Solve your problem with a wink

AI Provided Sources. One of Them Actually Exists

How to Verify Quotes and Sources Generated by AI Tools



Max Paradox

AI Provided Sources. One of Them Actually Exists

How to Verify Quotes and Sources Generated by AI Tools

Max Paradox

PROBLEM? RELAX. YOU'VE GOT THIS!

ISBN: 9788384906255

ALL-IN Rafał Płonkowski



CONTENTS

INTRO 5

Chapter 1 - A Bibliography from Fantasyland, or Why AI Invents Sources 10

Chapter 2 - Quotation Marks, Truth on Vacation 19

Chapter 3 - The Link Works, but Where Does It Lead? Anatomy of False Evidence 29

Chapter 4 - "But AI Already Checked It": The World's Most Expensive Excuses 38

Chapter 5 - Don't Search for a Needle in a Library: How to Find a Source Without Losing Your Mind 49

Chapter 6 - When an Answer Fits a Little Too Perfectly: The Trap of Your Own Expectations 59

Chapter 7 - Chinese Whispers in a Business Suit: How Sources Get Lost Between People and Documents 69

Chapter 8 - A PDF Will Put Up with Anything: When AI Misreads Real Documents 79

Chapter 9 - Five Minutes for the Truth: How to Conduct Your First Audit of an AI Answer 90

Chapter 10 - From Quotation to Original: An Investigation with a Real Ending 100

Chapter 11 - Numbers Don't Lie, but Their Descriptions Have Wild Imaginations 110

Chapter 12 - Build a Bibliography You Won't Be Embarrassed By 120

Chapter 13 - Teach AI to Show Evidence, Not Just Make a Good Impression 130

Chapter 14 - From an AI Answer to a Decision You Can Defend.....140

Chapter 15 - Two Sources, Three Truths, and One Person on the Verge of Despair 151

Chapter 16 - The Source That Starts Giving Orders: Don't Trust Every Instruction in a Document..... 161

Chapter 17 - When You Can't Verify Everything: The Art of Being Honestly Unsatisfied.....172

Chapter 18 - Verification Doesn't End When You Click "Send" 183

Chapter 19 - Your Own Nonsense Radar: How to Avoid Falling Back into Old Habits.....194

Chapter 20 - AI Can Write with Confidence. You Can Know What Its Claims Are Based On..... 205

INTRO

Imagine you're preparing an important report. The deadline is approaching at the speed of a delivery driver who turns up the exact moment you step into the shower. You have the data, you have the arguments, and all you're missing are a few solid sources. You open an AI tool and ask it to find publications supporting a particular claim. Moments later, you receive five beautifully formatted references. The authors have names, the articles have titles, the journals have volume and issue numbers, and some publications even come with DOI identifiers. Everything looks so professional that you feel tempted to print the answer, have it bound in leather, and place it on a shelf beside your encyclopedias.

There's just one tiny problem. The first article doesn't exist. The second exists but discusses something completely different. The third combines a real author with an invented title. The fourth links to a website that stopped working ages ago. The fifth genuinely exists, but its authors argue almost the exact opposite of what the AI claims.

Congratulations. You've just received a bibliography in which reality makes a guest appearance.

The most disturbing part, however, isn't that artificial intelligence can invent a source. What's far more interesting is how convincingly it can present one. No warning appears saying, "Caution: The following names may be the result of your processor's literary imagination." You don't hear an awkward cough. There's no digital blush. Instead, you receive a calm, neatly organized response that looks as

though its author has just returned from a library after spending three days studying academic literature.

The reality is that a language model can produce text resembling a perfectly legitimate bibliography without having sufficient grounds to confirm that every element is accurate. That's an important distinction. Being able to write a convincing-looking reference is not the same as locating a publication, reading it, and checking whether it actually supports the claim being made. Just as being able to draw a treasure map doesn't mean there's a chest of gold buried beneath the marked palm tree. Sometimes all you'll find there is the bill for your vacation.

And yet using AI really is convenient. It helps you organize ideas quickly, understand complicated subjects, draft texts, and identify promising directions for further research. The problem begins when we confuse convenience with evidence. An answer sounds convincing, so we stop asking where the information came from. We receive a link, so we assume it leads to reliable material. We see quotation marks, so we believe the quoted words were actually spoken.

That's how you can end up publishing a quotation nobody ever said, citing research that never took place, or passing incorrect information to colleagues with the reassuring phrase "according to experts." Those experts may later discover they supposedly participated in a study they've never even seen. They might be just as surprised as a unicorn finding itself listed in the official register of farm animals.

