# The Protection of Bystanders' Privacy Seminar on Privacy in Ubiquitous Computing

Mehmed Mustafa
*Institute of Computer Science*
*University of Göttingen*
mehmed.mustafa@stud.uni-goettingen.de

Chris Warin
*Institute of Computer Science*
*University of Göttingen*
chris.warin@stud.uni-goettingen.de

***Abstract*—The ubiquitous presence of privacy-threatening devices in nowadays' society is an issue that touches several domains, including economic, social, legal and technological. In this report, we present an overview of the situation regarding the privacy of bystanders, a topic that is getting more attention in research due to rising concerns. We describe the nature of devices that can threaten the visual, audio, and location privacy of bystanders in various situations: they include smartphones, drones, and, more recently, *Internet of Things* (IoT) devices such as smart-home devices. We then discuss over a number of past attempts at solutions that were unfruitful, before coming to our focus by presenting three technologies that aim to protect the privacy of bystanders. We denote two major approaches: one that takes action upon the collection of data (e.g. visual abstraction of bystanders when they are being photographed), one upon the online-sharing of interdependent data. The technologies adopting the former approach rely on user participation: bystanders register a privacy profile that is accessed by the system when it recognizes them. This gives a necessary flexibility to those systems, because of the different requirements regarding privacy across different individuals. The latter approach is perhaps even more crucial, as the sharing of interdependent data without the knowledge or the consent of the implied bystanders can have disastrous consequences. The limitations and challenges of the presented technologies are also discussed; the main challenge being the inclusion of such systems in everyday's life apps and websites, to efficiently start protecting bystanders' privacy.**

***Index Terms*—privacy, bystander, privacy enhancing technology**

## I. Introduction

Today's society is filled with technological devices that are capable of gathering data from people, such as smartphones, surveillance cameras, *Augmented Reality* (AR) devices or IoT devices [1]–[3]. Although these devices have been causing a number of concerns regarding the privacy of users, recent studies have shown that the privacy of bystanders (i.e. the individuals that are around users using these devices) is also often involved, as online-shared data often involves individuals other than the users sharing the data [4]. In other words, bystanders' personal information can be collected without their knowledge or consent [1].

A common real-life example could be a person taking a picture in a busy street, where the faces of bystanders are recognizable. The picture can later be posted on social media, and neither the posting or the taken picture have been made with the knowledge nor consent of the bystanders [2]. This example can easily be adapted with different mediums (i.e. pictures, videos, audio), and the collected data can be of different nature (i.e. face, voice, location). This high amount of devices capable of collecting data of different nature results in high pervasiveness in bystanders' privacy. This is a significantly harder problem to solve than regular user privacy, mainly because bystanders can be unaware that data involving them is shared in the first place [4].

In the past, several attempts with different approaches have been made in order to ensure the privacy of bystanders. Technological solutions exist: for protecting visual privacy, techniques such as Gaussian blur or pixelisation have been used in order to anonymize individuals—this is most commonly seen on TV [5]. However, with the framework they present in [5], Dufraux and Ebrahimi define Gaussian blur and pixelisation as naive methods to hide identity. Their results show high recognition rates (up to 56%) from the attacker (face recognition algorithms) on images that have been blurred or pixelised, even when privacy enhancing techniques are applied on the training set of images (recognition rate similar for Gaussian blur and up to 45% for pixelisation). Regarding video surveillance systems, the authors of [6] point out that the high deployment rate of these systems in public places leads to privacy invasion of the objects (i.e. individuals) being recorded. They mention several studies on privacy based on video surveillance: two studies proposed image processing methods in order to protect the privacy of objects [7], [8]. In two other studies, the privacy protection of objects depends either on the authority of objects or observers [9], [10]. Privacy information can be embedded by using digital watermarking technology in such way that only predefined authorized viewers will have access to it [11]. Later, proposals have been made to However, these solutions are not flexible enough, because the sense of privacy from different objects is not considered. Surveillance cameras are not the only threat for the visual privacy of people. The availability of cameras within modern mobile and wearable devices has also increased people's concerns about their visual privacy, as taking photos or recording videos before sharing them online is easier then ever nowadays. Moreover, sharing pictures and videos online could reveal more information than expected, especially when the data is publicly available. The usage of recognition technologies which are able to correlate the shared data to specific people, places, and things, could

make the data searchable [12], [13]. Even when taking photos is not involved, applications making use of the camera, such as AR apps, could still compromise the visual privacy of bystanders by leaking the captured visuals, maliciously or not. On the online-sharing side, Olteanu, Huguenin, Dacosta and Hubau [4] claim that few solutions exist for detecting and sharing interdependent data in a way that preserves privacy of interdependent users. They add that existing solutions are limited in efficiency, as they all assume that bystanders are aware of data including them being shared, and often disregard adversary models, such as the online service provider. On the legal side, simpler solutions have also been applied (e.g. forbidding the use of cameras, smartphones, or AR devices like Google Glass in certain places [2]. However, banning such devices is not a good and fundamental solution, because it takes away the possibility of people to capture and share happy moments, especially if there are not any bystanders around. As a result, there are increasing needs to design technical solutions which can protect visual privacy of individuals while not restricting the rights of other individuals.

