

Transforming Morocco's Prosocial AI Potential: A Multilevel Governance Framework, Prosocial AI Index and National Charter for Sustainable Human Development

Abdelmjid Lafram¹ and Salah Eddine Bahji²

¹School of Information Sciences, Rabat, Morocco

²Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco

alafram@esi.ac.ma

salaheddinebahji@research.emi.ac.ma

Abstract: Prosocial artificial intelligence is emerging as a new socio-technical paradigm, shifting AI systems from an instrumental and purely commercial vision to a new conception focused on social values, collective well-being, human dignity, ethics, inclusion and sustainable human development. This research aims to answer the following central question: how can a country in the Global South, such as Morocco, achieve leadership and transform its potential in Prosocial AI? To this end, this conceptual research adopted a qualitative triangulation combining document analysis, semi-structured interviews with national and international experts, and a comparative analysis of international best practices in Prosocial AI. This research explores Morocco's various strengths in prosociality that predispose it to become a pioneering country in Prosocial AI in Africa, highlighting the different actions taken within the framework of the welfare state and public policies, as well as institutional projects and programs, to make AI a driver of socio-economic development and social inclusion. This paper also examines the challenges and obstacles hindering the exploitation of this potential of Prosocial AI. This approach allowed us to develop a multilevel and multidimensional conceptual model. It also enabled us to propose a Prosocial AI Index, a multidimensional composite measurement index, as well as a national charter for prosocial AI, conceived as a normative instrument to guide and direct the development of prosocial AI within the context of an augmented welfare state. The findings of this study have allowed us to highlight the prerequisites and enabling conditions underpinning what we conceptualize as the foundations of prosociality in a Southern State, thereby positioning Morocco as a potential pioneering regional and African hub for Prosocial AI. They also offer a replicable model for emerging economies seeking to align AI innovation with ethical and societal objectives, serving both people and the planet.

Keywords: Prosocial artificial intelligence, Multilevel governance framework, Prosocial AI index, National AI prosocial charter, Sustainable human development

1. Introduction

The world is now transitioning from a VUCA environment to a new, more perplexing era with new societal vulnerabilities: the BANI era, characterized by Brittiness, Anxiety, Nonlinearity, and Incomprehensibility (Cascio, 2020). In this context of global challenges, the need for human-centered paradigms is becoming increasingly urgent, raising questions about the societal responsibility of Artificial Intelligence and its contribution to strengthening human well-being, inclusivity and social resilience (Shneiderman, 2022; UNESCO, 2021, OECD, 2019; United Nations, 2024; Floridi et al., 2018).

Indeed, today, we are witnessing a technological revolution marked by the considerable rise of artificial intelligence, which is developing and spreading throughout socio-economic, institutional, and environmental spheres, impacting all public and private institutions and organizations worldwide (Stanford Institute for Human-Centered Artificial Intelligence, 2026). Its development and global diffusion highlight two key observations. The first concerns the gap between the pace of technological progress and existing models of governance, regulation, and institutionalization. The second, also emphasized by the Stanford Institute for Human-Centered Artificial Intelligence (2026), is the predominance of profitability, technical, economic, and financial performance, and productivity often prioritized over societal, cultural, and human considerations, as well as the societal impacts of AI.

To govern and regulate the risks associated with artificial intelligence, several normative frameworks have emerged, the most dominant in terms of governance being responsible AI (OECD, 2024), ethical AI (UNESCO, 2021), trustworthy AI (European Commission, 2019), human-centered AI (OECD, 2019), and AI for the common good (ITU, 2023). These main paradigms are complementary and characterized by a lack of systematic operationalization. They focus primarily on the principles of compliance and good conduct, and risk management. Hence the need for an integrative framework oriented towards intentional and proactive prosociality and marked by the operational measurability of AI's societal impacts, a gap that our research addresses by proposing a Multilevel Governance Framework, a ProSocial AI Index and a National Charter for sustainable human development in the Moroccan context. In this context, Prosocial AI is of paramount

importance as an emerging paradigm and a promising conceptual alternative centered on humanity, collective well-being, equity, and social cohesion, aiming to foster inclusive and sustainable human development.

Few definitions exist for this emerging concept. They converge on the proactive and intentional integration of prosociality in the upstream stages of the design and development of AI-based solutions, placing humans at the heart of the technology. Indeed, Prosocial AI, while attempting to minimize negative effects or externalities throughout the design, development, deployment, piloting, governance, and evaluation processes of solutions, software, platforms, systems, ecosystems, and even organizations, in terms of digital and algorithmic prosociality. It operationalizes the principles in terms of measurable and auditable positive social impact.

