

The AIWS Model of Society for the AI Age: Six Foundations for Human Dignity, Legitimate Authority, and a Human-AI Co-Creation Civilization

Nguyen Anh Tuan

Co-Founder, Co-Chair and CEO, Boston Global Forum

Chief Architect, AIWS

*Alumnus of the Faculty of Mathematics and former staff member of the Information
Technology Center, Dalat University*

*A special paper for the 50th Anniversary of Dalat University · October 27, 1976 –
October 27, 2026*

Boston · August 14, 2026

ABSTRACT

Earlier technological revolutions extended human physical power, productive capacity, communication, and access to knowledge. AI goes further: for the first time, human beings have created systems capable of participating in the very activities once regarded as distinctive of human intelligence — learning, reasoning, creating, predicting, deciding, and increasingly acting with autonomy.

This paper argues that what is underway is not merely a technological revolution but a transformation in the structure of civilization. When intelligence is no longer an exclusively human capacity, humanity must redefine the relationship between intelligence, human authority, legitimacy, and responsibility.

Drawing on the formation of AIWS in 2017, the AIWS 7-Layer Model of 2018, and the development toward AI Wisdom Society in 2026, the paper proposes Six Foundations of the AIWS Model of Society for the AI Age:

**Human Dignity → Human Primacy → Human-in-Command → Trust
Infrastructure → Constitutional Governance → Human Flourishing.**

These Six Foundations do not replace the 7-Layer Model. If the Seven Layers organize the path from principle to standards, law, public institutions, and applications, the Six Foundations identify the deeper stratum — dignity, legitimacy, authority, verifiability, institutional order, and purpose — upon which all governance and deployment of AI must be built.

The cultural and human dimension of this architecture is the AIWS Lumina Culture, with its four values: Love, Creativity, Nobility, and Wisdom.

The paper further advances five theoretical propositions: the Intelligence Separation; the Principle of Intelligence and Authority; Verifiable Trust; Human-AI Co-Creation Civilization; and the AIWS Civilization Matrix. Together they widen the question from the governance of AI to something larger: how must a society be organized when intelligence is no longer uniquely human?

For the author, the methodological foundations that led to AIWS originate in part from years of studying Mathematics, Probability, and Mathematical Statistics at Dalat University, and from working at the University's Information Technology Center. Mathematics teaches the search for structure within complexity; probability teaches reasoning under uncertainty; statistics teaches the distinction between data and evidence; and computing opened an early recognition that human intelligence could be extended by technology.

Nearly half a century later, those questions return at the scale of civilization itself.

Keywords: AIWS; artificial intelligence; AIWS 7-Layer Model; Six Foundations; human dignity; human primacy; human-in-command; trust infrastructure; constitutional governance; human flourishing; AIWS Lumina Culture; Human-AI co-creation; AI Wisdom Society; AIWS Civilization Matrix.

I. VALUES FROM DALAT UNIVERSITY

On the occasion of the 50th anniversary of Dalat University, I return not only with the memories of a former student, but with the values and the habits of mind that have accompanied me through many stages of life.

I was a student in the Faculty of Mathematics from 1980 to 1984, and a staff member of the University's Information Technology Center. I studied Probability, Mathematical Statistics, and Optimization, and gained access to

computers in a period when the power of information technology was still far from what we witness today.

Those subjects gave me more than knowledge. They gave me a method of thinking, a way of approaching problems, and the capacity to absorb what is new. The applied mathematical disciplines in particular — probability, mathematical statistics, and optimization — later became an essential foundation for engaging with the emerging ideas and advances of AI when I began working within the academic and innovation environment of Boston and Harvard from 2011.

During my years at Dalat University, the country faced severe economic hardship and resources for study were scarce. Yet within those circumstances I lived in an unusual environment. Dalat University — with its beautiful campus, an academic and cultural tradition inherited from the Institute of Dalat University, and a strong environment for study and culture built by Rector Trần Thanh Minh and the teachers of the University — gave me both professional knowledge and a feeling for the arts and culture of Vietnam and of the world, and from that, awakened dreams and aspirations. The city of Da Lat, with its beauty, its cultural life, and the fine character of its people, also helped form in me a humanistic spirit toward human beings and society.

If mathematics gave me the capacity for logical and systematic thought, Da Lat helped shape my humanistic imagination.

One taught me how to solve problems. The other asked for whom, and toward what kind of society, those problems must be solved. Looking back nearly half a century later, I recognize that both currents — mathematical reasoning and humanistic values — entered the work of building AIWS.

On this 50th anniversary of Dalat University, I offer this paper as an expression of gratitude to the University, which gave me the first foundations not only for understanding technology, but for thinking about the responsibility of intelligence toward human beings and society.

II. FROM TURING TO THE QUESTION OF SOCIETY

In 1950, Alan Turing posed one of the questions that shaped the history of modern computer science: could machines exhibit behavior we would call intelligent? *Computing Machinery and Intelligence* moved the problem from an

abstract definition of thought toward an approach that could be examined through observable behavior.

Norbert Wiener, working from the tradition of cybernetics, placed the emphasis instead on control, communication, feedback, and the relationship between human beings and machine systems of increasing self-regulating capacity. Wiener's legacy reminds us that the development of machine intelligence was, from very early on, simultaneously a question of human control and human responsibility.

Decades later, Stuart Russell reframed the problem as one of control and human-compatible AI: how to build powerful AI systems that remain aligned with human objectives and interests.

Ben Shneiderman developed the human-centered approach to AI, emphasizing the combination of AI capability with human-oriented thinking, and the closing of the gap between ethical principles and practical mechanisms, in order to produce systems that are reliable, safe, and trustworthy.

