The Universal Middleware: AI Agents as the Architecture for Digital Equity

Daniel Luna

The University of Chicago
Chicago, Illinois, USA
lunad@uchicago.edu

Abstract—The digital divide is not only a matter of device ownership or connectivity; for millions of people navigating disability, language barriers, or unreliable networks, the modern web remains a maze of locked doors. Static accessibility, manual fixes to individual websites, cannot keep pace. Recent studies find that 94.8% of home pages fail basic accessibility standards, with an average of 51 distinct errors per page. This paper argues that AI agents can act as a Universal Middleware: an architectural layer that decouples the user from hostile, complex web content and adapts interface and content to the user's context. We trace three barriers that this middleware must address. First, the machine barrier: the web was built for human visual consumption; CAPTCHAs, anti-bot measures, and heavy JavaScript block automation. Adopting agentic standards such as LLM-LD, which expose semantic structure to agents, is a prerequisite. Second, the human barrier: agents can serve as visual interpreters (via DOM distillation), voice-first interfaces that bypass literacy demands, and universal translators for under-resourced languages. Third, the economic barrier: offloading parsing and rendering to the cloud revives the thin-client model, reducing data use and extending the life of older devices. We also consider ethical risks: algorithmic bias, cultural erasure, and opacity in agentic decision-making. We argue that the middleware should be framed as a temporary bridge to inclusion, supported as public digital infrastructure while the long-term goal remains universal design.

Index Terms—AI agents, digital equity, accessibility, universal design, middleware

I. Introduction: The Middleware Thesis

Imagine a user with limited vision attempting to apply for public benefits online. The page loads slowly on an aging phone connected to an unstable network. Form fields lack labels. Buttons are embedded in layers of JavaScript. A CAPTCHA demands that the user identify images of buses. For millions of people navigating disability, language barriers, or unreliable connectivity, the modern web is less a public utility than a maze of locked doors. AI agents offer a different possibility: a software intermediary that reads the page, interprets its structure, and presents only the essential information through voice or simplified text.

As a student of Computer Science, Mathematics, and Data Science at the University of Chicago, I am trained to see the world in layers. Data structures, algorithms, and interfaces connect human intention to machine execution. This interdisciplinary perspective is central to understanding the middleware thesis. My coursework in data science reveals that information flow is fundamentally about encoding and access: who can parse the data, and what gets lost in translation. Concurrently, studying human-computer interaction highlights the necessity of designing systems that serve the full spectrum of human abilities. Viewing both the technical substrate and its human impact has led me to a conviction. The digital divide is not merely about device ownership, but about who can actually use the web once they reach it. Together, these disciplines expose a sobering reality. The internet, for all its promise, remains a gatekeeper. Static accessibility has failed.

AI agents are not merely chatbots; they are a Universal Middleware, an architectural layer that decouples the human interface from digital content. In the words of IBM, an artificial intelligence (AI) agent is "a system that autonomously performs tasks by designing workflows with available tools" [1]. This definition points to something deeper: agents sit between the user and the hostile, complex web, translating and simplifying. They are the bridge that can bypass human barriers (literacy, language, disability) and economic barriers (bandwidth, hardware), provided we solve machine barriers via simple standards. The Universal Middleware thesis is simple. AI agents can act as a dynamic translation layer, adapting content and interface to the user's context. In doing so, they ensure that "all people are included in a state's economic system" [2, p. 2]. In the sections that follow, I first outline the systemic standards we must adopt, then

trace how this architecture addresses human and economic barriers. I then consider the ethical shadows and the pragmatic case for the agent as a temporary bridge to equity.

II. The Machine Barrier: Rearchitecting for Agentic Access

Solving the machine barrier is simpler than manual static accessibility; it is a systemic prerequisite. We, including standards bodies, governments, and major tech platforms, must adopt standards and reduce hostility to agents so that the universal middleware can function. Adopting these standards is the foundational step that enables the AI future.

A. The Hostile Web: CAPTCHAs and Complexity

The web was built for humans looking at screens, not agents parsing structure. CAPTCHAs, anti-bot measures, and complex JavaScript create a hostile environment that blocks automation, and thus any agent acting on behalf of a user. For example, a CAPTCHA that requires selecting images of buses prevents an automated accessibility agent from acting on behalf of a blind user. As the Agent-E team notes, "websites are primarily designed for human visual consumption," and "Complex widgets, such as date selectors, are easy for humans to operate but pose difficulties for agents" [3, p. 2]. Without reducing this hostility, the bridge cannot be built.

