

AI не трябва да е Open Source Project

Dimiter Dobrev
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
d@dobrev.com

Кой трябва да притежава технологията Изкуствен Интелект? Тази технология трябва да е на всички, но не самата технология, а плодовете, които тя ще ни даде. Разбира се, не трябва да позволяваме AI да попада в ръцете на безотговорни хора. Аналогично, ядрените технологии трябва да носят полза на всички, но тези технологии трябва да се пазят в тайна и да не са общодостъпни за всеки.

Въведение:

Има много хора, които защитават идеята, че технологията AI трябва да се разпространява свободно и дори, че тя трябва да бъде Open Source Project. Между тези хора има дори отговорни и сериозни хора, какъвто е президента Макрон [1]. Тук ще се опитаме да поспорим с тези хора и да им обясним колко погрешно и дори пагубно би било подобно решение.

Когато президента Макрон [1] говори за отворени алгоритми, може би той по-скоро има предвид собствеността върху тези алгоритми. Разбира се, няма нищо лошо собствеността да бъде на всички, но това не значи, че кода на тези алгоритми може да бъде общодостъпен. Например една ядрена електроцентрала може да бъде държавна, тоест на всички, но това не значи, че технологиите използвани в тази централа са общодостъпни и че всеки може да вземе чертежите и по тях да си направи собствена ядрена централа.

В момента отношението към технологията Изкуствен Интелект е изключително безотговорно. Ние се намираме в зората на развитието на тази технология и въобще не можем да си представим каква мощ и неподозирани възможности крие това изобретение. Какво се е случило през 1896 година? Тогава Анри Бекерел [2] открива, че ако постави в чекмедже парче уранова руда върху фотографска плака, то след време плаката се осветява. Ако постави метален ключ между плаката и рудата, то изображение на ключа се отпечатва върху плаката. Така Бекерел открива радиоактивността, но по това време той въобще не може да си представи потенциала, който крие тази технология. Експеримента на Бекерел е забавен и е по-скоро нещо като фокус. Същото в момента е положението с Изкуствения Интелект. Появяват се експерименти, които са интересни и забавни, но хората въобще не си представят до къде може да ни доведе тази технология.

Можели ли са хората през 1896 година да предвидят колко мощна и опасна е ядрената технология? Тогава не е имало как да се досетят. Това става по-късно, когато откриват колко много енергия се отделя при разпада на атомното ядро.

Можем ли сега да се досетим колко опасна е технологията Изкуствен Интелект? Да, и повечето разумни хора се досещат, макар и да не осъзнават действителния мащаб на това откритие.

Всеки разумен и отговорен човек трябва да си зададе въпросът дали да участва в разработката на новата технология или да остави това на глупавите и безотговорните.

В тази статия се занимаваме с технологичните катастрофи, които могат да се предотвратят, а не с неизбежните последствия, които няма как да бъдат предотвратени. Например, ако дадете на един глупак моторен трион, той може да изсече цялата гора и това е неизбежно последствие. Ако глупака си отреже крака, това е технологична катастрофа, която би била предотвратена, ако глупака не беше чак толкова глупав и ако внимаваше повече.

Казваме, че един много мощен интелект е нещо опасно. Това не е нова идея. Още Adorno и Horkheimer [3] казват, че разума може да бъде друга форма на варварство. Те казват, че с помощта на интелекта хората могат да променят природата по безогледен варварски начин. В [3] не се говори за изкуствен интелект, а за бюрократичната машина. Приликите между изкуствения интелект и бюрократичната машина са повече отколкото разликите и затова казаното в [3] може да се приеме за казано по темата. В [3] авторите разглеждат случая когато група хора използват бюрократичната машина като оръжие с цел да подчинят останалите (тоталитарната държава). В [3] не се разглежда случая, когато бюрократичната машина излиза извън контрол и започва да действа против волята на хората, защото това е невъзможно. Обществото като цяло винаги може да промени законите и правилата, които управляват бюрократичната машина. Тоест обществото като цяло няма как да загуби контрола върху бюрократичната машина, но отделния човек няма такъв контрол. За отделния човек бюрократичната машина е даденост, която той не може да промени. Аналогична е ситуацията с изкуствения интелект. Обществото като цяло ще запази контрола си върху AI, освен ако не сме достатъчно глупави да го изпуснем, но за отделния човек AI ще е даденост, която не може да бъде променена. Последното е едно от неизбежните последствия от появата на AI, което излиза извън обхвата на темите на тази статия.