Due to the inefficient past attempts to ensure bystanders' privacy, the issue remains to be solved. There can however not be a perfect solution, because of several critical aspects. First, individuals have different requirements in terms of privacy, and these requirements can change over their life [2]. According to Westin [14], they are categorized between privacy fundamentalists (i.e. distrustful regarding organizations requiring their data), pragmatics (i.e. deciding whether they want to obtain various services, opportunities, and more, in exchange of the potential pervasiveness caused by the organisation's information seeking), and unconcerned (i.e. trustful regarding organisations requiring their data). Although this categorization concerns the privacy requirements of individuals, the same can be applied when they become bystanders to the eye of others. In consequence, fundamentalists will seek solutions that will, for example, systematically anonymize them to ensure their privacy as bystanders. Pragmatists will decide whether they want to adopt a certain solution depending on certain aspects (e.g. the cost, how good the solution protects them, the usability of the solution) and with varying requirements in privacy (e.g. decide to not be anonymized to trusted users). Unconcerned will remain unaware of having their privacy compromised, or will not want to adopt a privacy enhancing solution because of the added complexity. This means that there cannot be one solution to fit everyone. Moreover, because bystanders can be unaware of being present in other individuals' data, most solutions require that they register themselves on the system to define their preferences. In other words, very few solutions can enhance the privacy of bystanders such that they don't need to care or take action about it.

Recent technologies for ensuring bystanders' privacy have had to adapt to these challenges. As a result, they are mostly based on user participation, meaning that users create a profile on a server used by the technology, where they indicate their preferences in terms of privacy [2], [6]. This allows bystanders with any privacy requirements to be satisfied; the unconcerned bystanders are not required to register themselves. These technologies are centred about one or more aspects of privacy: the majority aim to protect the visual privacy of bystanders (with again different approaches, e.g. privacy regarding drones, surveillance cameras, etc.), and/or sometimes, their location privacy [4] or audio privacy [15].

The protection of bystanders' privacy is a concern that touches several domains, namely economical, social, legal, and technological [1]. We have defined our focus by first selecting a variety of papers that evoked, or treated the topic of bystanders' privacy. From this basis, we refined our selection based on technological solutions that enhance the privacy of bystanders, which led us to further literature. As a result, we present a number of technologies that ensure bystanders' privacy around one or more aspects (e.g. visual privacy, location privacy, etc.), along with their limitations.

This report focuses on the different technologies that address the pervasiveness in the privacy of bystanders. Section II lists different real-life examples of bystanders' privacy being compromised. Section III goes over different technologies that ensure different aspects of the privacy of bystanders. Section IV describes the current limitations and challenges that these technologies are facing, whether they are technological or not. Finally, section V concludes this report.

## II. Bystanders' privacy pervasiveness

This section presents a number of examples where the privacy of bystanders was compromised in one way or another. These examples are divided into four categories: Visual Privacy, Audio Privacy, and Location Privacy. For each example, an explanation of why the situation is problematic is given.

### A. Visual Privacy

Visual privacy is the most common aspect of privacy that is threatened in Ubiquitous Computing, because of the never-ending increasing amount of devices containing cameras, e.g. smartphones, surveillance cameras, drones, laptops, dashcams, AR devices such as Google Glass, etc. [1], [6], [16]. Bystanders can be seen on pictures or videos without knowing it, and the contents can be shared online without their knowledge nor consent.

One of the most destructive forms of bystander privacy invasion is "revenge pornography", which consists of an individual publicly disclosing sexual content regarding their (usually) ex-partner, hence the term "revenge" [4]. Individuals having been recorded or photographed—sometimes without being aware—by their partner have no way of limiting the spread of the contents online. Because of the disastrous consequences this can have on victims, using a system that systematically requires the consent of bystanders to allow publication is crucial.

Examples of visual privacy being compromised by drones are described in [16]: small drones equipped with cameras can spy or stalk individuals without them noticing. The authors define a number of scenarios where drones are used and

may capture bystanders. For example, describe a scenario in a public park: "A drone controller is flying his drone in a public park and taking photos and videos for fun. [...] You and your family members may be captured in the pictures and videos taken by the drone". The authors point out that although technology-based mechanisms exist in drones in order to protect the privacy of bystanders, their perception by bystanders and individuals controlling the drones are unclear. Furthermore, they note different preferences depending on the context, further insisting on the need for configurable solutions.

More recently, IoT has taken more space in private spaces like homes. As such, individuals who are not the main users, such as family, friends, or domestic workers (e.g. childminders, babysitters and au-pairs), become bystanders whose privacy can be compromised by smart-home devices. For example, the perspectives of domestic workers and their employers (e.g. parents) regarding smart-home surveillance are studied in [17]. Some of the interviewed domestic workers indicated being bystanders to installed smart-surveillance systems, while others felt that they were primarily targeted on them. Some of the interviewed parents admitted having installed such systems for the specific reason of observing the domestic workers. The authors also mention how children may be targeted by smart-home devices. These examples underline the necessity of developing privacy policies to protect bystanders in the case of smart-home devices.

