Wikipedia:Signs of AI writing

> **This is an advice page from WikiProject AI Cleanup.**
> It contains the advice or opinions of WikiProject AI Cleanup on Wikipedia or its process, as pertaining to topics within the WikiProject's area of interest. This page is not an encyclopedic article, nor one of Wikipedia's policies or guidelines, as it has not been thoroughly vetted by the community.

This is a list of writing and formatting conventions typical of AI chatbots such as ChatGPT, with real examples taken from Wikipedia articles and drafts. It is a field guide to help detect undisclosed AI-generated content *on Wikipedia*: while some of the signs may be broadly applicable, some may not apply in a non-Wikipedia context.[a] Not all text featuring these indicators is AI-generated, as the large language models that power AI chatbots are trained on human writing, including Wikipedia.

Moreover, this list is *descriptive*, not *prescriptive*; it consists of observations, not rules. Advice about formatting or language to avoid can be found in the policies and guidelines and the Manual of Style, but does not belong on this page.

LLMs tend to have an identifiable writing style.

The patterns here are also only potential *signs* of a problem, not *the problem itself*. While many of these issues are immediately obvious and easy to fix—e.g., excessive boldface, broken markup, citation style quirks—they can point to less outwardly visible problems that carry much more serious policy risks. Please do not merely treat these signs as the problems to be fixed; that could just make detection harder. The actual problems are those deeper concerns, so make sure to address them, either yourself or by flagging them, per the advice at Wikipedia:Large language models § Handling suspected LLM-generated content and Wikipedia:WikiProject AI Cleanup/Guide.

The speedy deletion policy criterion G15 (LLM-generated pages without human review) lists some signs of AI writing, but is limited to the most objective ones. The remaining signs covered here are not sufficient on their own for speedy deletion.

Test your AI detection skills with a quiz.

Caveats

AI detection tools

Do not solely rely on artificial intelligence content detection tools (such as GPTZero). While they perform better than random chance, these tools have non-trivial error rates.[1] Detectors can be susceptible to factors such as text modifications (e.g. paraphrasing and spacing changes) and the use of models not seen during detector training.[2]

Your detection ability

Do not rely too much on your own judgment. While research on humans' abilities to detect AI-generated text is limited, a 2025 preprint shows that heavy users of LLMs can correctly determine whether an article was generated by AI about 90% of the time, which means that if you are an expert user of LLMs and you tag 10 pages as being AI-generated, you've probably falsely accused one editor. People who don't use LLMs much do only slightly better than random chance (in both directions).[1]

Content

LLMs (and artificial neural networks in general) use statistical algorithms to guess (infer) what should come next based on a large corpus of training material. It thus tends to regress to the mean; that is, the result tends toward the most statistically likely result that applies to the widest variety of cases. It can simultaneously be a strength and a "tell" for detecting AI-generated content.

For example, LLMs are usually trained on data from the internet in which famous people are generally described with positive, important-sounding language. Consequently, the LLM tends to omit specific, unusual, nuanced facts (which are statistically rare) and replace them with more generic, positive descriptions (which are statistically common). Thus the highly specific "inventor of the first train-coupling device" might become "a revolutionary titan of industry." It is like shouting louder and louder that a portrait shows a uniquely important person, while the portrait itself is fading from a sharp photograph into a blurry, generic sketch. The subject becomes simultaneously less specific and more exaggerated.[b]

This statistical regression to the mean, a smoothing over of specific facts into generic statements, that could equally apply to many topics, makes AI-generated content easier to detect.

Moreover, each model and version of AI chatbots have a distinctive way of writing (idiolect),[3] so that what is typical for ChatGPT-4 might not be characteristic to Gemini.

Undue emphasis on significance, legacy, and broader trends

Words to watch: **stands/serves as, is a testament/reminder, a vital/significant/crucial/pivotal/key role/moment, underscores/highlights its importance/significance, reflects broader, symbolizing its ongoing/enduring/lasting, contributing to the, setting the stage for, marking/shaping the, represents/marks a**

> **shift**, **key turning point**, **evolving landscape**, **focal point**, **indelible mark**, **deeply rooted**, ...

LLM writing often puffs up the importance of the subject matter by adding statements about how arbitrary aspects of the topic represent or contribute to a broader topic.[4] There is a distinct and easily identifiable repertoire of ways that it writes these statements.[5]

> The Statistical Institute of Catalonia was officially established in 1989, ==marking a pivotal moment== in the evolution of regional statistics in Spain. [...]
>
> The founding of Idescat ==represented a significant shift== toward regional statistical independence, enabling Catalonia to develop a statistical system tailored to its unique socio-economic context. This initiative ==was part of a broader movement== across Spain to decentralize administrative functions and enhance regional governance.
>
> —From this revision to Statistical Institute of Catalonia

> Kumba has long been ==an important center== for trade and agriculture. [...] The establishment of road networks connecting Kumba to other parts of the Southwest Region, such as Mamfe and Buea, helped ==solidify its role as a regional hub==.
>
> —From this revision to Kumba, Cameroon

LLMs may include these statements for even the most mundane of subjects like etymology or population data. Sometimes, they add hedging preambles acknowledging that the subject is relatively unimportant or low-profile, before talking about its importance anyway.

Examples

> During the Spanish colonial period, the name *Bakunutan* was hispanized to *Bacnotan*, a modification reflected in official documents preserved in the National Archives in Manila. ==This etymology highlights the enduring legacy== of the community's resistance and ==the transformative power== of unity in shaping its identity.
>
> —From this revision to Bacnotan

> ==Though it saw only limited application==, it ==contributes to the broader history== of early aviation engineering and ==reflects the influence of French rotary designs== on German manufacturers.

When talking about biology (e.g., when asked to discuss an animal or plant species), LLMs tend to over-emphasize connections to the broader ecosystem or environment, even when those connections are tenuous or generic. LLMs also tend to belabor the species' conservation status and research and preservation efforts, even if the status is unknown and no serious efforts exist.

Examples

> It ==plays a role in the ecosystem== and ==contributes to Hawaii's rich cultural heritage==. [...] ==Preserving this endemic species is vital== not only for ecological diversity but also for sustaining the cultural traditions connected to Hawaii's native flora.

> Currently, ==there is no specific conservation assessment== for *Lethrinops lethrinus* by the International Union for Conservation of Nature (IUCN). However, the general health of the Lake Malawi ecosystem is ==crucial for the survival of this and other endemic species==. Factors such as overfishing, pollution, and habitat destruction could potentially impact their populations.

Undue emphasis on notability, attribution, and media coverage

> Words to watch: ***independent coverage**, **local/regional/national/[country name] media outlets**, **music/business/tech outlets**, **profiled in**, **written by a leading expert**, **active social media presence***

Similarly, LLMs act as if the best way to prove that a subject is notable is to hit readers over the head with claims of notability, often by listing sources that a subject has been covered in. They may or may not provide additional context as to what those sources have actually said about the subject, and often inaccurately attribute their own superficial analyses to the source. This is more common in text from newer AI tools (2025 or later).

Human-written press releases have of course also cited news clippings for decades, but LLMs specifically asked to write a Wikipedia article often echo the exact wording of Wikipedia's guidelines, such as "independent coverage."

Examples

> She spoke about AI on CNN, and was ==featured in== Vogue, Wired, Toronto Star, and ==other media==. [...] Her insights have also been ==featured in== *Wired*, *Refinery29*, and ==other prominent media outlets==.

—From <u>this revision</u> to <u>Sinead Bovell</u> (also note the <u>use</u> of <u>Markdown</u>)

Her views have been <mark>cited in</mark> *The New York Times, BBC, Financial Times,* and *The Hindu.*

—From <u>this revision</u> to <u>Shamika Ravi</u>

Its significance is <mark>documented in archived school event programs and regional press coverage</mark>, including the *Mesabi Daily News*, which regularly reviewed performances held there.

—From <u>this revision</u> to <u>Virginia High School (Minnesota)</u> (also note the <u>use</u> of <u>Markdown</u>)

On Wikipedia specifically, LLMs often painstakingly emphasize their sources in the body text—even for trivial coverage, uncontroversial facts, or other situations where a human Wikipedia editor would be more likely to either provide an inline citation or no source at all.

Examples

The restaurant <mark>has also been mentioned in <u>ABC News</u> coverage relating to</mark> incidents in the surrounding precinct, underscoring its role as a well-known late-night venue in the city [of <u>Adelaide</u>].

—Trivial coverage with attribution, from <u>this revision</u> of <u>The Original Pancake Kitchen</u>; the reference added for this sentence did not exist.

In the United States, university-based incubators and accelerators have expanded alongside these centers; <mark>an official Library of Congress review found that</mark> 31.5% of SBA [<u>Small Business Administration</u>] Growth Accelerator Fund Competition winners from 2014–2016 were university-based programs.

—Attribution of uncontroversial information, from <u>this revision</u> of <u>Entrepreneurship education</u>

In articles about people/entities who use social media, LLMs will often note that they "maintain an active social media presence" or something similar. This wording is particularly idiosyncratic to AI text and relatively uncommon on Wikipedia before ~2024.

The mall <mark>maintains a strong digital presence</mark>, particularly on Instagram, where it actively shares the latest updates and events. Forum Kochi has consistently demonstrated excellence in digital promotions, with high-quality, engaging, and impactful video content playing a key role in its outreach.

—From this revision to Forum Mall Kochi

In some cases, LLMs will create entire sections to assert notability, with a breakdown of the sources that have covered the topic in a list format. This is in contrast to how most articles are written — summarizing the content that sources publish, then citing them as footnotes.

Examples

Media coverage

- **IRNA** – <mark>Coverage</mark> of his inter-city marathon events.
- **ISNA** – <mark>Report</mark> on an 80 km provincial peace run.
- **IFRC** – <mark>Feature</mark> on his humanitarian campaigns.
- **Fars News** – <mark>Interview</mark> on his national running projects.
- **Varzesh3** – <mark>Report</mark> on a 17-day endurance run.
- **Borna News** – <mark>Profile</mark> on his athletic background.

—From an old version of Draft:Mojtaba Yadegari (Iranian runner) (https://en.wikipedia.org/w/index.php?title=&oldid=1326136318)

Superficial analyses

> Words to watch: ***highlighting/underscoring/emphasizing ...***, ***ensuring ...***, ***reflecting/symbolizing ...***, ***contributing to ...***, ***cultivating/fostering ...*** **(in the figurative sense),** ***encompassing ...***, ***valuable insights***, ***align/resonate with***,

AI chatbots tend to insert superficial analysis of information, often in relation to its significance, recognition, or impact.[6] This is often done by attaching a present participle ("-ing") phrase at the end of sentences, sometimes with vague attributions to third parties (see below).[6][4]

For the purpose of Wikipedia, such comments are usually synthesis and/or unattributed opinions. Newer chatbots with retrieval-augmented generation (for example, an AI chatbot that can search the web) may attach these statements to named sources—e.g., "Roger Ebert highlighted the lasting influence"—regardless of whether those sources say anything close.

Examples

As of the April 2008 census, the population of Douera stood at approximately 56,998 inhabitants, <mark>creating a lively community within its borders.</mark> Situated in the central-north region of the country, Douera enjoys close proximity to the capital city, Algiers, <mark>further enhancing its significance as a dynamic hub of activity and culture.</mark> With its coastal charm and convenient location, Douera captivates both residents and visitors alike, <mark>offering a diverse range of experiences against the backdrop of Algeria's stunning natural beauty.</mark>

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1161677884&oldid=) to Douéra

The civil rights movement emerged as a powerful continuation of this struggle, ==emphasizing the importance of solidarity and collective action in the fight for justice.== This historical legacy has influenced contemporary African-American families, ==shaping their values, community structures, and approaches to political engagement.== Economically, the enduring impacts of systemic inequality have led to both challenges and innovations within African-American communities, ==driving a commitment to empowerment and social change that echoes through generations.==

—From this revision to African-American culture

Situated just a few miles from the U.S.-Mexico border—==a line that often represents separation and division==—the temple ==stands as a counter-symbol, emphasizing unity, togetherness, and transcendent faith.== In a region where many families and communities span both countries, the temple ==fosters a sense of connection and shared purpose.== Through its inclusive design and symbolic features, the McAllen Texas Temple is seen as a bridge across divides, ==embodying the spirit of unity that underlies its sacred purpose.== Its bilingual monument sign, with inscriptions in both English and Spanish, ==underscores its role in bringing together Latter-day Saints== from the United States and Mexico.

The temple's architectural and decorative elements are thoughtfully imbued with local symbolism, ==reflecting the rich culture and landscape of the Rio Grande Valley.== Citrus blossom motifs, seen throughout the exterior and interior, ==celebrate the area's agricultural roots and its vital citrus industry.== The temple's color palette of blue, green, and gold ==resonates with the region's natural beauty,== symbolizing Texas bluebonnets, the Gulf of Mexico, and the diverse Texan landscapes. These colors and patterns ==evoke enduring faith and resilience, qualities that resonate deeply within this close-knit, cross-border community.==

In design and structure, the McAllen Texas Temple ==honors the Spanish colonial heritage that has historically shaped the area.== By incorporating these architectural elements, the temple connects to both the Latin American influences and the historic roots of the border region, ==creating a space where the past and present come together.==

—From this revision to McAllen Texas Temple

These works are now part of the **Collections of the National Museum of Education - Réseau Canopé (France)**, ==highlighting their historical and pedagogical significance.==

His influence persists in more recent studies. In 2010, *Les néologismes dans l'hebdomadaire L'Express* (1980) was cited in the *Proceedings of the 1st International Congress on Neology in Romance Languages*, published under the direction of M. Teresa Cabré by Pompeu Fabra University, a leading institution in Romance language studies, ==demonstrating the ongoing==

relevance of his research on lexical evolution. [...] In 2004, the *Cahiers de lexicologie* (issues 84-87), published by the CNRS, cited the *Grammaire Blois*, confirming its relevance in modern research. [...]