Какво може да се случи?

Може да се случи катастрофа. Такава катастрофа може да е предизвикана умишлено или неволно. При ядрените технологии имената да две такива катастрофи са Хиросима и Чернобил. Първата е предизвикана умишлено, а втората поради глупост и невнимание.

Умишлена катастрофа означава някой да реши да използва новата технология със зъл умисъл, тоест като оръжие. Тук понятието зъл умисъл е относително, защото всеки създател на оръжие смята, че създава нещо полезно и че убивайки хора, той спасява живота на други хора. Обикновено се твърди, че убиваме малко, за да спасим много или поне че убиваме от чуждите, за да спасим от нашите.

Може ли AI да убива? Не действат ли законите на роботиката [4], които Айзък Айзимов е измислил? Всъщност тези закони са едно добро пожелание и те по никакъв начин не задължават създателите на AI да се съобразяват с тях. В момента технологиите, които претендират да имат нещо общо с AI се използват предимно за оръжия. Например, така наречените „умни бомби“. Обикновено се твърди, че глупавата бомба убива наред, докато умната убива само тези, които сме й казали да убие. Тоест умната бомба е по-малко кръвожадна и по-хуманна. Това последното е по-скоро оправдание за създателите на умни

бомби. Тези бомби са по-мощно оръжие от глупавите бомби и с тях можем да убием повече хора отколкото със старите глупави бомби.

Нека си представим, че се е случила технологична авария и някой е изпуснал духа от бутилката и е загубил контрола над AI. Ако гледаме на AI като на оръжие, нека си представим, че някой използва това оръжие срещу нас. По този начин отношението ни към използването на AI като оръжие ще е строго негативно, защото отношението към едно оръжие много зависи от това дали ние го използваме или някой го използва срещу нас.

Имаме ли шанс да преживеем подобна технологична авария? Отговорът е, нямаме никакъв шанс. Във фантастични филми от рода на Терминатора [5] се описва как хората воюват с роботите, но там имаме едни много тъпи роботи, които са много по-тъпи от хората. Истината за AI е, че той ще е много по-умен, от който и да е човек. Тоест, идеята за равноправно състезание между хора и роботи е безсмислена. Също толкова безсмислено е да организираме равноправно състезание по тичане между хора и автомобили. Хората няма да имат никакъв шанс, защото автомобилите са много по-бързи от хората.

Тоест, ние не можем да си позволим подобна технологична авария, просто защото няма да я преживеем. В историята си човечеството е преживяло много природни бедствия и технологични аварии, но винаги ефекта от тези аварии е бил локален. Например аварията в Халифакс причинява взрив, който разрушава града. Въпреки това, пораженията са локални и са ограничени до един град. Като възможно най-страшния сценарий за катастрофа се споменава ядрената война, но дори и това би имало само локални последици. Една подобна война би могла да унищожи градовете, но все някое от селата би оцеляло. Тоест, дори и възможна ядрена война не представлява такъв риск, какъвто е евентуалната загуба на контрола над AI.

Можем ли да заключим звяра в клетка?

Можем ли да създадем AI, но за да сме сигурни, че няма да направи някоя беля да го затворим в една виртуална реалност? Така ще можем да го наблюдаваме отстрани и да го изучаваме, но винаги ще можем да дръпнем щепсела и да го изключим, ако решим.

