Yoti Guardians Council Meeting

Attendance: see table at end

Date: 30/09/2024 - 15:00-17:00

Location: Google Meet

Recorder: AK

Agenda

- Welcome and approval of minutes from the previous meeting
- Update on funding and revenue
- Update on Yoti AI chatbot
- Update from the R&D Team
- Update on Digital ID Match
- Update from the Ethics & Trust Committee meeting
- Update on UN End Violence Against Children Compliance Scanner project
- Update from the Senior Privacy Counsel and Deputy DPO
- Discussion on how the Guardians can best add value to Yoti
- AOB
- Close

Welcome and approval of minutes from the previous meeting	Yoti staff opened the meeting and welcomed the new and latest Guardian to Yoti. A Chair was officially nominated. Minutes for the Guardian Council meeting in June were approved by the Guardians present.
Update on funding and revenue	Yoti CEO provided a company update including an update on our partnership with Lloyds, revenue updates, an outlook for the next few months and recent significant business opportunities. The Guardians asked for clarification on whether Lloyds has the option of a board seat at Yoti and whether Lloyds is aware of the Guardian Council and its role at Yoti, both of which Yoti confirmed.
Update on Yoti Customer Service chatbot POC	The Yoti team lead briefly explained the current state of Yoti's SAAS integrations. The team updated the Guardians about its recently developed AI chatbot which maintains the current customer support team, whilst allowing Yoti to scale for future growth to many more users and allows the Yoti human team to triage queries, handling the simple ones via the chatbot and allowing more indepth support for more complex issues. The team lead demonstrated the user flow. The team explained that currently the chatbot is only available during office hours, to enable close supervision; however the aim is for it to be available 24/7. During development, data considerations have been thought through. Data is anonymised and a privacy notice has been added in for users to explain how their data is used. The chatbot currently uses a private instance of OpenAI, though Yoti may move away from this once there is a viable internal solution. All responses are periodically reviewed by the Yoti team, with each outcome recorded by the team. The Guardians asked how Yoti ensures there are no hallucinations and that OpenAI does not use general information found from across the internet. The Yoti lead stated that it has tweaked the parameters available via the OpenAI API and confirmed that they only use the API built by OpenAI that is specifically for chatbots. Therefore, it is trained to only respond with what it knows, with prompt engineering being used to bring down that space. There is also a profanity filter which sends the bot into stop mode. However, this is why responses are being constantly monitored.

	The Guardians asked how Yoti guarantees that OpenAI does not use this information to train its models.[1] Yoti confirmed that it has opted out from this from the beginning. The Guardians asked for a copy of the privacy agreement for review (shared 1st October 2024).[2] Yoti also offered to share its privacy risk assessment for this project. The Guardians asked if Yoti has noticed any change in the demographics of people interacting with Customer Service vs. the chatbot. Since it doesn't ask for demographic data in order to maintain user privacy, Yoti is unable to track this. However, this would be a good longer term consideration.
Update from the R&D Team	Yoti gave an update on the latest September 2024 age estimation model. The team outlined the improvements in mean absolute error (MAE), where improvements can still be made and the reasons for the differences between Yoti's results and NIST's results. Yoti explained its approach to measurement, how the model's performance has improved over time, how Yoti tracks its accuracy and how the model is trained. Yoti outlined that the two key ways to improve the model is with 1) better data (cleanliness and quantity of data) and 2) a better model (including how it's trained, the functions being used and the objectives it fulfils). The team discussed if it's likely whether the model could get close to 'perfect', even though it's not possible to remove all bias re skin tone due to how darker skin tones reflect less light. They concluded that they'll try and get as close to perfect as possible. The team also raised a past issue when the facial age estimation model was reported to have been spoofed using an animal, in an academic paper. They discussed that this incident occurred with a demo version that did not include liveness detection. This is a different flow than the model that's currently in use by customers.
Update from the Ethics & Trust Committee meeting	The issue of the commercial use of breach or leaked data was presented to the Ethics & Trust Committee. The Committee raised concern about any use of stolen data, despite it being for a good cause. The Guardians agreed with the Committee's advice to avoid using this type of data.
Update on UN End Violence Against Children Compliance Scanner project	Yoti gave an update on this project which it has been working on for 9 months. The tool has been designed to help charities, NGOs and regulators speed up the process of scanning large volumes of websites to review for non-compliance. Phase 1 (collecting Terms and Privacy Policy information, providing a way to view changes at scale in percentage compliance over time) has been completed. The team is currently working on improving the UI and how information is presented to users. They are also exploring what else can be added as part of Yoti's broader Trust & Safety work. The Guardians asked how Yoti plans to measure the impact of this tool. Yoti stated they will gain feedback from the users of the tool. The Guardians recommended that Yoti document this process and track any unintended consequences. Yoti reiterated that regulation often places pressure on bigger companies to comply, which in turn drives users to smaller platforms that aren't required to comply. Automation of scanning technology could help level the playing field across all companies.
Update on Digital ID Match	Following on from the previous meeting, Yoti gave an update on the unintended consequences workshop, for Digital ID Match, as requested by the Guardians. The risks identified have been included in design updates, such as updating consent screens and ensuring that only necessary information can be shared, with consent, with third parties.

