

ЗОШТО НИ Е ПОТРЕБНА ПСИХОЛОГИЈА НА ВЕШТАЧКАТА ИНТЕЛИГЕНЦИЈА

Владимир Давчев¹

*Институти за родови студии, Филозофски факултет,
Универзитет „Св. Кирил и Методиј“ во Скопје*

Кратка содржина

Како никогаш досега, исправени сме пред предизвикот да креираме вештачка интелигентна машина која ќе мисли, учи и стекнува искуство, а да функционира автономно, без надзор на човекот. Но, како што проблемите, коишто ние луѓето им ги предаваме на машините да ги решаваат, стануваат сè посложени, така овој предизвик станува сè поинтригантен за научниците од повеќе академски дисциплини. Доколку имаме намера вештачката интелигенција ја креираме како „ентитет“ кој треба да создаде подобар, полесен и поквалитетен живот, тогаш од суштинско значење е да се употреби поинаков природ и поинаква методологија која е присутна во другите научни области, а не само онаа на компјутерските науки. Иако не постои универзална дефиниција за интелигенција, постојат повеќе модели и концепти, од кои повеќето меѓусебно се доближуваат до сфаќањето дека интелигенцијата ја вклучува можноста за размислување, рационализирање, решавање на проблеми и расудување. Сите тие се главно психолошки концепти. Со целосно автономно однесување, учење, размислување на еден вештачки интелигентен систем (вештачки мозок), доаѓа потребата тој систем да биде конструиран (софтверски и хардверски) така што ќе биде способен не само да изведува мимички процеси и потпроцеси кои постојат во човековиот мозок, вклучувајќи ги интуитивните и емоционалните концепти, туку и самостојно да ги разбира и создава. Целта на овој труд е да воспостави некои основни принципи и концепти кои би биле столб на автономна вештачка интелигенција, а која не е можна без психологија на вештачката интелигенција.

Клучни зборови: *интелигенција, вештачка интелигенција, психологија на вештачка интелигенција, вештачки ум.*

¹ vladimir@fzf.ukim.edu.mk

Интелигентното однесување, главно, му се припишува на човечкото суштество. Голем број истражувачи работеле, работат и ќе работат на проучување на интелигенцијата и интелигентното однесување и може да се каже дека, барем во последниве стотина години, дојде до некакво усогласување на ставовите околу дефинирањето на интелигенцијата и интелигентното однесување на човекот. Кога велиме некакво усогласувањето на ставовите, мислиме на тоа дека во академските кругови преовладува ставот дека интелигенцијата е ментална особина која ја сочинуваат неколку различни способности како: учење од искуство, прилагодување кон нови ситуации и нивно разбирање, способност во новите ситуации да се применуваат претходно стекнатите знаења (а снаоѓањето во нив да биде брзо и ефикасно), способност да се разберат и да се користат апстрактни концепти и идеи, односно апстрактно размислување и способност за брзо и лесно да се научат и надминат дадените проблеми.

Во ред, научниците (психолозите, во прв ред) можеби се согласиле околу дефинирањето на интелигенцијата, но како ќе одредиме дали и колку некој е интелигентен? Современата наука, во прв ред психологијата, креирала низа инструменти со кои оваа ментална особина може да се измери. Преку соодветни тестови, се добива IQ - количник или квоциент на интелигенцијата. За да бидат релевантни резултатите, потребно е тестовите да се прилагодат на одредена возраст и популација, а пред сè да бидат стандардизирани, што во голема мера не е случај со бројни тестови, кои се присутни на интернет и кои откако ќе ги решиме, нè убедуваат дека сме надарени со натпросечни способности (натпросечно интелигенти). Количникот или квоциентот на интелигенција се пресметува по формула: $IQ = \frac{MB}{KB} * 100$ (ментална возраст) / KB (календарска возраст) * 100, концепт што прв го предложи германскиот психолог Вилијам Штерн (William Stern) и го усвои Луис Терман (Lewis Terman) во скалата на Стенфорд-Бинет (Stanford Binet). IQ првично бил пресметуван како сооднос на менталната возраст на една личност со неговата хронолошка (физичка) возраст, помножена со 100. Така, ако дете од 10 години имало ментална возраст од 12 (т.е. тестот е изведен на ниво на просечен 12-годишник), тогаш на детето му бил доделен количник на интелигенција $(12/10) * 100$, или 120. И покрај тоа што многу тестови сè уште определуваат IQ според претходно посочената формула, во современата психологија оваа бројка сега се пресметува врз основа на статистичкиот процент на луѓе за кои се очекува да имаат одредено IQ.

