Analysing Privacy in Visual Lifelogging

Md Sadek Ferdous[a,], Soumyadeb Chowdhury[b], Joemon M Jose[c]

[a]Electronics and Computer Science, University of Southampton, UK
[b]Singapore Institute of Technology, Singapore
[c]School of Computing Science, University of Glasgow, UK

Abstract

The visual lifelogging activity enables a user, the lifelogger, to passively capture images from a first-person perspective and ultimately create a visual diary encoding every possible aspect of her life with unprecedented details. In recent years, it has gained popularities among different groups of users. However, the possibility of ubiquitous presence of lifelogging devices specifically in private spheres has raised serious concerns with respect to personal privacy. In this article, we have presented a thorough discussion of privacy with respect to visual lifelogging. We have re-adjusted the existing definition of lifelogging to reflect different aspects of privacy and introduced a first-ever privacy threat model identifying several threats with respect to visual lifelogging. We have also shown how the existing privacy guidelines and approaches are inadequate to mitigate the identified threats. Finally, we have outlined a set of requirements and guidelines that can be used to mitigate the identified threats while designing and developing a privacy-preserving framework for visual lifelogging.

Keywords: Privacy, lifelogging, privacy-by-design, OECD privacy guidelines

1. Introduction

The right to privacy is one of the fundamental human rights in any modern society. It advocates and facilitates mechanisms to uphold the privacy of all individuals within the society. However, what is private is highly debated. This is because privacy has social, legal, psychological, political and technical connotations [18]. Even more, privacy is of dynamic nature. What is considered private in a society can change considerably with time. Many of these changes are driven by technological advancements.

We are now facing a new wave of technical advancements that have far-reaching privacy implications than any existing technologies. One such technology is lifelogging. There are different types of lifelogging [18]. In this article, we exclusively focus on the visual lifelogging. The visual lifelogging enables a user, the lifelogger, to passively capture images from a first-person perspective and ultimately create a visual diary encoding every possible aspect of her life with unprecedented details. In the early stage of the lifelogging history, a lifelogging device was quite expensive. Understandably, the initial adoption of lifelogging was restricted among very few early adopters and enthusiasts who were quite keen to explore a novel technical innovation

Email addresses: `S.Ferdous@soton.ac.uk` (Md Sadek Ferdous), `soum.chowdhury@singaporetech.edu.sg` (Soumyadeb Chowdhury), `Joemon.Jose@glasgow.ac.uk` (Joemon M Jose)

and saw personal benefits in capturing their first-person perspectives in such a detailed fashion. However, with the introduction and availability of the next generation low-cost lifelogging devices, the lifelogging activity has been gaining a momentum in recent years. The possibility of ubiquitous presence of lifelogging devices specifically in private sphere has raised serious concerns with respect to personal privacy. Another factor that fuels this debate is the discreet size of many lifelogging devices which can enable a lifelogger to engage in lifelogging activities without raising any alert even in most private and sensitive places.

To illustrate the impact of a privacy breach in visual lifelogging, let us consider that a piece of a lifelog containing a picture of another person has been captured by a lifelogging device. Let us assume that such a lifelog along with GPS coordinates and timestamp embedded into it has been wrongfully exposed to an attacker. This is not a far-fetched assumption. Indeed, there are several commercial lifelogging devices in the market such as Narrative Clip [2] and Autographer [1]. In addition, augmented-reality smart-glasses such as Google Glass also has the potential to act as lifelogging devices when they are equipped with lifelogging apps. All these devices are equipped with sensors that allow to embed GPS coordinates of the location where the lifelog has been captured. The attacker can utilise ever-powerful and ever-accurate face detection and image search algorithms to identify the person in the lifelog and infer the whereabouts of the person using the GPS coordinates and the timestamp. This exhibits the danger to the invasion of privacy with just one single piece of lifelog. This is truly powerful compared to any textual information as one single piece of such textual information cannot be exploited to infer this type of knowledge.

This intrusive and invasive nature of lifelogging has understandably drawn the attention of advocates of personal privacy in different domains. There is an urge to understand the privacy implications, in the form of privacy threats, of such a ubiquitous technology. Different practitioners and researchers have explored different ways to understand these privacy implications and to propose, design and develop different frameworks to mitigate the identified threats. However, none of the existing works has considered a comprehensive threat model of privacy in lifelogging. Without such a model, it is difficult to comprehensively understand, identify, assess and address the risks associated with privacy threats. In addition, the concept of privacy has different dimensions influencing different stakeholders (actors) in multiple ways. None of the existing works has analysed this inter-relation; making it hard to analyse the privacy implications on these stakeholders. We aim to address these gaps in this paper.

With these gaps in mind, we have made the following contributions in our previous work [14]:

- We have formulated a novel definition of privacy with respect to lifelogging after analysing the effect of different dimensions of privacy in lifelogging.

- We have introduced a threat model which identifies different privacy threats in lifelogging.

The current paper is an extensive elaboration of the previous work with the following additional contributions:

- We have elaborated the previous contributions with additional contents and analysis in Section 3 and Section 4.

- We have analysed the existing privacy guidelines and approaches to investigate their adequacy in mitigating the identified threats.

- Finally, we have presented a set of requirements and a series of guidelines to mitigate the identified threats for designing and developing a privacy-preserving framework for visual lifelogging.

All in all, only around 25% of the current paper, contained within Section 2 and partially in Section 3 and Section 4, was included in the previous work and hence, the current paper provides a more thorough treatment on the issue.

Structure of the paper. An introduction to different aspects of lifelogging is presented in Section 2. Next, Section 3 formulates a definition of privacy in lifelogging after analysing the inter-relation of different privacy dimensions in lifelogging. The privacy threat model in lifelogging is introduced in Section 4. Next, we examine the mitigation of the identified threats by analysing two popular privacy guidelines in Section 5 and existing influential works within the scope of this article in Section 6. In Section 7, possible mitigation strategies that can be used for designing and developing a privacy-preserving framework for visual lifelogging are introduced. Finally, Section 8 concludes the paper.

2. Lifelogging

In general, lifelogging is a solipsistic activity that utilises pervasive computing technologies to capture the first-person view of the daily activities of a user in an automatic and continuous fashion. The main motivation for any user to engage in lifelogging is to create a digital representation of her daily experience that can be stored in a preferred storage medium for future recollecting, reminiscing, retrieving, reflecting, and remembering intentions [49] and/or for other purposes. To better understand and study the privacy implications in lifelogging, at first, we need to define the notion of lifelogging and study its different aspects.

Lifelogging is the process of creating a lifelog. There are a few definitions of a lifelog as well as of lifelogging in the literature. For example, in [13], a lifelog has been defined:"*as a form of pervasive computing consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive*". In [18], Gurrin et al. argue that this definition of a lifelog focuses only on the data-capture stage, ignoring other stages such as data gathering, storage, analysis and access. Then, a revised definition has been proposed in which lifelogging has been defined as:"*a form of pervasive computing which utilises software and sensors to generate a permanent, private and unified multimedia record of the totality of an individual's life experience and makes it available in a secure and pervasive manner*".

In this definition, a lifelog has been defined as a *permanent* multimedia record due to the fact that the cost of storing digital data including lifelogs is ever-decreasing. For example, it has been reported that images from wearable cameras spanning across six to eight years can be stored on a $100 hard drive [19]. However, we argue that providing a user with the capability to store massive amount of digital data at an ever-decreasing cost not necessarily guarantees the permanent storage of such data, since this is completely dependant on the reliability of the storage medium as well as the willingness of a lifelogger. In essence, there is a probability for any digital data (including lifelogs) to be stored permanently, however, not necessarily it is a permanent multimedia record.

Also, in this definition, a lifelog has been defined as a *private* multimedia record which echoes the viewpoint of the early adopters in which lifelogging was primarily considered to be a solipsistic activity captured only by the lifelogger for her own benefit. However, as suggested in [42], many users would be willing to share their lifelogs with others as it becomes a common practice and there is a technical capability to do so. It has been further studied how lifelogging can be an effective tool for memory recollection [31] and can be used as memory cues for people suffering from episodic memory impairment [34]. We predict that there will be other avenues in which a lifelog can be utilised in many more exciting ways in future which will require the lifelogger to share her personal lifelogs with others.

To rectify the stated shortcomings, we propose a revised definition of lifelogging, based on the definition of Gurrin et al. provided in [18], where our revised connotations are highlighted in bold.

Definition 1. *A lifelogging is a form of pervasive computing which utilises software and sensors to generate a* ***(potentially)*** *permanent, private* ***yet (potentially) shareable*** *and unified multimedia record, called a lifelog, capturing the totality of an individual's life experience and makes it available in a secure,* ***privacy-friendly*** *and pervasive manner.*

A lifelog can be of different types [18]. For example, a lifelog can be visual if it is recorded using a lifelogging device or smartphone in the form of an image or video. An aural lifelog is recorded via a lifelogging, audio or smartphone using its microphone in the form of an audio clip. An activity lifelog represents the monitored activities of a user such as the number of steps taken, distance travelled, caloric output, the duration of sleep, etc. recorded via an activity tracker, smartwatch or a smartphone. Also, a contextual lifelog can represent the location of the lifelogger at a certain time as well as capture the acceleration and movement of the lifelogger utilising different embedded sensors in a lifelogging device, activity tracker, smartwatch or a smartphone. However, in this article, we focus exclusively on the visual lifelog and, hence, from this point on, whenever we mention a lifelog or lifelogging, we will actually imply a visual lifelog or visual lifelogging.

Actors in lifelogging are the involved entities in the life-cycle of lifelogging (see below). Gurrin et al. have identified four different actors which are discussed below [18].

- **The Lifelogger.** A lifelogger is the entity which utilises a lifelogging device to capture and store lifelogs. Even though a lifelogger can be an inanimate object such as a robotic device as discussed in [17], we assume that a lifelogger is a person. This assumption is relevant yet non-restrictive with the focus of this paper since we are mostly concerned regarding the privacy of a person.

- **The Bystander.** A bystander is any person who is captured (intentionally, incidentally or accidentally) in a lifelog of another person (lifelogger). A bystander is not the person with whom the lifelogger has been interacting while capturing the lifelog. Examples of bystanders are strangers in an environment, family members, friends, colleagues, etc.

- **The Subject.** A subject is any person who is captured (intentionally or incidentally) in a lifelog of the lifelogger during their interaction.

- **The Host.** A host is the entity who bears the responsibility of storing a lifelog of the lifelogger. When a lifelog is stored by the lifelogger in a private storage medium (e.g. a hard disk) in a local setting (e.g. an office or home), the lifelogger acts as the host. However, when a lifelogger stores her lifelogs in a remote cloud storage medium, the cloud service provider acts as the host.

