

GRUPP TA' ESPERTI
**INDIPENDENTI TA' LIVELL GĦOLI DWAR
L-INTELLIGENZA ARTIFICJALI**
MWAQQAF MILL-KUMMISSJONI EWROPEA F'ĠUNJU 2018



**DEFINIZZJONI TAL-AI:
KAPAĊITAJET U DIXXIPLINI EWLENIN**

Definizzjoni żviluppata għall-fini tar-riżultati tangibbli tal-Grupp

Definizzjoni tal-AI: Kapaċitajiet u dixxiplini xjentifiċi ewlenin

Grupp ta' Esperti ta' Livell Għoli dwar l-Intelliġenza Artifiċjali

Dikjarazzjoni ta' Čaħda ta' Responsabbiltà u Użu ta' dan id-Dokument: Id-deskrizzjoni u d-definizzjoni li ġejjin ta' kapaċitajiet u ta' oqsma ta' riċerka tal-AI huma simplifikazzjoni kbira u bażika ħafna tal-istat attwali. L-intenzjoni ta' dan id-dokument mhix li jiddefinixxi b'mod preċiż u komprensiv it-tekniki u l-kapaċitajiet kollha tal-AI, iżda li jiddeskrivi fil-qosor il-fehim kongunt ta' din id-dixxiplina li l-Grupp ta' Esperti ta' Livell Għoli qed juża fir-riżultati tangibbli tiegħu. Minkejja dan, aħna nittamaw li dan id-dokument jista' jservi wkoll bħala punt tat-tluq edukattiv siewi għan-nies li mhumex esperti fl-AI, li mbagħad jistgħu jagħtu segwitu b'riflessjoni aktar estensiva u approfondita dwar l-AI biex jiksbu għarfien aktar preċiż dwar din id-dixxiplina u t-teknoloġija.

L-HLEG dwar l-AI hu grupp ta' esperti indipendenti li twaqqaf mill-Kummissjoni Ewropea f'Ġunju 2018.

Kuntatt Nathalie Smuha - Koordinatur tal-AI HLEG
Indirizz tal-posta elettronika CNECT-HLG-AI@ec.europa.eu

Il-Kummissjoni Ewropea
B-1049 Brussell

Dokument magħmul pubbliku fi X ta' April 2019.

L-ewwel abbozz ta' dan id-dokument inġareġ fit-18 ta' Diċembru 2018, flimkien mal-ewwel abbozz tal-Linji Gwida dwar l-Etika tal-HLEG dwar l-AI għal AI Affidabbli. Ġie rivedut fid-dawl tal-kummenti riċevuti permezz tal-Alleanza Ewropea għall-AI u permezz tal-konsultazzjoni miftuħa dwar l-abbozz tal-Linji Gwida. Aħna nixtiequ nroddu ħajr b'mod espliċitu u tassew mill-qalb lil dawk li kkontribwew bir-rispons tagħhom fuq l-ewwel abbozz tad-dokument.

La l-Kummissjoni Ewropea u l-ebda persuna li taġixxi f'isem il-Kummissjoni ma hija responsabbli għall-użu li jista' jsir mit-tagħrif li ġej. Il-kontenut ta' dan id-dokument ta' ħidma hu r-responsabbiltà unika tal-Grupp ta' Esperti ta' Livell Għoli dwar l-Intelliġenza Artifiċjali (HLEG dwar l-AI). Għad li l-persunal tal-Kummissjoni ffaċilita t-tnejn ta' dan id-dokument, il-fehmiet espressi fih jirriflettu l-opinjoni tal-HLEG dwar l-AI u fl-ebda ċirkustanza ma jistgħu jitqiesu li jirriflettu pożizzjoni uffiċjali tal-Kummissjoni Ewropea.

Aktar informazzjoni dwar il-Grupp ta' Esperti ta' Livell Għoli dwar l-Intelliġenza Artifiċjali hi disponibbli online (<https://ec.europa.eu/digital-single-maret/en/high-level-expert-group-artificial-intelligence>).

Il-politika tal-użu mill-ġdid tad-dokumenti tal-Kummissjoni Ewropea hi rregolata bid-Deciżjoni 2011/833/UE (ĠU L 330, 14.12.2011, p.39). Għal kull użu jew riproduzzjoni ta' ritratti jew materjal ieħor li mhuwiex taħt id-drittijiet tal-awtur tal-UE, għandu jintalab permess direttament mingħand id-detenturi tad-drittijiet tal-awtur.

DEFINIZZJONI TAL-AI:

KAPAĊITAJET U DIXXIPLINI XJENTIFIĊI EWLENIN

Aħna nibdew mid-definizzjoni li ġejja ta' Intelligenza Artifiċjali (AI), kif proposta fi hdan il-Komunikazzjoni tal-Kummissjoni Ewropea dwar l-AI¹:

"L-Intelligenza artifiċjali (AI) tirreferi għal sistemi li juru mgħiba intelligenti billi janalizzaw l-ambjent tagħhom u jieħdu azzjonijiet – b'ċertu grad ta' awtonomija – biex jiksibu għanijiet speċifiċi.

Is-sistemi bbażati fuq l-AI jistgħu jkunu purament ibbażati fuq softwer, billi jaġixxu fid-dinja virtwali (eż. assistenti bil-lehen, softwer ta' analiżi tal-immaginjiet, magni tat-tiftix, sistemi ta' rikonossiment tad-diskors u tal-wiċċ) jew inkella l-AI tista' tkun inkorporata f'apparati ħardwer (eż. robots avvanzati, karozzi awtonomi, droni jew applikazzjonijiet tal-Internet tal-Ogġetti)."

