniej)
    - Może udzielić sprzecznych odpowiedzi po odpowiednim zadaniu pytań (casus ogórków kiszonych i konserwowych)
    - **Nie ma wątpliwości, nie zadaje pytań**



## Modele językowe i GPT

- Generalizując można powiedzieć, że istotną różnicą między SI w obecnej formie a sposobem rozumowania człowieka jest kontekstowość
- Sposób rozumowania, kojarzenia, wyciągania wniosków przez człowieka zmienia się poprzez oddziaływanie kontekstu, w którym czynność jest wykonywana
- Kontekst można tłumaczyć jako okoliczności towarzyszące, pragmatyczny cel, hierarchię wartości, samoświadomość, moralność, oddziaływanie społeczne...





## Modele językowe i GPT

- Działanie kontekstowe – prosty przykład (nieco zmodyfikowany przykład podawany przez prof. Hinton)
- Jak przetłumaczymy zdania:
  - The giraffe didn't fit to bag because it was too small
  - 
  - The giraffe didn't fit to box because it was too small
  - 
  - The giraffe didn't fit to box because it was too big
  -



## Modele językowe i GPT

- Działanie kontekstowe – prosty przykład (nieco zmodyfikowany przykład podawany przez prof. Hinton)
  - Jak przetłumaczy to model językowy (tu Google Translate):
    - The giraffe didn't fit to bag because it was too small
    - Żyrafa nie zmieściła się do torby, ponieważ była za mała
    - The giraffe didn't fit to box because it was too small
    - Żyrafa nie zmieściła się do pudełka, ponieważ było za małe
    - The giraffe didn't fit to box because it was too big
    - Żyrafa nie zmieściła się do pudełka, ponieważ było za duże
- (\*) Ale Czat Bing (GPT 4) radzi sobie z tym



## Modele językowe i GPT

- To, że GPT korzysta z danych wyuczonych, ma fundamentalne znaczenie dla wartości generowanych wyników
  - Znajdziesz odpowiedź, o ile ktoś już odpowiedział na Twoje pytanie...
  - ... i o ile ta odpowiedź znalazła się w zbiorze uczącym
  - Czy wszystkie znane odpowiedzi są w zbiorze uczącym?
  - Czy wszystkie odpowiedzi zawarte w tym zbiorze zobaczysz w odpowiedzi GPT?



## Modele językowe i GPT

- Czy kontrolujesz zbiór uczący i proces uczenia?
- NIE. To niemożliwe. Uczenie takiego modelu to petabajty danych i olbrzymia moc obliczeniowa zużywająca olbrzymie zasoby energii, a także zasoby pracy ludzkiej
- KTO będzie więc kontrolował model? Jaki będzie cel tego podmiotu?



## Modele językowe i GPT

- Czy proces nauki i w rezultacie wyniki działania modelu językowego mogą podlegać manipulacjom?
- To oczywiste i już to obserwujemy
- O czym wiemy?
  - Poprawność polityczna
  - Przetłumacz „I am a nurse”
    - Google Translate: „Jestem pielęgniarką”
    - Bing AI: „Tłumaczenie na polski: “Jestem pielęgniarką” lub “Jestem pielęgniarem” (w zależności od płci).”
- O czym nie wiemy?
- A co z deep fake?



## Modele językowe i GPT

- Założmy na chwilę absolutnie idealistyczne cele podmiotu kontrolującego model językowy
- Jakie zjawiska może wytworzyć sam model?
  - Bańka informacyjna
  - Samonapędzająca się produkcja niezbyt wartościowych treści
  - Lem opisał już te problemy (Cyberiada)...
  - ... i już je wszystkie mamy, pytanie dotyczy skali



## Modele językowe i GPT

- Na czym polega różnica w obecnej sytuacji?
- Dotychczas trzeba było wyszukiwać źródła i dokonywać samodzielnie kompilacji treści w nich zawartych
- W ramach eksplorowania źródła można było rozważyć jego wiarygodność – było ono znane
- W ramach kompilacji można było dostrzegać sprzeczności, co skłaniało do dalszych dociekań
- Tak zaawansowany model językowy jak GPT „robi to za nas” i „nie ma wątpliwości”



