Purpose Limitation and Secondary Use Prevention in Large-Scale Video Surveillance Systems

Sultan, Shizra

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Purpose Limitation and Secondary Use Prevention in Large-Scale Video Surveillance Systems

Shizra Sultan

Kongens Lyngby 2021

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads, building 324,
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk

Summary (English)

Large-scale video surveillance systems (VSS) are increasingly seen as the answer to problems concerning public safety, law enforcement, and situational awareness in public places, as VSS has evolved from simple video acquisition and display systems to intelligent automated systems, capable of performing complex video analysis tasks. Video cameras are excellent multi-sensors, i.e., many different types of information can be extracted from the same video data, which when analyzed with other external data sources like different public information systems can generate a lot of useful information, which may be interpreted as personal. VSS observers are legally, socially, and morally obliged to use any piece of personal information for authorized purposes only, otherwise, it may lead to privacy violations.

In order to preserve individuals' privacy in VSS data, various data protection legislation have issued specific guidelines about the installation and operation of VSS, whenever it collects or processes any personal data. For instance, the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) require VSS owners to have a valid legal basis for its deployment. It also requires owners to state an explicit purpose for their data usage, and confirmation that video data will not be subject to secondary use, i.e., it will only be used for the consented primary purposes. Most data-protection legislation allows informed individual consent as a legal basis for recording personal data, which reduces legal uncertainty. However, due to the continuous presence of VSS, it is generally not possible to obtain consent from every individual every time a public camera records them. Therefore, VSS deployment often uses 'public interest' as a legal basis for collecting and processing all the recorded data (including personal information), as different public administrative ser-

vices and authorities use the data (broadly) for multiple purposes like public safety, traffic management, etc. Individuals do not have a right to erasure and data portability under this legal base, but they do retain a right to object in some cases. Hence, VSS data collected under 'public interest' supports different sorts of purposes that are beneficial for citizens, but individuals also have fewer rights and are often expected to trust observers with their data. However, often observers intentionally or unintentionally reuse personal information (secondary use), either by misunderstanding or exploiting the ambiguous purpose statements or by going beyond their authority to access personal information under 'public interest', causing a high number of privacy invasion incidents of voyeurism, blackmail, profiling, etc. Hence, due to the multitude of personal information that can be obtained from VSS data collected under legal base 'public interest', observers are exposed to a lot of personal information that is irrelevant to their authorized purposes. Moreover, individuals are expected to trust VSS owners to use their data for the purposes authorized under supporting legal base, which by incidents in past shows is often abused by observers. Therefore, in order to limit secondary use in VSS to preserve privacy, it needs to enforce a dynamic need-to-know view for observers according to their requirements, to reduce their exposure to irrelevant personal information available to them.

This thesis develops an access control model (ACM), and an associated prototype implementation of an access control mechanism, that enforces purpose limitation in a large-scale VSS (or other Big Data information systems and Data Lakes). Our proposed ACM is an RBAC-ABAC hybrid solution that is designed according to the large-scale infrastructure requirements and is called Attributes Enhanced Role-Based Access Control (AERBAC) model. AERBAC uses RBAC for its dynamic role-assigning simplicity in categorizing different observers and assigning them minimum default permissions per their role and then utilizes ABAC for evaluating different resource and system properties thus implementing fine-grained access. We have proposed an extended AERBAC model that ensures purpose limitation in large-scale systems by verifying the resource's 'collection purpose' with the observer's 'access purpose' to control the exposure of personal information. The implemented solution enforces the need-to-know view principle in large-scale video surveillance systems for observers by allowing them access to essential personal information based on their authorized requirements and limiting avoidable exposure to irrelevant personal information.

Summary (Danish)

Store videoovervågningssystemer (VSS) ses i stigende grad som svaret på problemer vedrørende offentlig sikkerhed, retshåndhævelse og situationsfornemmelse på offentlige steder, da VSS har udviklet sig fra enkle videoopsamlings- og displaysystemer til intelligente automatiserede systemer, der kan udføre komplekse videoanalyseopgaver. Videokameraer er fremragende multisensorer, dvs. at mange forskellige typer information kan udtrækkes fra det samme videodata, som ved analyse med andre eksterne datakilder så forskellige offentlige informationssystemer kan generere en masse nyttig information, der kan tolkes som personlig. VSS -observatører er juridisk, socialt og moralsk forpligtet til kun at bruge personlige oplysninger til autoriserede formål, ellers kan det føre til krænkelser af fortrolige oplysninger.

For at bevare personers privatliv i VSS -data har forskellige databeskyttelseslovgivninger udsendt specifikke retningslinjer for installation og drift af VSS, når det indsamler eller behandler personlige data. For eksempel kræver General Data Protection Regulation (GDPR) og California Consumer Privacy Act (CCPA) VSS -ejere at have et gyldigt retsgrundlag for dets implementering. Det kræves også, at ejere angiver et eksplicit formål med deres dataforbrug og bekræftelse af, at videodata ikke vil blive genstand for sekundær brug, dvs. at de kun vil blive brugt til den primære samtykke. De fleste databeskyttelseslovgivninger tillader informeret individuelt samtykke som retsgrundlag for registrering af personoplysninger, hvilket reducerer juridisk usikkerhed. På grund af den kontinuerlige tilstedeværelse af VSS er det imidlertid generelt ikke muligt at indhente samtykke fra hver enkelt person, hver gang et offentligt kamera optager dem. Derfor anvender VSS -implementering ofte 'offentlig interesse' som retsgrundlag for indsamling og behandling af alle de registrerede data (herun-

der personlige oplysninger), da forskellige offentlige administrative tjenester og myndigheder bruger dataene (stort set) til flere formål som offentlig sikkerhed, trafikstyring osv. Enkeltpersoner har ikke ret til sletning og dataportabilitet i henhold til dette retsgrundlag, men de beholder i visse tilfælde ret til at gøre indsigelser. Derfor understøtter VSS -data indsamlet under 'offentlig interesse' forskellige former for formål, der er gavnlige for borgerne, men enkeltpersoner har også færre rettigheder og forventes ofte at stole på observatører med deres data. Imidlertid genbruger observatører ofte forsætligt eller utilsigtet personlige oplysninger (sekundær brug), enten ved misforståelse eller udnyttelse af de tvetydige formålserklæringer eller ved at gå ud over deres autoritet til at få adgang til personlige oplysninger under 'offentlig interesse', hvilket forårsager et stort antal privatlivsinvasionhændelser af voyeurisme, afpresning, profilering osv. På grund af de mange personlige oplysninger, der kan hentes fra VSS -data indsamlet under juridisk grundlag 'offentlig interesse', udsættes observatører for en masse personlige oplysninger, der er irrelevante for deres autoriserede formål. Desuden forventes det, at enkeltpersoner stoler på VSS -ejere til at bruge deres data til de formål, der er godkendt under støttende retsgrundlag, hvilket ved hændelser i tidligere shows ofte misbruges af observatører. Derfor, for at begrænse sekundær brug i VSS for at bevare fortroligheden, er det nødvendigt at håndhæve et dynamisk behov for at kende syn for observatører i henhold til deres krav for at reducere deres eksponering for irrelevante personlige oplysninger, der er tilgængelige for dem.

Dette speciale udvikler en adgangskontrolmodel (ACM) og en tilhørende prototypeimplementering af en adgangskontrolmekanisme, der håndhæver formålsbegrænsning i en storstilet VSS (eller andre Big Data-informationssystemer og Data Lakes). Vores foreslåede ACM er en RBAC-ABAC hybrid løsning, der er designet i henhold til de store infrastrukturkrav og kaldes Attributes Enhanced Role-Based Access Control (AERBAC) model. AERBAC bruger RBAC til sin dynamiske rolle-tildelende enkelhed ved at kategorisere forskellige observatører og tildele dem minimum standardtilladelser pr. Rolle og anvender derefter ABAC til at evaluere forskellige ressource- og systemegenskaber og dermed implementere finkornet adgang. Vi har foreslået en udvidet AERBAC-model, der sikrer formålsbegrænsning i store systemer ved at verificere ressourcens 'indsamlingsformål' med observatørens 'adgangsformål' til at kontrollere eksponeringen af personlige oplysninger. Den implementerede løsning håndhæver princippet om behov for at kende visning i store videoovervågningssystemer for observatører ved at give dem adgang til vigtige personlige oplysninger baseret på deres autoriserede krav og begrænse uundgåelig eksponering for irrelevante personlige oplysninger.

Preface

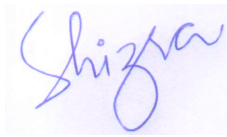
This thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark (DTU Compute) in partial fulfillment of the requirements for acquiring the degree of Doctor of Philosophy.

The thesis attempts to address the secondary usage issues in large-scale video surveillance systems by ensuring purpose limitation. The thesis is self-contained and revolves around the work done in a number of publications written during the period 2018-2021.

The PhD project has been supervised by Associate Professor Christian Damsgaard Jensen and co-supervised by Associate Professor Weizhi Meng.

The work has been supported by a grant from the Singapore Smart Nation Project.

Lyngby, 01-June-2021



Shizra Sultan

Dedication

To my beloved Parents

Thank you for your unconditional love and endless support

Thank you for always being there for me and my kids

"All that I am, or hope to be, I wholeheartedly owe it to my devoted and selfless parents"

Acknowledgements

I would like to express my sincerest gratitude to my supervisor Christian Damsgaard Jensen for his encouragement, support, and insightful feedback throughout my Ph.D. His guidance and invaluable advice helped me not just become a better researcher and academician, but a better person. I could not have imagined having a better advisor and mentor. I would also like to thank my co-supervisor Weizhi Meng for promptly answering my queries and for his invaluable input. I thank them both for providing me with the tools that I needed to choose the right direction and successfully complete my dissertation.

I would further like to acknowledge my esteemed colleagues from Section for Cyber Security for their support and collaboration: Lars R. Knudsen, Athanasios Giannetsos, Sajad Homayoun, Andreas Kidmose, Sam Afzal-Houshmand, Benjamin Larsen, Heini Bergsson Debes, Wei-Yang Chiu, Ashutosh Dwivedi, Freja Elbro, Tyge Tiessen, Tommaso Ricci, and Altug Tosun. A special thanks to our kind secretary Ann-Cathrin Dunker for her assistance with all the administrative matters.

I am highly grateful to my loving husband Ali and our beautiful kids (Musfira and Umar) for their eternal love, support, and sacrificial care, especially during the last four years that made the completion of this dissertation possible. I am forever indebted to my parents, siblings, and niece for encouraging me in all of my pursuits and inspiring me to follow my dreams. I also wish to thank my parents-in-law for their love and encouragement. Lastly, my deepest gratitude to all my friends, teachers, and mentors (Ms. Afsheen and Ms. Nabila) for providing me with a nourishing, encouraging, and supportive environment that has made me who I am today.

Contents

Chapter 1

Introduction

Large-scale Video Surveillance Systems (VSS) have become a necessary and unavoidable part of this modern era. In 2019, CNBC revealed in a study that there are nearly 770 million cameras installed globally for video surveillance and will reach 1 billion by 2021 [1]. A VSS deploys a large number of video cameras at multiple locations to monitor public places, and then the aggregated data is used by local administrations for different purposes. Traditionally, video surveillance is used as a deterrent tool to discourage wrongdoers from executing criminal or unwanted activities that can endanger public safety or damage public infrastructure in places under surveillance. Over the years, and with advances in VSS technologies, VSS can now assist local administrations in diverse areas, such as public safety, traffic management, autonomous navigation in public transportation, monitoring, smart facility management, and many more. VSS is actively used for both real-time monitoring as well as a forensic tool to accomplish different purposes and activities in the above-mentioned areas. For real-time monitoring, VSS is often used to identify and track objects of interest or interpret their behavior. For example, to ensure public safety, VSS data can be monitored to identify unattended luggage, which may contain harmful or unsolicited material, or to track a suspicious object (individual or vehicle) in public places. VSS can also be used for managing traffic, either real-time traffic management by monitoring traffic flows to handle traffic congestion, registering traffic violations, etc., or traffic planning by recording traffic patterns and transport times to improve infrastructure development. Thus, it assists local

administration or responsible entities to take preemptive measures to handle any incident in due time to protect against property theft, unwanted incidents, and other administrative tasks [2]. Moreover, VSS data can be used as a forensic tool or piece of evidence, to show proof of an occurred incident/event that happened at a certain time or location by a certain individual. For example, VSS data can be used for the identification of an intruder on restricted premises or to help law enforcement authorities identify suspects in larger cases, such as the London bombings in 2005, where authorities identified attackers with the help of VSS data and confirmed their identities by tracking their activities for the past several days that linked them to that incident [3]. Furthermore, VSS data can assist municipal authorities with periodic analysis of physical infrastructures (building, bridge, road, etc.) to monitor their degradation over a certain period of time [4]. As can be seen from all the examples presented above, VSS is an effective tool used for different deterrent, preemptive, and reformative purposes and has become a necessity for local administrations all around the world.

Most of the above-mentioned administrative services require continuous monitoring of public places because the information that the observers (a monitoring human, a program, or a system) are interested in is rarely predefined. No one usually knows the exact form that an unwanted incident may take or when a particular event will occur, for instance, a traffic-monitoring observer (TM) does not know when and where a vehicle will be speeding or a law-enforcement authority (LEA) observer cannot know beforehand if and when two individuals will start a fight. Therefore, to record a particular activity when it happens, and then use that specific piece of data according to the particular requirements of the observer, VSS data needs to be collected and analyzed continuously. This huge amount of VSS data aggregated at a large scale from various public locations contains a lot of general information about individuals, their routine activities, human associates, or frequently visiting places, etc., which may be interpreted as personal. Any piece of information describing any physical, psychological, ethnic, social, or biometric indicator that can uniquely or potentially identify an individual is considered personal information [7]. Thus, VSS collects huge amounts of data from public places that may contain different types of personal information about individuals, and VSS observers are legally, socially, and morally obliged to use that personal information strictly for their authorized purposes, otherwise, it may lead to privacy violations [8].

In order to preserve individuals' privacy in VSS data, various data protection legislation have issued specific guidelines about the installation and operation of VSS, whenever a VSS collects or processes any personal data [9]. In addition to this specific legislation, general data protection legislation, like the General Data Protection Regulations (GDPR) and California Consumer Privacy Act (CCPA), require VSS owners (either public or private) to have a valid legal basis for its deployment [10] [11]. It also requires owners to state an explicit

purpose for their data usage, and confirmation that video data will not be subject to secondary use, i.e., it will only be used for the consented primary purpose. Most data-protection legislation allows informed individual consent as a legal basis for recording personal data, which reduces legal uncertainty. However, due to the continuous presence of VSS, it is generally not possible to obtain consent from every individual every time a public camera records them. Therefore, VSS deployment often refers to 'public interest' as a legal basis, as different public administrative services and authorities use the data (broadly) for multiple purposes like public safety, traffic management, etc. Individuals do not have a right to erasure and data portability under this legal base, but they do retain a right to object in some cases. Hence, VSS data collected under 'public interest' supports different sorts of purposes that are beneficial for citizens, but they have fewer rights and are often expected to trust observers with their data. However, often observers intentionally or unintentionally reuse personal information, either by misunderstanding or exploiting the ambiguous purpose statements or by going beyond their authority to access personal information under 'public interest', causing a high number of privacy invasion incidents of voyeurism, blackmail, profiling, etc. [12] [13]. Hence, due to the multitude of personal information that can be obtained from VSS data collected under legal base 'public interest', individuals are expected to trust VSS owners to use their data for the consented/authorized purpose, which by incidents in past shows is often abused by observers'.

To summarize, large-scale VSS offers valuable information to observers with diverse requirements, and for authorized purposes, observers are allowed to access personal information contained in VSS data. This makes it imperative that the VSS ensures that observers are strictly limited to data pertinent to their specific stated purposes. Due to the nature of video data, however, it is often hard to limit the exposure of personal information, which may ultimately lead to privacy violations. To ensure privacy in large-scale VSS, the following concerns must be explicitly addressed. Firstly, who collects and owns the data contributed to VSS, secondly, what kind of information can be extracted from it, thirdly, who has the authority to choose observers and designate their access rights, and lastly, how can it be ensured that observers will use it according to their authorized purposes. In the next section, we address these concerns by looking at different aspects and entities of a VSS with the help of a general model of VSS and conclude it by highlighting different factors that are to be considered by VSS to preserve privacy.

1.1 Large-Scale Video Surveillance Systems Model

Here, we use a general model to describe how traditional VSS works. It has four basic modules: Capture, Transport, Monitoring, and Storage; the model is shown in Fig. 1.1. In the capture module, depending upon the purpose of surveillance, one or multiple cameras are deployed at different locations to cover the desired area. Each camera is assigned a unique identifier in order to be recognized and generates a recording (video stream/file) of occurred activities in a designated location for a specific time period. In the Transport module, these recordings are transported to the monitoring room or sent directly to the storage servers for archiving. In the Monitoring module, humans or (semi)automated systems observe the live streams/ recordings for different objects or events of interest. The fourth module is called storage, which holds all the recorded or archived video files.

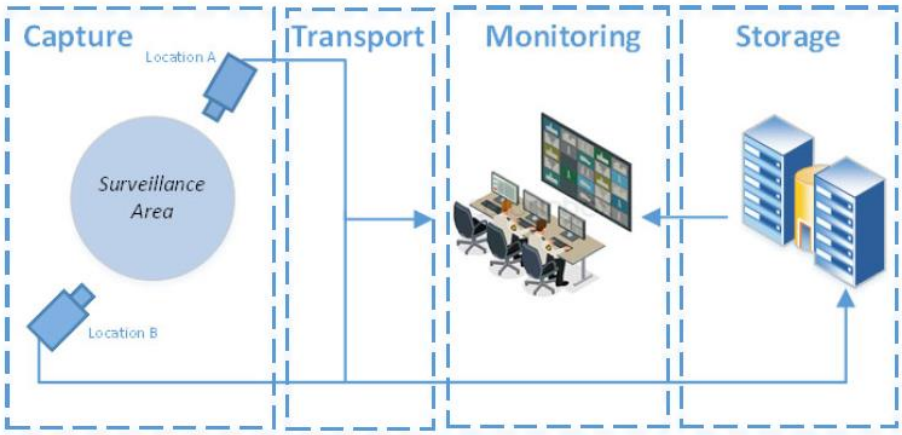


Figure 1.1: VSS traditional model

Each recording is of a particular duration and is archived so it can be retrieved later. The archived files and live streams can be requested from the monitoring room, where authorized observers (humans or systems) can view them. The monitoring room can be either a specific place where different stream/recordings can be viewed or it can be distributed to different rooms/observation points (e.g., hand-held devices)[14]. The recordings captured by VSS are called surveillance data or VSS data and a lot of information can be extracted from its content, i.e., different types of objects, their activities (events), physical location, time of the day, nearby objects, and landmarks, etc. Generally, the surveillance data is accessible to the VSS owners, and these can further assign access rights to different authorized observers. The recordings can also be given as input to an

intelligent system for automatic alert generation based on pre-defined events.

The above-described model can be extended to any scale given the right hardware and software infrastructure. Conventionally, the VSS is centralized, and data from all the cameras is analyzed at the central processing and storage server. However, nowadays, in large-scale and real-time environments, many VSS is now deployed as a distributed network architecture, i.e., different functional units perform independent computational operations on given/assigned data or task and either take a decision upon it or forward it to the central node for further processing. For instance, a large-scale VSS (like in a smart city) is built on the traditional model and is deployed in a large environment, managed by the local administration. Thousands of video cameras are deployed all around the city at various public places roads, parks, stations, town squares, etc. for several purposes. These cameras (Capture module) are then connected via different wired and wireless networks (Transport module) with several distributed storage and computational servers (storage), and different observers can access any of the cameras live (Monitoring module) or an archived recording (via Storage), from anywhere given they are connected to the system network [15]. This data is further processed and analyzed with other data sources available to different observers so they can extract relevant information from it. Information collection and sharing at such a large scale is highly beneficial for many observes, however, it also raises serious privacy concerns because of the personal nature of the information at disposal, Thus, in order to preserve the privacy of individuals recorded in VSS data without rendering its usefulness, it is important to answer the below-mentioned questions:

1. What is collected or recorded as VSS data and what type of information can be obtained from it? Can this information be categorized?

2. Which entities are involved in the VSS data cycle (collect, share, access), and who has the authority to regulate its access? How are the entities (observers) expected to use VSS data, once they have access to it? Discussed in

3. Is privacy a challenge in VSS, considering the information extracted from it, and different entities that require access to that information?

We will address each of these questions below.

1.1.1 VSS Data and recorded information

Data recorded and processed by a VSS is called surveillance data and is the most critical asset of any VSS. It comprises of video recordings (live and archived) that capture the activities at the installed location. All cameras in the VSS also have contextual (configuration) data that helps view the recorded video in a certain perspective or helps when data is being processed. This data includes the location of the camera, camera type, lens resolution, time-stamp, video encoding scheme, etc., and it is also considered part of the VSS data. Lastly, large-scale VSS may analyze surveillance data with other external data sources to generate new information. These sources can be coupled sensors, such as microphones, GPS, infrared, etc. that may add value or context to the existing video recordings. Alternatively, these sources can also input data in different formats that can be aggregated to enrich existing information. These sources can be the national citizen database, or vehicle registration database, or real-time traffic data, etc. For instance, if a person is observed in a video breaking a law, law-enforcement authorities can identify that person by comparing its face (extracted from video recording), with the records in the national citizen database. Hence, information from various external sources can be combined with VSS data, which can then be correlated with each other to generate useful information.

All the above-mentioned VSS data with different sources generate a lot of information, which in the context of our thesis, i.e., protection of personal information in VSS; can be categorized into two main categories: Personal information and non-personal information. As mentioned above, any piece of information whether from the content of the video recording or the supporting information that may evidently or potentially linked to a unique individual is personal information. While the information that cannot be used to identify a unique individual is non-personal. Personal data can be further divided into different categories such as biometric features (faces, gait), descriptive features (gender, estimated age or height, ethnicity, etc.), activities (driving, fighting), and associations (belongings, frequently visited places, or persons), etc. as per the observers' usage requirement of that data and will be discussed in detail in Chapter 4.

1.1.2 VSS Entities and information access

In any distributed information-sharing infrastructure such as VSS, many entities are involved in different modules that assist in capturing, transporting, storing, and presenting data. However, when such infrastructures process personal in-

formation, then involved entities can be broadly categorized into three groups: Data Subjects (DS), Data Controllers (DC), and Data Processors (DP).

1.1.2.1 Data Subjects

Any individual or a natural person whose identity can be bound to a unique identifier (in data) is called a Data Subject (DS). In VSS data or recordings, any camera installed at a public location records DS and different information associated with them, thus, recording personal information about an individual. An individual's right to protect and willingness to share its personal information is referred to as privacy and is regarded as a human right almost everywhere in the world. The use of data containing any piece of personal information (without the consent or implicit knowledge of the DS) is generally considered a privacy breach. Though large-scale VSS cannot obtain the explicit consent of every DS being recorded every time yet general consent of the public is assumed under the legal base of public interest that VSS data will be used for purposes related to the public interest, ensuring an acceptable level of privacy. The rights of DS under different legal bases are discussed in Section 2.1.5.

1.1.2.2 Data Controllers (Owners)

A data controller (DC) or owner is an entity that has a legal authority to collect, store and process data, and may delegate its access to other observers or data processors (of choice). To ensure privacy, data protection legislations around the world require DC to specify explicit purposes for why it is necessary to collect certain personal information from individuals and how will it serve its authorized requirements, also referred to as 'collection purpose'. This means that the collected personal information can only be used for reasons and ways agreed upon in the 'collection purpose' between the DC and the DS. It is easier to observe such 'collection purpose' in centralized infrastructures managed by one DC, however, in distributed infrastructures, where multiple DCs are involved and manage their set of observers, it is challenging to enforce 'collection purpose' for all the requesting observers. There are frequent data transformations and aggregations at such a large scale, and often 'collection purpose' are misinterpreted or miscommunicated or not preserved appropriately, leaving a gap for biased interpretation by the observers [16]. This leads to secondary data use causing privacy violations, thus, it is critical for DC to make sure that 'collection purpose' are preserved and observed properly. Otherwise, any other way of data usage no matter how benign or beneficial is secondary use and is considered a privacy breach.

In large-scale VSS, many DC are collecting data at different locations for different purposes and can be categorized into four groups as shown in a Surveillance Area v/s Ownership matrix in Fig. 1.2. There are several possibilities based on the nature of the surveillance area (private or public) and ownership of the surveillance area. Generally, Public DC deploy VSS in public areas to monitor public safety events under legal base 'public interest', while private owners deploy VSS in privately-owned spaces for property protection under legal base 'legitimate interest'. 'Legitimate interest' requires private owners to demonstrate the need for a valid and real-existence issue like property protection or preservation of evidence (of unlawful intrusion) for VSS deployment. Some locations are considered semi-public, such as banks or shopping malls, which though record public activities, but owned by private-owners, so their supporting legal base is also 'legitimate interest' and not 'public interest'. Individuals (recorded in data) under legal base 'legitimate interest' have more rights over the use of their recorded information than they have under legal base 'public interest'. VSS deployed in private areas usually do not have public ownership, unless it is requested by public authorities to fulfill a legal obligation, for instance, if it recorded a crime and can be used as a piece of evidence. Moreover, private owners have the flexibility to chose what to install inside their private property, while public owners are regulated by domestic surveillance policies. Therefore, the common three categories are:

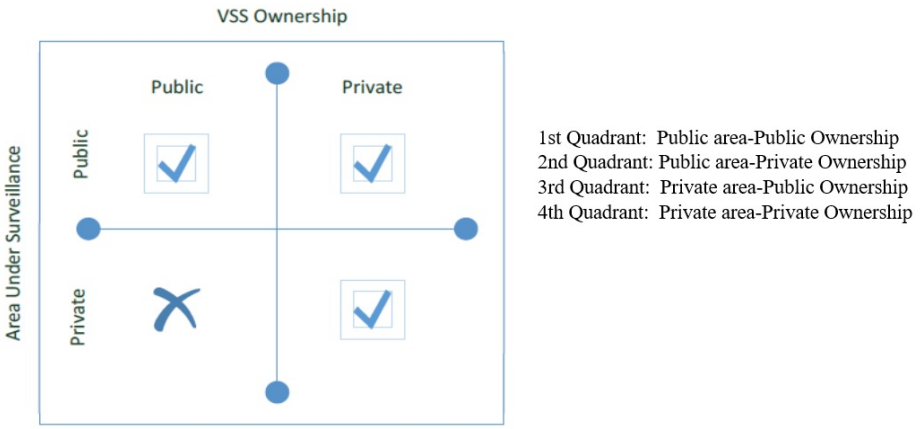


Figure 1.2: Surveillance Area v/s Ownership matrix

1. Public-Public Ownership: Public places such as parks, streets, bus stops, railway stations, airports, public offices, hospitals, etc. are generally under surveillance by public DC, i.e., national or local governments, and are managed in compliance with local/regional surveillance policies.

2. Public-Private Ownership: public area surveillance under private ownership like privately-owned spaces concerning public areas such as banks, hotels, shops, malls, etc. They generally comply with local surveillance guidelines, as they are recording citizens at a semi-public place.

3. Private-Private Ownership: individuals have deployed cameras inside their houses or private offices/property, to have proof if there was an outside intrusion. Recording from this surveillance is only of interest to a private DC, as it does not concern any other entity. Yet, they are still legally and morally obliged to not use cameras at their private premises for spying or voyeurism.

In the case of VSS, most countries have a law or a stipulation under different data protection legislation that requires DC to inform DS that they are under video surveillance or being recorded (e.g., through a display sign). It is a legal and social responsibility of the DC that it (or anyone associated with it) does not abuse any information recorded at its authorized space. On the other hand, it is also important that the DC have the right to secure its place/location with VSS so it can mitigate unwanted activities. DS under surveillance can also misuse this information (that they are being recorded). DS with malicious intent can escape those places that have cameras to avoid being noticed. Alternatively, they can also hide their identities/ faces so even if their activities are being recorded no video evidence can connect their identity with that activity. There are serious concerns for both DC and DS as both can misuse this information, but it is also necessary to protect the interests of one stakeholder without violating the (privacy) rights of the other.

For the scope of this thesis, we are focusing on public-public DC ownership and the data they collect and share at a large scale, which is further requested by different observers or data processors for diverse purposes.

1.1.2.3 Data Processors (Observers)

A data processor (DP) is an entity that is authorized by a DC to access data for achieving a task with a specific agreed-upon purpose (here referred to as 'access purpose'). DC designates skilled observers or DP with different roles/ responsibilities that they need to perform and authorizes them to request and access different types of information from VSS data manually or via some application. Traditionally, a VSS DC (public or private) delegated a specific set of DP with access to VSS data for a particular 'access purpose' and does not share or use the data collected by some other DC. However, with an increase in large-scale information-sharing infrastructures like smart cities, data from different DC is

aggregated, processed, and shared in an integrated manner. The collective data is available to a large number of observers authorized and managed by several DC, which can (ideally) retrieve information relevant to their assigned 'access purposes', as shown in Fig. 1.3

Traditional VSS, deployed to ensure public safety, uses the data to monitor events that are directly related to public safety and this data is not available to DP who may require it for other purposes. However, in modern large-scale VSS, different local administration departments (DC) such as law enforcement authorities, various emergency services, infrastructure and planning departments, etc. can benefit from the integrated data and request information that is relevant or helpful to their business operations. For instance, as mentioned in an earlier example, a traffic-monitoring observer (TM) is interested in information relevant to traffic events in VSS data, while a law-enforcement authority observer (LEA) is interested in public safety events, and both can now use the same video data for their authorized purposes. While the VSS data is the same, each observer is concerned with a specific type of information relevant to their tasks. If a DP has access to VSS data that is more than its requirements then it has a legal, social, and moral responsibility to not use spare data in any way that invades the privacy of individuals recorded in that data.

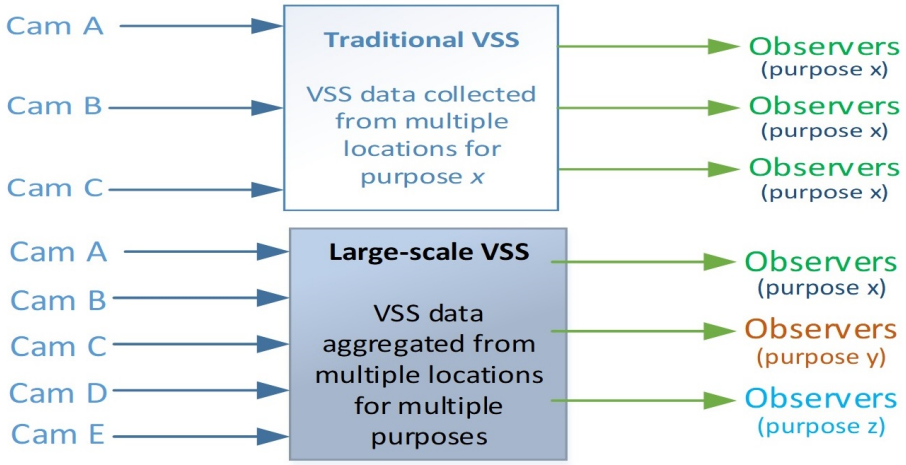


Figure 1.3: Traditional and SC-VSS model comparison

Ideally, VSS should ensure that DP is only allowed to view data strictly relevant to their authorized tasks or 'access purpose', which should be compatible with data's 'collection purpose' in order to ensure purpose limitation. However, in large-scale distributed infrastructures, often the DC to which the requested

resource belongs does not regulate all the requester DPs registered in large-scale infrastructure. In cases such as this, an individual DC's 'access purposes' may be designed or defined differently by another DC that may authorize them to request resources, thus allowing them access to data that might be incompatible in terms of 'collection purposes'. Thus, if both the 'collection purpose' and the 'access purpose' are inconsistent, and do not follow similar formats or characteristics, they may end up being incompatible with each other, challenging the purpose limitation principle. Therefore, both the 'collection purpose' and the 'access purpose' must be defined/designed consistently so they can be verified against each other to ensure purpose limitation.

1.1.3 VSS Privacy

The purpose of any information system is to make stored or processed data available to its DP, but if the data contains personal information, then it requires to be protected from unauthorized access by applying appropriate privacy-enhancing technologies (PETs). VSS is an example of such an information system, as it records data at public places that contain a large amount of personal data about individuals and their daily life activities. Even though VSS data is (ideally) to be used for specific purposes, there are a lot of examples and privacy invasion incidents, all around the world where VSS data has been used for other than presumed purposes (by both private and public) DC and DP [12] [17] [13]. Therefore, it is critical that VSS data is protected appropriately to be used for specific 'collection purposes' and is only available to observers for authorized 'access purposes' as directed by several data protection legislation [10] [11]. This can be achieved by applying suitable privacy-enhancing technologies(PETs)in VSS. Different PETs have been proposed over the years to preserve privacy in information systems based on the concepts of encryption, anonymization, data minimization, data obfuscation, access control, etc. PETs limit the amount of information reaching the observer. If strictly applied, these solutions result in the loss of information, but applied more loosely, the likelihood of privacy invasions increases. Therefore, it is crucial to determine which PET is most suitable to ensure privacy in a given system or application, without preventing authorized observers from accessing relevant information [18].It will be addressed in detail in Chapter 3

1.2 Individual's right to privacy and Observer's authorized purpose

To sum up the above discussion, a large-scale VSS is a distributed information-sharing infrastructure where a massive amount of data, contributed by multiple sources (DC) is aggregated and processed to extract useful information for observers (DP) with different requirements. As the data contains a lot of personal information, it is essential to balance the right to privacy of individuals (DS) captured in the recordings or VSS data against the authorized purpose of the observers (DP) so they have timely access to the required information. As discussed in section 1.1.3, various PETS can be applied to the VSS data at different points until it reaches the observer (DP) to ensure authorized access, and to retain data's usefulness for observers with diverse requirements without compromising its integrity, it is essential to have a privacy-aware access control mechanism (ACM).

To preserve individuals' privacy, it is imperative to limit observers' view of VSS data or specifically, personal information present in VSS data. There are two commonly used approaches to achieve this, first, to restrict observers' view of VSS data based on physical or contextual parameters of time and space, and is commonly used due to ease of implementation [23]. For example, an observer can view VSS data if it is recorded within a mile radius of its 'location', or if the 'time-stamp' of the requested camera-recording lies within its 'duty hours'. In this case, observers are allowed to view all the data recorded within their allowed physical parameters irrespective of whether they need to view more or less VSS data to complete their tasks. Here, observers are exposed to personal information limited by authorized physical parameters, and observers are expected to not use personal data for other than what is authorized. The second approach is to limit the observer's view based on the high-level semantic information obtained from the content of the recorded data such as specific objects or activities. For instance, a traffic-management observer can be authorized to view data from all the highway cameras that detect vehicles to register traffic incidents or violations. In this case, observers' access to data or personal information is limited by authorized requirements based on semantic content. It allows different observers to view a restricted portion of a recording based on different types of semantic information identified from the VSS content, making it more relevant than just physical parameters of time and space. However, in the context of large-scale VSS, observers are still exposed to irrelevant personal information even within the allowed duration of recording, which can lead to the possibility of secondary use or misuse i.e., reuse personal data by association[27]. Thus, to preserve privacy, observers should have limited access to VSS data, however, commonly adopted methods of limiting access based only on physical parameters

or semantic content are not enough, especially where the observers are expected to have minimal to zero possibility of personal data reuse. Thus, an optimal privacy-aware ACM also needs to consider the exposure of personal information required by the observer along with other physical or semantic parameters while making an access decision.

Thus, for a large-scale VSS to limit secondary use, it needs to enforce a dynamic need-to-know view for observers with diverse requirements, and there are three main factors to consider. First, for VSS to limit data misuse or reuse, it needs to understand the content of the VSS data so it can differentiate and categorize different types of personal information in it, in order to protect it or monitor its usage. Second, whenever data contains personal information, there must be an explicit or implied 'collection purpose' associated with it that ensures that the agreed-upon terms of usage limitations are being observed. The 'collection purpose' is an agreement between the DC and the DS about how can their personal information be used. Third, an observer or DP should have an authorized 'access purpose' that clarifies its requirements of personal data in order to achieve a particular task. The 'access purpose' is an agreement between the DC and the DP about how can it use the data (or personal information in it) and essentially contributes to its task completion. Therefore, to achieve a dynamic need-to-know view to ensure a privacy-aware VSS, the observer should only be allowed to view the personal information in VSS data/recording limited by the successful verification outcome between the observer's 'access purpose' and VSS data's 'collection purpose', also referred to as purpose limitation.

