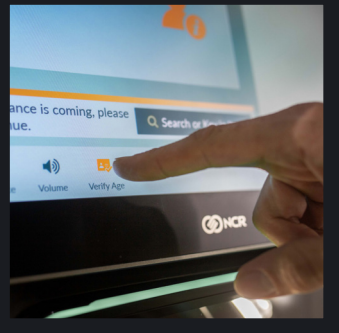


White paper

YAS (Yoti Age Scan) - Public Version





Doing things differently

The solutions we develop make it faster, simpler and safer to prove who you are, and who other people are, online and in person.

The way we build those solutions is a little different to other tech companies. With Yoti, you have full control of your own data. We will never mine your data and once your personal details have been verified by Yoti, only you can see them. Even we can't see them.

But we don't stop there. Using our business as a force for good is something everyone at Yoti is passionate about. That's why we're a certified BCorp - internationally awarded for our commitment to rigorous standards of social, environmental performance, accountability and transparency.

YAS (Yoti Age Scan) is just one innovative use of our digital identity technology. We built it to give everyone a secure and private way of proving how old they are in different everyday scenarios: from age checking in online stores and social platforms, to supermarket self checkouts, bars and clubs. In this whitepaper we'll explain everything you need to know about the new way to prove your age.

Contents

What is YAS and what can it do?	04
Data privacy and network security	06
How does it actually work?	07
Tackling the challenge of age determination	08
Human ability to determine age	10
More on how it works	11
Practical use	13
How accurate is YAS?	14
Public acceptance of AI technologies	16
Appendix	18
Data used to build the model	18
Accuracy across the entire dataset	18
Accuracy by size of dataset	19
Accuracy by gender	20
Accuracy by skin tone	21

What is YAS and what can it do?

YAS is a secure age-checking service that can estimate a person's age by looking at their face. We consider it to have wide application in the provision of any age-restricted goods and services, both online and in person.

YAS is designed with user privacy and data minimisation in mind. It does not require users to register with us, 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, not re-shared, not re-used and not sold on. It simply estimates their age.

In a retail setting, YAS can be used at a point-of-sale terminal with a dedicated camera, letting a consumer use a self-checkout without the need for staff assistance. This is not only quicker and less of a nuisance for shoppers, but can greatly reduce friction between them and retail staff.

For general online use, it can be embedded into web pages, and receive an image of the user's face from a webcam connected to their computer or the camera in their mobile device, ideal for controlling access to age-restricted gaming, gambling and other adult content (pornography).

A further application is use at the entrances to age-restricted premises such as bars, nightclubs and casinos. In this kind of application, YAS offers clear advantages — it does not get fatigued on a long shift¹, and it cannot show favour to personal friends or bias against individual customers.

YAS 'always forgets a face'

YAS is an emerging technology, and its age estimates are subject to a margin of error. To allow for this, the system is configurable to set whatever threshold a business or regulator requires for instance, requiring those over 18 to be estimated as at least 25 - a buffer of 7 years. Where someone is over 18 but appears to be under 25, they can use either the Yoti app, where their account is anchored with a verified ID document, or by a manual photo ID check from a member of staff).

1. Studies have shown that the objectivity of human judgement of this kind can be significantly affected by hunger and fatigue—see for instance Danziger, Levav, Avnaim-Pesso (2011) Extraneous factors in judicial decisions, *Proc Natl Acad Sci USA*, 108 (17): 6889–6892. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084045/>

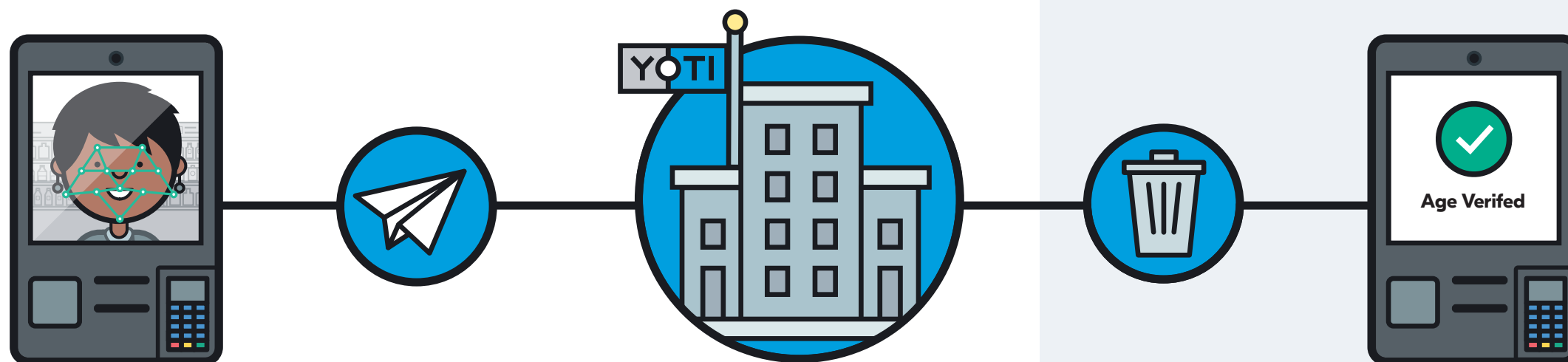
Data privacy and network security

YAS has been designed with data privacy and security as primary considerations.