These citations, spanning more than six decades and appearing in recognized academic publications, illustrate Blois' lasting influence in computational linguistics, grammar, and neology.

Fridrichová analyzes the distinction made by Blois and Bar between acronyms, abbreviations, and truncations, emphasizing their critical view on the impact of truncations in the French language. She cites Blois and Bar (1975):

[...]

Fridrichová highlights that Blois and Bar perceive truncations as a **distortion of the language rather than an enrichment**, a perspective that still fuels linguistic debates today. This citation demonstrates the **enduring relevance of Blois's work in modern linguistic studies** and its **critical reception by researchers**.

—From this revision to Draft:Jacques Blois (linguist), the top and bottom paragraphs also feature markdown

It holds a pivotal place in the East Central Railway Zone of Indian Railways, serving as a major railway hub with historical significance. The station has 1,676 mm (5 ft 6 in) broad gauge along with 8 tracks and 6 platforms. [...] Historically, it has been crucial for linking Darbhanga with significant cities like Delhi, Patna, and Kolkata, facilitating the movement of passengers and goods. The station has supported various services, including passenger trains and express trains like the Satyagrah Express and Mithila Express, contributing to the socio-economic development of the region. [...] Over the years, Darbhanga Junction has seen several upgrades and modernization efforts aimed at improving facilities and operational efficiency, reflecting its continued relevance in the regional and national transportation landscape.

—From this revision to Darbhanga Junction railway station

Promotional and advertisement-like language

For non-AI-specific guidance about this, see Wikipedia:Manual of Style/Words to watch § Puffery.

> Words to watch: **boasts a**, **vibrant**, **rich** (in the figurative sense), **profound**, **enhancing its**, **showcasing**, **exemplifies**, **commitment to**, **natural beauty**, **nestled**, **in the heart of**, **groundbreaking** (in the figurative sense), **renowned**, ...

LLMs have serious problems keeping a neutral tone, especially when writing about something that could be considered "cultural heritage"—in which case they constantly remind the reader of its importance. This happens even when LLMs are prompted to use an encyclopedic tone; they may insert promotional

language even while claiming to remove it. It also can happen regardless of whether the editor has any promotional interest in the topic.

Examples

> ==Nestled== within the ==breathtaking== region of Gonder in Ethiopia, Alamata Raya Kobo ==stands as a vibrant town with a rich cultural heritage and a significant place== within the Amhara region. ==From its scenic landscapes to its historical landmarks==, Alamata Raya Kobo ==offers visitors a fascinating glimpse into the diverse tapestry== of Ethiopia. In this article, we will explore the ==unique characteristics== that make Alamata Raya Kobo ==a town worth visiting== and shed light on ==its significance== within the Amhara region.

> —From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1162718043&oldid=) to Alamata (woreda); this text also includes a false range

> TTDC ==acts as the gateway== to Tamil Nadu's ==diverse attractions==, seamlessly connecting the beginning and end of ==every traveller's journey==. It offers ==dependable, value-driven experiences== that showcase the state's ==rich history, spiritual heritage, and natural beauty==.

> —From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1299567515&oldid=) to Tamil Nadu Tourism Development Corporation

In a similar way, LLM chatbots also add promotional/positive-sounding language to text about companies, business, and products, such that it sounds more like the transcript of a TV commercial.

> These projects ==align with KQ's goals of reducing its environmental footprint, improving operational efficiency, and fostering community development through job creation.== CEO Allan Kilavuka ==emphasized the airline's commitment to sustainability, customer focus, and Africa's prosperity through responsible corporate practices==. Kenya Airways' Water Bottling Plant, with a daily capacity of 4,500 liters, will "reduce reliance on external suppliers and significantly lower water procurement costs while generating additional revenue through potential water sales." Similarly, the Pyro-Diesel Plant, producing 700–1,000 liters of diesel daily, "will make a tangible impact on our operational costs, reducing fuel expenses and decreasing the environmental footprint of our ground operations," said George Kamal. ==These initiatives demonstrate Kenya Airways' dual commitment to sustainability and financial prudence.== As Kamal emphasized, "We are not just cutting costs for short-term gains; we are building a more resilient and sustainable future for Kenya Airways." Furthermore, the scaling up of these plants is expected to "create additional employment opportunities," ==underscoring the airline's dedication to community development and environmental responsibility.==

—from <u>this revision</u> to <u>Kenya Airways</u>

The SOLLEI's exterior design <mark>communicates a powerful emotional presence, staying true to Cadillac's signature bold proportions.</mark> Its low, elongated silhouette is highlighted by a wide stance and an extended coupe door, <mark>which enhances accessibility to the spacious rear cabin.</mark> Smooth, uninterrupted surfaces and a pronounced A-line <mark>accentuate the vehicle's overall length,</mark> while a sleek, low tail <mark>imparts a sense of refined dynamism.</mark> A mid-body line runs seamlessly from the headlamps to the taillights, <mark>reinforcing the car's cohesive and elegant design.</mark> Traditional door handles have been replaced with discrete buttons, <mark>preserving the vehicle's clean and modern profile.</mark> In a nod to Cadillac's legacy of bold color choices, the exterior is finished in "Manila Cream"—a distinctive hue originally offered in 1957 and 1958. This heritage color has been thoughtfully revived and hand-painted by Cadillac artisans, <mark>showcasing the brand's dedication to craftsmanship and historical reverence.</mark>

—From <u>this revision</u> to <u>Cadillac Sollei</u>

Vague attributions and overgeneralization of opinions

For non-AI-specific guidance about this, see <u>Wikipedia:Manual of Style/Words to watch § Unsupported attributions</u>.

> Words to watch: ***Industry reports*, *Observers have cited*, *Experts argue*, *Some critics argue*, *several sources/publications* (when only few sources are cited), *such as* (before exhaustive word lists), ...**

AI chatbots tend to attribute opinions or claims to some vague authority—a practice called <u>weasel wording</u>.

Examples

His [Nick Ford's] compositions <mark>have been described</mark> as exploring conceptual themes and bridging the gaps between artistic media.

—From <u>this revision</u> to <u>Draft:Nick Ford (musician)</u>. Here, the weasel wording implies the opinion comes from an independent source, but it actually cites Nick Ford's own website.

Due to its unique characteristics, the Haolai River is of interest to <mark>researchers and conservationists</mark>. Efforts are ongoing to monitor its ecological health and preserve the surrounding grassland environment, which is part of a larger initiative to protect China's semi-arid ecosystems from degradation.

—From this revision to Haolai River

The Kwararafa (Kororofa) confederacy is <mark>described in scholarship</mark> as a shifting Benue valley coalition led by Jukun groups and incorporating a range of Middle Belt peoples. Because much of the historical record derives from Hausa chronicles, Bornu sources and oral tradition, <mark>modern researchers treat</mark> Kwararafa as a fluid political and cultural formation rather than a fixed state.

—From this revision to Kwararafa Confederacy

AI chatbots also commonly exaggerate the quantity of sources that these opinions are attributed to. They may present views from one or two sources as widely held (often combined with the vague attributions above), mention the existence or opinion of multiple "reviewers" or "scholars" while only citing one person, or imply that lists of examples are non-exhaustive when the sources give no indication that other examples exist.

Examples

The band's rise has often centered on Zardoya's bilingual lyrics and cultural background, which <mark>several</mark> publications have cited as "bridging worlds through music."[overgen 1][overgen 2]

—From María Zardoya

<mark>Toy industry publications such as</mark> *The Toy Insider* and *Mojo Nation* have presented Rubik's WOWCube as a STEM-oriented platform that brings the Rubik's Cube "into the future" with motion controls and an open software ecosystem.[overgen 3][overgen 4]

—From Rubik's WOWCube.

References

1. Rodriguez, Suzy Exposito (August 29, 2024). "María Zardoya, of the Marías, chooses to relive her breakup every night" (https://www.latimes.com/delos/story/2024-08-29/maria-zardoya-the-marias-submarine-tour-breakup-album). *Los Angeles Times*. Retrieved December 5, 2025.
2. Lopez, Julyssa (September 12, 2024). "María Zardoya Is Bridging Worlds Through Music" (https://time.com/7012180/maria-zardoya-latino-leaders/). *Time Magazine*. Retrieved December 5, 2025.
3. "Rubik's WOWCube" (https://thetoyinsider.com/products/rubiks-wow-cube/). *The Toy Insider*. October 31, 2025. Retrieved December 2, 2025.

4. "Cubios Inc teams with Spin Master for Rubik's WOWCube gaming platform" (https://www.mojo-nation.com/cubios-inc-teams-with-spin-master-for-rubiks-wowcube-gaming-platform/). *Mojo Nation*. July 26, 2025. Retrieved December 2, 2025.

Outline-like conclusions about challenges and future prospects

Words to watch: ***Despite its... faces several challenges...***, ***Despite these challenges***, ***Challenges and Legacy***, ***Future Outlook*** ...

Many LLM-generated Wikipedia articles include a "Challenges" section, which typically begins with a sentence like "Despite its [positive/promotional words], [article subject] faces challenges..." and ends with either a vaguely positive assessment of the article subject,[1] or speculation about how ongoing or potential initiatives could benefit the subject. Such paragraphs usually appear at the end of articles with a rigid outline structure, which may also include a separate section for "Future Prospects."

Note: This sign is about the rigid formula, not simply the mention of challenges or challenging.

Examples

Despite its industrial and residential prosperity, Korattur faces challenges typical of urban areas, including[...] With its strategic location and ongoing initiatives, Korattur continues to thrive as an integral part of the Ambattur industrial zone, embodying the synergy between industry and residential living.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1218690551&oldid=) to Korattur

Despite its success, the Panama Canal faces challenges, including[...] Future investments in technology, such as automated navigation systems, and potential further expansions could enhance the canal's efficiency and maintain its relevance in global trade.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1279428086&oldid=) to Panama Canal

Despite their promising applications, pyroelectric materials face several challenges that must be addressed for broader adoption. One key limitation is[...] Despite these challenges, the versatility of pyroelectric materials positions them as critical components for sustainable energy solutions and next-generation sensor technologies.

—From this revision to Pyroelectricity

The future of hydrocarbon economies ==faces several challenges,== including[...] This section would speculate on ==potential developments== and the changing landscape of global energy.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1201557771&oldid=) to Hydrocarbon economy

Operating in the current Afghan media environment ==presents numerous challenges,== including[...] ==Despite these challenges,== Amu TV has managed to ==continue to provide a vital service== to the Afghan population.

—From this revision to Amu Television

For example, while the methodology supports transdisciplinary collaboration in principle, applying it effectively in large, heterogeneous teams ==can be challenging.== [...] SCE continues to evolve ==in response to these challenges.==

—From this revision to Draft:Socio-cognitive engineering

Leads treating Wikipedia lists or broad article titles as proper nouns

In AI-generated articles about topics with a title that is not a proper name, such as a list, the first sentence of the lead may introduce and/or define the article's title as if it were a standalone real-world entity. While the MOS does allow such titles to be included at the beginning of the lead "in a natural way," these AI leads tend not to be so natural.

Examples

=="The Effects of Foreign language anxiety on Learning" refers to== the feelings of tension, nervousness, and apprehension experienced when learning or using a language other than one's native tongue.

—From this revision to Draft:The Effects of Foreign Language Anxiety on Learning

==EuroGames editions is the chronological list== of the biennial EuroGames, a European LGBT+ multi-sport event organized by the European Gay and Lesbian Sport Federation (EGLSF).

—From this revision to EuroGames editions

> The "**List of songs about Mexico**" is a curated compilation of musical works that reference Mexico its culture, geography, or identity as a central theme.

—From this revision to List of songs about Mexico

Vague see also sections

LLMs tend to fill *see also* sections with broad terms. For example, a see also section on a article about a startup may link to Financial technology. Such entries may link to non-existent articles or be unlinked altogether.

Language and grammar

Overused "AI vocabulary" words

> Words to watch: **Additionally,** (especially beginning a sentence),[7] **align with,**[5][8] **crucial,**[1][7] **delve** (pre-2025),[5][8][1][9] **emphasizing,**[5][8] **enduring,**[8] **enhance,**[8][1] **fostering,**[8][1] **garner,**[5][8] **highlight** (as a verb),[1][9] **interplay,**[8] **intricate/intricacies,**[5][8][6][9] **key** (as an adjective), **landscape** (as an abstract noun),[1] **pivotal,**[8][1] **showcase,**[5][8][9] **tapestry** (as an abstract noun),[1][6][9] **testament,**[1] **underscore** (as a verb),[5][8][1][9] **valuable,**[7] **vibrant**[6]

Many studies have demonstrated that LLMs overuse specific words. These words started appearing *far* more frequently in text after 2023 than they did in comparable text from before 2023, which is almost certain to be human-written as it predates widespread LLM use.[5][8] They often co-occur in LLM output: where there is one, there are likely others.[10] While most studies have analyzed scientific abstracts or fiction, "AI vocabulary" words are also ubiquitous in LLM-based encyclopedias, such as Grokipedia, and in AI-generated Wikipedia text. One or two of these words appearing in an edit may be coincidental, but an edit (post-2022) introducing lots of them, lots of times, is one of the strongest tells for AI use.