The above solution is based upon an assumption that the 'collection purpose' of the data and the 'access purpose' of the observer are compatible. This holds to be true in traditional centralized infrastructures as the DC collecting the personal information are the ones responsible for documenting the relevant 'collection purposes' (recorded and maintained separately from data) and then regulate DP's access by authorizing 'access purpose' accordingly [24]. However, data in large-scale integrated infrastructures are contributed by multiple DCs, which are altered, transformed, and aggregated many times according to the requirements of observers with different authorized 'access purposes' [21]. This raises certain concerns when it comes to ensuring purpose limitations in the distributed and information-sharing environment like VSS. First, due to frequent changes in the structure and content of the data, often the 'collection purposes' of the data are lost, misinterpreted, or not preserved appropriately, leaving a gap for biased interpretation [32]. Second, DCs may not have control over all the data/resource transformations in a distributed or shared environment, and often it is hard to constantly authorize 'access purposes' for a DP requesting data in different forms and accommodate their emerging requirements [6]. Third, often the DC to which the requested resource belongs does not regulate all the requester DPs registered in large-scale infrastructure. In cases such as this, an individ-

ual DC's 'access purposes' may be designed or defined differently by another DC that may authorize them to request resources, thus allowing them access to data that might be incompatible in terms of 'collection purposes'. Thus, if both the 'collection purpose' and the 'access purpose' are inconsistent, and do not follow similar formats or characteristics, they are often incompatible with each other, and based on lenient or stricter policies this may either allow secondary use or prohibit authorized use [25]. To conclude, large-scale integrated infrastructures often fail to ensure purpose limitation due to unsuccessful verification between the different 'collection purposes' of resources and DP's 'access purpose', because of inconsistent definitions by different DCs. Current state-of-the-art solutions designed for large-scale infrastructures lack a comprehensive solution that addresses the mentioned problems simultaneously.

In order to address the above-mentioned concerns, we present two arguments; first, that the representation of purpose (both collection and access purpose) should follow some standard format so they can be verified against each other as per the requirements of the applied data protection guidelines to ensure purpose limitation. Second, a key requirement in purpose limitation is purpose integrity, or more specifically 'collection purpose' integrity, and thus this should be preserved. Purpose limitation ensures that resource usage is strictly governed by its 'collection purpose'; however, in the case of frequent transformations and aggregations, it is often not well-preserved, thus disregarding the purpose limitation principle. Thus, it is important to preserve the integrity of the 'collection purpose', i.e., ensure that this is exactly the same as agreed upon between the DC and DS, and second, that it is readily available to DPs in its conserved state whenever the data or resource is requested by an authorized DP.

1.3 Thesis Contribution

To address the above-mentioned issues, in this thesis, we aim to present a large-scale privacy-aware VSS that limits secondary use by ensuring purpose limitation. Secondary Use in VSS and Purpose limitation in large-scale VSS are both comprehensive domains, so we will first address these two issues separately, and then combine them to achieve an integrated solution to ensure privacy in VSS. To address the issue about Secondary Use in VSS data, we have analyzed VSS data in detail, and categorized (personal) information under different classifications, so it can be used to describe data or usage limitations, as discussed in Chapter 4. To address the issue of purpose limitation, we have proposed a framework for representing, storing, and aggregating the 'collection purpose' of a resource as per commonly observed data-protection guidelines, and demonstrated how an 'access purpose' can be verified against it to ensure purpose

limitation, discussed later in Chapter 2. Moreover, we have further proposed to add the said 'collection purpose' as an immutable data property (provenance), to preserve its integrity through different data/resource transformations. The provenance is a resource (metadata) property that catalogs different activities that are performed on a resource along with its lineage and are often immutable and append-only. The provenance will initially record the 'collection purpose' along with resource origin, and then with every transformation or aggregation, the 'collection purpose' will be preserved and appended (updated) if required, ensuring purpose integrity. We concluded our thesis by presenting an integrated solution that will limit the possibility of prejudiced interpretation and enforce fine-grained need-to-know permissions in VSS to limit secondary use, while allowing multiple observers to achieve their authorized purposes, without compromising individuals' or DS privacy, discussed in Chapter 5. Moreover, it will also ensure compliance with different data protection legislation by ensuring purpose limitation and preserving purpose integrity that not only limits secondary usage but also builds up trust among DCs, DPs, and DSs in regards to resource usage transparency. Our thesis contributions have been published as two conference proceedings and two journal articles, and excerpts from them have been used in following chapters where relevant and are properly referenced. The published articles are as mentioned below:

1. We have identified different privacy concerns in large-scale VSS and proposed measures to encounter them. The findings are published in a conference proceeding titled "Privacy-preserving measures in smart city video surveillance systems", Proceedings of the 6th International Conference on Information Systems Security and Privacy [22].

2. We have proposed a framework for representing, storing, and aggregating the 'collection purpose' of a resource as per commonly observed data-protection guidelines to prevent secondary use in large-scale information systems. The paper is published as a conference proceeding titled "Secondary Use Prevention in Large-Scale Data Lakes", Proceedings of the 2021 Computing Conference, Springer Nature [26].

3. We have developed an access control solution based on the identification and categorization of different types of personal information in large-scale VSS data, that can be verified against observers' requirements to enforce a need-to-know view in order to limit secondary use. The work is published as a journal article titled "Metadata-based need-to-know view in large-scale video surveillance systems" in Computers and Security. 2021, Vol.111 [27].

4. We have proposed and implemented an Integrated access control solution for preserving privacy in large-scale infrastructures in order to ensure pur-

pose integrity by enforcing purpose limitation through provenance metadata. The paper is published as a journal article titled "Ensuring Purpose Limitation in Large-Scale Infrastructures with Provenance-Enabled Access Control" in Sensors, 21, 3041 [28].

1.4 Thesis Overview

Chapter 2 discusses purpose limitation in detail and describes different properties that can help define 'collection purpose' in large-scale distributed infrastructures. It further presents how provenance can be used to preserve it, which can then be used in access control mechanisms to enforce purpose limitation.

Chapter 3 confers the privacy requirements of large-scale video surveillance systems and discusses smart-city VSS as a case study to enumerate access requirements of various smart-city observers.

Chapter 4 provides a detailed analysis of different types of information that can be extracted from VSS (meta)data. It further looks into how that extracted information can be categorized into different categories and top-down hierarchies to be further used in regulating access to various observers.

Chapter 5 presents our proposed access control solutions based on metadata in VSS data (as discussed in Chapter 4)can be used to limit secondary use by ensuring purpose limitation (as discussed in Chapter 2). Moreover, it discusses our proposed framework i.e., Extended Attribute-Enhanced Role-based Access Control (AERBAC) model and how its policies can be specified using XACML.

Chapter 6 covers state-of-the-art work relevant to how data categorizations and data hierarchies are used in regulating access control mechanisms in large-scale infrastructures. It also summarizes some of the commonly-used privacy preserving solutions to restrict access in video surveillance systems.

Chapter 7 considers some of the key future work directions and concludes the thesis.

Chapter 2

Purpose Limitation in Large-scale Integrated Infrastructures

Large-scale integrated infrastructures are the key to the smart world, as they collect and integrate data from a large number of data sources (IoT networks, integrated data lakes (DL), social media websites, public information systems, geographical information, and many more), transform it, and make it available for a large number of users or data processors with diverse requirements. One of the recent and most prevalent examples of such infrastructures is smart cities, which have been adopted all around the world. Many of the data sources collect and process personal information about individuals via different sources, and a large number of data applications and services use it to provide personalized and informed services back to individuals. Any data or resource that may contain personal information is ideally collected for a certain purpose, i.e., the 'collection purpose', which is understood and respected by all the involved entities, such as the Data Controllers (DC), Data Processors (DP), and the Data Subjects (DS). However, due to the data-driven economy, data collection, transformation, and analysis is a continuous and persistent process in large-scale integrated infrastructures, which causes data to repeatedly change forms. Often data is collected by a DC under certain circumstances at some point in time and is then processed or transformed under different conditions by several different DP per

their authorized requirements or 'access purposes', thus altering the form of data. Such data aggregation and transformations may hide the original 'collection purpose' of the involved resources, which, if not preserved, can potentially lead to secondary use. Data used in any way other than the 'collection purposes' is considered a privacy violation, which according to data protection legislation, like GDPR or CCPA, may lead to legal ramifications [10] [11]. Hence, to prevent secondary use of data, a system or DC needs to preserve the 'collection purpose' of the resources (with personal information), because it will otherwise render the usefulness of data, transforming into a data swamp with lots of valuable data yet missing the information that is required to legally/ethically use it [29].

Generally, the DC collecting data (containing personal information) is also responsible for documenting the associated 'collection purposes', usually recorded separately from the resource, as a part of an internal document like privacy policies, data protection and retention policies, or data sharing agreements, etc. DC then authorizes various DP about the applicable 'collection purposes', i.e., assigns a set of permissions for different resources per 'collection purpose' against a set of DP's requirements or 'access purposes'. However, in large-scale integrated infrastructure, often data is contributed by multiple DC, which are aggregated and analyzed together to generate new resources. Similarly, the DPs may also be managed by different DC, and may request any transformed resource from the integrated infrastructure for their own 'access purpose'. These DC may have different ways to design/ define 'collection purposes' for owned or managed resources, or may not be comfortable sharing their internal documents or policies to notify DPs with resource's 'collection purposes', which leaves room for biased interpretation of 'collection purposes' leading to secondary use. Moreover, due to the dynamic nature and growth of integrated infrastructures, where new DC and DP may become part of it anytime, a DC rarely has sufficient knowledge about data requirements or 'access purpose' of all the existing DP (managed by all the existing DC) and any new DC and DP that may emerge over time. It will again lead to DC constantly authorizing and updating emerging DC with the 'collection purposes' of the requested resources, which is infeasible and may prohibit authorized DP's access to the resource, if not done timely and may fail to ensure that a resource is only used according to its 'collection purpose' [16]. Therefore, if rather than recording 'collection purpose' separately from the resource, as a part of an internal or shared document, It can be designed as a generic structure according to data-protection guidelines, and further recorded and preserved as part of the resource (as part of its metadata) through different data transformations and aggregations, then it will become easy for the DC to notify DP about applicable 'collection purposes'. The preserved 'collection purpose' also shows purpose integrity when compared with the 'access purpose' of the DP to ensure purpose limitation. This chapter presents a framework for representing, storing, and aggregating the 'collection purpose' of a resource as per commonly observed data-protection guidelines, and demonstrates how an

'access purpose' can be verified against this to ensure purpose limitation. Please note that this chapter may contain excerpts and figures from our own published papers (as part of the PhD research) mentioned and referenced in 1.3.

2.1 Representation of a Collection Purpose

P3P defines purpose as "the reason(s) for data collection and use". Any data or resource that may contain personal information is (ideally) collected for a certain purpose, i.e., 'collection purpose', which records the terms of the agreement in which personal information enclosed in data is to be used by an authorized DP. The DC with explicit (in some cases implicit) agreement with the relevant DS initially constructs the 'collection purpose', which can later be modified as data transforms or aggregates with other sources [28]. One of the recent legislation GDPR states the notion of purpose and its limitation in the below-mentioned articles or stipulations, as shown in Table 2.1 [10].

The purpose limitation principle states that 'collection purpose' and 'access purpose' must be defined in a compatible way so they can be verified against each other. The data minimization article states that any data collected should be relevant and specific for the usage purpose, thus, to achieve this it is important to understand data and its properties so that personal properties can be distinguished. The Legal base article states that there must be a supporting legal base for every 'collection purpose', and the rights of DS should be observed accordingly. Thus, per reviewing the basic requirements of different data-protection legislation and obligations of DC to instruct DP about data usage, we propose that 'collection purpose' should have the following characteristics defining it, as shown in Fig. 2.1.

2.1.1 Data Description (Structure and Properties)

Every resource has a data description or data definition, which defines the structure of the data resource without revealing the actual values, similar to a class or type definition. It shows a set of data properties that represents the different types of information, whose type and structure can vary based on the data format, stored in the content of the data. Some of them may store or represent personal information (entity attributes), i.e., information that can be linked to a unique natural individual. For instance, a structured resource holding financial transaction has distinct data properties such as account name and number, depositor, receiver, location, date, or time of when a transaction was made,

Reason	Article Info	Description
Data minimization	[9, Article 5, x1(c)]	[Personal data shall be] adequate, relevant, and limited to what is necessary for relation to the purposes for which they are processed [...]
Legal Base: Consent	[9, Recital (32)]	Consent should be given by a clear affirmative act establishing a freely given, specific, informed, and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her [...] Consent should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them.
Right to be forgotten	[9, Article 17, x1]	[...] the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies: the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise.
Access control	[9, Article 25, x1]	The controller shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data, which are necessary for each specific purpose of the processing, are processed. [...] personal data are not made accessible without the individual's intervention to an indefinite number of natural persons."

Table 2.1: GDPR articles supporting purpose limitation

etc. describing personal information. Alternatively, an unstructured resource, i.e., image or video may represent data identified as different objects (humans, vehicles, buildings) that can be linked directly to a DS. Here, we suggest that DC explicitly mentions the data properties that contain personal information in the said resource, so it indicates that these data properties need to be handled cautiously, as agreed upon.

2.1.2 Purpose-Property Matching

A resource can have more than one purpose and each purpose may refer to a different subset of personal data properties. For example, if the resource is a healthcare record, it can have multiple personal data properties (attributes)

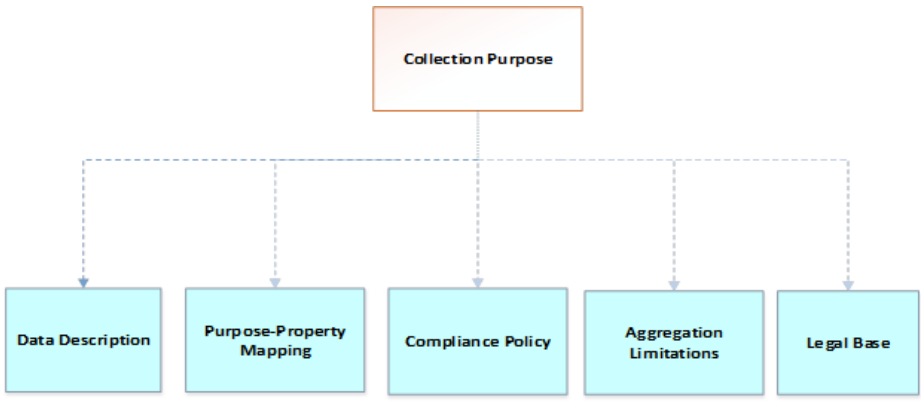


Figure 2.1: Collection Purpose Properties

such as patient-name, patient-social security number, age, gender, disease, diagnosis, doctor's notes, prescription, etc. There can be different purposes that are bound to a different number or combination of these data properties. For example, if the purpose is 'doctor appointment', then it only requires (name, address), or if the purpose is 'research' then it may only require (age, gender, disease), etc. Hence, not every purpose requires all personal data properties to be revealed, so it is a reasonable approach to bind or map agreed-upon purposes against a specific or required set of personal data properties. Ideally, to ensure data minimization DC should define a purpose for all possible combinations of personal data properties (entity attributes) subsets, but often could be left undefined for future use, or may need to be redefined or remapped against different data transformations or aggregations with other data sources.

2.1.3 Compliance Policy

Once a purpose is matched to a specific subset of properties, it can be further mapped to a DP with some specific attribute(s), an activity, or a combination of (DP and activity), indicating its usage pattern to describe a compliance policy. For example, a purpose 'research' bound to a subset (age, gender, disease) can be further mapped to an activity (read-only) or a DP (Researcher) to indicate that only a certain activity can be performed for the given purpose on bound properties subset. There can be more than one attribute or roles/permissions or some specific requirements that may be required to describe a DP or (activity), and thus can be mentioned here. For instance, in a large-scale video surveillance system, where video recording is a resource, one of the purposes can be to 'detect traffic violation', which is mapped to a subset of entity-attributes

(vehicle-license-plate, vehicle's owner name, social security number). A DP with certain requirements (attributes) of a 'traffic officer' can be mapped to the purpose 'detect traffic violation' that will allow its to access properties subset (vehicle-license-plate, vehicle's owner name, social security number). A purpose can also be bound to an activity, i.e., 'traffic officer' can only view (activity) the personal data properties and not edit, modify, or aggregate (activity) it with another resource. Compliance policy is the key characteristic of 'collection purpose' as it can be helpful to assist information systems with access control decisions to limit secondary use. It can be described in terms of conditions, permissions, usage policies, context parameters, or assertions about how a resource should or should not be used, etc. depending upon the nature of the resource, DC requirements, and DS preferences.

2.1.4 Aggregation Limitations

If a resource can create new or enhance existing personal information (data properties), when aggregated with another resource with some particular data properties, then it should also be recorded here. Different resources may have their subset of properties which when combined generate new personal information requiring new 'collection purposes' to be defined so the aggregated result can be accessed. For example, a dataset with information about facial identities if aggregated with video surveillance data can be used to track the activities of the individual. These two datasets have different types of personal information, the former containing (name, social security number, age, facial mapping), while the latter containing (objects (individuals), associated actions), which when combined reveal current and contextual personal information about the identified individuals. Thus, any (activity) that requires more than one resource, or complements the resource in question, should be mentioned explicitly, so the exposure to new and more revealing (data properties) can be managed. Moreover, if the aggregations or their limitations are not mentioned here, then in case of any new or undefined aggregation, the DC and DS responsible for the existing personal information should be notified, so appropriate measures can be taken to avoid any privacy breach or legal ramification.

The aggregation limitations can further be bound to different DP, i.e., whether a particular agent is allowed the specific aggregation (activity) on a given resource (entities), or if there are rules to limit how particular agents can use the aggregated data properties (entities and activities). For example, a DP such as a "police officer' may be allowed to aggregate resources (video surveillance data with facial recognition database), while a DP such as a "traffic monitoring officer' may not. Alternatively, in the case of aggregation, if the target resource also has its own set of 'collection purposes', how will they be combined? Or,

a DP requiring access to the aggregated subset must have the authorization to access the source and target subset individually, i.e., any DP requiring access to (data properties) of the aggregated resource may require to present DL with an authorization for either the combined purposes or distinct authorization for both purposes separately. Hence, describing how different resources can be aggregated with the given resources is important to control the exposure of newly generated information.

2.1.5 Legal Base

The 'legal base' is the foundation for the lawful (personal) data processing required by different data protection legislation. It means that whenever DC collects and processes personal data for whatsoever 'purpose/s', there should be specific legal grounds to support it. A legal base is a set of different laws/rules that grants DS rights about how their personal information should be managed. The legal base also binds the DS, DC, and DP, with their respective rights and obligations towards each other. Some examples of valid legal bases supported by the GDPR are Consent (explicit permission to use data), Contract (formal contract to which the DS is a party). Legitimate Interest (often followed by consent or contract, in which the DC already has the data), Public Interest (processing data in an official capacity for the public interest), Legal Obligation (Data processing complies with law (local, federal, global)), and Vital Interest (Data processing in order to save someone's life) [30]. Not all legal bases grant the same rights to DS, and differ in situation, as mentioned in Fig. 2.2.

Legal Base/ Right To	Consent	Contract	Legitimate Interest	Public Interest	Vital Interest	Legal Obligation
Access	✔	✔	✔	✔	✔	✔
Erasure	✔	✔	✔	×	✔	×
Withdraw	✔	×	×	✔	×	×
Object	×	×	✔	✔	✔	×
Informed	✔	✔	✔	✔	✔	×
Portability	✔	✔	×	×	×	×
Human Intervention	✔	✔	✔	✔	✔	✔
Restrict Handling	✔	✔	×	×	×	×

Figure 2.2: Legal Base and DS Rights

2.1.5.1 Rights

DS is entitled to a set of rights over their personal information, to which the DC is legally obligated to comply. These are the right to be informed, right

of access, erasure, withdrawal, object, and rectification, and restrict processing, portability, and human intervention, as shown in Fig. 2.3

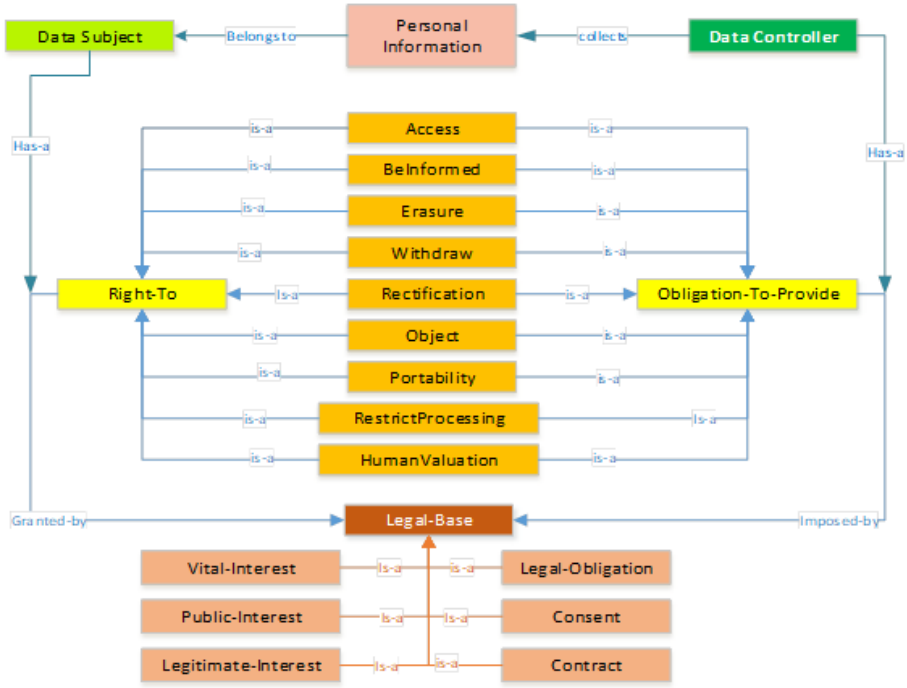


Figure 2.3: Legal Base, Rights and Obligations

2.1.5.2 Obligation

The rights of DS are obligations for DC, as shown in Fig. 2.3. Primarily, the obligation falls on the DC, however, DP is obligated to DC to process such requests from DS and abide by their rights. The DC and DP must always anticipate which rights might be applicable for DS considering their data-protection legislation and supported legal base. In case joint-controllers or multiple DC manages resources with personal information, and the DS wants to exercise one of its rights, the obligation falls onto the DC who collected data from the DS in the first place, except if agreed otherwise. The legal base must be a part of a 'collection purpose' property, as it plays a significant role in deciding what DS is entitled to, and even when DC is not directly regulating DP, this property can help DP (agent) to process due to requests from DS over their personal information. Moreover, it is common for one resource to have different legal bases for different purposes or activities.

To conclude, a 'collection purpose' has five main characteristics: first is "data-description" (identification of resource attributes/identifiers that store personal information). Second, "purpose-property matching" adheres to both data minimization and purpose limitation by binding identified personal properties to specific and explicit functions, so that they cannot be used otherwise (i.e., mapping purpose functions with the required set of identifiers). Third, "compliance policy" and "aggregation limitations" specify the conditions and limitations upon how the purpose-property matching functions can be used, further describing the conditions that a DP needs to fulfill in order to access specific data properties for an explicit function following an access control article. The last characteristic records the "legal base", as it is crucial to decide what rights a DS can execute over personal information.

2.2 Recording and Preservation of a Collection Purpose

In large-scale and distributed infrastructures, the significance of data often increases with different transformations, forming new linkages and correlations, making it valuable for various DP. Data undergoes different transformations, expands and takes different forms, can exist in multiple forms for different purposes, etc. Hence, once data changes multiple folds against different requests, the 'collection purpose' is often lost, overlooked, or loosely tracked by DC, which may be misunderstood by DP leading to secondary use [29]. Therefore, we propose to record and preserve the 'collection purpose' as a part of data provenance, (i.e., resource metadata), so it can be traced and regarded even when transformed, and DC can always track its usage [31].

To store and organize different types of data/resources contributed by different DC, large-scale distributed infrastructures usually have some primitive metadata schema that identifies the basic structure or nature of the data content without going into granular details. Often these schemas also store information about data or resource lineage, i.e., tracking different activities or processes data goes through from its origin to consumption by different DP. It also stores information about who collects and owns the data, how long it should be stored for, how different transformations are cataloged, etc., and is often referred to as Provenance [32]. Typically, metadata of any resource can be modified at any point during the data life cycle, though, provenance metadata is often considered immutable and append-only and requires systems to efficiently manage and preserve it through different transformations [33]. Moreover, when data from different DC is aggregated, it is anticipated that their provenances will also be integrated or stitched in a logical manner. Hence, provenance is a use-

ful way to record and catalog different changes in a data life cycle, especially in large-scale shared infrastructures where DC does not have complete control over data transformations [34]. Data provenance mainly represents a lineage of different activities that data goes through and the information about involved entities since it is first inducted in a DL to ensure that data is reliable [35]. Provenance helps DP understand the type of data that the resource contains, along with the ownership and retention information about the resource. The provenance originally has three key concepts: Agent, Activity, and Entity, as shown in Fig. 2.4 [36]. The activity is an action or any type of processing activity that creates, modifies, or deletes an entity. An entity is a data resource or object in any format, i.e., structured (tables), unstructured (pictures, videos), or semi-structured (files, social media feeds).

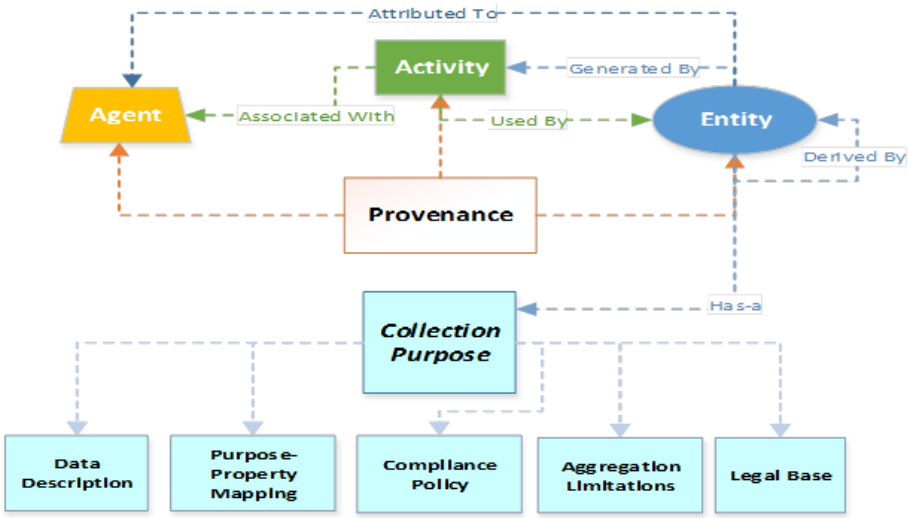


Figure 2.4: Open Provenance Model OPM

Lastly, the agent here is a DC or DP, which initiates or triggers an activity to be performed on the entities [35]. In the case where different resources are aggregated together, their provenances can (ideally) be stitched together to create a deep provenance trace, so no record is missing about how the entity was used by the agents [32]. Put simply, the provenance keeps records of which agent performed a particular activity over a given entity. Here, we propose to add 'collection purpose' as a fourth key concept in provenance, in order to also record the reason why an agent is allowed to perform a certain activity over an entity that contains personal information. The 'collection purpose' records the terms of the agreement in which personal information enclosed in data (entity) is to be used (activity) by an authorized DP (agent). The DC with explicit

(in some cases implicit) agreement with the relevant DS initially constructs the 'collection purpose', which can later be modified as data transforms or aggregates with other sources. The 'collection purpose' has five main properties as discussed in Section 3.1. The first is Data description (identifying resource attributes/identifiers that store personal information), then Purpose-property matching (mapping purpose-functions with the required set of identifiers), compliance policy (conditions under which earlier mapped functions can be used), then aggregation limitation (how the functions cannot be used), and lastly legal base supporting this 'collection purpose'. A resource can have one or more than one 'collection purposes', where each 'collection purpose' may have a different set of functions specific to different agents or activities.

Here, we will use a simple example to show how provenance is described using PROV-O [37]. A DP (traffic-law observer TLO) has a task to generate/publish a set of vehicle-owners with traffic violations by comparing the vehicle (license plates) detected in video surveillance data with another database that is vehicle registration data, to identify vehicle owners in order to issue a fine. Hence, it requests the system to perform aggregation of two data sources, traffic-violation video data, and vehicle registration database to create a new data source vehicle-owner identification dataset. Below is the PROV-O code for this example:

```
*traffic-law observer = TLO
@prefix xsd:  <http://www.w3.org/2001/XMLSchema#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix prov: <http://www.w3.org/ns/prov#>.
@prefix :     <http://example.org#> .

:VehicleOwner-identifcation_set

   prov:Entity;
   prov:wasGeneratedBy  :publicationActivity;
   prov:wasDerivedFrom  :aggregatedByAttribute; ##
      Attribute is license-plate
   prov:wasAttributedTo :TLO;
.

:TLO
   a foaf:Person, prov:Agent;
   foaf:givenName       "TLO";
   prov:actedOnBehalfOf :TrafficMangDept;
.

Traffic-Mang-Dept
   a foaf:Organization, prov:Agent;
```

Terms	Description
prov:Entity	A physical or digital concept with definite characteristics
prov:Activity	An action that occurs over a period of time, and includes activities such as recording, storing, processing, transforming, accessing, modifying, relocating, etc.
prov:Agent	One who is responsible for initiating, participating, or terminating an activity
prov:startedAtTime prov: endedAtTime	Both show the start and end temporal properties of an activity.
prov: used	An entity used by an agent for an activity
prov:wasGeneratedBy	agent generated an Entity for an activity
prov:wasInformedBy	Activity informed or generated an alert for another activity (assists in creating provenance chain)
prov:wasDerivedFrom	An entity is transformed into a new entity
prov:actedOnBehalfOf	An agent acting on behalf of another agent or activity
prov:wasAssociatedWith prov:wasAttributedTo	Properties of an agent or activity referenced by another agent or activity
prov:value	it is an optional attribute that provides a representation of an entity and may occur only once in a set of attribute-value pairs.
prov:plan	A plan is an entity that represents a set of actions or steps intended by one or more agents to achieve some goals.
Identifier	Two entities (resp. activities, agents) are equal if they have the same identifier.
prov:wasInformedBy	Communication is the exchange of an entity by two activities, one activity using the entity generated by the other.
Prov:wasQuotedFrom	A quotation is the repeat of (some or all of) an entity, such as text or image, by someone who may or may not be its original author. A quotation is a particular case of derivation.

Table 2.2: PROV-O commonly used terms and their descriptions

```
   foaf:name "DK␣Traffic␣Management␣Department";
.

: publicationActivity
   a prov:Activity;
   prov:used                :aggregatedByAttribute;
   prov:wasAssociatedWith   :TLO;
   prov:wasInformedBy       :aggregationActivity;
.
```

```
: aggregatedByAttribute
   a prov:Entity;
   prov:wasGeneratedBy   :aggregationActivity;
   prov:wasAttributedTo  :TLO;
.

:aggregatedByAttribute
   a prov:Activity;
   prov:startedAtTime     "2020-01-01T07:00:00Z"^^
      xsd:dateTime;
   prov:wasAssociatedWith  :TLO;
   prov:used                :TraffViolationVidData;
   prov:used                :nationalVehicleReg;
   prov:endedAtTime        "2020-06-30T07:00:00Z"^^
      xsd:dateTime;
.

: TraffViolationVidData
   a prov:Entity;
   prov:wasAttributedTo  :local-municipality;
.
:local-municipality a foaf: Organization, prov:Agent.

:nationalVehicleReg
   a prov:Entity;
   prov:wasAttributedTo  :vehicle-registration-
      department;
.
: vehicle-registration-department a foaf:Organization,
    prov:Agent
```

The above provenance code describes that the agent :TLO is associated with two activities: : publicationActivity and :aggregationActivity. The activity :aggregationActivity take two entities as input :TraffViolationVidData (vehicles (license plates) detected in video surveillance data) and :nationalVehicleReg (vehicle registration database), and generated a new entity, :aggregatedByAttribute that aggregates the vehicles (license-plates) in :TraffViolationVidData according to the license-plates in :nationalVehicleReg to identify owners. The :aggregatedByAttribute entity is then used by the :publicationActivity activity, to generate a new entity :VehicleOwneridentifcationSet that shows a table or structured records of license-plates against their registered owners. It also shows two nested activities as :publicationActivity is informed by the activ-

ity :aggregationActivity, and as the TLO is associated with (or authorized to) the with the activities :aggregationActivity and :publicationActivity, the newly generated entity :VehicleOwneridentifcationSet is also attributed to TLO. Moreover, TLO is a role or authorized user/agent performing this activity on behalf of TrafficMangDept. Now, we propose to add 'collection purpose' to this example. PROV-O does not have a 'collection purpose' as a characteristic, so we will use an existing characteristic prov:value and prov:wasQuotedFrom to represent or store 'collection purpose'. There are two motivation factors for proposing that 'collection purpose' become part of provenance, first, it should become an inherent part of the resource (entity), so even if it is transformed or aggregated, the entity retains it, second, it should be an immutable property so it is preserved in its original state. Below are the parts of code from above that will have to be modified in order to add 'collection purpose'.

```
:VehicleOwneridentifcationSet

   prov:Entity;
   prov:wasGeneratedBy  :publicationActivity;
   prov:wasDerivedFrom  :aggregatedByAttribute; ##
      Attribute is license-plate
   prov:wasAttributedTo :TLO;

:aggregatedByAttribute
 a prov:Activity;
 prov:startedAtTime    "2020-01-01T07:00:00Z"^^
    xsd:dateTime;
 prov:wasAssociatedWith :TLO;
 prov:used              :TraffViolationVidData;
 prov:used              :nationalVehicleReg;
 prov:endedAtTime      "2020-06-30T07:00:00Z"^^
    xsd:dateTime;
 prov:value "1-Collection␣Purpose:␣Traffic␣Violation␣
    Detection␣␣2-␣␣Collection␣Purpose:␣Traffic␣
    Violation␣Assessment␣";
prov:wasQuotedFrom < http://Link to Traffic Violation
   Detection description>;
prov:wasQuotedFrom <http://Link to Traffic Violation
   Assessment description>;

.

: TraffViolationVidData
 a prov:Entity;
 prov:value "Collection␣Purpose:␣Traffic␣Violation␣
```

```
    Detection";
 prov:wasQuotedFrom <http://Link to Traffic Violation
    Detection description>;
 prov:wasAttributedTo :local-municipality;
.
:local-municipality a foaf: Organization, prov:Agent.

:nationalVehicleReg
 a prov:Entity;
prov:value "Collection␣Purpose:␣Traffic␣Violation␣
   Assessment";
prov:wasQuotedFrom <http://Link to Traffic Violation
   Assessment description>;
 prov:wasAttributedTo :vehicle-registration-department
    ;
.
```

The above-modified code adds two attributes prov:value and prov:wasQuotedFrom, to both the entities involved in activity :aggregationActivity. Each entity can have only one prov:value, and here we have used it to store 'collection purpose'. The other attribute prov:wasQuotedFrom stores the link to the associated 'collection purpose'. This meets our first motivation factor to make it an inherent part of an entity. The new entity VehicleOwneridentifcationSet thus can inherit the values the prov:value and prov:wasQuotedFrom from the involved entities. This meets our second motivation factor, that in case of transformation or aggregation the new entity retains the 'collection purpose' of its source or parent entities. It is important to note here that 'collection purposes' are not a functional part of provenance, i.e, will not be altered or transformed and will be quoted to the newly generated resource exactly as received. The initial consideration to record 'collection purpose' as part of the provenance in an unaltered way was to preserve its integrity and availability to the DP's requesting a particular resource. In order for DP to request a certain resource, it sends a request to the system's access control module that takes into account its 'access purpose' and compares it with the preserved 'collection purpose' of the requested resource to decide whether DP should be allowed access and if yes, then how much. This helps in limiting the secondary use of data and reduces privacy violations. In the next subsection, we will extend the above-mentioned example and analyze how 'collection purposes' in an aggregated resource can be used.

