

## РЕЦЕНЗИЯ

върху дисертационен труд за придобиване на образователната и научна степен

„Доктор”

Автор на дисертационния труд: Андреас Четковски, докторант по направление 2.3.

Философия в СУ „Св. Климент Охридски”.

Тема на дисертационния труд: „Information and Entropy: Knowledge Structures in the Age of Artificial Intelligence“ („Информация и ентропия: структури на знанието в епохата на изкуствения интелект“)

Рецензент: доц. д-р Марина Иванова Бакалова

Андреас Четковски е предал дисертационен труд на английски език на тема „Information and Entropy: Knowledge Structures in the Age of Artificial Intelligence“ („Информация и ентропия: структури на знанието в епохата на изкуствения интелект“) в размер от 176 стандартни машинописни страници, както и автореферат на дисертацията.

Нека да започнем с няколко думи от професионалната автобиография на кандидата. В периода между 2003 и 2014 г. Четковски изучава немска филология и икономика в Техническия университет в Берлин, от където получава магистърска степен. Между 2014 г. и 2019 г., той е докторант в СУ „св. Климент Охридски“. В допълнение към академичните си занимания в този период, кандидатът преподава немски език и организира събития, свързани с немската култура. Четковски владее отлично английски, немски и полски език. Авторът има три публикации, посветени на херменевтиката на Дилтай и на автобиографичната памет в съвременните социални медии.

Дисертационният труд на Андреас Четковски е посветен на понятието за информация – фундаменталната съставка на всички форми на комуникация - и неговото детайлно историческо представяне. В основата на дискусиата е концепцията за информация на Клод Шанън, според който информацията е математическа мярка за едно и също съобщение, което може да се предава по различни начини или канали за свързка. Кодът за измерване на онова, което прави дадено съобщение едно и също се нарича ентропия. Един бит е най-малкото количество информация и то се свежда до отговор на въпроса „да“ или „не“. Ако има повече от две алтернативи, които могат да се случат, в системата има повече ентропия и съответно - повече информация. Идеята на Шанън е, че колкото и да е сложна една система за комуникация, информацията в нея винаги може

да бъде кодирана като избор между „да“ и „не“ или 0 и 1, заедно с определени вероятностни стойности да се случи едното или другото. Самата информация, според Шанън, се състои в изненадата или новото, което ни дава конкретния отговор.

Информацията е повсеместна, тя е в основата на всяко микро или макро действие – от ДНК и невронната комуникация до оралната и виртуална комуникация в нашето ежедневие. Намирайки начин да се компресират гигантски размери от информация, съвременната наука ни предоставя възможност за нечувана до този момент информативна ефективност. За един ден, ние сме в състояние да научим толкова неща, колкото един Средновековен човек – вероятно през целия си живот. Това до голяма степен се дължи на възможността за кодиране на информацията, открита от Шанън. Не на последно място - свеждането на информацията до 0 и 1 лежи в основата на изкуствения интелект (ИИ).

В своята дисертация, авторът се е опитал да постави под въпрос възможността за силен ИИ като отиде отвъд математическия анализ на понятието за информация и го сравни с онова, което се случва в субективното човешко съзнание при осмисляне на информацията. Авторът вижда двете като несъвместими.

В дисертацията си личи лингвистичния и икономически профил на Андреас Четковски и особено интересът му към етимологиите на отделни понятия и към събиране на изобилие от данни, които обаче не винаги следват някаква ясна мисловна нишка. Използвана е литература на английски и на немски език, като от текста ясно си личи, че авторът е направил доста мащабно проучване.

Дисертацията е разделена на три отделни глави.

Първата глава е посветена на концепцията за информацията и различните подходи към нея. Главата е изключително дълга и прекалено обзорна. Основната дихотомия в тази глава е между две най-общи понятия за информация: информацията като съобщение на някаква идея между подател и получател и информацията като статистически ново съобщение, което възниква от определени хаотични процеси като реализиране на конкретна вероятност. Предпочитаната от автора употреба на „информация“ е в термините на самоорганизация – вътрешен процес на системата, чрез който тя прави избор, за да превърне определена възможност в реалност. В глава 3, авторът ще се върне към тази дефиниция и ще я разгледа в хуманитарен контекст. Според него основната разлика между човешкия и изкуствения интелект се състои именно в начина, по който се осъществява въпросния избор, в единия случай чрез импровизирани правила, а в другия - чрез точно определен алгоритъм.

Интересен аспект от първа глава е опитът да се осмисли културната история посредством техническите формати за предаване на информация, характерни за отделните епохи. (глава 1, стр. 16)

Любимата ми част от тази глава е елегантният начин, по който авторът разграничава Просвещението от Ерата на информацията и изтъква връзката между информация и дезинформация. Дали ерата на информационните технологии е подобрене спрямо ерата на книгите – пита авторът на стр 18? Нетривиален е етическият проблем, който Четковски повдига по отношение на ИИ, а именно - дали възможността за компресиране и обменяне на толкова големи количества информация не ни отнема нещо? Много удачно авторът тръгва от мисълта на Сократ, че ако можеше да се записват мислите, това би намалило капацитета на човек да помни. Ако Сократ е бил прав - то в каква ли степен намаляват този капацитет съвременните информационни технологии? Мисля, че точно този аспект от взаимодействието на човека с компютъра не е съвсем тривиален и заслужава повече внимание.