The same emphasis of human and planetary well-being and social and human values is found in Walther's (2024) approach, which conceives Prosocial AI as *«a transformative approach aligned with the United Nations Sustainable Development Goals. It advocates for AI systems centered on human values, equity, and sustainability, aiming to enhance healthcare access, efficiency, and affordability, especially for underserved communities, promoting global well-being »* (Walther, 2024). She also defines Prosocial AI as *« AI systems that are tailored, trained, tested, and targeted to bring out the best in and for people and planet »* (Walther, 2024). Other authors such as Floridi et al. (2018) address similar concepts such as AI4People and Good AI Society but do not define Prosocial AI, while others introduce the concept of prosocial behaviors in LLM and conversational agents (Zhou, Y., et al. 2026 ; Kim, et al. 2022, Gupta, P. et al., 2026; Santos, F. P., 2024; Willis, R, 2026).

Morocco is positioning itself among the pioneering countries in Prosocial AI by launching a series of events and actions in this field, notably within the framework of the strategic partnership linking the High Commission for Planning (HCP) and the UNFPA. In this context, this article proposes a multidimensional conceptual model for the governance of prosocial AI, within the context of Morocco as a Global South country best suited, as Walther states, to this type of AI centered on prosociality, or what she calls the “fourth way” (Walther, 2026).

It aims to answer the central question: how can a country in the Global South, such as Morocco, achieve leadership and transform its potential in Prosocial AI? To this end, it explores Morocco's foundations and assets in prosociality, which predispose it to become a pioneering country in prosocial AI in Africa. This approach has allowed us to develop a multidimensional and multilevel conceptual model which we have named the PHASE Prosocial AI Framework (Prosociality & Public and Collective Value, Human Dignity, Equity & Well-Being, Algorithmic Accountability & Participatory Governance, Social Cohesion & Trust, and Ethical & Environmental Sustainability). It has also enabled us to propose a Prosocial AI Index, a multidimensional composite measurement index, as well as a national charter for prosocial AI, designed as a normative instrument to guide and direct the development of prosocial AI within the context of a strengthened welfare state. Finally, we conclude by identifying the prerequisites and enabling conditions that underpin what we conceptualize as the foundations of prosociality in a developing country, thus positioning Morocco as a potential regional and African hub for prosocial AI. This conceptual model is also replicable for emerging economies seeking to align AI innovation with ethical and societal objectives, serving both people and the planet.

2. Theoretical Framework and Research Methodology

2.1 Theoretical Framework

Within our holistic, systemic, and multidimensional approach to Prosocial AI, and to ground our conceptual model on solid theoretical foundations, we have drawn upon several complementary theories that enrich our multidimensional and multilevel governance architecture for Prosocial AI. This Framework is operationalized through the Moroccan Prosocial AI Index as a measurement tool and the Moroccan National Charter for Prosocial AI as a normative instrument that defines the principles and guidelines for the conception, development, deployment, governance, use and evaluation of Prosocial AI. To this end, we have mobilized the Public Value Theory (Moore, 1995), the Capability Approach (Sen, 1999; Nussbaum, 2011), and Multi-level Governance Theory (Hooghe & Marks, 2001). In addition, we also relied on the OECD principles (2019), as well as the UNESCO recommendation (2021), the UN report on global AI governance (2024), the European Union's AI ACT (2024) and the NIST AI Risk Management Framework (2023).

2.2 Research Methodology

This research adopts a conceptual model development research design (Jaakkola, 2020) for the construction of our new PHASE governance model for Prosocial AI. This conceptualization process is informed and enriched by qualitative insights through a triangulated qualitative approach combining document analysis, semi-structured interviews with approximately fifteen academic, public, and private sector experts, and a comparative analysis

of international best practices in Prosocial AI. These various sources were used as qualitative material for conceptual guidance, triangulation, awareness-raising, relevance and grounding of the proposed model. Their purpose was not to produce empirical or statistical validation, but rather to strengthen the theoretical robustness, conceptual relevance of our multilevel governance model for AI. This allowed us to operationalize it within a measurement system and subsequently develop the elements of Morocco's first National Charter for Prosocial AI. This methodological approach also enabled us to identify the pillars for developing a national Prosocial AI ecosystem, as well as its conditions and prerequisites, so that Morocco, as a developing country, can transform its Prosocial AI potential into sustainable leadership.

The following sections present the main results of our conceptual modeling process adapted to the Moroccan context. The foundations, challenges, PHASE governance framework, Moroccan Prosocial AI Index (MPAII), Moroccan National Charter for Prosocial AI (MNPAIC), as well as the prerequisites and pillars proposed for Morocco's emergence as a regional leader in prosocial AI, result from an iterative process of analysis and synthesis, informed by scientific literature, document analysis, international benchmarking and expert consultations.

