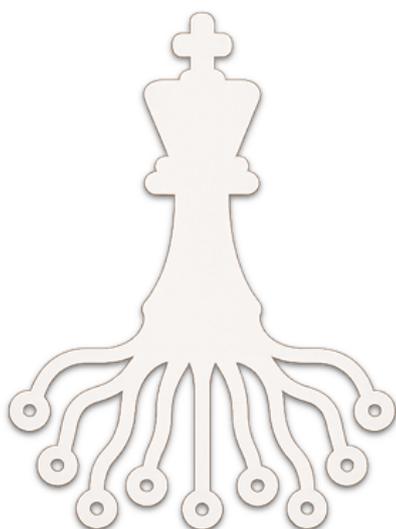
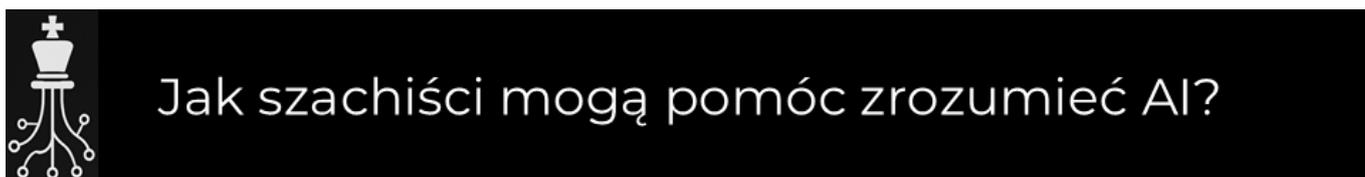


Chess XAI study

Weź udział w badaniu! Zostaw swój adres e-mail poprzez formularz. Badanie będzie **anonimowe** - wyślemy Ci odpowiednie instrukcje, gdy rozpoczniemy eksperyment.

<https://forms.gle/hEjpCTdZaka7sSsz8>

Dziękujemy!

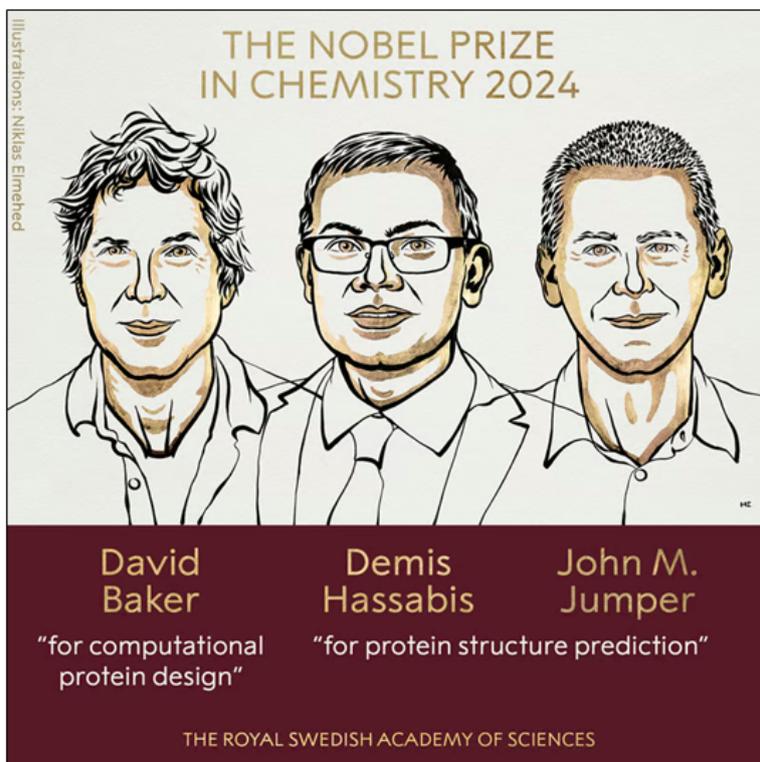


dr hab. Andrzej Siódmok, prof. UJ
Uniwersytet Jagielloński





Rok temu: AlphaFold święty Graal biologii



“Chess helped me win the Nobel Prize” D. Hassabis

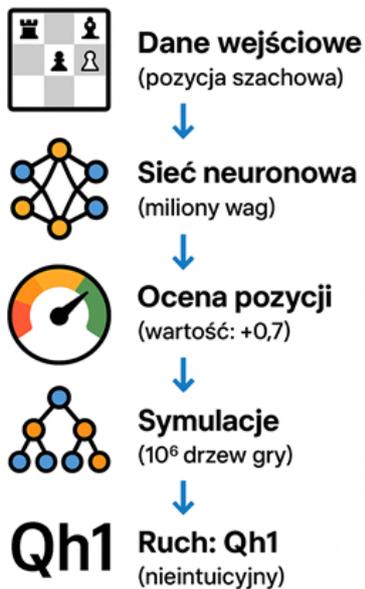


“so because of ... chess I started to think about thinking... and I started to improve my own thought processes... the computer play chess well I was intrigued and that got me into A.I.”