F'dan id-dokument aħna nespandu din id-definizzjoni biex niċċaraw ċerti aspetti tal-AI bħala dixxiplina xjentifika u bħala teknoloġija, bl-għan li nevitaw fehim ħażin, niksbu għarfien komuni kondiviż dwar l-AI li jista' jintuża b'mod utli anki minn dawk li mhumiex esperti fl-AI, u nipprovdut dettalji siewja li jistgħu jintużaw fid-diskussjoni kemm dwar il-linji gwida għall-etika fl-AI kif ukoll dwar ir-rakkomandazzjonijiet tal-politiki dwar l-AI.

1. Is-sistemi tal-AI

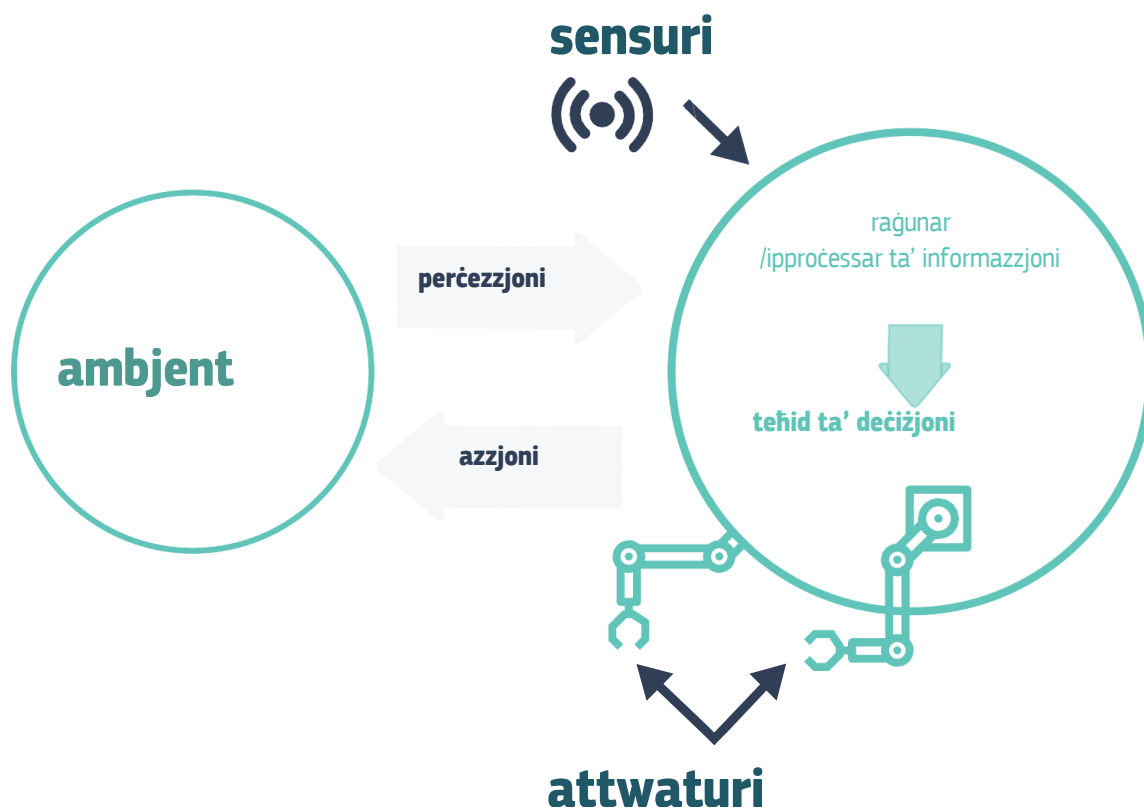
It-terminu AI fih referenza espliċita għall-kunċett ta' intelligenza. Madankollu, peress li l-intelligenza (kemm fil-magni kif ukoll fil-bniedem) hi kunċett vag, għad li għet studjata fil-fond minn psikologi, minn bijoloġisti u minn newroxjenzi, ir-riċerkaturi tal-AI jużaw l-aktar il-kunċett ta' razzjonalità. Dan jirreferi għall-ħila li tingħażel l-aħjar azzjoni li għandha tittieħed biex jinkiseb ċertu għan, fid-dawl ta' ċerti kriterji li għandhom jiġu ottimizzati u r-riżorsi disponibbli. Sintendi, ir-razzjonalità mhix l-uniku ingredjent fil-kunċett tal-intelligenza, iżda hi parti sinifikanti minnu.

F'dan li jsejwi, aħna se nużaw it-terminu *sistema tal-AI* biex infissru kwalunkwe komponent, softwer u/jew ħardwer ibbażat fuq l-AI. Tabilhaqq, normalment is-sistemi tal-AI jkunu *inkorporati* bħala komponenti f'sistemi akbar, minflok ma jkunu sistemi awtonomi.

