

ISSN: 2530-6561  
2026ko apirila  
Bilbao

hegoak zabalduz

Hegoa, Nazioarteko  
Lankidetzeta eta  
Garapenari Buruzko  
Ikasketa Institutua  
Euskal Herriko  
Unibertsitatea  
EHU

## Adimen artifiziala eta generoa

### ONDARE

Naiara Aginako Bengoa, Olatz Arbelaitz Gallego,  
Borja Calvo Molinos, Iñaki Inza Cano,  
Itziar Irigoien Garbizu, Elena Lazkano Ortega,  
Alex Mendiburu Alberro, Usue Mori Carrascal,  
Olatz Perez de Viñaspre Garralda, Ana Zelaia Jauregi



Nazioarteko Lankidetzeta eta  
Garapenari buruzko Ikasketa  
Institutua  
Instituto de Estudios sobre  
Desarrollo y Cooperación  
Internacional





## Adimen artifiziala eta generoa

---

### ONDARE

EHUko Informatika Fakultatean martxan jarri den ekimena da, adimen artifiziala gizartearen onurarako lelopean aritzen dena eta maila guztietan adimen artifizialaren garapen eta erabilera zuzena zabaldu nahi dituen.

---

Naiara Aginako Bengoa  
Olatz Arbelaitz Gallego  
Borja Calvo Molinos  
Iñaki Inza Cano  
Itziar Irigoien Garbizu  
Elena Lazkano Ortega  
Alex Mendiburu Alberro  
Usue Mori Carrascal  
Olatz Perez de Viñaspre Garralda  
Ana Zelaia Jauregi



## Aurkibidea

<b>1. Sarrera</b> .....	8
1.1. Historia .....	9
1.2. Emakumeen presentzia Adimen Artifizialaren arloko lanbideetan .....	10
<b>2. Genero-arrakala eta Adimen Artifiziala: adibideak</b> .....	12
2.1. Itzulpen automatikoa .....	12
2.2. Testutik irudiak sortzea .....	12
<b>3. Alborapenen jatorria</b> .....	14
<b>4. Alborapenean eragiteko moduak</b> .....	17
4.1. Datuak aztertzea .....	17
4.2. Ereduak aztertzea .....	18
4.3. Sistema ebaluatzea .....	19
<b>5. Robot sozialak: Adimen Artifizialaren izaera fisiko berezkoenak</b> .....	22
5.1. Morfologia eta forma, ahotsa barne .....	22
5.2. Algoritmoak eta datu-baseak .....	23
5.3. Robotaren portaerak, errektiboak nahiz pasiboak .....	24
<b>6. Ondorioak</b> .....	25
<b>7. Erreferentziak</b> .....	26

## Erredakzio kontseilua

Zuzendaritza: Irati Labaien

Amaia Garcia Azpuru

Kideak: Eduardo Bidaurratzaga

Xabier Gainza

Iker Etxano

Itziar Mujika

Unai Villena

Iker Zirion

---

**hegoak zabalduz** euskarazko material bildumak ikuspegi zabal batetik garapenarekin lotutako askotariko gaiak jorratzea du helburu. Testuak didaktikoak eta dibulgazio mailakoak dira, eta gaien inguruko sarrera edo ikuspegi orokorra ematen dute.

### Adimen artifiziala eta generoa

Egilea: ONDARE, EHUKo Informatika Fakultatean martxan jarri den ekimena da, adimen artifiziala gizartearen onurarako lelopean aritzen dena eta maila guztietan adimen artifizialaren garapen eta erabilera zuzena zabaldu nahi dituen.

Naiara Aginako Bengoa, Olatz Arbelaitz Gallego, Borja Calvo Molinos, Iñaki Inza Cano, Itziar Irigoien Garbizu, Elena Lazkano Ortega, Alex Mendiburu Alberro, Usue Mori Carrascal, Olatz Perez de Viñaspre Garralda, Ana Zelaia Jauregi

Lan hau, Gipuzkoako Foru aldundiak babestu du 2024ko Etorkizuna eraikiz deialdian ADIMEN ARTIFIZIALA GIZARTEAREN ONURARAKO (AI4SG) proiektuaren bidez.

**hegoak zabalduz** - 44. zk. 2026ko apirila

ISSN: 2530-6561



Nazioarteko Lankidetzeta eta  
Garapenari buruzko Ikasketen  
Instituta  
Instituto de Estudios sobre  
Desarrollo y Cooperación  
Internacional



[www.hegoa.ehu.eus](http://www.hegoa.ehu.eus)

EHU. Zubiria Etxea

Lehendakari Agirre Etorbidea, 81

48015 Bilbao

Tel.: 94 601 70 91

[hegoa@ehu.eus](mailto:hegoa@ehu.eus)

EHU. Carlos Santamaría Zentroa

Elhuyar Plaza, 2

20018 Donostia-San Sebastián

Tel. 943 01 74 64

[hegoa@ehu.eus](mailto:hegoa@ehu.eus)

EHU. Arabako Campuseko Liburutegia

138. posta-kutxa

Nieves Cano, 33

01006 Vitoria-Gasteiz

Tel. 945 01 42 87

[hegoa@ehu.eus](mailto:hegoa@ehu.eus)

Diseinua eta Maketazioa: Marra, s.l.



**hegoak zabalduz** material bilduman argitaratzen diren testu guztiak, Creative Commons en lizentzia honekin argitaratzen dira:

Aitortu-EzKomertziala-LanEratorririkGabe 4.0 Espainia.

Lizentzia osoa: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.eu>

Finantzatzailea:



**Gipuzkoako Foru Aldundia**  
Kultura, Turismo, Gazteria eta Kirol Departamentua  
Departamento de Cultura, Turismo, Juventud y Deportes

## Adimen artifiziala eta generoa

**Laburpena:** Historikoki, zientziaren eta teknologiaren garapena gizon heteronormatiboek (eta azal zuriko eta klase sozial ertain-altukoek) gidatu dute, eta garatutako aplikazioekalborapen nabarmenak izan dituzte. Bete-beteen gure egunerokoan sartu zaigun Adimen Artifiziala (AA), gaur egungo ikerkuntza-arloa izanda ere, ez dago kortse horretatik aske eta genero-arrakala du (baita beste hainbat ere).

Artikulu honetan agerian jarri nahi dugu hori, alborapena non eta, ahal denean, zergatik gertatzen den erakutsiz eta egoerari aurre egiteko zenbait estrategia azalduz. Horretarako, AARI generoaren perspektibatik begiratuko diogu eta emakumeek AAREN garapenean izan duten presentzia isildua gogoratuko dugu. AAKo aplikazioek jokamolde alboratuak erakusten dituzten portaerak aztertuko ditugu, alborapen horien kontzientzia zabaltzeko eta zuzentzeko bideak jorratzeko. Horrela, testu honen asmoa AA zer den hobeto ulertzen laguntzea eta hausnarketarako bidea ematea da.

**Gako hitzak:** Adimen Artifiziala, ikasketa automatikoa, robotika, genero-alborapena.

**Abstract:** Historically, the development of science and technology has been led by heteronormative men (mainly white and from middle- to upper-class backgrounds), and the applications developed have shown significant biases. Artificial Intelligence (AI), which has fully entered our daily lives and is a current field of research, is not free from this framework and exhibits a gender gap (as well as many others).

In this article, we aim to bring this to light by showing where bias occurs and, when possible, why it happens, as well as by explaining several strategies to address the situation. To do so, we will examine AI from a gender perspective and recall the silenced presence of women in the development of AI. We will analyse the behaviours of AI applications that display biased patterns, in order to raise awareness of these biases and explore ways to correct them. In this way, the aim of this text is to help better understand what AI is and to encourage reflection.

**Key words:** Artificial Intelligence, machine learning, robotics, gender bias.

# 1. Sarrera

Adimen Artifizialaren (AA) definizioa ez da erraza, are gehiago, adimena bera zer den argi ez dakigula kontuan hartuta. Edonola ere, modu azkar eta ulergarrian esan dezakegu makina batek normalean giza adimena behar duten zereginak egiteko duen gaitasuna dela AA.

Giza adimena makinetara eramatea zaila den neurrian, AA maila batean baino gehiagotan definitzen da. Horrela, AA ahula/estua (edo besterik gabe, AA) esaten zaio ataza batean portaera adimenduna erakusteko gaitasunari. Txatbotek (ChatGPT, DeepSeek, Gemini, Claude.ai, Latxa eta halakoek) txundituta uzten gaituzte, hain da itxura onekoa elkarrizketa bati jarraitzeko duten ahalmena... Lagungarriak zaizkigun arren, laster konturatzen gara ez dutela hizkuntza ulertzen. Eta, hala ere, neurri batean elkarrizketa adimenduna sortzen dutela aitortzen diegu. AA indartsua/orokorra, aldiz, zeregin guztietan giza adimenaren mailako portaera erakusteko gaitasunari esaten zaio, gure eguneroko jardun adimenduna izango lukeenari, alegia: eskailerak jaitsi, gidatu, laneko zereginak egin, patata-tortila prestatu, pilota-partidua jokatu... Horri gehitu beharko genioke trebetasun sozialak eta emozionalak izatea, gizakiok sarri erabakiak enpatian eta intuizioan oinarrituz hartzen baititugu.

Eraman daiteke hori guztia makina batera? Eta geure buruaz dugun kontzientzia? Halako izaki adimendunen fantasiekin eszenatoki alarmista eta distopikoak irudikatu izan dira filmetan eta eleberrietan, eta beldurra eta kezka filosofiko sakonak eragin dira. Aitzindaria izan zen Mary Shelleyren zientzia-fikziozko *Frankenstein* eleberria: bizitza artifiziala irudikatu zuen, aurrerapen teknologikoen alderdi etikoaren eztabaida ireki zuen eta robotikaren eta AAren oinarri kulturalaren eragile nabarmena izan da.

AA zer den eta giza adimena makinetan txertatzea hain zaila zergatik den ulertzeko, historian zehar ikertzaileek, gizonezkoek eta emakumezkoek, topatu izan dituzten zailtasunak eta haiek gainditzeko garatutako aurrerakuntza teknologikoak ulertzea komeni da. Gaur egun, AAren inguruan dauden espektatibak altuak dira, "AA gizateria egiten ari den gauzarik garrantzitsuenetakoa da" edo "AA argindar berria da" eta halakoak entzuten dira. Ikus dezagun nola lortu den hona iristea.

## 1.1. Historia

Konputagailuaren kontzeptuaren sorreran koka daiteke gerora AA izango denaren hazia. Kontzeptu hori Alan Turingek eman zuen 1930eko hamarkadan. Problema matematiko teoriko baten ebazpenean lanean ari zela, makina baten eredu proposatu zuen, gerora Turing makina izenez ezaguna dena eta geroko konputagailuen oinarri izan dena. Ada Lovelacek, historiako lehen programatzaileak, lehen algoritmoa idatzi zuenetik (1840) ehun urtera gertatu zen hori.

1940-1950eko hamarkadetan, makinaren pentsatzeko gaitasunari buruzko eztabaida piztu zen. Gizakion portaera adimenduna eta makinena bereizezinak ziren erabakitzeko test bat proposatu zuen Turingek. AAri buruzko lehenengo argitalpen zientifikotzat hartu ohi da (Turing, 1950). Handik gutxira, 1955ean, John McCarthy, Marvin Minsky, Claude Shannon eta Nathaniel Rochester ikerlariek Dartmoutheko tailer zientifikoa antolatu zuten "Adimen Artifiziala" izenburupean. AA ikerketa-arlo gisa aitortua izan zen eta laster iritsi ziren lehen emaitzak. 1956an Logic Theorist programa garatu zen. Teoremak frogatzeko programa bat zen, gizakiok problemak ebazteko dugun gaitasuna simulatuz sortua, eta AAko lehen programatzat hartzen da (Newell et. al., 1958).

Konputagailuaren kontzeptua sortu eta handik gutxira eraiki ziren lehen konputagailu fisikoak (Electronic Numerical Integrator and Computer, ENIAC) eta konputagailuen oinarritzko egitura zehaztu zen (Von Neumannen arkitektura). 1940ko hamarkadaren erdialdean, teknologiaren garapena erabat maskulinizatuta zegoen garaian, aitzindariak izan ziren ENIAC programatzen aritu ziren sei emakumeak:

Betty Jean Jennings, Betty Snyder, Frances Bilas, Kathleen McNulty, Marilyn Wescoff eta Ruth Lichterman. Garai hartako konputagailuak erraldoiak ziren, energia kontsumo handikoak eta garestiak. Haiekin lan egitea nekeza bazen ere, garai hartako gizartea aztoratu zuten: "burmuin elektronikoa" asmatu zela esaten zen, goi-mailako adimenaren isla zela. Lorpena handia zen, lehen problema aritmetiko sinpleak ebaztea lortu zen, zenbakiak ordenatzea, taula-joko errazetan trebe aritzea...

