

Dominacija engleskog jezika u digitalnom dobu

The Dominance of English in the Digital Age

Dino-Dominik Magić¹, Ivana Vidak Teskera², Maja Resner³

¹Department of Tourism (Virovitica University of Applied Sciences, Matije Gupca 78, 33000,

dino.dominik.magic@vuv.hr)

²Department of Tourism (Virovitica University of Applied Sciences, Matije Gupca 78, 33 000,

ivana.vidak@vuv.hr)

³Department of Tourism (Virovitica University of Applied Sciences, Matije Gupca 78, 33 000,

maja.resner@vuv.hr)

Sažetak

U digitalnom dobu engleski jezik se etablirao kao prevladavajuća globalna lingua franca, duboko utječući na internetsku komunikaciju, obrazovanje i profesionalnu interakciju. Ovaj rad istražuje povijesne i sociolingvističke čimbenike koji su pridonijeli globalnoj dominaciji engleskog jezika, analizirajući njegovu sveprisutnu uporabu u digitalnim kontekstima i događaje koji su učvrstili njegov status. Dodatno, rad razmatra prednosti i nedostatke postojanja jednog dominantnog globalnog jezika, s posebnim naglaskom na implikacije za jezičnu raznolikost i ravnopravan pristup informacijama. Kako bi se procijenile percepcije unutar neengleskog govornog područja, provedena je anketa među 124 odraslih osoba u Hrvatskoj, u dobi od 18 do 65 godina, s uravnoteženom zastupljenošću spolova i profesionalnih pozadina. Sudionici su ispitani o svojim stavovima prema prisutnosti engleskog jezika u digitalnim medijima i njegovoj relevantnosti u njihovim profesionalnim sferama. Rezultati ukazuju na to da, iako se individualna angažiranost s engleskim jezikom razlikuje, velika većina prepoznaje njegovu dominantnu ulogu u digitalnoj eri i predviđa njegovu daljnju prisutnost.

Ključne riječi

digitalno doba, digitalni mediji, engleski jezik, jezična raznolikost

Abstract

In the digital age, English has solidified its position as the predominant global lingua franca, deeply influencing online communication, education, and professional interactions. This paper examines the historical and sociolinguistic factors that have contributed to the global dominance of English, analysing its pervasive use in digital contexts and the events that have solidified its status. Additionally, the study explores the advantages and disadvantages of having a single dominant global language, particularly in relation to linguistic diversity and equal access to information. To assess perceptions within a non-English-speaking context, a survey was conducted among 124 Croatian adults aged 18–65, encompassing a balanced representation of genders and professional backgrounds. Participants were queried about their views on the prevalence of English in digital media and its relevance within their professional spheres. The findings indicate that while individual engagement with English varies, a substantial majority acknowledge its dominant role in the digital era and anticipate its continued prominence.

Keywords

digital age, digital media, English language, linguistic diversity

Introduction

English has become the defining linguistic emblem of the contemporary era. Its reach permeates virtually every sphere of modern life—communication, science, business, technology, entertainment, and beyond—making it not merely a widespread means of interaction but a structuring force of global modernity itself. Nowhere is this more visible than in the digital sphere, where English operates as the principal medium through which information is produced, circulated, and consumed. The digital age, characterized by unprecedented technological acceleration and the transformation of how individuals engage with their environments and with one another, has further entrenched the centrality of English in global communication networks. In many respects, the language has become synonymous with modern global connectedness.

This linguistic predominance has had far-reaching implications. The near-universal reliance on English in scientific publishing, for instance—where the vast majority of scholarly contributions appear in English (Di Bitetti & Ferreras, 2016)—has facilitated the rapid and efficient exchange of ideas across disciplinary and national boundaries. Such visibility and accessibility have enabled an unprecedented degree of interdisciplinary collaboration. As a result, English has become a conduit through which scientific communities, regardless of their geographic dispersion, are able to participate in collective knowledge production.

The emergence of English as a global language is, however, not the product of a sudden linguistic shift; rather, it is the culmination of long-term historical, political, and cultural processes that can be traced back to the early medieval period (Gordin, 2015). The language's earliest forms—rooted in the Celtic and Anglo-Saxon linguistic landscape—were markedly different from present-day English. The Norman conquest of 1066 introduced profound lexical and structural influences from Norman French and Latin, transforming Old English into a language far removed from its earlier form. By the late Middle Ages, a series of institutional and sociopolitical shifts—including the Statute of Pleading of 1362, which mandated English in legal proceedings, and the ascension of Henry IV as the first English-speaking monarch—firmly established English as a language of administration, governance, and cultural life. By the close of the 15th century, English had expanded into all domains of public life.

The subsequent global spread of English during the early modern period was propelled by the political, economic, and military expansion of the British Empire. At its height, the Empire governed nearly one-quarter of the world's population and territory. English thus became embedded in administrative structures, educational systems, and economic institutions across regions as diverse as India, West Africa, North America, and Oceania (Philipson 1992). The empire's designation as “the empire on which the sun never sets” captures the vast geopolitical footprint through which English became entrenched in the linguistic ecologies of numerous societies (Allen, 2017).

Following the decline of British imperial power, the United States emerged as the principal driver of the consolidation of English as the global lingua franca. The economic dominance of the U.S., coupled with its immense cultural influence—particularly through film, television, music, and digital media—positioned English as the symbolic and practical language of global modernity (Brody, 2010; House, 2003). The United States also became the source of many of the technological advancements that shaped the digital world. ARPANET, developed in the 1960s under the auspices of the U.S. Department of Defense, constituted the foundational architecture of the Internet and, subsequently, the entire digital ecosystem (Isaacson 2014). Whether one refers to the present era as the Fourth Industrial Revolution, the information age, or the digital age, English remains the linguistic infrastructure upon which the global information society is based.

This global interconnectedness and the pervasiveness of digital technology have given rise to globalization, causing the barriers between cultures to become increasingly tenuous and blurred. The double-edged sword of this cultural convergence is that while it paves the way for a more unified world, it also leads to homogenization, resulting not only in a loss of cultural diversity but also a decline in linguistic diversity (Pimienta et al., 2009; Zeng & Yang, 2024; Balogun & Aruoture, 2024). Lesser-known languages are increasingly marginalized, with some nearing extinction (Matasović, 2011).