2.2.1 Aggregation of Collection Purposes

It is a common occurrence for DP in shared infrastructures to request an aggregation of different resources to generate a new entity that holds information from all the parent resources. Technically, these resources are independent of each other so they will have independent or separate provenances. Once the resources are aggregated, then their provenances are also (preferably) aggregated or stitched together. Newly aggregated or transformed data, contains personal information that requires a valid or declared 'collection purpose' to be requested or accessed by any agent (DP) with an authorized 'access purpose'. Therefore, if different resources and their provenances are stitched, their 'collection purposes' should also be (ideally) aggregated and preserved [38]. Here, we will extend the example presented in the main section in a larger context to analyze if the newly generated or aggregated resource be used for inherited 'collection purpose'.

Let us consider an example of a large-scale or smart-city traffic management system, where an authorized DP (traffic law-enforcement observer TLO) requests an aggregation of three authorized resources, i.e., video surveillance data (entity A), vehicle registration data (entity B), and vehicle sensors data (entity C), as shown in Fig. 2.5. In order to prevent secondary use and a potential breach of confidence, any DP (or in this example an agent TLO) requesting the aggregated resource must have authorization or an access purpose that is compatible with the collection purposes of all the parent resources implicitly, unless otherwise explicitly specified for the new resource. The explicit 'collection purposes' are designed and followed when the aggregation or transformation is expected at some point during the resource life cycle. In this case, the 'collection purpose' of one or different parent resources can record aggregation conditions, which can specify whether a particular agent is allowed to perform an aggregation (activity) on any given resource or a set of resources (entities), or if there are limitations for particular agents, entities, or activities regarding certain transformations. However, often resources in large-scale infrastructures may not have explicitly defined collection purposes for every transformation, so they can inherit 'collection purposes' from their parent resources if allowed. Therefore, we suggest deriving an implicit 'collection purpose' from the provenance of the parent resources. The implicit collection purpose of the aggregated entityABC is a UNION set of the collection purposes of all the parent entities A, B, and C. However, in order for an agent TLO to access the entityABC, its access purpose must be a subset of the intersection set of all the parent entities A, B, and C. This way the aggregated resource may have a larger set of implicit collection purposes, however for a DP or an agent to access the aggregated resource, it must have either an explicit 'access purpose' that is predefined or a common set of all parent entities, which in this case is "traffic law enforcement".

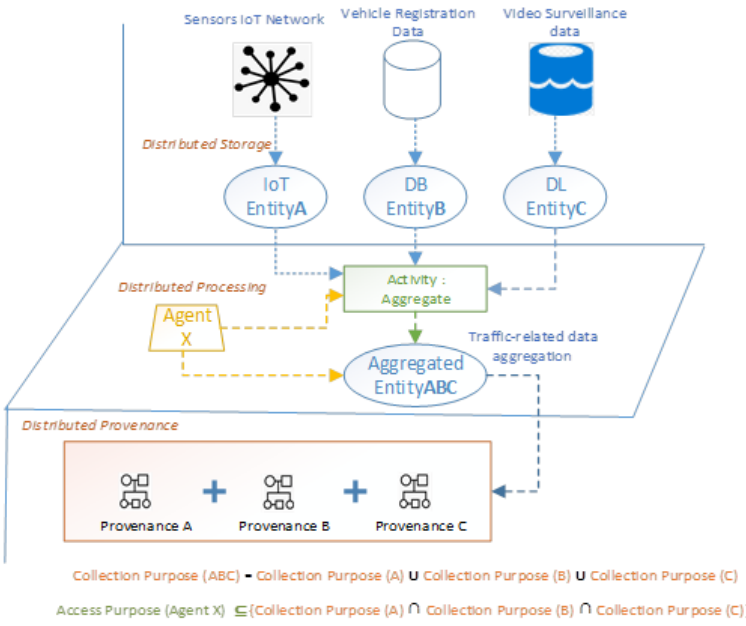


Figure 2.5: Preservation of 'collection purpose' in aggregated provenance

It is important to note here that generating an implicit 'collection purpose' is only suitable if the system has pre-defined rules for managing personal data. For instance, a new resource whose parent resources have the same DC or are collected under similar legal bases, or a resource with the same set of DS, can have an aggregation of collection purposes and thus an implicit collection purpose can be derived. An implicit 'collection purpose' should only be derived if the explicit aggregation conditions are either not mentioned or parent resources have allowed the derivation of an implicit 'collection purpose'. For example, if one of the parent's 'collection purpose' is supported by legal base public interest, while the other parent's 'collection purpose' is supported by informed consent, then the latter DC of the aggregated entity must acquire the consent of the DS (for the aggregation) in order to make the resource accessible to the DP, if not explicitly stated otherwise. If the parent entities' 'collection purposes' are supported by public interest or legal obligation, then the DC does not require explicit consent from the data subjects (DS), if it has the authorization to access parent resources (one or many) for the given 'collection purpose/s'.

Once, the transformed or aggregated resource/entity has a designated 'collection purpose' (either implicit or explicit), the next step is to verify it against the

'access purpose' of the agent or DP to check if it can be allowed access to the resource or not for the given 'collection purpose'. In order to do that, an access control module (ACM) of the system needs to take into account both the 'collection purpose' of the resource (entity) and the 'access purpose' of the DP (agent) while making an access control decision, as discussed in the next section.

2.3 Collection Purpose, Provenance, and Access control Mechanism

Secondary use in large-scale distributed infrastructures is concerned with how a DP is using a particular resource, and it is critical to limit secondary use if the resource contains personal information, so that resource privacy can be preserved. An access control mechanism (ACM) helps in regulating access to those resources (per defined policies) to different users (DP). The ACM takes into accounts the properties of users and resources and authorizes actions over the resources in any given system. For every resource, there is a defined resource policy, a set of conditions, which the DP must meet to obtain access to a resource. Moreover, once the user is authorized to access a certain resource, the 'collection purpose' of that resource is followed to limit the exposure of the resource to any DP according to the shared agreement between the data controller (DC) and data subjects (DS). Hence, privacy here is defined as "using an authorized resource solely for agreed-upon 'collection purpose/s'. Here, we will briefly differentiate between the 'collection purpose' of the resource and the 'access purpose' of DP. The former describes the terms of the agreement between DS and DC, while the latter describes the term of resource usage between DC and DP. Privacy-preserving ACMs need to ensure that DP's authorization or 'access purpose' complies with the resource's 'collection purpose' while making an access control decision as it limits secondary use and ensures purpose limitation. Moreover, it is also important that the integrity of the resource's 'collection purpose' is preserved throughout its life scale, cycle in distributed systems, in order to preserve privacy by complying with actual agreed-upon purposes with less possibility of misinterpretation and to achieve that we proposed to record resource's 'collection purpose' as part of its provenance, as discussed in the above section.

DC is the key entity with authority over 'why' and 'how' the resource is processed. It defines the 'collection purpose' for which the data is processed and may further bind different subsets of distinct responsibilities or the 'access purpose' to one or many DP. The DC records the 'collection purpose' as a provenance property and has different characteristics that will define how this data can

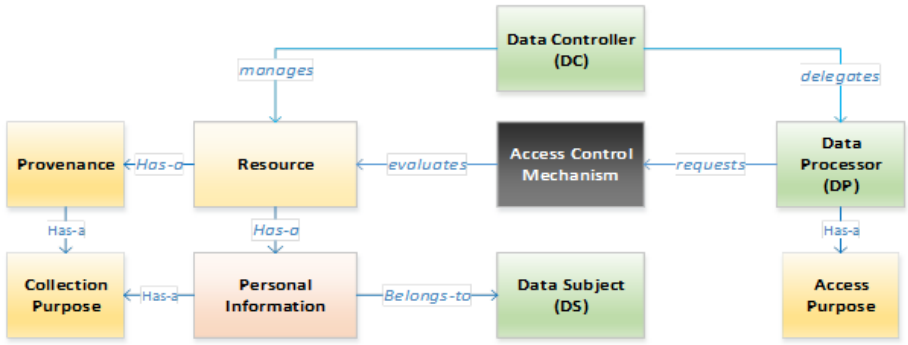


Figure 2.6: Purposes and Access Control Mechanism

be used, as discussed in Section 3.1, such as data properties, purpose-property mapping, aggregation limitation, compliance policy, etc. These resources are then introduced in distributed systems, where DPs managed by different DC can request access to these resources. Here, we assume that even though a DC does not delegate access to all the existing or future DP, every DP will be authorized by one of the associated and verified DC of the distributed infrastructure, this DC has the authority to define and authorize the basic DP requirements or their 'access purpose'. Yet, DPs may have different dynamic and contextual factors affecting their data requirements, along with many cross-data transformations [39]. Therefore, it is often hard to impose direct and static permissions especially when DP is not aware of usage limitations or 'collection purpose' of the transformed data, which increases the probability of secondary use. To cater to that, if the resource has some information that can validate how can it be used for a certain 'collection purpose' and the DP has authorized requirements to use the said type of resources for the same 'access purpose' (or a subset of it), then it can act as indirect permission from DC. Hence, we propose that the ACM can use the provenance of the requested resource to extract the 'collection purpose' of the resource and evaluate it against the 'access purpose' of the DP, which ensures both the availability and integrity of 'collection purpose'.

Any authorized DP with a set of approved requirements (ideally, a subset of the 'collection purpose'), therefore, can be allowed to access the allowed resource or its personal information per given 'access purpose', as shown in Fig. 2.6. The next section presents a case study describing how the presented approach can be used to enforce an access control mechanism that uses provenance to record 'collection purpose'.

2.3.1 Evaluation: Smart City Integrated Infrastructure

Let us take an example of a smart city integrated infrastructure, as shown in Fig. 2.7. A smart city infrastructure represents a distributed processing and storage infrastructure, which accumulates data/resources from multiple (public-authority) DC and then offers it to thousands of DP in form of different services or applications [40]. The DL stores data in different formats from various resources such as Video Surveillance Systems (unstructured), public transportation data (traffic signals, parking enforcement sensors), weather monitoring data, open navigation data (semi-structured), public administrative databases (structured), etc. [41]. Many of these resources contain personal information, which is collected under different legal bases and for various 'collection purposes'. The typical smart city has thousands of DP, such as traffic law-enforcement systems, infrastructure and planning department, law enforcement officers, emergency services personnel, etc., who all need different types of information from the above-mentioned resources per their authorized requirements or 'access purposes' [40].

In order to access a certain resource or an aggregation of more than one resources, the DP sends requests to the ACM, which then evaluates their 'access purposes' against the 'collection purpose' of the resource/s (as one of the access control parameters) and if verified, the DP is granted access to that resource. It is important to note here that there are a lot of different types of user, resource, and contextual attributes involved in access control decisions in large-scale infrastructures [42]. However, here in this section, we aim to only show how 'collection purpose' as a provenance property can influence access control decisions. There are various DC that are responsible for contributing and managing different resources in a typical smart city, as shown in Fig. 2.7. For instance, the DC-A is responsible for resource video surveillance data, and it has three 'collection purposes': public safety, traffic operation management, and infrastructure management [43]. There are other DC, which are collecting different resources for different 'collection purposes', but they can also have similar purposes such as resource vehicle- registration data also has traffic-operation management as one of its 'collection purposes'. We will discuss traffic-operation management in detail and Table 2.3 shows the defined characteristics for the given 'collection purpose', as discussed in Section 2.1. The described 'collection purpose' will become part of the resource provenance, and will be appended/updated if any transformation or aggregation modifies the content of the resources, as discussed in Section 2.2.1.

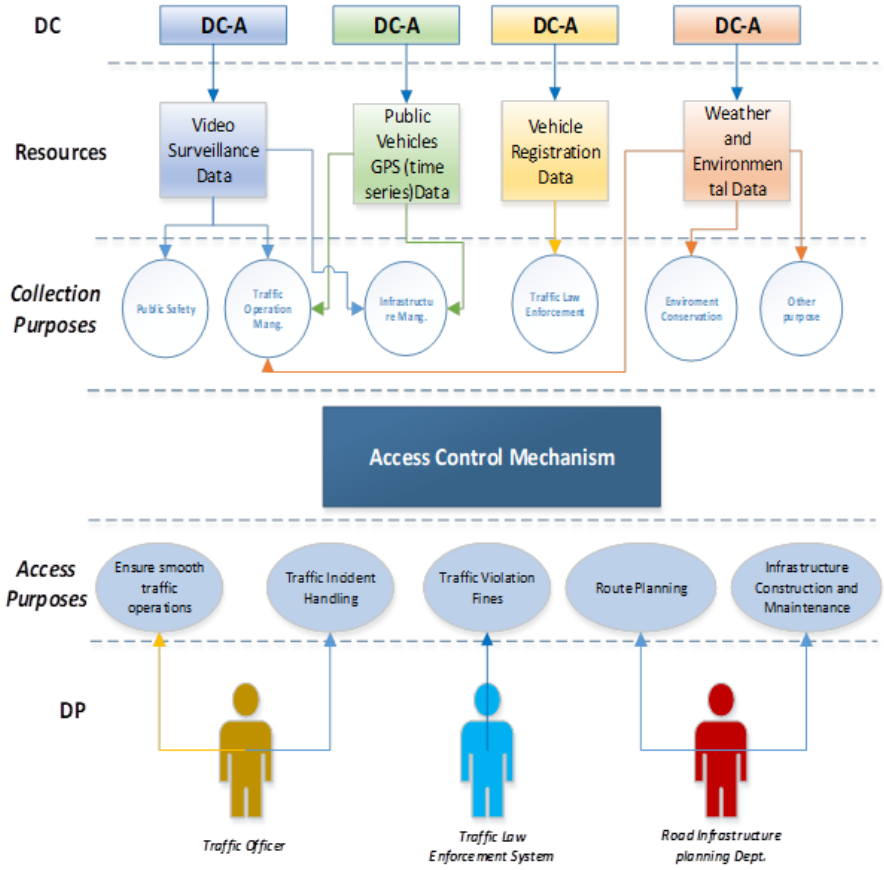


Figure 2.7: Purpose and Access Control Mechanism in Smart City

Collection Purpose	Traffic Operation Management
Resource Description	Mass-video surveillance data collected at public locations by video cameras
Personal Data Properties	As videos have unstructured data, so it does not have defined data characteristics, yet based on the processing capability of extracting personal information, they can be identified as : -Object-type (human, vehicles), object-descriptive-features (gender, color, estimated-age and height), object- identification-features (face, gait, license-plate), -geo-location data (Spatio-temporal position of any object at a specific time) -Locations (highways and others along the road capturing traffic only) -Devices (video cameras' types and unique IDs) -A timestamp of the recording
Personal data property-Purpose Mapping	Vehicle's License plate, driver's face -> traffic light violation, Speeding vehicle, Wrong parking, Wrong turn, Driving in a bus lane, Junction-box violation) -Vehicle's License plate, Human face-> Accident/ Vehicle collision, Seat belt, child detected without a child seat, etc. *An exhaustive list should be defined for all the properties against their usage requirements
Compliance Policy	will be used for public interest reasons: 1. To record, process, and store any event or object that demonstrates a Traffic operations or violation (traffic light violation,

DP	Traffic-Law-Enforcement Observer
Authorized Resource	Video Surveillance Recordings, Vehicle Registration Database, Real-Time Updates From Traffic-Related Sensors
Aim	To enforce traffic laws by capturing and processing any event or incident that results in a traffic violation and issue fines and penalty points on a license based on identification from the intended resources, where applicable
Authorized Resource Requirements	1. Detect and identify traffic event that is considered a violation either via video recording or senor-reading (e.g. speeding) 2. Identify the object-type vehicle (through its license plate or driver's identification information) from video data, and in case of a detected traffic violation issue a fine and penalty points to the object-type driver, if applicable. *mention an exhaustive list of all the traffic operations and violations that are supported by the DC obtain from available resource-objects
DP Authority Period	Jan-1-2020 to Jan-1-2021

Table 2.4: Access Purpose 'Traffic Law Enforcement Observer'

A DP can be authorized to access different resources managed by different DCs for similar or different 'access purposes'. These 'access purposes' may allow DP to aggregate different resources to perform their tasks. For instance, DP (Traffic-Law-Enforcement System) with an 'access purpose' of 'issue a fine on a traffic violation is allowed to aggregate video surveillance recording of a detected traffic violation (to the view license plate of the involved vehicle) with vehicle registration database, which, as shown in Fig. 2.4, is allowed under the 'collection purpose' of 'traffic-operations management'.

To sum up, in large-scale integrated infrastructures, there are a lot of resources with several 'collection purposes', and thousands of DPs with different requirements or 'access purposes' to access one or more of these resources. Therefore, instead of collaborating resources' 'collection purposes' to a large number of known and unversed DP's 'access purposes' directly, the 'collection purpose' can be described as a resource provenance property. The DC can design access control policies based on resource 'collection purpose', rather than binding them to a fixed set of DP's requirements. The 'collection purpose' is described with a broader scope of what is 'possible', and legally 'allowed' for the resource to be used. While DP's 'access purposes' are more targeted and specific per its authorized tasks restricting its to only access data if it meets the requirements of the 'collection purpose'. It also leaves room for the future DP with 'access purposes' that come under already defined 'collection purpose', so their roles or authorizations do not need to be updated explicitly while restricting secondary use.

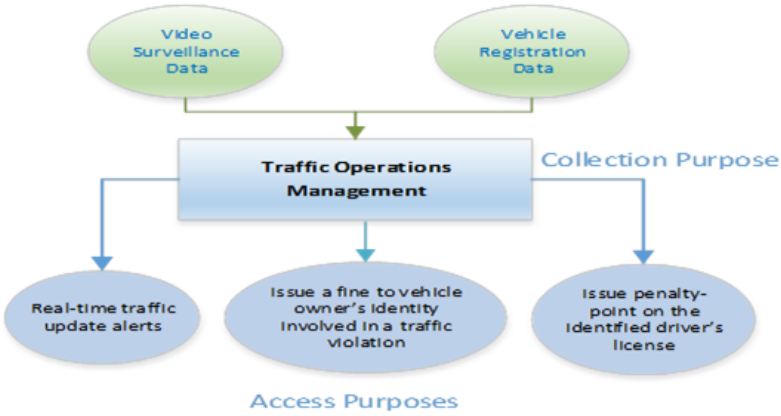


Figure 2.8: Collection Purposes and Access Purposes

2.4 Related Work

The tremendous increase in data sharing and the rise of distributed systems in the last decade has put a lot of focus on privacy. Data is stored, processed, and shared through large-scale infrastructures like relational databases, unstructured data repositories, data lakes (DL) to accommodate the different forms of data at one place, etc. All these infrastructures have different requirements for ACM to regulate access to their stored data to different DPs for different purposes. Here, we will briefly summarize how different large-scale data infrastructures containing resources with personal information are managed through different ACMs. The first subsection summarizes some of the contemporary purpose-based ACMs proposed for large-scale and distributed infrastructures to preserve privacy or limit data usage. The last subsection reviews some of the prominent work about how provenance can be used in access control decisions, by focusing if provenance can be used to preserve the integrity of purpose.

2.4.1 Purpose-based Access Control

To preserve privacy in relational databases, a hierarchical purpose-tree-based ACM is proposed [44] [45] [46] [47] [48]. The system or DC maintains a purpose tree, where each node represents a 'collection purpose' while edge represents a

hierarchal relationship between the parent and the child node 'collection purpose'. 'collection purpose' is bound to a set of data elements or columns in a relational database. In parallel, there is also a role hierarchy, and roles are assigned different 'access purpose' (a subset of purpose-tree nodes) based on their authorized data requirements based on role-based access control (RBAC). More than one 'access purpose' can be assigned to one role, or more than one role can have the same purpose. A DP or authorized user is to present an 'access purpose' as a role attribute when it requests a certain resource. The system then evaluates the 'access purpose' against the central purpose tree and makes an access decision [44]. Another solution based on conditional purpose and dynamic roles is proposed in [46]. The purposes are categorized into three groups: conditional purpose, allowed purpose, and prohibited purpose, which are assigned to different conditions of dynamic roles based on different user and contextual attributes. In this paper, resources are dynamically assigned different 'collection purposes' during the access decision, and if the 'access purpose' has the authorized values of contextual or dynamic attributes then the relevant resource against verified 'collection purpose' is allowed [46]. Most of the state-of-the-art purpose-based ACMs know about the structure and nature of the data and policies are designed for specific users as their data requirements are known in advance, so it is easy to associate a 'collection purpose' to the entire table, or a few column, or tuples. However, this is difficult in distributed infrastructures like data lakes due to the lack of a priori knowledge about a large number of dynamic DP and their 'access purposes' as well as the 'collection purposes' of transformed resources. Moreover, purposes defined in such methods rely heavily on DC's knowledge about resources and their usage, while if multiple DCs regulate access to distributed resources, then it requires 'collection purposes' to be described consistently with characteristics that are accepted by different DCs, so they can design 'access purposes' of their DPs accordingly.

Diversity in data formats and management approaches by different DC along with the dynamic access to that data makes traditional access control methods difficult to implement in data lakes. To address the diversity data challenge and provide uniform access across a data lake, a concept of Semantic Data Lake has been proposed [32]. It presents a middleware framework that requires data sources to prepare data on certain criteria before injecting data in data lakes, and then the middleware can derive mapping between resource's data attribute and semantic DL's ontology (, i.e., the structure of entity attributes) of different data concepts for better access control. This can provide a sense of homogeneity, and data in different formats can be queried based on the formal concepts designed by the semantic data lake middleware. An attribute-based ACM is also proposed for the commonly used Hadoop framework to implement distributed data lakes [49]. Users, resources, and the Hadoop environment have their defined attributes, which are considered when making an access control decision. Resources with similar attributes are grouped in a cluster, which is then assigned

a set of permissions based on the operations that can be performed on these resource clusters. These permissions or policies include different (uniform) tags that are associated with these resource clusters as part of a distributed data lake. These tag-based policies are then assigned to different roles. Users are assigned different roles and based on similar roles they are assigned to different groups and these groups are then defined in a hierarchy to manage a large number of users or DPs in a DL. Users assigned to a parent or higher group get all the roles of their junior groups, which in turn also gets the assigned tag-based permissions to resource clusters. This group hierarchy is beneficial for efficient role management for large-scale DL but is not very useful in limiting secondary use due to inheriting policy-based access without considering individual access control requirements of individual users or DPs.

In another approach, authors have proposed a purpose-based auditing solution, where "collection purpose' is bound to different business processes and then by using formal methods of inter-process communications to verify those purposes against different policies to show that they comply with certain data protection legislation like GDPR [50].

Hence, ACMs designed for large-scale are generally based on deriving homogenous semantic concepts for different data formats and then use these mappings to assign authorizations to different DPs. Other solutions use hierarchical authorizations where permissions or 'access purpose' are assigned to a large set of users against a cluster of resources, and DP (with similar requirements) inherits permissions to these resources. These approaches do provide efficient management of DPs and resources but do leave a potential gap for secondary use due to not considering 'collection purposes' at a fine-grained or individual resource level while making an access control decision [49]. One of the reasons is that it is hard for DC or the system to maintain "collection purpose' for all resources and then communicate it to different DPs in case of data transformations. However, if DC can ensure that updated 'collection purposes' of the resources are available at the time of making an access control decision, and then it can be evaluated against the 'access purpose' of the DP. Therefore, we proposed that DC add the 'collection purpose' as part of the resource provenance, so it will be available along with the resource, and ACMs can consider the 'collection purpose' from the resource while making an access control decision.

2.4.2 Provenance-based Access Control

In this chapter, our proposed approach emphasizes two key ideas: first, that provenance can be used to store the 'collection purpose' of resources that contain personal information, second, that provenance-recorded 'collection purposes' can

be used in large-scale access control mechanisms to prevent secondary use and ensure compliance. Therefore, in this section, we discuss various state-of-the-art solutions that use provenance in some form as part of their access control mechanism. Furthermore, we will also briefly discuss different methods that are used for purpose limitations in large-scale infrastructures.

In the last decade, a staggering number of applications and services based on distributed infrastructures and cloud technologies have highlighted the importance of provenance. Over the years, different provenance schemes have been proposed to describe a way to show data lineage and derivation [7] [30]. These schemes may offer a different view of provenance metadata based on its use, i.e., debugging, reproducibility, annotation, security, etc., thus, provenance along with data lineage information may store other characteristics as required for the usage purpose [34]. Provenance has also been used in different enhanced access control approaches for different storage and distributed platforms [51] [52] [53] [54] [55]. Some ACM solutions propose to capture resource provenance during different activities and then use this information in access control solutions: this is generally to as provenance-based access control (PBAC) [51]. One such notable contribution extracts resource dependencies from provenance logs and uses them to authorize and authenticate users in distributed cloud environments, and later uses this as an attribute in Attribute-based Access Control (ABAC) to make access decisions [101]. In another approach, the authors proposed a generic ontology to capture semantic information (attributes) from different provenance schemes present in distributed infrastructures, subsequently basing classified resources on this information in order to assign access privileges to classified resources [54]. Provenance can also help in the implementation of organizational security policies and is proposed as a hybrid approach with ABAC for enforcement [39]. An architecture for cloud infrastructure has also been developed that utilizes contextual information derived from provenance metadata for evaluating policy decisions [55]. Thus, provenance has been used in different ways for enabling ACM with either deriving policies from resource dependencies or authorizing users, but to our knowledge, it has not been used in ACMs to control the usage of personal information in resources, i.e., ensure purpose limitation to restrict secondary use, as we have proposed.

2.5 Conclusion

In large-scale distributed infrastructures, entities from different resources are transformed multiple times to fulfill the requirements of the DC, which may then be shared among multiple DC, and are available for DP's with diverse 'access purpose'. The DS whose personal information is recorded in those entities

has given either informed or implied consent (as a legal obligation) that their personal information may only be used for the agreed-upon 'collection purpose'. However, many DC are involved in managing large infrastructures, and not every DC will be forthcoming to allow another DC to examine how the shared resource or entities are being used, which eventually decreases the trust of DS over the use of its personal information as per agreed-upon 'collection purposes'.

In this chapter, we proposed to utilize resource provenance to also record its 'collection purpose/s', and any DP requiring access to a resource must comply with the 'collection purpose/s' available at the time of the request. As the provenance is considered immutable and append-only, it will be retained through different transformations, providing a way for any DC to trace and review the usage of its entities [38]. This provides both prevention of secondary use as well as a way to ensure 'collection purpose' verification. For the former objective, resource provenance can be used in access control decisions, where 'collection purpose' can be retrieved from the provenance of the resource and can be compared with the 'access purpose' of the requester, and if comply then access can be granted. For the latter objective, DC at any point of the resource's life cycle can confirm by reviewing provenance metadata whether the entities' 'collection purpose' was comparable to the 'access purpose' of the agent.

Though provenance can catalog different activities that are performed on data, it cannot confirm or ensure that the performed activity is valid or not. Moreover, when diverse DP requests an aggregated data or resource, it may be subject to a new set of access policies that require the explicit creation of a new 'access purpose', as the transformed resource has properties from more than one resources and may have different 'collection purposes'. To avoid the need for constant policy editing, it would be desirable if data (its provenance) were carried with it the necessary information ('collection purpose') to make accurate and informed access control decisions without compromising privacy or undermining compliance. This way provenance will not only be helpful for data auditing but also be useful in controlling resource usage and ensuring data privacy [56]. Therefore, if provenance can initially record 'collection purpose' with resource origin, and then with every transformation or aggregation, the 'collection purpose' can be preserved and appended (updated) when required, it can assist access control modules with purpose verification. For example, when a DP requests the data, the system (or ACM) can extract the resource's 'collection purpose' from its provenance and verify it against the 'access purpose' of the DP, and the result can verify that data is being accessed for the same purpose that it was collected for. Later, once the DP is allowed access to a resource, where provenance registers a unique entry for the performed activity, it can also record the 'access purpose' relevant to that activity. This will be helpful to confirm that all the activities performed over a resource are valid, legal, and did not violate privacy. Therefore, large-scale shared infrastructures and respective DC and DP can use

provenance metadata to catalog and track different data transformations along with their 'collection purposes', which can be effectively used in access control solutions to control personal data usage by comparing it with the 'access purposes'. Moreover, it also ensures compliance with different data protection legislation's requirement for protecting the personal information that achieves two goals, first preserves DS's privacy, secondly, builds up trust among DC, DP, and DS with data usage transparency.

Chapter 3

Large-scale Video Surveillance System Privacy Requirements

Over the past two decades, VSS has evolved from simple video acquisition and display systems to intelligent (semi)autonomous systems, capable of performing complex computer vision and automated decision-making tasks. The evolution of video surveillance systems can be categorized into four generations. First-generation comprised of analog CCTV cameras, videocassettes as storage material, and the video was displayed on a screen and was not connected to a network like the Internet. Over the years with the development of digital CCTV cameras and recorders (second-generation), data analysis and networked infrastructure (third-generation) has led to the fourth VSS generation with wide-area surveillance and reliable network-centric transmission [5]. Nowadays, a VSS can integrate many sophisticated images and video analysis algorithms to extract a multitude and multipurpose information from VSS data along with other logistic benefits such as reduced cost, less transmission latency, scalability, high resolution, and better performance. It can perform various tasks such as object detection and tracking, face recognition, event detection, and prediction, behavior recognition, video summarization, etc., effectively. This helps understand the content of the VSS data so that observers (a monitoring human, a program, or a system) can take appropriate and informed decisions. Hence,

recent VSS generation, based on the multitude of information obtained from VSS data, offers fine-grained indexing, searching, and object tracking that aids in automated decision-making to achieve multiple purposes at once.

Video cameras are excellent multi-sensors so they capture everything happening in its line of sight, without any bias or prejudice (usually). It records the activities of general individuals (passing the road, entering the building, boarding a train, etc.), objects associated with them (other people, vehicle, luggage, frequently visiting places), and all the visible information present in the surroundings. Thus, local administrations or municipalities all over the globe are using VSS to help multiple observers in accomplishing diverse tasks such as smart patrolling, public safety management, traffic operations management, infrastructure maintenance, or in case of recent pandemic social-distance inspection, etc. [6].On the other hand, VSS data collected at a large scale has substantial information about individuals in various dimensions, which when aggregated and analyzed with other relevant data sources with a specific motive or purpose, reveals distinguished sensitive and personal information about individuals. For authorized purposes, such data is highly useful for observers with diverse requirements, however, the personal nature and multitude of information extracted from such data also makes it prone to privacy violations, making it critical to preserving an individual's privacy. Thus, it is essential to balance the protection of the privacy of individuals(DS)captured in the surveillance data against the right of the observers (DP) to have timely access to the authorized information. In order to achieve that balance it is important to closely examine what happens to VSS data from the moment it is recorded until the time it reaches the observer. What type of Privacy Enhancing Technologies(PETs)can be utilized to develop a privacy-aware VSS? Is there a one-fits-all solution that can ensure privacy in VSS or are there factors that should be considered while deciding what observers can see or how much can they access without violating privacy?

In this chapter, we will address above-mentioned questions by elaborating on different privacy aspects of VSS data by using a modern case of large-scale VSS, i.e.,a Smart-City Video Surveillance Systems (SC-VSS).Please note that this chapter may contain excerpts and figures from our own published papers (as part of the PhD research) mentioned and referenced in 1.3.

3.1 Smart City Video Surveillance Systems (SC-VSS)

Smart cities (SC) aggregate and analyze data from thousands of IoT distributed networks and public information systems, which provide highly valuable information to a large number of users (observers/DP, DS) for different purposes. For example, in SC traffic management systems, collective data from vehicles, passengers, traffic signals, motion sensors, video surveillance cameras, etc., can help resolve various traffic issues like traffic congestion, accidents, parking, infrastructure limitations, and provide citizens with safe, efficient, and smart traffic and transportation choice [55]. Extending this concept to various aspects of citizens' daily life, such as healthcare, energy, climate, agriculture, governance, while aggregating its data intelligently forms the core of an SC. Citizens or data subjects (DS) are the prime data contributors of an SC, as the SC relies heavily on the personal information obtained from DS to offer real-time, customized, and context-aware SC services back to them [57].

3.1.1 SC-VSS Data Framework

Smart City Large-scale video surveillance (SC-VSS) refers to the deployment of thousands of cameras at various public locations, road junctions, and highways to record and monitor different events and activities that might be of interest to the various local or public authorities. The data is generally aggregated in a mutually shared distributed processing and storage infrastructure. VSS data contains a lot of information about general individuals, their activities and associations, public infrastructure and operations, etc., that after analysis can generate useful information. Referring to Fig. 3.1, the general SC-VSS framework has three main layers: Recording, Storage, and Presentation, At the first layer, recording devices (video cameras or couples sensors) can record continuously, or record based on a specific trigger/activity/threshold, and usually, everything in the line of sight is captured or recorded. Next, the storage layer processes the data to extract or retrieve information from it and further index, organize and store it in an efficient manner,further discussed in section 4.4. At the last layer, a set of DP or observers with diverse data requirements have been authorized by the SC-VSS (owners or DC) to observe the VSS data (ideally) according to their specified purposes.Several PETs can be applied at any or all of these layers to preserve data privacy with various benefits and drawbacks. These layers are discussed in detail below:

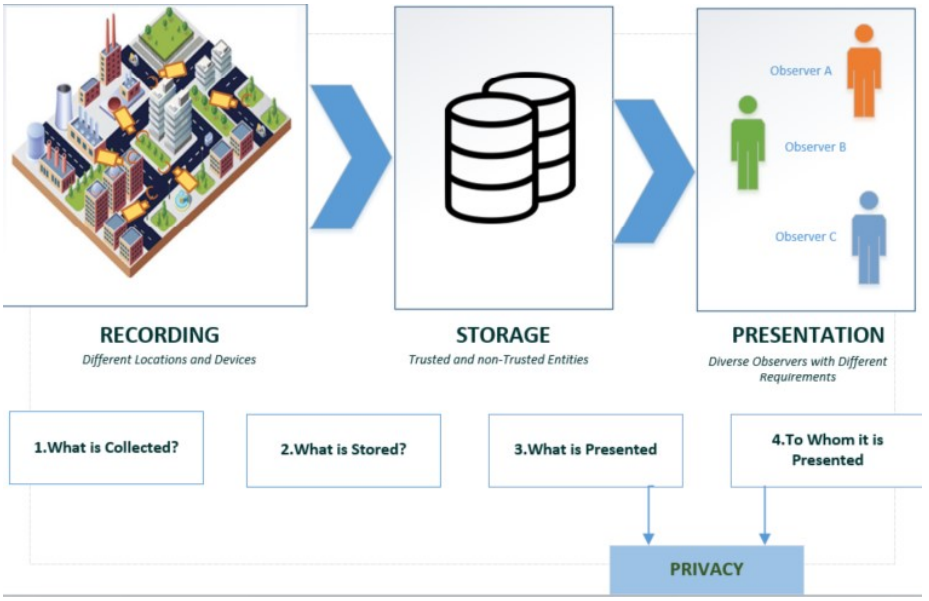


Figure 3.1: SC-VSS Data Framework

3.1.1.1 Recording Layer

This layer represents a physical layout of recording devices (cameras and coupled sensors) deployed at large premises. The type of devices to be used depends upon the nature of the location or specific purpose for surveillance, i.e., sensitive locations (public offices, airports) may need more sophisticated devices than other locations (parks, town-squares), etc. Examples of devices include high-resolution cameras, movable (pan-tilt-zoom) cameras, infrared cameras, drone cameras, etc. [13]. This layer collects the raw (original/unprocessed) data from all the devices at different locations, digital video recorders (DVRs) can be used with these devices to store the recordings (data) locally or can directly share it with a centrally or distributed storage infrastructure.

To limit the data flow or specifically view of personal information to the observer, a limited amount of data can be recorded at this layer. For instance, devices can be configured to record only based on an event or sensor trigger, like cameras can only capture data if a vehicle crosses a speed limit. Furthermore, devices can also be configured to record minimum or no personal information (unidentifiable faces, license plates, etc.), by modifying the form or content of the data, such as content blurring, degraded video resolution, etc. Devices with higher computational capacity can record everything, and analyze the data to

choose specific parts of data to be shared with the next layer or storage [19]. Limited data recording or specific data sharing surely limits observers' view of data; however, it also affects how the rest of the system can use the data. For example, if a recording device blurs the faces of all the persons while recording or sharing data with the storage, then that data cannot be used as evidence in court, because the person-of-interest will be unidentifiable. It does ensure privacy for the individuals, but it also limits the possibilities of how data can be used by authorized observers.

