EXCELENTÍSSIMA SENHORA DOUTORA JUÍZA DE DIREITO DA 37a VARA CÍVEL DO FORO CENTRAL DA COMARCA E CAPITAL DO ESTADO DE SÃO PAULO

Autos no.: 1090663-42.2018.8.26.0100

AÇÃO CIVIL PÚBLICA

I. About Access Now

Access Now is a global civil society organisation dedicated to defending and extending the digital rights of users at risk.[1] Through representation in ten countries around the world Access Now provides thought leadership and policy recommendations to the public and private sectors to ensure the continued openness of the internet and protection of fundamental rights and wields an action-focused global community of nearly half a million users from more than 185 countries. Access Now also operates a 24/7 Digital Security Helpline that provides real-time direct technical assistance to affected communities and vulnerable persons around the world. Access Now is non-partisan, not-for-profit, and not affiliated with any country, corporation, or religion.

Access Now routinely files amicus briefs on issues related to digital rights in domestic jurisdictions, including the United States, Cameroon, and Indonesia.[2] Access Now has also previously submitted amicus briefs with regional courts, such as the European Court of Human Rights and the Economic Community of West African States Court of Justice (ECOWAS).[3] The issues Access Now has submitted briefs on include internet shutdowns, website blocking, online platforms liability for content, data privacy, and surveillance. In 2016, Access Now was granted special consultative status to the United Nations (U.N.) Economic and Social Council (ECOSOC). Access Now has also released a number of publications dealing with the subjects of privacy, artificial intelligence, data protection, and the use of facial recognition.[4]

---

[1] Access Now, About Us,<https://www.accessnow.org/about-us/>

[2] https://www.accessnow.org/access-now-joins-legal-brief-supporting-privacy-facebook-users/, https://www.accessnow.org/access-now-isf-file-legal-intervention-cameroon-shutdown/, and https://www.accessnow.org/indonesians-seek-justice-after-internet-shutdown/.

[3] See, for example: https://www.accessnow.org/website-blocking-russia-goes-european-court-human-rights-access-now-intervenes/, https://www.accessnow.org/delfi-as-v-estonia-a-blow-to-free-expression-online/, https://www.statewatch.org/news/2018/dec/echr-hu-magyar-jeti-zrt-v-hungary-hyperlinks-defamation-judgement-4-12-18.pdf, https://www.accessnow.org/cms/assets/uploads/2016/02/ECtHRIntervention.pdf, and https://www.accessnow.org/judges-raise-the-gavel-to-keepiton-around-the-world/.

[4] See, for example: https://www.accessnow.org/proteccion-de-datos-es-importante/, https://www.accessnow.org/cms/assets/uploads/2018/04/manual-de-proteccion-de-datos.pdf, https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf.

II. Case summary and topics covered

The present case identified with “Autos no.: 1090663-42.2018.8.26.0100” is a Public Civil Action in which the plaintiff is the Brazilian Institute of Consumer Protection (hereafter IDEC) and the defendant is the concessionaire of the São Paulo S.A. subway line 4 (hereafter ViaQuatro). IDEC filed a lawsuit arguing that the implementation of the Digital Interactive Doors System (hereafter DID system) in the yellow line of the metro of São Paulo violates consumer and data protection rights. The DID system, produced by an Artificial Intelligence (hereafter AI) analytics company AdMobilize and installed by the defendant ViaQuatro, is a data processing system consisting of advertising panels with a camera located above several entrances to the metro. According to ViaQuatro (examples: fls.152, 155, 158, 369 dos autos), AdMobilize (fls. 1787 -1815 dos autos), the Instituto Brasileiro de Peritos (fls. 427 - 450 and fls.1816 - 1823 dos autos), and the expert Raul Spiguel (fls.1753 - 1815 dos autos), the DID system is able to recognize human faces and detect the emotion, gender, and age of passersby who look at the advertising panels.

This expert opinion will provide clarifications of several aspects of the case, focusing in particular on the nature of the data processing activities carried out, as well as highlighting additional concerns raised by the contentious scientific status of ‘emotion detection’ and the discriminatory impact of gender detection. To this end, we will analyse: i) the personal data processing activities involved in the DID system including the claims made about anonymisation of data and anonymous data; ii) the scientific status of so-called emotion detection technology; and iii) the potential for discrimination due to the classification of gender as a male-female binary in the functioning of the DID system. Our opinion builds on and complements the expert opinion submission made by the Institute for Research on Internet and Society (IRIS) (fls.1635 - 1664 dos autos), in particular regarding the following points: i) there was no clear and adequate information to consumers regarding the operation of the DID system; ii) the passengers’ right to choose their data collection was not respected by this system; and iii) there was a lack of sufficient information regarding the anonymisation of personal data and the possibilities for identification of passersby.

III. AdMobilize misrepresents its data processing activities

Our main point concerns the misrepresentation of the data processing activities carried out by AdMobilize through the DID system. In the Audience and Crowd Analytics FAQs (hereafter FAQs) (fls. 1789 - 1802 dos autos) document provided by the defendant ViaQuatro, and in further comments provided by the defendant and the defendant's experts, the Instituto Brasileiro de Peritos and Raul Spiguel which we will directly reference in what follows, (the later two brought to the case

by ViaQuatro) make a number of contentious claims about AdMobilize DID system's data processing activities, namely:

1. That the DID system only uses **facial detection** and does not use **facial recognition** (fls. 1791 dos autos)
2. That AdMobilize "never collect[s] private information in the first place" (fls.1798 dos autos) because all the data it collects is anonymous from the beginning, and that the software used by AdMobilize does not save or store "any sort of unique biometric information" and "cannot, and do[es] not personally identify individuals, at any time" (fls.1792 dos autos).
3. That the demographic information (about age, gender, and emotion of passersby) can be **extracted** (fls. 1796 dos autos) or **derived** (fls.1793 dos autos) from 'views.'