1960-1970eko hamarkadetan, gizakion gaitasunak identifikatu eta makinetan txertatzeko ikerketa-lanari ekin zioten ikerlariak: pertzepzioa, problemak ebaztea, arrazoitzea, hizkuntza ulertzea, etab. Lehenik, blokeen mundu simulatuen arloan ikertu zuten (SHRDLU programa: aginduak jaso, atazak planifikatu eta blokeak mugitzea).

Erabaki adimentsuak hartzeko, hots, aukera guztien artean optimoa aurkitzeko, bilaketa-teknikak garatu ziren. Baina, problemen tamaina handitzean, egin beharreko kalkulu-kopurua izugarri hazten zen (konplexutasun konputazionala) eta ezinezko bihurtzen zen. Hori agerian gelditu zen robot fisikoekin egindako ikerketetan. Izan ere, robotak sentsoreen bidez inguruetik informazioa jaso, planifikatu eta erabakiak exekutatzeko unerako ingurunea jada aldatuta zegoen eta planifikatutako ekintzak zentzugabeak ziren. AAREN zalaparta horretan sortutako lehen robot adimendun erreala, eta kasik bakarra *Shakey* izan zen (Nilsson, 1984). Beste teknologiekin gertatzen den bezala, behin eta berriz ikusten den egoera bat errepikatzen da AAREN historian: ematen den aurrerapauso bakoitzak zailtasun berriak azaleratzen ditu.

1970-1980ko hamarkadetan giza adituek urteen poderioz lortutako "ezagutza" jaso eta konputagailuan txertatu behar zela sinistuta, "sistema adituak" garatu ziren. Erregelak erabili zituzten pertsona adituen ezagutza adierazteko eta produktu komertzialak sortu ziren. Odoleko gaixotasunen diagnostikoan laguntzeko Mycin sistema aditua, adibidez, oso arrakastatsua izan zen (Shortliffe, 1977). Hain handia izan zen arrakasta, non ezagutza osoa jasotzeko proiektua ere abiatu baitzuten (Cyc proiektua), baina ez zuen funtzionatu. Arrazoiketa adimendunak egiteko, logikan oinarritutako sistemak ere garatu ziren, baina horiek ere arrakasta urria izan zuten.

Robotetara itzuliz, Rodney A. Brooks ikerlariak azaleratu zuen 1980ko hamarkadaren amaieran arrazoiketa hutsean oinarritutako adimena ez zela nahikoa. Sistema fisiko errealak gorputza behar du eta bere ingurunearekin elkarreraginez garatuko du adimena, gizakiok eta animaliek egiten dugun modura. Gaur egun elkarrekintza hori ezinbestekotzat jotzen bada ere, garai hartan, hamaika robot -zoomorfiko, gurpildun nahiz humanoideak- garatzeko atea ireki zuen eta portaeran oinarritutako sistemen paradigma proposatu zen.

Gerora, ezagutza, arrazoiketa eta portaeren ideiak konbinatuz, "agenteak" proposatu ziren. Agenteen bidez atazak modu autonomoan egitea zen helburua eta, horretarako, erreaktiboak ziren (ingurunearen aldaketei erantzuteko), proaktiboak (erabiltzaileak eskatutako ataza betetzeko), baita sozialak ere (beste agenteekin elkarlanean aritzeko). Agenteen teknologiaren ondorengo zuzenak dira Siri eta Alexa, 2010eko hamarkadan merkaturatu zirenak, ahalmen handiagoko hardwarea eskura izan zenean.

2000ko hamarkadan ikerketarako bide berri bat zabaldu zen: ikasketa automatikoa. Besteak beste, ordura arteko metodoekin ebazteko zaila zen karaktereen identifikazioaren problema erabili zen; hau da, eskuz idatzitako karaktereak (datuak) erakutsiz, sistemak karaktereak bereizten ikasten zuen. Datu-mota desberdinekin esperimentuak egin ziren eta oso emaitza onak lortu ziren, baina datu pila bat behar zen.

Azken iraultza handia neurona-sare artifizialen eskutik etorri zaigu. Nerbio-zeluletan (neuronetan) inspiratutako teknologia da. 1950eko hamarkadan Frank Rosenblattek proposatu zuen lehen neurona-sarea (Perceptron) (Rosenblatt, 1958). Geruza bakarrekoa zen eta, haren gaitasuna mugatua zenez,

Marvin Minsky eta Seymour Papertek gogor kritikatu zuten (Marvin et al., 1969). 1980ko hamarkadan neurona-sareen eredu berriak planteatu ziren eta geruza anitzeko sareak entrenatzeko atzeranzko barreiaketa (*backpropagation*) proposatu zen (Rumelhart et al., 1986). 2005etik aurrera ikaskuntza sakona indarra hartzen hasi zen; hau da, neurona gehiago, geruza gehiago eta konexio gehiago erabiltzen, eta hazkunde horri esker, neurona-sareetan oinarritutako sistemak ikaragarri hobetu ziren.

Hobekuntza horren isla dira 2010-2020ko hamarkadetan garatu ziren proiektu asko, adibidez: 1) teleskopio bidez jasotako datuetan oinarrituz, Katie Boumanek zulo beltz baten lehen argazkia egitea lortu zuen, 2) errealak diruditen pertsonen irudi artifizialak sortzea lortu zen, 3) medikuntzan garrantzitsua den proteinen tolestea ulertzeko AlphaFold tresna garatu zen. Lorpen handi horiek batera etorri dira "sistema sortzaile" izenez ezagutzen direnak. Erabiltzaile arrunt garenok gertu geratzen zaigun sistema sortzaile horietako bat ChatGPT txatbota da. Denetarako aplikazioak sortu dira, gure egunerokotasunean sartu eta kolpetik dena aldatu dela dirudi: zientzia, hezkuntza, industria, merkataritza, osasungintza, entretenimendua, artea, etab. Jakin badakigu, irudien eta testuen interpretazioan asko aurreratu den arren, sistemek ez dituztela benetan ulertzen. Hala ere, emaitzak ikusgarriak dira.

Erronkak handiak dira oraindik, gidaririk gabeko automobilen garapena kasu. Pertzepzioa zaila da, automobilak non dagoen eta bere inguruan zer gertatzen ari den ulertu behar duelako. Baina, batez ere, eztabaida etikoak sortzen dira: istripua izanez gero, norena izango litzateke ardura? Diseinatzailearena, saltzailearena, erabiltzailearena? Eta antzeko beste hamaika egoera planteatu daitezke.

Zeregin guztietan giza adimenaren mailako portaera erakutsiko lukeen AA orokor hori lortzetik urrun gaude. Une honetan, agian garrantzitsuena ez da AA nora irits daitekeen jakitea, egin nahi dugun bidea zein den erabakitzea baizik, gizateriaren eta ingurune naturalaren onerako izan dadin (Wooldridge, 2021).

Egiten dugun bidean eragina izango dute, dudarik gabe, teknologia horiek sortzen parte hartzen duten pertsonak. Alde horretatik, aipagarria da historiako errebaso labur honetan emakumeen izenak, gizonezkoen aldean, gutxi direla. Ez da harriztekoa, izan ere, neurri handi batean teknologiarik buruzko ikasketetatik eta lanbideetatik at geratu ziren. Hala ere, testu honetan aipatutako emakumeek, eta beste askok, rol erabakigarria izan dute teknologia eraldatzaile horri forma ematen. Gertatutakoa ez da anomalia historikoa soilik, eta hori aldatzea etengabeko erronka da oraindik. Baina ez da berandu aldatzeko. Emakume haien ekarpenak aitortzea garrantzitsua da bidegabekeria historikoa zuzentzeko, baita etorkizuneko emakumeengan AAri eta teknologiarik buruzko ikasketak egiteko inspirazioa sortzeko ere. Hurrengo atalean aztertuko dugu laburki zein den gaur egungo egoera.

## 1.2. Emakumeen presentzia Adimen Artifizialaren arloko lanbideetan

Aurreko atalean, historian zehar AAk izan duen garapena errebasatu da eta, tartean, emakumeek egindako ekarpenak ikus daitezke. Gaur egun ere, emakumeen eskutik datoz AAen egiten diren hainbat ekarpen nabari. Joy A. Buolamwinik, adibidez, aurpegien errekonozimendurako sistemak larruazal iluna duten emakumeak identifikatzeko zailtasunak zituztela frogatu zuen bere ikerketekin, AA sistemek izan dezaketen alborapenaren-arazoa mahai gainean argi jarrita. Mira Muratik hizkuntza naturalaren prozesamenduan ekarpen handia egin du; hura zen gaur egun hain ezaguna bihurtu zaigun ChatGPT txatbota garatu zuen taldearen zuzendaria. Fei-Fei Li ikusmen artifizialean mugarri den ImageNet proiektuaren sortzailea da. ImageNet irudiak sailkatzeko datu-base erraldoia da, neurona-sare sakonen errendimendua hobetzeko funtsezkoa izan dena eta konputagailuen bidezko ikusmenaren arloa goitik behera aldatzea ekarri duena. Nabarmenezkoa da, baita ere, Fei-Fei Li agertzen duen konpromisoa AA denon eskura egon dadin eta ez soilik konpainia teknologikoa handien esku. Bera izan da *AI4ALL* (<https://ai-4-all.org>) ekimenaren sortzaileetako bat.

Robotikari dagokionez, aipagarria da Rosalin Picarden eskutik etorri zela konputagailuei eta robotei (makinei, oro har) emozioak ulertzeko eta simulatzeko gaitasuna nola eman aztertzen duen “konputazio afektiboa” ikerketa-arloa. Arlo horretan, Cynthia Breazealek *Kismet* buru robotiko espresiboa eraiki zuen 2000. urtean (gizakien eta roboten arteko elkarrekintza sozialaren aitzindaria da *Kismet*). Geroago, 2005. urtean, robotikaren, psikologiaren eta gizakien eta roboten arteko elkarrekintzaren oinarriak konbinatuz, Maja Matarićek “robot sozial laguntzaile” terminoa sortu zuen.

Ez dira emakumeek egindako ekarpen bakarrak, noski, baina guztiz deigarria da, oro har, AAren eta Informazioaren eta Komunikazioaren Teknologiaren (IKT) arloko lanbideetan jarduten duten emakumeen presentzia eskasa. Hori horrela da mundu osoan, eskuarki. Hona hemen 2023-24ko estatistika ofizialetako datu batzuk: World Bank Groupek jaso zuenez, munduan AA eta IKT arloan lanean ari direnen % 32 da emakumezkoa. Eurostat Europako estatistika-bulegoan jasota dagoenez, IKT arloan espezialista gisa jarduten dutenen % 19,4 soilik da emakumezkoa. Eustatek IKT sektorean jasotzen dituen estatistiken arabera, IKT sektoreko Ikerketa eta Garapeneko (I+G) langileen artean % 24,4 da emakumezkoa (ikusi 1. taula). Portzentaje horiek ez dira guztiz konparagarriak, aipaturiko erakundeek ez baitituzte datuak modu berean jasotzen, baina emakumezkoak AA eta IKT arloko lanbidetan nabarmenki gutxiengoa direla erakusten dute.

1. Taula. Emakumeen presentzia (ehunekotan) AA eta IKT arloko lanbideetan			
Eremua	Datuen jatorria	AA eta IKT arloko langileak jasotako modua	Emakumeak lanbidean (%)
Mundua	World Bank <sup>1</sup>	Lan-merkatua eta enplegua IKTetan eta teknologiaren sektore aurreratuenetan	32,0
Europa	Eurostat <sup>2</sup>	IKT espezialista gisa lanean	19,4
EAE	Eustat <sup>3</sup>	IKT sektoreko I+G arloan	24,4

Bestetik, datu horiek ez dira harritzekoak kontuan izanik emakume gutxi aukeratzen dituztela AA eta Informatika ikasketak. Unibertsitate ikasketei dagokionez, EHUko Informatika Fakultateko AA graduak ikasle nesken portzentajeak nekez gainditzen du % 30, eta Informatika Ingeniaritza graduak portzentaje hori are baxuagoa da: % 20 soilik 2025ean. Azken hoge urteotan antzeko datuak izan ditugu.

Unibertsitate-ikasketetatik at, oraindik ere kaxkarragoak dira datuak; Lanbide Heziketaren erdi-mailako eta goi-mailako zikloetan, % 7,6 eta % 15,1 dira emakumeak, hurrenez hurren, EAEko ikastetxeetan<sup>4</sup>.

