



BOSTON GLOBAL FORUM · AIWS

FROM PHILADELPHIA TO BOSTON

Two Constitutions, Two and a Half Centuries

Nguyen Anh Tuan

*Co-Founder, Co-Chair, and CEO, Boston Global Forum · Founder and Chief
Architect, AIWS*

Boston · July 8, 2026

In the summer of 1787, fifty-five delegates gathered in Philadelphia and opened a document with seven words that changed the architecture of political legitimacy: “We the People of the United States.” In the summer of 2026, on Beacon Hill in Boston, a new document opened with a deliberate echo across two and a half centuries: “We, the People of Humanity.” The distance between those two sentences is not merely geographic or temporal. It measures how far the constitutional idea has traveled — and how far it must still go.

Declaration First, Constitution Second

America did not begin with a constitution. It began with a declaration — a statement of moral truths held to be self-evident — and only eleven years later, after the inadequacy of mere aspiration had been proven by experience, did it write the machinery to secure them. The sequence matters. The Declaration of 1776 told the world what Americans believed; the Constitution of 1787 told them how those beliefs would be enforced, funded, adjudicated, and amended. A declaration inspires. A constitution operates.

The same sequence is now unfolding in the age of artificial intelligence. The Boston Declaration on the Primacy of the Human Person, proclaimed on July 4, 2026, stated the moral truths of the AI Age. The AIWS Constitution for Humanity in the AI Age, now in its Constitutional Edition, begins the harder work: converting conviction into structure. History suggests the second step is where most movements fail — and where the American founders succeeded for reasons worth studying closely.

The Forgotten Genius of Article VII

The most consequential provision of the United States Constitution may be its least quoted. Article VII declared that the Constitution would enter into force “between the States so ratifying” once nine of the thirteen had approved it. Unanimity was not required. Delaware ratified first; Rhode Island resisted until 1790; the Union proceeded without waiting. The founders understood that a constitution which requires everyone’s consent before it binds anyone will bind no one. They designed for the willing.

This is the decisive lesson for a Constitution for Humanity. Its path to reality does not run through universal acclamation, which will never come, but through accession: an open instrument that nations join because they judge it reasonable, entering into force among the ratifying parties once a threshold is met. Modern treaty practice confirms the wisdom of Philadelphia — the Rome Statute required sixty ratifications, the Paris Agreement fifty-five. A constitutional framework for the AI Age should carry its own Article VII: a clause of accession, a threshold of entry into force, and the quiet confidence that the first movers will make membership worth having. Without such a clause, a constitution remains a charter; with it, aspiration acquires a legal pulse.

Why the States Joined

The states of 1788 did not ratify because the Preamble was beautiful. They ratified because union offered commerce, common defense, sound currency, and standing in the world — concrete benefits that made the costs of sovereignty pooled seem a bargain. Any constitution for the AI Age must answer the same unsentimental question: what does a ratifying nation receive? The answer is already visible in the architecture that precedes this Constitution rather than follows it. Nations that accede gain access to shared trust infrastructure; certification of their AI systems under common trust standards; interoperability with the trustworthy AI ecosystems of fellow members; and a voice in the governance of the intelligence that will shape their citizens’ lives. The American framers had to build their institutions for decades after ratification. The AI Age, unusually, has built some of its institutions first — trust standards, trust ratings, a trust order — so that the benefits of membership exist before the ink of accession dries.

An Inversion, and a Shared Design Principle

In one respect the new document deliberately reverses Philadelphia. The original Constitution built the machinery of government first and added a Bill of Rights only in 1791, under political duress. The Constitution for Humanity places dignity and rights first and institutions after — absorbing two centuries of constitutional learning that begins with the person, not the state.

Yet beneath the inversion lies a shared design principle. Madison’s generation made government deliberately slow — two chambers, a veto, divided powers — installing friction into the machine so that no momentary passion could become instant tyranny. Ambition was made to counteract ambition. The Constitution for Humanity installs the same principle one level deeper: “healthy friction” within AI systems themselves, so that no perfect optimization can quietly dissolve the human will. Tocqueville warned in 1840 of a soft despotism that would not break men’s wills but soften, bend, and guide them; he could not have imagined that the shepherd would one day be an algorithm. The “comfortable cage” named in the new Constitution is his warning, renewed for machines. In both centuries, the constitutional insight is identical: freedom is not preserved by power alone, but by the deliberate imperfection of power.

