

Regulatory Sandbox Final Report: Yoti

A summary of Yoti's participation in the ICO's Regulatory Sandbox

Date: April 2022

Contents

1. Introduction	3
2. Product description	5
3. Key data protection considerations	6
4. Ending statement	12

1. Introduction

- 1.1 The ICO introduced the Regulatory Sandbox ('Sandbox') service to support organisations who are developing products or services that use personal data in innovative and safe ways and where such products or services deliver a potential public benefit.
- 1.2 The ICO initially launched the Sandbox as a beta phase, for an initial group of participant organisations during 2019-2020. In August 2020, the ICO re-opened the Sandbox with a focus on projects involving one of two themes, children's privacy or data sharing. The ICO stated that projects submitted should be at the cutting edge of innovation and may be operating in particularly challenging areas of data protection, where there is genuine uncertainty about what compliance looks like.
- 1.3 Organisations who were selected for participation in the Sandbox following its reopening have had the opportunity to engage with us; draw upon our expertise and receive our advice on mitigating risks and implementing 'data protection by design' into their product or service, whilst ensuring that appropriate protections and safeguards are in place. Yoti was one of the candidates selected for participation in the Sandbox after it re-opened.
- 1.4 Since 2014 Yoti has provided digital identity services and biometric technologies to clients across the globe. One of Yoti's key products is their age estimation technology which estimates a person's age by algorithmically assessing their facial features. Yoti's aim in creating the age estimation was to create a privacy-friendly approach to age verification based only on a person's facial image, with all information being instantly deleted once an age is estimated.
- 1.5 Yoti entered the Sandbox to explore how they could extend the use of their age estimation technology to young people, aged between 6 and 12 years old. This would ensure that the providers of children's only services, such as gaming websites and forums, can create safe virtual environments and online spaces.

- 1.6 Yoti were accepted into the Sandbox on 6 November 2020. On 24 November 2020, the ICO and the Yoti engaged in a virtual scoping meeting to support the creation of the objectives and tasks of Yoti's bespoke Sandbox plan. Following the call in November 2020, the ICO and Yoti agreed the following objectives for Yoti's Sandbox engagement:
- **Objective 1:** Yoti will use data collected on its behalf by Go Bubble (and potentially other data sources) to train its age estimation technology to effectively estimate the ages of children aged between 6 and 12.
 - **Objective 2:** The British Esports Association (BEA) will roll out their new Membership Platform utilising Yoti's age estimation product and Go Bubble's 'Wrap' product to onboard new users. The platform itself should be designed using the standards outlined in the ICO's Age Appropriate Design Code (AADC).
- 1.7 The content of Yoti's Sandbox plan was agreed by Julie Dawson, Director of Regulatory and Policy for Yoti, as well as representatives from both Go Bubble and the British Esports Association (BEA) on 27 January 2021.
- 1.8 Yoti first applied to the Sandbox in collaboration with Go Bubble, a virtual classroom software provider, and the BEA, the national body for esports in the UK. The plan was to develop their age estimation technology for children aged between 6-12 years old using data gathered by Go Bubble and then trialled on a tournament platform created by BEA. However, due to the pressures of the COVID-19 pandemic and subsequent lockdowns both Go Bubble and BEA were unable to contribute to the work during the Sandbox period.
- 1.9 As Yoti's Sandbox work progressed the question as to whether Yoti's age estimation involved processing of biometric data arose and this was subsequently added to their Sandbox plan as an additional task under objective 1 to explore.
- 1.10 In February 2022 Yoti and the ICO completed the last piece of work in Yoti's Sandbox plan, bringing their participation in the ICO's Sandbox to an end.

