User Configurable Privacy Requirements Elicitation in Cyber-Physical Systems

Tope Omitola
Richard Gomer
Gary Wills
t.omitola@ecs.soton.ac.uk
r.c.gomer@soton.ac.uk
gbw@ecs.soton.ac.uk
Electronics & Computer Science, University of Southampton
Southampton, UK

Niko Tsakalakis
nt1n16@soton.ac.uk
School of Law, University of Southampton
Southampton, UK

Sophie Stalla-Bourdillon
sstalla-bourdillon@immuta.com
Immuta Inc.
USA

Ben Waterson
Tom Cherrett
bjw3@soton.ac.uk
tjc3@soton.ac.uk
School of Engineering, University of Southampton
Southampton, UK

ABSTRACT

The combination of our need for efficient mobility systems coupled with cyber-physical systems has brought about the evolution of Mobility-as-a-Service (MaaS), integrating transport services to provide one-stop access through a custom interface. Our interactions with these MaaS systems lead to a surfeit of data generation and consumption. And for MaaS growth to be sustained, users' trust in the system, especially in their data privacy, needs to be addressed. In this paper, we use LINDDUN privacy analysis framework to elicit privacy requirements of MaaS systems. We show how User-Dependent Analysis, i.e. modularizing complete use cases to different usage contexts and analysing these usages, can help guide us to discern that usage's privacy requirements, which can be enacted by relevant MaaS participants.

CCS CONCEPTS

• **Security and privacy → Information flow control**.

KEYWORDS

Privacy, Cyberphysical systems, cybersecurity, user configuration

UMAP '22 Adjunct, July 4–7, 2022, Barcelona, Spain

ACM ISBN 978-1-4503-9232-7/22/07…$15.00
https://doi.org/10.1145/3511047.3537683

ACM Reference Format:
Tope Omitola, Richard Gomer, Gary Wills, Niko Tsakalakis, Sophie Stalla-Bourdillon, Ben Waterson, and Tom Cherrett. 2022. User Configurable Privacy Requirements Elicitation in Cyber-Physical Systems. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22 Adjunct), July 4–7, 2022, Barcelona, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3511047.3537683

1 INTRODUCTION

With the increasing availability of the Internet of Things (IoT), there is increasing pervasiveness of networked computation in our environments, and data, the outputs of our interactions with these networked computation, are also now ubiquitous. These IoT devices are continually being embedded in many of the physical spaces that we inhabit and/or visit. These in turn has transformed such environments into cyber-physical systems. A **cyber-physical system** (CPS) is one in which there exists a coupling between the computational elements and the physical elements of a system and the environment around the system, with the interactions of these two subsystems leading to a profusion of generated data.

Much of these data are generated either when we are at home or when we are moving from one place to another within our towns and cities, i.e., when we are mobile. Mobility affords a range of societal and economic benefits, from access to services and employment to economic development and cultural exchange. The combination of our need of efficient mobility systems coupled with CPS has brought about the evolution of Mobility-as-a-Service (MaaS) into a concept that promotes the integration of transport services to provide one-stop access through a common interface [9]. MaaS capitalises on our need for effective mobility systems and on CPS to provide access to seamless multi-modal mobility to the end-user. Sometimes, it is often easy to lose sight of the fact that behind these generated data, are information of human beings. These human beings value their distinctiveness, their identity, and their privacy.

Numerous studies indicate that people are either unaware of what private information they are exposing or they do not understand what information they are consenting to share (e.g. [8]). Previous work, [7], has also identified that users' data privacy and fostering user trust in such systems are still to be addressed and are very paramount. In this paper, we investigate how the LINDDUN privacy threat analysis framework can be used for privacy requirements elicitation and analysis of MaaS systems.

2 MAAS AND ITS CONFIGURATIONS

MaaS systems exist in various configurations. In a bare-bones configuration, a MaaS system comprises the MaaS Provider, who offers transport services to customers. Different variations of possible configurations are possible. For example, a MaaS system where the provider of the service is different from the transport operators, supervised by one or more transport authorities, able to service international travel and interfacing with financial service providers, ticket service providers, and tracking service providers.

Such a configuration will involve a number of organisations: the MaaS provider, the supervising Transport Authority, the participating transport operators, the customers' payment providers, the ticketing providers that offer electronic tickets, the tracking providers that monitor the status of vehicles, stations, stops and luggage, and potentially cloud storage providers for data storage, mobile application providers for software engineering, insurance providers for travel insurance etc.

Participating organisations may operate more than one actor within the MaaS system. An actor is an entity that performs a data protection relevant action – most often producing or consuming personal data. For example, the MaaS provider may likely operate a mobile application that will be saving data on the customer's mobile phone but may also be transferring some data in its own premises, the 'back office'.

The data privacy of all participants of this ecosystem is an important issue to address to engender user trust.

3 MAAS AND DATA PRIVACY

Information privacy (and protection) is hardly a new concept. The need to protect valuable information is as old as mankind. Whether that information was the location of a rich hunting ground, a technology for producing weapons, or a knowledge of the divine, it had value and therefore needed protection. This is a truism for modern day MaaS systems, where data generated and consumed by the users of MaaS services have enormous value.

In 1987, Grace Murray Hopper, the coiner of the term "computer bug"[1], made a prescient statement, saying: "Some day, on the corporate balance sheet, there will be an entry which reads, 'Information'; for in most cases the information is more valuable than the hardware which processes it."[2]. Grace Hopper's statement is very pertinent today as the need to ensure the privacy of the data generated by MaaS is highly timely and important.

The data generated and consumed in a typical Maas system can be classified into three broad categories: (1) data pertaining to passengers, e.g. passengers' personal data, (2) data pertaining to the other stakeholders in the system, e.g. transport service providers, and (3) open data. Examples of such data include [13]: journey plans, passenger names, passenger location, etc. Data pertaining to the transportation system itself include route and schedule data, vehicles' location data, maintenance, staff and operations data, and companies' financial data.

There are benefits to data sharing in a typical MaaS. These benefits include: (a) Cost Savings to passengers and transport operators; (2) Transparencies of fare costs, useful for passengers; (3) Savings for both passengers and transport operators. There are also attendant risks. These risks include: (1) Intentional and Accidental Misuse of Data; (2) Security risks against passengers and the transport operators; and (3) Data Privacy.