Like any activity, lifelogging has different stages. Gurrin et al. identified five stages in lifelogging: *capture*, *storage*, *processing*, *access* and *publication*. We introduce another stage called *discard*. These six stages combinedly define the *Life-cycle* of lifelogging and are discussed below.

- **Capture.** In this stage, the lifelogging device captures lifelogs in an automatic and continuous fashion. The captured lifelogs are temporarily stored at the internal storage of the device.
- **Storage.** In this stage, captured lifelogs from the device are either stored in a computer storage or uploaded to a cloud storage. Stored lifelogs remain in the storage medium in a potential permanent state until they are discarded (see below).
- **Processing.** In this stage, stored lifelogs are analysed using image processing algorithms for extracting inner semantics which then can be used for temporal-spatial clustering, event segmentation and object/people detection and recognition.
- **Access.** In this stage, the lifelogger has the ability to access the captured and analysed lifelogs using a graphical user interface (GUI). The GUI generally organises the processed lifelogs in order to present them in a meaningful fashion. For example, the GUI can display the lifelogs based on spatio-temporal attributes or advanced visualisation mechanisms (e.g. clustering based on visual or event similarities) [54, 55].
- **Publication.** In this stage, either the lifelogger shares or delegates the task to another person to share (presumably a subset of) her lifelogs with other people.
- **Discard.** In this stage, the lifelogger deletes (a subset of) her lifelogs using the GUI. Once deleted, the lifelogs are either deleted from the local storage or from the cloud storage.

3. Privacy in Lifelogging

In this section, we develop a definition of privacy in lifelogging after an in-depth analysis of information systems, privacy, privacy dimensions and their associated inter-relationships.

3.1. Information System & Personally Identifying Information

An information system deals with information and is equipped with the functionalities of capturing, storing, processing and transferring of information to users [20]. An information system that is hosted privately in a user's computer can be regarded as a private information system. Examples of such systems are email clients, image, video or lifelog management software and so on. On the other hand, if an information system is hosted in a dedicated server and accessible to different users over the Internet, it can be regarded as a public information system. All online services that deal with information can be regarded as public information systems.

There are different types of information. However, to be in line with the focus of this paper, we are interested about *Personally identifiable information* (PII). A PII *"is any data that identifies or refers to a particular natural or legal person"* [24]. Another definition of PII as provided by NIST (National Institute of Standards and Technology) is: *"any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information."* [25].

A PII can be used to represent and identify, not necessarily uniquely, a person in an information system. It can take different forms in different systems. For example, it can be a value of an attribute (e.g. age, email address, telephone number, username and so on) or it can be a raw multimedia data such as an image, video or even a sound if the data can identify a person.

With its capability of identifying a person, a PII is one of the major sources that can affect the privacy of a user in an information system. The unauthorised access, use, or disclosure of PII can be exploited by an attacker for different types of attacks targeted towards people as well as organisations which store and mange such information. This is why it is crucial that special protection mechanisms are deployed to guard against the unauthorised exposure of any PII.

With our focus only on PII, whenever we talk about information in this article we will actually imply PII. Additionally, we consider a lifelog as a PII if it contains an image of at least one person since the lifelog can potentially be used to identify the person.

3.2. Privacy and Privacy Dimensions

What is the most appropriate definition of *Privacy* is highly debated. This is because privacy has social, legal, psychological, political and technical connotations [18]. A complex entanglement of these connotations dictates what can be considered private in a society. Interestingly, any perceived notion of privacy changes over time. With changes in social norms, political views and legal interpretations, what is considered private in a period of time may not be considered private in another period. The involvement of such different perspectives and their highly volatile dynamic nature make it harder to define a one-size-fits-all definition of privacy. The ultimate effect of this is the existence of a number of definitions of privacy from different perspectives and from different time periods. Next, we explore a few influential definitions and analyse their relevance and suitability in terms of lifelogging.

Motivated with the availability and popularity of modern photography and printing press and their implications on the privacy of people, Samuel Warren and Louis Brandeis wrote the seminal, influential paper *The Right to Privacy* in which they defined privacy as: *"the right to be alone"* [51]. It is thought to be the first definition of privacy [18] and devised with the motivation to protect people from nosy reporters who would take their photographs without their consent [33]. Unfortunately, this definition has lost its effectiveness in the modern day society where taking photographs of other people in public places is no longer considered a breach of privacy of those people, legally as well as socially. This notion of privacy is all about capturing the one's right to be in solitude and to protect him/her from intrusion in a physical domain. Hence, it is viewed as the privacy of personal sphere [46]. With the popularities of computers and computing systems and the possibilities of storing large amount of personal data into these systems and the capability of advanced data processing mechanisms, a new notion of privacy, called *Information Privacy*, in the domain of technology started to gain attention from 1960s onward. Since lifelogging lies inherently in a technical domain, we will restrict our attention to privacy definitions focusing on such technical domains.

In this regard, one of the most influential definitions of privacy was given by Alan Westin in [52] where privacy was defined as:*"the right to select what personal information about me is known to what people"*.

Similar to Westin's definition, several other definitions of privacy were proposed. A few of such definitions are presented below:

- **By I. Altman[4]:** *"Privacy is the selective control of access to the self or to one's group"*.
- **By Onn et al. [44]:** *"The right to privacy is our right to keep a domain around us, which includes all those things that are part of us, such as our body, home, thoughts, feelings, secrets and identity. The right to privacy enables us to choose which parts in this domain*

can be accessed by others, and control the extent, manner and timing of the use of those parts we choose to disclose".

- **By A. Moore [36]:** *"A right to privacy is a right to control access to and uses of - places, bodies, and personal information".*

The underlying notion in all these definitions is that privacy is all about controlling accesses of others to one's personally identifying information. This notion suits well with the focus of this paper as a lifelogging system essentially is an information system storing, processing, presenting and (potentially) sharing PII in the form of lifelogs.

Next, we explore how the concept of privacy is applied in the physical world using the concept of privacy dimensions. An understanding of this concept in the physical world will help to better analyse the concept in an information system since the application of the privacy concept in the physical world predates its digital counterpart.

Privacy dimensions denote different modes of privacy. Based on the four modes (*Solitude, Intimacy, Anonymity and Reserve*) of privacy introduced by Westin in [52], Pedersen conducted an empirical study and identified six dimensions of privacy in the social setting of the physical world [45]. The dimensions are explored below.

- **Reserve**. This represents the unwillingness of a person to be with and to interact with others, especially strangers.

- **Isolation**. This represents the desire of a person "*to be alone and away from others*".

- **Solitude**. This represents the state of a person when she is "*alone by oneself and free from observations by others*".

- **Intimacy with Family**. This represents the state of a person being alone with the members of the person's family.

- **Intimacy with Friends**. This represents the state of a person being alone with her friends.

- **Anonymity**. This represents the expectation of a person not being recognised or to remain unnoticed in a crowd and hence "*not wishing to be the centre of group attention*".

These six modes altogether define different aspects of privacy of a person in the social setting. Ensuring the gratification of these aspects can enable the right for a person to be private according to her needs. This is facilitated by social norms and legal practices. These social norms and practices draw a line, often imaginary, between what is private and what is public. However, advocates of personal privacy have witnessed a tension or even a threat to this imaginary line with the advent of modern technologies allowing devices, especially cameras, camcorders, mobile phones and tablets, to blur the distinction between what is private and public. One prime example of this is photography which once was considered intrusive and privacy-invasive around the beginning of 19th century (cf. Warren and Brandeis definition of privacy in [51]), now rarely is considered having privacy-invasive connotations in public places, neither with respect to social norms nor in legal practices.

3.3. Defining Privacy in Lifelogging

To define privacy with respect to lifelogging, we, at first, analyse the only one definition in the existing literature presented by Gurrin et al. [18] in which privacy in lifelogging has been defined as the "*the right to choose the composition and the usage of your lifelog and the right to choose what happens to your representation in the lifelogs of others*".

This is a very simple literal definition that captures the notion of user empowerment (especially data control) by enabling the lifelogger to capture a lifelog and the other actors (the subject and bystander) with the right to dictate what to do with the lifelogs in which they appear. However, this definition fails to embody other privacy dimensions. Even so, this definition lays down a solid foundation upon which a more elaborate definition can be formulated. This is what we aim to accomplish next. Importantly, our goal is to formulate an elaborate definition of privacy in lifelogging in an information system, mostly from an implementation perspective. This is because it is often difficult to concretise a literal definition of privacy as many fine-grained implementation details are often omitted. Next, we restrict our attention to an information system that primarily deals with lifelogs. The system may or may not deal with any other PII, however, we do not consider this in our definition. Therefore, from this point on, when we will talk about PII it will imply lifelogs.

With this goal and restriction in mind and based on the definitions of Moore in [36] and Gurrin et al. in [18], we have formulated an elaborate definition of privacy in lifelogging that embodies all privacy requirements of an information system and captures all dimensions of privacy in a physical world. The definition is presented below.

Definition 2. *The Privacy of captured lifelogs in an information system is the right that will enable all involved actors (lifeloggers, bystanders and subjects) to exercise anonymity when desired and to empower each respective actor by providing the required capability to exert privacy considering all (appropriate) dimensions while the lifelogs are stored in a storage medium, processed in a system, visualised in a graphical user interface and (optionally) shared among different users.*

Even though we exclusively focus on visual lifelogging, this definition of privacy is also applicable for other types of lifelogging.

This right can be enforced within an information system using the culmination of different technical capabilities which will be discussed later. It is easy to see that this elaborate definition embodies the need to address all privacy dimensions to ensure the privacy of each actor. However, the term *appropriate* in the definition deserves further attention. We argue that not all dimensions are meaningful for every actor in lifelogging. To examine this, we present the following analysis.

At first, let us concentrate on the lifelogger. As soon as a person equips herself with a lifelogging device and the device starts taking lifelogs, she assumes the role of a lifelogger until the lifelogger interrupts the device. This may continue even in the situations where she might be not be wilful (reserve) or desiring (isolation) to interact with other people in outdoor or even in indoor while in solitude. The similar argument applies to the *Intimacy with Family* and *Intimacy with Friends* dimensions as well. Interestingly, the anonymity dimension may not be always applicable as visual lifelogs dominantly represent the first-person view where the lifelogger is absent, unless, a lifelog contains her image (e.g. the device has been used to take a selfie or the lifelog contains the image of a lifelogger reflected via a mirror).

Next, we consider the case of a bystander. A bystander can appear in a lifelog when she, either alone or with family and friends, is in the same environment as the lifelogger even though she is unwilling or not desiring to interact with the lifelogger. Except when she is in solitude and not in the same environment as the lifelogger and hence, she has no possibility to appear in the lifelogs of the lifelogger. Thus, the dimension of *Solitude* does not apply to a bystander in lifelogging.