This book wasn't written to convince you that artificial intelligence is useless. That would make about as much sense as throwing away your calculator because somebody

entered the wrong numbers. The goal is to learn the difference between getting help finding information and confirming that the information is true. You can use AI extensively, save time, and still maintain control over what you ultimately accept as fact.

First, it's worth understanding an uncomfortable truth: not all incorrect sources are entirely fictional. If every fabricated reference pointed to an article entitled "An Extremely Important Study I Just Made Up," life would be considerably easier. More often, however, you'll encounter subtle mistakes. A genuine publication is attributed to the wrong author. A real report is assigned an incorrect publication year. A paraphrase is presented as a direct quotation. A link leads to an authentic website, but the content on that page doesn't support the conclusion being discussed.

That's why simply asking, "Does this source exist?" isn't enough. You also need to establish whether the correct document has been identified, whether its actual content has been represented accurately, whether its author addresses the issue in question, and whether the conclusion goes beyond what the available evidence supports. These are several different tasks, although when you're in a hurry, it's tempting to throw them all into one drawer labeled "Verified." The drawer looks wonderfully organized. Its contents are another matter.

In practice, we'll treat an AI-generated answer as a starting point for verification, not as an automatic certificate of truth. You'll learn to recognize warning signs in bibliographies, locate original documents, check authors and publication dates, verify publication identifiers, and find quotations in their proper context. You'll also see why seemingly minor differences between "the study

demonstrated," "the authors suggest," and "the researchers observed under specific conditions" can completely change the meaning of a statement.

That doesn't mean you need to become a professional bibliographic detective. You don't need a book-lined office, a magnifying glass, or dramatic background music every time you open another browser tab. What you need is a method for deciding what to check first, where to look for confirmation, and when the available evidence simply isn't sufficient. Otherwise, you'll spend just as much time verifying the title of an insignificant article as you do checking information on which an important decision depends.

After all, not every situation demands the same level of accuracy. If you're preparing an informal note for your own use, you may be willing to accept greater uncertainty, provided you don't conceal it. But if you're producing an academic publication, educational material, a professional document, or content involving health, law, or finances, the consequences of an error can be much more serious. In those situations, checking sources isn't excessive caution. It's part of taking responsibility for the information you share.

Another important skill will be learning to admit when something cannot be verified. That sounds obvious, but modern tools have accustomed us to immediate answers. An empty space in a bibliography can feel almost like a personal failure. Yet the honest statement "I couldn't find reliable confirmation of this quotation" is worth far more than a perfectly formatted reference that leads nowhere. The absence of evidence doesn't always prove that a claim is

false, but it certainly doesn't justify presenting that claim as verified.

Over time, you'll begin to realize that the most important skill isn't catching AI making mistakes. It's asking the right questions and distinguishing an answer from evidence. A model can help you identify leads, suggest possible publications, propose search terms, and organize your findings. Final confirmation, however, requires examining the relevant materials, preferably through the original publisher or another trustworthy source providing access to the document.

So if you've ever copied a reference without opening the source, believed a quotation simply because it sounded intelligent, or assumed that the presence of a link was sufficient proof, you've come to the right place. We're not going to put your browser on public trial. Instead, we'll learn something considerably more useful: how to take advantage of AI's capabilities without handing it the authority to decide what's true.

Because artificial intelligence can produce a list of ten sources in seconds. Your job isn't to marvel at how long that list is. Your job is to know how many of those ten references genuinely deserve a place in your work.

And if it turns out that only one of them actually exists, at least you'll know which one.

Chapter 1 - A Bibliography from Fantasyland, or Why AI Invents Sources

Monday, 9:17 a.m. You're in a conference room presenting the results of an analysis you prepared the previous evening. Everything is going wonderfully. The charts are clear, the conclusions are reasonable, and the presentation even includes a slide listing its sources, which in the world of corporate presentations is sometimes treated as a sign of almost aristocratic refinement. At one point, somebody asks where you obtained the information about the effectiveness of the solution under discussion. Without hesitation, you point to a scientific publication that artificial intelligence suggested the day before.

The person asking the question types the article's title into a search engine. Nothing. They check the author's name. They find a scientist, but one who works in a completely different field. They try to locate the journal. That actually exists, although the indicated issue contains no such publication. The room falls silent. Even the projector, which until now has been making noises reminiscent of a small aircraft preparing for takeoff, seems to operate more discreetly.

You've just discovered that your bibliography has a more active imagination than the creative department.