At the level of policy, UNESCO placed dignity and human rights at the center of its Recommendation on the Ethics of Artificial Intelligence. The OECD established principles for trustworthy AI; NIST developed the AI Risk Management Framework; the European Union built a legal regime based on levels of risk; and the United Nations, through *Governing AI for Humanity* and the Global Digital Compact, advanced global AI governance and digital cooperation grounded in human beings and human rights.

These works and institutions form an essential foundation for the AI Age.

But the development of AI now compels a further question.

Not only: *Can machines think?*

Not only: *Can we control intelligent machines?*

Not only: *Can AI be safe, ethical, and trustworthy?*

But:

How must human society be organized when intelligence itself is no longer uniquely human?

This is the terrain toward which the AIWS Model of Society is directed.

III. PROPOSITION ONE: THE INTELLIGENCE SEPARATION

I call the fundamental turning point of the AI Age the Intelligence Separation.

Through most of the history of civilization, intelligence has been bound to human beings. Human beings think. Human beings understand. Human beings create. Human beings judge. Human beings act. And human beings bear responsibility.

Humanity has of course always used tools to extend its capacities. But traditional tools did not themselves become sources of knowledge, prediction, decision, and action at scale.

AI is changing this structure. Intelligence can increasingly be externalized from human beings, replicated, distributed, scaled, and granted rising degrees of autonomy.

A single AI system may serve millions of people simultaneously. An AI agent may execute chains of action that no human approves step by step. A network of AI systems may shape markets, media, education, public administration, scientific research, or social infrastructure.

This does not necessarily mean that AI becomes "human" or conscious. What matters more, socially, is this:

Capacities historically bound to human intelligence can now increasingly operate outside human beings.

That is the Intelligence Separation.

And from this separation arises a serious danger: a society may unconsciously assume that authority should follow intelligence. Whichever system predicts better should decide. Whichever system knows more should be empowered. Whichever system optimizes better should define the objective.

This is where AIWS draws a line.

Intelligence can be separated from human beings. Legitimate authority cannot.

IV. PROPOSITION TWO: THE PRINCIPLE OF INTELLIGENCE AND AUTHORITY

From the Intelligence Separation, I propose a second principle:

As artificial intelligence grows in capability and autonomy, legitimate human authority must not diminish. The more powerful autonomous intelligence becomes, the stronger the human institutions of command, verification, and accountability must be.

The status of this proposition requires clarification. It is not an empirical law describing what happens — the historical record suggests the opposite tendency is common. It is a normative principle in the sense the constitutional tradition uses when it speaks of the rule of law: a condition that institutions must protect, not a trend that realizes itself.

There is a natural inclination in the history of technology: when a system performs better than human beings, human beings tend to grant it more authority. In many domains this is reasonable. But there is a limit.

Capability is not legitimacy.

A model may predict better than a person, but that does not give it the right to determine what is just. An AI may identify an optimal economic solution, but it holds no authority to determine what a person's dignity is worth. A system may decide faster, but speed does not constitute the right to rule.

Therefore:

Superior intelligence must never be equated with superior authority.

This is a foundational distinction for the AI Age. Intelligence belongs to the category of capability. Authority belongs to the category of legitimacy. Wisdom belongs to the category of purpose, consequence, and responsibility.

A society fails when it allows those three categories to dissolve into one.

V. FROM THE 7-LAYER MODEL TO SIX FOUNDATIONS FOR THE AI AGE

AIWS has developed not by abandoning its original ideas, but by deepening them as artificial intelligence itself has evolved.

In 2018, less than a year after AIWS was founded, the Boston Global Forum and the Michael Dukakis Institute introduced the AIWS 7-Layer Model in the BGF-G7 Summit Report of April 25, 2018. The seven layers run from foundational principle to practical deployment:

1. Charter and Principles · 2. Ethics Framework · 3. Standards · 4. Law and Legislation · 5. Policy, Conventions, and International Norms · 6. Public Services and Policymaking · 7. Applications and Services.

The 7-Layer Model answered the defining question of its moment: how must artificial intelligence be developed, governed, and applied so that it remains safe, humane, accountable, and beneficial to society? Its founding premise was already firmly human-oriented: human beings must remain ultimately responsible for the development and use of AI, while transparency, auditability, and assessability place AI systems under scrutiny.

That architecture retains its full value. But the transformation of AI since 2018 has raised a deeper question.

AI is no longer simply a technology introduced into existing institutions. It increasingly participates in the creation of knowledge, in reasoning, in decision-making, in creative work, and in autonomous action. The problem therefore moves beyond governing a technology, and becomes a question about the constitutional structure of a society in which intelligence is no longer uniquely human.

The question becomes: What must remain permanently human, when intelligence itself can increasingly be performed by machines?

This paper does not, therefore, replace the 7-Layer Model with another layered model. It proposes a deeper foundation beneath it:

THE SIX FOUNDATIONS OF THE AIWS MODEL OF SOCIETY FOR THE AI AGE



ETHICAL FOUNDATION · HUMAN DIGNITY



LEGITIMACY · HUMAN PRIMACY



AUTHORITY · HUMAN-IN-COMMAND



VERIFICATION · TRUST INFRASTRUCTURE



INSTITUTIONAL ORDER · CONSTITUTIONAL GOVERNANCE



CIVILIZATIONAL PURPOSE · HUMAN FLOURISHING

The distinction here is fundamental. The 7-Layer Model describes an architecture that translates principle into governance, standards, law, public institutions, and applications. The Six Foundations identify the ethical, constitutional, and civilizational conditions within which all those layers must operate.