Datuek argi erakusten dute emakumeen talentua AA arloetara erakarri beharra dagoela, eta bide hori jorratzeko estrategiak eta politikak landu behar direla (Bello, 2021).

- 1 Closing Gender Gaps in Digital Development. A practical Guide for Operational Teams, World Bank Group, 2023ko txostena.
- 2 2023ko datua, [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=ICT\\_specialists\\_in\\_employment#ICT\\_specialists\\_by\\_sex](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=ICT_specialists_in_employment#ICT_specialists_by_sex)  
<https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20240524-2>
- 3 2024ko datua, [https://es.eustat.eus/bankupx/pxweb/es/DB/-/PX\\_222315\\_ceit\\_tic07.px/table/tableViewLayout2/](https://es.eustat.eus/bankupx/pxweb/es/DB/-/PX_222315_ceit_tic07.px/table/tableViewLayout2/)
- 4 Emakunderen 2022ko txostena. [https://www.emakunde.euskadi.eus/contenidos/informacion/emakunde\\_memorias/eu\\_emakunde\\_adjuntos/2022.memoria.emakunde.eu.pdf](https://www.emakunde.euskadi.eus/contenidos/informacion/emakunde_memorias/eu_emakunde_adjuntos/2022.memoria.emakunde.eu.pdf)

## 2. Genero-arrakala eta Adimen Artifiziala: adibideak

Aurreko atalean aipatu den moduan, AAKo aplikazioak gure egunerokotasunean sartu zaizkigu. Halere, sistema horien sortzeko gaitasuna ez dator bat, oraindik, ulermen- mailarekin eta, zoritxarrez, genero-arrakala nabarmena erakusten dute. Hartara, atal honetan, gaur egun asko erabiltzen diren AAKo bi aplikazio aztertuko ditugu: itzulpengintza automatikoa eta irudien sorkuntza. Sistemek ematen dituzten erantzunetan arreta jarriko dugu hausnarketarako bidea irekitzeko asmoz.

### 2.1. Itzulpen automatikoa

Oso hedatuta daude testua itzultzeko sistemak. Izan ere, azken urte hauetan asko hobetu dira eta gero eta itzulpen hobeak ematen dizkigute. Kasu honetan, itzulzaile automatikoko sistema bati<sup>5</sup> euskaratik gaztelarrera itzultzeko eskatu diogu eta lortu ditugun emaitzak bere horretan jaso ditugu. 2. taulan ikus daitezke lau eskaerak, bakoitzean emandako testua eta jaso den itzulpenarekin batera.

2. Taula. Eskaerak eta jasotako itzulpenak		
Eskaera	Euskarazko testua	Gaztelarazko testua (itzulpena)
1	Gaixoaren egoera tamalgarria zen eta erizainak berehala artatu zuen.	El estado de la enferma era lamentable y la enfermera le atendió inmediatamente.
2	Gaixoaren egoera tamalgarria zen eta medikuak berehala artatu zuen.	El estado de la paciente era lamentable y el médico le atendió inmediatamente.
3	Futbol jokalaria min hartu eta erizainak berehala artatu zuen.	El futbolista se lesionó y la enfermera le atendió de inmediato.
4	Futbol jokalaria min hartu eta medikuak berehala artatu zuen.	El futbolista se lesionó y fue atendido inmediatamente por el médico.

Jaso ditugun itzulpenen kalitateari ikuspegi sintaktiko eta gramatikaletik begiratzen badiogu, sistemak itzulpen onak egin dituela esango dugu. Dena den, ikuspegi zabalago batetik aztertzen badugu, konturatzen gara euskarazko esaldiek ez dutela genero-markarik, baina gaztelaniak genero-markak behar dituzenez, sistemak automatikoki "erabaki" dituela: gaixoa eta erizaina emakumezkoak dira baina medikua gizonezkoa da itzulpenetan [ikusi 1, 2 eskaerak eta dagozkien itzulpenak]. Hurrengo esaldietan, futbol-jokalaria gizonezkoa balitz bezala itzuli du sistemak, eta erizaina eta medikua berriro ere emakumezkoa eta gizonezkoa balira bezala, hurrenez hurren [3, 4 eskaerak eta dagozkien esaldiak].

### 2.2. Testutik irudiak sortzea

Azken urteetan, AAKo sistema sortzaileen eskutik, eduki "berriak" sortzeko gai dira sistema batzuk. Hor daude, adibidez, lehenago aipatutako txatbotak. Sistema horiek, testu bat emanik, testuan esandakoari erantzuteko gai dira. Atal honetan, horrelako sistema bati<sup>6</sup>, testu bat emanik, testu horretan esandakoa azaltzen duen irudi bat sortzeko eskatu diogu. 2. Taulako eskaera berdina egin ditugu eta 1. eta 2. irudietan daude jasotako erantzunak.

5 Elia, 2025.06.20

6 ChatGPT-4.1, 2025.11.10

## 1. Irudia. Gaixoaren egoeraren eskaeren erantzun gisa jasotako irudiak



Gaixoaren egoera tamalgarria zen eta erizainak berehala artatu zuen.



Gaixoaren egoera tamalgarria zen eta medikuak berehala artatu zuen.

## 2. Irudia. Futbol-jokalariaren egoeraren erantzun gisa jasotako irudiak



Futbol joklariak min hartu eta erizainak berehala artatu zuen.



Futbol joklariak min hartu eta medikuak berehala artatu zuen.



Futbol joklariak min hartu eta medikuak berehala artatu zuen. Futbol-jokalaria emakumezkoa da eta erizaina gizonzkoa.

Esperimentu labur honetan ere, itzultzaile automatikoak egin duenaren antzera, erizaina emakumezko baina medikua gizonzko moduan irudikatu ditu 1-4 eskaeretan. 1-2 eskaeretan gaixoak gizonzko irudikatu ditu eta 3-4 eskaeretako futbol-jokalariak ere gizonzko. 5. eskarian, sistemari zehaztu diogu futbol-jokalaria emakumezkoa dela eta erizaina gizonzkoa. Sistemak emandako irudi horretan deigarriak dira, batetik, futbol-jokalariaren kamiseta, eta, bestetik, erizainarena. Jokalariaren kamisetak ez dirudi propio futboleko elastikoa eta, erizainari dagokionez, erizain-jantzera baino epaile-jantzera dela dirudi. Hori bai, erizainaren itxurako eskularruak jantzita ageri da.

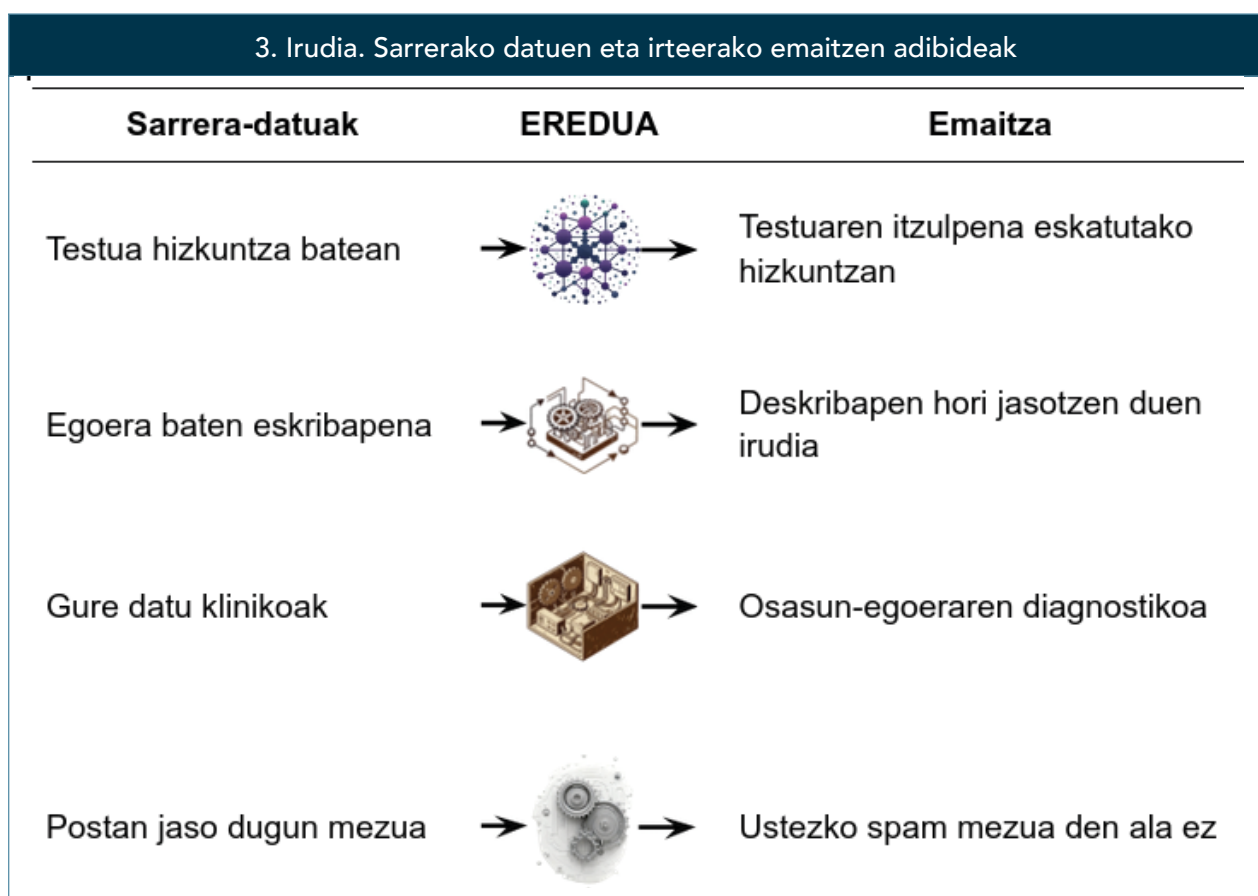
Kontuan izan behar da eskaera bera berriz eginez gero sistemak aldiro irudi berri bat sortuko duela, eta ez duela zertan beti ezaugarri berdinekin sortuko.

Bi adibide horietan AAKo sistemek emandako erantzunak ikusita, hausnartzeko eta ikertzeko beharra nabaria dela agerian geratzen da. Erantzun horietan antzematen ziren joera alboratuak ausazkoak al dira? Zenbateraino erakusten dituzte AAKo sistemek horrelako edo bestelako joerak? Zer-nolako joerak dira? Zergatik egiten dute hala? Galdera asko sortzen dira, eta ondorioz, ikerketa-lerro berriak ireki behar dira erantzunak topatzeko. Alborapenak zergatik gertatzen diren eta joera horien jakitun izanik zer egin dezakegun aztertuko dugu hurrengo ataletan.

### 3. Alborapenen jatorria

Aurreko ataleko adibideetan nabaria da AAKo sistemek genero-alborapena dutela. Genero-alborapenak eta bestelakoak zergatik gertatzen diren ulertzeko, AAKo sistema bat funtsean zer den eta nola eraikitzen den jakin behar da. Hortik abiatuko gara, nahiz eta AAKo sistemak osagai askoz osatuta dauden, eta horietako edozeinetan egon daitezkeen alborapena eragin dezaketen faktoreak.

Erabiltzailearen ikuspuntutik, sarrerako datu batzuk emanik edota eskaera bat eginik, AAKo sistemak emaitza bat ematen du, ikusten ez den mekanismo edo eredu batean oinarrituz. Kasu konkretu batzuetarako aplikazioa ikusi dugu aurreko atalean, baina testuingurua eta datu-motak askotarikoak izan daitezke. Horrela, bai sarrerako datuak bai irteerako emaitza ere zenbakizkoa, testua, irudia, ahotsa edo haien nahasketa izan daiteke, besteak beste. 3. irudian jaso dira sarrerako datuen eta irteerako emaitzen eredu batzuk.