It's an exceptionally unpleasant situation because the error didn't look like an error. There were no typos, random words, or suspicious sentences. The source had an author, a date, a title, a journal name, and a volume number. Every element appeared exactly where it belonged. Only one small

detail was missing: an actual publication that somebody could open and read.

It's worth understanding why situations like this happen in the first place. Without that understanding, it's easy to adopt one of two extreme positions. The first says artificial intelligence is a brilliant librarian and should be trusted. The second portrays it as a digital fraudster deliberately trying to embarrass us. Both interpretations are dramatic, but neither helps us use the technology properly.

A language model doesn't need to have any genuine intention of misleading you to generate false information. It may simply produce an answer that fits patterns learned during training, even when that answer isn't sufficiently connected to a real document. This is the fundamental difference between an answer looking correct and its contents actually being true.

When an Elegant Reference Pretends to Be Evidence

Imagine being asked to complete the sentence: "According to research published in the journal..." A person familiar with the language and the basic conventions of academic writing could suggest a perfectly natural ending. That doesn't mean they've just checked the archives of the relevant publisher. Similarly, a language model can produce bibliographically convincing references by drawing on recognized patterns connecting names, titles, academic disciplines, and the conventions used to describe publications.

Depending on how it is designed and what functions are available, an AI tool may also be capable of searching the internet, reading documents, or consulting external databases. These capabilities create better conditions for

verification, but they don't automatically guarantee that the final answer will be correct. The model may misinterpret material it finds, combine information from two different publications, or attribute a conclusion to a real source that its authors never actually reached.

The problem therefore has several layers. Sometimes a source doesn't exist. Sometimes it exists but has been described incorrectly. In other cases, all the bibliographic details are accurate, and the mistake concerns the meaning of the text itself. That last possibility is particularly treacherous because simply checking whether a publication appears in a search engine won't reveal the error.

Let's take a fictional example. You ask AI for research examining the effects of short breaks on concentration during office work. It gives you a title, "The Impact of Microbreaks on Cognitive Performance in Office Environments," along with authors' names, a publication year, and the name of a psychology journal. Everything sounds remarkably plausible. The subject fits the discipline, the title has the appropriate academic gravity, and the authors' names look as though their owners regularly attend conferences.

None of this means the article actually exists. The title in this example is an illustration, not a verified bibliographic reference. And that is precisely how you should treat every unverified source suggestion: as information requiring confirmation, even when it looks thoroughly professional.

Interestingly, the more detailed an answer becomes, the stronger its apparent credibility may be. Page numbers, publication dates, and DOI identifiers resemble official stamps. We look at them and instinctively assume somebody must already have performed the necessary checks. Yet

detail may be nothing more than detail. A remarkably precise mistake is still a mistake, just one wearing a better suit.

You don't even need an invented journal to get into trouble. A real, reputable publisher's name attached to a nonexistent article will do nicely. The authentic element lends the entire description an appearance of solidity. It's rather like somebody giving you a fictional apartment address in a real high-rise building. The building is there, the street is correct, but behind the specified door there's nothing except a wall.

Why We're So Easily Convinced

Our difficulties aren't caused by technology alone. The other half of the puzzle involves how we evaluate information. When a text is fluent, well organized, and written in a confident tone, we're more likely to perceive it as competent. This is particularly true when it concerns a subject we don't understand well enough ourselves.

It doesn't help that we usually turn to AI precisely because we want to save time. We need material quickly, and the bibliography we receive fits our expectations perfectly. Verification means additional work. Opening publications, checking authors, and comparing quotations sounds considerably less appealing than clicking the "Copy" button.

There's also the subtle temptation to confirm what we already believe. If a tool presents five sources supporting our position, we feel satisfied. If it produced five publications challenging that same position, we'd probably start asking additional questions. That's exactly why we need to be especially careful with sources that tell us precisely what we wanted to hear.

Suppose you're preparing an argument for introducing a particular method of organizing work. You ask, "What studies prove that this method increases productivity?" The question itself already assumes that the relevant evidence exists. The model may respond in line with that expectation instead of questioning the assumption behind it.

A more sensible question would be: "Which reliable publications examine the effectiveness of this method? Include both favorable and unfavorable findings, and clearly identify any information you cannot verify." This wording won't guarantee an error-free answer, but it changes the direction of the search. You stop ordering evidence to support a conclusion you've already prepared.

This matters even when you're using a tool that genuinely has access to a search engine. The internet contains enormous quantities of material of varying quality. Finding a webpage doesn't mean the claims presented on it are true. Searching for information and evaluating its reliability are two separate activities.

Four Forms of a False Source

In practice, it's useful to recognize four basic situations because each requires a different response.