Privacy is described as a "right to be left alone" [1], and has been defined as the ability to control information about oneself [4]. Therefore, privacy at its core is the ability to determine what information will be shared with others and when it will be shared [6]. In a MaaS, which has data sharing as its core, how these data are managed and analysed, together with the novel privacy vulnerabilities introduced by the data interacting with apps, APIs, the Cloud and hardware, is straining the ability of data subjects to control information about themselves, i.e. their privacy.

And today, the burden of privacy is shouldered primarily by MaaS end-users, i.e., passengers. Passengers, increasingly, have to know a great deal about the capabilities and behaviours of their devices, the operating systems and apps they are using, in addition to the devices owned by others or embedded in the cyber-physical system infrastructure. This is an enormous cognitive load for passengers to carry, in addition to performing their daily goals and tasks. We argue and believe that there is a need to share this burden of privacy amongst the participants of a MaaS ecosystem, and create an ecosystem for privacy [5]. [5] described a good analogy with spam email, where in the early 2000s, people had to spend a great deal of time manually deleting spam from their inbox. However, over time, email service providers started to deploy spam filters, Internet Service Providers started to coordinate in developing deny lists to filter out certain IP addresses, and law enforcement worked with industry to take down botnets and arrest individuals that were mostly responsible for spam. While spam is still an ongoing problem, but over time, end-users did not need to carry that cognitive load of ascertaining which email is spam and which is not, thereby freeing them up to do their job, as the rest of the email service provisioning ecosystem shoulder that burden.

In order for this burden of privacy to be shared amongst MaaS participants, there is a need to perform the privacy analysis of the ensuing subsystems.

4 RELATED WORK IN PRIVACY ANALYSIS

The PRIAM, Privacy RIsk Analysis Methodology [12], framework revolves around a collection of seven components, each of them being associated with (1) a set of categories from which the relevant elements have to be chosen for a given system and (2) a set of attributes which have to be defined and used for the computation of the risks. The seven components are the following: the system, the stakeholders, the data, the risk sources, the privacy weaknesses,

[1] McFadden Origin of the Term 'Computer Bug' https://interestingengineering.com/the-origin-of-the-term-computer-bug

[2] Grace Hopper, https://www.azquotes.com/quote/1043613

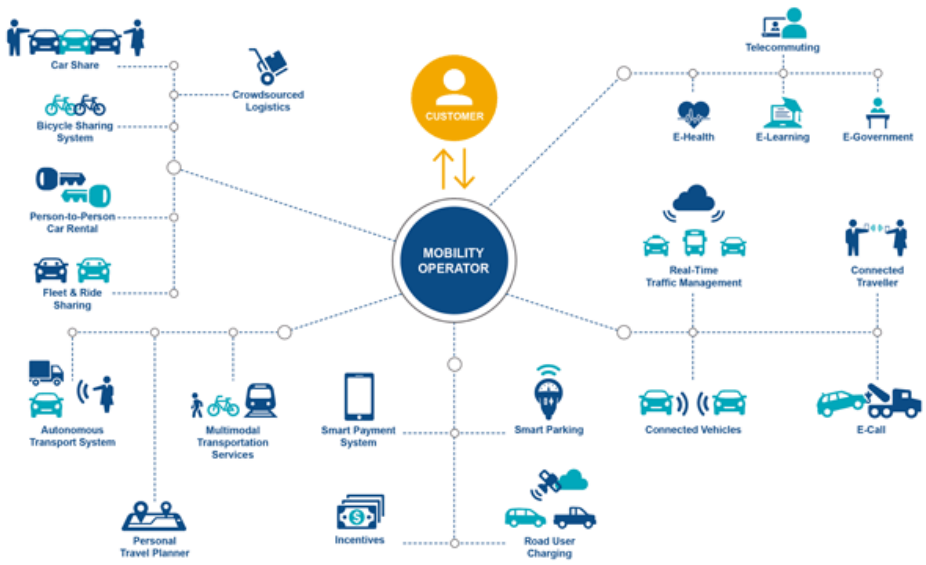

Figure 1: A Mobility as a Service Exemplar (from https://www.businessmaas.com/apps/mobility-service-overcoming-issues-get-critical-maas/).

the feared events and the harms. The system and the set of stakeholders are the inputs of the analysis. The PRIAM methodology consists of two phases: the Information gathering phase and the Risk assessment phase.

Although the PRIAM methodology is comprehensive, it offers little direction in how the analyst can discern the privacy properties of a system and accompanying privacy threats.

The CNIL, Commission Nationale de l'Informatique et des Libertes [2], methodology is intended for data controllers who wish to demonstrate their compliance approach and the controls they have selected to perform a Privacy Impact Assessment of the data generated by the objects they control. A data controller is an entity (individual or organization) that, alone or jointly with others, determines the purpose, conditions and means of processing of personal data. CNIL's principles include: (1) the assessment of the controls guaranteeing the proportionality and necessity of the processing of the data, by, for example, explaining and justifying the choice of the controls, and (2) the assessment of controls protecting data subjects' rights, by, for example, describing the controls chosen to protect data subjects' rights.

Although CNIL brings out many of the data types that may be involved in privacy impact assessments, it focusses majorly on three types of risks on personal data: (1) Illegitimate access to personal data, (2) Unwanted change of data, and (3) Disappearance of data. These risk types are analogues of the C-I-A triad of computer security, i.e., Confidentiality, Integrity, and Authorization. But, in the use of MaaS services, there are more than these three types of risks, some of which CNIL does not mention nor cover. In addition, CNIL does not bring out how data privacy can be regulated and controlled between an ecosystem of business participants, which is exemplified by a MaaS ecosystem, of which there are many participants.

In this work, we use the **LINDDUN** methodology for privacy analysis of MaaS. LINDDUN provides instructions of what privacy issues should be investigated and where in the system, including in the wider ecosystem, these issues could emerge.

4.1 LINDDUN Privacy Threat Analysis Framework

LINDDUN [3] is a model-based threat analysis technique, and relies on a model of the assets in the system-under-analysis. "LINDDUN" is an acronym reflecting the threat categories it aims to uncover: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, and Non-compliance. **Linkability** occurs when one can sufficiently distinguish whether two items of interest (IOI) are related; **Identifiability** occurs when it is possible to pinpoint the identity of a subject (e.g., a user); **Non-repudiation** occurs when it is possible to gather evidence so that a party cannot deny having performed an action; **Detectability** occurs when one can sufficiently distinguish whether an IOI exists, e.g., in a system; **Disclosure of information** is the exposure of information to individuals who are not supposed to have access to it; **Unawareness** occurs when the user is unaware of the information they are supplying to the system and the consequences of that act of sharing; and **Non-compliance** occurs when the system is not compliant with the (data protection) legislation, its advertised policies and the existing user consents

Table 1, from [14] lists these seven types of privacy threats that LINDDUN covers and the desired opposing privacy property to be assured and/or maintained for each threat.