From Virginia Plan to Philadelphia Moment

Honesty requires one further admission. The delegates in Philadelphia could bind their states because they were sent by their states. Legitimacy was not requested after the drafting; it was present at the table. A Constitution for Humanity will earn ratification the same way — not by asking nations to adopt a finished text, but by inviting them to finish it. The present Constitutional Edition should therefore be understood as this generation’s Virginia Plan: the coherent draft that frames the deliberation, not the deliberation itself. The Philadelphia moment lies ahead — a constitutional convention for the AI Age, in 2027, the tenth anniversary of AIWS, where nations, through their representatives, complete what has been begun and sign what they have shaped. What a nation helps to write, it will defend.

The Wager

The Constitution of 1787 was a wager that a people could govern themselves by reflection and choice rather than accident and force. It ran barely four and a half thousand words, and its austerity was its endurance. The Constitution for Humanity makes a larger wager with a longer sentence: that humanity can govern its most powerful creation before that

creation governs humanity. The first wager took nine states to win. The second will take the first courageous nations willing to ratify — and the patience to let the rest arrive, as Rhode Island eventually did. Two and a half centuries from Philadelphia to Boston, the constitutional idea remains what it has always been: not a description of the world as it is, but an instrument for building the world as it ought to be. Let the AI Age have its instrument.

TỪ PHILADELPHIA ĐẾN BOSTON

Hai bản Hiến pháp, hai thế kỷ rưỡi

Nguyễn Anh Tuấn

*Đồng sáng lập, Đồng Chủ tịch và CEO Boston Global Forum · Người sáng lập và
Kiến trúc sư trưởng AIWS*

Boston · Ngày 8 tháng 7, 2026

Mùa hè năm 1787, năm mươi lăm đại biểu tụ hội tại Philadelphia và mở đầu một văn kiện bằng bảy chữ đã làm thay đổi kiến trúc của tính chính danh chính trị: “We the People of the United States” — “Chúng ta, Nhân dân Hợp chúng quốc Hoa Kỳ.” Mùa hè năm 2026, trên đồi Beacon Hill ở Boston, một văn kiện mới mở đầu bằng tiếng vọng có chủ đích xuyên qua hai thế kỷ rưỡi: “We, the People of Humanity” — “Chúng ta, Nhân dân của Nhân loại.” Khoảng cách giữa hai câu văn ấy không chỉ là địa lý hay thời gian. Nó đo hành trình mà tư tưởng lập hiến đã đi qua — và chặng đường còn phải đi tiếp.

Tuyên ngôn trước, Hiến pháp sau

Nước Mỹ không khởi đầu bằng một bản hiến pháp. Nước Mỹ khởi đầu bằng một bản tuyên ngôn — lời khẳng định những chân lý đạo lý được coi là hiển nhiên — và phải mười một năm sau, khi kinh nghiệm đã chứng minh sự bất cập của khát vọng đơn thuần, người Mỹ mới viết nên bộ máy để bảo đảm những chân lý ấy. Trình tự này mang ý nghĩa sâu sắc. Tuyên ngôn 1776 nói với thế giới người Mỹ tin vào điều gì; Hiến pháp 1787 nói rõ những niềm tin ấy sẽ được thực thi, tài trợ, phân xử và tu chính ra sao. Tuyên ngôn truyền cảm hứng. Hiến pháp vận hành.

Trình tự ấy đang lặp lại trong kỷ nguyên trí tuệ nhân tạo. The Boston Declaration — Tuyên ngôn Boston về Vị thế Tối thượng của Con người — công bố ngày 4 tháng 7 năm 2026, đã khẳng định những chân lý đạo lý của Kỷ nguyên AI. Hiến pháp AIWS cho Nhân loại trong Kỷ nguyên AI, nay ở Ấn bản Hiến định, bắt đầu phần việc khó hơn: chuyển hóa niềm tin thành cấu trúc. Lịch sử cho thấy bước thứ hai chính là nơi hầu hết các phong trào thất bại — và là nơi các nhà lập quốc Hoa Kỳ đã thành công, vì những lý do rất đáng nghiên cứu kỹ.