B'hekk, qabel kollox, sistema tal-AI hi razzjonali, skont wieħed mill-aktar kotba użati dwar l-AI.² Iżda sistema tal-AI kif tikseb ir-razzjonalità? Kif indikat fl-ewwel sentenza tad-definizzjoni operazzjonali msemmija hawn fuq tal-AI, din tagħmel hekk billi tipperċepixxi l-ambjent li fih is-sistema tkun immersa permezz ta' ċerti sensuri, u b'hekk billi tiġbor u tinterpreta *data*, tirraġuna fuq dak li hu perċepit, jew tipproċessa l-informazzjoni miksuba minn din id-*data*, u tiddeċiedi xi tkun l-aħjar azzjoni, u mbagħad tagħxi kif xieraq, permezz ta' ċerti attwaturi, u b'hekk possibilmint timmodifika l-ambjent. Is-sistemi tal-AI jistgħu jew jużaw regoli simboliċi jew inkella jitgħallmu mudell numeriku, u jistgħu wkoll jadattaw l-imġiba tagħhom billi janalizzaw kif l-ambjent hu affettwat mill-azzjonijiet preċedenti tagħhom. L-illustrazzjoni ta' sistema tal-AI fl-Illustrazzjoni 1 tista' tgħin.

¹ Komunikazzjoni tal-Kummissjoni lill-Parlament Ewropew, lill-Kunsill Ewropew, lill-Kunsill, lill-Kumitat Ekonomiku u Soċjali Ewropew u lill-Kumitat tar-Reġjuni dwar L-Intelligenza Artifiċjali għall-Ewropa, Brussell, 25.4.2018 COM(2018) 237 final.

² "Artificial Intelligence: A Modern Approach", S. Russell u P. Norvig, Prentice Hall, it-tielet edizzjoni, 2009. "Artificial Intelligence: A Modern Approach", S. Russell u P. Norvig, Prentice Hall, it-tielet edizzjoni, 2009.



Illustrazzjoni 1: Turija skematika ta' sistema tal-AI.

Sensuri u perċezzjoni. Fl-Illustrazzjoni 1 is-sensuri tas-sistema huma murija bħala simbolu tal-wifi. Fil-prattika, dawn jistgħu jkunu kameras, mikrofoni, tastiera, sit web, jew apparat ieħor tal-input, kif ukoll sensuri ta' kwantitajiet fiżiċi (eż. temperatura, pressjoni, distanza, forza/torque, sensuri tattili). B'mod ġenerali, aħna rridu nipprovdha lis-sistema tal-AI b'sensuri li jkunu adegwati biex jipperċepixxu d-*data* preżenti fl-ambjent li tkun rilevanti għall-għan mogħti lis-sistema tal-AI mid-disinjatur uman tagħha. Pereżempju, jekk aħna rridu nibnu sistema tal-AI li tnaddaf b'mod awtomatiku l-art ta' kamra fejn tkun maħmuġa, is-sensuri jistgħu jinkludu kameras biex jieħdu ritratt tal-art.

Rigward id-*data* miġbura, spiss ikun siewi li ssir distinzjoni bejn *data* strutturata u dik mhux strutturata. *Data strutturata* hi *data* li hi organizzata skont il-mudelli predefiniti (bħal f'bażi ta' *data* relazzjonali), filwaqt li *data mhux strutturata* ma jkollhiex organizzazzjoni magħrufa (bħal fi stampa jew f'biċċa test).

Raġunar/proċessar ta' informazzjoni u Teħid ta' Deċiżjonijiet. Fil-qalba ta' sistema tal-AI jkun hemm il-modulu tar-raġunar/tal-proċessar tal-informazzjoni tagħha, li jieħu bħala input id-*data* li tiġi mis-sensuri u jipproponi azzjoni li għandha tittiehed, fid-dawl tal-għan li għandu jinkiseb. Dan ifisser li d-*data* miġbura mis-sensuri trid tiġi ttrasformata f'informazzjoni li l-modulu tar-raġunar/tal-proċessar tal-informazzjoni jkun jista' jifhem. Biex inkomplu fuq l-eżempju tagħna ta' sistema tal-AI tat-tindif, il-kamera tipprovdi stampa tal-art lill-modulu tar-raġunar/tal-proċessar tal-informazzjoni, u dan il-modulu jrid jiddeċiedi jekk inaddafx l-art jew le (jiġifieri x'inhi l-aħjar azzjoni biex jinkiseb l-għan fil-mira). Filwaqt li jista' jidher li donnu hu faċli għalina l-bnedmin biex immorru minn stampa ta' art għad-deċiżjoni ta' jekk għandhiex bżonn titnaddaf jew le, mhuwiex daqstant faċli għal magna, għaliex stampa hi sempliċiment sekwenza ta' sensiela ta' numri 0 u 1. Għaldaqstant, il-modulu tar-raġunar/tal-proċessar tal-informazzjoni jrid:

1. Jinterpreta l-istampa biex jiddeċiedi jekk l-art hix nadifa jew le. B'mod ġenerali, dan ifisser li jkun kapaċi jittrasforma d-*data* f'informazzjoni u jimmudella t-tali informazzjoni fil-qosor, li madankollu għandu jinkludi l-biċċiet ta' *data* rilevanti kollha (f'dan il-każ, jekk l-art hix nadifa jew le).

2. Ir-raġunar fuq dan l-għarfien jew il-proċessar ta' din l-informazzjoni biex jiġi prodott mudell numeriku (jiġifieri, formula matematika) biex tittieħed deċiżjoni dwar x'inhi l-aħjar azzjoni. F'dan l-eżempju, jekk l-informazzjoni miksuba mill-istampa tkun li l-art hi maħmuġa, l-aħjar azzjoni tkun li jiġi attivat it-tindif, inkella l-aħjar azzjoni tkun li kollox jibqa' wieqaf.

