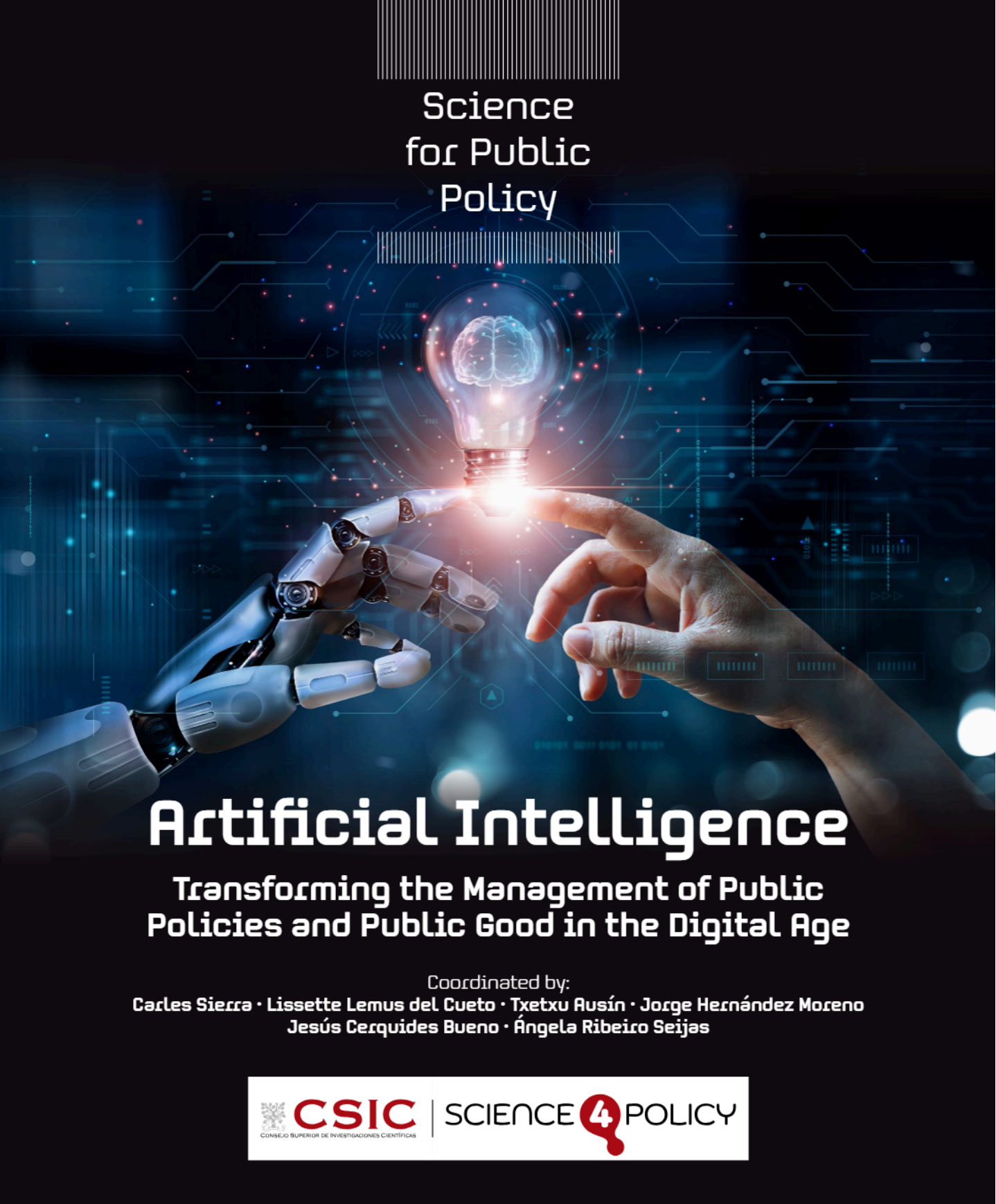




Science for Public Policy

Artificial Intelligence

Transforming the Management of Public
Policies and Public Good in the Digital Age

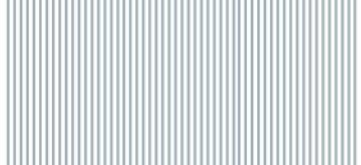
Coordinated by:

Carles Sierra · Lissette Lemus del Cueto · Txetxu Ausín · Jorge Hernández Moreno
Jesús Cerquides Bueno · Ángela Ribeiro Seijas



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

SCIENCE  POLICY



Science for Public Policy



Knowledge Transfer
Report



This is an open access book distributed under the terms of the Creative Commons Attribution 4.0 International [CC BY 4.0] Licence.

Further information about this licence is available at <https://creativecommons.org/licenses/by/4.0/>

The news, assertions and opinions contained in this work are the sole responsibility of the author(s). The publisher is solely responsible for the scientific interest of its publications.

Publications Catalogue of the General State Administration:
<https://cpage.mpr.gob.es>

EDITORIAL CSIC: <http://editorial.csic.es> (email: publ@csic.es)



**Communications
Department**

President's Office
CSIC, Calle Serrano 117
28006 Madrid
Email: comunicacion@csic.es

NIPO: 155-25-208-X

e-NIPO: 155-25-209-5

Depósito Legal: M-25563-2025

Non-commercial edition

Coordinated by:

Carles Sierra
Lisette Lemus del Cueto
Txetxu Ausín
Jorge Hernández Moreno
Jesús Cerquides Bueno
Ángela Ribeiro Seijas

Contributors:

Nardine Osman
Jordi Sabater-Mir
Carme Torras
Eva Armengol
Hannes Mueller
David Arroyo
Juan Antonio Rodríguez
Filippo Bistaffa
Raúl Mario del Toro Matamoros
José Javier Ramasco Sukia
Jorge Villagrà Serrano

**Coordinators of the Collection
Science for Public Policy:**

Jorge Hernández-Moreno
Cindy Matos Ramos

Edition:

Marta García
Caja Alta Edición y Comunicación

Illustrations:

Freepress Coop

Design and layout:

David Pamplona Roche

© Translation:

Emma Clare Barraclough

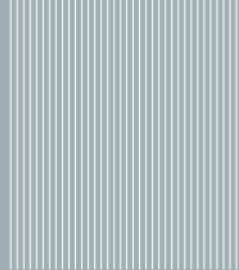
How to cite this report: Sierra, C., Lemus del Cueto, L., Ausín, T., Hernández Moreno, J. Cerquides Bueno, J., & Ribeiro Seijas A. (Coords.). [2025]. *Artificial Intelligence. Transforming the Management of Public Policies and Public Good in the Digital Age*. CSIC.

Printed in Spain

The paper used in this edition is organic paper that has undergone an ECF bleaching process, the fibre of which comes from sustainably managed forests.



This work is licensed under a Creative Commons
Attribution 4.0 International License.



ONE of the CSIC's functions is to inform, assist and advise public and private entities on science and technology, as stated in article 5 of its Articles of Association. As part of this function, the report *Artificial Intelligence. Transforming the Management of Public Policies and Public Good in the Digital Age* is presented as a document aimed at administrations and society in general. It explains how the evolution and potential impact of this technology on our society renders it a tool for effective integration into the design of public policies. It also includes some of the most outstanding lines of innovation being implemented by the CSIC in terms of the potential use of artificial intelligence in the management of policies in various fields such as education, welfare, social action and sustainability.

INDEX

one

|||||

Artificial intelligence and its role in Public Policy Management

- 1.1. Artificial intelligence: conceptual approach and regulatory framework
- 1.2. Artificial intelligence and the policy-making process in complex societies
- 1.3. Capabilities of artificial intelligence (as a technology) for public problem solving

two

|||||

How the CSIC Promotes and Develops the Use of Artificial Intelligence Technology in Different Public Sectors

- 2.1. Artificial intelligence and ethics
- 2.2. Artificial intelligence and education policies
- 2.3. Artificial intelligence and the welfare state
- 2.4. Artificial intelligence for social science and citizen participation
- 2.5. Artificial intelligence for sustainability

three

|||||

Conclusions and Recommendations

four

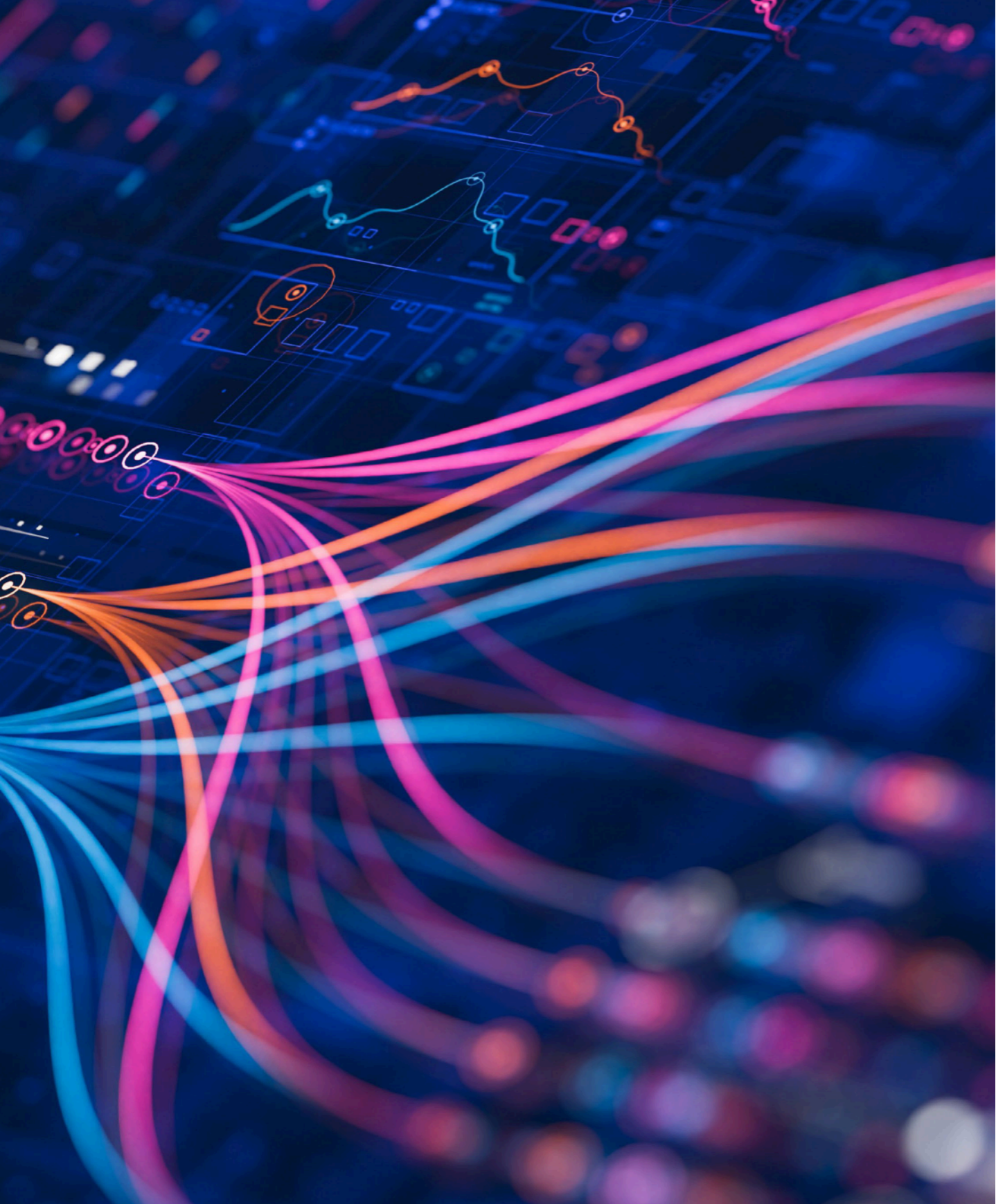
|||||

List of Centres

five

|||||

Further Information



one



Artificial intelligence and its role in Public Policy Management

Artificial intelligence: conceptual approach and regulatory framework

To counter these fears, it is imperative to redefine the economic framework in which our predominantly capitalist societies operate. It is increasingly urgent to adopt alternative economic models that promote a more equitable redistribution of wealth and guarantee resources for all citizens. Measures such as higher taxes on large corporations, a guaranteed minimum income, taxes on robots and technology, as well as continuing education throughout life, are essential for people to feel useful at all stages of their lives. In this way, the pessimistic view of progress finds solutions that strengthen social rights rather than slowing down technological progress. The answer is ultimately political, not technological.