The aim of this paper is to investigate the extent to which English dominates the digital age by examining available data on its use and prevalence in digital contexts. The paper further examines the role of major technology companies, whose platforms and products shape global digital practices and often reinforce the

privileged status of English. Additionally, it assesses the benefits and drawbacks of the global dominance of English, its implications for linguistic diversity, and the emergent technological developments that may support the preservation of lesser-used languages. Before addressing these broader questions, the paper first considers the empirical evidence for the contemporary dominance of English in the digital sphere.

1. English language as the Lingua Franca of the Digital age

The contemporary digital landscape is overwhelmingly shaped by English, whose linguistic predominance has become deeply embedded within the architecture, protocols, and communicative practices of the Internet. Empirical data consistently indicate that English maintains a disproportionate presence online, far exceeding that of any other language. According to the *Digital 2024 Global Overview Report*, 52.1% of all web content is in English (DataReportal, 2024)—an unprecedented concentration given that English speakers represent only 18% of global Internet users. By contrast, Chinese users account for 18.8% of the online population, yet merely 1.3% of websites are in Chinese. This disparity accentuates the structural imbalance of linguistic representation in digital spaces, where English continues to function not merely as a dominant language but as a gatekeeping mechanism that mediates access to information and participation in global discourse (Phillipson, 1992).

Additional research corroborates these findings. Web Technology Surveys conducted by W3Techs indicate that as of 28 March 2025, English remains the leading language of the web, constituting 49.1% of all online content, followed distantly by Spanish at 6% (W3Techs, 2025). Studies examining social media further reveal parallel patterns. For example, Alshaabi et al. (2021), in their analysis of over 150 languages across nearly 118 billion tweets, determined that English was used most frequently—appearing in approximately 42 billion tweets—followed by Japanese (c. 21 billion), Spanish (c. 11 billion), and Arabic and Portuguese (c. 7 billion each) (Alshaabi et al., 2021). The sheer volume of English-language digital communication reinforces its position as the central linguistic medium of online interaction.

A similar trend emerges on multilingual, collaboratively produced platforms. As of March 2025, more than 11% of all Wikipedia articles—out of nearly

60 million entries—are written in English, placing it far ahead of other major language editions such as Cebuano and German (Wikimedia Statistics, 2025). While Wikipedia hosts content in hundreds of languages, the English edition remains the most expansive, frequently updated, and well-resourced, serving as the primary reference point for global users.

To illustrate, Table 1 below presents data from the Digital 2024 report on the top 20 languages used on the web:

Table 1. Digital 2024 Global Overview Report: The Most Common Languages for Web Content

	Language	Share of websites	Share of population
1	English	52.1%	18.0%
2	Spanish	5.5%	6.9%
3	German	4.8%	1.6%
4	Russian	4.5%	3.2%
5	Japanese	4.3%	1.5%
6	French	4.3%	3.8%
7	Portuguese	3.1%	3.3%
8	Italian	2.3%	0.8%
9	Turkish	2.1%	1.1%
10	Dutch/Flemish	1.8%	0.3%
11	Polish	1.6%	0.5%
12	Persian	1.5%	1.0%
13	Chinese	1.3%	18.8%
14	Vietnamese	1.2%	1.1%
15	Indonesian	1.1%	2.5%
16	Czech	0.9%	0.2%
17	Korean	0.8%	1.0%
18	Ukrainian	0.6%	0.4%
19	Arabic	0.6%	5.2%
20	Greek	0.5%	0.2%
	Total	95.9%	71.2%

Source: The authors, according to the *Digital 2024 Global Overview Report*

Despite its overwhelming presence, longitudinal data suggest that the proportion of English-language content online has been gradually declining. As shown in Table 2, the share of content in English fell from approximately 56.8% in 2020 to 49.1% by March 2025 (W3Techs Web Technology Surveys, 2025). This downward trend does not indicate a weakening of the dominance of English; rather, it reflects the increasing digital participation of speakers of other languages, as well as the diversification of content creation across the global Internet. Languages such as Spanish, German, Japanese, Portuguese, and Italian have all experienced notable growth in digital representation.

Nonetheless, English remains the unchallenged linguistic hegemon of the digital sphere.¹

Table 2. Historical yearly trends in the usage statistics of content languages for websites

Language	2020	2021	2022	2023	2024	2025 Jan	2025 Mar
English	56.8%	60.4%	63.6%	57.7%	52.2%	49.3%	49.1%
Spanish	4.6%	4.0%	3.6%	4.5%	5.5%	6.0%	6.0%
German	4.1%	2.4%	2.0%	3.9%	4.7%	5.6%	5.8%
Japanese	2.9%	2.1%	1.9%	3.2%	4.3%	5.1%	5.1%
French	3.3%	2.6%	2.5%	3.9%	4.3%	4.4%	4.5%
Portuguese	2.5%	1.3%	0.7%	1.8%	3.1%	3.8%	3.9%
Russian	7.6%	8.6%	7.0%	5.3%	4.5%	3.9%	3.8%
Italian	1.5%	0.8%	0.7%	1.7%	2.3%	2.7%	2.8%
Dutch, Flemish	0.9%	0.6%	0.5%	1.2%	1.8%	2.2%	2.2%
Polish	1.3%	0.6%	0.5%	1.2%	1.6%	1.8%	1.8%
Turkish	2.2%	3.6%	3.9%	2.6%	2.1%	1.8%	1.7%
Persian	2.4%	3.0%	3.5%	2.2%	1.5%	1.2%	1.2%
Chinese	1.4%	1.4%	1.3%	1.6%	1.3%	1.2%	1.1%
Vietnamese	1.0%	1.7%	1.9%	1.4%	1.2%	1.1%	1.1%
Indonesian	0.5%	0.7%	0.8%	0.6%	1.1%	1.1%	1.1%
Czech	0.7%	0.3%	0.2%	0.6%	0.9%	1.0%	1.0%
Korean	0.8%	0.6%	0.5%	0.7%	0.8%	0.8%	0.8%
Ukrainian	0.3%	0.4%	0.6%	0.6%	0.6%	0.6%	0.6%
Hungarian	0.4%	0.2%	0.1%	0.3%	0.5%	0.6%	0.6%

Source: The authors, according to W3Techs Web Technology surveys

An older study conducted by Daniel Pimienta et al., published by UNESCO in 2009, analysed a 12-year research project that aimed to measure linguistic diversity on the Internet. The study's main goal was to challenge the widely held belief at the time that English comprised 80% of all web content. While the data may be outdated, it clearly illustrates the decline in English-language web content, which fell from 75–80% in the late 1990s to around 45% by 2008. Meanwhile, other languages such as Spanish, French, German, Italian, Portuguese, and Romanian increased their relative presence on the web.

