NEW HUMANISM IN THE TIME OF NEUROSCIENCE AND ARTIFICIAL INTELLIGENCE: AN INTERNATIONAL ACTION-RESEARCH PROJECT*

MATHIEU GUILLERMIN

Abstract:

This contribution presents NHNAI (New Humanism in the time of Neuroscience and Artificial Intelligence), an action-research project coordinated by Lyon Catholic University (UCLy) under the aegis of the International Federation of Catholic Universities. With a consortium of thirteen partners on five continents, NHNAI aims to harness the resources of partner institutions to strengthen the ethical orientation of all stakeholders, through collective exploration of what it means to be human in the age of AI and neuroscience. We outline the project's main philosophical orientations, as well as its organizational modalities, and we propose an overview of its first results.

Keywords: Artificial Intelligence, Neuroscience, Humanism, Ethics, Collective exploration.

Recent developments in the fields of Artificial Intelligence (AI) and Neuroscience (NS) raise a host of ethical issues. Some of these deeply question (or even disrupt and threaten) our societies, our lifestyles and the ways in which we relate to others and the world around us. How can we make sense of the ever-increasing insights provided by neuroscience into the biological aspects of our cognition, our free will and even our consciousness? Should we use this knowledge to capture people's attention and more or less guide their behavior? In commercial communication? Political communica-

^{*} This contribution is a translation, with the help of DeepL of a text in French, *Nouvel Humanisme au temps des Neurosciences et de l'Intelligence Artificielle : un projet international de recherche-action*, « RETM » CCCXXIX, Hors-série 2025 (in press).

tion? Should we be looking for an alternative to an economic model based on an apparent gratuity, but with the counterpart of a certain form of predation on attention time? How should we relate to robots or virtual personas capable of imitating the cognitive or affective behaviors of living beings, and humans in particular, in an increasingly convincing way? What role can AI technologies play in public administration and democratic life, in healthcare and education?

The project New Humanism in the time of Neuroscience and Artificial Intelligence (NHNAI)1 is a action-research project aimed at reinforcing the ethical orientation capacities of all players concerned by these types of issues linked to NS and AI. Under the aegis of the International Federation of Catholic Universities, the NHNAI project is coordinated by the CONFLUENCE: Sciences et Humanités (EA1598) research unit of the Lyon Catholic University (UCLy), and brings together eleven other partners from around the world². The aim of NHNAI is to put the resources of participating academic and scientific institutions at the service of collective reflection, in society, on what it means to be human, in order to better navigate in the age of AI and NS. The NHNAI project intends to contribute to strengthening collective capacities for ethical orientation along two structuring and intertwined axes: 1) contribute to the development of an "ethical compass" through reflection on the theme of humanism, and a renewed exploration of what it means to "be human" in the age of AI and NS. 2) to contribute to the development of this compass, not only through academic research, but also through an action-research initiative with the societal players concerned.

1. Humanism, and ambiguous though indispensable idea?

In which direction, to which magnetic north, should a compass that is supposed to support ethical orientation in the age of AI and NS point?

¹ For complementary information and more detailed results, see: https://nhnai.org/.

² In addition to the International Federation of Catholic Universities and the Lyon Catholic University, the NHNAI network currently includes the following partners: Santa Clara University and Notre Dame University (USA), Cégep de Sainte-Foy (Québec, Canada), Pontificia Universidad Católica de Chile (Chile), The Catholic University of Eastern Africa (Kenya), Fu Jen Catholic University (Taiwan), Université de Namur (Belgium), Universidade Católica Portuguesa (Portugal), Università di Roma LUMSA (Italia), Université Catholique de Lille and ICAM (France).

Faced with this major question, it would seem difficult to avoid an in-depth reflection on what "being humans" means, on what we are, but above all on what we want or need to be (hence the theme of humanism in its axiological sense). Numerous major principles of AI ethics appeal to the idea of the human: AI must be *human*-centric, at the service of *human* flourishing, the *human* must be kept on or in the loop, etc. More fundamentally, we can even see an almost organic link between ethics and the quest to understand what it means to be human. Ricoeur, for example, defines « ethical aim » as « the aim of the "good life" with and for others in just institutions " without at the same time reflecting on what it means to be human?