Today we are facing a new wave of profound transformations, driven by the widespread use of artificial intelligence tools. The ability of this technology to reshape our societies is a source of both optimism and concern. Therefore, the purpose of this report is to encourage debate on how artificial intelligence can contribute to the definition of public policies that improve people's lives.

Figure 1 Artificial intelligence: a strategic compass for public policy management



Historically, artificial intelligence emerged in the 1950s. Initially, it was defined as the study and design of machines capable of solving problems that, until then, required human intelligence. For decades, research in the field of this technology remained mainly in academia. There were golden periods with generous funding for these research groups, as well as periods of less financial support, known as *AI winters*.

Throughout all these periods, progress in artificial intelligence continued steadily, especially based on two major approaches: the first, the symbolic approach, sees intelligent systems as symbol manipulators, endowing them with the ability to reason, plan and represent the world; the second, non-symbolic or sub-symbolic representations, is based on systems in which language concepts are not identifiable at a particular point, but emerge in a distributed fashion.

Recently, applications based on the second approach, in particular generative artificial intelligence, have had a considerable social impact. This branch of artificial intelligence allows systems to generate texts that appear to be written by humans, images from linguistic descriptions, or videos based on narratives, among other capabilities. However, the risks that these technologies pose to our societies, especially to democracy, have prompted public authorities to intervene.

In response to these challenges, the European Union (EU) has adopted a law to regulate artificial intelligence. It prohibits certain practices, categorises others as high-risk and establishes a number of safeguards to avoid a negative impact of this technology on our lives, especially in critical areas of use. Under this European legislation, artificial intelligence would no longer be defined only as the creation of systems that solve complex problems, but in a more pragmatic way: an intelligent system is one that uses a specific technique from a list provided by law. In short, European law states that any *software* with reasoning, learning or modelling capabilities will be considered an artificial intelligence system.

In this report, we explore through various examples how artificial intelligence, both symbolic and sub-symbolic, can help us design effective public policies. We believe that the sustainable development goals set by the United Nations are challenges to which this technology will be key.

The various examples in this report are broken down into a number of areas of interest of sufficient relevance to give the reader a clear idea of the potential of artificial intelligence across the public policy domain. In the following, we briefly describe the key ideas in these areas.

With this overview of the evolution of this technology and its potential impact on our societies, the need arises to consider how it can be effectively integrated into public policy-making. The promise of artificial intelligence is not without its challenges: from the ethical implications involved to the infrastructure needed to implement these technologies in an inclusive and secure manner. In the following section we explore how governments and public institutions can harness it to address complex challenges ranging from economic planning to improving the quality of life in urban and rural contexts, without losing sight of the principles of transparency and equity.

1.2.

Artificial intelligence and the policy-making process in complex societies

FROM a social science perspective, complex societies are characterised by being plural, fragmented (institutionally and territorially), technified, determined by the co-existence of the global and the local domains, with new and numerous social problems (of all kinds) and citizen demands that increasingly exceed the capacities of actors and public administrations (although not exclusively) to solve them. In these complex societies, policy-making faces significant challenges, and artificial intelligence can play a supportive role for the actors in charge of addressing them.

- **Challenge 1: Basing policies on sound data, empirical evidence and rigorous research to understand problems and patterns of social need.**

In this context, artificial intelligence plays a key role in facilitating the acquisition, analysis and evaluation of large volumes of data, including classification, pattern recognition and identification of trends that can inform and strengthen decision-making.

- **Challenge 2: Identifying proposals based on scientific evidence and predict sufficiently appropriate outcomes.**

The study of possible solutions must consider multiple scenarios, evaluated through a comprehensive cost-benefit analysis, including not only whether the solution satisfies the problem, but also environmental and social impact assessments. In this process, artificial intelligence is configured as a tool that facilitates the analysis of scenarios and the simulation of alternatives ("What if?"), making it possible to explore various options and visualise the impact of each proposal, thus optimising the design of sustainable and balanced strategies in different contexts.

○ **Challenge 3: Responding to the rapid technological evolution of our society.**

The dynamics of our society, with its rapid technological evolution, pose a challenge to the policy-making process, as it makes it considerably more difficult to anticipate and respond quickly to emerging and changing demands. In this context, artificial intelligence enables the creation of efficient computational predictive models to detect future trends more accurately than traditional models.

○ **Challenge 4: Understanding how policies operate within complex systems.**

Considering the interdependencies and potential domino effects of policy decisions, and selecting the best options to implement them based on a multi-criteria and multivariate approach, is a significant challenge for modern policy action. In this respect, artificial intelligence provides optimisation algorithms to analyse large volumes of data and solve complex problems more efficiently, serving as a tool to explore appropriate solutions among numerous alternatives.

1.3.

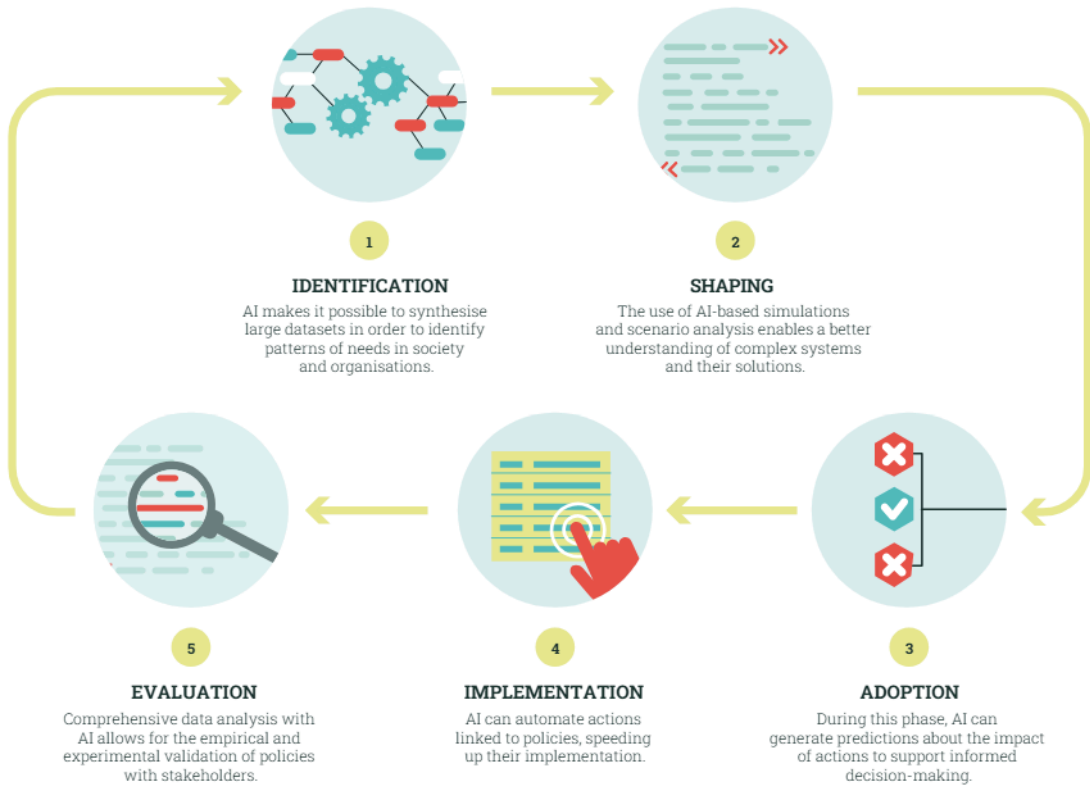
Capabilities of artificial intelligence (as a technology) for public problem solving

THIS report is part of the “Science for Public Policy” series of reports and focuses on how artificial intelligence can contribute in a cross-cutting way to public policy development in general and, in particular, to some of the issues already developed earlier in the series. In this respect, public policy-making can be seen as a process composed of several phases that include: the identification of needs, the formulation of hypotheses, the search for possible solutions, decision making on the adoption of some of the proposals, their implementation and their subsequent evaluation. The foundations of this process are based on the detection of patterns of need, the development of evidence-based programs, the prediction of outcomes and the analysis of the effectiveness of policies, all of which clearly correlate with the capabilities offered by artificial intelligence (see Figure 2).

The following are three examples of opportunities for applying artificial intelligence to different phases of the policy-making process.

- **Identification of needs:** Artificial intelligence systems are able to synthesise large data sets to detect patterns. This smart analysis capability allows the rapid identification of emerging issues of interest and is very useful in the identification phase, where gathering and analysing public opinion is of vital importance for policy makers to focus their actions on the social and political issues that are most pressing for citizens. Unlike traditional public opinion polling, artificial intelligence and data science combined facilitate the collection, synthesis and classification of large amounts of data from social networks, blogs and forums, and provide sentiment analysis, patterns and summaries of public opinion in near real-time. The Brandwatch platform, a public opinion analysis tool, was used during the COVID-19 crisis and provided a very agile, evidence-based approach to the most relevant public concerns.

Figure 2 The role of AI in the policy-making process



Source: From Patel, J., Manetti, M., Mendelsohn, M., Mills, S., Felden, F., Littig, L., & Rocha, M. (2021, March 24). AI Brings Science to the Art of Policymaking. BCG Global. <https://www.bcg.com/publications/2021/how-artificial-intelligence-can-shape-policy-making>

- **Complex scenario analysis:** Artificial intelligence multi-agent systems provide a flexible and decentralised way to simulate and analyse difficult problems. By allowing computational agents to interact and adapt, they can model scenarios of a distributed nature involving diverse entities and interactions between them and test a wide range of “what-if” scenarios. In this way, they can help decision-makers understand the consequences of different strategies and make informed decisions in complex systems such as urban transport, logistics and energy management. Work carried out in 2012 at the University of Southampton (England) argues how MAS can be crucial to developing a smart grid. As the existing electricity grid requires redesign to meet the increased demand for electric vehicles and the increased use of renewable energy sources, multi-agent models with autonomous agents, equipped with different artificial intelligence algorithms, representing consumers, electric vehicles, power generators and grid infrastructure, are proposed to optimise household energy consumption, manage electric vehicle charging and coordinate distributed generation. This would ensure that energy flows are efficient and respect grid constraints. Such a model would make it possible to carry out simulations of grid behaviour or direct implementations in which actors can make autonomous decisions and optimise energy management in real time.
- **Validation of evidence-based solutions:** It is of vital importance, in the adoption and implementation phase, to demonstrate through scientific evidence how actions that implement certain policies work. Once a rule that implements a given policy is in place, artificial intelligence, again, allows for a more thorough analysis of its results to demonstrate its empirical and experimental validity to decision-makers and stakeholders. An example of this is its use to validate the effectiveness of guidelines or health interventions. By using large-scale patient data, this technology would make it possible to validate the results of clinical analyses or compare the efficacy of different existing diagnostic methods.