B. Standardizing the Agentic Layer: From HTML to LLM-LD

HTML encodes presentation (layout, styling, scripts) for human eyes; it was never built for agents to parse. A new layer is emerging, one that exposes semantic structure directly so AI agents can read web content without reverse-engineering layout. Standards such as LLM-LD, a recently published open standard, formalize this approach, explicitly marking intent and available actions [4]. Consider the difference in practice. A government benefits page is a maze of nested containers, invisible labels, and date pickers that only function via complex JavaScript. An agent navigating this hostile page must guess which elements are form fields, often failing and leaving the user stranded. An agent-readable page, by contrast, exposes a structured representation that explicitly tags "eligibility criteria," "required fields," and the "submit action."

This allows the agent to reliably navigate and fill the form on behalf of a user relying on a voice interface or a low-bandwidth connection. The idea resembles standardized shipping containers, which transformed global trade. Work on the Internet of Agents explores fundamentals, applications, and challenges for a world where agents are first-class users of the web [5]. Measuring success cannot be only about benchmarks; it must be about "exploring equity of Internet experience across multiple dimensions" [6, p. 9]. Standardization is how we clear the path for the agent layer.

C. The Risk of Data Monoculture

If only well-resourced, agent-friendly sites are easily indexed, the datasets that train AI will reflect a narrow slice of humanity. Civil society reports on AI and the Global South highlight the risk of data monoculture if we fail to act, emphasizing the need for inclusive participation in AI decision-making [7]. Ensuring diverse, agent-accessible content is necessary for both equitable middleware and representative AI. There is a "consistent gap between legal requirements and socioeconomic optimization" [2, p. 3]. Closing that gap requires both technical standards, so that diverse sites are agent-accessible, and a commitment to ensuring that the middleware serves diverse communities.

III. The Human Barrier: Decoupling Interface from Content

A. The Failure of Static Accessibility

Static accessibility, manual fixes to individual websites, cannot keep pace with the scale of the web. Across the one million home pages evaluated in the most recent WebAIM study, "50,960,288 distinct accessibility errors were detected, an average of 51 errors per page", and "94.8% of home pages had detected WCAG (Web Content Accessibility Guidelines) 2 failures" [8], where a failure denotes a violation of a testable criterion, such as missing form labels or insufficient color contrast. For millions of users, whether relying on screen readers, navigating with cognitive disabilities, or struggling with dominant languages, these 51 errors per page are not mere inconveniences; they are locked doors that exclude them from vital services.

Design for inclusion, however, has historically lagged behind. As Pérez and Johnston observe, "many technologies we take for granted today, from captioning to audiobooks and text translated into speech, were created because of the lived experience of people with disabilities" [9, p. 5]. These innovations illustrate the curb-cut effect, where solutions designed for one population ultimately benefit everyone, while the absence of such design leaves individuals "penalized financially for inaccessibility" and forced to rely on costly, inefficient workarounds [2, p. 2]. Yet relying on every site owner to retrofit their pages has left the majority of the web behind. This intermediary does not wait for every publisher to comply, a process that has proven too slow and piecemeal. Instead, it interprets the page as it exists and

presents an accessible version to the user, offering a scale solution for a web-scale problem.

B. The Visual Interpreter: Agents as Eyes

For users who are blind or have low vision, the web is a forest of layout and visuals. Agents can act as eyes by interpreting the page and presenting a distilled version. This process is akin to taking a complex map and turning it into a simple set of directions, stripping away visual clutter to expose only the structure and actions a user or an assistive system needs. The agent first strips a webpage down to its structural skeleton; an approach researchers call *DOM distillation.* The Document Object Model (DOM) is the underlying structural blueprint that web browsers use to construct a page; distillation involves parsing this heavy HTML into a minimal, clean structure.

In Agent-E, a state-of-the-art web agent, "the planner agent is insulated from the overwhelming and noisy details of the website and DOM" [3, p. 3]. By stripping away this noise using the browser's accessibility tree, the structured representation of the page that assistive technologies use to navigate content, the agent reduces cognitive load and exposes a simplified HTML structure directly to assistive technologies. Projects like A2UI extend this idea by letting agents generate dynamic interfaces, sending the "UI spec as a sequence of messages" directly to the client [10]. The agent becomes the visual interpreter. It reads the hostile DOM and speaks a clearer, structurally sound language to the user.