### B. Audio Privacy

Along with the visual privacy usually goes audio, since ubiquitous devices are often equipped with a microphone. Therefore, previous examples can also be applied in the case of eavesdropping, e.g. a drone or a smartphone can record a video of bystanders having a conversation.

However, the case of audio privacy has been more impacted by IoT devices such as voice-assistants. In fact, cases have been reported where audio files were unwillingly sent to other individuals. For example, an Amazon customer received 1700 audio files of a stranger who used an Amazon Alexa device (i.e. Amazon's voice-assistants brand), when requesting his own archived data [18]. Not only did the audio contain the voices of the user, but also of the bystanders that were occasionally heard. Coupled with social media information, the identity of the person heard on the files could be established. Although this was the result of a human mistake on the side of the company, this pinpoints the lack of control over the information stored by such devices. More generally, individuals living with a user of such voice-assistants become bystanders and are exposed to these devices, which can lead to information disclosure, e.g. because of a misinterpreted set of instructions [18].

Another domain where audio privacy can be compromised is healthcare. In [15], the authors observe that cough detection systems disclose the audio recordings of both users and bystanders because of the lack of focus on maintaining the privacy of recordings. This let health professionals or physicians access to private speech from the user of the cough detection system and the bystanders around them. Therefore, more work has to be done in order to optimize the privacy of users and bystanders for such devices that record continuous audio streams.

### C. Location Privacy

The protection of location privacy is crucial for users. Because of the high amount of policies and regulations regarding the sharing of one's location, the location of bystanders cannot easily be disclosed without their consent. However, it is still possible for users to link bystanders with locations. For example, a Facebook user could post a picture of themselves that also contains the faces of bystanders, while indicating the location where the picture was taken. Further linking can also be done if the user tags the people appearing in the picture as other Facebook users, directly linking an identity to a location. Although Facebook users can adjust their privacy settings to allow or disallow others to tag them [19], the unregistered or untagged bystanders appearing on photos or videos that have a location linked to them still have their location information disclosed without their knowledge or consent. These issues regarding the non-consensual sharing of bystanders' information such as their location have to be addressed in a system that could become mandatory by law [4].

## III. Technologies for ensuring the privacy of bystanders

In this section, we present a simplified overview of several privacy enhancing technologies for bystanders. For each presented technology, we provide a typical run case and a system overview. Section III-A describes a technology that adds visual abstraction on recognized individuals at the time of data collection (in this case, when a video surveillance camera is recording). Section III-B presents a similar, but more modern, general approach that can be used with ubiquitous cameras attached to different devices. Section III-C describes a framework for managing interdependent data that is to be shared, in order to obtain the consent of all involved bystanders before sharing data.

### A. PriSurv System

The PriSurv system [6] can adaptively protect the privacy of objects and disclose their visual information according to the privacy policies of the different objects. PriSurv is a video surveillance system, which is defined by visual abstraction and protects the privacy of objects appearing within a video depending on who observes the video. The closeness between objects and observers determines the privacy policies to be used.

*1) Simple run case:* Figure 1 is an example run case of PriSurv, which shows how visual abstraction is used in order to protect the privacy of an object appearing within an image. Let o, a, v and S denote an object, an abstraction operator, a viewer and an original image, respectively. In the example, "Taro" is the object which appears within the original image

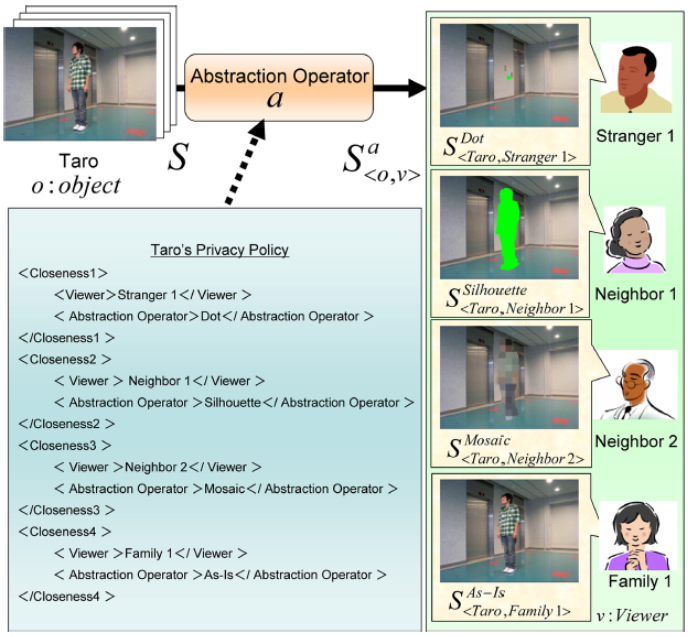


Figure 1. PriSurv simple run case [6].