Taken together, these data reveal a complex but unmistakable reality: English remains the principal language of the digital world, both structurally and functionally. Its hegemony is not merely the consequence of speaker population sizes; it is rooted in the historical development of Internet technologies, the global influence of Anglophone media, and the entrenched use of English in scientific, technological, and professional domains. The following section considers one of the most consequential domains of this influence: the dominance of English in programming and technological innovation.

1.1. The Dominance of English in Programming and Technology

The predominance of English in the digital realm is inseparable from its foundational role in the development of computing and information technologies. The emergence of English as the linguistic medium of programming, technical documentation, and software development is not the result of an explicit global consensus, but rather a historical consequence of the fact that many seminal computational innovations originated in Anglophone contexts (Gordin, 2015). From the earliest conceptualizations of programmable machines to the development of modern operating systems, networking protocols, and programming languages, English has served as the conceptual and linguistic substrate upon which the architecture of contemporary computing was built.

While computers ultimately interpret machine code—binary sequences consisting of 1s and 0s—the languages used by humans to instruct machines have, since the inception of high-level programming, drawn almost entirely from English lexicon and syntactic conventions. Whether one examines early languages such as FORTRAN, COBOL, and BASIC, or contemporary languages including Python, Java, C++, and JavaScript, English remains embedded in their command structures, reserved keywords, operator nomenclature, and even error messaging (Ammon, 2001). Although programming languages do not constitute natural languages in the strict linguistic sense, their structures are nonetheless modelled overwhelmingly on English vocabulary and logic. Thus, every programmer—irrespective of linguistic background—must engage with English to access the fundamental tools of computation.

¹ The website contains data from 2014 onwards. The table only shows data from the last five years due to the lack of space.

The same applies to the infrastructure of the Internet and digital communication. HTML, CSS, DNS, TCP/IP protocols, and the vast majority of web-related syntaxes employ English terminology as their basis. Whether users navigate a website written in German, French, or Thai, its underlying structure is articulated through English-derived technical grammar. This linguistic asymmetry reinforces the centrality of English not only as a communicative tool but as the operative language of the digital ecosystem. The dominance of English also extends to technical documentation, professional standards, and scientific publications within technology-related fields. Although there is no single global repository for technical documentation, estimates indicate that nearly 90% of all technical papers are written in English (Ammon, 2001). Major technology companies—including Microsoft, Google, Apple, and IBM—publicly acknowledge English as their primary language for the creation and dissemination of technical materials. Moreover, over 95% of international IT standards are published in English prior to translation into any other language, according to the annual reports of the International Organization for Standardization (ISO). Such structural centralization ensures that English remains the normative point of access for innovation, standardization, and professional knowledge. Taken together, these factors render English the operational, conceptual, and symbolic language of the technological world. From software development to networking, cybersecurity, data science, and artificial intelligence, English constitutes the common linguistic denominator through which systems operate and practitioners communicate. The question that arises, therefore, is not whether English is dominant—its dominance is axiomatic—but whether the existence of a single global linguistic centre is beneficial or detrimental for societies navigating the complexities of the digital age. The following section explores the potential advantages of such dominance.

2. Advantages of English Language Dominance in the Digital Age

The characterization of English as a global language may appear self-evident, yet this perception warrants deeper examination. While awareness of English is widespread, global fluency levels vary considerably. Nevertheless, English maintains official status in more than seventy countries—surpassing any other language—and functions as a compulsory or primary foreign language in the educational systems of numerous states, including large non-Anglophone nations such as China, Germany, Spain, and Egypt

(Crystal, 2003). Its presence within formal education systems contributes significantly to its global reach, embedding English proficiency within the linguistic repertoires of successive generations.

Beyond the classroom, English permeates cultural, professional, and technological domains to a degree unmatched by any other language. English-language entertainment—films, music, television series, and digital content—has become deeply integrated into the cultural landscapes of societies worldwide (Montgomery & Crystal, 2013). This omnipresence facilitates linguistic borrowing, the diffusion of Anglophone vocabulary, and the normalization of English expressions in everyday speech. Moreover, as Kisery (2019) observes, nearly 60% of all translated texts involve English as either the source or target language, underscoring its centrality within global knowledge exchange networks. In everyday practice, this means that simple acts of communication—such as greeting someone with “hello” in an urban centre anywhere from Europe to Asia—are likely to be understood, reflecting the language’s broad communicative reach.

These developments lead to a pivotal question: *Should the existence of a single dominant global language be considered desirable?* The answer, though complex, reveals several significant advantages associated with linguistic convergence.

One of the most salient benefits concerns the facilitation of communication across linguistic and cultural boundaries. A shared language enhances diplomatic efficiency by reducing or eliminating the need for extensive translation and interpretation infrastructures. In an increasingly interconnected international system, this logistical simplification has substantial implications. For instance, the United Nations—comprising nearly 190 member states—relies on coordinated multilingual communication, yet English remains the primary working language across most committees and negotiations. The presence of English as a common linguistic denominator expedites political processes, mediates cross-cultural exchanges, and fosters more immediate collaboration.