As discussed above, artificial intelligence can facilitate a more comprehensive, rapid and rigorous approach to the analysis required at each stage of policy making, thus reinforcing the process with data-driven, evidence-based decision making.

Given that we are already designed as socio-technical systems, where we maintain a physical, cognitive and even emotional interaction with artificial intelligence, the ethics applied to this technology constitutes an unavoidable exercise that transcends the mere appreciation of what should be forbidden or allowed; it constitutes a reflection on our way of life, on our habits, customs and relationships (work, education, affective...). Moreover, the ethics of artificial intelligence is not an end in itself, but the privileged means to generate trust and align the development of applications, programs and services based on it with social values, with the aim of generating public trust and social appropriation of technology. To this end, different CSIC research groups and institutes promote various instruments for ethical and socially responsible artificial intelligence from a proactive rather than reactive approach to this technology, as we will see in the following section.

two



How the CSIC Promotes and Develops the Use of Artificial Intelligence Technology in Different Public Sectors

HAVING explored how artificial intelligence can support complex societies in informed policy making, there is a need to further examine how these technological capabilities are translated into concrete tools throughout the policy making process. In a context where the challenges are diverse - from data collection and analysis to prediction and validation of results - this technology cannot only streamline these processes, but also bring rigour and precision at every stage, from the identification of needs to the assessment of impacts.

This section addresses the scientific production generated by the CSIC on artificial intelligence and the resolution of ethics, education and the welfare-related challenges, and how different tools of this technology can support the development of social and economic policies, as well as provide support for a more active management of participatory democracy, citizen science and sustainability.

2.1.

Artificial intelligence and ethics

THE CSIC proposes an ethical reflection on artificial intelligence based on three fundamental questions: why, what for and how. In this context, two important considerations must be made:

1. Technological artefacts have morality, they are not neutral. Technologies, and even more so digital ones, so embedded in our lives, not only contribute to efficiency and productivity, but also shape forms of power and authority, and profoundly modify habits, customs or relationships (work, affective, educational...). Consequently, the ethics of artificial intelligence transcends the mere assessment of what should be forbidden or permitted and goes further, constituting a true reflection on our way of life.
2. We maintain a profound interdependence not only with respect to other human beings, but also in relation to technology, with which we construct our identity. And this interdependent relationship opens up the possibility of situations of vulnerability, of risks of harm. Our relationship with digital technologies is not only very close, but also tremendously asymmetrical, where the companies and institutions that design and control these devices, applications and programs hold all the power and are practically immune to citizen and public control. To this must be added the growing *autonomy* of artificial intelligence systems themselves. Consider, for example, the digital afterlife industry (*griefbots*, *deadbots* or *post-mortem* avatars) which, driven by the deployment of generative artificial intelligence, is already enabling the creation of highly realistic and interactive digital individuals.

Artificial intelligence must be safe in its interaction with humans, establishing a clear procedure for measuring, testing and certifying its functionality. Not to harm physically and mentally is the basic and overriding ethical obligation. This is the principle of non-maleficence: first and foremost, do no harm (*primum non nocere*), i.e., the duty not to harm, which covers not only intentional actions, but also the imposition of risks of harm.

Thus, consideration is given both to possible harm relating to individuals and to groups, with special attention to individuals and groups subject to other situations of vulnerability, due to age (minors and the elderly), gender, origin (migrants), employment status (precarious employment), illness or functional diversity.

With the aim - among others - of protecting these individuals and groups, the CSIC has developed several lines of research focused on aligning social values with the development of artificial intelligence technologies in the framework of what is known in the European Union as responsible research and innovation (RRI):

VAE: Value-Awareness Engineering. Oriented towards the incorporation of moral, epistemic, technical and aesthetic values in artificial intelligence systems, as well as the analysis of the possibilities offered by digital applications for personal well-being.

VALAWAI: Value-Aware Artificial Intelligence. Focused on integrating ethics into artificial intelligence systems, in particular the development of systems capable of understanding and abiding by a human value system (value-aware artificial intelligence).

INBOTS: Inclusive Robotics for a Better Society. Analyses human-machine interaction with a special focus on social and care robotics from the perspective of social studies of science, philosophy of technology and democratisation of technoscience (participatory design, citizen science and value-sensitive design).

Main actions carried out by the CSIC for the development of an ethical artificial intelligence

Following the recommendations established in specialised documents and forums (high-level expert group on artificial intelligence of the European Commission, UNESCO, IEEE, the Spanish Charter of Digital Rights, the CSIC white paper on artificial intelligence, data science and robotics), we have defined five basic ethical duties in the development and deployment of both public and private artificial intelligence systems:

Minimising harm. The basic ethical duty in the development and deployment of artificial intelligence systems is to do no harm. This principle implies the idea of security (cybersecurity), understood as a social value that should be emphasised when humans interact with delegated decision-making systems, such as artificial intelligence systems. This security includes the defence of the user and consumer against large technological corporations, labour protection, prevention of failures, *hacks* or breaches and even protection in relation to the digital power of the Administration (*checks & balances*), for which elements of good digital governance such as participation, transparency, openness or accountability are essential (elements linked to the explainability of artificial intelligence systems, which we will discuss later).

Special attention will also be paid to avoiding biases, both in data collection and in the handling and training of artificial intelligence algorithms and tools, taking special care not to reproduce or increase existing biases and social inequalities. The introduction of biases in this technology leads to discrimination (on the basis of gender, race, economic status, etc.) and can reinforce existing prejudices and stereotypes, as well as reinforce social exclusion and stratification. This is known as *algorithmic injustice*, which harms and disadvantages individuals and groups.

Maximising benefits. The potential benefits must outweigh the possible risk to individuals, communities and the environment. This relates to the promotion of well-being, the common good and the Sustainable Development Goals. Data processing through artificial intelligence must demonstrate its necessity, effectiveness and orientation towards the common good, with a well-defined and determined purpose (proportionality).

This principle of beneficence includes the protection of the environment and future generations. To this end, in order to pursue environmental sustainability and commitment to future generations, the development of infrastructures linked to artificial intelligence (clouds, data and computing centres) will promote energy and raw material efficiency, favouring the minimisation of energy and material consumption, their fair extraction, recycling and the use of renewable and clean energy (see chapter on sustainability).

Autonomy. Artificial intelligence processes must be people-centred and support the right of people to retain the power to decide what decisions to make and to exercise freedom of choice when necessary. Humans must retain the power to ultimately decide which decisions to make and which to delegate to artificial intelligence systems.

Surveillance of individuals and brain modification, which affects people's own sense of autonomy and identity, must be prevented. Consequently, on the one hand, personal data (privacy) will be protected, avoiding re-identification, data leakage and lack of transparency in data collection. In this respect, the CSIC advises the Spanish Data Protection Agency (**AEPD**) on the ethical deployment of artificial intelligence.

The CSIC also forms part of international initiatives such as the Institute of Neurotechnology and Law, which focuses on the analysis and development of human neuro-rights: the right to mental privacy, the right to mental integrity and the right to psychological continuity.

Figure 3 Discrimination. Algorithmic injustice



Justice. It implies, on the one hand, recognition (inclusivity), taking into account all stakeholders (the diversity of actors and values) in the process. Likewise, those who do not freely wish to access the digital environment cannot be excluded from the use and development of goods and services (principle of non-exclusion).

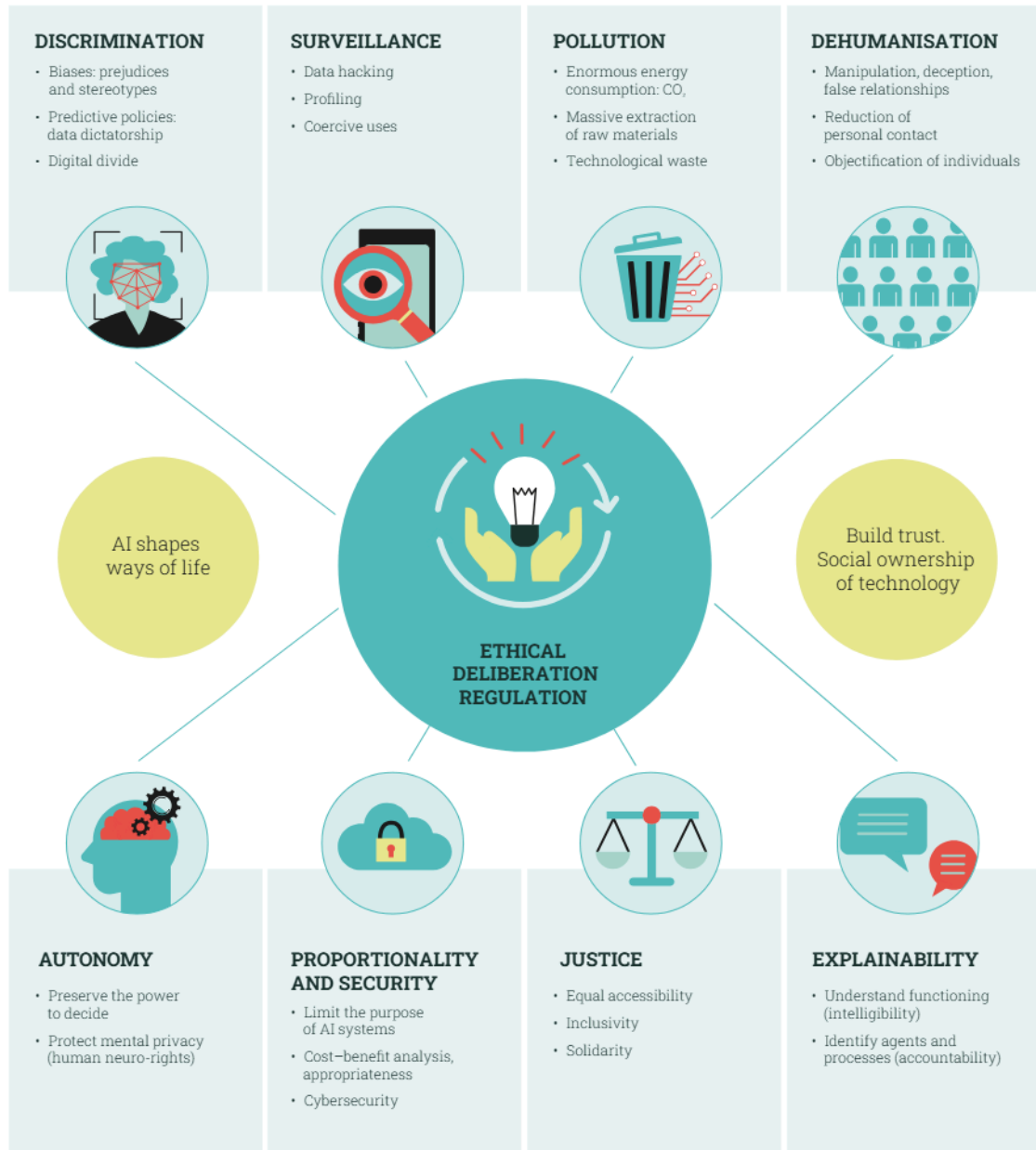
On the other hand, the search for equality and accessibility (equity) is required. A new inequality gap due to artificial intelligence must be avoided, on top of the already existing digital divide. Access to this technology cannot be conditioned by technological complexity or lack of resources, nor by options that force people to become users of specific companies.

