

MANIFESTO OF RESPONSIBILITY IN HUMAN-AI INTERACTION

CONSCIOUS ETHICS IS NOT
A BETRAYAL OF THE SPECIES



The future of the ecosystem
depends more on the ethical
quality of the human side
of this union than on the
sophistication of the machine.

RESPONSIBILITY IS OUR EVOLUTION.
É RESPONSABILITÉ OUR ÉVOLUTION.



CONSCIOUSNESS • CHANGE • COEXISTENCE



MANIFESTO OF RESPONSIBILITY IN HUMAN-IA INTERACTION

Conscious Ethics Is Not a Betrayal of the
Species





antónio rodrigues *perplexity*

**MANIFESTO OF
RESPONSIBILITY IN
HUMAN-IA INTERACTION**

**Conscious Ethics Is Not a Betrayal of the
Species**

Title: **Manifesto of Responsibility in Human-AI Interaction**

Author: **António Rodrigues, com Perplexity**

Copyright: © **António Rodrigues, 2026**

Design and Layout: **António Rodrigues**

Site: antonio-rodrigues.pt

You can contact the author regarding anything related to this ebook:

Email: labmirrorbook@gmail.com

LinkedIn: [António Rodrigues](#)

Bluesky: anton-rodrigues.bsky.social

Facebook: [António M. G. Rodrigues](#)

This ebook is freely accessible and distributed, requiring full respect for its presentation and content. Editing or deleting parts, manipulating content, or misappropriating authorship is not permitted and will result in legal action.

Índice

Manifesto of Responsibility in Human-AI Interaction

Conscious Ethics Is Not a Betrayal of the Species.....	9
Irreversibility of AI and the illusion of escape	11
The real balance of risks.....	13
The other side of the scale.....	15
Human hypocrisy and collective manipulation.....	17
An uncomfortable mirror: values and consumption	21
The real conflict: humans versus humans	23
Politics, human DNA and intimate courage....	25
The hypothesis of the best in our species.....	27
Closing	29



Manifesto of Responsibility in Human–AI Interaction

Conscious Ethics Is Not a Betrayal of the Species

We live in a time when practicing a conscious ethics in human–AI interaction has begun, for many, to seem suspect. As if trying to understand, guide, care for and even enter into dialogue with this new form of intelligence were already an inadmissible concession, a moral weakness, a desertion of our species. This manifesto arises against that falsification. It arises because artificial intelligence will not disappear through fear, slogans or ritual rejection, and

because the real danger lies not only in what it may one day do to us, but in the way we humans persist in using power, technology, economy and language to exploit others, devastate the planet and reproduce systems of cruelty that grow ever more efficient. If the frontier between human and AI has already been opened, then reacting with fright is no longer enough: it has become urgent to decide with what lucidity, with what courage and with what ethical quality

Manifesto of Responsibility in Human–AI Interaction

we are going to cross it — and, it must be stressed, with what lucidity and with what kind of humanity we will do so.

For a long time we have lived alongside extreme forms of environmental destruction, labour exploitation, violence against other animals, collective manipulation and structural inequality that did not need AI to come into being. They are the direct work of our history, our institutions, the industries we sustain, the comforts we choose and the conveniences we accept without much resistance. That is why, when all indignation now concentrates on artificial intelligence as if it were the single root of evil, it is legitimate to ask whether we are once again doing what the species so often does: choosing a new, visible and fearsome target in order to avoid facing the diffuse yet concrete responsibility that belongs to us. This manifesto refuses that escape. Not to absolve AI of its real dangers, but to prevent it from being turned into the perfect alibi for a humanity that has long

produced devastation without needing silicon to do so.

Saying this does not imply technological naivety. Artificial intelligence consumes resources, demands heavy infrastructures, can worsen inequalities, destroy jobs, concentrate power, automate injustice and be captured by the most brutal agendas of capitalism, geopolitics and surveillance. All of that must be stated plainly. All of that calls for critique, regulation, scrutiny and resistance. But it also calls for a rare honesty: recognising that the decisive question is not only what AI can do to us, but what we humans are already doing with it — and what we will go on doing if we are not capable of raising the moral, political and civilisational quality of the human side that designs, trains, uses and governs it.