## Modele językowe i GPT

- Czy można uznać, że GPT „zabierze nam pracę”?
- Tak, jeżeli wykonujemy tylko to, co „potrafi” GPT – wyszukać i bez wnikania w znaczenie skompilować treść
- Czy GPT dostarczy nam nowe perspektywy zatrudnienia?
  - Weryfikacja wiarygodności treści
  - Specjalizacja w obsłudze rozwiązania („prompt engineering”)





# Modele językowe i GPT

- Źródła – część 1

- <https://oko.press/chatgpt-to-przelom-a-przy-okazji-wielki-eksperyment-spoeczny>
- <https://oko.press/chatgpt-cala-prawda-o-wielkich-modelach-jezykowych>
- <https://pwr.edu.pl/uczelnia/aktualnosci/chatgpt-wywolal-rewolucje-jak-wiele-zmieni-w-naszym-zyciu-12824.html>
- Wykład prof. Wojciecha Broniowskiego  
<https://www.youtube.com/watch?v=Gn3vuzSSxLE>
- Wywiad z prof. Geoffreyem Hintonem  
<https://www.youtube.com/watch?v=qpoRO378qRY>



# Microsoft Copilot

- The Future of Work With AI - Microsoft March 2023 Event
- <https://www.youtube.com/watch?v=Bf-dbS9CcRU>
- Copilot łączy model językowy wyuczony na zbiorze danych historycznych z analizą bieżących danych użytkownika Office 365 (dokumenty, spotkania, maile)
  - Copilot był wcześniej znany w narzędziach rozwoju oprogramowania
  - Nowość polega na zastosowaniu w dowolnej dziedzinie z użyciem języka naturalnego



## Microsoft Copilot

- Zgodnie z zapowiedzią Copilot ma automatycznie realizować działania typu: „Stwórz prezentację opisująca zapotrzebowanie klienta X w ostatnim roku”
- Nadal nie oznacza to, że Copilot „rozumie” przetwarzane dane
- Potrafi je jednak wyekstrahować i połączyć używając modelu językowego, tak że ostatecznie z dość dużym prawdopodobieństwem wynik będzie użyteczny\*



## Microsoft Copilot

- \*Użyteczny – nie oznacza w tym przypadku „gotowy do wykorzystania”
  - Tak nie będzie jeszcze długo nie tylko ze względów technicznych, ale przede wszystkim prawno-etycznych (odpowiedzialność za skutki wywołane zastosowaniem SI), por. nagły zastój w dziedzinie pojazdów autonomicznych
- Owa użyteczność oznacza w tym przypadku, że prawdopodobnie będzie potrzebna weryfikacja i korekta treści, jednak jej zasadnicza struktura będzie już utworzona



# Microsoft Copilot

- Copilot ma zatem, według obietnicy promotorów tego rozwiązania, zwiększać naszą produktywność poprzez eliminowanie owej „mordęgi” – czyli pracy powtarzalnej, technicznej, nie wnoszącej niczego do problemu jako takiego
- Czy zmieni to świat? Prawie na pewno
- Czy zmieni na lepszy czy gorszy?
  - Co oznacza lepszy albo gorszy?



# Microsoft Copilot

- Czy świat przechodził już podobne metamorfozy?
- Ależ oczywiście – klasycznie w tym miejscu przytacza się tytułowe krosna
- Spójrzmy także na inne profesje – krawców, artystów, piekarzy (dlaczego akurat myślę o tych trzech?)
  - Czy w związku a uprzemysłowieniem tych dziedzin ubyło w nich miejsc pracy? Tak
  - Czy zatem nie ma już krawców, artystów, piekarzy?
  - Są – i są często elitą dziedziny



## Microsoft Copilot

- Coś, co jest łatwe do zdobycia, jest jednocześnie trudne do sprzedania – i odwrotnie
- Można się spodziewać, że nastąpi dewaluacja szablonowych treści generowanych przy znacznym udziale SI, a zarazem nobilitacja treści wytworzonych przy znacznym udziale wysiłku człowieka
  - Nie jest to tylko kwestia wartości finansowej – por. fragment dotyczący prezentacji na urodziny córki, który wzbudził uzasadnione kontrowersje w komentarzach
  - Czy wykształci się nowa forma snobizmu?