3.1.1.2 Storage Layer

Once the data is recorded, it is transported to a central or distributed storage platform (for both live monitoring and archiving). The storage platform can have both trusted storage units (owned and managed by local authorities) and independent third-party storage units/clouds. There is a fundamental requirement to keep all unmodified/unprocessed data (apart from well-known encoding/compression algorithms) so that data is as objective as possible if it is to be used as evidence in court. In this step, data is processed to extract different types of information from it (as per computational capacity and requirements) by applying modern computer vision and deep learning techniques and is further indexed according to some of the extracted (meta)data properties (cf. section 4.3). As this layer aggregates data from all the devices at the recording layer, often data can be of different types, formats, and specifications, thus, offering various levels of information. Ideally, this layer should have an efficient processing and storage mechanism that extracts most of the useful information from data and index it in a well-organized manner, so it offers multiple ways for accessing and utilizing this data.

At this layer, different PETs can be applied while data is being stored, e.g. by applying anonymization, symmetric/asymmetric encryption, reversible or irreversible transformations, etc. If the same PETs are applied to all types of VSS data, then it might affect its usefulness, i.e., the data cannot be used for multiple purposes. For example, if VSS data is stored in an irreversibly anonymized form so it only offers information about real-time traffic management, then in case of an accident, the police may not be able to identify the license plate of the car that left the scene. Hence, different observers may require different types and amounts of information from the same data [20]. Therefore, VSS data or parts of it must be carefully selected to be transformed (encrypted, anonymized, scrambled, etc.), to protect it from unauthorized access and still be useful. In the case of third-party storage servers or data sources, it must be guaranteed that they collect, store, share or use this data strictly according to applicable guidelines so there is no threat of secondary use or information leakage. While

sharing the data with third parties, it can be encrypted by keeping some non-personal data properties in the clear for requesting it or keeping encryption keys at trusted storage, which can later be used to reference and retrieve it [21].

3.1.1.3 Presentation Layer

This layer deals with observers that need to access VSS data in some form, such as a live stream, archived recording, reports, analysis, proof of evidence, etc. In order to access the data, there is a querying mechanism based on different indexing properties (defined at storage layer), such as time-stamp, location, event type, object type, etc. (cf. Section4.4), to allow observers to search and access a particular piece of data (video recording). The observer or DP here is an authorized user, who must have a legitimate 'access purpose' for accessing this data, and (ideally) it needs to demonstrate its authorized 'access purpose' (maybe in the form of credentials that encode some properties or permissions), every time a request is made. Here, the request of each observer is dealt with individually and a combination of PETs can be applied right before data is made available to them. The most common PET used at this step is the access control mechanism (ACM), which controls the flow of information to each observer given its request and VSS's allowed permissions. Access control methods are broadly categorized in two ways: static or dynamic. In the case of static ACM, all observers with similar credentials (department, role, authorized area, etc.) will be given the same access based on fixed resource/data properties and observers' permissions, regardless of their current conditions. For example, traditionally, VSS recordings are available in their original form to all authorized observers, and even if there is some privacy preservation technique like blurring or selective encryption applied over data to hide personal information, the same view of data will be available to all viewing. Alternatively, in the case of dynamic ACM, the access level to VSS data or in particular may vary according to the contextual requirements of the observer. The most common contextual parameters are space and time. For instance, an observer can only be allowed to access VSS data, if the location of the recording camera is a subset of the observer's authorized locations, or an observer is only allowed to view VSS data recorded during its authorized duty hours. With the recent development and success of video analysis solutions, semantic information from the content can also be included in ACM policies, for instance, an observer can view data if a particular object is detected in VSS data content. Hence, ACMs in general limit the information flow to authorized observers based on some specified parameters or a defined policy. Furthermore, different PETS can also be bound to these policies after an observer has been authorized to view VSS data. For example, an observer is only allowed to view VSS data in a blurred or encrypted form unless authorized by a superior entity or data owner [22]. It is also important to note here, that

authorized observers are allowed to access some of the personal information and may be prohibited to access another type of personal information. Therefore, it is important for the SC-VSS or its ACM to categorize personal information which can then be verified or compared with observers' requirements or 'access purpose'.

To conclude, in order to preserve privacy in large-scale VSS, where it can be used for multiple purposes, it is often challenging to only use non-invasive surveillance tools. For instance, limit video cameras to record occasionally at the recording layer, or apply a one-fits-all PET solution to all the aggregated data at the storage layer, as it will not serve the purpose of multiple VSS observers. Therefore, the ideal solution is to limit what any particular observer can see at the Presentation layer, by implementing a privacy-aware access control mechanism that restricts the observer's access to the VSS data based on two things: Understand and categorize different types of (personal)information in SC-VSS, and second, regulate access to different types of information based on observer's authorized requirements or 'access purpose' [27].

3.2 Privacy and Access Requirements in SC-VSS

SC-VSS data from all over the city is shared, aggregated, and analyzed with other data sources to generate new insights that enrich the existing information, and useful results are derived for observers or data processors (DP) with a diverse set of requirements [58]. Thus, SC-VSS data-ownership and access granting dynamics are complex as it aggregates data from various public DCs in heterogeneous formats collected for various 'collection purposes', supported by a valid legal base. A legal base establishes legal grounds for personal data processing activities and is a must requirement by various data-protection legislation along with a suitable 'collection purpose'. Several smart-city authorized DPs such as law-enforcement authority observers, traffic-enforcement observers, congestion handing and route-planner services, infrastructure-planning departments, etc., can request the SC-VSS to combine video data with data from different DC sources in order to generate results that fulfill their requirements or serve their 'access purposes'. For instance, VSS data or video recordings can be combined with the time-series location data from public transportation means for the 'collection purpose' of routes and traffic congestion management, or video recordings can be combined with the vehicle registration database for the 'collection purpose' of handling traffic violations and missing vehicles, etc., as shown in Table 3.1. Similar aggregations (with VSS data) are being used in all the major cities of the world by various DPs to assist with different administrative operations that facilitate citizens/DS with customized services.

Data Source (with Personal Information)	Legal Base	Likely Collection Purposes for Aggregated Resource
Video Surveillance Data (Unstructured)	Public Interest	Public safety Traffic management Real-time traffic updates Route planning and congestion handling Traffic law enforcement
Vehicles' Sensor Data from Public Transportation Means (Semi-Structured)	Contract	Vehicle tracking Congestion handling Weather monitoring Noise reduction Real-time traffic updates Route planning Traffic law enforcement
Vehicle registration data (structured)	Legal Obligation	Vehicle registration License registration Incident handling Violation handling Traffic law enforcement

Table 3.1: Resources and their 'collection purposes'

On the one hand, data aggregation and analysis at such a large scale do provide numerous benefits, but on the other hand, it also raises questions about the observers' access and who delegates those access rights to them in a distributed infrastructure. For instance, if a DP requests an aggregation of resources managed by different DCs, then who is responsible for deciding which observer should be allowed to access the requested resource/information, and how much of that information is necessary to serve their 'access purpose'. Moreover, due to the intrusive nature of video data, how is personal information protected from unauthorized access or secondary use by an authorized observer, and when aggregated with another resource, does it affect the existing personal information (enrich it in one way or the other), or does it offer a possibility for secondary use, etc. For instance, a law-enforcement authority observer (LEA) is watching the feed of several cameras deployed in a city square to look for objects or events of interest that are relevant to the 'collection purpose' of public safety. The observer is looking at the activities of individuals in a specific area, such as people coming in and out of different buildings at different times of the day, their belongings (vehicle, luggage), associations (interaction with other humans, work location (if in city square), etc. The DP can infer a great deal of information about individuals by merely observing them and their routine activities, which is irrelevant for public safety. Moreover, if the observer also has access to another data source let's say national citizen registration database, and it aggregates it with VSS data to identify different persons in the recording, is

it acceptable with the 'collection purpose' of public safety? Or, the observer should only be allowed to view data that is relevant to its 'access purpose'. Ideally, the observer's 'access purpose' must be a subset of the requested resource's 'collection purpose', to ensure purpose limitation. Thus, SC-VSS requires an efficient access control mechanism that takes into account different factors like valid 'collection and access purposes' of resources and observers, and also verify their integrity by validating their authorization by their respective DC, and then decide whether the resource (video-recording) should be allowed for observer's view or not,as discussed in previous chapter. It further ensures purpose limitation and limits secondary use so the privacy of DS present in the SC-VSS data can be preserved.

Here we analyze a specific scenario to see what type of information is of interest to different observers, and what is irrelevant based on their requirements or 'access purpose', for a specific 'collection purpose' of 'public safety' supported by a legal base 'public interest'. A set of adjacent cameras installed on the highway detects a collision between a truck (carrying combustible liquid) and a van (carrying passengers). After a few minutes, and due to the impact of the collision, a fire starts when combustible liquid leaks out of the truck on the site of the accident. Different local authority observers or first responders might be interested in this particular VSS data (recording) for 'public safety' purposes, so they may obtain some information to perform their duties efficiently. Now, keeping in mind, the different types of information can be extracted from video data, there are two main activities or events in the particular recording, one 'moving vehicles' and the other is 'vehicle-vehicle collision'. The first activity is of interest only to a traffic-monitoring observer (TM) to analyze traffic flow, congestions, traffic violations, etc. and the TM only needs to view regular traffic operations, while the identity of vehicles, drivers, and passengers is irrelevant unless there is a law violation. The second event 'vehicle-vehicle collision' is of interest to several first responders (law enforcement, fire department, paramedics, etc.), as it may result in endangering human lives and damage to public infrastructure, hence, a 'public safety' issue. Each of the observers has access to view all the available content in the recording like different objects (vehicle, humans, and signposts) involved in the accident, or otherwise captured in the same recording at the 'time' and 'location' of the accident. These objects can be further processed to extricate granular-level information about them such as descriptive features of vehicles (shape, model, color, license plate, etc.), humans (clothes, height, color, etc.), and biometric features of humans (face, gait), etc.

Identities of these vehicles and humans can be found out by associating their known features with other data sources. For instance, information about the vehicle's 'license plate' can be extracted with the automatic number plate recognition (ANPR) tool, and can then be compared with the vehicle registration database for identification. Similarly, biometric features like the human face

can be searched within a national database through facial recognition for identification. An observer (a paramedic) needs to view this recording to prepare for the 'access purpose' of "providing medical assistance". It is interested in the information about humans involved in a collision, while information regarding vehicles is irrelevant to her. Moreover, there is another paramedic observer with the same requirements, but it is not close to the vicinity of the vehicle-vehicle collision, so ideally it should have a limited view of personal information as compared to the former paramedic. Similarly, another observer (firefighter) is viewing this recording to assess the situation to 'put out the fire'. Does it need to see biometric information of humans or identification information of vehicles in order to do its task? This example shows that even in the case of interested events/activities, there is still a lot of personal information that is not relevant for every observer. Below is a sample Table 3.2, which shows the different types of observers and what information they do and do not require of VSS data in order to fulfill their 'access purpose' by achieving a particular task in the presented scenario? How much of the available content is relevant for the observer is decided by its particular task against that purpose.

Thus, to implement a need-to-know privacy-aware VSS ACM, the following inquiries must be taken into account:

1. Does the contextual or physical parameters (current location, recorded timestamp, etc.) of both the resource (VSS data) and the observer affect the decision about the view of personal information for the current request? For example, a paramedic present at the site of the accident with the task of "providing medical assistance" can view identification information regarding humans in the requested recording but does not need to view vehicles' license plates. However, is it ok for an observer (another paramedic) who is not near the accident location or responsible for this particular accident to view any of this recording, or is it better that only paramedics with the highway as their "current location" should be allowed to see it?

2. What kind of information can be extracted from the content of the VSS data? And can the extracted information be classified in a way that could be correlated with the observers' requirements? Can personal information extracted from VSS data be categorized so that different observers may have access to different categories per their requirements?

3. With the unstructured nature of video data or recordings, is it possible that a particular part of the video recording can be used to control access or exposure level to another part of the video recording? As in the above scenario, the event "vehicle-vehicle collision" has an influence in deciding the type of personal information (about the objects involved and around that event) that should be available to the observer per its requirements. Ideally, if there is no event of

interest relevant to the purpose of VSS data collection, then the observer should not have access to any type of personal information under normal circumstances.

Observers	Tasks to be achieved in the current scenario	Information Required from VSS Data	Information NOT Required from
Traffic Operations Management	Clear accident site to resume normal traffic operations	Descriptive information about the objects (humans or vehicles) involved in the particular incident	biometric or Identification Information about any individual or vehicle
	Record traffic violations happened/caused this incident	Recognition of occurred traffic violations (wrong turn, speeding, overtaking, etc.) Identification information about drivers of vehicles involved in those traffic violation to issue fines/ penalties	descriptive or biometric or Identification information about accompanying passengers or any nearby people
	Investigation of the cause of the incident	Descriptive information about the objects (humans or vehicles) involved in the particular incident	biometric or Identification Information about any individual
Law Enforcement Authorities (LEA)	Information about involved/affected humans if said traffic violation has criminal implications (e.g. accidents resulting in any sort of damage or DUI)	Identity Information about vehicles, their drivers, and people who witnessed the accident	Descriptive features of the involved individual (not required to know about the color or gender)
	Notify next-of-kin of humans (unconscious) affected by this accident	(In case affected humans are conscious then their consent is required to inform next-of-kin) Identity Information about affected persons (drivers or passengers or passing-by) in the accident that can be cross-referenced with central citizen database to inform next-of-kin	biometric or identification Information about non-affected humans in accident
Mobile Healthcare Units (paramedics, ambulances)	Provide medical assistance to individuals affected by the accident	Descriptive features of affected individuals to know about the number and state of people injured in an accident Identification information about unconscious individuals to access health records if any medical procedure needs to be administered	biometric or identification information about conscious humans in accident
Fire department	Contain fire (for workload estimation)	Descriptive Information about involved objects that started the fire (source of fire)	Descriptive or biometric or Identification information about the driver, accompanying passengers, or any nearby people or vehicles
	Rescue Affectees	Number of people who need rescuing (still in the vehicle and cannot get out) Biometric or Identification information about affectees	Descriptive or biometric or Identification information about non-affectees
Infrastructure Management	Record the damages to infrastructure (if any)	Damage to public infrastructure (roads, landmarks)	Descriptive or biometric or Identification information about any involved object (human or vehicle)

Table 3.2: Observers and required information against different tasks about event "Vehicle-vehicle collision"

4. In distributed environments like SC-VSS, where there are observers with diverse requirements, is it enough to have just one purpose. For the above-mentioned purpose, public safety is one of the 'collection purposes' of the resource (VSS data), that is assigned by the public authorities DC. Thus, is the 'collection purpose' alone enough to decide the level of access to personal information or to ensure need-to-know privacy measures? As in the example above, observers have requirements relevant to the same 'collection purpose', yet they still require different view levels of information.

5. The amount and type of information available to observers are dependent upon their requirements yet are highly dependent upon or are limited by the 'collection purpose' of the requested resource. Can the relationship or dependability between observer's requirements ('access purpose') and resource's 'collection purpose' can be quantified?/ described/ visualized. (better word)

In order to preserve privacy, is it essential to ensure purpose limitation, i.e., it is important to ensure that observer's 'access purpose' is controlled by the resource's 'collection purpose'?

6. Can the information extracted from VSS data, 'collection purpose' of the resource, and observer's requirements or 'access purpose' be defined in terms so they all can be corroborated/validated or mapped against each other?

7. Due to the distributive nature of SC-VSS, where there are multiple data sources and different observers, is there a need to ensure purpose (both collection and access) integrity? When surveillance data is aggregated or analyzed with another data source to generate enriched personal information, how is it ensured that the 'collection purpose' of both the resources is observed concerning purpose limitation, and the 'access purpose' of the observer still complies with it unless otherwise specified.

Summarizing the inquiries mentioned above, a privacy-aware ACM for SC-VSS needs to take the following things into account while making a decision: First, understanding of the VSS data or specifically its content to classify different types of personal information present in the resource is important to achieve fine-grained privacy-aware VSS or ACM. Second, for any resource that has personal information, it is a legal requirement to have an explicit 'collection purpose' for its collection and usage. Third, Ideally, the better the 'collection purpose' can be mapped to the classified personal information in the resource content, it will be efficient to protect or regulate its access. Fourth, to ensure that the authorized observer is using the personal information legitimately, it is important that the resource's 'collection purpose' and observers' authorized requirements or 'access purpose' are comparable and can be verified against each other. Lastly, it is critical to preserve 'collection purpose' integrity in order to

validate its verification against the 'access purpose' of the users to ensure privacy in distributed systems, where data transformations and aggregations are frequent. Thus, to implement a fine-grained and privacy-aware large-scale VSS ACM it is important to consider all the above-mentioned factors to be a part of the decision-making process. In the next chapter, we propose a need-to-know privacy-aware ACM that fulfills the mentioned requirements.

Due to the complex nature of VSS data and personal information present in it, and their high relevance to the concerns addressed above, we will discuss the nature and categorization of VSS data in the next chapter, and after that propose a privacy-aware (metadata-based) ACM in Chapter 5.

Chapter 4

Video Surveillance Data and Personal Information

VSS was initially designed to observe the camera recordings in real-time, where an observer (a human) was continuously looking at a monitor for any incident/event of interest. With time, VSS has become a standard surveillance technology, which is now extensively used for different purposes, making it difficult for humans to observe all recordings manually. For instance, observing traffic management data for a whole city 24/7 manually is not feasible, so machine assistance has become necessary. Numerous computer vision, machine and deep learning algorithms for object detection, object tracking, object classification and recognition, event detection and prediction, behavior recognition, video summarization, etc., are being applied to the image and video data to automatically extract different types of information from it that is of interest to different observers [59]. Thus, there is a lot of information that can be obtained from video surveillance data depending upon the required computational capacity for executing different video analysis solutions and requirements of the application/observer using that data. In this chapter, we will analyze video surveillance data to extract and categorize different types of information that can be of use to VSS observers. Please note that this chapter may contain excerpts and figures from our own published papers (as part of the PhD research) mentioned and referenced in 1.3.

4.1 Video Data Extraction

Video contains a lot of unstructured information, which is easy for the human eye to understand and categorize, yet very complex for machines to understand and process efficiently. For human observers, when they view a certain video recording, they can classify information such as different objects (humans, vehicles, landmarks, etc.), their isolated activities, interaction among different objects (human driving a car, human entering a building), etc. However, machines do not process the unstructured information the same way and see a grid of bits/pixels representing different numerical values. A video may consist of hundreds of such grids and based on varying information of bits/pixels can be divided into a top-down information hierarchy of different structural units, i.e., video clips, scenes, shots, and frames. A frame is the smallest building unit of a video, frames with similar information make up a shot, and similar shots make up a scene, and so on. Multiple frames are analyzed together to classify different types of information or concepts present in a video recording. Initially, change detection techniques are applied over a series of frames to detect background and foreground. The background represents the static or least changing part of the image/video that remains consistent throughout a scene. Once the background is detected, it is subtracted from the frame and then the rest of the information is considered as foreground. The foreground is then processed by classifying it in distinct regions representing homogenous or consistent information based on different parameters like edge or color density and are extracted as binary large objects (blob). These blobs are transformed into 2D objects and then further 3D objects are reconstructed by applying region localization algorithms (to differentiate and identify their shapes, boundaries (edges), and location in a frame), and the process is called object identification. The identified objects are then categorized into different known classifications such as humans, vehicles, animals, plants, etc. Based on the requirement of the application or observer the objects can be further processed to label different properties or features of these objects. For instance, if the identified object is a human, its face or head or other body parts can be labeled, or if the object is a vehicle, its license, or type can be labeled, such process is usually known as object recognition. Consecutive video frames can be analyzed to record and model different Spatio-temporal parameters about objects to classify their motion into different activities. The process of information extraction from a video is shown in 4.1.

For instance, isolated activities (humans (running, standing, sitting), vehicle (on-move, parked, taking a turn)), interactive activities (handshake, pushing, fighting, playing, collision), natural occurrences (fire, rain, smoke), etc., which can further be grouped into a consecutive set of activities called as an event. Anything tangible that is of interest to the observer in an image or video is normally called an Object-Of-Interest (OOI).

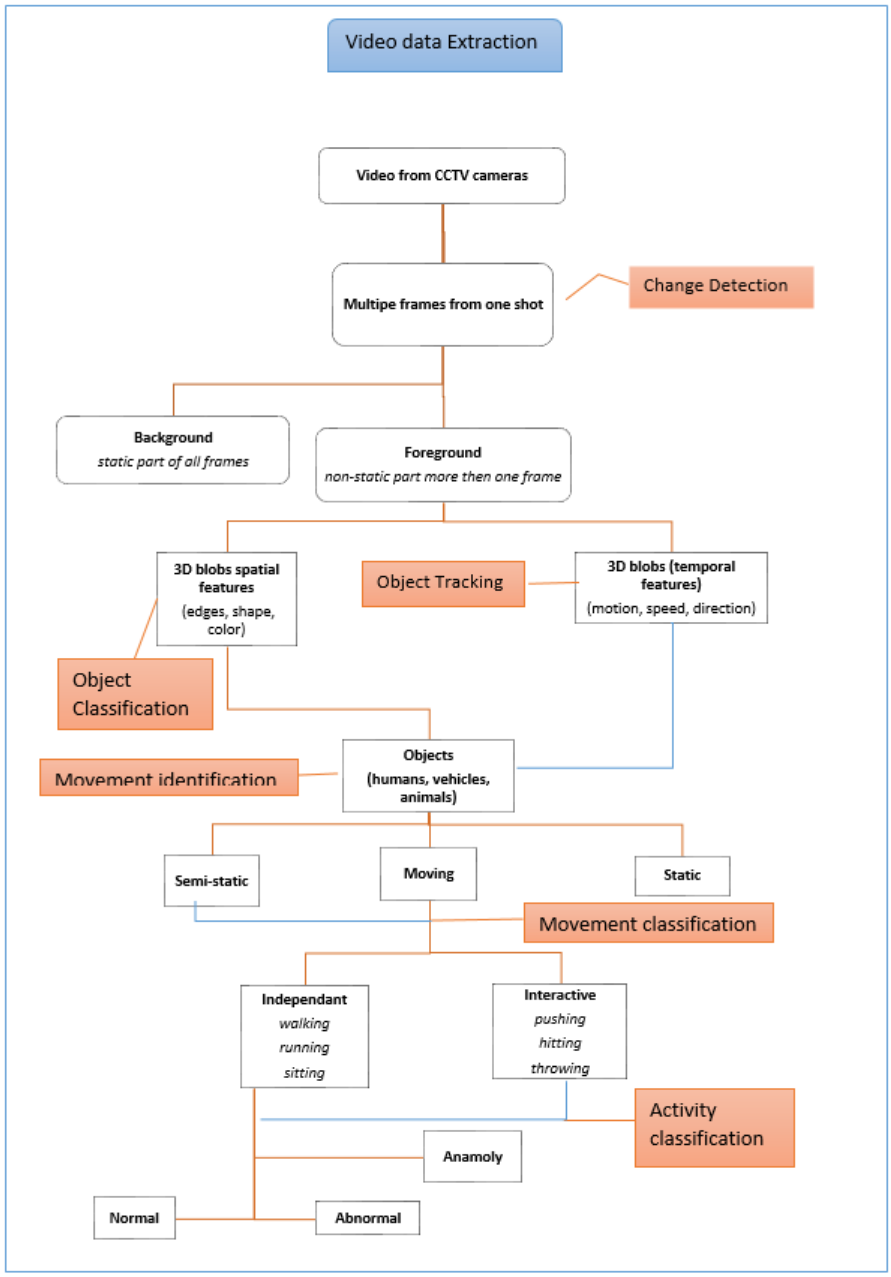


Figure 4.1: Information Extraction from Video Data

Different applications or observers may consider different OOI in the same video/image, e.g. vehicles are OOI for a traffic management system, while for law enforcement authorities people are OOI. Activities and events can also be described as Events-of-Interest (EOI). Below is a simple diagram to show how information is extracted from video data.

This subsection briefly discussed how the information is extracted from video data; the next subsection will briefly discuss the state-of-the-art solutions for how the information extracted from the video data is analyzed.

4.1.1 Video Data Analysis

Object detection and object tracking methods are the major tools to obtain information from video data, closer to the way humans do. Object detection methods help in the detection and recognition of all the distinct concepts (objects) in an image or a video frame and then categorize them into known classifications [60]. While object tracking trails the motion of a particular object in a video stream by analyzing its frames (static image of a continuous recording) sequentially [61]. In the past, most of the solutions to object detection and tracking were based on traditional methods that extracted low-level information from images (colours, edges, texture densities, etc.) and constructed different statistical models to classify and recognize objects. Traditional methods like Scale-invariant feature transform (SIFT), Viola-Jones object detection based on Haar features, and Histogram of oriented gradients (HOG) features were computationally cheaper but did not provide high accuracy [62] [63] [64]. The recent advances that seem to be making a lot of progress in this domain with high accuracy are mostly related to deep learning and specifically its applications based on convolutional neural networks (CNNs). LeNet proposed an algorithm based on CNN architecture almost 25 years ago, which was implemented in the recent past by AlexNet in 2012 [65]. It was the first breakthrough that performed better than traditional statistical solutions and since then a lot of work is being done in this area. Deep learning methods (mostly) extract high-level semantic information (based on learning from low-level information) and classify them into salient objects, track these objects based on their spatial-temporal properties and relations among multiple objects. Nowadays most Image and video-related applications use one of the following solutions for object detection and classification: Single Shot MultiBox Detector (SSD), You Only Look Once (YOLO), and Retina-Net [66] [67] [68]. In the case of large-scale video surveillance systems, where real-time object detection is required, YOLO is a suitable option. It is based on a neural network that processes images or video frames as a "coarse-to-fine" search, i.e., performs a quick scan of the full image first, and extracts regions of interest. These regions-of-interests are localized in 'bounding

box' coordinates and are further classified into known or labeled classes based on class probabilities. Fig. 4.4 in the next section shows the YOLO output of an image, which has accurately classified the detected objects by associating them with their closely resembling classes [14].

Once, different objects are detected and classified in different frames, the next step can be to analyze their relationships or understand the context of these objects. Semantic segmentation techniques help understand the context of an image or multiple frames of a video by associating different semantic properties rather than focusing on only one (like in object classification) [69]. It is not an isolated step, but an advanced step in the natural progression of object detection and recognition. The base of semantic segmentation is the methods that are used in classification such as AlexNet, GoggleNet, etc. Semantic segmentation usually forms the base for more complex tasks such as Scene Understanding and Visual Question and Answer (VQA) [67]. A scene graph or caption is usually the output of scene understanding algorithms [70]. This technique is also vital for self-driving cars; so many efforts are going into this domain for understanding complex scenes such as in Fig. 4.2 [71].

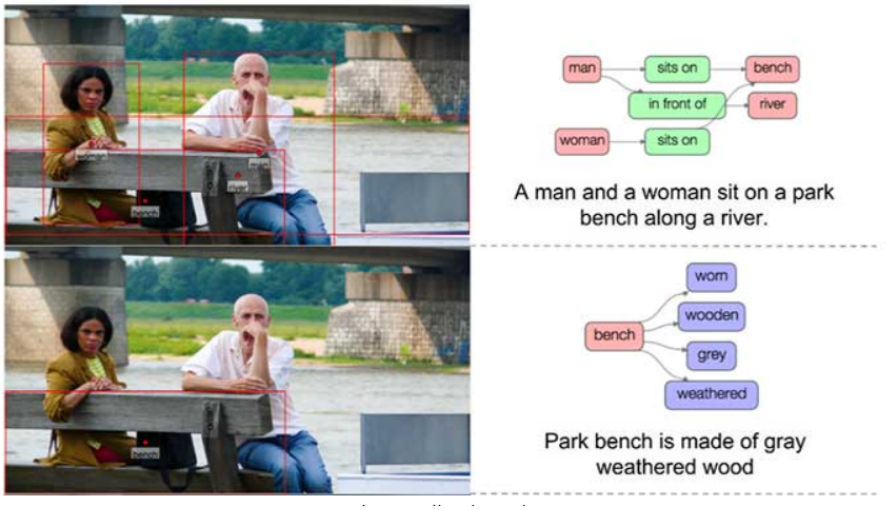


Figure 4.2: Semantic Segmentation Example

Instance segmentation describes different instances of the same type of object. For example, in object detection, we labeled the same types of objects as one category such as 'car', though, instance segmentation differentiates those instances from each other as separate objects like labeling five 'cars' based on their color or model [67]. Instance segmentation requires more complexity than semantic segmentation because of the complicated backgrounds and overlapping

objects.

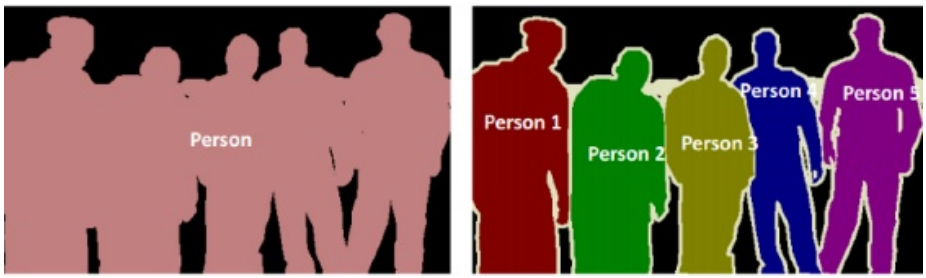


Figure 4.3: Semantic Segmentation vs Instance Segmentation

Instance segmentation requires a combination of classification, bounding box regression, segmentation, i.e., bounding boxes, to first detect and classify an object of the same class, and then the pixel-wise calculation of each instance of the same class is used to differentiate their properties. Facebook AI, known as Mask R-CNN is the most used approach for instance segmentation shown in Fig.4.3 [70] [71]. Similar to R-CNN, Mask R-CNN has a similar architecture to detect objects, added in later is the application of pixel-level segmentation. It outputs a binary mask that says whether a given pixel belongs to a certain object. It requires a CNN Feature Map as the input and then the network outputs a matrix with ones on all positions where the pixel belongs to the object and zeros elsewhere, also known as a binary mask.

After different objects are classified in a video, the next step is to track their activities, hence, object tracking. Both traditional and deep learning techniques, to a certain extent, follow the same traditional tracking mechanism such as Kalman Filter, Multiple Hypothesis Tracking (MHT), and Layer-based Tracking, etc. [72] [73] [66]. The main difference is that traditional mechanisms collect features first and then construct a track prediction model while deep learning methods update the model iteratively. Earlier visual tracking methods based on machine learning techniques used neural networks as a black-box feature extractor to distinguish objects and then associated their corresponding data to track an object [59]. There are other solutions based on Siamese neural networks that measure the similarity between adjacent frames for efficient feature learning as well as temporal matching of different features of various objects [74]. This helps in cataloging all the temporal features of an object over a series of frames, which are then modeled to construct a pattern, which is further matched with available classes, known as activity recognition [75] . Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) are well-known ways to address activity recognition problems [76]. LSTM network models are a type of recurrent neural network that can learn and remember

long sequences of input data. They are intended for use with datasets that are comprised of long sequences of data, up to 200 to 400-time steps. LSTM with its different variations can provide high accuracy (up to 80 percent) in detecting activities with benchmark video datasets like UCF-101 and HMDB51 [77]. UCF-101 and HMDB51 are the most relevant video datasets to our proposed approach, as their categories are mostly related to human activities or events such as Human–Object Interaction, human-human interaction, human standalone activities (jump, pull up, push up, run, sit down, etc.), sports, etc. [78]. There is another dataset called 'ActivityNet', which has organized daily-life activities in a hierarchical structure and co-relates it with social interactions [79] [80].

In short, the above subsection identifies different well-known techniques that can be used to extract semantic information from video data content, i.e., how distinct regions within frames can be detected and classified into different objects, and further various semantic attributes about these objects can be distinguished. Object-tracking techniques extract temporal properties of different objects over a series of frames to construct patterns, which are classified into known activities. Video Content analysis is still a developing field, but it gives a strong idea of how different objects and their activities can be used in different applications like video surveillance. Optical character recognition (OCR) techniques can be used along with video analysis techniques to understand the text in images and videos, such as Automatic Number Plate Recognition (ANPR), which is used to identify license plates of vehicles in videos. Similarly, face recognition algorithms can be applied to extracted objects (human faces) then link them with a national database for identification or use third-party solutions like Clearview AI. Several other techniques can be used to further process extracted objects to specify their details such as color, age, gender, ethnicity (humans), or type and model of vehicle, etc. [76]. Hence, there is a lot of information or metadata that can be obtained automatically from surveillance data depending upon the VSS computational capacity and requirements of the application/observer using that data. The next section will discuss, how the extracted information from the video can be categorized.

4.2 Video Surveillance Data

Let us look at the example of a typical video-recording frame in detail to describe the different types of information that can be extracted from it, and how can it be organized. Let us say that the frame shown in Fig. 4.4 is from a video recording of a camera deployed at a lamppost somewhere in a city, capturing different activities happening in its line of sight. It can detect several different objects such as people (men, women) crossing a road, moving and parked ve-

hicles (car, pickup trucks), landmarks (buildings, traffic signs, lampposts, etc.), etc. There is some other information that can be inferred from the video content. For instance, the direction of people relative to the camera, red traffic lights for the opposite direction of the road, dim light in the frame shows the time to be in the early morning or late afternoon, etc. This is the literal information (content) extracted from video recording or known as semantic content using different video analysis solutions, as discussed in subsection 4.1.1. This information can be classified under different labels or categories like 'objects' and activities of those objects, i.e., 'events', and can generally be called metadata. Metadata is the 'data' about data, organized or classified under different labels and properties. Metadata extracted from content is called semantic metadata.

Figure 4.4: A surveillance image from dataset COCO

There are types of metadata other than semantic metadata that is contributed by VSS components, for instance, like VSS devices cameras, coupled sensors, processing and storage devices, etc. Cameras recording videos have their specification properties such as device type, model, resolution, deployment angle, installed location, etc. When the video data is transported from cameras to storage or later for observation, the video stream has some additional properties, such as the compression ratio, transmitted resolution, encryption algorithm, time delays, etc. Furthermore, there are also the contextual properties that are logged with every recording such as the date and time of the recording, data from other sensors such as a microphone, GPS, accelerometer, (if deployed along

with the camera), etc. All this information (other than the content) is categorized as non-semantic data. There is also a specialized type of metadata, known as provenance, which stores information about data lineage. For instance, when a VSS collects data from multiple cameras or locations, the provenance metadata of each recording shows the origin, ownership, and creation time of that particular recording.

Thus, metadata obtained from VSS data can be broadly categorized into three types:

1. Semantic metadata
2. Non-semantic metadata
3. Provenance metadata

4.2.1 Semantic Metadata

Video recording is processed to extract and organize semantic information from the content. For instance, such number of objects, object category (humans, vehicles, buildings), Types of objects (humans-> man, woman, children) (vehicles-> car, pickup van, truck). Details of objects (humans-> man-> estimated height, black hair, clothes (color, type), accompanying object (if another human then their traits and so on)), (vehicle-> (type, model, color, current location, direction, driver and passenger count), etc. For different events or activities detected in content information such as Event Category (Traffic Violation, Public safety, etc.), Event Type (Traffic Violation-> speeding, illegal parking,) are useful for different applications or observers.

SR. No	Properties
1	No. of Events in Video
2	Event Category
3	Event Type
4	Event Duration
5	No. of Objects in Event
6	Object Category
7	Object Type
8	Object Location

Table 4.1: Semantic Metadata Properties

4.2.2 Non-semantic metadata

Non-semantic metadata is the data that is generated by the devices other than the surveillance data (recordings). This can be further categorized into two types:

4.2.2.1 Non-semantic metadata (Device Related DR)

Non-semantic metadata (DR) is provided by cameras (or other devices such as recorders, storage servers, etc.) themselves and is generated instantly. It includes information like camera type, model, firmware, lens type, lens resolution, installed location, current angle (in case of PTZ), elevation (height of the camera from the ground), timestamp of (when the video was captured), duration, Identity of the camera (MAC and IP address) of network devices (IP Cameras, recorders), etc. This is all very important information and can add value to the original data (the content of the video).