2. Product description

- 2.1 Yoti's facial age estimation is an age-checking service that can estimate a person's age by algorithmically assessing their facial features. It is an example of an age assurance product, the use of which may help other data controllers comply with their obligations under Standard 3 of the ICO's Children's code. Standard 3 requires organisations to:
- “Take a risk based approach to recognising the age of individual users and ... effectively apply the standards in this code to child users. Either establish age with a level of certainty that is appropriate to the risks to the rights and freedoms of children that arise from your data processing, or apply the standards in this code to all your users instead.”
- 2.2 The product does not require users to register with Yoti, nor to provide any documentary evidence of their identity. It neither retains any information about users, nor any images of them. The images are not stored, shared, re-used or sold on. It simply estimates their age.
- 2.3 Yoti states that facial age estimation is based on a computing technique known as a 'neural network'. Yoti has trained the neural network to be able to estimate human age using a process of 'machine learning', a form of artificial intelligence (AI). The product works by first identifying the human face within the image to ensure the product is assessing a genuine human face.
- 2.4 Next, the facial age estimation product breaks the image down into its component pixels. To the product, each pixel is just a set of numbers. These numbers are fed into the artificial neural network. This is a network of mathematical processing nodes, arranged in layers, that is roughly analogous to the connections in the human brain. Whilst a typical brain has around 100 billion neurons, the artificial neural network has just hundreds of thousands of nodes. The numbers (pixel data) are then fed in, and they percolate through the neural net. Each node performs a mathematical function on the pixel data and passes the result on to nodes in the next layer, until a number finally emerges out the other side. This number is an age

estimate. Facial age estimation works quickly, returning an age estimate in around 1 to 1.5 seconds. The user needs to present their face to the camera uncovered (although glasses do not usually present a problem).

- 2.5 When training their product to work on 6 to 12 year olds, Yoti fed the algorithm a large number of diverse facial images, for which they already knew the subject's age with confidence. The neural network keeps digesting the pixel data from each image, processing the numbers, and trying to get a result which matches the right answer. It keeps repeating the process, adjusting the processing, keeping the variations which bring it closer to the right answer and rejecting the variations which do not help. In other words, it is 'learning'. After repeating the process a significant number of times, it arrives at sets of processing formulae which work best.

3. Key data protection considerations

- 3.1 Yoti's Sandbox participation focused on its development of its age estimation product as the organisation sought to expand the product's usage to children aged between 6 and 12 years of age. In order to aid Yoti with this work, the ICO and Yoti agreed the objectives of Yoti's Sandbox plan, the details of which are outlined below.

Objective 1: Yoti will use data collected on its behalf by Go Bubble (and potentially other data sources) to train its age estimation technology to effectively estimate the ages of children aged between 6 and 12.

- 3.2 In order to complete Yoti's first Sandbox objective, we advised Yoti to consider the data protection risks associated with extending the age range of their age estimation technology in a Data Protection Impact Assessment (DPIA). Yoti did this by revising their existing age estimation DPIA to include a view of the risks associated with processing children's data to retrain their model to work effectively on children aged between 6 and 12 years of age.

- 3.3 In February 2021, Yoti decided to move away from relying on Go Bubble for training data for their purposes. This was due to the evolving global situation during the COVID-19 pandemic as Go Bubble's work pivoted away from the development of the 'Wrap' platform to focus on virtual learning solutions instead. Following this, Yoti worked to procure training data from several sources in order to obtain the required data to train the model, these include:
- Collecting data using a web portal. This collection method involved requesting adult users to provide an image of their child, their child's birth month and year, and explicit parent or guardian's consent for this data to be processed for the purposes of training Yoti's age estimation algorithm for 6 to 12 year olds.
 - Collecting data via a family digital wellness organisation ([Be In Touch](#)) located in South Africa. This collection method involved using Be In Touch's relationship with a network of schools in South Africa. Yoti sent a letter to the parents or guardians of the children attending these schools with a written consent form and suitable privacy information to collect and process the students' data. The ICO did not advise on the data protection implications of this collection of personal data.
- 3.4 During the course of the data collection, the ICO provided ad hoc steers and advice to Yoti as they developed consent and privacy notices. The ICO advised on how to make the content more accessible for children while preserving the detail necessary for UK GDPR compliant consent and privacy notices.
- 3.5 In October 2021, Yoti published its facial age estimation white paper which provided details about its success in creating a viable age estimation model for 6-12 year olds. In the white paper Yoti claimed the model was accurate to within 1.28 years and worked effectively across gender and skin tone.

Objective 2: The British Esports Association (BEA) will roll out their new membership platform utilising Yoti's age estimation product and Go Bubble's 'Wrap' product to onboard new users. The platform itself should be designed using the standards outlined in the ICO's Age Appropriate Design Code.

