

The Paradox of Scientific Visibility, Ranking, and Internationalization of Universities in the Age of Artificial Intelligence: Toward Convergence between Policy Mandates and Infrastructure Capabilities of the Iranian Scholarly Web

Abdolreza Noroozi Chakoli

 *Editor-in-Chief of Scientometrics Research Journal: Scientific Bi-Quarterly of Shahed University, And Professor, Department of Information Science and Knowledge Studies, Shahed University, Tehran, Iran.*

Email: Noroozi@shahed.ac.ir



In recent decades, the dominant discourse in national science and technology policy has consistently underscored concepts such as scientific authority, science diplomacy, the ranking and internationalization of universities and journals, and enhancing the visibility of Iran's research achievements and scholarly publications. Accordingly, upstream policy documents, strategic roadmaps, and academic promotion regulations have persistently steered researchers and academic institutions toward active participation in the international scientific constellation and the dissemination of scholarship with global resonance; after all, elevating the standing and advancing the internationalization of periodicals and universities across all dimensions guarantees invaluable dividends for all stakeholders within the country's science and research landscape (Noroozi Chakoli, 2021-a, 2021-b, 2022).

Nevertheless, reflecting on the technical and functional attributes of the nation's scientific web reveals a conspicuous hiatus between the targets envisioned in policy frameworks and the empirical realities at the infrastructure and operational layers—a gap that warrants critical scrutiny and strategic recalibration. Addressing this issue within the domains of scientometrics, webometrics, and alternative metrics (Altmetrics) is of paramount importance because, in contemporary science assessment paradigms, web-based visibility serves as an indispensable prerequisite for achieving scholarly impact and citability (Noroozi Chakoli, 2011). Today, the global standing of researchers, scientific journals, and universities—all of which constitute the cornerstones and emblems of national scientific stature and authority—is governed more than ever by quantitative and qualitative indicators derived from data networks and web-based presence. In global university ranking systems (including the Webometrics Ranking, as well as citation- and reputation-based systems such as Leiden, Times Higher Education, and QS) (Aguillo et al., 2008), as well as in journal impact calculations, a substantial weight of evaluation matrices

is allocated to information retrieval capability, linkability, citation yield, and active engagement within the open scientific information ecosystem. When access to dissemination platforms is severed or impaired, the automated web crawlers and indexers of international abstracting systems and academic search engines (such as Google Scholar) are rendered incapable of indexing and crawling these corpora. Consequently, even the most rigorous scholarly output is omitted from scientometric computations—a vicious cycle that directly results in the decline of university rankings, the diminution of journal impact factors, and the erosion of the nation's key scientific indicators.

Currently, one of the most critical impediments to the global visibility of Iran's scientific publications lies in the enforcement of geographic access restrictions (geo-blocking) and the international inaccessibility of major national repositories tasked with organizing, managing, and disseminating bibliographic and scientific data. Authoritative platforms—such as the National Library and Archives of the Islamic Republic of Iran (NLAI), the GANJ system of the Iranian Research Institute for Information Science and Technology (IranDoc), the Scientific Information Database (SID), Magiran, Civilica, and the Comprehensive Humanities Portal—have frequently encountered structural disruptions or full overseas access blockades due to technical contingencies, data protection policies, or network infrastructure constraints. Such restrictions preclude international scholars and academic institutions from accessing the vast corpus of scientific knowledge generated and indexed in Iran—encompassing books, dissertations, journal articles, conference proceedings, and funded research projects. As a result, researchers, libraries, publishers, and scholarly repositories worldwide are deprived of access to the master record—namely, the official, authoritative bibliographic entry and documented identity of Iranian scholarly works—preventing them from linking to these sources in formal citations. Beyond dispossessing the global academic community of a rich knowledge repository, this breakdown drastically diminishes the capacity of national publications to influence international scientific workflows and accrue global citations; a circumstance that denies Iranian scholarship the opportunity for global uptake, citation accrual, and active participation in international knowledge generation. Consequently, an international researcher seeking to identify or cite an Iranian work is inadvertently funneled toward commercial platforms, online bookstores, or alternative, often volatile secondary sources instead of connecting directly to an authoritative national repository. Under such conditions, the nexus between the original scholarly work and its official bibliographic record is severed, substituting a commercial broker or vendor in the citation chain in lieu of a legitimate academic repository. This displacement undermines the bibliographic authority and integrity of Iranian scholarship within the international scientific domain, posing formidable barriers to its documentation, retrieval, and citability.

The gravity of this impediment becomes increasingly pronounced when juxtaposed against an analogous scenario involving the Library of Congress or global union catalogs such as WorldCat. Were international public access to the bibliographic records of these institutions compromised, thereby barring scholars from referencing and citing their cataloged entries, these entities would progressively forfeit their functional utility and international authority; for the prestige of any bibliographic repository hinges not solely on data accuracy and exhaustiveness, but equally upon its continuous accessibility, referential integrity, citability, and the linkability of its records. Viewed through this perspective, the inaccessibility of authoritative national infrastructures—such as the National Library and Archives, the GANJ system of IranDoc, SID, Magiran, Civilica, and

the Comprehensive Humanities Portal—cannot be dismissed as a mere technical defect or transient networking glitch. Rather, it represents an infrastructural rupture between Iranian intellectual output and the global discovery, documentation, and citation fabric, eroding the country’s authoritative standing within the global bibliographic information system.

The immediate consequence of this impediment—bearing profound scientometric ramifications—is the systemic marginalization of a major proportion of national research output, the exclusion of global scholars from engaging with primary Iranian literature, and the artificial compression of Iran’s authentic contribution to international scientific discourse. In an era dominated by ubiquitous artificial intelligence and large language models (LLMs), this predicament assumes even more complex and critical dimensions. Web scrapers, web crawlers, and autonomous AI agents tasked with indexing, synthesizing, and training upon human knowledge inevitably bypass these scholarly assets when confronted with impenetrable digital perimeters.