Finally, we analyse the case of a subject. A subject is the actor who interacts with the lifelogger. We argue if she would be reserved or in isolation, she would not interact with the lifelogger. Hence, we conclude that the dimensions *Reserve* and *Isolation* do not apply to the subject. Furthermore, using the similar argument for a bystander, we conclude that the dimension of *Solitude* does not apply to a subject. The other three dimensions (*Intimacy with Family*, *Intimacy with Friends* and *Anonymity*) are applicable for a subject in the setting of lifelogging.

Table 1 and Figure 1 summarise our analysis by illustrating which dimensions apply to which actor in lifelogging.

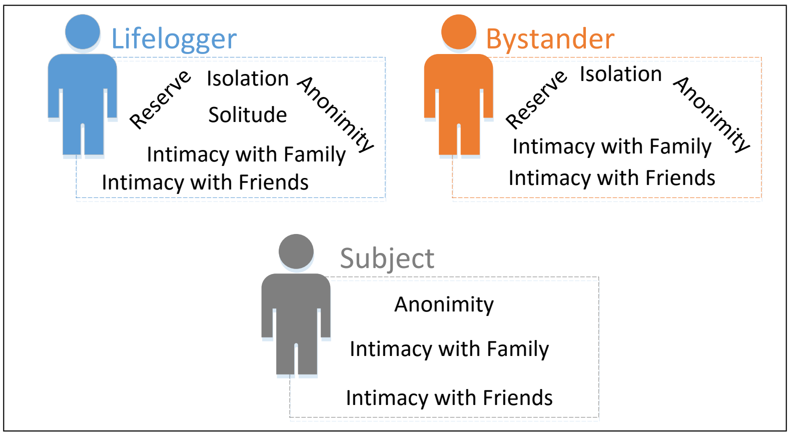

Figure 1: Privacy dimensions for each actor.

Table 1: Lifelogging actors and the privacy dimensions

	Lifelogger	Bystander	Subject
Reserve	✓	✓	–
Isolation	✓	✓	–
Solitude	✓	–	–
Anonymity	✓	✓	✓
Intimacy with friends	✓	✓	✓
Intimacy with families	✓	✓	✓

4. Privacy Threat Model in Lifelogging

Threat modelling is an integrated process of designing and developing a secure and privacy-friendly system. A well-defined threat model helps to identify security and privacy threats on different assets of an information system. In order to tackle such threats, different mitigation strategies need to be outlined by formulating security and privacy requirements [37]. In essence, a threat modelling consists of the following steps [12, 10]:

- enlisting assets of the system,

- identifying possible security and privacy threats on those assets with the assoiated risks and
- outlining mitigation strategies.

Since each information system has different assets, threat modelling process of one system will be considerably different than that of another system. Here, we would like to restrict our focus on modelling privacy threats in the setting of an information system that deals with visual lifelogs. There exists no paper in this context. However, there are a few papers focusing on the issue of threat modelling in the setting of web services [12, 10]. Based on these works, each single step of our threat modelling process is described in the following sections.

4.1. Enlisting Assets

An asset is the abstract or physical resource in an information system that needs to be protected from an adversary (attacker) [37]. It is the resource for which a threat exists and represents the target of the adversary in the system. The motivation behind this step is to highlight those assets in the system which can be the target of an adversary so that the threats for these assets can be identified. In this regard, we enlist the following assets:

Lifelog. In a lifelogging system, the captured lifelogs are the core assets since the main purpose of such a system is to deal with captured lifelogs.

Identity of a user. The identity of a user is defined as a representation of the user in a specific application domain [15]. In an application domain this representation, in reality, is encoded using different profile attributes such as (e.g., age, gender, email, address and so on) [15]. Each of these attributes can be leveraged to identify a user, no necessarily uniquely. For example, the age attribute having a value of 35 years can be used to identify (possibly) a group of users having the same value in the application system. On the other hand, a social security number attribute can be utilised to uniquely identify a user. Since lifelogs can be used to (probably uniquely) identify users, they can be linked with other profile attributes, representing the identity, of a user. Hence, the identity of a user should be considered as a crucial asset.

Information embedded within a lifelog. Not only the lifelog, but also information and meta-information (e.g. GPS coordinates) presented within each lifelog represents a valuable asset. This is because such embedded information offers a lucrative mechanism for an adversary to infer unforeseen knowledge about a user.

Access control mechanisms. The deployed access control mechanism in a public information system using either an Access Control List (ACL) or Role Based Access Control (RBAC) [29, 48] also represents an important asset. This is because the deployed access control mechanism determines which lifelog is exposed to which user(s).

4.2. Identifying Threats

A threat represents the activity or capability of an adversary onto an asset of a system with an intention to invade the security of the system or invade the privacy of a user in the system [37]. The main motivation behind this step is to identify possible threats on different assets of the system along with the associated risks so that proper mitigation techniques can be carefully planned. Based on the threat modelling process presented in [12], we identify the following threats. For each identified threat, the associated risk is also analysed.

T-1: Unnoticed Capture. A lifelogging device can be quite discreet in nature. This will allow a lifelogger to carry on the unnoticed capture of lifelogs.

- The associated risk is that other actors may appear in the captured lifelogs even without their knowledge and/or consent. This risk might be aggravated if such lifelogs are captured in sensitive places and/or within intimate environments which can be exploited if the lifelogger acts as the attacker.

T-2: Unaware Identification and unforeseen inference. An attacker can identify a person using a lifelog and ever-powerful image search online services.

- The associated risk is the unforeseen inference. Combining the identity of the person with other information (especially GPS coordinates) embedded inside a lifelog the attacker can create a profile of the person without her knowledge. Such profile can be used to create inferences for future occurrence of events which otherwise were not possible.

T-3: Lack of control. Many lifelogging devices (e.g. Narrative Clip) require the lifelogger to upload data in a cloud server maintained by the manufacturer even before the lifelogs are accessible to the lifelogger. The very nature of the lifelogging process which captures lifelogs in a continuous stream makes it very difficult for other actors such as subjects or bystanders to express their consent explicitly regarding their presence in the captured lifelogs.

- There are two associated risks. The first risk is from the perspective of the lifelogger. Once lifelogs are uploaded to the server, the lifelogger has limited control over them and she may not be aware how such lifelogs are being abused by the manufacturer. The second risk is from the perspective of other actors. Since other actors cannot express their consent explicitly while they are being captured in a lifelog, it becomes challenging to create access control rules which would allow these actors to exercise preferences while such lifelogs are processed, presented and shared in a system.

T-4: Inaccessibility of lifelogs. Lifelogs are inaccessible to subjects and bystanders until they are shared by the lifelogger. A private information system has no provision of sharing. Sharing in a public information system depends entirely on the goodwill of the lifelogger. If the lifelogger is the adversary, she understandably will not share the captured lifelogs with anyone.

- Here, the risk is the combination of other risks associated with T-2 and T-3. This is because if other actors cannot access the captured lifelogs, they cannot exercise the required control (T3) and hence, there is always a chance of unaware identification and unforeseen inference (T2).

T-5: Determining sensitivity. Sensitivity in a lifelog will determine if the lifelog can be considered private. For example, a lifelog containing only the image of the environment without the presence of any image of a person may be considered as of low sensitivity. On the other hand, a lifelog captured in a private and/or intimate setting can be considered as of high sensitivity. Technically, lifelogs are created in large volume. For example, modern lifelogging devices such as Narrative Clip allow capturing nearly 3000 lifelogs each day. This sheer volume of lifelogs makes it extremely difficult even for the lifelogger to pinpoint each sensitive lifelog. To make things worse, there is a possibility that lifelogs are inaccessible to the subjects and bystanders and hence, they may be entirely deprived of any opportunity to determine the sensitivity of lifelogs in which they appear.

- The risks here are twofolds. Sensitive lifelogs might be exploited by any attacker to gain monetary or other advantages. If the process of determining sensitivity relies on a manual intervention approach, it might not work properly because of the large collection of lifelogs.

T-6: Security. There is a strong inter-relationship between security and privacy. In many ways, different security measures safeguard the privacy of users in an information system. The threats related to security are presented below.

- **T-6.1: Secure storage.** The information system should take great care to securely store the captured lifelogs so that the attacker cannot access them inappropriately.
- **T-6.2: Confidentiality and integrity.** An attacker can intercept shared lifelogs while being transmitted, allowing the attacker to get hold of lifelogs in an unauthorised fashion and may alter a lifelog before they are transmitted to the destination system.
- **T-6.3: Unauthorised disclosure.** Lifelogs are disclosed to another unauthenticated and/or unauthorised user allowing the second user to get hold of such lifelogs inappropriately.
    - The risk for these threats is that sensitive lifelogs might be accessible to the attacker who then could exploit these lifelogs to gain illicit advantages.

Some threats are applicable for all actors whereas others apply to a specific actor. Also, as identified in Figure 1 and Table 1, the dimension of *Solitude* is not applicable for a subject and a bystander and the the dimensions of *Reserve* and *Isolation* are not applicable for a subject. By combining these two arguments, we summarise, in Table 2, which threats are applicable to which actor within which dimensions.

In Table 2, *R* represents *Reserve* dimension whereas *I*, *S*, *I-Fm*, *I-Fr*, *A* represent *Isolation*, *Solitude*, *Intimacy with Family*, *Intimacy with Friends* and *Anonymity* respectively. Furthermore, the symbol (✓) has been used to indicate that a threat is applicable to a particular actor within a dimension and the symbol (–) has been used to indicate the respective threat does not apply to the particular actor within a dimension.

As the third step of the threat modelling process, mitigation strategies for the identified threats need to be outlined. Within this scope, we explore the existing approaches. At first, we analyse the existing privacy guidelines against the identified threats. Then, we analyse the existing work with respect to privacy in lifelogging against the identified threats. The main motivation is to understand if the privacy guidelines and the existing approaches are adequate enough to mitigate the identified threats.

5. Threat Model and Privacy Guidelines

To counteract different threats pertaining to privacy of personal data, different guidelines have been proposed. In this section, we explore two popular privacy guidelines, namely OECD (Organisation for Economic Co-operation and Development (OECD)) Personal Privacy guidelines [41] and Privacy-by-design guidelines [5], and analyse their effect on undermining the threats identified in the previous section.