Kun af li t-terminu "deċiżjoni" għandu jiġi kkunsidrat b'mod ġeneriku, bħala kwalunkwe att ta' għażla tal-azzjoni li għandha tittieħed, u mhux bilfors ifisser li s-sistemi tal-AI huma awtonomi għalkollox. Deċiżjoni tista' tkun ukoll l-għażla ta' rakkomandazzjoni li għandha tiġi pprovduta lil bniedem, li jkun dak li jieħu d-deċiżjoni finali.

L-attwazzjoni. Ladarba tiġi deċiża l-azzjoni, is-sistema tal-AI tkun lesta li twettaqha permezz tal-attwaturi disponibbli għaliha. Fit-tpingġja ta' hawn fuq, l-attwaturi huma rappreżentati bħala idejn artikolati, iżda dawn mhux bilfors ikunu fiżiċi. L-attwaturi jistgħu jkunu softwer ukoll. Fl-eżempju tagħna tat-tindif, is-sistema tal-AI tista' tipproduċi sinjal li jattiva vacuum cleaner jekk l-azzjoni tkun li titnaddaf l-art. Bħala eżempju ieħor, sistema konverżazzjonali (jiġifieri, bot tat-taħdit) tagħxi billi tiġġenera testi biex twieġeb għal dak li jgħid l-utent.

L-azzjoni mwettqa jista' jkun li timmodifika l-ambjent, allura d-darba li jmiss is-sistema jerga' jkollha bżonn tuża s-sensuri tagħha biex tipperċepixxi possibbilment informazzjoni differenti mill-ambjent modifikat.

Sistemi razzjonali tal-AI mhux dejjem jagħżlu l-aqwa azzjoni għall-għan tagħhom, u b'hekk jiksbu biss *razzjonalità ristretta*, minhabba limitazzjonijiet fir-riżorsi bħall-ħin jew is-saħħa komputazzjonali.

Sistemi razzjonali tal-AI huma verżjoni bażika ħafna tas-sistemi tal-AI. Dawn jimmodifikaw l-ambjent iżda ma jadattawx l-imġiba tagħhom maż-żmien biex jiksbu l-għan tagħhom b'mod imtejjeb. *Sistema razzjonali li titgħallem* hi sistema razzjonali li, wara li tieħu azzjoni, tevalwa l-istat il-gdid tal-ambjent (permezz tal-perċezzjoni) biex tiddetermina kemm irnexxiet l-azzjoni tagħha, u mbagħad tadatta r-regoli dwar ir-raġunar u l-metodi ta' teħid ta' deċiżjoni tagħha.

2. L-AI bħala dixxiplina xjentifika

Dik ipprovduta hawn fuq hi deskrizzjoni astratta u sempliċi ħafna ta' sistema tal-AI, permezz ta' tliet kapaċitajiet ewlenin: il-perċezzjoni, ir-raġunar/it-teħid ta' deċiżjoni, u l-attwazzjoni. Madankollu, mhuwix biżżejjed biex tippermettilna nintroduċu u nifhmu ħafna mit-tekniki u mis-subdixxiplini tal-AI li jintużaw bħalissa biex jinbnew sistemi tal-AI, peress li dawn ilkoll jirreferu għad-diversi kapaċitajiet tas-sistemi. F'termini ġenerali, dawn it-tekniki kollha jistgħu jiġu raggruppati f'żewġ gruppi ewlenin li jirreferu għall-kapaċità tar-raġunar u tat-tagħlim. Ir-robotika hi dixxiplina oħra rilevanti ħafna.

Ir-Raġunar u t-Teħid ta' Deċiżjoni. Dan il-grupp ta' tekniki jinkludi r-rappreżentazzjoni tal-għarfien u r-raġunar, l-ippjanar, l-iskedar, it-tiftix, u l-ottimizzazzjoni. Dawn it-tekniki jippermettu biex jitwettaq ir-raġunar fuq id-*data* mis-sensuri. Biex ikun jista' jagħmel dan, wieħed ikollu bżonn jittrasforma d-*data* għal għarfien, u b'hekk qasam minnhom tal-AI għandu x'jaqsam ma' liema hu l-aħjar mod kif jiġi mmudellat it-tali għarfien (*rappreżentazzjoni tal-għarfien*). Ladarba l-għarfien jiġi mmudellat, il-pass li jmiss ikun li jsir raġunar bih (*raġunar tal-għarfien*), li jinkludi inferenzi permezz ta' regoli simbolici, attivitajiet ta' *ppjanar* u ta' *skedar*, *tiftix* permezz ta' sett kbir ta' soluzzjonijiet, u *ottimizzazzjoni* fost is-soluzzjonijiet possibbli kollha għal problema. Il-pass finali hu li tittieħed deċiżjoni fuq liema azzjoni għandha tittieħed. Il-parti tar-raġunar/tat-teħid ta' deċiżjoni ta' sistema tal-AI normalment tkun kumplessa ħafna u teħtieġ taħlita ta' diversi teknoloġiji minn fost dawk imsemmija hawn fuq.