And it is indeed an effort to explore and deepen that is at hand. The notion of humanism is far from being clear and consensual, unproblematic and ready to serve the purpose of ethical orientation. Even taken only in the context of its emergence (in Europe with the Renaissance, then Modernity and the Enlightenment), the notion already incorporates multiple, disparate and sometimes conflicting dimensions⁴. Subsequently, many currents have opposed and continue to oppose the humanism of modernity head-on (anti-, post-, trans-humanism), highlighting its difficulties and limitations. Are there really precise characteristics that distinguish the human from the non-human? Are they universal? Isn't the almost absolute primacy granted by modern humanism to human autonomy and rationality problematic? Hasn't it led to the myth of the human as master and possessor of nature, with an automatic link between techno-scientific development and human progress? Today, these difficulties seem to be reinforced by the theoretical and scientific contributions of AI and NS, as well as by their technological spin-offs.

But perhaps the solution is not to reject the notion of humanism outright. It is also possible to put the notion back to work, through a renewed exploration capable of preserving and deepening the most fruitful contributions of the humanism of Modernity and the Enlightenment. It is common, for example, to retain from Kant only the idea of a human reason that can reach *a priori* conclusions in the spheres of science (about phenomena) and morality (with the categorical imperative) - pure theoretical and practi-

³ P. RICOEUR, *Soi-même comme un autre*, Éd. du Seuil, Paris, 1990, p. 202, our translation : « visée éthique » is defined as « la visée de la "vie bonne" avec et pour autrui dans des institutions justes ».

⁴ See for instance the illuminating historical and philosophical presentation of the notion of humanism: T. Davies, *Humanism*, Routledge, New York 1997.

cal reason. It is therefore all too common to point to the failure of Kantian epistemology (for example, the overcoming of Newtonian physics) and to discredit Modernity and humanism. But this would be to ignore the heart of Kantianism, with its central idea of the passage of humanity from minority to majority (the individual can and must think for himself, "Sapere Aude") and the importance of the fundamental couple freedom-responsibility, which leads Kant to place practical reason, and above all the faculty of judgment, at the center of his entire philosophical system⁵. A « critical » rather than dogmatic modernity is conceivable⁶.

Depriving ourselves of this type of input in the age of AI and NS would only fuel the difficulties. It seems far more fruitful to deepen these resources, and to couple them with the exciting insights of AI, NS and cognitive science⁷ in order to outline the contours of a new humanism, opening up to a renewed understanding of our freedom, our intelligence, or our capacity to judge. From the limited epistemological and philosophical field in which I work, I see all the fruitfulness and necessity of an understanding that goes beyond the formal and algorithmic aspects to the living, lived dimension of who we are⁸. But it's more broadly an interdisciplinary exploration that we need to embark on, without prejudice as to which disciplines are likely to make a fruitful contribution. Anthropology, philosophy, sociology, NS, cognitive science, psychology, computer science, theology, law... and many more besides. Such diversity is essential if we are to address the issue of humanism in the age of AI and NS.

2. Structuring philosophical and epistemological commitments

As we have just seen, an interdisciplinary approach seems necessary if we are to successfully explore the notion of humanism. But the NHNAI

⁶ B. Feltz, La science et le vivant. Philosophie des sciences et modernité critique, De

Boeck supérieur, Louvain-La-Neuve 2014.

⁷ See for instance: H. Mercier et D. Sperber, *L'énigme de la raison*, Odile Jacob, Paris 2021. A. Damasio, *Sentir et savoir : Une nouvelle théorie de la conscience*, Odile Jacob, Paris 2021.

⁸ It seems to me, moreover, that the contributions of a phenomenology of life such as the one by Michel Henry could prove particularly valuable in this perspective. See: M. Henry, *Incarnation. Une philosophie de la chair*, Éd. Du Seuil, Paris 2000.