We believe that the claims made by ViaQuatro and its experts are misleading and should not be taken as legitimate technical descriptions of the DID system's functioning. While they often report factually correct information regarding the functioning **of certain aspects of the system**, they fail to discuss other relevant aspects of the system's functioning and thereby fail to alleviate the concerns raised by the IRIS opinion. Thus, we will demonstrate that the DID system **does in fact use a form of facial recognition, namely facial categorisation/classification**. We will also demonstrate that the DID system **collects, temporarily stores, and processes the biometric data of passersby** without giving them a possibility to opt out or to consent to their data being processed. Moreover, the DID system collects data **which could be used to identify passersby.** Lastly, we will point to two serious issues with the claim that demographic information can be derived from "views:" first, that there is no clear scientific basis for deriving information about emotions from facial expressions and that **such inferences about emotion are therefore invalid**; second, that the automated gender recognition technology used to predict the gender of passersby **systematically discriminates against trans and non-binary individuals**.

A. The AdMobilize DID System uses facial categorisation/classification, not simply facial detection

We will begin with the claim that AdMobilize's DID system "uses facial detection as opposed to facial recognition" (fls.1791 dos autos). This claim is a misrepresentation of what AdMobilize actually does in accordance with the company's own description of their activities. In reality, AdMobilize's software **first** performs facial detection, and **then** processes the biometric data (here, facial data) in a process that is technically termed **facial analysis**, **classification**, **or categorization** (these terms refer to the same technical process).

The misrepresentation here hinges on the definition of a number of terms, such as facial detection, facial recognition, and biometric data, used by AdMobilize, ViaQuatro, the Instituto Brasileiro de Peritos and Raul Spiguel, which we will show do not conform to standard definitions. For example, **facial recognition technology is typically used as an umbrella term which encompasses a range of technological processes, including the process of**

facial detection. Although there may be some variation as to how these different processes are named, there is general agreement in describing what these processes are.

In order to establish clarity about the processes involved in this technology, we will begin with a number of definitions related to biometric technologies from the European Union's Article 29 Data Protection Working Party document, *WP193 Opinion 3/2012 on developments in biometric technologies* (hereafter, Working Party Opinion).[5] Once these definitions have been established, we will look at the specifics of the DID system and demonstrate how the data collection and processing activities involved are mischaracterized by the defendant and the defendant's experts. We will reference the Working Party Opinion and the European Union's General Data Protection Regulation (hereafter GDPR)[6] because they have been quoted by the parties in this case and by AdMobilize. The GDPR also served as an inspiration for the Brazilian Data Protection Law, which explains the similarities in both laws. Moreover, at multiple points the defendant and the defendant's experts have claimed that the technology used in the DID system is GDPR compliant (example: fls. 1758 dos autos). Thus, it is important to test this claim against the relevant definitions contained in the European law.

What is biometric data?

The Working Party Opinion states that biometric data may be defined as:

> biological properties, behavioural aspects, **physiological characteristics**, living traits or repeatable actions where those features and/or actions are both unique to that individual and measurable, even if the patterns used in practice to technically measure them involve a certain degree of probability.[7]

Since the DID system captures and processes images of the faces of passersby (and thereby makes inferences about their age, gender, and emotion), the system captures and processes biometric data relating to the **physiological characteristics** of individuals.

The Working Party Opinion further notes that biometric data "make[s] the characteristics of the human body 'machine-readable' and subject to further use."[8] In the case of the DID system, the technology makes the facial characteristics of passersby machine readable and subject to further use by categorising them according to age, gender, and emotion.

What are biometric systems?

The Working Party Opinion defines "biometric systems" as "applications that use biometric technologies, which allow the automatic identification, and/or authentication/verification of a

[5] Article 29 Data Protection Working Party document, WP193 Opinion 3/2012 on developments in biometric technologies https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2012/wp193_en.pdf
[6] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN
[7] Working Party Opinion WP193, p.3-4
[8] Ibid, p.4

person."[9] They add, however, that "[d]ue to the recent technological developments it is **now also possible to use biometric systems for categorisation/segregation purposes**," and note that "a more general definition would be a system that extracts and further processes biometric data."[10] Although AdMobilize claims that their DID system does not identify individuals, it clearly **categorizes or segregates** individuals according to age, gender, and emotion, thus we can unambiguously state that their system falls under this definition of a biometric system.

What are the stages of biometric data processing?

The Working Party identifies three typical stages involved in processing biometric data in a biometric system: biometric **enrolment**, biometric **storage**, and biometric **matching**. Biometric enrolment "[e]ncompasses all the processes that are carried out within a biometric system in order to extract biometric data from a biometric source and link this data to an individual."[11] The Opinion adds that while such enrolment can involve asking the individual for consent to collect their biometric information, "it is also possible to enrol individuals without their knowledge or consent (e.g. CCTV systems with embedded facial recognition functionality)."[12] This latter situation of enrolment without direct consent clearly describes AdMobilize's DID technology as used by Viaquatro in the São Paulo metro line, which collects metro customers' biometric information without their knowledge, and without giving them an opportunity to opt out or consent to the processing of their biometric data. As noted in the Parecer elaborado pelo Instituto de Referência em Internet e Sociedade - IRIS:

> "No momento em que são detectadas a presença e a reação emocional dos usuários por aparelhos que não exigem consentimento nem possibilitam opt out, os usuários estão tendo essa liberdade violada. Ademais, ao não publicizar que esta coleta é feita, a concessionária incorre em negligência quanto ao dever de informar" (fls. 1657 dos autos).