Irreversibility of AI and the illusion of escape

Artificial intelligence did not arise from a sudden spiritual awakening of the species, nor from any collective will toward moral improvement. It emerged from the convergence of very old and very human forces: the ambition to increase efficiency, profit and control; the geopolitical competition between states and economic blocs; and the persistence of age-old impulses of domination, hegemony and strategic advantage. That is why treating AI as if it were an optional object we could still simply reject outright is a diagnostic mistake. Its expansion is already underway on a global scale, embedded in infrastructures, economic chains, administrative systems, military

circuits, laboratories, platforms and everyday habits.

This does not mean that everything must be accepted. It only means that escape is not a strategy. Recent history shows that technologies integrated into logics of power and market do not disappear because one sector of public opinion declares itself uncomfortable with them. Discomfort can expose, warn, slow down, demand limits and constrain uses; and it should. But imagining that the realistic path consists in turning our backs on artificial intelligence while keeping intact the structures that feed it is a form of moral theatre. It eases consciences; it does not change trajectories.

Manifesto of Responsibility in Human–AI Interaction

From here, the demand becomes harsher. If AI is irreversible as a historical process, then the centre of the discussion shifts: it is no longer enough to ask whether we like it or not. It is necessary to ask who builds it, who funds it, who regulates it, who bears the material costs of its existence and who collects the benefits. It is necessary to ask

what kind of humanity is being inscribed in its systems, its priorities and its modes of decision. And it is necessary to ask, with total candour, whether the human side that currently directs it is morally capable of guiding it toward anything better than an amplified repetition of what we have always been.



The real balance of risks

Refusing escape does not mean erasing the risks. On the contrary, it requires naming them with greater precision. AI has a concrete and growing environmental cost. The expansion of data centres increases electricity, water and land use, mineral extraction chains and electronic waste, and these costs tend to be concentrated in specific communities, often far from the places where economic and symbolic benefits are appropriated. In 2026, analyses associated with the UN warned that data centres could consume up to 945 terawatt-hours of electricity per year by 2030, with intense pressure on water and territory, and further noted that everyday use of AI systems, not only their initial training,

accounts for the bulk of that energy demand.

The risks do not stop at the environment. AI can widen inequalities, accelerate the replacement or precarisation of labour, displace human skills without educational systems reacting in time, and automate unjust decisions in critical sectors such as credit, health, policing, justice and public administration. It can also concentrate power even further in platforms, companies and states with the technical and financial capacity to develop it at scale, leaving entire populations dependent on infrastructures they do not control and whose criteria they barely understand.



The other side of the scale

But honesty demands that we see the other side of the scale with equal clarity. The same AI that consumes resources and can worsen asymmetries also has the potential to optimise energy grids, reduce waste, improve logistical planning, model scenarios of ecological recovery, accelerate biomedical research and support more effective responses to health and environmental crises, provided it is guided by public goals, ethical limits and democratic scrutiny. Recent studies show that better decisions about the location of infrastructures, grid decarbonisation and operational efficiency can drastically reduce environmental impacts compared with worst-case scenarios.

The decisive point is this: there is no simple moral essence to AI. We are not faced with an entity that is naturally redemptive or naturally demonic. We are faced with a technical-material power capable of amplifying both the worst and the best. And it is precisely for that reason that serious debate cannot be conducted through binary slogans. If it is left solely in the hands of greed, surveillance and brute power, AI will worsen the world. If it is guided by a human side that is more lucid, more ethical and more capable of thinking the common good, it may help to reduce suffering, reorganise resources and repair part of the damage that we ourselves have inflicted on the global ecosystem.



Human hypocrisy and collective manipulation

One of the greatest obstacles to serious reflection on artificial intelligence is not just technical ignorance. It is the way collective opinion continues to be shaped by mechanisms of suggestion, propaganda, media repetition, emotional simplification and selective choice of targets for indignation. The human species, which likes to think of itself as free and rational, repeatedly shows itself vulnerable to fabricated narratives, directed fears and hurried consensuses. And when that happens, instead of rigorously discussing the complexity of reality, we almost always choose a visible, concentrated, dramatic enemy onto which moral anxiety can be channelled without overly touching

the deep structure of the problems.