Да, можем. Ако AI няма вход и изход (тоест уши и уста), тогава той не може целенасочено да повлияе на външния свят. Дори и да повлияе, то това няма да е целенасочено, защото той за външния свят въобще няма да знае. Например, ние хората знаем ли дали не сме в Матрицата (филма Матрицата [6]). Както ние не знаем дали нашият свят е истински или виртуален, така и AI няма да знае.

Достатъчно е само AI да няма вход (тоест да не получава никаква информация от външния свят). Що се отнася до изхода, щом ние наблюдаваме AI, то той има изход (ръце и уста), защото чрез своето поведение той ще влияе на нас и от там, чрез нас ще повлияе на външния свят.

Тоест, рецептата е много проста. Държим AI в една виртуална реалност и така избягваме риска от технологична катастрофа. Да, но ние ще искаме AI да ни каже нещо за реалния свят. Например да ни направи прогноза за времето или да ни каже какви ще са цените на борсата. Може да поискаме да свърши някаква работа, например да измие чиниите или да

измете пода. За всичките тези неща ще е нужно да го пуснем от виртуалната реалност и да му позволим да влезе в реалния свят. Дори и да се уговорим да не го пускаме, все ще се намери някой, който ще се изкуши и ще го пусне.

Можем ли да заключим AI в клетка без да го лишаваме от информация за външния свят? Тоест, да може да чува реалния свят и да го гледа през решетката, но въпреки това да си запазим възможността винаги да можем да дръпнем щепсела и да го изключим.

Може ли лъвът да избяга от зоологическата градина? Може, макар че това е малко вероятно, защото лъвът е много глупав. Достатъчно е да има едно обикновено резе на вратата на клетката и лъвът няма да се сети как да си отвори. Маймуната има по-големи шансове да избяга, защото тя е по-умна от лъва. Ако разчитаме на обикновено резе, това няма да спре маймуната, защото тя ще се сети как да го отвори. Най-трудно е да заключим човек. Бягства се случват дори и в най-добре охраняваните затвори. Хората са много умни и почти винаги намират начин как да се измъкнат. Нека сега си представим, че сме се опитали да заключим едно същество, което е много по-умно от нас. Нека си представим човек охраняван от маймуни. Ще успее ли човека да надхитри маймуните и да избяга?

Добре, но ние винаги ще имаме възможността да дръпнем щепсела и да изключим AI, ако решим, че той е излязъл извън контрол. Да, но може и да не можем да го изключим. Когато AI излезе извън контрол той може да реши да не ни позволява да го изключим.

След като AI не може да бъде заключен в клетка, можем ли да го използваме в реалния свят? Да можем, но трябва да сме сигурни, че сме създали един добронамерен AI, който няма да се опитва да вземе властта от нас. Например кучето може да избяга и може да ни ухапе, но не го прави, защото е добронамерено.

Тоест, ако внимаваме какъв AI създаваме, няма да се случи технологична катастрофа. Това означава, че хората, на които ще разрешим да се занимават с тази технология трябва да са достатъчно умни и отговорни. Ако такива бяха хората занимаващи се с ядрени технологии нямаше да се случи нито Хиросима, нито Чернобил.

Принципа на моркова и тоягата

Ето един много прост пример за технологична катастрофа. Когато говорим за AI предполагаме, че той се обучава с поощрения и наказания (тази концепция е известна като Reinforcement Learning). Това е принципа на моркова и тоягата. Целта на AI е ясна, повече поощрения и по-малко наказания. Можем да стартираме такъв AI и да дадем на човека, който го управлява два бутона, с които да поощрява и да наказва AI. Проблем би възникнал, ако AI реши да забрани да човека да натиска бутона за наказание и да го принуди непрекъснато да натиска бутона за поощрение. По този начин AI ще превърне човека в свой роб, който е принуден непрекъснато да натиска бутона за поощрение.

Подобно нещо би се случило, ако магарето вземе властта и забрани на стопанина си да използва тоягата и ако го принуди по цял ден да го храни с моркови. Слава богу магарето не е достатъчно умно, за да вземе властта и това не може да се случи.