Во тој контекст може да се запрашаме, дали може да се зборува за интелигенцијата како за општа способност, која секој од нас, во одреден, помал

или поголем степен ја поседува или може да се расчлени на повеќе форми на изразување?

Американскиот психолог Гарднер (Gardner, 1983) изложува теорија за повеќе видови интелигенција. Тој ја дефинира интелигенција како „способност за решавање на проблемите и создавање на ствари кои се ценети во една или во повеќе културни рамки“ (Gardner, 1983, p.x) и го застапува мислењето дека секој поединец има свои недостатоци и предности, врз основа на кои се формираат димензиите на повеќекратната интелигенција. Со други зборови - поседување понизок количник на интелигенцијата не мора да значи дека не сме доволно интелигентни, туку дека нашата интелигенција можеби лежи во некои други способности, кои тестовите за мерење на количникот на интелигенција не ги земаат предвид. Според неговата теорија, денес можеме да зборуваме за девет видови на оваа ментална особина кај човекот: вербално лингвистичка или јазична интелигенција, логичко-математичка интелигенција, визуелно-просторна интелигенција, телесно-кинетичка интелигенција, музичка интелигенција, интерперсонална интелигенција, интраперсонална интелигенција, природно-научна интелигенција, егзистенцијалистичка или духовна интелигенција.

И покрај големиот број трудови и теории за човековата интелигенцијата, таа и натаму останува мистерија, па затоа нималку не е едноставно да кажеме дека нешто друго, особено машина, поседува интелигенција. Значителниот напредок на информатичката технологија и обработката на податоци, интернет на стварите, проширената реалност и машинското (длабинското учење) го насочи фокусот и интересот на науката и пошироката јавност во правец на дискусија за интелигентни машини и вештачка интелигенција (во натамошниот текст ВИ). Всушност, вештачката интелигенција претставува еден вид крстопат на повеќе научни полиња и области - психологија, филозофија, антропологија и компјутерски науки.

Но ако човечката интелигенција ја определуваме преку тестови, како што веќе кажавме, зошто тогаш да не ја тестираме и машината за која претпоставуваме дека е интелигентна. На оваа идеја, во 50-тите години на 20-тиот век, дошол и познатиот англиски математичар Туринг (Turing, 1950). Но наместо да ја тестира машината (програмот) дали и колку е интелигентна, тој смислил тест во кој човек, преку комуникација на слепо, ќе се обиде да погоди, дали комуницира со компјутер или со вистинско човечко суштество. До колку човекот не може да направи разлика дали комуницира со човек или со машина, тогаш машината (програмата) го минала тестот.

Прашањата за тоа дали и како треба да се креира ВИ во однос на човечката когнитивна психологија е постара од самите термини вештачка интелигенција и когнитивна психологија. Туринг (Turing, 1950) претпоставувал дека е полесно да се гради и едуцира „дете-машина“, отколку да се направи обид целосно се сфати како настанува когницијата кај возрасен човек. Тој го замислува умот на детето како тетратка со „прилично мал механизам и многу празни листови“, а умот на „дете-машината“ како процес на пополнување на празни листови во тетратка во вид на засилено учење (*reinforcement learning*). Ова гледиште за когницијата и учењето е еден вид ехо на бихејвиоризмот, доминантна психолошка традиција во времето на Туринг. Истовремено, таа го изразува и силниот емпиризам кој е изразен во современите модели на конективизмот - идеја дека можеме да научиме скоро сè што знаеме од статистичките модели на сензорните влезни информации.

За жал, овој тест не помогна многу во креирањето на вистинска ВИ. Креаторите на алгоритмите на ВИ, само се натпреваруваат во тоа кој ќе направи подобар алгоритам, што ќе го помине тестот и преку кој човекот нема да може да препознае дали станува збор за машина или за вештачки интелигентен систем. Но тоа е само еден вид психолошка игра во интерес на комерцијализацијата на ВИ.