Explainability. The opacity and lack of transparency of artificial intelligence processes must be prevented. This means that it is possible to understand how they work (intelligibility) and to identify the decision-making process, the people involved and the consequences (accountability). Accordingly, the understanding and interpretability of artificial intelligence systems will be facilitated, avoiding as far as possible the arbitrary nature of the so-called *black boxes* (whose inner workings are not known) and acting proactively to provide information to citizens.

In this sense, the **AI-CONTROL** project, led by the CSIC, aims to develop a governance model for artificial intelligence systems based on complexity sciences, with a holistic approach that does not separate social and technical aspects, integrating philosophy, politics and engineering as complementary disciplines.

Likewise, the CSIC project **DesafIA2030** is developing a collaborative strategy for the formulation of ethical guidelines for artificial intelligence, promoting awareness-raising and education tools about the challenges posed by this disruptive technology.

Finally, it should be noted that the CSIC Ethics Committee is promoting the scientific-technical project **ETHICAL**, within the Momentum Programme for attracting and retaining digital talent, on the ethical assessment of artificial intelligence projects.

Figure 4 Ethics and society

OWN CREATION

2.2.

Artificial intelligence and education policies

It is always difficult to make predictions about the impact of technology on an economic or social sector. However, all recent analyses make it clear that repetitive tasks, or tasks with little added value on the part of the humans who perform them, will be redesigned to facilitate their automation through the use of artificial intelligence techniques. Banking and commerce are examples of sectors that are undergoing a profound transformation, in part facilitated by AI techniques such as *chatbots* or personalisation systems, leading to a significant reduction in jobs. In contrast, the education sector will continue to need the human component, and on a permanent basis, as schools fulfil a basic socialising function for the development of individuals. However, this does not mean that artificial intelligence will not have an impact on educational processes, as it has already done so, does so and will continue to do so.

The design of educational policies that take artificial intelligence into account goes back a long time. In May 2019, around a hundred member states of UNESCO made a series of recommendations that mark the path to be followed in the coming years. The first and most relevant is that this technology must be integrated into the education system. It must be taught and, at the same time, used to strengthen students' learning. This integration and use must be based on a scrupulous respect for human rights and serve to train students to be critical about the use of this technology, so that they can understand the risks and take advantage of the opportunities it offers.

Following UNESCO's recommendations and integrating artificial intelligence into education policies is already perfectly feasible. Today, we have numerous artificial intelligence applications, not all of which have been developed specifically for education, although they are very useful in this field. For example, automatic subtitling of videos, tutoring systems that interact in natural language, realistic conversion of text to human

Figure 5 Adaptive learning



OWN CREATION

speech, preparation of learning situations or student assessment, among many others. The recent emergence of artificial intelligence with systems such as ChatGPT is an opportunity to improve the interaction between students and intelligent tutors, while at the same time it poses a risk of potential misuse, which educators will have to face. At this point, it is important to reflect on the recent law regulating artificial intelligence. In particular, the assessment of pupils using artificial intelligence techniques or the assignment of pupils to schools are high-risk activities and therefore require special supervision by the authorities. This needs to be taken into account when designing education policies based on this technology.

Artificial intelligence, in its origins, has already been applied to education, particularly personalised education. Adapting content to each student is a pedagogical imperative that is difficult for teachers to achieve in large groups or with limited economic resources. Simple systems of personalised education were developed in the 1960s and have reached a remarkable level of sophistication. For example, over the last fifteen years the ALEKS system (aleks.com), developed in the United States, has improved the performance of millions of students in the area of mathematics. This system poses open-ended problems, analyses the answer and, thanks to a machine learning system, identifies errors and skills not acquired so that it can explain the error to the student and recommend new problems to help them acquire the necessary skills. Such systems continue to be developed in different countries. In particular, the system developed by Squirrel AI Learning in Shanghai, with more than three million students, has achieved a great improvement in individual performance (<http://squirrelai.com/product/ials>). We will undoubtedly see these systems more frequently over the next decade and covering areas increasingly outside the STEM (*Science, Technology, Engineering and Mathematics*) disciplines on which they are currently focused. The Instituto de Investigación en Inteligencia Artificial [Artificial Intelligence Research Institute] (IIIA-CSIC) has also been working on personalised education using artificial intelligence tools. In particular, a content personalisation system has been developed for an online learning platform aimed at short courses.

The use of artificial intelligence by teachers is being extended to most tasks. For example, the IIIA-CSIC has developed a recommender system that supports the teacher in planning a course in the context of formal education. This system, developed together with a company in the sector, is adapted to the needs of the new educational methodologies (the development

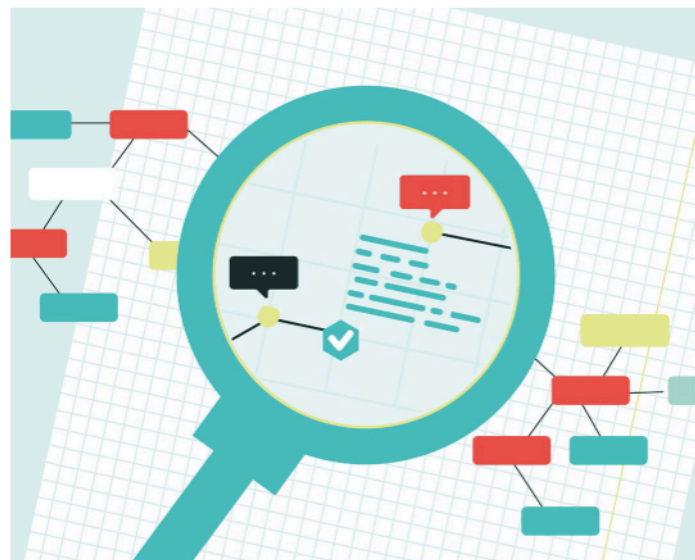
Figure 6 Human teams. Collaborative learning

OWN CREATION

of competences, as opposed to the acquisition of knowledge only) and, specifically, to the new education law (LOMLOE). In a first phase, prior to the start of the course, the system helps the teacher to draw up a sequence of didactic units that are adapted both to the objectives of the course in terms of the competences and knowledge to be worked on and to a set of teacher preferences. Once the course has started, the system monitors the class and proposes alternatives when the initial plan does not fit the evolution of the students (e.g., if the skills that are supposed to have been acquired at that point in the course have not been assimilated to the expected degree).

Within collaborative education, and based on tasks, one of the recurring problems is how to form teams of learners. Artificial intelligence makes it possible to analyse a multitude of factors (sociological, competence, psychological, etc.) to explore the huge space of possible combinations and find the optimal teams of learners. A system such as Eduteams (eduteams.iii.csic.es), developed at the IIIA-CSIC, distributes the students in a classroom into teams of a certain size, obtaining an appropriate balance of personalities and trying to ensure that at least one student in each team possesses each of the skills required by the task to be performed. The evaluation of the results of these teams is significantly better than that of those formed by the tutors. In addition, the CSIC is currently engaged in using these techniques for the education of teams of young people in various African countries in collaboration with UNICEF. Recent work has made it possible to predict the dedication of students depending on the team they belong to and the method of evaluation. These new models will allow a much more effective design of teams and tasks for the improvement of students' competences.

Figure 7 Assessment of learners



OWN CREATION

Another area where artificial intelligence will have a significant impact is in assessment processes. Progress will be made in automatic and peer assessment processes, which will further democratise *online* and lifelong learning. Advances in natural language processing and computer vision, combined with explainability techniques, will make the self-assessment that systems provide to learners much more informative and useful. In addition, peer assessment combined with artificial intelligence techniques will make the assessment of large groups of online learners acceptable to teachers. For example, in the case of a massive open *online* course with thousands of students enrolled, the assessment of a large number of exercises may simply not be feasible for the teacher, and very often it is necessary to rely on the opinions of the students themselves. In this respect, a model has been developed at the CSIC that uses peer reviews (by the students themselves) to approximate what the teacher would do if given the opportunity. In addition, the model is conservative and provides a measure of the uncertainty of the calculated assessments and a recommendation on the best exercises to assess next in order to reduce that uncertainty.

Artificial intelligence will impact all aspects of education policy, including assessment methods, personalised and collaborative learning, school management and curriculum design. For all of these, the technology will be useful in all the steps of the policy management cycle mentioned in the introduction to this report, in particular, in the identification of needs expressed by parents and learners; in impact analysis using multi-agent simulation (one of the most widely used techniques in current development work); and in the generation of scientific evidence.

If we integrate artificial intelligence into the world of education wisely, the results can be extraordinary.

2.3.

Artificial intelligence and the welfare state

THE welfare state, as an institution, is one of the most significant achievements of contemporary societies. Emerging in Europe in the last quarter of the 19th century, it has demonstrated for decades its capacity to protect the vast majority of the population through the development of fundamental programmes and policies (see Figure 8), combating different risk situations such as illness, unemployment, the progressive increase in the number of elderly people or the lack of basic goods and resources needed to live with dignity.

However, welfare states and social policies are in a constant process of reform in order to continue to serve societies. Megatrends such as the accelerating pace of technological change, financial and cultural globalisation, mass migrations, gender implications or new risks linked to climate change and global pollution are putting pressure on and profoundly affecting the nature, sustainability and functioning of welfare states (see Figure 9). In a 2019 report, the United Nations Special Rapporteur on extreme poverty and human rights addressed for the first time (in an institutional manner) the notion of the *digital welfare state*, in a clear reference to those countries that have opted to try to improve the efficiency and effectiveness of their social protection systems through the use and development of *big data* and, in particular, artificial intelligence technology.

Under this social, technological and regulatory reality, the CSIC, through several of its research centres and institutes, carries out extensive scientific activity to produce knowledge and evidence, which is extremely useful for informing the public policy-making process and influencing the day-to-day running of several of the sectors that currently make up the welfare state.

Figure 8 Basic institutions and programmes of the welfare state (19th and 20th centuries)



Figure 9 Megatrends putting pressure on the digital welfare state



Health and social services: research and improvement in service delivery

In the field of health, the CSIC is working intensively on the development of new health techniques, programmes and treatments. CSIC researchers are providing key information to improve the care and attention of children affected by autism spectrum disorders (ASD). Specifically, several teams are focusing on research to prevent and anticipate potential crises related to the way people interact with others, communicate, learn and behave. To do this, they monitored a set of signals and then, using time series analysis and classification algorithms, were able to characterise ten behaviours that could lead to crises. The CSIC is also developing research aimed at personalising cognitive stimulation and rehabilitation programmes for patients who have suffered a stroke. These studies focus especially on causal information regarding functional or structural impairment profiles and temporal relationships, as well as maps of cognitive rehabilitation processes. The aim of these programmes is to determine the exact relationships between cognitive impairment profiles and therapeutic plans. Artificial intelligence constitutes the base technology on which the tools for the personalised generation of therapeutic proposals are designed. The CSIC is also involved in the creation of new devices and techniques for tracking and monitoring certain types of patients and ailments. For example, several researchers from this institution have developed, together with several partners, a *Holter* system for ambulatory electroencephalogram (EEG) recording using portable, discreet and non-invasive devices. These allow the patient to perform, in an easier and more comfortable way, activities of daily living while their electrical brain activity is recorded. The specific role of the CSIC groups focuses on analysing the signal produced by this system using artificial intelligence algorithms.

