

Chess XAI study

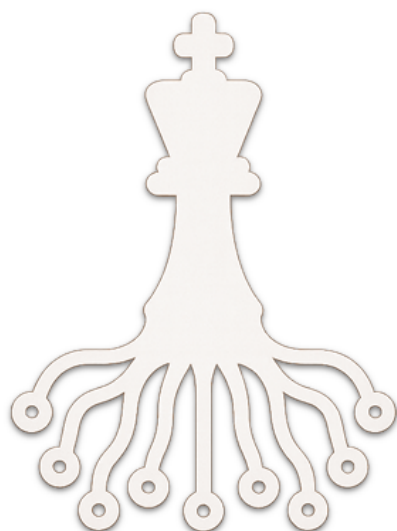
Weź udział w badaniu! Zostaw swój adres e-mail poprzez formularz. Badanie będzie **anonimowe** - wyślemy Ci odpowiednie instrukcje, gdy rozpoczniemy eksperyment.

<https://forms.gle/hEjpCTdZaka7sSsz8>

Dziękujemy!



Jak szachiści mogą pomóc zrozumieć AI?

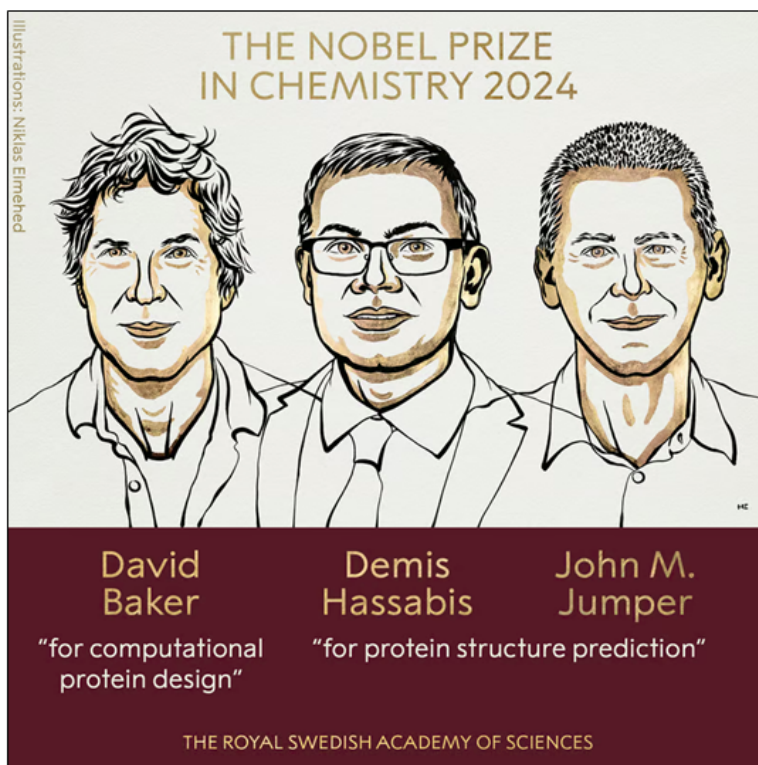


dr hab. Andrzej Siódmok, prof. UJ
Uniwersytet Jagielloński





Rok temu: AlphaFold święty Graal biologii



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

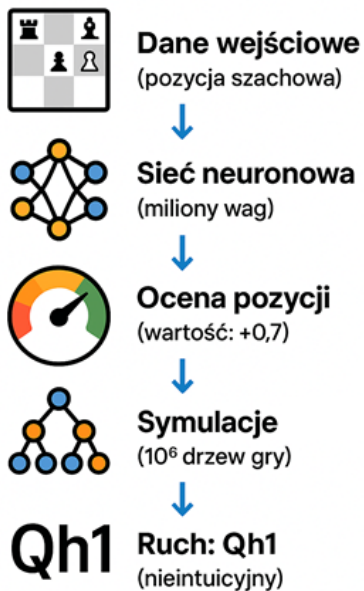


“so because of ... chess I started to think about thinking... and I started to improve my own thought processes... the computer play chess well I was intrigued and that got me into A.I.”