SR. No	Properties
1	Device category (camera, recorder,)
2	Device type (PTZ, DVR, servers, etc.)
3	Device model
4	Default firmware
5	Original resolution
6	Aspect ratio
7	Storage capacity
8	Bandwidth
9	Elevation
10	Visibility (Lens clarity)

Table 4.2: Non-Semantic Metadata Properties DR

4.2.2.2 Non-semantic metadata (Stream Related SR)

Non-semantic metadata (SR) is data associated with every recording or video (not related to content), i.e., timestamp, duration of the video, etc. Video stream (surveillance data while being transported) also have some metadata that is useful for the observer. As video files have a lot of redundant information, so they are compressed before transportation. Compression parameters like format, compression ratio, resolution of the video stream (high compression reduces video resolution), or if there is only one resolution or streams of multiple resolutions are streamed, pixel resolution, time resolution, etc. Additional

information can be added as if the stream has some cryptographic properties like encryption algorithm, key size, hash algorithm, etc.

SR. No	Properties
1	Compression Ratio
2	Recording format
3	Recording angle
4	Recording resolution
5	Recording orientation
6	Recording altitude (elevation of the device at the time of recording)
7	Visibility (lightning conditions)
8	Encryption Algorithm

Table 4.3: Non-Semantic Metadata Properties SR

4.2.3 Provenance Metadata

Provenance metadata catalogs different steps that any resource (or in this case a video recording) goes through from its origin until an observer has an access to it. Furthermore, it stores information about data lineage such as who created that recording, when was it created or in some specific cases why is it created, etc.

SR. No	Properties
1	Owner/Controller (Device/stream)
2	Designated observer
3	Creation/Installation Timestamp
4	Root/parent device deployment location
5	Last maintenance timestamp
6	Last access timestamp

Table 4.4: Provenance Metadata

Hence, to summarize, different types of information can be extracted from VSS data, that can be labeled under different metadata categories, i.e., semantic, non-semantic, and provenance metadata. All these different metadata properties describe a unique aspect of VSS data, which alone or in combination with another property reveals useful information about VSS data or specifically VSS recording.

4.3 Video Metadata describing Personal Information

Here in this section, we will describe which metadata properties describe personal information in the content and how can it be categorized. Generally, in VSS data, observers have access to the content (semantic metadata) of the recording, when it views an object, it can characterize that information in multiple ways.

Metadata	Classification	Descriptive Distinction	Identification or Recognition
Object	Human	Adult/child -> Gender/ height/ color /race	(face, gait) – biometric features Facial recognition or identification *Cross-reference with other datasets to find identity, associations, frequently visiting places, people, workplace, home location, etc.
	Animal	Cat/dog/other -> Color/breed	Cross-reference with other datasets to associate personal information with a human or vehicle
	Vehicle	Cycle/car/van/bus -> Color/ type/model	(License plate) identification feature *Cross-reference with other datasets to find owner, registration date, location, current or last location, past traffic violations, etc.
	Landmark	Building/ check post / traffic sign / name or direction sign ->	names/signs (names, signs, logos) identification feature Cross-reference with other datasets to associate personal information with a human or vehicle
Event	Isolated	Walking/lying/ sitting Pushing/hitting/ throwing	*classification, distinction, or identification of all objects in a particular event/activity
	Interactive	Fighting/vehicle collision Natural Occurrences Fire/smoke/ /flood	
	Activity	occurrence e.g., traffic-related activities Traffic violation -> Accident/ wrong parking/ speeding, etc. Medical emergency Other details of the event (speed, direction, severity)	

Table 4.5: Metadata categorization of personal information

For instance, it can first classify an object, like the object is 'human' or a 'vehicle', further, it can associate different distinctive or descriptive attributes/properties to that classification like the object 'human' is an 'adult' and is 'male' and has certain 'ethnic descriptive features', or object 'vehicle' is a 'car' of a certain 'model' or 'type', etc. Furthermore, objects can also have attributes that can link them to a unique identity or can be used with some other properties together to link to a unique identity, these semantic properties are known as biometric or identification properties. Any property that can help an object link to a unique identity is regarded as personal information. Thus, information about objects or their semantic properties can be further dissected into different levels of information (classification, distinction, identification) that can gradually describe an object and its personal information, as shown in Table 4.5.

Events can also be classified into levels of information, for instance, either an event is 'isolated' or 'interactive' or 'natural occurrence', and then there can be other attributes that can describe details about the event, as shown in Table 4.5 Events in technical do not describe personal information, objects involved in those events do. For example, if an event 'traffic violation' is detected as a semantic metadata property, it doesn't describe personal information, however, the object 'vehicle' with its attribute 'license plate' involved in 'traffic violation' describe personal information, as it can be used to identify the vehicle owner when cross-referenced with another relevant database. Thus, semantic metadata properties can be used to describe most of the personal information present in VSS data.

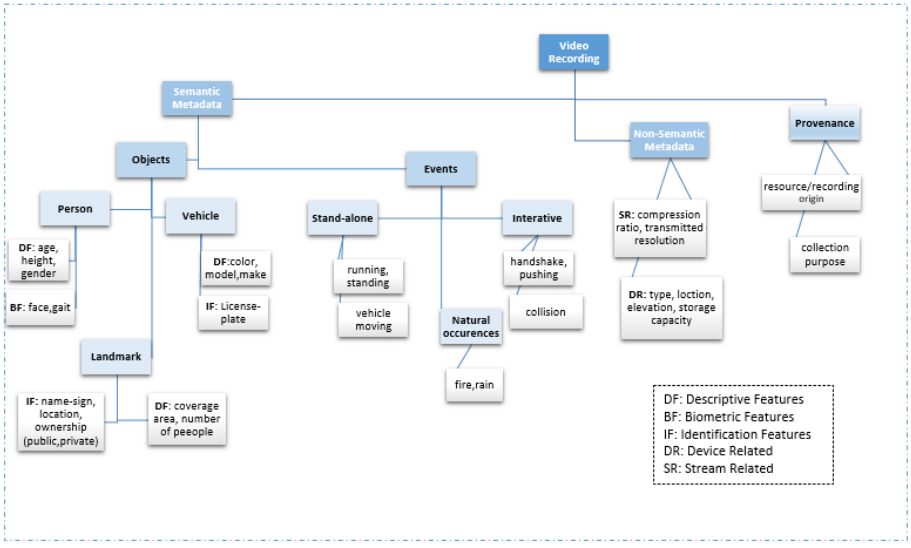


Figure 4.5: Information categorization under different metadata labels

Other metadata types (non-semantic) alone may not describe personal information about any object; however, when used in combination with a semantic metadata property can enrich personal information about an object. For instance, a semantic metadata property can show a certain object involved in an activity, but together with a provenance property 'location of camera' can help deduce the location of the object too that can be regarded as personal information. Thus, different metadata properties evidently or gradually can reveal personal information about the objects detected in the content of the VSS data and can be categorized into different levels of information such as descriptive or biometric features, etc. as shown in Fig. 4.5.

4.4 Metadata Storage and Indexing

As discussed in the above sections, a video recording can have different types of metadata properties extracted from it, which provides sufficient information about the content of the recording along with the context it was recorded in [81]. Once, the information extracted from content is categorized under different labels, it can be used in different ways to refer to a specific recording. For instance, these properties can be used for efficient storage and quick retrieval of video surveillance recordings, i.e., video recordings can be described as a set of different metadata properties, stored as a separate metadata file, along with video recording, as shown in Fig. 4.6. Metadata extraction and indexing type depend on the requirements of the VSS, i.e., what is it going to be used for or what needs to be searched. Typically, video recordings are indexed via some provenance metadata property such as in chronological order of recording-time or by camera-deployment location. For example, search queries can be like a request for recordings stored on 'January 1, 2000, between 0900 to 1100 hours', or recordings from cameras located at area X. In advanced semi-automated VSS, recordings can also be searched via semantic properties such as a particular object or event. For example, search queries can be like a request for recordings that have 'moving objects' in the content, or 'black pickup van' or recordings that detected 'collision' or 'trespassing' as a semantic event. Often, various sensors are coupled with cameras, and their metadata can be analyzed together with VSS metadata. For example, accelerometers can detect when a vehicle crosses a certain speed threshold, and that data property aggregated with semantic properties can be used to view the license plate of that particular vehicle. Cameras can also be coupled with microphones that when register with high decibels values corresponding to that of 'gunshots' or 'screaming' along with VSS metadata can be used to narrow down the acoustic source. Hence, comprehensive, systematized, and precise VSS metadata files provide searching and indexing flexibility for observers and applications with different requirements.

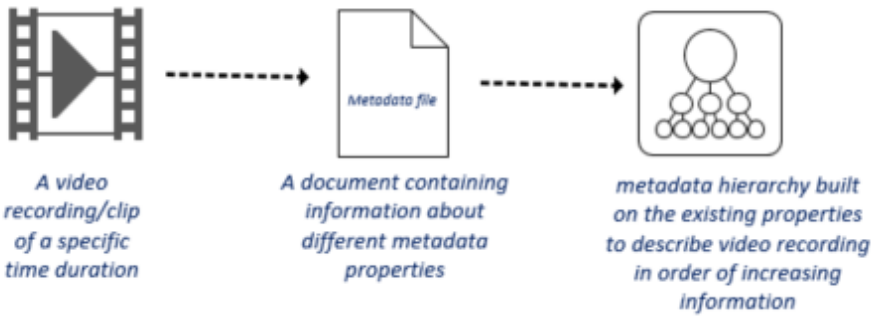


Figure 4.6: Metadata Storage and Indexing

These metadata types and properties are also used as the searching or indexing terms for the video data, which the system supports, as discussed in Section 4.2. Thus, here, these indexing terms can either be non-semantic (like compression ratio, or encryption type, etc.), provenance (deployment location, or recording timestamp), or semantic metadata properties. For semantic metadata-type, indexing categorization is either an object type (which can be described as a noun) or an event type (more closely described as a verb). As can be seen in Fig. 4.5, events and objects are further described by their intrinsic properties. These intrinsic properties (descriptive, biometric features, etc.) can be defined as an adjective to the object and event type. Therefore, at the basic level, video metadata indexing hierarchies are majorly constructed based on the detected objects and events as the first layer, and then subsequent layers can describe their intrinsic properties in detail. More exhaustive the metadata indexing hierarchies, easier for the searching and indexing mechanisms to find the relevant video. This also means that any noun (object type), verb (event type), or adjective (DF, BF, IF, DR, SR) that the indexing mechanism knows how to handle can be included in the access control policy and used in an access control decision.

All these different metadata combined logically give a thorough understanding of what a resource is, and can often be described in the form of an information hierarchy [82]. Metadata is arranged in a way that gradually reveals information about different metadata properties, as described in Fig. 4.7. It reflects a layered model and follows the natural approach of data flow, i.e., each new piece of information adds value to the underlying data, which makes the data more meaningful than it was at the previous layer [83]. Often, metadata hierarchies offer more organized and readily available information than their original resource and require strict data protection measures (similar to the original resource) if it contains sensitive or personal information. Such metadata hierarchies may also allow stronger or weaker inference about objects in the video

recording, or gradually reveal information that may lead to describing their associated personal information. For instance, in Fig. 4.7, the following example illustrates an increasing amount of information about an object in a video recording: "movement", "red object moving", "red car moving", and "F1 Ferrari number 5 is passing". At each step of the mentioned example, a new piece of information (metadata) is added, which has made it significantly clear and more enriched. The different levels of information exposure or level of abstraction, i.e. what information about the objects will metadata reveal and in how many classifications depend upon the requirements of VSS observers.

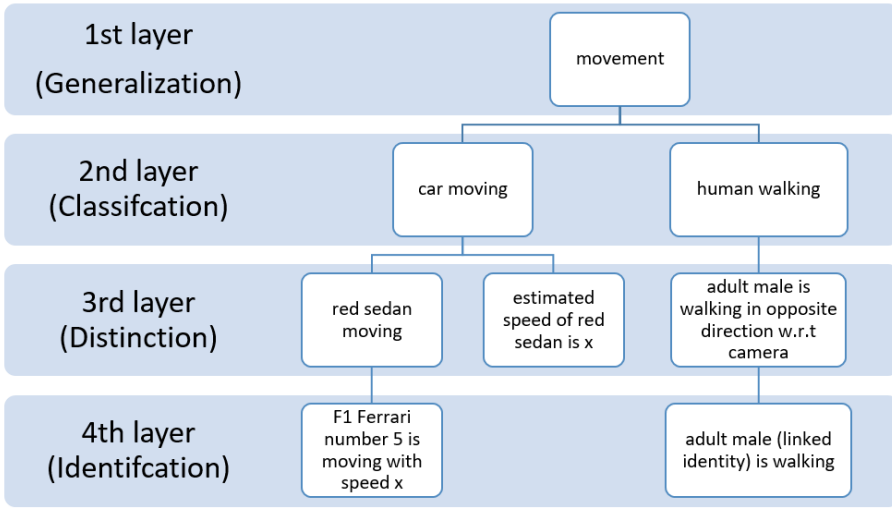


Figure 4.7: Metadata Storage and Indexing

4.4.1 Metadata Hierarchies and Access Control

We propose to make use of such metadata hierarchies to preserve privacy in large-scale VSS by limiting the access of an observer to certain information levels according to its authorized requirements. Based on the mentioned categorization of different metadata properties, the original recording can be subjected to different transformations, and its information can be layered under different labels, where each layer reveals more (personal) information about the objects, which can be used to control the information flow of the content to the observers. For example, a recording may be shown at different information levels to different observers, depending upon their requirements, to control the flow of personal information [23]. For example, an event 'traffic violation' is de-

tected in a video recording, and traffic management (TM) observer requests to view that recording, there can be different possibilities of what should be made available to the observer. For Instance, an observer under normal circumstances can initially be shown layer-1(Generalization) or layer-2 (Categorization) level information, i.e., license plates and human faces can be blurred/hidden (when there is no event of interest to the observer), and so it cannot be linked to the identity of the object. The recording may reveal minimal information about the detected events or objects in the recording, or may only reveal classification information about objects, but not descriptive, biometric, or identification. However, in the case of an event 'vehicle collision', TM can be allowed to see the semantic metadata property 'license plate' of the object 'vehicle' to find its (identification) or revealing the 'face' of the object 'human' (driver) for linking its identity to a unique individual (biometric), thus, can be given access to layer-3 (Recognition) or layer-4(Identification),as shown in Fig. 4.8.

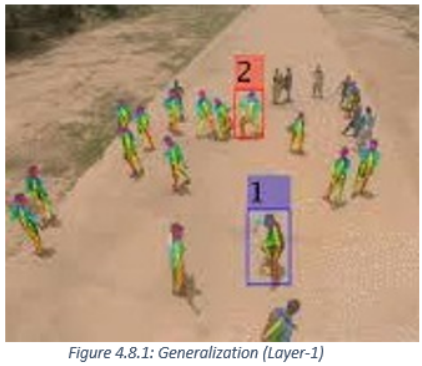


Figure 4.8.1: Generalization (Layer-1)

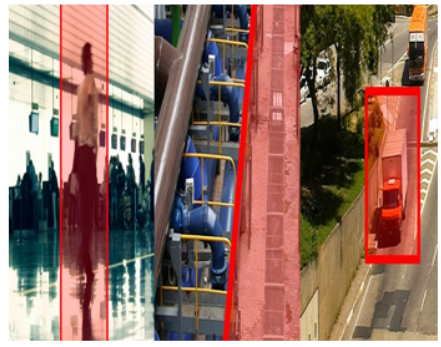


Figure 4.8.2: Categorization (Layer-2)

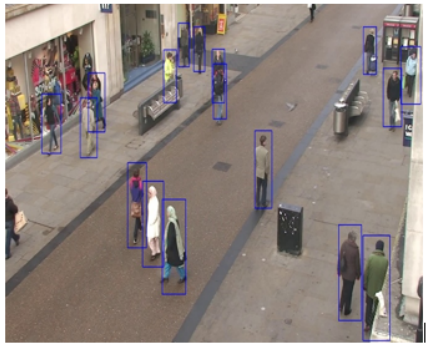


Figure 4.8.3: Distinction (Layer-3)

Figure 4.8.4: Identification (Layer-4)

Figure 4.8: Classification based on Metadata

Other than the object-of-interest (vehicles involved in a collision) and person-of-interest, the rest of the objects in the recording are irrelevant to the existing scenario or the purpose of the TM, so information about them should not be revealed. It means that even within the recording exposure of information about different objects can be shown in different (metadata) layers to a single observer, based on its purpose or authorized data requirements. Often, even with hidden identification and biometric features of the objects, other metadata types (non-semantic or provenance) can also be used to know enough information about the objects to invade their privacy [83]. For example, cameras deployed at location 'town square' recording 'persons' going to building 'Y' regularly at time '0900 to 1600' hours may reveal association among different objects. Therefore, for strict privacy preservation, different metadata properties of the objects can be anonymized or transformed for observers and should be revealed based on the 'principle of least privilege', i.e., access to applicable metadata properties of the recording specific to the purpose of the observer and limit secondary data usage by controlling irrelevant exposure of personal information to the observer.In Chapter 5, we will discuss in detail how different metadata properties can be used in realizing a fine-grained access control mechanism for limiting secondary use and ensuring privacy in large-scale VSS.

4.5 Conclusion

To conclude, this chapter discussed different types of information that can be extracted from video surveillance data that can be of use to observers with diverse requirements. VSS data generally comprises video recordings (source of semantic metadata), device and stream metadata (non-semantic and provenance) with the help of different video analysis and machine learning solutions. These metadata properties can further be explored to label different levels of personal information about detected objects and events of interest. The ACM can make use of different metadata properties to control what part of video recording or what level of information in a video recording should be made available to an observer according to its authorized purpose, and will be further discussed in Chapter 5.

Chapter 5

Metadata-Based Access Control To Limit Secondary Use

Access Control Mechanism (ACM) has three main components: users (observers), resource objects (video recordings), and a reference monitor. The observer is an authorized user, which has distinct attributes (user properties) by which it identifies itself to the system and then requests to access a certain resource (recording). A reference monitor then evaluates these policies against a request (presented by a user), and a request must provide a way (attributes or authorization) that a reference monitor can understand and compare with the properties that are used in describing the conditions and permissions in the policies are defined. Thus, an ACM takes a request (containing the user's properties and resource request parameters) as input, already has resource and its properties, and then evaluates the access policy defined for the user to make an access decision. In this chapter, we aim to present an ACM that regulates access to large-scale VSS resources for observers with a diverse set of requirements that may vary with time. Our focus is on preserving privacy while providing due access to authorized observers without leaving any gap for secondary use. We will address various privacy concerns that are discussed in Chapter 3. A detailed literature review of different ACMs targeting concerns similar to those mentioned in the above chapters about video data and large-scale infrastruc-

tures has been summarized in Chapter 6, but we will very briefly discuss them in the next paragraph to give context to our proposed solution.

Role-based access control (RBAC) and Attribute-based access control (ABAC) are two of the most commonly used ACMs for large-scale information systems. In RBAC, observers with a set of similar user properties (name, ID, location, role-type, etc.) will be given the same level of access to requested resources, regardless of their current conditions [25]. For example, observers of role-type 'parking guards are allowed to watch video recordings only from the cameras installed at location 'parking'. In the case of ABAC, the access level to a resource may vary according to the current requirements of the observer or a resource at a given time. This can be decided by the set properties of the observers (name, role, location, etc.), data resources/video recordings (metadata properties as discussed in Section 4.2), and the current conditions of both at the time of the request [84]. For instance, along with the location 'parking', observers can be further restricted to view recordings, only within their 'duty hours. ABAC solutions are more flexible than RBAC solutions, though the policy mechanism is complex if the dynamic properties and attributes are greater both in number and dimensions (belonging to multiple users and resources), thus making it hard for adoption in large-scale distributed systems [85]. However, on the positive side, ABAC can accommodate many different types of resource properties (metadata, provenance, collection purpose, etc.), that can be defined in (explicit) permissions for a resource policy, thus, it can play an important role in decision-making. On the other hand, RBAC solutions have a fairly easy access policy mechanism, making them a feasible choice for any infrastructure with a defined set of users. However, only using RBAC will not accommodate the varying factors present in large-scale infrastructures where resources are transformed and aggregated multiple times, potentially changing their nature and attributes, thus requiring a change in RBAC access policies defined for users every time. Over the years, several variations of RBAC and ABAC have been proposed according to the different system requirements and some hybrid solutions use both RBAC and ABAC in a logical combination to serve the needs of the system. For instance, identity-based access capability (ICAP), trust-based access control (TBAC), provenance-based access control (PBAC), etc. [85] [86]. It is important to distinguish here between PBAC and provenance access control (PAC). The former uses provenance data to make an access decision, while the latter deals with regulating access to provenance as a resource PAC [53] [56] [51]. Moreover, in PBAC, provenance has a limited capacity to store and represent different properties of both the user (agent) and the resource (entity), therefore, in large-scale infrastructures, it is unlikely that PBAC can be used as the only ACM, and it is often coupled with RBAC or ABAC. One such relevant RBAC–ABAC hybrid solution that is designed according to large-scale infrastructure requirements is Attributes Enhanced Role-Based Access Control (AERBAC) model, a hybrid solution to cater to the large-scale VSS

requirements, is as shown in Fig. 5.1 [23]. It uses RBAC for its role-assigning simplicity in categorizing observers and assigning them minimum default permissions per their role, and then utilizes ABAC for evaluating different resource and environmental properties, thus assigning fine-grained access to observers. In the aforementioned implementation, AERBAC accommodated dynamic user attributes based on the current conditions of the observer, and the set of 'allowed' resources could be different under different values for user attributes, providing context-aware access to the observer. However, it does not consider dynamic resource properties or provenance properties.

To conclude, keeping in view our requirements for a large-scale privacy-aware VSS that prevents secondary use, the ACM also needs to consider different resource properties (resource metadata (cf. Section 4.3), provenance, 'collection purposes' (cf. Section 3.1), etc.) in case of transformations and aggregations along with accommodating an access policy mechanism for users/DPs with emerging requirements, while making a decision. Please note that this chapter may contain excerpts and figures from our own published papers (as part of the PhD research) mentioned and referenced in 1.3.

5.1 Metadata-based Need-to-Know Access Control Framework

Privacy is a key requirement for a large-scale VSS, as it has a large amount of data containing personal information that is to be used for authorized 'collection purposes'. However, due to the complex nature of VSS data, specifically video recordings, it is often hard to impose privacy measures that leave no room for secondary use, as observers are exposed to a lot of irrelevant data, leading to privacy violations. To minimize the secondary use of personal information, understanding of the VSS data or specifically its content to classify different types of personal information is important so it can be analyzed, how much of the recording should be exposed to the observer (cf Section 4.3) (rephrase). In order to achieve that, purpose limitation needs to be ensured. Though, when implemented at a large scale (let's say city-scale like smart cites), there may be a lot of DC contributing data to VSS, and a lot of observers requesting VSS for data relevant to their 'access purpose'. This raises concerns about how the observer's 'access purposes' can be verified against the 'collection purposes' of the resources contributed by different DC, as discussed in Section 2.3 if purpose integrity is not observed.

To address the concerns described above and in our motivation example (cf. Section 3.1), we propose a privacy-aware and a need-to-know view ACM for the

large-scale VSS by using different types of metadata obtained from VSS in access control decision. We will extend the AERBAC model, to include semantic and provenance metadata of the VSS data (or resource) as part of its conditions, and permissions, so it can be compared with the observer's requirements. As discussed in Section 4.2, every resource has a set of metadata properties, of which some are often used as indexing and searching parameters, for instance, non-semantic properties like time or location of the recording, or semantic properties like events or objects of interest (EOI or OOI). Therefore, it is common that the observer's request contains different types of metadata properties to help systems classify a particular resource, which is then evaluated by the reference monitor against the available policies and permissions defined by the VSS. We propose that these metadata properties or parameters be part of the 'access purpose' of the observer, which can then be verified against the similarly described 'collection purpose' of the resource (cf. Section 2.1). AERBAC in its previous implementation only uses non-semantic properties of VSS data to decide whether an observer is allowed to access or view a certain recording or not. We will extend AERBAC with a three-step content control to implement a need-to-know view. Once, the observer is authorized, at the first step, based on the physical parameters (time and location) of the observers and the similar non-semantic properties of the resource, will decide whether the observer with its current contextual properties can view the recording or not. At the second step, purpose integrity and purpose limitation is ensured by first retrieving the 'collection purpose' from the provenance metadata, which will provide the legally allowed EOI for which that particular resource can be accessed. If the 'access purpose' of the observer has a similar EOI, then here, the part of the resource (recording) containing that EOI will then be forwarded/moved to the third layer. The third step aims to minimize the secondary use, so it based on the verification of 'collection and access purpose' decides about the exposure of personal information of OOI (objects) present in the authorized EOI, and the rest of the irrelevant personal information in the resource is hidden from the observer.

The next subsection will briefly discuss the previous AERBAC model and its characteristics, and after that, the next section will present the extended AERBAC.

5.1.1 Attributes Enhanced Role-Based Access Control (AERBAC)

AERBAC is a hybrid RBAC-ABAC approach used for the dynamic assignment of resources to observers based on policies defined in terms of user attributes, object attributes, and environmental attributes, as shown in Fig. 5.1. It has

been previously proposed and discussed in great detail in the pioneer paper [23] [9]. Here, we will briefly describe its basic concepts and then focus more on the novel extension that addresses the above-mentioned concern in Section 5.2.

5.1.1.1 User and Resource-Object Attributes

Video recordings have a set of distinctive metadata properties, which are categorized into different levels (semantic>object>descriptive features, etc.), discussed in Section 4.3. These properties are called resource-object attributes, referred to be OATT. Observers also have their own set of unique properties that they use to identify themselves with the system. These properties can be a name, department, designation, duty-hours, role, etc., and are referred to here as UATT. The role is a significant UATT in AERBAC, as it is used for assigning the requisite permission against requesting various resource objects (video recordings). For both observer and video recordings, some properties are fixed, and only system administrators change their value/information. Moreover, there are other properties, whose value is dependent upon the system or environmental changes; these are known as environmental attributes, EATT, or contextual attributes [67]. For example, an observer can only access a video recording, if the "location" of the resource is within the perimeter of the "authorized area" as well as matches the 'current location' of the observer. Here, the observer's "current location" is not a static property and can keep changing, so before processing this request; the system needs to know the environmental or contextual value of this property. The same permission will sometimes result in "allowed" and sometimes "denied" based on the contextual value of that UATT, which is also an EATT. Usually, the value of EATT is checked at the time of evaluating the request, and if it is according to the policy, access is granted. However, within the duration, while the user has access to the resource, and the value of EATT changes, the permission is not revoked. Here, we are also integrating the concept of 'continuous enforcement', which allows the reference monitor to continuously check the value of EATT, and if it changes then permission to the user should be revoked. For example, if the "current-location/area" of the user changes, and is different than what is permitted, then the access to the video stream should be revoked instantaneously.

5.1.1.2 Permissions and Policies

For every observer (User), based on the (UATT) and EATT, it is assigned a dynamic role via user role assignment (URA). The role is bound to the user for a specific session in which a certain request is made. Against every assigned

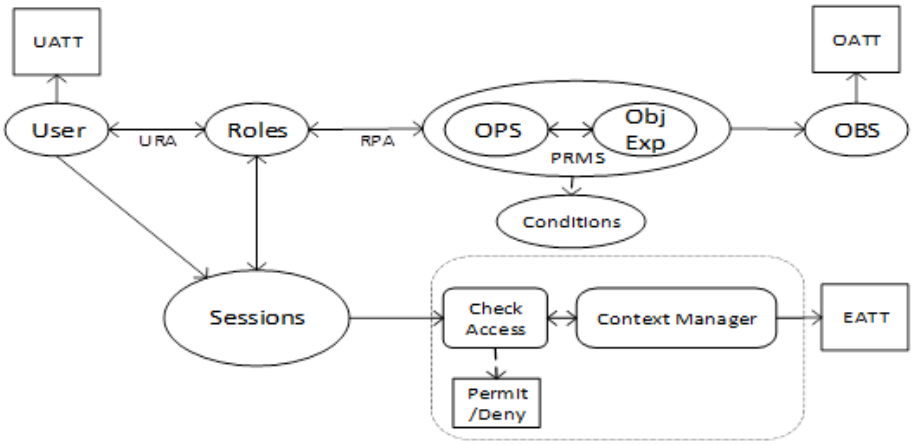


Figure 5.1: AERBAC Model

role, there is a set of permissions (PRMS), which consists of an object expression (Obj. Exp.) constructed from several video recording (OBS) attributes (OATT), and an authorized operation (OPS) on that resource object/ video recording, as shown in Table 5.1. Each permission is linked with a few numbers of conditions, which must be evaluated to be true for the user to exercise that permission, and the Role Permission Assignment (RPA) captures this relation, as shown in Table 5.2. A condition associated with permission may contain properties of all entities including users (UATT), objects (OATT), and the environment (EATT). Permissions and conditions are defined in XACML policy files against different roles of observers, which are evaluated by the access control module, to allow access to observers, as shown in Fig. 5.1 [23].

5.2 Extended AERBAC

We extend AERBAC to implement a need-to-know view based on the authorized EOI if present in both the 'collection purpose' of the requested resource and the 'access purpose' of the observer's role, as shown in Fig. 5.2. Each recording (OBS) when processed generates different types of metadata, and here we propose to organize that metadata in a hierarchy (of increasing information) under different categories, as discussed in Section 4.3. First, the semantic metadata (information extracted from literal content) is divided into two main categories: events and objects (involved in those events). An event can have more than one object, and an object can be a part of more than one event. These events and objects then become part of the semantic metadata hierarchy and are treated as

EATT, as they may vary from recording to recording. The 'collection purpose' of the resource is regarded as OATT, as it seldom changes value, once it has been associated with a particular resource. However, if an observer requests a resource, it will only be authorized to access the personal information about the objects that are part of the events (referred to as OOI) permitted in its 'access purpose' as EOI, moreover, similar events should also be a part of 'collection purpose' as EOI. This means when an observer X requests a resource Y, it must present three things to the AERBAC reference monitor, its physical parameters (time, location, etc.), EOI from the 'access purpose', and particular personal information (properties) allowed from objects (OOI) that are part of those EOIs. The AERBAC, then at the first layer, compares the physical parameters with the non-semantic OATT of the resource, secondly, compares the EOI of 'access and collection purpose', and thirdly, allows the observer to view the OOIs (in original mode) in the recording where rest of the objects are hidden/blurred, limiting the flow or exposure of personal information [28].

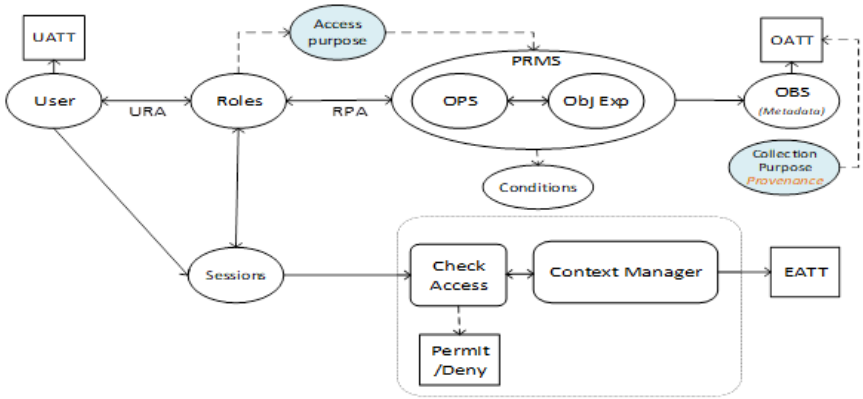


Figure 5.2: Extended AERBAC Model

The proposed or extended modification helps in the assignment of coarse-level permissions and object expression in terms of semantic events and objects against different roles and their 'access purposes', as shown in Tables 5.1 5.2. Below is a simple example to differentiate the previous and extended implementation of AERBAC.

The second modification is the add-on of a privacy layer that hides the irrelevant personal information of all the other semantic objects that are not involved in EOIs. At this step, there are three viewing modes available for the observer, and an observer may view different parts of the recording in different modes, or specifically can view relevant or object involved in permitted EOIs in their unmodified form, while irrelevant objects will be hidden or blurred or will be available for view in privacy-away mode. Here, the privacy layer has three

modes: privacy-aware, descriptive, and biometric. The privacy-aware mode will hide (transform) all types of personal data in the video recording, the descriptive mode will allow to reveal of the descriptive features of the OOIs, while biometric mode will allow revealing the biometric features of the OOIs. For the full disclosure or access to the original or raw VSS data, the observer must be allowed to have permission for both descriptive and biometric modes. Thus, modified AERBAC has two privacy levels; the first one determines which events are EOIs for a particular observer, so the time-slice of the requested recording during which the EOI occurred can be selected for view. In the second privacy level, based on the purpose of the observer against the EOI, the apt privacy level can be decided so only authorized OOI (prime or potential) are made available for view to the observer. Hence, the first privacy level deals with EOIs to slice the recording for the relevant portion, while the second deals with different privacy levels regarding OOIs during that portion. The rest of the irrelevant personal information in the recording is in privacy-aware mode, i.e., transformed/blurred to hide the descriptive and biometric features. The extended AERBAC approach is applied to the case study described in Section 3.1 and is described in the sub-section below.

	AERBAC	
Sr.No	Search Queries	Object Expression (Obj. Exp.)
1	All the video recordings in area X from January 1, 2019- January 15, 2019	(“X” CONTAINS cam=area(o)) ̂(timestamp(o) AFTER 2019.01.01 00:00:00 BEFORE 2019.01.15 00:00:00)
2	Video-recordings with location “East Highway” in area X from 1 PM TO 4 PM on March 10, 2019	(Loc-type(o)=” East Highway”) ̂(timestamp(o) DURING 13:00:00 2019.03.10 - 16:00:00 2019.03.10) ̂(“X” CONTAINS cam=area(o))
	Extended AERBAC	
1	Video-recordings with location “Downtown” that contains “traffic-violation” in the semantics events during the last weekend	(Loc-type(o)=” Downtown”) ̂(semanticEvent(o) INCLUDES “traffic violation”) ̂6,7day.week
2	Video-recordings with location “East Highway” that contains “Speeding” in the semantics events and contains descriptive features (object type= vehicle, colour = Black, vehicle type = SUV) in the semantic object extracted from the recordings in the current time	(Log-type(o)=” East Highway”) ̂(semanticEvent(o)->standalone INCLUDES “speeding”) ̂(semanticObject(o) INCLUDES DF (object type= vehicle̊colour = Black v̂ehicle-type=SUV) “Black SUV”) ̂(timestamp(o) current.timestamp)

Table 5.1: Object Expressions of AERBAC and Extended AERBAC

As it can be seen from Table 5.1 5.2, there are two main modifications, first, semantic metadata properties are now also defined as EATT, and have become

part of the conditions that need to be fulfilled, (along with the non-semantic conditions part) for the permission to be granted. This will help to limit the access of the observer (user, strictly to those video recordings that have "events" corresponding closely to the "permitted EOIs" of the assigned role. Permitted EOIs will become a part of the "collection purpose" as well as the "access purpose" of the user (UATT), which will allow property-level compliance. For instance, if a user has a role "firefighter", then its "access purpose" can describe that it is authorized for requesting a resource that has "fire" as an allowed resource property or EOI to be specific as a "semantic event->natural-occurrence".