LINDDUN is composed of the following steps: (1) Define the Use Case and use the Use Case to build a Data Flow Diagram (DFD) of the system. The DFD describes the key functionalities within the scope of analysis, and it is a structured, graphical representation

Table 1: LINDDUN Privacy Threat Categories; from[14]

Privacy properties	Privacy Threats	Privacy properties	Privacy Threats
Unlinkability	Linkability	Anonymity and pseudonymity	Identifiability
Plausible deniability	Non-repudiation	Undetectability and unobservability	Detectability
Confidentiality	Disclosure of information	Content awareness Content	Unawareness
Policy and consent compliance	Policy and Consent Non-compliance		

of the system using four major types of building blocks: external entities, data stores, data flows, and processes. It describes the way information flows throughout the system via data flows that connect external entities, to Process nodes, and data stores; (2) Map DFD Elements to LINDDUN Threat Types; (3) Elicit privacy threats, by using threat trees and misuse cases, per DFD element type; and (4) Elicit Privacy Requirements. Here, we use threat trees and especially misuse cases to identify privacy requirements.

The next section describes how we applied LINDDUN to MaaS privacy analysis.

5 PRIVACY ANALYSIS OF MAAS

There are two major ways we can use to analyse the data privacy characteristics of a system or use case:

- Monolithic analysis of the whole system or of the complete use case, which we term **LINDDUN-per-element**. This is a simplified approach to identifying threats and suitable for an incipient privacy analysis, or
- Usage Dependent analysis, which we term **LINDDUN-per-interaction**. In a MaaS, we broaden the term **users** to also include all participants of a MaaS ecosystem, including passengers and service providers. And, here, the whole system can be broken down into subsystems, or the complete use case can be modularised into contexts or usages, with each usage analysed on its own, based on the interactions of the DFD elements in that usage. In LINDDUN-per-interaction, we are acknowledging that threats do not stand in a vacuum, but are results of interactions of the elements of the system. LINDDUN-per-interaction also allows a deeper privacy analysis of each usage context, which can help us discern more threats, when compared to the monolithic analysis method

By providing a privacy analysis of MaaS, and of its individual usage scenarios, allows the different partners responsible for each scenario to enact the implementation of the privacy requirements helping to secure and protect the data privacy of all the members MaaS ecosystem.

The next section describes a complete use case and of its modularization into usage contexts.

6 A MAAS USE CASE STUDY: A TRIP FROM SOUTHAMPTON, UK TO LYON, FRANCE

This case study is constructed around a one-way trip of a passenger, Alex, from the city of Southampton in the UK to the city of Lyon in France.

Alex would like to take advantage of public transport as much as possible, and of the Channel Tunnel from Kent in the UK to Coquelles in France. The Southampton City Council, the transport authority for Southampton where Alex lives, is participating in a MaaS scheme. Using the application of the MaaS platform, Alex can organise and book their journey, from start to finish. The MaaS application 'the App' was developed by the MaaS provider 'the MaaS Provider', which also operates the platform. The Southampton City Council ('the Council') has commissioned the MaaS Provider to include transport methods and bookings for the Southampton area in its platform. Route options, journey times, ticket fares and bookings can all be found through the App.

Through the App, Alex can find the journey times that suit them, arrange the bookings and pay a sum price. Alex has selected to take a train from Southampton Airport Parkway (a UK train station). At Southampton Airport Parkway they will take a train service to London Waterloo, where they will take the Tube to London King's Cross. They will transfer to a Eurostar service that will take them from London King's Cross to Lille Europe in France. There, they will transfer to a service from Lille Europe to Lyon Part Dieu, provided by CNSF Voyagers – the operator that provides long-distance travel in France. At Lyon Part Dieu, Alex will reach their hotel in Lyon on foot, where their journey ends.

6.1 Use Case Modularised Into Different Usages

We have modularized the aforementioned complete use case, above, into different usages. These usages are: (1) Registration, this usage, enacted by Alex, the passenger, is used to register with the MaaS Provider; (2) Journey Requested, this usage is enacted by Alex to request route options from point of departure to point of destination; (3) Journey Availability, usage enacted by Transport Operators can be used to provision the journey; (4) Journey Options, usage enacted to present offered route options to Alex; (5) Booking a Journey – payment, enacted to book and pay for the journey; (6) Starting the Journey, enacted to help Alex commence the journey; (7) Journey In Progress – Monitoring, usage enacted to keep Alex updated of journey's progress; and (8) Processing for Analytics and Planning Purposes, usage enacted whenever any one of MaaS ecosystem participants wishes to perform additional processing, such as analytics, planning, or troubleshooting.

Showing the results of the 8 usages will take up a lot of space, so we will focus on one usage, **Usage 7 Journey In Progress – Monitoring**, it describes a usage that a MaaS business participant needs to enact as part of their service level agreement.

6.1.1 Usage 7 Journey In Progress – Monitoring. During the journey, Alex is able to monitor their journey's progress on the App. They will be able, for example, to monitor their train approaching the platform on a map. Five actors take part in this stage: Alex, the passenger; the App, interfacing with Alex and the providers (in some configurations, the Provider may mediate the communication

between the App and the Service Providers); the Transport Operator(s), offering the transport service; the Tracking Service Provider, providing tracking of the fleet to the Transport Operator; and the vehicle servicing the journey. The data that may be exchanged between the parties are: the Journey Identifier or the particular leg (of the journey) identifier, and the train vehicle GPS data.

7 APPLYING THE LINDDUN STEPS TO MAAS PRIVACY ANALYSIS

As a reminder, the LINDDUN steps are: (1) Build a Data Flow Diagram of the system; (2) Map LINDDUN Privacy Threats to DFD Element types; (3) Elicit privacy threats, by using threat trees and misuse cases; and (4) Elicit Privacy Requirements, by using the outputs from Step 3.

7.1 Step 1: Build a Data Flow Diagram of the System

Both the Monolithic analysis and the Usage Dependent analysis methods can utilize the same DFD, depicted in figure 2.

