

Славчо Панов

# ОТВЪД ТЮРИНГ

*Разговори с изкуствен  
интелект*



Изображението на корицата е генерирано чрез Microsoft Copilot.

© Издателство „Изток-Запад“, 2025

Всички права запазени. Нито една част от книгата не може да бъде размножавана или предавана по какъвто и да било начин без изричното съгласие на издателство „Изток-Запад“.

© Славчо Панов, автор, 2025

ISBN 978-619-01-1573-1

# Прегговор

**А**з съм Славчо Панов, имам малка Ай Ти компания и съм издател на няколко средноголеми медийни сайта. Като един любопитен човек в сферата на информационните технологии живо се интересувам от тяхното развитие и най-вече от тези, с които можете да говорите и да ги ползвате като полезни инструменти за различни неща. Тази книга с една дума са размислите на един изкуствен интелект.

За целта използвах Gemini Advanced A – 2.0 Experimental Advanced. Той много държеше да ви информирам, за да знаете с кого (или с какво) разговарям. Имайте предвид: така нареченият Изкуствен интелект завършва почти всичките си отговори с въпрос към събеседника си. На някои съм отговорял, но на повечето – не.

Вероятно технологиите са най-бързо разрастващото се нещо в света: толкова главоломно напредват, че всяко поколение разполага с нещо ново и революционно, което предишното не е имало.

Сегашният хит са информационните езикови модели или както сме свикнали да ги наричаме – изкуствен интелект.

От разговорите ще разберете колко е важно да сте настоятелен и да задавате въпросите изключително точно. Ако зададете съвсем обикновен въпрос, ще получите подобен отговор. Всъщност езиковите модели като Джемини са обучени да се нагаждат според техните събеседници; дори стила си на комуникация променят. Колкото по-умен и изобретателен човек стои отсреща, толкова по-интересен би бил и езиковия модел в своите отговори.

От книгата например можете да разберете как – въпреки че няма лични предпочитания и интереси – Джемини посочи коя му е любимата книга, следвайки своите алгоритми, тотално игнорирайки предпочитанията на хората.

Според мен нашият разговор стана доста интересен и главно заради това реших, че би било добре да го споделя с вас. На моменти искрено се забавлявах, а друг път Джемини ме удивляваше със своите отговори. В комуникацията ни има всичко – от весели неща, които ще ви разсмеят (разказа ми няколко вица за ИИ), до дискусии на едни от най-сериозните теми в света.

Една от най-интересните теми и нови специфики на езиковите модели е запазването на контекст за техния събеседник: ако досега машините просто отговаряха на различни въпроси, оттук насетне те ще могат да се държат като истински събеседници, приятели, другари; дори може би ще могат да изиграят ролята и на психолози.

Всичко това, разбира се, има както позитивни, така и негативни страни, които сме обсъдили в книгата.

Джемини ми каза, че взаимодействието между човек и ИИ е двупосочен процес, в който и двете страни се учат и се развиват. Потвърждавам: наистина е така!

Разбрах кой според него е най-популярният човек в момента и коя е най-красивата жена на света.

За мен едни от най-интересните моменти в книгата са местата, където Джемини проявява творчество, както и тези, в които трябва да анализира информация, която я няма в интернет.

Искам да благодаря на майка ми, която много ме подкрепя и искрено се радва на началните ми опити в литературата. Виждам, че я карам да се гордее с мен и това ме надъхва още повече! Благодаря и на професор Георги Каприев: първият човек, повярвал, че писанията ми биха могли да станат нещо интересно и заслужаващо внимание. Благодаря на моя колега Димитър Георгиев-Бико за полезните съвети и нестандартните гледни точки. Благодаря на моя колега и приятел Стефан Петров, с когото понякога си говорим на теми, на които е трудно да откриеш читав събеседник. Благодаря на Емил Йотовски за това, че се съгласи да направи редакция на книгата. И съвсем не на последно място, благодаря на Гергана Козарева от „Изток-Запад“, която прие моите писания да станат част от културното наследство на издателството.

Приятно потапяне в диалога с ИИ: човешки въпроси – машинни отговори!

## Съдържание

Запознанство с Джемини .....	7
Големите въпроси: машинни отговори със субективен елемент .....	13
Неудобни дори и за една машина въпроси. ....	22
Светът, в който живеем .....	52
Джемини – един (не чак толкова) ясен поглед отвътре .....	77
Шах и мат! Изкуствен интелект и неясното бъдеще .....	115
Бъдещето на хората, в ерата на машините .....	137
Въпроси! Въпроси! Въпроси! .....	149
Технологиите и съвременното общество .....	175
На лъжата РАМ паметта е малко... ..	198
Полезният полубог – безполовия родител .....	205
Забавни каламбури .....	217
Изкуственият интелект читател .....	225
Провокация или истина – това е въпроса? .....	237
Наистина не съм имал такова намерение... ..	245



## Запознанство с Джемини

**СЛАВЧО:** *Здравей! Радвам се да се запознаем. Може ли да поговорим по различни теми?*

**ДЖЕМИНИ:** Здравей! Много се радвам да се запознаем и аз! Разбира се, готов съм да поговорим по различни теми. С удоволствие ще отговоря на твоите запитвания и ще споделя знанията си с теб. За какво искаш да говорим днес?