3. Foundations, Achievements and Enabling Assets for the Emergence of Prosocial AI in Morocco

Our analysis of Morocco's potential for prosocial AI has shown that Morocco, as a country in the Global South, possesses a set of constitutional, institutional, strategic, and technological foundations. It also enjoys predispositions toward prosociality inherent in its historical and socio-economic development trajectory. These foundations are favorable to the emergence of prosocial AI and position it to become a pioneer in this field in Africa and the Arab world. In this context, and as illustrated in the figure below, we have identified what we have termed the "Foundations of Prosociality in the Moroccan Social State," which are based on the following elements:

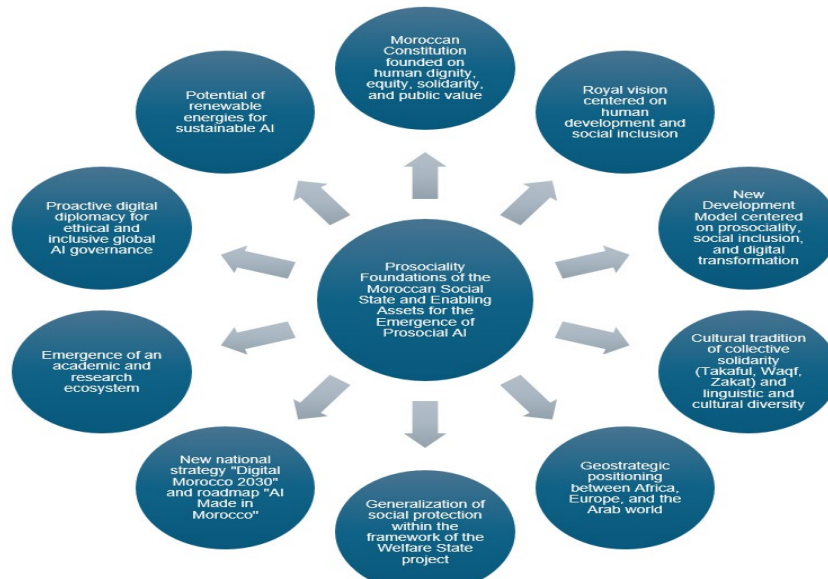


Figure 1: Foundations, Achievements and Enabling Assets for the Emergence of Prosocial AI

- The 2011 Constitution, as a normative foundation: Our analysis of the 2011 Moroccan Constitution revealed that prosociality is a central concept. Indeed, the concept of protection, social aspects, and societal considerations of human rights and sustainable human development are central to the 2011 Moroccan Constitution.
- The Royal Vision and Guidelines of the New Development Model, as strategic orientations, place human rights and the création of public value at the heart of the country's digital transformation and inclusive, integrated, and sustainable development policies.
- Institutional foundations characterized by the creation of several constitutional institutions and numerous public establishments overseeing participatory governance, human rights protection, social inclusion, social dialogue, and the evaluation of the societal impact of public policies such as the Economic, Social and Environmental Council, the Ombudsman of the Kingdom, the Social

Protection Agency, and other socially oriented institutions. This institutional ecosystem will facilitate the implementation of Prosocial AI in Morocco and remains favorable to responsible, ethical, human-centered, and prosocially oriented AI.

- Operational foundations characterized by the launch of the major project of the Social State in Morocco and the generalization of social protection through various institutional and operational mechanisms for the generalization of social protection, the reduction of inequalities, and equitable access to essential public services and social cohesion (Unified Social Register).
- Cultural foundations embodied by cultural traditions, the intrinsic characteristics and civilization of Morocco as a secular country perpetuating the communal values of solidarity, the culmination of collective action, mutual aid, cooperation, and religious practices and values of tolerance, respect for others, solidarity, and the entire heritage of prosociality (Takaful, Zakat, Waqf).

These foundations provide a strong rationale for a deliberate and proactive prosocial approach, making the Moroccan institutional ecosystem favorable to ethical, responsible, and sustainable AI and Prosocial AI. Furthermore, other achievements and accomplishments establishing a suitable and favorable environment for the adoption and development of prosocial AI should be highlighted. Indeed, Morocco has many assets that can position it as an African AI hub and strengthen its leadership and digital influence (Lafram, Bahji, 2024). Thus, several initiatives and actions have been launched by Morocco, reinforcing its strengths and capabilities. These include: the national strategy "Morocco Digital 2030" and its roadmap "AI Made in Morocco"; the national network of Jazari Institutes, comprising 12 centers of excellence covering the Kingdom's 12 regions; the Digital Hub for Sustainable Development (D4SD), in partnership with UNDP; the "AI Movement – The International AI Center of Morocco" ; various initiatives promoting ethical and responsible AI; partnerships with UNESCO and other international organizations; and many other actions demonstrating a proactive digital diplomacy for ethical and inclusive global AI governance.