The stakeholders in this scenario are: (a) Passenger(s), (b) The MaaS App (possibly provided and run by the MaaS Provider), (c) Train Operator A, Train Operator b, and The 3^{rd}-party Ticketing and payment service provider.

In this work, we have chosen the aforementioned stakeholders, focussing on their relationships in order to highlight the salient data privacy issues involved in a next generation MaaS system.

7.1.1 Service workflow at Ecosystem Boundaries. Some interactions take place at the interfaces between the entities of the ecosystem, and they are: (1) Passenger, using the MaaS App plans to make a journey, and uses the 3^{rd}-party Ticketing and payment service provider to plan their journeys, using the 3^{rd}-party ticketing service provider's Journey Planner process, (2) The details of these queries are sent to the 3^{rd}-party ticketing service provider's Customer Relationship Management (CRM) process, (3) The CRM, after running the queries, returns the journey plans details back to the Passenger, (4) If the Passenger is happy with the journey plans, payment is made via the Payment process of the 3^{rd}-party ticketing service provider (4a), and payment acknowledgement is sent to the passenger (4b), (5) Passenger, travelling on the Train, will present their tickets, via the MaaS App, for verification (5a) and verification result(s) sent back (5b), (6) Should the Passenger present multiple or shared journey tickets, these are checked and verified via the Shared Ticketing processes of the transport operators and the 3^{rd}-party ticketing service provider. Also included will be timetable and live service data that will allow the Journey Planner process to function, (8) If the Passenger is in possession of Vouchers, these can be redeemed at either the Tram operator or the Bus operator (7a) and result(s) of redemption sent back to passenger (7b)

7.1.2 Service workflow - Internal. The following interactions and data flows occur within each entity: (9) The CRM process stores data into the CRM database (CRM DB), (10) At set intervals, these data are pulled by the Data Aggregator, (11) ... into the Aggregated database (Aggregated DB) [the data in this database may be used to perform activites such as quality of service assessments, service level agreements audits, etc.]

7.2 Step 2: Mapping Privacy Threats to Element types

This step is composed of two sub-steps: (1) Step 2A: Select the DFD elements from the complete use case and from the modular usages; and (2) Step 2B: Map LINDDUN Privacy Threats to DFD Element types.

7.2.1 Step 2A: Select the DFD elements from the complete use case and from the modular usages. Here, we select DFD elements from the complete use case or from one of the usages.

Step 2A (i) Select the DFD elements from the complete use case. Here, we show, in a tabular format (Table 2), the DFD elements we discover from analysing the complete use case in section 6.

Step 2A (ii) Select the DFD elements from Usage 7 Journey In Progress – Monitoring. For the aforementioned 8 usages, we have been able to observe more elements than the ones enumerated in table 2. We continue to focus our analysis on Usage 7, and table 3 lists the elements pertinent to this usage.

7.2.2 STEP 2B: Map LINDDUN Privacy Threats to DFD Element types. Here, we apply both the Monolithic Analysis and Usage Dependent Analysis to this Usage.

Step 2B (i) Monolithic Analysis. Selecting the relevant entities, processes, dataflows, and datastores from table 2, for Usage 7, we have: (1) Entities: Passr, App, Prov, TSO, TSP; (2) Processes: JM; (3) Dataflows: Passr-App, App-Prov, Prov-Tso, Tso-Tck; (4) Datastores: AppdB, ProvdB, TsodB, and TckDB. **We observe** that Entity TrainVcl and Dataflow TkSP-TrainVcl, **are missing** from the DFD elements enumerated above, **but are present** in the Usage Dependent Analysis of Usage 7 (Table 3). **Being able to modularize a bigger use case to usage dependent scenarios allowed us to notice more DFD elements**, which is a good aid to having a more complete and robust threat model.

After selecting the relevant DFD elements from the complete use case that are pertinent to Usage 7, we then map the LINDDUN privacy threats to the DFD element that can exhibit these privacy threats. This mapping is shown in Table 4 (Note: a cross, X, means the element exhibits the threat for this context, while a dash, -, means the threat cannot be observed in this element for this context).

Step 2B (ii) Usage Dependent Analysis. For this particular part of Step 2B, we are focussing on the use case described in section 6.1.

Table 3 shows the DFD elements for Usage 7, while table 5 shows the mapping of LINDDUN components, i.e. privacy threats, to these DFD elements. We observe noticeable differences between table 4 and that of table 5. **By focussing on one particular usage, we are able to discern, more clearly, how threats can emerge from these elements**, which helped us to specify the presence of possible privacy threats, as exemplified in table 5.

7.3 STEP 3: Elicit Privacy Threats

Instead of enumerating DFD elements by name, we are enumerating DFD elements by type, i.e. Datastore, Process, Dataflow, and