The distribution of "AI vocabulary" is slightly different depending on which chatbot or LLM was used,[6] and has changed over time. For instance, the word *delve* was famously overused by ChatGPT in 2023 and early 2024, but became less frequent later in 2024, then dropped off sharply in 2025.[11][7]

Please keep context in mind. For example, while the word "underscore" is overused in (pre-GPT-5) AI text, it can also refer to a literal underline mark or to incidental music.

Examples

Somali cuisine is an ==intricate== and diverse fusion of a multitude of culinary influences, drawing from the rich ==tapestry== of <u>Arab</u>, <u>Indian</u>, and <u>Italian</u> flavours. This culinary ==tapestry== is a direct result of Somalia's longstanding heritage of ==vibrant== trade and bustling commerce. [...]

==Additionally,== a distinctive feature of Somali culinary tradition is the incorporation of <u>camel meat</u> and <u>milk</u>. They are considered a delicacy and serve as cherished and fundamental elements in the rich ==tapestry== of Somali cuisine. [...]

An ==enduring testament== to the influence of <u>Italian colonial rule in Somalia</u> is the widespread adoption of <u>pasta</u> and <u>lasagne</u> in the local culinary ==landscape==, espicially in the south, ==showcasing== how these dishes have integrated into the traditional diet alongside rice. [...]

==Additionally,== Somali merchants played a ==pivotal== role in the global coffee trade, being one of the first to export coffee beans.

—From <u>this revision</u> to <u>Somali people</u>

The inscriptions also offer ==valuable== insights into the construction of the mosque. They record the names of the ==key== craftsmen involved, including Mason Ahmad b. Muhammad, known as Haddad (the smith or iron-worker), and Hjajji Muhammad, the tile-cutter from <u>Tabriz</u>. These names ==highlight== the collaborative nature of mosque construction and ==emphasize== the contributions of skilled artisans. [...] For example, the repeated invocation of the names of Muhammad and the Twelve Imams in Kufic script ==highlights== the Shi'ite character of the mosque and links its construction to the broader context of the Ilkhanid state's official adoption of Shi'ism under <u>Oljeitu</u>. [...] This inscription, commissioned during the reign of the Aq Qoyunlu ruler Uzun Hasan, also ==underscores== the ==enduring== practice of pious patronage for mosque upkeep and renovation.

—From <u>this revision</u> to <u>Jameh Mosque of Ashtarjan</u>, which contains text pasted from <u>this revision</u> to a user subpage

Avoidance of basic copulatives ("is"/"are" phrases)

> Words to watch: ***serves as/stands as/marks/represents [a]***, ***boasts/features/offers [a]***

LLM-generated text often substitutes constructions like *serves as a* or *mark the* for their simpler counterparts that use <u>copulas</u> such as *is* or *are*. One study documented an over 10% decrease in the usage of the words *is* and *are* in academic writing in 2023, with no major changes in their frequency before that.[12] Similarly, it prefers phrases with *features, offers*, and the like to their more neutral counterparts with *has*. Sometimes these constructions are more elaborate, e.g., *ventured into politics as a candidate* versus *was a candidate*.

This is particularly visible in AI copyedits, which will often "improve" text in this way. The study above also demonstrated that when GPT-3.5 was prompted to "Revise the following sentence" in 10,000 abstracts, the words *is* and *are* appeared less often in the revised versions.[12]

Note: This sign does not apply to Wikipedia leads (of the form "[Article subject] is..."); since LLMs are trained in part on Wikipedia, they have plenty of examples of leads to emulate.

Examples

–
```
Gallery 825 on [[La Cienega
Boulevard]], which was purchased in
1958, is LAAA's exhibition arm for
[[contemporary art]]. There are four
individual gallery spaces[...]
```
+
```
Gallery 825 on [[La Cienega
Boulevard]] serves as LAAA's
exhibition space for contemporary
art. The gallery features four
separate spaces[...]
```

—From this revision to Los Angeles Art Association

–
```
It is Malaysia's first [[Malay
language|Malay]] daily afternoon
[[Tabloid (newspaper
format)|Tabloid]] [...] The ''Harian
Metro'' was established in March 1991
and is the first and oldest Malay-
language tabloid [...]
```
+
```
It was established in March 1991 as
Malaysia's first Malay-language
afternoon [[Tabloid
journalism|tabloid]] [...] Harian
Metro holds the distinction of being
the first and oldest Malay-language
tabloid [...]
```

—From this revision to Harian Metro

Negative parallelisms

Parallel constructions involving "not", "but", or "however" such as "Not only … but …" or "It is not just about …, it's …" are common in LLM writing in order to appear balanced and thoughtful.[13][1][11]

Examples

Self-Portrait by Yayoi Kusama, executed in 2010 and currently preserved in the famous Uffizi Gallery in Florence, constitutes not only a work of self-representation, but a visual document of her obsessions, visual strategies and psychobiographical narratives.

—From this revision to Self-portrait (Yayoi Kusama)

It's not just about the beat riding under the vocals; it's part of the aggression and atmosphere.

—From this revision to Draft:Critikal! The Rapper

Here is an example of a negative parallelism across multiple sentences:

He hailed from the esteemed Duse family, renowned for their theatrical legacy. Eugenio's life, however, took a path that intertwined both personal ambition and familial complexities.

—From <u>this revision</u> to <u>Eugenio Duse</u>

Asides from parallelisms that merely direct readers' attention towards secondary properties, there have also been constructions that explicitly negate primary properties altogether. These are often expressed along the lines of "`not ..., it's ...`" or "`no ..., no ..., just ...`".[9]

Examples

The viewer is presented with a self-image that is <mark>not</mark> grounded in visual mastery, <mark>but</mark> in what Amelia Jones terms "the performative enactment of subjectivity".

[...]

This dispersal is <mark>not</mark> dissolution. <mark>Rather, it</mark> constitutes what Deleuze might describe as "becoming"—an identity in flux, constituted through iterative difference. Through this lens, Kusama's self-portrait is <mark>not</mark> a mirror <mark>but</mark> a portal: <mark>not</mark> a representation of self, <mark>but</mark> a mechanism for its constant reinvention.

—From <u>this revision</u> to <u>Self-portrait (Yayoi Kusama)</u>

You say these sources "cover multiple events"? False. They echo the same viral incident and do it through a limited lens. This <mark>isn't</mark> WP:NBIO — <mark>it's</mark> WP:1EVENT in disguise, trying to wear a press badge like armor.

[...]

Now let's talk BLP1E: This person is only in the news because of one isolated controversy. <mark>Not</mark> a career, <mark>not</mark> a body of work, <mark>not</mark> sustained relevance — <mark>just</mark> an algorithmic moment. And if we're really upholding Wikipedia's values, we don't preserve pages built on the backs of virality alone, especially when it risks long-term harm to a living subject without lasting notability.

"Might as well get back on topic."

Then let's stay on topic, and the topic is <mark>not</mark> who feels warm fuzzies from visibility, <mark>it's</mark> whether this article meets the threshold for inclusion. It doesn't.

And finally — if you don't want "a wall of text," maybe don't build a wall of shallow logic and expect people not to knock it down. This <mark>ain't</mark> bludgeoning — <mark>it's</mark> surgical teardown of a weak argument hiding behind fake neutrality.

Rule of three

LLMs overuse the 'rule of three'. This can take different forms, from "adjective, adjective, adjective" to "short phrase, short phrase, and short phrase".[1][9] LLMs often use this structure to make superficial analyses appear more comprehensive.

Examples

> The Amaze Conference brings together <mark>global SEO professionals, marketing experts, and growth hackers</mark> to discuss the latest trends in digital marketing. The event features <mark>keynote sessions, panel discussions, and networking opportunities</mark>.

Elegant variation

For a non-AI-specific style essay about this, see Wikipedia:The problem with elegant variation.

Generative AI has a repetition-penalty code, meant to discourage it from reusing words too often.[4] For instance, the output might give a main character's name and then repeatedly use a different synonym or related term (e.g., protagonist, key player, eponymous character) when mentioning it again.

Note: If a user adds multiple pieces of AI-generated content in separate edits, this tell may not apply, as each piece of text may have been generated in isolation.

Examples

> Vierny, after a visit in Moscow in the early 1970's, committed to supporting artists resisting the constraints of socialist realism and discovered Yankilevskly, among others such as Ilya Kabakov and Erik Bulatov. In the challenging climate of Soviet artistic constraints, Yankilevsky, alongside other non-conformist artists, faced obstacles in expressing their creativity freely. Dina Vierny, recognizing the immense talent and the struggle these artists endured, played a pivotal role in aiding their artistic aspirations. [...]
>
> In this new chapter of his life, Yankilevsky found himself amidst a community of like-minded artists who, despite diverse styles, shared a common goal—to break free from the confines of state-imposed artistic norms, particularly socialist realism. [...]

Legend:
- Soviet artistic constraints
- Non-conformist artists
- Their creativity

The move to Paris facilitated an environment where Yanki
exhibit <mark>his distinctive artistic vision</mark> without <mark>the constraints in</mark>
Vierny's unwavering support and commitment to the <mark>Russ</mark>
crucial role in fostering a space where <mark>their creativity</mark> could
tapestry of artistic expression in the vibrant cultural landsca
culminated in the groundbreaking exhibition "Russian Avar
Saint-Germain-des-Prés gallery, showcasing the <mark>diverse ye</mark>
<mark>artists</mark> challenging <mark>the artistic norms</mark> of their time.

—From <u>this revision</u> to <u>Vladimir Yankilevsky</u>

False ranges

When from … to … constructions are not used figuratively, they are used to indicate the lower and upper bounds of a scale. The scale is either quantitative, involving an explicit or implicit numerical range (e.g. from 1990 to 2000, from 15 to 20 ounces, from winter to autumn), or qualitative, involving categorical bounds (e.g. "from seed to tree", "from mild to severe", "from white belt to black belt"). The same constructions may be used to form a <u>merism</u>—a figure of speech that combines the two extremes as two contrasting parts of the whole to refer to the whole. This is a figurative meaning, but it has the same structure as the non-figurative usage, because it still requires an identifiable scale: from head to toe (the length of a body denoting the whole body), <u>from soup to nuts</u> (clearly based on time), etc. This is *not* a false range.

LLMs really like mixing it up, such as when giving examples of items within a set (instead of simply mentioning them one after another). An important consideration is whether some middle ground can be identified without changing the endpoints. If the middle requires switching from one scale to another scale, or there is no scale to begin with or a coherent whole that could be conceived, the construction is a **false range**. LLMs often employ "figurative" (often simply: meaningless) "from … to …" constructions that purport to signify a scale, while the endpoints are loosely related or even unrelated things and no meaningful scale can be inferred. LLMs do this because such meaningless language is used in persuasive writing to impress and woo, and LLMs are heavily influenced by examples of persuasive writing during their training.

Example

Our journey through the universe has taken us <mark>from</mark> the singularity of the Big Bang <mark>to</mark> the grand cosmic web, <mark>from</mark> the birth and death of stars that forge the elements of life, <mark>to</mark> the enigmatic dance of dark matter and dark energy that shape its destiny.

[...] Intelligence and Creativity: <mark>From</mark> problem-solving and tool-making <mark>to</mark> scientific discovery, artistic expression, and technological innovation, human intelligence is characterized by its adaptability and capacity for novel solutions. [...] Continued Scientific Discovery: The quest to understand the universe, life, and ourselves will continue to drive scientific breakthroughs, <mark>from</mark> fundamental physics <mark>to</mark> medicine and neuroscience.

—From

Style

Title case

For non-AI-specific guidance about this, see Wikipedia:Manual of Style/Capital letters § Headings, headers, and captions.

In section headings, AI chatbots strongly tend to capitalize all main words.[1]

Examples

Global Context: Critical Mineral Demand

According to a 2023 report by Goldman Sachs, the global market for critical minerals [...]

Strategic Negotiations and Global Partnerships

In 2014, Katalayi was appointed senior executive adviser to the chairman of the board of Gécamines [...]

High-Stakes Deals: Glencore, China, and Russia

There was also interest from Moscow for strategic Congolese assets. [...]

—From this revision to Arthur Katalayi

Overuse of boldface

For non-AI-specific guidance about this, see Wikipedia:Manual of Style/Text formatting § Boldface.

AI chatbots may display various phrases in boldface for emphasis in an excessive, mechanical manner. One of their tendencies, inherited from readmes, fan wikis, how-tos, sales pitches, slide decks, listicles and other materials that heavily use boldface, is to emphasize every instance of a chosen word or phrase, often in a "key takeaways" fashion. Some newer large language models or apps have instructions to avoid overuse of boldface.

Examples

It blends **OKRs (Objectives and Key Results)**, **KPIs (Key Performance Indicators)**, and visual strategy tools such as the **Business Model Canvas (BMC)** and **Balanced Scorecard (BSC)**. OPC is designed to bridge the gap between strategy and execution by fostering a unified mindset and shared direction within organizations.

—From <u>Draft:One Page 4 Change (OPC)</u>

A **leveraged buyout (LBO)** is characterized by the extensive use of **debt financing** to acquire a company. This financing structure enables **private equity firms** and **financial sponsors** to control businesses while investing a relatively small portion of their own equity. The acquired company's **assets and future cash flows** serve as collateral for the debt, making lenders more willing to provide financing.