5.1. OECD Privacy Guidelines

One of the earliest examples of a considerably successful privacy guidelines is the OECD privacy guidelines. Introduced in 1980, it is known as the *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* [41]. Having presented in a technologically neutral manner, the guidelines exhibit flexibility in the sense that they can be adopted for any scenario where personal data is involved. It is interesting to note that these guidelines are not legally binding, meaning that any (particularly OECD) country can adopt it fully, partially or even ignore it altogether. Despite this, the guidelines have been proved to be influential in the policy

Table 2: Applicability of threats for each actor within different dimensions

		Lifelogger						Bystander						Subject					
		R	I	S	I-Fm	I-Fr	A	R	I	S	I-Fm	I-Fr	A	R	I	S	I-Fm	I-Fr	A
T-1: Unnoticed Capture		–	–	–	–	–	–	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
T-2: Unaware Identification and unforseen inference		–	–	–	–	–	–	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
T-3: Lack of control		✓	✓	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
T-4: Inaccessibility of lifelogs		–	–	–	–	–	–	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
T-5: Determining sensitivity		✓	✓	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
T-6: Security	T-6.1: Secure storage	✓	✓	✓	✓	✓	✓	–	–	–	–	–	–	–	–	–	–	–	–
	T-6.2: Confidentiality and Integrity	✓	✓	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓
	T-6.3: Unauthorised disclosure	✓	✓	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	–	–	–	✓	✓	✓

legalisation process for all OECD countries as well as in other countries and have served as the foundation in all existing data protection frameworks around the world. The guidelines are presented below [41, 43].

OG-1: Collection limitation. This guideline proposes that personal data should be collected legally and by fair means and, where appropriate, the consent of the involved person should be taken before such data is collected.

OG-2: Data quality. This guideline proposes that personal data should be relevant, accurate and complete. It also suggests that proper measures should be taken to ensure the accuracy of such data.

OG-3: Purpose specification. This guideline proposes that the purpose of data collection should be specified during data collection and such data should be destroyed if it cannot serve the specified purpose.

OG-4: Use limitation. This guideline proposes that data should only be used for the specified purpose. If it needs to be used for other purposes, the consent of the involved personnel must be taken.

OG-5: Security safeguard. This guidelines proposes that personal data must be protected using the state-of-the-art security safeguard mechanisms.

OG-6: Openness. This guideline proposes that data collection and processing need to be transparent to all personnel involved in personal data.

OG-7: Individual participation. This guideline proposes that involved personnel must have the right to access, update and or even delete any personal data.

OG-8 Accountability. This guideline proposes that the data controller (host) should be accountable for ensuring that all these principles are complied with.

Since the introduction of these guidelines in 1980, the landscape of personal data has changed dramatically. With the emergence of powerful and ubiquitous smart mobile devices, cheaper storage capacity and advanced analytic mechanisms, the ways the personal data is created, curated, stored and shared have been revolutionised. This has challenged the perception on the privacy of personal data and ultimately, an urge has been felt to revise the OECD privacy guidelines. In light of this urgency, the OECD privacy guidelines have been revised in 2013 [40]. The previous eight guidelines have been considered to be sound enough and hence, no fundamental change in those guidelines has been proposed. Instead, to address the novel privacy challenges, new measures have been proposed in such a way that they can be accommodated within the existing privacy guidelines. The more relevant new measures are presented below [40, 32]:

OM-1: Privacy management programmes. The revised guidelines advocate the introduction of a privacy management program which can be used for assessing privacy risks, for oversighting mechanisms to ensure the implementation of other privacy guidelines and for incident response mechanisms in case there is a breach of privacy. It also emphasises the need for streamlining any interaction between data controller and data owner via this privacy management program.

OM-2: Security breach notification. The revised guidelines advocate the need to notify security breach incidents to the appropriate national authorities as well as to affected individuals whose personal data is in question.

OM-3: Risk-based approach. The revised guidelines propose a risk-based approach to determine the sensitivity of personal data, assess risks based on the sensitivity and feed the assessment back to the privacy management program.

5.1.1. Analysis

OG-1, OG-3 and *OG-6* can be combinedly used to mitigate threat *T-1* as these guidelines would require the lifelogger to inform and take consent from any bystander or subject before the lifelogger captures them in her lifelogs. Threat *T-2* is unfolded when advanced image processing algorithms are conducted on the captured lifelogs. *OG-6* can be used to inform a bystander or a subject which data processing techniques are carried out in the lifelogs in which they appear. This, in a way, can be used to mitigate *T-2*. There are two aspects with respect to threat *T-3*. The lack of control from the viewpoint of the lifelogger when she uploads her lifelogs to the cloud storage medium of the device manufacturer and the lack of control from the viewpoint of the bystander/subject as she is captured in lifelogs of the lifelogger. The first aspect can be addressed with *OG-1, OG-3, OG-4 and OG-6 – OG-8* as these guidelines will enable the lifelogger to exert control over her captured lifelogs even when they are stored by the manufacturer. The second aspect can be addressed with *OG-1, OG-3, OG-4, OG-6 and OG-7* as these would allow the bystander or subject to express their consent or to enable them to exert control after lifelogs have been captured and stored in a system. Finally, *T-4* can be mitigated using *OG-7* whereas threats *T-5* and *T-6* can be addressed using *OM-3* and *OG-5* respectively.

5.2. Privacy-by-design (PbD) Guidelines

Another well known privacy guideline is called the *Privacy-by-design (PbD)* guidelines which outlined seven foundational principles for integrating privacy into the design and development of software systems [5]. PbD is defined "*as an engineering and strategic management approach that commits to selectively and sustainably minimize information systems' privacy risks through technical and governance controls*"[6]. In essence, PbD advocates the need for the recognition of privacy threats of personal data using a risk management process and then integrating privacy principles at the earlier stage of a system's design so that such threats can be minimised. The PbD guidelines are presented below.

PBD-1: Proactive not Reactive; Preventative not Remedial. This guideline outlines that PbD aims to mitigate threats pro-actively even before they are materialised. In other words, it aims to prevent the ways a privacy threat can be unfolded so that there is no need for remedy of privacy violations.

PBD-2: Privacy as the Default Setting. This guideline states that the privacy setting of a software should be enforced by default so that a user of the system has no need to take any action to guarantee the privacy of her personal data while using the system.

PBD-3: Privacy Embedded into Design. This guideline states that privacy enhancing techniques should be embedded into the design in such a way that they become an integral part of the system, not as an add-on in the event of a privacy threat.

PBD-4: Full Functionality - Positive-Sum, not Zero-Sum. This guideline states that PbD seeks to realise all required privacy functionalities without compromising any security functionalities so that no trade-off between security and privacy is required.

PBD-5: End-to-End Security - Full Lifecycle Protection. This guideline states that PbD emphasises the need for end-to-end strong security measures of personal data while it is stored, being transferred between different entities or being destroyed.

PBD-6: Visibility and Transparency - Keep it Open. This guideline states that the components and operations of a system handling any personal data must be open, transparent and verifiable to all involved stakeholders so that anyone can verify how their data is handled.

PBD-7: Respect for User Privacy - Keep it User-Centric. This guideline states that PbD upholds the interest of a person regarding her data with appropriate mechanisms such as default privacy, appropriate notice and user-empowerment.

5.2.1. Analysis

The PbD guidelines are more general and vague than the OECD privacy guidelines. This is because most of the guidelines in PbD are hypothetical in nature and lack any concrete recommendation by which they can be realised in any context. For example, it might be easy to understand that *PbD-6* can be used to mitigate threat *T-4* and *PbD-5* can be used to deter threat *T-6*. Unfortunately, it is difficult to perceive the concrete steps that make a system proactive and preventative (*PbD-1*); not only for lifelogging, but also for other scenarios. The same difficulties arise while analysing how other PbD guidelines can be used to mitigate the outlined threats with respect to lifelogging. Therefore, we conclude that PbD guidelines can only mitigate threats *T-5* and *T-6*, leaving all other threats exposed. A summary of threats that can be mitigated by the OECD and PBD guidelines is presented in Table 3.

Table 3: Threats and privacy guidelines

	OECD	PbD
T-1	OG-1, OG-3, OG-6	–
T-2	OG-6	–
T-3	OG-1, OG-3, OG-4, OG-6, OG-7, OG-8	–
T-4	OG-7	PbD-6
T-5	OM-3	–
T-6	OG-5	PbD-5

6. Threat Model and Existing Work

In this section, we analyse the existing influential works in this domain. For our analysis, we group the existing works in three categories: i) analytical, ii) systematical and iii) user study based. In the first category, we analyse the works which were analytical in nature. In the second category, we analyse those works in which a system, framework, model or an architecture have been presented in order to address privacy issues in lifelogging. In the third category, we analyse the works in which a user study has been conducted to understand different aspects of privacy in lifelogging.

6.1. Analytical

In [3], the author outlines several privacy issues in lifelogging. The main two concerns raised by the author are: (1) the ability for lifelogging technology to be used for pernicious records giving a lifelogger the ability to recall, replay, and remember for short, pernicious memory; and (2) pernicious surveillance - allowing a state to abuse the lifelogging technology for massive-scale surveillance where the lifelogger acts as the sibling of the "*big brother*". Regarding the first aspect, the author outlines several scenarios where lifelogging could be abused. The first scenario involves the capacity of the lifelogging technology to capture and store an embarrassing moment (or an action of committing a crime) of a person (or a convicted) for a longer period of time, maybe forever. We argue this to be true for other available technologies such as photography or video-taping using digital cameras or smartphones, hence the worry should not be raised for lifelogging only. However, there is a stark difference between lifelogging and other technologies, as the former has the capability to capture moments in unobtrusive and ubiquitous manner.

This gives it the opportunity to capture images in a places where no image should be taken. The extreme example is the toilet, however, more general examples may be those places where photography is not allowed. The second scenario is the "*voluntary, but pathological rumination*": allowing the lifelogger to live in the past which may have detrimental effect in the mental health as it is tangential against our natural process of being forgetful and moving forward. We argue that this could be the case; however, how many lifeloggers would like to revisit their painful memories and how such revisit may have detrimental effect on their mental health is an open question which is hard to predict without any proper psychological analysis or experiment. And also, it is arguable if this is a concern regarding privacy or health. We ask: Can restricting someone reminiscing their own memory be considered as the issue of privacy? The second concern is the capacity of lifelogs to be used as the surveillance medium by the Government as the Government has the capability to force anyone to release his/her lifelogging to the authority. The author argues that, in this manner, the lifelogger could be used as the sibling of the *big brother* and thus, the lifelogging technology could be abused as a tool of surveillance. In a way, it could be used as the supplement of the already established video surveillance. Especially, because of its capability to capture moments in an environment where there is no video surveillance, e.g. inside a private building, flat, house, etc.

A thorough discussion of the inter-relation of identities and privacy with respect to lifelogging is presented in [42]. This paper discusses a little bit of history and technological possibilities and constraints with lifelogging tools. In addition, the article seems to be a reply to what privacy concerns were raised by Allen's paper (cf. [3]). The authors argue that most of the privacy issues raised in Allen's paper are somewhat far-fetched. Their main argument is: lifelogging data should not be considered as private, instead should be treated as public data since most of these lifelogs will be made available publicly, just like a photo taken at a public place are not considered private even though there might be other people in it.

The authors in both these works have mostly concentrated on the implications of lifelogging in personal domains as well as how this activity may drive changes in attitude and behaviour in societies at large. Even though there are discussions regarding different privacy issues that reflect the underlying threats, there is no discussion on how such threats can be mitigated.