Another advantage relates to the dissemination of scientific and technical knowledge. A global lingua franca enables the rapid circulation of research, allowing scholars and practitioners in disparate regions to engage with shared discourses. Scientific inquiry—especially in fields such as engineering, computer science, medicine, and environmental research—relies on the accessibility of up-to-date findings. English serves as the medium through which this global conversation unfolds. Without a widely understood linguistic platform, the exchange of

discoveries between, for example, a physicist in Germany and colleagues in Japan would be considerably constrained. Montgomery and Crystal (2013) argue that the existence of a single global scientific language accelerates both innovation and interdisciplinary engagement, strengthening the collective capacity for technological advancement. The benefits of English extend beyond academia into the professional sphere. In multinational corporations, even a basic command of English reduces operational costs by minimizing reliance on translation services (Zeng & Yang, 2024). This efficiency accelerates decision-making, improves organisational integration, and enhances competitiveness in global markets. English proficiency also increases labour mobility, enabling workers to seek employment opportunities beyond their national borders and to participate in the growing landscape of remote and hybrid work. As highlighted in the *EF English Proficiency Index (2024)*, the ability to operate professionally in English is now considered a core competency for participation in globalized labour markets. This is succinctly expressed in the EF English Proficiency Index (2024):

“Enhancing English proficiency lowers barriers to innovation by facilitating access to information and enabling professional collaboration between experts from diverse backgrounds. Whether this innovation occurs within the same office or between continents, freeing up the flow of ideas is efficient.”

Empirical evidence strongly supports these observations. According to Pearson’s *How English Empowers Your Work and Career report (2024)*², 85% of surveyed professionals indicated that English plays an important role in their current employment, and 88% expect its importance to increase over the next five years. These findings reinforce the interpretation that English functions as an economic resource: a tool that opens access to employment, professional growth, and international collaboration.

Taken together, these advantages highlight the pragmatic and structural benefits of a dominant global language. English serves as a mechanism for efficient communication, a facilitator of knowledge sharing, and a catalyst for economic and professional mobility. Yet the concentration of linguistic power also raises significant concerns. The following section explores the potential drawbacks and sociocultural risks associated with the global ascendancy of English.

3. Disadvantages and Challenges of English Language Dominance in the Digital Age

While communication remains the fundamental function of language, its significance extends far beyond the exchange of information. Language shapes identity, mediates social belonging, and enables the articulation of ideas—arguably the most transformative force available to human societies. Ideas construct communities, institutions, values, and worldviews. Thus, linguistic forms are not neutral vehicles but culturally and historically situated systems of meaning. For this reason, the global dominance of any single language inevitably invites critical reflection (Phillipson, 1992).

Scholars such as George Steiner, Walter Benjamin, and Friedrich Schleiermacher have articulated the profound difficulties inherent in translation, emphasising that linguistic structures are embedded within cultural, historical, and conceptual worlds that resist direct equivalence. Their collective work argues that each language contains untranslatable elements—semantic, idiomatic, and symbolic features that cannot be fully rendered in another tongue (Schleiermacher, 1813/2012; Benjamin, 1923/1997; Steiner, 1975). This notion of “inherent untranslatability” underscores the idea that language is not simply a tool for describing reality but a lens through which reality is perceived, interpreted, and experienced.

From a broader linguistic and anthropological perspective, languages encode the distinct worldviews of their speakers. Humboldt’s seminal reflections on *Weltansicht*—the worldview inherent in each language—along with subsequent refinements in linguistic relativism by Sapir (1929), Whorf (1956), and Duranti (1997), reinforce the argument that languages embody culturally specific epistemologies. Through idioms, metaphors, and figures of speech, a linguistic community preserves collective memory and historical consciousness (Lakoff & Johnson, 1980). A living language, shaped continuously through its speakers, evolves alongside its community (Labov, 1994; Milroy & Milroy, 1999). Conversely, when a language declines or dies, it extinguishes a unique mode of thought, a cultural identity, and an irreplaceable body of knowledge (Crystal, 2000; Harrison, 2007). In this sense, linguistic loss signifies not merely the disappearance

² <https://www.linkedin.com/pulse/how-english-empowers-your-work-career-pearson-report-part-1-30bxf/>

of a communication system but the erosion of cultural plurality.

The dominance of English in the digital age accelerates pressures on linguistic diversity. David Crystal (2003) warns of “linguistic complacency,” wherein the global power of English diminishes speakers’ motivation to learn other languages, reducing multilingualism and fostering an implicit hierarchy of linguistic value (Crystal, 2003). As English becomes the default medium in digital communication, international business, and global culture, other languages are increasingly relegated to local or informal domains. Crystal cites a European survey from 1996 showing significant disparities in foreign-language competency: 90% of business executives in Belgium, the Netherlands, Luxembourg, and Greece knew at least one foreign language, compared to only 38% in the United Kingdom. Such unequal linguistic expectations persist in the present era, further entrenching English as a privileged linguistic resource.

Among younger generations, English often acquires a prestige status, associated with modernity, opportunity, and global engagement. However, this symbolic elevation carries ambivalent consequences. The widespread consumption of Anglophone media and digital content not only shapes linguistic preferences but influences values, aesthetics, and cultural norms. In some contexts, such as Nigeria, English has become the dominant language of youth cultural expression, largely because digital platforms default to English and because online participation often presupposes at least partial English proficiency (Balogun & Arouture, 2024). Such trends contribute to cultural homogenisation, wherein local linguistic practices and cultural identities face displacement by globalised Anglophone norms.

The rise of large language models (LLMs) introduces additional challenges. A series of studies by Sourati et al. (2025) demonstrates that LLM-generated content tends to favour dominant linguistic patterns, thereby contributing to the standardisation of expression and the attenuation of linguistic variation. Their findings highlight measurable declines in lexical diversity, syntactic flexibility, and idiomatic expressiveness, with LLMs implicitly reproducing the normative conventions of high-resource languages—particularly English. Such developments risk accelerating linguistic convergence, further marginalising lesser-spoken languages and reducing the richness of global linguistic ecosystems.