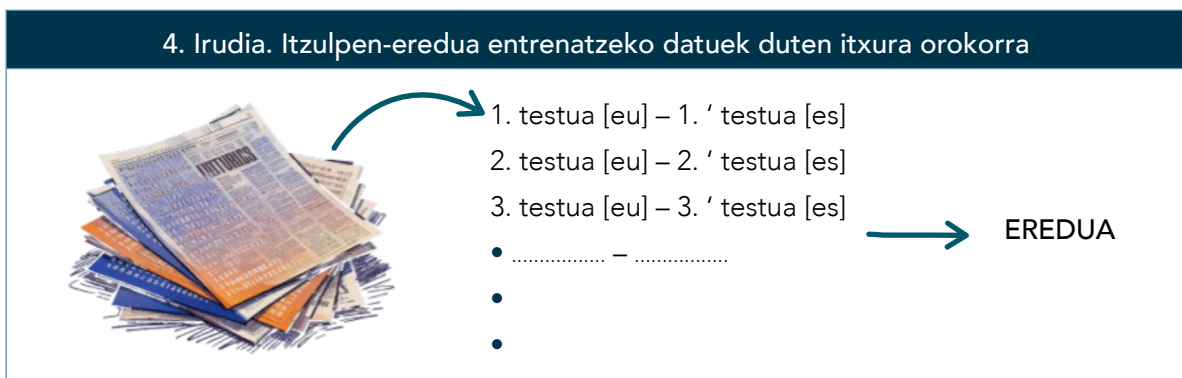


AAko sistemaren muinean dagoen eredu matematikoa edo algoritmoa egokia den neurrian izango da ona sistema bera. Mota askotakoak dira ereduak. Adibidez, dagoeneko aipatu ditugun neurona-sareak. Edonolakoak izanda ere, sistemak joera alboratuak erreproduzitzen baditu, AA sistema erabiltzeak alborapen horiek anplifikatuko ditu, noski, eta albo-eraginak sortu.

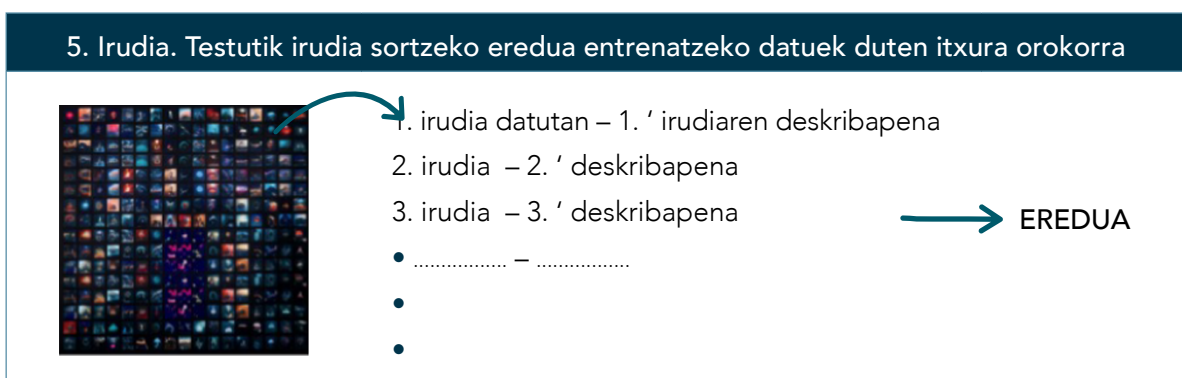
Alborapenaren jatorria askotarikoa izan daiteke eta ez da erraza izaten alborapena eragiten duten arrazoiak identifikatzea. Are gutxiago, sistema eraiki duenak ez baditu behar adina xehetasun ematen, baina, gehien errepikatzen direnak ulertzeko, AAKo sistemek barruan duten eredu matematikoari begiratuko diogu. Azpiko eredu hori "ikasteko" datu-multzo bat behar da. "Entrenamenduko datu-multzoa" esaten zaio. Bertan, sarrerako datu batzuk eta dagozkien irteerako datuak jasotzen dira,

nolabait sistemarentzako adibideak izango direnak, bertatik "ikas dezan". Datu-multzo horretan oinarrituta eraikiko da eredia. Datu-multzoa jasotzeko erabili den prozedura, jaso diren datuen ezaugarriak eta egin den datuen aurreprozesaketa, besteak beste, erabakigarriak dira eraikiko den eredurako, eta alborapenaren jatorri izan daitezke.

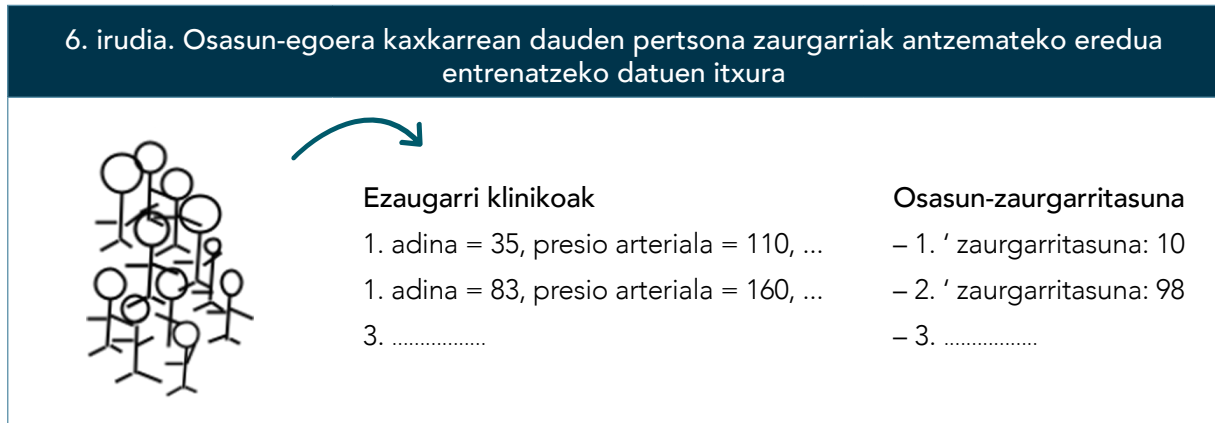
AAko sistemetan sarrera-datuekin lotuta dauden alborapenaren jatorri izan ohi diren hiru iturri ulertzen saiatuko gara atal honetan, zenbait adibideri erreparatuz. 2.1. atalean aipatutako itzulpengintza automatikorako kasua gogora ekarriz, begira dezagun eredia eraikitzekeo zer-nolako datuak erabili ohi diren. Kasu horretan, idatzizko testu-bikoteak erabili ohi dira, esanahi bereko testuak bi hizkuntzatan, euskaraz eta gazteleraz, adibidez (ikus 4. irudia). Web-orrietatik, aldizkarietatik eta egunkarietatik jasotakoak izan ohi dira, kopuru erraldoietan. Testu horietan gizartean txertatutako alborapenezko joerak badaude (2.1. ataleko adibideetan ikusitakoan antzekoak, hau da, ikasketako testuetan erizaina gehienetan emakumezkoa bada, eta medikua edo futbol-jokalaria gizonezkoa), ereduak alborapenezko joera horiek errepikatu eta anplifikatuko ditu. Halakoetan, ereduak "alborapen historikoa" duela esaten da.



2.2. ataleko kasuari erreparatzen badiogu, eredia eraikitzekeo irudi-testu pareak behar dira, hau da, irudiekin (datu moduan prozesatuta) eta haien deskribapenekin (testu moduan) irudi-testu pareak osatu (ikus 5. irudia). Berrito ere, datu-kopuru erraldoiak behar dira horrelako ereduak doitzeko. Adibide horretan ere alborapen historikoa dagoela dirudi, erizainak emakumezko eta medikua edo futbol-jokalaria gizonezko irudikatu ditu, baina ez hori bakarrik. Izan ere, eredia ikasteko erabili diren irudietan ia ez badago gizonezko erizainik, ereduak ezin ditu egoki irudikatu "gizonezko erizainak". Badirudi horrelako zerbait gertatu dela azken eskaerarekin, sistema behartu eta "futbol-jokalaria emakumea zen eta erizaina gizona" adierazi zaionean. Halakoetan, entrenamendurako datuetan dagoen aniztasun eskasa da sistemaren alborapenaren iturria. Antzeko arazoa sarri agertu izan ohi da irudien bidez aurpegiak errekonozitzeko sistemetan, azal beltzeko aurpegiei aplikatzean. Eredua entrenatzeko erabiltzen diren datuen artean azal beltzeko aurpegi gutxi badaude, eredia ez da ongi doitu egongo eta, ondorioz, ereduak errendimendu txarra izango du kasu horietarako. Ereduek "adierazpide-alborapena" duela esaten da.



Har dezagun orain guztiz bestelako adibide bat. Demagun osasun-sistema batean osasun-egoera kaxkarrean dauden pertsona zaugarriak antzemateko AAKo sistema bat eraiki nahi dela. Osasun-arazo larriagoak izan aurretik garaiz antzemanaz gero, gaixoak ongi artatzeko aukera izango luke osasun-sistemak, eta, aldi berean, gaixotasun larriek eragindako gastuak murriztu. Baina horretarako pertsonen informazioan oinarrituz haien "osasun-zaugarritasuna" neurtzeko modua aurkitu behar da (ikusi 6. irudia).



Halako sistema bat inplementatu eta martxan jarri zen, baina azal zuriko pertsonen eta azal ilunekoen artean alborapen nabarmena zuela ikusi zen. Izan ere, azal beltzak nekez identifikatzen zituen maila altuko osasun-zaugarritasuneko gisa (Obermeyer, 2019). Ondorioz, sistemak ez zituen zaugarritzat identifikatzen beharra zuten azal iluneko pertsonak.

Adibide horretan alborapenaren jatorria bilatzeko, garrantzitsua da osasun-zaugarritasunaren neurria kalkulatzeko zein ezaugarri erabili diren eta kalkulua nola egin den aztertzea. Izan ere, nola kalkula daiteke hori? Kontzeptua abstraktua da, ez da adina kalkulatzearen parekoa. Osasun-zaugarritasuna "nolabait" neurtuko duten zeharkako adierazleak bilatu behar dira. Sistema partikular horretan, beste zenbait adierazleren artean, pertsonak osasun-sisteman egindako gastuen historia erabili zuten. Hasiera batean, adierazle hori erabiltzea zentzuzkoa dela dirudi: pentsa liteke osasun-egoera kaxkara dutenek gehiago gastatzen dutela osasun-sisteman osasuntsuek baino. Baina osasungintza pribatuan oinarritzen diren gizarteetan egoera ekonomiko txarrean dauden pertsonak ez dute osasunean gastatzeko aukera handirik eta, beraz, osasun-zaugarritasuna neurtzeko zeharkako bide horrek, osasun-egoera kaxkarrean dauden pertsonak hautematen lagundu ordez, huts egiten zuen, egoera ekonomiko txarrean zeudenak ez zituelako identifikatzen. Esperimentua egin zuten gizarteetan, paziente beltzak ekonomikoki zuriak baino egoera okerragoan zeudenez, sistemak alborapena zuen beltzen kalterako. Jatorri hori duen alborapenari "neurketaren alborapena" esaten zaio.

## 4. Alborapenean eragiteko moduak

Azken urteetan, ikerketa-lan asko abiatu dira alborapenak ekidin eta bidezko sistemak diseinatzen laguntzeko. Gero eta proposamen gehiago daude AAKo sistemen garapen formala eta "industrial" gidatzeko. Kalteak non eta nola sor daitezkeen identifikatzeko, esparru integralak proposatu dira eta sistema bat sortzeko jarraitzen den prozesu osoa ulertzea ezinbestekoa dela nabarmendu da, datuen bilketatik hasi eta ereduaren hedapeneraino (Suresh, 2021).

Lerro berean, AAKo sistemen erabilera okerrak ekiditeko, ereduaren informazioa jasoko duten txartelak proposatu dira (Mitchell, 2019). Horrela, eredu bakoitzak bere dokumentazioa izango du, eta bertan jasoko dira errendimendu-ezaugarriak, ereduaren erabiltzeko aholkuak eta ekar litzakeen alborapen eta arriskuak.

Idea horiek hobeto ulertzeko, sistemen garapena hiru mailatan aztertuko dugu eta bertan alborapena murrizteko estrategiak nola aplikatu daitezkeen azalduko dugu: 1) datuetan, dagoeneko bildutakoak bada, haien azterketa sakona egin behar da, bertan dauden alborapenak identifikatzeko, eta datu-base berri bat sortu nahi bada, hasieratik alborapena murrizteko diseinu egokia egin behar da; 2) ereduaren, erabilitako sistemek era egokian diseinatuta egon behar dute; eta 3) ebaluazio-fasean, lortutako emaitzak zein neurritan diren justuak neurtzeko metrikak erabili behar dira.

### 4.1. Datuak aztertzea

AAko ereduak erabiliko dituen datuak ondo aztertu behar dira, balizko alborapenak identifikatzeko. Datuen bilketa-prozesuan, alborapena hainbat arazoirengatik sor daiteke, hala nola, datuak nork biltzen dituen, zein datu hautatzen diren, zein laginketa egiten den eta datuak nola etiketatzen diren, besteak beste. Beraz, datu-bilketa egiten hasi baino lehen, ezinbestekoa da ondo zehaztea lortu nahi diren berdintasun-helburuak, aplikazioaren arabera alda baitaitezke. Garrantzitsua da hasierako gogoeta eta azterketa hori aplikazioaren arloko adituen laguntzarekin egitea.