Problemy z interpretacją AI

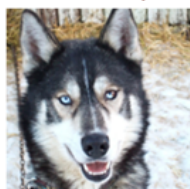
Jak AI „myśli”?



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

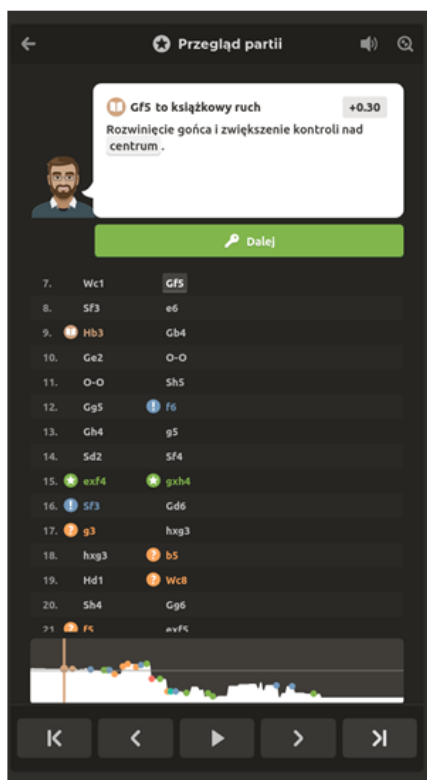
Pytanie *Husky czy wilk?*

Odpowiedź *Husky*

Dlaczego? *Wykryto śnieg*

Co chcemy zrobić?

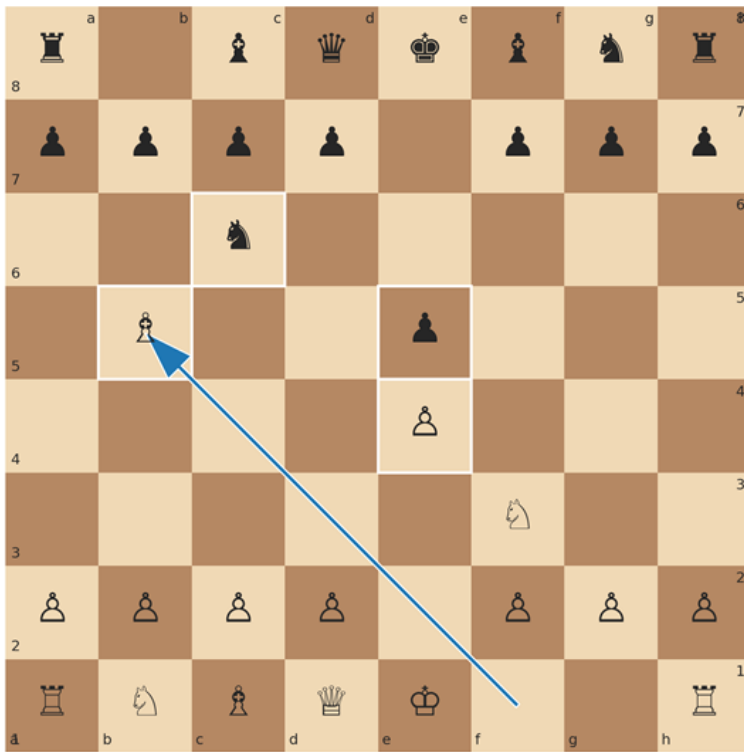
Jak objaśnić wnioskowanie?