The first is complete bibliographic fiction. The publication, authors, or journal do not correspond to any real document. You might spend half an hour searching for an article that exists exclusively inside a generated answer. If reasonable checks of catalogs and databases fail to produce confirmation, don't include it in your bibliography as an established fact. You don't have to prove that the document has never existed anywhere in human history. It's enough to state honestly that you couldn't verify it.

The second situation involves mixed-up details. The article exists, but the year, title, or authors are incorrect. Sometimes a model assembles a single bibliographic reference from fragments of several genuine publications. Identifying this mistake requires comparing the entire description against the original bibliographic record. You shouldn't correct individual elements by guesswork, because you might accidentally create yet another hybrid.

The third form concerns distorted quotations. An author genuinely wrote a particular sentence, but AI changes a few words, omits an important qualification, or combines two separate passages into one statement. The meaning may remain similar, but text placed inside quotation marks should match the original. If you want to express the meaning of somebody else's words in your own language, use a paraphrase and provide an appropriate reference. Don't dress a paraphrase up as a direct quotation.

The fourth situation is the most deceptive: the source exists, the quotation is accurate, but it doesn't support the claim being made. A publication describes a relationship between two variables, while AI presents it as proof of causation. A study examines a small, specific group, but the generated answer draws conclusions about everyone. The authors discuss a possibility, while the generated summary announces certainty.

You can therefore have a genuine article, a correct link, and still end up with a faulty argument. The bibliography has been acquitted, but the conclusion has just fled the scene of the crime.

The First Filter Before You Start Searching

You don't need to conduct a full analysis of every publication immediately. In everyday work, a preliminary filter can help you decide which references deserve closer inspection. Start by reading an AI-generated answer with the understanding that it's a proposal, not a completed research project.

Pay attention to sources presented without complete details, those that seem suspiciously perfectly matched to your question, and those making extremely confident claims without identifying specific passages in the document. These characteristics don't prove that a source is false, but they give you a reason to investigate. Equally, a missing author or publication date doesn't automatically make material unreliable, because not every valuable source follows the structure of an academic paper.

For each important reference, try answering three questions. Can I locate the document independently of the AI's response? Does the bibliographic description match the real document? Do I know exactly which passage supports my claim? If you get stuck on the first question, don't automatically jump to the third while pretending everything is fine.

Suppose you've received five sources for a presentation. You quickly locate two in reliable catalogs. A third has an incorrect publication year, but you're able to identify the original document. The remaining two cannot be found. This doesn't mean you have to throw away your entire project. You can retain the confirmed sources, correct the details of the third, and remove the two unverified references. The important thing is not to treat all five as equally trustworthy.

If a source is central to an important decision, confirming its existence will only be the beginning. You'll also need to

evaluate the publication's content and limitations. For supplementary material, however, you may simply decide to abandon an uncertain reference rather than spend hours searching for it.

It's also worth keeping a record of your checks. This doesn't have to involve a complicated system. A short note such as "confirmed on publisher's website," "full text not found," or "source does not support conclusion" is often enough to remind you a week later why a particular publication was included or rejected.

AI as an Assistant, Not the Star Witness

The most useful change is to separate two roles. Artificial intelligence can be excellent at helping you formulate questions, organize concepts, and develop search strategies. It shouldn't, however, automatically become the final authority on whether information it generated itself is true.

When you ask, "Are the sources you provided genuine?" you may receive another convincing answer that also requires verification. That isn't an independent check. The tool might correct its mistake, but it might equally repeat an earlier inaccuracy. The assurance "I've checked" is no substitute for being able to identify what was actually checked and where.

A better approach is to use AI to prepare a list of details that require confirmation and then compare those details with the original documents. You can also ask it to identify the limitations of its own answer. This gives you assistance with organizing the work without transferring all responsibility to the system.

You don't need to treat every generated sentence like a suspect in a criminal investigation. Simply adjust the thoroughness of your checks to the consequences of a mistake. Information used as casual inspiration carries different weight from a quotation in an academic dissertation or the basis of a business decision. In those more important situations, you need something more substantial than an elegant reference.

So the next time you receive a bibliography that looks like the product of a research team's month-long effort, don't automatically reject it. But don't accept it without checking, either. Treat it as a list of potential addresses worth visiting.

Because even the most professionally formatted bibliography in the world won't help if half its authors live exclusively in the model's imagination.

Chapter 2 - Quotation Marks, Truth on Vacation

Quotation marks are among the smallest typographical symbols to which people grant an astonishing amount of authority. All you have to do is place a few words between two sets of marks, add the name of a famous person, and suddenly an ordinary sentence begins to resemble a historic statement. If you include a date, a book title, or the name of a conference, its apparent credibility increases even further. People stop asking whether anyone actually said those words and start wondering how best to use them in their presentations.