⁵ A. Philonenko, L'œuvre de Kant - La Philosophie Critique Tome I : La Philosophie Precritique Et La Critique de La Raison Pure, VRIN, Paris 1983. A. Philonenko, L'œuvre de Kant - La philosophie critique : Tome II : Morale et politique, VRIN, Paris 1997.

project does not limit itself to such an interdisciplinary effort in a purely academic configuration. The NHNAI network partners in the nine countries involved are striving to contribute to strengthening the capacities for ethical orientation of all players concerned by the challenges raised by NS and AI, by putting academic and university resources at the service of collective reflection with societal players on what it means to be human in the age of AI and NS. Indeed, an academic and interdisciplinary approach to the question of humanism is essential, but certainly not sufficient to meet the ethical challenges raised by AI and NS.

2.1 « Wicked problems » and citizen involvement

The NHNAI project thus adopts a resolutely anti-technocratic approach, distancing itself from the traditional, highly unidirectional model of science as the holder of absolute truths which, in a position of epistemic surplomb, is tasked with informing society9. In this traditional model, scientific expertise is universally recognized as « the crucial component of decision-making 10 ». Even supposing that a college of experts, albeit interdisciplinary, could determine what it means (must mean) to be human in the age of AI and NS, this solution would still have to be understood, accepted and implemented. But more fundamentally, this type of relationship between (scientific) expertise and political decision-making needs to be called into question, especially for epistemological reasons. Generally speaking, the idea that science is neutral and radically independent of politics and ethics is illusory¹¹. And the illusion becomes particularly dangerous when dealing with what we might call « wicked problems 12 » or « socially sensitive issues », such as « the nuclear phase-out, the dangerousness of GMOs, global warming », that is to say, « fuzzy problems relating to

 $^{^9}$ M. Gibbons, Science's new social contract with society, in « Nature » CDII, 1999, pp. C81-C84.

¹⁰ S.O. Funtowicz and J.R. Ravetz, *Science for the post-normal age*, in « *Futures* » XXV, 1993, p. 739-755.

¹¹P. KITCHER, Science, Truth, and Democracy, Oxford University Press, New York 2001. ¹² C. Pohl, B. Truffer and G. Hirsch-Hadorn, Addressing Wicked Problems through Transdisciplinary Research, in R. Frodeman (ed.), The Oxford Handbook of Interdisciplinarity, Oxford Academics, 2017, pp. 319-331.

decision-making¹³ ». These problems are as much scientific (including the human sciences) as they are political.

To tackle this type of thorny challenge, the contribution of technical and scientific expertise is essential, but not sufficient. A response will never be purely technical, but will also inevitably correspond to a political commitment (possibly implicit), to the adoption (voluntary or otherwise) of a way of living together. Just think of the algorithms used to recommend content on the Internet, designed above all to capture attention? How can we do things differently? With what economic model? Enlightened collective reflection and deliberation seem to be indispensable, on pain of tipping over into a technocratic approach that would be as ineffective as it would mutilate the freedom and autonomy of each and every individual. As Pope Francis affirms in his message for the fifty-seventh World Day of Peace, ensuring that AI serves « the cause of human fraternity and peace », it is not « the responsibility of a few but of the entire human family 14 ». It's a question of freedom, but also of responsibility. It is the responsibility of each and every one of us to recognize, to welcome, to make ourselves available to the « ethical goad », to the « sincere concern for our common home and [the] real preoccupation about assisting the poor and the needy discarded by our society » (in the words of Pope Francis *Laudate Deum*¹⁵).

This affirmation, not only of the right (notably in the name of human freedom), but also of the responsibility of every member of the human family to engage in a collective effort of ethical orientation is highly significant for thinking about the specific responsibility of scientific, academic and university communities and expert groups. Faced with wicked problems such as those raised by AI and NS, the contribution of scientific and technical expertise must be humble and contribute to collective efforts in response to this ethical goad. It is in this spirit that the dimension of service to society (action-research) is placed at the heart of the NHNAI project. Academic exchanges within the network have great intrinsic value.

¹³ C. Chauvigné et M. Fabre, *Questions socialement vives : quelles approches possibles en milieu scolaire ?*, in « *Carrefours de l'éducation* » LII, 2021, pp.15-31, our translation.