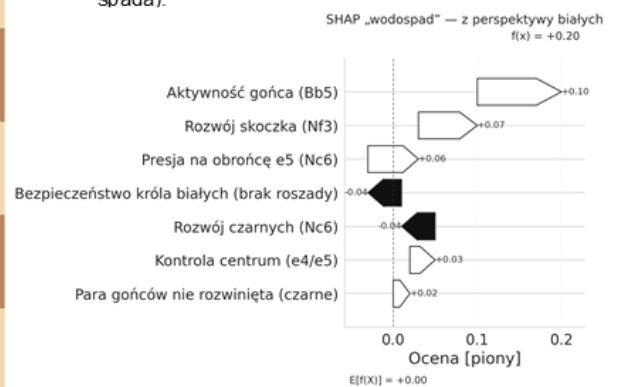
Krótko o eXplainable AI

Ruy Lopez (Hiszpańska) — po 3.Bb5



SHAP (Shapley Additive Explanations)

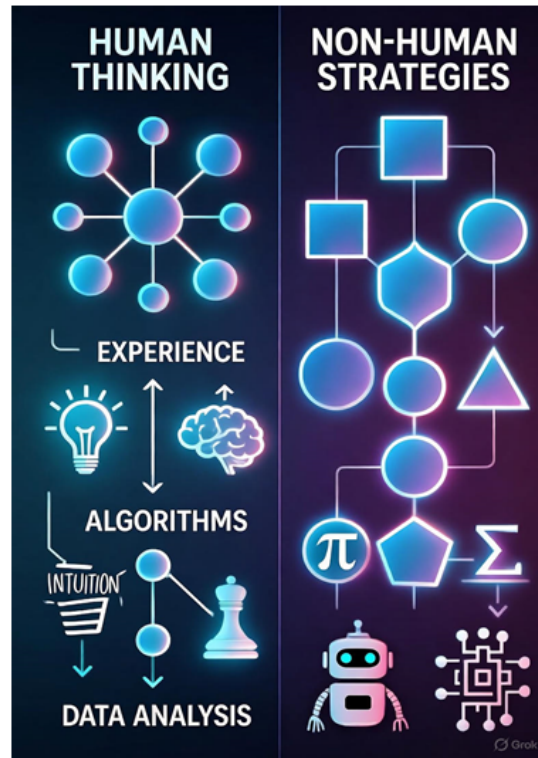
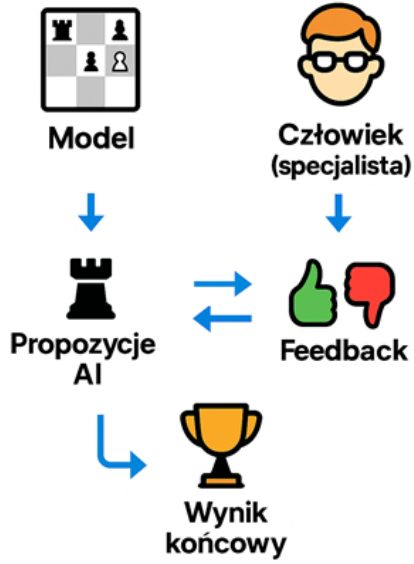
- O ile każda cecha podnosi lub obniża wynik modelu dla konkretnego przykładu.
- Wykorzystuje wartości **Shapleya** z teorii gier: „sprawiedliwie” dzieli wpływ między cechy, biorąc pod uwagę ich współdziałanie.
- Ma dwie kluczowe własności: **dokładność lokalną** (suma wkładów = wynik) i **spójność**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



The image shows a chess game interface with an AI analysis overlay. The main board is a standard chessboard with pieces in their starting positions. The AI analysis overlay on the right side of the board shows a list of moves and a highlighted move: **Kh8** is the best move, with a score of -2.84. The text says "Kh8 jest najlepsze" and "To jest najsilniejsza opcja." Below this, there are buttons for "Pokaż kontynuację", "Spróbuj ponownie", and "Dalej". A list of moves follows, with **Kh8** highlighted in green. At the bottom of the interface, there is a navigation bar with a play button and a date "8.11.2025".

