Algorithmic Sovereignty and Standards: Building the Architecture of Trust

Authored by: Angelo Valerio Toma

Angelo Valerio Toma is a writer and international affairs analyst specializing in digital sovereignty, algorithmic governance, emerging technologies, and the Global South. His work examines how power, inequality, and control take shape through technological infrastructures and cross-border algorithmic systems.

Artificial intelligence has become part of the core infrastructure of public communication. Systems built in one country now shape what people see in another—deciding what enters public debate and what disappears from view. Meta's algorithms did exactly that in Ethiopia, amplifying divisive content into the center of public life. This influence is structural, not intentional, born of how models are built, trained, and scaled.

Governments are trying to respond, but the gap between law and technology keeps widening. Regulatory tools move too slowly, creating uncertainty and pushing the policy environment toward fragmentation.

The U.S. system of voluntary, private-sector-led standards offers a more flexible path. Grounded in technical expertise and consensus, it evolves with the technology and provides practical methods for transparency, oversight, and risk management. When it comes to AI, it offers something regulation alone cannot: a realistic way to preserve trust as these systems become embedded in the civic infrastructure of multiple jurisdictions.

AFI as Structural Influence and the Limits of Traditional Regulation

The technical dynamics behind Algorithmic Foreign Influence (AFI) start in the architecture of modern AI systems. Artificial intelligence functions as a transnational layer of information infrastructure. Systems developed in one jurisdiction help determine what becomes visible, searchable, or meaningful in another. Even small changes in model behavior can shift what users encounter in civic spaces, altering informational visibility across borders.

This is the environment in which AFI is situated. In earlier work, I defined AFI as "the measurable, cross-jurisdictional impact of algorithmic systems—trained, hosted, or governed abroad—that shape political discourse, suppress civic expression, or restructure cultural and linguistic visibility within a sovereign state, absent direct foreign intent." The essential point is structural: influence emerges from system design and deployment practices, not targeted actions, and informational environments can shift faster than governments can detect or respond.

Traditional regulation cannot keep pace with this dynamic. Rulemaking cycles extend across years, while frontier models change continuously. The resulting gap creates uncertainty and forces governments to respond to systems that have already evolved. National measures aimed at mitigating cross-border risks, including data localization and platform-level restrictions, may address specific concerns but also contribute to fragmentation.

AFI sits at the center of these challenges. It stems from fast-moving, cross-border technical dynamics that exceed what national regulation can manage alone. Dealing with AFI demands governance mechanisms that can evolve iteratively and still apply consistently across jurisdictions. This is where standards become indispensable alongside regulation. Unlike static regulatory tools, standards and conformity assessment provide adaptable methods for documentation, testing, and transparency—and create a shared technical language through which governments, developers, and institutions can evaluate cross-border algorithmic impacts.

Standards and Conformity Assessment as the Architecture of Algorithmic Sovereignty

Standards offer a different foundation—one built for rapid technical development without prescriptive controls. The U.S. voluntary, private-sector-led model relies on technical expertise and consensus to define shared expectations for safety, reliability, and interoperability. It aligns with the reality of algorithmic systems, which update iteratively rather than in discrete cycles.

As model architectures shift more frequently, systems require flexible documentation, testing, and evaluation methods.

A particular strength of the U.S. standards ecosystem is its integration with conformity assessment. Accreditation, testing, inspection, and certification provide mechanisms for determining whether systems meet agreed-upon criteria. As the coordinator of the U.S. standards and conformity assessment system, the American National Standards Institute (ANSI) supports this infrastructure across sectors where transparency and traceability are critical. In the context of AI, these tools make assessment operational: evaluation occurs within development workflows rather than waiting for regulatory triggers.

That logic echoes the emerging idea of algorithmic due process. Oversight does not require governments to micromanage technology; it requires clear procedures for documenting system behavior, assessing impacts, and enabling independent review. Standards provide the technical expectations for these procedures, while conformity assessment bodies deliver the external evaluations needed to build trust. Together, they let institutions examine algorithmic systems at the structural level—where influence actually originates.

Voluntary standards also scale internationally without requiring regulatory uniformity. Because AI systems cross borders by default, interoperability is essential. Open, consensus-based standards allow jurisdictions to measure and communicate risk using shared concepts, reducing fragmentation even when national laws diverge.