Problemy z interpretacją AI

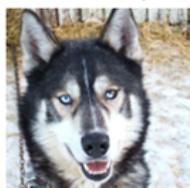
Jak AI „myśli”?



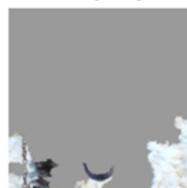
Nieintuicyjne cechy opisujące dane

```
relative_castling_rights
relative_piece_connectivity
wk_fgh_l2_bk_fgh_78
wk_not_l2_bk_de_78
wk_not_l2_bk_fgh_78
difference_king_safety_own_pieces
difference_castling_rights
.
```

Wnioski na podstawie błędnych przesłanek = ryzyko błędu



(a) Husky classified as wolf



(b) Explanation

Pytanie *Husky czy wilk?*

Odpowiedź *Husky*

Dlaczego? *Wykryto śnieg*



Co chcemy zrobić?

Jak objaśnić wnioskowanie?



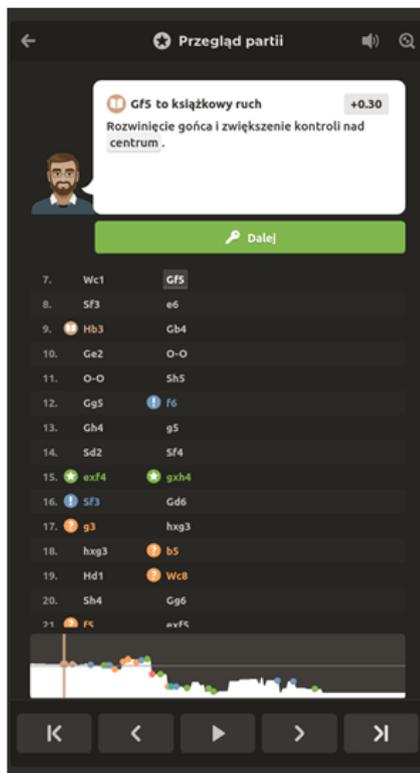
Czarne pudełko
(black box)
→ 10 mln wag → decyzja



XAI: algorytmiczne metody objaśnialności
→ Ruch Qh1: 0,4 za linię h



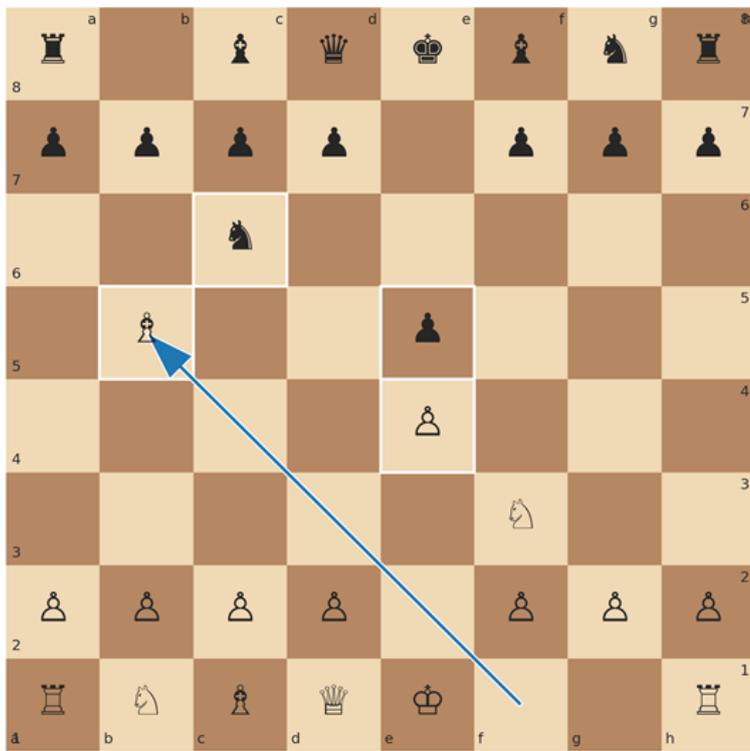
Tłumaczenie na język ludzi
→ AI widzi długoterminowy atak na króla