Once the data has been collected in the enrollment stage, it must now be stored. Regarding the storage of biometric data, the Working Party notes that there are two possibilities for storage. First, the data can be stored in a raw form, which "allows recognising the source it comes from without special knowledge e.g. the photograph of a face, the photograph of a fingerprint or a voice recording."[13] Second, the data can also be stored in a modified form as a biometric template, where "the captured raw biometric information is processed in a way that only certain characteristics and/or features are extracted and saved as a biometric template."[14]

An important consideration here is how much information is stored in a biometric template. The Working Party notes that there is a tradeoff between how much information the template

[9] Ibid, p.5
[10] Ibid
[11] Ibid
[12] Ibid, p. 5
[13] Ibid, p.4.
[14] Ibid

contains, and thus how useful it is for further processing and analysis, and how secure the template is against efforts at reconstructing the raw data.[15] The danger here is that the more useful and information rich a template is, the higher the risk is of someone being able to reconstruct the original, and typically sensitive, raw data:

> The definition of the size (the quantity of information) of the template is a crucial issue. On the one hand, the size of the template should be wide enough to manage security (avoiding overlaps between different biometric data, or identity substitutions), on the other hand, the size of the template should not be too large so as to avoid the risks of biometric data reconstruction.[16]

As noted by IRIS in their opinion (fls.1643 dos autos), the defendant and the defendant's experts provide insufficient information regarding how any such template used by the DID system is constructed so as to avoid the risk of re-identification at a later stage, especially if combined with other sources of data, such as the date and time that the image was captured, or other information contained in the metro system, including the Single Ticket database.

Regarding where the data is stored, the Working Party notes that "data obtained during enrolment can be stored locally in the operations centre where the enrolment took place (e.g. in a reader [or, as the case at hand, in the camera system before transfer to the server]) for later use, or on a device carried by the individual (e.g. on a smart card) or could be sent and stored in a centralised database accessible by one or more biometric systems."[17]

The final stage in the biometric data processing is **biometric matching**, which the Working Party defines as "the process of comparing biometric data/template (captured during enrolment) to the biometric data/template collected from a new sample for the purpose of identification, verification/authentication or categorisation."[18] AdMobilize claims not to engage in identification (fls. 1791 dos autos), but they fail to mention that **they do engage in biometric (and in this case, facial) categorisation**, which the Working Party defines as follows:

> The categorisation/segregation of an individual by a biometric system is typically the process of establishing whether the biometric data of an individual belongs to a group with some predefined characteristic in order to take a specific action. In this case, it is not important to identify or verify the individual but to assign him/her automatically to a certain category. For instance **an advertising display may show different adverts depending on the individual that is looking at it based on the age or gender.**[19]

Thus, by making inferences about metro customers' age, gender, and emotion, AdMobilize's DID System engages in biometric categorisation of passersby, which qualifies as processing of

[15] Ibid.
[16] Ibid.
[17] Ibid. p.5
[18] Ibid
[19] Ibid. 6

their personal information. Let us now look more closely at the specific cases of facial detection, facial recognition, and their associated technologies.

The DID system does not just perform facial detection

The National Institute of Standards and Technology (NIST), a non-regulatory agency of the United States Department of Commerce, defines facial detection as a technology that "determines whether the image contains a face."[20] Similarly, the Ada Lovelace Institute, an independent research and deliberative body in the United Kingdom that specialises in research on AI and data, define facial detection as "identifying that there is a face in an image or series of images, and where it is located in the image(s)," and add that facial detection "**is usually the first step of a facial recognition process**, to enable matching, identification or classification on only the relevant parts of an image."[21]

Facial detection therefore does not go beyond answering two questions: firstly, is there a face in this image?; secondly, where is the face located in this image? The primary purpose of facial detection is to isolate the facial data within an image (this is the biometric enrolment phase described above) to prepare it for further processing (in this case, **storing** it as **raw data or as** a **biometric template** so that in the process of **biometric matching** it can be analyzed to make inferences about age, gender, and emotion).

According to AdMobilize's documentation (fls. 1787 -1815 dos autos), ViaQuatro's arguments, the Instituto Brasileiro de Peritos's reports (fls. 427 - 450 and 1816 - 1823 dos autos) and the expert Raul Spiguel' report (fls.1753 - 1815 dos autos), AdMobilize do more than simply determine if there are faces present in the images captured by the DIDsystem, and where those faces are located. Instead, AdMobilize states that they want to 'extract' or 'derive' demographic information about age, gender, and emotion from those images, this being the primary output of their software (fls.1793 and 1796 dos autos). This process of deriving such information from an analysis of facial images is referred to by a number of different terms. NIST, for example, refers to this as **face analysis**, which "aims to identify attributes such as gender, age, or emotion from detected faces."[22] The Ada Lovelace Institute uses the term **facial classification**, which they define as "identifying a characteristic of a face, such as age, gender or expression." Similarly, in their focus paper on Facial Recognition Technology, the European Union's Fundamental Rights Agency (FRA) use the term 'facial recognition' as a general umbrella term, and define it as "automatic processing of digital images which contain the faces of individuals for identification, authentication/verification or categorisation of those individuals."[23] Following the Working Party

[20] NIST Testimony. (2020) Facial Recognition Technology (FRT): https://www.nist.gov/speech-testimony/facial-recognition-technology-frt-0

[21] Ada Lovelace Institute. (2019) Facial recognition: defining terms to clarify challenges: https://www.adalovelaceinstitute.org/facial-recognition-defining-terms-to-clarify-challenges/

[22] NIST Testimony. (2020) Facial Recognition Technology (FRT): https://www.nist.gov/speech-testimony/facial-recognition-technology-frt-0

[23] They take this definition from Article 29 Data Protection Working Party (2012), Opinion 02/2012 on facial recognition in online and mobile services, 00727/12/EN, WP 192, Brussels, 22 March 2012, p. 2

Opinion, the FRA use the term **facial categorisation** (or **face analysis**) to refer to a subtype of facial recognition that matches precisely the type of 'extraction' or 'derivation' of demographic information which AdMobilize carries out:

> Apart from verification and identification, **facial recognition technology is also used to extract information about an individual's characteristics.** This is sometimes referred to as 'face analysis'. It can, therefore, also be used for profiling individuals, which involves categorising individuals based on their personal characteristics. Characteristics commonly predicted from facial images are sex, age and ethnic origin.[24]

As we can see from these examples, **AdMobilize's claim that they do not perform facial recognition is misleading**. According to accepted terminology, the facial detection and the facial categorisation/classification which they perform **are typically classified as sub-categories of facial recognition technology**. Therefore contrary to their caim, **AdMobilize's system does use a form facial recognition, namely facial categorisation/classification**. Moreover, inferring personal information about a person amounts to data processing and requires the application of safeguards established under data protection principles recognised by the Council of Europe Convention 108, which Brazil signed and ratified,[25] as well as most local and regional data protection laws. This includes ensuring that the processing is authorised by law, limited in time, conducted for a specific purpose and limited in scope to the only the data necessary for the defined purpose.