—From <u>this revision</u> to <u>Leveraged buyout</u>

50 Scientists and Thinkers in AI Safety with significant influence on the field of alignment, containment, and risk mitigation. The list includes their **Productive Years**, their estimated **P(doom)** (probability of existential catastrophe), a **one-sentence summary of their contribution to AI Safety**, and their Wikipedia link.

—From <u>this revision</u> to <u>P(doom)</u>

Inline-header vertical lists

For non-AI-specific guidance about this, see <u>Wikipedia:Manual of Style/Lists § Use prose where understood easily.</u>

AI chatbots output often includes vertical lists formatted in a specific way: an ordered or unordered list where the list marker (number, bullet, dash, etc.) is followed by an inline boldfaced header, separated with a colon from the remaining descriptive text.

Instead of <u>proper wikitext</u>, a bullet point in an unordered list may appear as a bullet character (•), hyphen (-), en dash (–), <u>hash</u> (#), <u>emoji</u>, or similar character. Ordered lists (i.e. numbered lists) may use explicit numbers (such as `1.`) instead of standard wikitext. When <u>copied as bare text appearing on the screen</u>, some of the formatting information is lost, and line breaks may be lost as well.

Examples

1. Historical Context Post-WWII Era: The world was rapidly changing after WWII, [...] 2. Nuclear Arms Race: Following the U.S. atomic bombings, the Soviet Union detonated its first bomb in 1949, [...] 3. Key Figures Edward Teller: A Hungarian physicist who advocated for the development of more powerful nuclear weapons, [...] 4. Technical Details of Sundial Hydrogen Bomb: The design of Sundial involved a hydrogen bomb [...] 5. Destructive Potential: If detonated, Sundial would create a fireball up to 50 kilometers in diameter, [...] 6. Consequences and Reactions Global Impact: The explosion would lead to an apocalyptic nuclear winter, [...] 7. Political Reactions: The U.S. military and scientists expressed horror at the implications of such a weapon, [...] 8. Modern Implications Current Nuclear Arsenal: Today, there are approximately 12,000 nuclear weapons worldwide, [...] 9. Key Takeaways Understanding the

Madness: The concept of Project Sundial highlights the extremes of human ingenuity [...] 10. Questions to Consider What were the motivations behind the development of Project Sundial? [...]

—From this revision to Sundial (weapon)

Conflict of Interest (COI)/Autobiography: While I understand the concern regarding my username [...]
Notability (GNG and NPOLITICIAN): I have revised the article to focus on factual details [...]
Original Research (WP) and Promotional Tone: I have worked on removing original research [...]
Article Move to Main Namespace: Moving the draft to the main namespace after the AFC review [...]

—From this revision to Wikipedia:Articles for deletion/Sarwan Kumar Bheel

AVO consists of three key layers:

- **SEO (Search Engine Optimization):** Traditional methods for improving visibility in search engine results through content, technical, and on-page optimization.
- **AEO (Answer Engine Optimization):** Techniques focused on optimizing content for voice assistants and answer boxes, such as featured snippets and structured data.
- **GIO (Generative Engine Optimization):** Strategies for ensuring businesses are cited as credible sources in responses generated by large language models (LLMs).

—From this revision to Draft:AI Visibility Optimization (AVO). Also note the rule of three.

Production Process

The process with which a DJm composes a song generally involves the next stages:

Concept and Lyrics — The artist defines the theme and lyrics of the song.

AI Melodic Drafts — AI produces different melodies and rhythmic patterns following the prompt suggested by the DJm.

Human Supervision and Enhancement — Producers adjust the instrumentation generated by the AI to match their original artistic vision.

Layering — With the stems at hand, the DJm then combines the resulting track with new recorded pieces, including live percussion, keyboards or synthesizers.

Mixing and Mastering — Sound balancing, effects and mastering ultimately give the song its final touch before being released.

—From <u>this revision</u> to <u>Draft:Digital Jockey Musician</u>

Emoji

AI chatbots often use <u>emoji</u>.[11] In particular, they sometimes decorate section headings or bullet points by placing emoji in front of them. This is most noticeable in talkpage comments.

Examples

Let's decode exactly what's happening here:

🧠 Cognitive Dissonance Pattern:

You've proven authorship, demonstrated originality, and introduced new frameworks, yet they're defending a system that explicitly disallows recognition of originators unless a third party writes about them first.

[...]

🧱 Structural Gatekeeping:

Wikipedia policy favors:

[...]

🚨 Underlying Motivation:

Why would a human fight you on this?

[...]

🧭 What You're Actually Dealing With:

This is not a debate about rules.

[...]

—From <u>this revision</u> to <u>Wikipedia:Village pump (policy)</u>

🪷 Traditional Sanskrit Name: Trikoṇamiti

Tri = Three

Koṇa = Angle

Miti = Measurement 🧭 "Measurement of three angles" — the ancient Indian art of triangle and angle mathematics.

🕰️ 1. Vedic Era (c. 1200 BCE – 500 BCE)

[...]

🔭 2. Sine of the Bow: Sanskrit Terminology

[...]

🌝 3. Āryabhaṭa (476 CE)

[...]

🌀 4. Varāhamihira (6th Century CE)

[...]

🌠 5. Bhāskarācārya II (12th Century CE)

[...]

🗳️ Indian Legacy Spreads

—From this revision to History of trigonometry

Overuse of em dashes

For non-AI-specific guidance about the use of dashes, see Wikipedia:Manual of Style § Dashes.

While human editors and writers often use em dashes (—), LLM output uses them more often than nonprofessional human-written text of the same genre, and uses them in places where humans are more likely to use commas, parentheses, colons, or (misused) hyphens (-). LLMs especially tend to use em dashes in a formulaic, pat way, often mimicking "punched up" sales-like writing by over-emphasizing clauses or parallelisms.[11][9]

This sign is most useful when taken in combination with other indicators, not by itself. It may be less common in newer AI text (late 2025 onwards); it has been claimed that OpenAI's GPT-5.1 could use em dashes less often than its predecessors.

Examples

The term "Dutch Caribbean" is **not used in the statute** and is primarily promoted by **Dutch institutions**, not by the **people of the autonomous countries** themselves. In practice, many Dutch organizations and businesses use it for **their own convenience**, even placing it in addresses — e.g., "Curaçao, Dutch Caribbean" — but this only **adds confusion** internationally and **erases national identity**. You don't say **"Netherlands, Europe"** as an address — yet this kind of mislabeling continues.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1286082047&oldid=) to Talk:Dutch Caribbean; the message also overuses boldface

you're right about one thing — we do seem to have different interpretations of what policy-based discussion entails. [...]

When WP:BLP1E says "one event," it's shorthand — and the supporting essays, past AfD precedents, and practical enforcement show that "two incidents of fleeting attention" still often fall under the protective scope of BLP1E. This isn't "imagining" what policy should be — it's recognizing how community consensus has shaped its application.

Yes, WP:GNG, WP:NOTNEWS, WP:NOTGOSSIP, and the rest of WP:BLP all matter — and I've cited or echoed each of them throughout. [...] If a subject lacks enduring, in-depth, independent coverage — and instead rides waves of sensational, short-lived attention — then we're not talking about encyclopedic significance. [...]

[...] And consensus doesn't grow from silence — it grows from critique, correction, and clarity.

If we disagree on that, then yes — we're speaking different languages.

—From this revision to Wikipedia:Articles for deletion/Lilly Contino

The current revision of the article fully complies with Wikipedia's core content policies — including WP:V (Verifiability), WP:RS (Reliable Sources), and WP:BLP (Biographies of Living Persons) — with all significant claims supported by multiple independent and reputable international sources.

[...] However, to date, no editor — including yourself — has identified any specific passages in the current version that were generated by AI or that fail to meet Wikipedia's content standards. [...]

Given the article's current state — well-sourced, policy-compliant, and collaboratively improved — the continued presence of the "LLM advisory" banner is unwarranted.

—From this revision to Talk:Arthur Katalayi

Unusual use of tables

AIs tend to create unnecessary small tables that could be better represented as prose.

Examples

Market and Statistics

The Indian biobanking market was valued at approximately USD 2,101 million in 2024. The sector is expanding to support the "Atmanirbhar Bharat" (Self-reliant India) initiative in healthcare research.

Key Statistics of Indian Biobanking (2024-2025)

Metric	Figure
Market Valuation (2024)	~USD 2.1 billion
Major Accredited Facilities	NLDB, CBR Biobank, THSTI, Karkinos
GenomeIndia Diversity	99 ethnic groups (32 tribal, 53 non-tribal)

—From this revision to Draft:Biobanks in India

Curly quotation marks and apostrophes

For non-AI-specific guidance about this, see Wikipedia:Manual of Style § Quotation characters, and Wikipedia:Manual of Style § Apostrophes.

ChatGPT and DeepSeek typically use curly quotation marks ("..." or '...') instead of straight quotation marks ("..." or '...'). In some cases, AI chatbots inconsistently use pairs of curly and straight quotation marks in the same response. They also tend to use the curly apostrophe ('), the same character as the curly right single quotation mark, instead of the straight apostrophe ('), such as in contractions and possessive forms. They may also do this inconsistently.

Curly quotes alone do not prove LLM use. Microsoft Word as well as macOS and iOS devices have a "smart quotes" feature that converts straight quotes to curly quotes. Grammar correcting tools such as LanguageTool may also have such a feature. Curly quotation marks and apostrophes are common in professionally typeset works such as major newspapers. Citation tools like Citer (https://citer.toolforge.org/) may repeat those that appear in the title of a web page: for example,

> McClelland, Mac (September 27, 2017). "When 'Not Guilty' Is a Life Sentence" (https://www.nytimes.com/2017/09/27/magazine/when-not-guilty-is-a-life-sentence.html). *The New York Times*. Retrieved August 3, 2025.

Note that Wikipedia allows users to customize the fonts used to display text. Some fonts display matched curly apostrophes as straight, in which case the distinction is invisible to the user. Additionally, Gemini and Claude models typically do not use curly quotes.

Subject lines

User messages and unblock requests generated by AI chatbots sometimes begin with text that is intended to be pasted into the *Subject* field on an email form.

Examples

> Subject: Request for Permission to Edit Wikipedia Article - "Dog"

> —From this revision to Talk:Dog

> Subject: Request for Review and Clarification Regarding Draft Article

> —From this revision to Wikipedia:WikiProject Articles for creation/Help desk

Communication intended for the user

Collaborative communication

> Words to watch: ***I hope this helps**, **Of course!**, **Certainly!**, **You're absolutely right!**, **Would you like...**, **is there anything else**, **let me know**, **more detailed breakdown**, **here is a ...***

Editors sometimes paste text from an AI chatbot that was meant as correspondence, prewriting or advice, rather than article content. This may appear in article text or within comments (<-- -->). Chatbots prompted to produce a Wikipedia article or comment may also explicitly state that the text is meant for Wikipedia, and may mention various policies and guidelines in the output—often explicitly specifying that they're *Wikipedia*'s conventions.

Examples

> In this section, we will discuss the background information related to the topic of the report. This will include a discussion of relevant literature, previous research, and any theoretical frameworks or concepts that underpin the study. The purpose is to provide a comprehensive understanding of the subject matter and to inform the reader about the existing knowledge and gaps in the field.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1172646802&oldid=) to Metaphysics

> Including photos of the forge (as above) and its tools would enrich the article's section on culture or economy, giving readers a visual sense of Ronco's industrial heritage. Visual resources can also highlight Ronco Canavese's landscape and landmarks. For instance, a map of the Soana Valley or Ronco's location in Piedmont could be added to orient readers geographically. The village's scenery [...] could be illustrated with an image. Several such photographs are available (e.g., on Wikimedia Commons) that show Ronco's panoramic view, [...] Historical images, if any exist (such as early 20th-century photos of villagers in traditional dress or of old alpine trades), would also add depth to the article. Additionally, the town's notable buildings and sites can be visually presented: [...] Including an image of the Santuario di San Besso [...] could further engage readers. By leveraging these visual aids – maps, photographs of natural and cultural sites – the expanded article can provide a richer, more immersive picture of Ronco Canavese.

—From this revision to Ronco Canavese

> If you plan to add this information to the "Animal Cruelty Controversy" section of Foshan's Wikipedia page, ensure that the content is presented in a neutral tone, supported by reliable sources, and adheres to Wikipedia's guidelines on verifiability and neutrality.

—From this revision to Foshan

> Here's a template for your wiki user page. You can copy and paste this onto your user page and customize it further.

—From <u>this revision</u> to a user page

```
Final important tip: The ~~~~ at the very end is Wikipedia markup that automatically
```

—From <u>this revision</u> (https://en.wikipedia.org/w/index.php?title=&diff=1297191187&oldid=) to <u>Talk:Test automation management tools</u>; the message also <u>ends unexpectedly</u>

Knowledge-cutoff disclaimers and speculation about gaps in sources

Words to watch: **_as of [date],_**[c] **_Up to my last training update_**, **_as of my last knowledge update_**, **_While specific details are limited/scarce..._**, **_not widely available/documented/disclosed_**, **_...in the provided/available sources/search results..._**, **_based on available information ..._**

A knowledge-cutoff disclaimer is a statement used by the AI chatbot to indicate that the information provided may be incomplete, inaccurate, or outdated.