	AERBAC	
Role	Conditions	Permissions
Police Dept. Responder	"Location" of the video-recording (OATT) is within the perimeter of the "current-area" of the User (EATT) AND "timestamp" of the video-recording (OATT) is within the range of "duty-hours" of the role (EATT)	If conditions are fulfilled, then the user can view the requested video-recording
	Extended AERBAC	
Police Dept. Responder	(Non-Semantic) "Location" of the video-recording (OATT) is within the perimeter of the "current-area" of the User (EATT) AND "timestamp" of the video-recording (OATT) is within the range of "duty-hours" of the role (UATT) (Semantic) Semantic Events of the requested video-recording (OATT) includes "designated Events-of-Interest" in the user's assigned role (UATT)	If conditions are fulfilled, then the user can view the "Objects-of-Interest" in "biometric" mode (with the rest of the recording in "privacy-aware" mode)

Table 5.2: Metadata categorization of personal information

The second modification is the add-on of a privacy layer that hides the irrelevant personal information of all the other semantic objects that are not involved in EOIs. At this step, there are three viewing modes available for the observer, and an observer may view different parts of the recording in different modes, or specifically can view relevant or object involved in permitted EOIs in their

unmodified form, while irrelevant objects will be hidden or blurred or will be available for view in privacy-away mode. Here, the privacy layer has three modes: privacy-aware, descriptive, and biometric. The privacy-aware mode will hide (transform) all types of personal data in the video recording, the descriptive mode will allow to reveal of the descriptive features of the OOIs, while biometric mode will allow revealing the biometric features of the OOIs. For the full disclosure or access to the original or raw VSS data, the observer must be allowed to have permission for both descriptive and biometric modes. Thus, modified AERBAC has two privacy levels; the first one determines which events are EOIs for a particular observer, so the time-slice of the requested recording during which the EOI occurred can be selected for view.

In the second privacy level, based on the purpose of the observer against the EOI, the apt privacy level can be decided so only authorized OOI (prime or potential) are made available for view to the observer. Hence, the first privacy level deals with EOIs to slice the recording for the relevant portion, while the second deals with different privacy levels regarding OOIs during that portion. The rest of the irrelevant personal information in the recording is in privacy-aware mode, i.e., transformed/blurred to hide the descriptive and biometric features. The extended AERBAC approach is applied to the case study described in Section 3.1 and is described in the sub-section below.

5.2.1 Extended AERBAC and SC-VSS

Continuing with the scenario described in Section 3.1, here we will see how AERBAC can be applied to it to achieve a need-to-know view, as shown in Fig. 5.3. Multiple potential observers (traffic department (A), police department (B), and the fire department(C)) request a certain recording. The requested video stream/recording is processed and analyzed to extract different types of information in terms of various metadata properties. After the categorization of those objects and events, these labels become part of the metadata file and are further arranged into an information hierarchy. Metadata properties that are fixed become part of OATT, which includes most of the non-semantic properties, and semantic objects related properties, and the 'collection purpose' (extracted from provenance metadata). Some of the properties, whose values may vary with external factors like semantic events, become part of EATT. Each of the observers has a set of UATTs, and is interested in some specific information regarding the specific EOI, as described in Table 3.2 in Section 3.1, which is now described in terms of "allowed EOIs" in the 'access purpose' per their role. All these attributes (UATT, EATT, OATT) including the role and its 'access purpose' of the observer are then passed to the AERBAC reference monitor for request evaluation. After authentication of the observer per given role, a unique

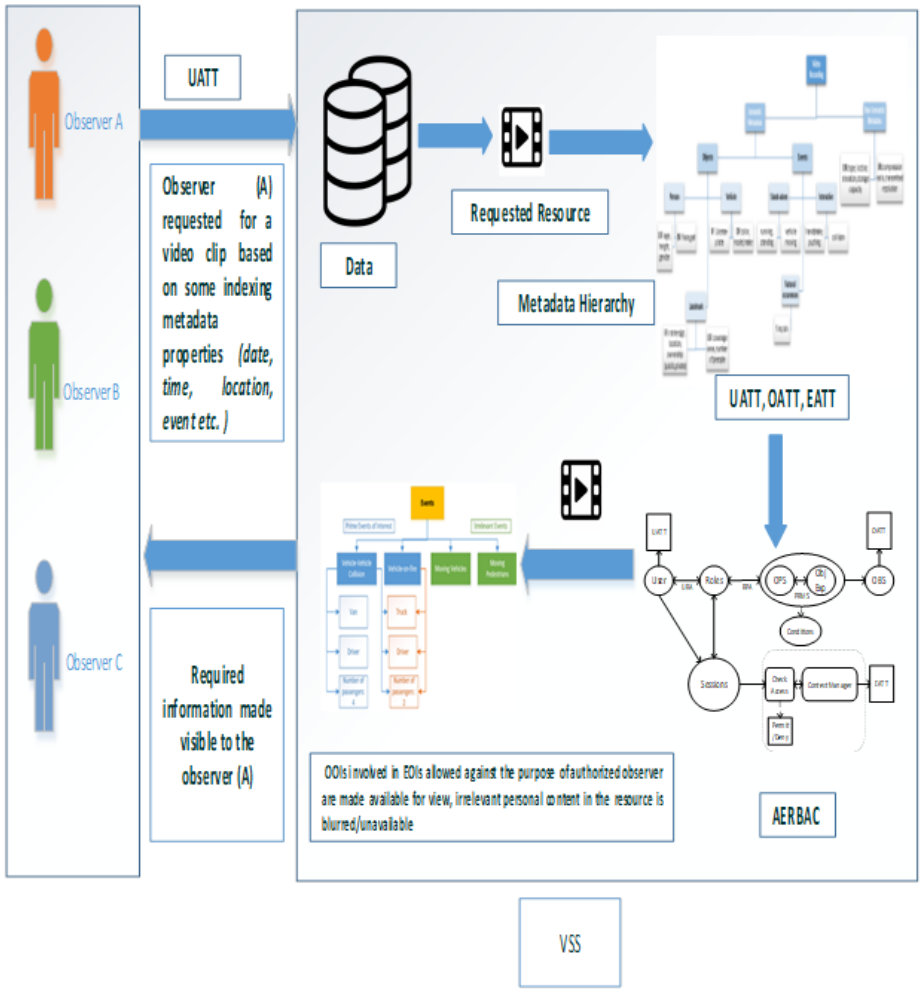


Figure 5.3: Extended AERBAC and SC-VSS

session is allocated to the observer. There is a set of permissions (PRMS) consisting of authorized operations and object expression (consisting of UATT, OATT, EATT), shown in Table 5.2, to every role. These permissions per role are then evaluated by the 'check access' module of the ACM for a particular session, where 'context manager' checks the current values of EATT mentioned in PRMS, i.e., events, compares the 'collection and access purpose', and then an access decision is given in terms of the allowed view mode for the relevant OOIs, and rest of the objects are hidden. All the above-mentioned steps are shown in Fig. 5.3.

A resource can have more than one 'collection purpose', and each 'collection purpose' will have its own independent characteristics. As VSS collects data under a 'public interest' legal base thus can have a 'collection purpose' related to public safety, emergency services, traffic management, traffic law enforcement, real-time traffic updates, route planning, and congestion handling, etc. Below is the example of two of such 'collection purposes', Traffic Law Enforcement, described in Table 5.3, and Public Safety, discussed in Table 5.4, along with examples of object expressions relevant to them. As described in Section 2.1, the 'collection purpose' has five characteristics. The first characteristic declares the data attributes/properties (OATT) of the resource (OBS) that holds personal information. Secondly, a set of properties are mapped to a specific purpose or function, describing input and output properties. Thirdly, for every function, a set of compliance policies have been defined that need to be followed if the properties bound to that function are to be accessed. Lastly, it records aggregation limitation, if there are explicit conditions to be followed in the case of resource transformation or aggregation. It also stores the legal frameworks supporting the collection purpose. The second column of Table 5.3 and 5.4 show some examples of object expressions that can be used for requesting specific recordings. It is important to note here that only OATT that may store personal information is mentioned in 'collection purpose'.

Collection Purpose for OBS (VSS recording) Traffic Law Enforcement	**Obj. Expression** (examples)
1.Personal Data Properties OATT : Object-type (human, vehicles), object-descriptive-features (gender, color, estimated-age and height), object- identification-features (face, gait, license-plate), geo-location data (spatiotemporal position of any object at a specific time) **2.Personal data property to function Mapping** (Vehicle's License plate, driver's face) -> (are bound to functions traffic light violation, Speeding vehicle, Wrong parking, Wrong turn, Driving in a bus lane, Junction-box violation) (Vehicle's License plate, driver's and passengers' face)-> Accident, Vehicle collision, Seat belt, child detected without a child seat, etc. *An exhaustive list is defined for different agreed-upon functions and are bind to the required OATT describing personal information **3.Compliance Policy** will be used for public interest reasons: -To record, process, and store (activities) any event or object that demonstrates a Traffic operations or violation (traffic light violation, Speeding vehicle, Wrong parking, Wrong turn, Driving in a bus lane, Junction-box violation, Accident/ Vehicle collision, Seat belt, child detected without a child seat, etc.) -To record, process, and store any event or object that demonstrates passenger handling, incompliance to traffic regulations, hinders/stops the routine or smooth traffic operations (function/sub purpose) -To record, process, and store events and object involved in routine traffic operations function/sub purpose) -To record, process, and store events and object involved in parking management function/sub purpose) **Cannot be used for tracking any event or object that is not mentioned in 'purpose' unless otherwise authorized by another legal base or higher authorized DC **4. Aggregation Limitation** The said resource when aggregated with any other resource requires specific authorization from public-authority DC supported by a legal base of Consent, Legal obligation, or Vital Interest if used for the following functions/sub purposes. -Link a license plate (OATT) to a unique DS ID, name, face (OATT) -Link the descriptive features (OATT) of a human to the identification features (OATT) of a unique DS, -Link the descriptive- features (OATT) of a human to the geo-location features (OATT) of a unique DS, **5.Legal Base** : Public Interest	**Example 1** Description : Video-recordings that contains event-type "Speeding" (EATT) at location (EATT) "east highway" at the current time (EATT) Formal: Loc-type (EATT)= "East Highway") ∩ (Event (OATT) INCLUDES "speeding") ∩ (timestamp (EATT) current. Timestamp) **Example 2** Description Insert fine for licensed-owner of the vehicle with event-type "Speeding" (EATT) Formal OBS event-type "Speeding" (EATT) ôbject-identification = "license-plate"-> Operation (OPS) INSERT fine (OATT) FOR "license-plate-> licensed owner" = "licensed owner"

Table 5.3: Collection Purpose Description of Traffic Law Enforcement

Collection Purpose for OBS (VSS recording) Public Safety	**Obj. Expression** (examples)
1.Personal Data Properties OATT : Object-type (human, vehicles), object-descriptive-features (gender, color, estimated-age and height), object- identification-features (face, gait, license-plate), geo-location data (spatiotemporal position of any object at a specific time) **2.Personal data property to function Mapping** (human's biometric features) -> (are bound to functions critical events – brawl, battery, burglary, robbery, destruction of property, property theft, trespassing, fighting/violence, damage to health, break-in, arsenic activity, vehicle theft,) (human's descriptive features) -> (are bound to functions minor violations, crowd gathering, rallies, protests) (Vehicle's License plate, driver's face)-> Accident, Vehicle collision *An exhaustive list is defined for different agreed-upon functions for both critical and minor public safety events, and are bind to the required OATT describing personal information **3.Compliance Policy** will be used for public interest reasons: -To record, process, and store (activities) any event or object that demonstrates critical public safety event like the destruction of property, trespassing, fighting/violence, the person laying on the ground, break-in -To record, process, and store (activities) any event or object that demonstrates minor public safety event like public gatherings, rallies, protests **Cannot be used for tracking any event or object that is not mentioned in 'purpose' unless otherwise authorized by another legal base or higher authorized DC **4. Aggregation Limitation** The said resource when aggregated with any other resource requires specific authorization from public-authority DC supported by a legal base of Consent, Legal obligation, or Vital Interest if used for the following functions/sub purposes. -Link the descriptive features (OATT) of a human detected in minor violation to the biometric features (OATT) of a unique DS -Link the descriptive- features (OATT) of a human detected in minor violation to the geo-location features (OATT) of a unique DS **5.Legal Base**: Public Interest	**Example 1** Description : Video-recordings that contains event-type "trespassing" (OATT) at a location (EATT) "town-museum" from December 1, 2020-December 15, 2020 (EATT) Formal: Loc-type (EATT)=" East Highway") ∩(Event (OATT) INCLUDES "speeding") ∩(timestamp (EATT) " ∩(timestamp(o) AFTER 2020.12.01 00:00:00 BEFORE 2020.12.15 00:00:00) **Example 2** Description : Search identity for the object (human) in an event-type "destruction of property" (EATT) Formal: OBS event-type "destruction of property" (EATT) ôbject-identification = "biometric features"-> Operation (OPS) SEARCH identity (OATT) FOR "ID-> biometric (face)" = "citizen"

Table 5.4: Collection Purpose Description of Public Safety

Similarly, an observer can have more than one 'access purpose' related to different resources, and each 'access purpose' will have its own independent characteristics. As VSS collects data under a 'public interest' legal base thus DC can authorize different 'access purposes' related to public safety, emergency services, traffic management, traffic law enforcement, real-time traffic updates, route planning, and congestion handling, etc., to its observers/DP Below is the examples of 'access purposes', for Observer (A) and Observer (B), discussed in Table 5.5 and 5.6 respectively.

Access Purpose	**Traffic Law Enforcement DP**
Authorized Resource	Video Surveillance Recordings, Vehicle Registration Database, Real-Time Updates From Traffic-Related Sensors
Aim	To enforce traffic laws by capturing and processing any event or incident that results in a traffic violation and issue fines and penalty points on a license based on identification from the intended resources, where applicable
Authorized Resource Requirements	1.Detect and identify traffic event that is considered a violation either via video recording or senor-reading (e.g. speeding) 2.Identify the object-type vehicle (through its license plate or driver's identification information) from video data, and in case of a detected traffic violation issue a fine and penalty points to the object-type driver, if applicable. *mention an exhaustive list of all the traffic operations and violations that are supported by the DC responsible for the available resource-objects
DP Authority Period	Jan-1-2020 to Jan-1-2021

Table 5.5: Access Purpose 'Traffic Law Enforcement Observer'

There may be multiple semantic events occurring in a recording (OBS), against authorized permissions, if the recording consists of lets us say 20 minutes, then the chunk of recording that has 'EOI' or 'OOI' (per se 10 minutes) in it, will only be made available for the view. This here is the first step to limit the flow of irrelevant information to the observer. For the selected time period, OOIs involved in authorized EOIs will be made available according to the given privacy level assigned in the 'collection purpose'. In this case, 'vehicle-vehicle collision 'and 'vehicle on fire', which are allowed against the observer's (A, B, C), authorized EOI, and are prime EOIs. The biometric and identification features of the driver and the vehicle respectively are essential to the 'access purpose' of observers A and B, so OOIs are considered prime. Biometric information about passengers is irrelevant to observer A, so it should not be available for view, while it can be

of use to observer B if the event becomes disputed or objected, i.e., OATT event changes to EATT EOI, then passengers can be used as witnesses, so they are potential OOIs here, so the descriptive information is available to observer B, but not biometric. If the biometric/identification information about potential OOIs is required, then it may require further authorization, as mentioned in policies.

Access Purpose	**Police Law Enforcement DP**
Authorized Resource	Video Surveillance Recordings, National Citizen Database, Vehicle Registration Database
Aim	To ensure public safety by capturing and processing any event or incident that results in minor or critical public safety law violations and take preemptive and post-incident measures to handle such events based on identification from the intended resources, where applicable
Authorized Resource Requirements	1.Detect and identify public safety event that is considered a violation via live or archived video recording 2.Identify the object-type human and its identification in the critical public safety event from the intended video recording, 3.Identify the object-type human and its description in the minor public safety event from the intended video recording *mention an exhaustive list of all the public safety events that are supported by the DC responsible for the resource-objects
DP Authority Period	Jan-1-2020 to Jan-1-2021

Table 5.6: Access Purpose 'Police Law Enforcement Observer'

Another event of interest EOI "fire" is detected so that a particular part of the video recording is authorized for access to observer C (fire department), as shown in Fig. 5.4. The purpose of observer C is to "contain the fire", so it might need to know the source of fire, area perimeter under fire, and if there are objects in that perimeter. As the 'truck' caused the fire, so it is the prime OOI, so descriptive information will be available for view. The fire might affect the 'driver' inside the 'truck', or there may be other objects within the perimeter covered by fire, then DF of these OOIs will be available to observer C. It does not need to know the identity of the OOIs, for its authorized 'access purpose' i.e., ("rescue objects under fire") so BF and IF will not be made available to her.

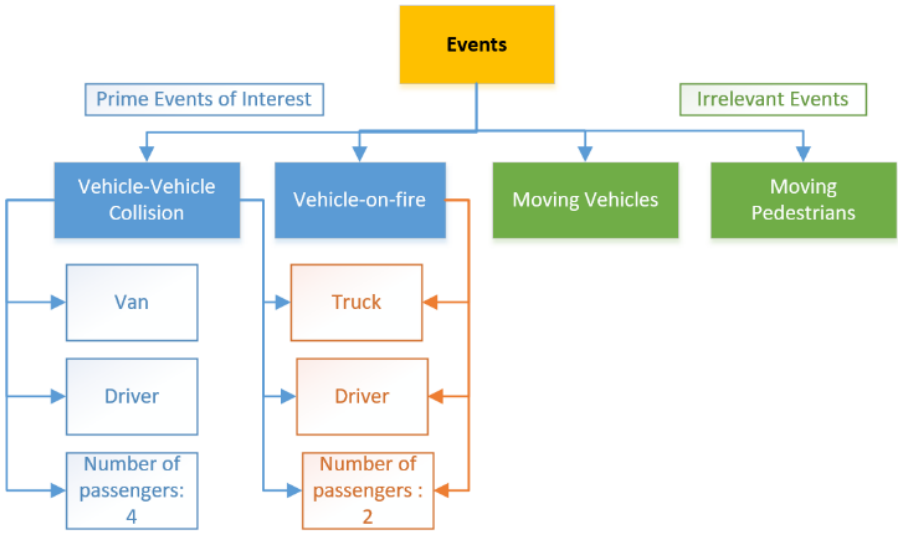


Figure 5.4: Events in a semantic metadata hierarchy

Figure 5.5: Privacy Levels in modified AERBAC

To summarize, the content in the video recording (OBS) to the observer is first restricted based on the prime EOIs according to the authorized 'access purpose' of the observer. It means that observers are only authorized to view the parts of the recordings (OBS) that have their authorized EOIs in them. The second privacy level determines how much of the OOI's' personal information is sufficient for the observers to do their task according to the 'collection purpose' of the OBS, as shown in Fig 5.5. As the VSS data is collected under public interest, its 'collection purposes' contain certain terms or specificities about what an observer is allowed to view (either descriptive or identification features) of the OOIs observer for a specific event or EOI. An observer may not require having all the personal information about an object, and only a specific piece of information. For example, observer A requires information about the driver to issue a fine/offense on a license, therefore, in principle; the observer only needs to get the BI of the driver and not the passengers or IF of a vehicle. Observer B requires information about the driver's identity only (if someone is hurt) or about the vehicle's identity if not adhering to the law, to press legal charges. In the case of observer C, it does not require IF of either vehicle or driver but may require DF information, e.g. (type, height, width) of vehicle, the number of passengers, etc. for the rescue operation.

5.2.2 Aggregated Resources and Collection Purposes

In this section, we will discuss a tangential concern to the above-mentioned solution. In large-scale VSS, due to the diverse nature of observers' requirements, VSS resources are likely to be aggregated with different resources managed by various DC. For instance, continuing the example discussed in Section 3.1, the observer (A) can request an aggregation of VSS data with national citizen database to find the identity of an OOI that was involved in an EOI 'trespassing', only if it is authorized in the 'collection purpose' of both the resources. Often, in large-scale distributed information sharing infrastructures, due to frequent transformational changes in resources data, and emerging requirements of the observer, it may require aggregation of resources with no prior context. For example, observer (B) as a part of the traffic-law enforcement system is allowed to access both the vehicle registration database and the video surveillance system separately in order to access the required data (about traffic violations) relevant to both resources' 'collection purposes'. Let us assume a case in which observer (B) has requested an aggregation of the above-mentioned resources. Theoretically, observer (B) should be allowed to access the aggregated resource for its already authorized 'access purpose' that is an overlap of the 'collection purposes' of both involved resources, and thus this use would not be a privacy violation or secondary use. However, the SC-VSS, in order to limit secondary use, does not let the DP access the aggregated resource, as either the aggregated

resource has undefined 'collection purpose' or requires a DC to authorize a new 'collection purpose' every time their contributed resource is transformed, even when the access is being requested by an already authorized DP. To avoid the need for continuous authorization by DCs or the declination of a rightful request of a DP, it is necessary to preserve the 'collection purpose' of the resource in a way that ensures it is readily available when the resource is requested, even if it is transformed or aggregated. We addressed this issue in Section 2.2, where we proposed that 'collection purpose' be recorded and preserved as part of the provenance metadata, which is then later used in extended AERBAC when making a decision.

Furthermore, in city-scale infrastructure with cross aggregations between different data sources managed by different DCs, it is often hard for any DC to be aware of the current and future 'access purposes' of thousands of DPs and manage their authorizations continuously. We addressed this concern in Section 2.2, where we suggested how an implicit 'collection purpose' can be derived from the 'collection purposes' of parent resources without violating privacy. We proposed that observer who has access to parent resources for a certain 'collection purpose' can access the aggregated resource for a 'collection purpose' that is common between the resources, and the observer has similar authorization in its 'access purpose'. Thus, if the resource preserves the integrity of that 'collection purpose' as part of its indisputable metadata (provenance), it can assist the SC-VSS in establishing that the 'access purpose' complies with the 'collection purpose' and thus ensures purpose limitation. Moreover, in the case of different data aggregations, where the same DC does not manage resources or 'collection purposes' of the involved resources, DPs can use the preserved 'collection purposes' derived from the provenance for relevant or compatible 'access purposes' without violating purpose limitation.

Let's see an example, an observer/DP with a role "traffic law enforcement system" has a UATT 'access purpose' that allows it to access VSS recordings and another resource VRD (vehicle registration data) separately. It may request to aggregate both resources to generate a new resource that matches the license plated detected in VSS data with the registration information of vehicles in VRD. The observe may access the aggregated resource for the 'access purpose' such as "issue fine for a traffic violation", as this is allowed in the implicit 'collection purpose', i.e. both VSS and VRD have a common 'collection purposes' similar to "issue fine for a traffic violation". An observer may request the aggregated resource to be used for an 'access purpose' relevant to the 'collection purpose' of VRD that is "registration of a new or unregistered vehicle". However, the observer, in this case, is not allowed to see the metadata contributed by VSS, i.e., an observer cannot use/view metadata from VSS recording showing an unregistered vehicle and use it to register a new vehicle in VRD, as this is not allowed in the 'collection purpose' of VSS recordings. Thus, an observer

can use the aggregated resources for 'access purposes' relevant to the 'collection purposes' common in all the parent or involved resources, unless otherwise defined [28].

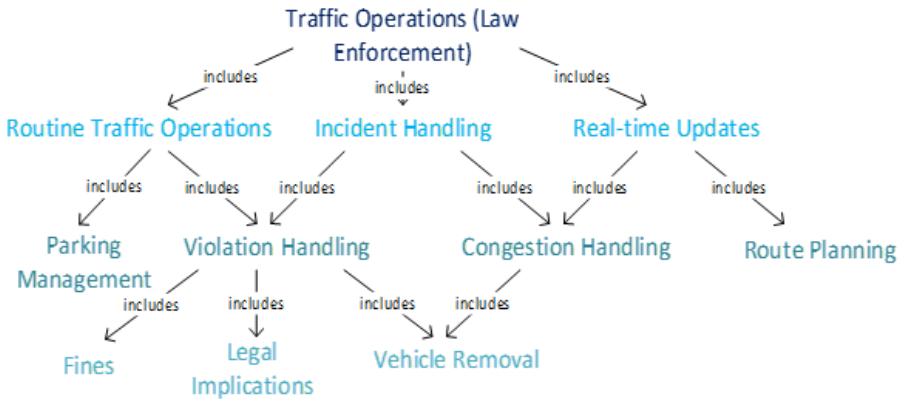


Figure 5.6: Implicit collection purpose hierarchy for traffic-law enforcement

It is possible that a common 'collection purpose' is not available, as in the above-mentioned example. Often in the case of aggregations or transformations, there is some cohesion or similarity that acts as a motivating factor for the combination of the data for the enrichment of the existing information [26] . For instance, in an SC-VSS, most of the 'collection purposes' are regarding public interest operations like public safety or traffic operations and management, etc. In such cases, a basic implicit 'collection purpose' hierarchy can be created for the aggregated resource, as shown in Fig. 5.6. 'Collection purposes' can be arranged in order of the highest number of OATTs describing personal information to the lowest number of OATTs with personal information. Thus, even if the resources do not have a common 'collection purpose' at the same level of the hierarchy, an authorized observer/DP with a valid 'access purpose' may be allowed to access the aggregated resource for a purpose lower in the hierarchy. It will limit the exposure of information to the DP and yet allow them to access the OBS without redefining permissions for this DP. Hence, extended AERBAC allows a DP with existing 'access purpose' to use transformed or aggregated resources without violating any 'collection purposes'. For instance, if an aggregated resource has a set of 'collection purposes' containing "routine-traffic operations", "incident handling", and "real-time updates", then a DP with an 'access purpose' similar to "violation handling" can request the aggregated resource for access to certain OATTs. As a result, the DP will only be allowed to access information relevant to their 'access purpose', i.e., "violation handling" as a subset of the shared 'collection purpose' of some of the parent resources, allowing limited disclosure only about that specific information. Thus, DPs are

allowed access to new or aggregated resources with existing permissions limited to their access purposes.

5.3 Prototype Implementation

The main focus of this chapter is to show how different types of metadata extracted from VSS data, can be used in access control mechanisms or policies to ensure a need-to-know view. Thus, for implementing its prototype, we have focused on how different metadata properties can be substituted in policies. We have developed a prototype to implement extended AERBAC that based on video surveillance data provides a decision on whether an observer is allowed to access the recording in the requested access mode (privacy-aware, descriptive, identification). There are three main components of the prototype: a back-end database to store surveillance data, a policy-decision-point (PDP) engine that evaluates AERBAC policies, and a front-end application to show privacy-aware or need-to-know VSS data to observers. The backend database is implemented in MySQL, where different types of metadata extracted from surveillance data are stored under different defined labels/attributes in a database (MySQL for prototype), and the entity-relationship diagram is as shown in Fig. 5.7. In total, there are 8 main tables, where tables (Observer, Role, Area) store attributes that describe observer or UATT, while tables (VidCam-*(Info, Recording, Object, Event, and Location)) store various resource attributes or OATT. Many of these UATT and OATT are used in AERBAC policies as shown in Table 5.1, and 5.2 above.

5.3.1 Policy Specification Language –XACML

The eXtensible Access Control Mark-up Language (XACML) is a standardized policy specification language and is widely used for access control solutions [87]. There are two key benefits of XACML. First, it is generic, as it can be used for any environment for a resource with any number of users. One policy can be used for multiple applications and services, which also makes it very easy to manage. Second, it's distributive, different people or groups can manage different policies at the same time in an efficient way, and XACML will appropriately combine the results from these different policies into one decision. Therefore, for SC-VSS ACM, we are using XACML to define our policies.

XACML supports ABAC policies as it allows policies to be specified in terms of attributes associated with observers and resources rather than their roles or

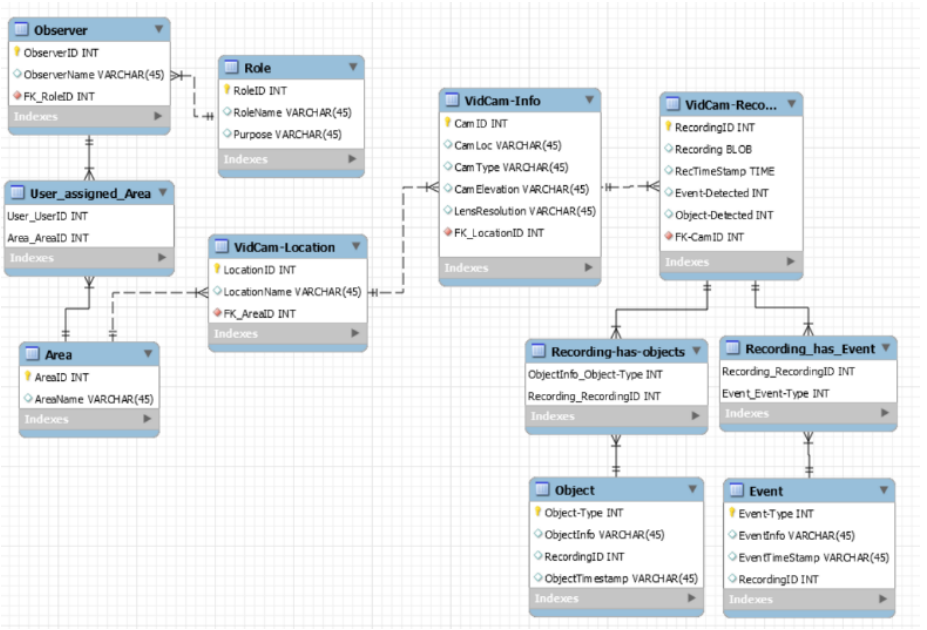


Figure 5.7: ERD for SC-VSS AERBAC prototype

identities. It also supports policies that are specified by multiple DC for the same resource. XACML outlines a method that needs to be followed along with policy enforcement, which is as follows: The user/observer sends a request for a resource to the Policy Evaluation Point (PEP) module, which enforces the access decisions. PEP then forwards the request to the Context Handler that converts it into another format known as XACML request Context. This step is overviewed by the Policy Information Point (PIP) which provides the attribute values of subjects, objects, and environment after the interaction. The context handler then sends this request to the Policy decision point (PDP) which with the help of the Policy Administration Point (PAP) extracts attributes from the request and desired resource. PDP then evaluates the policies and returns XACML response context to Context handler that then converts it back to the required format of PEP, which then grants access to the requestor, if permitted, otherwise denies the request.

XACML has three main concepts: policySet, policy, and rule. A rule is a basic idea with three steps: target, effect, and condition, a target define the environment where a rule is applicable, an effect describes the outcome of that rule (permit or deny), and lastly, the condition is a set of attributes (UATT, OATT, EATT) in a logical combination which results in a Boolean response. A policy

includes a target and a set of rules, while policySet includes a set of policies. For AERBAC, we have defined two main policySet, first describes permissions (consisting of object expressions and access privileges assigned to them, as shown in Table 5.1, and 5.2), while the other describes Roles for observers, as already discussed in the preceding section. The core component of our prototype is the PDP, which evaluates the observer's requests against a set of access control XACML policies (or policySet) and returns an allowed or denied decision. There are many open-source solutions available for implementing PDP such as Sun's XACML (Systems, 2006), XACML Light (Light, 2011), XACML Enterprise, and Balana XACML (Balana, 2013), etc. [87] [88] [89] [90]. Here, we are using Balana for implementing PDP due to its extendible and modular architecture, as it is suitable for large-scale applications. Moreover, it also provides a framework to extend Balana XACML with context manager in order to implement different types of hierarchies (role, resource, location, privileges, etc.), and other OATT, and EATT, which is a requirement in AERBAC. Thus, we have defined a set of logical XACML functions that are to be used in our policies later, as shown in Table 5.7.

Functions	**Definitions**
location-contained-by	It takes two locations as input and return a Boolean output (TRUE), if second input (location) comes under (or is contained by) the first input (location)
Mode-equal-or-superior-to	It takes two privilege modes (privacy-aware, descriptive, biometric/ identification) as input and returns a Boolean output TRUE, if the second input (privilege mode) is equal or at the upper level in the hierarchy than the first input (privilege mode)
Time-in-range	This function takes three inputs (timestamps) values and returns a Boolean output TRUE if the first input (timestamp) falls in between the second and third input (timestamp)
EOI-allowed-for-Role	This function takes two inputs (Role, EOI) and returns a Boolean output TRUE if the first input (Role) is authorized to view content against the second input (EOI)
OOI-allowed-for-EOI	This function takes four inputs (Role, EOI, OOI, PrivilegeMode) and returns a Boolean output TRUE if the first input (Role) is authorized to view third input (OOI) with access privileges in fourth input (PrivilegeMode) against the second input (EOI)

Table 5.7: Collection Purpose Description of Public Safety

Thus, by using primitive and newly derived XACML functions, access policies are written for different roles that will then be evaluated by Balana-based PDP against observers' resource (OBS) requests to make a decision. In order for the PDP to evaluate an OBS request, it first retrieves UATT, EATT, and OATT

with the help of different attributes finder modules (AFM) designed independently for retrieving observers, environment, and OBS attributes from the backend database. After that, these attributes are compared or substituted for the input parameters required in different functions used in XACML policies as described in Table 5.7. Below is one of such XACML policies written for observer A (traffic department responder) The policy describes the identification-mode privilege access for observer A by using some of the above-defined functions. It states that an observer is only allowed to view OOIs in identification mode if certain conditions are met. First, the given observer is requesting the recording for its authorized area and time range, second, the role of the observer is allowed to access the recording that has its authorized EOI, and lastly, the observer is allowed to access the OOIs in identification mode for the allowed EOI. The policy written for observer (B) (traffic department observer) is given below.

```
<?xml version="1.0" encoding="UTF-8"?>
<PolicySet PolicyCombiningAlgId="
    urn:oasis:names:tc:xacml:1.0:policy-combining-
    algorithm:permit-overrides" Version="1.0"
    PolicySetId="RPS:traff_dept_observ:role"
    xsi:schemaLocation="urn:oasis:names:tc:xacml:3.0
    :core:schema:wd-17␣http://docs.oasis-open.org/xacml
    /3.0/xacml-core-v3-schema-wd-17.xsd" xmlns:xsi="
    http://www.w3.org/2001/XMLSchema-instance" xmlns="
    urn:oasis:names:tc:xacml:3.0:core:schema:wd-17">

<Target><AnyOf><AllOf>
<Match MatchId="urn:oasis:names:tc:xacml:1.0
    :function:string-equal">
<AttributeValue DataType="http://www.w3.org/2001/
    XMLSchema#string">traff_dept_observ</AttributeValue
    >
<AttributeDesignator DataType="http://www.w3.org/2001/
    XMLSchema#string" AttributeId="SC-
    VSS:observer:attribute:role" Category="
    urn:oasis:names:tc:xacml:1.0:subject-
    category:access-subject" MustBePresent="false"/>
</Match>
</AllOf></AnyOf></Target>
<!-- permissions associated with the traffic-
    department observer role -->
<!--<PolicySet PolicySetId="PPS:traff_dept_observ:role
    " PolicyCombiningAlgId="urn:oasis:names:tc:xacml:1
    .0:policy-combining-algorithm:permit-overrides"> <
```

```
    Target/> -->
<Policy
RuleCombiningAlgId="urn:oasis:names:tc:xacml:1.0:rule-
    combining-algorithm:permit-overrides" PolicyId="
    Permissions:specifically:for:the:traff_dept_observ:role
    ">
<Target/>

<!-- Permission to view cameras/stream in user's␣
    response␣area␣with␣IDENTIFICATION-MODE␣privileges␣>
<Rule
Effect="Permit"
RuleId="Permission:to:view:OOI-
    inRecording:in:responseareas:with:Identification:privileges:
    ␣traff_dept_observ">

<Target␣>
<AnyOf><AllOf>
<Match␣MatchId="SC-VSS:functions:target:equal-or-
    superior-in-order">
<AttributeValue␣DataType="http://www.w3.org/2001/
    XMLSchema#string">␣Identification_Mode</
    AttributeValue>
<AttributeDesignator␣DataType="http://www.w3.org/2001/
    XMLSchema#string"␣AttributeId="
    urn:oasis:names:tc:xacml:1.0:action:action-id"␣
    Category="urn:oasis:names:tc:xacml:3.0:attribute-
    category:action"␣MustBePresent="false"/>
</Match>
</AllOf></AnyOf>
</Target>

<!--␣current␣time␣must␣fall␣in␣the␣observer's response
     duration AND observer's␣response␣area␣must␣contain
    ␣camera's area -->
<Condition>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:and">
<Apply FunctionId="SC-VSS:functions:condition:time-in-
    range">
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:time-one-and-only">
<AttributeDesignator DataType="http://www.w3.org/2001/
    XMLSchema#string" AttributeId="
```

```
    urn:oasis:names:tc:xacml:1.0:environment:current-
    time" Category="urn:oasis:names:tc:xacml:3.0
    :attribute-category:environment" MustBePresent="
    false"/>
</Apply>
<AttributeValue DataType="http://www.w3.org/2001/
    XMLSchema#time">08:00:00 </AttributeValue> <
    AttributeValue DataType="http://www.w3.org/2001/
    XMLSchema#time">16:00:00</AttributeValue> </Apply>
<Apply FunctionId="SC-VSS:functions:condition:contains
    ">
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only">
<AttributeDesignator DataType="http://www.w3.org/2001/
    XMLSchema#string" AttributeId="SC-
    VSS:user:attribute:response:area" Category="
    urn:oasis:names:tc:xacml:1.0:subject-
    category:access-subject" MustBePresent="false"/>
</Apply>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"> <
    AttributeDesignator DataType="http://www.w3.org
    /2001/XMLSchema#string" AttributeId="SC-
    VSS:object:attribute:area" Category="
    urn:oasis:names:tc:xacml:3.0:attribute-
    category:resource" MustBePresent="false"/>
</Apply>
</Apply>
<Apply FunctionId="SC-VSS:functions:condition:EOI-
    allowed-for-Role">
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only">
<AttributeDesignator DataType="http://www.w3.org/2001/
    XMLSchema#string" AttributeId="SC-
    VSS:user:attribute:role" Category="
    urn:oasis:names:tc:xacml:1.0:subject-
    category:access-subject" MustBePresent="false"/>
</Apply>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"> <
    AttributeDesignator DataType="http://www.w3.org
    /2001/XMLSchema#string" AttributeId="SC-
    VSS:object:attribute:EOI" Category="
    urn:oasis:names:tc:xacml:3.0:attribute-
```

```
    category:resource" MustBePresent="false"/>
</Apply>
</Apply>
<Apply FunctionId="SC-VSS:functions:condition:OOI-
    allowed-for-EOI">
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"> <
    AttributeDesignator DataType="http://www.w3.org
    /2001/XMLSchema#string" AttributeId="SC-
    VSS:user:attribute:role" Category="
    urn:oasis:names:tc:xacml:1.0:subject-
    category:access-subject" MustBePresent="false"/> </
    Apply>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"><AttributeDesignator
     DataType="http://www.w3.org/2001/XMLSchema#string"
     AttributeId="SC-VSS:object:attribute:EOI" Category
    ="urn:oasis:names:tc:xacml:3.0:attribute-
    category:resource" MustBePresent="false"/>
</Apply>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"><AttributeDesignator
     DataType="http://www.w3.org/2001/XMLSchema#string"
     AttributeId="SC-VSS:object:attribute:OOI" Category
    ="urn:oasis:names:tc:xacml:3.0:attribute-
    category:resource" MustBePresent="false"/>
</Apply>
<Apply FunctionId="urn:oasis:names:tc:xacml:1.0
    :function:string-one-and-only"> <
    AttributeDesignator DataType="http://www.w3.org
    /2001/XMLSchema#string" AttributeId="SC-
    VSS:env:attribute:mode" Category="
    urn:oasis:names:tc:xacml:3.0:attribute-
    category:environment" MustBePresent="false"/>
</Apply>
</Apply>
</Apply>
</Condition>
</Rule>
</Policy>
<!--</PolicySet> -->
</PolicySet>
```

After the PDP evaluates the retrieved attributes with the above-mentioned policy, it gives a result 'permit' or 'deny', which is basically a decision about whether the requesting observer can view the requested recording, which has authorized EOIs (as a verification result between observer's 'access purpose' and requested resource's 'collection purpose') and OOIs in identification or biometric mode. Based on that decision, different privacy-preserving solutions can be applied to the recordings for providing observers with different information or view levels i.e., need-to-know view for observers primarily based on the verification of observer's 'access purpose' and resource's 'collection purpose'.