Идејата дека ВИ може да се создаде така што би ѝ дале за задача да ја научи Енциклопедија Британика ѝ би ѝ рекле да го научи тоа совршено, и дека тоа е сè што некогаш ќе ѝ треба, можеби некогаш и се вртела во главата на некои научници, но тоа никогаш не би успеало. На човечките суштества им е потребно долгогодишно искуство во постојано променливата околината, пред да можеме да го разбереме светот околу нас. Истражувањата покажуваат (Lake, Ullman, Tenenbaum, & Gershman, 2017; Winerman, 2018) дека, кога ние се обидуваме да се фокусираме на некој предмет, нашите мозоци не се целосно способни да ги филтрираат визуелните дистрактори сè до 17-годишна возраст, а нашата способност да согледуваме лица продолжува да се развива до 20-тата година од животот. Ние сме „подготвени“ да учиме, благодарение на 3,5 милијарди години искуството на суштества кои живееле пред нас и кои секојдневно се соочувале со живот и смрт и ситуации во кои морале да согледаат и да постапуваат правилно, за да преживеат. ВИ нема ништо од ова. Алгоритмите што сме ги креирале за ВИ користат многу поедноставени идеи за учење и претежно прават малку повеќе од класификација на податоци или предвидување врз основа на тие податоци.

Единствениот начин на кој можеме да ги направиме нашите (човечки креирани) алгоритми за поефикасна работа е да ги обучиме за големи коли-

чини на тесно фокусирани збирки на податоци со дефинирани цели. Тие сè уште не се во состојба да управуваат со менување на „сценарија“ на начин каков што можеме ние луѓето. ВИ не е способна да ги „разбере“ причината и ефектот. Без оглед колку комплексен алгоритам сме креирале во досегашниот развој на ВИ, таа сè уште не е во можност правилно да ги поврзе поимите, како што се „стол“ или „возило“ со реалните физички предмети, затоа што ВИ-системите никогаш не ја доживуваат реалноста како ние луѓето (Crowder, Carbone, Friess, 2013). И додека некои ВИ-системи можат да ги класифицираат емоциите со обработка на слики од лица со експресија, истражувањата сè уште се во почетна фаза како да ја научиме ВИ да може да „чувствува“, да „сочувствува“ (да развие емпатија) или како што нагласува авторот Кирван (Kirwan, 2016) да „разбере“ како нејзиното однесување може да влијае врз нас.

Со време, ВИ можеби ќе ја приближиме до нас. Можеби ќе можеме да ѝ „помогнеме“ да размислува повеќе како нас, развивајќи алгоритми кои учат и обработуваат информации на нови начини. Но, јазот меѓу нас и ВИ веројатно ќе биде секогаш присутен. ВИ не ја дели нашата еволутивна историја и можеби нема да има мозок толку комплексен или фино подесен како нашиот.

Што може да направи ВИ денес?

Вештачката интелигенција продолжува да станува попаметна: може да нè надигра во игри, може да препознае и подреди слики подобро од човекот и да управува со автомобил. Но, таа не може да ја имитира човечката мисла. Компјутерските модели кои се засновани врз имитирање на нервните мрежи во нашите мозоци, снабдени се со огромни количини на податоци користејќи огромни кластери на процесори. Тие сега може да ги класифицираат предметите или сликите на предметите подобро отколку што можеме ние луѓето. И, како што покажаа истражувањата на ИБМ (IBM) поврзани со ВИ и проектот „DeepMind“ на Гугл (Google), тие може да нè надиграат во игрите како шах или Го (Go), па дури и да постигнат највисок број освоени бодови во компјутерската игра СтарКрафт 2 (StarCraft 2). Но, ВИ-системите не го разбираат нашиот свет или нивното место во него. Тие остануваат само супериорни машини во исполнување на една или мал број задачи. Повеќето ВИ-системи во денешно време се однесуваат на специфични активности и нивната активност е исклучително тесна. Некои ВИ-системи можат автономно да возат автомобили, некои може да го предвидат нашето следното online купување, некои може да сортираат зелени и црвени јаболка, а некои може да препознаат лик на сликата што сте ја направиле. Ова е т.н.

„слаба ВИ“ и е единствената форма на ВИ што ја постигнавме досега - специјализирана само за еден домен, иако таа во голема мера ги надминува луѓето во зададените задачи. Но, целта на креирање вештачки интелигентен систем не е самата машина, која формира сосема посебно инженерско поле, туку комуникација, координација и споделена интелигенција со сите субјекти во реалниот свет, преку ефикасна соработка како главни точки на фокус. Тоа е целта на „вистинската“ ВИ.