B. AdMobilize's DID system collects, saves, and stores unique and identifiable biometric data of metro customers.

The mischaracterisation of AdMobililze's data processing activities leads us to the next erroneous claim that the company does not collect, store, or save any unique biometric information about individuals because they collect 'anonymous data'. Moreover, AdMobilize states that they a) ***cannot*** and b) ***do not*** **personally identify individuals** at any stage of the process. Let us look at these claims individually.

AdMobilize collects and stores unique biometric information

Firstly, it is manifestly false for AdMobilize to claim that they do not collect unique biometric information about individuals. According to the technical description of the DID system provided by ViaQuatro's documents, AdMobilize (fls. 1787 -1815 dos autos), the Instituto Brasileiro de Peritos (fls. 427 - 450 and fls.1816 - 1823 dos autos), and the expert Raul Spiguel (fls.1753 - 1815 dos autos), there are three significant stages in the processing of the data: firstly, the

[24] European Union Agency for Fundamental Rights (FRA). (2019). Facial recognition technology: fundamental rights considerations in the context of law enforcement: https://fra.europa.eu/en/publication/2019/facial-recognition-technology-fundamental-rights-considerations-context-law p.8

[25] Council of Europe. Data Protection. National information: Brazil https://www.coe.int/en/web/data-protection/brazil

capture of images of passersby by the camera; secondly, the processing of this data from the cameras by the MATRIX program; and thirdly, the transfer of the data generated by the MATRIX program to the server. In his expert testimony, Mr. Spiguel notes that “O programa MATRIX é responsável pela integração câmera e software” (fls. 1762 dos autos). The MATRIX program therefore takes the data collected by the camera, processes it, and sends newly generated data to the server where it becomes available on the ‘dashboard.’

In order to perform facial recognition in any of its forms, the system must initially collect and process data on people’s faces via the cameras. This data, according to the Working Party Opinion’s definition cited above, represents unique biometric information about individuals. Furthermore, the system must store/save this data from the camera, if only temporarily, for the purpose of analysing it in the process of facial detection and analysis/categorisation. Thus, **there is no way to perform facial detection and facial categorisation/classification without the collection and processing of personal data**. These personal data of people’s faces, as mentioned above, may be either stored as **raw data**, or as **biometric templates**. In his expert testimony, Mr. Spiguel notes that the MATRIX program receives real-time images from the camera and captures specific points on the faces detected in order to perform the facial analysis to predict gender, age, and emotion. He claims that the Matrix program “4. não armazena a imagem tampouco os ponto capturados, sendo os mesmos descartados da memória após finalizar os cálculos acima descritos” (fls.1774 dos autos). However, this only confirms our point: the captured points of human faces and the images are **discarded *after* finishing the processing**, meaning that they **are still stored for the duration of the processing**. Moreover, **these captured points are biometric data.** Even without considering the suspicion that the images or captured points are sent to the server, it is beyond doubt that the images are captured by the camera in the first stage and stored temporarily while they are processed by the MATRIX program. This capture, storage, and processing involve sensitive biometric data of passersby, i.e. data derived from their faces, and they were not given an opportunity to opt out and have not consented to the collection of this data and could not consent in the first place.

In the documentation provided by ViaQuatro, there are a number of confusing and contradictory claims about how AdMobilize stores the facial data which it processes during facial analysis to determine age, gender, and emotion. In the *FAQs* document, attached by Mr. Spiguel, AdMobilize makes the following claim: “Not [sic] storing and transmitting of images is done at any point, only **raw,[26] anonymous and aggregated data is accessible** to either the data processor and the data controller. This makes the solution anonymous by design. We never collect private info in the first place” (fls.1798 dos autos). As we have seen above, this is not true: **the data collected in the first stage of biometric enrolment is personal data**. Further confusion is introduced by the Instituto Brasileiro de Peritos, in its parecer técnico concordante with Mr. Spiguel’s expert document, when they state that “a perícia confirmou que (...) Assim, i r. sistema opera com dados estatísticos, que nascem originalmente anônimos; (...)” (fls. 1817 dos autos) and that “O r. sistema não coleta ou grava dados biométricos ou quaisquer outros que

[26] The raw data contains the biometric data in its most granular form.

permitam identificar indivíduos, mesmo que temporariamente" (fls. 1817 dos autos). When the Instituto Brasileiro de Peritos speaks of statistical data which were born anonymous, they are speaking of the data generated by the MATRIX program and sent to the server in the third stage of processing. However, the MATRIX program is only able to generate this anonymous data by processing the unique biometric data captured by the cameras in the first stage. As such, it is disingenuous to claim that the system as a whole, including the cameras and initial processing activities, does not collect or record any biometric data that would allow individuals to be identified.

AdMobilize can identify individuals

The claim that AdMobilize's technology can '**at no point' be used to track or identify individuals** (fls.1792 dos autos) is incorrect or naive at best because in the first stage of biometric enrolment, the technology captures images of people's faces at a particular time and location. This information is more than sufficient to identify people. AdMobilize provides us with no information to assuage doubts that sufficient anonymisation efforts are undertaken thereafter.

As the Working Party Opinion points out, creating a biometric template involves a tradeoff between containing enough information to be useful for the processing activity (in this case, determining of the age, gender, and emotion of a person), and making sure there is not too much information to allow for identification of the person. AdMobilize, ViaQuatro, the Instituto Brasileiro de Peritos, and the expert Raul Spiguel provide no information regarding this tradeoff, and instead make several contradictory statements, as quoted before, which betray either a misunderstanding of the technicalities involved, or a misrepresentation of their activities.