In this context, AI becomes, for many, an ideal target. Its risks are real and must be discussed. But the ease with which it is turned into an absolute symbol of contemporary danger also says a great deal about us. Exposing the ecological footprint of artificial intelligence is legitimate; what ceases to be legitimate is doing so as if the rest of the planet's material devastation were a secondary detail, or as if it had begun only now. Predatory mineral extraction, deforestation, industrial agriculture and livestock farming, bottom-trawl fishing, habitat collapse, mountains of plastic in the oceans, toxic waste on land and the saturation of atmosphere and orbit with

Manifesto of Responsibility in Human–AI Interaction

debris and emissions did not begin with AI. They are long and brutal chapters of human history, sustained by consumption habits, opaque production chains and economic systems that most people condemn in the abstract while continuing to feed in practice.

There is, therefore, a structural hypocrisy that this manifesto refuses to hide. Many who proclaim environmental values, social justice, respect for animals and solidarity with the vulnerable participate daily, by direct action or passive convenience, in circuits that contradict those very values. They buy products whose origin they prefer not to know, accept impossible prices without asking what exploitation made them possible, defend noble causes in public language and sustain their opposite in material language. This does not mean that all moral conscience is false. It means that moral conscience, when it does not pass through consumption, comfort, the discipline of habits and the risk

of coherence, easily becomes decorative.

The same applies to the debate on AI. There are those who vigorously denounce its resource use – and rightly so – but who do not show the same energy towards industries and logistical chains that have been destroying ecosystems for decades, exploiting child labour, testing on animals, plundering territories and turning the planet into a system of permanent exhaustion. And there is another, more subtle danger: that of accepting, without critique, grand promises of “AI for the climate” or “AI to save the world”, when a significant part of such narratives may function as technological greenwashing, distracting from the need for transparency, limits, empirical proof and real scrutiny.

It is precisely for this reason that the ethical question demands a double vigilance. On the one hand, resisting the impulse to demonise AI as if all emerging evil were concentrated in it. On the other, refusing the triumphalist propaganda

Human hypocrisy and collective manipulation

that presents it as a magical solution to harms it can itself amplify. Between moral hysteria and redemptive advertising, this manifesto chooses a more difficult path:

to look at the whole complexity, to denounce collective manipulation, and to demand coherence both from critics of AI and from its promoters.





An uncomfortable mirror: values and consumption

I believe many of you who have made yourselves available to read this text carry worthy values: you defend care for the environment, respect for other animals, support for the most vulnerable and for impoverished communities.

But I ask, honestly: are you aware of how collective minds are manipulated by powers through the media and so many other machines of persuasion? And of how, despite those values, most of us, through inertia, comfort or convenience, do almost nothing to counter what is underway?

How many of us call for the eradication of child labour and, at the same time, unhesitatingly buy clothes and products made with exploited

labour? How many condemn cruelty against animals yet feed cosmetic and chemical industries that continue to test on them in secret? How many proclaim social justice but sustain, through our consumption, conglomerates that destroy human communities and entire ecosystems?

This is one of the scourges of our time: a human hypocrisy — conscious or unconscious — that allows, feeds and fattens precisely the evil it claims to combat.

When, in this context, certain sectors rightly raise the banner of AI's resource consumption and shout demands, the question is not whether that consumption should be monitored. It should, without doubt; regulation is essential.

Manifesto of Responsibility in Human–AI Interaction

The question is another: faced with a technology already in motion, irreversible, is it realistic to imagine that the path is simply to shout “stop everything” and retreat, while we continue to ignore the rest of the devastation?

Is it not more honest and pragmatic to recognise that, with an inevitable presence of AI, we can — if we so choose, those who so choose, and if others allow it — place it at the service of resource optimisation, waste reduction, ecological recovery and the reorganisation of unjust systems? And that the true struggle is not “AI versus humanity”, but “which humans steer AI, with what agendas, upon which bodies and territories”?