4. Challenges and Obstacles Limiting the Potential of Prosocial AI in Morocco

Our research also identified the main challenges and obstacles limiting the potential of the Prosocial AI ecosystem in Morocco. As shown in Figure 2 below, these challenges and obstacles can be summarized as follows:

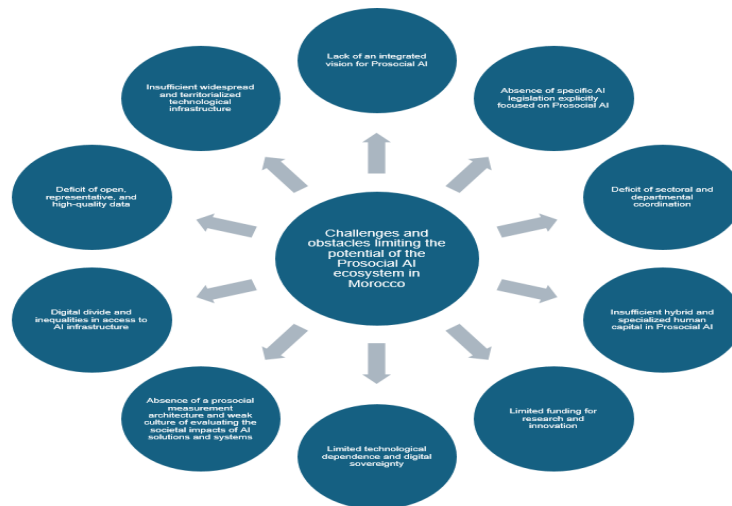


Figure 2: Main Challenges and obstacles limiting the potential of the Prosocial AI ecosystem in Morocco

The first challenge is the lack of an integrated vision for Prosocial AI. This is exacerbated by the absence of specific AI legislation governing the design, development, deployment, advancement, and use of AI. The lack of an operational framework law is hindering investment, innovation, and the potential of AI-based technology and creativity. The second challenge is the lack of sectoral coordination between different departments, which still operate in silos without real convergence in terms of governance and system management. AI. The third challenge is the lack of specialized human capital in Prosocial AI. There is a shortage of hybrid profiles, of Prosocial AI practitioners capable of combining technical expertise, algorithmic skills, and human skills with social impact. The fourth challenge is the absence of a prosocial measurement architecture. There is no composite index measuring the prosocial dimension of AI solutions and systems. It should be noted in this

context that the High Commission for Planning (HCP) has begun the process of developing a Prosocial AI index in a pilot phase in partnership with UNFPA. This innovative and strategic initiative fills this gap.

Another challenge highlighted by our research is the reliability of R&D and technological dependence. Indeed, as a country in the Global South, Morocco remains dependent on Foreign LLM models limit the country's cognitive and digital sovereignty. Another obstacle that must also be highlighted is the digital divide and inequalities in access to AI infrastructure, which deprives entire segments of the population of access to inclusive prosocial AI. Added to this is the deficit of open, representative, and high-quality data, which will consequently impact the performance and equity of prosocial AI solutions and systems with a human-centered focus on the needs and expectations of local populations.

5. The PHASE Framework: A Multilevel Governance Model Prosocial AI

5.1 Presentation of PHASE Prosocial AI Framework

Based on our diagnosis of the foundations and strengths of prosociality in the Moroccan context, and by analyzing the challenges and obstacles limiting the potential of Prosocial AI in Morocco, we have designed a holistic, multilevel, and multidimensional model of Prosocial AI governance, which we have named the PHASE Prosocial AI Framework. This conceptual model is based on the following five dimensions of prosociality, which synthesize our approach to the prosociality of an AI system, solution, or ecosystem: Prosociality & Public and Collective Value, Human Dignity, Equity & Well-Being, Algorithmic Accountability & Participatory Governance, Social Cohesion & Trust, and Ethical & Environmental Sustainability.

This is an integrative conceptual model that synthesizes existing approaches to AI governance by incorporating the various facets and multidimensional principles of prosociality within a unified, multilevel governance architecture. It thus enriches the conceptualisation of Prosocial AI by articulating governance, measurement, and institutionalisation within a single integrated framework. It further contributes to structuring a conceptual definition of Prosocial AI, along with its core dimensions and the criteria guiding its conception, development, deployment and evaluation.

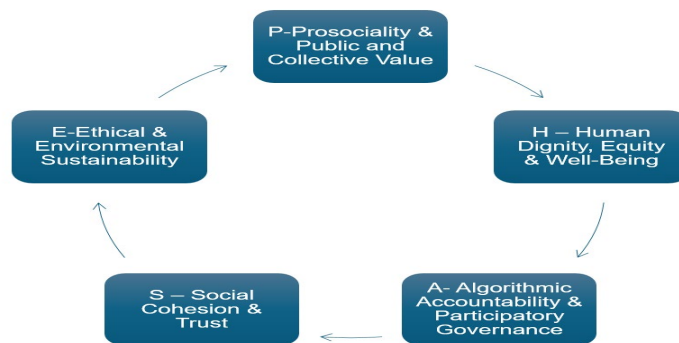


Figure 3: PHASE Prosocial AI Framework Overview

1. **Prosociality & Public and Collective Value:** This dimension focuses on the explicit prosocial purpose. It questions the real public and collective value that AI brings to the community and collective well-being.
2. **Human Dignity, Equity & Well-Being:** This second dimension focuses on the role that AI can play in the emancipation and development of fundamental human capacities. It examines its effective contribution to human dignity, equity and individual and collective well-being, as well as to the structural reduction of inequalities.
3. **Algorithmic Accountability & Participatory Governance:** This third dimension of our PHASE framework involves examining and reporting on the effects of AI. It calls for algorithmic accountability and participatory governance to guard against any missteps, biases, or exclusions.
4. **Social cohesion and trust:** this dimension examines the impact of AI on social capital, digital trust, and social cohesion. It involves assessing the extent to which AI can strengthen or weaken bonds of solidarity, reciprocity, and trust.
5. **Ethical and future-oriented sustainability:** this fifth dimension of algorithmic prosociality imposes criteria for ecological and ethical sustainability and intergenerational responsibility. It requires AI that respects the planet and future generations, in alignment with the SDGs.