The user does not have to register to use the service, and does not have to provide any information about themselves. They simply present their face in front of the camera. Their image is not stored locally on the point of sale terminal. It is securely transmitted to the Yoti backend server (currently hosted in the United Kingdom), secured by TLS 1.2 encryption. After the age estimate is performed, the captured facial image is deleted from Yoti's backend servers.

How does it actually work?

YAS is based on a computing technique known as a 'neural network', which we have trained to be able to estimate human age using a process of 'machine learning'. This is a form of artificial intelligence (AI), and is increasingly used in a wide variety of applications, from driverless cars to medical diagnosis, from tailoring online advertising to detecting credit card fraud. We discuss machine learning in more detail below, but first some context on the problem we are using it to solve.



Tackling the challenge of age determination

Determining a person's exact age in the absence of documentary evidence of their date of birth is a difficult task. Indeed, the truism that 'age is just a number' could be said to have a sound scientific basis. By 'ageing' in a medical sense, we mean the physiological changes which occur when individuals develop and grow from juvenile to mature forms, and then the types of damage that progressively accumulate within the human body as time passes. The important point is that the rate at which human bodies 'age' in this way is influenced by numerous external factors other than simple passage of time. Factors that affect the aging process, both in the long and short term, can include: quality of diet and nutrition, exposure to disease, adverse environmental conditions, use of narcotics, physical labour, stress and lack of sleep. Clearly, there are large variations throughout populations as to how different individuals are exposed to these aging factors. The more extensively we look through different countries, ethnicities, and socio-economic groups, the wider these variations in exposure to aging factors become.

It may be surprising to learn that there are currently no entirely reliable medical or forensic methods to determine human age. Two of the more commonly attempted medical techniques focus on trying to ascertain whether the subject is above or below the legal age of maturity. These are X-ray or Magnetic Resonance Imaging of bone structure in the wrists (the degree to which the cartilage between the carpal bones has ossified) and dental X-rays (examining the maturity of wisdom teeth). However both of these methods have a typical margin of error of at least two or three years, and for individuals with an atypical history to the general population, the error can be significantly worse. Due to this unreliability, their use has proved controversial — for instance, their use by immigration authorities to attempt to differentiate between child and adult refugees who have no documentation.

Other medical techniques examine 'biomarkers' taken from blood or tissue samples. Examples include measuring the degree of DNA methylation present, the length of the 'telomere' portion of chromosomes, or the serum levels of the metabolite C-glycosyl tryptophan.

Whilst these biomarker techniques tend to provide good indicators of ageing processes in an individual, they do not correlate reliably with their chronological age from date of birth.

Ultimately, it could be argued that much of the difficulty in trying to measure 'age' (that is, a person's chronological age from their date of birth) arises because 'age' defined this way is a rather arbitrary quantity that does not mean anything definite in physiological terms. Science can accurately measure the extent to which a person's body has aged (that is, how to what extent it has developed, grown, matured and decayed), but cannot always reliably determine how many years it took for their body to arrive at that state.

Human ability to determine age

Notwithstanding the difficulty in devising an accurate forensic test for age, people still possess a reasonably good ability to guess someone's age simply by looking at them. We can all do it, usually coming within a few years of the right answer. How do we manage it? In terms of facial features, what are the tell-tale signs we look for?

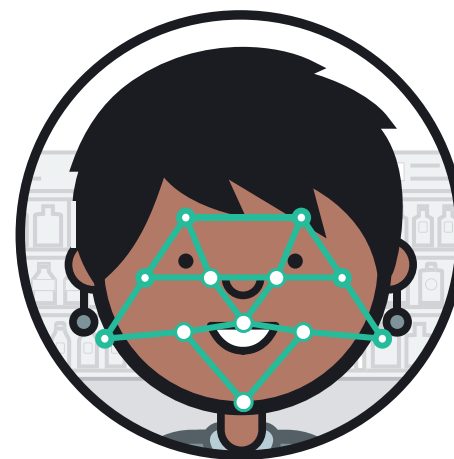
The most obvious visual cues include bone structure (bones grow and develop as we pass from child to adulthood), skin tone (wrinkles, elasticity) and hair colour (greyness), male baldness or facial hair after puberty. We could add dozens more cues to this list. However, whatever the detailed nature of the visual cues, the more general point is this: as humans, we simply learn "that's what people of a particular age look like". As we go through life, we encounter other people, we see what they look like and we learn how old they are, with varying degrees of precision (e.g. "a baby", "14", "mid-40s", "79" and so on). We accumulate this information and experience throughout our lives, and our brains can use it to make quick intuitive judgements. The