- 3.6 When Yoti first applied to the Sandbox they applied in collaboration with Go Bubble, a virtual classroom software provider, and the BEA, the national body for esports in the UK. This would allow Yoti to develop their age estimation technology for children aged between 6-12 years old using data gathered by Go Bubble and trailed on a tournament platform created by BEA. However, due to the pressures of the COVID-19 pandemic and subsequent lockdowns both Go Bubble and BEA were unable to contribute to the work in the Sandbox.
- 3.7 If in the future BEA do choose to create their tournament platform utilising Yoti's age estimation technology, they should give due consideration to:
- The standards of the [Children's code](#), ensuring their platform is designed to take into account the best interests of children using it.
 - The principles of data protection outlined in Article 5 of the UK GDPR. BEA should ensure that they conform with these principles and of course their wider obligations under the UK GDPR as they create and run their platform to ensure children's personal data is processed fairly and securely.

Considerations around the nature of biometric personal data

- 3.8 During the course of Yoti's Sandbox participation questions were raised about the nature of Yoti's age estimation and whether the processing activity involved the use of biometric data as defined by Article 4 (14) UK GDPR which states:
- “biometric data’ means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data”

This resulted in an additional task being added to Objective 1 of Yoti's Sandbox work and was explored as follows:

3.9 **Facts as understood by the ICO**

The ICO understands from Yoti that the age estimation tool detects and isolates the human face within an image, and that only this part of the image is processed by the age estimation algorithm. Yoti have stated that each node in the neural network used performs a mathematical function on the pixel data and passes the result on to nodes in the next layer, until a number finally emerges at the other side. This number is an age estimate.

Yoti has stated that the end result age estimation is not based on the neural network looking for age identifiers as a human being would (eg, grey hair and wrinkles) as the computer is just being fed numbers. It doesn't understand or know what the numbers represent or what they mean. The ICO understands that Yoti doesn't try to express the meaning behind the numbers to the computer. When training the neural network, Yoti told the computer what the right answers were within its training database and not what these answers meant.

Yoti has emphasised that its age estimation is not 'facial recognition'; where a computer system is trying to match a particular face against a database, to confirm that person's identity. Rather it is detecting whether there is anything in the captured image that looks like a human face. The facial image the age estimation tool receives is made up of pixels. To the age estimation tool, each pixel is just a set of numbers. These numbers are fed into the artificial neural network.¹

3.10 **Biometric Data**

The ICO took the decision to focus analysis on whether Yoti is likely to be processing special category biometric data in using the age estimation tool. The ICO did not look at the question of whether age estimation will result in the processing of

¹ Further detail can be found at [Yoti age estimation White Paper - October 2021 - 20211026](#)

biometric data. The ICO followed this approach because our view is that the question relating to special category data has more significant practical implications for Yoti – given the additional steps that would be needed to comply with data protection law if age estimation results in the processing of special category data. On the basis that Yoti will be processing personal data, our view is that whether that data is also biometric data is likely to have a limited impact on the data protection compliance obligations that will apply in the context of Yoti's processing activities as we understand them.

For the purposes of the Sandbox, in considering the question of whether Yoti is processing special category data, the ICO has proceeded on the basis that age estimation will result in the processing of biometric data, though noting our comments above.

3.11 **Will age estimation result in the processing of special category data as defined by Article 9 (1) UK GDPR?**

Special category data is defined by Article 9(1) UK GDPR:

“Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited”.

Yoti has expressed the view that the use of their age estimation tool to determine a person's age is not used for the purpose of uniquely identifying that person. Rather it is used to provide an age estimate in relation to that person.

Having considered how the age estimation tool works (as explained to us by Yoti) we have concluded that it can be distinguished from other facial recognition technology (FRT). It appears that Yoti is not using the tool for the purpose of uniquely identifying the individuals whose images are captured using the age estimation tool. Instead, it is being used to categorise them by age without uniquely identifying them.

3.12 **ICO's position as stated in live facial recognition Opinion**

The Commissioner's recent Opinion on the use of live facial recognition technology (LFR) in public places considers the definition of "uniquely identifying" a natural person. It states as follows:

"Biometric data constitutes special category data whenever it is processed "for the purpose of uniquely identifying a natural person...."