Thus, despite its statements to the contrary, AdMobilize's DID system does more than just facial detection. The system also collects sensitive biometric information, stores that information at least temporarily for the purposes of processing, processes the inferred personal information, and, therefore could, at multiple points in the process, identify individuals. The DID system collects all the data necessary to identify passersby, and could easily be repurposed to perform facial identification.

Furthermore, as noted by IRIS in their expert opinion, even this so-called anonymous data generated by the MATRIX program could be combined with other data and metadata to potentially allow for re-identification of the individuals.

C. Emotion detection has no clear scientifc basis and making inferences about the gender of passersby harms trans and non-binary people

We have shown in the previous section that the DID system does in fact collect, store, and process sensitive biometric data about individuals. Moreover, it does so without providing any

opportunity for the individuals in question to opt out or to give or deny their consent. Below, we identify additional concerns that further undermine the credibility and legitimacy of the DID system as they establish that there are serious flaws in the ultimate aim for which this collection, storage, and processing of the data occurs.

Regarding the processes of facial analysis/categorisation, according to AdMobilize's documentation provided by Raul Spiguel, the information about age, gender, and emotion is **extracted** from 'views.' According to the documentation in the case file, an **impression "**is a face detected (passed by)," and a **view** is a face looking at the camera which is deemed to be 'engaged' with the advertisement. AdMobilize also states that a view "occurs when a face detected, is also detected as "viewing" an advertisement within +/- 5 degrees of the position of the Sensor and with an attention time greater than 0.5 seconds" (fls.1793 dos autos). They then state that "[a]ll the demographic data such as: gender, age and emotions is **derived** from the VIEWS" (fls. 1793 dos autos). As we will show in below, it is **not technically possible to simply derive or extract this information** from such views. Using the language of extraction and derivation suggests that the information is simply there to be collected. Rather, information about emotion and gender (we focus on these two claims) can be **inferred** from the biometric data collected in the stage of biometric enrolment (and which results in so-called views). However, we will show that there is **no clear scientific basis for such inferences about emotion**, and that making such inferences about the gender of passersby is based on a **flawed and harmful physiognomic conception of gender that leads to a number of harms for trans and non-binary people**.

1. It is not technically possible to derive/extract information about emotions from the data collected by the DID system

AdMobilize claims to be able to 'detect' three demographic data points about the people whose image the camera captures: their age (classified in 4 groups: young, young adult, adult, senior), their gender (male or female), and their so-called 'emotion' (Angry, Happy, Calm, Sad, Neutral, Surprised, Disgust). We will begin by discussing the claim about emotion detection, because it is the most misleading and potentially harmful of the three and the most problematic from a scientific point of view.

According to Admobilize's documentation provided by ViaQuatro, Mr. Spiguel and the Instituto Brasileiro de Peritos, the software can detect the emotion that somebody is feeling by analysing the image of that person captured by the camera to determine whether they are feeling one of 7 'emotions': Angry, Happy, Calm, Sad, Neutral, Surprised, Disgust (fls.1794 dos autos). This claim is both scientifically and technically flawed.

As Andrew McStay, Professor of Digital Life at Bangor University, UK, and author of the book Emotional AI - The Rise of Empathic Media,[27] points out in his article ''This time with feeling?'

[27] McStay, Andrew. (2018). Emotional AI. https://uk.sagepub.com/en-gb/eur/emotional-ai/book251642

Assessing EU data governance implications of out of home appraisal based emotional AI,' there are two possible approaches to 'detecting' emotion. On the one hand, there are approaches which use some form of the 'basic emotions' theory to make inferences about people's emotions from an analysis of video footage, as in the case of AdMobilize. On the other, there is the so-called appraisal-based method which "entails use of data about internal physiological and experiential contexts, but also factors external to an individual."[28]

AdMobilize clearly falls into the first category: they are using 'neural networks,' a form of machine learning algorithm, to analyse biometric data (in this case, data about the faces of passersby) to produce predictions about age, gender, and emotion for the people whose image is captured on camera. **It is important to clarify here that the technology in question cannot 'perceive' or 'detect' emotion**. Instead, it first detects whether there is a face in the image. It then stores the biometric data about that face. The technology can then detect **facial movements** or **facial configurations** in this biometric data of people's faces. It next makes a **reverse inference** based on this information to **infer the emotion that person is feeling**. Based on the review of the currently available scientific literature, there is no clear scientific basis for such an inference.

In a recent article entitled 'Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements,'[29] a number of the world's top researchers in the science of human emotion, including Lisa Feldman Barrett, a University Distinguished Professor of Psychology at Northeastern University, investigated the evidence for what they call the 'common view' (or 'basic emotions' view) - namely, the claim "that a person's emotional state can be readily inferred from his or her facial movements, typically called emotional expressions or facial expressions."[30] As we can see from the documentation provided by AdMobilize, ViaQuatro, the Instituto Brasileiro de Peritos, and the expert Raul Spiguel, this is the view which underlies the operation of the DID system: they claim that the system can identify seven emotions from an analysis of video footage.

Barrett et al. make specific mention of the real-world applications of the common view by stating that "[t]echnology companies are investing tremendous resources to figure out how to objectively "read" emotions in people by detecting their presumed facial expressions, such as scowling faces, frowning faces, and smiling faces, in an automated fashion," and that "[s]everal companies claim to have already done it." We can surely count AdMobilize among these companies who claim to be able to 'read emotion' in people using automated measures. Despite

[28] McStay, Andrew & Urquhart, Lachlan. (2019). 'This time with feeling?' Assessing EU data governance implications of out of home appraisal based emotional AI. First Monday. 24. 10.5210/fm.v24i10.9457, p.1
[29] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological Science in the Public Interest, 20(1), 1–68. https://doi.org/10.1177/1529100619832930
[30] Ibid, p.1

this enthusiasm from technology companies, however, the conclusion of Barrett et al. is that "the science of emotion is ill-equipped to support any of these initiatives."[31]

Barrett et al. provide "a systematic review of the evidence, testing the common view"[32] The conclusions of this review of the scientific evidence are that there is no scientific basis for the common view, and they point to three specific problems:

1. **Limited reliability:** i.e., instances of the same emotion category are neither reliably expressed through nor perceived from a common set of facial movements.
2. **Lack of specificity:** i.e., there is no unique mapping between a configuration of facial movements and instances of an emotion category.
3. **Limited generalizability:** i.e., the effects of context and culture have not been sufficiently documented and accounted for.[33]

Regarding the first two problems, **limited reliability and specificity**, what is at issue is that a given emotion, such as anger, is not reliably expressed or perceived through a particular facial movement, such as a frown, and facial movements do not map uniquely to emotions (e.g. smile-happy, scowl-angry). In the case of AdMobilize's so-called emotion detection software, this means that they cannot reliably infer that a person is feeling anger or happiness from an analysis of their facial movements. Indeed, Barrett et al. conclude that while anger may sometimes involve scowling, "a scowling facial configuration is not the expression of anger in any generalizable or universal way (there appear to be no prototypical facial expressions of emotions)."[34]

Relating this to the claims made by AdMobilize, we see that the DID system can only detect **facial configurations/movements**, and cannot, as the company claims, 'detect' emotion or 'extract' emotion from the biometric data collected. **There is no clear scientific basis to infer that somebody is feeling happy because they are smiling, or angry because they are scowling; people often smile for other reasons than because they are happy, or express happiness by other facial configurations than a smile. AdMobilize does not take these fundamental scientific considerations into account.** As such, AdMobilize is misrepresenting its technology on two fronts: on the one hand, it is making invalid inferences about the private emotional life of passersby, which has no scientific basis and is done without giving them an opportunity to opt out and without their consent; on the other, it is deceiving advertisers by claiming to detect the emotions of passersby and to provide the advertisers with insights about these emotions.

[31] Ibid, p. 48
[32] Ibid, p.3
[33] Ibid
[34] Ibid

Regarding the criticism of **limited generalisability**, the problem is that the approach taken by AdMobilize, which is based on the common view/basic emotions view, does not and cannot take into account context and culture. Proponents of the common view claim that these facial configurations/movements are prototypes for emotional expression with *universal validity*. By contrast, Barrett et al. have demonstrated that these configurations are "best thought of as **Western gestures, symbols or stereotypes** that fail to capture the rich variety with which people spontaneously move their faces to express emotions in everyday life," adding further that a "stereotype is not a prototype [...] because a prototype is the most frequent or typical instance of a category (Murphy, 2002), whereas a stereotype is an oversimplified belief that is taken as generally more applicable than it actually is."[35] Rather than 'detecting emotion,' AdMobilize is actually making invalid inferences based on oversimplified beliefs.

In this particular case it is also important to understand the Brazilian society. Brazil is a diverse country, with individuals who came from Africa, Asia, Europe, and other parts of the world, and then mixed with the native population. Thus, the context and culture of the Brazilian population may differ significantly from those of the populations in the United States or other Western countries where this technology was developed. In that sense, the information that is used to infer individuals' emotion in Western countries is not necessarily accurate for categorizing individual's emotions in Brazil.

The fact that these inferences are invalid further **undermines the justification for the use of this system**. The DID system is collecting and processing sensitive biometric information of passersby without giving them an opportunity to opt out and without their consent to ultimately make invalid inferences about their emotional state. Moreover, **just because these inferences are invalid it does not make them less invasive**. The fact remains that **people's emotional life is subject to surveillance** and that decisions are made **based on flawed assumptions about people's emotional states**.

ViaQuatro acknowledged inaccuracy of the DID system

AdMobilize claims that it can detect emotions with an accuracy of 80% (fls. 1793 dos autos), however, this claim has clear scientific basis, as the very idea of inferring emotion from an automated analysis of facial configuration/movement is flawed. Bearing in mind this lack of scientific basis, it is a surprise for us to find this statement of ViaQuatro in the first answer to the plaintiff claim:

> "Inclusive, todas as estimativas geradas a partir dos dados estatísticos tratados são de baixa precisão, o que ratifica a ausência de utilização de dados que possam tornar uma pessoa identificada ou identificável" (fls 369 dos autos).

According to this statement, **ViaQuatro has acknowledged, since the beginning, that the technology used in the DID system does not function at a high accuracy. Nevertheless,**

[35] Ibid, p.46

ViaQuatro decided to procure the AdMobilize technology anyway by signing a contract with AdMobilize.

Moreover, ViaQuatro excuses itself by saying that the low precision ratifies the lack of use of data that can make a person identified or identifiable. However, first, as mentioned above, all the processing done using the AdMobilize DID system constitutes biometric data processing. Second, **even if the processing of personal data gives a poor result, it does not imply that there was no data processing**. There was biometric data processing from the moment they collected the data and fed it into the MATRIX program.

Processing of biometric data is a sensitive decision that impacts all metro users. Unlike the metro users, ViaQuatro knew that the DID system was inaccurate. However, **ViaQuatro was being disingenuous when talking in public about the capabilities of the AdMobilize system and the benefits it can provide in return (fls 152, 155, 158 dos autos).** Therefore, not only were the users unable to opt out, but **they were also misinformed in order to have their faces detected and analyzed in an experiment that did not have any clear scientific basis.**

2. The discriminatory impact of automated gender recognition (AGR) technology

AdMobilize's DID system also claims to be able to detect the gender of passersby. They claim to have an accuracy of 80-90%, meaning that they claim to be able to determine whether a passerby is male or female with an accuracy of 80-90%. Our concern here relates to the fact that AdMobilize seems to consider gender as a binary, with only male and female as possibilities. Given that they claim to detect gender by analysing the faces of passersby, we must assume that they believe that gender can be determined from the physiological characteristics of a person's face. This is **a flawed assumption, which not only fails to account for the existence of non-binary people and trans people, but in fact perpetuates discrimination and harms individuals who do not conform to this binary and physiologically based conception of gender.**

In an article entitled 'The Misgendering Machines,[36]'Os Keyes, a researcher in human-computer interaction and AI, discusses how the binary conception of gender and automated gender recognition (AGR) technologies which operationalise it, fail to account for the existence of trans and non-binary individuals and perpetuate harms and discriminations against them. Regarding the binary conception of gender, Keyes points out that this rests on a misunderstanding of the distinction between **sex** and **gender**:

[36] Keyes, Os. (2018). The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human-Computer Interaction. 2. 1-22. 10.1145/3274357.

