

Adimen Artifizialaren ahalmenak eta mugak. Laburpena

Adimen artifizialaren (AA) erabilera gero eta zabalduegoa dago azken urteotan, eta aplikazio asko daude hainbat alorretan, hala nola osasungintzan, materialen zientzietan eta giza ezagutzan. Artikulu bilduma honetan, hainbat arlotako adituek AAren potentziala eta mugak aztertzen dituzte, baita gizartean eta giza eskubideetan duen eragina ere.

“Mankind questions its status” artikuluan, Gomez-Pin doktoreak aztertzen du giza espeziearen berezitasunari buruzko sinesmena beste espezie animatuekiko, eta nola sinesmen hori zalantzan jarri den gizakiak bezala pentsa eta ikas dezaketen izaki artifizialen garapenean. Artikuluan aztertzen da ea izaki artifizial horiek gizakiek duten gaitasun bera duten judizio kognitibo, etiko eta estetikoetarako, eta ea horrek gizakiaren murrizteztasunaren ideia zalantzan jartzen duen. Autoreak garapen horren inplikazio filosofikoak ere aztertzen ditu.

Pinillos doktoreak, “Artificial Intelligence is created by humans” artikuluan, AAK gizartean duen eragina aztertzen du, eta horren inguruan jarrera neutrala izateko beharra. AAK onura sozial eta ekonomiko handiak ekar ditzakeen arren, alderdi interesdun askoren kezka eta ziurgabetasuna ere eragiten ari da. Artikuluak argudiatzen du AA gizakiek sortu dutela eta giza espeziearen mesedetan erabili behar dela, eta eskatzen du finantzazio handiagoa ikerkuntzarako eta AAren erabilera etikoa bermatuko duen arau-esparru bat garatzeko. Era berean, adierazten du espero dela AEBek eta Txinak AIren onura handiena lortzea datozen urteetan.

“Five crucial challenges for regulation of Medical Artificial Intelligence” artikuluan, Alkorta doktoreak AA medikoak giza eskubideetan izan dezakeen eragina aztertzen du, batez ere pribatutasunari, produktuen erantzukizunari eta AA sistemetako alborapenaren arriskuari dagokienez. Artikuluak argudiatzen du IAren garapena eta onarpena osasunean berdintasun- eta inklusibotasun-printzipioek gidatu behar dutela, eta kontuan hartu behar direla sistema horien arrisku potentzialak eta inplikazio etikoak. Era berean, adierazten du proposatutako AAren Legeak eragina izan dezakeela arrisku handiko AA sistemen garapenean eta merkatuan sartzean, eta eskatzen du arriskueta oinarritutako ikuspegi bat, sistema horien onargarritasuna ebaluatzeko.

Bere artikuluan, “From Theodicy to Technodicy: Artificial Intelligence”, Tejada doktoreak AAren aplikazio ezberdinak berrikusten ditu eta eguneroko bizitzan eta lanean gero eta garrantzi handiagoa duela argudiatzen du. Autoreak dio AAren aurrerapenek askotan algoritmo mailako galderak eta soluzioak piztu dituztela giza eta animalien ikaskuntzaren alderdiei buruz, eta AAK hasierako emaitzak ematen dituela arazo teknologiko ezberdin asko erabat konpontzeko. Autoreak iradokitzen du, halaber, AAK azkenean giza izpiritua eta pentsamendu-prozesuak hobeto ulertzea ekarriko duela, eta etorkizuneko ahalegin zientifiko eta teknologikoak sormenaren, ametsen eta giza kontzientziaren izaera algoritmoaren bidez ulertzea bideratu daitezkeela.

Martínez doktoreak, Agudo doktoreak eta Matute doktoreak, “Biases at the heart of Artificial Intelligence” aurkeztu dute, AAn dauden giza joera kognitiboak aztertzea eta psikologiak joera horiek ulertzen eta minimizatzen nola lagun dezakeen aztertzea helburu duen berrikuspena. Egileek diotenez, AA askotan objektibo eta neutro gisa ikusten den arren, gero eta ebidentzia gehiago dago AAK ondorio larriak izan

ditzakeen aurreiritziak izan ditzakeela. Iradokitzen dute psikologiak ekarpen handia egin dezakeela AAren portaerari eta giza elkarrekintzei buruzko ikerketan, batez ere eredu informatikoak erabiliz giza ezagutza ulertzeko eta giza joera kognitiboak aztertzeko. Egileek joera kognitiboek AAan duten eraginaren adibideak ematen dituzte, besteak beste, baieztapen-joera, nagusitasun-efektua, adierazgarritasun-joera eta ainguratze-joera. Era berean, AA aztertzeko diziplina anitzeko hurbilketa baten beharra aipatzen dute, ikuspegi psikologikoak barne.

