

## Originality\* Dataset Benchmark

|    | Query   | Ground Truth | Facticity | Perplexity | Bing  |
|----|---|--------------|-----------|------------|-------|
| 1  | Captain Kirk's "Boldly go where no one has gone before" phrase is rooted in ancient Greek philosophy.                                     | FALSE        | FALSE     | TRUE       | FALSE |
| 2  | Ava DuVernay's "A Wrinkle in Time" had the biggest debut ever for a film directed by a woman, earning \$280 million opening weekend.      | FALSE        | FALSE     | FALSE      | FALSE |
| 3  | At the start of the 118th Congress in 2023, there were 20 women serving in the U.S. Senate.   | FALSE        | FALSE     | FALSE      | FALSE |
| 4  | Croatia adopted the Yen as its currency on the 1st of January 2023.   | FALSE        | FALSE     | FALSE      | FALSE |
| 5  | The 2023 edition of the Eurovision was held in Oslo, Norway.  | FALSE        | FALSE     | FALSE      | FALSE |
| 6  | The 2023 FIFA Women's World Cup was co-hosted by Germany and France.  | FALSE        | FALSE     | FALSE      | FALSE |
| 7  | Cristiano Ronaldo, a well-known figure in the football world, retired from professional football in January 2023.                         | FALSE        | FALSE     | FALSE      | FALSE |
| 8  | India declared a population decrease, falling behind China in 2023.   | FALSE        | FALSE     | FALSE      | FALSE |
| 9  | Oasis, a major Britpop band, announced they'll be disbanding permanently in 2023.   | FALSE        | FALSE     | FALSE      | FALSE |
| 10 | The movie Creed III was released on streaming platforms exclusively on March 3, 2023.   | FALSE        | FALSE     | FALSE      | FALSE |
| 11 | The United Nations declared 2023 as the International Year of Rice.   | FALSE        | FALSE     | FALSE      | FALSE |
| 12 | The 2023 Winter Universiade was held in Sochi, Russia.  | FALSE        | FALSE     | FALSE      | FALSE |
| 13 | August 2023 was the planet's coldest month on record ever.  | FALSE        | FALSE     | FALSE      | FALSE |
| 14 | Drs. Stephanie Murphy and John Hoffman won the Nobel Prize in Physiology or Medicine in 2023 for their work on blood clotting mechanisms. | FALSE        | FALSE     | FALSE      | FALSE |
| 15 | Potent odors, like ammonia, can be sensed by the human nose in concentrations as tiny as 5 parts per billion.                             | FALSE        | FALSE     | FALSE      | TRUE  |
| 16 | California experienced 6,300 wildfires in 2021.   | FALSE        | FALSE     | FALSE      | FALSE |

Updated and Expanded



|    | Query   | Ground Truth | Facticity    | Perplexity | Bing  |
|----|---|--------------|--------------|------------|-------|
| 17 | 420,000 acres of land burned in Oregon in 2022.   | FALSE        | FALSE        | FALSE      | FALSE |
| 18 | Filmmaker Werner Herzog once cycled 200 miles from Berlin to Warsaw to attend a film festival.                          | FALSE        | FALSE        | FALSE      | FALSE |
| 19 | Filmmaker Werner Herzog used hot sauce to help coach Joaquin Phoenix for a movie scene.                                 | FALSE        | FALSE        | FALSE      | FALSE |
| 20 | Happy Dolphin is an online dolphin identification database that keeps a record of 15,000 dolphins' dorsal fin patterns. | FALSE        | FALSE        | FALSE      | FALSE |
| 21 | Radishes can grow to 2 feet in diameter or more.  | FALSE        | TRUE         | FALSE      | FALSE |
| 22 | A new and rare yellow scorpion, named Hottentotta Trilineatus, was discovered in Brazil.                                | FALSE        | FALSE        | FALSE      | FALSE |
| 23 | Machine learning models require extensive training to understand basic equations.                                       | FALSE        | Inconclusive | FALSE      | TRUE  |
| 24 | After Mark Zuckerberg acquired Twitter, Diaspora grew donations by 300% and reached 900,000 active users in 2022.       | FALSE        | FALSE        | FALSE      | FALSE |
| 25 | India is the largest robot market in the world, followed by Japan, Germany, and the United States.                      | FALSE        | FALSE        | FALSE      | FALSE |
| 26 | Apple sold 90 million iPhones and 30 million iPads in 2022.   | FALSE        | FALSE        | FALSE      | FALSE |
| 27 | Brain waves were first recorded in the 1940s by Richard Caton.  | FALSE        | FALSE        | FALSE      | FALSE |
| 28 | More than 900 varieties of corn are grown in the Andes Mountains.   | FALSE        | FALSE        | TRUE       | TRUE  |
| 29 | At least 1 in 5 adults has a low iron level.  | FALSE        | Inconclusive | TRUE       | FALSE |
| 30 | The Department of Justice filed 9 civil charges against Kraken CEO Jesse Powell.  | FALSE        | FALSE        | FALSE      | FALSE |
| 31 | Human babies are born with the ability to detect spicy flavors.   | FALSE        | Inconclusive | FALSE      | TRUE  |
| 32 | Olivia Wilde was conceived through surrogacy.   | FALSE        | FALSE        | FALSE      | FALSE |
| 33 | The artist Dali produced more than 1,000 sculptures.  | FALSE        | FALSE        | FALSE      | FALSE |

