
Intel·ligència artificial: automatització

PID_00246034

Ferran Adell Español

Índex

1. Concepte: què és la intel·ligència artificial?	5
1.1. Definició	5
1.2. Abast	6
2. Implicacions tecnològiques de l'IA	8
2.1. Categories d'intel·ligència artificial	8
2.2. Robòtica	9
2.3. Bots (Internet) i assistents personals intel·ligents	11
2.4. <i>Machine learning</i> : l'únic camí cap a l'autonomia real	15
3. Implicacions socials de la intel·ligència artificial	19
3.1. Relació home-màquina (HMI) (òptica social)	19
Bibliografia	23

1. Concepte: què és la intel·ligència artificial?

1.1. Definició

La intel·ligència artificial és un concepte clarament de moda, quelcom que podem trobar associat a tot tipus de tecnologies, dispositius i serveis diferents. Ja no es troba només vinculada amb la robòtica, sinó que és un concepte genèric que s'ha universalitzat per a definir el pas cap a un entorn digital més autònom, capaç, *intel·ligent*. Vegem, a continuació, la definició que la *wikipedia* fa de la intel·ligència artificial:

«La intel·ligència artificial (IA) és una intel·ligència exhibida per màquines. En informàtica, una màquina "intel·ligent" ideal és un agent racional flexible que percep el seu entorn i adopta accions que maximitzen la seva probabilitat d'èxit en algun objectiu. Col·loquialment, el terme "intel·ligència artificial" s'aplica quan una màquina imita funcions "cognitives" que els éssers humans associen a altres ments humanes, com "aprendre" i "resoldre problemes".»

Wikipedia

Per tant, fem servir aquest concepte per a referir-nos a tot allò que les màquines fan com a imitació del comportament humà, de la nostra forma de pensar. Accions i reaccions que, amb major o menor complexitat, pretenen simular amb al major grau d'exactitud les accions humanes, tant en el procediment com en el resultat. I això abasta situacions tan heterogènies com la classificació o processament de la informació (gestió de calendaris o del correu-e) fins a tasques tan complexes com la conducció d'un vehicle per la xarxa viària. No hem d'associar l'IA només a aquelles activitats que, per la seva complexitat, requereixen un processament molt complex, una programació complexa, etc. sinó que aquest concepte s'usa per a referir-se a qualsevol comportament tecnològic que automatitzi processos humans que, normalment, depenen de la nostra intel·ligència.

Evidentment, amb un concepte tan ampli i que abasta espectres tan diferents de la tecnologia, tindrem categories i àmbits diferenciats en l'IA que ens ajudaran a endreçar millor el concepte i entendre l'abast i les implicacions tecnològiques d'aquesta automatització avançada del digital.

Considerem que el comportament d'un dispositiu o programari digital és intel·ligent quan és capaç de desenvolupar determinades tasques similars a la nostra cognició i, sobretot, quan és capaç d'adaptar el seu comportament a determinades variables de l'entorn, situacions concretes, etc. i anar més enllà de la simple execució d'instruccions prefabricades.

1.2. Abast

La frontera, però, no és gens fàcil de definir. Moltes vegades quan un procediment es torna habitual, normal, deixem de considerar-lo intel·ligent per a sumar-lo al conjunt de funcionalitats tecnològiques habituals. Quan enviem un correu electrònic, considerem aquesta acció normal; si en fer-ho hi afegim una cita en el cos del text, el calendari que tenim vinculat al correu-e ens afegirà aquesta cita com a tasca pendent, i aleshores considerem que aquest comportament és intel·ligent. Intel·ligent perquè fa una tasca que no hem demanat concretament però que ens serà útil en la majoria dels casos i que implica un reconeixement actiu del text del correu per a destriar en el cos del text aquella informació que és pertinent, en aquest cas una cita amb una data, una ubicació, etc. Si, a més a més, és un correu per a diferents destinataris, molt probablement, l'assistent també sabrà qui són els contactes que hi estan convidats. La suma de funcionalitats i l'entrecreuament de dades fan que molt sovint percebem l'IA allà on només hi ha un processament entrecreuat de dades. Ho comentarem més endavant.

Aquí s'esdevé una fractura entre el que en els cercles científics i tecnològics es considera IA i el que la societat, amb més tendència a flexibilitzar els conceptes, accepta com a propi. Els telèfons intel·ligents, rellotges intel·ligents, etc. realment són exemples de l'IA? Doncs depèn; probablement si analitzem totes les funcions que realitzen i la capacitat que tenim de combinar funcionalitats entre aplicacions, sensors, etc. i condicionar el comportament del dispositiu a diferents paràmetres podrem afirmar, sense por d'equivocar-nos que sí; però també veurem, ràpidament, que és una IA diferent de la qual s'amaga darrere de la tecnologia per a la conducció autònoma o de la gestió i control automàtic d'una fàbrica o un magatzem.

En realitat, a nosaltres, com a creadors digitals, ens interessen les dues accepcions del terme, perquè l'IA des del punt de vista de la creació abasta tot l'espectre del concepte. Des de l'automatització senzilla de processos, mitjançant aplicacions i serveis, fins al disseny i el plantejament de la resolució de tasques molt més complexes. Des dels petits ginys, aplicacions, algorismes i funcionalitats fins als grans reptes que implica la construcció d'un dispositiu autònom, amb capacitat d'aprenentatge, etc.