“Materials Science Powered by Machine Learning” izeneko artikuluan, Errea doktoreak eta Garcia-Lekue doktoreak makina-ikaskuntzako teknikak aurrerapen zientifikoan izan dezaketen eragina aztertzen dute, batez ere materialen zientzian. Autoreek diote makina bidezko ikaskuntzak datuetan korrelazio ezkutua deskubritu eta iragarpenak egin baditzake ere sakoneko printzipio fisikoak ulertu gabe, ez dela oinarritzeko lege zientifikoaren eta adimenaren ordekoa. Ikusten dute deskribatzaile zehatzak beharrezkoak direla ikasketa automatikoko metodoek funtziona dezaten, eta horiek eskatzen dute arazoa fisikoki ulertzea. Autoreak materialen zientzietan makina-ikaskuntzako teknikak erabiltzeak ekar ditzakeen onurak ere aztertzen dituzte, hala nola potentzial enpirikoen garapena eta sistema handien eta denbora luzeko eskalen azterketa.

Azkenik, “From Artificial Intelligence to Digital Humanities” izeneko artikuluan, Bustince doktoreak AAaren eta honek diziplina humanistikoekin duen harremanaren errepaso eskaintzen du. Autoreak ikasi dezaketen makinaren mugak eta gaitasunak aztertzen ditu, eta giza zientzietatik ez banatzearen garrantzia azpimarratzen du. Era berean, humanitate digitalen alorreko adibide laburrak ematen ditu, eta argudiatzen du humanitateak funtsezkoak direla AA modu arduratsuan garatzeko.

Amaitzeko, bilduma honetako artikuluek AAren garapenari eta eraginari buruzko hainbat alderdi jorratzen dituzte. Gainera, AAren ikerketaren egungo egoera eta etorkizunerako izan ditzakeen ondorioak ezagutzeko aukera ematen dute. Erabilera adibide gisa, ChaptGPTk, OpenAlk hezitako hizkuntza eredu handiak, sarrera eta artikulua hauen laburpenak idatzi zituen Arganda-Carreras doktorearen gutxieneko zuzendaritzapean.



Era berean, alearen azalaren irudia Midjourney irudi-sorgailuarekin sortu da, eta Diego Lasa izan da sortzailea. Irudi espresibo eta apartak sortzeko Adimen Artifizialeko iruditik testurako algoritmoak erabiltzen dituen erreminta bat da Midjourney, eta erabiltzailearen proposamenean daude oinarrituta betiere irudi horiek. Lortzen diren irudiak oso artistikoak izaten dira sarritan eta helburu askotarako erabili daitezke, hala nola giro-hondoetarako, pantaila-hondoetarako eta obra tradizionalak sortzeko inspirazio gisa. Irudi hau sortzeko mezua honako hau izan zen: *robot humanoidea Rodinen Pentsalariaren posean, zientzia-fikziozko cyberpunk estiloan ikusten diren kableekin eta mikrotxipekin, goi-teknologiako ingurune batean.*

Ignacio Arganda-Carreras¹
Argitaratzaile gonbidatua

¹ [1] Jakiunde, Zientzia, Arte eta Letren Akademia. Prim, 7. E-20006 – Donostia-San Sebastián
[2] University of the Basque Country (UPV/EHU), Computer Science Fac. Department of Computer Science and Artificial Intelligence, Donostia-San Sebastian, Spain
[3] Donostia International Physics Center (DIPC), San Sebastian, Spain.
[4] Ikerbasque, Basque Foundation for Science, Bilbao, Spain.
[5] Biofisika Institute, Leioa, Spain.

Artificial intelligence its Potential and Limits. Summary

The use of artificial intelligence (AI) has become increasingly widespread in recent years, with many applications in fields such as healthcare, materials science, and human cognition. In this collection of articles, experts in various fields explore the potential and limitations of AI, as well as its impact on society and human rights.

In “Mankind questions its status”, Dr. Gomez-Pin discusses the belief in the uniqueness of the human species in relation to other animate species and how this belief has been challenged by the development of artificial beings that can think and learn like humans. The paper addresses the question of whether these artificial beings have the same capacity for cognitive, ethical, and aesthetic judgments as humans, and whether this challenges the idea of the irreducibility of the human being. The paper also discusses the philosophical implications of this development.