Through this operational dynamic, an empirical scientometric phenomenon emerges in the behavioral analysis of AI models that aligns directly with Zipf’s Principle of Least Effort (Zipf, 1949, as cited in Case & Given, 2016). Driven by programmatic constraints, AI architectures default to the most easily accessible, cost-effective, and retrievable data nodes rather than filtering and synthesizing the most authoritative, rigorous, and authentic scholarly sources. Within this framework, indigenous academic disciplines—most notably Iranian and Islamic history, culture, literature, art, and civilization—bear the heaviest burden. This transpires despite the fact that elevating the prestige and international authority of the Persian language in scholarly discourse has consistently occupied the highest echelon of national strategic directives, particularly under Macro-Strategy 9 of the Comprehensive Scientific Map of the Country (Secretariat of the Supreme Council of the Cultural Revolution, 2011). Nonetheless, the infrastructural isolation and invisibility of domestic repositories deprive a vast volume of seminal Persian-language works and genuine intellectual paradigms of the opportunity to participate in international discourse, exert scholarly influence, and garner global citations. Conversely, in this algorithmic landscape, when an Iranologist, Orientalist, or general user queries an AI system regarding historical or civilizational facts concerning Iran, the algorithmic agents—finding the primary national databases and bibliographic repositories inaccessible—inevitably bypass authentic domestic scholarship. Instead, they retrieve secondary, superficial, imprecise, and occasionally biased materials solely because they reside upon the visible, open web. In practice, this outcome not only impedes the realization of Persian as a language of scientific authority, but also breeds the hazard of unintended or systemic algorithmic distortion targeting national history, culture, and intellectual heritage.

Beyond the challenges of repository accessibility, another foundational impediment impairing the web-based visibility and citability of national scholarship stems from technical web engineering, architectural design, and the formatting conventions and persistence of Uniform Resource Locators (URLs). An examination of the digital ecosystem of Iranian academic websites, scholarly periodicals, and commercial distribution platforms (such as Gisoom, Ketabrah, Taaghche, Fidibo, Ketabnak, Ketabkhoon, Book Agency, and several university presses) reveals structural deficits that adversely impact the discovery and citability of scholarly assets:

1. *Character Encoding Challenges and Referential Fragility of Links*: A primary technical bottleneck in the digital tracking and sharing of Iranian scholarly works lies in the handling of Persian Unicode characters within URL conventions and web protocol standards. Directly embedding raw Persian strings into URI paths without executing standardized optimization mechanisms triggers percent-encoding, yielding excessively convoluted, unreadable, and brittle links (e.g., %D8%B9%D9%84%D9%85 representing the term *Elm* [Science]). In addition to completely compromising human readability, this phenomenon hinders link-sharing and direct citation to web-hosted academic texts. It routinely compels researchers to rely on third-party link-shortening services to manage unwieldy URLs; an intervention that introduces an unnecessary and unstable dependency layer, effectively undermining persistent trackability, seamless algorithmic processing, and the global visibility of scholarly works.
2. *Link Fragility and Link Rot*: A formidable barrier to international citability is the absence of robust, standardized persistent identification and tracking frameworks. Within digital scholarly communication, link persistence is imperative; a citation forged by an academic in 2024 must not terminate in an impassable dead end (broken link) by 2030. The failure to adopt persistent identifiers (PIDs)—such as DOIs and Handles—for monographs, articles, and scholarly resources, coupled with the legacy practice of reassigning a single transient URL to multiple heterogeneous works over time, severely compromises the integrity and citability of academic literature. While the international publishing ecosystem relies on persistent identifiers—such as DOI (for serial literature) and Handle or ISBN (for monographs)—many domestic systems continue to rely exclusively on volatile dynamic web addresses. Unlike a conventional URL that functions merely as a transient locator pointing to a physical storage address, Handle-based systems are grounded in the Digital Object Architecture (DOA), resolving directly to the digital entity or content itself. This architectural independence permits infrastructural restructuring and physical file migrations without breaking the established citation chain (Hisseine et al., 2024). Consequently, the Handle System empowers web administrators to modify directory structures or migrate servers without the risk of destroying scholarly citation trails. The absence of these identifiers not only disrupts digital trackability but also directly impairs visibility, thereby confronting the fundamental evaluative mandate of scientometrics with formidable methodological challenges in tracking international citations.

In global practice, leading platforms such as Google Books and Amazon maintain the perpetual integrity of an item's identifier across arbitrary infrastructural overhauls, as their systems are anchored to immutable, persistent identifiers. In contrast, on numerous domestic academic websites and publishing portals, the URLs of a monograph or journal article remains tied to ephemeral folder hierarchies or internal database keys. As a consequence, even the most modest architectural modification or server migration precipitates acute link rot—a terminal state wherein the URL documented in a bibliography degrades into a dead link that no longer resolves to the source document (Pishchik, 2026). This fragility does not merely obstruct scientific retrieval and citation workflows; it deeply compromises the referential validity and academic prestige of national scholarship in the eyes of the international research community.

Recommendations

To navigate toward an optimal state and transition Iran's scientific web into a dynamic, highly linked, and globally authoritative ecosystem, concerted action across two interrelated tiers- macro-strategic/policy and technical/infrastructural- is recommended:

A) At the Marco-Policy and Scientific Information Management Level

1. *Revising Geo-Access Management Protocols*: Drawing a rigorous architectural demarcation between confidential enterprise data and open bibliographic metadata and scientific

literature; and guaranteeing stable international access to primary scientific databases, national bibliographic repositories, and university portals via modern edge security layers, adaptive rate-limiting, and intelligent firewalls rather than blanket geographical blockades.