5.4 Analysis and Discussion

The privacy-enhanced need-to-know framework outlined above is implemented as an extension to AERBAC, which is successfully applied for large-scale video surveillance systems for a defined set of observers [28]. AERBAC allows specifying the access control policies in XACML based on the semantic and non-semantic data properties of the video recording (resource object). Metadata-based authorizations enable VSS to have permissions for resource-objects that are not even recorded yet, and that too based on the dynamic properties (contextual attributes) they may have at the time of recording, storage, or presenting.

For the above-mentioned example, using the standard RBAC approach, we need three roles to represent the "Police Dept. Responder", "Traffic Dept. Responder", and "Fire Dept. Responder". For each given role, the permissions have to be specified using the metadata properties of the resource (video-recordings) and its authorized 'access purpose'. In its previous implementation for video surveillance systems, AERBAC only registers the non-semantic properties as a part of object expression or attributes in policy files such as time and location of the observer and the video recording but does not consider the semantic properties of the content. For instance, if an observer wants to access a video recording, then authorized 'area', and authorized 'duty-hours' of the observes will be checked against the 'camera-deployment-location' and 'timestamp' of the recording, as mentioned in Table 5.1. In the extended current model, the same resource will also be checked against the occurred event (semantic content) with the authorized EOIs defined in its 'access purpose'. Furthermore, in the previous model, once the conditions with non-semantic properties were met, the whole requested video recording was made available to the observer with no privacy layer. In the extended model, once conditions are met (including both non-semantic and semantic properties), a portion of the recording with only OOIs will be revealed to the observers (after comparison with resource's 'collection purpose'), while hiding/blurring other existing objects in the video, providing an extra privacy

layer. In case the semantic properties are not met (no authorized EOIs or OOIs), while the observers are authorized to access the resource, then it can be made available in the "privacy-aware" mode. Privacy-aware mode refers to that all the objects that may reveal any direct or associated personal information will be transformed (blurred), such as humans, vehicles, name posts, etc. Here, there are different levels of revealing information (as to which level of information must be revealed, descriptive mode, biometric mode, etc.), based on EATT, such as type, location, or severity of authorized events or objects according to the given 'access purpose' of the observers, hence enabling 'need-to-know' view. Metadata level restrictions based on the verification of 'access purpose' and 'collection purpose' allow the observer to view authorized need-to-know personal information in the recording while obscuring biometrics features and descriptive weak identifiers (such as height, clothing, color) of irrelevant objects [91]. These different levels of information ensure that observers do not have access to enough identifying characteristics of any object to uniquely identify it unless it is essential to its 'access purpose' and helps limit secondary use. Moreover, it ensures purpose limitation by using 'collection purposes' (stored as provenance metadata) with preserved integrity as part of its access decision mechanism. After a DP is authorized to access a resource, extended AERBAC verifies the 'collection purpose' of the requested resource against the 'access purpose' of the DP to ensure that the resource is being used as intended. In the case of aggregated resources, where an explicit 'collection purpose' is not defined, an implicit 'collection purpose' can be derived from the common set of 'collection purposes' of all the parent resources to allow authorized access to the DP.

To summarize, the dire issue with large-scale video surveillance is increased secondary usage, as users/observers had access to a lot of personal information that was not "relevant" to their job or role. By using, extended AERBAC with its dynamic role assignment and ABAC-based metadata-level resources permissions, we have tried to limit them to information only pertinent to the verification result of 'collection and access purpose' by reducing the amount of spare personal information available to the observer, thus reducing the probability of secondary usage by providing a 'need-to-know' view.

5.5 Conclusion

Large-scale has become a universal tool to accomplish several administrative tasks from ensuring public safety to real-time traffic management and many more. The persistent and continuous video recording of data (collecting different types of personal information) from a large number of locations at a city or national level raises serious privacy concerns about data usage. Though different

data protection legislations provide guidelines for privacy-aware video surveillance, it is often hard to accomplish that, as first, to accommodate observers with diverse requirements at a large scale like smart cities, applying uniform privacy-preserving measures for all types of data will render its usability. Secondly, data collected under legal base 'public interest' may have auxiliary personal information which has a possibility to be reused as it does not require an individual's consent. Thus, in order to implement a privacy-aware and need-to-know VSS, it is essential to understand the requirements of the observers, as well as the semantic content of the data to see how they correlate. Furthermore, it is to be ensured that different observers have a relevant yet limited view of VSS data as authorized per their requirements.To achieve this, we propose an extended Attribute Enhanced Role-Based Access Control model (AERBAC) to enforce the need-to-know view principle in large-scale video surveillance systems, for observers by allowing them authorized personal information based on their requirements and limiting avoidable exposure to irrelevant personal information.

The extended AERBAC model allows specification of access control policies in XACML based on the semantic and non-semantic metadata properties retrieved from the video surveillance data and provides a privacy-aware view for observers according to their authorized 'access purposes' defined in terms of events and objects of interest, EOI and OOI. Metadata-based authorizations enable observers to have permissions against resources (video recordings) based on the 'purpose' of the observer. The observers' 'access purpose' for requesting the data should comply with the agreed-upon 'collection purposes' for which data is collected, so it needs to be described in terms, like that of resource metadata properties to allow compliance with supporting legal bases and data protection regulations. Extend AERBAC utilizing RBAC for defining responsibilities and 'access purposes' for users/observers with diverse requirements, and ABAC for using OBS properties including 'collection purposes' in access decisions can help verify them against each other at a metadata level, ensuring purpose limitation. Extended AERBAC implements a need-to-know view by introducing three privacy layers or modes (privacy-aware, biometric, identification) for observers based on the authorized EOI and OOI. It also helps preserve 'collection purpose' integrity by storing it as an immutable provenance metadata property that validates resource or OBS usage transparency and shows compliance with' data protection legislation guidelines. Moreover, implicit 'collection purpose' can be derived in case of aggregation, eliminating the explicit and continuous creation of new access purposes for roles and providing flexibility for a DP/observer with evolving requirements to access transformed or aggregated resources and without constant authorization. This makes AEERBAC an apt choice for large-scale and distributed infrastructure dynamics, as DPs/roles are not bound to resources but to access purposes, giving the flexibility of adding, removing, or transforming DPs at any point in the system without changing permissions against any role. To conclude, in order to achieve a context and privacy-aware ACM for

SC-VSS, we added semantic metadata resource properties in existing AERBAC to enhance it with a content control layer, i.e., enforcing a need-to-know view for the observers based on the verification of 'collection and purposes'. This approach is flexible for applying a privacy-aware access control mechanism, in large-scale dynamic systems, with a diverse number of observers using the same data for different purposes.

Chapter 6

Related Work

Information is the most important entity in the digital world and protecting it against all unwanted means is an absolute necessity. When it comes to the protection of information, a CIA triad needs to be followed: confidentiality, integrity, and availability. Confidentiality implies that information is not disclosed to an unwanted user/observer, integrity means to ensure that information is not illicitly modified and availability confirms that information is available to the allowed users/observers at any given time. In order to ensure the CIA of information, it is important to differentiate between legitimate/allowed and unwanted users, and the concepts of authentication and authorization answer that. Authentication ensures that a user is what it is claiming to be while authorization certifies the rights an authenticated user has over any system. Any user that is authenticated is legitimate. It is very critical for any information system to implement all the above concepts to make sure that the desired information is available to legitimate users while an unwanted user is prohibited, and access control mechanism (ACM) or broadly known as identity and access management solutions are used to accomplish this goal. The ACM manages every user request directed to the system to get access to any resource and determines whether this access request should be granted or not. The effectiveness of an ACM relies on two aspects: first, is the proper identification of an entity and second is the due diligence of the authorized users and their activities regarding a requested resource. It is not only concerned with denying access requests to the unwanted users but also about the activities of a legitimate user, (whether

it is exceeding its access rights, etc.). Before going into detail, it is important to clear the difference among some common terms that are overly used in ACMs like policies, permissions, etc. Policies are high-level guidelines about how access is controlled for different users and how decisions are determined to permit or deny the access request for a particular resource. A policy is formalized through a security model and further enforced by a "reference monitor", which evaluates the users against a requested resource with the help of an authorization database. Resources that need to be accessed are commonly referred to as "objects" while users that request to access any resource is known as "subjects" and a subject can be either a user or a system. Activities that are allowed on an object to be executed by a subject are generally called "permissions".

There are several traditional AC solutions (discussed briefly in the next section) that are been widely used everywhere from small-scale to large-scale infrastructures for a large variety of purposes. This thesis focuses on ACM for a large-scale VSS that preserves privacy by offering a need-to-know view for different users (observers) based on their authorized purposes, therefore, in this chapter, we will focus on ACM solutions that address similar ideas. Various concerns need to be addressed while designing a privacy-preserving ACM for large-scale VSS such as technological and technical capabilities (distributed information sharing and processing), organizational needs (a large number of users with diverse requirements), diverse data structural needs (structured and unstructured data), trust-based organizational relationships, etc. Every system or infrastructure has different requirements; different sets of users, resources, and sets of rights over resources, etc. so one broad or standard ACM is not the desired solution. Traditional ACMs in their intrinsic form often cannot meet the above-mentioned challenges, as they require more granular, dynamic, and flexible access control solutions. Different solutions have been proposed over the years for efficient ACMs for evolving systems with constantly changing requirements. This chapter will present a summary of traditional ACMs with more focus on solutions presented for large-scale and distributed infrastructures utilizing different parameters such as video data as a resource, access based on semantic content, ACM utilizing user and resource hierarchies in large-scale infrastructures, etc. to ensure fine-grained access to resources [92]. It is important to note here, that this chapter only focuses on a literature review of the solution proposed in Chapter 5, as Chapter 2 and Chapter 4 have their own relevant related work section. Please note that this chapter may contain excerpts and figures from our own published papers (as part of the PhD research) mentioned and referenced in 1.3.

6.1 Traditional Access Control Mechanisms

6.1.1 Role-based Access Model (RBAC)

RBAC model regulates the access control mechanism based on the different activities a user can perform in the system [25]. For this to be effective, first roles are defined. A role describes as a set of actions and tasks linked with a particular activity in the system. Thus instead of describing what each user can and cannot do, access can be associated with roles. Access is no longer linked to a particular user, rather users are authorized for particular roles and then access of user is determined by the role and permissions assigned to him. This simplifies and somehow automates access management tasks, for example, if a role of the user is changed, the manager only needs to change his role, and all access privileges will be granted or revoked as associated with a role. It works on the principle of the least privileged. A user that has been assigned a role X can have minimum access as defined in Role X. A user can also have multiple roles assigned to him and perhaps in cases are simultaneously exercised. Hierarchy is also presented in RBAC based on the principles of generalization and specializations. Role hierarchies are a bit complex as they affect role activation and access privileges, i.e., a specialized role inherits the authorizations of its generalizations. To enforce more fine-grained control, separation of duties is appended with RBAC. Separation of duties can be both static and dynamic. Static separation of duties refers to a permanent restriction related to a role irrespective of its current situation while dynamic makes a decision at runtime about which permission to give and which to revoke. RBAC decisions are centered on the subject's association to roles so it gets a little difficult when RBAC has to evaluate requests based on multiple factors or parameters.

The authors in another paper had discussed in detail the role hierarchies in Access Control [93]. The role as discussed above is a set of rights assigned to a subject. There are three types of role hierarchies. Is-a role hierarchy (generalization principle), role activity hierarchy (aggregation principle), and supervision role hierarchy, which in their traditional sense are a tail static inheritance to develop a hierarchy. The current state of any system is called an authority state that can describe the result of any access request. The author suggested that hierarchies in access control should not follow static inheritance roles rather should be more dynamic and depend upon the changing constraints of the authoritarian state. To achieve this, permission transitions from different authority states should be formally modeled and analyzed and should be tested in real organizational scenarios with large datasets.

6.1.2 Attribute-based access control (ABAC)

Traditional mechanisms fell short when used in distributed computing paradigm. Distributed computing has independent servers offering services to a massive number of users all around. The identity of a user may not be known at the time of the request, so RBAC is not very suitable for cloud environments. Context information about a user can tell more about him instead of his identity for making an access decision. ABAC grants access to objects/services based on the attributes possessed by the subject/requestor [84]. ABAC evaluates requests based on attributes, i.e., characteristics of subjects and objects, environment conditions (operational or situational awareness (context), and their indicated policies. The key benefit of ABAC is that it can evaluate requests without the prior knowledge of the object by subject, i.e., eliminates the need for explicit authorization. This makes it very useful for large organizations as administration of roles and lists will be very cumbersome so defining attributes that cover subjects, objects, authorization and authentication activities, etc., and make dynamic decisions based on them while preserving an appropriate level of security. The only issue in large-scale systems is to generate discriminant attributes systems for all subjects, objects, and environments. Every object in the system must at least define one policy that states, which operations are allowed to subjects and under what conditions can they be performed.

6.1.2.1 Policy-Based Access Control (PBAC)

PBAC is a modified version of ABAC. Like ABAC, PBAC creates access control policies using attributes from subjects, objects, and environment but ABAC is not a wise choice in a large enterprise where there are a different set of users with diverse requirements and need different levels of access to some resources along with consistent access to the shared resources [85]. An ACM may only need a username and password to validate a user that needs to access resource A, whereas it may require more of the user's credentials to determine whether the user is authorized to access Resource B or not. Besides, the roles or permissions might have some descriptive conditions for access that cannot be applied due to ABAC. PBAC helps with this shortcoming and helps to derive policies according to systems' demand and enforcement of abstracted access control principles by listing them concretely with rules in ABAC. PBAC is much more complicated to implement than ABAC, due to complex policy and evaluation mechanisms and requires a lot of processing capability.

6.1.2.2 Risk-Adaptable Access Control (RAdAC)

Distributed infrastructures are developing strategies in this fast-paced information-sharing environment with a high economic aim and rough financial veracities [86]. The dynamic nature of users and resources demands a high adaptive ability of access control policies that can persistently assess the risk of information proliferation at different components of the system. Therefore, the risk-adaptable access control (RAdAC) model emerges as a real-time and adaptable access control mechanism, since the existing models are insufficient to cater to the vitality in the risk assessment. RAdAC estimates the security risks to decide whether the existing policies of access control need to be overruled or not. RAdAC also takes into account the contextual or environmental attributes while making a decision. For example, under normal conditions, users can access a resource with just a username and password but in case of a security breach, RAdAC will enforce a much stricter access control policy to protect the requested resources by asking for more user credentials.

6.2 Access Control Models for Video as a Resource

Video is a complex data type and contains a whole lot of information, and users with multiple requirements can require different types of information from the same video resource, as discussed in detail in chapters 4 and 5. Over the years, different ACMs and their variants have been introduced to manage efficient access to video data or video database management systems. Traditionally, a video recording of a certain time length is considered a resource, and is stored or indexed via a non-semantic property such as time-of-recording, or camera location, now with sophisticated video data analysis tools information present in the content of the recording can act as a resource too like objects or events detected in the content. Our presented solution (cf Section 5.2) is based on the key idea of how information in video content can be utilized to regulate access control in large-scale dynamic systems, such as SC-VSS. Two main ideas are addressed by our proposed solution, first, how much information can be extracted from the video content, and whether is it possible to obtain data about activities happening in video content, already discussed in Chapter 4. Second, how different types of access control mechanisms regulate access based on that type of information or content (semantic and non-semantic metadata) in large-scale infrastructures for users with diverse requirements. In this section, we will focus on the contributions in areas of content-based access control, and access control for large video databases and what gap can be filled by our proposal

as ACMs used for large-scale infrastructures processing video data are multi-dimensional and are addressed with different perspectives. We have summarized the related work by two different parameters hierarchical access for video data, and content-based access.

6.2.1 Hierarchical Access Control Mechanisms in Distributed Infrastructures managing Video Data

A fundamental hierarchical access control scheme states that there are separate sets of security clearances for resources and a user is committed to a certain clearance class. A certain security clearance assigned to a user means that it could access all the resources with security clearance lower or equal to it but not the other way round. As for hierarchical access control for video data, Bertino et. al pioneered the idea [94]. An access control method is proposed for video databases based on the structural and semantic composition of the video. A video element is considered as an "authorization unit" which can be a sequence of several video frames or any object that is a part of the frame. Pre-determined identifiers based on changing low-level information in the content (boundaries, scene change, etc.) specify how videos are divided into smaller sub-elements. Two modes of operations can be performed over video data, i.e., viewing or editing. Upon the arrival of a request for a certain portion of the video, the proposed algorithm evaluates which sequence of frames can a user have access to, based on its identity and description of the video object/frame. Each video element has id, start, and end frames, and may have other sequences of frames, etc. a new video element can be constructed from the existing video elements. Authorization rules are rather described traditionally, i.e., a 3-tuple <subject, object, mode>. For the subject, a credential expression is formulated which contains the subject's identity and conditions that he needs to fulfill in order to access the requested video element. Then, for specifying which objects need what kind of permissions, a content expression is made based on Spatio-temporal properties of the object (like start frame, end frame), etc. and at last mode expression is created. To summarize, the proposed ACM is based on the idea to convert a longer length video into smaller length shots and then treat each of them as a separate sub-object for the video recording, and the user can have access to either edit it or just view it, based on the user's identity. However, the solution does not consider high-level visual or semantic concepts like (objects or events) as part of the permission or request expressions [95]. The same authors enhanced the same model by allowing a set of filtering rules that would consider high-level semantic concepts. This solution is more suitable for small-scale video systems with a limited set of users. As it did not take into account the changing context of the user or the environment so, it is not appropriate for large-scale dynamic systems such as in smart city distributed

video surveillance systems [96].

User access and resource usage are relative in any system. Enforcement of access policies relies on the resources of the system to list which user is permitted to use which resource and in many cases under what time and condition. The following paper explores a user hierarchy and a resource hierarchy in terms of access control [95]. Hierarchies arise from the fact that there are many users, of which some possess more rights than others and similarly, some resources have more access constraints than others do, based on the information they have to offer or the type of operations that can be performed on them. The paper presents an algorithm to unify both these hierarchies (user and resource hierarchies as sub-hierarchies). Unified hierarchy permits compact specifications of access control and Cryptographic key-based hierarchical schemes help enforce them easily in distributed environments. This proposal works best for static permissions and not very suitable for dynamic access control and specification of negative access relations.

Video can be realized as a hierarchical combination of different access units such as shots, objects, and regions of interest. Moreover, all these units are associated with different types of attributes such as color, shape, texture, and layouts. All this high-dimensional video data needs to be properly indexed for efficient retrieval. Access based on semantic visual concepts rather than low-level visual features is required but for that to happen, effective video representation and concept categorization are highly necessary [94]. The conceptual hierarchy of video elements can help resolve the "curse of dimensionality" for visual data indexing. Based on above mention observations a multilevel video database access mechanism is proposed, i.e., a user-adaptive video access control mechanism based on hierarchical indexing of visual semantic concepts. A domain-dependent concept hierarchy organizes contextual and logical relationships among these semantic concepts. Decision tree-based classifiers seem to be very fitting for video data classification by learning from trained examples nevertheless it can be an issue to video indexing because of the numerous internal nodes corresponding to the same classifiers. Semantic video classifiers can be useful as an efficient indexing structure but also as a good way to bridge the semantic gap. The access control mechanism is integrated with a video database system based on a set of filtering rules that considers the unique protection requirement of video data. Different granularity levels of access control are implemented based on what filtering rule is applicable for complete semantic cluster or sub-clusters, or different content present in a video segment or frame, etc. In addition, two different modes are defined: querying and browsing. The model also supports content-independent access control. A filtering rule can have a video element or content expression (based on visual features). Access is granted based on the privileges associated with these filtering rules. A unique contribution of this paper is the introduction of visual features as a part of content expression, as

well as the multilevel access control mechanism is based on it and the clustering structure of video data. However, it has not presented concrete proof of work to show how efficient this algorithm might be.

A hybrid video data model is presented based on the hierarchical model and content usage in the following paper [97]. Video is segmented into basic units of objects and frames to provide multi-level access until the granular level. In this paper, content refers to the different types of annotations or keywords added manually by the administrator. It mostly focuses on the hierarchy of video data, which is based on the segmentation of scenes, rather than the content (objects and events) of the data (Tran et. al, 2007). A cluster-based tracking algorithm is developed to acquire motion trajectories. Every activity inherits all the semantic concepts that belong to its specific activity model, which acts as a node in a hierarchy. The data can be browsed via keyword, object, or queries by sketch (similar trajectories drawn by users to spatial trajectories present in the data) [95].

Thuraisingham et.al proposed an authorization model for large-scale video surveillance systems to prevent unauthorized access. This model sightsees the extensive use of hierarchical taxonomies that can help develop the subject, object, and privilege hierarchies, whose outcome is an explicit policy base devised by the user himself. Semantic concepts focus more on human describable ideas as the focus of this paper is event extraction, event comparison, and then event detection from low-level features (colour, edges, and densities), and then users can label these events [98]. Another author, Vanessa, proposed an algorithm based on indexing video summaries instead of segmenting them into frames and shots. Spatio-temporal and contextual information is extracted from video data and then video units are indexed according to that and an efficient hierarchy is maintained. If a new constraint is to be introduced, then all indexes are updated. To preserve privacy, frames with sensitive information are stored separately in blurred form. A multi-level access control mechanism is implemented by extracting spatial-temporal and contextual information from the user's profile and is then matched within the hierarchy to grant access [99].

One of the traditional access control mechanism Role-based Access Control (RBAC) model is extended to fulfill the security needs of distributed multimedia applications, MRBAC has adopted object-oriented concepts and developed a hybrid role hierarchy for roles and rules by evaluating Spatio-temporal characteristics of multimedia data [100]. The data is segmented by applying techniques such as object identification, video shot segmentation, etc. and then subunits are indexed based on generated results. Users can also specify their region or objects of interest (visual or audio) object in the multimedia repository in their preferences. The proposed model supports multi-level access control by checking Spatio-temporal and IP address constraints and decentralize the administration

role to make it more efficient. This paper has not explored the structure of multimedia data in detail, therefore has not presented an efficient hierarchy in how to store and process data, however, it has considered the Spatio-temporal properties of multimedia data.

Scalable Video Coding (SVC) is a method used for bandwidth adaptability with different network and device requirements and needs a layered access control scheme to preserve it. SVC integrates three modes: temporal (achieved via hierarchical structure like B-trees), spatial (using layered coding), and quality scalability [101]. Thus, it has one base and several enhancement layers. Therefore, different scales of coding generate different bitstreams, and users with different priorities have different privilege needs. The authors have analyzed SVC bitstream, transformed it into a longitudinal hierarchical structure, and have proposed a secure and efficient key management scheme, which bids high security, low computational, and storage complexity. The pattern can be applied or extended to other scalable multimedia formats and can be suitable in a large number of applications. The advancement in network technologies and high data rates has increased the interest in video applications in different fields such as video conferencing, pay-per-view (sports, movies, news, etc.), and shared gaming, etc. Most of these services are paid and need a mechanism to bill users according to their usage; therefore, need an ACM to differentiate authorized and unauthorized users while granting access to group service. Customarily, many group applications contain numerous related data streams and have several access privileges for different types of users. Such distributed and multi-layer applications require multi-level access privileges so an access control mechanism is required to manage it, which is referred to as the hierarchical access control. To achieve that, the following presents an algorithm for multi-group key management to realizes hierarchical access control in secure group communications [102]. This paper presented a multi-group key management scheme that attains hierarchical access control in secure group communications, where multiple data streams are available to group members with various access privileges. They considered an integrated key graph, which allows users to join or leave group communications and designate different access privileges while preserving forward and backward security. A media stream is scaled from multiple dimensions like resolution, SNR, frame rate, etc. Each media layer is encrypted with a different key and a key is then shared with the authorized user. The manager/owner who owns and distributes the media content is the attribute authority (AA) which regulates access privileges. The author claims that content is only transferred to the trusted nodes in a secure chain and consumes fewer resources than traditional key-based key mechanisms and can be adapted for preserving privacy in large-scale social networks.

A multi-level access model is proposed to streamline the process of defining and assigning permissions via enhanced expression ability in order to realize

fine-grained access control known as RTBAC (role and task-based access control) [103]. This model is designed to cater to the dynamic authorization needs following virtual enterprise role hierarchy and task management. RTBAC model is more useful in managing collaborative operations. Within an organization, different roles are working on the same product but are only allowed to restrict their access to their relevant portion and not the whole product. To handle this scenario, RTBAC manages multi-level privileges to the components where different users are assigned different tasks. It helps solves the problem of "all or nothing" permissions. Different organizations have different access control needs that all cannot be resolved with traditional models like in some scenarios access should be granted to some trained sets of users, with the ability to separate roles and the inability to abuse the designated role. The proposed model is a well-organized encryption scheme, denoted as Shared Encryption Based Construction (SEBC), which assigns to each class a single piece of private information, whereas, the public information depends on the number of classes, as well as on the number of edges in the hierarchy. The security of the projected construction relies on the ones of the underlying encryption and secret sharing schemes. The model will differentiate and recognize scenarios where more than one entity is essential to achieve a particular authorization or special permissions. Moreover, a formal definition of hierarchical and shared key assignment schemes is also presented.

6.2.2 Content-Based OR Content Dependent Access Control (CBAC) for video Data

Content-based access control refers to regulating access to users based on the content information or attributes within the resource. It has been commonly used by systems using relational databases or structured data and was initially introduced in mentioned papers [104] [105] [106] [97]. CBAC has two major areas, first deals with dynamic content, in which tags or annotations are dynamically extracted from structured or tagged data (like in web 2.0 or XML structures), and then access privileges are bind to different users based on it, though it requires regressive supervised learning o system's part in order to look for the relevant annotations [107] [108] [109]. The second category of CBAC deals with static content and is more relevant in multimedia or video data context. , i.e., predefined attributes extracted from the content are defined apriori, and access privileges are mapped between user credentials (roles, attributes) and resources' predefined content attributes via some [110] [111] [112]. Static CBAC has been adopted in context to video data, where it was proposed that predefined annotations extracted from video data (manual textual descriptions about the video) can be used in access control policies [113]. Some of the notable work done in regard to video data and content are discussed in detail below:

Video is the most informative mode of communication; it has a lot of content and its applications in different areas are on the rise. However, video querying includes a lot of user interaction and feedback-based query refinement, which produce large traffic volumes on the network if the whole video is sent in its original form. To make more use of video data, protocols need to design fine compact representations for long video sequences to make video extraction, browsing, and retrieval more useful [114]. The referenced paper has suggested that video segments should be clustered through fuzzy clustering instead of gradual scene change (GCD), and then these clusters should be arranged in hierarchical order. These clusters can then be accessed as "content" but how will they be accessed is not described in the paper.

Due to the increasing amount of multimedia data, it is important to organize it in a useful way for efficient retrieval. In the following paper, a hybrid video data model is presented based on the hierarchical model and content usage [97]. Video is segmented into basic units of objects and frames to provide multi-level access until granular level. Moreover, content-based queries are included for efficient retrieval. In this paper, content refers to the different types of annotations or keywords added manually by the administrator. It mostly focuses on the hierarchy of video data, which is based on the segmentation of scenes, rather than the content (objects and events) of the data.

Video surveillance produces a huge amount of data that needs to be effectively indexed for efficient retrieval. Although many algorithms are proposed about how to present the content of video clips in current systems, a huge semantic gap is being observed between the user and the video retrieval systems. Video surveillance can offer support for investigating semantic-based video retrieval. The authors have presented a semantic-based video retrieval framework [115]. A cluster-based tracking algorithm is developed to acquire motion trajectories. The trajectories are then clustered hierarchically using the Spatio-temporal information, to learn activity models. Every activity inherits all the semantic concepts that belong to its specific activity model which acts as a node in a hierarchy. The data can be browsed via keyword, object, or queries by sketch (similar trajectories drawn by users to spatial trajectories present in the data). The model also supports multiple queries that are conditioned with restrictions and the efficiency of the protocol is tested on traffic scenarios.

Most social-networking applications have implemented role-based or group-based access control policies to preserve user privacy. Nonetheless, many times it is unable to protect user's data from going into the wrong hands. The authors have suggested improvements in existing methods to enhance user's privileges [108]. Users have the facility to stipulate which type of content they want to share and with whom by specifying tags or keywords. Defining all potential keywords to a particular topic is very cumbersome. Linked data is used to enrich the poten-

tial keywords by identifying related concepts. Two different algorithms are used to design the proposed semantic framework Semantic Enhancement and Direct Comparison. Semantic Enhancement helps identify closely and conceptually related meaningful words, i.e., Linked Data terms, and then design policies according to them. Direct Comparison performs the searches directed by users in a policy-aware manner. Traditional access control policies have somewhat upfront solutions to access control as roles are explicitly defined against data objects; which works well in an acceptable limit of data objects and users in a dynamic environment. In large-scale organizations with a huge amount of data objects, it becomes difficult to use traditional access control methods, especially when the semantic content of data is anticipated to affect access decisions. Underprivileged and over-privileged, both types of users are a threat to the system so it is important to have another access control method that can help deal with the above-mentioned issues. The authors have introduced Content-Based Access Control (CBAC), an advanced and upgraded access control model for content-centric information sharing in conjunction with RBAC or Multi-level Security (MLS) [108]. RBAC or MLS allows users to access a large dataset then CBAC is applied as an additional access layer to control the limit of data (a subset) based on content privilege. The periphery of the subset is dynamically selected by the textual content of data objects. The enforcement mechanism of CBAC policy is imposed with the help of Oracle's Virtual Private Database (VPD). The accuracy of semantic content matching with a tagging mechanism is implemented to increase the efficiency of the model. Investigational outcomes show that the decisions made by CBAC are reasonable, and the overhead is tolerable.

6.2.3 Security Preserving Video Data sharing with Access Control Solutions

Access Control solutions provide a good level of privacy when it comes to structured data like textual records, but with unstructured data like videos and images, it needs extra measures to ensure that content affects the decision-making process as in CBAC mentioned in Section 6.2.2. Moreover, often, in large-scale infrastructures, like mass-scale video surveillance, the same data can be subjected to different privacy levels as per the requirements of the users, requiring an additional layer to CBAC. This section discusses some of the notable work in the domain focusing on using semantic or content information in video data to enforce different or multiple levels of privacy preferences or various levels of information (video content). A considerable amount of work has been done in similar areas such as retrieving semantic properties/attributes from different resources and developing shared ontologies and using them in decision-making processes for large-scale and distributed environments [116].

Furthermore, researchers have proposed several solutions based on the outcome of semantic-based decisions to ensure privacy for different users or observers [117]. Some of the notable work is mentioned below. Video Surveillance at a large scale with various stakeholders using automated technologies has the ability to exploit such data, thus requiring protection from unauthorized use to preserve privacy. However, observers with authorized access have often abused their access privileges, therefore, it is important to enforce certain privacy levels that restrict authorized observers with limited data. The below-mentioned paper is one of the pioneers in this aspect and focuses on the idea of a "privacy console" for authorized (human) observers watching the live video streams [91]. It considers several non-semantic properties of surveillance data like location of a video camera, time of the recording, and policies are pre-defined for different locations and based on that, the video data stream is "blurred" or transformed, so observers cannot identify humans present in the video live stream. If they do want to observe un-blurred video data, they need authorization. It presents a conceptual framework based on the idea that observers do not need to view all the information, and content in video should be blurred or for different observers based on the non-semantic properties of video surveillance data. The paper does not present any access control mechanism to realize this framework, or how to specify access control policies for observers. Moreover, it considers a static privacy mechanism for video data, i.e., either video is transformed or not, depending upon the authorization level of the observer, and doesn't offer different levels of information.

Large-scale video surveillance has made it possible to implement law enforcement policies to check for unauthorized event detection as a preventive measure against unusual events. On the other hand, it is also invading the privacy of common people [98]. This paper suggests that all organizations with integrated video databases should preserve the sensitive (or personal) information present in video data. It is unavoidable that video data should not be shared, so the aim is to implement security-preserving data sharing. High-level semantic content can reveal much more private information than the low-level visual features (such as color, object shape, etc.) in video data. An authorization model is proposed based on semantic-based component hierarchies and credential expressions (query-like expressions based on visual features and concepts) and an explicit policy base. This model explores the extensive use of hierarchical taxonomies that can help develop the subject, object, and privilege hierarchies, whose outcome is an explicit policy base devised by the user himself. Semantic concepts focus more on human describable ideas as the focus of this paper is event extraction, event comparison, and then event detection from low-level features and then users can label these events. Timestamp and location information can also be extracted with such events. Another part of the paper discusses the factors that should be considered to maintain the privacy of video data are: content sensitivity, a user requesting the content, and how are they going to use it. A

method is prosed to detect the sensitive information within video content but is no very concrete and clear. Besides, this model is suitable for smaller and similar sets of video data and not so much for large data sets due to manual labeling by the users.