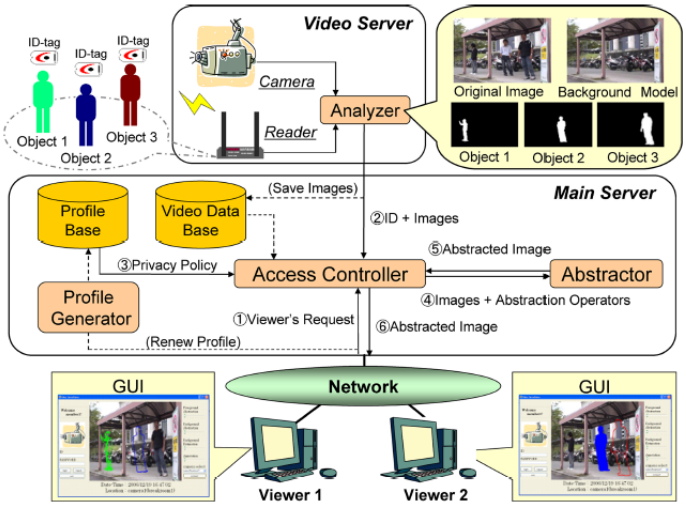


Figure 2. PriSurv system overview [6].

and is monitored by "Stranger 1", "Neighbor 1", "Neighbor 2" and "Family 1", which are the different viewers. Since the closeness between the object and viewers is different, different abstraction operators are used for hiding the visual information of the object. In the example, these operators are "Dot", "Silhouette", "Mosaic" and "As-Is", which are 4 of the 12 possible operators. Each viewer then receives a privacy protected version of the original image which is denoted by $S^a_{<o,v>}$. A simplified version of Taro's privacy policy, which is a part of the abstraction operator, is also available to give a better understanding of how the closeness is defined.

*2) System overview:* Figure 2 shows the architecture of the PriSurv system. It consists of three main parts: Video Server, Main Server and Network. It also has six different components: **Analyzer**, **Profile Generator**, **Profile Base**, **Access Controller**, **Abstractor** and **Video Data Base**.

The **Analyzer**, which is a part of the video server, is responsible for identifying different objects. Each identifiable object must have its own *Radio-frequency identification* (RFID)-tag. The surveillance area is divided into smaller N x N areas and the location area of each RFID-tag is determined by an RFID-reader. Each object inside the original image is extracted and then identified separately by comparing the obtained visual data with the binary images of each object.

The **Profile Generator** is used for setting up privacy policies for different members of the system. Each member's personal information such as name, age, gender, address and privacy policy is stored securely inside the **Profile Base**. The Profile Generator is also responsible for converting data taken from the GUI to *Extensible Markup Language* (XML)-based syntax. The profile of each member can only be updated by them and other members have no access to this data.

The **Access Controller** determines the closeness relationship between a requesting viewer and an object to be monitored by reading the XML-based privacy policies of the object stored inside the Profile Base. Once the types of abstraction operators are extracted, the Access Controller sends them to the Abstractor.

The **Abstractor** is an image processor that performs visual abstraction by using abstraction operators.

The **Video Data Base** stores past video data and makes it available to viewers after appropriate visual abstractions are performed.

### B. Cardea Framework

As discussed in section I, there have been several unfruitful past attempts at protecting visual privacy regarding the pervasiveness of cameras. Later, there were proposals for using visual markers [20], [21] or colourful hats [22] with which people can declare their disagreement to be captured. These approaches, however, are not flexible enough, because people should be able to control, modify and express their individual privacy preferences naturally—without the need of any extra facilities. The PriSurv system, which was discussed in the previous section, also relies on extra facilities such as RFID-tags and readers. Thus, it is not a feasible solution for built-in cameras.

The Cardea framework [2] does not rely on such or similar extra facilities. Instead, it takes advantage of computer vision techniques, which are effective and reliable. Moreover, it allows people to change their privacy preferences dynamically, e.g. by using predefined hand gestures. The framework also specifies context elements, such as scenes and the presence of others. People can set their personal privacy profiles, the hand gestures to be recognised by cameras for flexible interactions, and their context related privacy preferences. For example, some people might prefer their visual appearance to be hidden inside bars, but not in parks. But in case they change their mind and, for example, decide to appear in a photo taken inside a bar, they can indicate it with a hand gesture. Devices using the

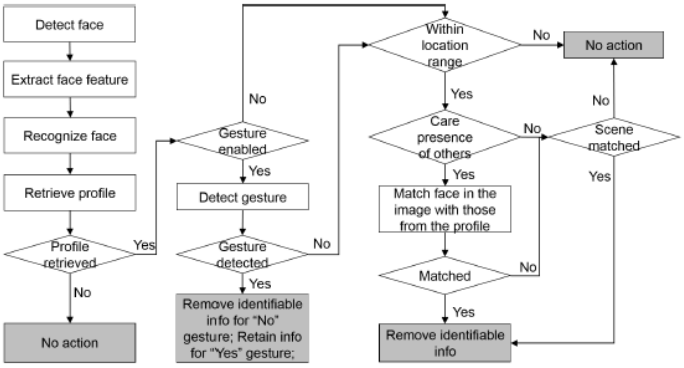


Figure 3. Cardea framework workflow [2].