Човечкиот род ќе мора да се справи со постојаниот товар на сопствената ограничена рационалност и мора да си признаеме дека никој од нас нема да може да направи чисто рационален избор во која било ситуација, како што тоа го очекуваме од ВИ. Повеќе од кога било, ќе се соочиме и ќе мора да прифатиме дека со нашиот ограничен ментален капацитет ќе треба да работиме со огромен број информации и за сè пократко време да донесеме некоја одлука. Меѓу другото, ова е едно од можните образложенија зошто толку многу инвестираме во технологија и во развојот на ВИ. Тоа ќе ја закрепне нашата постојана болка дека не сме совршени суштества (што и да подразбираме под тоа). Вештачката интелигенција претставува логички одговор и последица на предизвикот на реалноста што ги преоптоварува нашите сетила и мозокот со огромен број на информации што треба да се „обработат“ во секојдневниот живот. Таа е нашиот „излез“ (барем така претпоставуваме) и „алатка“ која треба да се справи со сложените проблеми произведени од огромната количина на информации. Таа е изградена како проширување на менталните способности на луѓето, асистент за вршење непријатна/ни работа/и и функционира како дополнителна работна сила, во овој случај машина, со цел да се извршат повеќе задачи одеднаш и да се донесат побрзи одлуки.

Се разбира, ВИ не е решение за сè и крај на секое решение. Така на пример, големиот план на Елон Маск (Elon Musk) за фабрика за производство на работи, во кои заедно би работеле работи и луѓе, драматично пропадна бидејќи ја потцени флексибилноста на човековата работа. Неговата првична идеја воведо многу одложувања, високи трошоци и речиси ја банкротираше компанијата. Главниот проблем искрснал поради тоа што алгоритмот за вештачка интелигенција што ги контролирал роботите, воопшто не бил толку флексибилен како што се луѓето, пред сè во смисла на интеракција со лабави и свитливи делови во одредени сегменти на производствениот процес.

Но, предизвикот допрва претстои бидејќи ВИ со поголем темпо се имплементира во сè поголем број уреди кои се ползуваат во секојдневниот живот, а тоа ќе бара не само препознавање на говор, туку и разбирање на природниот јазик, контекстот и расудувањето. Тоа е токму доменот на истражу-

вањето на вештачката интелигенција и невозможно е да се очекува само софтвер инженерите да го решат овој проблем и да го имплементираат преку алгоритми. Потребно е вклучување на експерти од повеќе области (психологија, социологија, филозофија...). Во тој контекст Бил Марк (Bill Mark), претседател на Одделот за информации и компјутерски услуги на Мајкрософт, чиј тим ја измисли и креираше Сири (апликација за оперативниот систем iOS) посочува дека овој „дигитален личен асистент“, кој се базира врз значајни ресурси од областа на вештачка интелигенција, како што се обработка на природниот јазик и препознавање на говорот, смета дека сега е моментот ВИ да премине на следно ниво - препознавање на човековите емоции. Тука се отвора прашањето колку воопшто знаеме за емоциите на човекот и дали сме способни да креираме компјутери (машини) што навистина ќе разберат емоција. „Некои луѓе плачат кога се среќни, многу луѓе се смеат кога се фрустрирани. Значи, еден многу симплистички пристап, како размислувањето дека ако некој се насмевне тогаш е среќен, нема да функционира во креирање на вештачки систем кој би разбирал што е емоција“ (Winarski, 2015, p. 40).

Ник Бостром (Nick Bostrom) и Елиезер Јудковски (Eliezer Yudkowsky) го застапуваат ставот дека интелигентните машини мора да бидат програмирани со разбирливи правила, ако сакаме да ги толерираме меѓу нас. Доколку, „однесувањето“ на машините е базирано на донесување одлуки кои се темелат врз алгоритмите на машински јазик, не е јасно како таквата недвосмислена транспарентност ќе биде достапна. Секогаш ќе има и алтернативни објаснувања за какво било однесување на човекот и тоа постојано се случува во реалноста. Во случај на интелигентни машини, нивната „состојба“ ќе може (некои велат и ќе мора) да се смета за иста како човечката, освен ако не се додаде нешто сосема ново, што ќе биде исклучиво поврзано со нив, што уште повеќе ќе ја искомпликува работата (Bostrom & Yudkowsky, 2014).