Втората глава е посветена на ентропията като информационен безпорядък или шум. Ентропията е потенциална информация, чиято вероятност може да се измери количествено. В теорията на Шанън, високото ниво на хаос води потенциално до генерирането на повече информация. В тази глава авторът се опитва да покаже, че механизмите по които възниква информацията от хаоса според теория на информацията не са същите като тези, които използва човешкото субективно съзнание. Хората използват импровизирани схеми, за да осмислят реда в хаоса. Съответно, информацията зависи от наблюдателя – някой може да вижда ясен ред, а друг пълна бъркотия в определена ситуация. Според примера на Четковски, една чистачка, която влиза в работно помещение вижда бюрото в пълен безпорядък, докато работещият на това бюро, може да вижда всичко, което се намира там в точно определен ред. В този смисъл, както казва и Фред Дрекке: „Информацията, както и красотата е в окото на наблюдателя. (“Information, like beauty, is in the eye of the beholder.”) –1983 p. 55

Третата глава е посветена на разликата между начина, по който човешкото съзнание обработва и съхранява информация и начина, по който компютърът го прави. Тук основната идея е, че човекът променя и обогатява спомените си чрез въображението си, докато компютърът извлича съхраненото в паметта му едно към едно. Четковски изтъква, че субективната гледна точка, която предизвиква личностен избор е

невъзможно да бъде обективирана в компютър. Все пак, в последната главата отсъства дискусия, която обяснява защо това е принципно невъзможно.

### **Оценка:**

Кандидатът се е запознал с доста обширна литература. В резултат на това, в дисертацията има голям брой интересни идеи, които не са тривиални и до които човек не може да достигне ако не е изчел въпросната литература. За съжаление, обаче тези идеи са някак в насипно състояние.

Основната ми критика е, че дисертацията е много компилативна и основната идея е структурно недоразвита в нея, често няма логическа нишка между отделните абзаци. Поради тази причина, текстът се чете трудно, дори бих казала мъчително. Много компоненти в дисертацията са свързани със свободни асоциации и читателят трябва да гадае защо те се намират там. Например, не става ясно защо в трета глава се минава през детайлна експозиция на Лайбниц и Хайдегер за разума и връзката му с истината когато става дума просто за обработка на информация? Освен това се очаква още в началото на първа глава да има ясно представяне на теория на информацията на Шанън, от която да се види, че авторът наистина разбира тази теория.

Втората ми критическа бележка е следната. Тъй като докторантът свързва различни традиции, в които един и същи термин има несъвместими употреби – авторът трябва да покаже, че осъзнава разликата и въпреки това аргументите му са валидни. В увода тези разлики трябва да бъдат пояснени. Най-фрапиращ за мен е терминът „субективно знание“. Какво означава това? Ако нещо е знание, то трябва да е истинно, а всяка истина е обективна. Какво е субективната истина, какво я прави истина? Знанието не може да бъде субективно, мнението може. Може да имаме знание за вътрешни субективни състояния, но за да бъде знание - то трябва да е истинно. Имам чувството, че докторантът не прави разлика между „знание“ и „информация“ и не е ясно защо заглавието съдържа фразата „структури на знанието“. Може би има предвид „когнитивни структури“, но това е различно. Според мен, авторът категорично трябва да премахне терминът „знание“ от дисертацията си или да го замени с „когнитивен процес“. Знание и информация са две различни неща, а в дисертацията става дума единствено за информация.

По време на защитата очаквам кандидатът да отговори на следните въпроси:

1. Защо е *принципно* невъзможно субективното съзнание да се пресъздаде в компютър, например чрез оптимални, макар и все още не реални програми за мета-когнитивен мониторинг и програми за емоционална интелигентност? Кое е онова, което прави невъзможно субективния избор да бъде пресъздаден в ИИ?
2. Авторът говори за субективното състояние на човека, но в ИИ също има понятие за състояние, което се променя в зависимост от командите. Каква е разликата между двете?
3. Как ентропията определя че няма ясна предсказуемост на законите на природата и същевременно е базирана върху теория на вероятностите? (края на стр. 10)

В заключение бих искала да кажа, че адмиравам сериозните усилията и мотивация на Андреас Четковски в провеждането на това изследване. Имам сериозни проблеми с начина, по който е написан текстът обаче. Но предвид това, че докторантът се е потрудил достатъчно усърдно и е достигнал до някои интересни, оригинални и потенциално значими идеи, аз ще гласувам „за“ присъждане на научната и образователна степен „доктор“ на Андреас Четковски.

София, 2.10. 2024 г.

доц. д-р Марина Бакалова