extent of our previous experiences will be an important factor in how good our guesses are. We will be more accurate at guessing the age of someone from our own familiar peer group than from one we've not encountered.

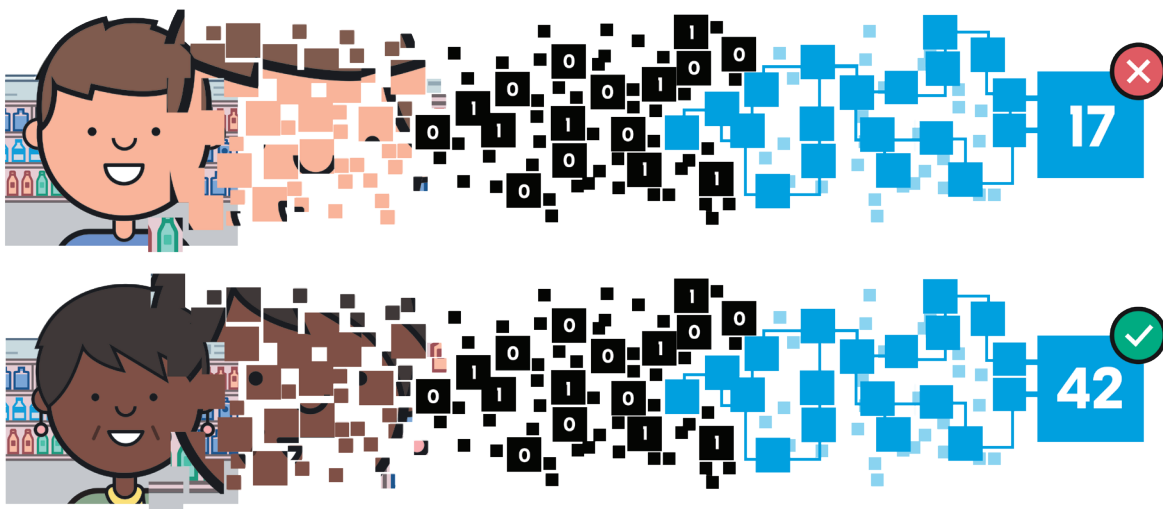
It is worth emphasizing that, although we might be able to retrospectively rationalise or refine our guess at someone's age, our initial judgement is more or less intuitive. We are not consciously following some step by step, rule-based method (for instance "add five years if there are wrinkles", or "add ten years for grey hair"). In effect, we don't 'know how we do it' — generally, our brains process the image and form an instinctive judgement, in line with what we've learnt from past experience, faster than any conscious deliberation or systematic evaluation of facial features. It turns out that this 'black box' approach to describing our cognitive process (that is, simply training our brain with data, without worrying too much about how it works) can actually be employed as a successful technique in machine learning too.

More on how it works

The first challenge for YAS is 'face detection'. It has to examine the image it gets from the camera, and work out which bit of it is an actual human face. Only this portion of the image is then fed into the neural network to get an age estimate. This stage also allows for basic error checking: if the system can't find a face in the image (for example, because a customer didn't position themselves properly in front of the camera, or some inappropriate object is put there) then the system can return an error message instead.



We now come to the interesting bit. The facial image is made up of pixels. To the computer, 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. We feed numbers (pixel data) 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.



More on how it works

It's an obvious question to ask 'how is the neural network processing the data? What is it looking for — wrinkles? grey hairs?' and so on. However this is a rather human way of thinking about it, and it's not really a very useful question to ask: to the computer, it is just being fed numbers. It doesn't 'know' what the numbers represent or what they mean. We don't try to tell it that. What we have told it, in the training phase when YAS was being developed, was what the right answers were. In the training phase, we fed it tens of thousands of diverse facial images, for which we 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, rejecting the variations which don't help — in other words, it is 'learning'. After repeating the

process a huge number of times, it arrives at sets of processing formulae which work best. To a human, these formulae would be bafflingly long and complex, and next to meaningless (and no, we're not going to print them here...for one thing, they wouldn't fit on the page!). However it has effectively created a very complex model of age determination that is far superior to relying on a set of hand-crafted instructions that a human programmer might supply.