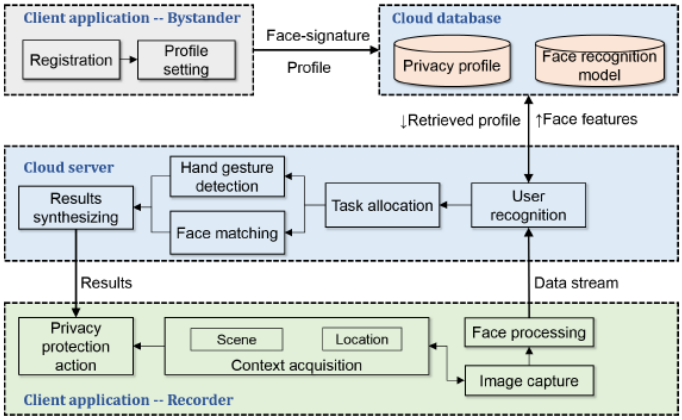


Figure 4. Cardea framework system overview [2].

framework will automatically compute factors related to the context, check people's privacy preferences, and protect their visual appearance by blurring their face.

*1) Key Concepts:* A few key concepts have to be briefly described before proceeding to the system overview. A person who has a device with an integrated camera and is taking photos is referred to as the **recorder**. People who may appear in these photos are the **bystanders**. Bystanders who do not want to be identifiable in the photos can express their privacy preferences by using the Cardea framework. On the other hand, the recorder who cares and respects the privacy preferences of the bystanders can use the Cardea framework to take photos. Both the recorder and the bystanders are classified as **users** of the Cardea privacy protection framework. Each bystander sets their **privacy profile**. Each profile contains information such as **context elements**, **gesture interaction** options, and the **privacy protection action** to be made. Bystanders have different privacy concerns depending on the context elements which can be locations, scenes and nearby people. The framework supports two gesture interactions: a victory sign and a palm, which stand for "Yes" and "No", respectively, to the question "Would you like to appear in the taken photo?". If a bystander does not use any of these gestures while the photo is taken, then the default privacy settings set in their privacy profile apply. The taken action in order to protect the visual privacy is face blurring. But further methods such as full body blurring, blending the body region into the background, or replacing the face with an average face are possible to be integrated. All of the bystanders' preferences are stored inside a **cloud** which accepts requests from the recorder and processes images in such way that the privacy preferences of the bystanders appearing in the images are respected. Figure 3 visualizes the decision workflow of the Cardea framework. Decisions are made according to the computing results from the camera and remote cloud, then the original image is processed, in case it contradicts the privacy preferences of the bystanders.

*2) System Overview:* Figure 4 shows the interactions between the major components. Cardea is composed of the client app and the cloud server. Bystanders can register, setup their profile settings and upload their photo by using the user interface of the client application. The face features of the bystander are extracted and stored into the cloud server's database to be used for training the face recognition model. The privacy profile is also stored in the database. Bystanders can change their privacy preferences at any time later.

Recorders use the client application to take photos. The application detects faces in the taken photos, extracts face features and computes context elements on the device. The face matching, user recognition and hand gesture detection processes are done on the cloud server, because they are computationally-intensive and infeasible on most devices. Once the results from the cloud are received, protection action will be performed.

The cloud server is responsible for storing user profiles and processing detection tasks upon requests from the recorder side. The user recognition process is divided into three steps. The first step consists in detecting face regions on the image before filtering the results in order to reduce false positives. The second step is to extract face features in a compact, but still discriminative way, so the face detection is done using these features rather than comparing just raw pixels. The last step is to match these representative features to faces stored on the cloud database.

The location filtering feature helps users defining specific locations in which they may worry more about their privacy. The location of the device taking a picture is obtained directly from the *Global Positioning System* (GPS). Scene classification is a more complex concept, because it does not only rely to places, but also gives information about the current activities of people. The framework covers 9 general scene groups which are further divided into more specific scenes. In total, there are 98 different scenes. Users can more easily specify their privacy preferences by selecting from these general scene groups.

*3) Example run case:* Figure 5 shows a privacy protection example of the Cardea framework. The framework detects three registered users, two hand gestures and a scene. The scene is classified correctly as Park & street. The face of Hao is detected and blurred by default in order to match his visual privacy preferences. The face of Zerry, under normal circumstances, would not have been blurred, but because he


Figure 5. Cardea framework example run case [2].