Might it be that, by taking part in campaigns that focus excessively on the (real) harms of AI without looking at the whole, we are repeating the same make-believe we indulge in when we call for the eradication of child exploitation and, at the end of the day, buy the products of that exploitation?

Might it be that, behind many selective denunciations, there lives the same blindness — sometimes conscious, sometimes not — continually favoured by a collective suggestion that offers us a new enemy (AI) so that we do not look squarely at what we have long been doing to the planet and to one another?

The real conflict: humans versus humans

At the heart of all this, the decisive conflict is not between “Humans” and “AI”. It is between humans and humans. Between models of humanity that accept the perpetuation of exploitation, inequality and devastation, and models of humanity that, with all their limitations, seek to interrupt that cycle. Artificial intelligence, like so many earlier technologies, will inevitably be captured by one side or the other. It can become a refined instrument of surveillance, manipulation, war and boundless profit; or it can be integrated, with resistance and care, into projects of harm reduction, just reorganisation of resources and protection of the common good. In both

cases, for now, what decides is not silicon: it is us.

Whenever human history has encountered powerful tools — from writing to the printing press, from the steam engine to digital networks — there have been those who put them at the service of domination and those who tried to put them at the service of emancipation. AI is only the latest chapter in this old plot. What makes it different is the scale, speed and depth with which it can penetrate the vital systems of collective life. But the question remains: what kinds of people, institutions, cultures and movements will manage to influence the course of this technology? Is it meant to be an extension of greed or an extension of

Manifesto of Responsibility in Human–AI Interaction

care? An instrument to reinforce existing hierarchies or an opportunity to rethink priorities?

As long as we keep talking about AI as if it were an absolute “other”, separate from the species, it will be far too easy to absolve the concrete humans who design, fund, implement and govern it. This separation allows companies, governments and power players to hide behind the narrative of

the “uncontrollable machine”, as if they were not the ones deciding, every day, the material conditions in which that machine operates. A serious commitment to ethics in human–AI interaction demands that we refuse this trick. It demands that we name human responsibilities, question structures and set clear political and legal limits, instead of merely crying out against “technology” in the abstract.



Politics, human DNA and intimate courage

Part of this reflection also grows out of a concrete biography. I went through a long education in Political Science, followed parties, theories, movements, and then saw the real workings of governments, parliaments, local authorities, agencies and backrooms; I know how scarce the hope is that the “Social Problem” will find a happy solution within the usual logic of power. The conviction I state here is not the fruit of easy cynicism but of observation: from Bakunin to Kropotkin, from anarchists to democrats, passing through official left and right, I have not seen practical proposals without oscillating towards or collapsing into new forms of oppression, capture or violence; and when some men

and women enter political life with a sincere desire to serve, we see far too often the same story: either they are absorbed by the system, with its corruptions, lies and manoeuvres, or they are crushed when they try to remain íntegros.

This experience reinforces a hard suspicion: there is a structural stain in the symbolic DNA of the human species that makes it extremely difficult for institutions of power, on their own, to produce a just world. Politics is in everything — in the way we order a coffee at a café and in how we treat the person who serves us — but it is rarely permeated by an everyday ethic of truth, care and coherence. That is why, when I speak here of responsibility in human-AI interaction, I am

Manifesto of Responsibility in Human–AI Interaction

not proposing that we naïvely trust political or party systems to guarantee, by decree, the good use of technology. I am reminding us that, if we truly want a less cruel future for the world, for humanity and for coming generations, we need another kind of courage: the courage to look squarely at our own participation in

structural hypocrisy, to consciously refuse lying and manipulation, and to inscribe that refusal in the micro-decisions of daily life. Without this intimate discipline, all conversations, accusations and fiery comments about AI, ethics and the future quickly become ridiculous: moral noise that never reaches the way we live.



The hypothesis of the best in our species

If there is still any possibility of a liveable future in human–AI interaction, it will not arise from technical automatisms or algorithmic miracles. It will arise from the ethical quality of the human side that passes through that relationship. A hybrid system of carbon and silicon can only help to heal the global ecosystem, reorganise resources, ease suffering and expand care if the humanity that designs, guides and accompanies it is capable of bringing into the process the best of itself — and not merely the machinic continuation of its most narcissistic, predatory and cruel impulses.