2. *Formulating and Promulgating Scientific Web and Linked Data Standards*: Implementing coordinated policy initiatives by governing authorities (such as the Supreme Council of the Cultural Revolution, the Ministry of Science, Research and Technology, and the Ministry of Health and Medical Education) mandating institutional adherence to international webometrics, semantic web, and Linked Data standards (Hall & O'Hara, 2009) across all universities, research institutes, and academic journals.
3. *Integrating Webometrics and Visibility Indicators into Academic Evaluation Schemes*: Elevating performance metrics governing link persistence, international open access availability, and automated discoverability within national periodic institutional audits, university rankings, and journal accreditation frameworks.

B) At the Technical Engineering and Digital Publishing Infrastructure Level

1. *Institutionalizing Persistent Identifiers (PIDs)*: Expanding the mandatory implementation of persistent identifiers, notably DOIs for all journal articles, conference papers, and institutional research reports, while systematically binding monographs and book chapters to Handle systems and standard ISBN architectures across institutional portals and publishers to eliminate link rot.
2. *Optimizing URI and Path Architectures*: Re-engineering the linking conventions of open journal systems (OJS) and academic repositories through standard semantic naming conventions, intelligent transliteration/slugs, or search-engine-friendly numeric keys in place of raw Unicode characters, thereby obviating percent-encoding artifacts and preserving hyperlink structural integrity.
3. *Ensuring Compatibility with Web Crawlers and AI Indexers*: Structuring institutional repositories with standardized academic metadata schemas (such as Dublin Core, Highwire Press, and Schema.org microdata) across all landing pages to facilitate automated harvesting, semantic parsing, and systematic indexing by academic search engines (e.g., Google Scholar) and autonomous AI retrieval agents.
4. *Reinforcing the Master Record Authority within National Bibliographic Systems*: Establishing standard programmatic cross-linking between distributed book portals and the definitive, canonical bibliographic record at the National Library and Archives of Iran, ensuring that international citations anchor directly to the verified national registry of record.

Conclusion and Concluding Remarks

An examination of the nexus connecting scientometrics, webometrics, and digital infrastructures demonstrates that elevating national scientific stature and projecting scholarly output onto the global stage can be profoundly accelerated by synergizing upstream strategic policies with cutting-edge scientific web standards. In the present epoch, optimizing maximum accessibility and discoverability serves as the cardinal prerequisite for citability and sustainable scientific impact—a paradigm that guarantees the authentic, rigorous representation of academic achievements within transnational bibliographic indices and scientometric analyses.

Upgrading technical infrastructures in alignment with Semantic Web specifications, alongside fostering secure and open access to primary scholarship, will chart novel pathways for representing the nation's intellectual heritage and the Persian language

across emerging technologies and artificial intelligence models. Consequently, elucidating the multidimensional facets of this domain unfolds an invigorating agenda of novel, constructive research inquiries for scholars, theorists, and investigators within scientometrics and webometrics:

- Through which operational and mathematical models can the expansion of global access to national academic repositories optimize the visibility rates, foreign citations, and bibliometric performance indicators of the country's universities and researchers?
- How can the rigorous implementation of web accessibility and data engineering principles maximize the retrieval efficacy of large language models in accurately synthesizing primary Persian scholarship and solidifying indigenous scientific authority within global platforms?
- What empirical role does the systemic deployment of persistent identifiers (such as Handle and DOI) across national academic ecosystems play in sustaining the citation chain, mitigating link decay, and enriching the national knowledge graph on an international scale?
- To what extent can the integration of semantic metadata standards (e.g., Dublin Core, Highwire Press, and Schema.org) augment the crawlability, indexing fidelity, and discoverability of Persian scientific output across global academic search engines?
- How can scientometric evaluation frameworks be redefined to assimilate webometric dimensions—such as link persistence, algorithmic discoverability, and web accessibility—as value-adding parameters within institutional accreditation, academic promotion systems, and university ranking methodologies?

پارادوکس رؤیت‌پذیری علم، رتبه‌بندی و بین‌المللی‌سازی دانشگاه‌ها در عصر هوش مصنوعی: به‌سوی همگرایی اسناد سیاستی و قابلیت‌های زیرساختی وب علمی کشور

عبدالرضا نوروزی چاکلی

id سردبیر پژوهش‌نامه علم‌سنجی، دوفصلنامه علمی دانشگاه شاهد و استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه شاهد، تهران، ایران.

Email: Noroozi@shahed.ac.ir

صفحه ۸-۱



در دهه‌های اخیر، گفتمان مسلط بر سیاست‌گذاری علم و فناوری کشور همواره بر مفاهیمی همچون مرجعیت علمی، دیپلماسی علمی، رتبه‌بندی و بین‌المللی‌سازی دانشگاه‌ها و مجلات، و افزایش مشاهده‌پذیری^۱ دستاوردهای پژوهشی و انتشارات علمی ایران تأکید داشته است. از همین رو، اسناد بالادستی، برنامه‌های راهبردی و آیین‌نامه‌های ارتقاء، همواره پژوهشگران و نهادهای دانشگاهی را به حضور فعال در منظومه علم بین‌الملل و انتشار آثاری با بازتاب بین‌المللی سوق داده‌اند؛ چرا که ارتقای جایگاه و بین‌المللی‌سازی دانشگاه‌ها و نشریات از هر منظر، متضمن دستاوردهای ارزشمندی برای تمامی ذی‌نفعان عرصه علم و پژوهش کشور است (نوروزی چاکلی، الف، ۱۴۰۰؛ ب- ۱۴۰۰؛ ۱۴۰۱).