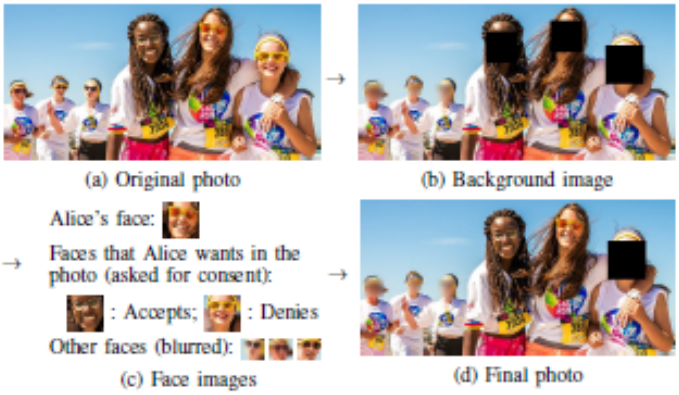


Figure 6. ConsenShare framework example run case [4].

is doing a "No" gesture while this particular picture is taken, his face is blurred. Jiayu has the opposite case of Zerry. Under normal circumstances, her face would have been blurred, but because of her "Yes" gesture, her face is identifiable.

### C. ConsenShare

Over the last decade sharing of variety of *Interdependent Personal Data* (IPD) and *Multiple-Subject Personal Data* (MSPD) online has drastically increased due to emergence of *Online Social Networks* (OSN)s such as Facebook. The shared data can include very diverse content such as contact data, location data, multimedia data (audio, photo, video) and genomic data. Such data often involves information not only about the individual who shares them online, but also for other data subjects [23]. Some data types, such as genomic data, photos and videos (especially ones including sexually-explicit content) attract more attention than others. Although there are some attempts to decrease the exposure of such data, typically for photos, they are not enough. For example, Facebook notifies users if they are tagged on some content. The users can then remove the tag or report the whole content for removal. However, this solution is still limited because users have to be either explicitly referenced (e.g. tagged) or they have to find the content themselves. Moreover, even after the content is eventually removed, it could be late, because some other users or service providers might already have seen the shared content.

*1) ConsenShare Overview:* The ConsenShare framework [4] considers three major entities: The users, the *Content Management Service* (CMS) (e.g. OpenSNP for genomic data, Flickr for photos, or YouTube for videos) and the *Identity Management Service* (IMS) (responsible for identification of the users involved in the content). In practice, agencies or popular OSNs, such as Facebook, could play the role of the IMS. Responsibility could be also split across several entities. Content uploaded to the CMS will be first sent for verification to all users whose privacy could be compromised, after they are first identified by the IMS. The users will also receive contextual information about the upload content, such as the identity of the uploader, target audience, upload time, etc. The content will be visible to the target audience only after it is confirmed by all the users.

Developers of ConsenShare framework assume that the CMS and IMS are honest, trustable and independent parties. However, they could still be curious and try to infer sensitive information from the data observed. Moreover, conclusion about social ties among users could be made. For instance, if a user named Bob often accepts that other user Alice shares content regarding him, it is likely they are good friends. Other threats such as creation of fake profiles and taking the identity of another individual still remain. Developers also assume that secure two-way communication channels, typically over HTTPS, have been established between all entities. A malicious user might still try to gather data and monitor communications among the different entities in the framework.

*2) Example run case:* Figure 6 shows an example application of the ConsenShare framework. (a) is the original photo taken by Alice. The application of Alice will produce the Background image (b) by removing the detected faces and extracting them into separate face images (c). The faces which appear in the background will be automatically blurred. The background and face images will be then sent to the CMS and IMS. IMS will identify other two individuals on the photo. They will receive notification and will be asked for consent. After all users choose their privacy settings, the final photo (d) will be produced and hosted by the CMS. In the example, one of the individuals accepts and the other one denies consent. Alice's consent is automatically granted by default, as she is the photo's uploader.

*3) Possible extensions for ConsenShare:* The ConsenShare framework can be extended for other data, such as audio, video and co-location. Video and Audio extensions are straightforward, because they are similar to the photo solution. The only difference is that the users have to be identified by the IMS by using different solutions. Different solutions for identification in audio and video data have been proposed [24], [25]. Co-location data privacy is a more complex topic, because the data itself can introduce dependencies among the different users. For example, if Alice and Bob appear in the same picture and they shared the location, it is uncertain how to protect the

location data of their future posts. Alice's future location posts could affect Bob's location privacy and vice versa.

## IV. Limitations and challenges of privacy ensuring technologies

Although the proposed technologies try to protect the privacy of bystanders, there are still some limitations and challenges. PriSurv, for example, relies on extra facilities: RFID-tags and RFID-readers. These requirements limits the efficiency of this framework for massive usage.

Cardea and ConsenShare frameworks are more feasible for massive usage. However, these frameworks have other challenges. Although the both frameworks have a good detection and matching rate of faces, there is still a small margin of false positives and false negatives. False positive is the case when a face is detected in an area where none exists. False negative, on the other hand, is the case when an area having a face is not detected. This case causes more privacy related problems, because a face appearing on a picture would never be protected.

All of the discussed technologies require creation of a privacy profile and uploading private data to the system of the particular system. This itself could be considered counter-productive because the users have to give out private data about themselves before their privacy can be protected. The privacy of non-registered individuals is not taken into consideration. Only the ConsenShare framework is able to blur unrecognised individuals, but only if they appear in the background of the photo. Moreover, the proposed solutions are centralized. Maybe a decentralized peer-to-peer solution could potentially increase the trustfulness and make users more comfortable with sharing private data.