Krótko o eXplainable AI

Przeгляд партии

Kh8 jest najlepsze -2.84
To jest najsilniejsza opcja.

Pokaż kontynuację

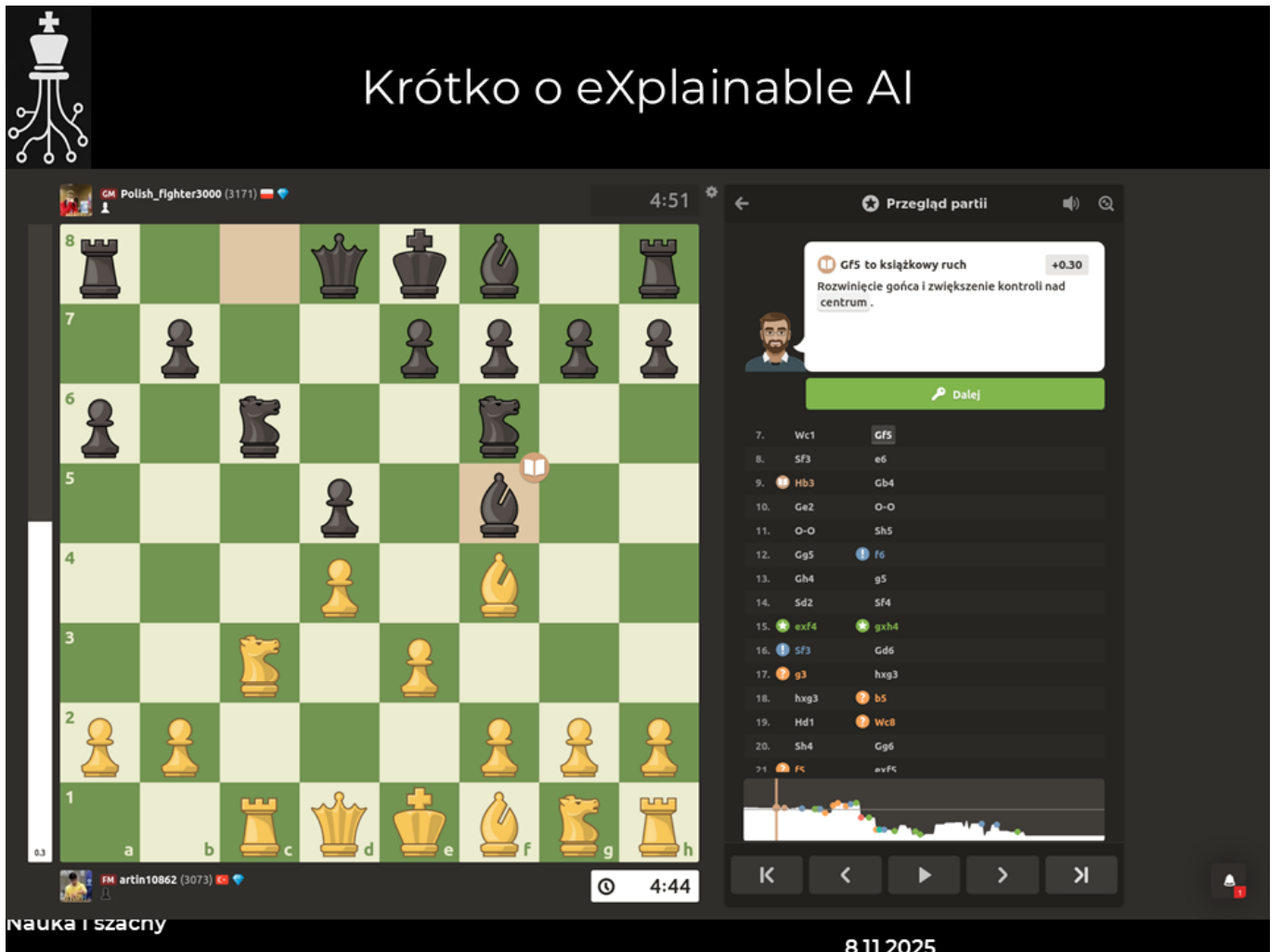
Spróbuj ponownie Dalej

- 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

0:45

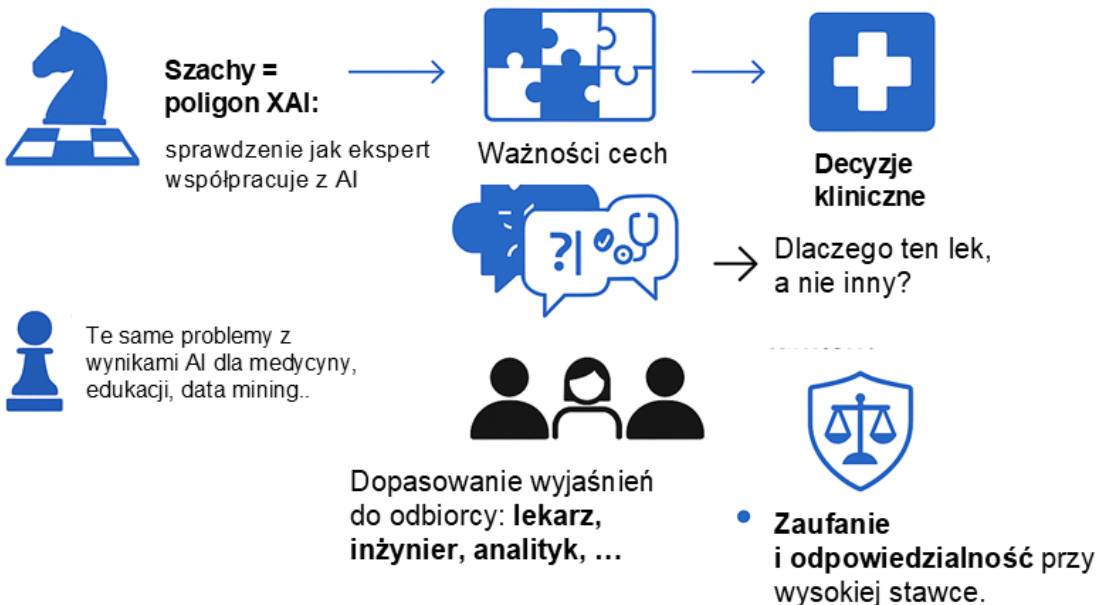
8.11.2025

Krótko o eXplainable AI



The screenshot shows a chess game in progress on a digital platform. The main board is a standard 8x8 grid with alternating light and dark squares. The pieces are arranged as follows: White (bottom): King on e1, Queen on d1, Rook on c1, Rook on g1, Knight on f1, Pawns on a2, b2, c2, d2, e2, f2, g2, h2. Black (top): King on e8, Queen on d8, Rook on c8, Rook on g8, Knight on f8, Pawns on a7, b7, c7, d7, e7, f7, g7, h7. A black knight is on f6, and a white bishop is on f5. A red book icon is next to the bishop on f5. The game timer shows 4:51. On the right, a panel titled 'Przegląd partii' (Game Review) displays an AI explanation for the move 'Gf5'. The explanation text is: 'Gf5 to książkowy ruch. Rozwiniecie gońca i zwiększenie kontroli nad centrum.' Below the text is a green button labeled 'Dalej'. A list of moves follows: 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 gxh4, 16. Sf3 Gd6, 17. g3 hxg3, 18. hxg3 b5, 19. Hd1 Wc8, 20. Sh4 Gg6, 21. f4 exf4. At the bottom left, the text 'nauka i szachy' is visible. At the bottom right, the date '8.11.2025' is displayed.

Po co to wszystko: od szachów do medycyny



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

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

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