6.2. Systematical

The authors in [27] have presented a system, called *DARKLY*, which provides a privacy-preserving layer in a system where different (including potentially malicious) applications have unrestricted access to sensitive perceptual sensors such as camera and microphone. The DARKLY system acts as a middleware between such sensors and applications of the system. It consists of two components: 1) a local server and ii) a client library. The local server is a privileged application which has unrestricted access to the perceptual sensors. All other applications need to leverage the client library to interact with the sensors using the local server. When an application would like to access a perceptual sensor (e.g. a camera), it will need to submit a request to the server using the client library. The server would then collect the raw data from the sensor (for a camera it means a photo being captured) and apply image analysis mechanisms to identify objects such as a face. Then, the captured image will be shown to the user via a GUI called console where the user can apply different levels of filters which can obscure faces or other sensitive information according to the privacy choice of the user. After this, the obscured image would be released to the requested application. DARKLY will ensure that any untrusted application cannot capture images in a secretive manner and highly sensitive images will be released to an application only after privacy filters have been applied to it according to the privacy need of the user. DARKLY can be utilised to mitigate threats T-1 and T-3 for the lifelogger and T-2 and T-5 for any actor. Unfortunately, it cannot mitigate threats T-4, and T-6. Furthermore, it will be impractical to deploy the DARKLY system in any dedicated lifelogging device as there is no graphical interface for users to review the permission process. Hence, this approach may be suitable only for smart devices which are used as lifelogging devices.

A similar work as DARKLEY has been presented in [26] which provides an abstraction layer of privacy between raw sensor data and applications which request such sensor data. The abstraction layer is implemented with a recogniser which leverages image analysis mechanisms to identify and obscure sensitive objects. This is then reported back to the users using a GUI called privacy goggle in order to inform users what information will be delivered to the requested application and then aid them to grant permissions. Like DARKLEY, this approach can be utilised to mitigate threats T-1 and T-3 for the lifelogger (since other actors have no way to express their consent) and T-2 and T-5 for any actor, depending on the goodwill of the lifelogger. Unfortunately, it cannot mitigate threats T-4, and T-6. Like DARKLY, it will be impractical to deploy this approach in any dedicated lifelogging device as there is no malicious application in such devices and there is no GUI for users to review the permission process. Hence, this approach may be suitable only for smart devices which are used as lifelogging devices.

The authors in [50] have presented a probabilistic framework, called PlaceAvoider, to avoid taking lifelogs (or automatic images) in sensitive places by a malicious app in a smartphone. The system consists of a policy that a user can set to determine where not to take pictures or to apply different actions based on a location or the context of an image. It also has an image classifier that classifies a captured image and determines the context in the image and a policy enforcer. To set policies, users are required to capture and share the images, locations and/or visual models of sensitive places. Based on the images of these sensitive places, the image classifier uses a probabilistic algorithm to determine if the captured lifelogs are sensitive. Depending on the outcome of the algorithm and the policy, the policy enforcer enforces the policy by prohibiting taking images in such places or drawing the user's attention. The framework mostly focuses on the privacy issues that may arise while taking lifelogs in sensitive places. Also, the privacy of other actors is not considered; neither is there any discussion regarding different security issues. Even so, the framework can be used to mitigate threat *T-5*, i.e., to determine the sensitivity of each lifelog. Hence, the framework will remain ineffective against all other threats. It will also be impractical to deploy such an approach in any dedicated lifelogging device as there is no GUI for users to set a privacy policy based on previously captured images. Hence, this approach may be suitable only for smart devices which are used as lifelogging devices.

In [47], the authors have presented a novel approach, called World-driven Access Control (WDAC), that enables real-world objects and entities to define and disseminate access policies which dictate what should be sensed or recorded by the sensors of mobile devices. This is a similar approach as PlaceAvoider and DARKLY with the similar motivation where applications in mobile devices are considered untrusted. Hence, a middleware is used to detach the direct interaction between the such applications and sensors. Any real-world objects (a door in a bathroom) or entities (people) can express their desired access policies using visual tokens such as QR Codes using a novel type of certificate called Passport. An example of such a policy by a bystander can be expressed like this: "Blur my image". Such policies will be sensed and decoded by the middleware which will ensure that the policy is enforced. For example, any captured image in which the entity appears will be modified in such a way that the entity cannot be identified. An effective implementation of this approach can be utilised to mitigate threat T-1 and T-3 as any actor can control and express consent when she is captured in a lifelog. Similarly, WDAC can apply mechanisms to ensure that an entity is not identified which can be used to mitigate T-2. Since any entity express her own sensitivity, it can mitigate Threat T-5 as well. Unfortunately, like the previous three systems, it will be impractical to deploy the WDAC approach in dedicated lifelogging devices as these devices need to be equipped with additional processing capabilities for decoding QR codes and enforcing privacy policies. Hence, this approach may be suitable only for smart devices which are used as lifelogging devices. Also, WDAC burdens other actors to define their own access policies which may be impractical for different groups of people, e.g. aged people.

In [35], authors describe a framework for privacy-preserving lifelogging which enables a user to capture images of bystanders using a privacy policy. The policy dictates when and when not as well as where and where not to capture lifelogs. The condition may consist of different constraints such as temporal, spatial, spatio-temporal and other users. For example: a user can define where and when to take a picture of a particular user. A smartphone is simulated as a lifelogging device which can share privacy policies of different users using bluetooth. Each device is given a unique ID and it is assumed that each user will have a specific device with that ID. In such, each user is identified with the unique device id. Each smartphone is also equipped with an IR transceiver that transmits a device id to another device when both devices are in the line of sight of each other. Once a device receives a device id from another device using the IR channel, all policies in the devices that have been accumulated via bluetooth are checked to determine if an image of the nearby user can be taken. Based on the policy evaluation, an image is taken. The main drawbacks of this approach are: i) this approach is not feasible to be applied with any lifelogging device, ii) the default mode for the lifelogging app in the smartphone is to take pictures, this means that if the privacy policy is not found for a specific user or the IR transmitter of the bystander is not in the line of sight or obstructed, it will take pictures. The approach can be improved by toggling the default mode to disable taking pictures, meaning that the app will only take pictures if the IR transmitter of the bystander can transmit the device id to the user and the receiver of the user's smartphone can receive the device id and the policies and this results in a positive evaluation. This approach might allow a subject or a bystander to assert a certain amount of control (*T-3*) by exchanging privacy policies with the lifelogger, via their lifelogging devices, to allow or disallow capturing lifelogs with their images on them. Unfortunately, other threats cannot be mitigated using this approach.

One of the most comprehensive works with respect to privacy in lifelogging is presented by Gurrin et. al. [18]. In this work, the authors have introduced a novel definition of lifelogging and presented four actors involved in lifelogging activities: the lifelogger, bystander, subject and host. After a brief analysis on different existing definitions of privacy with respect to a digital system, the authors, then, have proposed a definition of privacy in lifelogging. The authors have argued that there are five-stages in lifelogging, namely Capture, Storage, Processing, Access and Publication and among these five stages, only in the stages of access and publication, restrictions must be enabled to protect the privacy of third parties. Finally, the authors have presented one of the first frameworks with respect to privacy-aware lifelogging. The framework is inspired by the principles of privacy-by-design and evolves around the following idea. A lifelogger is assumed to create a set, let us assume that it is called the *recognisable set*, of images containing a series of face models of individuals from whom the lifelogger has taken consent to capture and store lifelogs. Then, the lifelogger can capture lifelogs in an indiscriminate manner which are then securely stored in a storage medium. The captured lifelogs go through an array of image processing mechanisms to identify faces of individuals among the captured lifelogs. The privacy-by-design approach of the framework depends on a privacy policy which seeks to match the patterns of faces between the recognisable set and the stored lifelogs. If a face is identified in a lifelog and whose pattern matches with any face model in the recognisable set, the lifelog can be viewed in the lifelog explorer. If a face is identified in a lifelog and the face does not match any face model in the recognisable set, the face is automatically blurred so that the individual is unidentifiable while showing it in the lifelog explorer. If a lifelog does not contain any face or if a face is not detected, the lifelog is shown to the explorer without any restriction. One of the major limitations of this framework is the need to build face models by taking images of of all users who presumably have given their consent to appear in the lifelogs of the lifelogger. Depending on the time a lifelogging device is used to capture lifelogs, a lifelogger may come across unprecedented amount of people during the lifelogging activity which might make it quite difficult, if not impossible, to gather consent explicitly in such an evasive fashion. Even with this limitation, we believe that this work has laid down a solid foundation in the domain of lifelogging privacy. The current paper is highly motivated by the work of Gurrin et. al. and in many ways, have extended a few ideas (definitions and life-cycle of lifelogging) presented in their work. The PbD-inspired lifelogging framework can effectively mitigate threat *T-2* as it will blur every face in a lifelog for which there is no consent in the system. This will ensure that such lifelogs cannot be used for unaware identification and unforeseen inference. Similarly, it will provide a degree of control (*T-3*) to a subject and bystander in a sense that the lifelogs in which faces appear will be blurred by default unless a consent is present in the system. The framework also handles different levels of sensitivity to determine which lifelogs need to be blurred and which need not and hence can mitigate *T-5*. Even though it has not been explicitly considered, the framework also identifies the need for secure storage of lifelogs as well as keeping their integrity intact and hence, it can be concluded that the framework can be used to mitigate *T-6.1* and *T-6.2*.

The authors in [28] have presented a mathematical model for privacy in lifelogging allowing a user to map lifelogging data to him/herself and to access his/her lifelogging data under certain conditions. Then, the model is extended to include bystanders so that a relation between lifelogging data captured by another user and the bystanders who are in the lifelogging data can be established. However, this paper lacks in discussing any guidelines or a framework that could be used to mitigate the identified threats.

6.3. User study based

The authors in [23] have presented the result and analysis of a user study which investigates the way the location of a user can be shared with others in his/her social networks. For this, the authors initially conducted interviews with several participants mainly via questionnaire to understand the attitude of the interviewed users in sharing their locations with other users. Based on the response, they built a prototype which was then used to initiate a pilot study. Based on the responses, a final prototype was created and evaluated. The evaluation results have been used to create a series of privacy guidelines for any location sharing app. The study is not directly related to lifelogging. However, there are discussions on several similar privacy issues that can be applied for lifelogging as well. Importantly, it has been reiterated that people are willing to sacrifice a certain amount of privacy to a system only if the system and the released information provide substantial usefulness. The authors have also suggested to offload the burden of mundane security and privacy tasks to machines. The guidelines also have also emphasised the need for more control (T-3) of disclosing sensitive information (location) from the sensor. Other identified threats have not been considered in this study.