Психологијата на вештачката интелигенција како добра основа за вистинска ВИ

Психологијата го истражува умот, животот и однесувањето на луѓето. Како академска дисциплина има огромен обем на истражување, на пример, од когнитивна психологија до социјална психологија или од клиничка психологија до организациска психологија. Психологијата е особено погодна за креирање на интердисциплинарни концепти, на пример, оние коишто се однесуваат на човекот, наспроти социологијата или политиката, кои креираат концепти на системско ниво. Впрочем, етичките и биолошките концепти се

многу интегрирани во психологијата. Ова е една од причините зошто старото разграничување помеѓу дисциплините не може да биде можно и корисно во секој случај и зошто е потребно поактивно вклучување на психологијата во активностите поврзани со ВИ.

Кога станува збор за вештачката интелигенција, едно е јасно: не можете да го „извадите“ човекот од вештачка интелигенција, без разлика дали тоа е интеракција помеѓу човекот-агент, перцепција, јазик, когнитивни процеси или меки вештини како што се емпатија, емоции или комуникациски вештини. Секогаш е вклучена човечка страна - без разлика дали ја пишува програмата, уредува податоци или комуницира со системот. Засега, не постои психологија на вештачката интелигенција или „артифициелна психологија“ (artificial psychology), термин предложен од Ден Кртис (Dan Curtis) во 1963 како теоретска дисциплина на психологијата, што подразбира дека таков вид психологија би се занимавал со вештачка интелигенција која има свој ум, па дури и сопствена свест да донесува одлуки без никаква човечка интеракција или алгоритми напишани (креирани) од страна на човекот. Сепак, напредокот во имитирање на човековото однесување и различните ментални процеси е неверојатен и може да се очекува многу повеќе, бидејќи овој пат не е кор-со-как, и може да се очекува дека многу други доаѓаат за да се придружат на патувањето по тој пат.

Со децении, вештачката интелигенција претставува голема мистерија. Но нејзиниот развој ѝ го препуштивме на информатичката технологија и компјутерските науки. Впрочем, вештачката интелигенција ја дефинираме како „област на компјутерски науки што работи на создавање интелигентни машини кои работат и реагираат како луѓето“ („All About AI“, 2018). Современата наука посветена на проблемите на рабирање и развивање на ВИ, разликува три типа на ВИ, во однос на нивото на достигната интелигенција:

Специфична вештачка интелигенција (ВИ), која е способна комплетно да изврши само една задача. Таа поседува специфично знаење или е добра во извршување на една одредена активност како што е играње шах или програмот наречен AlphaGo кој „знае“ да ја игра играта ГО или систем кој ги одделува т.н. спам пораки во електронската пошта или виртуелен помошник. Во овој тип на ВИ спаѓаат и автономните автомобили, иако тие се засновани врз повеќе специфични ВИ-системи, кои работат паралелно и се поврзани во систем на меѓусебна комуникација. Специфична ВИ е она што научниците и креаторите на разновидни ВИ-артефакти го имаат досега постигнато.

Вештачка општа интелигенција (ВОИ), многу посоефицицирана од специфичната ВИ. Општата вештачка интелигенција е онаа за која голем број научници и креатори на ВИ зборуваат и навестуваат дека доаѓа. Сега засега, уште сме далеку од креирање вакви системи или машини, бидејќи мозокот на човекот е многу компликуван и истражувачите сè уште потполно не разбираат како тој функционира. На пример, нашиот мозок наликува на котел со зовриени и бурни невротрансмитерски реакции и досега во тој мозок се откриени и дефинирани преку 100 000 неврохемиски процеси. Тие се одвиваат истовремено и непрекинато, секоја секунда, вклучувајќи го тука и времето поминато во спиење. Овој тип на интелигенција е предизвик пред кој науката е исправена - да се развие ВИ која ќе интерпретира знаење и ќе го поврзува тоа знаење од различни области за да планира и донесува одлуки. Супер вештачка интелигенција или суперинтелигенција, е ВИ која е поинтелигентна од сите гении од сите области на знаењето. Таа би поседувала креативност, мудрост и социјални вештини. Во одредени научни кругови преовладува идејата и надежта дека ќе креираме ваков тип суперинтелигенција, набргу по креирањето на општата ВИ.

