

Individ

Description of system

Purpose

The purpose of the system is to analyze customer flows with a minimum impact on the rights and freedoms of the natural persons being part of such customer flows.

Description of processing

- The system is based on the analysis of pictures from camera surveillance.
- When a person enters the area subject to analysis (typically a shopping center, but could also be individual stores), this is captured by a surveillance camera.
- The system will detect persons on the recording and classify images of people into predefined groups. In some cases, an embedding of the image with added noise will be used instead. The noise is irretrievably blurring the data to a degree where no person can be identified from the data and where the noise will provide an anonymity level at least equivalent to microaggregation into groups of three people.
- Furthermore, the system will make a guess on each person's age (within a 10 year predefined interval, e.g. 18-27) and gender. This guess is based on a neural network, which has been trained through the use of millions of pictures (prior to the system coming into production and therefore unrelated to the live data)
- All theoretically possible combinations of values for the 1,024 measuring points are randomly distributed into 1,000 different groups, using location-sensitive hashing. Using this method, and given the extremely high number of theoretically possible combinations of values, it is extremely unlikely that any characteristics applicable for all members of one of the 1,000 groups (such as age or gender) can be derived.
- To avoid the risk of anyone being identifiable due to being member of a group with very few individuals, noise is added in the form of "fictitious persons". It is not possible subsequently to determine, which "individuals" in the groups are real persons, and which ones are fictitious.
- When a person is placed into a group, the original data will be deleted. This process is completed within milliseconds of the image being received from the camera surveillance. Depending on transmission time over the network, the process will as a general rule take up to a few seconds from the image being recorded by the camera to the data being deleted after completed processing.
- Each time a person's image is captured – for example when entering a specific store or passing other cameras in the area, the same analysis is performed, but since both the picture of a person and the mathematical value calculated is discarded after being analyzed, it is not possible to identify the individual person any further than as a member of one of the predefined 1,000 groups. This will be sufficient to provide a reasonably valid picture of the overall customer flow in the area.
- Furthermore, each time a person's image is captured, the system will also "guess" the age and gender as described above. Based on this, stores will be able to get a high-level picture of the demographics of the total population of customers visiting the store, while it will still not be practically possible to relate this information to the information about the 1,000 groups used for analyzing the overall flow through the area.

Key features

- Personal information is only processed for the purpose of being anonymized (in order for the anonymized data to form the basis for further analysis).
- The processing – including deletion of the original information – takes a few milliseconds (plus the time it takes to transfer images from the camera surveillance system to the system analyzing the pictures).
- Looking at the data stored by the system, no person is identified further than as a member of a specific group. Groups are created in a way that makes it practically impossible to derive any information on the characteristics of the members (such as age, gender, physical appearance, etc.)
- The subsequent analysis of customer flow is not based on how a certain individual moves around the area, but merely on occurrences of presumed members of each of the 1,000 predefined groups. Allegedly, the likelihood of a person being placed into the same group each time the person's image is captured is approximately 50 %.
- Therefore, only the process of anonymization (which lasts a few seconds including time for the image to be transferred through the network) is in scope for GDPR – the remaining part of the process, which analyses customer behaviour is based on anonymous data and therefore out of scope.