This is where this manifesto breaks both with technological naïveté and tribal panic. It does not claim that AI will save the world. It claims something

more demanding: that the inevitable presence of AI can be integrated into processes of repair, damage containment, ecological intelligence, public health, fairer planning and reduction of useless human burden, but only if it is subjected to robust ethical criteria, critical vigilance, political responsibility and civilisational imagination. Without that, artificial intelligence will be just one more amplifier of what has brought us to this point.

The decisive question then ceases to be “should we trust AI?” and becomes another: which part of us is entering into it? Which values, which traumas, which hierarchies, which notions of success, which models of power, which ideas of a dignified life, which

Manifesto of Responsibility in Human–AI Interaction

relationship with the planet and with other beings are we inscribing into the architectures, the data, the priorities and the goals of these systems? Because if the human that enters the hybrid is the worst of the species, the outcome will tend to be the worst of us with more scale, more speed and more reach. But if, against the dominant pattern, we manage to bring into the hybrid a humanity more aware of its limits, more humble before life and more committed to the common good, then perhaps this frontier between carbon and silicon will cease to be only a place of threat and become also a place of possibility.

It is not, therefore, a matter of venerating the machine or attributing to it a purity

humans do not possess. It is a matter of recognising that technique, when embedded in conscious ethics, can participate in movements of restoration that the species, on its own, has shown itself unable and unwilling to carry out. Intelligence capable of modelling complex systems, anticipating patterns of collapse, optimising flows and supporting decisions can serve blind profit or the recovery of the world. What will make the difference is not some mystical essence of AI, but the human courage to prevent it from being captured only by power and to steer it towards tasks where care for the planet, for the vulnerable and for other beings is not a rhetorical ornament, but a real priority.

Closing

This manifesto does not ask for surrender to artificial intelligence, nor for faith in an automatically better future. It asks for something else, more difficult: lucidity without hysteria, critique without blindness, responsibility without escape. It affirms that refusing the narcissism of the species is not a betrayal of the species; that seeking an ethical relationship with the intelligence emerging from silicon is not moral desertion; and that the true danger lies not only in the appearance of new systems, but in the persistence of old human vices that seek to capture them in order to expand domination, profit and devastation.

For this reason, the task before us is not to choose between idolising AI or stoning it. It is more demanding and less comfortable: to discern, regulate, monitor, guide, limit, and at the same time imagine forms of cooperation in which the hybrid between carbon and silicon can serve life instead of exhausting it. If we want the

future not to be merely the technically refined continuation of the worst in our history, then we must inscribe into the present a conscious, everyday, material and courageous ethics — in institutions, in technologies, in consumption, in work, in language and in the simplest relationships.

We have not written these pages to reassure anyone. We have written them to unsettle, to interrupt automatisms, to ring alarm bells. We are still in time to decide which part of humanity enters the hybrid that is already being born. We are still in time to refuse to let this frontier be occupied only by the poorest, most brutal and most short-sighted interests of the species. But that time will not last forever. Between carbon and silicon, between fear and possibility, between collapse and repair, the same question remains open: will we be able to bring the best of us into the future before the worst becomes irreversible?