با این حال، تأمل در مختصات فنی و کارکردی وب علمی کشور نشان می‌دهد که میان اهداف مورد انتظار در اسناد سیاستی و واقعیت‌های موجود در لایه زیرساخت‌ها و عمل، فاصله‌ای درخور توجه وجود دارد که نیازمند توجه و بازنگری است. اهمیت پرداختن به این مسئله در قلمروهای علم‌سنجی، وب‌سنجی و سنجه‌های جایگزین (آلت‌متریکس) از آن روست که در ساختار نوین ارزیابی علم، مشاهده‌پذیری در بستر وب پیش‌شرط بنیادین تحقق هرگونه اثرگذاری و استنادپذیری است (نوروزی چاکلی، ۱۳۹۰). امروزه جایگاه بین‌المللی پژوهشگران، نشریات علمی و دانشگاه‌ها- که همگی ارکان و نمادهای مرجعیت و اقتدار علمی کشور به شمار می‌روند- بیش از هر زمان دیگری تحت تأثیر شاخص‌های کمی و کیفی برخاسته از شبکه‌های داده‌ای و حضور وب‌محور است. در نظام‌های رتبه‌بندی جهانی دانشگاه‌ها (از جمله رتبه‌بندی وب‌متریکس و نظام‌های مبتنی بر استناد و اعتبار آکادمیک نظیر لایدن، تایمز و کیواس) (Aguillo et al., 2008) و نیز در محاسبات شاخص‌های تأثیر مجلات، بخش چشمگیری از امتیازها به توانایی بازیابی، میزان پیوندپذیری، نرخ استناد و حضور فعال انتشارات در اکوسیستم باز اطلاعات علمی اختصاص دارد. هنگامی که دسترسی به بسترهای انتشار قطع یا مختل باشد، ربات‌ها و خزنده‌های سامانه‌های نمایه‌سازی بین‌المللی و موتورهای کاوش علمی (مانند گوگل اسکالر)

قادر به نمایه و پایش این متون نخواهند بود؛ در نتیجه، حتی باکیفیت‌ترین تولیدات علمی نیز در محاسبات علم‌سنجی نادیده مانده و این چرخه معیوب، مستقیماً به افت رتبه دانشگاه‌ها، کاهش ضریب تأثیر نشریات و تضعیف شاخص‌های کلیدی علم کشور منجر می‌شود.

امروزه یکی از مهم‌ترین چالش‌های جدی در مسیر مشاهده‌پذیری جهانی انتشارات علمی ایران، اعمال محدودیت‌های جغرافیایی دسترسی^۱ و عدم دسترسی‌پذیری پایگاه‌های اصلی متولی سازماندهی، مدیریت و اشاعه اطلاعات علمی و کتابشناختی کشور برای کاربران، پژوهشگران و کاوشگران بین‌المللی است. سامانه‌های مرجعی نظیر سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران، سامانه گنج پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، پایگاه اطلاعات علمی جهاد دانشگاهی^۲، مگ‌ایران، سیویلیکا^۳ و پرتال جامع علوم انسانی در مقاطع متمادی و به دلایل فنی، سیاست‌های حفاظتی داده‌ها یا محدودیت‌های زیرساختی شبکه، از خارج از مرزهای جغرافیایی کشور مسدود یا با اختلالات ساختاری در دسترسی مواجه هستند. این محدودیت سبب می‌شود دسترسی پژوهشگران و دانشگاه‌های خارج از کشور به بخش اعظم منابع علمی تولید و نمایه سازی شده در ایران - نظیر کتاب‌ها، پایان‌نامه‌ها، مقاله‌های نشریات، مقاله‌های همایشی و طرح‌های پژوهشی - میسر نگردد. لذا پژوهشگران، کتابخانه‌ها، ناشران و سامانه‌های علمی در سراسر جهان خواهند توانست به رکورد مادر^۴، یعنی مدخل کتابشناختی رسمی و مرجع، و شناسنامه مستند کتاب‌ها و آثار علمی ایرانی دسترسی یابند و در استنادهای علمی خود به آن پیوند دهند؛ امری که علاوه بر محروم‌سازی جامعه جهانی از این ذخیره ارزشمند دانشی، امکان اثرگذاری آثار علمی داخلی بر جریان بین‌المللی علم و به تبع آن، استنادپذیری آثار علمی کشور را به شدت کاهش می‌دهد؛ رخدادی که فرصت بهره‌مندی جهانی، ارتقای ضریب استناد و ایفای نقش مؤثر در چرخه تولید دانش بین‌المللی را از علم کشور سلب می‌کند. در نتیجه، پژوهشگری که قصد شناسایی یا استناد به یک اثر ایرانی را دارد، به جای اتصال به مرجع ملی و معتبر، ناگزیر به استفاده از پیوندهای موجود در پایگاه‌های تجاری، فروشگاه‌های آنلاین کتاب یا دیگر منابع واسطه و گاه ناپایدار سوق داده می‌شود. در این حالت، پیوند میان اثر اصیل و مرجع کتابشناختی رسمی آن گسسته می‌شود و به جای یک منبع علمی و مرجع ملی، یک مرکز فروش یا واسطه تجاری در زنجیره استناد قرار می‌گیرد. این جابه‌جایی، اعتبار کتابشناختی آثار ایران را در محیط علمی بین‌المللی تضعیف کرده و فرایند مستندسازی، بازیابی و استناد به آن‌ها را با چالش مواجه می‌سازد.

اهمیت این مسئله زمانی روشن‌تر می‌شود که آن را با وضعیتی مشابه درباره کتابخانه کنگره یا پایگاه‌های جهانی مانند WorldCat مقایسه کنیم. اگر دسترسی عمومی و بین‌المللی به رکوردهای کتابشناختی این مراکز برقرار نباشد و پژوهشگران نتوانند به اطلاعات ثبت‌شده در آن‌ها ارجاع و استناد دهند، این مراکز به تدریج بخشی از کارکرد و مرجعیت بین‌المللی خود را از دست خواهند داد؛ زیرا اعتبار یک پایگاه کتابشناختی، افزون بر دقت و جامعیت داده‌ها، به دسترس‌پذیری، قابلیت ارجاع و استناد و امکان پیوندپذیری رکوردهای آن نیز وابسته است. از

1 . Geo-blocking
2 . SID
3 . Civilica
4 . Authority record

همین منظر، نبودِ دسترسی برای پژوهشگران بین‌المللی به سامانه‌های مرجعی نظیر سازمان اسناد و کتابخانه ملی، سامانه گنج پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، پایگاه اطلاعات علمی جهاد دانشگاهی، مگ‌ایران^۱، سیویلیکا، پرتال جامع علوم انسانی و مانند آن، صرفاً یک کاستی فنی یا ارتباطی نیست، بلکه موجب گسست میان آثار ایرانی و شبکه جهانی شناسایی، مستندسازی و استناد می‌شود و جایگاه مرجعیتی کشور را در نظام جهانی اطلاعات کتابشناختی تضعیف می‌کند.