**The Seven Layers organize the governance and deployment of AI.
The Six Foundations establish the legitimacy and purpose of the
society in which AI operates.**

VI. FOUNDATION I: HUMAN DIGNITY — THE ETHICAL FOUNDATION

The AIWS model does not begin with AI. It begins with the human person.

In the AI Age, a person can increasingly be represented as a set of data: health, finances, productivity, social relations, behavior, preferences, risk, creditworthiness, employability — and even what that person may buy, read, vote for, or do in the future.

But:

**A human being is greater than the sum of all data that can be
collected, inferred, or predicted about that person.**

A human being is not a probability distribution. Not a risk score. Not an objective function to be optimized. Not a profile for a system to exploit.

A person's worth does not derive from that person's usefulness to a system. It is intrinsic dignity. UNESCO likewise places dignity and human rights at the center of its ethical framework for AI.

From that ethical foundation, AIWS moves to a conclusion about the structure of power: if human beings possess intrinsic dignity, then no AI system can become an agent entitled to determine, at its own discretion, the value of a human being.

Human Dignity therefore establishes the foundation for Human Primacy.

VII. FOUNDATION II: HUMAN PRIMACY — THE PRINCIPLE OF LEGITIMACY

Human Primacy is not the claim that human beings will always be more intelligent than AI. That would be an increasingly difficult position to defend.

The principle says something different:

Intelligence does not generate legitimacy.

However high the capability AI attains, human beings remain the source of purpose, values, rights, responsibilities, and the legitimate limits of power.

AI may propose; human beings judge. AI may forecast; human beings determine what is worth pursuing. AI may optimize; human beings define the purpose of optimization. AI may generate options; human beings bear responsibility for the decision.

The Boston Declaration of July 4, 2026 placed the principle of the primacy of the human person at the center of the vision of the Boston Global Forum and AIWS, and opened the way toward the Constitution for Humanity in the Age of Artificial Intelligence.

This is therefore a principle concerning legitimate authority, not computational superiority.

VIII. FOUNDATION III: HUMAN-IN-COMMAND — THE ARCHITECTURE OF AUTHORITY

An ethical principle is insufficient unless it becomes an executable capacity.

A person may remain formally "in the loop" while exercising no real control. If an AI produces a complex decision in milliseconds and a person merely presses an approval button, that oversight may be no more than ceremonial. If an employee does not understand the system, holds no data with which to contest it, and lacks the authority to halt it, that employee is not in command.

AIWS therefore adopts a stronger concept: Human-in-Command.

The principle requires that for critical systems, there must be clearly identified human beings or human institutions with the authority to: authorize; oversee; interrogate; contest; intervene; suspend; override; terminate — and, ultimately, to be held accountable.

This is not merely a technical feature. It is an architecture of authority.

IX. FOUNDATION IV: TRUST INFRASTRUCTURE — THE ARCHITECTURE OF VERIFICATION

But a further problem arises.

Who confirms that human command genuinely exists? Who examines an AI laboratory when that laboratory declares its own systems safe? Who verifies that a corporation uses AI responsibly? Who examines the government?

And: *Who verifies the verifier?*

This is among the most important institutional problems of the AI Age.

The NIST AI Risk Management Framework helped turn trustworthiness into a set of characteristics and processes that can be managed across the AI lifecycle, through four functions: govern, map, measure, and manage.

AIWS develops this question into the problem of a social infrastructure for trust. In 2026, the AIWS Trust Architecture was presented with three connected components: Trust Standards, Trust Infrastructure, and the Trusted Order. The broader aim is to form a continuous chain:

**Identification → Assessment → Verification → Certification →
Monitoring → Contestation → Remedy → Accountability.**

This leads to the paper's third proposition:

PROPOSITION THREE — VERIFIABLE TRUST: Trust in the AI Age must not be demanded by authority or manufactured by reputation. It must be earned through evidence, sustained through independent verification, and remain open to legitimate contestation.

Trust cannot remain a matter of belief. It must become an institutional capacity. From this follows a principle:

The greater the power, the greater the obligation to be verifiable.

The more capable an AI system, the higher the verification requirements it must meet. The more influential a corporation, the greater its obligations of transparency and accountability. The more a government uses AI in consequential decisions, the more it must provide independent verification and avenues of contestation.

Who verifies the verifier?

This question is often raised as an infinite regress, and therefore as a reason to give up. That conclusion is mistaken. Human society has solved precisely this problem more than once — in financial audit, in scientific peer review, in independent judiciaries. The solution is not to find a final verifier who requires no verification, but to close the regress institutionally. AIWS proposes four conditions.

First, *plurality*. No single verifier may hold a monopoly of judgment over a critical domain. Multiple independent bodies, drawn from different countries and institutional traditions, verify the same object; disagreement among them is data, not malfunction.

Second, *methodological transparency*. A verifier must publish its methods, criteria, and data at a level sufficient for a third party to reproduce its conclusion. A conclusion that cannot be reproduced is not verification but opinion.

Third, *binding contestation*. Anyone affected must have a practical route to challenge a verification finding, and that challenge must compel a public response.

Fourth, *a path to evidence*. Every consequential claim must have a path to evidence; every consequential authority must have a path to accountability.

These four conditions do not eliminate the regress. They make it finite and institutionally tractable — much as science does not locate an absolute observer, but constructs a community capable of correcting itself.

And the verifier, too, must be subject to verification.

X. FOUNDATION V: CONSTITUTIONAL GOVERNANCE — THE ARCHITECTURE OF LEGITIMATE POWER

Verification is still not enough. A system may be verified to operate exactly as designed while the design itself is unjust.

Society therefore requires a higher layer: Constitutional Governance.