The quality of the training data is crucial to any machine learning process. We use many thousands of images from Yoti users (as explained to them at onboarding²). Yoti users want Yoti to make their lives safer and simpler, and they understand that using their data for research and development purposes is how we are able to improve and develop the products and technology to achieve this. We will publish whitepapers that demonstrate such applications. For YAS,

these research images are tagged with only two attributes taken from a verified ID document that they have uploaded: their gender and their year of birth. Supported documents include passports, driving licences and national ID cards. We believe the size, diversity and verified age accuracy of this training data set gives YAS an advantage over competing solutions.

Practical use

YAS works quickly, returning an age estimate in around 1 to 2 seconds. The user needs to present their face to the camera, uncovered (although glasses do not usually present a problem). Dim lighting is not helpful; bright ambient light works best. Beards and facial disfigurement are further areas of research.

How accurate is YAS?

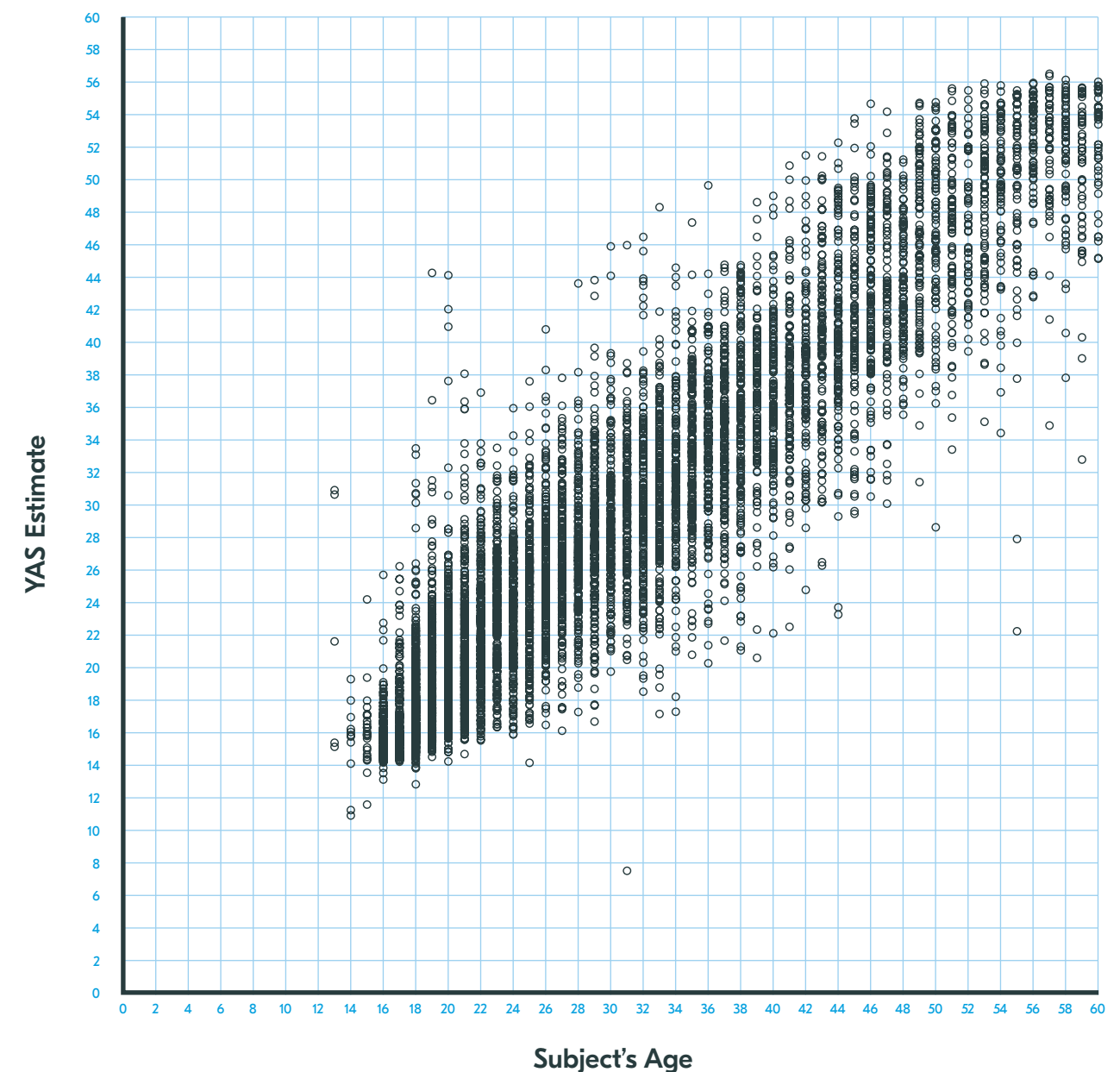
We believe that when presented with a clear facial image, YAS's ability to estimate age compares favourably with human abilities.