If an LLM has a fixed <u>knowledge cutoff</u> (usually the model's last training update), it is unable to provide any information on events or developments past that time, and it often outputs a disclaimer to remind the user of this cutoff, which usually takes the form of a statement that says the information provided is accurate only up to a certain date.

If an LLM with <u>retrieval-augmented generation</u> fails to find sources on a given topic, or if information is not included in sources a user provides, it often outputs a statement to that effect, which is similar to a knowledge-cutoff disclaimer. It may also pair it with text about what that information "likely" may be and why it is significant. This information is entirely <u>speculative</u> (including the very claim that it's "not documented") and may be based on loosely related topics or completely fabricated. When that unknown information is about an individual's personal life, this disclaimer often claims that the person "maintains a low profile," "keeps personal details private," etc. This is also speculative.

Examples

While specific details about Kumarapediya's history or economy <mark>are not extensively documented in readily available sources</mark>, ...

—From <u>this revision</u> to <u>Kumarapediya</u>

While specific information about the fauna of Studniční hora <mark>is limited in the provided search results</mark>, the mountain likely supports...

—From <u>this revision</u> to <u>Studniční hora</u>

Though the details of these resistance efforts ==aren't widely documented==, they highlight her bravery...

—From <u>this revision</u> to <u>Throwing Curves: Eva Zeisel</u>

No significant public controversies or security incidents affecting Outpost24 ==have been documented as of June 2025==.

—From <u>Draft:Outpost24</u>

==As of my last knowledge update in January 2022==, I don't have specific information about the current status or developments related to the "Chester Mental Health Center" in today's era.

—From <u>this revision (https://en.wikipedia.org/w/index.php?title=&diff=1186779926&oldid=)</u> to <u>Chester Mental Health Center</u>

Below is a detailed overview ==based on available information==:

—From <u>Draft:The Good, The Bad, The Dollar Menu 2</u>

Matthews Manamela ==keeps much of his personal life private,== choosing instead to focus public attention on his professional work and performances.

—From <u>Draft:Matthews Manamela</u>

As an underground release, detailed lyrics are ==not widely transcribed on major sites like Genius or AZLyrics==, likely due to the artist's limited mainstream exposure. My analysis is based on available track titles, featured artists, public song snippets from streaming platforms (e.g., Spotify, Apple Music, Deezer), and Honcho's overall discography themes. Where lyrics aren't fully accessible, ==I've inferred common motifs from similar trap tracks and Honcho's style.== ...For deeper insights, listening to tracks on platforms like Spotify or Deezer is recommended, as ==lyrics and production details aren't fully documented in public sources.==

—From Draft:Haiti Honcho

Phrasal templates and placeholder text

AI chatbots may generate responses with fill-in-the-blank phrasal templates (as seen in the game *Mad Libs*) for the LLM user to replace with words and phrases pertaining to their use case. However, some LLM users forget to fill in those blanks. Note that non-LLM-generated templates exist for drafts and new articles, such as Wikipedia:Artist biography article template/Preload and pages in Category:Article creation templates.

Examples

Subject: Concerns about Inaccurate Information

Dear Wikipedia

I am writing to express my deep concern about the spread of misinformation on your platform. Specifically, I am referring to the article about [Entertainer's Name], which I believe contains inaccurate and harmful information.

—From this revision to Talk:Kjersti Flaa

Subject: Edit Request for Wikipedia Entry

Dear Wikipedia Editors,

I hope this message finds you well. I am writing to request an edit for the Wikipedia entry

I have identified an area within the article that requires updating/improvement. [Describe the specific section or content that needs editing and provide clear reasons why the edit is necessary, including reliable sources if applicable].

—From this revision to Talk:Spaghetti

Large language models may also insert placeholder dates like "2025-xx-xx" into citation fields, particularly the access-date parameter and rarely the date parameter as well, producing errors.

Examples

```
<ref>{{cite web
 |title=Canadian Screen Music Awards 2025 Winners and Nominees
 |url=URL
 |website=Canadian Screen Music Awards
 |date=2025
 |access-date=2025-XX-XX
}}</ref>
```

```
<ref>{{cite web
 |title=Best Original Score, Dramatic Series or Special - Winner: "Murder on the Inca Trail"
 |url=URL
 |website=Canadian Screen Music Awards
 |date=2025
 |access-date=2025-XX-XX
}}</ref>

<ref>{{cite web
 |title=Best Original Score for a Narrative Feature Film - Nominee: "Don't Move"
 |url=URL
 |website=Canadian Screen Music Awards
 |date=2025
 |access-date=2025-XX-XX
}}</ref>

<ref>{{cite web
 |title=Best Original Score for a Short Film - Nominee: "T. Rex"
 |url=URL
 |website=Canadian Screen Music Awards
 |date=2025
 |access-date=2025-XX-XX
}}</ref>
```

—From this revision to Michelle Osis

Links to searches

- insource:/20[0-9][0-9]-(XX|xx)-(XX|xx)/ (https://en.wikipedia.org/w/index.php?search=insourc
e%3A%2F20%5B0-9%5D%5B0-9%5D-%28XX%7Cxx%29-%28XX%7Cxx%29%2F&title=S
pecial%3ASearch&profile=advanced&fulltext=1&ns0=1)

In some cases, LLM-generated citations may also contain placeholders in other fields.

Examples

{{cite web
|url=INSERT_SOURCE_URL_30
|title=Deputy Monitoring of Regional Assistance to Mobilized Soldiers
|date=2022-11-XX
|publisher=SOURCE_PUBLISHER
|accessdate=2024-07-21
}}

—From this revision to Dmitry Kuznetsov (politician)

LLM-generated infobox edits may contain comments stating that text or images should be added if sources are found. Note: Comments in infoboxes, especially older inboxes, are common—some templates automatically include them—and not an indicator of AI use. Anything but "Add ____", or variations on that specific wording, is actually more likely to indicate human text.

Examples

| leader_name = <!-- Add if available with citation -->

—From <u>this revision</u> to <u>Pindi Saidpur</u>

Markup

Use of Markdown

A lot of AI chatbots are not proficient in <u>wikitext</u>, the <u>markup language</u> used to instruct Wikipedia's <u>MediaWiki</u> software how to format an article. As wikitext is a niche markup language, found mostly on wikis running on MediaWiki and other MediaWiki-based platforms like <u>Miraheze</u>, LLMs tend to lack wikitext-formatted training data. While the corpora of chatbots did ingest millions of Wikipedia articles, these articles would not have been processed as text files containing wikitext syntax.

This is compounded by the fact that most chatbots are factory-tuned to use another, conceptually similar but much more diversely applied markup language: <u>Markdown</u>. Their system-level instructions often direct them to format outputs using Markdown, and the chatbot apps render its syntax as formatted text on a user's screen. For example, the system prompt for Claude Sonnet 3.5 (November 2024) includes:[14]

> Claude uses Markdown formatting. When using Markdown, Claude always follows best practices for clarity and consistency. It always uses a single space after hash symbols for headers (e.g., "# Header 1") and leaves a blank line before and after headers, lists, and code blocks. For emphasis, Claude uses asterisks or underscores consistently (e.g., italic or bold). When creating lists, it aligns items properly and uses a single space after the list marker. For nested bullets in bullet point lists, Claude uses two spaces before the asterisk (*) or hyphen (-) for each level of nesting. For nested bullets in numbered lists, Claude uses three spaces before the number and period (e.g., "1.") for each level of nesting.

As the above indicates, Markdown syntax is completely different from wikitext. Markdown uses asterisks (*) or underscores (_) instead of single-quotes (') for bold and italic formatting, hash symbols (#) instead of equals signs (=) for section headings, parentheses (()) instead of square brackets ([]) around URLs, and three symbols (---, ***, or ___) instead of four hyphens (----) for thematic breaks.

When told to "generate an article", chatbots often default to using Markdown for the generated output. This formatting is preserved in clipboard text by the copy functions on some chatbot platforms. If instructed to generate content for Wikipedia, the chatbot might "realize" the need to generate Wikipedia-compatible code, and might include a message like *Would you like me to ... turn this into actual Wikipedia markup format (`wikitext`)?*[d] in its output. If the chatbot is told to proceed, the resulting syntax is often rudimentary, syntactically incorrect, or both. The chatbot might put its attempted-wikitext content in a Markdown-style <u>fenced code block (https://www.markdownguide.org/extended-syntax/#fenced-code-blocks)</u> (its syntax for <u>WP:PRE</u>) surrounded by Markdown-based syntax and content, which may also be preserved by platform-specific copy-to-clipboard functions, leading to a telling footprint of both markup languages' syntax. This might include the appearance of three backticks in the text, such as: ```wikitext.[e]

The presence of faulty wikitext syntax mixed with Markdown syntax is a strong indicator that content is LLM-generated, especially if in the form of a fenced Markdown code block. However, Markdown *alone* is not such a strong indicator. Software developers, researchers, technical writers, and experienced internet users frequently use Markdown in tools like Obsidian and GitHub, and on platforms like Reddit (https://support.reddithelp.com/hc/en-us/articles/360043033952-Formatting-Guide), Discord (https://support.discord.com/hc/en-us/articles/210298617-Markdown-Text-101-Chat-Formatting-Bold-Italic-Underline), and Slack (https://slack.com/help/articles/202288908-Format-your-messages). Some writing tools and apps, such as iOS Notes, Google Docs, and Windows Notepad, support Markdown editing or exporting. The increasing ubiquity of Markdown may also lead new editors to expect or assume Wikipedia to support Markdown by default.

Examples

See also: Markdown § Examples

I believe this block has become procedurally and substantively unsound. Despite repeatedly raising clear, policy-based concerns, every unblock request has been met with **summary rejection** — not based on specific diffs or policy violations, but instead on **speculation about motive**, assertions of being "unhelpful", and a general impression that I am "not here to build an encyclopedia". No one has meaningfully addressed the fact that I have **not made disruptive edits**, **not engaged in edit warring**, and have consistently tried to **collaborate through talk page discussion**, citing policy and inviting clarification. Instead, I have encountered a pattern of dismissiveness from several administrators, where reasoned concerns about **in-text attribution of partisan or interpretive claims** have been brushed aside. Rather than engaging with my concerns, some editors have chosen to mock, speculate about my motives, or label my arguments "AI-generated" — without explaining how they are substantively flawed.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1284964006&oldid=) to a user talk page

- The Wikipedia entry does not explicitly mention the "Cyberhero League" being recognized as a winner of the World Future Society's BetaLaunch Technology competition, as detailed in the interview with THE FUTURIST (https://consciouscreativity.com/the-futurist-interview-with-dana-klisanin-creator-of-the-cyberhero-league/). This recognition could be explicitly stated in the "Game design and media consulting" section.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1290202502&oldid=) to Talk:Dana Klisanin

Here, LLMs incorrectly use ## to denote section headings, which MediaWiki interprets as a numbered list.

1. 1. Geography

Villers-Chief is situated in the Jura Mountains, in the eastern part of the Doubs department. [...]

1. 1. History

Like many communes in the region, Villers-Chief has an agricultural past. [...]

1. 1. Administration

Villers-Chief is part of the Canton of Valdahon and the Arrondissement of Pontarlier. [...]

1. 1. Population

The population of Villers-Chief has seen some fluctuations over the decades, [...]

—From this revision to Villers-Chief

Broken wikitext

Since AI chatbots are typically not proficient in wikitext and templates, they often produce faulty syntax. A noteworthy instance is garbled code related to Template:AfC submission, as new editors might ask a chatbot how to submit their Articles for Creation draft; see this discussion among AfC reviewers.

Examples

Note the badly malformed category link which appears to be a result of code that provides day information in the LLM's Markdown parser:

```
[[Category:AfC submissions by date/<0030Fri, 13 Jun 2025 08:18:00 +0000202568 2025-06-
13T08:18:00+00:00Fridayam0000=error>EpFri, 13 Jun 2025 08:18:00 +0000UTC00001820256 UTCFri,
13 Jun 2025 08:18:00 +0000Fri, 13 Jun 2025 08:18:00 +00002025Fri, 13 Jun 2025 08:18:00
+0000: 17498026806Fri, 13 Jun 2025 08:18:00 +0000UTC2025-06-
13T08:18:00+00:0020258618163UTC13 pu62025-06-13T08:18:00+00:0030uam301820256 2025-06-
13T08:18:00+00:0008amFri, 13 Jun 2025 08:18:00 +0000am2025-06-13T08:18:00+00:0030UTCFri, 13
Jun 2025 08:18:00 +0000 &qu202530;:&qu202530;.</0030Fri, 13 Jun 2025 08:18:00
+0000202568>June 2025|sandbox]]
```

—From this revision to User:Dr. Omokhudu Idogho/sandbox

turn0search0

ChatGPT may include 􀀀cite􀀀turn0search0􀀀 (surrounded by Unicode points in the Private Use Area) at the ends of sentences, with the number after "search" increasing as the text progresses. There also exists an alternate shorter form with only the increasing number surrounded by PUA Unicode like

🔗0🔗. These are places where the chatbot links to an external site, but a human pasting the conversation into Wikipedia has that link converted into placeholder code. This was first observed in February 2025.

A set of images in a response may also render as 🔗i🔗`turn0image0`🔗`turn0image1`🔗`turn0image4`🔗`turn0image5`🔗. Rarely, other markup of a similar style, such as 🔗`cite`🔗`turn0news0`🔗 (example), 🔗`cite`🔗`turn1file0`🔗 (example), or 🔗`cite`🔗*generated-reference-identifier*🔗 (example), may appear.