"Any biometric data processed for this purpose will constitute special category data, regardless of whether the individual is identified. For example, all biometric facial templates collected and compared to a watchlist will constitute special category data regardless of whether there is a match. As such, biometric data will be special category data in the majority of cases"

- 3.13 Again, the LFR position set out above can be distinguished here from the technology used in Yoti's age estimation tool, because the purpose of the LFR discussed in the ICO's Opinion is to uniquely identify a natural person.
- 3.14 The ICO also considered our guidance on special category data which, goes further than that set out above. In our guidance we currently state that:
- "All biometric data is personal data, as it allows or confirms the identification of an individual. Biometric data is also special category data whenever you process it "for the purpose of uniquely identifying a natural person". This means that biometric data will be special category data in the vast majority of cases. If you use biometrics to learn something about an individual, authenticate their identity, control their access, make a decision about them, or treat them differently in any way, you need to comply with Article 9."
- 3.15 Having reconsidered our guidance in the context of our engagement with Yoti, we have concluded that the above guidance needs to be updated. This is because Yoti's age estimation tool has demonstrated that it is, in some contexts, possible to use biometrics to make a decision about an individual or treat them differently without using that biometric data for the purpose of uniquely identifying that person. We have therefore concluded that it will be appropriate to revise the guidance above to make this clear.

3.16 The revised wording will now read as follows:

“All biometric data is personal data, as it relates to an identified or identifiable individual. Biometric data is also special category data whenever you process it “for the purpose of uniquely identifying a natural person”. This means that biometric data will be special category data in many cases.

If you use biometric data to learn something about an individual, authenticate their identity, control their access, make a decision about them, or treat them differently in any way, it is likely that this will be processing for the purpose of uniquely identifying that individual and it will involve processing special category data which requires compliance with Article 9.

If you believe you have a specific use case where you are processing biometric data for one of the purposes outlined above but not for the purpose of uniquely identifying a natural person, such that you are not processing special category data, you should document your organisation’s rationale alongside a risk based analysis and evidence for this decision in your DPIA. In doing so you should be able to clearly demonstrate how you are compliant with applicable data protection law and why your processing should not be seen as for the purposes of unique identification. If you identify a high risk that you cannot mitigate, you must consult the ICO before starting the processing.”

4. Ending statement

- 4.1 Yoti’s participation in the Sandbox has given the ICO a valuable opportunity to gain insights into the UK’s innovative age assurance sector and how age assurance providers are working to develop tools to children and young people online operationalise the advice in the ICO’s [Children’s code](#).

- 4.2 We have concluded that Yoti's age estimation tool will not result in the processing of special category data. As above, for the purposes of our analysis we have operated on the assumption that the tool will be processing biometric data, but a detailed assessment of this question falls outside the scope of this Sandbox project.
- 4.3 It should be noted that biometric data will nonetheless be special category data in many cases. For accountability purposes, Yoti should document their rationale for the conclusion that Article 9 doesn't apply in the specific context set out above, alongside a risk-based analysis and evidence for this decision in its DPIA. In doing so Yoti should be able to demonstrate clearly how it is compliant with data protection legislation and why the processing is not for the purposes of unique identification. If, during this process, Yoti identifies a high risk that cannot be mitigated, it must consult the ICO before starting the processing.
- 4.4 Informing our existing and future guidance is a positive outcome from the Sandbox process. We have updated our special category guidance to recognise that there may be specific use cases where technology can be used to estimate age without uniquely identifying an individual. Our conclusions will also feed into our continued work in relation to biometric data as our thinking in this area develops.
- 4.5 Yoti is taking meaningful steps to ensure its compliance with UK data protection legislation as it develops its age assurance product. The ICO continues to develop its thinking on biometrics and the use of age assurance technologies.
- 4.6 Moving forwards Yoti should ensure that it continues to follow the ICO's steers, relevant ICO guidance and Yoti's wider data protection obligation when processing personal data.
- 4.7 Yoti's participation in the ICO Regulatory Sandbox demonstrates our role as a trusted regulator in supporting stakeholders to integrate data protection compliance into the fabric of their projects. This helps to ensure that UK data protection legislation is not seen as a barrier to the development of innovative technologies.