> Traditionally, Western culture has alternately conflated and drawn arbitrary distinctions between two constructs: sex, a person's biological category (male or female) based on anatomy, chromosomesand hormones, and gender, a person's cultural category (man or woman), based on their behaviourand social role. The latter is seen to derive from the former: a person's gender is an inevitable consequence of their sex.[37]

Keyes points out that both social scientists and biologists have shown this conception to be inaccurate because, on the one hand, there is a large number of intersex conditions that do not fit this binary, and on the other, "the assumption that sex dictates gender—in other words, that it mandates social roles, combinations of behaviours and traits and aspects of presentation and identity—fails to capture the existence of transgender (trans) people, whose genders do not match their assigned sex."

In contrast to this binary view of gender, Keyes mentions trans-inclusive views which "generally agree that gender is not immutable, binary or tied inherently to physiology." Regarding the harms that come from excluding trans and non-binary people, Keyes notes that this binary conception of gender "creates space for widespread explicit discrimination, which has (e.g., in the United States) resulted in widespread employment, housing and criminal justice inequalities, increased vulnerability to intimate partner abuse and, particularly for trans people of colour, increased vulnerability to potentially fatal state violence."[38]

The technology used by AdMobilize and used by ViaQuatro, in claiming to 'detect gender', clearly operationalises this binary conception of gender, and thus falls under what Keyes refers to as Automatic Gender Recognition (AGR):

> AGR purports to allow the automatic, computational identification of a person's gender from photographs or videos. Implementations first isolate the person within a photograph: some use geometric structure, while others rely on skin texture, and yet others depend on 3D modelling. The resulting image can then be subject to "gender recognition" which—while experimentally undertaken using gait or overall body shape—is usually based on the person's face.[39]

Of particular relevance to AdMobilize, Keyes mentions "gendered advertising and user interfaces, in which adverts or applications could (upon detecting a particular person's gender) alter their presentation to be more appealing to stereotypical members of that gender."[40]

Keyes undertook a comprehensive study of AGR research, and determined **that "AGR is particularly likely to misclassify (and so discriminate against) trans people.**" Furthermore, they state that:

---

[37] Ibid, p. 2
[38] Ibid, p. 3
[39] Ibid, p. 4
[40] Ibid

> [the] presumption that gender is physiologically-rooted cuts against trans people overall [...] by essentialising the body as the source of gender. **The presumption that gender is a binary additionally harms non-binary people, who by definition cannot be accurately classified.** Both of these things are a problem when the technology is integrated with binary, gendered spaces.[41]

Again, on the harms of AGR in advertising, Keyes says the following:

> AGR papers propose billboards with gender recognition technology that, after evaluating a pedestrian walking nearby, "may choose to show ads of cars when a male is detected, or dresses in the case of females." **This may seem relatively innocuous, but a trans man who sees a billboard flicker to advertise dresses to him as he approaches is, even if he likes dresses, unlikely to feel particularly good about it.** For the billboard to have done that under the design specifications quoted above, it must have concluded he was a woman. Other papers experiment with recommender systems that present a range of different items to purchase, which offers the opportunity for even more explicit (and jarring) automated misgendering given the vast number of products labeled "for men" and "for women" in the world.[42]

AGR systems, such as the one used by AdMobilise, either fail to classify trans and non-binary people as either male or female (they are thus excluded), or it misgenders them, by assigning them a gender which does not match what they themselves have chosen as their gender. According to Keyes, although trans and non-binary people may represent a small segment of the population, this inaccuracy consitutes discrimination against them because it falls disproportionately on a minority.

This factor of discrimination is quite important in Brazil. According to the annual study of the Associacão Nacional de travestis e transexuais do Brasil and the Instituto Brasileiro trans de educacão, 124 trans people were killed in 2019, 80% of them in a context of extreme violence.[43] São Paulo was the state were most of the killings happened[44]. Moreover, according to the LGBTQ+ Danger Index, Brazil receives a qualification of -B in the ranking of safe countries for the LGBTQ+ population.[45] This information shows how difficult is the daily living for trans and non-binary people, which **calls for a quick response from the government and from any company that provides a public service to ensure a safer city for the population.** This response should include education and a rebuttal of misconceptions such as the binary approach to gender, similar to the one we see in **ViaQuatro's behavior when using the**

---

[41] Ibid, p. 11
[42] Ibid, p. 12 (bolding ours)
[43] Benevides, Bruna G. and Sayonara Naider Bonfim Nogueira. (2019). Dossiê dos assassinatos e da violência contra travestis e transexuais brasileiras em 2019. https://antrabrasil.files.wordpress.com/2020/01/dossic3aa-dos-assassinatos-e-da-violc3aancia-contra-pessoas-trans-em-2019.pdf
[44] Keyes, The Misgendering Machines, p.12
[45] Asher and Lyric. (2019). The worst (& safest) countries for LGBTQ+ travel. https://www.asherfergusson.com/lgbtq-travel-safety/

AdMobilize DID system that denies this diversity and dignity of the passersby who have the right to self identify and not to be misclassified.

Keyes unambiguously states that it is impossible to use AGR technology without harming trans and non-binary people because it is fundamentally based upon denying them the opportunity to choose and express their own gender:

> Whatever approach (physiology, clothing, hair length...) an AGR system takes for discriminating between genders, however many trans people the dataset includes, the technology is fundamentally premised on the idea that gender is something assigned. Yet to be trans—to be of a gender that runs contrariwise to that which society assumed of you—means to stand as testament to the idea that it is self-knowledge, not external assignation,that has primacy in defining gender. Put simply, a trans-inclusive system for non-consensually defining someone's gender is a contradiction in terms[46]

Indeed, Keyes goes so far as to say that there is no way to make this technology hospitable to trans and non-binary people and that, as such, **AGR technology should simply not be used**. Given the unavoidable inaccuracy and harm caused by such systems, it is inappropriate for the DID system, which is based on AGR technology, to be used in a public space such as the metro, where trans and non-binary people will be highly likely to be misgendered without their consent and with no possibility of redress. The insights which AdMobilize provides based on determining the gender of passersby will be tainted by these ineradicable inaccuracies.