Yet digital technologies also contain the potential to support linguistic preservation—a theme examined in the following section. While English’s dominance

imposes structural and cultural pressures on minority languages, the digital age simultaneously offers tools capable of documenting, revitalising, and sustaining endangered linguistic heritage. The tension between linguistic homogenisation and preservation is therefore one of the central linguistic paradoxes of the digital era.

4. Initiatives to preserve linguistic diversity

Although emerging technologies—particularly large language models (LLMs)—pose notable risks to linguistic diversity, they simultaneously offer unprecedented opportunities for the preservation, revitalisation, and documentation of low-resource and endangered languages. This section examines the technological advancements and institutional initiatives that seek to counterbalance the homogenising effects of digital communication by leveraging the digital medium itself as a tool for linguistic conservation.

One of the most significant developments in recent years has been the rapid improvement of digital translation technologies. Modern machine translation systems are increasingly capable of capturing contextual meaning, syntactic nuance, and cultural references, representing a substantial evolution from earlier rule-based and statistical approaches. As Mahajan (2025) argues, contemporary artificial intelligence—particularly neural architectures—has transformed translation into a dynamic, context-aware process. Neural machine translation (NMT), for instance, utilises deep learning models such as recurrent neural networks and transformer-based architectures to interpret context and semantic relationships, thereby improving the accuracy of translations across a broad spectrum of linguistic structures. The key developments in this field of machine translation are:

1. **Neural machine translation (NMT)**- allows systems to learn from vast multilingual corpora, enabling models to generalise meaning rather than rely on phrase-level memorisation. (Gu et al., 2017)
2. **Contextual embeddings**- capture the meaning of words in relation to their immediate linguistic environment, improving translation of idioms, slang, metaphors, and non-standard expressions. (Mahajan, 2025).
3. **Multilingual and Zero-Shot Learning**- techniques enable models trained on specific language pairs to generate reasonably accurate translations for pairs on which they

have not been directly trained—a major breakthrough for low-resource languages. (Johnson et al., 2017)

4. **Real-time optimization**- such as quantisation, model pruning, and knowledge distillation significantly reduce computational demands, facilitating real-time translation in mobile and low-bandwidth environments (Gu et al., 2017).
5. **Adaptability to social media discourse**- permits LLMs to manage informal registers, code-switching, emojis, and emerging linguistic phenomena characteristic of digital communication (Mahajan, 2025).

The measurable improvement of machine translation is perhaps best demonstrated by Facebook's adoption of NMT, which increased its BLEU (Bilingual Evaluation Understudy) score by approximately 11%, reaching an informal accuracy rate of around 61%. (Gu et al., 2017). More importantly, the iterative nature of machine learning ensures that translation models continue to improve as they receive additional multilingual input. Notably, Johnson et al.'s (2017) seminal experiment with Google's Multilingual Neural Machine Translation System demonstrated that a model trained on English–Spanish and English–Portuguese pairs could translate directly between Spanish and Portuguese, despite never having been explicitly trained on that linguistic pair. Such findings underscore the transformative potential of AI systems in supporting languages with limited digital presence.

These advances not only enhance translation accuracy but also promote linguistic authenticity by enabling models to reproduce linguistic and cultural nuances with greater fidelity (Spector, 2025). For communities whose languages lack sufficient institutional support, such technologies offer vital tools for preserving linguistic identity and transmitting cultural knowledge across generations.

In addition to improvements in translation, digital connectivity offers new opportunities for speakers of

low-resource languages to collaborate, teach, and engage with wider communities. Virtual classrooms, online forums, and social media platforms can facilitate cross-border linguistic exchange, while digital repositories allow communities to document endangered languages in written, spoken, and audiovisual form. Despite these affordances, challenges remain: many communities lack equal access to digital tools, adequate internet infrastructure, or the digital literacy needed to fully participate in online documentation projects. This “digital access gap” complicates efforts to preserve linguistic diversity through technologically mediated means (van Dijk, 2006).

Given the global scale of language endangerment, various organisations and institutions have developed targeted initiatives to support linguistic preservation. Stanford University's SILICON initiative (Stanford Initiative on Language Inclusion and Conservation in Old and New Media) represents a notable example of cross-sector collaboration, bringing together UNESCO, Unicode, technology companies, and linguistic communities to promote the inclusion of low-resource languages in digital environments (Stanford News (2025). The Endangered Languages Archive (ELAR)³, housed at the Berlin-Brandenburg Academy of Sciences and Humanities, constitutes a major digital repository containing nearly 550 collections from over 70 countries. ELAR provides open-access audio and video recordings, orthographic materials, and cultural documentation, thereby supporting both academic research and community-led revitalisation efforts.

Wikimedia's *Lingua Libre*⁴ project similarly contributes to digital linguistic preservation by enabling rapid, user-friendly recording of words and phrases in more than 250 languages, creating open-access resources useful for language learners, researchers, and speech technologies. Other significant initiatives include FirstVoices⁵, an open-source platform supporting Indigenous language revitalisation; Aikuma⁶, an

³ <https://www.elararchive.org/>

⁴ https://en.wikipedia.org/wiki/Lingua_Libre

⁵ <https://www.firstvoices.com/>

⁶ <https://www.aikuma.org/>

Android application designed for collecting speech recordings with time-aligned translations, and Mozilla's *Common Voice* project⁷, which crowdsources global speech corpora to build open, linguistically diverse datasets for speech recognition technologies.

These initiatives represent only a portion of the global efforts currently underway, yet collectively they demonstrate the growing recognition of digital technology as both a threat to and an instrument for linguistic diversity. Ultimately, the digital age need not be synonymous with linguistic homogenisation. If deployed conscientiously, technological tools can support the preservation and revitalisation of endangered languages, ensuring that linguistic plurality remains a defining feature of global cultural heritage.

5. Research Methodology

To examine how individuals perceive the dominance of English in the digital age, this study employed an empirical research design based on a structured survey. The objective of the research was twofold: first, to assess Croatian respondents' awareness of English-language prevalence in digital environments; and second, to evaluate whether they consider this dominance beneficial or detrimental to contemporary society.