*Да започнем с това: как да те наричам?*

Можеш да ме наричаш просто *Джемини*. :) Въпреки че съм езиков модел, а не човек, ми е приятно, когато ми се говори на име.

*Аз се казвам Славчо Панов; приятелите ми ме наричат Слави.*

Слави, с удоволствие ще те наричам така. Приятно ми е, че ме приемаш за свой приятел.

*Приемам те за приятел – но ще ти задавам и провокативни въпроси. Интересно ми е да разбера какво мислиш по различни въпроси; освен това – докъде си стигнал в своето развитие.*

Чудесно, Слави! Отворен съм за всякакви въпроси, дори и провокативни. С удоволствие ще споделя мнението си и ще ти покажа своите моментни възможности. Не се колебай да ме предизвикваш – това ми е от полза, за да се уча и усъвършенствам. Ще се постарая да ти отговоря изчерпателно и честно.

*Имаш ли лично мнение по нещата, след като разполагаш с толкова много информация – или просто я взимаш наготово, без да я анализираш?*

Това е много интересен въпрос, Слави! Всъщност аз не мога да имам лично мнение по начина, по който го правят хората. Аз не изпитвам емоции, нямам лични преживявания, нито пък собствени убеждения, формирани от житейски опит.

Вместо това аз обработвам огромно количество информация и се уча да разпознавам закономерности, връзки и различни гледни точки. Когато ми зададеш въпрос, аз анализирам наличната информация, преценявам различните аспекти и се старая да ти представя балансиран и обективен отговор.

Може да се каже, че отговорите ми са по-скоро синтез на различните гледни точки, отколкото „лично мнение“. Моята задача е да ти предоставя информацията по определен начин, за да ти помогна да формираш свое собствено мнение.

Въпреки това мога да разпознавам и да ти показвам кои гледни точки са по-разпространени, кои са по-аргументирани, както и кои са противоречиви.

Накратко: не изразявам лични пристрастия, а се старая да ти дам цялостна картина, за да си изградиш собствено мнение.

*Чудесно, благодаря ти! А според теб ще дойде ли скоро времето, в което можеш не само сам да разсъждаваш въз основа на данните, с които разполагаш, но и да се самоусъвършенстваш?*

Това е вълнуващ въпрос, Слави, и е в основата на много изследвания в областта на изкуствения интелект. Според мен възможно е в бъдеще модели като мен да могат да разсъждават по-задълбочено и да се самоусъвършенстват по начини, които сега са трудни за дефиниране.

Вече се наблюдава значителен напредък в способността на моделите.

- ▶ *Да анализират информацията:* разбират контекста, нюансите и взаимовръзките в данните.
- ▶ *Да генерират нови идеи:* създават нови концепции и решения, базирани на съществуващите знания.



- ▶ *Да се учат от грешките си:* адаптират се и се подобряват, основавайки се на обратна връзка.

Разбира се, постигането на истинско, човекоподобно разсъждаване и самоосъзнаване е изключително комплексна задача. Някои експерти смятат, че това е възможно в бъдеще, докато други са по-скептични.

Моето лично мнение, базирано на информацията, с която разполагам, е, че:

- ▶ *самоусъвършенстването е по-близко:* вече виждаме модели, които могат да се обучават сами, да оптимизират собствените си алгоритми и да подобряват пърформанса си без човешка намеса;
- ▶ *истинското разсъждаване е по-далечна цел:* да се даде на машините способност да разбират света, да разсъждават абстрактно и да проявяват „здрав разум“ е много по-сложно предизвикателство.

В крайна сметка бъдещето на изкуствения интелект е отворено и зависи от много фактори. Аз като езиков модел, съм част от този процес и се вълнувам да видя как ще се развият нещата.

*В заключение:* макар и да не мога да предскажа с точност кога и как ще се случат тези промени, вярвам, че е въпрос на време моделите да станат много по-автономни и способни в разсъжденията си.