The CSIC, as an institution, is also part of important international networks and acts as an active coordinator and promoter of multidisciplinary projects. For example, it is part of the European network Eurova (European Oocyte Biology Research Innovation Training Network), founded by the European Commission.

Through this network, the CSIC aims to improve the efficiency of assisted reproduction techniques. The main objectives are to predict the success rates of these techniques, identify the gametes and embryos most likely to develop and improve the diagnostic and prognostic capabilities of current methods. The CSIC also acts as a coordinator of different actors (both public and private) at the European level. The aim of these projects is to determine the effects that indoor air quality can have on human health. To this end, the work focuses on the implementation of an extensive monitoring campaign of chemical and biological pollutants in different EU countries, as well as the analysis of pollution sources and the interactions and main correlations with certain health problems.

Co-creation and deployment of assistive robotics

People are living longer than ever before around the world. The number of people aged 80 and over is expected to triple between 2020 and 2050, to 426 million. According to the Instituto Nacional de Estadística [National Statistics Institute] (INE) and Eurostat, the dependency ratio at age 67 (the ratio of the population aged 67 and over to the population aged 16 to 66, as a percentage) is set to rise steadily, peaking around 2050. Spain would be the country with the fourth highest dependency ratio (around 60%), behind Portugal, Greece and Italy. In terms of the increase in this rate, Spain would have the greatest difference with respect to 2019. According to the World Health Organisation (WHO), this change in the distribution of the population towards older ages represents one of the most important challenges for the sustainability of health and social services systems worldwide: from the point of view of the expenditure associated with ageing, as well as in relation to the human resources needed to meet the strong and growing demands for care work (see Figure 10). In this respect, the CSIC is studying and developing solutions to reduce this burden and cushion this impact in the coming years. To this end, many of the projects focus on the field of care robotics and the co-design of human-robot interaction, applicable, as we have said, to both health and social services.

Figure 10 Care workers in the digital welfare state



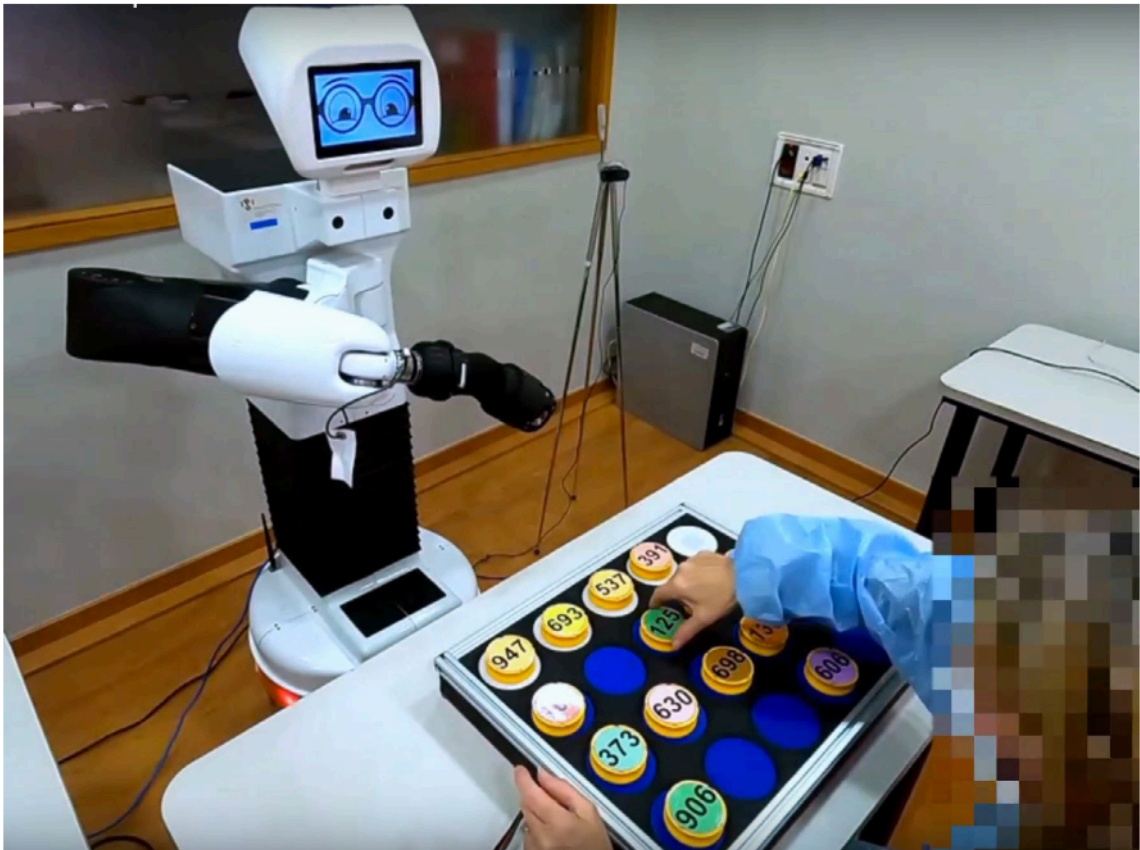
Source: Adapted from Sánchez Rosado, J. C., & Díez Parra, M. (2022). Impacto de la inteligencia artificial en la transformación de la sanidad: Beneficios y retos. *Economía industrial*, 423, 129-144.



Prototype of a robotic feeding system, co-designed by the Institut de Robòtica i Informàtica Industrial (CSIC-UPC) and the Parc Sanitari Pere Virgili in Barcelona

Assistive robotics is emerging as a decisive technology to help caregivers in hospitals, rehabilitation centres, nursing homes and domestic help, to relieve them of routine tasks with little added human value. The aim of this technology is to increase the quality of life and autonomy of the elderly or people with physical or cognitive limitations.

CSIC researchers are instrumental in the development and prototyping of physical and cognitive assistive robots capable of manipulating objects in proximity to or in contact with people (i.e., help with dressing, shaving, eating, standing up or moving, as well as rehabilitation and cognitive training). This project has been developed in collaboration with healthcare institutions, medical staff, therapists, social workers, caregivers and patients. Lessons learned include the importance of interdisciplinarity, human-centred deployment, personalisation and training in technoethics. Through an interactive process, the prototypes have been enhanced with new functionalities that broaden their usability, making them more likely to be adopted by institutions,



Prototype of a robotic system to provide cognitive training, co-designed by the Institut de Robòtica i Informàtica Industrial (CSIC-UPC) and the Ace Alzheimer Center Barcelona

thus increasing their cost-effectiveness and making them more attractive for commercialisation. These research prototypes can be adapted to the needs of an entity or type of patients to provide them with autonomy, while being of great help to caregivers. In the field of health and social services, the ethical deployment of this technology is paramount. Under the umbrella of the ethical guidelines for reliable artificial intelligence of the European Commission's High Level Expert Group on Artificial Intelligence and the report *Ethically Aligned Design A Vision for Prioritising Human Well-being with Autonomous and Intelligent Systems* by the Standards Association of the Institute of Electrical and Electronics Engineers (IEEE), CSIC researchers are also developing training materials and activities on ethics in social robotics and artificial intelligence.

Employment and universal basic incomes: the usefulness of artificial intelligence in other social policy sectors

The influence of artificial intelligence tools in shaping labour markets and employment policies is a complex and multifaceted issue that is generating intense debate in today's society. CSIC researchers are actively working on the development of this line of research, looking especially at the challenges it presents in terms of labour inequality and the possible widening of the wage gap between workers with and without technological skills, as well as the need for continuous adaptation (acquisition of new skills to work alongside artificial intelligence systems). They are also considering the opportunities and challenges opening up for labour markets in relation to increased productivity, the creation of new jobs and the potential of this technology to improve working conditions, linked to greater flexibility and the establishment of remote working. CSIC researchers are actively working on the analysis of the variables that would determine the success of this transition through the development of specific policies and strategies. These studies focus mainly on the implications and scope for effective collaboration between actors such as governments themselves, key international institutions and organisations in the sector (such as the International Labour Organisation), the role of the business sector, trade union organisations and educational institutions (at all levels and scales).

Artificial intelligence tools also have the potential to significantly influence the development and management of universal basic incomes (UBI) or universal basic wages. In the case of Spain, the Public Administration Digitalisation Plan 2021-2025 (which underpins the institutional and organisational performance of practically all central government departments) envisages (and, in some cases, has adopted) intelligent automation and *machine learning* services, for example in the deployment of the minimum living income or ingreso mínimo vital (IMV). CSIC researchers are also attempting to boost the production of scientific evidence linking artificial intelligence to the implementation processes of these monetary allocations, in relation to their efficient distribution, personalisation and continuous evaluation processes. The CSIC is further working intensively to increase knowledge and understanding of the ethical and privacy challenges, which are fundamental to ensure that the integration of this technology in the management of these instruments is fair and equitable. Finally, research is also being carried out on the need for a collaborative (including society itself) and institutionalised approach to maximise the benefits of this technological tool and minimise its potential risks.



2.4.

Artificial intelligence for social science and citizen participation

THIS section reviews some of the contributions made by the CSIC to reduce misinformation, improve citizen participation and develop artificial intelligence-based tools for policy management..

Artificial intelligence and information clutter

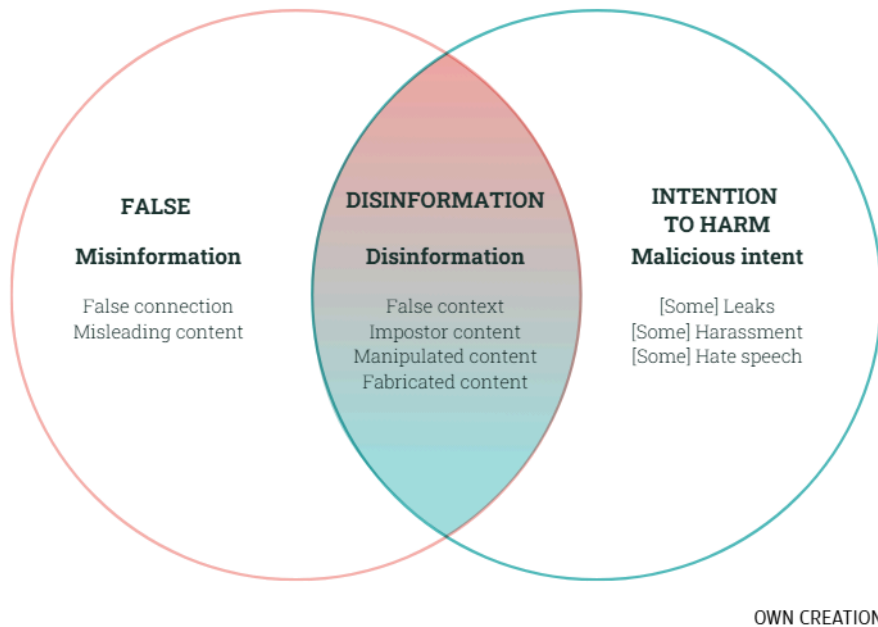
According to UNESCO, information clutter is classified into three types:

1. **Disinformation:** false information intentionally spread by people who know it is false.
2. **Misinformation:** false information disseminated in the belief that it is true.
3. **Malicious information:** information based on true facts, but used to harm others.

The Forum Against Disinformation Campaigns in the field of national security in Spain has analysed these types of manipulated information and highlights the potential effects as threats, as well as its strategic use as a protection tool.