It-tagħlim. Dan il-grupp ta' tekniki jinkludi t-tagħlim awtomatiku, in-netwerks newrali, l-apprendiment profond, is-siġar tad-deċiżjonijiet, u ħafna tekniki oħra ta' tagħlim. Dawn it-tekniki jippermettu biex sistema tal-AI titgħallem kif issolvi problemi li ma jistgħux jiġu speċifikati bi preċiżjoni, jew li l-metodu ta' soluzzjoni tagħhom ma jkunx jista' jiġi deskritt b'regoli dwar ir-raġunar simboliku. Eżempji ta' problemi bħal dawn huma dawk li għandhom x'jaqsam ma' kapaċitajiet ta' perċezzjoni bħal *fehim tat-taħdit* u *tal-lingwa*, kif ukoll *il-viżjoni tal-kompjuter* jew *it-tbassir tal-imġiba*. Kun af li dawn il-problemi jidhru faċli, għaliex tabilhaqq normlment ikunu faċli għall-bniedem. Madankollu, mhumix daqstant faċli għas-sistemi tal-AI, peress li dawn ma jistgħux jibbażaw fuq ir-raġunar bis-sens komun (almenu mhux għalissa), u huma partikolarment diffiċli meta s-sistema jkollha bżonn tinterpreta *data* mhux

strutturata. Hu hawn fejn it-tekniki li jsegwu l-approċċ tat-*tagħlim awtomatiku* jkunu siewja. Madankollu, it-tekniki tat-tagħlim awtomatiku jistgħu jintużaw għal ħafna kompiti oħra u mhux biss għall-perċezzjoni. It-tekniki tat-tagħlim awtomatiku jipproduċu mudell numeriku (jiġifieri, formula matematika) użat biex tiġi kkalkolata d-deċiżjoni mid-*data*.

It-tagħlim awtomatiku fih diversi forom. L-aktar approċċi użati spiss huma t-*tagħlim issorveljat*, it-*tagħlim mhux issorveljat*, u t-*tagħlim tat-tiżi*.

Fit-tagħlim awtomatiku ssorveljat, minflok ma jingħataw regoli dwar l-imġiba lis-sistema, aħna nipprovdulha eżempji ta' mġiba ta' input-output, bit-tama li din tkun kapaci tiġġeneralizza mill-eżempji (li tipikament jiddeskrivu l-imġiġoddi) u jgħibu ruħhom sew anki f'sitwazzjonijiet mhux murija fl-eżempji (li jistgħu jinqalgħu fil-ġejjieni). Fl-eżempju li qed nużaw, aħna kieku nagħtu lis-sistema ħafna eżempji ta' stampi ta' art u l-interpretazzjoni korrispondenti (jiġifieri, jekk l-art hix nadifa jew le f'dik l-istampa). Jekk aħna nagħtu biżżejjed eżempji, li jkunu varjati u inkluzivi biżżejjed għal ħafna mis-sitwazzjonijiet, is-sistema, permezz tal-algoritmu tat-tagħlim awtomatiku tagħha, tkun kapaci tiġġeneralizza biex issir taf ukoll kif tinterpreta sew stampi ta' artijiet li qatt ma tkun rat qabel. Ċerti approċċi tat-tagħlim awtomatiku jadottaw algoritmi li huma bbażati fuq il-kunċett ta' *networks newrali*, li hu ispirat b'mod ġeneriku mill-moħħ tal-bniedem peress li għandu network ta' unitajiet żgħira ta' proċessar (b'mod analoġu għan-newroni tagħna) b'ħafna ta' konnessjonijiet peżati bejniethom. Network newrali għandu bħala input id-*data* li tiġi mis-sensuri (fl-eżempju tagħna, l-istampa tal-art) u bħala output l-interpretazzjoni tal-istampa (fl-eżempju tagħna, jekk l-art hix nadifa jew le). Matul l-analiżi tal-eżempji (il-fażi tat-*taħriġ* tan-network), il-ponderazzjonijiet tal-konnessjonijiet jiġu aġġustati biex jaqblu kemm jista' jkun ma' dak li jgħidu l-eżempju disponibbli (jiġifieri, biex jiġi mminimizzat l-iżball bejn l-output mistenni u l-output ikkalkolat min-network). Fi tmiem il-fażi tat-*taħriġ*, fażi tal-ittestjar tal-imġiba tan-network newrali fuq eżempji li qatt ma jkunu dehru qabel tivverifika li l-kompitu ġie mifhum sew.

Importanti li jiġi nnotat li dan l-approċċ (bħal kull teknika ta' tagħlim awtomatiku) dejjem ikollu ċertu perċentwal ta' iżball, għad li normalment dan ikun żgħir. B'hekk, kunċett essenzjali hu l-*preċiżjoni*, kejl ta' kemm hu kbir il-perċentwal tat-tweġibiet korretti.

Hemm diversi tipi ta' *networks newrali* u ta' approċċi tat-tagħlim awtomatiku, li bħalissa wieħed mill-aktar li rnexxa minnhom hu l-*apprendiment profond*. Dan l-approċċ jirreferi għall-fatt li n-network newrali għandu diversi saffi bejn l-input u l-output li jippermettu li tintfieh hem ir-relazzjoni input-output ġenerali fi stadji suċċessivi. Dan jagħmel l-approċċ ġenerali aktar preċiż u b'anqas bżonn ta' gwida umana.

In-*networks newrali* huma biss għodda waħda tat-tagħlim awtomatiku, iżda hemm ħafna oħrajn, bi proprjetajiet differenti: foresti aleatorji u siġar imsaħħa, metodi ta' raggruppar, fattorizzazzjoni tal-matriċi, eċċ.