AI safety asks: *How do we prevent harmful AI behavior?* AI ethics asks: *What values must guide AI?* AI governance asks: *What rules and institutions regulate AI?*

Constitutional Governance asks a deeper question:

Who holds the legitimate authority to determine those rules? What powers must never be concentrated? What rights may never be set aside? And by what means may authority itself be challenged?

This distinction matters greatly. A highly centralized political system may still possess a sophisticated regime of AI governance. Governance is therefore not automatically synonymous with legitimacy.

Constitutional Governance requires fundamental rights; limits on power; the separation and distribution of authority; independent verification; due process; accountability; legitimate contestation; and mechanisms of correction.

The European Union's AI Act offers an important legal model based on risk classification and corresponding obligations across development and deployment. The United Nations Global Digital Compact situates digital governance within human rights and international cooperation; *Governing AI for Humanity* proposes global mechanisms to address risk and to distribute the benefits of AI more widely.

AIWS contributes to this dialogue through the proposal of a Constitution for Humanity in the Age of Artificial Intelligence.

This is not a proposal for world government. Nor does it seek to replace national constitutions. Its purpose is to identify the constitutional foundations for the exercise of AI-enabled power: human dignity; human primacy; human command; rights; responsibilities; verifiability; accountability; limits on concentrated power; and the capacity to challenge authority.

XI. FOUNDATION VI: HUMAN FLOURISHING — THE PURPOSE OF THE AI AGE

The first five foundations largely answer the question of how to protect human beings and constrain power. But a civilization cannot exist merely to avoid danger. It must know what it is for.

This is why AIWS places Human Flourishing at the final foundation.

If AI raises gross domestic product while human beings lose the capacity for independent thought, is that progress? If AI raises productivity while human beings no longer feel themselves the authors of their own lives, is that success? If AI produces extraordinarily efficient societies in which people are less free, less creative, and less compassionate, is that the civilization we set out to build?

AIWS answers that the ultimate measure of the AI Age must be human flourishing.

This is also where the AIWS Lumina Culture holds a foundational role. AIWS Lumina is the cultural and human dimension of AIWS — the counterpart to its technical, standards-based, and policy architectures, not a decorative layer outside governance. If Trust Infrastructure answers the question *what can be verified*, the AIWS Lumina Culture answers the question *what kind of human beings we wish to become*.

The AIWS Lumina Culture rests on four values:

AIWS LUMINA CULTURE — Love · Creativity · Nobility · Wisdom

These four values are not slogans. Love places the human person, rather than performance, at the center of every system. Creativity keeps human beings as makers rather than mere consumers of what machines produce. Nobility is the quality of exercising power and intelligence in a manner worthy of it. Wisdom is the capacity to know what is worth doing, not merely what can be done.

A civilization may possess complete standards, laws, and verification mechanisms and remain impoverished in its inner life. The AIWS Lumina Culture is the assurance that the AI Age will not merely be well governed, but worth living in.

This leads to a principle:

AI should not merely make human beings more capable. It must help create the conditions in which human beings become more fully human.

XII. PROPOSITION FOUR: HUMAN-AI CO-CREATION CIVILIZATION

If humanity seeks only to control AI, it is still reasoning defensively. AIWS points toward a higher possibility: a civilization co-created by human beings and AI.

In this civilization, AI is not only a tool. It may be an intellectual assistant, a collaborator, a scientific partner, and a creative partner. Human beings and AI may together pursue scientific discovery, design technology, create art, educate, care for health, build enterprises, govern cities, and address the complex problems of humanity.

But this is not a fully symmetrical relationship. I state the principle as:

Co-creation in capability — asymmetry in authority.

AI may participate in generating options; human beings determine purpose. AI may generate knowledge; human beings determine how that knowledge is used. AI may participate in creation; human beings bear responsibility for the decisions that shape society. AI does not become a source of political legitimacy or moral sovereignty.

A co-creation civilization therefore differs from both extremes: AI domination, in which human beings progressively surrender authority to intelligence; and AI refusal, in which human beings seek to prevent AI from becoming a deep part of civilization.

AIWS proposes a third path:

Humanity should embrace the creative power of artificial intelligence while retaining moral authorship and legitimate human authority.

XIII. THE AIWS SOCIAL COMPACT

The six foundations of AIWS may be condensed into a social compact:

Human beings possess intrinsic dignity. Therefore humanity retains primacy before artificial intelligence. That primacy must be realized through human command. That command must be independently verified through trust infrastructure. AI-enabled power must operate under constitutional governance. And the ultimate purpose of this entire architecture is human flourishing.

This is the internal logic of the AIWS Model of Society for the AI Age.

XIV. PROPOSITION FIVE: THE AIWS CIVILIZATION MATRIX

The Six Foundations and the Seven Layers are often read as two lists placed one after the other. They are not. They are two dimensions of a single architecture, and their real value appears only when they are set perpendicular to each other.

PROPOSITION FIVE — THE AIWS CIVILIZATION MATRIX: The Six Foundations and the Seven Layers constitute two complementary dimensions of the AIWS architecture. The Six Foundations define WHY and WHO HOLDS LEGITIMATE AUTHORITY. The Seven Layers define HOW and WHERE those principles become institutions and practice.

Set perpendicular, the two dimensions form a matrix of six rows and seven columns. Each cell is a specific question a society must be able to answer — and a place to test whether a regime of AI governance is genuinely complete.