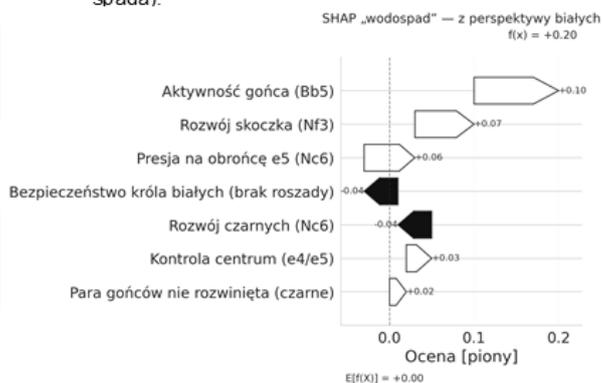
Krótko o eXplainable AI

Ruy Lopez (Hiszpańska) — po 3.Bb5



SHAP (Shapley Additive Explanations)

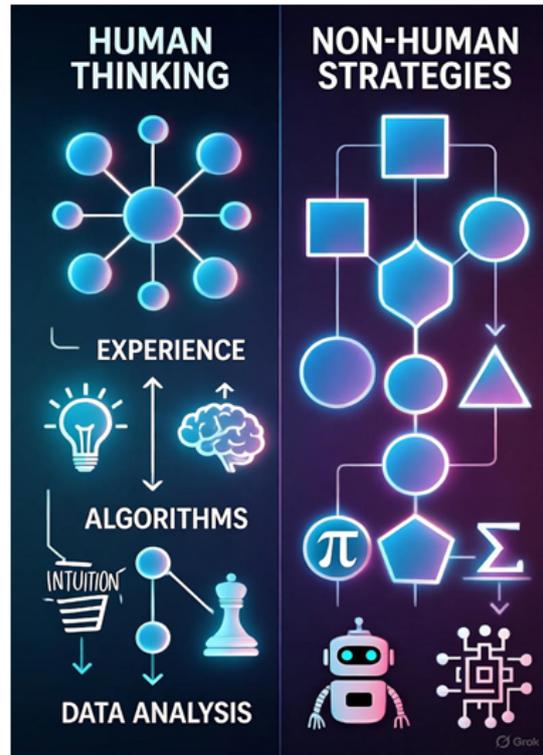
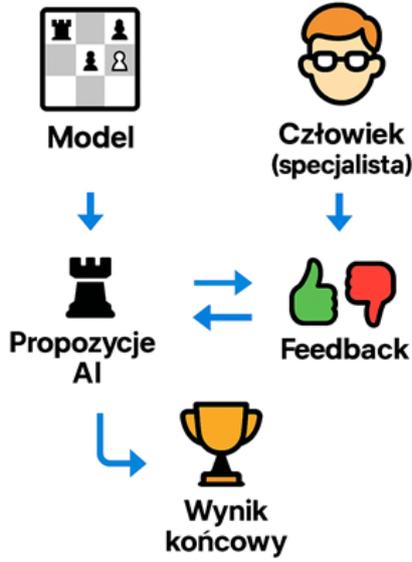
- O ile każda cecha podnosi lub obniża wynik modelu dla konkretnego przykładu.
- Wykorzystuje wartości **Shapleya** z teorii gier: „sprawiedliwie” dzieli wpływ między cechy, biorąc pod uwagę ich współdziałanie.
- Ma dwie kluczowe własności: **dokładność lokalną** (suma wkładów = wynik) i **spójność** (jeśli cecha pomaga bardziej w modelu, jej wkład nie spada).





Jak to osiągnąć?

HITL w szachach (Human-in-the-Loop)



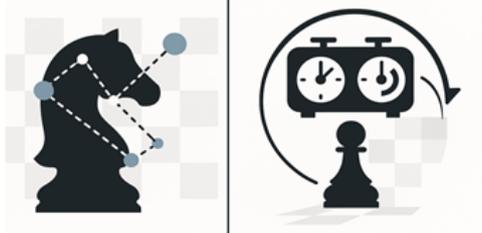


Dlaczego potrzebujemy właśnie Was

- **Ekspercka intuicja i plan:** jak formułować objaśnienia. "Ask an expert"



- **Domain knowledge:** które objaśnienia są użyteczne, co realnie „działa” przy zegarze i presji





Krótko o eXplainable AI

The screenshot shows a chess game in progress on a green and white board. The game is between GM Polish_fighter3000 (3171) and FM artin10862 (3073). The time is 2:03. An AI explanation overlay is visible on the right side of the board, highlighting a move. The overlay text reads: "Kh8 jest najlepsze To jest najsilniejsza opcja." with a score of -2.84. Below the text are buttons for "Pokaż kontynuację", "Spróbuj ponownie", and "Dalej". A list of moves is shown below the overlay, with the selected move Kh8 highlighted. At the bottom of the interface, there is a navigation bar with a play button and a timestamp of 0:45. The text "Nauka i szachy" is visible in the bottom left corner, and "8.11.2025" is in the bottom right corner.