Another user study, reported in [38], has investigated the attitudes and privacy concerns of the users “surrounding” day-to-day tracking and recording technologies, namely, credit cards, loyalty cards, electronic toll systems, web server logs, CCTV cameras and RFID tags. Participants in the study have reported serious concerns regarding information privacy, yet they seem to be less concerned regarding the privacy issues of these tracking technologies. According to this study, participants have expressed their concerns regarding information privacy in tracking and recording technologies, especially with respect to unnoticed collection (T-1), unauthorised secondary use, e.g. inference (T-2) and improper access (T-6.3). Participants in one state (California) have been more concerned regarding their privacy than their counterparts in another state (Louisiana). In addition, female participants have been more concerned regarding their privacy than their male counterparts. The majority of the participants have commented that it is unrealistic to expect any level of privacy in public places. They argue that such data are usually stored by different organisations or even Governments (for CCTV surveillance) and they are bound by legal systems not to abuse such data. However, it does not apply to lifelogs since they can be stored by lifeloggers who have no such obligation, legally and ethically, not to abuse lifelogs. There is a discrepancy in attitudes of users using such technologies and their expectation of privacy. The results presented in the paper have identified two fundamental findings:

- people struggle to identify the possible threats associated with these technologies, and
- they fail to assess their capabilities and options to minimise these threats and allow them to negotiate how different private information is disseminated to other parties.

In [30], the authors have presented the result of a study regarding privacy concerns. The study has been carried out via interview among 24 participants in which the participants used pervasive devices to track their physical activities. To provide more accurate inference of their physical activities, the authors have presented the idea of using GPS devices to record GPS tracks and audio devices to record audio tracks in addition to other sensors in the particular pervasive device used in the study. Their results suggest that the privacy concern arising from continuous sensing is mixed and largely depend on three factors: what is being recorded, where it is being recorded and the perceived advantage of ubiquitous sensing. From example, data recorded by sensors such as accelerometer and barometer has been hardly considered privacy invasive whereas data recorded via GPS and Audio Mic has been considered highly sensitive. 42% of the participants have been against the concept of continuous recording GPS data and the others have been also uncomfortable on this idea. On the other hand, only 2 participants among 24 (8.3%) would allow a sensor to continuously record audio data because of the huge privacy implications. One of the findings of the study is that users can make informed privacy trade-off only if they understand what the technology is doing along with their underlying security and privacy implications. They have suggested to add mechanicms in a system which would allow users to see and control (T-3) any data being recorded. This work has not reported any concern regarding other threats.

In [39], the authors have presented a user study in two North American and two European cities involving one of the pioneering lifelogging devices called SenseCam. The main motivation of the study is to understand the responses of people, especially bystanders and subjects in the environment, and their implications with respect to continuous recording carried out by SenseCam. The participants of the user study have pointed out the importance of the visibility of lifelogging devices (concerning T-1) while it it being used to record lifelogs so that other people have the visual cue that they are being recorded. They participants have expressed a high degree of concern regarding the lack of control (T-3) and the sensitivity (T-5) of what is being recorded. They have expressed their desire to be informed or for consent being asked before any recording takes place. In addition, they have been worried about the inaccessibility (T-4) and unauthorised disclosure (T-6.3) of the captured lifelogs.

An evaluation of a user study focusing the attitude of lifeloggers is presented in [22]. For the study, a smartphone is simulated as a lifelogging device where the lifelogging functionality is achieved via an app. The user of the app has the option to start and pause a lifelogging session as well as delete a session recorded in the last 15/30/60 minutes. The user study has been conducted among 36 users within a week. The app would record and upload the lifelog recorded by the device of each user and the user has to evaluate their uploaded images at the end of each day. The user has the option to delete any sensitive image and then he/she has to state a reason for deleting an image. The user also is also provided with the option to label different places and later the labelling has been used to determine the correlation among deleted images. Based on the evaluation, the authors have drawn the following conclusions: i) some people prefer to manage privacy pro-actively by disabling lifelogging captures so that they can avoid the burdensome review of all collected images later; ii) the sensitivity of a lifelog can be determined by a combination of factors including time, location, and the objects and people appearing in the lifelog and iii) some lifeloggers have reported their concerns about the privacy of bystanders, even though no bystander opposed capturing lifelogs during the study. The main drawback of the paper is that privacy is not fully defined. The notion has been explored literally. Also, the emotion and attitude of bystanders have not been explored. They have retroactively concluded that the bystanders have had no problem to be recorded in other's lifelogs as none of the lifeloggers have been confronted by any bystanders. We argue that this is an imprecise conclusion for two reasons: i) a bystander may not have enough understanding regarding the lifelogging technology and even may not be aware what the term means and ii) every bystander should be exclusively interviewed to understand their reactions regarding this. The authors have reported results, analysis or suggestions considering the possibility of identification (T-2), control (T-3) and sensitivity (T-5).

The paper in [21] presents an extension of the previous paper in which the authors have analysed the images captured by different lifeloggers to find a correlation between the analysis of the images and the survey result. The motivation of the paper is to understand the following questions: 1) what makes a lifelog sensitive?; 2) what makes a lifelog private from the viewpoint of subjects in the study?; and 3) what can be learned about privacy by analysing the correlation? The study has found that lifeloggers are worried about privacy implications in lifelogging and they would like to ensure that private information, e.g specific objects (computer displays and physical documents) and activities in lifelogs, are suppressed before they are released or shared. Lifeloggers would most likely share their lifelogs, however, they would like to ensure that the privacy of bystanders is satisfied. Furthermore, the study concludes that the lifelogger may be overwhelmed with the large number of captured lifelogs which may cause accidental release of sensitive lifelogs (T-5). To mitigate this, the authors have emphasised the need for a privacy-preserving framework for lifelogging activities.

A user study examining the privacy perspectives of bystanders when they are within the vicinity of augmented reality (AR) glasses such as Google Glass is presented in [11]. The user study has been carried out in cafes and consists of 12 sessions where 31 bystanders have been interviewed. The reported result shows that the perspective of users change according to the context in which they are captured. Many participants have not noticed the AR device yet showed concerns regarding unnoticed capture (T-1). The participants have expressed worries regarding unware identitication (T-2) and lack of control (T-3) as well as the sensitivity of captured images (T-5). Moreover, participants have expressed interest in the idea of taking their permission before such a recording takes place and emphasised the need for countermeasures against unauthorised disclosure (T-6.3).

The authors in [53] have explored different aspects of lifelogging ranging from the motivation of lifelogging activities and where to place a lifelogging devices to capture the best possible lifelogs. In addition, authors have also touched upon the legal and ethical issues related to lifelogging in public places. The authors have highlighted the discrepancy of legal requirements in different countries regarding the consent for taking photographs of an individual. Then, the authors argue that even if the lifelogging may not be legally prohibited, many people would be uncomfortable with the idea of lifelogging. The authors have also pointed out the lack of control (T-3) in terms of sharing and consent and security issues (mostly T-6.1 and T-6.2) of the captured lifelogs.

Memory augmentation is an idea of aiding people to trigger recall using different mechanisms. The implications of pervasive lifelogging and data collection from different sensors on the mermory augmentation of people using different lifelogging devices have been analysed in [9]. Especially, different security and privacy issues have been explored in the context of pervasive memory augmentation. The authors have highlighted the importance of security aspects such as secure storage (T-6.1), confidentiality and integrity (T-6.2) and access control (T-6.3). The privacy aspects of bystanders with respect to pervasive sensing has also been emphasised, unfortunately, their examination fails to underscore any of our other identified threats.

The authors in [7, 8] presented two supplementary user-studies to understand the reactions of bystanders from the perspective of lifeloggers and bystanders respectively. In the first study ([7]), 40 lifeloggers have been recruited to use a lifelogging device, Autographer, for a period of 3 days (6-8 hours every day). At the end of the study, semi-structured interviews have been carried out to understand the reactions of the bystanders. 29 out of 40 lifeloggers have reported that people they lived with, i.e. either their family members or flatmates enquired about the device and around 85% of them were not in favour of using the device inside the shared flat. In addition, 18 lifelogger have reported that their colleagues in the workplace enquired about the device, and requested for not using the device during office hours. Especially, there were 9 instances when the line manager explicitly asked the subjects to avoid using the device in office settings. It has also been reported that 12 bystanders enquired about the device, and showed discontent and concern after hearing its characteristics. They politely requested the lifeloggers to delete their photographs, because they did not like to the idea of being randomly photographed. These bystanders also told that they felt as if they were being tracked. In summary, the bystanders were concerned about: (1) self-impression management and sensitivity of lifelogs (T-5), how they were being represented in the lifelogs, (2) being tracked, i.e. identified (T-1) being at a certain place engaged in certain activities; (3) how the lifelogs could be manipulated (T-2) to give a negative impression about them (or activities).

In a supplementary user study ([8]), the authors have reported the results of a user study conducted within two contexts/scenarios to understand the privacy perspectives of bystanders. In their study, two scenarios have been simulated: (i) the use of wearable cameras during a student presentation (formal meeting where a group of students are demonstrating their projects to their project supervisors) and (ii) the use of wearable cameras in a gathering among friends in the personal space of the experimenter (indoor informal meeting). In both settings, bystanders have been worried about of the sensitivity of lifelogs (T-5) as they may capture highly sensitive information during such a formal arrangement. All the subjects have reported that they tend to become anxious, and have been more conservative and serious than usual. Moreover, they have been worried if lifelogs captured and shared without their consent (T-3) might be a breach to their privacy. They have also commented that they have forgotten quickly about the presence of

Table 4: Threat mitigation in existing work

	T-1	T-2	T-3	T-4	T-5	T-6		
						T-6.1	T-6.2	T-6.3
Anita Allen [3]	–	–	–	–	–	–	–	–
O'Hara et. al. [42]	–	–	–	–	–	–	–	–
Jana et. al [27]	L	✓	L	–	✓	–	–	–
Jana et. al [26]	L	✓	L	–	✓	–	–	–
Templeman et. al. [50]	–	–	–	–	✓	–	–	–
Roesner et. al. [47]	✓	✓	✓	–	✓	–	–	–
Memon et. al. [35]	–	–	✓	–	–	–	–	–
Gurrin et. al [18]	–	✓	✓	–	✓	✓	✓	–
Johansen et. al. [28]	–	–	–	–	–	–	–	–
Iachello et. al. [23]	–	–	✓	–	–	–	–	–
Nguyen et. al. [38]	✓	✓	–	–	–	–	–	✓
Klasnja et. al. [30]	–	–	✓	–	–	–	–	–
Nguyen et. al. [39]	✓	–	✓	✓	✓	–	–	✓
Hoyle et. al. [22]	–	✓	✓	–	✓	–	–	✓
Hoyle et. al. [21]	–	–	–	–	✓	–	–	–
Denning et. al. [11]	✓	✓	✓	–	✓	–	–	✓
Wolf et. al. [53]	–	–	✓	–	–	✓	✓	–
Davies et. al. [9]	–	–	–	–	–	✓	✓	✓
Chowdhury et. al. [7]	✓	✓	–	–	✓	–	–	–
Chowdhury et. al. [8]	✓	–	✓	–	✓	–	–	–

the camera in the informal setting which may lead to the situation of unnoticed capture (T-1) of lifelogs.