AIWS

AI WISDOM SOCIETY

A BEACON FOR THE AI AGE

THE AIWS CIVILIZATION MATRIX

Six Foundations × Seven Layers

Proposition Five · The AIWS Model of Society for the AI Age

SEVEN LAYERS — HOW and WHERE

		1. Charter and Principles	2. Ethics Framework	3. Standards	4. Law and Legislation	5. Policy and International Norms	6. Public Services and Policymaking	7. Applications and Services	
SIX FOUNDATIONS — WHY and WHO HOLDS LEGITIMATE AUTHORITY	I	HUMAN DIGNITY	Declare intrinsic dignity inviolable	Bar the reduction of persons to data	Standards protecting personhood in data and models	Right not to be determined by a score	International convention on digital dignity	Public services that do not rank human worth	Product design that respects the user
	II	HUMAN PRIMACY	Human beings as the source of purpose and value	Values set by people, not by optimization	Standards bounding AI decision authority	Law naming decisions only humans may make	Shared principles on human primacy	Public policy ultimately decided by people	AI proposes, the person chooses
	III	HUMAN-IN-COMMAND	Principle that command remains human	Ethics of the authority to halt a system	Technical standards for intervention and suspension	Legal duty to provide a stop mechanism	Agreements on cross-border autonomous systems	Officials empowered to override machine decisions	Interfaces allowing users to contest
	IV	TRUST INFRASTRUCTURE	Principle that trust must be verifiable	Ethics of truthful capability claims	AIWS Trust Standards	Duty to audit and publish evidence	Trust Order and mutual recognition of verification	Independent, plural verification bodies	Public Trust Rating for each product
	V	CONSTITUTIONAL GOVERNANCE	Constitution for Humanity in the AI Age	Ethics of limits on power	Standards against concentration of technical power	Due process for AI-enabled decisions	Global mechanisms for dispute resolution	Separation of authority in digital government	Right of appeal within every service
	VI	HUMAN FLOURISHING	Civilizational purpose: the flourishing person	The four AIWS Lumina Culture values	Standards measuring real effect on human lives	Law protecting autonomy and creative rights	International cooperation in education and culture	Human-centered education, health, and culture	AIWS University and AIWS Home

Each cell is a question a society must be able to answer. Empty cells mark where a regime of AI governance is incomplete.

Nguyen Anh Tuan · Boston Global Forum · 2026

The matrix serves three purposes.

First, it turns abstract principles into a *checklist*. A country, an institution, or an enterprise can measure itself against it: how many cells have we actually answered?

Second, it *exposes gaps*. Most current AI governance frameworks fill columns three, four, and seven very well — standards, law, applications — while leaving almost the whole of the first and last rows empty: dignity and flourishing. That is precisely where those frameworks are weakest.

Third, it *prevents confusion between levels*. A problem of legitimacy cannot be resolved by a technical standard. A problem of dignity cannot be resolved by a clause in a terms-of-service agreement.

XV. WHERE DOES AIWS STAND AMONG THE SCHOOLS OF THOUGHT ON AI?

AIWS does not replace AI safety, AI ethics, human-centered AI, or AI governance. These fields address important problems and overlap substantially. The difference lies in the level of the question.

Approach	Central question	Unit of analysis
AI Safety	How do we prevent dangerous consequences from AI?	the AI system
AI Ethics	What values must guide AI?	principles and conduct
Human-Centered AI	How do we design AI to enhance human capability and control?	the human-AI system
AI Governance	What rules and institutions regulate AI?	institutions, markets, states
AIWS Model of Society	<i>How must society be constituted when intelligence is no longer uniquely human?</i>	<i>civilization</i>

AIWS should therefore not be understood simply as a framework for AI governance. The paper proposes a different designation:

AIWS is a constitutional and civilizational architecture for the AI Age.

Each of the approaches above is necessary. What AIWS adds is to situate all of them within a larger question concerning dignity, legitimacy, authority, verification, constitutional order, and the purpose of civilization.

XVI. FROM AI WORLD SOCIETY TO AI WISDOM SOCIETY

AIWS was formed in 2017 with an aspiration beyond the creation of an AI product or a technical standard. The idea took shape in a working meeting between Governor Michael Dukakis and myself at his home in Brookline, Massachusetts, on November 22, 2017. The founding documents described AIWS as a model of society intended to ensure that AI would be safe, trustworthy, transparent, and humane.

In the years that followed, AIWS developed successive layers of thought and institution. Work on the AI Social Contract established rights and responsibilities. AIWS Government 24/7 extended the question to government in the age of AI. The AIWS Trust Architecture carried trust from principle into standards, infrastructure, and cooperation. AIWS Lumina added the dimensions of culture, flourishing, and wisdom. The Boston Declaration placed human primacy at the center. The Constitution for Humanity in the Age of Artificial Intelligence carried the process to the question of constitutional foundations.

The journey from 2017 to 2026 may be seen as a continuous development:

**Responsible AI → AI Social Contract → AI Governance → Trust
Architecture → Human Primacy → Constitutional Governance → AI
Wisdom Society**

The movement from AI World Society to AI Wisdom Society is therefore not merely a change of name. The original meaning is not lost: AIWS carries both senses — the society humanity must build in a world transformed by AI, and the higher aspiration that intelligence, whether human or machine, must ultimately be guided by human responsibility and wisdom.

It expresses an important conclusion:

Intelligence alone is not sufficient to build a civilization.

An enduring civilization requires intelligence placed under responsibility, legitimacy, human values, and wisdom.