C. The Literacy Bypass: Voice as the Primary Interface

For billions with low or no literacy in the language of a given site, text-heavy interfaces are a barrier. Voice-first and agentic interfaces can bypass that barrier. The user speaks or gives simple instructions; the agent navigates forms, reads dense text, and returns answers in plain language. Research on agentic artificial intelligence for low-resource languages shows that agents can mediate between the user's spoken or simple input and the complexity of online services [11]. By automatically extracting the necessary fields from a complex government form and prompting the user verbally for just the essential information, the agent handles the literacy demand on the user's behalf. It acts as the primary interface, shielding the user from bureaucratic jargon and dense layouts. This is not a replacement for education or language rights; it is a temporary bridge making critical information accessible today.

D. The Universal Translator: Bridging the Linguistic Divide

Unicode proved that technical fragmentation could be solved through a "unique, unified, universal" framework, creating an encoding system that "exists solely to support the various processes that act upon text" [12, p. 3]. Today, AI agents are the active layer that utilizes this encoding to deliver true cross-language access, transforming static text into dynamic, multilingual dialogue.

The next step is to support the process of access across languages. Initiatives such as Masakhane [13] and MENYO-20k [14] advance machine translation and corpora for under-resourced African languages. By providing the foundational data required to train localized models, these efforts enable agents to summarize, translate, or simplify content directly in the user's native tongue. They extend that universal layer, acting as the translator between the global web and the local user, with the translation layer functioning as a linguistic bridge.

IV. The Economic Barrier: The Agent as the Computer

A. Digital Inequity in the Global South and the Urban Core

Access is not only about connectivity; it is about adoption and affordability. In Chicago, "Broadband adoption varies between 58-93% across community areas" [15, p. 1]. This massive spread indicates that a significant portion of the urban population remains cut off from essential digital infrastructure. Research shows that "income can affect the ability to get a broadband connection" and that education shapes the "perceived utility of the Internet" [15, p. 8]. In least developed, landlocked, and small island nations, the economic impact of broadband is well documented [16]. Where devices are old and bandwidth is scarce, the traditional model of shipping heavy web pages to local browsers fails entirely. Here, the agent can shift work to the cloud. The user's device becomes a thin client, a simple interface that mainly displays output and sends input using low-bandwidth text or voice, while the agent does the heavy lifting of rendering and parsing in the datacenter.

B. The Renaissance of the "Thin Client"

Offloading processing to remote agents revives the thin client idea. Digitally accessible services and universal design ensure "that all people are included in a state's economic system" [2, p. 2]. When the agent fetches, parses, and summarizes a page, the user's device need not render bloated DOMs or run heavy JavaScript. Because running AI models in the cloud incurs compute costs, leaving this layer entirely to the free market could replace a technical barrier with a financial one for low-income users. Therefore, this deployment model must be supported by governments or NGOs

and provided as public digital infrastructure. For governments and organizations, investing in such architectures is a "strategic opportunity for economic development" and more efficient service delivery [2, p. 2]. This model fits the picture of sustainable computing. As AI diffusion in low-resource language countries accelerates [17], shifting the processing burden to the cloud extends the life of older devices and lowers the cost of participation in the digital economy.

C. Efficiency in Scarcity: Agent-Based Data Reduction

In regions where connectivity is expensive or unstable, every byte counts. Reports on the economic impact of disruptions to Internet connectivity underscore how critical resilience and efficiency are [18]. Equity in internet experience is not only about raw speed. Latency, loss, and stability vary by geography and income. Studies that compare neighborhoods find stark disparities: for instance, "the median loss rate for South Shore is 0.54%, while that of Logan Square is 0.012%" [6, p. 9]. In communities facing high loss rates and fragile connections, attempting to load a modern, multimegabyte website is an exercise in frustration. The authors note that existing performance metrics often fail to capture the reality of these marginalized populations [6, p. 1]. Agent-based distillation directly addresses this scarcity by extracting only the essential text and functionality, dramatically reducing data use. This functions as a form of bandwidth hygiene, ensuring that even the most fragile connections can successfully retrieve the necessary information.

V. ETHICAL IMPLICATIONS: THE SHADOW OF THE AGENT

A. Algorithmic Bias and Cultural Erasure

When agents mediate access, they also mediate representation. Models trained on skewed data can perpetuate bias. Low-resource languages and dialects are especially at risk. Work on voices unheard, such as resources and models for regional dialects, underscores how easily marginalized varieties are erased from both data and interfaces [19]. Because a single agentic model may serve millions, training biases are amplified, forcing users to adapt to the machine's expectations. Regional dialects face the highest risk of this cultural flattening. Making the bridge equitable requires inclusive data and participatory design: communities depending on this middleware must shape its boundaries.