Tip utli ieħor ta' approċċ ta' tagħlim awtomatiku jissejjaħ *tagħlim tat-tiżi*. F'dan l-approċċ, aħna nħallu s-sistema tal-AI libera biex tiegħu d-deċiżjonijiet tagħha, maż-żmien, u f'kull deċiżjoni aħna nipprovdulha sinjal premju li jgħidilha jekk kinitx deċiżjoni tajba jew ħażina. L-għan tas-sistema, maż-żmien, hu li timmassimizza l-premju pożittiv riċevut. Dan l-approċċ jintuża, pereżempju, f'sistema ta' rakkomandatur (bħad-diversi sistemi ta' rakkomandatur online li jissuggerixxu lill-utenti dak li jista' jkun jinteressahom li jixtru), jew anki fil-kummerċjalizzazzjoni.

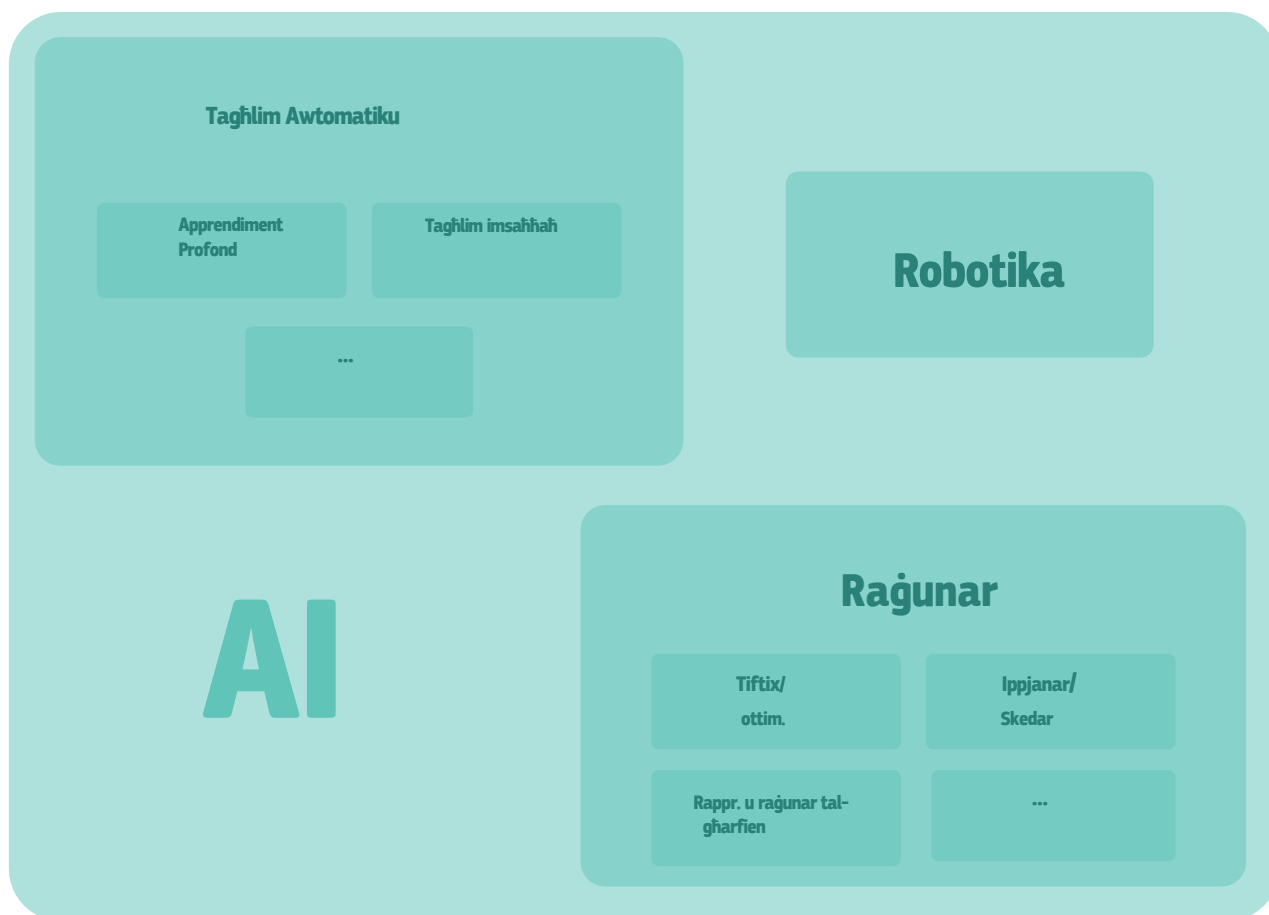
L-approċċi tat-tagħlim awtomatiku huma siewja mhux biss fil-kompiti ta' perċezzjoni, bħall-viżjoni u l-fehim tat-test, iżda f'dawk il-kompiti kollha li hu diffiċli li jiġu definiti u ma jistgħux jiġu deskritti b'mod komprensiv minn regoli dwar l-imġiba simbolika.

Kun af li d-distinzjoni bejn l-approċċi tat-tagħlim awtomatiku biex jinftehem kompitu għdid li ma jistax jiġi deskritt sew b'mod simboliku, u l-aġenti razzjonali tat-tagħlim (imsemmija fit-taqsima ta' qabel) li jistgħu jadattaw l-imġiba tagħhom maż-żmien biex jiksibu aħjar l-għan partikolari. Dawn iż-żewġ tekniki jistgħu jirkbu fuq xulxin jew jikkooperaw, iżda mhux bilfors ikunu l-istess.