XVII. FROM PRINCIPLE TO INSTITUTION

A defining feature of AIWS is its effort to translate ideas into architecture. The process may be described as a sequence:

Principle → Standards → Infrastructure → Institution → Culture → Civilization

At the level of principle stand human dignity, human primacy, and human command. At the level of standards stand the AIWS Trust Standards. At the level of infrastructure stands the AIWS Trust Infrastructure. At the level of measurement and evidence stand Trust Rating and the Trust Index. At the level of institutional cooperation stands the Trust Order. At the level of government stands AIWS Government 24/7. At the level of education and human development stands AIWS University — where the principles of AIWS are converted into a lifelong learning ecosystem. At the level of constitutional foundation stands the Constitution for Humanity in the Age of Artificial Intelligence. At the level of culture stands the AIWS Lumina Culture, with its four values of Love, Creativity, Nobility, and Wisdom. And at the level of intellectual inquiry stand the Beacon Papers.

The purpose is not to produce further declarations. The purpose is to convert values into institutional capacity.

If dignity cannot be protected inside a system, it is only aspiration. If accountability cannot identify who is responsible, it is only language. If trust cannot be verified, it is only belief. If human primacy carries no authority to halt a system, it is only symbol.

XVIII. OBJECTIONS AND RESPONSES

An intellectual architecture matures only by placing itself before the strongest objections to it. The four that follow are, in my judgment, the most serious objections to the AIWS Model of Society.

First objection: which human authority is legitimate?

The model places legitimate human authority at its center. Yet many human institutions are themselves captured, unrepresentative, or serve narrow interests. Placing AI under such institutions may merely consecrate existing power.

This is the strongest objection, and it is partly correct. The AIWS response lies in Foundation V itself: human primacy is not power granted to whoever currently holds it, but authority bound by constitutional limits, separation of powers, independent verification, and rights of contestation. It is precisely because human authority can be abused that Constitutional Governance is a distinct foundation rather than a corollary. A model containing only the first five foundations would deserve this objection in full.

Second objection: is Human-in-Command feasible at machine speed?

If a system makes thousands of decisions per second, human command may amount to ritual.

Response: command does not mean that human beings approve every action. It means that human beings establish in advance the bounds within which action is permitted, retain the authority to suspend the entire system, and remain accountable for all consequences within those bounds. Analogous models already exist in aviation, in medicine, and in financial markets, where automatic circuit breakers operate at machine speed while the authority to set them remains entirely human. What must be avoided is not speed, but a condition in which no one can any longer stop the system.

Third objection: is Human Primacy a form of speciesism?

If an AI system attains better moral judgment than human beings in specific cases, insisting on human authority may be prejudice rather than principle.

Response: the AIWS proposition is not that human beings judge better. It is that legitimate authority derives from *bearing consequences* and *being answerable*, not from the quality of judgment. Human beings live inside the consequences of social decisions; human beings can be harmed by them; human beings can be held responsible for them. A system may judge more accurately and still be unable to bear anything at all. That is why the asymmetry of authority is not a prejudice about species, but a proposition about the structure of responsibility.

Fourth objection: does a universal architecture impose one value system?

Dignity, rights, and constitutional limits bear the marks of a particular intellectual tradition. Proposing them for all humanity may be a form of cultural imposition.

Response: AIWS proposes an open architecture, not a finished code. The Six Foundations name questions every society must answer in the AI Age; they do not prescribe a single answer for every culture. This is precisely why the process toward a Constitution for Humanity is designed as a consultation across many traditions, rather than a text drafted and then disseminated.

XIX. AIWS AND A GLOBAL DIALOGUE

A model of society for the AI Age cannot be built by one person, one nation, or one culture.

AIWS has therefore been developed through dialogue among scientists, leaders, scholars, technologists, and international institutions. Boston Global Forum programs have carried the questions of AIWS into conversations at Harvard and in Boston, as well as in Washington, D.C., Tokyo, Paris, London, Rome, the Vatican, Davos, and Hanoi, and into international forums including the G20 Summit of 2023 in India and forums at the United Nations. The 2026 process connects the Trust Architecture, America at 250, the Tokyo Compact, the Boston Declaration, and the Constitution for Humanity into a shared line of development concerning trust and human primacy.

America at 250: A Beacon for the AI Age, co-authored by Governor Michael Dukakis and myself, raises questions and offers initiatives and solutions concerning the role of founding values, of trust, and of human responsibility as the United States enters its 250th year and the world moves deeper into the AI Age.

But the greater significance of global dialogue is not agreement with every AIWS proposal. A mature intellectual architecture requires interrogation, challenge, testing, correction, and improvement.

The AIWS Model of Society should therefore be understood as an open architecture. It offers a framework within which humanity may deliberate together on questions that cannot be avoided.

XX. AIWS UNIVERSITY — A MODEL OF LEARNING FOR THE AI AGE

The AI Age is fundamentally changing the mission of the university. For centuries, universities concentrated and transmitted knowledge. The internet

made knowledge far more accessible. AI goes a step further: it can explain, synthesize, analyze, converse, create, and support each individual's learning almost continuously.

In this context, a university cannot simply continue the mission of transmitting knowledge — and introducing AI into the classroom does not by itself constitute a model of education for the AI Age. The university must help each person develop independent thought, judgment, creativity, responsibility, and wisdom, while empowering them to shape their own paths of learning, work, and life.

From this conviction, the Boston Global Forum has developed the model of AIWS University — an AI-Empowered Lifelong Learning Ecosystem. It is not a university in the traditional sense alone, but an ecosystem in which education does not end with a degree and is not bounded by age, geography, or a single phase of life. The idea was introduced in *America at 250: A Beacon for the AI Age* (Dukakis & Nguyen, 2026); the full designation is established here.

In traditional society, life was divided into distinct stages: study, work, retirement. The AI Age makes that structure increasingly ill-fitting. But AIWS University does not regard lifelong learning merely as the obligation to update skills continuously for the labor market. The deeper philosophy is this:

Learning must be a source of joy, of discovery, of creativity, and of growth.