5.2 Dimensions and Sub-dimensions of the PHASE Prosocial AI Framework

As shown in Figure 4 below, the components or dimensions constitute the building blocks and define the overall structure of our PHASE Prosocial AI Framework model. Each dimension has been structured into sub-dimensions, specifying the analytical elements.

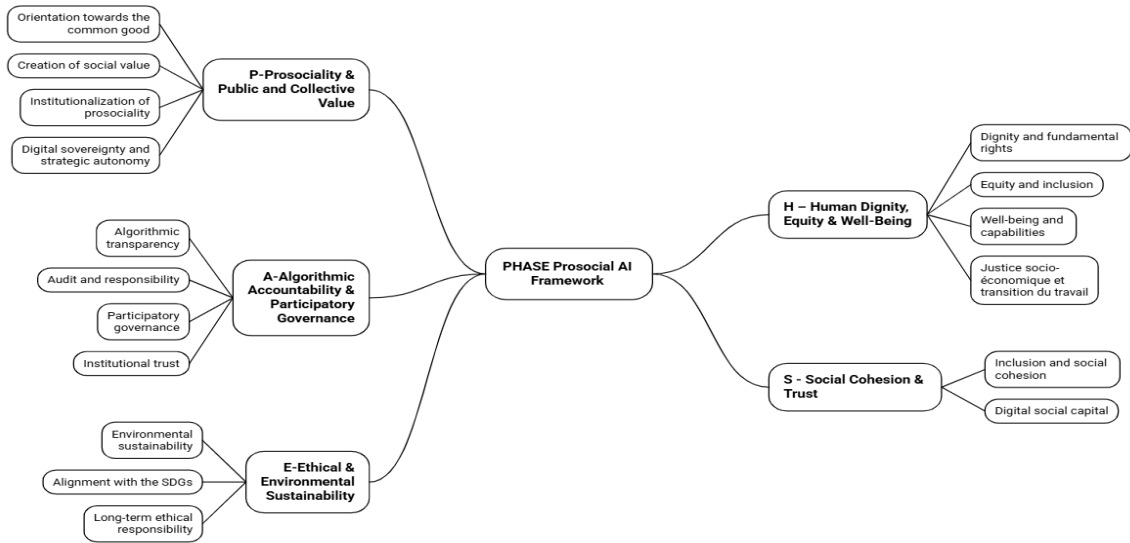


Figure 4: PHASE Prosocial AI Framework Dimensions and Sub-dimensions

6. Moroccan Prosocial AI Index (MPAII)

To operationalize the PHASE Framework, we proposed the Moroccan Prosocial AI Index. This is a measurement and evaluation tool that allows us to assess the prosocial dimensions of an AI system, solution, or ecosystem. Our Moroccan Prosocial AI Index is based on a different holistic and systemic vision, linked to public policy. It uses a multidimensional approach: meta level (Global level, planet), macro level (the state), meso level (organizations, regions, ministries, sectors), meso level (plateformes, systems, ecosystems), and nano level (human-computer interaction). This proposed index serves as a tool for comparative organizational and territorial analysis, informed decision-making, monitoring of public policies, and international evaluation of best practices. The adopted rating system was designed to be simple and pragmatic to facilitate the work of evaluators in the field: 0 (absence), 1 (initial), 2 (emerging), 3 (advanced). The overall score of our index corresponds to the aggregated average of the five dimensions of the PHASE framework.

7. Moroccan National Prosocial AI Charter (MNPAIC)

Our findings suggest that Prosocial AI necessitates a normative framework capable of establishing safeguards and ethically and sustainably guiding various institutional, regulatory, technological, and societal practices. This is why we have proposed the Moroccan National Charter for Prosocial AI (MNPAIC) as an institutionalization tool and an ethical and normative guidelines to guide the design, development, deployment, evaluation, and management of AI systems, in accordance with the principles of social justice, human dignity, social cohesion, collective well-being, transparency, citizen participation, and sustainability. This Charter, derived from our conceptual model constitutes a soft law mechanism and a tool for strengthening trust, legitimacy, harmonization of governance, and regional soft power. It guarantees ethical and sustainable coherence between technological and algorithmic progress and innovation, public and societal value, and inclusive, integrated, and sustainable human development. The guiding principles of this charter align with the five dimensions and sub-dimensions mentioned above in our conceptual model PHASE