6.4. Summary

In this section, we summarise the result of our analysis of the existing approaches.

At first, a summary of our findings is presented in Table 4 which exhibits the inadequacy of existing approaches in mitigating the identified threats. In addition to the symbols – and ✓, we have used the symbols *L, S,* or *B* to indicate if the approach can mitigate the corresponding threat for only a particular actor.

Next, for each of the presented research in this section, we have counted the frequency of each threat that has been analysed in the respective approach. A frequency distribution of each threat is illustrated in Figure 2. This figure again highlights the concern of the users as well as the gaps prevalent in the existing research. From Figure 2, T-3 and T-5 are the most highlighted threats whereas T-4, T-6.1 and T-6.2 are the least highlighted threats.

7. Privacy in Information Systems: requirements and guidelines as mitigation strategies

In this section, we present a set of security and privacy requirements and guidelines for a lifelogging (information) system that can be used to mitigate the identified threats. It is to be noted that mitigating these threats with the identified security and privacy requirements would also minimise the associated risks.

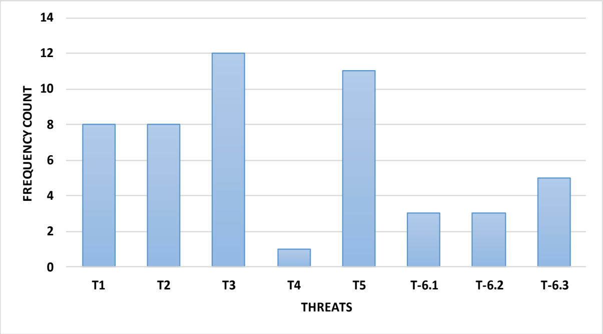

Figure 2: Threat frequency count.

To ensure the privacy of its users, an information system formulates different privacy requirements. Then, the system is designed, developed and implemented using one of the access control mechanisms in such a way that these requirements are satisfied. In [16], different privacy requirements for an ideal identity management system has been presented. Since, an identity management system is essentially an information system that shares PII and other information between different entities, we can use their privacy requirements as a basis to formulate requirements for a lifelogging system. The formulated privacy requirements are presented below:

Support for Anonym and Pseudonym (SAP). This requirement ensures that a user can interact with another entity either anonymously or pseudonymously.

User Empowerment (UE). This requirement ensures that a user is fully empowered over her data when data is shared with other entities. It consists of the following requirements.

- **Data Control (DC).** This is to ensure that a user has full control over her data and has the right to choose the entities to which her data will be released.
- **Selective Disclosure (SD).** This requirement will allow any user to choose specific PII before the information is released to another entity.
- **Explicit Consent (EC).** This requirement will enable a user to provide explicit consent before any PII is released to another entity.

Next, we analyse the adequacy of these requirements to mitigate the identified threats. To analyse the SAP requirement, it is needed to define what anonymity means with respect to lifelogging. We define anonymity with respect to lifelogging is *the state that enables any actor within a lifelog to be anonymous*. One of the mechanisms by which anonymity in lifelogging can be achieved is by identifying the face of an actor and then blurring the face in such a way that it remains unidentifiable. Enabling an information system with this capability will mitigate threat *T-2*. Interestingly, an anonymous lifelog has the provision of decreasing sensitivity even if the lifelog is captured in a private or intimate setting and hence can be used to mitigate threat *T-5*. The sub-requirements (especially *DC* and *EC*) of the *UE* requirement can be used to mitigate Threats *T-1* and *T-3* as these requirements will enable the actors to exercise data control and to express their consent for each lifelog. This, in a way, ensures that corresponding actors have accessibility over the captured lifelogs to exercise their control, thereby mitigating threat *T-4*.

None of the privacy requirements can be used to mitigate any security threat. That is why we need to formulate security requirements. Based on the security requirements of an ideal identity management system as outlined in [16], we formulate the following security requirements for a lifelogging information system.

- **Authentication & Authorisation (AA).** This is to ensure only the authenticated and authorised users in a system can access their corresponding lifelogs.
- **Secure Storage (SS).** This is to ensure that lifelogs are securely stored in the system in such a way that their integrity remains intact and they are never exposed to attackers or any authorised entities.
- **Secure Transmission (ST).** If the system requires transmission of lifelogs via the Internet, the transmission must be carried out via a secure channel so that the confidentiality and the integrity of transmitting lifelogs are guaranteed.

It is easy to analyse that the *AA* requirement can be used to mitigate threat *T-6.3*, *SS* requirement can be used to mitigate threat *T-6.1* whereas the *ST* requirement can be used to mitigate threat *T-6.2*.

7.1. Implementation Guidelines

One might assume that realising a system that satisfies these privacy and security requirements will be enough to mitigate all identified threats. However, implementing a system that satisfies the formulated, especially privacy, requirements can be challenging.

The security mechanisms can be realised using the current state of practice. For example, *AA* can be realised by utilising an identity management system which can also take care of the *ST* requirement as all existing IMSs rely on the secure transmission of data between different entities. Similarly, the *SS* requirement can be satisfied by ensuring that lifelogs are stored in encrypted format.

However, it might be extremely difficult, if not impossible, to design a lifelogging system that can collect explicit consents, during the lifelogging process, of all subjects or bystanders whom the lifelogger might never meet afterwards in her life. Similarly, how a subject or bystander can exercise control over the lifelogs in which they appear can be challenging since it would require the system to identify the actors and then contact them to transfer their right of data control. To counteract these difficulties we propose the utilisation of a visual cue or marker that will enable a bystander or subject to express their agreement or disagreement for capturing lifelogs with them in those lifelogs. Also, such a cue could also be used to encode their consent (e.g. via a QR code) using a policy of how such lifelogs should be treated as utilised in [47]. This visual marker or cue can be attached to a place on the body (preferably attached in a cloth in the upper portion of their body) in such a way that they are clearly visible in a lifelog so that the lifelog can be treated in a special way. For any sensitive place, the cue can be attached at the entrace of such place as reported in [50]. In addition, the lifelogger can also utilise a visual marker, as highlighted in [39], to annouce and alert any bystander or subject that a lifelogging session is being captured.

To facilitate the design and development of a privacy-preserving lifelogging system that meets the outlined privacy as well security requirements, we have prepared a series of guidelines which are presented below.

- A privacy-preserving lifelogging system should be equipped with an image processing capability to detect faces or sensitive places and objects in the captured lifelogs. Such capabilities have already been reported in the existing works, e.g. in [39, 26, 50, 47] with satisfactory performance. These capabilities would allow the lifelogger to offload the burden of fully manual identification of sensitive lifelogs to a computing system. In addition, such a system should offer the flexibility in providing the users with a means of defining privacy parameters for sensitive lifelogs based on different metrics, e.g., a location, or a specific event. This would allow to create a semi-automatic system where the user can adjust her privacy preferences to determine the sensitivity of lifelogs applicable for different scenarios.

- Such a system should have the capability to identify a special type of visual markers as discussed previously.
- Such a system should have the capability to decode, extract policies and interpret consents, if available. The policy itself can be either expressed using XML or JSON and then encoded using a QR code. Alternatively, it can just be a special type of visual cue such as a predefined "thumbs up" or "thumbs down" image encoding the corresponding consent of a bystander or subject.
- If a sensitive object (a face, place or an object) is detected in a lifelog:
  - the object must be blurred, unless a cue or policy indicating the consent of the corresponding object can be found. This will ensure that the *SAP* and *EC* requirements are fulfilled.
  - the lifelog must not be shareable without the respective object is blurred, unless a cue indicating the consent of the corresponding person can be found. This will ensure that the *DC*, *SD* and *EC* requirements are fulfilled.
- Such a system should ensure that captured lifelogs are stored securely in encrypted format. This will ensure that the *SS* requirement is fulfilled.
- Such a system must ensure that the captured lifelogs are exposed only to the authenticated and authorised users. This will ensure that the *AA* requirement is fulfilled.
- In case such a system requires transmission of lifelogs between different entities, it must be carried out via a secure channel. This will ensure that the *ST* requirement is fulfilled.

We believe that incorporating these guidelines into a system will be challenging and require a holistic approach considering the privacy issues of all actors.

8. Conclusion

In this paper we have provided a thorough discussion on different aspects of privacy in visual lifelogging. We have adjusted the existing definitions of lifelogging to reflect the current state of lifelogging. We have formulated a definition of privacy in lifelogging by exploring different avenues of privacy in different domains. We have presented a threat model of privacy in (visual) lifelogging which is the first of its kind. We have shown the inadequacy of existing privacy guidelines and approaches in mitigating the identified threats. Then, we have explored the ways the identified threats can be mitigated by formulating different security and privacy requirements for a lifelogging system which deals with (visual) lifelogs. Finally, we have presented a series of guidelines that can be utilised to implement such a privacy-preserving visual lifelogging system.

Our guidelines rely on the concept of utilising a visual marker that can be used to determine if a subject or lifelogger has consented to be in a lifelog or to share such lifelogs with others or to declare a place or an object as sensitive. Such a marker can encode a privacy policy that a privacy-preserving lifelogging system must enforce. One may argue how such a visual marker or cue can be designed and developed which is universally acceptable. A community driven approach will certainly generate different types of such markers which can cause inconsistencies across different systems. A better approach would be to standardise such a marker from a standard body.

Another issue is how to incentivise software developers so that they develop such a privacy-preserving system using a holistic approach. One way to handle this is to generate enough demand which will drive the incentives for software developers. However, it must be noted that lifeloggers will be the primary customer for a lifelogging system, unless it has a sharing capability. Hence, the demand for a privacy-preserving lifelogging system that mostly addresses the privacy of subjects and bystanders may not be as attractive to lifeloggers as to other actors. This may, unfortunately, decrease the demand for a privacy-preserving lifelogging system which in turn lessen the motivation of software developers in developing such a system. One solution could be to legislate rules and norms with respect to privacy of lifelogging by regulatory bodies that will essentially necessitate the creation of such a privacy-preserving system.

Being a nascent technology, it is still not clear how the lifelogging technology will be shaped and what privacy implications it will expose in future. One thing is for sure that there will be many more interesting use-cases of lifelogging apart from being a tool of personal recollection and ramification. As such, it has the potential to gain mainstream traction just like photography. If this happens, the attitude of different actors towards lifelogging may change significantly and many of these privacy concerns could be relaxed. Nevertheless, there is a need to understand the privacy implications of lifelogging, identify privacy threats for all relevant actors and then mitigate those threats. This paper aims to meet these goals and lay out the foundation for subsequent research to design and develop a privacy-preserving lifelogging system.