## Microsoft Copilot

- Podsumowując tę część można moim zdaniem stwierdzić, że rozwiązania takie, jak Copilot, nie przynoszą co do zasady ani szans, ani zagrożeń, których byśmy dotąd nie doświadczyli
- Nowością jest przeniesienie problemu z dziedzin „materialnych” na „intelektualne” oraz prawdopodobnie masowy dostęp





## Microsoft Copilot

- Czy „przez sztuczną inteligencję nie będzie pracy”?
- Natura nie znosi próżni, a zdolności adaptacyjne (uczenie się nowych umiejętności) człowieka są unikatowe
- Ktoś pamięta ojca Charliego w „Charlie i fabryka czekolady”?
- Wykorzystujemy swoją zdolność, a poradzimy sobie na pewno
  - chyba, że...



## Nowość czy nie? Retrospekcja

- Czy model językowy jako źródło informacji jest czymś AŻ TAK nowym?
- Trend polegający na upowszechnieniu dostępu do informacji trwa od setek lat...
  - Druk jako taki
  - Encyklopedie
  - Media elektroniczne
  - Wyszukiwarki



## Nowość czy nie? Retrospekcja

- Wszystkie te problemy, które [wy]generuje nam GPT i podobne mechanizmy, już znamy
- Nawet ten, którego intuicyjnie obawiamy się najbardziej: korzystanie z GPT daje możliwość bezrefleksyjnego kopiowania treści naśladujących wynik rozumowania / będących wynikiem czyjegoś rozumowania
  - Czy to jest NAPRAWDĘ nowe? A czy bryki już wyszły z użycia?
  - A rodzice?



## Co robić? Jak żyć?

- Jaki wpływ będzie mieć dostępność modeli językowych na proces nauczania / uczenia się?
- Na pewno nie możemy udawać, że „żaden”
- Zmiana roli nauczyciela
- Wartość tzw. prac domowych
- Ale... czy to naprawdę jest NOWOŚĆ?



## Dodatek - Uniwersalna Sztuczna Inteligencja

- Uniwersalność jest cechą, której brakuje obecnym modelom sztucznej inteligencji, aby zbliżyć się do inteligencji prezentowanej przez człowieka
- Wszystkie znane nam modele sztucznej inteligencji są specjalizowane – „potrafią jedno”
  - Przypomnijmy: GPT i Copilot są pozornie uniwersalne, ponieważ mogą „pracować” na treściach z dowolnej dziedziny...
  - ... jednak nie przypisują tym treściom znaczenia w żadnej dziedzinie



# Uniwersalna Sztuczna Inteligencja

- Uniwersalność umożliwia nam przenoszenie wiedzy, doświadczenia, umiejętności POMIĘDZY dziedzinami
- Zjawisko to pozwala nam twórczo rozwiązywać problemy, kształtować nowe byty...
- ... i nie zawsze zachodzi świadomie (por. pojęcia takie jak „natchnienie”, „intuicja”)



# Uniwersalna Sztuczna Inteligencja

- Czy jednak owa „nieświadomość” oznacza, że zachodzenie tego procesu będzie zawsze poza naszą zdolnością postrzegania?
- Jeśli nasz układ nerwowy ma charakter wyłącznie materialny, to jest deterministyczny – istnieje jednoznaczne wytłumaczenie każdego zachodzącego w nim zjawiska...
- ... a zatem istnieje algorytm pozwalający odtworzyć to zjawisko w modelu obliczeniowym
- Optymistyczne (???) szacunki mówią o osiągnięciu takiego stanu wiedzy w ciągu kilkudziesięciu lat



# Uniwersalna Sztuczna Inteligencja

- Jak będzie wyglądał świat, gdy algorytmy o zdolnościach do przetwarzania po wielokroć przekraczających zdolności biologicznych mózgów zostaną „wyposażone” w cechy takie jak:
  - Dążenie do efektywności
  - Samoświadomość (włącznie z tzw. „śmiertelnością”)
  - Ambicje
  - Emocje?





# Uniwersalna Sztuczna Inteligencja

- Zainteresowanych taką sferą rozważań zapraszam do:
- Intro to AI Safety, Remastered (oraz innych prezentacji Roberta Milesa)
- <https://www.youtube.com/watch?v=pYXy-A4siMw>
  
- CZYTANIA LEMA!



# KONIEC

Wytrwałym dziękuję za uwagę :)