¹⁴ Pope Francis, *Message for the 57th world day of peace*, 1 January 2024, https://www.vatican.va/content/francesco/en/messages/peace/documents/20231208-messaggio-57giornatamondiale-pace2024.html

¹⁵ Pape François, *Laudate Deum*, Apostolic Exhortation, 4 October 2023, https://www.vatican.va/content/francesco/en/apost_exhortations/documents/20231004-laudate-deum.html, sections 29-33.

Encouraging them is an objective in itself for the project. But it's crucial that these exchanges are put at the service of collective reflection, to enrich the societal discussions and debates organized by the various partners in the nine countries involved.

2.2 More than a juxtaposition of opinions, a joint elaboration respectful of complexity

Another distinctive feature of the NHNAI project is the way in which societal discussions are approached and organized. To put it succinctly, the NHNAI project aims to go beyond the simple juxtaposition of the individual opinions of the participants in the discussions, by encouraging a collective and rational elaboration of a shared understanding of the question of humanism in the age of AI and NS. In this respect, the epistemological background is of great importance. NHNAI stands in firm opposition to a culturally widespread conception that draws a principled demarcation between the realm of facts, susceptible to rational (scientific) investigation, and the realm of ought-to-be, politics or ethics, over which pure individual freedom (the right to think what you like, pure freedom of opinion) would preside. This dichotomous approach is partly based on the ideal of pure inquiry and an understanding of rationality as neutrality. In this perspective, rationality is achieved by relying solely on absolutely indubitable elements (such as raw empirical data and logic or mathematics), to the exclusion of any arbitration, choice or subjective feeling. From the outset, this conception condemns any possibility, in the political, ethical or moral fields, of a collective, rational effort to establish a shared understanding recognized as valid by the participants. Although rooted in a legitimate concern for emancipation and the preservation of democratic pluralism, this approach, by failing to lay down any conceptual or epistemological marker for reflection on what ought to be, opens a royal road to the abuse of freedom, to solipsistic or communitarian alienation, to a sterile pluralism which, being only "window-dressing", weaves nothing.

If only this dichotomous conception (factual rationality - pure freedom of ethical and political opinion) were founded philosophically and epistemologically, we could do no more than deplore this unsurpassable cognitive condition. But everything points to the fact that it doesn't hold. The neutrality of factual and scientific inquiry, which would confer on it

the status of pure inquiry, is an illusion, culturally tenacious perhaps, but easy to deconstruct philosophically¹⁶. No process of reflection or inquiry, be it scientific, factual or any other, can dispense with judgment, with being rooted in feeling, in a largely intuitive ability to recognize the obvious, to distinguish between what is reasonable and what is not¹⁷. There is therefore no reason in principle, no epistemological reason, to banish rationality and the relationship to truth from reflection on *ought-to-be* (from moral, ethical and political reflection). In this field too, we can agree on the validity of at least certain assertions, either because they are based on sufficiently solid and shared evidence, or because they are supported by reasoning that is itself rooted in such evidence. And so, perhaps more than the possibility – and even *since* this possibility exists – we actually have a duty and a responsibility to seek out the greatest possible agreement also in this area of *ought-to-be*.

It is in this spirit that the NHNAI project encourages a collective and rational elaboration of a shared understanding of the question of humanism in the age of AI and NS. The aim is to try and agree on statements that are sufficiently solid to create a kind of common ground from which to move forward and orientate ourselves. And once again, it's not *simply* a question of pooling subjective opinions that are shared, leaving aside divergences. The epistemological stance adopted in NHNAI allows us to go much further. While the aim of the exchanges is indeed to establish a kind of consensus, this consensus relates first and foremost to the validity of the contributions to the reflection. And, once freed from the straitiacket of pure inquiry, nothing prevents a contribution recognized as valid by the participants from coming into tension or conflict with another contribution also recognized as valid for other reasons. So, instead of simply tolerating others' differing opinions in the name of abstract democratic pluralism, we can try to reach a consensus on divergences as a sign of legitimate tensions inherent to the question being explored (and not of divergences between individuals and their respective subjective opinions).