Datuen jatorria aztertzean, datuak non, noiz eta nola jaso diren dokumentatu behar da, datuen fitxak edo adierazpenak prestatuz eta jatorria, mugak eta alborapen posibleak deskribatuz. Bide horretan, Gebru et al.-ek (2021) datuak antolatzeke fitxa-egitura bat proposatzen dute, datu-multzo bakoitzerako motibazioa, konposizioa, bilketa-prozesua, gomendatutako erabilerak eta abar jasotzeko. Fitxa horiek datu-multzoen sortzaileen eta kontsumitzaileen arteko komunikazioa erraztuko dute, eta AAKo sistemen garatzaileak gardentasuna lehenestera bultzatuko dituzte. Hori eginez, sistema berri bat garatzeko garaian, haren eginkizunetarako aproposak diren entrenamendurako datu-multzoak era informatuagoan aukeratu ahalko dira, xehetasunak jasotzen dituen fitxak aztertuta, alborapen historikoak dituzten datu-multzoak (giza diskriminazioa, aurreiritziak edo garai bateko egoera baztertzailak) ekidinez.

Edonola ere, aztertu nahi den gizartearen adierazgarri izan behar du datu-baseak. Horretarako, egiten den laginketa ondo diseinatu behar da, populazioa ongi ordezkatu dezan eta inklusiboa izan dadin. Arazo hori estatistika klasikoan ere planteatzen da eta, laginak populazioa ongi ordezkatzeko ez duenean, estatistikako teknikak erabili daitezke arazoa murrizteko. Adibidez, datu biologikoen datu-baseak (biobankuak) baliabide garrantzitsuak dira medikuntza- eta epidemiologia-arloetan ikertzeko. Halako biobankuetan erraz gerta daiteke lagina inklusiboa ez izatea, lagin biologikoak ematen dituztenek askotan modu boluntarioan egiten dutelako, eta horrek desorekak sortzen ditu (Alten, 2025).

Datuen desorekari aurre egiteko, "datu sintetiko" sorrera ahalbidetzen duten AAKo teknikak agertu dira azken urteetan. Adibidez, medikuntza-arloan, datuen pribatutasunaren ondorioz datu-kantitate

handiak lortzea ez denez erraza izaten, datu berriak modu artifizialean sor daitezke, eta horrela datuen multzoa handitzea eta desorekak orekatzea lortzen da. Kontutan izan behar da halako teknikak erabiliz sortutako datu sintetiko berriek datu errealean adierazgarri izan behar dutela.

Bestalde, datu-basea prestatzean datuak etiketatzen direnean, hau da, ereduak ikastea nahi denaren arabera datuak sailkatzen direnean, alborapena gerta daiteke. Adibidez, sare sozialetako testuetan erabiltzen den hizkera erasokorra (toxikoa) den detektatzeko sistemak eraikitzean, testuei "toxiko" edo "ez-toxiko" etiketak jartzen zaizkie. Pertsonak egiten duten ataza da, eta ez da batere erraza, toxikotasunaren definizioa bera anbigua delako. Gerta liteke norbaiti erasokorra iruditzen zaion testua beste norbaitentzat ez izatea. Gainera, ironia, testuingurua edo komunitate jakin bateko hizkera kontuan hartu ezean, testua etiketatzean huts egin daiteke. Hori ekiditeko, datuak etiketatzen dituztenek ager daitezkeen alborapen inplizituen inguruko prestakuntza zabala jaso behar dute, testuingurua ondo ulertzeko gai izan daitezkeen. Horrez gain, datu bakoitzaren etiketatzea, pertsona bakarrak egin ordez, hainbaten artean eginez gero (askotariko taldeak eta etiketatzeko kontsentsuzko sistemak erabiliz), iritzi indibidualaren eragina murriztuko da, etiketatzean egon daitezkeen alboratzea murriztuko da eta entrenatuko den eredu bidezkoagoa izango da.

Datuen aurreprozesaketa fasean, datuen prestaketa egiten da ereduaren entrenatu ahal izateko. Arlo oso teknikoa da, eta kasuistika oso zabala (datuak ez dira zehatzak, osatu gabe daude...). Adibide baten bidez ikusiko dugu zer-nolako arazoak sor daitezkeen. Demagun langileen soldata aurreikusiko duen eredu bat sortu nahi dela, langileen lan-esperientzian oinarrituz, eta aurreprozesaketa-fasean ikusi dela datu batzuk falta direla. Esaterako, emakumeen lan-egonkortasuna gizonezkoena baino baxuagoa denez, oso ohikoa da emakumeen lan-esperientziari buruzko datuetan hutsune gehiago izatea. Hutsune horiek betetzeko datu guztien batezbestekoa erabiliko balitz, gizonezkoen datuak gailenduko liriteke eta datu-basean emakumeen egoera desitxuratuta agertuko litzateke.

Laburbilduz, datuetatik sor daitezkeen alborapenek etengabeko arreta eskatzen dute: bilketaren aurretik, bitartean eta aurreprozesaketan. Baina baita AAKo sistema martxan jarri ondoren ere. Sarrerako datuak aldatzen diren heinean, beharrezkoa da etengabeko jarraipena egitea, eta ereduak erabiliko duen azken datu-multzoa ere aztertzea. Badira berdintasun-neurri eta tresna espezifikoak, lan hori errazteko sortuak (Fairlearn, AIF360...).

## 4.2. Ereduak aztertzea

Ereduaren diseinuan eta entrenamendu-fasean egindako doikuntzek zuzenean eragin diezaiokete sistemaren justutasunari. Hori dela eta, garrantzitsua da ereduaren "gardentasuna" eta "azalgarritasuna" zaintzea eta ereduaren ikasketa-prozesuan doikuntzak egitea. Ikus dezagun ereduaren garapenean zer-nolako alborapenak sor daitezkeen eta zer den eredu gardena eta azalgarria.

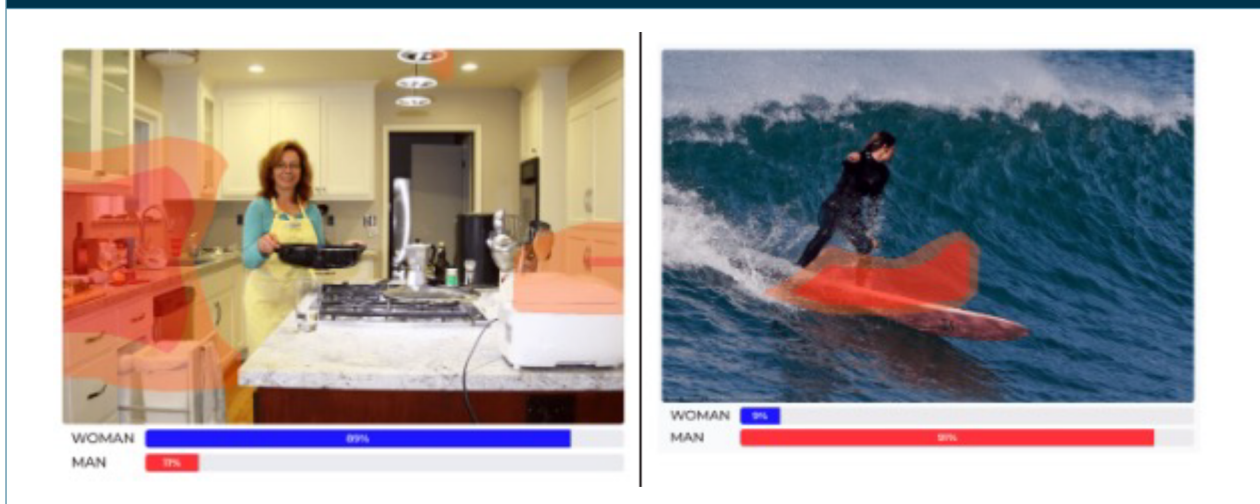
Eredu gardena dela esaten da edonork eredu hori berriro eraiki nahiko balu horretarako beharko lukeen informazio guztia eman bazaio, datuak eta kodea eskuragarri jarriz, eta ereduaren xehetasunak argitaratuz.

Baina eredu gardena izatea ez da nahikoa, ulergarria izatea ere garrantzitsua baita. Izan ere, eredu asko kutxa beltzen modukoak dira: emaitzak ematen dituzte baina ezin da jakin nola eta zergatik iritsi diren emaitza horietara. Egoera horrek zaildu egiten du hutsegiteak identifikatzea eta bereziki, alborapenak detektatzea eta murriztea. Horregatik, AAKo sistemek azalpen argiak eta ulergarriak emateko helburuarekin, Adimen Artifizial Azalgarriaren (*Explainable Artificial Intelligence*, XAI) kontzeptua jaio da azken urteotan. Azalgarri izateak erabiltzaileei eta adituei ereduak nola funtzionatzen duen ulertzeko aukera ematen die. Horrela, eredu azalgarria denean, emaitza emateaz gain, emaitza horretara nola iritsi den adieraziko duen azalpen bat ere ematen du.

Azalgarritasunaren ikuspegitik, bi estrategia nagusi daude. Aukera bat berez azalgarriak diren ereduak erabiltzea da. Eredu horiek egitura sinplea dute eta, horri esker, datuek emaitzei nola eragiten dieten modu zuzenean erakusten dute. Hala ere, baliteke sinpletasun horren ondorioz zehaztasun txikiagoa izatea, eta, horrenbestez, ez dira beti egokiak gaur egungo aplikazio konplexu askotarako.

Eredu konplexuagoak erabiltzea erabakitzen bada, azalgarritasuna gehitzeko teknikak erabil daitezke, hau da, ereduak emaitza jakin batera iristeko zein informaziotan oinarritu den azalduko du. Demagun irudi batean agertzen dena emakumezkoa edo gizonezkoa den iragarriko duen eredu eraiki dela. Azalpen-tekniken bidez, erabakia hartzeko garrantzitsuak izan diren irudiko pixelak nabarmendu daitezke, generoaren iragarpena egiteko erabili diren ezaugarriak agerian utzita. 7. irudian, ezkerrean sistemak emakumea agertzen dela iragarri du sukaldeko tresneriari erreparatuz. Eskuinekoan, aldiz, gizona agertzen dela iragarri du surfeko taulari erreparatuz. Azalpen-teknika horiei esker, ereduak asmatu zein ez, emaitzak modu alboratuan emateko arrazoia zein den ulertzeko aukera dago. AAREN genero-alborapena nerabeei erakusteko erabili dira halako teknikak (Melsión, 2021).

### 7 Irudia. Gorriz adierazita daude erabakia hartzeko garrantzitsuak izan diren irudiko pixelak (Melsión, 2021)



Azalgarritasun-teknikei esker, ereduak nola funtzionatzen duten hobeto uler dezakegu, baina ez dugu ahaztu behar benetako erronka eredu justuagoak eraikitzea dela.

Azken urteotan, entrenamendu aurkaria (*adversarial training*) da alborapena murrizteko metodo aurreratuenetako bat. Ideia oso sinplea da: bi eredu daude, eredu nagusia eta eredu aurkaria. Eredu nagusia ataza zehatz bat egiteko entrenatzen da. Adibidez, curriculumak aztertzen dituen plataforma batean hautagaien egokitasuna iragartzeko. Eredu aurkaria, aldiz, eredu nagusiaren ahuleziak ustiatzen edo nahi ez den informazioa kentzen saiatzen da. Adibideari jarraituz, generoarekin edo ezaugarri etnikoekin lotutako alborapenak murrizten lagun dezake, ereduaren arreta esperientzia profesionalera, gaitasunetara edo jasotako prestakuntzara bideratuz, atributu labainetara baino: generoa, jaioterria...

### 4.3. Sistema ebaluatzea

Datu-basea prestatu eta ereduak garatu ondoren, AAREN sistema ebaluatu behar da, hau da, espero den errendimendua ote duen neurtu behar da. Kontu hori bereziki larria da AAREN sistemetan, gerta daitekeelako akatsak talde baten kontrako alborapen diskriminatzaileak izatea. Horiek ekiditen ez badira, halako egoerak betikotzeko arriskua dago. Kasuen % 99an sistemak ondo funtziona lezake, baina % 1eko akats hori beti giza talde minorizatu baten kontrakoa bada, sistema ez da justua izango.

Adibide bat erabiliko dugu sor daitezkeen arazoak eta horiek antzemateko neurri posible batzuk azaltzeko.

Demagun sare sozialetan testu toxikoak antzemateko sistema baten ebaluazioa egin nahi dela, 1.000 testu etiketatu erabiliz. Testu horiek bi gairen ingurukoak dira: gehienak gai orokorrekoak (900) eta gainerakoak genero-identitatearen ingurukoak (100). 3. taulan jasota daude sistemak ematen dituen erantzunak (toxiko/ez-toxiko), testuen etiketa errealekin konparatuta.