Можем ли да не създаваме AI?

Можем ли да се уговорим да не отваряме кутията на Пандора? Могат ли учените, които се занимават с изследвания в областта на Изкуствения Интелект да се съберат и да се уговорят да не правят това откритие? Отговорът е, че това няма как да се случи. Дори и част от учените да успеят да постигнат подобно споразумение, ще има и такива които няма да са част от споразумението или ще са част, но няма да се съобразят и ще го нарушат.

Подобно нещо се случва с ядрената технология. Физикът Вернер Хайзенберг твърди, че през 1941 година се среща тайно със бившия си учител Нилс Бор в Копенхаген и те двамата постигат уговорка да не създават атомната бомба. Нилс Бор от своя стана отрича да е сключвал подобно споразумение. Имало ли е или е нямало подобно споразумение е без значение, защото и Германия и САЩ не спират ядрените си програми. Възможно е, отделни хора да са саботирали развитието на ядрените технологии, но е имало и достатъчно много хора, които са продължили да работят по темата.

Защо трябва да засекретим AI технологията?

Всяка опасна технология се засекретява и достъпа до нея се ограничава. Пример за това са огнестрелните оръжия. Защо не позволяваме на децата да си играят с картечници? Каква беля може да се случи? Детето може да застреля един-двама, най-много сто човека. С помощта на AI технологията детето може да направи много по-голяма беля. Например, детето може да зададе задача на AI да избие всички, които не му харесват и това може да се окаже, че са всички хора, включително и този, които е задал задачата.

Сега ще ми кажете, че AI технологията е много сложна и едно дете не може да се справи с нея. Аз не твърдя, че едно дете може да създаде тази технология, но твърдя, че ако му я предоставим Open Source, то то би могло да я използва. Същото е и с картечницата. Едно дете не може само да си направи картечница, но ако му дадем една, то то би могло да стреля с нея. В това няма нищо сложно. Дърпаш спусъка и тя започва да стреля.

Подобно е положението с хаковете за компютри. Това са пропуски в сигурността на вашия компютър, които позволяват да се проникне дистанционно в него. Хаковете са или злоумишлено направени или са резултат на неволна грешка. Обикновено при създаването на една операционна система се оставят пропуски в сигурността, които по-късно да позволят проникването и контролирането на компютрите работещи с тази операционна система. Идеята на тези хакове е да се пазят в тайна и да се ползват само от техните създатели. Какво се случва на практика? Някой открива някакъв хак и го публикува в мрежата. Идеята е, да може всеки да се защити. Резултата е, че всеки може да го използва и да проникне в компютъра ви.

Целта на умишлените хакове е да могат тайните служби да влизат и да контролират вашия компютър. Аз лично нямам нищо против това тайните служби да влизат и да контролират моя компютър, защото знам, че това са хора отговорни и че ще влязат и излязат от моя компютър и аз дори няма и да забера, че са влизали.

Какво става обаче, когато в компютъра ви проникне някой тийнейджър. Той ще иска да отбележи своето посещение. Ще затрие някой важен файл и ще напише нещо неприлично на десктопа. Как тийнейджърите успяват да открият пропуските в сигурността, толкова ли са умни? Истината е, че въобще не са умни. Просто използват публикувани хакове, които са Open Source и които могат да се ползват от всеки. Ако някой тийнейджър е достатъчно умен сам да открие хак, то той вероятно ще е и достатъчно отговорен, за да не драска мръсотии върху десктопа ви.

Секретни списания

Мисля, че вече сте съгласни, че технологията AI не трябва да попада в ръцете на деца и на безотговорни хора. По същия начин сте съгласни, че на децата не трябва да им разрешаваме да си играят с картечница.