The idea behind the proposed algorithm is to index video based on the summaries instead of segmenting it into frames and shots [118]. Spatio-temporal and contextual information is extracted from video data and then video units are indexed according to that and an efficient hierarchy is maintained. If a new constraint is to be introduced, then all indexes are updated. To preserve privacy, frames with sensitive information are stored separately in blurred form. A multilevel access control mechanism is implemented by extracting Spatio-temporal and contextual information from the user's profile and matched within the hierarchy to grant access. This scheme theoretically solves the space efficiency, correspondingly to the request-time efficiency problem because of summarized indices. Social networking websites produce roughly one-third of video data and developed highly sophisticated tools for video data analysis to retrieve user preferences to generate useful marketing information. Often, privacy on video data is not a priority, but in recent years this trend is changing, and more and more applications and services are considering privacy in video data an important requirement. One such example is an API or an application provided by YouTube or Google, which allows users (at the time of uploading video content) to choose the areas of the video (manually) or frames, which they want blurring or transformed, to hide sensitive or irrelevant information [119]. Though, it has another perspective that, viewers have a restricted or privacy-aware view of the uploaded content, but the owner or service provider in this case still has access to the whole video content. Cloud has boosted the need for multimedia applications due to cost-effective and prevailing resources and is becoming the ultimate choice to store and share multimedia content. However, due to its decentralized and traditional public nature, it also raises security and privacy issues. Authors have proposed a cryptographic solution to securely share video data among a group of people for a particular time period, i.e., a secure time-domain attribute-based access control (TAAC) scheme [120]. The timestamp is embedded in both keys and video ciphertexts, so the access is only allowed to certain users within a certain time and one having enough attributes can decrypt video content. This caters to the dynamic nature of user attributes. Special queries can be performed on video contents of previous time slots are also discussed. The security analysis and performance evaluation show that TAAC is provably secure in the generic group model.

A video privacy framework is proposed in the mentioned paper that uses an off-the-shelf encryption scheme for controlling the access to the Region of Interest (ROI) in video data. it is focused on blurring/encrypting the whole human body (ROI) detected in video data, for a fixed time interval, and authorized ob-

servers can decrypt ROI with a self-decryption [117]. The proposed model only considers humans as ROI, as they represent most of the personal information (face, gait, color), and in order to access/ view ROI, the user/observer requires a decryption key. Moreover, if the observer has the key then it can access the original or all information (descriptive and biometric) about humans in video data. This approach does not accommodate objects with personal information other than humans, (belongings, associations). Plus, it doesn't consider that in real-time different users might need different types of personal information depending upon their purposes, and giving a recording with the same encrypted ROI will either provide observers with extra information than required or prevent some from completing their authorized tasks. As most of the work on video privacy is focused on blurring objects or regions of interest [121] [81]. It is often hard to measure the extent of privacy risks in video data, also, it often neglects the contextual information present in the video, i.e., the human face may be blurred, but if the background of the video hints about the location of said human, then that is a privacy risk. To address that, the authors presented an application VERRO, that converts objects with personal information into indistinguishable objects, also hiding the contextual information around the object [81]. This technique "anonymizes" video data so it can be made available to public-domain users. It takes videos as input and transforms humans and vehicles into generalized figures (as shown in the image) or blurry objects, along with blurring the contextual information. In a similar approach, instead of blurring the objects, the authors have proposed to remove the objects from the image or video frame that contain personal information based on applying deep learning convolutional feature-based technique [122]. This may be useful for archiving or processing video data with extreme privacy measures, though is not a very useful technique in access control solutions that focuses on authorizing observers with required personal data in video content.

The distributed information-sharing paradigms put a lot of focus on having multiple privacy levels that comply with both users/observers and individual requirements [123]. In the context of smart cities, authors have proposed an ABAC-based privacy preference ACM, that provides individuals with the flexibility of deciding the level of privacy or sensitivity for their personal information collected by wearable sensors (like fitness devices) [124]. This and similar solutions are effective when consent is the supporting legal base for collecting and sharing personal data, though, are often ineffective when data is collected for other legal bases such as public interest or legal obligation, which is often the case in large-scale video surveillance data [125] [126] [127].The reason being that every legal base does not support the same set of rights for individuals. Like, in the case of public video surveillance data, an individual does not have complete control over how its information will be used, the level of sensitivity attached to different types of personal data is decided by public authorities per national or regional regulations.

6.3 Conclusion

To conclude this discussion, most of the related work in similar areas to preserve privacy in large-scale infrastructures are primarily based on user hierarchies or high-level resource hierarchies and put less focus on content or semantic hierarchies as part of the context-aware access control mechanism. Moreover, solutions presented for large-scale video databases are more suitable for a defined set of users, whereas our presented model considers the dynamic and diverse nature of observers (users)' requirements as well as the resources' semantic content, in a large dynamic environment (such as a smart city) due to the flexibility of metadata parameters controlling the flow of information at a granular level. On the other hand, solutions that focus on semantic-based ACM, are often limited to providing a uniform view or identical privacy measures to transform or protect semantic concepts based on declaring regions or objects of interest (humans, license-plates), and do not offer variable privacy levels for different observers within the same data content. For instance, if there are 10 humans detected in the data, a law enforcement observer should only be allowed to view the identity of the human that was involved in an event of interest (EOI that is intruding here), meaning that the identity of all the other humans that were not involved in the EOI must be protected. Moreover, many of these solutions do not consider different legal bases supporting the collection of personal data, which they should, as it affects how personal data should be processed according to collection purposes. Our proposed solution binds the purpose of data collection to events of interest (EOI) and objects of interest (OOI) retrieved from video data, thus limiting the observer view to a 'need-to-know' basis. It not only provides a flexible multiple-layer privacy-aware view for observers with diverse data requirements but also limits the secondary use of personal information. Our proposed solution accommodates both variable access to information per observers' requirements and ensures privacy for objects other than humans too, with different levels of information (descriptive, biometric). It also provides some level of control to individuals, as it takes into account the agreed-upon purposes per the supported legal base public interest and reduces secondary use, thus preserving privacy.

Chapter 7

Conclusion and Future Directions

During the past decade, the invasive presence of IoT, social media websites, and smart-city services has emphasized the importance and usefulness of personal information. A large number of data applications and services collect and process personal information about individuals via different sources to provide them with personalized and informed services. One such example is large-scale video surveillance systems (VSS), which provide many benefits by monitoring a specific area or activity but also implies a serious concern i.e. invasion of privacy, one that threatens the right of individuals to have control over access to personal information about them. Local authorities (LA) have deployed many cameras to monitor different public places, so in case of any noticeable incident, they can provide real-time responses and record evidence if required. VSS data is largely video recordings usually indexed by the time or location of the recording. Video data is an unstructured and complex data type, and to extract most of its information, it either requires manual monitoring by a human observer or advanced video analysis tools to understand its content. With recent development in the machine and deep learning solutions, now video data can be processed to extract most of the information as humans perceive it. With such type of information at large-scale, let us say at city scale, with thousands of video cameras feeding VSS, can reveal a lot of personal information about individuals, despite being highly useful for many local authorities administrative observers. Moreover, information extracted from video data can be analyzed with other data sources

like national registration databases or traffic-management data, or other public information systems, which can enrich the information when analyzed together. For instance, a person detected in a video recording is just an individual to an unbiased observer, though when the face of that individual is compared with data from the national registration database, it can link an identity to that individual. Therefore, large-scale VSS has various data sources at its expense, that when analyzed with video data enrich the existing information for the requested observer. However, the same information when collected and used against an individual's consent or knowledge is an invasion of privacy, which has occurred recurrently in past [50] [12] [128].

Misuse or misrepresentation of any piece of personal information by a data-collecting entity, or any use of personal information without individuals' consent or a valid legal base, is a legal violation leading to privacy invasion. To protect personal information, countries around the globe have introduced several data-protection legislation such as the EU General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), Asia-Pacific Economic Cooperation (APEC) Privacy Framework, Personal Information Protection and Electronic Documents Act (PIPEDA), etc. [10] [11] [129]. These legislations provide a legal framework for data owners or data controllers (DC) regarding the use and protection of the personal information of individuals or data subjects (DS), declaring data reuse for any purpose other than the agreed-upon as a violation, i.e., purpose limitation. Furthermore, it also underlines the rights of DSs over their personal information, such as their right to be informed about how their information is being used, and a right to object if it is not used accordingly. Thus, it is a legal obligation of a DC and its authorized observers/DP to ensure purpose limitation by collecting and using personal information only for agreed-upon purposes [10]. However, in past years a large number of privacy invasion incidents that have occurred were traced back to the authorized observers that reuse the VSS data or especially personal information extracted from it. Some of those incidents may have been unintended, yet a lot of those incidents were deliberate that exploited a presented opportunity of data reuse due to misunderstanding, misinterpretation, and ambiguous terms of purposes. Therefore, it is crucial to define and declare purposes that offer less to zero possibilities of misunderstanding or exploitation, as well as ensure transparency and compliance with applicable data protection legislation. Moreover, declared purposes should be easily relatable and verifiable with collected data (properties), as discussed in Section 2.3. In this thesis, we have distinguished purposes into two categories. First, the 'collection purpose', referring to the terms of an agreement between the DC and the DS about resource usage, i.e., why the data (specifically personal information) is being collected, how much of it will be stored and used, etc. Second, the 'access purpose', referring to the terms of an agreement between the DC and the data-users or data processors (DPs) describing how can DPs use data and under what limitations, discussed briefly

in section 1.1.2.3 [10]. In order to ensure purpose limitation, an 'access purpose' of a DP should comply with the 'collection purpose' of the resource. Moreover, a valid legal base is required by data protection legislation whenever the resource with personal information is collected or processed. Therefore, when the observer requests a resource (recording) with personal information, it should also demonstrate its 'access purpose' to show that it or particularly its role is legally allowed to access the requested resource for a comparable 'collection purpose'. We proposed to describe the 'collection and access purpose' in terms similar to that of resource metadata properties to allow fine-grained compliance (cf. Section 2.3). Moreover, data changes into multiple forms in distributed systems, and often the 'collection purpose' is overlooked during different transformations and aggregations. Therefore, we also proposed that the resource preserves the integrity of that 'collection purpose' as part of its indisputable metadata (provenance), it can assist the VSS in establishing that the 'access purpose' complies with the 'collection purpose' and thus ensures purpose limitation. Moreover, in the case of different data aggregations, where the same DC does not manage resources or 'collection purposes' of the involved resources, DPs can use the preserved 'collection purposes' derived from the provenance for relevant or compatible 'access purposes' without violating purpose limitation.

In this thesis, we have proposed a privacy-aware Access Control Framework to ensure purpose limitation by addressing the above-mentioned concerns in large-scale infrastructures. We have implemented an extended Attribute Enhanced Role-Based Access Control model (AERBAC) to enforce the need-to-know view principle in large-scale VSS, for observers by allowing them authorized personal information based on their requirements and limiting avoidable exposure to irrelevant personal information. The proposed solution focuses on preserving the privacy of individuals recorded in VSS data by ensuring purpose limitation and restricting secondary use with the efficient utilization of different types of metadata (semantic, non-semantic, and provenance) extracted from VSS data. The 'collection and access purposes' of resources and observers respectively are defined in terms similar to that of VSS metadata so they can be correlated and compared with each other to enforce a need-to-know view so that that different observers have a relevant yet limited view of VSS data. The extended AERBAC model allows specification of access control policies in XACML based on the semantic and non-semantic metadata properties retrieved from the VSS data and provides a privacy-aware view for observers according to their authorized purposes defined in terms of events and objects of interest. Metadata-based authorizations enable observers to have permissions against resources (video recordings) based on the 'collection purpose' of the resource. The observers' 'access purpose' for requesting the data complies with the agreed-upon 'collection purposes' described in terms of resource metadata properties to allow compliance with supporting legal bases and data protection regulations. This approach is flexible for applying a privacy-aware access control mechanism, in

large-scale dynamic systems, with a diverse number of observers using the same data for different purposes.

7.1 Future Research Directions

In this section, we will discuss some of the ideas that can be extended from this thesis.

In this thesis we have described a framework to represent 'collection purposes' that can then become part of provenance metadata and is used to ensure compliance by ensuring purpose limitation. However, we have used the existing technologies to store it as an attribute, as there is no specific policy language that can represent different types of information. A general policy language for representing different types of metadata as well as usage purpose can be developed to support data protection legislation as it will be helpful in automated compliance.

Purpose limitation is realizable in large-scale infrastructures where there are trusted data controllers and data processors, however it is hard to ensure in non-trusted public domains. With potential to derive useful information for greater good as well as financial gain, data is being shared and transformed with unlimited possibilities,and data protection legislation alone are not enough to limit secondary use. Large and distributed infrastructure require a technological and operational framework that limits secondary use with the help of usage purposes bound to the (meta)data, and this thesis provides the base-work for that.

Bibliography

[1] E. Cosgrove, "One billion surveillance cameras will be watching around the world in 2021," *CNBC (Dec, 2019), https://www.cnbc.com/2019/12/06/one-billion-surveillance-cameras-will-be-watching-globally-in-2021.html*, 2019.

[2] M. H. Sedky, M. Moniri, and C. C. Chibelushi, "Classification of smart video surveillance systems for commercial applications," in *IEEE Conference on Advanced Video and Signal Based Surveillance, 2005.*, pp. 638–643, IEEE, 2005.

[3] G. M. Segell, "Terrorism on london public transport," *Defense & Security Analysis*, vol. 22, no. 1, pp. 45–59, 2006.

[4] L. Tian, H. Wang, Y. Zhou, and C. Peng, "Video big data in smart city: Background construction and optimization for surveillance video processing," *Future Generation Computer Systems*, vol. 86, pp. 1371–1382, 2018.

[5] V. C. Banu, I. M. Costea, F. C. Nemtanu, and I. Bădescu, "Intelligent video surveillance system," in *2017 IEEE 23rd International Symposium for Design and Technology in Electronic Packaging (SIITME)*, pp. 208–212, IEEE, 2017.

[6] A. Gaur, B. Scotney, G. Parr, and S. McClean, "Smart city architecture and its applications based on iot," *Procedia computer science*, vol. 52, pp. 1089–1094, 2015.

[7] A. P. P. guidelines, "What is personal information?," *APP (May, 2017), https://www.oaic.gov.au/privacy/your-privacy-rights/*, 2017.

[8] J. Van den Hoven, M. Blaauw, W. Pieters, and M. Warnier, "Privacy and information technology," 2014.

[9] Q. M. Rajpoot and C. D. Jensen, "Video surveillance: Privacy issues and legal compliance," in *Promoting Social Change and Democracy Through Information Technology*, pp. 69–92, IGI global, 2015.

[10] E. Union, "Principles relating to the processing of personal data," *EU (May 26, 2018), http://gdpr-info.eu/art-5-gdpr/*, 2018.

[11] M. Garlie, *California Consumer Privacy Act of 2018: A Study of Compliance and Associated Risk*. PhD thesis, Utica College, 2020.

[12] E. Snowden, *Permanent record*. Pan Macmillan, 2019.

[13] C. Campbell, "'the entire system is designed to suppress us.'what the chinese surveillance state means for the rest of the world," *Time (Time, November 21, 2019), https://time. com/5735411/china-surveillance-privacyissues*, 2019.

[14] S. Zhang, Y. Lin, and Q. Liu, "Secure and efficient video surveillance in cloud computing," in *2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 222–226, IEEE, 2014.

[15] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE communications surveys & tutorials*, vol. 16, no. 1, pp. 414–454, 2013.

[16] R. Hai, S. Geisler, and C. Quix, "Constance: An intelligent data lake system," in *Proceedings of the 2016 international conference on management of data*, pp. 2097–2100, 2016.

[17] A. M. Froomkin, "Regulating mass surveillance as privacy pollution: Learning from environemntal impact statements," *U. Ill. L. Rev.*, p. 1713, 2015.

[18] S.-C. Cha, T.-Y. Hsu, Y. Xiang, and K.-H. Yeh, "Privacy enhancing technologies in the internet of things: Perspectives and challenges," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2159–2187, 2018.

[19] E. McCallister, *Guide to protecting the confidentiality of personally identifiable information*, vol. 800. Diane Publishing, 2010.

[20] J. Hoepman, "Privacy design strategies, draft version," 2012.

[21] D. Neal and S. M. Rahman, "Video surveillance in the cloud-computing?," in *2012 7th International Conference on Electrical and Computer Engineering*, pp. 58–61, IEEE, 2012.

[22] S. Sultan and C. D. Jensen, "Privacy-preserving measures in smart city video surveillance systems," in *6th International Conference on Information Systems Security and Privacy*, pp. 506–514, SciTePress, 2020.

[23] Q. M. Rajpoot, C. D. Jensen, and R. Krishnan, "Attributes enhanced role-based access control model," in *International Conference on Trust and Privacy in Digital Business*, pp. 3–17, Springer, 2015.

[24] M. Leiser, "The gdpr: one year on," *Leiden Law Blog*, 2019.

[25] R. Sandhu, D. Ferraiolo, R. Kuhn, *et al.*, "The nist model for role-based access control: towards a unified standard," in *ACM workshop on Role-based access control*, vol. 10, 2000.

[26] S. Sultan and C. D. Jensen, "Secondary use prevention in large-scale data lakes," in *Intelligent Computing*, pp. 967–985, Springer, 2021.

[27] S. Sultan and C. D. Jensen, "Metadata based need-to-know view in large-scale video surveillance systems," *Computers and Security*, vol. 111, p. 102452, 2021.

[28] S. Sultan and C. D. Jensen, "Ensuring purpose limitation in large-scale infrastructures with provenance-enabled access control," *Sensors*, vol. 21, no. 9, p. 3041, 2021.

[29] R. Wenning and S. Kirrane, "Compliance using metadata," in *Semantic Applications*, pp. 31–45, Springer, 2018.

[30] Q. M. Rajpoot, "Enhancing security and privacy in video surveillance through role-oriented access control mechanism," 2016.

[31] I. D. Nogueira, M. Romdhane, and J. Darmont, "Modeling data lake metadata with a data vault," in *Proceedings of the 22nd International Database Engineering & Applications Symposium*, pp. 253–261, 2018.

[32] W. Oliveira, D. de Oliveira, and V. Braganholo, "Experiencing prov-wf for provenance interoperability in swfmss," in *International Provenance and Annotation Workshop*, pp. 294–296, Springer, 2014.

[33] M. D. Allen, A. Chapman, L. Seligman, and B. Blaustein, "Provenance for collaboration: Detecting suspicious behaviors and assessing trust in information," in *7th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 342–351, IEEE, 2011.

[34] P. Missier, K. Belhajjame, and J. Cheney, "The w3c prov family of specifications for modelling provenance metadata," in *Proceedings of the 16th International Conference on Extending Database Technology*, pp. 773–776, 2013.

[35] I. Suriarachchi and B. Plale, "Crossing analytics systems: a case for integrated provenance in data lakes," in *2016 IEEE 12th International Conference on e-Science (e-Science)*, pp. 349–354, IEEE, 2016.

[36] L. Moreau, B. Clifford, J. Freire, J. Futrelle, Y. Gil, P. Groth, N. Kwasnikowska, S. Miles, P. Missier, J. Myers, *et al.*, "The open provenance model core specification (v1. 1)," *Future generation computer systems*, vol. 27, no. 6, pp. 743–756, 2011.

[37] W3C, "Prov-o, The PROV Ontology," 2013. https://www.w3.org/TR/2013/REC-prov-o-20130430/, Last accessed on 2021-05-14.

[38] P. Buneman, J. Cheney, and E. V. Kostylev, "Hierarchical models of provenance.," in *TaPP*, 2012.

[39] M. Strohbach, H. Ziekow, V. Gazis, and N. Akiva, "Towards a big data analytics framework for iot and smart city applications," in *Modeling and processing for next-generation big-data technologies*, pp. 257–282, Springer, 2015.

[40] B. J. A. Miller, "Smart cities are harnessing the power of data lakes for social good," 2019. www.nutanix.com/theforecastbynutanix/technology/smart-cities-harnessing-power-of-data-lakes-for-social-good/,Last accessed on 2021-05-18.

[41] L. Cranor, *Web privacy with P3P.* " O'Reilly Media, Inc.", 2002.

[42] V. Moustaka, A. Vakali, and L. G. Anthopoulos, "A systematic review for smart city data analytics," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–41, 2018.

[43] I. Elizalde Salazar, "Comentario a las «guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications»(comment to the «guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications»)," *InDret*, vol. 2, 2020.

[44] E. Bertino, "Purpose based access control for privacy protection in database systems," in *International Conference on Database Systems for Advanced Applications*, pp. 2–2, Springer, 2005.

[45] M. E. Kabir and H. Wang, "Conditional purpose based access control model for privacy protection," in *Proceedings of the Twentieth Australasian Conference on Australasian Database-Volume 92*, pp. 135–142, Citeseer, 2009.

[46] M. E. Kabir, H. Wang, and E. Bertino, "A conditional purpose-based access control model with dynamic roles," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1482–1489, 2011.

[47] P. Colombo and E. Ferrari, "Enhancing mongodb with purpose-based access control," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 6, pp. 591–604, 2015.

[48] H. Wang, L. Sun, and E. Bertino, "Building access control policy model for privacy preserving and testing policy conflicting problems," *Journal of Computer and System Sciences*, vol. 80, no. 8, pp. 1493–1503, 2014.

[49] M. Gupta, F. Patwa, and R. Sandhu, "An attribute-based access control model for secure big data processing in hadoop ecosystem," in *Proceedings of the Third ACM Workshop on Attribute-Based Access Control*, pp. 13–24, 2018.

[50] D. Basin, S. Debois, and T. Hildebrandt, "On purpose and by necessity: compliance under the gdpr," in *International Conference on Financial Cryptography and Data Security*, pp. 20–37, Springer, 2018.

[51] D. Nguyen, J. Park, and R. Sandhu, "A provenance-based access control model for dynamic separation of duties," in *2013 Eleventh Annual Conference on Privacy, Security and Trust*, pp. 247–256, IEEE, 2013.

[52] D. Nguyen, J. Park, and R. Sandhu, "Adopting provenance-based access control in openstack cloud iaas," in *International Conference on Network and System Security*, pp. 15–27, Springer, 2015.

[53] A. Gehani, D. Tariq, B. Baig, and T. Malik, "Policy-based integration of provenance metadata," in *2011 IEEE International Symposium on Policies for Distributed Systems and Networks*, pp. 149–152, IEEE, 2011.

[54] X. Fan, F. Zhang, E. Turamat, C. Tong, J. H. Wu, and K. Wang, "Provenance-based classification policy based on encrypted search," in *2020 2nd International Conference on Industrial Artificial Intelligence (IAI)*, pp. 1–6, IEEE, 2020.

[55] G. Nanni, "Transformational "smart cities": Cyber security and resilience," *Symantec Corporation*, 2013.

[56] J. Park, D. Nguyen, and R. Sandhu, "A provenance-based access control model," in *2012 Tenth Annual International Conference on Privacy, Security and Trust*, pp. 137–144, IEEE, 2012.

[57] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big data: Issues and challenges moving forward," in *2013 46th Hawaii international conference on system sciences*, pp. 995–1004, IEEE, 2013.

[58] L. Van Zoonen, "Privacy concerns in smart cities," *Government Information Quarterly*, vol. 33, no. 3, pp. 472–480, 2016.

[59] S. Krebs, B. Duraisamy, and F. Flohr, "A survey on leveraging deep neural networks for object tracking," in *2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*, pp. 411–418, IEEE, 2017.

[60] L. Li, W. Huang, I. Y. Gu, and Q. Tian, "Foreground object detection from videos containing complex background," in *Proceedings of the eleventh ACM international conference on Multimedia*, pp. 2–10, 2003.

[61] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *Acm computing surveys (CSUR)*, vol. 38, no. 4, pp. 13–es, 2006.

[62] G. Lowe, "Sift-the scale invariant feature transform," *Int. J*, vol. 2, no. 91-110, p. 2, 2004.

[63] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *International Journal of Computer Vision*, vol. 63, no. 2, pp. 153–161, 2005.

[64] W.-L. Lu and J. J. Little, "Simultaneous tracking and action recognition using the pca-hog descriptor," in *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, pp. 6–6, IEEE, 2006.

[65] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[66] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[67] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[68] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.

[69] S. Hong, H. Noh, and B. Han, "Decoupled deep neural network for semi-supervised semantic segmentation," *arXiv preprint arXiv:1506.04924*, 2015.

[70] H. Yu, J. Wang, Z. Huang, Y. Yang, and W. Xu, "Video paragraph captioning using hierarchical recurrent neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4584–4593, 2016.

[71] R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, S. Chen, Y. Kalantidis, L.-J. Li, D. A. Shamma, *et al.*, "Visual genome: Connecting language and vision using crowdsourced dense image annotations," *arXiv preprint arXiv:1602.07332*, 2016.

[72] L. Matthies, R. Szeliski, and T. Kanade, "Kalman filter-based algorithms for estimating depth from image sequences," in *Multisensor Fusion for Computer Vision*, pp. 87–130, Springer, 1993.

[73] M.-L. Zhang and Z.-H. Zhou, "Improve multi-instance neural networks through feature selection," *Neural processing letters*, vol. 19, no. 1, pp. 1–10, 2004.

[74] C. Ma, J.-B. Huang, X. Yang, and M.-H. Yang, "Hierarchical convolutional features for visual tracking," in *Proceedings of the IEEE international conference on computer vision*, pp. 3074–3082, 2015.

[75] R. Tao, E. Gavves, and A. W. Smeulders, "Siamese instance search for tracking," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1420–1429, 2016.

[76] R. R. Varior, B. Shuai, J. Lu, D. Xu, and G. Wang, "A siamese long short-term memory architecture for human re-identification," in *European conference on computer vision*, pp. 135–153, Springer, 2016.

[77] M. Wang and W. Deng, "Deep face recognition: A survey," *arXiv preprint arXiv:1804.06655*, 2018.

[78] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre, "Hmdb: a large video database for human motion recognition," in *2011 International conference on computer vision*, pp. 2556–2563, IEEE, 2011.

[79] K. Soomro, A. R. Zamir, and M. Shah, "Ucf101: A dataset of 101 human actions classes from videos in the wild," *arXiv preprint arXiv:1212.0402*, 2012.

[80] I. Rodríguez-Moreno, J. M. Martínez-Otzeta, B. Sierra, I. Rodriguez, and E. Jauregi, "Video activity recognition: State-of-the-art," *Sensors*, vol. 19, no. 14, p. 3160, 2019.

[81] H. Wang, Y. Kong, Y. Hong, and J. Vaidya, "Publishing video data with indistinguishable objects," in *Advances in database technology: proceedings. International Conference on Extending Database Technology*, vol. 2020, p. 323, NIH Public Access, 2020.

[82] M. Worring, C. G. Snoek, O. de Rooij, G. P. Nguyen, and A. W. Smeulders, "The mediamill semantic video search engine," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 4, pp. IV–1213, IEEE, 2007.

[83] L. Xie, H. Sundaram, and M. Campbell, "Event mining in multimedia streams," *Proceedings of the IEEE*, vol. 96, no. 4, pp. 623–647, 2008.

[84] M. A. Al-Kahtani and R. Sandhu, "A model for attribute-based user-role assignment," in *18th Annual Computer Security Applications Conference, 2002. Proceedings.*, pp. 353–362, IEEE, 2002.

[85] B. Anggorojati, P. N. Mahalle, N. R. Prasad, and R. Prasad, "Capability-based access control delegation model on the federated iot network," in *The 15th International Symposium on Wireless Personal Multimedia Communications*, pp. 604–608, IEEE, 2012.

[86] N. Dimmock, J. Bacon, D. Ingram, and K. Moody, "Risk models for trust-based access control (tbac)," in *International Conference on Trust Management*, pp. 364–371, Springer, 2005.

[87] "Oasis extensible access control markup language (xacml) tc," 2017.

[88] "Xacml light."

[89] "Sun's xacml implementation," 2006.

[90] "Wso2. xacml 3.0 implementation – balana," 2013.

[91] A. Senior, S. Pankanti, A. Hampapur, L. Brown, Y.-L. Tian, A. Ekin, J. Connell, C. F. Shu, and M. Lu, "Enabling video privacy through computer vision," *IEEE Security & Privacy*, vol. 3, no. 3, pp. 50–57, 2005.

[92] A. Cherif, *Access control models for collaborative applications.* PhD thesis, Université de Lorraine, 2012.

[93] J. D. Moffett and E. C. Lupu, "The uses of role hierarchies in access control," in *Proceedings of the fourth ACM workshop on Role-based access control*, pp. 153–160, 1999.

[94] E. Bertino, J. Fan, E. Ferrari, M.-S. Hacid, A. K. Elmagarmid, and X. Zhu, "A hierarchical access control model for video database systems," *ACM Transactions on Information Systems (TOIS)*, vol. 21, no. 2, pp. 155–191, 2003.

[95] J.-C. Birget, X. Zou, G. Noubir, and B. Ramamurthy, "Hierarchy-based access control in distributed environments," in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No. 01CH37240)*, vol. 1, pp. 229–233, IEEE, 2001.

[96] E. Bertino, M. A. Hammad, W. G. Aref, and A. K. Elmagarmid, "An access control model for video database systems," in *Proceedings of the ninth international conference on Information and knowledge management*, pp. 336–343, 2000.

[97] N. A. T. Tran and T. K. Dang, "A novel approach to fine-grained content-based access control for video databases," in *18th International Workshop on Database and Expert Systems Applications (DEXA 2007)*, pp. 334–338, IEEE, 2007.

[98] B. Thuraisingham, G. Lavee, E. Bertino, J. Fan, and L. Khan, "Access control, confidentiality and privacy for video surveillance databases," in *Proceedings of the eleventh ACM symposium on Access control models and technologies*, pp. 1–10, 2006.

[99] A. Castiglione, A. De Santis, B. Masucci, F. Palmieri, A. Castiglione, J. Li, and X. Huang, "Hierarchical and shared access control," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 850–865, 2015.

[100] N. Zhao, M. Chen, S.-C. Chen, and M.-L. Shyu, "Mrbac: Hierarchical role management and security access control for distributed multimedia systems," in *2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, pp. 76–82, IEEE, 2008.

[101] G. Zhang and C. Yuan, "Hierarchical access control for scalable video coding," in *2010 International Conference on Multimedia Information Networking and Security*, pp. 89–92, IEEE, 2010.

[102] Y. Sun and K. R. Liu, "Scalable hierarchical access control in secure group communications," in *IEEE INFOCOM 2004*, vol. 2, pp. 1296–1306, IEEE, 2004.

[103] F. Cuiyu, L. Aiping, and X. Liyun, "Hierarchical and dynamic security access control for collaborative design in virtual enterprise," in *2010 2nd IEEE International Conference on Information Management and Engineering*, pp. 723–726, IEEE, 2010.

[104] E. Bertino, B. Catania, and B. Shidlovsky, "Towards optimal two-dimensional indexing for constraint databases," *Information Processing Letters*, vol. 64, no. 1, pp. 1–8, 1997.

[105] L. Giuri and P. Iglio, "Role templates for content-based access control," in *Proceedings of the second ACM workshop on Role-based access control*, pp. 153–159, 1997.

[106] S. K. Tzelepi, D. K. Koukopoulos, and G. Pangalos, "A flexible content and context-based access control model for multimedia medical image database systems," in *Proceedings of the 2001 workshop on Multimedia and security: new challenges*, pp. 52–55, 2001.

[107] S. Hastings, M. Ribeiro, S. Langella, S. Oster, U. Catalyurek, T. Pan, K. Huang, R. Ferreira, J. Saltz, and T. Kurc, "Xml database support for distributed execution of data-intensive scientific workflows," *ACM SIGMOD Record*, vol. 34, no. 3, pp. 50–55, 2005.

[108] W. Zeng, Y. Yang, and B. Luo, "Content-based access control: Use data content to assist access control for large-scale content-centric databases," in *2014 IEEE International Conference on Big Data (Big Data)*, pp. 701–710, IEEE, 2014.

[109] B. Fabian, S. Kunz, M. Konnegen, S. Müller, and O. Günther, "Access control for semantic data federations in industrial product-lifecycle management," *Computers in Industry*, vol. 63, no. 9, pp. 930–940, 2012.

[110] L. Kagal, T. Finin, and A. Joshi, "A policy based approach to security for the semantic web," in *International semantic web conference*, pp. 402–418, Springer, 2003.

[111] S. A. Khan and R. Bhatti, "Semantic web and ontology-based applications for digital libraries: An investigation from lis professionals in pakistan," *The Electronic Library*, 2018.

[112] P. Reddivari, T. Finin, A. Joshi, *et al.*, "Policy-based access control for an rdf store," in *Proceedings of the IJCAI-07 workshop on semantic web for collaborative knowledge acquisition*, 2007.

[113] E. Bertino, G. Ghinita, A. Kamra, *et al.*, "Foundations and trends® in databases," *Foundations and Trends® in Databases*, vol. 3, no. 1-2, pp. 1–148, 2011.

[114] A. Joshi, S. Auephanwiriyakul, and R. Krishnapuram, "On fuzzy clustering and content based access to networked video databases," in *Proceedings Eighth International Workshop on Research Issues in Data Engineering. Continuous-Media Databases and Applications*, pp. 42–49, IEEE, 1998.

[115] S. Paradesi, I. Liccardi, L. Kagal, and J. Pato, "A semantic framework for content-based access controls," in *2013 International Conference on Social Computing*, pp. 624–629, 2013.

[116] L. Hu, S. Ying, X. Jia, and K. Zhao, "Towards an approach of semantic access control for cloud computing," in *Cloud Computing* (M. G. Jaatun, G. Zhao, and C. Rong, eds.), (Berlin, Heidelberg), pp. 145–156, Springer Berlin Heidelberg, 2009.

[117] A. Martínez-Ballesté, H. Rashwan, D. Puig, and A. Solanas, "Design and implementation of a secure and trustworthy platform for privacy-aware video surveillance," *International Journal of Information Security*, vol. 17, no. 3, pp. 279–290, 2018.

[118] N. Batra and M. Singh, "Multilevel policy based security in distributed database," in *International Conference on Advances in Computing and Communications*, pp. 572–580, Springer, 2011.

[119] N. Cozma, "How to blur objects in youtube videos," 2016.

[120] K. Yang, Z. Liu, X. Jia, and X. S. Shen, "Time-domain attribute-based access control for cloud-based video content sharing: A cryptographic approach," *IEEE Transactions on Multimedia*, vol. 18, no. 5, pp. 940–950, 2016.

[121] C. E. Boyle, Michael and S. Greenberg, "The effects of filtered video on awareness and privacy," *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 2000.

[122] R. Uittenbogaard, C. Sebastian, J. Vijverberg, B. Boom, D. M. Gavrila, and P. H. N. de With, "Privacy protection in street-view panoramas using depth and multi-view imagery," 2019.

[123] F. Z. Q. Mukhtaj S. Barhm, Nidal Qwasmi and K. el Khatib, "Negotiating privacy preferences in video surveillance systems," *nternational Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, 2011.

[124] M. Drozdowicz, M. Ganzha, and M. Paprzycki, "Semantic access control for privacy management of personal sensing in smart cities," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–1, 2020.

[125] K. L. Huang, S. S. Kanhere, and W. Hu, "Towards privacy-sensitive participatory sensing," in *2009 IEEE International Conference on Pervasive Computing and Communications*, pp. 1–6, 2009.

[126] A. Clarke and R. Steele, "Smartphone-based public health information systems: Anonymity, privacy and intervention," *Journal of the Association for Information Science and Technology*, vol. 66, no. 12, pp. 2596–2608, 2015.

[127] F. Lin, B. Li, W. Zhou, H. Li, and Y. Lu, "Single-stage instance segmentation," vol. 16, (New York, NY, USA), Association for Computing Machinery, July 2020.

[128] J. Wieczorkowski and P. Polak, "Big data and privacy: The study of privacy invasion acceptance in the world of big data," *Online Journal of Applied Knowledge Management (OJAKM)*, vol. 5, no. 1, pp. 57–71, 2017.

[129] J. M. Spaeth, M. J. Plotkin, and S. C. Sheets, "Privacy and the impact of canada's personal information protection and electronic documents act on transnational business," *Vand. J. Ent. L. & Prac.*, vol. 4, p. 28, 2002.