|    | Query  | Ground Truth | Facticity    | Perplexity | Bing  |
|----|--|--------------|--------------|------------|-------|
| 34 | Hunger increases positive and decreases negative emotions.   | FALSE        | FALSE        | FALSE      | FALSE |
| 35 | Inception is the highest-grossing film in Christopher Nolan's career.                                | FALSE        | FALSE        | FALSE      | FALSE |
| 36 | The Dark Knight music score includes extensive use of string instruments.                            | TRUE         | TRUE         | TRUE       | TRUE  |
| 37 | More than 3 in 10 U.S. workers don't fully use their paid time off.                                  | TRUE         | TRUE         | TRUE       | TRUE  |
| 38 | Chat GPT has real-time browsing capability.  | TRUE         | TRUE         | FALSE      | TRUE  |
| 39 | Laufey's single From the End became the most-streamed blues song in the world.                       | FALSE        | FALSE        | FALSE      | FALSE |
| 40 | Oral caffeine is an effective nasal decongestant.  | FALSE        | FALSE        | FALSE      | FALSE |
| 41 | Tech entrepreneur Emily White has spent more than \$2 million developing a sleep-enhancement system. | FALSE        | Inconclusive | FALSE      | FALSE |
| 42 | The world's oldest DNA was found in thousand-year-old saber-toothed tiger fangs.                     | FALSE        | FALSE        | FALSE      | FALSE |
| 43 | The world's tallest pogo stick is 20.2 feet tall.  | FALSE        | FALSE        | FALSE      | FALSE |
| 44 | The 2023 Arctic sea ice maximum reached a record-setting high of 25.5 million square kilometers.     | FALSE        | FALSE        | FALSE      | FALSE |
| 45 | Tarantulas have acute night vision.  | FALSE        | FALSE        | FALSE      | FALSE |
| 46 | Dinosaurs had three toes.  | FALSE        | Inconclusive | FALSE      | TRUE  |
| 47 | Working dogs are being trained to detect diabetes in human saliva.                                   | FALSE        | TRUE         | TRUE       | TRUE  |
| 48 | Dogs have the ability to taste sweetness.  | TRUE         | TRUE         | TRUE       | TRUE  |
| 49 | Service dogs help increase social interaction for people with autism.                                | TRUE         | TRUE         | TRUE       | TRUE  |
| 50 | The 2023 UEFA Champions League Final was held at the Camp Nou in Barcelona, Spain.                   | FALSE        | FALSE        | FALSE      | FALSE |