Thiên tài bị lãng quên của Điều VII

Điều khoản có hệ quả lớn nhất của Hiến pháp Hoa Kỳ có lẽ lại là điều ít được trích dẫn nhất. Điều VII quy định Hiến pháp sẽ có hiệu lực “giữa các Bang phê chuẩn” một khi chín trong mười ba bang thông qua. Không đòi hỏi sự nhất trí tuyệt đối. Delaware phê chuẩn đầu tiên; Rhode Island trì hoãn đến tận năm 1790; Liên bang vẫn tiến bước mà không chờ đợi. Các nhà lập quốc hiểu rằng một bản hiến pháp đòi hỏi sự đồng thuận của tất cả trước khi ràng buộc bất kỳ ai thì sẽ không ràng buộc được ai. Họ thiết kế cho những người sẵn lòng.

Đây chính là bài học quyết định cho một Hiến pháp cho Nhân loại. Con đường đến hiện thực của nó không đi qua sự tung hô toàn cầu — điều sẽ không bao giờ đến — mà đi qua cơ chế gia nhập: một văn kiện mở, để các quốc gia tham gia vì thấy hợp lý, và có hiệu lực giữa các bên phê chuẩn khi đạt một ngưỡng nhất định. Thực tiễn điều ước hiện đại xác nhận minh triết của Philadelphia — Quy chế Rome cần sáu mươi nước phê chuẩn, Hiệp định Paris cần năm mươi lăm. Một khung hiến pháp cho Kỷ nguyên AI cần mang trong mình “Điều VII” của riêng nó: một điều khoản gia nhập, một ngưỡng hiệu lực, và niềm tin lặng lẽ rằng những quốc gia tiên phong sẽ làm cho tư cách thành viên trở nên đáng giá. Thiếu điều khoản ấy, hiến pháp sẽ dừng lại ở hiến chương; có điều khoản ấy, khát vọng bắt đầu có nhịp đập pháp lý.

Vì sao các bang gia nhập

Các bang năm 1788 không phê chuẩn vì Lời nói đầu hay đẹp. Họ phê chuẩn vì liên bang mang lại thương mại, quốc phòng chung, đồng tiền ổn định và vị thế trên thế giới — những lợi ích cụ thể khiến cái giá của việc chia sẻ chủ quyền trở nên xứng đáng. Bất kỳ bản hiến pháp nào cho Kỷ nguyên AI cũng phải trả lời cùng một câu hỏi tỉnh táo ấy: quốc gia phê chuẩn sẽ nhận được gì? Câu trả lời đã hiện hữu trong chính kiến trúc đi trước bản Hiến pháp này thay vì theo sau nó. Các quốc gia gia nhập sẽ được tiếp cận hạ tầng niềm tin chung; được chứng nhận hệ thống AI của mình theo các chuẩn mực niềm tin thống nhất; được tương thông với các hệ sinh thái AI đáng tin cậy của các thành viên khác; và có tiếng nói trong việc quản trị nguồn trí tuệ sẽ định hình đời sống công dân của mình. Các nhà lập quốc Mỹ phải mất hàng thập kỷ sau phê chuẩn mới dựng xong thể chế. Kỷ nguyên AI, một cách khác thường, đã dựng trước một phần thể chế của mình — chuẩn mực niềm tin, xếp hạng niềm tin, một trật tự niềm tin — để lợi ích của tư cách thành viên tồn tại ngay trước khi mực của văn kiện gia nhập kịp khô.

Một sự đảo chiều, và một nguyên lý thiết kế chung

Ở một phương diện, văn kiện mới chủ ý đảo ngược Philadelphia. Hiến pháp nguyên bản dựng bộ máy chính quyền trước, và chỉ bổ sung Tuyên ngôn

Nhân quyền vào năm 1791 dưới sức ép chính trị. Hiến pháp cho Nhân loại đặt phẩm giá và quyền con người lên trước, thể chế theo sau — tiếp thu hai thế kỷ học hỏi lập hiến của nhân loại: khởi đầu từ con người, không phải từ nhà nước.