References

[1] Autographer. http://www.autographer.com/#home. Accessed on 15 February, 2016.

[2] Narrative Clip. http://getnarrative.com/. Accessed on 15 February, 2016.

[3] Anita L Allen. Dredging up the past: Lifelogging, memory, and surveillance. *The University of Chicago Law Review*, pages 47–74, 2008.

[4] I. Altman. *The environment and social behavior: privacy, personal space, territory, crowding*. Brooks/Cole Pub. Co., 1975.

[5] Ann Cavoukian. Privacy by Design: The 7 Foundational Principles. https://www.ipc.on.ca/images/resources/7foundationalprinciples.pdf, August 2009. Accessed on 15 January, 2016.

[6] Ann Cavoukian. Operationalizing Privacy by Design: A Guide to Implementing Strong Privacy Practices. https://www.ipc.on.ca/images/Resources/operationalizing-pbd-guide.pdf, December, 2012. Accessed on 15 January, 2016.

[7] Soumyadeb Chowdhury, Md Sadek Ferdous, and Joemon M Jose. Bystander privacy in lifelogging. In *30th British Human Computer Interaction Conference*. BCS e-WIC, 2016. http://goo.gl/jT65nE.

[8] Soumyadeb Chowdhury, Md Sadek Ferdous, and Joemon M Jose. Lifelogging user study: Bystander privacy. In *30th British Human Computer Interaction Conference*. BCS e-WIC, 2016. http://goo.gl/EshZug.

[9] Nigel Davies, Adrian Friday, Sarah Clinch, Corina Sas, Marc Langheinrich, Geoff Ward, and Albrecht Schmidt. Security and privacy implications of pervasive memory augmentation. *IEEE Pervasive Computing*, 14(1):44–53, 2015.

[10] Danny De Cock, Karel Wouters, Dries Schellekens, Dave Singelee, and Bart Preneel. Threat modelling for security tokens in web applications. In *Communications and Multimedia Security*, pages 183–193. Springer, 2005.

[11] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2377–2386. ACM, 2014.

[12] Lieven Desmet, Bart Jacobs, Frank Piessens, and Wouter Joosen. Threat modelling for web services based web applications. In *Communications and multimedia security*, pages 131–144. Springer, 2005.

[13] Martin Dodge and Rob Kitchin. " Outlines of a world coming into existence": Pervasive computing and the ethics of forgetting. *Environment and Planning B*, 34(3):431–445, 2007.

[14] Md Sadek Ferdous, Soumyadeb Chowdhury, and Joemon M Jose. Privacy Threat Model in Lifelogging. To appear in *the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp'16, 2016. ACM. http://goo.gl/klRZOK.

[15] Md. Sadek Ferdous, Gethin Norman, and Ron Poet. Mathematical modelling of identity, identity management and other related topics. In *Proceedings of the 7th International Conference on Security of Information and Networks*, SIN '14, pages 9:9–9:16, New York, NY, USA, 2014. ACM.

[16] Md Sadek Ferdous and Ron Poet. A comparative analysis of identity management systems. In *High Performance Computing and Simulation (HPCS), 2012 International Conference on*, pages 454–461. IEEE, 2012.

[17] C. Gurrin, Hyowon Lee, and J. Hayes. iForgot: A model of forgetting in robotic memories. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, pages 93–94, March 2010.

[18] Cathal Gurrin, Rami Albatal, Hideo Joho, and Kaori Ishii. A Privacy by Design Approach to Lifelogging. *Digital Enlightenment Yearbook 2014: Social Networks and Social Machines, Surveillance and Empowerment*, page 49, 2014.

[19] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. LifeLogging: Personal Big Data. *Found. Trends Inf. Retr.*, 8(1):1–125, June 2014.

[20] F. Heylighen. Web Dictionary of Cybernetics and Systems. http://pespmc1.vub.ac.be/ASC/indexASC.html. Accessed on 15 July, 2015.

[21] Roberto Hoyle, Robert Templeman, Denise Anthony, David Crandall, and Apu Kapadia. Sensitive lifelogs: A privacy analysis of photos from wearable cameras. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1645–1648. ACM, 2015.

[22] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. Privacy behaviors of lifeloggers using wearable cameras. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 571–582. ACM, 2014.

[23] Giovanni Iachello, Ian Smith, Sunny Consolvo, Mike Chen, and Gregory D Abowd. Developing privacy guidelines for social location disclosure applications and services. In *Proceedings of the 2005 symposium on Usable privacy and security*, pages 65–76. ACM, 2005.

[24] MODINIS IDM. Common Terminological Framework for Interoperable Electronic Identity Management. https://www.cosic.esat.kuleuven.be/modinis-idm/twiki/bin/view.cgi/Main/GlossaryDoc, 23 November, 2015. Accessed on 15 July, 2015.

[25] Erika McCallister, Tim Grance, and Karen Scarfone. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII) . http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf, April, 2010. Accessed on 15 November, 2016.

[26] Suman Jana, David Molnar, Alexander Moshchuk, Alan Dunn, Benjamin Livshits, Helen J Wang, and Eyal Ofek. Enabling fine-grained permissions for augmented reality applications with recognizers. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, pages 415–430, 2013.

[27] Suman Jana, Arvind Narayanan, and Vitaly Shmatikov. A scanner darkly: Protecting user privacy from perceptual applications. In *Security and Privacy (SP), 2013 IEEE Symposium on*, pages 349–363. IEEE, 2013.

[28] Håvard Johansen, Cathal Gurrin, and Dag Johansen. Towards consent-based lifelogging in sport analytic. In *MultiMedia Modeling*, pages 335–344. Springer, 2015.

[29] Geon Woo Kim, Deok Gyu Lee, Jong Wook Han, Sang Choon Kim, and Sang Wook Kim. Security framework for home network: Authentication, authorization, and security policy. In *Emerging Technologies in Knowledge Discovery and Data Mining*, pages 621–628. Springer, 2007.

[30] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. Exploring privacy concerns about personal sensing. In *International Conference on Pervasive Computing*, pages 176–183. Springer, 2009.

[31] S. Kodama, H. Akaike, and H. Kakuda. Lifelog sharing system for memory recollection. In *Consumer Electronics (GCCE), 2014 IEEE 3rd Global Conference on*, pages 81–83, Oct 2014.

[32] Monika Kuschewsky. What does the revision of the OECD Privacy Guidelines mean for businesses? https://www.cov.com/~/media/files/corporate/publications/2013/10/what_does_the_revision_of_the_oecd_privacy_guidelines_mean_for_businesses.pdf, 22 October, 2013. Accessed on 9 January, 2016.

[33] Marc Langheinrich. Privacy by design—principles of privacy-aware ubiquitous systems. In *Ubicomp 2001: Ubiquitous Computing*, pages 273–291, 2001.

[34] Matthew L. Lee and Anind K. Dey. Lifelogging Memory Appliance for People with Episodic Memory Impairment. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, UbiComp '08, pages 44–53, 2008.

[35] Mohsin Ali Memon and Jiro Tanaka. Ensuring privacy during pervasive logging by a passerby. *Journal of Information Processing*, 22(2):334–343, 2014.

[36] Adam Moore. Defining privacy. *Journal of Social Philosophy*, 39(3):411–428, 2008.

[37] Suvda Myagmar, Adam J Lee, and William Yurcik. Threat modeling as a basis for security requirements. In *Symposium on requirements engineering for information security (SREIS)*, volume 2005, pages 1–8, 2005.

[38] David H Nguyen, Alfred Kobsa, and Gillian R Hayes. An empirical investigation of concerns of everyday tracking and recording technologies. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 182–191. ACM, 2008.

[39] David H Nguyen, Gabriela Marcu, Gillian R Hayes, Khai N Truong, James Scott, Marc Langheinrich, and Christof Roduner. Encountering sensecam: personal recording technologies in everyday life. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 165–174. ACM, 2009.

[40] OECD. THE OECD PRIVACY FRAMEWORK. `http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf`, 2013. Accessed on 9 January, 2016.

[41] OECD. Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. `http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=114&InstrumentPID=312&Lang=en`, 23 September, 1980. Accessed on 9 January, 2016.

[42] Kieron O'Hara, Mischa M Tuffield, and Nigel Shadbolt. Lifelogging: Privacy and empowerment with memories for life. *Identity in the Information Society*, 1(1):155–172, 2008.

[43] D.E. O'Leary, S. Bonorris, W. Klosgen, Yew-Tuan Khaw, Hing-Yan Lee, and W. Ziarko. Some privacy issues in knowledge discovery: the oecd personal privacy guidelines. *IEEE Expert*, 10(2):48–59, Apr 1995.

[44] Yael Onn, Michael Geva, Y Druckman, A Zyssman, R Lev Timor, et al. Privacy in the Digital Environment. *Haifa Center of Law & Technology*, pages 1–12, 2005.

[45] Darhl M Pedersen. Dimensions of privacy. *Perceptual and Motor Skills*, 48(3c):1291–1297, 1979.

[46] Stefanie Pötzsch. Privacy awareness: A means to solve the privacy paradox? In *The future of identity in the information society*, pages 226–236. Springer, 2009.

[47] Franziska Roesner, David Molnar, Alexander Moshchuk, Tadayoshi Kohno, and Helen J Wang. World-driven access control for continuous sensing. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1169–1181. ACM, 2014.

[48] Ravi S Sandhu, Edward J Coyne, Hal L Feinstein, and Charles E Youman. Role-based access control models. *Computer*, (2):38–47, 1996.

[49] Abigail J Sellen and Steve Whittaker. Beyond total capture: a constructive critique of lifelogging. *Communications of the ACM*, 53(5):70–77, 2010.

[50] Robert Templeman, Mohammed Korayem, David Crandall, and Apu Kapadia. Placeavoider: Steering first-person cameras away from sensitive spaces. In *Network and Distributed System Security Symposium (NDSS)*, 2014.

[51] Samuel D Warren and Louis D Brandeis. The right to privacy. *Harvard law review*, pages 193–220, 1890.

[52] Alan F Westin. Privacy and freedom. *Atheneum, New York*, 1967.

[53] Katrin Wolf, Albrecht Schmidt, Agon Bexheti, and Marc Langheinrich. Lifelogging: You're wearing a camera? *IEEE Pervasive Computing*, 13(3):8–12, 2014.

[54] Soumyadeb Chowdhury, Philip J. McParlane, Md. Sadek Ferdous, and Joemon Jose "My Day in Review": Visually Summarising Noisy Lifelog Data. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR)*, pages 607–610. ACM, 2015.

[55] Soumyadeb Chowdhury, Md. Sadek Ferdous, and Joemon Jose A User-Study Examining Visualization of Lifelogs. In *Proceedings of the 14th International Workshop on Content-based Multimedia Indexing (CBMI2016)*. IEEE, 2016.