

# Sztuczna inteligencja: czym jest i w czym jej siła

Jacek Koronacki

Instytut Podstaw Informatyki PAN, 14 stycznia 2019

## SI - czym jest?

Nie ma jednej, powszechnie przyjętej definicji sztucznej inteligencji,

ale na pewno pomocne jest takie jej operacyjne określenie,

zaproponowane przez jednego z pionierów tej dyscypliny, Nilsa J. Nilssona:

„Sztuczna inteligencja jest efektem ludzkiego działania, zmierzającego do uczynienia maszyn inteligentnymi; zaś inteligencja jest tą jakością, która pozwala danemu bytowi działać właściwie i przewidująco w jego otoczeniu.”

## SI - czym jest, cd

Takie operacyjne i ogólne opisanie samej inteligencji pasuje tak samo dobrze do inteligentnej maszyny, jak i do człowieka.

Tyle że maszynę stwarza człowiek.

Chcąc ją uczynić inteligentną, człowiek wyposaża maszynę w pamięć, w której zapisana zostaje informacja z przeszłości

(np. wszystkie dotąd spisane dzieła oraz tysiące rozegranych w przeszłości gier w szachy czy w pochodzącą z Chin go)

oraz w czasie bieżącym zapisywane są kolejne informacje docierające z otoczenia do maszyny.

## SI - czym jest, cd

Maszyna musi być także wyposażona w algorytmy pozwalające przetwarzać posiadaną przez nią informację,

np. uczyć się grupowania obiektów podobnych do siebie, czy odróżniania jednej klasy obiektów od innych, uczyć się przetwarzania języka naturalnego, czy też uczyć się gry w szachy.

Maszyna powinna umieć doskonalić swe działanie na podstawie własnego doświadczenia.

## SI - czym jest, cd

Dzisiejsze maszyny są doskonale zdolne do pisania nowych powieści czy wierszy.

Zaczynają również być zdolne do modyfikowania wbudowanych w nie algorytmów uczenia się.

To ostatnie na razie chyba (ale to i tak bardzo dużo!) ogranicza się do uczenia się reprezentacji danych (jak w głębokim uczeniu się sieci neuronowych).

## SI - czym jest, cd

Wszystko to jest owocem pracy naukowców i inżynierów zajmujących się od dekad dyscypliną o nazwie

„maszynowe uczenie się” (dziś częściej mówi się o „uczeniu maszynowym”).

I nie ma w sobie nic tajemniczego to, że budowane przez naukowców i inżynierów maszyny wykonują takie same zadania jak człowiek, przy tym coraz częściej lepiej od człowieka.

Podobnie rzecz się już ma - lub wkrótce będzie miała - z humanoidalnymi robotami.

## „Tajemnicza” SI, czy „zwykłe” uczenie maszynowe, ale przydatne (jeśli mamy dane i wiedzę dziedzinową)

Dziennikarze i politycy, zaskoczeni wyborem na prezydenta Donalda Trumpa, rzucili się szukać jakiejś „Russia-gate” .

Bardzo niewielu zwróciło swoją uwagę na piękną pracę Cambridge Analytica (najpierw Michal Kosinski i jego The Big Five Model - OCEAN).

[Tytułem wtrętu kilka ciekawych nazwisk i nazw: Robert i Rebekah Mercerowie, Renaissance Technologies, Peter Thiel, Palantir Technologies, Steve Bannon, Breitbart, Cambridge Analytica.]

## ... przydatne, ale groźne

China Social Credit - wprowadzany od 2014 ma objąć cały naród do 2020;

8 korporacji ma licencję na budowę algorytmów dla CSC, w tym Rapid Finance z 850 mln użytkowników tej sieci społecznej, Sesame Credit („małe” odgałęzienie Alibaby), Ali Pay (chiński PayPal z 530 mln użytkowników).

[Tytułem wtrętu: rozkwita czarny rynek usług podnoszących mierzony poziom zaufania do obywatela.]



## Rewolucja technologiczna umożliwiła ...

... analizę wielkich i niejednorodnych zbiorów danych (jezior danych), które charakteryzuje tzw. HACE Theorem (2013):

- large volume Heterogeneous and Autonomous data sources with distributed and decentralized control, and these used to explore Complex and Evolving relationships among data.
- 3V Douga Lanely'ego (2001): Volume, Velocity, Variety, i z czasem dodawane Value, Variability, Veracity ...

[Tytułem wtrętu: amerykańska bankowość to: 70 / 35 / 17 mld transakcji na kartach debetowych / kredytowych / czekowych; do 2020 r. nasze zasoby wyniosą  $35 \times 1000^7$  bajtów (35 zetabajtów).]

## Rewolucja technologiczna umożliwiła także ...

... nauczenie komputerów współzawodnictwa z człowiekiem, vide np. takie programy jak:

- Dr Watson (IBM)
- AlphaGo (Deep Mind)
- mechaniczny Turek (Facebook; projekt „mikro”, ale ciekawy, ponieważ ...)
- GPT-2 (OpenAI; generatywny model języka)

... oraz rozmawiania z nim, jak np. chatboty pracujące w działach HR

... i także czyni nowe technologie wszechobecnymi za sprawą IoT.

## SI odczarowana

- Powtórzmy, SI w niczym dziś nie wykracza poza to, co daje zwykły, jakkolwiek imponująco szybki, postęp technologiczny.
- „Inteligencja” nie odwołuje się do niczego, co istniałoby jako takie - w sobie i z siebie. Niejako tym bardziej odnosi się to inteligencji „sztucznej”. Inteligencja jest względną miarą efektywności wykonywania danych zadań.