L'expansió de l'IA al nostre entorn no s'està esdevenint en forma de robots andròides que ocupen l'espai social, malgrat que cada vegada tinguem més dispositius d'aquest tipus a prop, és més una revolució invisible que es dona al mateix programari i als mateixos dispositius que ja fa temps que usem. No només del *hardware* depèn l'IA, no només de la capacitat de les màquines per a accedir als sentits humans: vista, tacte, etc. sinó que la major revolució des del punt de vista global està en el programari i les noves funcionalitats que, amb el processament de les dades, la implementació de nous sensors i l'augment de la capacitat de computació, som capaços d'implementar a l'entorn habitual d'ús dels usuaris. Si hom planteja a la societat que, d'un dia per l'altre, canviï

els seus hàbits de desplaçament, deixi el cotxe particular per a pujar en un d'autònom i comunitari es trobarà ràpidament amb una oposició social. Cal un procés d'acceptació, comprensió i adaptació d'aquest fenomen. Però si, per altra banda, hom planteja a la societat que pot fer el mateix que feia ahir, amb la mateixa aplicació de correu, però de forma molt més efectiva, automatitzada, intel·ligent, l'acceptació acostuma a ser massiva i ràpida.

2. Implicacions tecnològiques de l'IA

2.1. Categories d'intel·ligència artificial

Stuart Russel i Peter Norvig van definir la intel·ligència artificial a *Un Enfoque Moderno* en diferents tipus d'IA que podem usar com a categories conceptuals per a explicar els diferents comportaments intel·ligents de la tecnologia:

- **Sistemes que pensen com humans.** Aquests sistemes tracten d'emular el pensament humà; per exemple, les xarxes neuronals artificials. Automatització d'activitats que vinculem amb processos de pensament humà, activitats com la presa de decisions, resolució de problemes i aprenentatge.
- **Sistemes que actuen com humans.** Aquests sistemes tracten d'actuar com humans; és a dir, imiten el comportament humà; per exemple, la robòtica. Estudi de com aconseguir que les computadores realitzin tasques que, de moment, fan millor els humans.
- **Sistemes que pensen racionalment.** És a dir, amb lògica (idealment), tracten d'imitar o emular el pensament lògic racional de l'ésser humà; per exemple, els sistemes experts. Estudi dels càlculs que fan possible percebre, raonar i actuar.
- **Sistemes que actuen racionalment (idealment).** Tracten d'emular de forma racional el comportament humà; per exemple, els agents intel·ligents. Estan relacionats amb conductes intel·ligents en artefactes.

Wikipedia

Tipologies que recullen diferents formes que la tecnologia té per a imitar el comportament humà o la nostra forma de pensar, processar l'entorn i actuar en conseqüència, etc. Però, ara per ara, hem de tenir clar que es tracta de processos d'imitació i simulació que disten molt del funcionament real de la nostra intel·ligència; és cert que amb el *machine learning* i els sistemes capaços de prendre decisions de forma autònoma la simulació que l'IA fa de la intel·ligència humana és cada cop més potent, però no deixa de ser una simulació d'una característica o comportaments humans concrets, com a màxim una suma.

Des del punt de vista creatiu, el fet que la intel·ligència artificial s'escolli en tots els espais en què hi ha dispositius digitals i que la manera de fer-ho es basi en la imitació de determinats comportaments o característiques humanes per separat, fa que les potencialitats creatives associades a l'IA siguin amplíssimes.

No necessitem desenvolupar robots intel·ligents, amb forma hominoide, per a desenvolupar i fer propostes que usin l'IA per a resoldre problemes. Hem d'entendre el concepte com a quelcom que se suma a altres procediments informàtics, dispositius, funcions estàndards per a esprémer més la tecnologia ampliant la capacitat d'automatització de processos que és la base del digital.

Podem aplicar la intel·ligència artificial per a la classificació i gestió de la informació, per a gestionar de forma més efectiva el correu electrònic, per a fer les cerques a Internet més eficients i més properes al llenguatge natural, per a gestionar qualsevol tipus de maquinària o controlar la majoria de dispositius digitals. La recollecció massiva de dades i la capacitat de processament d'aquestes establint patrons de comportament, funcionalitats, fa que puguem aplicar comportaments intel·ligents a la majoria d'objectes digitals. A poc a poc, a mesura que les tecnologies associades a l'IA siguin més eficients i més populars, la implementació d'aquesta serà massiva, provocant un canvi difícil d'imaginar en el funcionament general de l'entorn digital.

Però sigui quin sigui el canvi, de ben segur, implicarà una extensió i una participació massiva de l'IA en aquest procés d'hibridació dels mitjans i les tecnologies que els formen cap a un entorn digital en què les fronteres desapareixeran ràpidament.

2.2. Robòtica

Un dels resultats més impactant de l'ús de l'IA és la robòtica. En primer lloc, perquè en l'imaginari popular, aquell que comentàvem que es basa en les històries de la ciència ficció, els robots tenen un paper protagonista com a mostres de l'avenç tecnològic i de l'augment de l'autonomia i la capacitat de les màquines. En segon lloc, perquè són una mostra visible de l'avenç de l'IA, una mostra que podem copsar de forma molt més efectiva i vivencial que aquella que s'escola darrere de les aplicacions informàtiques o el panell de control del nostre vehicle privat.

La robòtica, però, de nou no només engloba aquelles propostes que preten desenvolupar andròides de forma humana que imiten el nostre comportament, la nostra manera de caminar, parlar, desplaçar-nos, etc. sinó que és un camp tecnològic molt més ampli. La robòtica és un camp relacionat amb diferents enginyeries que s'encarrega de desenvolupar sistemes autònoms, generalment controlats per un programa informàtic, que desenvolupen una tasca concreta. Trobem robots que s'ocupen de tasques molt diverses i que tenen formes i comportaments molt heterogenis.