Through its action-research effort in the service of collective reflection in society, NHNAI therefore intends to contribute to the development of

¹⁶ M. GUILLERMIN, Non-neutralité sans relativisme? Le rôle de la rationalité évaluative, in L. Brière, M. Lieutenant-Gosselin and F. Piron (eds.), Et si la recherche scientifique ne pouvait pas être neutre?, Éd. science et bien commun, Québec 2019, chapter 15, pp. 315-338.
¹⁷ M. GUILLERMIN, Bulletin Humanisme et rationalité, in « Revue des sciences philosophiques et théologiques » CVI(3), 2022, pp. 489-543.

such a shared understanding of what it means to be human in the age of AI and NS, an understanding that builds consensus in particular on the complexities of this question of humanism. NHNAI aims to support the collective effort to identify and reflect on such points of tension and complexity intrinsic to the topic of humanism in the age of AI and NS in order to bring out commons capable of strengthening the ethical orientation capacities of those involved.

3. NHNAI project implementation

The current phase of the NHNAI project runs from January 2022 to December 2025. It is punctuated by the following stages:

- January December 2022: Academic experts from the NHNAI network came together for a workshop in March 2022. The aim of this workshop was for everyone to contribute their vision of what, in their discipline, would be particularly important to mobilize as part of societal reflections on what it means to be human in the age of AI and NS. At the same time, the partners worked in collaboration with representatives of local societal communities to bring out the issues to be prioritized for discussion. The meeting of these two activities generated a synthesis of the expert contributions proposed by the network's researchers in response to the major groups of questions defined by societal actors ¹⁸.
- January 2023 April 2024: NHNAI network members organized a first wide wave of debates and exchanges in society. In each of the nine countries involved, face-to-face workshops with the various societal communities were organized. Discussions focused in particular on what it means to be human in the age of AI and NS in the fields of education, health and democracy. Each workshop was restituted in the local language on an online discussion platform, to enable further exchanges¹⁹.
- May 2023 July 2024: NHNAI network members analyzed the various local corpuses of contributions to the societal discussion in

¹⁸ This first major deliverable of the NHNAI project is accessible here: https://nhnai.org/output_1-2/.

¹⁹ The various online debates can be found here: https://cartodebat.org/nhnai.

order to first generate local syntheses (for each country involved), and then international syntheses. Special attention was paid to the complexities and tensions emerging from the discussions. These initial results have been made available to those involved in the discussions (academic and extra-academic), in particular to enable them to delve deeper into these points of complexity²⁰.

 Ongoing - December 2025: A second wave of social debates and exchanges is currently underway. Its analysis will enrich the syntheses already produced. The network of researchers is continuing its work in support of societal reflection, in particular by endeavoring to provide relevant elements of expertise to shed light on the points of tension and complexity emerging from the discussions.

Let's add a few important figures to this panorama. The NHNAI network mobilizes some seventy academic actors from a wide range of disciplines, from the thirteen institutions currently participating in nine countries on five continents. The face-to-face workshops of the two waves of societal discussion brought together some two thousand participants. These included over seventy associations or organizations working with vulnerable or excluded people. The number of contributions restituted online or produced during the continuation of the exchanges amounted to more than three thousand seven hundred.

Through this organization, NHNAI strives to put the resources of its partner institutions at the service of societal reflection and capacity-building for ethical orientation of all stakeholders (including researchers in the network). The contributions of academic actors are made at different levels, always from a position of support and not of surplomb in relation to the discussion in the agora. A first type of contribution is organizational, enabling the opening of spaces (face-to-face and online) for exchange geared towards the collective and rational elaboration of a shared understanding of the question of humanism in the age of AI and NS. his notably involves an effort to mediate and facilitate exchanges. In particular, discussion organizers emphasize to participants the importance of giving reasons in support of assertions submitted to the discussion collective, as well as the distinction to be made between a willingness to express divergence from an idea that is nonetheless recognized as valid (a sort of "ves,

 $^{^{20}}$ The syntheses of this first wave of discussions are available online: https://nhnai.org/2023-results/.

but ...", possibly signalling a point of complexity) and the expression of disagreement as to the validity of an idea expressed by others (in this case, it's a question of expressing something like "I don't understand this idea, I don't understand why we should accept it in the shared understanding we're trying to build"). A second type of contribution concerns the synthesis of exchanges, with special attention paid to points of complexity. Finally, a third type of contribution consists in providing participants with insights from the researchers in the network, based on their disciplinary expertise (both to support the discussions and with regard to the points of complexity identified in the syntheses).