Ir-robotika. Ir-robotika tista' tiġi definita bħala "l-AI fl-azzjoni fid-dinja fiżika" (imsejja wkoll *AI inkorporata*). Robot hu magna fiżika li trid tlaħhaq mad-dinamika, mal-inċertezzi u mal-kumplessità tad-dinja fiżika. Il-perċezzjoni, ir-raġunar, l-azzjoni, it-tagħlim kif ukoll il-kapacitajiet ta' interazzjoni ma' sistemi oħra normalment ikunu integrati fl-arkitektura tal-kontroll tas-sistema robotika. Minbarra l-AI, dixxiplini oħra jaqdu rwol fid-disinn u fl-operar tar-robots,

bħall-iġinerija mekkanika u t-teorija tal-kontroll. Eżempji ta' robots jinkludu manipolaturi robotiċi, vetturi awtonomi (eż. karozzi, droni, taxis li jtiru), robots umanojdi, vacuum cleaners robotiċi, eċċ.

L-Illustrazzjoni 2 turi ħafna mis-subdixiplini tal-AI msemmija hawn fuq, kif ukoll ir-relazzjoni tagħhom. Madankollu, importanti li jiġi nnotat li l-AI hi ferm aktar kumplessa milli turi din l-istampa, peress li tinkludi ħafna subdixiplini u tekniki oħra. Barra minn hekk, kif innotat aktar 'il fuq, ir-robotika tibbaża wkoll fuq tekniki li jaqgħu barra mill-ispazju tal-AI. Madankollu, aħna nemmnu li dan hu biżżejjed biex jiġu informati b'mod siewi l-kondiviżjoni, l-għarfien, u d-diskussjoni dwar l-AI, l-etika tal-AI, u l-politiki dwar l-AI, bid-diskussjoni li trid isseħħ fi ħdan il-grupp ta' esperti ta' livell għoli ferm multidixiplinari u b'diversi partijiet ikkonċernati.



Illustrazzjoni 2: Ħarsa ġenerali ssimplifikata lejn is-subdixiplini tal-AI u r-relazzjoni tagħhom.

Kemm it-tagħlim awtomatiku kif ukoll ir-raġunar jinkludu ħafna tekniki oħra, u r-robotika tinkludi tekniki li jaqgħu barra mill-AI. L-AI kollha kemm hi taqa' taħt id-dixiplina tax-xjenza tal-kompjuters.

3. Kunċetti u kwistjonijiet importanti oħra tal-AI

AI dejqa (jew dgħajfa) u ġenerali (jew b'saħħitha). Sistema ġenerali tal-AI hi maħsuba biex tkun sistema li tkun tista' twettaq ħafna mill-attivitàjiet li jagħmel il-bniedem. Min-naħa l-oħra, sistemi dojoq tal-AI huma sistemi li jistgħu jwettqu komputu speċifiku wieħed jew ftit kompiti speċifiċi. Is-sistemi tal-AI mnedija bħalissa huma eżempji ta' AI dejqa. Fil-jiem bikrin tal-AI, ir-riċerkaturi użaw terminoloġija differenti (AI dgħajfa u b'saħħitha). Għad hemm ħafna sfidi etiċi, xjentifiċi u teknoloġiċi miftuħa biex jinbnew il-kapaċitajiet li jkunu meħtieġa biex tinkiseb AI ġenerali, bħar-raġunar bis-sens komuni, il-koxjenza, u l-ħila li l-magna tiddefinixxi l-iskop tagħha stess.

Kwistjonijiet u preġudizzju tad-data. Peress li ħafna sistemi tal-AI, bħal dawk li jinkludu komponenti tat-tagħlim awtomatiku ssorveljat, jibbażaw fuq ammonti enormi ta' data biex jaħdmu sew, importanti li jiġi mifhum kif id-data tinfluwenza l-imġiba tas-sistema tal-AI. Pereżempju, jekk id-data tat-taħriġ tkun preġudikata, jiġifieri, ma tkunx

ibbilancjata jew inkluziva biżżejjed, is-sistema tal-AI m'harrġa fuq it-tali *data* ma tkunx kapaċi tiġġeneralizza sew u possibbilment tista' tiegħu deċiżjonijiet iġusti li jistgħu jiffavorixxu lil ċerti gruppi fuq oħrajn. Dan l-aħħar, il-komunità tal-AI kienet qed taħdem fuq metodi għad-detezzjoni u għall-mitigazzjoni tal-preġudizzji f'settijiet ta' *data* tat-taħriġ kif ukoll f'partijiet oħra ta' sistema tal-AI.

AI kaxxa sewda u spjegabbiltà. Ċerti tekniki tat-tagħlim awtomatiku, għad li rnexxew ħafna mil-lat tal-preċiżjoni, huma opaki ħafna f'termini ta' fehim ta' kif jieħdu d-deċiżjonijiet. Il-kunċett ta' *AI kaxxa sewda* jirreferi għal xenarji bħal dawn, fejn ma jkunx possibbli li tiġi traċċata lura r-raġuni wara ċerti deċiżjonijiet. L-ispegabbiltà hi proprjetà ta' dawk is-sistemi tal-AI li, minflok, jistgħu jipprovdu forma ta' spjegazzjoni għall-azzjonijiet tagħhom.

AI orjentata lejn l-għanijiet. Is-sistemi kurrenti tal-AI huma orjentati lejn l-għanijiet, u dan ifisser li huma jirċievu l-ispeċifikazzjoni ta' għan li għandu jkiseb mingħand bniedem u jużaw xi tekniki biex jiksibu t-tali għan. Dawn ma jiddefinixxux l-għanijiet tagħhom stess. Madankollu, ċerti sistemi tal-AI (bħal dawk ibbażati fuq ċerti tekniki ta' tagħlim awtomatiku) jista' jkollhom aktar libertà biex jiddeċiedu liema perkors jieħdu biex jiksibu l-għan partikolari.

