Privately

Age Appropriate Design Code Engagement Report – Executive Summary

September 2021

Information Commissioner's Office

Background & Scope

Under section 123(1) of the Data Protection Act 2018 (DPA18), the Information Commissioner is required to produce a code of practice on standards of age appropriate design ("the Code"). The Code applies to "relevant information society services which are likely to be accessed by children" in the UK. This includes many apps, programs, connected toys and devices, search engines, social media platforms, streaming services, online games, news or educational websites and websites offering other goods or services to users over the internet. It is not restricted to services specifically directed at children.

The Code sets out 15 headline standards of age appropriate design that companies need to implement to ensure their services appropriately safeguard children's personal data and process children's personal data fairly. The Code came into force on 2 September 2020, with a 12 month transition period. Organisations are required to conform to the Code by 2 September 2021.

Following the publication of the Code, the ICO Assurance Team have developed an Age Appropriate Design Code (AADC) Audit Toolkit to support the intended future work by the Office.

Privately agreed to assist the ICO in the development of an AADC Toolkit, including a period of engagement to review the practicality of the toolkit measures. The primary purpose of the engagement was to provide the ICO with further practical information and expertise to finalise their AADC Audit Toolkit, and to provide Privately with assurances regarding their compliance with data protection (DP) legislation and relevant sections of the ICO's AADC.

The engagement primarily focussed on Privately's 'real time' multimodal age estimation technology. This technology is built using on-device edge-artificial intelligence (AI) and uses a combination of voice, image and text, which can be integrated into apps, games and devices to assess a users age. For the most part, Privately operate as business to business (B2B); their technology is directly integrated into the client's console firmware or existing application. The age estimation model operates within the end-user's device and as a result, Privately are currently processing very limited personal data. As Privately operate for the most part as B2B, they are likely to be a data processor for the majority of their client relationships and as such, a number of controls within the AADC Toolkit were not applicable to their processing activities. The report has been produced with this in mind.

The engagement focussed on compliance with DP legislation and the Code in the following areas:

- Data protection governance and accountability
- Due diligence
- Data protection impact assessments
- Transparency
- Detrimental use
- Default privacy settings
- Data minimisation
- Data sharing

A series of presentations were led by Privately to showcase their current working practices and how these comply with DP legislation and the AADC. Prior to each presentation Privately shared key documents with the ICO team, and these have been reviewed against the above mentioned scope areas. The content of these documents were also used to facilitate discussion throughout the engagement. The report makes reference to these documents where applicable. In addition to their own presentations, Privately organised additional calls with their client's to further showcase how their multimodal age estimation technology works in practice. The ICO team would like to thank Superawesome, AGEify and Scandit for their contribution to the engagement.

Overview of Service

Privately's Age Estimation solution is based on a computing technique known as "deep learning", which uses real-life examples to learn complex mathematical models to detect age. The technology to estimate age is based on face pattern (image), voice pattern (audio) and writing (text). The age estimation models operate within the end-user's device and make a privacy preserving age estimation of the user's age. No personal data of the user leaves the device. There are currently two types of solutions: SDK and On-browser.

Privately's age-estimation SDK is directly integrated into the client-specific console firmware and also on the Client's existing applications on mobile devices, PCs and Macs. Age estimation is performed on a user device, thus user biometric data never leaves the user device. The SDK solution can perform both open and silent estimation. Furthermore, it performs model download/update independently from the time of estimation, therefore reducing the delay of the age estimation.

The On-browser integration delivers its age estimation technology to client browsers, so again, user biometric data never leaves the end-user device. Estimation is done real time, on-device. No new software is installed on end-user devices. This solution does not permit silent estimation.

Privately’s AI models are deployed in mobile phones as lightweight browser extensions. In order to do so, Privately opts to compress models into deployable sizes. Privately employs a unique combination of knowledge distillation (effectively learns a small student model from a large teacher model) and quantization (reducing the precision of inner model parameters), both of which are widely used in edge computing applications.

It is possible to build-in age estimation into the user’s console or app and perform estimation ‘silently’ as often as might be required without disturbing the user or extracting their data. A user age label can be a metadata shared to the Client server in every session to dynamically adapt the digital environment to the age of the user. This is especially relevant when many people use the same account.

Their models are trained to estimate an age range (0-7, 8-12, 13-17, 18-25, 26-34, 35-49, and 50+). In case detailed annotations are not available or not reliable, they perform their tests on binary classification (child, adult) . They measure the performance of the models in terms of classification accuracy, using the F1-score metric2. The model estimates the age range of the user using a ‘confidence rating’. In other words, from the 7 different age ranges that can be classified the model calculates a % confidence you fall in the one displayed based on the highest likelihood / probability. This is not communicated to the end user in the live environment. One of the age ranges might be selected with only a low confidence rating, based on probability. The service is flexible to change the age ranges, or produce exact calculations e.g. probability of person being 16+ or under 25. The range / service is driven by clients requirements.

In developing the system, the data is stratified: in other words, each gender, skin tone, language, and other characteristics are fairly represented in both the training and test datasets, with similar proportions.

To assure further reliability, they have synthetically augmented these datasets by introducing random blurs and brightness modifications on original images and by introducing random background noises to the original sound files.

The machine learning models are currently trained ‘offline’ from data sources that have been acquired but not on the user data itself. This data comes from a range of open, licensed sources or from manifestly public sources such as Mozilla common voice, Google audioset, Wikipedia and YouTube video uploads. These data sets include a variety of human faces from different ethnicities and age ranges.

Academic researchers are then employed to test the training model outputs and establish a baseline / benchmark.

There are plans to more widely use synthetic data sets, i.e. images of real people that are technically modified whilst retaining human like properties, to augment and expand existing training data sets using a small selection of original users. In other words, using one photo to create multiple synthetic images by augmenting the image by, for example, partial blurring or exclusion of bits of face or changing skin tones or the main image to look younger or merging multiple faces into a third face to create a synthetic face. This technology is still in development.

Areas for Improvement

- Privately should ensure they have properly documented all necessary information relating to their identified lawful bases for processing, in line with their Article 5 (2) Accountability obligations.
- Privately should ensure they have identified the privacy and AADC related training needs of their staff have been identified and documented, and that plans are put in place to fulfil those needs.

Good Practice

- The Zero Data Principle, which appears to fundamentally underpin Privately's approach to development and commercialisation, is an excellent approach which helps to ensure that Privately's final product is as privacy conscious as possible, and seriously reduces the risk of a data subject's rights being breached, or any aspect of data protection legislation being infringed by the operation of this service.

Disclaimer

The matters arising in this report are only those that came to our attention during the course of the engagement and are not necessarily a comprehensive statement of all the areas requiring improvement. The responsibility for ensuring that there are adequate risk management, governance and internal control arrangements in place rest with the management of Privately.

We take all reasonable care to ensure that our report is fair and accurate but cannot accept any liability to any person or organisation, including any third party, for any loss or damage suffered or costs incurred by it arising out of, or in connection with, the use of this report, however such loss or damage is caused. We cannot accept liability for loss occasioned to any person or organisation, including any third party, acting or refraining from acting as a result of any information contained in this report.

This report is solely for the use of Privately. The scope areas and controls covered have been tailored to this engagement and, as a result, the report is not intended to be used in comparison with other ICO report.