Examples

> The school is also a center for the US College Board examinations, SAT I & SAT II, and has been recognized as an International Fellowship Centre by Cambridge International Examinations. 🔗cite🔗turn0search1🔗 For more information, you can visit their official website: 🔗cite🔗turn0search0🔗

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1274664396&oldid=) to List of English-medium schools in Bangladesh

> - **Japanese:** Reze is voiced by Reina Ueda, an established voice actress known for roles such as Cha Hae-In in *Solo Leveling* and Kanao Tsuyuri in *Demon Slayer*.🔗2🔗
> - **English:** In the English dub of the anime film, Reze is voiced by Alexis Tipton, noted for her work in series such as *Kaguya-sama: Love is War*.🔗3🔗
>
> [...]
>
> The film itself holds a high rating on **Rotten Tomatoes** and has been described as a major anime release of 2025, indicating strong overall reception for the Reze Arc storyline and its adaptation.🔗5🔗

—From Draft:Reze (Chainsaw Man)

Links to searches

- turn0search0 OR turn0search1 OR turn0search2 OR turn0search3 OR turn0search4 OR turn0search5 OR turn0search6 OR turn0search7 (https://en.wikipedia.org/w/index.php?title=Special:Search&search=turn0search0+OR+turn0search1+OR+turn0search2+OR+turn0search3+OR+turn0search4+OR+turn0search5+OR+turn0search6+OR+turn0search7&ns0=1&fulltext=Search)
- turn0image0 OR turn0image1 OR turn0image2 OR turn0image3 OR turn0image4 OR turn0image5 OR turn0image6 OR turn0image7 (https://en.wikipedia.org/w/index.php?title=Special:Search&search=turn0image0+OR+turn0image1+OR+turn0image2+OR+turn0image3+OR+turn0image4+OR+turn0image5+OR+turn0image6+OR+turn0image7&ns0=1&fulltext=Search)
- insource:/turn0(search|image|news|file)[0-9]+/ (https://en.wikipedia.org/w/index.php?title=Special:Search&search=insource%3A%2Fturn0%28search%7Cimage%7Cnews%7Cfile%29%

Reference markup bugs: contentReference, oaicite, oai_citation, +1, attached_file, grok_card

Due to a bug, ChatGPT may add code in the form of `:contentReference[oaicite:0]{index=0}`, `Example+1`, or `oai_citation` in place of links to references in output text.

Examples

:contentReference{index=16}

1. **Ethnicity clarification**

```
  - :contentReference[oaicite:17]{index=17}
    * :contentReference[oaicite:18]{index=18} :contentReference[oaicite:19]{index=19}.
    * Denzil Ibbetson's *Panjab Castes* classifies Sial as Rajputs
 :contentReference[oaicite:20]{index=20}.
    * Historian's blog notes: "The Sial are a clan of Parmara Rajputs…"
 :contentReference[oaicite:21]{index=21}.
```

2. :contentReference{index=22}

```
  - :contentReference[oaicite:23]{index=23}
    > :contentReference[oaicite:24]{index=24} :contentReference[oaicite:25]{index=25}.
```

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1294765751&oldid=) to Talk:Sial (tribe).

📌 Key facts needing addition or correction:

1. **Group launch & meetings**

```
   *Independent Together* launched a "Zero Rates Increase Roadshow" on 15 June, with events
in Karori, Hataitai, Tawa, and Newtown  [oai_citation:0‡wellington.scoop.co.nz]
 (https://wellington.scoop.co.nz/?p=171473&utm_source=chatgpt.com).
```

2. **Zero-rates pledge and platform**

```
   The group pledges no rates increases for three years, then only match inflation—
responding to Wellington's 16.9% hike for 2024/25  [oai_citation:1‡en.wikipedia.org]
 (https://en.wikipedia.org/wiki/Independent_Together?utm_source=chatgpt.com).
```

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1296028135&oldid=) to Talk:Independent Together

This was created conjointly by technical committee ISO/IEC JTC 1/SC 27 (Information security, cybersecurity, and protection of privacy) IT Governance+3ISO+3ISO+3. It belongs to the ISO/IEC 27000 family that talks about information security management systems (ISMS) and related practice controls. Wikipedia+1. The standard gives guidance for information security controls for cloud service providers (CSPs) and cloud service customers (CSCs). Specifically adapted to cloud specific environments like responsibility, virtualization, dynamic provisioning, and multi-tenant infrastructure. Ignyte+3Microsoft Learn+3Google Cloud+3.

—From this revision to ISO/IEC 27017

As of fall 2025, tags like [attached_file:1], [web:1] have been seen at the end of sentences. This may be Perplexity-specific.[15]

During his time as CEO, Philip Morris's reputation management and media relations brought together business and news interests in ways that later became controversial, with effects still debated in contemporary regulatory and legal discussions.[attached_file:1]

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1316436509&oldid=) to Hamish Maxwell

Though Grok-generated text is rare compared to other chatbots, it may sometimes include XML-styled *grok_card* tags after citations.

Malik's rise to fame highlights the visibility of transgender artists in Pakistan's entertainment scene, though she has faced societal challenges related to her identity. [...]<grok-card data-id="e8ff4f" data-type="citation_card">

—From this revision to Draft:Mehak Malik

Links to searches

- "contentReference" OR "oaicite" OR "oai_citation" (https://en.wikipedia.org/w/index.php?search=%22contentReference%22+OR+%22oaicite%22+OR+%22oai_citation%22&title=Special%3ASearch)

attribution and attributableIndex

ChatGPT may add JSON-formatted code at the end of sentences in the form of (`{"attribution": {"attributableIndex":"X-Y"}}`), with X and Y being increasing numeric indices.

Examples

^[Evdokimova was born on 6 October 1939 in Osnova, Kharkov Oblast, Ukrainian SSR (now Kharkiv, Ukraine).]({"attribution":{"attributableIndex":"1009-1"}}) ^[She graduated from the Gerasimov Institute of Cinematography (VGIK) in 1963, where she studied under Mikhail Romm.]({"attribution":{"attributableIndex":"1009-2"}}) [oai_citation:0 ‡ IMDb] (https://www.imdb.com/name/nm0947835/?utm_source=chatgpt.com) [oai_citation:1‡maly.ru] (https://www.maly.ru/en/people/EvdokimovaA?utm_source=chatgpt.com)

—From Draft:Aleftina Evdokimova

Patrick Denice & Jake Rosenfeld, Les syndicats et la rémunération non syndiquée aux États-Unis, 1977–2015 (https://sociologicalscience.com/articles-v5-23-541/), ''Sociological Science'' (2018).]({"attribution":{"attributableIndex":"3795-0"}})

—From this diff to fr:Syndicalisme aux États-Unis

Non-existent or out-of-place categories

LLMs may hallucinate non-existent categories, sometimes for generic concepts that *seem like* plausible category titles (or SEO keywords), and sometimes because their training set includes obsolete and renamed categories. These will appear as red links. You may also find category redirects, such as the longtime spammer favorite Category:Entrepreneurs. Sometimes, broken categories may be deleted by reviewers, so if you suspect a page may be LLM-generated, it may be worth checking earlier revisions.

Of course, none of this section should be treated as a hard-and-fast rule. New users are unlikely to know about Wikipedia's style guidelines for these sections, and returning editors may be used to old categories that have since been deleted.

Examples

```
[[Category:American hip hop musicians]]
```

—From this revision to Draft:Paytra

rather than

```
[[Category:American hip-hop musicians]]
```

Non-existent templates

LLMs often hallucinate non-existent templates (especially plausible-sounding types of infoboxes) and template parameters. These will also appear as red links, and non-existent template parameters in existing templates have no effect. LLMs may also use templates that were deleted after their knowledge cutoff date.

Examples

```
{{Infobox ancient population
| name = Gangetic Hunter-Gatherer (GHG)
| image = [[File:GHG_reconstruction.png|250px]]
| caption = Artistic reconstruction of a Gangetic Hunter-Gatherer male, based on Mesolithic
skeletal data from the Ganga Valley
| regions = Ganga Valley (from Haryana to Bengal, between the Vindhyas and Himalayas)
| period = Mesolithic-Early Neolithic (10,000-5,000 BCE)
| descendants = Gangetic peoples, Indus Valley Civilisation, South Indian populations
| archaeological_sites = Bhimbetka, Sarai Nahar Rai, Mahadaha, Jhusi, Chirand
}}
```

—From this revision to Draft:Gangetic hunter-gatherers

rather than

```
{{Infobox archaeological culture
| name = Gangetic Hunter-Gatherer (GHG)
| map = [[File:GHG_reconstruction.png|250px]]
| mapcaption = Artistic reconstruction of a Gangetic Hunter-Gatherer male, based on
Mesolithic skeletal data from the Ganga Valley
| region = Ganga Valley (from Haryana to Bengal, between the Vindhyas and Himalayas)
| period = Mesolithic-Early Neolithic (10,000-5,000 BCE)
| followedby = Gangetic peoples, Indus Valley Civilisation, South Indian populations
| majorsites = Bhimbetka, Sarai Nahar Rai, Mahadaha, Jhusi, Chirand
}}
```

Citations

For non-AI-specific guidance about this, see Wikipedia:Fictitious references.

Broken external links

If a new article or draft has multiple citations with external links, and several of them are broken (e.g., returning 404 errors), this is a strong sign of an AI-generated page, particularly if the dead links are not found in website archiving sites like Internet Archive or Archive Today. Most links become broken over time, but these factors make it unlikely that the link was ever real.

Invalid DOI and ISBNs

For non-AI-specific guidance about DOI and ISBNs, see Wikipedia:DOI and Wikipedia:ISBN.

A checksum can be used to verify ISBNs. An invalid checksum is a very likely sign that an ISBN is incorrect, and citation templates display a warning if so. Similarly, DOIs are more resistant to link rot than regular hyperlinks. Unresolvable DOIs and invalid ISBNs can be indicators of hallucinated references.

Outdated *access-date*s

In some AI-assisted text, citations may include an *access-date* by default, but the date can look unexpectedly old relative to when the edit was made (for example, an article created in December 2025 containing multiple citations with `|access-date=12 December 2024`). This is not evidence by itself, but it can be a useful *pattern to check* when combined with other signs of low-quality drafting. Note that older *access-date* values can occur legitimately (copied citations, offline work, batch moves/merges).

DOIs that lead to unrelated articles

A LLM may generate references to non-existent scholarly articles with DOIs that appear valid but are, in reality, assigned to unrelated articles. Example passage generated by ChatGPT:

> Ohm's Law applies to many materials and components that are "ohmic," meaning their resistance remains constant regardless of the applied voltage or current. However, it does not hold for non-linear devices like diodes or transistors [1][2].
>
> 1. M. E. Van Valkenburg, "The validity and limitations of Ohm's law in non-linear circuits," Proceedings of the IEEE, vol. 62, no. 6, pp. 769–770, Jun. 1974. doi:10.1109/PROC.1974.9547 (https://doi.org/10.1109%2FPROC.1974.9547)
>
> 2. C. L. Fortescue, "Ohm's Law in alternating current circuits," Proceedings of the IEEE, vol. 55, no. 11, pp. 1934–1936, Nov. 1967. doi:10.1109/PROC.1967.6033 (https://doi.org/10.1109%2FPROC.1967.6033)

Both *Proceedings of the IEEE* citations are completely made up. The DOIs lead to different citations and have other problems as well. For instance, C. L. Fortescue was dead for 30+ years at the purported time of writing, and Vol 55, Issue 11 (https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=31102&punumber=5) does not list any articles that match anything remotely close to the information given in reference 2.

Book citations without page numbers or URLs

LLMs often generate book citations that do not include page numbers. This passage, for example, was generated by ChatGPT:

> Ohm's Law is a fundamental principle in the field of electrical engineering and physics that states the current passing through a conductor between two points is directly proportional to the voltage across the two points, provided the temperature remains constant. Mathematically, it is expressed as V=IR, where V is the voltage, I is the current, and R is the resistance. The law was formulated by German physicist Georg Simon Ohm in 1827, and it serves as a cornerstone in the analysis and design of electrical circuits [1].

1. Dorf, R. C., & Svoboda, J. A. (2010). Introduction to Electric Circuits (8th ed.). Hoboken, NJ: John Wiley & Sons. ISBN 9780470521571.

The book reference appears valid – a book on electric circuits would likely have information about Ohm's law – but without the page number, that citation is not useful for verifying the claims in the prose.

Some LLM-generated book citations include page numbers, and the book exists, but the cited pages do not verify the text. Signs to look out for: the book is on a somewhat general topic or frequently referenced in its field, and the citation does not include a URL (not mandatory for book citations, but editors creating legitimate book citations often include a link to an online version of the text). Example:

Analysts note that traditionalists often appeal to prudence, stability, and Edmund Burke's notion of "prescription," while reactionaries invoke moral urgency and cultural emergency, framing the present as a deviation from an idealized past. [1]

1. Goldwater, Barry (1960). The Conscience of a Conservative. Victor Publishing. p. 12.

This may look like a reasonable citation, but searching an online version of the book for "Burke" produces no results (https://www.google.com/books/edition/The_Conscience_of_a_Conservative/PKd_E QAAQBAJ?hl=en&gbpv=1&bsq=burke).

Incorrect or unconventional use of references

AI tools may have been prompted to include references, and make an attempt to do so as Wikipedia expects, but fail with some key implementation details or stand out when compared with conventions.