B. The Transparency Gap in Agentic Decision-Making

Agents make choices: what to show, how to summarize, which link to follow. Those choices can be opaque. The UNESCO Recommendation on the Ethics of Artificial Intelligence stresses transparency, accountability, and human oversight [20]. Similarly, analyses highlight the black-box problem and the need for transparency in agentic decision-making [21]. Consider a user accessing government services through an agent translating into an under-resourced language, such as those supported by Masakhane [13]. A mistranslation or an over-aggressive summarization could easily misstate eligibility criteria, omit a critical deadline, or drop a required document field from a spoken prompt. Because the user cannot easily verify the agent's output against the original hostile web page, these errors can lead to rejected applications and lost benefits. This opacity is especially concerning given the nature of the temporary bridge: vulnerable users are asked to trust an intermediary system whose internal logic they cannot inspect or verify. Without robust transparency, the intermediary system risks replacing one barrier, an inaccessible website, with another, an unaccountable gatekeeper.

VI. THE NATURE OF THE SOLUTION: PRAGMATISM VS. PERFECTION

A. The "Handmade" Ideal vs. the Scale of Exclusion

The ideal remains a web where every site is natively accessible, multilingual, and lightweight. That ideal is not yet reality. The curb-cut effect reminds us that "solutions created for one target population provide universal benefits" [2, p. 2]. Furthermore, "Digitally accessible platforms and universal design standards create solutions that work for PWDs but also improve usability for every constituent." [2, p. 2]. There is an inescapable tension between this handmade, structural ideal and the sheer scale of modern exclusion, where 94.8% of home pages currently fail basic accessibility standards. Universal design is the ultimate goal, but waiting for millions of independent developers to retrofit their sites leaves marginalized users stranded today. Until the scale of the web matches that ideal, agentic middleware is the pragmatic bridge we must build.

B. Objections to Agentic Mediation

Critics may argue that AI middleware lets developers off the hook for native accessibility. However, acknowledging agents as a temporary bridge maintains the pressure for universal design without leaving users stranded today. Others warn that agentic infrastructure creates dependencies on tech companies. This underscores why middleware must be supported as public digital infrastructure rather than a private gatekeeper. Finally, while skeptics question if agents can handle the web's full complexity, emerging standards

like LLM-LD are making the web machine-readable, closing this technical gap.

C. The Temporary Bridge: AI as a Human Right

UNESCO's framework supports the view that access to the benefits of AI should be aligned with human rights and inclusion. Ensuring equitable access to digital services is both an economic and a moral imperative [20]. Calling the middleware a "temporary bridge" is a deliberate framing. It acknowledges that AI translation layers are a stopgap measure, providing essential inclusion today while our long-term efforts remain focused on native structural change. The bridge does not replace the ideal of universal design; rather, it makes digital equity possible in the present while we do the hard work of rearchitecting the web for the future. The temporary bridge is the pragmatic path to digital equity while we work toward structural change.

VII. Conclusion: Toward an Agentic Architecture for Equity

AI agents, acting as a universal middleware layer, have the potential to decouple the user from the hostile web and align digital participation with true economic and social inclusion. Just as the browser once separated users from raw HTML, AI agents may become the next interface layer between humans and the increasingly complex web. Realizing this architecture depends on adopting agentic standards. The machine barrier is not one the middleware solves, but a simpler policy issue we must solve so the middleware can function. Furthermore, deploying these agents ethically requires strict transparency, accountability, and attention to algorithmic bias.

Looking into the next decade, widespread agent adoption could make agent-mediated access the default for essential services. Voice and simplified interfaces may become as common as traditional browsers. However, this future requires vigilance: regulatory frameworks must evolve to mandate agent-readable public sites and oversee algorithmic transparency. If we successfully adopt agentic standards and invest in public middleware infrastructure, we can ensure that language, disability, and income no longer dictate digital citizenship. By treating agents as essential public infrastructure, we assert our digital rights. This middleware is the necessary next layer of the stack, finally making the web readable, usable, and equitable for all.