---

[1] https://openai.com/policies/data-processing-addendum/

[2] https://openai.com/policies/business-terms/

	The Guardians asked if there were any notable unintended consequences. Yoti shared the unintended consequences review.
Update from the Senior Privacy Counsel and Deputy DPO	The Senior Privacy Counsel and Deputy DPO updated the Guardians on the use of certain UK online platform data for R&D projects. Yoti has performed a privacy assessment and some privacy risks were highlighted. The privacy team, as a result, has implemented a couple of mitigation measures to help make this processing more compliant and more fair. These measures are as follows: - Yoti has updated its website privacy notice to provide general information about this processing, including a way for individuals to opt-out via email. - Yoti has limited retention of this data to 6 months only which is much smaller than standard. - Yoti has recorded this processing in its formal R&D data retention policy which is reviewed annually by the DPO team and R&D.

Discussion on how the Guardians can best add value to Yoti	The Guardians asked where Yoti sees potential strategic issues arising in the next 18 months and how the Guardian Council fits into this in order to support Yoti's growth. Yoti stated that its main focus is to protect users online, including, but limited to, supporting age-appropriate experiences, protecting against fraud and protecting against non-consensual intimate image abuse. Yoti gave examples of how the Guardian Council has previously helped to highlight risks that the Yoti team may have missed. Yoti stressed that it welcomes any new suggestions of potential risk alongside suggestions to support its growth.
	
	Yoti explained that key staff from teams such as R&D and Legal give updates to the Guardian Council as the work from these teams is usually that which is more ethically challenging. Yoti highlighted that some of the challenges it faces includes resistance from privacy groups and sceptics as well as parties who try to perform physical attacks on Yoti's live infrastructure.
	
	Yoti said that it hopes the process of bringing topics to the Guardians forces self reflection across the wider team and builds a collective conscience.
	
	Yoti explained that there are a range of channels through which staff can raise ethical issues. These lead to a triage of topics to consider to bring to the Guardians.
	
	The Guardians reflected on the role of the Chair. The following was agreed:
	- The Chair will help review the agenda and pre-read.
	- Each meeting will now have **2-3 longer focus topics** where the pre-read is not reiterated during the meeting. This is in the interest of balancing frequency of Guardian Council meetings with the length of each meeting.
	- Presenters will make it clear what support they'd like from the Guardian Council.
	- Where possible, the Guardian Council will see projects at their inception, or as early on in the development process as possible.
	- Presenters will classify their sessions into one of the following:
	  - Level 1 Educational (non-controversial, simple topics)
	  - Level 2 Informational (for mid level of complexity topic the Guardians will be asked to consider the proposed next steps and any flagged issues or difficulties)
	  - Level 3 For in-depth review of complex topics (the Guardians will be asked to consider the risks, benefits and proposed mitigations)
	- Where required, Yoti staff can give short updates on relevant topics, limited to 5 minutes each.
	- Since meetings are quarterly, the Guardians confirmed that they are happy to receive interim background materials between meetings to reduce the workload immediately prior to each quarterly meeting.

Meeting Attendance

● = in attendance   O = absent/apologies   ☒ = Not scheduled to attend

Guardians							Yoti staff								

Seyi Akiwowo	●						AB	●						
Dorothy Gordon	●						FCV	●						
Sindhu Joseph	●						JC	●						
Jerry Michalski	●						ED	●						
							JD	●						
							AH	●						
							EH	●						
							AK	●						
							ML	○						
							OR	●						
							RT	●						