Standards and conformity assessment therefore serve as the architecture of algorithmic sovereignty. They do not replace regulation, but they supply the technical and procedural groundwork that enables institutions to respond to evolving risks, including AFI-related risks, in a coordinated, predictable way. In practice, they offer a stable reference point in an environment where systems shift rapidly, allowing evaluation to keep pace with technology instead of trailing behind it.

The stakes here are not only technical; they are geopolitical. If standards fail to keep pace, global AI governance will default to whichever jurisdictions or platforms move first, regardless of whether their approaches support openness or democratic oversight. That outcome would constrain

innovation and give powerful platforms or regulatory blocs an outsized say over the informational conditions far beyond their borders. Voluntary, market-driven standards that preserve innovation are also the best defense against techno-nationalist silos and de facto algorithmic firewalls between jurisdictions. For the United States, leadership in standards has become a strategic requirement. Without agreed-upon standards, cross-border AI governance splinters along regulatory lines, and the informational sovereignty of open societies erodes by default rather than by design.

International Coordination and the Role of Standards in Preventing Fragmentation

As artificial intelligence systems globalize, national regulatory approaches are diverging. Jurisdictions are developing their own definitions, documentation requirements, and risk-classification schemes, each moving at a different pace. The result is a growing mismatch in how countries expect developers to demonstrate safety, transparency, and reliability.

These differences shape deployment. Many developers now maintain separate release workflows across regions because disclosure expectations, testing procedures, and documentation formats differ. That produces operational friction: multiple compliance pathways, inconsistent evidence requirements, and duplicated assurance practices. For systems built on global data and cross-border engineering teams, procedural divergence rapidly becomes technical fragmentation.

Technical standards push back against this dynamic by creating shared, non-regulatory baselines that apply across jurisdictions. When core concepts—such as documentation practices, terminology, and approaches to risk evaluation—align, interoperability becomes possible even when national laws do not. Standards give developers and auditors a common reference point for understanding and communicating system behavior, preventing procedural divergence from hardening into structural barriers.

Through ANSI, the U.S. is an active player in maintaining this alignment in the AI space. As the U.S. representative to ISO and IEC, and through its leadership as secretariat of ISO/IEC JTC 1—the primary venue for international IT and AI standards—ANSI connects domestic expertise to global processes. U.S. stakeholder contribution help ensure that U.S. risk-management approaches and technical practices are reflected in the frameworks countries use to evaluate models, document behavior, and conduct assurance.

International standardization has started to concentrate on the areas where fragmentation bites the hardest: cross-border testing, documentation, and verification. Shared technical methods are often the only practical way to maintain alignment as jurisdictions adopt different assurance expectations. Standards provide common structures for reporting and evaluation, reducing duplication and creating consistent processes even when regulatory environments remain in flux.

Global standards provide procedural common ground—shared expectations and validation methods that remain usable even when national policies diverge. While they cannot eliminate political differences, they can significantly reduce operational friction. For AI systems that operate globally, that reduction is essential: it allows cooperation to continue, transparency to scale, and oversight to function in a world where influence—and risk—cross borders by default.

Integrating AFI Considerations into ISO/IEC 42001

ISO/IEC 42001 introduced the first global management system standard for artificial intelligence, defining governance processes, documentation practices, and continuous improvement cycles. Although it does not explicitly address cross-border informational impacts, its structure offers a natural entry point for integrating AFI-related considerations.

The most direct integration point is documentation. Management systems rely on structured reporting to make technical decisions transparent. Extending documentation fields to include geographic deployment, linguistic coverage, and context-specific performance would require minimal adjustment while capturing factors that shape visibility and reliability across borders. When performance varies by language or region, the potential for structural influence increases, yet these differences often remain undocumented. Incorporating them into routine reporting would give organizations early visibility into AFI-relevant risks.

Impact assessment is another obvious fit. While 42001 does not mandate a specific methodology, it establishes the governance framework within which assessments occur. Adding informational risk categories—such as linguistic variation, context-specific accuracy, or cross-jurisdictional generalization—would broaden the lens through which risk is understood without altering the structure of the standard. In this context, AFI becomes one of the structural risk sources to be monitored, and 42001 can be extended incrementally to cover AFI-related risks without disturbing its core architecture.