Usage of these frameworks is not compulsory or regulated with laws, so an individual should care for the privacy of others in order to bother using one of the frameworks. In the end, despite the good results these frameworks offer, it is all only in research, nothing is in today's apps.

## V. Conclusion

This report presented technologies that ensure the privacy of bystanders in various situations. An overview of past attempts at solving privacy issues has been seen, with the conclusion that they were mainly insufficient or inefficient. This leads back to tackling the problem today, with the challenge of providing flexible solutions for all users, depending on their placement on the privacy requirement spectrum. For this, modern technologies rely on user participation by asking users to create profiles with their personal privacy preferences. Some revolve around certain privacy threatening devices (e.g. drones, IoT devices...) [16], [17], others are more generic (e.g. Cardea [2]). A distinction can be made between technologies that protect the privacy at time of data collection (e.g. when a picture or video is taken), and technologies that protect upon sharing of the data.

Several privacy enhancing tools have been presented. PrivSurv [6], a video surveillance system, relies on RFID tags to detect individuals. It then processes the recorded videos accordingly to the privacy preferences defined by the concerned bystanders in the system's profile base: bystanders can choose between different abstraction operators that are selected by the access controller upon determination of the closeness between the bystander and the viewer. Despite not being the state of the art and relying on extra facilities, it remains an interesting first step and brings good ideas such as having profiles defined by the users themselves to match their personal requirements.

The same idea was taken almost ten years later in Cardea [2], which is the second technology we presented. The strength of Cardea lies in the absence of external facilities: bystanders can indicate their consent for being photographed with simple hand gestures. The two client applications (i.e. one for the users that with to take pictures, the other for bystanders to register their privacy profile) are linked through the cloud server, where the resource-heavy calculations (i.e. user recognition, hand gesture detection) are performed. As a result, bystanders can easily manage their preferences, thanks to the scene management, their privacy profile, and the possibility of indicating a change in preferences with hand gestures.

The third presented technology, ConsenShare [4], takes a different approach for protecting bystanders' privacy. Instead of acting at the time of data collection, it monitors data where bystanders are involved and necessitates the consent of said bystanders before users are allowed to share the data. This approach gives control to bystanders who become systematically aware if content they appear on is planned to be published. This way, bystanders can not only control their visual privacy, but also their location privacy. This kind of system is critical to address big privacy and ethical issues where sensible aspects of bystanders are disclosed, such as revenge pornography.

Despite our research, we could not find audio-centred technologies that aim to protect the privacy of bystanders, despite the existence for tools for protecting the audio privacy of users (e.g. in [15], [26]). This underlines the need for tools that aim to protect the audio privacy of bystanders.

One issue that remains, and is likely to stay until made mandatory by law, is that individuals are never required to use any of the proposed privacy enhancing tools. This means that bystanders' privacy can only be qualitative if a high number of users adopt these systems. Ideally, ubiquitous devices should have privacy by design, and bystanders should not have to take action to protect their privacy. In [4], the authors estimate that the use of systems that check for the consent of bystanders before online sharing could become mandatory by law in a few years. Another limitation comes in the perception that creating privacy profiles on a cloud for ensuring one's own protection is a paradoxical, counter-productive action. Lastly, and more importantly, all of the presented technologies, although promising and efficient, stayed in the field of research. Privacy protection for bystanders is still not addressed in everyday-use end-user applications and devices, despite the growing interest in research. However, the growth of clouds facilitate solutions like Cardea, making the use of such tools in everyday life a

hopefully soon reality.

## Abbreviations and Acronyms

**IoT** *Internet of Things*
**AR** *Augmented Reality*
**RFID** *Radio-frequency identification*
**XML** *Extensible Markup Language*
**GPS** *Global Positioning System*
**OSN** *Online Social Networks*
**MSPD** *Multiple-Subject Personal Data*
**IPD** *Interdependent Personal Data*
**CMS** *Content Management Service*
**IMS** *Identity Management Service*

## References


[1] Y.-H. Lu, A. Cavallaro, C. Crump, G. Friedland, and K. Winstein, "Privacy protection in online multimedia", in *Proceedings of the 25th ACM international conference on Multimedia*, 2017, pp. 457–459.

[2] J. Shu, R. Zheng, and P. Hui, "Cardea: Context-aware visual privacy protection from pervasive cameras", *arXiv preprint arXiv:1610.00889*, 2016.

[3] T. Denning, Z. Dehlawi, and T. Kohno, "In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 2377–2386.

[4] A.-M. Olteanu, K. Huguenin, I. Dacosta, and J.-P. Hubaux, "Consensual and privacy-preserving sharing of multi-subject and interdependent data", in *Proceedings of the 25th Network and Distributed System Security Symposium (NDSS)*, Internet Society, 2018, pp. 1–16.

[5] F. Dufaux and T. Ebrahimi, "A framework for the validation of privacy protection solutions in video surveillance", in *2010 IEEE International Conference on Multimedia and Expo*, IEEE, 2010, pp. 66–71.