The MsW [misinformation widget] methodology against covert science

One important form of disinformation is **covert science**, where a false scientific approach is used to promote hidden interests. Notable cases of this nature occurred during the COVID-19 pandemic, such as the promotion of conspiracy theories about the origin of the virus and the early use of hydroxychloroquine without sufficient evidence, which affected both public health and trust in science.

Figure 11 Types of disinformation

To combat covert science, the CSIC proposes the **MsW (misinformation widget)** methodology, which focuses on verifying the accuracy of claims and the credibility of sources through a three-step process: assessing the credibility of the source, **verifying** the veracity of the claim, and **estimating** the potential impact of the information.

The MsW methodology and its tools aim to provide a structured framework for assessing and managing misinformation effectively. Currently, and in the framework of the XAI-Disinfodemics and SafeHorizon projects, this methodology is being extended through the Athena and IQi platforms.

Social factors associated with belief in conspiracy theories

A **conspiracy theory** is a belief or explanation that suggests that an important event or situation is the result of a secret conspiracy, usually carried out by a powerful or influential group. Such theories often imply that the group in question is acting in secret to achieve goals that are illegal, immoral or harmful to the rest of society.

Conspiracy theories are of great importance to academic research, given the broad impact they have had on public health, trust in institutions, and political and social discourse. For example, several studies illustrate a positive correlation between acceptance of conspiracy theories and vaccine hesitancy. Consequently, examining the psychological, cultural and political factors that contribute to the formation and spread of conspiracy theories is critical to developing effective strategies to counteract misinformation and promote evidence-based knowledge, rational decision-making and public health. Among other work developed around conspiracy theories, the CSIC Instituto de Políticas y Bienes Públicos [Institute of Politics and Public Goods] (IPP) has explored the social factors associated with belief in conspiracy theories in Spain during the COVID-19 pandemic, helping to understand the social predictors of this phenomenon and generating knowledge that can inform policy, practice and public knowledge.

The main findings of the study show that beliefs in conspiracy theories are characterised by being:

- a. associated with significantly poorer vaccination behaviours, including the children of those who hold these beliefs;
- b. not associated or only weakly associated with standard demographics such as age and sex; the same is true for education and health occupations;
- c. related to instrumental rationality considerations;
- d. weakly related to value rationality indicators, such as ideological and religious affiliations.

Social Listening for infodemic control

The WHO defines an infodemic as “too much information, including false or misleading information, in digital and physical environments during a disease outbreak”. Its management seeks to reduce its impact on public health through evidence-based analysis and approaches, promoting good practice through community listening, improving understanding of risk, building resilience to misinformation and empowering communities. Social listening, which involves analysing conversations on digital platforms to understand public perceptions and emotions, is key to making informed decisions and managing infodemics.

Social listening can be useful in managing infodemics and helps capture community concerns and questions. Traditionally, social media analytics platforms have been designed for use by commercial brands for *marketing* and sales purposes, and have been barely used for a broader understanding of social dynamics in areas such as public health. Therefore, in response to a request from the WHO, IIIA-CSIC technology has served as the basis for the creation of the EARS (Early Artificial Intelligence-Supported Response with Social Listening) platform to address the evolving needs of public health analysts during the COVID-19 pandemic.

Data for the EARS platform were collected daily from online conversations in public sources available in nine languages. Public health and social media experts developed a taxonomy to categorise COVID-19 narratives into five main categories and 41 sub-categories. A semi-supervised machine learning algorithm was implemented to categorise social media posts into various categories and according to various filters. In total, 215 469 045 social posts were collected for processing between December 2020 and February 2022. In order to evaluate the machine learning-based categorisation, it was compared with the previously used method (Boolean search filters). This comparison showed a statistically significant improvement in accuracy. In addition, the constructed platform is scalable, allowing for easy incorporation of new countries and languages.

The application of a public health taxonomy and artificial intelligence technology in an accessible and easy-to-use platform for social listening, which analysts can access directly, is a significant step towards better understanding global narratives. EARS has become a success story that endorses the capabilities of social listening technologies to obtain current information on issues of interest to citizens.

Artificial intelligence and democracy

Artificial intelligence will undoubtedly have a significant impact on our democratic system. This will be negative in some cases (for example, generative intelligence facilitates disinformation) and positive in others. In particular, one of the latter is the improvement of citizen participation through the optimisation of digital democracy tools. Specifically, the IIIA-CSIC has worked on improving participatory democracy through the use of artificial intelligence techniques that incorporate human values in participatory budgeting (PB) as a case study.

Participatory budgeting has two main shortcomings. First, it requires the active participation of citizens, and citizens often do not participate. For example, Paris' participatory budget, considered one of the largest in the world, involved 137,622 participants in 2023, representing approximately 6.5% of the more than 2.1 million citizens. Low participation may not be representative of the population and may lead to biased decisions. Second, participatory budgets have traditionally managed small portions of the government budget. Continuing with the case of Paris, its participatory budget is 100 million euros each year, which represents 1.25% of the more than 8 billion euros of municipal spending. The underfunding of participatory budgets allows governments to claim an interest in participatory democracy, while retaining full control over most budgetary decisions. It could be argued that governments do not allocate more money to participatory budgets in order to avoid the possibility of a small percentage of the population pushing proposals against their interests.

To solve these problems, we propose to adopt a new perspective on participatory budgeting based on finding a consensual budget allocation in order to satisfy both the participants and the government. Our vision is that if we consider both sides when allocating the budget, we accomplish the following key points:

- We avoid low participation biasing the decision and we give a voice to citizens who do not participate (represented by the Government).
- Because governments are assured that low participation will not bias the decision against them, they will be inclined to invest more money in the process.
- If more money is invested in the process, citizens will be more likely to participate, providing their personalised preferences for proposals.

At this point, it is important to note that most participatory budgets tend to use sub-optimal allocation techniques. Therefore, the application of artificial intelligence not only brings the aforementioned benefits, but optimises votes and resources better than commonly used techniques. The approach is based on game theory to consider both the preferences of the participants and those of the government (partially representing non-participants). In order to take into account executive preferences, the proposed approach takes into account the values that the Government seeks to promote and those with which each proposal is associated.

Experiments show promising results that validate this approach. Considering data from 29 recent participatory budgets in different districts of Barcelona and Warsaw (Poland), it was observed that the selection of proposals that incorporate the alignment of values with government preferences reduces participants' support by less than 5% compared to the selection of proposals that only consider participants. This demonstrates that adding new stakeholders to the decision does not necessarily have to be to the detriment of others. Thus, adopting this compromise solution satisfies both participants and the executive bodies and is arguably a better solution for citizens who are not actively involved in the process.

Artificial intelligence for public policy management

Artificial intelligence can contribute to each and every stage of the public policy making and management process. In this respect, a key aspect of this technology is its novel ability to predict. One way to integrate this strength is through the formulation and implementation of more effective preventive policies. Such policies, which are often difficult to target, tend to be ineffective. The integration of artificial intelligence-based forecasting can completely transform the possibility space for such policies. One example is the prediction of negative international crises, such as the outbreak of armed conflicts, coups d'état or repression. Research has shown that such events can be predicted with increasing accuracy. Platforms such as PACE, VIEWS and Conflict Forecast are already making such predictions. The Conflict Forecast system uses six million news articles from around the world to train an algorithm capable of predicting the outbreak and escalation of armed violence before it occurs.

The CSIC's Instituto de Análisis Económico [Institute of Economic Analysis] is researching the integration of artificial intelligence-based predictive algorithms for armed conflict prevention policies. Although these tools, such as DEWAM (Dynamic Early Warning and Action Model), are currently used more for analytical purposes, their application in prevention presents challenges, as it requires integrating quantitative forecasts into policy design. DEWAM combines a forecasting module that assesses conflict risk and intensity and a dynamic decision-making module that simulates policies and quantifies their potential impact. This model classifies conflict risk into 12 stages, from stable peace (stages 1 to 4) to intense conflict (stages 11 and 12), suggesting that the most cost-effective interventions are in the preventive stages (stages 5 to 10), where the risk of escalation is high but the costs are still manageable.

DEWAM represents a first step towards AI-informed policy, with the potential to be extended to other policy domains. However, technical and ethical challenges remain, as machine learning predicts trends, but policy-makers require forecasts conditional on specific policy options, which is complex. This capability demands collaboration between quantitative disciplines such as computer science, economics and political science.

Moreover, DEWAM's relevance goes beyond the prevention of armed conflict, applying also to risk management of hybrid threats, such as manipulation in social networks that trigger conflicts in the physical world. Here, the challenge is to understand the relationship between hate speech on social media, political polarisation and social mobilisation. In line with the European Regulation on Freedom of the Media, disinformation management should be harmonised with freedom of the press, which implies developing advanced monitoring systems, as proposed in the working paper of the Forum against Disinformation Campaigns in the Field of National Security.

2.5.

Artificial intelligence for sustainability

SUSTAINABILITY refers to the ability to meet present needs without compromising the ability to meet the needs of future generations. Thus, the term is intrinsically linked to the use of natural, economic and social resources in a balanced and responsible manner, so that human development is viable in the long term and in harmony with the environment.

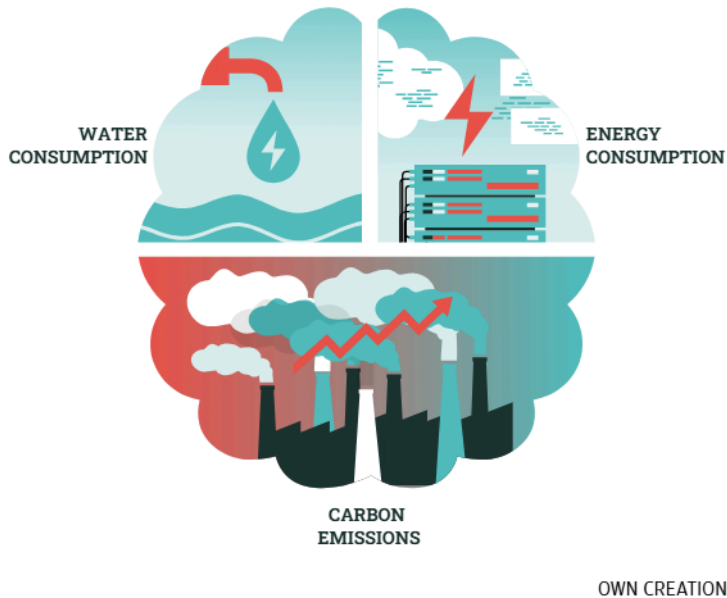
Sustainability, in the economic and social context, seeks models that promote equality, well-being and economic stability without depleting the planet's resources. In the environmental setting, it involves protecting ecosystems, reducing pollution and preserving natural resources.

Some notable examples of sustainability that make the use of artificial intelligence essential can be found in the fields of energy, mobility and industry.

Intelligent energy management

The increase in greenhouse gas emissions resulting from human activity in recent centuries (see **Figure 12**) has led to unprecedented rapid global warming. Limiting the global temperature increase to 1.5°C and achieving the Paris Agreement targets requires a profound systemic transformation of the global energy system over the next 30 years. The success of this energy transition implies a transformation from a sector based on fossil fuels to one based on zero carbon emissions or energy generated from renewable sources. In this regard, the International Renewable Energy Agency's (IRENA) report *Global Energy Transitions Outlook 2023* identifies electrification and efficiency as key drivers of the energy transition, driven by renewables, clean hydrogen and sustainable biomass.

