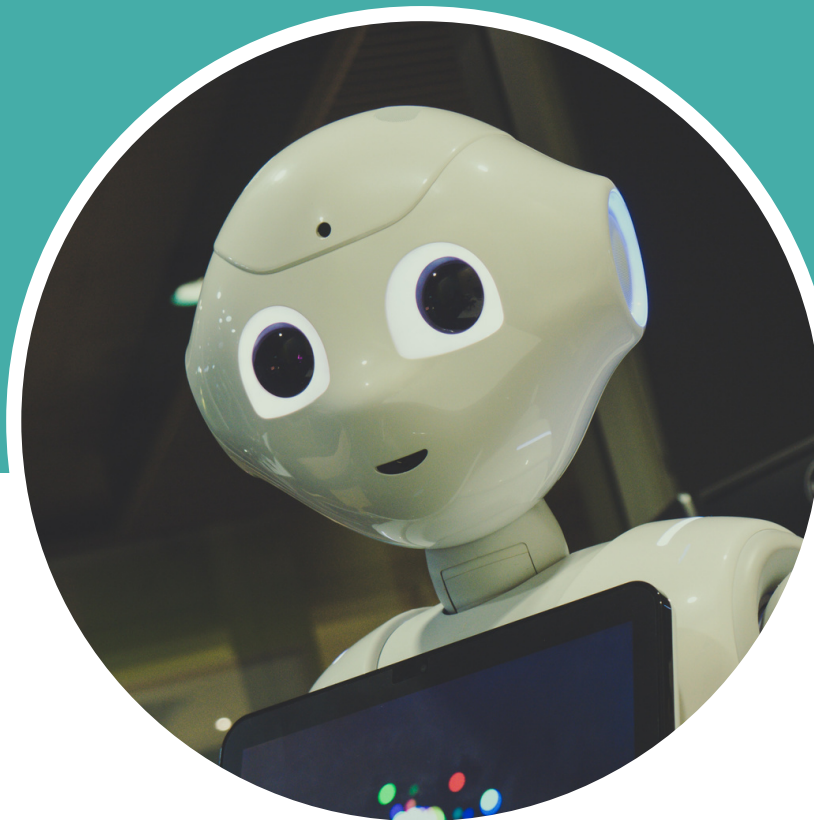
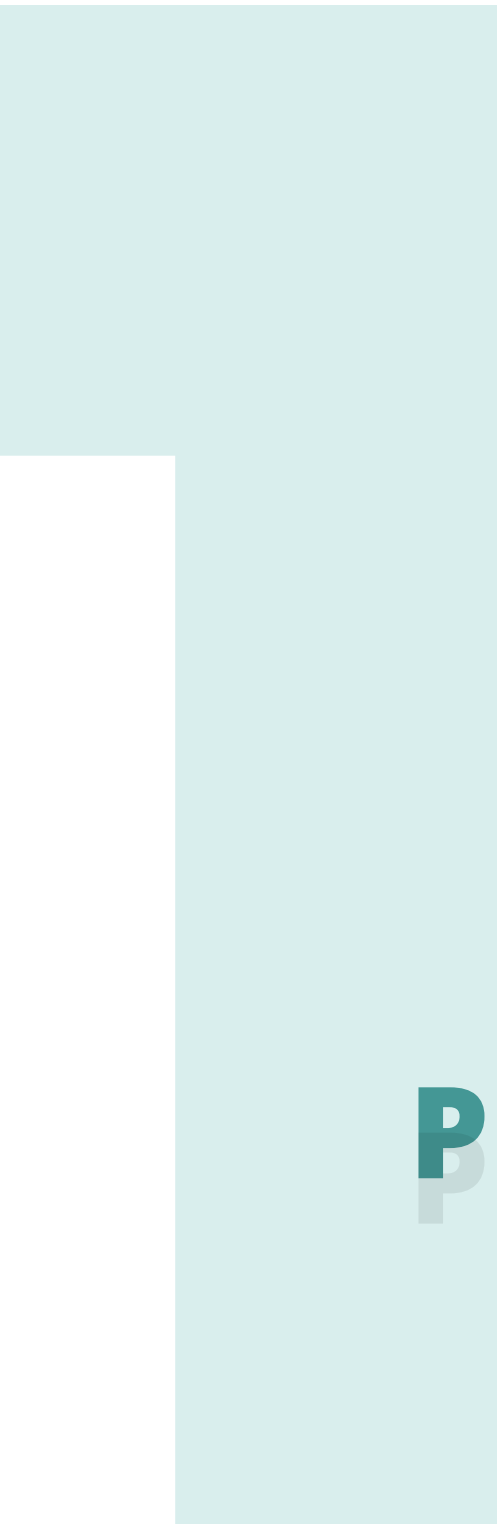