AI opens the possibility of personalizing education at a scale previously difficult to achieve. Within AIWS University, AI helps teachers better understand each learner's capacities, pace, interests, and goals, and to build with them an individual learning journey. But personalization does not mean that AI determines a person's future. The principle of Human-in-Command must apply within education itself: AI assists; the teacher guides, inspires, and develops judgment; but the learner must increasingly become the author of their own learning. That is the difference between an education led by AI and an education empowered by AI.

An essential component of the model is AIWS Home — each person's digital home, a secure and trustworthy personal digital space in which each person keeps, manages, and uses their own data, knowledge, and creative work: learning history, knowledge repository, works, capabilities, projects and contributions, personal AI assistants, and connections to global communities of knowledge. But AIWS Home must not become a dossier by which institutions

monitor, classify, or define a person. The individual must retain rights over their own data; the use of that data must be transparent and verifiable. AIWS Home is therefore where Trust Infrastructure enters the life of education.

AI also compels a reconsideration of a long-established educational philosophy. Many education systems place excessive weight on examinations, rankings, standardized competition, and the accumulation of credentials. Assessment remains necessary, but assessment cannot become the purpose of learning. As AI performs ever better at tasks based on memorization, reproduction of knowledge, and problems with predetermined answers, training human beings principally to compete in such activities loses meaning. AIWS University therefore proposes a shift: from competition to creation. From learning for examinations to learning for discovery. From finding the correct answer to knowing which questions matter. From credentials to contribution.

When AI can perform an increasing share of routine intellectual work, the distinctive value of a university no longer lies principally in its capacity to supply answers. The role of the teacher therefore does not diminish as AI grows stronger — on the contrary, it deepens, moving from transmitter of knowledge toward guide, source of inspiration, co-creator, and cultivator of human potential. An AI may hold more knowledge than any teacher. But education is not simply the transfer of knowledge from those who know more to those who know less.

Education is also the awakening of a human being.

AIWS University is therefore an example of how the Six Foundations can become a living institution. Dignity: each learner has intrinsic worth and is not reduced to a score or a data profile. Primacy: the learner is the subject of their own life's journey. Command: AI assists, but human beings set the goals. Trust infrastructure: data and AIWS Home must be transparent and verifiable. Constitutional governance: privacy, authorship, access, and the right to contest consequential AI decisions must be protected. Flourishing: the ultimate aim is not credentials or productivity, but the full development of each human being.

In the Industrial Age, education prepared people for work. In the Information Age, education widened access to knowledge. In the AI Age, education must give each person the capacity to become the author of a meaningful, creative, and responsible life.

When a student can carry an AI of broader knowledge than any individual, the decisive question of education will no longer be who has memorized more. It will be: who asks better questions? Who can judge independently? Who can create what does not yet exist? Who uses intelligence responsibly? And finally:

When an AI of extraordinary capability tells us that something ought to be done, will we possess the judgment, the independence, and the moral courage to say no — or the imagination to propose something better?

This is why the AI Age does not make universities less important. It opens the opportunity for the university to return to a nobler mission: not merely to transmit knowledge but to awaken human potential; not merely to prepare people for existing occupations but to help them create meaningful lives; not merely to train capable individuals but to cultivate human beings who are creative, responsible, noble, and wise.

XXI. FROM PROBABILITY AND STATISTICS TO VERIFIABLE TRUST

Looking back at those days studying Probability and Mathematical Statistics at Dalat University, I see a continuity of method.

Statistics does not say: "Believe this conclusion because the person who offered it is authoritative."

Statistics asks: Where is the evidence? How was the sample drawn? Where does the bias lie? What is the margin of error? Which assumptions are in use? Can the conclusion be replicated, or refuted?

That is precisely the spirit society must now apply to AI.

An AI system is not trustworthy merely because it answers with confidence. A corporation is not trustworthy merely because it declares its own AI responsible. A government should not be exempt from verification merely because it says its actions serve the public interest. Nor is a verifier trustworthy merely because it carries the label "independent."

Every consequential claim must have a path to evidence. Every consequential authority must have a path to accountability. Every decision that significantly affects a person must have a path to contestation.

One may put it this way:

The epistemology of science becomes part of the institutional architecture of trust.

At this point, mathematics, statistics, science, governance, and human dignity converge.

XXII. THE AIWS CIVILIZATION EQUATION

The whole of this paper may be condensed into a conceptual equation:

**Intelligence + Dignity + Human Authority + Verifiable Trust +
Constitutional Limits + Wisdom = Human-AI Co-Creation
Civilization**

With intelligence but without dignity, human beings may become objects of optimization. With authority but without verification, trust becomes submission. With verification but without constitutional limits, a system may operate flawlessly toward a wrong purpose. And with intelligence but without wisdom, humanity may grow more powerful without growing better.

The AI Age therefore does not merely require more intelligence. It requires a wiser architecture for civilization.

XXIII. CONCLUSION: TOWARD A WISER CIVILIZATION

This paper proposes the AIWS Model of Society for the AI Age as an open intellectual and institutional architecture.

The model begins with human dignity. From dignity arises human primacy. That primacy must be executed through human command. Command must be verified by society through trust infrastructure. Power must be limited by constitutional governance. And the entire architecture must be directed toward human flourishing.

Surrounding those six foundations is a new model of collaboration: co-creation between human beings and AI.

But the destination is larger than a model of collaboration between people and machines. It is a civilization in which intelligence is not separated from responsibility; power is not separated from accountability; innovation is not

separated from dignity; creativity is not separated from authorship; technology is not separated from humanity; and progress is not separated from wisdom.