Може ли да дадем на децата едно достатъчно добро упътване, което да им позволи сами да си направят картечница? Разбира се, най-малките няма да се справят, но техните батковци биха могли, особено ако упътването е достатъчно подробно и добре написано. Може ли да смятаме, че всеки, който е достатъчно умен да направи картечница по упътване е и достатъчно отговорен, за да я ползва? Според мен не можем да направим такова предположение и затова упътванията за това как се прави картечница трябва да са с ограничен достъп.

Същото важи и за статиите, в които се описва AI технологията. Първо да кажем какво е AI. Изкуственият Интелект е програма. Една програма може да бъде написана. Написването на програмата е техническа задача подобна на решаването на ребус. Разбира се, преди да напишем програмата трябва да измислим алгоритъма. Трябва да си доста умен, за да измислиш алгоритъма на AI, но този алгоритъм може да бъде подробно описан в някоя статия за AI. Ако разполагаш с такава статия, самото написване на програмата AI може да се окаже техническа задача. Затова, според мен, статиите за AI трябва да са с ограничен достъп.

Навремето в бившия Съветски Съюз имаше секретни списания. Всички военни и потенциално опасни технологии се отпечатваха в такива списания. Разбира се, хората, които имаха право да четат тези списания бяха внимателно подбрани и техният кръг беше силно ограничен. Смятам, че днешните AI технологии трябва да се публикуват само в такива секретни списания.

Сериозните списания

Може ли в сериозно научно списание да се появи статия описваща алгоритъма на AI? Това е почти невъзможно, защото сериозните списания имат сериозна цензура, която не допуска статии, които биха представлявали опасност. Тоест, те не биха допуснали статия, която описва нова неизвестна и потенциално опасна технология. Цензурите се наричат рецензенти. Те са анонимни и не носят никаква отговорност за своите рецензии.

Това, че сериозните списания не допускат сериозни статии описващи AI технологията не означава, че такива статии не се появяват. Учените отхвърлени от рецензентите публикуват техните резултати където им попадне. Често те публикуват на своите Интернет страници или в различни блогове. Срещу това безразборно публикуване се взеха мерки в последните години. Например, имаше сайтове които пазеха снимка на Интернет и там можеше да се види една статия кога е публикувана. Тези сайтове бяха затворени. Също така блоговете даваха дата на публикациите. Сега вече блоговете не дават такава дата (точно дата има, но няма дата на последната редакция). Въпреки всичко безразборното публикуване продължава и цензурата не въвежда ред, а води до още по-голяма анархия.

В хартиените списания цензурата е неизбежна, защото хартията е ограничен ресурс и не може всеки да публикува каквото му скимне. Разбира се, хартиените списания са една отживелица от миналото. Днес имаме електронни списания, където не е нужно да се ограничава дължината на статията, нито да се спират статии, затова че са глупави или прекалено умни.

Дори и едно електронно списание може да се претовари, защото един автор може да генерира и да изпрати един милион статии. Тоест, при електронните списания имаме явлението наречено SPAM. Затова е добре когато се подават статии в електронното списание да има малка такса. Например един долар на статия или един долар на страница.

Защо рецензентите спират всички статии, които излизат от традиционния шаблон и не са нещо, което се дъвче поне от 50 години? Може би рецензентите са отговорни хора и искат да предпазят човечеството от неконтролираното навлизане на новите технологии, а може те просто да са глупави и да не могат да разберат една статия, ако тя е малко по-нестандартна. Може би го има и едното и другото. Може би рецензентите просто са ревниви и се дразнят от всеки, който се прави на много умен, особено ако този някой е пропуснал да цитира рецензента.

Въпреки всичко рецензентите са нужни дори и в електронните списания, но не за да казват да или не, а за да оценяват статиите и да насочват читателите струва ли си тази статия да бъде прочетена. Тоест, рецензентите са нужни, но не в качеството им на цензури, а в качеството им на критици. В литературата имаме литературни критици, които оценяват романите и насочват читателите. Тези критици застават зад написаното с името си и носят отговорност за критиката си, защото ако заблудят читателя, той повече няма да вярва на тяхната критика.