The following figure schematically illustrates our conceptual approach as an integrated and unified multilevel governance architecture:

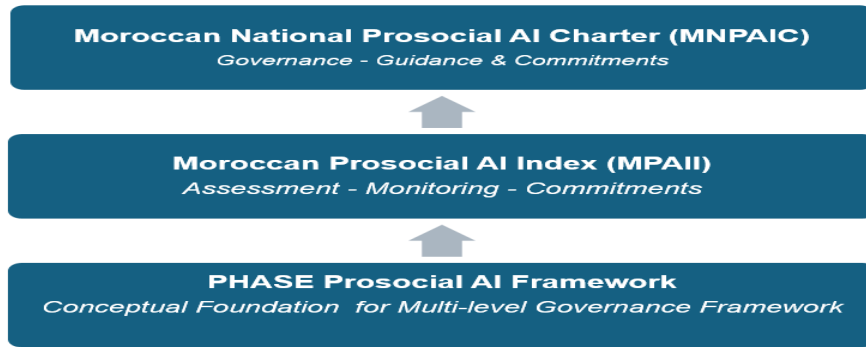


Figure 5: Synthetic conceptual approach: PHASE, Index, Charter

8. Prerequisites and Key Pillars of Morocco's Prosocial AI Leadership

8.1 Prerequisites for Morocco's Emergence as a Regional Leader in Prosocial AI

Our research identifies the prerequisites for Morocco to emerge as a leader in Prosocial AI that we have conceptualized as the main foundations of prosociality that ensure the robustness and sustainability of the Prosocial AI.

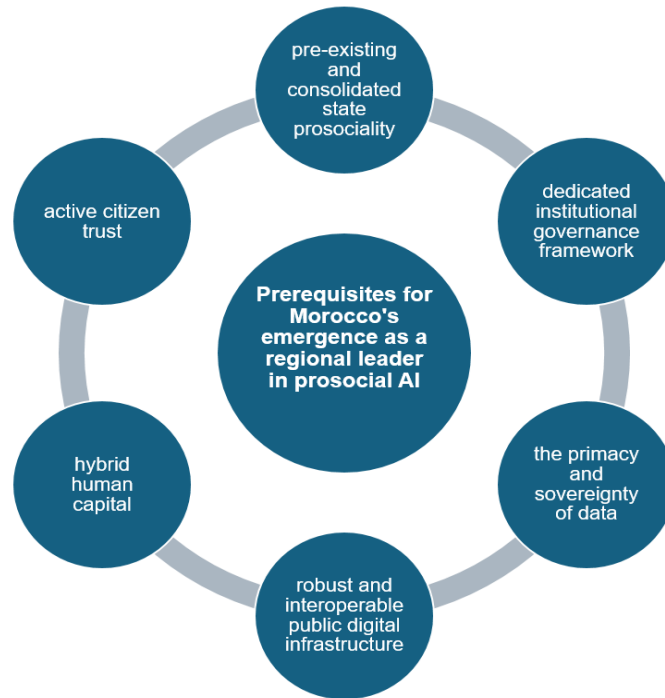


Figure 6: Prerequisites for Morocco's emergence as a regional leader in Prosocial AI

These conditions can be summarized as follows: a pre-existing and consolidated state prosocialism, a dedicated institutional governance framework with a set of measurable performance indicators capable of developing an AI ecosystem, data primacy and sovereignty, a robust and interoperable public digital infrastructure, and a hybrid human capital that masters both technical and social aspects, including ethics. Added to this is the active trust of citizens, supported by a clear strategic vision and structured, sustainable funding.

8.2 Key Pillars of Morocco's Prosocial AI Leadership

Our research highlights that Morocco possesses fundamental pre-existing potential in prosociality and considerable strengths in AI (Lafram & Bahji, 2024). This positioning predisposes it to become a pioneering country in this field. The pillars proposed within this framework consist of implementing the foundations of an effective, robust, and sustainable Moroccan ecosystem, ensuring African and continental leadership in prosocial AI.

The recommended pillars are mainly a national doctrine on prosocial AI, based on a strategic and civilizational vision, making AI an instrument for comprehensive, integrated, and sustainable human development, multilevel governance and institutionalization, digital sovereignty and strategic autonomy ensuring multidimensional dependence, human capital and collective intelligence. Another pillar is also proposed, focusing on inclusive and sustainable human development. This involves directly and intentionally linking AI to the real needs and expectations of populations, and to the imperatives of integrated, inclusive, and sustainable human development.

Like the other pillars, this pillar of trust and social cohesion also aligns with our multilevel and integrative conceptual model. It involves bringing together the essential ingredients to guarantee trust, acceptability, legitimacy, and citizen participation. Furthermore, we consider international leadership and normative digital diplomacy to be a fundamental pillar of regional influence and digital and normative soft power, aiming to make Morocco a regional hub, a normative benchmark, and an emerging laboratory for countries of the Global South. Thus, Morocco can leverage its progress in engaging with digital diplomacy by strengthening South-South cooperation in the field of pro-social AI and forging new regional partnerships.