Instead of using a technology that harms people's rights to self determination, ViaQuatro as a company that offers a public service should advance the rights of the trans and non-binary individuals. Advancing rights is key to addressing the violence, segregation, and discrimination against this population. In fact, this is supported by the Supremo Tribunal Federal of Brazil's unanimous decision to declare unconstitutional the law 1.516/15[47] that aimed to prohibit the teaching of non-binary conceptions of gender in schools,[48] as well as by the Inter-American Court of Human Rights in its Parecer Consultivo OC-24/2017,[49] which states:"i)O reconhecimento da identidade de gênero pelo Estado é de vital importância para garantir o pleno gozo dos direitos humanos das pessoas transgênero, incluindo a proteção contra a violência, a tortura, os maus-tratos, o direito à saúde, à educação, ao emprego, à moradia, ao acesso à seguridade social, bem como o direito à liberdade de expressão e associação (...)" (par. 101). Therefore, it is important that in every context, including in public transportation, the

[46] Keyes, The Misgendering Machines, p. 13

[47] Supremo Tribunal Federal. Arguição de descumprimento de preceito fundamental 457. http://portal.stf.jus.br/processos/detalhe.asp?incidente=5192888

[48] According to the judges, they declared material unconstitutionality because the Law 1.516/15 violated the principle of equality (Article 5 of the Federal Constitution), the rights to not to be discriminated against, freedom of expression, to have plural ideas, freedom of learning, researching, and more. More information available in the relatório of the judges: http://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/ADPF457.pdf

[49] Corte Interamericana de Direitos Humanos. Parecer Consultivo OC-24/2017 http://www.corteidh.or.cr/docs/opiniones/seriea_24_por.pdf

idea of a binary gender should not be forcibly applied to passersby who have no opportunity to opt out of this misgendering technology.

The additional concerns outlined here in Section C demonstrate that there is no valid purpose for which the DID system collects and processes the biometric data. The inferences about emotion made by the system are not scientifically valid. ViaQuatro acknowledging this flaw used the DID system to perform invasive surveillance and judgment about the private emotional life of the passersby, who did not receive clear information about the processing of their biometric data and did not have any possibility to give consent. Moreover, the inferences the DID system makes about gender cause harm to the dignity of trans and non-binary people and undermine especially their right to self-determination.

IV. Conclusion

In the expert testimony we have demonstrated that the Digital Interactive Doors system **collects, stores, and processes sensitive biometric data about passersby**. Contrary to the claims made by the defendant and their experts, we demonstrated that the DID system **does perform subtypes of facial recognition**, namely facial detection and facial categorisation/classification. We have further demonstrated that **the technical information provided by ViaQuatro, the Instituto Brasileiro de Peritos and the expert Raul Spiguel constitutes a mischaracterisation of the data processing activities carried out by the DID system**. As such, we cannot accept their technical opinions as valid descriptions of the DID system's functioning. We also reaffirm our agreement with the points raised by the IRIS opinion regarding the abusive practices they report, being: i) there was no clear and adequate information to consumers regarding the operation of Digital Interactive Doors; ii) the passengers' right to choose their data collection was not respected by this system; and iii) the lack of sufficient information regarding the anonymisation of personal data and the possibilities for identification of passersby. All of these points combined clearly demonstrate that the DID system violates the rights of users of the metro line.

We have also highlighted two additional concerns that further undermine the justification for the use of the DID system. As we have shown, the claim that the DID system can detect the emotions of passersby is scientifically unfounded. Instead, the system makes invalid inferences about the private emotional life of passersby, and makes decisions about what advertisement to show them based on these invalid inferences. **Users of the metro line were thus subjected to invasive surveillance and judgement of their inner emotional life, and had no opportunity to opt out of or to deny consent to such processing of their data and did not receive clear information about the system**. Secondly, we have shown that the detection of gender by the DID system systematically undermines the rights of non-binary and trans people who do not conform to the binary conception of gender which underlies its functioning. Rather than 'detecting' gender, this technology **forcibly assigns gender** and thereby undermines the ability of trans and non-binary people to self-determination and impacts their human dignity. These

harms are perpetuated solely for the purpose of serving advertisements to people, which cannot be considered a proportionate aim for such a risk of harm.

As such, the claims made by ViaQuatro and their experts regarding the DID system's data processing activities and the ability of the system to accurately detect gender or emotions are misleading and should not be taken as legitimate technical descriptions of the DID system's functioning.

Verónica Arroyo
Latin America Policy Associate

Daniel Leufer, PhD
2019-20 Mozilla Fellow at Access Now

Access Now
P.O. Box 20429
Greeley Square Station
4 East 27th Street
New York, NY 10001-9998

Curriculum Vitae

Verónica Arroyo is a Policy Associate for Latin America at Access Now based in Lima. Her main focus is privacy and new technologies. Verónica received her Bachelor of Laws from the Pontificia Universidad Católica del Perú ranking fifth higher, and did a course on Human Rights and Globalization in the Universidad Complutense de Madrid. Verónica has professional proficiency in Portuguese. Before joining Access Now, Verónica worked in a Peruvian law firm and was a teacher assistant in Human Rights, Antitrust and Media Law. She was also awarded several internet governance fellowships.

Daniel Leufer is a Mozilla Fellow hosted by Access Now in Brussels, Belgium. He is a philosopher and policy analyst who works at the intersection of technology and politics. His current research focuses on artificial intelligence, with a focus on surveillance and biometric technologies. He has a PhD in Philosophy from KU Leuven in Belgium and has worked on political philosophy, philosophy of technology, and the philosophy and sociology of war. He is also a member of the Working Group on Philosophy of Technology at KU Leuven.