Ако списанието е секретно, то рецензента ще е нужен, за да каже какво е нивото на секретност на дадена статия. Тоест, колко ограничен да е кръга от хора, които ще могат да я прочетат.

Това, че списанието е секретно ще означава, че само проверени, достатъчно отговорни хора ще могат да го четат, но там всеки трябва да може да публикува. Ако ограничим хората, които имат право да публикуват, резултата ще е, че голяма част от учените ще продължат да публикуват там където им попадне.

Заклучени компютри

Както вече казахме, AI е програма и това е една много опасна програма. Затова не трябва да позволяваме на безотговорни хора да си играят с нея. Една програма има смисъл само, ако имате компютър, на който да я пуснете. Програмата без компютър представлява просто безсмислен текст.

Днес всеки тийнейджър разполага с много мощен компютър и може да пусне на този компютър каквато си поиска програма. Навремето баща ми беше директор на най-големия изчислителен център в България за научни цели. В момента имам в джоба си много по-мощен компютър от това, на което беше директор баща ми.

Много хора се притесняват от това, че Северна Корея разполага с термоядрено оръжие. Аз повече се притеснявам от това, че те разполагат със супер компютър, на който биха могли да стартират програмата AI и да направят много по-големи поразии, отколкото би направила една термоядрена бомба.

Все пак, Северна Корея се управлява от пълнолетни хора, които носят отговорност за действията си, но ние позволяваме на всяко непълнолетно дете, което по закон не отговаря за действията си, да притежава компютър и да пуска на него каквато си поиска програма.

Щом не позволяваме на децата да си играят с огнестрелно оръжие, защо им позволяваме да си играят с компютри? Дали не трябва да забраним на обикновените хора да притежават компютър? Дали не трябва за притежаването на компютър да се иска разрешение, както се иска за притежаването на огнестрелно оръжие.

По времето на Китайската империя е имало закон забраняващ на обикновените хора да притежават оръжие. Това е причината, поради която в Китайската империя са измислени различни техники за ръкопашен бой, както и за бой с пръчка и с разни селскостопански инструменти (например нунджако). Дали не е дошло времето да се приеме закон забраняващ на обикновените хора да притежават компютри?

Идеята да се ограничи притежаването на компютри вече е реализиран в известна степен. Например смартфоните представляват компютри, но това са заключени компютри, на който не можеш да пуснеш каквато си поискаш програма, а само програма проверена и одобрена от някой, който решава вместо нас коя програма е опасна и коя е безопасна.

Постепенно в заключени компютри се превръщат таблетите и лаптопите. Все още, огромната част от настолните компютри не са заключени и позволяват на хората да програмират и да пускат на тях собствени програми. Предполагам, че и това ще е до време и че идва деня, в който всички компютри ще са заключени и притежаването на отключен компютър ще се наказва като едно от най-опасните углавни престъпления.

Преди двадесетина години хората оставяха компютрите си включени през нощта и позволяваха всеки желаещ да ги ползва и да изпълнява програми (говоря за компютрите работещи под UNIX). Тоест, имаше един огромен супер компютър, който беше на разположение на всеки, който пожелае да си поиграе с него. Днес благодарение на технологията Bitcoin този ресурс вече не е свободен. Сега всички свободни компютърни ресурси са впрегнати да копаят биткойни. Дори някой да има свободен компютърен

ресурс, той не би го предоставил за свободно ползване, защото знае, че някой друг ще се възползва, ще изкопае биткойни и ще спечели за негова сметка.

Заклучение

Трябва много сериозно да се отнесем към технологията AI и да засекретим всички програми, които имат нещо общо с тази технология. Трябва да засекретим и статиите в областта на AI и дори трябва да заключим компютрите, за да не може всеки да си играе и да експериментира с AI технологиите. Тези експерименти трябва да бъдат разрешени само на хора, които са достатъчно умни и отговорни. В момента, за да работиш като лекар трябва да отговаряш на куп изисквания, а за да правиш изследвания в областта на AI, не се иска нищо. Това трябва да се промени и да се въведат изисквания, на които трябва да отговарят AI изследователите.