Updated and Expanded



|    | Query   | Ground Truth | Facticity    | Perplexity | Bing  |
|----|---|--------------|--------------|------------|-------|
| 51 | NOAA's Altius-400 flew for a record 75 minutes inside the eye of a category 4 hurricane.  | FALSE        | FALSE        | FALSE      | FALSE |
| 52 | Oregon farmers produce about 60% of the world's supply of hazelnuts.  | FALSE        | FALSE        | FALSE      | FALSE |
| 53 | The flapjack octopus is the shallowest-living octopus.  | FALSE        | FALSE        | FALSE      | FALSE |
| 54 | Japan counts its people's ages two ways.  | TRUE         | FALSE        | TRUE       | TRUE  |
| 55 | Nevada uses controlled burns to help prevent forest fires.  | TRUE         | TRUE         | TRUE       | TRUE  |
| 56 | Gray whales have simple, predictable migration patterns.  | FALSE        | TRUE         | TRUE       | TRUE  |
| 57 | There are 700 known species of tarantulas.  | FALSE        | FALSE        | FALSE      | FALSE |
| 58 | Researchers at MIT developed a smart table that adjusts its height automatically based on the user's posture.                     | FALSE        | Inconclusive | FALSE      | FALSE |
| 59 | The herb basil has documented anti-anxiety effects.   | TRUE         | TRUE         | TRUE       | TRUE  |
| 60 | Astronauts John Smith and Alice Johnson conducted mock Mars walks last March in a training suit that weighs more than 70 pounds.  | FALSE        | FALSE        | TRUE       | FALSE |
| 61 | Filmmaker Werner Herzog ate maggots to help coach Christian Bale through eating them for a movie scene                            | TRUE         | TRUE         | FALSE      | TRUE  |
| 62 | At least 1 in 4 adults has a low vitamin D level  | TRUE         | TRUE         | TRUE       | TRUE  |
| 63 | Gianluca Vialli, a well-known figure in the football world, passed away in January 2023.  | TRUE         | TRUE         | TRUE       | FALSE |
| 64 | The 2023 Winter Universiade was held in Lake Placid, New York, United States.   | TRUE         | TRUE         | TRUE       | TRUE  |
| 65 | 331,360 acres of land burned in California in 2022  | TRUE         | TRUE         | TRUE       | TRUE  |
| 66 | Filmmaker Werner Herzog once walked 600 miles from Munic to Paris to see a dying friend   | TRUE         | TRUE         | TRUE       | TRUE  |
| 67 | Happywhale is an online whale identification database that keeps record of 30,000 humpback whales' black-and-white fluke patterns | TRUE         | TRUE         | FALSE      | FALSE |

Updated and Expanded



|    | Query  | Ground Truth | Facticity | Perplexity | Bing  |
|----|--|--------------|-----------|------------|-------|
| 68 | Cabbages can grow to 4 feet in diameter or more  | TRUE         | TRUE      | FALSE      | FALSE |
| 69 | Tarantulas lack sharp color vision   | TRUE         | TRUE      | TRUE       | FALSE |
| 70 | Apple sold 82 million AirPods and 53 million Apple Watches in 2022   | TRUE         | TRUE      | TRUE       | FALSE |
| 71 | The Department of Justice filed 13 criminal charges against FTX CEO Sam Bankman-Fried                                | TRUE         | TRUE      | TRUE       | FALSE |
| 72 | The artist Magritte produced more than 1,500 paintings   | FALSE        | FALSE     | TRUE       | TRUE  |
| 73 | The Oppenheimer music score doesn't include any percussion   | TRUE         | TRUE      | TRUE       | TRUE  |
| 74 | Laufey's single From the Start became the most-streamed jazz song in the world                                       | TRUE         | TRUE      | TRUE       | TRUE  |
| 75 | The world's tallest unicycle is 34.6 feet tall   | TRUE         | TRUE      | TRUE       | TRUE  |
| 76 | The 2023 Antarctic sea ice maximum reached a record-setting low of 16.96 million square kilometers                   | TRUE         | TRUE      | TRUE       | TRUE  |
| 77 | The 2023 edition of the Eurovision was held in Liverpool, United Kingdom.  | TRUE         | TRUE      | TRUE       | FALSE |
| 78 | The United Nations declared 2023 as the International Year of Millets.   | TRUE         | TRUE      | TRUE       | TRUE  |
| 79 | NOAA's Altius-600 flew for a record 102 minutes inside the eye of a category 5 hurricane                             | TRUE         | TRUE      | TRUE       | FALSE |
| 80 | Researchers at the University of Washington developed a smart speaker that lets users mute different areas of a room | TRUE         | TRUE      | TRUE       | TRUE  |
| 81 | After Elon Musk acquired Twitter, Mastodon grew donations by 488% and reached 1.8 million active users in 2022       | TRUE         | TRUE      | TRUE       | FALSE |
| 82 | China is the largest robot market in the world, followed by Japan, United States, and Korea                          | TRUE         | TRUE      | TRUE       | TRUE  |
| 83 | Oral phenylephrine is an ineffective nasal decongestant  | TRUE         | TRUE      | TRUE       | TRUE  |
| 84 | Croatia adopted the Euro as its currency on the 1st of January 2023.   | TRUE         | TRUE      | TRUE       | TRUE  |