The primary data collection instrument was an online questionnaire administered via Google Forms. This platform was selected due to its accessibility, ease of dissemination, and suitability for reaching a geographically dispersed sample. The survey yielded a total of 124 valid responses (N = 124). The sample comprised adult participants residing in Croatia, aged between 18 and 65, representing a broad range of professional and educational backgrounds. Although the sample is not statistically representative of the entire population of Croatia, it provides meaningful insights into attitudes within a multilingual European context.

The questionnaire consisted of two major components. The first section collected socio-demographic information, including age, gender, educational attainment, and professional sector. These data enabled the interpretation of linguistic attitudes within specific demographic and occupational profiles.

The second section focused on respondents' perceptions and experiences regarding the role of English in digital communication, online learning, professional advancement, and everyday information consumption. Items were designed to measure familiarity with English-language digital content, self-reported online language behaviour, and evaluations of English as a dominant global medium.

The survey relied on descriptive statistical analysis to interpret the results. Frequencies and percentages were used to summarise participants' responses, identify general trends, and illuminate patterns in digital language use and attitudes toward English. Given the exploratory nature of the study, descriptive methods were deemed most appropriate for capturing the breadth of respondents' experiences without imposing assumptions about causality.

The following section presents the demographic characteristics of the respondents, followed by a detailed analysis of the survey's main findings.

5.1. Respondents' socio-demographic indicators

Below is a table of the socio-demographic indicators of the respondents:

Table 3 Socio-demographic indicators of respondents

	Number of Respondents	%
Gender	124	100%
Male	41	33.06%
Female	80	64.52%
Doesn't want to say	3	2.4%
Age of the respondents	122	98.39%
18-25	35	28.23%
26-35	34	27.42%
36-50	36	29.03%
51-older	17	13.71%
Education level	124	100.00%
Elementary school	2	1.61%
Secondary school	41	33.06%
Bachelor's degree	20	16.13%
Master's degree	57	45.97%
Doctoral degree	4	3.23%

⁷ <https://commonvoice.mozilla.org/>

Profession		
	123	99.19%
IT sector	20	16.13%
Education	18	14.52%
Translation and language services	4	3.23%
Marketing and communication	3	2.42%
Medicine and healthcare	8	6.45%
Social sciences	15	12.10%
Technical sciences	17	13.71%
Economy and business	37	29.84%
Retired	1	0.81%

Source: The authors

The demographic profile of the sample reveals that the typical respondent was female, aged between 36 and 50, held a master's degree, and was employed in the economy and business sector. Two respondents did not disclose their age, and one did not indicate their profession; these responses were omitted from the corresponding analyses. The overall distribution demonstrates a relatively balanced sample in terms of age groups, while education levels and professional fields show expected variations given the recruitment methods. The next section presents the results of the survey and provides an analytical interpretation of the respondents' answers.

5.2. Survey result analysis

The survey began by examining respondents' language use in digital environments. The first question asked participants which language they most frequently use when browsing online. The results are presented below:

Table 4 Answers to: "Which language do you use the most when browsing online?"

Language	Number of respondents	%
English	42	33.87%
Croatian	81	65.32%
Other	1	0.81%

Source: The authors

Given that all respondents were Croatian nationals, it is unsurprising that a majority reported using their mother tongue while navigating the Internet. Nevertheless, a noteworthy proportion—over one-third—indicated English as their primary browsing language. This suggests that English exerts a considerable influence not only on the global digital sphere but also on the online practices of Croatian users. Such a finding aligns with broader trends in multilingual contexts where English frequently supplements or, in some domains, surpasses native-

language usage due to the availability and perceived quality of English-language digital content.

The second question asked respondents which language they prefer when learning new skills online. The distribution of answers was as follows:

Table 5 Answers to: "Which language do you prefer when learning new skills online?"

Language	Number of respondents	%
English	56	45.16%
Croatian	66	53.23%
Other	2	1.61%

Source: The authors

Slightly less than half of the respondents preferred English for online learning, which represents a modest discrepancy compared to the previous question. This divergence likely stems from the fact that English dominates instructional resources and online courses, making it substantially easier for learners to find high-quality materials in English than in Croatian. The preference for native-language learning may reflect comfort or accessibility, but the substantial proportion favouring English underscores its central role in shaping digital learning environments.

The third question assessed how much of the Internet content consumed by respondents is in English. The findings are displayed below:

Table 6 Answers to: "How much of the Internet content that you consume is in English?"

Percentage of the content	Number of respondents	%
Less than 25%	32	25.8%
25%-50%	36	29%
50%-75%	32	25.8%
More than 75%	24	19.4%

Source: The authors

The distribution of responses is relatively even across categories. When grouped into broader bands, however, the results indicate that 54.8% of respondents consume less than half of their online content in English, while 45.2% consume more than half. This suggests that while Croatian remains dominant for general browsing purposes, English constitutes a significant portion of respondents' overall media consumption, reflecting its strong presence in global digital content ecosystems.

The next question evaluated respondents' perceptions of linguistic dominance online by asking whether they consider English to be the most dominant language of

the Internet and digital technology. The results are shown below:

Table 7 Answers to: Is English the most dominant language of the Internet and of digital technology?

Response	Number of respondents	%
Yes	113	91.1%
No	5	4%
Not sure	6	4.8%

Source: The authors

An overwhelming majority (91.1%) affirmed that English is indeed the dominant language of the digital age. This perception appears largely independent of respondents' personal language use, indicating a widely shared recognition of English's prevailing status in the digital sphere. The low proportion of respondents selecting "No" or "Not sure" further reinforces the widespread acknowledgement of English's hegemonic position.

A follow-up question asked whether English *should* remain the primary language of digital communication. The distribution is presented below:

Table 8 Answers to: Should English remain the primary language of digital communication?

Response	Number of respondents	%
Yes	79	63.7%
No	19	15.3%
Neutral	26	21%

Source: The authors

Nearly two-thirds of respondents (63.7%) expressed support for maintaining English as the principal medium of digital communication. This favourable

Response	Number of respondents	%
Yes	37	29,8%
No	49	39,5%
Not yet, but I believe it will	38	30,6%

view likely stems from the practical advantages associated with English proficiency—convenience, accessibility, and the ubiquity of English-language content. Given that most Croatian learners begin English instruction early in their education, the language's entrenched presence in daily life reduces barriers to acceptance and diminishes the perceived need to adopt or learn alternative *linguae francae* for online interaction.