References

[1] A. Gutowska, "What are AI agents?." Accessed: Mar. 15, 2026. [Online]. Available: https://www.ibm.com/think/topics/ai-agents

[2] K. Young-Gibson, "The Economic Case for Digital Accessibility," NASCIO, 2025. [Online]. Available: https://www.nascio.org/wp-content/uploads/2025/08/NASCIO_The-Economic-Case-for-Digital-Accessibility_2025_a11y.pdf

[3] T. Abuelsaad, D. Akkil, P. Dey, A. Jagmohan, A. Vempaty, and R. Kokku, "Agent-E: From Autonomous Web Navigation to Foundational Design Principles in Agentic Systems," July 2024, 2407.13032v1. [Online]. Available: https://arxiv.org/abs/2407.13032v1

[4] CAPXEL, "LLM-LD — The Open Standard for AI-Readable Websites." Accessed: Mar. 15, 2026. [Online]. Available: https://llmld.org/about

[5] Y. Wang *et al.*, "Internet of Agents (IoA) - Fundamentals, Applications, and Challenges," 2025, 2505.07176. [Online]. Available: https://arxiv.org/abs/2505.07176

[6] R. Sharma, T. Mangla, J. Saxon, M. Richardson, N. Feamster, and N. P. Marwell, "Benchmarks or Equity? A New Approach to Measuring Internet Performance," *A New Approach to Measuring Internet Performance (August 3, 2022)*, 2022.

[7] Care International UK, "AI and the Global South - Exploring the Role of Civil Society in AI Decision-Making," 2024. [Online]. Available: https://careinternationaluk.ams3.cdn.digitaloceanspaces.com/media/documents/AIthe_Global_South_Exploring_the_Role_of_Civil_Society_in_AI_Decision-Making.pdf

[8] WebAIM, "The WebAIM Million - The 2025 report on the accessibility of the top 1,000,000 home pages." Accessed: Mar. 15, 2026. [Online]. Available: https://webaim.org/projects/million/

[9] L. F. Peréz and S. C. Johnston, "Creating Disability-Friendly and Inclusive Accessible Spaces in Higher Education," The National Academies Press, 2024. [Online]. Available: https://nap.nationalacademies.org/resource/27245/Johnston_and_Perez_Creating_Disability-Friendly_and_Inclusive_Accessible_Spaces_in_Higher_Education.pdf

[10] Google A2UI Team, "Introducing A2UI - An open project for agent-driven interfaces." Accessed: Mar. 15, 2026. [Online]. Available: https://developers.googleblog.com/introducing-a2ui-an-open-project-for-agent-driven-interfaces/

[11] N. Kshetri, "From Challenges to Opportunities - Agentic Artificial Intelligence for Low-Resource Languages," *IEEE*, 2024, [Online]. Available: https://ieeexplore.ieee.org/document/11029704

[12] J. D. Becker, "Unicode 88," Aug. 1988. [Online]. Available: https://www.unicode.org/history/unicode88.pdf

[13] W. Nekoto, V. Marivate, T. Matsila, and T. Fasubaa, "Participatory Research for Low-resourced Machine Translation - A Case Study in African Languages," in *Findings of the Association for Computational Linguistics (EMNLP 2020)*, 2020.

[14] D. I. Adelani, D. Ruiter, J. O. Alabi, and D. Adebonojo, "MENYO-20k - A Multi-domain English-Yoruba Corpus for Machine Translation and Domain Adaptation," in *AfricanNLP Workshop, EACL 2021*, 2021.

[15] T. Mangla, U. Paul, A. Gupta, N. P. Marwell, and N. Feamster, "Internet inequity in chicago: adoption, affordability, and availability," *Affordability, and Availability (August 5, 2022)*, 2022.

[16] International Telecommunication Union, "New study shows economic impact of broadband on least developed, landlocked and small island nations." Accessed: Mar. 15, 2026. [Online]. Available: https://www.itu.int/hub/2020/05/new-study-shows-economic-impact-of-broadband-on-least-developed-landlocked-and-small-island-nations/

[17] A. Misra, S. W. Zamir, W. Hamidouche, I. Becker-Reshef, and J. Lavista Ferres, "AI Diffusion in Low Resource Language Countries," 2025, 2511.02752. [Online]. Available: https://arxiv.org/html/2511.02752v1

[18] Global Network Initiative, "The Economic Impact of Disruptions to Internet Connectivity," 2016. [Online]. Available: https://globalnetworkinitiative.org/wp-content/uploads/GNI-The-Economic-Impact-of-Disruptions-to-Internet-Connectivity.pdf

[19] O. Ahia, A. Aremu, D. Abagyan, and H. Gonen, "Voices Unheard - NLP Resources and Models for Yoruba Regional Dialects," in *EMNLP 2024*, 2024.

[20] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," 2022. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137

[21] M. Hosseini, M. Murad, and D. B. Resnik, "Benefits and Risks of Using AI Agents in Research," *Hastings Center Report*, 2026, [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/hast.70025