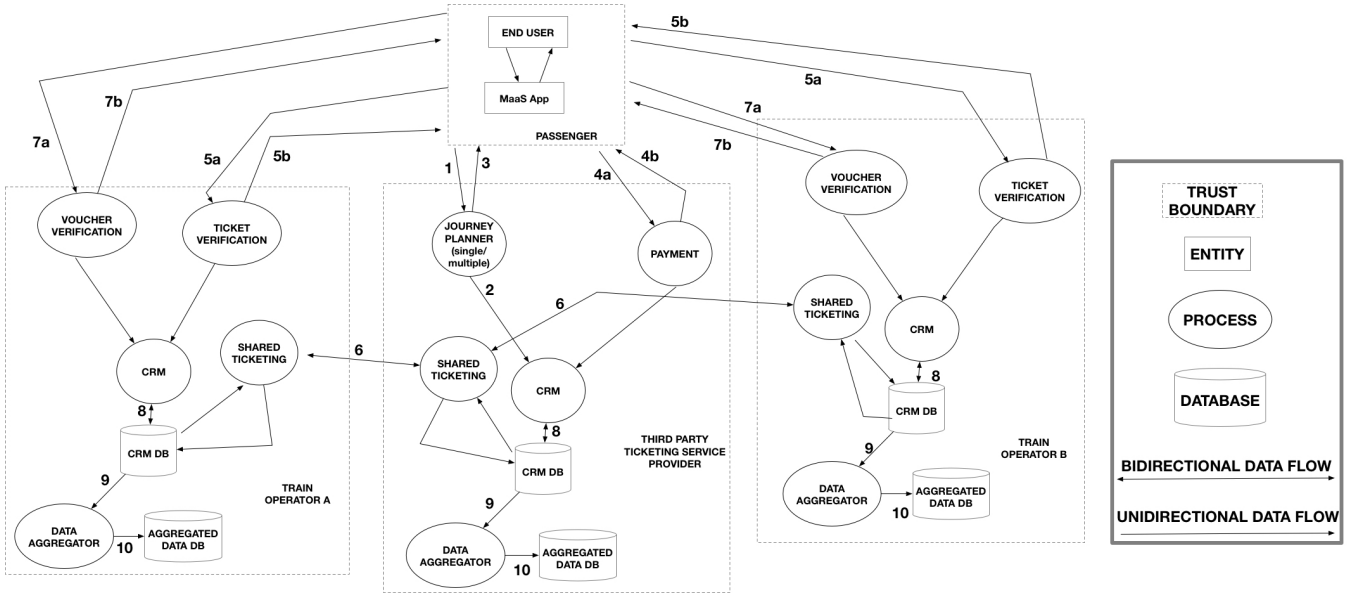

Figure 2: System Data Flow Diagram (labels on data flows are further expanded on in Sections 7.1.1 & 7.1.2)

Table 2: The DFD Elements of The Complete Use Case

DFD Element Type	DFD Element Members	DFD Element Type	DFD Element Members
ENTITY	Passenger (Passr), MaaSApp (App), MaaSProvider (Prov), 3rdPartySoftwareService, FinancialServicesProvider (FSP), TransportServicesOperator (TSO), TransportServicesProvider (TSP), TrackingServicesProvider (TckSP), eTicketInspector, MetropolitanAuthority	PROCESS	PassengerRegistration (PR), StoreRegistrationData (SRD), JourneyPlanner (JP), Booking (Bk), Payment (Pyt), StartingTheJourney (STJ), JourneyMonitor (JM), FinishingTheJourney (FTJ), Analytics (AL)
DATASTORE	MaaSAppDB (AppDB), MaaSProviderDB (ProvDB), TransportServiceOperatorDB (TSOdB), 3rdPartySoftwareServiceDB (3rdPdB), PassengerDeviceDB (PassrDB), TrackingServicesProviderDB (TckDB), FinancialServicesProviderDB (FspDB), TransportServicesProvider (TSPdB), MetropolitanAuthorityDB (MetroAuthDB)	DATAFLOW	Passenger-AppDataStream (Passr-App), App-Provider (App-Prov), Provider-3rdParty (Prov-3rd), App-Fsp, App-TSP, Tsp-Tso, Prov-Tso, Tso-Tck, Prov-MetroAuth

Table 3: The DFD Elements of Usage 7

DFD Element Type	DFD Element Members	DFD Element Type	DFD Element Members
ENTITY	Passr, App, Prov, TSO, TCkSP, TrainVcl (Train vehicle)	PROCESS	JourneyMonitor (JM)
DATASTORE	PassrDevDB (Passenger's device data store), AppdB, ProvdB, TSOdB, TckSPdB	DATAFLOW	Passr-App, App-Prov, Prov-TSO, TSO-TckSP, TckSP-TrainVcl

Entity, and we elicit corresponding privacy threats. This allows us to abstract threats that may be applicable to the class type.

This elicitation involves 2 sub-steps: (1) Step 3A, using threat trees to discern the privacy threats, for the DFD element types, and (2) Step 3B, using misuse cases to document the threat scenarios of these elements.

7.3.1 Step 3A, using threat trees to discern the privacy threats, per DFD element type. Threat trees, also known as attack trees, are conceptual diagrams showing how an asset, or target, might be

Table 4: Map of LINDDUN privacy threats to relevant DFD elements of Complete Use Case to Usage 7

Threat Categories	Passr	App	Prov	TSO	TSP	JM	Passr-App	App-Prov	Prov-Tso	Tso-Tck	AppdB	ProvdB	TsodB	TckdB
Linkability	X	X	X	-	-	X	-	-	X	X	X	X	X	X
Identifiability	X	X	X	-	-	-	X	-	X	X	X	X	X	X
Non-Repudiation	X	X	X	X	X	-	X	-	X	X	X	X	X	X
Detectability	X	X	X	-	-	-	X	-	X	X	X	X	X	X
Information Disclosure	X	X	X	-	-	X	X	-	X	X	X	X	X	X
Content Unawareness	X	X	X	-	-	X	X	-	X	X	X	X	X	X
Consent Noncompliance	-	-	-	-	-	-	X	-	X	X	-	-	-	-

Table 5: Map of LINDDUN privacy threats to the DFD elements of Usage 7

Threat Categories	Passr	App	Prov	TSO	TCkSP	TrainVcl	JM	Passr-App	App-Prov	Prov-TSO	TSO-TckSP	TckSP-TrainVcl	Passr DevDB	AppdB	ProvdB	TSOdB	TckSPdB
Linkability	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Identifiability	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Non-Repudiation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Detectability	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Information Disclosure	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Content Unawareness	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Consent Noncompliance	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

attacked [10]. They provide a formal, methodical way of describing the security and privacy of systems, based on varying attacks and threats, and reflect common attack patterns helping application designers think about privacy conditions in the system. Due to lack of space, we will show threat trees for Linkability privacy threat for the 3 element types of a Dataflow, a Datastore, and an Entity.

Linkability of a data flow threat tree. The linkability of a data flow threat tree, figure 3, suggests two preconditions: (1) the transmitted data between the members of MaaS ecosystem are linkable, which can lead to such data being available to untrusted parties, which ultimately can lead to information disclosure of the data flow; or the data content, itself, is linkable; (2) the other precondition is that contextual data, such as location data, may be linkable. These may be linkable because non-anonymous communications systems are used, which ultimately may allow the contextual data to be linked to the sender's or receiver's IP address, session identifier, or device identifier; or probably the infrastructure has deployed insecure anonymity systems, which may lead to traffic analysis, and passive as well as active attacks on the infrastructure.

Linkability of a data store threat tree. Three conditions correspond to the threat of linkability of a data store (figure 4). These are: (1) there is weak access control to the data store, which can lead to information disclosure; (2) there is weak data anonymization, which can make re-identification possible or the data to be linkable to other data in the data store; (3) the minimization method employed is not sufficient.