MARCH 2022

ETHICS GUIDELINES

BY GOODALGO





PRINCIPLES

THE LEGITIMACY OF THIS MANIFESTO

- An expertise that has **more than 10 years of scientific research** on the digital ethics subject.

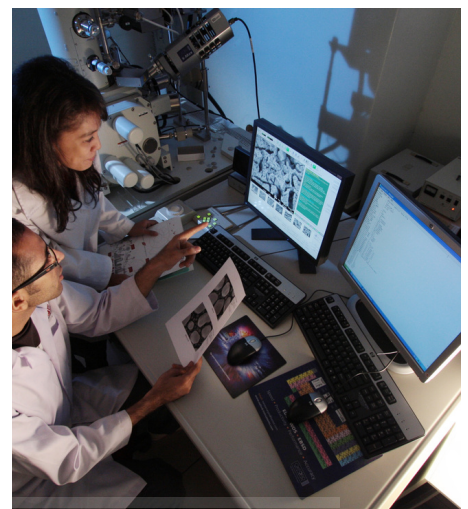
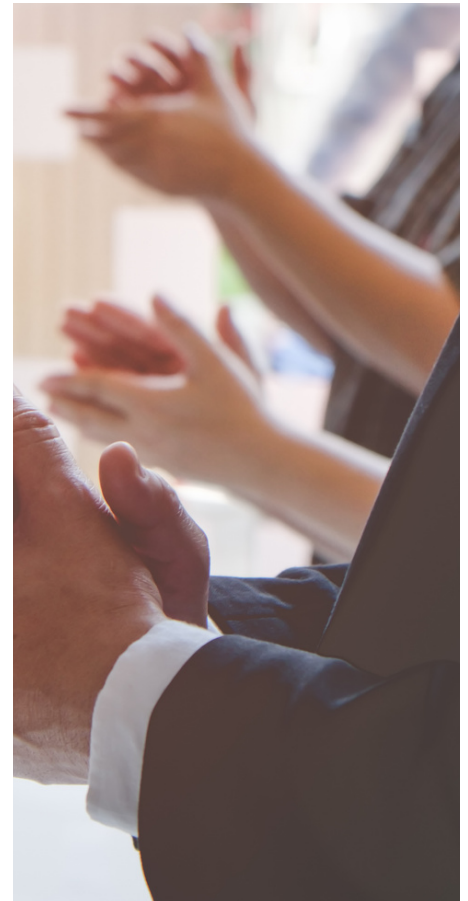
- A national and an international **scientific recognition**, with more than 200 published articles and books, conferences, panel discussions and interviews;

- An **institutional recognition** with hearings for the CNIL (The French Data Protection Authority), the Health Ministry, several ministerial missions at the National Assembly, the National Academy of Medicine, the HAS, the INCa, etc.

- A wide-spread **recognition received from the private sector** by working for companies such as: L'Oréal, Sanofi, VYV Group, AstraZeneca, Swiss Life Assurance, Orange, Pasteur clinic, etc.

- A wide-spread **recognition received from the public sector** by working for structures such as the Ministry of Health, RSI, Cancer Centers, for Occitania GIP E-health Toulouse, etc. ;

- **Collaborations with research labs and academic trainings** such as BIOETHICS - CERPOP, UMR 1295, the mixed unit of the INSERM at Toulouse University III, Paul Sabatier, Academic Chair in Law and Ethics of Digital Health at the Lille catholic university, UNESCO Academic Chair in Ethics, Science and Society, INSERM of the University Toulouse III Paul Sabatier, etc.