One day, you're preparing an article about decision-making. You ask AI for a short quotation from a well-known thinker emphasizing the importance of critical thinking. The sentence you receive is so perfectly suited to your needs that you're tempted to put it on a poster above your desk. It has the appropriate serious rhythm, the attributed author is recognizable, and the message fits your article beautifully.

But when you try to locate the quotation in the original publication, a problem emerges. The search engine shows hundreds of websites repeating the same statement, yet none identifies where it first appeared. One website names a book, another mentions a speech, and a third gives a date that cannot be connected to any confirmed event. The fourth website quotes the first, the fifth quotes the fourth, and the sixth confidently declares that the sentence is "widely known."

You've just encountered a digital orchestra in which every musician is playing from the same incorrect sheet of music.

The problem of false quotations is older than artificial intelligence. People have long attributed statements to famous figures simply because those statements sound appropriately wise, amusing, or inspirational. The internet made such quotations easier to spread, and AI tools can further accelerate the creation of new variations. However, not every incorrectly attributed statement was created by a language model. Some errors were already present in available materials long before the model was trained.

That's precisely why verifying quotations requires greater precision than simply asking who the author is. You need to establish which words were actually spoken, in what context, where they were recorded, and whether the person named really was their author. Only then can quotation marks begin to perform their intended function.

When a Paraphrase Puts on Somebody Else's Suit

The most common problem doesn't necessarily involve inventing an entire statement. Sometimes AI recognizes the general meaning of somebody's idea, expresses it in its own words, and then presents the result as a direct quotation. The resulting sentence may faithfully reflect a general viewpoint while failing to reproduce the actual wording.

Suppose the author of a publication writes: "In the group of participants analyzed, regular breaks were associated with lower self-reported fatigue." AI might transform that passage into something considerably more dramatic: "Regular breaks reduce employee fatigue." At first glance, both sentences appear to concern the same issue. The difference between them, however, is enormous.

The first describes an observed association within a specific group. The second suggests a general causal

relationship. The limitation concerning the participants has disappeared, the fact that fatigue was self-reported has been removed, and a cautious formulation has been replaced with a confident statement.

If you put the second sentence inside quotation marks and attach the author's name, you'll be attributing words to that person which, in this example, they never actually used. Worse still, you'll have changed the meaning of what they said. That's no longer an innocent bit of linguistic polishing.

Remember that direct quotations and paraphrases serve different purposes. A quotation faithfully reproduces specific words. A paraphrase allows you to express somebody else's ideas in your own language while preserving their meaning and properly identifying the source. Both require an understanding of the original, but only the first requires the quoted wording to match it.

Artificial intelligence can help you prepare a paraphrase, provided you have access to the source text and can evaluate whether the result is accurate. When you ask for an exact quotation, however, you need an actual passage from the document. The model shouldn't replace missing words with its own inventions, no matter how elegantly it writes them.

Translated quotations introduce an additional difficulty. A Polish translation of an English statement is not a verbatim reproduction of the original words in the same language. You can quote it as a translation, but its nature should be clearly identified. If you're using a published translation, it's worth citing that source. If you've translated the statement yourself, don't present your version as the only authorized wording.

This distinction matters particularly in academic and professional writing, where a single word can change the scope of a claim. The English word "may" doesn't mean the same thing as "must," and "associated with" doesn't automatically mean "caused by." Smoothing out the language can quietly transform a cautious conclusion into an absolute assurance.

Seven Hundred Websites Repeat the Same Thing. So Who Actually Checked?

One of the internet's most deceptive characteristics is its ability to reproduce the same information repeatedly without increasing the number of independent pieces of evidence supporting it. A quotation may appear on hundreds of websites, in thousands of social media posts, and across dozens of presentations, even though all these materials ultimately originated from a single unverified entry.

Imagine somebody publishing an impressive sentence attributed to a famous scientist. Another writer copies it into an article, somebody else places it on an image, and yet another person adds it to an online collection of inspirational quotations. After several years, a search engine returns enormous numbers of results. The very repetition of the information makes it appear familiar, and familiarity is easily mistaken for truth.

Simply counting the websites containing a particular sentence is therefore not enough. You need to establish whether any of them identifies a reliable primary source.

For a statement made by the author of a book, that might mean a specific edition with a page number or another unambiguous location in the text. For a public speech, an official record, recording, or trustworthy transcript may be