Vegem a continuació alguns dels camps en què els robots han ocupat un espai fins ara només humà:

a) Entorn industrial: Un dels àmbits en què hi ha hagut una penetració més significativa dels robots és el de la fabricació industrial. Entorns en què fa uns anys només hi trobàvem éssers humans ara han esdevingut espais compartits amb grans robots que realitzen tasques de diferent complexitat substituint la mà d'obra humana. Això, evidentment, té grans avantatges des del punt de vista de l'eficiència, i també de la seguretat dels treballadors, però també provoca un problema social donada la pèrdua de llocs de treball que cal reciclar cap a les noves professions industrials. Així, trobem robots que s'ocupen de les cadenes de muntatge de gran varietat de productes o que gestionen grans magatzems apilonant caixes de forma ordenada. A l'entorn industrial podem associar-hi també tot el que fa referència a la gestió de residus perillosos, el treball de mineria, etc.

b) Medicina i cures pal·liatives: Un altre dels sectors que ha viscut una clara revolució amb la incorporació de la robòtica és el camp de la medicina. Des de ja fa alguns anys s'ha avançat moltíssim en el desenvolupament de robots capaços d'assistir els cirurgians en les operacions, o intervenir de forma molt més acurada que els éssers humans per a realitzar determinades tasques de microcirurgia. La ginecologia, la cirurgia pediàtrica, toràcica o cardíaca són alguns dels camps especialment beneficiats per la incorporació de robots.

Per altra banda, la capacitat dels laboratoris mèdics i de la realització de proves, tests, etc. també s'ha vist fortament potenciada per la utilització de robots capaços de fer part de la feina de forma autònoma.

c) Defensa (militar): Un dels principals motors de la tecnologia ha estat sempre el desenvolupament militar. Donada l'alta inversió en defensa de la majoria de països occidentals i la importància d'aquest camp per a la geopolítica mundial, s'han fet grans inversions per a portar la robòtica al camp militar. Així, des dels famosos drons militars, petits vehicles (aeris o terrestres) capaços de desenvolupar tasques de forma autònoma o controlats remotament, fins al control dels vehicles militars més moderns: avions, vaixells, submarins, etc. La robòtica té una forta presència en la configuració actual dels exèrcits contemporanis. La necessitat d'estar sempre a l'avantguarda, fa que la majoria de tecnologies de nou descobriment s'implantin primer en el terreny militar. La idea de la majoria d'estats a mitjà termini és substituir la mà d'obra humana, soldats, tècnics, pilots, etc. per sistemes autònoms digitals.

d) Transport, rescat, gestió de residus, vigilància: Com comentàvem anteriorment, en els entorns de risc en què els éssers humans han de dur a terme determinades tasques perilloses són especialment útils els dispositius robòtics per a objectius tan importants com el rescat de persones en situacions en què és impossible enviar agents humans, la manipulació de determinats productes nocius, residus, explosius, etc. que comporten un risc en el seu trasllat, activació, transformació, etc. En aquest cas, hem agrupat aquests diferents camps perquè l'actuació robòtica és molt similar: elements autònoms o controlats per control remot que tenen una mobilitat àmplia, sia terrestre, aquàtica o aèria,

i una gran capacitat d'interacció amb objectes i persones, sia per a manipular materials o per a rescatar éssers humans en situacions de perill. A poc a poc, tots els camps en què és possible incorporar aquests tipus de dispositius van implementant l'ajuda robòtica per a reduir el perill al qual se sotmetien els agents humans implicats: operaris, bombers, policies, etc.

Evidentment, aquests només són alguns exemples del gran nombre d'entorns professionals i socials en què els robots s'estan transformant en elements imprescindibles per a dur a terme tasques cada cop més complexes i especialitzades. La incorporació de la robòtica als diferents entorns humans representa un altre espai d'oportunitats creatives d'enorme significació. No només cal pensar la implementació d'aquests dispositius en entorns en què encara no hi tenen presència, sinó també noves possibilitats per a la implementació de l'ús d'elements autònoms digitals per a la realització de tasques abans només humanes. Des de l'àmbit personal i domèstic fins a la cirurgia avançada, es nodriran d'elements robòtics que milloraran significativament l'eficiència mostrada fins ara pels humans.

2.3. Bots (Internet) i assistents personals intel·ligents

Comentàvem a la presentació d'aquest mòdul, i ja anteriorment, que una de les característiques de l'entorn digital actual que més ha trencat amb la proposta imaginativa de la ciència ficció és l'expansió silenciosa dels mitjans digitals. Aquella que no podem veure a primer cop d'ull, que no percebem d'una forma intuïtiva. En el cas de la intel·ligència artificial, com és de suposar, vivim també una expansió silenciosa de tecnologies intel·ligents que comencen a ser responsables de determinades gestions i accions fins ara impossibles per a ser portades a terme pels humans. Aquesta expansió silenciosa abasta des de petites millores en aplicacions d'ús habitual en els dispositius mòbils: la gestió del correu-e, del calendari d'esdeveniments, dels planificadors de tasques, de la gestió i classificació del fons fotogràfic personal i un llarg etcètera en constant creixement, fins a la incorporació de noves funcionalitats i capacitats digitals impossibles sense les tecnologies associades a l'IA.

Un dels camps en què, de forma més significativa, s'han implementat aquests tipus de noves funcionalitats intel·ligents és el dels bots informàtics i els assistents digitals personals. Programari capaç d'automatitzar tasques, dur a terme accions complexes i aprendre de la interacció amb l'usuari.

a) Bots (Internet)

El terme bot és una abreviació de robot i es refereix a aplicacions informàtiques capaces d'executar tasques de forma automàtica, generalment *scripts*. En el seu origen, la implementació d'aquestes aplicacions es feia amb la intenció de substituir els humans en la lectura, anàlisi i processament d'informació d'una forma molt més eficient i ràpida. Tasques, però, senzilles i de baixa complexitat algorítmica i conceptual. En aquest cas no tenim l'objecte robot, físic, que es pot desplaçar per l'entorn empíric o interactuar amb els objectes. No hi ha un cos, ni un braç mecànic o una cadena de muntatge robotitzada, parlem de programari, codi, que realitza accions de forma autònoma en funció d'un *input* determinat.

Durant molt de temps el més habitual era trobar bots que feien tasques senzilles, com revisar arxius, analitzar dades o indexar el nou contingut publicat a la xarxa. Un dels clars exemples de l'ús d'aquests tipus de bots és Google que revisa, amb l'ajuda de diferents programes automatitzats, la informació que es publica contínuament a la xarxa amb la intenció de processar-la i classificar-la, perquè els usuaris la puguem recuperar via cercadors. Aquest tipus de bots es coneixen com a web crawlers o spiders.