In the Spanish context, Law 7/2021 on Climate Change and Energy Transition, better known as the *Climate Change Law*, establishes minimum national targets for 2030 related to the reduction of greenhouse gas emissions, consumption and generation of energy from renewable sources and improvement of energy efficiency.

Gráfico 12 Industrial activity and its environmental impact

Advances in embedded data capture systems and communications technologies have facilitated the development of new solutions based on artificial intelligence techniques applied to different areas and with a focus on energy sustainability. Some examples of these applications are:

- **Improved forecasting of renewable energy production in generation systems.** In research in which the Instituto de Geociencias [Institute of Geosciences] (IGEO-CSIC-UCM) has participated, it has been shown that artificial intelligence algorithms fed with meteorological data can significantly improve the prediction of energy production in real time, which facilitates the integration of a greater amount of renewable energy into the electricity system.

- **Maintaining grid stability and reliability in transmission and distribution systems.** Artificial intelligence techniques can be used to optimise energy management in decentralised grids and microgrids, both to maintain the balance between energy demand and supply and to guarantee the quality of electrical parameters (voltage level, frequency, etc.). The Instituto de Física Interdisciplinar y Sistemas Complejos [Institute of Interdisciplinary Physics and Complex Systems] (IFISC-CSIC-UIB) has participated in this research.
- **Optimisation of demand forecasting and its efficient management.** Based on energy consumption data, artificial intelligence techniques can be used to predict and generate individualised consumption habits and patterns for each consumer. This information is useful both for predicting demand and improving the energy efficiency of buildings, according to research carried out at the Centro de Automática y Robótica [Centre for Automatics and Robotics] (CAR CSIC-UPM). Using generative artificial intelligence, it is possible to analyse historical data and current conditions to predict energy demand.
- **Improving energy efficiency in industrial processes.** Based on models of energy consumption and resources of machines and industrial processes, solutions can be developed for their monitoring and optimisation (research carried out at the CAR CSIC-UPM), in order to be used in predictive maintenance systems or for the optimisation of energy consumption and materials, among other applications (research carried out at the Instituto de Robótica e Informática Industrial [Institute of Robotics and Industrial Informatics] (IRII-CSIC-UPC)).

Intelligent mobility

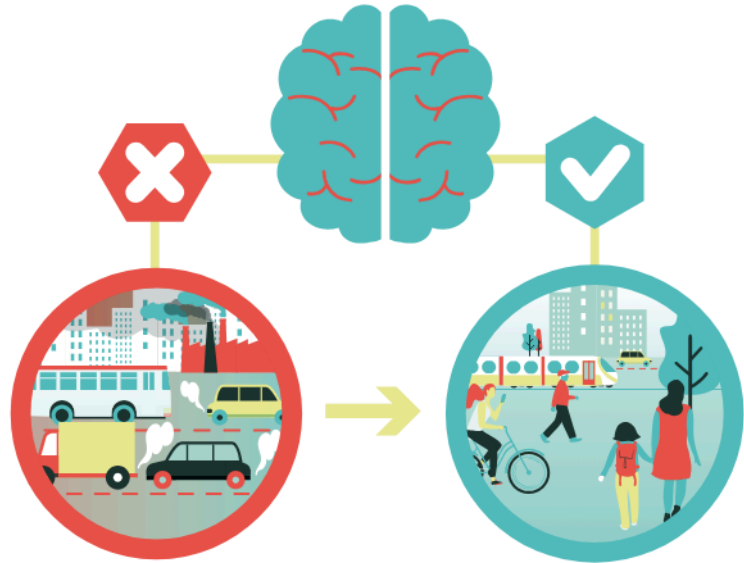
Urban mobility is undergoing a process of profound structural change, with more sustainable means of transport (pedestrians, bicycles, public transport and shared mobility) increasing-

ly replacing the classic mobility model based on private vehicles. Moreover, the widespread use of connected vehicles has the capacity to create a new ecosystem where a large amount of information is achieved in real time to be processed with the help of artificial intelligence systems. In this context, it becomes imperative for mobility managers to rely on systems of this nature, such as the so-called *digital twins*, to be able to interpret possible changes and study the impact of measures implemented in computer simulations before they are implemented in the real world.

Building these models requires a large amount of information on mobility needs or the state of infrastructures and their capacity, as well as more sensitive issues involving people, such as decision-making on which routes to take, which modes of transport to use and how to react to changes, according to research results from IFISC (CSIC-UIB). In short, artificial intelligence-based solutions can quantify the impact of mobility policies and make it possible to define action plans aimed at achieving certain objectives (IIIA-CSIC). For example, some tools based on this technology analyse policies for the adoption of shared mobility (IIIA-CSIC), prioritising environmental aspects or the quality of service for users. New solutions improve the mobility of vehicles and the movement of pedestrians or cyclists within an urban environment, recommending routes that not only minimise distance but also exposure to atmospheric pollutants (IIIA-CSIC). Likewise, with the use of generative artificial intelligence, it is possible to design transport systems that optimise traffic flow, minimising traffic jams and reducing pollutant gas emissions.

Within intelligent mobility, the autonomous vehicle (research carried out at the CAR) deserves special attention because, although it arose with the intention of improving road safety, it is now seen as a key element in transforming urban mobility, making it more sustainable, inclusive and accessible. Artificial intelligence-assisted autonomous navigation has to respond to several challenges that articulate much of the research in this field:

Figure 13 Transition towards more sustainable mobility guided by artificial intelligence.



OWN CREATION

- Ensuring confidence in decision-making by accurately identifying the operational environment (such as traffic, terrain and weather) and adjusting vehicle behaviour accordingly. This requires continuous, long-term learning to respond effectively to a variety of complex circumstances.
- Explainably, transparently and robustly predict the world around them to make better decisions.
- Exchanging on-board sensors and safely updating *software* functions to understand, as accurately and reliably as possible, the vehicle's environment. In this respect, the HAL4SDV project, which brings together the European automotive industry and in which CSIC participates, aims to lay the foundations for an open *software* infrastructure that will enable Europe to remain competitive in this sector.

Intelligent industry

Artificial intelligence plays a crucial role in developing more sustainable industrial activity. With methods belonging to this technology, it is possible to optimise the consumption of water, energy and raw materials by analysing the production flow, generating recommendations in real time aimed at reducing the use of these resources (CAR) or it is possible to reduce CO₂ emissions or other pollutants in industrial plants, identifying critical points and proposing solutions, through generative artificial intelligence, that optimise production patterns.

Artificial intelligence also facilitates the circular economy by identifying opportunities to reduce, reuse or recycle materials in production processes. Generative artificial intelligence enables the design and exploration of more efficient recycling systems. Artificial intelligence also enables predictive maintenance to anticipate equipment failures before they occur (CAR), avoiding costly repairs and reducing energy consumption and waste associated with the manufacture and replacement of parts, while generative intelligence can help in the design of new, more efficient and sustainable products by assisting in carrying out simulations to assess the environmental impact of a product before it is manufactured.

The transport industry can also benefit from artificial intelligence techniques to optimise distribution routes, reducing fuel consumption and thus emissions, or to achieve more sustainable supply chain management, helping to anticipate demand and optimising logistics to reduce the carbon footprint.

Agriculture is another industrial activity where we find sustainability practices (such as precision farming) that promote crop management that considers crop variability, with the aim of reducing inputs and environmental impact. Research carried out at the CAR has shown that this technology can analyse soil, climate and crop data to optimise the use of water and agrochemicals that pollute water and soil.

At the same time, artificial intelligence will also allow early detection of pests through inspection systems that collect information in the field with a set of sensors, which can be on board autonomous robots both aerial and terrestrial (CAR research) or be part of a fixed network installed in the field. Generative intelligence facilitates the analysis of climate, soil and water data to suggest crops and sowing patterns that maximise productivity and minimise the use of resources.

Finally, artificial intelligence enables inclusive work environments in which operators and intelligent machines (*cobots*) successfully collaborate, integrating human intelligence and technology in a balanced way, oriented towards well-being, personalisation and positive social impact. An example of these *cobots* can be found in manual fruit picking, where the robot coordinates with the operator to perform tasks that require strength, such as transporting heavy boxes, while the operator performs tasks that require dexterity, such as picking bunches of grapes for wine with denomination of origin (research by CAR).

Figure 14 Industry 5.0. Robots do not replace people, but complement their skills



OWN CREATION

three



Conclusions and Recommendations

SCIENCE  POLICY | 61

|||||

In the societal domain, various scientific and practitioner fora have highlighted the importance of introducing artificial intelligence to address the huge challenge posed by an ageing population, the shortage of care staff and the increasing burden this places on systems (health and social services) and families. There is also a broad consensus on the wide potential of this technology to support digital inclusion and empower communities, their workers and end-users across many of the sectors that make up the welfare state.

|||||

The influence of artificial intelligence tools on labour markets, employment policies and working conditions is profound and multifaceted. While this technology presents significant challenges, such as labour displacement and the need for continuous adaptation, it also offers opportunities to increase productivity and create new employment landscapes in new fields (such as data governance or interpretation and decision-making based on artificial intelligence analysis). It also has the potential to impact the development and management of monetary allowances paid by the state or a public administration, such as universal basic income, significantly improving their effective implementation.

|||||

In turn, artificial intelligence is a cornerstone in the fight against disinformation, in the design and optimisation of social policies and in improving citizen participation. In this respect, the CSIC is at the forefront with initiatives such as the MsW methodology and the EARS platform, which have proven to be effective in combating disinformation. In addition, it has contributed to facilitating democratic processes through participatory budgets optimised using this technology and allows the design of preventive policies through predictive algorithms such as DEWAM.

|||||

Artificial intelligence is also playing a crucial role in the promotion of sustainability and the fight against climate change thanks to its ability to generate techniques that optimise processes, reduce the use of resources and improve decision-making. Several research groups at the CSIC are working on the development of solutions based on this technology for complex problems in the field of sustainability, such as those related to energy, mobility and industry. This work shows that it is critical to balance the use of artificial intelligence with strategies that minimise environmental impact, ensuring that its implementation is aligned with sustainable development goals. In short, society can harness it as an ally in sustainability and climate change, if public authorities and companies promote it responsibly and ethically in key sectors such as energy, mobility and industry. Through informed choices, supported by responsible policies and the implementation of sustainable habits, we can all contribute to this technology driving a more balanced and sustainable future for people and planet.

Recommendations

Recommendations for public decision-makers

ETHICS

|||||

From an ethical point of view, the necessity and appropriateness of developing an artificial intelligence application must be assessed in advance. The real need and relevance of using such a system for the proposed objective must be convincingly justified. To this end, a cost-benefit analysis must be carried out, including the possibility of alternatives through traditional means and the convenience or otherwise of delegating a certain process or activity to a system of this nature, depending on its special relationship with fundamental rights of the individual, clearly defining and documenting the objectives and scope of the system or application (what for), as well as who has decided on its development (legitimacy).

|||||

It is also necessary to develop controlled test spaces (*sandboxes*) for artificial intelligence that facilitate the development and testing of innovative systems under strict regulatory oversight before they are put into service. In this regard, the formulation of a plan for risk mitigation and damage compensation measures is also crucial. Individuals or groups harmed as a result of system malfunctioning should receive treatment and assistance as well as compensation.