Research in this area³ suggests that the root mean square error in human guesses across an age range of 7 to 70 approaches ± 8 years. Furthermore, when viewing a succession of faces, a person's judgement tends to be influenced by the preceding faces they have just seen, which is not a problem that affects YAS. Humans tend to systematically underestimate the ages of older people, and over estimate the age of younger people. The latter problem clearly has particular implications for provision of age-restricted goods and services, where we need to check whether teenagers are above or below a required legal age.

The scatter graph overleaf shows how our model performs with our test data set. Each circle represents a single human subject; the graph plots the subject's actual age (x axis) against the age estimate produced by YAS (y axis). If YAS always guessed correctly, the plot would be a straight diagonal line. As can be seen, whilst not perfect, there is a good correlation between actual age and estimated age, and the deviation is reasonably consistent throughout the entire range. Currently, the mean error

across the entire data set is ± 3.45 years. Further detail accuracy, broken down by gender, skin tone and age range, is presented in this paper's appendix. We believe this accuracy will improve still further in years to come, as our Age Scan is trained on an ever greater set of data from Yoti users. We intend to continue comparing YAS's accuracy against that of ordinary human estimators, and against people who believe they have a special aptitude at estimating age, to demonstrate that YAS is usually a more accurate approach (and cheaper and faster).

For this reason, Yoti recommends using YAS as part of a strategy such as the UK Wine and Spirit Trade Association's 'Challenge 25'. This type of retail strategy works as follows: Certain goods and services can only be sold to customers over a particular age (e.g. 18 years old). However it is difficult for human staff to be sure whether someone is over 18 just by looking at them. Conversely though, it is fairly easy to tell if someone is significantly older than 18, and customers in this age range would find it an unjustifiable inconvenience to have to show ID to prove their age. Therefore, the store's policy is to only require customers to prove their age if they appear to be under 25.



Graph displays results obtained with Yoti age estimation model as of 3-Jan-2019

3. Clifford CWG, Watson TL, White D. (2018) Two sources of bias explain errors in facial age estimation. R. Soc. open sci. 5:180841. <http://dx.doi.org/10.1098/rsos.180841>

How accurate is YAS?

YAS can be configured to work with legal age thresholds in a similar way: choose a suitable buffer that is comfortably outside YAS's margin of error, and configure the system to estimate whether customers are above or below that threshold. At present we would not permit a buffer of less than five years.

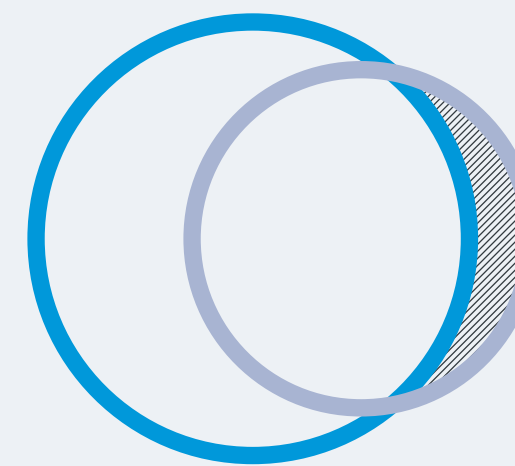
For instance, selling alcohol in a jurisdiction with a legal age restriction of 21, Yoti suggests choosing an initial threshold of 30. If YAS estimates that the customer is at least 30 years old, then no further age checking is required. If YAS estimates that the customer is below 30, then they will be directed into a user flow where they need to present documentary proof of their age (for example, using their Yoti app that is anchored to their passport, driving licence or national ID card).

This is great news for the 30 plus population — they will not need to provide ID document evidence of their age and they will be able to happily leave their documents at home. In some countries regulators may initially look for a high buffer of over ten years. With a 15 year buffer we would currently have a nil per cent error rate and with a 10 year buffer we would have a 0.86% error rate. Over time regulators will be able to set lower buffers with confidence.

Public acceptance of AI technologies

When discussing the accuracy of YAS, it is worth considering a general point about machine learning and the public's attitude to AI technologies of this kind: namely, how unforgiving humans tend to be in regard to mistakes made by AI.