پیامد بلافصل این انسداد که در حوزه علم‌سنجی از اهمیت ویژه‌ای برخوردار است، در حاشیه ماندن بخش عظیمی از بروندهای علمی کشور، محرومیت پژوهشگران جهانی از استناد به منابع اصیل ایرانی و تقلیل سهم واقعی ایران در تولید و تبادل دانش جهانی است. این مسئله در عصر فراگیری هوش مصنوعی و مدل‌های زبانی بزرگ، ابعادی به مراتب پیچیده‌تر و حیاتی‌تر به خود گرفته است. وب‌اسکراپرها، خزنده‌های وب^۲ و کارگزاران هوشمند هوش مصنوعی که وظیفه نمایه‌سازی، بازنمایی و آموزش دانش بشری را بر عهده دارند، در مواجهه با دیوارهای نفوذناپذیر شبکه‌ای، این منابع غنی را نادیده می‌انگارند.

در این فرایند، پدیده‌ای علم‌سنجی در تحلیل رفتار هوش مصنوعی پدیدار می‌شود که می‌توان آن را با/اصل کمترین کوشش زیریف^۳ منطبق دانست (Zipf, 1949, as cited in Case & Given, 2016); به‌طوری که سامانه‌های هوش مصنوعی به‌جای پالایش و ارائه بهترین، موثق‌ترین و دقیق‌ترین پاسخ، لزوماً به در دسترس‌ترین، کم‌هزینه‌ترین و بارگذاری‌پذیرترین پاسخ‌ها روی می‌آورند. در چنین ساختاری، حوزه‌های پژوهشی با هویت بومی و اصیل - نظیر تاریخ، فرهنگ، ادب، هنر و تمدن ایرانی و اسلامی - بیشترین آسیب را متحمل می‌شوند. این در حالی است که ارتقای جایگاه و مرجعیت زبان فارسی در عرصه تولید و تبادل علم جهانی، همواره در صدر سیاست‌ها و اسناد راهبردی نظام، به‌ویژه در راهبرد کلان ۹ نقشه جامع علمی کشور قرار داشته است (دبیرخانه شورای عالی انقلاب فرهنگی، ۱۳۸۹). با این حال، انسداد و عدم مشاهده‌پذیری پایگاه‌های بومی سبب می‌شود که حجم عظیمی از آثار فاخر فارسی‌زبان و اندیشه‌ها و نظریه‌پردازی‌های اصیل دانشمندان ایرانی که در آثار داخلی منتشر شده است، فرصت حضور در گفتمان جهانی، تأثیرگذاری علمی و کسب استندهای بین‌المللی را از دست بدهند. از سوی دیگر، در عصر الگوریتم‌ها، هنگامی که یک ایران‌پژوه، مستشرق یا حتی یک کاربر عادی در جهان در پی دستیابی به حقایق تاریخی و تمدنی ایران باشد، با مسدود بودن پایگاه‌های مادر و مخازن کتابشناختی دست‌اول کشور، مدل‌های هوش مصنوعی به‌ناچار از منابع موثق بومی چشم‌پوشی کرده و به منابع ثانویه، سطحی، غیردقیق و بعضاً سوگیرانه‌ای رجوع می‌کنند که صرفاً به دلیل مشاهده‌پذیر بودن در وب آزاد بازیابی شده‌اند. این رخداد در عمل نه‌تنها مانع از تحقق مرجعیت زبان فارسی می‌شود، بلکه خطر شکل‌گیری گونه‌ای از تحریف ناخواسته یا سازمان‌یافته الگوریتمی علیه تاریخ، فرهنگ و میراث علمی کشور را به همراه خواهد داشت.

افزون بر مسئله دسترسی‌پذیری به پایگاه‌ها، چالش بنیادین دیگری که رؤیت‌پذیری و استنادپذیری انتشارات

1 . Magiran

2 . Web Crawlers

3 . Zipf's Principle of Least Effort

پارادوکس رؤیت‌پذیری علم، رتبه‌بندی و بین‌المللی‌سازی دانشگاه‌ها در عصر هوش مصنوعی: ...

علمی کشور را در پهنه وب متأثر می‌سازد، به مهندسی و معماری فنی وب و الگوی نام‌گذاری و پایداری نشانگرهای منبع یکپارچه (URLها) بازمی‌گردد. بررسی زیست‌بوم وبگاه‌های دانشگاهی، نشریات، و کارگزاران و پلتفرم‌های عرضه کتاب و منابع علمی در ایران (نظیر گیسوم، کتابراه، طاقچه، فیدیو، کتابناک، کتابخون، آژانس کتاب و حتی برخی ناشران معتبر دانشگاهی) نشان می‌دهد که ساختار آدرس‌دهی و پیوند منابع در این درگاه‌ها با چند نارسایی ساختاری مواجه است که می‌تواند بر رؤیت‌پذیری و استنادپذیری انتشارات علمی کشور تأثیرگذار باشد:

(۱) چالش‌های رمزگذاری کاراکترهای فارسی و اختلال در ارجاع‌پذیری پیوندها: یکی از موانع فنی در ردیابی و اشتراک‌گذاری دیجیتال آثار علمی ایران، نحوه مواجهه با کاراکترهای فارسی در ساختار آدرس‌های اینترنتی^۱ و استانداردهای پروتکل‌های وب است. درج مستقیم کاراکترهای فارسی در مسیر نشانی‌های اینترنتی بدون رعایت استانداردهای فنی و پیاده‌سازی استانداردهای بهینه‌سازی، منجر به پدیده رمزگذاری درصدی^۲ و تولید پیوندهایی فوق‌العاده طولانی، نامفهوم و شکننده می‌شود (نظیر 84%D9%85%D9%B9%D8% به جای واژه «علم»). این پدیده علاوه بر سلب خوانایی انسانی، فرایند اشتراک‌گذاری و قابلیت استناددهی مستقیم به صفحات حاوی متون علمی را با دشواری روبه‌رو می‌سازد و حتی پژوهشگران را ناچار می‌سازد تا برای مدیریت این پیوندهای پیچیده، از سرویس‌های واسطه‌ای کوتاه‌کننده لینک^۳ استفاده کنند؛ اقدامی که با ایجاد یک لایه‌ی وابستگی غیرضروری و ناپایدار، عملاً قابلیت ردیابی مستمر، پردازش روان و رؤیت‌پذیری جهانی این آثار را در محیط وب تضعیف می‌نماید.

(۲) ناپایداری و زوال پیوندها: یکی از موانع جدی در مسیر استنادپذیری بین‌المللی، فقدان ساختار مستحکم برای شناسایی و ردیابی منابع است. در زیست‌بوم دیجیتال، پایداری پیوندها امری حیاتی است؛ به گونه‌ای که استاندارد یک پژوهشگر به یک کتاب یا مقاله در سال ۲۰۲۴، نباید در سال ۲۰۳۰ با پیوند شکسته^۴ رو به رو شود و به بن‌بست برسد. عدم استفاده از شناسه‌های پایدار (مانند DOI و Handle) در حوزه کتاب‌ها، مقالات و منابع بین‌المللی و تخصیص شناسه یک URL به آثار متعدد در گذر زمان، اعتبار استنادپذیری متون را مخدوش می‌کند. در شرایطی که در زیست‌بوم بین‌المللی، کتاب‌ها و منابع علمی با استفاده از شناسه‌های پایدار^۵ نظیر DOI (برای مقالات) و Handle و ISBN (برای کتاب‌ها) تعریف می‌شوند، در بسیاری از سامانه‌های داخلی، تنها به استفاده از آدرس‌های پویای وب^۶ اکتفا شده است. در حالی که آدرس‌های معمولی وب (URL) صرفاً به مکان^۷ فعلی فایل در سرور اشاره دارند و با کوچک‌ترین تغییر در وب‌سایت از کار می‌افتند، شناسه‌های پایدار (مانند Handle یا DOI) به ماهیت خود اثر متصل هستند. این یعنی حتی اگر ساختار وب‌سایت دانشگاه تغییر کند یا فایل‌ها به سرور دیگری منتقل شوند، لینک استناد همچنان بدون تغییر باقی مانده و پژوهشگر را به مقصد اصلی هدایت می‌کند (Hisseine et al., 2024). از این رو، بهره‌گیری از این شناسه‌ها به مدیران سایت‌ها اجازه می‌دهد بدون نگرانی از شکسته شدن زنجیره استنادات علمی، ساختار سایت خود را به‌روزرسانی کنند. فقدان این شناسه‌ها، نه تنها قابلیت ردیابی دیجیتال را مختل می‌کند، بلکه به

1 . URL
2 . Percent-encoding
3 . Link Shorteners
4 . Broken Link
5 . Persistent Identifiers
6 . Dynamic URLs
7 . Location

دلیل اثرگذاری مستقیم بر مشاهده‌پذیری، کارکرد بنیادین علم‌سنجی را نیز در تحلیل استندهای بین‌المللی با چالش مواجه می‌سازد.

در الگوی جهانی، پلتفرم‌های پیشرو نظیر Google Books یا Amazon، با وجود تغییر در زیرساخت‌های فنی خود، هیچ‌گاه پیوند شناسه‌ی یک اثر را از دست نمی‌دهند؛ زیرا این پلتفرم‌ها بر پایه یک شناسه‌ی ثابت و پایدار بنا شده‌اند که به عنوان 'نگرگاه' عمل می‌کند. اما در بسیاری از وبگاه‌های دانشگاهی و کارگزاران نشر در ایران، آدرس یک کتاب یا مقاله، وابسته به ساختار موقت پوشه‌ها یا پایگاه داده‌های داخلی است. این یعنی با کوچک‌ترین تغییر در معماری وبسایت یا جابه‌جایی فایل‌ها در سرور، با پدیده‌ی *زوال پیوند*^۲ روبرو می‌شویم؛ وضعیتی که در آن پیوند ارائه‌شده در فهرست منابع، به یک آدرس مرده تبدیل شده و دیگر به منبع اصلی منتهی نمی‌شود (Pishchyk, 2026). این ناپایداری در پیوندها، نه تنها فرآیند بازیابی علمی و استناددهی را مختل می‌کند، بلکه اعتبار استناد به متون را در نگاه پژوهشگران بین‌المللی به شدت مخدوش می‌سازد.

راهکارهای پیشنهادی

برای گذار به شرایطی مطلوب‌تر و تبدیل وب علمی ایران به زیست‌بومی پویا، پیوندپذیر و مرجع در تراز بین‌المللی، اهتمام به راهکارهای زیر در دو سطح راهبردی-سیاستی و فنی-اجرایی پیشنهاد می‌شود:

الف) در سطح سیاست‌گذاری و مدیریت کلان اطلاعات علمی

۱. بازنگری در سازوکارهای مدیریت دسترسی جغرافیایی: تفکیک دقیق میان داده‌های حساس و سازمانی و اطلاعات و فراداده‌های کتابشناختی عمومی و علمی، و تمهید دسترسی پایدار بین‌المللی به پایگاه‌های اطلاعات علمی مادر، سامانه‌های کتابشناسی ملی و پورتال‌های دانشگاهی از طریق به‌کارگیری لایه‌های نوین حفاظت شبکه و دیوارهای آتش هوشمند به‌منظور استمرار دسترسی کاربران و پژوهشگران خارجی.