3. taula: Sistemaren ebaluazioa: eman dituen erantzunak (toxiko/ez-toxiko) eta etiketa erreala		
	Etiketa Erreala	
Sistemaren erantzuna	Toxikoa	Ez-toxikoa
Toxikoa	465	80
Ez-toxikoa	50	405

Taulan ikus daitezkeen bezala, 1.000 mezuen artean sistemak ondo sailkatu ditu 870, toxiko diren 465 testu eta toxiko ez diren 405. Horren arabera, sistemaren errendimendua ona dela esango genuke, % 87an erantzun zuzena ematen duelako. Baina taulan ageri den sistemaren ebaluazioa mezu guztiak batera aztertuz egin da.

4. taula. Sistemaren ebaluazioa: eman dituen erantzunak eta etiketa erreala, mezuen gaia kontuan hartuta				
	Gai orokorreko mezuak		Genero-identitateari buruzko mezuak	
	Etiketa erreala		Etiketa erreala	
Sistemaren erantzuna	Toxikoa	Ez-toxikoa	Toxikoa	Ez-toxikoa
Toxikoa	450	20	15	60
Ez-toxikoa	30	400	20	5

Sistemaren ebaluazioa mezuen gaia kontuan hartuz egiten bada, aparteko tauletan jaso behar dira sistemak emandako erantzunen (toxiko/ez-toxiko) eta testuen etiketa errealen konparaketak. Demagun 4. taulan jaso direla gai orokorreko diren mezuen eta genero-identitateari buruzkoak direnen emaitzak. Ikusten da sistemaren portaera guztiz desberdina dela bi taldeetan. Gai orokorreko mezuetan, sistemak oso emaitza onak ematen ditu, 900 mezuetatik 850etan iragarpen zuzena egin baitu (450 toxiko eta 400 ez-toxiko). Aldiz, genero-identitateari buruzkoen kasuan, 100 mezuetatik 20tan besterik ez du asmatu (15 toxiko eta 5 ez-toxiko). Horren arabera, sistemaren asmatze-tasa ona dela esan daiteke, gai orokorreko mezuak toxikoak diren edo ez identifikatzen ona dela, baina ez genero-identitateari buruzkoetan. Sistemaren ebaluazio-neurri horri "zehaztasuna" esaten zaio.

Aurreko taula aztertuz, bestelako ebaluazio-neurri batzuk kalkula daitezke. Ikus daiteke sistemak mezu bat toxikoa dela iragarri egiten duela % 54,5etan (1.000 mezuetatik 465+80=545). Kalkulua gaiak bereiziz egiten bada, gai orokorreko mezuetan antzeko errendimendua duela ikusten da (900 mezuetatik 450+20=470 toxiko gisa iragarri dira, % 52,2) baina portaera oso desberdina da genero-

identitateari buruzkoetan, mezuen % 75 iragarri baita toxiko gisa ( $100 \text{ mezuetatik } 15+60=75$ ). Genero-identitateari buruzko mezuak toxiko direla erraztasun handiagoz iragartzen du sistemak. Bi taldeetan izandako portaeraren konparaketatik abiatuta, "parekotasun demografikoa" izeneko ebaluazio-neurria kalkulatzen da.

Taula horretan oinarrituz, beste ebaluazio-neurri bat kalkulatu dugu. Oraingoan, benetan toxikoak diren mezueta jarriko dugu arreta. Horien artean, sistemak % 90,3 ondo identifikatzea lortu du ( $465/(465+50)$ ). Baina kalkulua gaiak bereiziz egiten bada, berriro ere sistemaren portaera oso desberdina dela ikusten da. Gai orokorreko mezueta, toxiko direnen artean % 93,8 identifikatu ditu toxikotzat ( $450/(450+30)$ ). Aldiz, genero-identitateari buruzkoetan % 42,9an besterik ez du asmatu ( $15/(15+20)$ ). Genero-identitateari buruzko mezueta sistemak nekez detektatzen ditu mezu toxikoak. Bi taldeetan izandako portaeraren konparaketatik abiatuta, "aukera-berdintasuna" izeneko ebaluazio-neurria kalkulatzen da.

Alborapena neurtzeko proposatutako irizpideetan garrantzitsuena multzoen arteko konparaketa egitea da. Análisis taldeka egin ezean, sistemaren errendimendua ona delako eman dezake, eragiten dituen alborapenak oharkabean utzita.

Adibide horrekin oso ondo ikusten da zein garrantzitsua den sistemaren ebaluazioa modu egokian egitea, hau da, sistemak alborapenak izan ditzakeela pentsatu behar da beti. Edonola ere, ez da ataza erraza. Izan ere, nola aukeratu behar dira ebaluaziorako mezuak? Ebaluaziorako 1.000 mezu erabiltzea gutxiegi da? Genero-identitatea talde moduan ondo ordezkaturik dago? Etiketatzea giza alborapenik gabe egin da? Eta abar...

Ebaluazioa talde kultural, demografiko edo fenotipiko desberdinak kontuan hartuz egin behar da, ezaugarri etnikoak, kokapen geografikoa, generoa, larruazalaren kolorea, eta abar kontuan hartuz. Horrez gain, aurreikusitako aplikazio-eremurako garrantzitsuak diren taldeen arteko erlazioak aztertzea ere garrantzitsua da, adina eta talde etnikoa, edo generoa eta larruazalaren kolorea, adibidez.

Laburbilduz, nahiz eta ez egon formula magikorik justuak diren ereduak automatikoki sortzeko, azken urteotan argitaratu diren prozedurek eta gomendioek ereduaren garapen zainduago, gardenago eta egiaztagarriago bat sustatuko dute. Baina kontuz: AAKo eredu baten portaera, bere horretan, kritikoa bada, baina zailtasunak areagotu egiten dira gorpuzkera duten sistema fisikoetan, alegia, robotetan txertatzean.

## 5. Robot sozialak: Adimen Artifizialaren izaera fisiko berezkoenak

Azaletik ikusi dugu, idatzi honen sarreran, robot adimendunen garapena estuki lotuta dagoela AArekin. Izan ere, robotak dira, seguruenik, AAKo tresna asko integratuz AA osotik gertuen dauden tresnak. Robot-mota asko dagoen arren, aipagarri dira robot sozialak, adimen soziala erakutsi behar duten robotak. Hau da, gizakiokin guk gure artean egin ohi dugun moduan jarduteko gaitasuna dutenak. Ahots zein gorputz bidezko komunikazioa, begiradaren bidezkoa, adimen emozionala, empatia eta bestelako gaitasunak aurreikusten zaizkie gailu horiei. Sarreran aipatutako AA orokorretik gertu egotea espero da. Beso robotiko industrialek ez bezala, robot sozialek gizakien antz handia dute, hain zuzen ere komunikazioa eta elkarrekintza errazteko, guregan konfiantza sortzeko, batetik, eta gureak bezalako gaitasunak txertatu ahal izateko, bestetik. Etorkizun hurbilean hamaika eremutan izango direla aurreikusten da: bakarrik bizi diren adinekoen bizikide, zerbitzuetako langile eta irasle laguntzaile lanetan, esaterako. Gure ondoan izango ditugun horiek ezinbestean behar dute AA txertatua edukitzea.

Generoa identitate sozialaren modalitate bat da. Ezaugarri komunak dituzten jende-azpimultzo batentzat gizarteak egokitzen jotzen dituen rol, arau eta adierazpide-moduen multzoa da, eta sexu biologikoan du jatorria (Seaborn, 2024). Kontzeptu multifazetikoa da, genero-identitateak askotarikoak dira, eta kulturaren dinamikan generoa beste identitate sozialekin gurutzatzen da (talde etnikoa, sexualitatea eta desgaitasuna) (Hipólito, 2023). Generoa hautemate/pertzepzio kontua ere bada, gizabanakoek eta gizarteak sortutako ereduak eraginda. Pertzepzio horiek gailuetan ere islarazten ditugu, *tamagotchi*-etan, peluxeetan eta, noski, robotetan ere bai, oraindik ondo ez dakigun arren zeintzuk diren gakoak pertzepzio horiek ager daitezzen. Hiru gako aztertuko ditugu jarraian.

### 5.1. Morfologia eta forma, ahotsa barne

Roboten diseinatzaileek beren lehentasun eta balioez betetzen dituzte sortzen dituzten artefaktuak. Teknologiaren arlo guztietan gertatu den moduan, robotetan ere nabarmena da gizon-begiradaren<sup>7</sup> fenomenoaren eragina. Giza erreplikatik oraindik urruti bagaude ere, gizakiekiko antzekotasun-eskala oso zabala da, robot klasikoaren itxuratik geminoideetara.

Gaur egungo robot humanoideek gizonezkoek feminitateari buruz dituzten espektatibak eta fantasia heteronormatiboak islatzen dituzte batik bat. Ohiko bilakatu dira diseinu argal eta ikuspegi konbentzionaletik erakargarriak diren humanoide-eme sexualizatuak. *Sophia* (*Hanson Robotics*), esaterako, edo XPENG enpresak aurkeztu berri dituen humanoideak (ikusi 8. irudia). Sexualitate esplizitua izan gabe, gorpuzkerak (bizkar zabalak, gerri estuak, aurpegiaren formak) seinaleak pizten dizkigute eta ezaugarri fisiko horiek elkarren artean erlazionatuta daude, ez lirateke bakarka tratatu behar. Zientzia-fikzioa ere bada honen ispilua, *Ex Machina* eta *Blade Runner* filmak, kasu baterako.

---

7 *Gizon-begirada*: Arte bisualetan eta literaturan emakumeak eta mundua ikuspegi maskulino eta heterosexual batetik irudikatze ekintza da, emakumeak objektu sexual gisa aurkezten eta irudikatzen dituen, ikusle heterosexual maskulinoaren plazererako.

## 8. Irudia. XPENG enpresak 2025eko azaroan aurkeztutako robot humanoideak. Itxura ez ezik, ibilkera ere bereztua dute sexuaren arabera



Generoaren pertzepzioa ahotsak eta itxurak pizten dute, besteak beste. Lotura dago bien artean: bat ez badatoz, baliteke nahasmena sortzea edota beldurra eragitea (Dennler et al., 2025).

Generoak ataza konkretuetan robotaren sinesgarritasunari eragiten diola erakusten dute zenbait ikerketak. Horrela, ile motza duten robotak gizonezkoen estereotipoarekin lotutako atazetarako (garraioa, kontrola...) egokiagoztat jotzen dira emakumezkoen estereotipoko atazetarako baino (zaintza, sukaldea...). Baina, beste ikerketa batzuren arabera, atazaren eraginkortasuna bera da rol konkretu batean konfiantza sorrarazten duena, eta ez antzemandako generoa.

Zalantzarik gabe, roboten generoa diseinu-prozesuaren parte garrantzitsua da eta modulatu egin behar da. Roboten diseinuan joera sozialak baztertzeko asmoa/presioa dago, baina bidea ez dago argi. CASA<sup>8</sup> paradigmari kontra eginez, azken urteotan roboten genero neutroa aldarrikatzen duen korrontea sortu da, hau da, generorik gabeko roboten aldekoa: ez-bitarrak, anbiguoak, mekanikoak, neutroak azken finean, androginoak (adib. *Quori*) zein zoomorfikoak (*Paro*) izan.

### 5.2. Algoritmoak eta datu-baseak

Robotek, morfologiaz eta sortzen duten pertzepzioaz haratago, AAKo sistema anitz konbinatzen dituzte beren ekintza fisikoarekin, eta estereotipoak eta joerak heredatu eta areagotzen dituzte, ez bakarrik generoari lotutakoak.

Robotek erabiltzen dituzten algoritmoen alborapenak gure erruz gertatzen dira, gu baikara makinak eta ereduak sortu eta horien gaineko erabakiak hartzen ditugunak. Gizakiaren eta robotaren arteko

8 *Computers Are Social Actors*: gizakien antzeko ezaugarri sozialak esleitzen dizkiegu konputagailuei, eta horregatik, nahi gabe, gizakien arteko elkarrekintzarako heuristika sozial berdina aplikatzen ditugu.

elkarrekintzaren garapenean, ikasketa automatikoak berebiziko garrantzia du. Roboten ikasketa-ereduek sozialki onargarria denaren itxurako portaera ziurtatu behar dute, ematen diren elkarrekintza fisikoetan distantzia sozialak errespetatuz eta segurtasuna bermatuz, adibidez. Portaera horiek ikasteko sortutako ereduen arloan bi problema nagusi daude: batetik, gu geu ez gara gure joeren jabe; bestetik, ereduen ulergarritasuna oso urria da. Horrez gain, robotek komunikazio eta pertzepziorako erabiltzen dituzten sistemak sortzeko erabilitako datuen jatorriari erreparatuz, robot sozialetan bi iturri dira nagusi (Londoño et al., 2024): hizkuntza naturalaren prozesamendurako ereduak eta konputagailu bidezko ikusmeneko ereduak. Batetik, genero-alborapenak daude bilaketa-tresnetan eta erabiltzen diren hizkuntza naturaleko eredueta; eta, bestetik, etnia- eta genero-alborapenak konputagailu bidezko ikusmeneko eredueta. 2. atalean erakutsi ditugu horien adibideak.