[6] N. Chinomi and B. Ito, "Prisurv: Privacy protected video surveillance system using adaptive visual abstraction", in *Advances in Multimedia Modeling*, 2008, pp. 144–154.

[7] A. Cavallaro, O. Steiger, and T. Ebrahimi, "Semantic video analysis for adaptive content delivery and automatic description", *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 15, pp. 1200–1209, Nov. 2005. DOI: 10.1109/TCSVT.2005.854240.

[8] I. Kitahara, K. Kogure, and N. Hagita, "Stealth vision for protecting privacy", vol. 4, Sep. 2004, 404–407 Vol.4, ISBN: 0-7695-2128-2. DOI: 10.1109/ICPR.2004.1333788.

[9] J. Wickramasuriya, M. Alhazzazi, M. Datt, S. Mehrotra, and N. Venkatasubramanian, *Privacy-protecting video surveillance*. Feb. 2005. DOI: 10.1007/978-1-84882-301-3.

[10] A. Senior, S. Pankanti, A. Hampapur, L. Brown, Y. Li, A. Ekin, J. Connell, C.-F. Shu, and M. Lu, "Enabling video privacy through computer vision", *Security & Privacy, IEEE*, vol. 3, pp. 50–57, Jun. 2005. DOI: 10.1109/MSP.2005.65.

[11] W. Zhang, S.-c. Cheung, and M. Chen, "Hiding privacy information in video surveillance system", Jan. 2005, pp. 868–871. DOI: 10.1109/ICIP.2005.1530530.

[12] A. Acquisti, R. Gross, and F. Stutzman, "Face recognition and privacy in the age of augmented reality", *Journal of Privacy and Confidentiality*, vol. 6, Dec. 2014. DOI: 10.29012/jpc.v6i2.638.

[13] R. Shaw, "Recognition markets and visual privacy", Jan. 2006.

[14] M. Langheinrich, "Privacy in ubiquitous computing", in *Ubiquitous Computing*, CRC Press Boca Raton, FL, 2009, pp. 95–160.

[15] E. C. Larson, T. Lee, S. Liu, M. Rosenfeld, and S. N. Patel, "Accurate and privacy preserving cough sensing using a low-cost microphone", in *Proceedings of the 13th international conference on Ubiquitous computing*, 2011, pp. 375–384.

[16] Y. Yao, H. Xia, Y. Huang, and Y. Wang, "Privacy mechanisms for drones: Perceptions of drone controllers and bystanders", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 6777–6788.

[17] J. Bernd, R. Abu-Salma, and A. Frik, "Bystanders' privacy: The perspectives of nannies on smart home surveillance", in *10th {USENIX} Workshop on Free and Open Communications on the Internet ({FOCI} 20)*, 2020.

[18] Npr. (2018). "Amazon customer receives 1,700 audio files of a stranger who used alexa", [Online]. Available: https://www.npr.org/2018/12/20/678631013/amazon-customer-receives-1-700-audio-files-of-a-stranger-who-used-alexa (visited on 09/03/2020).

[19] Facebook. (2020). "How do I tag my friends at a location on Facebook?", [Online]. Available: https://www.facebook.com/help/201009576609790/ (visited on 09/04/2020).

[20] F. Roesner, D. Molnar, A. Moshchuk, T. Kohno, and H. Wang, "World-driven access control for continuous sensing", *Proceedings of the ACM Conference on Computer and Communications Security*, pp. 1169–1181, Nov. 2014. DOI: 10.1145/2660267.2660319.

[21] C. Bo, G. Shen, J. Liu, X.-Y. Li, Y. Zhang, and F. Zhao, "Privacy.tag: Privacy concern expressed and respected", *SenSys 2014 - Proceedings of the 12th ACM Conference on Embedded Networked Sensor Systems*, pp. 163–176, Nov. 2014. DOI: 10.1145/2668332.2668339.

[22] J. Schiff, M. Meingast, D. Mulligan, S. Sastry, and K. Goldberg, "Respectful cameras: Detecting visual markers in real-time to address privacy concerns", Oct. 2007, pp. 971–978. DOI: 10.1109/IROS.2007.4399122.

[23] S. Ahern, D. Eckles, N. Good, S. King, M. Naaman, and R. Nair, "Over-exposed? privacy patterns and considerations in online and mobile photo sharing", Jan. 2007, pp. 357–366. DOI: 10.1145/1240624.1240683.

[24] D. Reynolds, "An overview of automatic speaker recognition technology", *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, vol. 4, Jan. 2002. DOI: 10.1109/ICASSP.2002.5745552.

[25] A. H.Mansour, G. Zen Alabdeen Salh, and K. Mohammed, "Voice recognition using dynamic time warping and mel-frequency cepstral coefficients algorithms", *International Journal of Computer Applications*, vol. 116, pp. 34–41, Apr. 2015. DOI: 10.5120/20312-2362.

[26] S. Ahmed, A. R. Chowdhury, K. Fawaz, and P. Ramanathan, "Preech: A system for privacy-preserving speech transcription", in *29th {USENIX} Security Symposium ({USENIX} Security 20)*, 2020, pp. 2703–2720.