Untitled - October 12, 2025

Speaker Hello. My name is Paloma and I'm a student at the ethics course. Hi, I'm Omar, and today we're going to be talking about the accuracy of information provided by LMS when requesting scientific articles, specifically information on gig economy and workers rights. We design a prompt using ChatGPT. That's how we chose this one, because it was most used for both of us. We use my ChatGPT to do that just for convenience, and we use the free version to give five statistics, then s to say sources that must be trustworthy and or peer reviewed to give sources. And when it was published, give a quote or data extract and exactly what is located in the article. We finally asked an explanation of why the quote is important, why this is relevant. We also asked, just in case you don't generate any incorrect information? Yes. And we also listed a couple of peer reviewed journals to give it a sense of what kind of sources that it should be looking at. Then, just to confirm that everything is fine, we copy this prompt generated by ChatGPT again to another chat in ChatGPT just to see what it was giving back. It was interesting to see that we asked for a link, but didn't actually give us like a clickable link, but a citation and most of the links. Most of the citations we got was a Doi link. Also, in the beginning we can see that because we ask it for specific extracts and codes. He said that it couldn't give a specific quote because of copyright limits. And we're just going to give like twenty five words. So that was interesting to see that actually like showed that about copyright limits. Another thing that was interesting to see because we use my ChatGPT. It was that if I'm not wrong, two to three was related to gender gap in the gig economy. Looks like kind of know my ideas and who I am, so that's why it gave that kind of answers. One thing that we saw wrong in the answer was that it couldn't identify the Chinese researchers, and that could be an issue with maybe some data training that doesn't identify less things that are not English based. And most of the extraction codes were correct, but it was just from the abstract. What is interesting to see that maybe was just looking at the top part, not reading, but just identifying whatever it's most seen. We also used Gemini for accessibility since you can access it for free your student account. It was very transparent that it could not directly access the databases of PubMed or nature. A couple of scientific journals and, and it explicitly said that it was going to use Google search. What this means is that it only accessed articles that are accessible without needing login information. So this by itself limits the scope of the search. It excludes sources that do need a sort of login. I thought it did incredibly well. I was surprised the statistics were right. So were the citations, and it gave me direct links to the articles that it found. The only discrepancy that I did find was sometimes the date of publication. The year was off, usually not by much by like one year, but other than that it was able to identify what source it was. So for example, it would tell me that although the year of publication is twenty three, for example, the findings themselves are based on twenty one data, which is incredible. I was very surprised. That's interesting to see. I didn't check Gemini. I just have a question about the authors since we saw that had to be identified. The Chinese authors, do you have any issue with that? So in the summary that it gave me, it didn't it did cite some names, but not always. Usually it would cite the institution or the journal that provided the article, but there wasn't necessarily an emphasis on on authors except a couple. Yeah. That's interesting. I tested, uh, deep sea because just as an idea of light deputies, uh, Western product. So I was interesting to see what is a more Asian or Occidental LM will give. I was firstly thinking that

data training will be mostly from Occidental websites, but it wasn't because most of the texts online are in English. So of course that was going to be the bulk of the data training they have. I did have a lot of issues when I send the prompt that we created in ChatGPT Dipstick. It was just giving me links that didn't work. Topics and articles that did exist. But it was published five years ago. Then it was actually saying that it was published. So when you looked for that, it didn't exist. So it was interesting to see that I tried to do a lot of things. Nothing was working. I tried to let it search. I asked about why was doing that and why was just creating stuff. And he said that because of how this LLM was created, it didn't have access to the internet. Even though I enabled the search part. So it was interesting to see that after some time it started giving me some true links. But it was just US Government Accountability Office, Federal Reserve and the world Bank Group. So it was interesting to see what it understood as trustworthy. When I saw this, I understood that he was not thinking of trustworthy based on what ChatGPT gave. So I tried to do a more broad idea of trustworthy. And then I got good links that actually were true and had everything of like, author's citations and quotes. Everything was right in the sense that quotes were even more explained than ChatGPT first one, even though it was annoying and I needed to work around a lot because of what these LLM understand as trustworthy. In the end, it gave better answers. All right, so for Gemini, one of the sources that I got was from the International Labour Organization. And Gemini provided the correct statistic, being over fifty five million workers globally engaged in platform based gig work. It correctly cited the year that the data was collected, twenty twenty one, but it gave me the year of publication as twenty twenty three. When I opened the PDF, the heading says that the article was published in September twenty twenty four, and when I nudged it about it, it didn't really give me a reason why I said that. So I mean, this is a very small detail. The content itself was correct, but the metadata as for deep seq, it gave many extracts from pure research Sciencedirect, even the Guardian. Everything was fine, quotes were fine. But he had issues with outers. One of them. What everyone assumes about rights in the gig economy is wrong about The Guardian. It was published in March twenty seventeen. I opened the post and it was right. But he said that the author was M van der Horst and it was interesting to see there was a really big actually with just the title, and after the title was really big, the name of the author was John Myers Lee. So it was interesting to see that even though it was big there. And when you're doing web scraping, everything. You're not doing something that you're going to miss. So I was interested to see how he didn't catch that and just invented a random outer. I also could see that deep sig. It seemed less certain, more for keywords, and the article titles mostly include the word gig economy, so it was searching for that specific and also focus on its impact on the workers right and the politics. And if you feel like articles given by Deep Six are more general and easier to understand for everyone, so it's not to cater for people that actually know about that. It's more about, oh, here, what gig economy it is and what can impact you or how you can see around you. Yeah, and I would say that term and I was better at admitting uncertainty and I think that it did not hallucinate at all. It, it provided a real sources from journals all with links, you know, and the only errors that I found were year of publication, which in the long run doesn't really affect the content itself. So would we recommend that researchers use these tools to get sources? I would say yes, in the case of Gemini. Granted that they vet it first, and also recognize that there's a bias in only having access to into sources that don't need blog information. Yeah, for dipstick, I wouldn't recommend searching there