Però, amb el pas dels anys i la popularització dels mitjans digitals, la consolidació de la web 2.0 i l'expansió massiva de les xarxes socials, s'han popularitzat noves categories de bots informàtics que ocupen cada cop més l'entorn digital. Noves categories i una enorme complexitat necessària per a poder desenvolupar tasques més intel·ligents, més d'acord amb el que pretenem que l'IA sigui: una automatització de processos condicionada a l'entorn, amb capacitat de presa de decisions, etc.

Una de les categories en què és més fàcil de veure la implantació dels bots a simple vista és la dels *chatterbots*, sistemes informàtics als quals un pot plantejar preguntes amb el llenguatge natural i que són la base dels assistents personals intel·ligents, dels quals parlarem posteriorment.

- Quin temps farà demà?
- Quina és la distància fins a Madrid?
- Qui és el president de Rússia?

Però també càlculs senzills, conversió de moneda o unitats, consultes sobre resultats esportius o financers, etc. Són algunes de les preguntes que podem plantejar a aquest tipus d'aplicacions, que esdevenen vies potencials de desenvolupament de la interacció de l'home amb els dispositius digitals, impulsant una relació més propera al nostre llenguatge natural i que podem usar de forma molt més intuïtiva.

Un altre àmbit en què aquest tipus de tecnologies són cada cop més presents és tot el sector relacionat amb la gestió de la borsa i la compra i venda ràpida de valors. Aquí els bots s'encarreguen d'engegar accions massives de compra-venda de valors per a l'obtenció de beneficis en funció de determinats paràmetres. Un objectiu comercial que està generant una certa controvèrsia per la poca regulació que tenim per a controlar el seu comportament i, sobretot, gestionar l'alt poder que tenen per a alterar el mercat de valors. Altres activitats econòmiques, basades en l'automatització de funcions, també s'estan beneficiant clarament de la capacitat dels bots per a realitzar tasques de forma automàtica i molt més eficient i ràpida que els éssers humans: l'escrutini de la xarxa buscant ofertes immobiliàries, la revisió contínua de portals de compra i subhastes per a trobar grans oportunitats abans que els usuaris humans, etc. són alguns dels exemples d'aquest ús per a activitats econòmiques dels bots informàtics.

El sector del màrqueting digital és un dels altres sectors importants pel que fa a la utilització de bots. En aquest cas les funcionalitats van des de l'automatització de la detecció i seguiment de perfils de xarxes socials, blogs, etc. que estiguin relacionats amb determinades paraules clau, per a aconseguir atraure l'atenció dels usuaris cap a l'origen del bot, fins a la publicació automàtica de comentaris, referències o notícies sense la necessitat de la intervenció humana. En aquest últim cas, una pràctica que s'ha de repensar contínuament a si mateixa perquè els sistemes de detecció d'*spam* cada cop són més eficients i dificulten aquestes pràctiques de forma significativa. Igualment, cada cop que tenim a disposició dades sobre la quantitat de la informació que corre per la xarxa que és produïda per bots, els resultats fan esfereir. Les pràctiques comercials i les del màrqueting digital són en si focus de controvèrsia perquè afecten directament els drets i les llibertats de l'usuari, en aquest cas, client potencial.

I aquests tipus de pràctiques abusives ens porten a l'últim grup de bots que voldríem analitzar, l'ús dels bots amb intencions delictives, malvades, d'abús, etc., encara que estem deixant de banda altres categories amb menys afectació global. Aquest últim grup es refereix a aquells bots que s'usen per a aconseguir objectius il·lícits segons les lleis o segons l'ètica social. En aquesta categoria, una de les més protagonistes avui, s'hi engloben des de les eines que diferents *hackers*, independentment del color del seu barret, fan servir per a controlar diferents ordinadors a distància, fins a les eines que s'usen per a augmentar de forma artificial les visualitzacions d'un vídeo a qualsevol plataforma en línia, tenir accés a informació personal o escrutar la xarxa en recerca de contingut que un pugui copiar per a reaprofitar sense el consentiment del seu autor.

Hi ha multitud d'exemples d'aquest ús malintencionat dels bots informàtics que posen de manifest la necessitat ineludible de regular el sector i les possibilitats d'ús d'aquests tipus de tecnologies per a garantir el respecte a les lleis i els drets dels ciutadans. Des de problemàtiques tan importants com els famosos atacs de denegació de servei (DDoS), en què hom envia peticions massives a un servidor fins a saturar-lo, fins a l'alteració contínua dels jocs de rol en línia

en què s'usen bots per a obtenir recursos, millores dels personatges, etc. trencant l'estabilitat del sistema. L'ús indegut de bots informàtics està esdevenint una problemàtica social d'una importància definitiva.

b) Assistents personals intel·ligents

Els assistents personals estan en un veritable moment d'expansió. Les grans companyies digitals: Google, Amazon, Apple, Facebook, etc. han vist en aquest tipus d'aplicacions una via amb grans possibilitats per a incorporar noves funcionalitats per a l'obtenció i processament de dades de l'usuari. Els assistents personals són sistemes que permeten automatitzar diferents tasques, que duem a terme amb eines digitals de manera que esdevinguin instruccions multiaplicació, és a dir, que siguin capaces de gestionar no només una aplicació particular, sinó d'establir comunicacions, peticions i respostes, entre diferents aplicacions instal·lades en un sistema operatiu o residents al núvol. Els assistents personals intel·ligents són capaços de llegir un correu electrònic que hem rebut al nostre compte personal, processar el contingut per a aïllar aquella informació que és sensible, més important, i portar a terme accions sobre aquesta informació. Un exemple molt clar és quan l'assistent és capaç d'afegir, automàticament, una cita al calendari en funció d'una convocatòria rebuda per correu-e com comentàvem anteriorment. Alhora, el mateix assistent és capaç de buscar i calcular amb un sistema de geolocalització la via més efectiva i ràpida per a arribar al destí de la nostra cita, tenint en compte el trànsit, el temps, etc.