Whilst we feel it is fair to claim that the accuracy of YAS generally compares favourably with human judgement in the broad majority of cases, there will inevitably be rare occasions where it 'makes mistakes'. Of course, humans make mistakes too. However, sometimes machine learning systems make mistakes that no human would have made. This is illustrated in the Venn diagram below:



■ Errors made by humans ■ Errors made by machines
 ▨ Errors humans react more badly to

As can be seen, typically, humans make errors, just as a well-trained machine learning system does. Furthermore, in most of the cases where the machine system gets it wrong, a human would make the same mistake. However, humans tend to be much more bothered by the small percentage of cases on the right of the diagram — these are cases where the machine learning system makes a mistake, but a human would not have been fooled. It can be argued that this is an irrational reaction, and objectively, the machine learning system is no worse than the human judgement it is replacing (sometimes it may even be better overall!). Nevertheless, the general public may often unduly focus their attention on the machine failings, until they become comfortable with the new technology.

Yoti has signed the Safe Face Pledge⁴, which encourages companies using artificial intelligence to ensure that facial recognition technology is not misused.

4. <https://www.safefacepledge.org>

Appendix

This appendix provides further detail on the current accuracy of YAS's estimates. We expect these figures to continue to improve as the volume and diversity of our dataset increases.

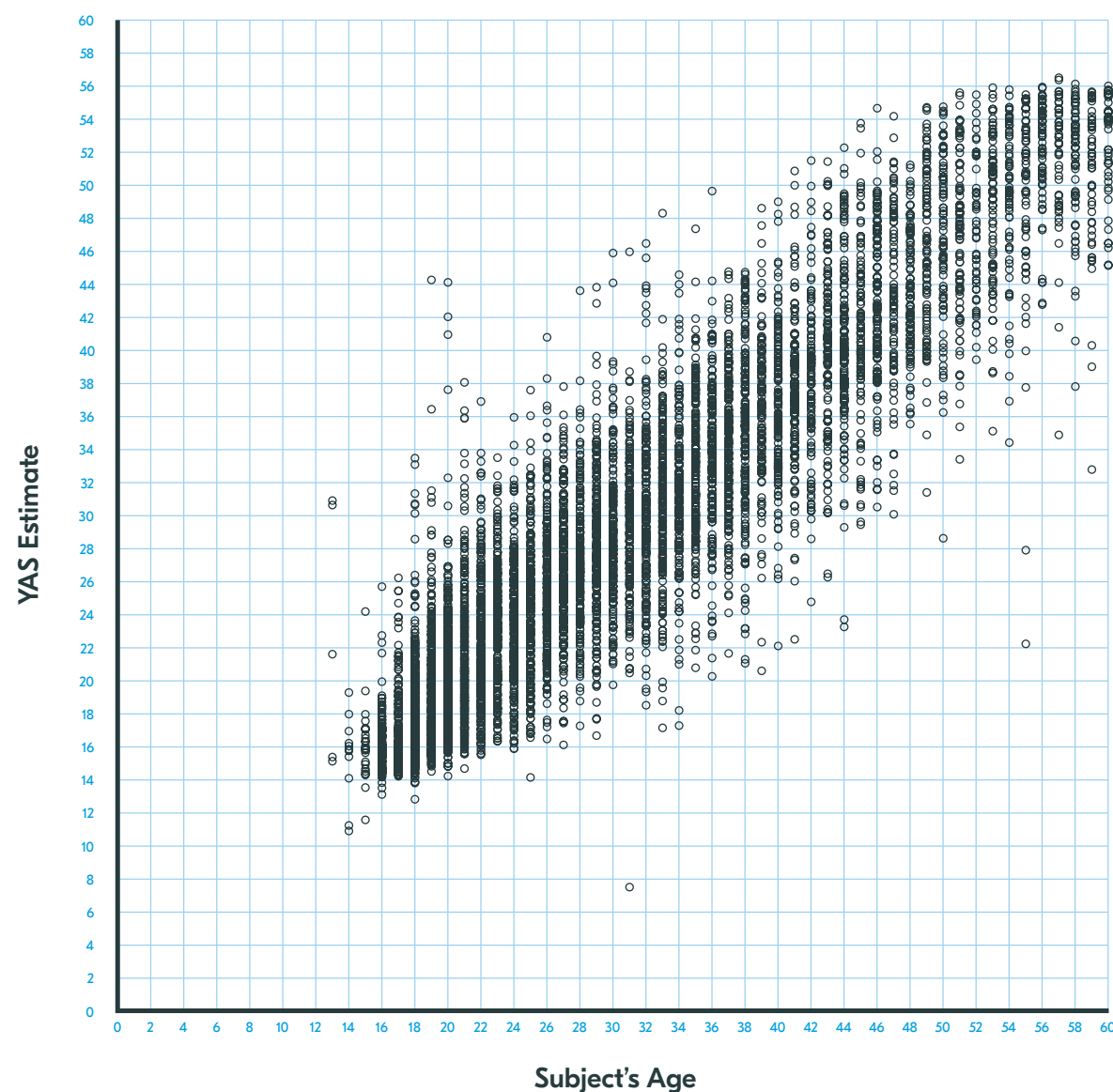
Data used to build the model

Yoti has invested significantly in building a leading R&D team since early 2015 working on a variety of AI initiatives.