Linkability of an entity threat tree. Linkability of an entity (figure 5 refers to when an attacker can sufficiently distinguish if two or more entities are related within the system. This sometimes implies that different pseudonyms are linkable to one another. The threat tree pattern for linkability of the Passenger is shown in Figure C. The passenger identity is linkable if: (1) they login using untrusted communication systems, which can lead to their (different) login interactions being linkable and to untrust-able communications between them and other parties; or (2) passenger personal data is linkable, based on their login IP address or presence of linkable metadata in the (possibly encrypted) personal data.

The next (sub-)step is the application of MisUse Cases. MisUse Cases are employed in LINDDUN to document many of the threat scenarios developed using Threat Trees.

7.3.2 STEP 3B: MisUse Cases. A use case usually describes some function that the system is required to perform. Use cases are used to elicit functional requirements, but not necessarily non-functional ones [11]. Use cases, with actors that initiate them, focus on what a system should do and allow.

A misuse case is the inverse of a use case, it describes a function that the system should not allow. We can define a misuse case as a sequence of actions, initiated by a misactor, resulting in loss for a

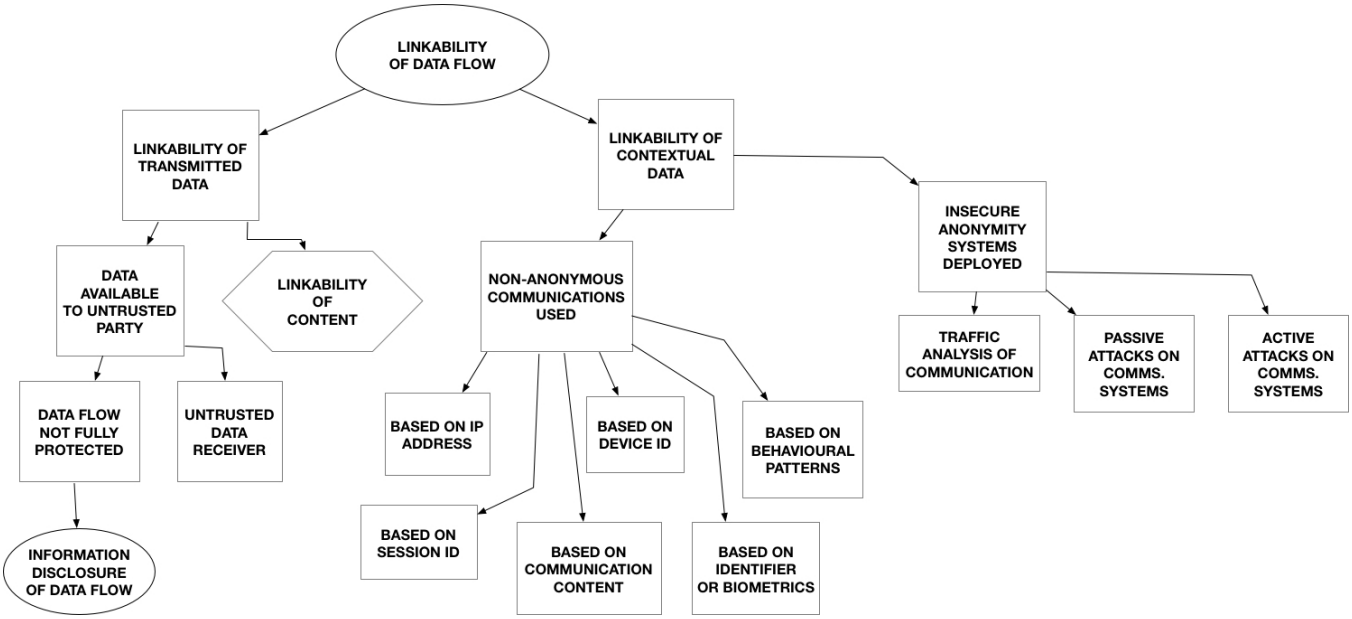

Figure 3: Threat tree for Linkability privacy threat of a Data Flow

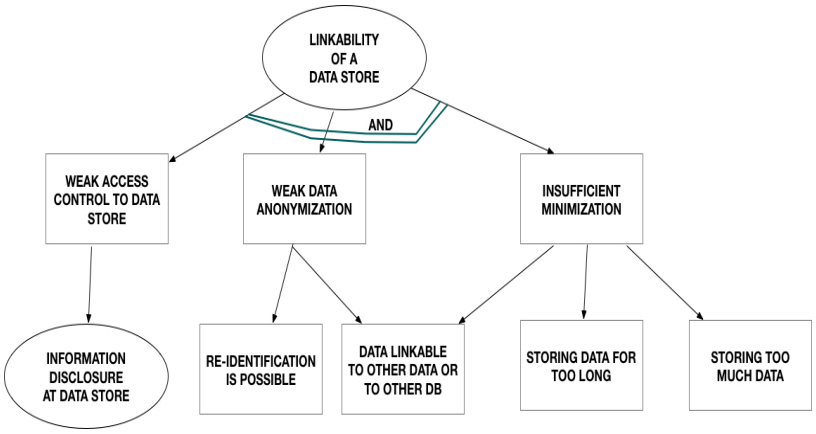

Figure 4: Threat tree for Linkability privacy threat of a data store

stakeholder. A misactor is an inverse of an actor, it is an entity that initiates misuse cases.

The structure of a misuse case that we use in this paper, based on the structure provided by Sindre and Opdahl [11] is: (i) Summary of the misuse case, a description of assets, and if available, the stakeholders and threats; (ii) the primary misactor, i.e., the actor performing the misuse case; (iii) the basic flow of actions, and (iv) alternative flows, if any; (v) a trigger for the misuse case, i.e., how the misuse case may be initiated; and finally (vi) any required preconditions that system must meet for the attack to be feasible.

In this paper, we will describe the misuse cases for Linkability of the Passenger, the MaaS Provider, and the MaaS App Database (AppDB), as these three entities are some of the core actors in a MaaS. We will also describe the misuse case of the Identifiability of AppDB, in order to show how this can be done.

