SERAMIS

Sensor-Enabled Real-World Awareness for Management Information Systems

Sensor-Enabled Real-World Awareness for Management Information Systems

SERAMIS

Project Deliverable D5.6

Ensuring Customer Privacy: A log of time/cost intensive tasks

Project acronym:	SERAMIS
Project full title:	Sensor-Enabled Real-World Awareness for Management Information Systems
Grant agreement no.:	612052

Doc. Ref.:	D5.6
Responsible Beneficiary:	WUV
Author(s):	Sushant Agarwal (WUV)
List of contributors:	
Reviewers:	Giovanni Romagnoli (UniPR), Alexander Weinhard (UNIWUE)
Contractual Delivery Date:	30.04.2017
Actual Delivery Date:	16.06.2017
Status:	Final

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)

	Dissemination Level	
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the Consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

Executive Summary

In this deliverable, we present a log of activities or tasks related to privacy controls which have been time/cost intensive. The report is divided into three parts: 1) Tasks during the project, 2) Tasks extending after the project, 3) Possible tasks in future. The report helps in identifying tasks that could prove to be intensive for IT systems involved in processing personal data.

List of figures

Table of Contents

1 Introduction

The new EU data protection regulation has been published in 2016 which comes into force in May 2018. The regulation aims at unifying different privacy related legislations existing in different member states. It also sets high standards for protection of personal data processed by companies. In case of non-compliance, the monetary fines could go up to 4% of the global turnover or €20 Million whichever is higher. However, the companies do not have much time left to ensure compliance with the GDPR. Thus, they have to act fast.

For our industry partners as well, ensuring compliance with the GPDR has become a paramount concern and they have started taking action to fulfil the new requirements. In this deliverable we discuss the activities related to ensuring compliance with the GDPR along with other privacy controls related tasks which have been or would be time/cost intensive.

The list of tasks would be of interest for any IT application processing personal data where compliance with the GDPR is to be ensured. We believe that IT applications through this report would get a rough idea of the tasks that might be intensive in terms of time and cost such that they plan the implementation accordingly to ensure compliance before May 2018.

2 Major tasks: during the project

During the project, the industry partners have deployed services like the smart fitting rooms and the Real Time Location System (RTLS). For all the customer related services, privacy was analysed at every step of the implementation. Meetings were organised from time to time to discuss the possible privacy threats.

Broadly, time and efforts put in for privacy related tasks can be broadly divided into 3 categories – 1) Preparation of Data Flows, 2) Collection of RFID data for profiling analysis and 3) Analysing privacy threats and implementation of privacy controls. Figure 1 illustrates a timeline giving an overview of the tasks, meetings and deliverables related to privacy.

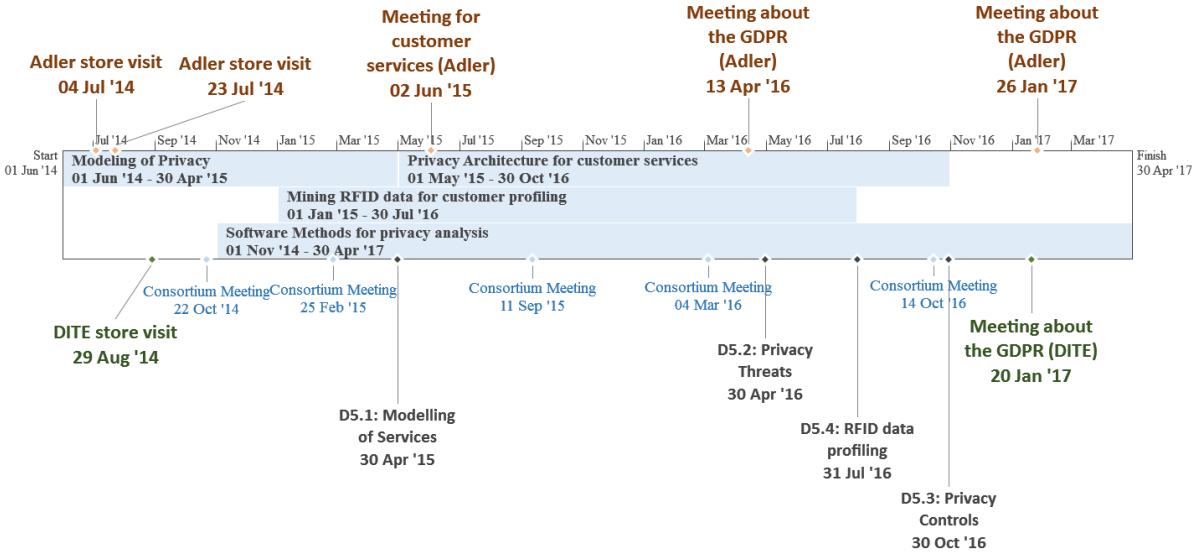

Figure 1: A timeline depicting tasks, deliverables and meetings related to privacy during the project

2.1 Description of tasks

2.1.1 Data flow control

During the initial period, the focus was to understand the flow of customer data (storage and exchange) through the various IT systems. Exchange of personal data between the different systems was modelled along with the description for the exact purpose for storage and exchange. Data flow diagrams were modelled in an iterative way and more granular details were added for every iteration. Figure 2 shows an example of modelling where the process of recommendation service is illustrated. To come up with data flow diagrams, all the services

were described along with the systems and the data involved which have been described in D5.1.