4. An overview of first results

This collective exploration of the question of humanism in the age of NS and AI has already led to the syntheses mentioned above. To illustrate these results, let's mention a few examples of places of complexity that have emerged from societal debates.

4.1 How to relate to machines?²¹

Some participants in the discussions (mainly in Portugal) point out that, as AI progresses, we will tend to develop machines (robots, chatbos) capable of imitating or simulating behaviors and capacities specific to humans and living beings, such as empathy, assertiveness, emotional and affective life. As a result, it will become increasingly tempting to relate emotionally to this type of machine (such as artificial companions or assistants, or robots for the care of people).

At the same time, many contributions to the discussions (in most of the participating countries) stressed the importance of not losing sight of the specificity of the living and the human as compared to machines. Machines are not conscious, do not feel emotions, cannot be wise, creative, critical or autonomous, and are not capable of spirituality, in the usual sense of these terms that implies rootedness in a lived experience, in a biological body.

 $^{^{21}\,\}mathrm{More}$ details available online: https://nhnai.org/focus-on-nexuses-of-complexity-transversal/.

At best, they can simulate convincing behaviors in these registers (notably through conversation), behaviors that human beings or living beings would have in given circumstances.

Also emerging (mainly in France) is the question of the rights to be granted to advanced robots or intelligent systems. From this point of view, many participants agree that AI cannot be a subject of law. The issue is widely described as speculative or science-fictional, without being less than interesting.

Thus, it is quite widely expressed in the discussions that it is necessary to resist the (increasingly real and powerful) temptation to perceive certain robots or AI systems as genuine people and try to relate to them affectively (as we would to a human, or even to another living being). We must resist the temptation to substitute interactions with machines for genuine human relationships.

4.2 The challenge of sovereign AI capabilities²²

Only participants from Kenya explored this theme in depth. The contributions show an enthusiastic recognition of the potential benefits of AI (drones, translation, help for the most vulnerable...). The prospects for economic development were also highlighted. Nevertheless, concerns are expressed about the possibility of losing control of locally-generated data and not developing national and local technological capabilities due to Kenya's limited sovereign digital infrastructure. There is therefore a risk of technological dependence, in some ways akin to a new form of colonization (technocolonialism or digital colonialism), with a possible drain on AI-related economic benefits.

4. Towards commons in support of ethical orientation

Identifying places of complexity, such as those briefly mentioned above, is an important milestone in the effort to develop commons that could strengthen the ethical orientation capacities of the actors involved, through

 $^{^{22}}$ More details available online: https://nhnai.org/focus-on-nexuses-of-complexity-democracy-5/.

a shared exploration of what it means to be human in the age of AI and NS. It's important to insist on this: the idea is not simply to report on the perceptions and opinions of the participants, particularly with regard to international and intercultural contrasts. It is also (and perhaps primarily) about mutual enrichment. Certain ideas expressed only in certain countries or by a small fraction of participants (such as the issue of technological sovereignty raised in Kenya) are likely to inspire other participants. A major challenge for the NHNAI network is now to deepen these commons and make them live. To this end, the network's researchers, drawing on their disciplinary expertise, are striving to shed light on these areas of complexity in ways that are relevant to collective reflection. The NHNAI network partners then propose to pursue societal exchanges by further exploring these areas of complexity. Any person of good will who feels concerned is warmly invited to take part in the exchanges and contribute to the collective effort towards a shared and refined understanding of what it means and ought to mean to be human in the age of AI and NS²³.

Acknowledgements: This work was carried out within the framework of the NHNAI project (https://nhnai.org/), supported by Porticus and the CONFLUENCE: Sciences et Humanités research unit (EA 1598) of the Catholic University of Lyon.

²³ Those wishing to find out more can find all the points of complexity identified to date and links to the debates at this address: https://nhnai.org/complexities/.