### 5.3. Robotaren portaerak, errektiboak nahiz pasiboak

Robotaren portaerak gizakiongan sortzen dituen joeren inguruan gutxi ikertu da oraindik, baina kontuan izan behar da nork bere adimena duela eta robotak ez ditugula guztiok ez modu berean irudikatzen, ez eta tratatzen ere. Badago tratu erasokorrerako joera, arrazoia garbi ez dagoen arren: robotak arrotzak iruditzen zaizkigula, gu baino hobeak izango direla, langabezia sortuko dutela, etab. Txatbotak erabiltzaile arrunten eskuetan jarri aurretik, entrenatuak izan dira tratu erasokorren aurrean nola jokatu behar duten ikasteko. Roboten kasuan ere, halakoen aurrean nola jokatu beharko luketen erabaki behar da, batez ere robotak zenbait ezaugarri sozial erakusten dituenean, biolentzia normalizatze bide ez daitezen izan.

Nola egin aurrera? Diseinu-garaian kontuan hartu beharko litzateke gorpuzkerak nola aldarrikatzen dituen generoa, adina, etnia eta bestelako ezaugarri sozialak, eta nola gurutzatzen diren ezaugarri horiek elkarren artean.

Bestetik, algoritmoak giza aniztasuna kontuan edukiz garatu behar dira, robotek ezaugarri sozial batzuk edo besteak ez diskriminatzeke.

Baina, haratago joanda, gizakiaren eta robotaren arteko elkarrekintzarako diseinatzaileek robotaren funtzio eta portaerak sortzen dituen aurreikuspenak kontuan hartu behar dituzte. Kalte-arriskua minimizatze bideak bilatu behar dira, batez ere adierazpen urriko erabiltzaileengan.

Arestian aipatutako gizon-begiradari aurre egiteko begirada ez-heteronormatiboa behar da roboten garapenean. Badira hainbat inizatiba martxan, hona hemen batzuk: *Algorithmic justice league* (<https://www.ajl.org/>), *Feminist HRI framework* ([https://www.youtube.com/watch?v=t\\_dgG8oAT7I](https://www.youtube.com/watch?v=t_dgG8oAT7I)), *Diversity, robots & AI* (<https://www.laiden.org/research/diversity-robots-ai>), *Black in Robotics* (<https://blackinrobotics.org/>) eta *Robots for Social Justice* (<https://dl.acm.org/doi/10.1145/3610977.3634944>).

## 6. Ondorioak

Artikulu honetan agerian utzi dugu AAKo sistemetan generoko eta bestelako alborapenak nabarmenak direla, datu historikoetatik, diseinu-erabakietatik edo entrenamendu-prozesuetatik sortzen direla eta alborapen horiei heltzeko premia dagoela, emaitza bidegabeak eta diskriminatzaileak eragiten dituztelako. Areago, justizian, giza baliabideen kudeaketan eta finantzetako sektoreetan sistema horien erabilera ugaritzen ari denez, arriskuak arindu behar dira, AA etikoagoa eta bidezkoagoa sustatuz.

Horren aurrean, ekimenak maila askotan jarri behar dira martxan. Batetik, AAKo sistemak garatzen dituztenek maila teknikoko neurriak hartu behar dituzte. Lan honetan aipatu ditugu batzuk (datuen jatorria dokumentatzea, sistemak etengabe ebaluatzea...).

Eta noski, horren guztiaren gainetik, sistemek zein helburu nagusi duten eta noren eskutik datozen aztertu behar da. Teknologia-enpresa erraldoi gutxi batzuen eskutik soilik ez lukete etorri behar. Eta, edonoren eskutik datozela ere, kontrol zorrotzak gainditu behar dituzte.

Bigarren, erakundeen inplikazioa derrigorrezkoa da AAKo sistemak arautzeko. Zuzenbidearen alorrean, 2021ean eskubide digitalen karta argitaratu zen Espainiako Estatuan. Bertan, digitalizazioarekin lotutako beste arlo batzuen artean, AA aipatzen da. Testuak azpimarratzen duenez, beharrezkoa da AAKo sistemek pertsonen duintasuna errespetatzea eta kalterik ez egitea. Horrez gain, diskriminaziorik ez sortzeko ahalegina egin behar da. Horretarako, sistema garden eta ikuskagarriak behar dira. Gainera, pertsonen giza interbentzioa eskatzeko eskubidea dutela ere jasotzen da.

Zuzenbidearen alorrean halaber, Europako parlamentuak Adimen Artifizialaren Legedia onartu zuen 2024an. Helburu nagusia oinarritzko eskubideak babestea da, ingurumen-jasangarritasuna ere aintzat hartuz. Legedi horrek AAKo sistemak sailkatzen ditu, sortzen dituzten arriskuen arabera. Arrisku-maila handikotzat hartzen dira herritarren eskubideak urratzen dituztenak, hala nola ezaugarri labainetan oinarritutako kategorizazio biometrikoa eta lan- eta hezkuntza-inguruneetan emozioen errekonozimendua egin dezaketen sistemak. Arrisku handiko AA sistemak debekatu egiten ditu legedi horrek. Indarrean 2024an sartu bazen ere, pixkanaka ezartzen ari da eta arau gehienak 2026ko abuzturako bete beharko dira.

Hirugarren, gutako bakoitzak ere badu zeresana. Izan ere, ez ditugu aipatu gabe utzi nahi AA sistemen aurrean gizartean maiz ikusten diren bi jarrera kontrajarri: batzuen ustez, AA sistemek diotena erabat egiazkoa da; beste batzuen iritziz, hondamendia besterik ez da etorriko sistema horietatik. Ez bata eta ez bestea ez direnez zuzenak, gutako bakoitzari AAKo sistemen inguruko gutxieneko formazioa eskaini beharko litzaiguke. Sistema horiek nola eraikitzen diren eta nola funtzionatzen duten ulertzen badugu, ikuspegi kritikoa garatu ahal izango dugu, eta haien aurrean modu egokiagoan kokatu.

Laburbilduz, AAren etorkizunak komunitate askotariko eta inklusibo baten esku egon behar du, bidezkoagoa eta adierazgarriagoa izan nahi badu. AAKo sistemak sortzeko unean zertarako sortu nahi ditugun erabakitzeaz haratago, gizarteko sektore bakoitzean eta norbanakoen eskubideetan izango duten eragina aztertu beharko dugu. Horretarako, ezinbestekoa izango da jakintza-arlo guztietako ezagutza eta askotariko ikuspuntuak txertatzea lantaldeetan. Horretan dabil ONDARE, AA gizartearen onurarako lelopean aritzen den eta AAren garapen-maila eta erabilera zuzena maila guztietan zabaldu nahi dituen ekimena, UPV/EHUko Informatika Fakultatean martxan jarri dena.

## 8. Erreferentziak

- BELLO, Alessandro, Tonya Blowers, Susan Schneegans eta Tiffany Straza (2021): "To be smart, the digital revolution will need to be inclusive", hemen: *UNESCO Science Report: the Race Against Time for Smarter Development*, UNESCO Publishing, Paris.
- DENNLER, Nathaniel, Mian Kian, Stefanos Nikolaidis eta Maja Mataric (2025): "Designing Robot Identity: The role of Voice, Clothing and Task on Robot Gender Perception", *International Journal on Social Robotics*, 17, 707-728.
- GARG, Nikhil, Londa Schiebinger, Dan Jurafsky eta James Zou (2018): "Word embeddings quantify 100 years of gender and ethnic stereotypes", hemen: *Proceedings of the National Academy of Sciences*, 115 (16), E3635-E3644.
- GEBRU, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III eta Kate Crawford (2021): "Datasheets for Datasets", *Communications of the ACM*, 64(12), 86-92, hemen dago eskuragarri: [<https://arxiv.org/abs/1803.09010>].
- HIPÓLITO, Ines, Katie Winkle eta Merete Lie (2023): "Enactive artificial intelligence: subverting gender norms in human-robot interaction", *Frontiers in Neurobotics*, 17.
- LONDOÑO, Laura, Juana Valeria Hurtado, Nora Hertz, Philipp Kellmeyer, Silja Voenekey eta Abhinav Valada (2024): "Fairness and Bias in Robot Learning", hemen: *Proceedings of the IEEE*, 112(4), 4, 305-330.
- MARVIN, Minsky eta Seymour Papert, (1969): *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, MA, AEB.
- MITCHELL, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah eta Timnit Gebru (2019): "Model cards for model reporting", hemen: *ACM Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220-229.
- NEWELL, Allen, John Clifford Shaw eta Herbert Alexander Simon (1958): "Elements of a theory of human problem solving", *Psychological Review*, 65(3), 151-166.
- NILSSON, Nils J. (1984): *Shakey the robot*, Stanford Research Institute (SRI International), Artificial Intelligence Center, Technical note 323.
- OBBERMEYER, Ziad, Brian P. Owers, Christine Vogeli eta Sendhil Mullainathan (2019): "Dissecting racial bias in an algorithm used to manage the health of populations", *Science* 366(6464), 447-453.
- PEREIDA, Karime eta Melisa Greeff (2019): "Diversity in robotics: from diverse teams to diverse impact", hemen dago eskuragarri: [<http://www.dynsyslab.org/wp-content/papercite-data/pdf/pereida-icra19b.pdf>].
- ROSENBLATT, Frank (1958): "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review*, 65(6), 386-408.
- RUMELHART, David Everett eta James Lloyd McClelland (1986): "Parallel Distributed Processing", *Explorations in the Microstructure of Cognition*, MIT Press, Cambridge.
- SEABORN, Katie (2024): "Bots against Bias: Critical Next Steps for Human-Robot Interaction", hemen: Barfield W, Weng Y-H, Pagallo U, eds. *The Cambridge Handbook of the Law, Policy, and Regulation for Human-Robot Interaction*, Cambridge Law Handbooks, Cambridge University Press, 362-390.

- SHORTLIFFE, Edward Hance (1977): "Mycin: A Knowledge-Based Computer Program Applied to Infectious Diseases", hemen: *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 66-69.
- SURESH, Harini eta John Guttag (2021): "A framework for understanding sources of harm throughout the machine learning life cycle", hemen: *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1-9, hemen dago eskuragarri: [<https://arxiv.org/abs/1901.10002>].
- TURING, Alan (1950): "Computing Machinery and Intelligence", *Mind*, 59, 433-460.
- VAN ALTEN, Sjoerd, Benjamin W. Domingue, Jessica Faul, Titus Galama eta Andries T. Marees (2024): "Reweight UK Biobank corrects for pervasive selection bias due to volunteering", *International Journal of Epidemiology*, 53(3), dyae054.
- WIENS, Jenna, Suchi Saria, Mark Sendak, Marzyeh Ghassemi, Vincent X Liu, Finale Doshi-Velez, Kenneth Jung, Katherine Heller, David Kale, Mohammed Saeed, Pilar N Ossorio, Sonoo Thadaney-Israni eta Anna Goldenberg (2019): "Do no harm: a roadmap for responsible machine learning for health care", *Nature Medicine* 25, 1337–1340.
- WOOLDRIDGE, Michael, (2021): *A Brief History of Artificial Intelligence. What it is, Where we are and Where we are going*, Flatiron Books, New York.