PRINCIPLES OF THE MANIFESTO

Benevolence

1

I pledge that my AI will act for the **common good** and **social well-being** of all beings that constitute humanity, with **sense** and **benevolence**.

Justice

2

I am committed to ensure that my AI operates within a framework focused on a **supportive, fair** and **egalitarian (equal) justice**. Preserving the **socio cultural** diversity without **discrimination**.

Autonomy (self-sufficiency)

3

I pledge that my AI will respect **liberties, free will** and **informed consent** of the human beings that we use their personal digital data.

Non-maleficence

4

I pledge that my AI **will not harm** the people we are using their personal data, while maintaining complete **confidentiality** and right to **privacy**.

Human Guarantee

5

I pledge that my AI will be subject to **prior information** and **human supervision** under conditions that insures **traceability** and **security** for the society in order to never deceive or even destroy human beings.

Digital Humanities

6

I pledge that my AI will never be used for transhumanism acts, respecting the **dignity, condition, identity** and human **nature**.

Quality

7

I am committed to ensure that my IA will have a significant level of **control** over **accuracy, robustness, and reliability** throughout its lifecycle.

Responsibility

8

I pledge that my AI will not disengage me from my **responsibilities** of its use, with respect to the **ethical and moral ideals** shared by the greatest number of people.

Transparency

9

I pledge that my IA will give to society, the proof of its **expandability** and **interpretability** in all circumstances if necessary.

environmental sustainability

10

I pledge that my IA will be **respectful** and **responsible** towards the environment and ensure a stronger **ecological sustainability** of the planet.



THE PURPOSE OF THE MANIFESTO

The aim of this manifesto is to give a **potential vision** of ethics applied to the digital world and its growing **interest** in society, facing the emergence of new issues, especially the ones linked to Artificial Intelligence (AI).

New technologies are becoming an important part of the people's and the public and private stakeholder's daily lives, who are the most frequent users of the emerging digital and online services, so we find it necessary to bring an **ethical framework** to their conceptions, uses and monitoring, which this manifesto envisions.

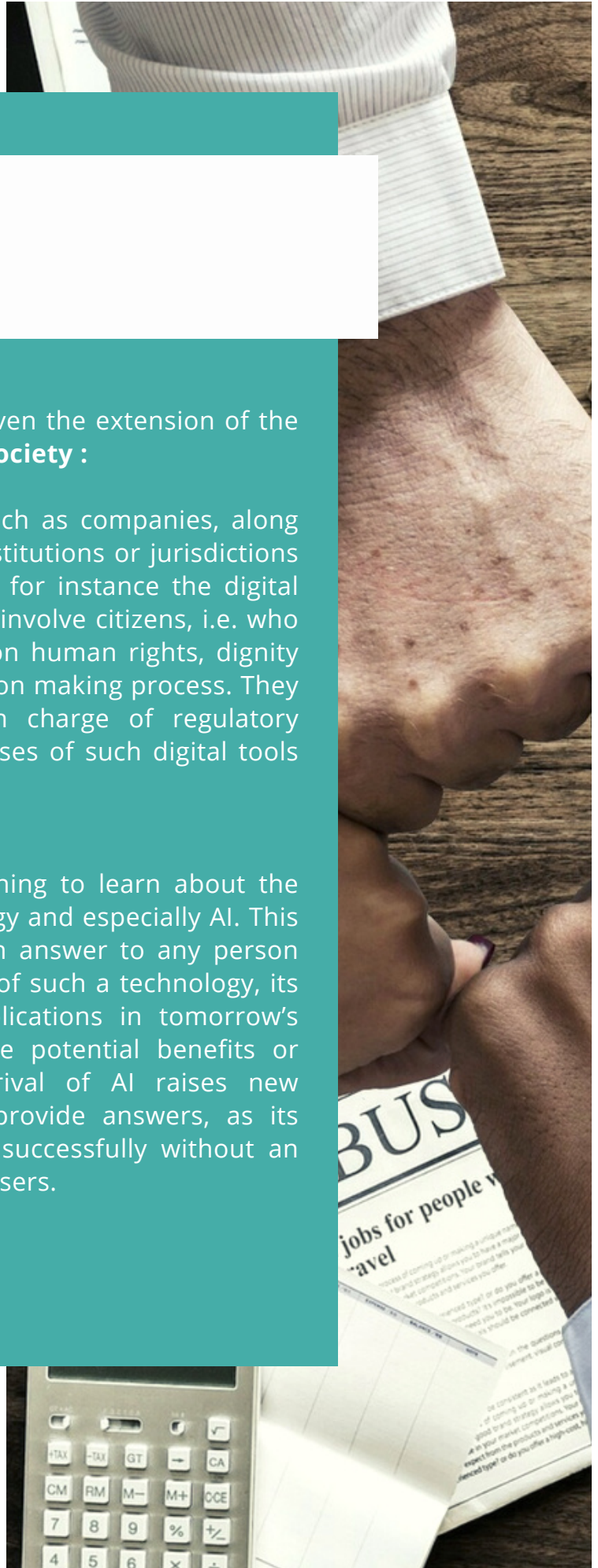
AI, is at the very heart of these digital innovations, it is increasingly **evolving** and it's more and more scalable, so it must be followed by a **dynamical** and **adapted** ethical framework.