MisUse Case 1 (MUC 1): Linkability of Passenger. The MisUse Case, MUC 1, of the Passenger entity involves the following:

- **Summary** A passenger planning a journey and paying for the journey, probably using a pseudonym, can be linked to their real physical identity.
- **Assets** (Assets of focus, here, are the passenger's personal data): (i) Train tickets' data can be linked to one another which may reveal real life physical identities of passengers, (ii) Attacker can build a profile of a passenger's journey patterns
- **Primary misactor** A skilled insider or a skilled outsider
- **Basic Flow**: (i) Misactor may perform traffic analysis of the interactions/communications between the passenger and the MaaS App, (ii) Misactor may intercept the traffic of the

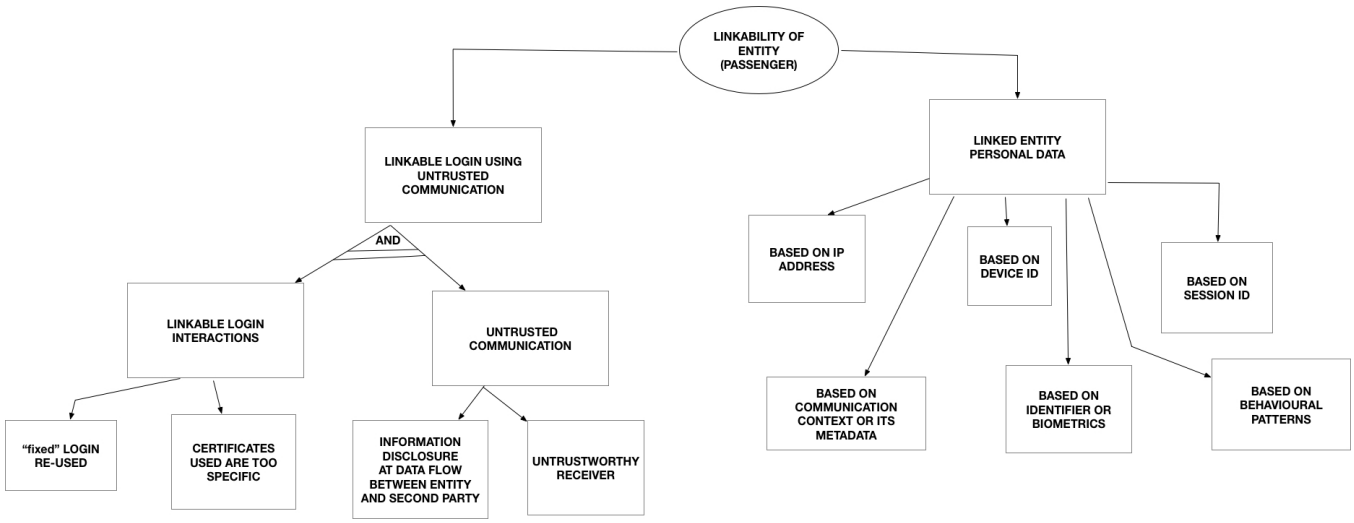

Figure 5: Threat tree for Linkability privacy threat of an entity (e.g. that of passenger)

interactions between the passenger and the MaaS App, (iii) Misactor can link the pseudonyms under which the passenger is known, with passenger's physical identity
- **Trigger** by misactor, can happen whenever there is an interaction between the passenger and the MaaS App
- **Preconditions**: (i) Untrusted and/or insecure communications permitting linkable passenger logins, (ii) Different passenger "pseudonyms" are linked with one another based on passenger interactions with the MaaS App/Provider

MUC2 – Linkability of MaaS Provider. The MisUse Case of the Provider's Linkability is as follows:

- **Summary** The MaaS Provider, by providing a one-shop interface for journey provisioning, communicates with many Passengers, and other data subjects and data processors. The data entries of these different assets in its datastore can be linked to the same subject, without necessarily revealing the subject's identity
- **Assets**: (i) Personal data of passengers, their bank details, and itineraries, (ii) Identifiers of other assets, such as vehicle identifiers of train, (iii) Misactor is able to link passenger identifier with the transportation service(s) they are using, which can be used to track passenger
- **Primary Misactor**: A skilled insider / a skilled outsider
- **Basic Flow**: (i) Misactor gains access to the MaaS Provider, (ii) Misactor gains access to passenger's personal data, including bank details, (iii) Misactor links passenger with transportation service(s) passenger uses
- **Trigger** By misactor, and can always happen
- **Preconditions**: (i) Insufficient or no protection of MaaS Provider, (ii) No or weak data anonymization techniques

MUC3 – Linkability of the AppDB. The MisUse Case of the Linkability of AppDB involves the following:

- **Summary** The App communicates with different subjects, such as the Provider, the Transportation Operators (TOs), and the Financial Services Providers (FSPs); data entries can be linked to the same subject (e.g. an entity, a process, or a stream), without necessarily revealing the subject's identity
- **Assets**: (i) Personal data of passenger, their bank details, and their itineraries, (ii) Event histories of the processes that interface with AppDB, (iii) The misactor can build a profile of the passenger's travel patterns
- **Primary Misactor** A skilled insider or a skilled outsider
- **Basic Flow**: (i) The misactor gains access to AppDB, (ii) The misactor can link processes' event histories together, (iii) The misactor can gain access to passenger's personal data, including their bank details
- **Alternative Flow**: (i) The misactor gains access to AppDB, (ii) Data entries of MaaSProvider, TO, and FSP are linked to pseudonyms, (iii) Based on the pseudonyms, the misactor can link the different pseudonyms together, inside the AppDB or outside of it (such as, on the internet), (iv) Based on the pseudonyms, the misactor can link the different data entries together within AppDB
- **Trigger** By misactor, and can happen whenever there is interaction with AppDB
- **Preconditions**: (i) Insufficient or no protection of AppDB, (ii) No or weak data anonymization techniques

MUC4 – Identifiability of AppDB. The fourth MisUse Case we will describe is the Identifiability of AppDB, and this includes:

- **Summary** The subjects (MaaSProvider, FSP, TO) communicating with AppDB can be identified
- **Assets**: (i) Personal data of passenger, including bank details, (ii) Subjects' data in AppDB
- **Primary Misactor** skilled insider / skilled outsider
- **Basic Flow**: (i) The misactor intercepts (or observes) the movement of data between AppDB and the Database Process Interface (DPI), (ii) This movement contains data and/or data proxies of subjects communicating with AppDB, (iii)

Table 6: Privacy Requirements for Un-linkability and Non-Identifiablity of MaaS elements