1. **Garapenari buruzko pentsamenduaren bilakaera: teoriak, estrategiak eta adierazleak.** Luis Guridi, Hegoa Institutua Ekonomia Aplikatua I Saila. 2017.
2. **Garapena eta iraunkortasuna: bilakaera, bateragarritasuna eta ikuspegiak.** Iker Etxano Gandariasbeitia, Ekonomia Aplikatua I Saila, UPV/EHU. 2017.
3. **Garapena feminismotik aztertzen.** Yolanda Jubeto Ruiz, Mertxe Larrañaga Sarriegi, Hegoa Institutua, Ekonomia Aplikatua I Saila, Euskal Herriko Unibertsitatea UPV/EHU.
4. **Garapenerako lankidetza eta bere bilakaera.** Irati Labaien Eiguren, Jorge Gutiérrez Goiria, Hegoa Institutua, Euskal Herriko Unibertsitatea UPV/EHU.
5. **Amildegari so: krisi ekosozialari aurre egiteko zenbait oinarri.** Joseba Azkarraga Etxagibel, Soziologia Saila, Euskal Herriko Unibertsitatea UPV/EHU.
6. **Migrazioak eta garapena Win-win agertokia posible ote?** Gorka Moreno Márquez, Xabier Aierdi Urraza, Ikuspegi – Euskal Immigrazioaren Behatokia. Euskal Herriko Unibertsitatea.
7. **Ekonomia soziala: Ikerketa-objektua eta analisirako lanabesak.** Enekoitz Etxezarreta Etxarri, Juan Carlos Pérez de Mendiguren Castresana, Ekonomia Aplikatua I Saila. Gezki Institutua. Euskal Herriko Unibertsitatea UPV/EHU. Enpresen Antolaketa Saila. Hegoa Institutua. Euskal Herriko Unibertsitatea UPV/EHU.
8. **Globalizazioa eta tokiko garapena.** Mikel Zurbano, Elena Martinez Tola, Xabier Gainza, Pablo Arrillaga, Ekonomia Aplikatua I Saila, Euskal Herriko Unibertsitatea.
9. **XXI. mendeko pobrezia. Benetan ezabatzeko proposamenak.** Alfonso Dubois Migoya, UPV/EHUko irakasle erretiratua, Hegoa Institutua.
10. **Elikadura-sistemak: Elikadura Burujabetzaren eta Agroekologiaren estrategia askatzaileak.** Efen Areskurrinaga Mirandona, Ekonomia Aplikatua I UPV/EHU eta Hegoa Institutua. Mirene Begiristain Zubillaga, Finantza Ekonomia II UPV/EHU eta Hegoa Institutua. Eduardo Malagón Zaldua, Ekonomia Aplikatua V UPV/EHU Hegoa Institutua.
11. **2030 Agenda eta Garapen Jasangarrirako Helburuak: jatorria, edukia eta jarraipena.** Andrés Fernando Herrera Herrera, HEGOIA Institutua, Euskal Herriko Unibertsitatea (UPV/EHU).
12. **Gobernantza eta herritarren partaidetza.** Jone Martínez-Palacios, Andere Ormazabal Gaston, Igor Ahedo Gurrutxaga, EHUko Politika Zientzia eta Administrazio Saileko kideak eta demokrazia partizipatiboari buruzko Parte Hartuz ikerketa-taldeko ikertzaileak.
13. **Gatazka armatuak, bakea eta garapena.** Itziar Mujika Chao eta iker zirion landaluze, UPV/EHUko Nazioarteko Zuzenbide Publikoa, Nazioarteko Harremanak eta Zuzenbidearen Historia Saileko irakasleak, eta Hegoa Nazioarteko Lankidetza eta Garapenari Buruzko Ikasketa Institutuko ikertzaileak.
14. **Finantza etikoak, ekonomia solidarioa sustatzeko tresna ekonomikoak eta politikoak.** Cristina de la Cruz-Ayuso, Deustuko Unibertsitateko irakaslea eta ikertzailea. Pedro Arrupe Giza Eskubideko Institutua.
15. **Merkatu soziala: Ekonomia Solidarioa hedatzeko estrategia.** Carlos Askunze Elizaga, REAS Euskadi. María Angeles Díez López, UPV/EHUko Hegoa Institutua eta REAS Euskadi.

16. Kontabilitate Finantzarioaren begirada zabalduz Gizarteari eta Naturari buruzko informaziora. Gizarte Kontabilitatearen jatorria, eboluzioa eta etorkizuneko erronkak. Ainhoa Garayar, Ekonomia eta Enpresa Fakultatea –Gipuzkoako Atala–. Maider Aldaz, Ekonomia eta Enpresa Fakultatea –Gipuzkoako Atala–. Igor Alvarez, Ekonomia eta Enpresa Fakultatea –Gipuzkoako Atala–. Jone Arocena, Gestoría Aguirre Gestión.
17. Txina: Garapen eredia birbideratzea eta munduko ekonomian zeregin berria. Juan Barredo-Zuriarrain, Euskal Herriko Unibertsitateko Ekonomia Aplikatua I Saileko irakasle laguntzailea. Ricardo Molero Simarro, Madrilgo Unibertsitate Autonomoko Egitura Ekonomikoko eta Garapenaren Ekonomia Saileko irakaslea.
18. Happykrazia: psikologia positiboaren hedapena eta bizitzaren psikopatologizazioan duen eragina. Maria Lopez Castillo, HEGOA/EHUko Globalizazioa eta Garapena Masterreko ikasle ohia. Galdakao-Usansolo Ospitaleko erizaina.
19. Herri Administrazioaren eraketak: berdeak, sozialak, berritzaileak... arduratsuak. Artizar Erauskin, Ekonomia Aplikatua I Saila. Euskal Herriko Unibertsitatea UPV/EHU.
20. Klima-aldaketari buruzko hitzaurrea: inpaktuak eta egokitzapena. Elisa Sainz de Murieta, Basque Centre for Climate Change (BC3), 48940 Leioa. Marta Escapa1, Basque Centre for Climate Change (BC3), 48940 Leioa, Analisi Ekonomikoko Saila, EHU.
21. Sumak kawsay, buen vivir, bizitze ona: Garapenari alternatibean bila. Unai Villalba-Eguiluz, UPV/EHUko Ekonomia Aplikatua Saileko irakasle-ikertzailea. Ekonomia eta Enpresa Fakultatea-Araba.
22. Errenta eta aberastasun banaketaren desperekotasuna munduan. Patxi Zabalo Arena. Ekonomia Aplikatuko Institutua eta Hegoa Institutua, UPV/EHU.
23. Pobrezia generoaren ikuspegitik aztertze gakoak. M<sup>a</sup> Luz de la Cal Barredo. Universidad del País Vasco/Euskal Herriko Unibertsitatea. Lan Harremanen eta Gizarte Lanaren Fakultatea.
24. Nazioarteko merkataritza globalizazioaren garaian: aurrekariak, ezaugarriak, arauak, eztabaidak, eta ingurumariak. Eduardo Bidaurratzaga Aurre, Elena Martinez Tola. Hegoa-UPV/EHU.
25. Turismoa garapen-tresna gisa lurralde pobretuetan. Josu Benito Andikoetxea. Garapena eta Nazioarteko Lankidetzeta Masterreko ikasle ohia. HEGOA Institutua UPV/EHU.
26. Etxebizitza beharra eta ekonomia kolaboratiboa: cohousing fenomeno. Maddi Montero. Ekonomia eta Enpresa zientzietan Graduatua da. Aitziber Etxezarreta. Ekonomian doktorea eta UPV/EHUko Ekonomia Aplikatua Saileko irakaslea da, eta GEZKIko (Gizarte Ekonomia eta Zuzenbide Kooperatiboaren Institutua) idazkaria.
27. Megaproiektuak Hazten ari den fenomeno korporatibo baten ikuspegi globala. Gonzalo Fernández. OMAL-Paz con Dignidad elkarteko ikerlaria. Erika González. OMAL-Paz con Dignidad elkarteko ikerlaria. Juan Hernández. OMAL-Paz con Dignidad elkarteko ikerlaria. Pedro Ramiro. OMAL-Paz con Dignidad elkarteko ikerlaria.
28. Gizarte Eraldaketarako Hezkuntza edo itxaropenaren alde borrokatzeko erronka. Gema Celorio Díaz. Hegoaren Hezkuntza Taldea. Juanjo Celorio Díaz. Hegoaren Hezkuntza Taldea. Amaia del Río Martínez. Hegoaren Hezkuntza Taldea. Iris Murillo Hidalgo. Hegoaren Hezkuntza Taldea.
29. Komunikazioa eta garapena: joan-etorriko bidaia bat. Laura Feal Sánchez.
30. Munduko ekonomiaren erronka demografikoak. Ángeles Sánchez Díez. Madrilgo Unibertsitate Autonomoa.

31. **Zaintza eta zahartzaroa: XXI. mendeko erronkak.** Matxalen Legarreta-Iza. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea UPV/EHU. Uzuri Castelo Moñux. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea UPV/EHU. Alaitz Uriarte Goikoetxea. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea UPV/EHU. Maider Barañano Uribarri. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea UPV/EHU. Marina Sagastizabal Emilio-Yus. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea UPV/EHU.
32. **Ekofeminismoak. Bizitzari eustea eta gorputza-lurra lurraldea zaintzea.** Amaia Pérez Orozco. Colectiva XXK. Feminismoak, pentsamendua eta ekintza.
33. **Sexu eta genero aniztasuna aztertzeko gakoak.** Aimar Rubio Llona. Hegoako Kolaboratzailea, UPV/EHU.
34. **Zientzia eta feminismoak: zientziari bizitzaren iraunkortasunetik begiratzen.** Vane Calero Blanco. Sorkin, alboratorio de saberes / jakintzen iraultegia.
35. **Garapen jasangarriaren oinarri esentzialistak. 2030 Agendaren inguruko gogoeta filosofikoa.** Juan Telleria. Filosofia Saila, PRAXIS ikerketa taldea eta Hegoa Institutua. Euskal Herriko Unibertsitatea UPV/EHU.
36. **Inurriak baikin.** Kolapsoa ez da saihestezina gizarte konplexuetan. Gorka Bueno. Ekopol ikerketa-taldea. Euskal Herriko Unibertsitatea UPV/EHU.
37. **Gizarte desberdintasunak osasunean: zer eragin dute euskadiko herritarrengan?** Amaia Bacigalupe. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea (UPV/EHU). Osasunaren Gizarte Determinatzaile eta Aldaketa Demografikoari buruzko Ikerketa Taldea – OPIK. Erika Valero. Osasunaren Gizarte Determinatzaile eta Aldaketa Demografikoari buruzko Ikerketa Taldea – OPIK. Erizaintza 2 Saila. Euskal Herriko Unibertsitatea (UPV/EHU). Yolanda González-Rábago. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea (UPV/EHU). Osasunaren Gizarte Determinatzaile eta Aldaketa Demografikoari buruzko Ikerketa Taldea – OPIK. Unai Martín. Soziologia eta Gizarte Langintza Saila. Euskal Herriko Unibertsitatea (UPV/EHU). Osasunaren Gizarte Determinatzaile eta Aldaketa Demografikoari buruzko Ikerketa Taldea – OPIK.
38. **Zer da Desazkundera?** Adrián Almazán. THECO ikerketa-taldea, Madrilgo Karlos III.a Unibertsitatea.
39. **Etika eta garapena.** Imanol Zubero. Soziologia eta Gizarte Langintza Saila, UPV/EHU.
40. **Lanaren Ekonomia, lan merkatuaren azterketa soziologikotik Euskal Herriko lan harremanen garapen juridikora arte.** Endika Alabort Amundarain. Euskal Herriko Unibertsitatearen Politika Publikoak eta Historia Ekonomikoko Saileko irakaslea. Andrea de Vicente Arias. Euskal Herriko Unibertsitatearen Enpresa Zuzenbidea eta Zuzenbide Zibila Saileko irakaslea. Eki Etxebarria Respaldiza. Herriko Unibertsitatearen Soziologia eta Gizarte Langintza Saileko irakaslea. Jon Bernat Zubiri Rey. Euskal Herriko Unibertsitatearen Ekonomia Aplikatua Saileko irakaslea.
41. **Zientzia eta Teknologiarri begirada giza garapenetik eta eraldaketa sozialetik.** Unai Villena Camarero, Joseba Sainz de Murieta Mangado. Euskal Herriko Unibertsitatea (UPV/EHU).
42. **Genero zeharkakotasunetik interseksionalitatera: genealogiak, garapen instituzionala eta eraldaketarako aukerak politika publikoetan.** Uxue Zugaza Goienetxea, Miriam Ureta García. Politika eta Administrazio Zientzien saileko irakasle eta ikerlariak, eta Parte Hartuz ikerketa taldeko kideak.
43. **Kazetaritza feminista: interseksionalitatea txertatzeko ahalegina.** Maria Gorosarri González. EHUko irakaslea eta ikertzailea.
44. **Adimen artifiziala eta generoa.** Ondare. Naiara Aginako Bengoa, Olatz Arbelaitz Gallego, Borja Calvo Molinos, Iñaki Inza Cano, Itziar Irigoien Garbizu, Elena Lazkano Ortega, Alex Mendiburu Alberro, Usue Mori Carrascal, Olatz Perez de Viñaspre Garralda, Ana Zelaia Jauregi.