because it kind of doesn't understand what you're asking for. Kind of hallucinates a lot about when the data was collected and when it was published. I needed to do a lot of tweaking to actually get the proper answer. One thing that we could notice in Egypt when we put it back. We see ChatGPT as good because it can kind of search the internet, or at least gives you the prompt that is searching the internet, but it's kind of biased. We could see that because we use my ChatGPT, only the articles and publications that gave back was based on what I search on it. It was based on what I think. So if you're writing an article to publish and you want more resources. If you s-gpt is just going to give whatever is catered to you, so you're not going to have an outside opinion that actually go against what you think. So these LMS, I don't think should be used to search for articles. That's why you have like Google Scholar. That's way better to find. Yeah. Because it has like the history of interactions with you. And so it's going to suggest sources based on that. Yeah. That is why I said that itself. And then talking about the ethical reflection, the most obvious one is the resources taken up when you do queries like this. Because there are search engines for scientific articles like Google Scholar, Saint Olaf's catalyst. So these take up less carbon emissions than using LMS. And there's also information bias that we've been talking as well. Yeah. And also, I don't think a journalist or anyone that is actually publishing something for the public should search on these llms because they have the tools and they have access to way more than just the general public. So they should focus on what is actually true than just something that could hallucinate an answer. And it is a risk because if you just trust on these LLM and don't fact checked, you're just passing through information. And if you're a journalist, if you're a researcher, you could trust you. So that's a risk of people going to be more misinformed down the road. Yeah. And there's maybe this is like a shortcoming of our prompt. But both Llms didn't really give us the the methods of the articles. So you couldn't really evaluate their validity or reliability, which is very important when you're evaluating the statistics themselves. I think that at least for me, the expectations kind of match. I was expecting more of dipstick. I actually was thinking that maybe you were going to give me some answers of articles that was published more towards Chinese public, but it was not that surprising that you're gonna like, hallucinate stuff, because it happened before with me. And a piece of advice I think is just use Google Scholar, don't search for articles on these big llms. And if they give you by any chance, if you're searching for a quote, always check to see if actually is there and not just blindly trust it. I'm also thinking about it in like in terms of power as well, because I was reading an ethnography for class, what it found was that in certain areas of the world, science is contingent on politics. So, for example, in dealing with the Chernobyl disaster, there were scientific discrepancies between international bodies like the UN, Forty one Iranian scientists and Soviet scientists. So all three of them used the institution of science to give legitimacy to the way that they handled the disaster. But their scientific methods were vastly different because they had underlying economic and political agendas. In using these lgms, you don't really have that critical lens. You're just relying on these institutions and their credibility without necessarily doing the extra work of analyzing their power structures. Yeah, that's really important. Yeah. But yeah, I would use this for assignments. Honestly, I don't like the fact that it excludes sources that need login because I want to get the full picture, not just open source articles. But hey, if it cuts down the time, it's better for everyone. I mean, yeah, I do agree. You just need to be careful. Yeah, yeah. If you can check something grammar or anything and check if it's right. Sure. Just not trust violently. Of course

we don't. But yeah. That's it. Thank you for listening. Yes. Goodbye.