Misuse Case	Precondition	Privacy Requirement	Misuse Case	Precondition	Privacy Requirement
MUC1	Untrusted and/or insecure communications permitting linkable logins	**PR1**. The system shall use and maintain trusted and secure communications, at all times	MUC1	Different passenger "pseudonyms" are linked with one another	**PR2**. The system shall ensure strict separation of "pseudonyms", at all times
MUC2	Insufficient or no protection of MaaS Provider	**PR3**. The system shall ensure and maintain strong protection of the named (DFD) element	MUC2	No or weak anonymization techniques	**PR4**. The system shall use and maintain strong anonymization techniques
MUC3	Insufficient or no protection of AppdB	Same as PR3	MUC3	No or weak anonymization techniques	Same as PR4
MUC4	Weak or no deniable encryption	**PR5**. The system shall use and maintain strong encryption, at all times	MUC5	Weak access control in AppdB	**PR6**. The system shall deploy, use, and maintain strong access control in named (DFD) element
MUC6	Subjects wanting deniability of actions are able to do so	**PR7**. In addition to PR6, the system shall deploy, use, and maintain strong non-deniability mechanisms			

Misactor can gain access to Passenger's personal data from observing data movements at the DPI

- **Trigger** by misactor, can happen whenever the DPI reads from or writes to AppDB
- **Preconditions**: (i) Weak or no deniable encryption, (ii) Weak access control in AppDB, (iii) Subjects wanting deniability of actions in AppDB are able to do so

The MisUse Cases are now available to be used to elicit privacy requirements.

7.4 STEP 4 Elicit Privacy Requirements

By modularizing the complete use case into different usages, it is easier, for the analysts and designers, to focus on a particular usage and use the results from Steps 1, 2 and 3 to elicit that usage's privacy requirements. Here, we continue to focus on Usage 7, using the results from Steps 1 to 3 to discern and state its privacy requirements.

While MisUse Cases describe the relevant threat scenarios for the system, their preconditions are based on threat tree patterns, and the basic flows are inspired by the system's use cases. The preconditions of the misuse cases (and also that of the threat trees) can be used as bases for stating the privacy requirements of the system. Therefore, we focus on the MisUse Cases described in section 7.3.2 and especially on the preconditions stated therein, use these as the bases of privacy requirements satisfying the unlinkability and non-identifiability of the identified DFD elements.

Table 6 shows the privacy requirements that can form part of the privacy requirements that a MaaS system needs to assure.

The privacy requirements enumerated in table 6 are meant to be descriptive (and not prescriptive). This allows different MaaS configurations to ground these descriptive requirements, into relevant prescriptions, and implement them to suit their respective MaaS configurations. In addition, by performing Usage Dependent Analysis which allows us to perform the privacy analysis of one particular usage, the business operator(s) enacting and/or materializing that usage can use the analysis, especially its results, i.e. the privacy requirements, in provisioning privacy in their business operations.

8 CONCLUSIONS AND FUTURE WORK

This paper showed how LINDDUN privacy analysis framework can be applied to the privacy analysis of a Mobility-as-a-Service, MaaS, system (an exemplar cyber-physical system). We showed how LINDUN's threat categories of Linkability, Identifiability, Non-Repudiation, Detectability, Information Disclosure, Content Unawareness, and Consent Non-compliance, can help the system's designer to discern the privacy properties of a system. We showed how modularizing a complete MaaS use case to different usages, and by a carrying out a Usage-Dependent Analysis of each usage, the result is a more robust and comprehensive threat analysis. We also described how to derive the system's privacy requirements by applying LINDDUN's systematic approach.

In future work, we will explore how to use the outputs of the four major steps described in this paper to designing mitigating strategies for the elicited privacy threats.

ACKNOWLEDGMENTS

This work has been supported by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1

REFERENCES

[1] A.C. Breckenridge. 1970. *The Right to Privacy.* University of Nebraska Press, Lincoln.
[2] Commission Nationale de l'Informatique et des Libertes. 2015. *Privacy Impact Assessment (PIA) Methodology (how to carry out a PIA).* Technical Report.
[3] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen. 2011. A privacy threat analysis framework: Supporting the elicitation and fulfillment of privacy requirements. *Requir. Eng.* 16 (03 2011), 3–32. https://doi.org/10.1007/s00766-010-0115-7
[4] E.L. Godkin. 1890. The Rights of the Citizen To His Own Reputation. *Scribners* (JULY 1890), 58–67. https://www.unz.com/print/Scribners-1890jul-00058/
[5] Jason Hong. 2017. The Privacy Landscape of Pervasive Computing. *IEEE Pervasive Computing* 16, 3 (2017), 40–48. https://doi.org/10.1109/MPRV.2017.2940957
[6] F.S. Lane. 2009. *American Privacy: The 400-year History of Our Most Contested Right.* Beacon Press.
[7] Carsten Maple, Susan Wakenshaw, and Mariarosaria Taddeo. 2019. *Privacy and Trust Stream.* Technical Report.
[8] C. Miller. 2014. *The Upshot - Americans Say They Want Privacy, but Act as if They Don't.* http://www.nytimes.com/2014/11/13/upshot/americans-say-they-want-privacy-but-act-as-if-they-dont.html
[9] Dalia Mukhtar-Landgren, Marianne Karlsson, Till Koglin, Annica Kronsell, Emma Lund, Steven Sarasini, Göran Smith, Jana Sochor, and Björn Wendle. 2016. Institutional conditions for integrated mobility services (IMS): Towards a framework for analysis. K2 Working Papers 2016:16.

[10] B Schneier. 1989. Attack Trees. *Dr. Dobb's Journal* (December 1989). https://www.schneier.com/academic/archives/1999/12/attack_trees.html

[11] Guttorm Sindre and Andreas Opdahl. 2005. Eliciting security requirements with misuse cases. *Requirements Engineering* 10 (01 2005), 34–44. https://doi.org/10.1007/s00766-004-0194-4

[12] Joyee De Sourya and Daniel Le Métayer. 2016. *PRIAM: A Privacy Risk Analysis Methodology RR-8876.* Technical Report.

[13] Helen Treharne, Stephan Wesemeyer, Steve Schneider, Tracy Ross, Andrew May, Stuart Cockbill, Raja N. Akram, Konstantinos Markantonakis, Simon Blainey, James Pritchard, and Matthew Casey. 2017. Personalised rail passenger experience and privacy. https://eprints.soton.ac.uk/414982/

[14] Kim Wuyts. 2015. Privacy Threats in Software Architectures. https://lirias.kuleuven.be/retrieve/295669$$Dwuyts2014_thesis_online.pdf