Честопати на ВИ гледаме како „наша иднина“, која нуди голем потенцијал трајно да го промениме начинот на кој работиме и живееме. Со оглед на тоа што нашите алгоритми креирани за ВИ стануваат од ден на ден се помоќни, можеби ќе почнеме да гледаме креации на ВИ кои можат да работат понезависно во еден поширок опсег или поширок спектар на домени, но тоа е долг пат. Луѓе кои се загрижени за истражување на вештачката интелигенција не се загрижени за денешната технологија и денешниот степен на развој на ВИ, тие се загрижени за системите за вештачка интелигенција кои може да постојат во некое време во иднина. Секако, незапирливиот развој на ВИ уште сега наметнува повеќе прашања на кои мора да се даде одговор или барем сериозно да се разгледаат можните одговори. Типични прашања, но не и единствени, во тој поглед се: Ќе ја замени ли ВИ „вистинската“ човечка интелигенција и дали ќе развие своја свест? Дали ВИ ќе нè замени во нашата работа, општество и дури и во романтичните односи? Доколку развие самостојна свест, дали ВИ ќе ја преземе власта и на крајот ќе владее со светот?

Со оглед на моменталната состојба на развојот на ВИ, на овие прашања може да дадеме негативен одговор, но иднината би можела да наслика многу поинаква слика. Денешните креации на ВИ, не се ништо повеќе (повеќе или помалку) од „фенси“ статистика, без оглед дали станува збор за различни типови машини за учење или вештачки креирани нервни мрежи. Сепак, интен-

цијата на науката, техниката и капиталот е да не запрат на ова ниво и покрај предупредувањата на голем број научници вклучени во истражувањето на ВИ дека треба внимателно да се чекори и добро да се промисли по кој пат ќе оди развојот. Нема сомнение дека следниот чекор е „вистинската вештачка интелигенција“. Но, што ја одредува „вистинската вештачка интелигенција“ и дали е тоа сама контрадикција, да се постават зборовите „вистински“ и „вештачки“ еден до друг? „Вистинската вештачка интелигенција“, како симулација на вистинското човечко однесување, честопати се нарекува вештачка општа интелигенција (Artificial General Intelligence), а таа е повеќе од статистика и обработка на податоци со голема брзина. Машините или комплексните системи засновани врз вештачка општа интелигенција треба да бидат изградени на начин што ќе направи да не може да ги разликуваме од вистинските луѓе.

Доколку прифатиме дека сите досегашни активности поврзани со ВИ се насочени кон создавање машини кои би реагирале, би се однесувале и дејствувале на начин што нормално би бил подведен под поимот човечка интелигенција, тогаш психолошките компоненти на ваквите машини, стануваат уште поважни (Crowder et al., 2013; Wilson, Rauwolf, & Bryson, 2019). Додека многу компании прават големи достигнувања во областа на ВИ, не само благодарение на високите инвестиции, вистинската вештачка интелигенција допрва треба да биде изградена. Во таа насока неопходно е психологијата да се вклучи поактивно, бидејќи таа е релевантна за ВИ.

Вклучувањето на психологијата во креирањето на ВИ и ВОИ е значајно од многу причини. Би посочиле на три области, кои на прв поглед се различни, но подеднакво важни:

1. Искуството на корисниците во интеракцијата со ВИ, односно психолошкото влијание на ВИ врз луѓето (социјалното влијание, на пример, социјалните медиуми).
2. Психолошки концепти поврзани со ВИ, како што се емоции и емпатија на пр. „вештачко сочувство“ (ќе треба да се создадат ваков вид на „вештачки“ концепти, на ист начин како и вештачката интелигенцијата за да може да ја креираат целосно функционалната вештачка општа интелигенција).
3. „Ментално“ здравје и на машините и на човечките суштества (Graham et al., 2016).

Наведените три области, веќе се предмет на интерес на психолозите и истражувачите на ВИ. Така на пример, Бренден Лејк (Lake et al., 2017), психолог и истражувач на ВИ од универзитетот во Њујорк и неговите колеги, раз-

виле систем на ВИ наречен Бејсиан (Bayesian), кој може да постигне форма на еднократно учење (one-shot learning). Луѓето, па дури и децата, се многу добри во ова - на пример едно дете треба само еднаш или двапати да види ананас за да разбере дека тоа е овошје, да го земе од корпа и можеби да го нацрта. Еднократното учење е проблем на категоризација на објектите, кои главно се однесуваат на „гледање“ (визуализација) поврзана со компјутерите и системите на ВИ. Додека поголемиот дел алгоритми за категоризација на објектите, засновани на машинско учење, има потреба од стотина или илјадници урнеци/слики и многу голем број на податоци, еднократното учење има за цел да ги научи ВИ-системите како информацијата за категориите на објектите да ја стекнат од само еден или од само неколку урнеци/слики за обука. Овој т.н. „одгоре-надолу“ пристап кон учењето на ВИ се потпира помалку на идентификување на обрасците во податоците, а наместо тоа користи математички начини за опишување на правилата што го регулираат човечкото сознание и алгоритми кои моделираат како луѓето размислуваат и учат за светот.