4. Definizzjoni aġġornata tal-AI

Aħna nipproponu li tintuża d-definizzjoni li ġejja tal-AI:

“Sistemi tal-intelliġenza artifiċjali (AI) huma sistemi ta' softwer (u possibbilment ħardwer ukoll) imfassla mill-bniedem³ li, wara li jingħataw għan kumpless, jaġixxu fid-dimensjoni fiżika jew diġitali billi jipperċepixxu l-ambjent tagħhom permezz ta' akkwizizzjoni tad-*data*, interpretrazzjoni ta' *data* strutturata jew mhux strutturata, raġunar dwar l-għarfien, jew proċessar tal-informazzjoni, miksuba minn din id-*data* u teħid ta' deċiżjoni dwar l-aħjar azzjoni(jiet) li għandha/għandhom tittieħed/jittieħdu biex jinkiseb l-għan partikolari. Is-sistemi tal-AI jistgħu jew jużaw regoli simboliċi jew inkella jitgħallmu mudell numeriku, u jistgħu wkoll jadattaw l-imġiba tagħhom billi janalizzaw kif l-ambjent hu affettwat mill-azzjonijiet preċedenti tagħhom.

Bħala dixxiplina xjentifika, l-AI tinkludi diversi approċċi u tekniki, bħat-tagħlim awtomatiku (li fosthom l-apprendiment approfondit u t-tagħlim tat-tisħiħ huma eżempji speċifiċi), ir-raġunar awtomatiku (li jinkludi l-ippjanar, l-iskedar, ir-rappreżentazzjoni u r-raġunar tal-għarfien, it-tiftix, u l-ottimizzazzjoni), u r-robotika (li tinkludi l-kontroll, il-perċezzjoni, is-sensuri u l-attwaturi, kif ukoll l-integrazzjoni tat-tekniki l-oħra kollha f'sistemi ċiberfiżiċi).”

u li jsir riferiment għal dan id-dokument bħala sors ta' informazzjoni addizzjonali b'sostenn ta' din id-definizzjoni.

³ Il-bniedem jiddisinja s-sistemi tal-AI b'mod dirett, iżda jista' juża wkoll tekniki tal-AI biex jottima d-disinn tiegħu.

Dan id-Dokument thejja mill-membri tal-Grupp ta' Esperti ta' Livell Għoli dwar I-AI

elenkati hawn taħt f'ordni alfabetiku

Pekka Ala-Pietilä, President tal-HLEG dwar I-AI AI Finland, Huhtamaki, Sanoma	Pierre Lucas Orgalim – Europe's technology industries
Wilhelm Bauer Fraunhofer	Ieva Martinkenaite Telenor
Urs Bergmann Zalando	Thomas Metzinger JGU Mainz u l-Assoċjazzjoni Ewropea tal-Universitajiet
Mária Bielíková L-Università Slovakkja tat-Teknoloġija fi Bratislava	Catelijne Muller ALLAI Netherlands u EESC
Cecilia Bonefeld-Dahl DigitalEurope	Markus Noga SAP
Yann Bonnet ANSSI	Barry O'Sullivan, Viċi President tal-HLEG dwar I-AI University College Cork
Loubna Bouarfa OKRA	Ursula Pahl BEUC
Stéphan Brunessaux Airbus	Nicolas Petit L-Università ta' Liège
Raja Chatila L-Inizjattiva tal-IEEE Ethics of Intelligent/Autonomous Systems u L-Università ta' Sorbonne	Christoph Peylo Bosch
Mark Coeckelbergh L-Università ta' Vjenna	Iris Plöger BDI
Virginia Dignum L-Università ta' Umea	Stefano Quintarelli Garden Ventures
Luciano Floridi University of Oxford	Andrea Renda Il-Fakultà tal-College of Europe u CEPS
Jean-Francois Gagné Element AI	Francesca Rossi* IBM - Ġestjoni Integrata tal-Fruntieri
Chiara Giovannini ANEC	Cristina San José Il-Federazzjoni Bankarja Ewropea
Joanna Goodey L-Aġenzija għad-Drittijiet Fundamentali	George Sharkov Digital SME Alliance
Sami Haddadin Munich School of Robotics and MI	Philipp Slusallek Iċ-Ċentru Ġermaniż tar-Riċerka dwar I-AI (DFKI)
Gry Hasselbalch It-thinkdotank DataEthics u L-Università ta' Copenhagen	Françoise Soulié Fogelman Konsulent tal-AI
Fredrik Heintz L-Università ta' Linköping	Saskia Steinacker Bayer
Fanny Hidvegi Access Now	Jaan Tallinn Ambient Sound Investment
Eric Hilgendorf L-Università ta' Würzburg	Thierry Tingaud STMicroelectronics
Klaus Höckner Hilfsgemeinschaft der Blinden und Sehschwachen	Jakob Uszkoreit Google
Mari-Noëlle Jégo-Laveissière Orange	Aimee Van Wynsberghe TU Delft
Leo Kärkkäinen Nokia Bell Labs	Thiébaut Weber KETU
Sabine Theresia Köszegi TU Wien	Cecile Wendling AXA
Robert Kroplewski Avukat u Konsulent għall-Gvern Pollakk	Karen Yeung The University of Birmingham
Elisabeth Ling RELX	

*Francesca Rossi aġixxiet bħala relatur għal dan id-dokument.