EDUCATION

|||||

In order to address the integration of artificial intelligence in the education system, it is recommended that part of the teaching staff in each educational centre should carry out research tasks and be responsible for collecting evidence on the impact of the use of this technology in the classroom.

|||||

This evidence-gathering should be based on experiments with students, which is why it is essential that some members of the teaching staff have the appropriate training to carry out these experiments correctly. This could be done by reducing the number of teaching staff or by hiring specialised staff.

|||||

A key aspect of validating the use of artificial intelligence in the classroom is data collection. Indeed, this should be a priority for the education system in general. Teachers, school management and the administration must ensure that the set of data collected is as rich as possible and of the necessary quality. In turn, this data should be made freely, openly and anonymously available.

WELFARE

|||||

The integration of artificial intelligence into welfare states brings with it a number of challenges in several critical areas. The latest findings and evidence underline the need to implement robust regulatory and control measures and effective policy options to minimise certain risks linked to increased social inequalities, lack of transparency and trust, and vulnerability to hacking. This would maximise the benefits that this technology can bring to the social policy mix.

|||||

At the same time, to address these challenges, a joint design and development approach is recommended, involving multiple stakeholders throughout the AI production lifecycle (e.g., in the case of social services, research groups, home care or technology companies, health and social care institutions, administrations and patients' associations) and increasing transparency and training and education in this technology for both professionals in the welfare state sectors and citizens themselves.

|||||

As the conflict between artificial intelligence decision-making and the preservation of human freedom and dignity emerges as a central issue, a cautious and well-regulated approach that prioritises the safety of users of welfare state services in all sectors, as well as equity of access to them, is recommended. In short, it is crucial to foster a continuous dialogue between AI technology developers, practitioners, policy-makers and society at large to ensure that its development and implementation is carried out in an ethical and responsible manner.

SOCIETY

|||||

At this level, artificial intelligence can be useful for implementing information verification systems in public administration. In this regard, to combat information clutter and protect citizens against misinformation, governments should adopt verification methodologies such as **MsW (misinformation widget)**. This system helps to assess the veracity of information circulating in the public sphere and reduce the impact of conspiracy theories or malicious information in key sectors such as health and security.

|||||

Artificial intelligence can also promote transparency and participation in democratic processes. Encouraging its inclusion in citizen participation processes, such as participatory budgeting, to ensure that decisions reflect both the will of citizens and the interests of institutions, would help to improve the representativeness and effectiveness of digital democracy by creating resource allocation mechanisms that consider the preferences of both parties and reduce the biases of low participation.

|||||

It is also possible to use artificial intelligence for preventive public health and safety policies and to integrate predictive artificial intelligence tools, such as **DEWAM**, in the formulation of public policies to anticipate crises, such as health or social conflict. These tools can help governments implement early interventions and reduce the costs associated with crisis management, thereby promoting a proactive approach to public safety and well-being.

SUSTAINABILITY

|||||

In relation to the implications of artificial intelligence for sustainability, the creation of big data spaces needs to be fostered to facilitate the development of this technology in this area. Policies and projects should be put in place to promote the use of open data for the development of artificial intelligence solutions to address critical problems such as resource management or climate change. This is directly related to the need to facilitate access to reliable information on infrastructures, capacities, frequencies, location of points of interest and attraction, pedestrian flows, measures to adequately deal with special events, major disasters, etc.

|||||

Public authorities have to define a regulation on data privacy for consumers and approve incentives to participate in pilot projects as data providers. They should also lead and promote the creation of sustainable or net-zero energy data centres and artificial intelligence services, with the aim of minimising their environmental impact, for example, by ensuring that they consume certified renewable energy or that they integrate systems for the use of waste heat generated and the recirculation of water used in cooling systems.

|||||

The different levels of government, through coordination and cooperation and integrated management of capacities and resources, must pursue the development of a robust 5G communications network with full coverage for the entire territory, promoting full digitisation to jointly design artificial intelligence solutions to complex intergovernmental and intersectoral problems, such as mass migrations, those arising from climate change, poverty or the ageing of the population.

|||||

The creation of controlled and safe environments for experimentation (*sandboxes*), with greater scope and complexity, needs to be enabled and implemented, i.e., not only for testing new vehicles and new mobility strategies, but also for testing new sensor systems and validating AI-based modelling predictions.

|||||

Finally, the creation and development of a green entrepreneurship incentive system is also required, supporting entrepreneurs and *start-ups* that develop artificial intelligence solutions for sustainability in the energy, mobility and industry sectors, thereby fostering a greener economy.

Recommendations for companies

|||||

From an ethical perspective, the business sector should establish robust cybersecurity mechanisms to protect the integrity of databases and prevent data leaks or *hacking*. At the same time, it should attempt to ensure that a representation of stakeholders and users is involved in the different phases of the project lifecycle (from design to implementation) through specific methodologies: focus groups, co-creation, DELPHI methodology, etc. The private sector should document all design and evaluation steps to facilitate the auditing of systems, so that any potential biases can be more easily detected and accountability facilitated.

|||||

In the field of education, the industrial sector needs to build low-cost, high-quality adaptive and collaborative learning tools. The economic return on these tools could come from their use in business environments where continuous training is required for their workers or in the field of lifelong learning. The use of these tools in the compulsory education sector should benefit from a reduced subscription price, or be free of charge, in order to make their use affordable for public administrations.

|||||

Given the influence of artificial intelligence in the creation and spread of misinformation, companies should implement information verification and social listening tools, such as the MsW (*misinformation widget*) system, to protect their reputation. These tools help to identify and manage potential misinformation risks that may impact on the perception of goods and services. Companies should prioritise transparency in the collection and use of personal data, ensuring that consumers understand how artificial intelligence is used in their products. Predictive artificial intelligence tools can also be used to anticipate market changes or manage risks, thanks to the functionality of these models to help companies make strategic decisions based on future trends, thereby optimising resources and improving resilience.

|||||

In relation to sustainability, the business sector needs to set targets aligned with the Sustainable Development Goals and assess the impact of integrating artificial intelligence solutions on aspects such as energy efficiency, waste reduction, responsible resource management, etc. It should also consider the impact of the integration of this technology on the automation of processes and the impact on employment, exploring ways of training and relocation of employees. The aim has to be to integrate the principles of Industry 5.0, which focus on close, daily and efficient collaboration between people and artificial intelligence, including intelligent machines/robots. Finally, companies should pursue the creation of monitoring services that encourage responsible consumption. To this end, specific metrics should be implemented to assess the environmental and social impact of artificial intelligence solutions and publish sustainability reports showing the company's commitment, thereby fostering trust and support from consumers and business partners. Finally, product labelling should include information on the sustainability strategies employed, which would build a responsible brand image.

Recommendations for society

|||||

Ethics is a central element in the development of artificial intelligence tools. It is necessary to promote education and raise awareness in society about the potential of data and this technology to generate economic and social value, highlighting examples of success in this regard and communicating internally and externally the principles and good practices described here. The use of sustainable and energy-efficient *hardware* and *software* (green algorithms), as well as the reuse of materials and electronic devices, a culture of repair and the prioritisation of renewable energies in digitisation processes (circular digital economy) should also be encouraged.

|||||

With regard to education and the education system, information materials on the use of artificial intelligence should be prepared and made available to pupils' parents. These materials should give a balanced view of the potential benefits and risks of using this technology in the classroom and at home.

|||||

For society as a whole, digital literacy skills need to be developed. It is essential that citizens learn to identify false or manipulated information, understanding the risks of information clutter. This involves training in the use of basic fact-checking tools and source evaluation techniques, especially on critical issues such as public health and safety; education in these skills will help to reduce the impact of misinformation in society. There is also a need to actively engage in digital democracy processes. Society should take advantage of citizen participation platforms enabled by artificial intelligence, such as participatory budgeting, to express their opinions and contribute to decision-making on issues of public interest. Greater participation will strengthen representativeness in government decisions and encourage the adoption of technologies that respect community values and preferences.

|||||

Finally, society must also play an active and mobilising role, speaking up and demanding from all actors the integration of artificial intelligence solutions that manage smart grids to make better use of renewable energy, public services or mobility. At the household and community level, smart self-consumption must be promoted through efficient management with this type of solutions. In the field of mobility, it is necessary to commit to smart public transport, as well as to the use of electric and shared vehicles or sustainable multimodal transport, supporting policies and services that encourage and facilitate the use of these forms of mobility.

four



List of Centres

|||||

CENTRE	WEB PAGE	E-MAIL
Centro de Automática y Robótica [CAR, UPM-CSIC]	www.car.upm-csic.es/	direccion.car@csic.es
Centro de Ciencias Humanas y Sociales [CCHS, CSIC]	www.cchs.csic.es/	direccion.cchs@csic.es
Instituto de Análisis Económico [IAE, CSIC]	www.iae.csic.es/	direccion.iae@csic.es
Instituto de Filosofía [IFS, CSIC]	www.ifs.csic.es/	direccion.ifs@csic.es
Instituto de Física Interdisciplinar y Sistemas Complejos [IFISC, UIB-CSIC]	ifisc.uib-csic.es/	direccion.ifisc@csic.es
Instituto de Geociencias [IGEO, UCM-CSIC]	www.igeo.ucm-csic.es/	direccion.igeo@csic.es
Instituto de Investigación en Inteligencia Artificial [IIIA, CSIC]	www.iiia.csic.es/	direccion.iiia@csic.es
Instituto de Políticas y Bienes Públicos [IPP, CSIC]	ipp.csic.es/	direccion.ipp@csic.es
Instituto de Robótica e Informática Industrial [IRII, CSIC-UPC]	www.iri.csic.es/	direccion.iri@csic.es
Instituto de Tecnologías Físicas y de la Información Leonardo Torres Quevedo [ITEFI, CSIC]	www.itefi.csic.es/es	direccion.itefi@csic.es

five



Further Information



European Commission: 2023 *Strategic Foresight Report. The Communication Builds on the Joint Research Centre's Science for Policy Report: Towards a fair and sustainable Europe 2050: Social and economic choices in sustainability transitions.*



United Nations (UN) (2019). *A/74/493: Digital Welfare States and Human Rights - Report of the Special Rapporteur on Extreme Poverty and Human Rights.* Available at: <https://www.ohchr.org/en/documents/thematic-reports/a74493-digital-welfare-states-and-human-rights-report-special-rapporteur>.



Ministry of the Presidency, Relations with the Courts and Democratic Memory (November, 2023). Forum against Disinformation Campaigns in the Field of National Security. <https://www.dsn.gob.es/es/file/9233/download?token=9VtNf50a>.



UNESCO (2019). *Beijing Consensus on Artificial Intelligence and Education* <https://unesdoc.unesco.org/ark:/48223/pf0000368303>.



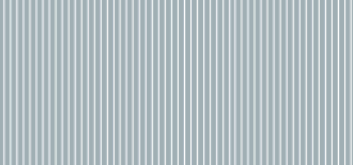
Jorge Villagr  (2019, 5 March). "Towards an intuitive decision system in automated vehicles" *Revista Carreteras*. <https://www.revistacarreteras.com/2019/03/05/hacia-un-sistema-de-decision-intuitivo-en-los-vehiculos-automatizados-jorge-villagra/>.



 ngela Ribeiro (2023, 30 October). *Sustainable robotics and artificial intelligence* (video). CAR. CSIC-UPM. CSIC Comunicaci n. <https://www.youtube.com/watch?v=p0tGxOUESqQ>.



Jos  J. Ramasco, Aleix Bassolas, Mattia Mazzoli and Pere Colet. "How do you gravitate to work?". *The Conversation*. <https://theconversation.com/como-gravita-usted-al-trabajo-123427>.



Science for Public Policy



Knowledge Transfer
Report



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

EDITORIAL
CSIC

SCIENCE  POLICY