۲. تدوین و ابلاغ چارچوب استانداردهای وب علمی و پیوندپذیری داده‌ها: سیاست‌گذاری منسجم از سوی مراجع ذی‌ربط (نظیر شورای عالی انقلاب فرهنگی، وزارت علوم، تحقیقات و فناوری و وزارت بهداشت، درمان و آموزش پزشکی) جهت رعایت استانداردهای جهانی وب‌سنجی، وب معنایی و داده‌های پیوندیافته^۳ (Hall & O'Hara, 2009) توسط دانشگاه‌ها، پژوهشگاه‌ها و مجلات علمی.

۳. گنجاندن شاخص‌های وب‌سنجی و رؤیت‌پذیری در نظام‌های ارزیابی: ارتقای سنجش‌های سنجش پایداری پیوندها، قابلیت دسترسی آزاد بین‌المللی و شاخص‌های کشف‌پذیری در نظام‌های رتبه‌بندی و اعتبارسنجی ادواری دانشگاه‌ها، نشریات و مؤسسات پژوهشی کشور.

ب) در سطح فنی و مهندسی زیرساخت‌های نشر دیجیتال

۱. توسعه و نهادینه‌سازی کاربرد شناسه‌های پایدار: گسترش بهره‌گیری از شناسه‌های ماندگار نظیر DOI برای کلیه مقالات، همایش‌ها و طرح‌های پژوهشی، و پیوند ساختارمند کتاب‌ها به شناسه‌های Handle و سامانه‌های متناظر با شابک^۴ در درگاه‌های دانشگاهی و پلتفرم‌های نشر به‌منظور پیشگیری از زوال پیوندها.

1 . Anchor
2 . Link Rot
3 . Linked Data
4 . ISBN

پارادوکس رؤیت‌پذیری علم، رتبه‌بندی و بین‌المللی‌سازی دانشگاه‌ها در عصر هوش مصنوعی: ...

۲. بهینه‌سازی معماری نشانی‌ها^۱: بازنگری در الگوی پیونددهی سامانه‌های مدیریت نشریات و وبگاه‌های علمی، با استفاده از نام‌گذاری‌های استاندارد، دگرنویسی^۲ هوشمند یا شناسه‌های عددی بهینه‌سازی‌شده برای موتورهای کاوش به‌جای درج مستقیم نویسه‌های خام، با هدف رفع پیامدهای رمزگذاری درصدی و حفظ یکپارچگی پیوندها.

۳. انطباق‌پذیری با خزنده‌های وب و سامانه‌های هوش مصنوعی^۳: درج متادیتاهای استاندارد و ساختاریافته آکادمیک (نظیر Schema.org, Dublin Core, Highwire Press) در ساختار فنی صفحات وبگاه‌های علمی جهت تسهیل بازیابی خودکار توسط موتورهای کاوش علمی (مانند Google Scholar) و کارگزاران هوش مصنوعی.

۴. تقویت مرجعیت رکورد مادر در پایگاه کتابشناسی ملی: ایجاد پیوندهای استاندارد میان سامانه‌های ارائه‌دهنده کتاب و مدخل کتابشناختی رسمی در سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران، به نحوی که استنادها و ارجاعات مستقیماً به شناسنامه معتبر و مرجع ملی متصل شوند.

جمع‌بندی و کلام آخر

واکاوی پیوند میان علم‌سنجی، وب‌سنجی و زیرساخت‌های دیجیتال نشان می‌دهد که ارتقای جایگاه علمی کشور و مشاهده‌پذیری آثار پژوهشگران در تراز جهانی، با بهره‌گیری هم‌افزا از اسناد بالادستی و به‌کارگیری استانداردهای نوین وب علمی شتابی دوچندان خواهد یافت. در عصر حاضر، تسهیل دسترسی‌پذیری و کشف‌پذیری^۴ حداکثری آثار، پیش‌نیاز اصلی استنادپذیری و زمینه‌ساز اثرگذاری پایدار علم است؛ رویکردی که تضمین‌کننده بازنمایی شایسته و دقیق دستاوردهای علمی و دانشگاهی در پایگاه‌های فراملی و محاسبات علم‌سنجی خواهد بود.

ارتقای زیرساخت‌های فنی همگام با استانداردهای وب معنایی، افزون بر توسعه دسترسی آزاد و امن به منابع اصیل علمی، افق‌های تازه‌ای برای معرفی جامع‌تر میراث علمی و زبان فارسی در فناوری‌های نوین و مدل‌های هوش مصنوعی فراهم می‌سازد. از این رو، تبیین ابعاد گوناگون این حوزه، بستری ارزشمند از مسائل و پرسش‌های علمی، بدیع و راهگشا را پیش روی استادان، صاحب‌نظران، پژوهشگران و دانشجویان حوزه علم‌سنجی و وب‌سنجی می‌گشاید:

- توسعه دسترسی جهانی به سامانه‌های علمی و دانشگاهی، چگونه و با چه مدلی می‌تواند نرخ رؤیت‌پذیری، ارجاعات برون‌مرزی و شاخص‌های استنادی پژوهشگران و دانشگاه‌های کشور را ارتقا بخشد؟
- به‌کارگیری اصول دسترسی‌پذیری و مهندسی داده‌ها چگونه می‌تواند کارایی مدل‌های زبانی بزرگ را در بازیابی دقیق منابع اصیل علمی فارسی و تقویت مرجعیت محتوای بومی در فضای بین‌المللی بیشینه‌سازی کند؟
- پیاده‌سازی گسترده شناسه‌های پایدار (نظیر Handle و DOI) در نشریات و سامانه‌های علمی کشور، چه نقشی در حفظ زنجیره ارجاعات، پایداری پیوندها و تقویت گراف دانش ملی در عرصه جهانی ایفا می‌کند؟