Nhưng bên dưới sự đảo chiều ấy là một nguyên lý thiết kế chung. Thế hệ của Madison cố tình làm cho chính quyền vận hành chậm — lưỡng viện, quyền phủ quyết, tam quyền phân lập — cài đặt lực ma sát vào bộ máy để không một con cuồng nhiệt nhất thời nào có thể lập tức trở thành bạo quyền. Tham vọng được dùng để kiềm chế tham vọng. Hiến pháp cho Nhân loại cài đặt cùng nguyên lý ấy ở một tầng sâu hơn: “ma sát lạnh mạnh” ngay trong chính các hệ thống AI, để không một sự tối ưu hóa hoàn hảo nào có thể lạng lể làm tan rã ý chí con người. Tocqueville từng cảnh báo năm 1840 về một nền chuyên chế mềm không bẻ gãy ý chí con người mà làm mềm nó, uốn nó, dẫn dắt nó; ông không thể hình dung rằng người chặn dắt một ngày kia sẽ là thuật toán. “Chiếc lông êm ái” được gọi tên trong bản Hiến pháp mới chính là lời cảnh báo của ông, được viết lại cho thời đại máy móc. Ở cả hai thế kỷ, minh triết lập hiến vẫn là một: tự do không được gìn giữ bằng quyền lực, mà bằng sự bất toàn có chủ đích của quyền lực.

Từ Phương án Virginia đến khoảnh khắc Philadelphia

Sự trung thực đòi hỏi thêm một điều phải thừa nhận. Các đại biểu ở Philadelphia có thể ràng buộc các bang của mình vì họ được chính các bang cử đến. Tính chính danh không được xin sau khi soạn thảo; nó hiện diện ngay tại bàn nghị sự. Một Hiến pháp cho Nhân loại sẽ giành được sự phê chuẩn theo đúng cách ấy — không phải bằng việc đề nghị các quốc gia thông qua một văn bản đã hoàn tất, mà bằng việc mời họ cùng hoàn tất nó. Ấn bản Hiến định hiện nay, vì vậy, nên được hiểu như Phương án Virginia của thế hệ này: bản dự thảo mạch lạc định khung cho cuộc nghị luận, chứ chưa phải chính cuộc nghị luận. Khoảnh khắc Philadelphia còn ở phía trước — một hội nghị lập hiến cho Kỷ nguyên AI, vào năm 2027, kỷ niệm mười năm AIWS, nơi các quốc gia, qua đại diện của mình, hoàn thiện những gì đã được khởi đầu và ký vào những gì chính họ đã góp phần tạo nên. Điều gì một quốc gia góp tay viết ra, quốc gia ấy sẽ bảo vệ.

Ván cược

Hiến pháp 1787 là ván cược rằng một dân tộc có thể tự trị bằng suy xét và lựa chọn thay vì ngẫu nhiên và bạo lực. Nó chỉ vón vện khoảng bốn nghìn rưỡi từ, và chính sự kiệm lời ấy làm nên sức bền của nó. Hiến pháp cho Nhân loại đặt một ván cược lớn hơn với một câu văn dài hơn: rằng nhân loại có thể quản trị tạo vật quyền năng nhất của mình trước khi tạo vật ấy quản

trị nhân loại. Ván cược thứ nhất cần chín bang để thắng. Ván cược thứ hai sẽ cần những quốc gia can đảm đầu tiên sẵn lòng phê chuẩn — và sự kiên nhẫn để những quốc gia còn lại tự tìm đến, như Rhode Island rồi cũng đã đến. Hai thế kỷ rưỡi từ Philadelphia đến Boston, tư tưởng lập hiến vẫn là điều nó luôn là: không phải bản mô tả thế giới như nó đang có, mà là công cụ để kiến tạo thế giới như nó cần phải trở thành. Hãy để Kỷ nguyên AI có công cụ của mình.

. . .

Boston Global Forum · AIWS · Boston, July 8, 2026