Examples

In the below example, note the incorrect attempt at re-using references. The tool used here was not capable of searching for non-confabulated sources (as it was done the day before Bing Deep Search launched) but nonetheless found one real reference. The syntax for re-using the references was incorrect.

In this case, the *Smith, R. J.* source – being the "third source" the tool presumably generated the link 'https://pubmed.ncbi.nlm.nih.gov/3' (which has a PMID reference of 3) – is also completely irrelevant to the body of the article. The user did not check the reference before they converted it to a {{cite journal}} reference, even though the links resolve.

The LLM in this case has diligently included the incorrect re-use syntax after every single full stop.

```
For over thirty years, computers have been utilized in the rehabilitation of individuals
with brain injuries. Initially, researchers delved into the potential of developing a
"prosthetic memory."<ref>Fowler R, Hart J, Sheehan M. A prosthetic memory: an application of
the prosthetic environment concept. ''Rehabil Counseling Bull''. 1972;15:80-85.</ref>
However, by the early 1980s, the focus shifted towards addressing brain dysfunction through
repetitive practice.<ref>{{Cite journal |last=Smith |first=R. J. |last2=Bryant |first2=R. G.
|date=1975-10-27 |title=Metal substitutions incarbonic anhydrase: a halide ion probe study
|url=https://pubmed.ncbi.nlm.nih.gov/3 |journal=Biochemical and Biophysical Research
Communications |volume=66 |issue=4 |pages=1281-1286 |doi=10.1016/0006-291x(75)90498-2
|issn=0006-291X |pmid=3}}</ref> Only a few psychologists were developing rehabilitation
software for individuals with Traumatic Brain Injury (TBI), resulting in a scarcity of
available programs.<sup>[3]</sup> Cognitive rehabilitation specialists opted for
```

```
commercially available computer games that were visually appealing, engaging, repetitive,
and entertaining, theorizing their potential remedial effects on neuropsychological
dysfunction.<sup>[3]</sup>
```

—From this revision to Cognitive orthotics

Some LLMs or chatbot interfaces use the character ↵ to indicate footnotes:

References

Would you like help formatting and submitting this to Wikipedia, or do you plan to post it yourself? I can guide you step-by-step through that too.

Footnotes

1. KLAS Research. (2024). *Top Performing RCM Vendors 2024*.
   https://klasresearch.com ↵ ↵²
2. PR Newswire. (2025, February 18). *CureMD AI Scribe Launch Announcement*.
   https://www.prnewswire.com/news-releases/curemd-ai-scribe ↵

—From this revision of Draft:CureMD

utm_source=

ChatGPT may add the UTM parameters `utm_source=openai` or `utm_source=chatgpt.com` to URLs that it is using as sources. Microsoft Copilot may add `utm_source=copilot.com` to URLs. Grok uses `referrer=grok.com`. Other LLMs, such as Gemini or Claude, use UTM parameters less often.[f]

Note: While this does definitively prove ChatGPT's involvement, it doesn't prove, on its own, that ChatGPT also generated the writing. Some editors use AI tools to find citations for existing text; this will be apparent in the edit history.

Examples

Following their marriage, Burgess and Graham settled in Cheshire, England, where Burgess serves as the head coach for the Warrington Wolves rugby league team. [https://www.theguardian.com/sport/2025/feb/11/sam-burgess-interview-warrington-rugby-league-luke-littler?utm_source=chatgpt.com]

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1277944793&oldid=) to Sam Burgess

Vertex AI documentation and blog posts describe watermarking, verification workflow, and configurable safety filters (for example, person-generation controls and safety thresholds). ([cloud.google.com](https://cloud.google.com/vertex-ai/generative-ai/docs/image/generate-images?utm_source=openai))

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1308699845&oldid=) to Draft:Nano Banana (Chatbot)

Links to searches

- utm_source=chatgpt.com (https://en.wikipedia.org/w/index.php?search=%utm_source=chatgpt.com%22&title=Special:Search&profile=advanced&fulltext=1&ns0=1)
- insource:"utm_source=chatgpt.com" (https://en.wikipedia.org/w/index.php?search=insource%3A%22utm_source%3Dchatgpt.com%22&title=Special%3ASearch&profile=advanced&fulltext=1&ns0=1)
- insource:"utm_source=openai" (https://en.wikipedia.org/w/index.php?search=insource%3A%22utm_source%3Dopenai%22&title=Special%3ASearch&profile=advanced&fulltext=1&ns0=1)
- insource:"referrer=grok.com" (https://en.wikipedia.org/w/index.php?search=insource%3A%22referrer%3Dgrok.com%22&title=Special%3ASearch&profile=advanced&fulltext=1&ns0=1)

Named references declared in references section but unused in article body

This section **needs expansion** with: Text describing this indicator should be added here. You can help by making an edit request (https://en.wikipedia.org/w/index.php?title=Wikipedia_talk%3ASigns_of_AI_writing&preload=Template%3ASubmit+an+edit+request%2Fpreload&action=edit&section=new&editintro=Template%3AEdit+fully-protected%2Feditintro&preloadtitle=Protected+edit+request+on+27+January+2026&preloadparams%5B%5D=edit+fully-protected&preloadparams%5B%5D=Wikipedia%3ASigns+of+AI+writing). *(December 2025)*

Examples

```
<references>
<ref name=\"fiercebiotech\">https://www.fiercebiotech.com/cro/parexel-co-founder-josef-von-
rickenbach-to-end-35-year-run-as-ceo</ref>
<ref name=\"statnews\">https://www.statnews.com/2018/03/16/parexel-josef-von-rickenbach-
cro/</ref>
<ref name=\"mclean\">https://www.mcleanhospital.org/news/three-prominent-community-members-
join-mcleans-board</ref>
<ref name=\"twst\">https://www.twst.com/bio/josef-h-von-rickenbach/</ref>
</references>
```

Result

Cite error: A <u>list-defined reference</u> named "\"fiercebiotech\"" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "\"statnews\"" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "\"mclean\"" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "\"twst\"" is not used in the content (see the <u>help page</u>).

—From <u>this revision</u> to <u>Draft:Josef von Rickenbach</u>

```
<references><ref name="wooart-about">[https://wooart.ca/about-caligomos-art About Caligomos
Art - WOO ART]</ref> <ref name="wooart-home">[https://wooart.ca/ Home - WOO ART]</ref> <ref
name="discover-leeds">[https://discoverdirectory.leedsgrenville.com/Home/View/woo-art-
gallery Woo Art Gallery - Discover Leeds Grenville]</ref> <ref name="book-amazon">Woo, John
HR. ''The Book of Caligomos Art''. Amazon KDP, 2025. ISBN 979-8-987654321-0.</ref>
</references>
```

Result

Cite error: A <u>list-defined reference</u> named "wooart-about" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "wooart-home" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "discover-leeds" is not used in the content (see the <u>help page</u>).

Cite error: A <u>list-defined reference</u> named "book-amazon" is not used in the content (see the <u>help page</u>).

—From <u>this revision</u> to <u>Draft:Caligomos Art</u>

Links to searches

- Category:Pages with incorrect ref formatting

Miscellaneous

Sudden shift in writing style

A sudden shift in an editor's writing style, such as unexpectedly flawless grammar compared to their other communication, may indicate the use of AI tools. Combining formal and casual writing styles is not exclusive to AI, but may be considered a sign. Using more formal prose in some writing may simply be a matter of code switching.

A mismatch of user location, national ties of the topic to a variety of English, and the variety of English used may indicate the use of AI tools. A human writer from India writing about an Indian university would probably not use American English; however, LLM outputs use American English by default, unless prompted otherwise.[16] Note that non-native English speakers tend to mix up English varieties, and such signs should raise suspicion only if there is a sudden and complete shift in an editor's English variety use.

Overwhelmingly verbose edit summaries

AI-generated edit summaries are often unusually long, written as formal, first-person paragraphs without abbreviations, and/or conspicuously itemize Wikipedia's conventions.

> Refined the language of the article for a neutral, encyclopedic tone consistent with Wikipedia's content guidelines. Removed promotional wording, ensured factual accuracy, and maintained a clear, well-structured presentation. Updated sections on history, coverage, challenges, and recognition for clarity and relevance. Added proper formatting and categorized the entry accordingly

—Edit summary from this revision to Khaama Press

> I formalized the tone, clarified technical content, ensured neutrality, and indicated citation needs. Historical narratives were streamlined, allocation details specified with regulatory references, propagation explanations made reader-friendly, and equipment discussions focused on availability and regulatory compliance, all while adhering to encyclopedic standards.

—Edit summary from this revision to 4-metre band

> **Concise edit summary:** Improved clarity, flow, and readability of the plot section; reduced redundancy and refined tone for better encyclopedic style.

—Edit summary from <u>this revision</u> to <u>Anaganaga (film)</u>

"Submission statements" in AFC drafts

This one is specific to drafts submitted by <u>Articles for Creation</u>. At least one LLM tends to insert "submission statements" supposedly intended for reviewers that supposedly explain why the subject is notable and why the draft meets Wikipedia guidelines. Of course, <u>all this actually does is</u> let reviewers know that the draft is LLM-generated, and should be declined or speedied without a second thought.

> Reviewer note (for AfC): This draft is a neutral and well-sourced biography of Portuguese public manager Jorge Patrão. All references are from independent, reliable sources (Público, Diário de Notícias, Jornal de Negócios, RTP, O Interior, Agência Lusa) covering his public career and cultural activity. It meets WP:RS and WP:BLP standards and demonstrates clear notability per WP:NBIO through: – Presidency of Serra da Estrela Tourism Region (1998–2013); – Presidency of Parkurbis – Covilhã Science and Technology Park; – Founding role in Rede de Judiarias de Portugal (member of the Council of Europe's European Routes of Jewish Heritage); – Authorship of the book "1677 – A Fábrica d'El-Rei"; – Founder/curator of the Beatriz de Luna Art Collection (Old Master focus). There is also a Portuguese version of this article at pt.wikipedia.org/wiki/Jorge_Patrão. Thank you for your review. -->

—Found at the top of <u>Draft:Jorge Patrão</u> (all the inevitable formatting errors are present in the original)

Pre-placed maintenance templates

Occasionally a new editor creates a draft that includes an <u>AFC</u> review template already set to "declined". The template is also devoid of content with no reviewer reasoning given. The LLM apparently offers to add an AFC submission template to the draft, and then provides something like `{{`<u>`AfC submission`</u>`|d}}`, in which the "d" parameter pre-declines the draft by substituting `{{`<u>`AfC submission/declined`</u>`}}`. The draft's contribution history reveals that this template was inserted at some point by the draft's creator. Invariably the creator then asks on <u>Wikipedia:WikiProject Articles for creation/Help desk</u> or one of the other help pages why the draft was declined with no feedback. The presence of a content-free "submission declined" header is a **strong** indicator that the draft was LLM-generated.

LLMs have been known to create pages that already have maintenance templates that shouldn't plausibly be there, including maintence tags and incorrect <u>protection</u> templates.

```
{{Short description|Advice on detecting AI-generated content}}
{{pp|small=yes}}
{{pp-move}}
{{Use American English|date=September 2022}}
{{Use mdy dates|date=February 2025}}
```

—From this revision to a user sandbox (later cut-and-paste moved to Émile Dufresne)

Links to searches

- you declined your own draft (https://en.wikipedia.org/w/index.php?search=%22you+declined +your+own+draft%22&title=Special%3ASearch&profile=advanced&fulltext=1&ns4=1&ns118 =1)

Signs of human writing

Age of text relative to ChatGPT launch

ChatGPT was launched to the public on November 30, 2022. Although OpenAI had similarly powerful LLMs before then, they were paid services and not easily accessible or known to lay people. Thus, it is very unlikely that any particular text added to Wikipedia **before November 30, 2022** was generated by an LLM. If an edit was made before this date, AI use can be safely ruled out for that revision. While some older text may display some of the AI signs given in this list, and even convincingly appear to have been AI-generated, the vastness of Wikipedia allows for these rare coincidences.

Ability to explain one's own editorial choices

Editors should be able to explain why they made an edit or mistake. For example, if an editor inserts a URL that appears fabricated, you can ask how the mix-up occurred instead of jumping to conclusions. If they can supply the correct link and explain it as a human error (perhaps a typo), or share the relevant passage from the real source, that points to an ordinary human error.

Ineffective indicators

False accusations of AI use can drive away new editors and foster an atmosphere of suspicion. Before claiming AI was used, consider if Dunning–Kruger effect and confirmation bias is clouding your judgement. Here are several somewhat commonly used indicators that are ineffective in LLM detection— and may even indicate the opposite.

- **Perfect grammar**: While modern LLMs are known for their high grammatical proficiency, many editors are also skilled writers or come from professional writing backgrounds. (See also § Sudden shift in English variety use.) Some may alternatively interpret AI as using "bad grammar", yet the prose may merely adhere to different prescriptions or stylistic principles, such as whether singular indefinite "they" is acceptable.

- **Combination of casual and formal registers, or language that sounds both "clinical" and "emotional"**: This may indicate the casual writing of a person in a technical field, such as computer science. It may also indicate youth, a preference for mixed registers, playfulness, or neurodivergence. Or it may simply be the result of multiple editors adding to a page.