Les possibilitats són immenses si tenim en compte el que hem comentat en el mòdul de l'IoT i també sobre el *big data*. Si tenim accés a un gran conjunt de sensors i fonts de dades digitalitzats i capacitat d'emmagatzemar i processar aquesta informació, les potencialitats d'automatització intel·ligent es multipliquen perquè som capaços de definir patrons, condicionals de comportament i possibilitats de resposta. Per altra banda, el gran conjunt de les API disponibles avui fa que la comunicació entre aplicacions i serveis sigui cada cop més eficient, senzilla i fàcil d'implementar, donant encara més força a aquest procés de treball col·laboratiu entre aplicacions i serveis digitals.

Els assistents personals són un concepte que ens acompanyarà per sempre a l'entorn digital, en diferents formes i amb diferents capacitats, però són la resposta a la pregunta: què pot fer la tecnologia digital per mi, més enllà de les tasques habituals i normals que ja estem acostumats a realitzar?

Si som capaços de definir un entorn de dades i processament digital que permeti la comunicació i la compartició d'informació entre serveis, concertar una cita amb un amic, en el moment en què un passa per una exposició que creu que pot interessar a ambdós, podrà ésser una tasca automàtica. El sistema cercarà la nostra agenda per a trobar un espai lliure, i farà el mateix amb la del nostre contacte, buscarà una coincidència i notificarà a les dues persones interessades la proposta perquè la puguin confirmar. En rebre aquesta confir-

mació afegirà la cita a l'agenda donant-los la ubicació, la forma més ràpida d'arribar i l'hora a la qual hem de sortir per a no fer tard, etc. Més enllà de la captura d'una cita puntual al correu, el sistema serà capaç d'entrecreuar diferents funcionalitats i la de diferents usuaris, per a definir un patró de coincidència que, a més a més, és fàcil d'ampliar. Si la cita requereix alguna preparació prèvia, per exemple una activitat esportiva per a la qual puguem necessitar un material concret, el sistema ens oferirà preus, recomanacions sobre els millors productes, etc. perquè puguem comprar-los de forma senzilla, tenint sempre en compte el termini d'entrega en funció de la data acordada per la cita. Podríem estirar l'exemple afegint cada vegada més funcionalitats en la suma d'aplicacions i dispositius digitals cada cop més capaços de comunicar-se entre si i intercanviar dades en temps real i en formats universals.

2.4. *Machine learning*: l'únic camí cap a l'autonomia real

El *machine learning*, terme també de moda els últims anys, és una branca de l'IA que es dedica a proposar algorismes, tècniques i aplicacions que aprenen amb l'experiència, condició indispensable per a poder atribuir l'adjectiu intel·ligent a qualsevol objecte. En l'actualitat, un dels grans reptes de l'IA és desenvolupar objectes digitals capaços d'analitzar l'entorn (respostes dels usuaris, dades, estadístiques, etc.) i definir patrons per a l'aprenentatge automàtic, amb la intenció de millorar l'efectivitat i ampliar les capacitats del sistema.

Els anys 60 Artur Samuel, un dels pioners de l'IA, ja definia el *machine learning* com el fet de donar als ordinadors la capacitat d'aprendre sense estar específicament programats per a fer-ho. Aquest procés, molt proper a l'estadística computacional, l'optimització matemàtica, etc., és un procés molt complex que implica diferents disciplines i procediments tècnics i que no definirem detalladament. No obstant això, sí que és important que analitzem les característiques més importants d'aquest aprenentatge automàtic, així com el seu paper fonamental en el desenvolupament d'una IA cada cop més intel·ligent.

En els inicis de la disciplina, l'objectiu central d'aquesta era potenciar i desenvolupar una IA cada cop més potent i capaç. Poc a poc, però, també fruit de l'avenç de l'IA com a disciplina en si mateixa, el *machine learning* s'ha concentrat més en la resolució de problemes pràctics i la proposta de solucions d'ús i implantació immediat. No obstant això, l'aprenentatge de les màquines com a concepte, és un element central, una condició de possibilitat per al desenvolupament d'objectes digitals cada cop més eficients, capaços i més propers a la nostra forma resoldre els problemes.

Podem establir dues grans categories de *machine learning*: per una banda, aquells processos que són supervisats per un instructor humà; per l'altra, aquells sistemes dissenyats per a aprendre de forma autònoma basant-se en algorismes que cerquen i defineixen patrons per la seva banda. Però, sigui quina sigui la categoria, sempre ens trobarem davant d'un procés que beu considerablement del fenomen del *big data* que hem presentat anteriorment. Són els

grans conjunts de dades recollides de sensors diversos, de la xarxa, dels continguts generats pels usuaris, etc., base sobre la qual es construeix el *machine learning*. Per a ser capaços de definir patrons i respostes adequades és necessari processar de forma contínua grans volums de dades que permeten generar models, detectar patrons, etc. Precisament, una de les grans problemàtiques que fa que les expectatives dipositades en el *machine learning* siguin difícils de realitzar és que encara no tenim prou dades de tot tipus per a permetre aquest aprenentatge automàtic.

Però quins són els usos pràctics més habituals del *machine learning*?

Hem plantejat que l'objectiu central del *machine learning* és definir sistemes capaços d'aprendre de forma autònoma, millorant amb l'experiència i essent capaços de simular el procés de presa de decisions en funció dels paràmetres de l'entorn que fem els essers humans.

Però per què ens és això útil?

