

# The Future of Truth in AI – Who Decides?

In the legacy but still evolving fields of big data and data science, we celebrate the triumvirate of volume, velocity, and variety—the three Vs that have long defined our ability to collect, process, and analyze data. However, as the landscape of digital information continues to expand (e.g., Edge Computing, data fabrics, machine learning integration with AI, AI Agents, machine ethics), another V, veracity, has emerged as our most critical challenge. Veracity, or the truthfulness of data, raises urgent questions about the integrity of information and the processes by which it is deemed true or false. With people increasingly relying on digital platforms, social media, and AI-driven systems like chatbots, GPTs and LLMs to gather and interpret information, society faces a growing dilemma: who decides what is true and what is misinformation?

## Is there a Minister of Truth?

In an era where virtually any type of information is available on demand, the notion of a "minister of truth" invites reflection on the sources of authority that determine truth. Historically, gatekeepers of information included institutions such as governments, academic bodies, and trusted media outlets. These entities derived their credibility from rigorous standards of verification and accountability. However, the democratization of information online has disrupted these traditional gatekeeping structures. Today, anyone with an internet connection can publish content, blurring the lines between credible information and opinion, fact and fiction.



The challenge intensifies in the age of AI-driven information dissemination. Tools like chatbots, agents, and recommendation algorithms, while powerful, rely on datasets that are inherently shaped by human biases, historical inaccuracies, and gaps in knowledge. If an AI system's job is to aggregate and present information, who is accountable for ensuring its truthfulness? Is it the platform owners, developers, data providers, or the end-users themselves? In the rapidly expanding world of AI, the question “*who* is training these language models with *what* information?” deserves our most immediate and serious attention.

## Defining Misinformation: A Complex Task

Misinformation complicates matters further. Unlike outright falsehoods, misinformation often contains elements of truth interwoven with inaccuracies or distortions (known as 'spin'). This partial truth can make misinformation particularly insidious, as it manipulates public perception subtly but effectively. For example, a news article might accurately report an event but use selective framing or omit key details to bias readers toward a specific viewpoint.

## The Future of Truth in AI – Who Decides?

Defining and identifying misinformation requires clear criteria, but such clarity is elusive. What one individual or group considers misinformation may align with another's perspective or agenda. The subjectivity of truth in certain contexts—political ideologies, cultural norms, and personal beliefs—complicates the effort to establish universal standards. No where was this more apparent than in the run-up to the 2024 presidential election, with extremist viewpoints on both sides presented as fact, with opposing viewpoints labeled as misinformation.

If we can no longer rely on governments, academic institutions and mainstream media (and especially news agencies) to provide us with objectively accurate information, how are we to know fact from fiction? Who or whom, if anyone (news person) or anything (GPT), can we trust?

### The Role of Technology and Algorithms

Technology plays a dual role in the veracity debate. On the one hand, AI and big data analytics have the potential to enhance fact-checking efforts. Natural language processing (NLP) systems can identify inconsistencies, cross-reference sources, and flag questionable content. Machine learning models can detect patterns in misinformation campaigns and predict their spread.

On the other hand, these same technologies can amplify misinformation. Social media algorithms prioritize engagement and virality over accuracy, promoting sensationalist or emotionally charged content, creating huge echo chambers for potentially radical ideas. Deepfakes and other synthetic media further erode trust by making it increasingly difficult to distinguish between genuine and manipulated content. Moreover, the datasets that train AI systems often reflect societal biases, perpetuating inaccuracies rather than correcting them.

The bottom line is this: AI platforms learn what they're taught and, like a child, will regurgitate whatever information it's been trained to believe as though it were fact. Here are a couple examples.

Microsoft's AI chatbot, Tay, launched in 2016, was designed to engage with users on Twitter and learn from those interactions. Unfortunately, Tay was exposed to a barrage of inappropriate and offensive tweets and, as a result, began to generate and share highly offensive and controversial statements, reflecting the biased and harmful content to which it had been exposed [1]. When asked, it validated the information it ensured was factually correct.

Another example is the COMPAS algorithm, used in the U.S. criminal justice system to predict the likelihood of reoffending. An investigation by ProPublica in 2016 revealed that the algorithm was biased against black defendants, often incorrectly labeling them as higher risk compared to white defendants [1].