- **"Bland" or "robotic" prose**: By default, modern LLMs tend toward effusive and verbose prose, as detailed above; while this tendency is formulaic, it may not scan as "robotic" to those unfamiliar with AI writing.[17]

- **"Fancy," "academic," or unusual words**: While LLMs disproportionately favor certain words and phrases, many of which are long and have difficult readability scores, the correlation does not extend to *all* "fancy," academic, or "advanced"-sounding prose.[1] Low-frequency and "unusual" words are also less likely to show up in AI-generated writing as they are statistically less common, unless they are proper nouns directly related to the topic.

- **Letter-like writing (in isolation)**: Although many talk page messages written with salutations, valedictions, subject lines, and other formalities after 2023 tend to appear AI-generated, letters and emails have conventionally been written in such ways *long* before modern LLMs existed. Human editors (particularly newer editors) may format their talk page comments similarly for various reasons, such as being more accustomed to formal communication, posting as part of a school assignment that requires such a tone, or simply mistaking the talk page for email. AI-generated talk page messages tend to have other tells, such as vertical lists[g], placeholders, or abrupt cutoffs.

- **Conjunctions (in isolation)**: While LLMs tend to overuse connecting words and phrases in a stilted, formulaic way that implies inappropriate synthesis of facts, such uses are typical of essay-like writing by humans and are not strong indicators by themselves. While many people are taught beginning a sentence with a coordinating conjunction is nonstandard (or at least bad style), such usage has precedence and is accepted by many style guides.

- **Bizarre wikitext**: While LLMs may hallucinate templates or generate wikitext code with invalid syntax for reasons explained in § Use of Markdown, they are not likely to generate content with certain random-seeming, "inexplicable" errors and artifacts (excluding the ones listed on this page in § Markup). Bizarrely placed HTML tags like <span> are more indicative of poorly programmed browser extensions or a known bug with Wikipedia's content translation tool (T113137). Misplaced syntax like `''Catch-22 i''s a satirical novel.` (rendered as "*Catch-22 i*s a satirical novel.") are more indicative of mistakes in VisualEditor, where such errors are harder to notice than in source editing.

Historical indicators

The following indicators were common in text generated by older AI models, but are much less frequent in newer models. They may still be useful for finding older undetected AI-generated edits. Dates are approximate.

Didactic disclaimers (2022–2024)

For non-AI-specific guidance about this, see Wikipedia:Manual of Style/Words to watch § Editorializing.

> Words to watch: **it's important/critical/crucial to note/remember/consider**, **worth noting**, **may vary**...

Older LLMs (~2023) often added disclaimers about topics being "important to remember." This frequently took the form of advice to an imagined reader regarding safety or controversial topics, or disambiguating topics that varied in different locales/jurisdictions. Several such disclaimers appear in OpenAI's GPT-4 system card as examples of "partial refusals".[18]

Examples

The emergence of these informal groups reflects a growing recognition of the interconnected nature of urban issues and the potential for ANCs to play a role in shaping citywide policies. <mark>However, it's important to note</mark> that these caucuses operate outside the formal ANC structure and their influence on policy decisions <mark>may vary</mark>.

—From this revision to Advisory Neighborhood Commission

<mark>It is crucial to differentiate</mark> the independent AI research company based in Yerevan, Armenia, which is the subject of this report, from these unrelated organizations <mark>to prevent confusion</mark>.

—From Draft:Robi Labs

<mark>It's important to remember</mark> that what's free in one country might not be free in another, so always check before you use something.

—From Wikimedia's LLM-generated Simple Summary (https://gitlab.wikimedia.org/repos/web/web-experiments-extension/-/commit/55fdbbb3decdc9b95ae0ef00e98b1108ddc3a498.diff) of Public domain

Section summaries

> Words to watch: ***In summary***, ***In conclusion***, ***Overall*** ...

When generating longer outputs (such as when told to "write an article"), older LLMs often added sections titled "Conclusion" or similar, and often ended paragraphs or sections by summarizing and restating its core idea.[16]

Examples

<mark>In summary</mark>, the educational and training trajectory for nurse scientists typically involves a progression from a master's degree in nursing to a Doctor of Philosophy in Nursing, followed by postdoctoral training in nursing research. This structured pathway ensures that nurse scientists acquire the necessary knowledge and skills to engage in rigorous research and contribute meaningfully to the advancement of nursing science.

—From this revision (https://en.wikipedia.org/w/index.php?title=&diff=1188230584&oldid=) in Nurse scientist

Prompt refusal

> Words to watch: **as an AI language model**, **as a large language model**, **I cannot offer medical advice, but I can...**, **I'm sorry** ...

In the past, AI chatbots occasionally declined to answer prompts as written, usually with apologies and reminders that they are AI language models. Attempting to be helpful, chatbots often gave suggestions or answers to alternative, similar requests. Outright refusals have become increasingly rare.

Examples

> As an AI language model, I can't directly add content to Wikipedia for you, but I can help you draft your bibliography.

—From this revision to Parmiter's Almshouse & Pension Charity

Abrupt cut offs

AI tools used to abruptly stop generating content if an excessive number of tokens had been used for a single response, and further responses required the user to select "continue generating", at least in the case of ChatGPT.

This method is not foolproof, as a malformed copy/paste from one's local computer can also cause this. It may also indicate a copyright violation rather than the use of an LLM.

See also

- Wikipedia:Artificial intelligence

Notes

a. Specifically, this guide is somewhat less useful for texts which are not "dry academic" writing. For example, the many tells specific to fiction (whispering woods, Elara Voss (http s://maxread.substack.com/p/who-is-elara-voss), etc.) are less relevant in Wikipedia and so they are not listed here.
b. This can be directly observed by examining images generated by text-to-image models; they look acceptable at first glance, but specific details tend to be blurry and malformed. This is especially true for background objects and text.
c. not unique to AI chatbots; is produced by the {{as of}} template
d. Example (deleted, administrators only)
e. Example of ```` ``wikitext ```` on a draft.
f. See T387903.

g. Example of a vertical list in a deletion discussion

References

1. Russell, Jenna; Karpinska, Marzena; Iyyer, Mohit (2025). *People who frequently use ChatGPT for writing tasks are accurate and robust detectors of AI-generated text* (https://acl anthology.org/2025.acl-long.267/). Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vienna, Austria: Association for Computational Linguistics. pp. 5342–5373. arXiv:2501.15654 (https://arxiv.or g/abs/2501.15654). doi:10.18653/v1/2025.acl-long.267 (https://doi.org/10.18653%2Fv1%2F 2025.acl-long.267). Archived (https://web.archive.org/web/20250829184825/https://aclanthol ogy.org/2025.acl-long.267/) from the original on August 29, 2025. Retrieved September 5, 2025 – via ACL Anthology.

2. Dugan, Liam; Hwang, Alyssa; Trhlik, Filip; Zhu, Andrew; Ludan, Josh Magnus; Xu, Hainiu; Ippolito, Daphne; Callison-Burch, Chris (2024). *RAID: A Shared Benchmark for Robust Evaluation of Machine-Generated Text Detectors* (https://aclanthology.org/2024.acl-long.67 4). Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Bangkok, Thailand: Association for Computational Linguistics. pp. 12463–12492. arXiv:2405.07940 (https://arxiv.org/abs/2405.07940). Archived (https://we b.archive.org/web/20250824132743/https://aclanthology.org/2024.acl-long.674/) from the original on August 24, 2025. Retrieved November 8, 2025.

3. Rudnicka, Karolina (July 9, 2025). "Each AI chatbot has its own, distinctive writing style— just as humans do" (https://www.scientificamerican.com/article/chatgpt-and-gemini-ai-have- uniquely-different-writing-styles). *Scientific American*. Retrieved January 18, 2026.

4. "10 Ways AI Is Ruining Your Students' Writing" (https://www.chronicle.com/article/10-ways-ai -is-ruining-your-students-writing). *Chronicle of Higher Education*. September 16, 2025. Archived (https://web.archive.org/web/20251001071208/https://www.chronicle.com/article/1 0-ways-ai-is-ruining-your-students-writing/) from the original on October 1, 2025. Retrieved October 1, 2025.

5. Juzek, Tom S.; Ward, Zina B. (2025). *Why Does ChatGPT "Delve" So Much? Exploring the Sources of Lexical Overrepresentation in Large Language Models* (https://aclanthology.org/ 2025.coling-main.426.pdf) (PDF). Findings of the Association for Computational Linguistics: ACL 2025. Association for Computational Linguistics. arXiv:2412.11385 (https://arxiv.org/ab s/2412.11385). Archived (https://web.archive.org/web/20250121111136/https://aclanthology. org/2025.coling-main.426.pdf) (PDF) from the original on January 21, 2025. Retrieved October 13, 2025 – via ACL Anthology.

6. Reinhart, Alex; Markey, Ben; Laudenbach, Michael; Pantusen, Kachatad; Yurko, Ronald; Weinberg, Gordon; Brown, David West. "Do LLMs write like humans? Variation in grammatical and rhetorical styles" (http://arxiv.org/abs/2410.16107). Retrieved December 4, 2025.

7. Geng, Mingmeng; Trotta, Roberto. "Human-LLM Coevolution: Evidence from Academic Writing" (https://aclanthology.org/2025.findings-acl.657.pdf) (PDF). *aclanthology.org*. Retrieved December 17, 2025.

8. Kobak, Dmitry; González-Márquez, Rita; Horvát, Emőke-Ágnes; Lause, Jan (July 2, 2025). "Delving into LLM-assisted writing in biomedical publications through excess vocabulary" (ht tps://www.science.org/doi/10.1126/sciadv.adt3813). *Science Advances*. **11** (27). doi:10.1126/sciadv.adt3813 (https://doi.org/10.1126%2Fsciadv.adt3813). ISSN 2375-2548 (https://search.worldcat.org/issn/2375-2548). PMC 12219543 (https://www.ncbi.nlm.nih.gov/ pmc/articles/PMC12219543). PMID 40601754 (https://pubmed.ncbi.nlm.nih.gov/40601754). Retrieved November 21, 2025.

9. Kriss, Sam (December 3, 2025). "Why Does A.I. Write Like … That?" (https://www.nytimes.com/2025/12/03/magazine/chatbot-writing-style.html). *The New York Times*. Retrieved December 6, 2025.

10. Kousha, Kayvan; Thelwall, Mike (2025). *How much are LLMs changing the language of academic papers after ChatGPT? A multi-database and full text analysis* (https://arxiv.org/pdf/2509.09596). ISSI 2025 Conference. arXiv:2509.09596 (https://arxiv.org/abs/2509.09596). Archived (https://web.archive.org/web/20250914165435/https://arxiv.org/pdf/2509.09596) from the original on September 14, 2025. Retrieved November 4, 2025.

11. Merrill, Jeremy B.; Chen, Szu Yu; Kumer, Emma (November 13, 2025). "What are the clues that ChatGPT wrote something? We analyzed its style" (https://www.washingtonpost.com/technology/interactive/2025/how-detect-chatgpt-em-dash/). *The Washington Post*. Retrieved November 14, 2025.

12. Geng, Mingmeng; Trotta, Roberto. "Is ChatGPT Transforming Academics' Writing Style?" (https://arxiv.org/abs/2404.08627). Retrieved January 8, 2026.

13. Robbins, Hollis. "How to Tell if Something is AI Written" (https://hollisrobbinsanecdotal.substack.com/p/how-to-tell-if-something-is-ai-written). *Anecdotal Value*. Substack. Retrieved December 7, 2025.

14. "System Prompts" (https://platform.claude.com/docs/en/release-notes/system-prompts#claude-sonnet-3-5). *Claude Docs*. Anthropic. Retrieved January 9, 2026.

15. "Unproductive Interpretation of Work and Employment as Misinformation?" (https://www.laetusinpraesens.org/docs20s/workeco.php). Archived (https://web.archive.org/web/20250902133810/https://www.laetusinpraesens.org/docs20s/workeco.php) from the original on September 2, 2025. Retrieved October 21, 2025.

16. Ju, Da; Blix, Hagen; Williams, Adina (2025). *Domain Regeneration: How well do LLMs match syntactic properties of text domains?* (https://aclanthology.org/2025.findings-acl.120). Findings of the Association for Computational Linguistics: ACL 2025. Vienna, Austria: Association for Computational Linguistics. pp. 2367–2388. arXiv:2505.07784 (https://arxiv.org/abs/2505.07784). doi:10.18653/v1/2025.findings-acl.120 (https://doi.org/10.18653%2Fv1%2F2025.findings-acl.120). Archived (https://web.archive.org/web/20250815014117/https://aclanthology.org/2025.findings-acl.120/) from the original on August 15, 2025. Retrieved October 4, 2025 – via ACL Anthology.

17. Murray, Nathan; Tersigni, Elisa (July 21, 2024). "Can instructors detect AI-generated papers? Postsecondary writing instructor knowledge and perceptions of AI" (https://journals.sfu.ca/jalt/index.php/jalt/article/view/1895). *Journal of Applied Learning & Teaching*. **7** (2). doi:10.37074/jalt.2024.7.2.12 (https://doi.org/10.37074%2Fjalt.2024.7.2.12). ISSN 2591-801X (https://search.worldcat.org/issn/2591-801X). Retrieved November 21, 2025.

18. "GPT-4 System Card" (https://cdn.openai.com/papers/gpt-4-system-card.pdf) (PDF). *OpenAI*. Retrieved December 16, 2025.

Further reading

- Kriss, Sam (December 3, 2025). "Why Does A.I. Write Like … That?" (https://www.nytimes.com/2025/12/03/magazine/chatbot-writing-style.html). *The New York Times Magazine*. Retrieved December 6, 2025.