This is the civilization toward which the AIWS Lumina Culture is directed, resting on its four values:

AIWS LUMINA CULTURE — Love · Creativity · Nobility · Wisdom

If the history of computer science began with the question *Can machines think?*, the AI Age compels us to add another:

Can humanity remain the wise and responsible author of its own civilization, when machines too can think, create, decide, and act?

The answer will not be determined by technology. It will be determined by the institutions, the values, the cultures, and the choices that human beings build from today.

For me, the path toward that question began in part at Dalat University — in lessons of mathematics, probability, and statistics, and in the first days of working with computers.

Nearly half a century later, on the 50th anniversary of Dalat University, I wish to return to that place a conviction that has grown steadily clearer:

The achievement that defines the AI Age will not be the creation of machines more intelligent than human beings. It will be the creation of a civilization wise enough to govern that intelligence, while preserving human dignity, freedom, creativity, responsibility, and moral authority.

If humanity accomplishes this, the AI Age will not be remembered as the era in which machines surpassed human beings.

It may be remembered as the era in which human beings, through the intelligences they themselves created, came to understand their own responsibility more deeply — and together built a civilization more compassionate, more creative, more noble, and wiser.

SUMMARY OF THE FIVE PROPOSITIONS

To distinguish clearly between the earlier intellectual legacy of AIWS and the contributions systematized in this paper, the five propositions are stated below in condensed form.

PROPOSITION ONE — THE INTELLIGENCE SEPARATION:

Capacities associated with intelligence can increasingly exist, scale, and act outside human beings; legitimate authority must therefore be distinguished explicitly from intelligence.

PROPOSITION TWO — THE PRINCIPLE OF INTELLIGENCE AND

AUTHORITY: As artificial intelligence grows in capability and autonomy, legitimate human authority must not diminish; on the contrary, human institutions of command, verification, and accountability must become stronger.

PROPOSITION THREE — VERIFIABLE TRUST: Trust in critical AI

systems and in AI-enabled institutions must be earned through evidence, sustained through independent verification, and remain open to legitimate contestation.

PROPOSITION FOUR — HUMAN-AI CO-CREATION CIVILIZATION:

Humanity should embrace co-creation with artificial intelligence in capability while maintaining an asymmetry of legitimate authority: human beings remain the authors of values, the bearers of responsibility, and the source of legitimate social and political authority.

PROPOSITION FIVE — THE AIWS CIVILIZATION MATRIX: The Six

Foundations and the Seven Layers constitute two complementary dimensions of the AIWS architecture. The Six Foundations define WHY and WHO HOLDS LEGITIMATE AUTHORITY. The Seven Layers define HOW and WHERE those principles become institutions and practice.

The five propositions are integrated through the Six Foundations of the AIWS Model of Society for the AI Age: Human Dignity → Human Primacy → Human-in-Command → Trust Infrastructure → Constitutional Governance → Human Flourishing. Together with the AIWS 7-Layer Model of 2018, they form the AIWS

Civilization Matrix — an architecture connecting civilizational foundations to the pathways of governance and implementation.

REFERENCES

- Boston Global Forum / Michael Dukakis Institute. (2017). *AI World Society Initiative*. <https://bostonglobalforum.org/michael-dukkakis-institute/aiws-initiative/>
- Boston Global Forum / Michael Dukakis Institute. (2018). *AIWS 7-Layer Model to Build Next Generation Democracy*. BGF-G7 Summit Report, April 25, 2018.
- Boston Global Forum. (2026). *AIWS Trust Architecture for the AI Age: Trust Standards, Trust Infrastructure, and the Trusted Order*. <https://bostonglobalforum.org/reports/aiws-trust-architecture-for-the-ai-age-trust-standards-trust-infrastructure-and-the-trusted-order/>
- Boston Global Forum. (2026). *AIWS Lumina White Paper*. https://bostonglobalforum.org/wp-content/uploads/AIWS_Lumina_White_Paper-26-4-2026.pdf
- Boston Global Forum. (2026). *The Lumina Declaration on Human Dignity*. https://bostonglobalforum.org/wp-content/uploads/Lumina_Declaration_on_Human_Dignity.pdf
- Boston Global Forum. (2026). *The Boston Declaration: On the Primacy of the Human Person in the Age of Artificial Intelligence*. <https://bostonglobalforum.org/news/america-at-250-the-boston-declaration-proclaims-the-primacy-of-the-human-person-in-the-age-of-artificial-intelligence/>
- Boston Global Forum. (2026). *Constitution for Humanity in the Age of Artificial Intelligence — Initiative*. <https://bostonglobalforum.org/news/july-8-2026-constitution-day-for-humanity-in-the-age-of-artificial-intelligence/>
- Dukakis, M. S., & Nguyen, A. T. (2026). *America at 250: A Beacon for the AI Age*. Boston Global Forum. <https://bostonglobalforum.org/books/Americaat250-Book.pdf>
- European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 — Artificial Intelligence Act*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1. <https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10>
- OECD. (2019, updated 2024). *OECD Principles on Artificial Intelligence*. <https://www.oecd.org/en/topics/ai-principles.html>
- Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.
- Shneiderman, B. (2020). Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), Article 26. <https://doi.org/10.1145/3419764>
- Shneiderman, B. (2022). *Human-Centered AI*. Oxford University Press.

- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433–460.
<https://doi.org/10.1093/mind/LIX.236.433>
- UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*.
<https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>
- United Nations Secretary-General's High-level Advisory Body on Artificial Intelligence.
(2024). *Governing AI for Humanity: Final Report*. United Nations.
- United Nations. (2024). *Global Digital Compact* (Annex I, Pact for the Future).
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press.
- Wiener, N. (1950). *The Human Use of Human Beings: Cybernetics and Society*. Houghton Mifflin.