The same logic carries over to evaluation and testing. Management systems require organizations to verify whether system outputs meet expectations. Encouraging testing across relevant linguistic and cultural domains—without prescribing fixed metrics—would help identify informational asymmetries before deployment. The goal is not to create new performance requirements but to ensure that uneven model behavior is observed and documented consistently.

None of these adjustments require changes to the architecture of ISO/IEC 42001. AFI considerations fit naturally within existing concepts such as context of use, intended purpose, and risk sources. Enhancing documentation, impact assessment, and evaluation practices would allow organizations to identify cross-border informational effects early and integrate them into governance processes. Additions could take the form of annexes, guidance documents, or cross-references to related standards, preserving 42001's flexibility while making it more capable of addressing the structural nature of algorithmic influence.

Conformity Assessment as a Catalyst for AFI-Relevant Action

Understanding whether AI systems generate cross-border informational effects begins with the capacity to evaluate how they behave across different linguistic, cultural, and operational environments. Conformity assessment provides the structure for that evaluation. Through defined testing methods, documentation requirements, and criteria for examining system behavior, it enables organizations to identify where informational variation may arise without making assumptions about intent or political context.

ANSI's role is to ensure that the standards infrastructure supporting this work is coherent and technically reliable. By facilitating U.S. participation in ISO, IEC, and ISO/IEC JTC 1, ANSI helps maintain a standards environment in which terminology, documentation expectations, and evaluation concepts develop in a coordinated and predictable way. When existing standards are clarified or expanded—whether through improved terminology, clearer documentation guidance, or more precise descriptions of evaluation conditions—they become more useful for identifying how AI systems perform across regions, languages, and deployment contexts.

ANSI can also strengthen the connection between risk management and operational assurance.

NIST's AI Risk Management Framework emphasizes robust documentation, evaluation grounded in context of use, and ongoing monitoring throughout the AI lifecycle—elements that align naturally with AFI-related concerns. Accreditation bodies can help translate these principles into voluntary assessments that examine multilingual performance, deployment environments, and cross-border variation. This transforms abstract risk categories into concrete review practices and gives organizations repeatable methods for identifying informational risks early.

Pilot programs focused explicitly on AFI-related risks are a natural next step. ANAB-accredited evaluators can conduct exploratory assessments focused on cross-border informational dynamics—reviewing documentation quality, examining model behavior in underrepresented languages, or evaluating variation across regions. Such pilots reveal gaps, show which practices actually work, and generate evidence that can inform future standards.

International compatibility is just as critical, because as jurisdictions adopt different reporting and documentation expectations, fragmentation risks grow. U.S. participation in multilateral standardization efforts helps anchor AFI-relevant work in a shared technical vocabulary and ensures that U.S. practices remain interoperable with developments in Europe, Canada, and Asia. This work also progresses in coordination with other national and international standards bodies in like-minded democratic nations, ensuring that AFI-relevant methods develop through open, consensus-based processes.

Across these activities, ANSI facilitates the technical groundwork that makes informed governance possible. For AFI—where risks arise from structural dynamics rather than intent—this groundwork is essential, so that institutions can actually detect and manage cross-border informational effects using methods that are credible, interoperable, and able to keep up with rapid technological change.

Conclusion

The cross-border effects of AI systems expose a gap that traditional regulation cannot close. These systems evolve faster than statutory definitions can track and generate informational impacts shaped by context, deployment, and linguistic variation across jurisdictions. Dealing with this gap demands governance tools that can keep up with rapid technical change while still working transnationally. Standards and conformity assessment provide that foundation: shared terminology, structured documentation, and evaluation practices that remain reliable even as national policies diverge.

The U.S. standards system is well suited to this environment. Its voluntary, expertise-driven structure evolves with technology rather than trailing behind it. Through international standardization, accreditation, and alignment with NIST's risk-management work, U.S. stakeholders can help build the technical infrastructure needed to understand and evaluate how AI systems shape information across borders.

For challenges like AFI—where influence emerges from structural dynamics rather than intent—this infrastructure is essential. It allows institutions to identify and manage informational risks before they harden into fragmentation and provides the common ground required for cooperation in a digital environment where sovereignty depends less on controlling borders and more on governing the infrastructures that cross them.