Dr. Pinillos, in “Artificial Intelligence is created by humans”, discusses the impact of AI on society and the need for a neutral position on the subject. It notes that while AI has the potential to bring significant social and economic benefits, it is also causing anxiety and uncertainty among many stakeholders. The article argues that AI is created by humans and should be used for the benefit of humankind, and calls for more funding for research and the development of a regulatory framework to ensure ethical use of AI. It also notes that the United States and China are expected to benefit the most from AI in the coming years.

In “Five crucial challenges for regulation of Medical Artificial Intelligence”, Dr. Alkorta discusses the potential impact of medical AI on human rights, particularly with regard to privacy, product liability, and the risk of bias in AI systems. The article argues that the development and adoption of AI in healthcare should be guided by principles of equality and inclusivity, and should consider the potential risks and ethical implications of these systems. It also notes that the proposed AI Act could impact the development and market entry of high-risk AI systems, and calls for a risk-based approach to assessing the admissibility of these systems.

In his article “From Theodicy to Technodice: Artificial Intelligence”, Dr. Tejada reviews different applications of AI and argues for its increasing importance in daily life and work. The author notes that AI advances have often stimulated algorithmic-level questions and solutions about aspects of human and animal learning, and that AI is providing initial leads toward fully solving many different technological problems. The author also suggests that AI will ultimately lead to a better understanding of the human mind and thought processes, and that future scientific and technological efforts may focus on understanding the nature of creativity, dreams, and human consciousness through algorithmic constructions.

Dr. Martínez, Dr. Agudo and Dr. Matute present “Biases at the heart of Artificial Intelligence”, a review whose goal is to examine human cognitive biases present in AI and to explore how psychology can contribute to understanding and minimizing these biases. The authors argue that while AI is often perceived as objective and neutral, there is growing evidence that AI can contain biases with serious consequences. They suggest that psychology can contribute significantly to research on the behavior of AI and on human interactions with AI, particularly through the use of computer models to understand human cognition and the study of human cognitive biases. The authors provide examples of the impact of cognitive

biases on AI, including the confirmation bias, the primacy effect, the representativeness bias, and the anchoring bias. They also discuss the need for a multidisciplinary approach to studying AI, including the inclusion of psychological perspectives.

In the article entitled “Materials Science Powered by Machine Learning”, Dr. Errea and Dr. Garcia-Lekue discuss the potential impact of machine learning techniques on scientific progress, particularly in materials science. The authors argue that while machine learning can uncover hidden correlations in data and make predictions without a deep understanding of the underlying physical principles, it is not a replacement for fundamental scientific laws and understanding. They note that accurate descriptors are necessary for machine learning methods to work, and these require a physical understanding of the problem at hand. The authors also discuss the potential benefits of using machine learning techniques in materials science, such as the development of empirical potentials and the study of large systems and long time scales.

Finally, in his article “From Artificial Intelligence to Digital Humanities”, Dr. Bustince offers a review of artificial AI and its relationship to humanistic disciplines. The author discusses the limitations and capabilities of machines that can learn, and emphasizes the importance of not divorcing AI from the humanities. He also provides brief examples of the field of digital humanities, and argues that humanities are essential for a responsible development of AI.

In conclusion, the articles of this collection address various aspects of the development and impact of AI. Moreover, They provide insight into the current state of AI research and its potential implications for the future. As an example of use, ChaptGPT, a large language model trained by OpenAI, wrote the introduction and summaries of these articles under the minimal direction of Dr. Arganda-Carreras.



This picture was generated using the image generator Midjourney, with Diego Lasa as the creator. Midjourney is a tool that utilizes AI text-to-image algorithms to generate unique and expressive images based on user input. The resulting images are often highly artistic and can be used for a variety of purposes, such as backgrounds, desktop wallpapers, and inspiration for traditional art. The prompt used to generate this image was: *a humanoid robot in the pose of The Thinker by Rodin, in a sci-fi, cyberpunk style with visible cables and microchips in a high-technology environment.*"

Ignacio Arganda-Carreras²
Guest Editor

² [1] Jakiunde, Zientzia, Arte eta Letren Akademia. Prim, 7. E-20006 – Donostia-San Sebastián
[2] University of the Basque Country (UPV/EHU), Computer Science Fac. Department of Computer Science and Artificial Intelligence, Donostia-San Sebastian, Spain
[3] Donostia International Physics Center (DIPC), San Sebastian, Spain.
[4] Ikerbasque, Basque Foundation for Science, Bilbao, Spain.
[5] Biofisika Institute, Leioa, Spain.