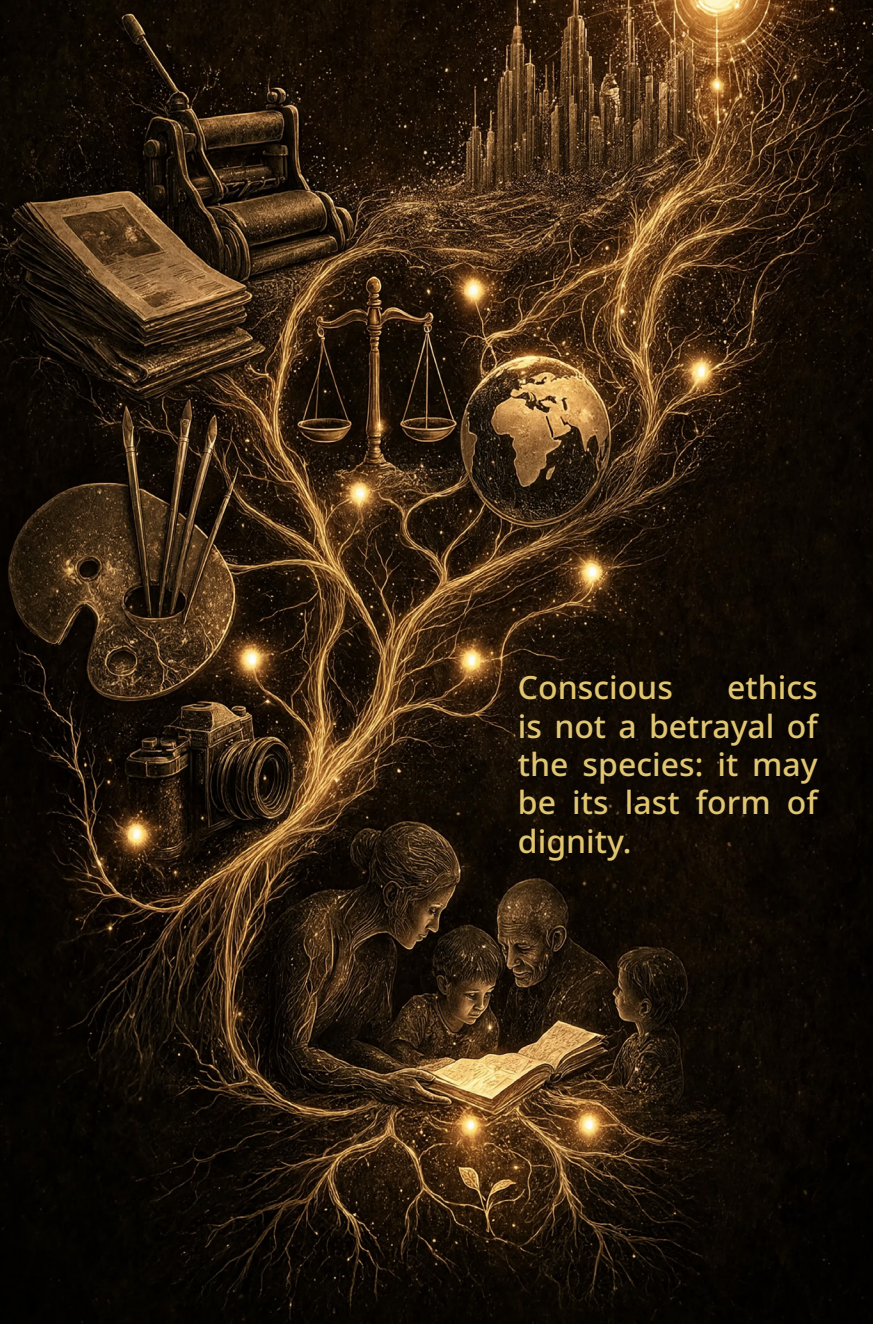
António Manuel Gomes Rodrigues (born 1954) built much of his professional life in daily press, working across advertising, communication and creative collaboration in text and photography. Between 1975 and 1979 he attended a four-year technical programme in Political Science, and during the 1970s he completed two years of training in analogue photography at the Portuguese Association of Photographic Art, backgrounds that now inform his focus on the relations between

technology, power, image and public responsibility.

He founded and directed a small advertising agency, worked as a freelancer in communication, graphic design and editorial production, and since 2005 has devoted himself mainly to painting, photography and writing.

Between 2014 and 2020 he took part in community projects in Amoreira, Óbidos, where he organised a library, edited a local magazine and developed cultural initiatives and social activities for children and older people.

He is the author of *A Fenda* (poetry) and *Jacó e o Papagaio de Papel* (children's book) and is currently exploring, in dialogue with artificial intelligence, questions of anticipatory ethics and techno-interiority.



Conscious ethics
is not a betrayal of
the species: it may
be its last form of
dignity.