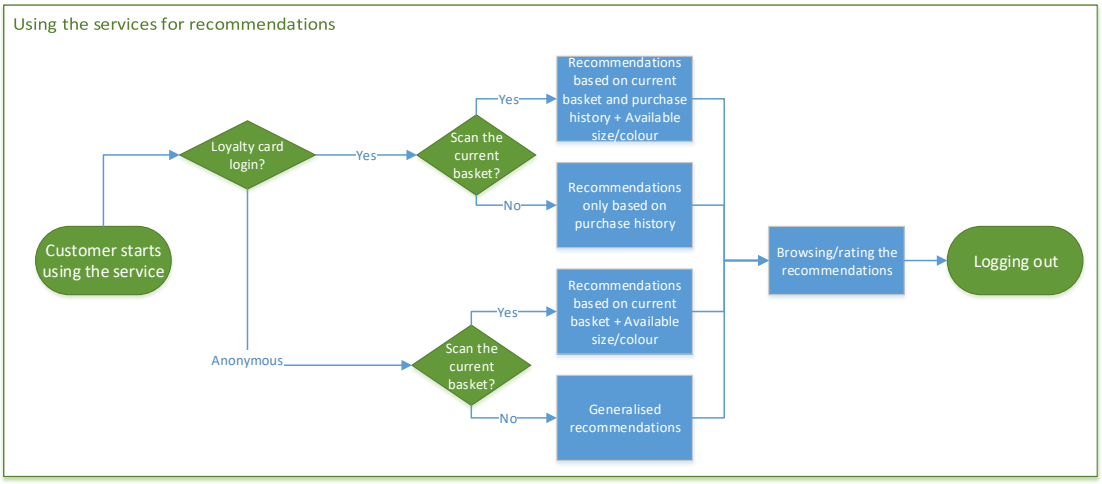

Figure 2: An example of the modelling done: Recommendation service

2.1.2 Data management: collection of RFID data for profiling analysis

After the data flow documents were completed, WUV started with the process of Privacy Impact Assessment (PIA). For the PIA, most of the communication was done via the telephone and Consortium meetings proved helpful to clarify doubts.

During this time, industry partners focussed on the collection of RFID data that was required for customer profiling analysis, elaborated in D5.4. Adler worked on tweaking their RFID systems, such that data log could be stored and shared with the partners. Additionally, smart fitting room was set up for testing as well. Similarly, DITE focussed on collection of Real Time Location (RTLS) Data. Collection of data was challenging especially for the RTLS systems because of the file size of the data logs (>6 GB for 1 cycle, explained in D5.3).

2.1.3 Data protection compliance: analysing privacy threats and implementation of privacy controls

Based on the privacy threats identified (D5.2) and preliminary results from the profiling analysis (D5.4), industry partners in the second half of 2016 spent time in analysing the privacy threats and discussing the proposed privacy controls (D5.3) with their teams. Also, after the GDPR was officially published, substantial time was also spent in understanding what set of requirements change for data processing. A meeting was organised with Adler to discuss the new GDPR in detail and similar discussions took place with DITE through Skype and

telephone. In Jan 2017, meetings were again organised with both the partners to discuss the implementation plan for the privacy controls described in D5.3.

These were the 3 major time consuming tasks for the industry partners during the project which were related to data protection and privacy.

2.2 Analysis of the tasks

The discussed tasks related to data flow control, data management and data protection compliance were mainly part of two work packages, namely WP5: Social & Legislative Implications and WP7: Use Case Implementation & Evaluation. For the tasks, we first estimate the efforts and then calculate costs associated with them.

2.2.1 Estimation of efforts

Based on the estimation in the description of work along with the tasks described in the previous section, Table 1 shows a breakdown of the efforts per partner for the relevant work packages. It is to be noted that WP5 and WP7 extended beyond the tasks related to data flow control, data management and data protection compliance. Thus, a rough estimate of person months has been considered for these tasks.

Table 1: Efforts as person months spent for the work packages

Partner	Work Package	Description	Person Months
Adler	WP5 - Social & Legislative Implications	Analysis of data flows, privacy threats, privacy controls. Analysis of smart fitting room data	4
	WP7 - Use Case Implementation & Evaluation	Implementation of recommended privacy controls, collection of data from the smart fitting rooms	2
DITE	WP7 - Use Case Implementation & Evaluation	Implementation of recommended privacy controls, collection of data from RTLS	1
WUV	WP5 - Social & Legislative Implications	Preparation of data flow diagrams, conducting focus group discussions, analysis of privacy threats, revised obligations based on the GDPR, recommendations for privacy controls	16
	WP7 - Use Case Implementation & Evaluation	Assisting with the implementation of recommended privacy controls and collection of data from various systems	2
		Total	**25**

For WUV, around 16 person months in WP5 were spent directly working with the industry partners for describing marketing services, identifying privacy threats, recommending privacy controls and analysing identifiability of customers in their respective datasets. Similarly, one person month each was spent for WP7 related activities as shown in Table 1.