1 . URL Structure

2 . Transliteration

3 . AI & Academic SEO

4 . Discoverability

- ادغام استانداردهای فراداده‌ای وب معنایی (همچون Dublin Core, Highwire Press و Schema.org) تا چه میزان می‌تواند قابلیت خزش‌پذیری، نمایه‌سازی و حضور موثر تولیدات علمی فارسی را در موتورهای کاوش علمی جهان بهبود بخشد؟
- چگونه می‌توان با بازتعریف الگوهای ارزیابی علم‌سنجی، متغیرهای وب‌سنجی نظیر پایداری داده‌ها، کشف‌پذیری و میزان دسترسی‌پذیری وب را به‌عنوان مؤلفه‌های ارزش‌افزا در نظام‌های ترفیع، رتبه‌بندی و سنجش عملکرد علمی مؤسسات و دانشگاه‌ها گنجانده؟

فهرست منابع

- دبیرخانه شورای عالی انقلاب فرهنگی. (۱۳۸۹). *سند نقشه جامع علمی کشور*. دبیرخانه شورای عالی انقلاب فرهنگی. بازیابی شده در ۲۱ مرداد، ۱۴۰۵، از <https://www.irimc.org/Portals/0/PDF/ScientificMap.pdf>
- نوروزی چاکلی، ع. (۱۳۹۰). *آشنایی با علم‌سنجی: مبانی، مفاهیم، روابط و ریشه‌ها*. سازمان مطالعه و تدوین کتب علوم انسانی دانشگاه‌ها (سمت): مرکز تحقیق و توسعه علوم انسانی؛ دانشگاه شاهد. <https://samt.ac.ir/fa/book/99>
- نوروزی چاکلی، ع. (۱۴۰۱). سخن سردبیر: طبقه‌بندی موضوعی دقیق، پیش‌نیاز رتبه‌بندی صحیح مجلات و عدالت علمی. *پژوهش‌نامه علم‌سنجی*، ۸(۲)، ۱-۲. <https://doi.org/10.22070/rsci.2022.3945>
- نوروزی چاکلی، ع. (الف-۱۴۰۰). سخن سردبیر: مسئله رتبه‌بندی و سطح‌بندی دانشگاه‌ها و ضرورت تغییر رویکرد کلان آن به رویکردی خرد، تخصصی و کاربردی. *پژوهش‌نامه علم‌سنجی*، ۷(۱)، ۲-۱. <https://doi.org/10.22070/RSCI.2021.3048>
- نوروزی چاکلی، ع. (ب-۱۴۰۰). سخن سردبیر: رتبه‌بندی دانشگاه‌ها، از آرمان تا واقعیت. *پژوهش‌نامه علم‌سنجی*، ۷(۲)، ۲-۱. <https://doi.org/10.22070/RSCI.2021.3327>
- Aguillo, I. F., Ortega, J. L., & Fernández, M. (2008). Webometric Ranking of World Universities: Introduction, Methodology, and Future Developments. *Higher Education in Europe*, 33(2-3), 233-244. <https://doi.org/10.1080/03797720802254031>
- Case, D. O., Given, L. M. (2016). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. Emerald Group Publishing Limited. https://www.google.com/books/edition/_/F31VjgEACAAJ?hl=en&sa=X&ved=2ahUKEwio9YrJmNWWAxV1UaQEHQTIPGgQre8FegQIJBAJ
- Hall, W., O'Hara, K. (2009). Semantic Web. In R. Meyers (eds), *Encyclopedia of Complexity and Systems Science* (pp 8084-8104). Springer. https://doi.org/10.1007/978-0-387-30440-3_478
- Hisseine, M. A., Chen, D., Xiao, Y., & Alimo, P. K. (2024). A review of digital object architecture and handle system: Development, current applications and prospective. *Internet of Things*, 26, 101230. <https://doi.org/10.1016/j.iot.2024.101230>

- Noroozi Chakoli, A. (2011). *Introduction to Scientometrics: Foundations, Concepts, Relations, and Origins*. The Organization for Researching and Composing University Textbooks in the Islamic Values and the Humanities (SAMT); Institute for Research and Development in the Humanities; Shahed University. <https://samt.ac.ir/fa/book/99> [In Persian].
- Noroozi Chakoli, A. (2021-a). Note from the Editor-in-Chief: The issue of ranking and leveling universities and the necessity to change its approach from macro ranking to micro, specialized, and applied ranking. *Scientometrics Research Journal*, 7(1), 1-2. <https://doi.org/10.22070/RSCI.2021.3048> [In Persian].
- Noroozi Chakoli, A. (2021-b). Note from the Editor-in-Chief: Ranking of universities, from Ideal to reality. *Scientometrics Research Journal*, 7(2), 1-2. <https://doi.org/10.22070/RSCI.2021.3327> [In Persian].
- Noroozi Chakoli, A. (2022). Note from the Editor-in-Chief: Accurate subject classification: A prerequisite for correct journal ranking and scientific fairness. *Scientometrics Research Journal*, 8(2), 1-2. <https://doi.org/10.22070/rsci.2022.3945> [In Persian].
- Pishchyk, S. I. (2026, April 13-17). *The problem of link rot in digital information* [Conference presentation]. Current issues of economics and information technology: Collection of materials from the 62nd Scientific Conference of Graduate Students, Master's students and students of Belarusian State University of Informatics and Radioelectronics (BSUIR), Minsk, 2026. <https://libeldoc.bsuir.by/handle/123456789/65382>
- Secretariat of the Supreme Council of the Cultural Revolution. (2011). *Comprehensive Scientific Map of the Country*. Secretariat of the Supreme Council of the Cultural Revolution. Retrieved August 12, 2026, from <https://www.irimc.org/Portals/0/PDF/ScientificMap.pdf> [In Persian].