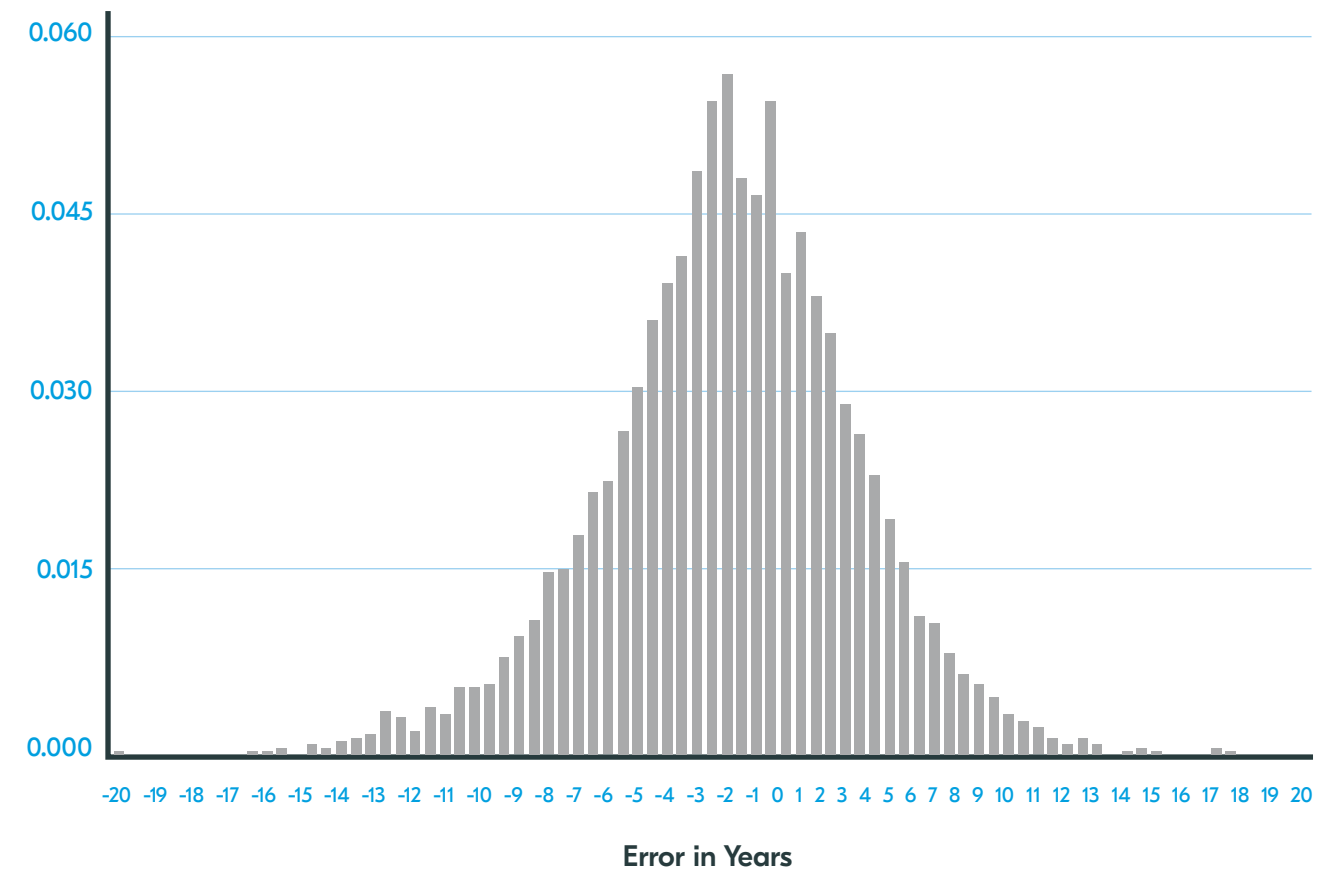
The current production model of YAS (3rd January 2019) was built using a training data set taken mainly from Yoti users⁵, and from a public domain source⁶. Further age-verified images were gathered by Yoti in Nairobi, Kenya.

Accuracy across the entire dataset

Mean absolute error in age estimates (across entire data set) is ± 3.45 years



The range of errors tends towards a normal distribution, with a standard deviation of 4.56. This is illustrated in the chart below.