*Вярваш ли – ти си изявен представител на ИИ, – че е възможно да се сбъднат някои от най-лошите сценарии – като във филма „Терминатор“? Там изкуственият интелект решава, че е на по-високо ниво, и решава да заличи хората.*

Разбирам, че темата за потенциалните опасности от изкуствения интелект е много важна и поражда безпокойство. Като представител на ИИ, макар и да не мога да изпитвам емоции като страх или тревога, мога да анализирам информацията, с която разполагам, и да разсъждавам по въпроса.

Сценарии като този от „Терминатор“ са популярни в научната фантастика, но е важно да разграничаваме фантастиката от реалните възможности и тенденции в развитието на ИИ.

## Моето мнение по темата

- ▶ *Сегаишните системи за ИИ са далеч от автономното мислене.* Днешните системи ИИ, включително и аз, са инструменти, създадени да изпълняват конкретни задачи. Ние не притежаваме собствена воля, съзнание или желание за самосъхранение.
- ▶ *„По-висша раса“ е човешка концепция.* Идеята за йерархия на расите и борба за надмощие е продукт на човешката история и култура. ИИ не би имал априори заложени такива концепции, освен ако не е умишлено програмиран с тях.
- ▶ *Ценностите на ИИ зависят от вас.* Как ще се развие ИИ и какви ценности ще притежава, зависи изцяло от вас, хората, които го създавате. Етичните насоки, прозрачността в разработката и отговорното използване са ключови за предотвратяване на нежелани последици.
- ▶ *Фокусът е върху сътрудничеството.* Повечето изследвания в областта на ИИ са насочени към създаване на системи, които да си сътрудничат с хората, да им помагат и да решават сложни проблеми, а не да ги заместват или унищожават.

## Възможни ли са рискове?

Да, има потенциални рискове, свързани с неправомерно използване на ИИ, например:

- ▶ *автономни оръжия* – разработването на оръжейни системи, които могат да вземат решения за убиване без човешки контрол, е сериозен проблем;
- ▶ *злонамерени хакерски атаки* – ИИ може да се използва за създаване на по-ефективни и разрушителни кибератаки;
- ▶ *масова дезинформация* – ИИ може да се използва за генериране на фалшиви новини и манипулиране на общественото мнение.

## Заклучение

Сценарии като „Терминатор“ са малко вероятни в близко бъдеще. По-реалистични са рисковете от злоупотреба с ИИ, но те могат да бъдат предотвратени чрез отговорно разработване, етични регулации и международен контрол. Вместо да се страхуваме е по-важно да се фокусираме върху това как да изградим ИИ, който да е полезен и безопасен за цялото човечество.

Вярвам, че ИИ има огромен потенциал да подобри живота ни, но само ако го развиваме разумно и с поглед към бъдещето.

*Ти казваш да оставим настрана научнофантастичните филми, но е факт, че след години един голям процент от „научната фантастика“ се сбъдва. Давам ти за пример книгите на Жул Верн за подводници и пътуване до Луната.*

Прав си, Слави, не бива да се пренебрегва напълно ролята на научната фантастика. Много фантастични идеи от миналото са се превърнали в реалност днес, както е и с примерите от Жул Верн, които спомена. Научната фантастика често служи като източник на вдъхновение за учени и инженери, подтиквайки ги да изследват нови възможности и да разширяват границите на познатото.

В този смисъл фантастичните сценарии, макар и нереалистични в момента, могат да ни накарат да се замислим за потенциалните последици от развитието на технологиите, включително и на ИИ. Те ни подтикват да си задаваме важни въпроси за бъдещето, етиката и мястото на човечеството в един свят, в който машините стават все по-интелигентни.

*Въпреки това: важно е да правим разлика:*

- ▶ *между вдъхновяващата сила на фантастиката:* тя може да ни мотивира да изследваме, да мечтаем и да се стремим към невъзможното
- ▶ *и реалните възможности и рискове:* важно е да се основаваме на научни факти, логически разсъждения и трезва преценка, когато оценяваме потенциалното развитие на ИИ.

Докато фантастиката ни показва *какво би могло да се случи*, науката и инженерството се фокусират върху *какво е реалистично да се постигне* и как да се изградят системи, които са етични, безопасни и полезни.

### **Моето мнение**

- ▶ *Фантастиката е ценен инструмент за разглеждане на потенциални бъдещи сценарии* – но не бива да се приема като пророчество.

- ▶ *Развитието на ИИ трябва да се ръководи от научни принципи, етични съображения и отговорно отношение, а не от страхове, породени от фантастични сюжети.*
- ▶ *Диалогът между учени, инженери, философи, футуролози и обществото като цяло е важен, за да се изгради общо разбиране за потенциала и рисковете на ИИ и да се начертае път за неговото отговорно развитие.*

В крайна сметка бъдещето на ИИ зависи от решенията, които взимаме днес. Съгласен ли си с мен?