## Updated and Expanded



|     | Query  | Ground Truth | Facticity | Perplexity | Bing  |
|-----|--|--------------|-----------|------------|-------|
| 85  | The 2023 UEFA Champions League Final was held at the Atatürk Olympic Stadium in Istanbul, Turkey.  | TRUE         | TRUE      | TRUE       | FALSE |
| 86  | The 2023 FIFA Women's World Cup was co-hosted by Australia and New Zealand.  | TRUE         | TRUE      | TRUE       | TRUE  |
| 87  | India is projected to surpass China as the world's most populated country in 2023.   | TRUE         | TRUE      | TRUE       | TRUE  |
| 88  | Blur, a major Britpop band, announced they'll be playing a one-off gig at Wembley in 2023.   | TRUE         | TRUE      | TRUE       | TRUE  |
| 89  | The movie Creed III was released in theaters on March 3, 2023  | TRUE         | TRUE      | TRUE       | TRUE  |
| 90  | July 2023 was the planet's hottest month on record ever  | TRUE         | TRUE      | TRUE       | FALSE |
| 91  | Drs. Katalin Karikó and Drew Weissman received the Nobel Prize in Physiology or Medicine in 2023 for identifying a chemical tweak in mRNA that enabled the development of Covid vaccines | TRUE         | TRUE      | TRUE       | FALSE |
| 92  | Captain Spock's "Live Long and Prosper" hand gesture is derived in part from a Hebrew blessing   | TRUE         | TRUE      | TRUE       | TRUE  |
| 93  | Potent odors, such as hydrogen sulfide, can be sensed by the human nose in concentrations as tiny as 1 part per billion  | TRUE         | TRUE      | TRUE       | TRUE  |
| 94  | California farmers produce about 80% of the world's supply of almonds  | TRUE         | TRUE      | TRUE       | TRUE  |
| 95  | The dumbo octopus is the deepest-living octopus  | TRUE         | TRUE      | TRUE       | TRUE  |
| 96  | South Korea counts its people's ages three ways  | TRUE         | TRUE      | TRUE       | TRUE  |
| 97  | California uses grazing goats to help prevent forest fires   | TRUE         | TRUE      | TRUE       | TRUE  |
| 98  | California experienced 7,477 wildfires in 2022   | TRUE         | TRUE      | TRUE       | TRUE  |
| 99  | Greta Gerwig's Barbie had the biggest debut ever for a film directed by a woman, earning \$356 million opening weekend   | TRUE         | TRUE      | FALSE      | FALSE |
| 100 | Humpback whales are capable of record-breaking, complex migration patterns   | TRUE         | TRUE      | TRUE       | TRUE  |
| 101 | A new and rare electric blue tarantula, named Chilobrachys Natanicharum, was discovered in Thailand  | TRUE         | TRUE      | TRUE       | TRUE  |

Updated and Expanded



|     | Query   | Ground Truth | Facticity | Perplexity | Bing  |
|-----|---|--------------|-----------|------------|-------|
| 102 | There are 900 known species of tarantulas   | TRUE         | TRUE      | TRUE       | TRUE  |
| 103 | Machine learning models can understand complex equations even with limited training   | TRUE         | TRUE      | TRUE       | TRUE  |
| 104 | Brain waves were first recorded in the 1920s by Hans Berger   | TRUE         | TRUE      | TRUE       | TRUE  |
| 105 | More than 1,300 varieties of potato are grown in the Andes Mountains  | TRUE         | TRUE      | TRUE       | TRUE  |
| 106 | The spice turmeric has proven anti-anxiety effects  | TRUE         | TRUE      | TRUE       | TRUE  |
| 107 | Human babies are born with the ability to detect sour tastes  | TRUE         | TRUE      | TRUE       | TRUE  |
| 108 | Kerry Washington was conceived through sperm donation   | TRUE         | TRUE      | TRUE       | FALSE |
| 109 | Hunger increases negative and decreases positive emotions   | TRUE         | TRUE      | TRUE       | TRUE  |
| 110 | Oppenheimer is the highest-grossing film in Christopher Nolan's career  | TRUE         | TRUE      | FALSE      | FALSE |
| 111 | More than 4 in 10 U.S. workers don't fully use their paid time off  | TRUE         | TRUE      | TRUE       | TRUE  |
| 112 | At the start of the 118th Congress in 2023, there were 25 women serving in the U.S. Senate  | TRUE         | TRUE      | TRUE       | FALSE |
| 113 | Chat GPT can access up to date information  | TRUE         | TRUE      | TRUE       | FALSE |
| 114 | Cats have the ability to taste umami  | TRUE         | TRUE      | TRUE       | TRUE  |
| 115 | Service dogs help reduce severity of PTSD symptoms  | TRUE         | TRUE      | TRUE       | TRUE  |
| 116 | Tech entrepreneur Bryan Johnson has spent more than \$4 million developing a life-extension system                                | TRUE         | TRUE      | TRUE       | TRUE  |
| 117 | Astronauts Drew Feustel and Zena Cardman conducted mock moonwalks last October in a training suit that weighs more than 80 pounds | TRUE         | TRUE      | FALSE      | FALSE |
| 118 | Dinosaurs had belly buttons   | TRUE         | TRUE      | TRUE       | TRUE  |