Przeгляд партии

Kh8 jest najlepsze -2.84
To jest najsilniejsza opcja.

Pokaż kontynuację

Spróbuj ponownie Dalej

32.	Wf6	Kh8
33.	Sf3	Gh5
34.	Wd3	Hh6
35.	Kg2	Gxf4
36.	Hf2	Hg7+
37.	Kf1	Gg3
38.	Wxe7	Gxf2
39.	Wxg7	Kxg7
40.	Kxf2	Gxf3
41.	Kxf3	Wf6
42.	Kf4	Kg6
43.	Wg3+	Kf7
44.	Wd3	h6
45.	Wd4	Kg6
46.	r4	hvr4

Nauka i szachy

8.11.2025

Krótko o eXplainable AI

The screenshot shows a chess game in progress on a digital platform. The board is in a standard orientation with White on top and Black on bottom. The game is at move 21. An AI explanation overlay is visible on the right side of the board, highlighting a specific move. The explanation text reads: "Gf5 to książkowy ruch. Rozwinięcie gońca i zwiększenie kontroli nad centrum." Below this text is a green button labeled "Dalej". A list of moves is shown below the explanation, with the 15th move, "gxf4", highlighted in green. The interface includes player names and ratings (GM Polish_fighter3000 (3171) and FM artin10862 (3073)), a timer (4:51), and a "Przeгляд partii" (Game Review) button. At the bottom left, the text "nauka i szachy" is visible, and at the bottom right, the date "8.11.2025" is displayed.

GM Polish_fighter3000 (3171) 4:51

Przeгляд partii

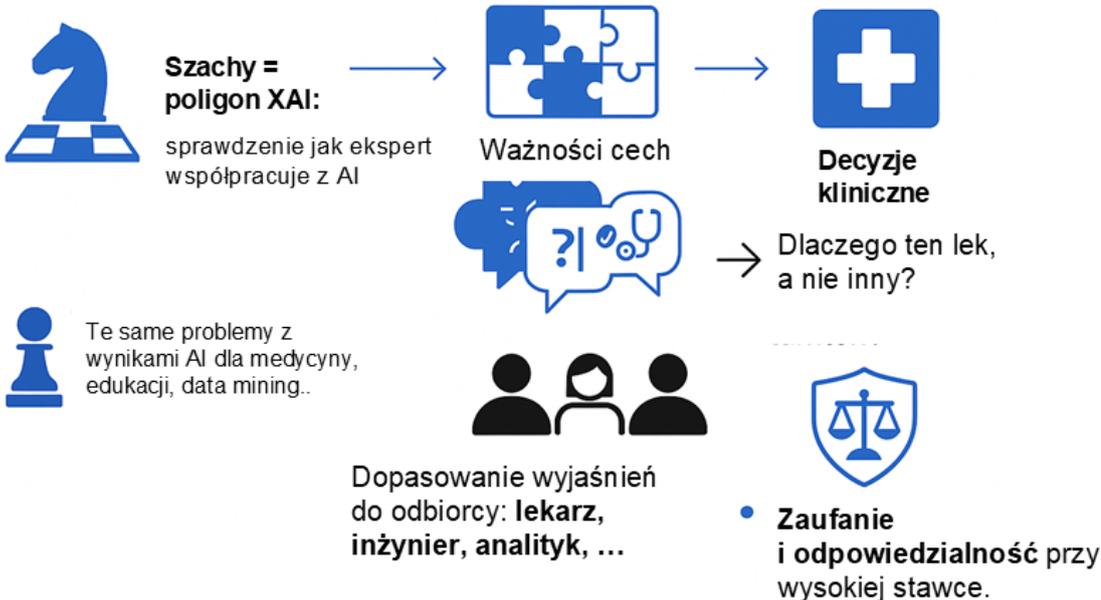
Gf5 to książkowy ruch +0.30
Rozwinięcie gońca i zwiększenie kontroli nad centrum.

Dalej

7. Wc1 Gf5
8. Sf3 e6
9. Hb3 Cb4
10. Ge2 O-O
11. O-O Sh5
12. Gg5 f6
13. Gh4 g5
14. Sd2 Sf4
15. exf4 gxf4
16. Sf3 Gd6
17. g3 hxg3
18. hxg3 b5
19. Hd1 Wc8
20. Sh4 Gg6
21. f4 exf4

nauka i szachy 4:44 8.11.2025

Po co to wszystko: od szachów do medycyny



From: <https://www.geist.re/> - **GEIST Research Group**

Permanent link: <https://www.geist.re/pub:projects:chessxai:start?rev=1762553996>

Last update: **2025/11/07 22:19**