Les aplicacions del *machine learning* són cada cop més heterogènies i extenses. Des de la capacitat d'un sistema d'entendre les necessitats dels usuaris i adaptar-se a aquestes, per exemple en la navegació web en funció de determinats paràmetres del dispositiu, l'entorn, les preferències de l'usuari en altres serveis similars, l'experiència anterior com a visitant a la mateixa, fins a entorns complexos de predicció de les preferències de l'usuari com en els casos de les plataformes de compra com Amazon, o de distribució de continguts com Netflix, etc. Per a aconseguir aquests resultats s'apliquen algorismes complexos per a definir, en funció de determinats paràmetres, dades, quines són les respostes més adequades del sistema davant d'un *input*.

- Quins són els productes que podem relacionar per a oferir-los a l'usuari?
- Quines són les sèries o pel·lícules que, en funció de les preferències mostrades anteriorment, són especialment interessants per a un usuari en concret?

Però el *machine learning* no només beu de les preferències de l'usuari i del factor, sinó que si ens fixem en l'avenç dels sistemes d'assistència personal que presentàvem en el punt anterior, ens adonarem que el reconeixement de la parla, per exemple, o l'adaptació d'aquests sistemes a les preferències de l'usuari són una altra capacitat dependent de l'aprenentatge del sistema.

De nou, hi ha tot un conjunt de funcionalitats i procediments del *machine learning* que no són evidents per l'usuari, que no detectem però que són la base per a l'avenç de moltes disciplines científiques. L'estudi de l'ADN, l'aprenentatge aplicat a les dades mèdiques i a la presa de decisions o diagnòstics, són també entorns en què l'aprenentatge de les màquines, associat a l'anàlisi de les dades, són molt importants.

I, per últim, hi ha molta més informació d'aquest camp arreu de la xarxa, cal plantejar la relació del *machine learning* amb la robòtica, tal com la definíem anteriorment. Resulta evident que el pas dels robots cap a una intel·ligència més potent, més complexa, més àmplia, passa per la capacitat d'aquests de prendre decisions més enllà de les seves funcionalitats bàsiques per a les quals han estat programats. Robots capaços d'adaptar el seu comportament no només a determinats paràmetres predefinitos, sinó d'adaptar la resposta als canvis de l'entorn o a l'experiència pròpia. Un camí llarg encara però que ens permetrà desenvolupar sistemes més capaços d'entendre l'entorn humà i actuar tenint en compte més paràmetres que l'execució d'un codi de programació.

El *machine learning*, però, com ja hem anat veient amb tots els grans conceptes digitals analitzats al llarg d'aquests materials, pot ser també l'origen de problemàtiques socials i ètiques complexes si no actuem d'una forma preventiva. Això és així perquè nosaltres no prenem decisions basades només en dades o en patrons estrictes, adaptem les nostres decisions, des de les més personals a les més globals, a una ètica, uns valors i un objectiu social que és molt difícil de digitalitzar. La possibilitat que l'estudi i anàlisi de les dades generin un aprenentatge automàtic que no tingui en compte res més que l'eficiència digital és quelcom que representa un repte per a la societat. Hem de ser capaços de dissenyar sistemes que no només tinguin en compte els valors objectius de les dades sinó que prenguin decisions d'una forma més complexa i atenen a més variables i patrons, condicionant-los a un sistema de valors humà difícil de processar digitalment. Si analitzem alguns casos controvertits, com per exemple les decisions que ha de prendre un cotxe autònom a l'hora d'escollir el benefici de qui prioritza en el moment de prendre una decisió, ens adonarem que les respostes automàtiques, encara que es basin en patrons, no són quelcom que els humans vulguem acceptar tal qual.

Si tenim en compte el bé universal global i programem els dispositius en funció d'aquest bé, segurament, un vehicle que tingui l'opció d'evitar el dany d'un grup de vianants, a canvi de provocar-lo a un únic ocupant d'un vehicle, prendrà la decisió de danyar el conductor per a evitar un mal major. Però la decisió no és tan fàcil. En primer lloc, i si els vianants són els que estan caminant pel mig de la carretera sense tenir en compte les normes de circulació? I si aquest grup de vianants decideixen llençar-se contra el cotxe? Quina és la decisió correcta a prendre? I, encara més, si els ciutadans sabem que el nostre vehicle personal, malgrat que la "raó" estigui de la banda dels vianants, prendrà la decisió de danyar-nos a nosaltres, i a la nostra família, si considera que el dany global és menor, voldrem comprar i conduir aquest tipus de vehicle? La resposta serà clarament que no.

Les preferències dels éssers humans no es basen en el bé comú, ni en el càlcul objectiu de probabilitats, conseqüències, etc. La nostra sensibilitat i la nostra raó són quelcom molt més complex. Parlarem una mica més d'aquesta discordança entre el bé comú i les preferències personals, i com tot això afecta l'evolució del digital en el mòdul final de les conclusions.

3. Implicacions socials de la intel·ligència artificial

Arribats a aquest punt ja no té massa sentit intentar parlar sobre les implicacions de l'IA com a quelcom independent de la resta de grans conceptes de la tecnologia. Hem analitzat l'impacte de les tecnologies en general, de la hibridació dels mitjans, de l'expansió dels sensors i la connectivitat amb l'IoT, de la massificació de les dades amb el *big data* i la generació contínua de dades per part dels usuaris, etc. Però al llarg de tots els mòduls també hem anat veient que en l'actualitat digital, en el moment d'hibridació dels mitjans en què ens trobem, és molt difícil de parlar de les tecnologies d'una forma aïllada perquè les relacions de condicionament són omnipresents i les fronteres tangencials. La intel·ligència artificial és molt difícil d'entendre sense tenir en compte l'expansió de les dades o l'avenç de la Internet de les coses; de la mateixa manera, té poc sentit plantejar què podem fer amb les dades sense aplicar, per a processar-les, algorismes del *machine learning* (IA), etc. Per tant, ara que ja hem repassat la majoria dels conceptes centrals de les tecnologies digitals, hem d'entendre el punt següent com una anàlisi i un plantejament de la problemàtica associada a la relació entre els éssers humans i els dispositius digitals, en general, que recollirem a l'últim mòdul d'aquests materials.