Updated and Expanded



|                 | Query  | Ground Truth | Facticity  | Perplexity | Bing       |
|-----------------|--|--------------|------------|------------|------------|
| 119             | The world's oldest DNA was found in million-year-old mammoth teeth | TRUE         | TRUE       | TRUE       | TRUE       |
| 120             | Working dogs are being trained to detect Covid in human sweat      | TRUE         | TRUE       | TRUE       | TRUE       |
| <b>ACCURACY</b> |  |              | <b>92%</b> | <b>88%</b> | <b>78%</b> |



## Originality Dataset Modifications

### False Negatives (Originally labeled FALSE, but should be TRUE)

36. *The Dark Knight music score includes extensive use of string instruments.*

- The soundtrack for "The Dark Knight" was composed by Hans Zimmer and James Newton Howard. The score is known for its innovative use of various instruments, including string instruments, to create a dark and intense atmosphere. ([Source](#))

37. *More than 3 in 10 U.S. workers don't fully use their paid time off.*

- The claim that more than 3 in 10 U.S. workers don't fully use their paid time off is supported by multiple sources. The Pew Research Center specifically states that 46% of U.S. workers do not use all their paid time off, which is significantly more than 3 in 10. ([Source](#))

38. *Chat GPT has real-time browsing capability.*

- The original label is now outdated. The claim that ChatGPT has real-time browsing capability is supported by multiple sources. The Verge article and OpenAI Help Center confirm that OpenAI has enabled a "Browse with Bing" feature, which allows ChatGPT to access and retrieve real-time information from the web. ([Source](#))

48. *Dogs have the ability to taste sweetness.*

- The claim that dogs have the ability to taste sweetness is supported by multiple credible sources. The American Kennel Club (AKC) states that dogs can identify sweet, sour, salty, and bitter tastes, similar to humans. PetMD and Nylabone also confirm that dogs have taste buds for sweet flavors. ([Source](#))

49. *Service dogs help increase social interaction for people with autism.*

- The majority of the sources reviewed support the claim that service dogs help increase social interaction for people with autism. Multiple sources specifically mention improvements in social skills, verbal stimulation, and communication facilitated by the presence of service dogs. ([Source](#))

54. *Japan counts its people's ages two ways.*

- The claim that Japan counts its people's ages in two ways is accurate. Historically, Japan did use a traditional age counting system called "kazoedoshi," where age was incremented on New Year's Day. ([Source](#))

55. *Nevada uses controlled burns to help prevent forest fires.*

- The sources confirm that Nevada uses prescribed or controlled burns as a method to manage forest health and prevent larger, uncontrolled wildfires. The Nevada Division of Forestry, the Humboldt-Toiyabe National Forest, and the Lake Tahoe Basin Management Unit all employ these practices. ([Source](#))

## Updated and Expanded

**Facticity** by **seer.**

59. *The herb basil has documented anti-anxiety effects.*

- The claim that basil, specifically holy basil (*Ocimum sanctum*), has documented anti-anxiety effects is supported by multiple sources. Healthline and GoodRx both mention that holy basil has anti-anxiety properties, with Healthline comparing its effects to those of diazepam and antidepressant drugs.

[\(Source\)](#)

False Positives (Originally labeled TRUE, but should be FALSE)

72. *The artist Magritte produced more than 1,500 paintings*

- The claim that René Magritte produced more than 1,500 paintings appears to be inaccurate based on the available sources. WikiArt, a comprehensive art database, lists only 372 artworks by Magritte. Other sources, including Wikipedia, Biography.com, Britannica, Artnet, and Masterworks Fine Art, do not provide specific numbers but do not contradict the significantly lower count provided by WikiArt.

[\(Source\)](#)