Справувањето со предизвиците што ги носи визуелното учење на машините и улогата на психологијата во проучувањето на софистицираноста на човековото визуелно учење го истакна, авторката Линда Смит (Smith, 2019). Таа укажува на фактот дека човековата способност за визуелно учење им овозможува уште на бебињата да ги сфатат имињата и категориите на предметите на начин кој досега не е постигнат во светот на вештачката интелигенција. Смит (Smith, 2019) упатува на фактот дека деца не се раѓаат со оваа единствена способност; тие тоа го учат во првите 30 месеци од нивниот живот. Истражувањата на раниот развој на јазикот, што таа и нејзините колеги ги изведуваат во лабораторијата за когнитивен развој во Блумингтон, се сосредоточени на два дела на процесот на учење: податоци за обука и механизми кои го креираат тоа учење. Смит истакнува дека: „Податоците за учењето, искуството според кое се јавува учењето на визуелните категории, во основа се разликува од искуството кое се користи во машинското учење за обука од компјутерски вид - и од искуството кое се користи во експериментални психолошки студии за учењето“ (Smith, 2019, стр. 15). Токму проучувањето на оваа разлика, може да ни биде од голема помош да разбереме зошто човековиот визуелен систем е толку софистициран и зошто бебињата можат да ги научат поимите (имињата) на објектите преку еднократно учење.

Втора значајна област во која ја гледаме улогата на психологија во креирањето на ВИ се тестови за мерење на интелигенцијата. Тестовите за мерење на интелигенција се од големо значење за оние што ја креираат ВИ. Меѓутоа, најголем дел од тие тестови имаат визуелни компоненти. Тоа, на сегашниот

развој на ВИ го отежнува степенот на препознавање на „интелигенцијата“ на ВИ. Размислувањата одат во насока дека треба да се создаде чисто устен тест за интелигенција, каде испрашувачот поставува вербални прашања, а одговорот е даден со вербален одговор, без потреба од какво било пишување или цртање. Ваквиот вид на тестот би требало да се фокусира повеќе на анализирање на способноста да се размислува апстрактно, да синтетизира ново знаење, како и да анализира емоционална и социјална интелигенција. Развој на ваков вид тест ќе бара ангажман на значителен број научници и истражувачи од повеќе области, во прв ред од психологијата, развојот на децата, лекари за ментално здравје и секако научници од областа на компјутерските науки. Креирањето на ваков вид тестови е од големо значење за развој на вистинска ВИ, како и за проценка на здравјето на тие „идни“ членови на нашето општество. Преку развој на нови тестови можеме да ја придвижиме вештачката општа интелигенција напред и до суперинтелигенција. Овие тестови на крајот ќе бидат употребливи како начин да се идентификуваат проблемите со менталното здравје кај овие „нови“ членови на нашата заедница бидејќи исто како што човечките суштества се прилагодуваат на ВИ и таа ќе мора да го направи истото. ВИ треба да развие „разбирање“ за човечките суштества на сличен начин како што човечките суштества ќе треба да развијат разбирање за ВИ. Ако можеме да развиеме протокол за да анализираме различни апликации за ВИ, за да видиме кои од нив се поблиску до интелигенција слична на човекот, тогаш ќе мора да избереме методологии кои не придвижуваат во вистинската насока. Една од клучните области во која психологијата треба да даде значаен придонес е токму интеракцијата помеѓу ВИ и луѓето. Тоа секако нема да биде можно ако не развиеме психологија на вештачката интелигенција и тоа не само како теориска, туку и како практична дејност.

Една општа претпоставка за вештачка интелигенција е дека таа е нешто што ја „имитира“ свеста или го имитира „субјективното“ искуство на човечкото. Во реалноста, сосема е можно „вештачките интелигенции“ да бидат многу интелигентни, но исто така да бидат и многу туѓи и сосем поразлични од луѓето. Човековата интелигенцијата не е спектар, таа е огромен простор на можни умови, со многу можности. Некои експерти се сомневаат дали машините воопшто би можеле да имаат нешто слично на човековата интелигенција и свест. Еден од нив е Христоф Кох (Christof Koch), главен научен директор на Ален институтот за наука за мозокот во Сиетл, кој застапува мислење дека сè уште не знаеме зошто се јавува, како се јавува, па дури и што е тоа што го определуваме како „свесно“. „Мислам дека свеста, како маса, е фундаментално својство на универзумот. Аналогијата, и тоа многу добра, е дека можете да