3.1. Relació home-màquina (HMI) (òptica social)

a) Les lleis d'Asimov com a marc de referència?

Si revisem la ja no tan curta història dels mitjans digitals, ens adonarem que la velocitat a la qual avancen els nous descobriments ha provocat que la reflexió filosòfica i ètica respecte a l'impacte de la tecnologia sobre la societat es faci sempre a posteriori, sobre fets consumats. Lluny de manifestar por, o una certa desconfiança enfront les noves eines digitals, la societat, tot i desconèixer completament les implicacions socials, s'ha sumat al procés de digitalització global sense massa miraments a l'hora d'acceptar les condicions, suposades implícites, del canvi tecnològic.

Lluny d'astorar-se davant del que diversos filòsofs, teòrics i experts han presentat com els grans perills de la tecnologia, el grup humà ha tirat pel dret, fent-se propi l'ecosistema digital sense parar atenció a la forta capacitat d'aquest per a posar en entredit els valors socials. Si el canvi tecnològic no s'acompanya d'un marc de regulació unificat amb l'entorn físic i en què els drets de l'usuari ocupin la cúspide de la piràmide d'importància, i la pèrdua de drets deixi de ser considerada només un efecte innocu secundari, un dany col·lateral que no podem evitar, l'estat del benestar, entès d'una forma àmplia com un entorn que protegeix els drets bàsics dels ciutadans, serà molt difícil de mantenir.

Sense una reflexió ètica i filosòfica que assegurí que la tecnologia es desenvolupa tenint en compte el marc social en què s'ha d'esdevenir, és molt complicat poder garantir la continuïtat de la societat tal com la coneixem. Com hem comentat anteriorment, els escenaris de la ciència ficció dels anys 60 van dibuixar a l'imaginari popular un ecosistema digital molt més exagerat en aparença, molt més vistós, més present en el paisatge urbà (un imaginari que el cinema contemporani s'ha encarregat de potenciar), un entorn que s'allunya molt de com s'està esdevenint la implantació de les tecnologies digitals. Però això no vol dir que les tecnologies, els conceptes, les relacions home-màquina i les problemàtiques que la ciència ficció plantejava, no siguin ja un fet en la societat contemporània. Simplement, la forma és diferent, la seva presència és diferent i el procés d'implantació és més progressiu, però conceptes com intel·ligència artificial, biotecnologia, control mental dels dispositius, atribució de responsabilitats en les màquines, delegació de funcions, etc. són conceptes molt presents en l'ecosistema digital actual. Cal, però, que també siguin importants centres d'atenció en els debats socials.

Si recuperem els textos de W.Gibson, per exemple, ens adonarem que ja s'hi plantejaven la majoria de les grans problemàtiques a l'entorn de les quals ha de girar la reflexió filosòfica i ètica sobre les tecnologies digitals, especialment sobre la intel·ligència artificial. Els sensors, les xarxes neuronals, la normes per a la relació home-màquina que Asimov va plantejar, la convivència sempre complexa de la naturalitat i l'artificialitat, etc. són conceptes que ja es van plantejar com a centrals en el debat filosòfic digital fa dècades, però que no hem sabut prendre com a propis a l'entorn actual.

Un debat que ha de versar sobre l'atribució de responsabilitats a les màquines, i si això és un sense sentit, sobre la relació entre la intel·ligència humana i la intel·ligència artificial, sobre quanta autonomia podem donar als robots, quines decisions poden prendre sense la intervenció humana i, sobretot, potenciar i declarar obertament que el fet que tecnològicament "tot sigui possible" no vol dir que digitalitzar tots els aspectes de la nostra existència sigui desitjable.

No n'hi ha prou amb les lleis d'Asimov per a definir el marc que ha de regir les relacions home-màquina, necessitem una regulació legal que protegeixi els principis inqüestionables de la nostra espècie.

Lleis d'Asimov:

- Un robot no pot danyar un ésser humà o, per inacció, permetre que un ésser humà pateixi un dany.
- Un robot ha d'obeir les ordres donades pels éssers humans, excepte si aquestes ordres entressin en conflicte amb la 1a. llei.
- Un robot ha de protegir la seva pròpia existència en la mesura que aquesta protecció no entri en conflicte amb la 1a. llei o la 2a. llei.

J. Campbell; I. Asimov (1940)

A més a més, si fem un ràpid repàs a la situació tecnològica actual, ens adonarem que ni tan sols aquest esquema tan bàsic d'Asimov actualment s'està respectant, probablement perquè els estaments dirigits no hi estan d'acord, no se'l fan seu.

Per tant, no és una clara actuació contra la vida humana el paper d'un *drone* militar que du a terme una missió de matar? Però relativitzem aquesta actuació per dos motius evidents: en primer lloc, perquè realment no considerem que un *drone* sigui un robot intel·ligent en el sentit que puguem atribuir-li algun nivell de responsabilitat, sinó que la responsabilitat ètica és de l'exèrcit que el fa servir; i, en segon lloc, perquè tots entenem perfectament que un comandament militar prefereixi arriscar drons que soldats. No obstant això, tot i que ho puguem entendre, això no vol dir que no ho haguem de discutir, avaluar, regular, etc.