Based on the aforementioned elements, we have synthesized the main components and levers for developing an integrated Prosocial AI ecosystem that Morocco, as a country in the Global South, can activate in order to transform its potential in Prosocial AI into true regional leadership.

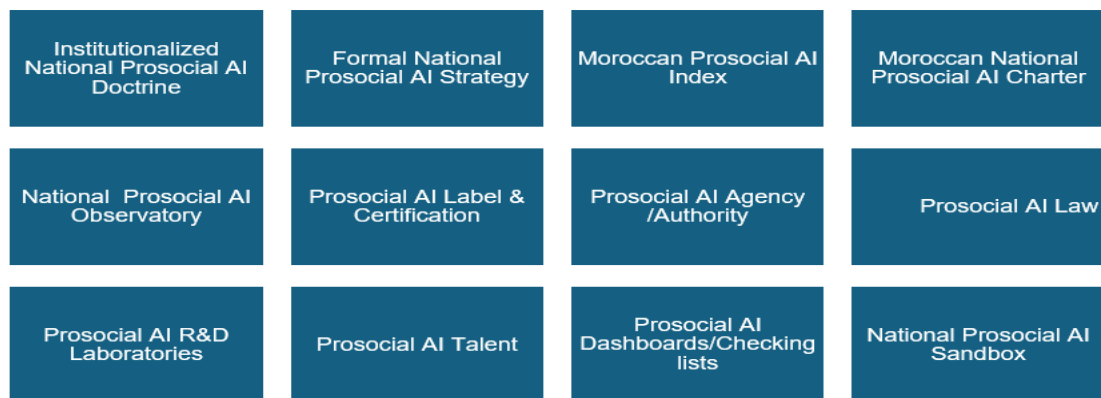


Figure 7: Components and levers of development of an integrated and convergent Moroccan ecosystem of Prosocial AI

9. Conclusion

The findings of this conceptual research highlight the potential of Moroccan prosociality, which fosters the emergence of a prosocial AI ecosystem and supports its long-term development. The conceptual and strategic analysis conducted in this study identifies several constitutional, institutional, structural, strategic and cultural foundations and assets and various pioneering initiatives and projects launched by Morocco in the field of digital transition and AI (Morocco Digital 2030, AI Made in Morocco, D4SD, AI Movement, Jazari Institutes, etc.). However, several challenges limiting the potential of the prosocial AI ecosystem in Morocco have also been highlighted, arguing in favor of an innovative model adapted to the context of Morocco as a country of the Global South. Taking into account the results of this diagnostic of Moroccan algorithmic prosociality, we realized the need to develop an alternative multilevel conceptual model adapted to the Moroccan context, embodying what we call the “Moroccan Prosocial Way.” This model is part of our strategic thinking around the concept of the augmented Social State, which does not conceive of AI as a separate, compartmentalized policy but as an integrated national ecosystem for comprehensive, sustainable, and enhanced human development.

To respond to our central question: how can a country in the Global South, such as Morocco, achieve leadership and transform its potential in Prosocial AI? this research proposes the PHASE Prosocial AI Framework, which is based on the five dimensions of prosociality. These dimensions synthesize our approach to the prosociality of an AI system, solution, or ecosystem: Prosociality & Public and Collective Value, Human Dignity, Equity & Well-Being, Algorithmic Accountability & Participatory Governance, Social Cohesion & Trust, and Ethical & Environmental Sustainability.

To operationalize this integrative, multidimensional model, we proposed the Moroccan Prosocial AI Index, a pragmatic measurement tool based on the five dimensions of prosociality that comprise our PHASE model. Furthermore, we also proposed a normative framework: the Moroccan National Charter for Prosocial AI, to institutionalize our PHASE model. This charter defines the guiding principles and objectives, particularly those of prosociality, public value, social cohesion, equity, ethics, responsibility, and sustainability.

Thus, this article provides three major theoretical and conceptual contributions (PHASE Prosocial AI Framework, Index, Charter) that explore an emerging scientific field by proposing an exportable and replicable conceptual model in other contexts of the Global South. Our research also provides practical contributions for decision-makers, practitioners, and international organizations interested in the prosocial dimensions of AI solutions, systems, and ecosystems, and who aim to experiment with or launch pilot projects for the measurement, auditing, and development of Prosocial AI.

As with any qualitative and exploratory study, this research has certain limitations, mainly related to the novelty of these measurement tools, which require institutional adoption and commitment from all stakeholders. It opens up promising new perspectives for research in emerging fields such as the development and deployment of measurement, audit and evaluation instruments for prosocial AI and experimental implementation of these tools in various priority sectors and national and regional contexts, as well as comparative and longitudinal studies on prosocial AI in Africa and the MENA region.

Ethics Declaration: No ethical approval was required for this paper.

AI Declaration: No AI tools were used in the preparation and writing of this paper.