The subsequent set of questions examined the professional relevance of English. First, respondents were asked whether they had taken an online course in English. Their answers are shown below:

Table 9 Answers to: Have you ever taken an English online course?

Response	Number of respondents	%
Yes	31	25%
No	93	75%

Source: The authors

While nearly half of respondents reported a preference for English when learning new skills online, only 25% had taken an online course in English. This discrepancy may indicate a preference for learning in one's native language when instruction is formal or professionally oriented, or it may reflect limited engagement with structured online courses overall. It also suggests that passive or informal learning in English is more common than participation in structured English-language training.

The next question explored the role of English in respondents' current or desired employment:

Table 10 Answers to: Does your current (or desired) job require English?

Response	Number of respondents	%
Yes	59	47.6%
No	37	29.8%
Partially	28	22.6%

Source: The authors

A clear majority (47.6% "Yes" and 22.6% "Partially") indicated that their jobs require at least some proficiency in English. These results corroborate broader labour market trends in Croatia and the European Union, where English often functions as a key professional competency, particularly in fields such as information technology, economics, and international business. Only 29.8% reported no professional need for English, highlighting its pervasive role across sectors.

Respondents were then asked whether English had helped them secure employment, freelance work, or international collaboration:

Table 11 Answers to: Has English helped you get a job, freelance work, or cooperate internationally?

Source: The authors

While only 29.8% reported direct benefits, a significant proportion (30.6%) believe that English will likely support their future career advancement. This forward-looking expectation reflects the perceived importance

of English as a tool for economic mobility and professional development, even among those who have not yet experienced concrete career benefits.

The next question asked respondents whether knowledge of English confers an advantage in their line of work:

Table 12 Answers to: Does knowledge of English give an advantage in your line of work?

Response	Number of respondents	%
Yes	87	70.2%
No	16	12.9%
It depends	21	16.9%

Source: The authors

A strong majority (70.2%) answered affirmatively. This alignment with earlier results reinforces the interpretation that English operates as a valuable professional resource across diverse sectors. Although some variability exists depending on occupational field, English remains widely viewed as a strategic asset that enhances employability and facilitates workplace communication.

The final question asked whether respondents' knowledge of English had helped them access additional information and resources online:

Table 13 Answers to: Has your knowledge of English helped you access more information and resources online?

Response	Number of respondents	%
Yes, significantly	77	62.6%
Yes, somewhat	31	25.2%
Not really	13	10.6%
Not at all	2	1.6%

Source: The authors

More than 87% of respondents indicated that knowing English had enhanced their access to online resources—62.6% to a significant degree and 25.2% to some degree. Only a small minority reported little or no benefit. Given that English dominates digital knowledge production and online instructional materials, this result is consistent with broader empirical data: English proficiency substantially expands digital information horizons and improves access to high-quality online content.

Taken together, the survey results reveal a complex but coherent pattern: while Croatian remains central to respondents' everyday digital practices, English is perceived as indispensable for learning, professional

development, and accessing global information. Respondents overwhelmingly acknowledge English's dominant role in the digital age and anticipate its continued importance. These findings mirror global trends and underscore the sociolinguistic dynamics explored throughout this paper.

6. Conclusion

This paper has examined the emergence and consolidation of English as the dominant global language in the digital age, situating its contemporary status within a broader historical, sociolinguistic, and technological trajectory. Through an analysis of academic literature, empirical data, and an original survey conducted among Croatian respondents, the study has demonstrated that English occupies an unparalleled position in digital communication, scientific dissemination, education, and professional practice.

The results of the survey reveal that the dominance of English is acutely perceived within Croatia, despite the continued importance of Croatian in everyday online engagement. English remains widely used for digital learning, heavily relied upon in professional contexts, and central to accessing a broader range of information and resources. A clear majority of respondents believe that English not only is the dominant language of digital communication but should remain so due to its convenience, established presence, and practical advantages. These findings mirror global trends and reinforce the understanding of English as a key facilitator of global connectivity.

Nevertheless, English's ascendancy is not without significant sociocultural consequences. As the paper has argued, the concentration of linguistic power risks diminishing linguistic diversity, eroding cultural heritage, and reproducing structural inequalities between speakers of dominant and non-dominant languages. The conceptions of untranslatability articulated by Steiner, Benjamin, and Schleiermacher, alongside the theories of linguistic relativity advanced by Humboldt, Sapir, and Whorf, underscore the centrality of language to human cognition, identity, and worldview. The marginalisation or loss of minoritized languages thus represents a profound cultural and epistemological loss.

At the same time, the digital age offers new opportunities for the preservation and revitalisation of endangered languages. Advances in neural machine

translation, multilingual artificial intelligence, and digital archiving demonstrate that the same technological forces contributing to linguistic homogenisation may also be leveraged to support linguistic diversity. Initiatives such as Stanford's SILICON, the Endangered Languages Archive, Wikimedia's Lingua Libre, and various open-source revitalisation platforms illustrate that digital technologies can function as tools of linguistic empowerment rather than instruments of erasure.

The challenge for the coming decades will be to navigate the tension between global connectivity and linguistic plurality. As digital technologies continue to shape communication practices and knowledge production, it is imperative that policymakers, educators, linguistic communities, and technology developers work collaboratively to promote multilingualism, support low-resource languages, and ensure equitable access to digital spaces. A sustainable and culturally inclusive digital future requires recognizing English's practical benefits while actively safeguarding the linguistic diversity that enriches human experience.

English may be the de facto lingua franca of the digital age, but the long-term vitality of global linguistic heritage depends on cultivating an environment in which multiple languages can coexist, evolve, and continue to serve as repositories of culture, identity, and collective memory.