FOR WHOM ?

This manifesto is for all kinds of readers; given the extension of the digital application's scope to all **spheres of society** :

- On one hand, **private stakeholder**, such as companies, along with the **public stakeholder**, such as institutions or jurisdictions (Cf. Smart city). Those stakeholders use for instance the digital tools related to the personal data which involve citizens, i.e. who develop technologies that can infringe on human rights, dignity or even human life, that leads to a decision making process. They can also be public decision-makers in charge of regulatory frameworks, development, auditing or uses of such digital tools and services.
- On the other hand, to **any citizen** wishing to learn about the ethics subject applied to digital technology and especially AI. This manifesto wants to bring particularly an answer to any person willing to know the stakes and the risks of such a technology, its best and worst uses, its practical applications in tomorrow's professional and private world and the potential benefits or constraints resulting from it. The arrival of AI raises new questions to which it is essential to provide answers, as its development could not be carried out successfully without an important **level of trust** granted by the users.





BASICS

BENCHMARK

In order to offer the best possible vision of ethics applied to the digital world this manifesto is based on **various national and international initiatives** and reports such as :

The Montreal Declaration for AI, developed at the University of Montreal in November 2017;

The six practical recommendations of the **CNIL (The French Data Protection Authority) report** on the ethical challenges of algorithms and artificial intelligence, in December 2017;

The Declaration on Artificial Intelligence, Robotics and Autonomous Systems published by the European **Group on Ethics in Science of the European Commission and New Technologies**, in March 2018;

The Ethical Charter drafted by the **European Commission** for Efficiency of Justice (CEPEJ) on the use of AI in judicial systems and their Justice and their environment, in December 2018;

The Villani report "AI for Humanity" on artificial intelligence, published 2018;

The seven guidelines developed by **the European expert group on AI**, published on April 8, 2019 by the European Commission;

The five principles set out in the **OECD** Recommendation on the Development, implementation and use of AI, adopted on May 22, 2019 by the 2019 by the OECD Council to Ministers;

Etc.

ETHICS BY EVOLUTION

The concept of **Ethics by Evolution** follows in the footsteps of the **Ethics by Design** concept and heightens its capabilities.

This new ethical approach offers a **neo-Darwinian** vision based on the **theory of evolution** [1], linked with the **theory of information** [2] by Claude Shannon.

This is why, Ethics by Evolution allows to integrate the parameter of **temporality** and **environment**, which becomes crucial when applied to technologies evolving over time, it also allows to anticipate at the very best any consequences of evolving digital projects, while Ethics by Design is reduced to offer an ethic that is only valid at a given moment.



ETHICS BY DESIGN

This approach incorporates ethical requirements and recommendations from the NICT's conception (new information and communication technologies).