### **A co można powiedzieć o jej przyszłości?**

**W połączeniu z innymi działami IT, nano- i bio-technologią zbuduje człowieka-plus ...**

## Przyszłość SI, na jednym, ważkim przykładzie

- Transhumaniści marzą o cyber-nieśmiertelności.

Zakładają, że „ja” jest niczym innym, jak pewną ilością odpowiednio zorganizowanej informacji zdolnej do samomodyfikacji oraz modyfikacji pod wpływem bodźców zewnętrznych.

Na takie „ja” można spojrzeć jak na powieść, jakkolwiek modyfikowalną,

taką samą bez względu na to, czy jest zapisana na stronach papierowej książki, ma postać książki elektronicznej, czy też jest zapisana w pamięci komputera.

## Przyszłość SI, na jednym, ważkim przykładzie, cd

Gdy zatem nauka odpowie nam na pytanie o tę informację i sposób jej zorganizowania,

będziemy osiągniętą wiedzę mogli załadować do stosownie skonstruowanego superkomputera i tak otrzymać postczłowieka,

albo też przenieść świadomość człowieka do owego superkomputera, najlepiej umieszczając w nim dokładny model ludzkiego mózgu.

---

*Mamy tu do czynienia z dualizmem psychofizycznym,*

*zamkniętym w radykalnie naturalistyczno-materialistycznych ramach.*

## Przyszłość SI, na jednym, ważkim przykładzie, cd

*Pierwszym, który opowiedział się za dualizmem psychofizycznym,  
ale zarazem metafizycznym (ponad-naturalistycznym),*

*czyli radykalnym, substancjalnym rozdzieleniem ciała i umysłu  
(duszy),*

*był Kartezjusz.*

*[Tytułem wtrętu: W taki sposób odrzucił Kartezjusz realizm  
Arystotelesa i św. Tomasza z Akwinu.]*

*100 lat potem przyszedł dualizm natury i historii J.J. Rousseau,  
który otworzył drogę filozofiom końca historii.*

---

## Przyszłość SI, na jednym, ważkim przykładzie, cd

Transhumanizm chce przenieść człowieka (a przynajmniej jego świadomość) do cyberprzestrzeni.

Jest programem ucieczki przed naturą i śmiercią, i w tym sensie programem końca historii człowieka.

- **Powstaje pytanie:**

Czy cyber-nieśmiertelność dzięki załadowaniu naszych świadomości do maszyny jest programem spełnialnym?

## Przyszłość SI, na jednym, ważkim przykładzie, cd

- **Pytanie to jest równoważne jest takim pytaniom:**

Czy iluzją jest miłość – osoby do osoby oraz miłość prawdy, które są powiązane z naszą śmiertelnością, od której w jakimś sensie są uzależnione i która ich nie unieważnia?

Czy iluzją jest również samoświadomość, która czyni możliwymy praktykowanie cnót i doświadczanie specyficznie ludzkich radości, i która jest uzależniona od naszych podlegających zniszczeniu ciał?

- **W odpowiedzi możemy rzec:**



## Przyszłość SI, na jednym, ważkim przykładzie, cd

Filozofowie umysłu nie potrafią przekonująco umieścić go w naturalistyczno-materialistycznych ramach –

pewne fundamentalne pytania pomijają albo uznają wybrane przez nich odpowiedzi za obowiązujące a priori.

Odważniejsi, jak ateista Thomas Nagel albo Roger Scruton, przekonująco i precyzyjnie uzasadniają, iż

to co mentalne i to co fizyczne istnieje obiektywnie, ale jedno jest doskonale nieredukowalne do drugiego –

są to rzeczywistości różne, z tym co mentalne nie dającym się opisać przez współczesne nauki szczegółowe.

## Przyszłość SI, na jednym, ważkim przykładzie, cd

Ale jeśli tak, to marzenia transhumanistów są rojeniami i możemy za Peterem Augustinem Lawlerem powtórzyć, że

„Ludzka dusza

- a zatem owa niematerialna zasada fundująca nasze motywacje i aspiracje, które odróżniają nas od otaczającego świata przyrody -

przemienia i rozświecła wszystkie nasze myśli i działania,

również te, jak nie byłyby piękne,

które daremnie usiłują wyzwolić nas z natury i [poszukiwania sensu, a może] Boga.”

## Przyszłość SI, na jednym, ważkim przykładzie, cd

Zaś Leszek Kołakowski tak kończył swój esej „Horror metaphysicus”:

I czyż nie narzuca się podejrzenie,

że gdyby ‘być’ nie było ‘do czegoś’, a świat pozbawiony sensu,

to nie tylko nigdy nie zdołalibyśmy wyobrazić sobie, że jest inaczej,

lecz nawet pomyśleć nie bylibyśmy w stanie tego właśnie,

że być nie jest w istocie ‘do czegoś’, a świat pozbawiony sensu?

## Trochę szerzej o tym samym i parę lektur poważniejszych

JK, *Średniowiecze, Kartezjusz i sztuczna inteligencja*, ukaże się w: „Wszystko co Najważniejsze”, <https://wszystkoconajwazniejsze.pl/> ; wątki związane z historią filozofii są szerzej omówione w:

<https://teologiapolityczna.pl/prof-jacek-koronacki-o-koncu-historii-i-transhumanizmie/>

oraz <https://teologiapolityczna.pl/prof-jacek-koronacki-idee-maja-konsekwencje>

Peter Augustine Lawler, p. zwłaszcza: *Postmodernism Rightly Understood*, Rowman & Littlefield 1999, *Stuck with Virtue*, ISI Boks 2005, *Modern and American Dignity*, ISI Books 2010, *Restless Souls*, (artykuł w) *The New Atlantis* 2004

Thomas Nagel, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*, Oxford University Press 2012

Roger Scruton, *On Human Nature*, Princeton University Press

*Dziękuję za uwagę*