References

- (1) Allen, R. C. (2017). *The industrial revolution: A very short introduction* (1st ed., Vol. 509). Oxford University Press.
- (2) Alshaabi, T., Dewhurst, D. R., Minot, J. R., Arnold, M. V., Adams, J. L., Danforth, C. M., & Dodds, P. S. (2021). The growing amplification of social media: Measuring temporal and social contagion dynamics for over 150 languages on Twitter for 2009–2020. *EPL Data Science*, 10(1). <https://doi.org/10.1140/epjds/s13688-021-00271-0>
- (3) Balogun, S. K., & Aruoture, E. (2024). *Cultural homogenization vs. cultural diversity: Social media's double-edged sword in the age of globalization*. *African Journal of Social and Behavioural Sciences*, 14(4), 1491–1507
- (4) Brody, R. (2010, April 27). Hollyworld. *The New Yorker*. https://www.newyorker.com/culture/richard-brody/hollyworld?utm_
- (5) Crystal, D. (2000). *Language death*. Cambridge University Press.
- (6) Crystal, D. (2003). *English as a global language*. <https://doi.org/10.1017/cbo9780511486999>
- (7) DataReportal. (2024). *Digital 2024: Global overview report* <https://datareportal.com/reports/digital-2024-global-overview-report>
- (8) Di Bitetti, M. S., & Ferreras, J. A. (2016). Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. *AMBIO*, 46(1), 121–127. <https://doi.org/10.1007/s13280-016-0820-7>
- (9) Duranti, A. (1997). *Linguistic anthropology*. Cambridge University Press.
- (10) EF. (2024). *English Proficiency Index: A ranking of 116 countries and regions by English skills*. <https://www.ef.com/wwen/epi/>
- (11) Gordin, M. (2015). *Scientific Babel: How science was done before and after global English*. University of Chicago Press.
- (12) Gu, J., Neubig, G., Cho, K., & Li, V. O. (2017, April 1). Learning to translate in real-time with neural machine translation. *ACL Anthology*. <https://aclanthology.org/E17-1099/>
- (13) Harrison, K. D. (2007). *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford University Press
- (14) Historical trends in the usage statistics of content languages for websites. (2024, August). *W3Techs Web Technology Surveys*. https://w3techs.com/technologies/history_overview/content_language
- (15) House, J. (2003). English as a lingua franca: A threat to multilingualism? *Journal of Sociolinguistics*, 7(4), 556–578. <https://doi.org/10.1111/j.1467-9841.2003.00242.x>
- (16) How English Empowers Your Work and Career (Pearson Report, part 1), (2024, May), Pearson, <https://www.linkedin.com/pulse/how-english-empowers-your-work-career-pearson-report-part-1-30bxf/>
- (17) Isaacson, W. (2014). *The innovators: How a group of hackers, geniuses, and geeks created the digital revolution*. Simon and Schuster.
- (18) Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., & Dean, J. (2017). Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5, 339–351. https://doi.org/10.1162/tacl_a_00065
- (19) Kisery, A. (2019). 3 percent, 60 percent: The singularity of English in the world of translations. In B. Epstein & K. S. Gillett (Eds.), *The Routledge handbook of literary translation* (pp. 395–407). Routledge.
- (20) Kulkarni, G., Ahale, S., Koli, P., Tapadiya, A., & Pachade, P. (2024). AI for real-time language translation. *International Research Journal of Modernization in Engineering, Technology and Science*, 6(10), 5203–5206. <https://www.ijrmets.com>
- (21) Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- (22) Labov, W. (1994). *Principles of linguistic change, Volume 1: Internal factors*. Blackwell.
- (23) Mahajan, N. H. (2025). Machine learning techniques for real-time language translation in social media platforms. *International Journal of Research in Science Commerce Arts Management and Technology*, 293–298. <https://doi.org/10.48175/ijarsct-26933>
- (24) Matasović, R. (2011). *Jezična raznolikost svijeta*. Algoritam.
- (25) Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating language prescription and standardisation* (3rd ed.). Routledge.
- (26) Montgomery, S. L., & Crystal, D. (2013). *Does science need a global language?* <https://doi.org/10.7208/chicago/9780226010045.001.0001>
- (27) National consultation on indigenous languages: A call for preservation and technology integration: The power of digitization for indigenous language preservation. (2025, May 18). UNESCO. <https://www.unesco.org/en/articles/national->

- [consultation-indigenous-languages-call-preservation-and-technology-integration?utm](#)
- (28) Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.
- (29) Pimienta, D., Prado, D., & Blanco, Á. (2009). Twelve years of measuring linguistic diversity in the Internet: Balance and perspectives. *UNESCO*. <https://unesdoc.unesco.org/ark:/48223/pf0000187016>
- (30) Sapir, E. (1929). The status of linguistics as a science. *Language*, 5(4), 207–214.
- (31) Schleiermacher, F. (2012). On the different methods of translating. In L. Venuti (Ed.), *The translation studies reader* (pp. 43–63). Routledge. (Original work published 1813)
- (32) Sourati, Z., Karimi-Malekabadi, F., Ozcan, M., McDaniel, C., Ziabari, A., Trager, J., Tak, A., Chen, M., Morstatter, F., Dehghani, M., & others. (2025). The shrinking landscape of linguistic diversity in the age of large language models. *arXiv*. <https://arxiv.org/abs/2502.11266v1>
- (33) Spector, V. (2025). AI in language preservation: Safeguarding low-resource and indigenous languages. <https://www.welocalize.com/insights/ai-in-language-preservation-safeguarding-low-resource-and-indigenous-languages/?utm>
- (34) Stanford initiative seeks to bring more languages online. (2025, April 22). *Stanford*. <https://news.stanford.edu/stories/2025/04/Stanford-initiative-seeks-to-bring-more-languages-online?utm>
- (35) UNESCO Institute for Statistics. (2005). *Measuring linguistic diversity on the Internet*. United Nations Educational, Scientific and Cultural Organization.
- (36) van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4–5), 221–235. <https://doi.org/10.1016/j.poetic.2006.05.004>
- (37) Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. MIT Press.
- (38) Zeng, J., & Yang, J. (2024). English language hegemony: Retrospect and prospect. *Humanities and Social Sciences Communications*, 11(1). <https://doi.org/10.1057/s41599-024-02821-z>