Submission in the consultations of the European Data Protection Board Guidelines 8/2020 on the targeting of social media users

Panoptykon Foundation is a Warsaw-based NGO with a mission to protect fundamental rights in the context of growing surveillance and fast-changing information technologies.

We welcome the EDPB's effort to issue guidelines on the challenging issue of targeting of social media users. In our view, the Guidelines offer much-needed clarifications on the most disputed provisions of the GDPR and they fully encompass and realise data protection principles enshrined in the law. However, we believe there are still some aspects of targeting of social media users that deserve further clarification.

1. Access to user profiles

In paragraph 84 the EDPB mentions that "data subjects should be informed in an easily understandable language if a profile will be built based on their online behaviour on the platform or on the targeter's website". In paragraph 93 it is further clarified that "controllers may want to consider implementing a mechanism for data subjects to check their profile, including details of the information and sources used to develop it".

In our experience and in the experience of other digital rights organisations in Europe, social media providers tend to only reveal parts of the profile that are uncontroversial and are built on the basis of data provided by the user or data observed by the platform in relation to direct interactions with content. This is by no means an exhaustive list of the results of constant behavioural observation and algorithmic analysis that social media providers engage in. Facebook for instance provides users only with parts of the profile that reveal users' interests, while hiding other attributes that advertisers can select in the advertising interface (e.g. those related to particular life events, demographics such as income brackets, or inferred education level)[1]. As documented by extensive research, including facts revealed by whistleblowers in the Cambridge Analytica scandal, **characteristics assigned to users can go way beyond interests, and may be sensitive or reveal users' vulnerabilities**.

Therefore, we recommend that the EDPB clarifies that the profile which social media providers should give users access to (upon request or via a special mechanism) should contain **any and all personal data processed about this person**, including all inferred data, regardless of its category.

---

[1] A. Andreou et al., *Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations*, https://hal.archives-ouvertes.fr/hal-01955309/document

2. Transparency of targeting criteria

The EDPB notes in paragraph 93 that controllers must facilitate access to information regarding the targeting, including the targeting criteria that were used.

Experience with existing mechanisms that offer this kind of explanation to users (e.g. Facebook's "Why am I seeing this ad" tool) shows that they do not provide enough transparency on targeting that would enable data subjects to understand targeting in their particular case. Empirical research[2] demonstrated that these explanations are:

(a) incomplete, as they:

- present only one or two targeting criteria even when a targeter selects more;

- do not explain the social media provider's role in ad delivery, i.e. why the user was qualified to meet the targeter's criteria in the first place (e.g. which data provided by the user or which observed behaviour was taken into account and how it was interpreted by the social media provider);

(b) misleading:

- by presenting only the most common targeting criterion: for example, if an advertiser selects two targeting criteria — interest in "medicine" (potential reach of 668 million users) and interest in "pregnancy" (potential reach of 316 million users) — the user's explanation will only contain "medicine" as the criterion that applies to a higher number of users, even if it was interest in pregnancy that in their case determined the targeting;

- by preferring criteria related to demographics over other types of targeting: researchers found that whenever the advertiser used one demographic-related criterion (e.g. education level, generation, life events, work, relationships), in addition to other criteria (e.g. recent travel, particular hobbies), the demographic-based criterion would be the one presented in the explanation.

Therefore, we propose to amend the second part of this paragraph in the following way (changes in bold):

> *The data subject is entitled to learn of the identity of the targeter, and controllers must facilitate access to information regarding the targeting, including **all targeting criteria used by the targeter and any additional criteria used by the social media provider to deliver the ad**, as well as the other information required by Article 15 GDPR.*

3. Special categories of data and optimisation algorithms

In relation to paragraphs 116-118, we would like to point out challenges related to social media providers' use of algorithms which optimise targeting.

The role of these algorithms is to select a group of users that will eventually see an ad from a larger group of all eligible users who meet requirements selected by targeters. For example, a targeter may specify that they would like to reach users who are over 18 and live in Brussels. Given the size

---

[2] Idem.

of the city, there are probably nearly two millions of eligible users who meet these criteria and it would cost a targeter a lot of money to reach them all. This is why social media providers optimise ad delivery in a way that makes the best use of the available budget and reaches those users that are the most likely to achieve the targeter's goal (e.g. click on the ad). The outcome of optimisation (i.e. which users are eventually reached with an ad) is **determined by how relevant the social media provider deems a particular ad for particular users**. Relevance is calculated by the social media provider's optimisation algorithms which analyse both individual users' previous interactions with ads and statistical correlations resulting from combining data about all users. For a more detailed explanation of this process please refer to Panoptykon Foundation's report "Who (really) targets you. Facebook in Polish election campaigns"[3].

The optimisation process may lead to a situation where **the algorithm recognizes individuals as similar on the basis of their behavioural patterns, and treats them as such, without identifying the sensitive nature of these patterns** and labelling users accordingly, e.g. as "white" and "black" or "liberal" and "conservative".

Concrete examples of this problem were presented by researchers from the Northeastern University in Boston[4] who run a series of experiments on Facebook by indiscriminately targeting a broad category of users with ads related to job offers or housing. They observed that despite their non-discriminate targeting criteria, Facebook delivered ads in a way that discriminated against people on the basis of race and gender, e.g. ads for cleaning jobs were delivered mostly to black women, while lumber job offers were delivered to white men; ads related to hip hop music predominantly reached black people, while ads related to country music – white people. In other words, **race was a distinctive feature in how ads were delivered, even though the targeter did not intend this**.

This problem is inherent to how optimisation algorithms work. If social media providers wanted to protect these sensitive characteristics, they would have to first define them and program the algorithm to – paradoxically - actively search for them in order to exclude them. As such, it poses an immense challenge in terms of interpretation of the GDPR. In our view EDPB's guidance in this regard would be incredibly useful, as it is a key practical challenge both for data controllers and data subjects.

---

[3] https://panoptykon.org/political-ads-report#part-1-4

[4] M. Ali et al., *Discrimination through Optimization: How Facebook's ad delivery can lead to skewed outcomes*, https://arxiv.org/abs/1904.02095