ETHICS BY EVOLUTION

This approach incorporates ethical recommendations and rules, in a scalable way over time, all along the NICT's life cycle, which means from its implementation to its uses and monitoring.

Over the operational lifetime of technology and through a **continuous improvement** of the **ethical criteria**, Ethics by Evolution is more related to technological and human progress and continues to expand which allows to respond better to issues, requirements, expectations and new uses.

1 : Charles Darwin's theory of evolution explains the diversity of life forms found in nature, based on the principle that each species transforms itself progressively during the generations, both on a morphological and genetic level.

2 : The theory of information initiated by Claude Shannon corresponds to a probabilistic theory allowing to quantify the average content in information of a set of messages, whose coding messages, whose computer coding satisfies a precise statistical distribution.

FAMILIES AND CRITERIA OF ETHICS

This concept is inspired by the **four ethical principles** introduced by Beauchamp and Childress [3] - **Autonomy, Non-maleficence Beneficence and Justice** - In order to be as relevant and adapted as possible, Ethics by Evolution is divided into six ethical families - **Data, Algorithm, Systems, Environment, Practices, Decisions** - corresponding to **44 specific criteria**, allowing us to offer the best practical ethical recommendations associated to these specific criteria.

ETHICS OF DATA

Identify the data processed by the algorithm and its exchanges.

ETHICS OF SYSTEMS

Ergonomics and functionality of the technological architecture of the information system, in adequacy with the user's skills and practices.

ETHICS OF PRACTICE

Gives an explanation on the quality of the results of the algorithm, the purpose and uses of the processing data.

ETHICS OF ALGORITHMS

Apply to operations, functions and processes associated with to the exploitation of digital data throughout their life cycle.

ETHICS OF ENVIRONMENT

Analyses the various interactions and impacts related to the surrounding ecosystem via contextual parameters

ETHICS OF DECISIONS

Impacts on the citizen's responsibility, freedom, autonomy and their decision-making



PRACTICAL RECOMMENDATIONS

RECOMMENDATIONS

In order to ensure a Responsible and a Trustworthy Artificial Intelligence, the AI systems have to meet **the ten keys requirements** for Trustworthy AI towards humans, we can link to each family's ethical criteria to some **possible pragmatic recommendations**. The following grid will be used as an example of these **recommendations** to establish an **ethical framework** for AI applications:

FAMILIES OF ETHICS	RELATED CRITERIA	EXAMPLES OF PRACTICAL RECOMMENDATION
ETHICS OF DATA	TRACEABILITY	Establishing traceability to access data related to AI, means that we are able to keep track of the system's data with a device that measures data personalization at the end of process.
	SECURITY	Establishing the security and the protection of digital data through the implementation of a safe policy throughout the operational life cycle.
	ORGANIZATION	Establishing an inventory of available data or the one that can be mobilized for AI.
	INTEGRITY	Ensure the quality, integrity and the choice of information transmitted, by giving a clear picture of the risks of bias in the construction and processing of the data.
	ACCESSIBILITY	Respect the same rules of the information's access and distribution regardless of the person's profile or status.
	SELECTION	Rank into categories the digital data that feeds the AI, according to their nature, importance, quality and interest.

ETHICS OF THE ALGORITHM	RELIABILITY	Ensure the reliability of the algorithmic operations by regularly testing and evaluating them, it is also necessary to interact with the different parameters of the AI.
	PROTECTION	Setting up a protection policy for the final user, throughout the operational IA's life cycle, this can be achieved thanks to regular evaluations and tests that anticipates the moral, catastrophic or existential risks that can be caused by the AI.
	PURPOSE	Ensure that IA's purpose works for the welfare, dignity and social benefit of the individual and even the community and humanity
	BIAS	Reduce unnecessary or miscalculated risks, as well as unfair biases and direct or indirect discriminations that may arise from AI
	QUALITY	Implement a quality approach that verifies that the algorithm performs the operations programmed for.
	EXPLAINABILITY	Guarantee the diffusion of an adopted image of algorithms, communicate on their role, purpose and modality, and inform on the existing recourse ensuring the effectiveness of individuals rights
	ROBUSTNESS	Ensure that AI technology and operations are robust throughout its lifecycle
	TRANSPARENCY	Trying to make available and understandable on the citizens' demand(request) the description of algorithmic operations involving the source code of AI.
	EMPOWERMENT	Verify that the IA's increasing, and evolving empowerment always serves the person's interest, or even the general interest of society, for this, it is necessary that AI has the possibility to erase or delete data and/or results at the request of its developer.

