Confidential documents and conversational AI

Description

Confidentiality is key in any profession, especially as it related to client-consultant relationships. Iâ??m hard pressed to find confidentiality foregrounded in architectural codes of practice, but it is crucial in law and financial services. The Handbook of the Financial Conduct Authority, for example, states that a financial advisor (a â??skilled personâ?•)

> â??may not pass on confidential information without lawful authority, for example, where an exception applies under the Financial Services and Markets Act 2000 â?¦ or with the consent of the person from whom that information was received and (if different) to whom the information relates.â?•

Can providers of professional services deploy conversational AI platforms to supplement or even replace some of their consultation roles. In most real-world professional services delivered entirely by humans the â??skilled personâ?• exists within a network of relationships. Not every agent in that network needs to have full information. Nor, in the interests of client or proprietary confidentiality, should they. Hereâ??s an example of the challenge.

Asking for a loan

Imagine that a loan company says it will issue me with a reverse mortgage (equity release) on my house if I have over Â£200k share in the property, if I am over the age of 55, I have a regular income above Â£20k per year, and Iâ??m not supporting any dependents. The data showing that I meet those criteria is confidential. I donâ??t want creditors, associates or family members to see it. Imagine that I meet the criteria for the reverse mortgage. I could ask the loan company to confirm that I am eligible, but please donâ??t disclose the conditions or the facts to anyone else. I will take the firmâ??s reply as certification of my eligibility and just show that to a relevant third party, e.g. a vendor, debtor, bank, etc.

Thatâ??s not so difficult for human agents to manage in a professional contexts, as long as the players know their roles, what they can and cannot say and are committed to appropriate confidentiality protocols.

Iâ??ve put the scenario to a conversational AI (ChatGPT3) session, stating the conditions for the reverse mortgage, delivered my data, and asked it in its reply to please confirm that I am eligible, but without disclosing any of my data in that reply. I would treat its reply as certification to show to a third party.

It was a complicated request to put to an AI, which would benefit from the ability to do some rudimentary calculation and rule-based logic, but the GPT3 platform balked at the idea that it was being asked to preserve confidentiality, making the obvious case that the platform relies on training from corpora of texts from diverse sources, including (possibly) conversations such as this one. (Iâ??ve not yet tried this with GPT4.)

Zero-knowledge proof

What I was doing was attempting to invoke from the conversational AI platform aspects of the performance of a zero-knowledge proof (ZKP) service. According to a helpful blog by Lexie.

> Zero-knowledge proofs are encryption schemes used to prove that you know something without revealing what it is. For example, you can show without a doubt that you know the answer to a puzzle without actually disclosing the solution.

I was asking the AI to confirm that something was true without disclosing how it knew that. Conversational AI might indeed benefit from ZKP strategies to secure its data, but as yet seems unable to perform that professional function in conversational mode, as would a (human) professional agent.

Conversational AI is so far modelled on the performance of two agents (the user and the platform) interacting with one another, and as yet does not handle multiple synchronous and asynchronous interactions between groups of agents, amongst whom varied protocols of confidentiality apply.

Top secret documents

Another example is the requirements of a jury to decide on a case that involves the theft of top secret and highly classified documents. How can jurors and lawyers reason about the content of those documents without having the necessary security clearance, i.e. without knowing the content or even the value of those classified documents?

I asked ChatGPT how it could help in such a case. Helpfully, it outlined six strategies to address the challenge, including document redaction, use of expert testimony and extending security clearances. It drew attention to its potential role in research, definitions, advice, procedures, and sorting documents in relation to the case, but abrogated any role as juror or legal agent:

> â??while a conversational AI platform can provide valuable assistance, it cannot replace the expertise and judgment of human legal professionals. Its role is to provide information, support research, and offer guidance within the legal framework established by the court and relevant authorities.â?•

When I put to it the potential role of zero-knowledge proof it provided helpful information about how ZKP could be deployed for this confidential documents case, but did not offer it as a service that could be accomplished by current conversational AI. I pressed it on prospects and it explained the benefits of generating, verifying, explaining, integrating and protecting ZKP measures in future iterations of conversational AI.

Bibliography

- EU. â??EU AI Act: first regulation on artificial intelligence.â?• *News: European Parliament*, 15 June 2023, 2023. Accessed 8 June 2023. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
- Lexie. â??Zero-knowledge proofs explained: Part 1.â?• *Express VPN*, 16 October, 2017. Accessed 17 June 2023. https://www.expressvpn.com/blog/zero-knowledge-proofs-explained/
- Madiega, Tambiama. â??Artificial Intelligence Act.â?• *Briefing, EU Legislation in Progress*, 2022. Accessed 16 June 2023. https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.p
- Wagner, Alex. â??â??Goldilocks documentsâ??: How to try an Espionage Act case without spilling national secrets.â?• *MSNBC*, 15 September, 2023. Accessed 17 June 2023. https://www.youtube.com/watch?v=S7BSy-ugYxM

Category

1. Artificial Intelligence

Tags

1. confidentiality
2. conversational AI
3. finance
4. law
5. professionalism
6. security

Date Created
June 24, 2023
Author
rcoyne99