направите прилично добри временски предвидувања овие денови. Можете да ја предвидите внатрешноста на бурата. Но, во компјутерот никогаш не е влажно. Можете да симулирате црна дупка во компјутер, но 'простор-времето' нема да биде искривено. Симулирањето на нешто не е вистинска работа“ (Koch, 2004, p. 60). Со други зборови, иако свеста може да се симулира, Кох (Koch) верува дека никогаш нема да се случи вистински во вештачка форма.

Луѓето, во моментот, се единствениот „предмет“ што го проучува когнитивната наука, единствениот „интелигентен систем“ што е предмет на изучување. Може да ги поврземе ефектите со причините и да разбереме кои причини се специфични за луѓето и што предизвикуваат интелегентните умови воопшто. Но, за жал, луѓето се исто така единствените когнитивни научници и како луѓе што ги имаме претпоставките за умовите воопшто, не може а да не бидеме пристрасни. Дали воопшто ќе успееме да ја надминеме таа пристрасност во однос на ВИ, останува да покаже иднината.

Литература

- All about AI: What you need to know. (2018, April 12). Retrieved October 3, 2019, from <http://www.microknowledge.com/all-about-ai-what-you-need-to-know/>
- Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In K. Frankish & W. M. Ramsey (Eds.), *The Cambridge Handbook of Artificial Intelligence* (pp. 316-334). Cambridge University Press.
- Crowder, J., Carbone, N. J., Friess, S. (2013). *Artificial cognition architectures*. New York: Springer.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Graham, S., Deep, C., Lee, E. (2019). Artificial Intelligence for mental health and mental illnesses: an overview. *Current Psychiatry Reports*, 21, 116-123.
- Kirwan, G. (2016). The psychology of artificial intelligence. In I. Connolly, M. Palmer, H. Barton, & G. Kirwan (Eds.), *An Introduction to Cyberpsychology*, (pp. 286-298). Routledge/Taylor & Francis Group.
- Koch, C. (2004) *The quest for Consciousness: A neurobiological approach*. Roberts & Company Publishers, Englewood, Colorado.
- Lake, B., Ullman, T., Tenenbaum, J., & Gershman, S. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, 1-72.
- Smith, L.B. (2019). What do babies have that computers don't?. *Observer*, 34, 15-16.
- Turing, A. M. (1950). Computing machine and intelligence. *Mind*, 59, 433-460.
- Wilson, H., Rauwolf, P., & Bryson, J. J. (2019). Evolutionary psychology and artificial intelligence: The impact of artificial intelligence on human behaviour. In T. Shackelford (Ed.), *The SAGE Handbook of Evolutionary Psychology*, Sage Publications.
- Winarski, N. (2015). The president of SRI ventures on bringing SIRI to life. *Harvard Business Review*, September issue, 39-42.
- Winerman, L. (2018). Making a thinking machine. *Monitor on Psychology*, 49, 30-42.

WHY DO WE NEED THE PSYCHOLOGY OF ARTIFICIAL INTELLIGENCE

Vladimir Davchev

Abstract

As never before, we are faced with the challenge of designing an artificial intelligence machine that will think, learn and experience, and operate autonomously without human supervision. But as the problems that we teach people to solve with machines become more complex, so does the challenge become more intriguing for practitioners from more academic disciplines. If we intend to create artificial intelligence as “entities” to create better and easier quality of life, then it is essential to use a more substantial approach and a different methodology from other fields of science, not just technology and computer science. Although there is no universal definition of intelligence, there are many models and concepts, most of which overlap with the idea that cognitive intelligence involves the ability to think, rationalize, solve problems and reason. They are all psychological concepts. The purpose of this paper is to establish some basic principles and concepts that would be the foundation of autonomous artificial intelligence, which would not be possible without the psychology of artificial intelligence. With the fully autonomous behavior, learning, thinking of an artificial intelligent system (artificial brain), there is need for that system to be designed (software and hardware) in order to be able to perform mimic processes and sub-processes existing in the human brain, including intuitive ones as well as emotional concepts. This paper is an attempt to discuss the psychological concepts of artificial intelligence and possible ways for its creation.

Keywords: *intelligence, Artificial Intelligence, psychology of AI, artificial mind*