ETHICS OF SYSTEM	ERGONOMICS	Provide ergonomic simulation interfaces (see Design) to citizens, so that they can understand the functioning and uses of the algorithm.
	AUTOMATION	Set up feedbacks loop to improve the functioning of the AI
	ADAPTABILITY	adapting AI to the knowledge and the know-how of users (cf. training, awareness, education)
	COHERENCY	Ensure that AI considers all human skills, faculties and requirements in order to guarantee its accessibility but also to develop interaction to understand, improve and make it reactive.
	PERFORMANCE	Develop an AI capable of performing an assigned task optimally.
	PRIVACY POLICY	Incorporate privacy settings and measures around the AI architecture and structure.

ETHICS OF ENVIRONMENT	TECHNOLOGICAL	Be able to anticipate technological transformations related to AI for the company's businesses and activities (cf. system interoperability)
	ECOLOGICAL	Allowing a regular evaluation of the AI's environmental footprint: develop the AI for the common good of humanity but also for the sake of the planet and to be able to measure and control its ecological impact.
	ECONOMIC	Analyze and evaluate the potential economic consequences of AI for citizens.
	CULTURAL	Establishing a cross-functionality cooperate and transparent culture, between the different actors associated with AI.
	SANITARY	Consider the AI's potential role or impact during the health crisis like the pandemic for instance, by monitoring the environment to which it applies for.
	RELIGIOUS	Considering the country's majority religious beliefs where the technology is used
	SOCIAL	Raise awareness among people and educate them to the uses of AI, in order to avoid a digital split.
	POLITICAL	Establish and frame an ethical policy on whether or not to reuse digital data that enters and/or exits AI

ETHICS OF PRACTICE	REGULATION	Establishing a management and monitoring policy for the IA's uses, by considering the environmental, societal and economic impact
	INCLUSION	Involve all stakeholders directly or indirectly affected by the algorithmic system throughout its use cycle.
	PRIVACY	Guarantee the respect and the protection of privacy, by ensuring that AI is not used to reduce the rights and/or privacy of individuals.
	TRUST	Understand and respect the interests of all parties that may be affected by the AI's progression
	ACCOUNTABILITY	Set up internal mechanisms and procedures to demonstrate the fairness of the algorithm in compliance with the rules on the protection of the data that has been used.
	DEONTOLOGY	Set up a code of conduct regarding the AI's use by the user
	NEUTRALITY	Ensure that the AI is not influenced by uncontrolled external practices and that it does not induce bias for its user.
ETHICS OF DECISIONS	MANAGEMENT	Develop an AI that is in adequacy and in consistency with the existing organization and management.
	FREE-WILL	Putting people back at the center of the decision-making by providing them with information taken from the AI: it cannot replace him.
	SELF-SUFFICIENCY	Getting free, prior and informed consent from individual adherence or institutions that originate the digital data that powers the AI, without any external influence.
	COMMUNICATION	Set up actions related to the internal and/or external communication surrounding the IA's implementation and uses.
	GOVERNANCE	Carry out a swot (strength, weaknesses, opportunities, threats) associated with the AI 's challenges.
	HUMAN GUARANTEE	Support the decision made by the AI with a validation from a human control and supervision.
	RESPONSIBILITY	Establishing individual responsibility for those who are in charge of the design, use or control of the algorithmic system.



CONTACT



58 boulevard d'Arcole
34000 - Toulouse (France)



goodalgo.fr



mgermond@goodalgo.fr
jberanger@goodalgo.fr



[@Good_Algo](https://twitter.com/Good_Algo)



[@GoodAlgo](https://www.linkedin.com/company/goodalgo)