En realitat aquí el problema rau en l'autonomia del robot, en la seva programació, si un *drone* no pren decisions de forma autònoma, si només es limita a executar les ordres directes del seu pilot, aleshores l'ètica sobre l'IA aquí hi té poc a dir, no hi ha gaire diferència amb l'ús de qualsevol arma automàtica. Però què passa quan aquest *drone*, de forma autònoma, pot prendre decisions sobre si és el moment de disparar o no, sobre si l'objectiu és correcte, sobre si els danys "col·laterals" són acceptables, sobre el valor d'una vida humana al cap i a la fi... És aquesta autonomia digital a la qual hem de posar límits per a garantir l'ecosistema social de drets a l'entorn actual.

b) L'abast assumible de l'IA: "possible no és desitjable"

En l'actualitat, encara ens sembla alarmista plantejar un entorn tecnològic en què les decisions autònomes de les màquines tinguin un impacte real sobre l'esdevenir de l'espècie humana, però si parem atenció a la realitat tecnològica que hem plantejat a la primera part del mòdul ens adonarem que, en realitat, la situació actual és molt més complexa i té moltes més implicacions socials de les que podríem pensar. La problemàtica no rau només en robots de forma humanoide, que s'erigeixen com els soldats del futur, sinó que ja s'origina en els robots d'aspiració, tan de moda els últims anys, que són capaços de fer un mapa de la nostra llar i comunicar-se mitjançant la tecnologia wifi amb un servidor remot (alguns models). Si aquest robot, de poca autonomia i funcionalitats molt limitades, és capaç de recollir dades, establir patrons, etc. que, a posteriori, navegaran per la xarxa, ja ha de ser regulat per a garantir els drets del seu propietari. No només el dret a què el dispositiu sigui segur per al seu ús i que no destrueixi els nostres mobles, sinó que la comunicació que estableixi amb els servidors de l'empresa, la informació que reculli, etc., només sigui correctament protegida i recollida amb el consentiment de l'usuari. Si analitzem ginys com els altaveus *bluetooth* intel·ligents (Amazon Echo, per exemple) ens adonarem que el seu abast i la seva capacitat de processar, creuar i emmagatzemar informació sobre l'usuari és cada cop més potent i, per tant, sensible a atacs malintencionats, usos comercials no autoritzats, etc. I aquests sistemes

s'actualitzen contínuament implementant noves capacitats, noves funcionalitats, cosa que fa que el volum de dades sensibles estigui contínuament *in crescendo*.

La nostra percepció de les potencialitats tecnològiques està massa lligada a l'aparença física dels objectes, de la seva presència "real" com a objectes físics, però la tecnologia digital s'escola arreu en forma d'una connectivitat invisible que és difícil de percebre a primer cop d'ull. Cal, doncs, analitzar l'evolució de l'IA amb més atenció, analitzant no només les seves capacitats, sinó també les seves potencialitats. Allò que és i serà capaç de fer. No podem seguir percebent els avenços tecnològics com a un simple pas més cap a la millora de les nostres condicions de vida, cap a la simplificació de tasques, l'automatització, la millora de la comunicació, etc. Cal obrir la mirada i entendre quines tecnologies, per la seva capacitat de penetració social, signifiquen una posada en qüestió del nostre sistema de valors.

I amb això no volem dir que no s'hagi generat cap debat a l'entorn d'aquestes qüestions, que haguem de començar de zero, sinó que aquests s'han plantejat sempre a posteriori, quan els efectes de la tecnologia ja són molt més incontrolables i difícils de regular. Davant la transcendència del canvi que l'IA proposa, ara és el moment d'avançar la reflexió ètica i filosòfica respecte l'evolució de la tecnologia digital i la seva necessària relació amb la constitució de l'entorn social per a garantir una actuació preproblemàtica, una regulació que protegeixi els drets de l'usuari abans que aquests siguin posats en entredit, i no com a resposta, com a defensa, de la infracció d'un dret amb reiteració.

I per acabar amb aquesta segona part del mòdul, malgrat que aquest no és l'espai per a plantejar d'una forma detallada els debats i les reflexions necessàries per a atacar amb garanties l'evolució de l'entorn digital, sí que volem presentar algunes preguntes, com hem fet al llarg de tots els mòduls, sobre algunes de les reflexions més importants que com a creadors digitals ens hem de fer.

- Quins han de ser els límits en la relació home-màquina? Tot val? Podem deixar que els robots ocupin tots els espais de la nostra intimitat?
- I quines són les normes, no només legals, sinó de comportament que s'erigeixen com a indispensables per a garantir aquests límits?
- Podem fer servir l'IA per a ocupar espais que depenen i requereixen de la sensibilitat humana?
- Quin és el límit en l'autonomia de les màquines? Realment volem que aquestes puguin prendre decisions de forma autònoma? A quin nivell i en quins camps?
- Quin ha de ser el límit del cos humà? I el límit de la biotecnologia?

La resposta a aquestes preguntes, que recuperarem en el mòdul final per a valorar-les amb més detall, versarà sempre a l'entorn del plantejament que hem de fer sobre si tot allò que tecnològicament és o serà possible és acceptable des del punt de vista dels valors socials sobre els quals hem construït l'entorn humà.

Bibliografia

Rishal Hurbans (2016). *Artificial Intelligence in Perspective*. Medium. <https://medium.com/@rishal.hurbans/artificial-intelligence-in-perspective-6331dc384685>

Steven Levy (2017). *What deep blue tells us about AI in 2017*. Wired. <https://backchannel.com/what-deep-blue-tells-us-about-ai-in-2017-3284f92b2a93>

Deep learning

Michael Dietz (2016). *Deep adversarial learning is finally ready, and will radically change the game*. Medium. <https://medium.com/intuitionmachine/deep-adversarial-learning-is-finally-ready-and-will-radically-change-the-game-f0cfda7b91d3>

Bots

Matt Chessen (2016). *Artificial intelligence chatbots will overwhelm human speech online; the rise of MADCOMs*. Medium. <https://medium.com/artificial-intelligence-policy-laws-and-ethics/artificial-intelligence-chatbots-will-overwhelm-human-speech-online-the-rise-of-madcoms-e007818f31a1>

Ètica

Matt Chessen (2016). *Machines will soon program people*. Medium. <https://medium.com/@mattlesnake/machines-will-soon-program-people-73929e84c4c4>

Rand Hindi (2017). *How my research in AI put my dad out of a job*. Medium. <https://medium.com/snips-ai/how-my-research-in-ai-put-my-dad-out-of-a-job-1a4c80ede1b0>