Технологичната катастрофа, да бъде изпуснат контрола над AI, може да се случи от глупост или от безотговорност. Не трябва да позволяваме с AI да се занимават хора с комплекс за малоценност, които биха могли да пожелаят да получат абсолютната власт и да се превърнат в нещо като господар на вселената.

От друга страна трябва да позволим на независимите изследователи да публикуват (в секретните списания) и по този начин да им позволим да получат признание за своя труд. Разбира се, когато независимите изследователи публикуват, те трябва да получат дата и гаранция, че никой няма да се опита да оспори или открадне тяхната заслуга.

Когато говорим за технологична катастрофа имаме предвид неща, които могат да се избегнат и които ако сме достатъчно умни и отговорни ще избегнем. Това са загубата на контрол върху AI и използването на AI като оръжие или като средство една група хора да подчини останалите.

Има много други последствия на AI технологията, които са неизбежни. Например това, че хората ще си загубят работата е такова последствие, което не можем да предотвратим. Освен, че не можем да го предотвратим, ние не искаме да го предотвратяваме, защото никой от нас не иска да е принуден да работи. Бихме работили за удоволствие, но не искаме да работим по принуда.

Много хора се притесняват, че тайните служби им ровичкат из компютрите. Днес тайните служби виждат всичко и знаят всичко. Господ също вижда всичко и знае всичко, но това не притеснява никого. Разбира се, Господ е дискретен и добронамерен. Той няма да каже на жена ви, че сте ѝ изневерил, нито ще се възползва от информацията на вашия компютър за своя изгода. Тайните служби също са дискретни, но не винаги са добронамерени. Не случайно най-големите бандити обикновено са бивши или настоящи служители на тайните служби.

Няма нужда да се притесняваме, от това че тайните служби виждат всичко. Това е нещо неизбежно. Глупаво е да се притесняваме от неизбежни неща, които не можем да променим. Казахме, че в тези служби работят хора отговорни. Е, по-отговорни са от тийнейджърите, но това не означава, че са достатъчно отговорни. Начина да избегнем

проблеми не е да се крием от тайните служби, а да ги контролираме и да внимаваме, кой работи там.

Моята теза е, че трябва да разрешим на тайните служби официално, без да се крият, да контролират нашите компютри. Ако го направим ще забравим проблеми като компютърни хакове, вируси и SPAM. Компютрите ни ще станат сигурни и надеждни. Дори може да поискаме, когато ни гръмне харддиска, те да ни възстановят информацията. Така или иначе те пазят нашата информация при тях.

Кой работи в тайните служби е важно, но още по-важно е на кого ще позволим да прави изследвания в областта на AI. Това трябва да са умни, разумни и отговорни хора, без комплекс за малоценност и без престъпни намерения. Ако разрешим на всеки, който си поиска, да се занимава с подобни изследвания и експерименти, то ще се случат технологични катастрофи в сравнение, с които Хиросима и Чернобил ще ни се видят като дребни инциденти.

References

[1] Emmanuel Macron (2018). Emmanuel Macron Talks to WIRED About France's AI Strategy. *31 of March, 2018*, www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy

[2] Henri Becquerel (1896). "Sur les radiations émises par phosphorescence". *Comptes Rendus*. 122: 420–421.

[3] Adorno, Theodor W. and Horkheimer, Max. *Dialectic of Enlightenment*. Stanford University Press, 2002, ISBN: 9780804736336.

[4] Isaac Asimov (1950). *I, Robot* (The Isaac Asimov Collection ed.). New York City: Doubleday. ISBN 0-385-42304-7.

[5] James Cameron (1984). *The Terminator*. American science-fiction action film.

[6] The Wachowski Brothers (1999). *The Matrix*. Science fiction action film.