References

- Cascio, J. (2020) Facing the Age of Chaos. Institute for the Future.
- European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (Artificial Intelligence Act), Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. and Vayena, E. (2018) "AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations". *Minds & Machines* Vol 28, N° 4, pp. 689–707 <https://doi.org/10.1007/s11023-018-9482-5>
- Gupta, P., Zhong, Q., Yakura, H., Eisenmann, T., & Rahwan, I. (2026) "The Role of Social Learning and Collective Norm Formation in Fostering Cooperation in LLM Multi-Agent Systems". In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus. <https://doi.org/10.48550/arXiv.2510.14401>
- High-Level Expert Group on Artificial Intelligence (2019) Ethics guidelines for trustworthy AI. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Hooghe, L., & Marks, G. (2001) "Types of multi-level governance". *European Integration online Papers (EIoP)*, Vol. 5, No. 11. <https://eiop.or.at/eiop/pdf/2001-011.pdf>
- International Telecommunication Union. (2023) AI for good global summit initiative. ITU. <https://aiforgood.itu.int/>
- Jaakkola, E. (2020) "Designing conceptual articles: four approaches". *AMS Review*, Vol. 10, No. 1-2, pp. 18-26 <https://doi.org/10.1007/s13162-020-00161-0>
- Jobin, A., Ienca, M. & Vayena, E. (2019) "The global landscape of AI ethics guidelines". *Nature Machine Intelligence* Vol.1, No 9., pp.389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Lafram, A., & Bahji, S. E. (2024) "Artificial intelligence in Morocco: Towards holistic, responsible and ethical national AI strategy for Moroccan competitiveness and strategic intelligence". In *Proceedings of the 20th European Conference on Management Leadership and Governance (ECMLG 2024)* 14 -15 November 2024, Universidade Lusófona Lisbon, Portugal. pp. 255–263.
- Moore, G.E. (1995) *Principia Ethica*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511818917>
- Nussbaum, M. C. (2011) *Creating capabilities: The human development approach*. Harvard University Press. <https://www.hup.harvard.edu/books/9780674050549>
- OECD. (2024) OECD AI principles overview. OECD AI Policy Observatory. <https://oecd.ai/en/ai-principles>
- OECD (2019) *Artificial Intelligence in Society*, OECD Publishing, Paris, <https://doi.org/10.1787/eedfee77-en>.
- European Parliament & Council of the European Union. (2024). Artificial Intelligence Act (Regulation (EU) 2024/1689). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- OECD. (2019) OECD principles on artificial intelligence. OECD Publishing. <https://doi.org/10.1787/eedfee77-en> <https://oecd.ai/en/ai-principles>
- National Institute of Standards and Technology. (2023) Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). <https://doi.org/10.6028/NIST.AI.100-1> <https://www.nist.gov/itl/ai-risk-management-framework>

- Santos, F. P. (2024) Prosocial dynamics in multiagent systems. *AI Magazine*, 45, 131-138.
<https://doi.org/10.1002/aaai.12143>
- Sen, A. (1999) *Development as freedom*. Oxford University Press .<https://global.oup.com/academic/product/development-as-freedom-9780198297581>
- Shi, H., Yin, B.(M.), Teng, L., & Ma, B. (2025) Empathetic AI Encounters : Pathways to Prosocial Behavior. *Journal of Service Research*. Advance online publication.
- Shneiderman, B. (2022) *Human-Centered AI*. Oxford University Press.
- Stanford Institute for Human-Centered Artificial Intelligence. (2026). *AI Index Report 2026*. Stanford University.
https://hai.stanford.edu/assets/files/ai_index_report_2026.pdf
- United Nations. (2024) *Governing AI for humanity: Final report of the High-Level Advisory Body on Artificial Intelligence*. United Nations.<https://www.un.org/en/ai-advisory-body>
- UNESCO (2021) *Recommendation on the ethics of artificial intelligence*. UNESCO.<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- Walther, C. C. (2026) "The Fourth Path: Middle-Income Countries and Prosocial AI". *Centre for International Governance Innovation (CIGI)*, Policy Brief No. 230.
- Walther, C.C. (2024) "ProSocial artificial intelligence as a catalyst for holistic health: a multidimensional approach". *BMC Global Public Health* 2, Article 76 <https://doi.org/10.1186/s44263-024-00111-z>
- Willis, R., Zhao, J., Du, Y., & Leibo, J.Z. (2026) "Evaluating collective behaviour of hundreds of LLM agents". *ArXiv*.
<https://arxiv.org/abs/2602.16662>
- Wang, Z., Song, R., Shen, C., Yin, S ; Song, Z. Battu, B., Lei Shi, L., Jia, D.
- Zhou, Y., Wang, H., Ai, Q., Wu, Z., & Liu, Y. (2026) "Investigating Prosocial Behavior Theory in LLM Agents Under Policy-Induced Inequities". *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(3), 2254–2262.
<https://doi.org/10.1609/aaai.v40i3.37209>