Accuracy by size of dataset

Following the initial pilot study, we envisage continual periodic retraining of the model on an ever-expanding data set, as we continue to add further age-verified images taken from Yoti users at onboarding. We expect the accuracy of the model to continue to improve on this basis.

The diversity of the dataset (for instance in terms of gender, skin tone and age) is important, as age estimates tend to be less accurate for any group that is under-represented in the training data. We are optimistic that we will be able to broaden the diversity of our training data. Where appropriate we shall endeavour to undertake further targeted fieldwork in this regard.

5. The Yoti app explains to new users when they are in the process of creating an account that their biometric data (specifically, their photo and year of birth) may be used by our R&D team for internal research. More detail is set out in our Privacy Policy <https://www.yoti.com/privacypolicy/>

6. Images were taken from the Computer Vision Center and University of Barcelona's APPA-REAL Database, <http://chalearnlap.cvc.uab.es/dataset/26/description/>

Accuracy by gender

We performed a test of the current production model using 7,574 age-verified images from Yoti users, each tagged with the user’s gender as present on an uploaded identity document. We consider it a reasonable assumption that the gender and age distribution of these test images was similar to that of the original training images used to create the model. That being the case, we conclude that the slightly lower accuracy of age estimates obtained for females in the 35–60 age band is simply a reflection of that group currently being under-represented in the training data. Moreover we feel justified in having some confidence that the noticeably higher accuracy obtained throughout the 13–35 age band will be achievable for all bands and genders once we have increased their numbers in our training set. Additionally we may hypothesize that the mean error rate will generally tend to be higher for older people than younger people in any case, because older people will have been exposed to various environmental factors (such as the effects of tobacco, alcohol, sunlight) for longer.

Age Band

	13 – 35		35 – 60		All	
	Mean error/years	Number of subjects	Mean error/years	Number of subjects	Mean error/years	Number of subjects
Male	2.99	3,586	3.83	2,233	3.31	5,819
Female	3.50	1,267	5.02	476	3.92	1,743
All genders	3.12	4,853	4.03	2,709	3.45	7,562 ⁷

Accuracy of YAS by gender, across different age bands

Accuracy by skin tone (ST)

Just as with human ability to estimate age, machine learning works best when it can be trained on a diverse set of subjects. If the training dataset does not contain sufficient diversity of skin tones, then age estimate will tend to be less accurate on skin tones that were under-represented in the training data. For the purposes of analysis we have classified facial skin tone on a four degree scale, from ST1 (lightest) to ST4 (darkest).

In the test exercise below, we took 463 images of verified age, gathered from students at Queen Mary’s University, London, and manually tagged their skin tone and gender, then obtained an age estimate for them from the current YAS production system. The subjects were predominantly between 18 and 30.

The results are tabulated below, giving the mean error in the age estimates, broken down by skin tone and gender. They show that there is little difference in accuracy between skin tones for this 18–30 age range.

Age Range 18–25*

	Skin Tone							
	ST1 (lightest)		ST2 / ST3		ST4 (darkest)		All	
	Mean error/years	Number of subjects	Mean error/years	Number of subjects	Mean error/years	Number of subjects	Mean error/years	Number of subjects
Male	2.34	32	2.57	111	2.69	43	2.56	186
Female	2.81	62	2.39	164	2.50	51	2.50	277
All genders	2.65	94	2.46	275	2.59	94	2.53	463

*Accuracy of YAS by skin tone, mainly 18-25 age range (university students in London)

7. The test set contained 12 images of “Other” gender, which is why this number is lower than the 7,574 images referenced previously.

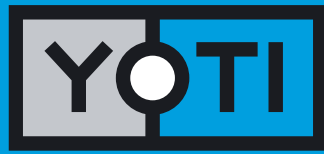
Additionally, we performed testing on a further set of ST4 individuals of verified age (1,625 images collected from both Yoti users and from individuals in Nairobi, Kenya), as we believe this may currently be our least well represented demographic. The results are tabulated below, broken down by age and gender.

Whilst accuracy for ST4 over this wider age range is not as good as for the above university-age subjects from London, we believe this is largely due to these ages being underrepresented in the training data set. We expect to see continuing improvement as the diversity of our training data set increases

Skin Tone ST4

	Age Range					
	13–34		34–60		All	
	Mean error/ years	Number of subjects	Mean error/ years	Number of subjects	Mean error/ years	Number of subjects
Male	4.03	405	4.62	795	4.42	1,200
Female	5.36	285	6.25	140	5.65	425
All genders	4.58	690	4.86	935	4.74	1,625

Accuracy of YAS for skin tone ST4, all ages (including volunteers in Nairobi)



To find out more visit
yoti.com