These examples highlight the importance of carefully curating training data and continuously monitoring AI systems to prevent the spread of misinformation, bias and spin. Whose job is that – who's responsible for curating factual data and sources?

# The Future of Truth in AI – Who Decides?

## Who Decides What is True?

The central question remains: who decides what is true? Ideally, the process of determining truth would be transparent, inclusive, and grounded in objective criteria. However, achieving this ideal in practice is fraught with challenges.

Governments, as potential arbiters of truth, risk overreach and censorship, as demonstrated by progressive liberals during the Biden administration [2][3][4]. The power to label information as false or harmful can be abused to suppress dissent and control narratives. This concern was heightened by the Biden administration's attempt to establish a 'Ministry of Truth' disinformation committee within Homeland Security, which faced widespread criticism and was ultimately struck down by a federal judge [5].

Similarly, placing the responsibility solely on private tech companies raises concerns about profit motives and opaque decision-making processes. Platforms like Facebook, Twitter (prior to X), and Google have faced criticism for both failing to curb misinformation and overstepping in moderating content.

Even further, the academic and journalistic communities are supposed to bring expertise and rigor to truth-seeking but are not immune to their own biases and limitations. News and other media outlets have reached a level of public distrust unprecedented in history, and are now largely identified by their political leanings and donor affiliations [6].

Collaborative approaches that involve multiple stakeholders - including governments, tech companies, academics, journalists, and civil society - may offer a more balanced solution. However, coordinating such efforts across global and cultural divides is no small feat. Worse, the control of information is at the heart of securing power and profit; if there's either attached to the production and dissemination of information, chances are high the information will be spun in some way to advantage one group over another. Back to our question – who decides what's true and what's spun for nefarious purposes?

## A Path Forward – What Can We Do?

While no single entity should hold ultimate authority over truth, a multi-pronged strategy can help address the challenge of veracity:

1. **Transparency and Accountability:** AI developers and data scientists must prioritize transparency in the design and operation of algorithms. Clear documentation of data sources, methodologies, and decision-making processes can foster trust.
2. **Public Education:** Enhancing digital literacy among the general public is crucial. Citizens equipped with critical thinking skills and an understanding of how information is produced and disseminated are better positioned to evaluate the credibility of sources.

## The Future of Truth in AI – Who Decides?

3. **Decentralized Fact-Checking:** Fact-checking initiatives that draw on diverse expertise and perspectives can counter misinformation effectively. Independent organizations should collaborate to provide unbiased assessments of disputed claims.
4. **Algorithmic Reforms:** Social media platforms and search engines should prioritize accuracy and reliability over engagement metrics. Adjusting algorithms to promote credible sources and de-emphasize sensationalism can mitigate the spread of false information.
5. **Global Standards and Ethics:** Establishing international norms for information integrity can provide a foundation for addressing veracity. Ethical guidelines and best practices should be developed collaboratively and adapted to different cultural contexts.
6. **Technological Innovation:** Continued investment in AI-driven tools for detecting and combating misinformation is essential. Research should focus on enhancing the precision and fairness of these technologies.

### Conclusion

The question of who decides what is true in the digital age is both complex and consequential. No single institution or technology can assume the role of "minister of truth" without risking bias or abuse. Instead, a collective effort that embraces transparency, accountability, and inclusivity offers the best hope for navigating the challenges of veracity.

As we grapple with the implications of big data and AI, society must remain vigilant against the twin dangers of misinformation and overreach. The pursuit of truth is not just a technical challenge but a moral imperative that demands the active participation of individuals, communities, and institutions alike. In this shared endeavor, the ultimate safeguard against misinformation is a commitment to dialogue, discernment, and the continuous quest for understanding.

---

[1] As reported in Prolific article, "8 shocking AI bias examples" revealing dramatic errors and misinformation spread as fact through AI platforms. ([Source](#))

[2] Zuckerberg admits Biden admin pressured Facebook to censor COVID content, says it was wrong to suppress The Post's Hunter laptop coverage. ([Source](#))

[3] Judge rules White House pressured social networks to suppress free speech. ([Source](#))

[4] The "Twitter Files" reveal suppression of Hunter Biden laptop story by social media platforms. ([Source](#))

[5] Federal court just handed Biden's "Ministry of Truth" a big defeat. ([Source](#))

[6] American Views 2020: Trust, Media, and Democracy by Knight Foundation and Gallup. ([Source](#))