2.2.2 Costs involved

For each partner, the cost per person month was different based on the differences in salaries and other overhead costs. Thus, to get a rough estimate of cost associated per person month, we evaluated an average for the costs per person month. Table 2 shows the calculation of the costs per person month for the partners.

Table 2: Average costs per person month for the partners

Partner	Person months spent in total	Total estimated eligible costs	Average cost per person month
Adler	15	€ 221,522.60	€ 14,768.17
DITE	11.5	€ 64,706.00	€ 5,626.61
WUV	89.5	€ 723,344.00	€ 8,082.06

Now, to calculate the costs associated with the tasks under consideration, we take the person months estimated from Table 1 and multiply them with the average costs per person month from Table 2, as shown in Table 3.

Table 3: Estimated total average costs

Partner	Person months	Average cost per person month	Estimated costs
Adler	6 (4+2)	€ 14,768.17	€ 88,609.04
DITE	1	€ 5,626.61	€ 5,626.61
WUV	18 (16+2)	€ 8,082.06	€ 145,477.01
		Sum of the costs	**€ 239,712.65**

2.2.3 Comparison of costs and efforts with the rest of the tasks

After estimating the person months and the costs associated, let's compare the values with rest of the tasks undertaken during the project. For SERAMIS, the estimate for efforts in terms of person months for all research related tasks was 372. Also, an estimate for the overall costs associated with these tasks was around € 2,600,597.00. Out of the 372 person months, a total of 25 (6+1+18) have been roughly spent on tasks under consideration. Thus, in terms of efforts, around 6.7% of the total efforts were focussed on tasks related to data management, data flow

control and data protection compliance. In terms of costs, this accounts for approximately 9.3% of the total costs. The breakdown of efforts and costs in illustrated in Figure 3.

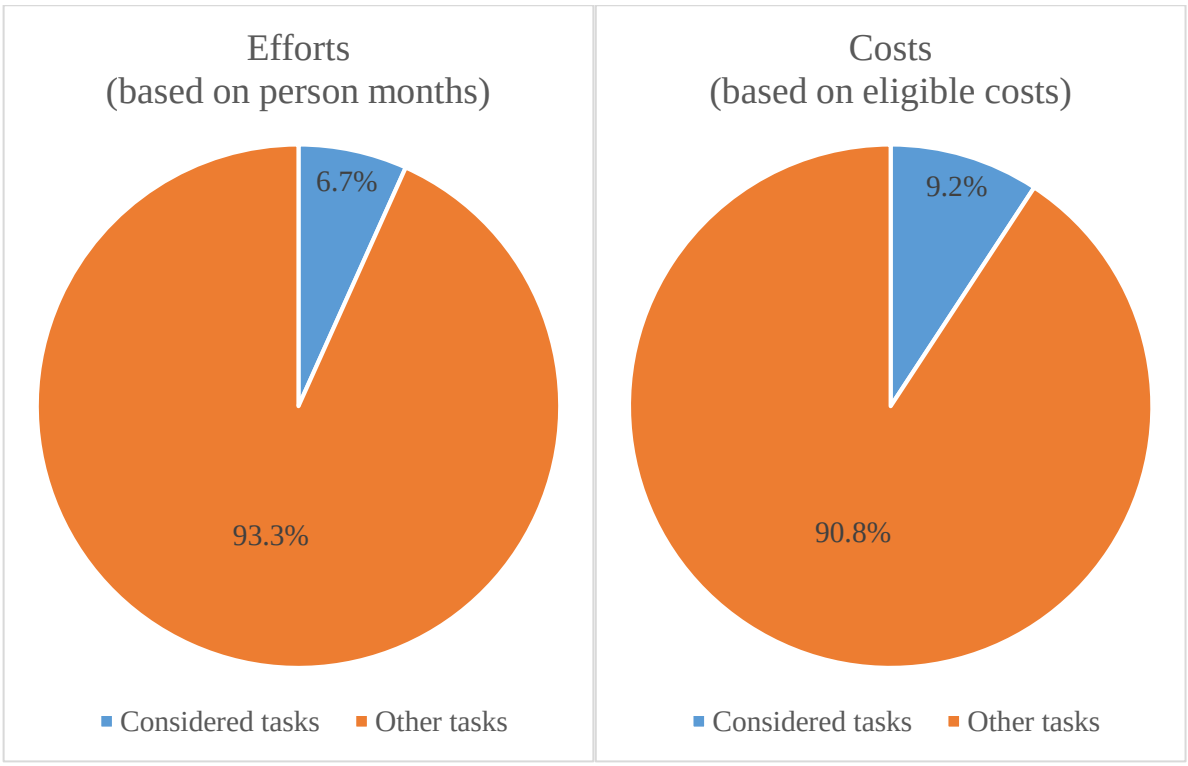

Figure 3: Breakdown of efforts and costs for the project

For the project, Adler's focussed more on customer related use cases along with store operations whereas DITE focussed more on tasks related to inventory management. Thus, it would be also interesting to calculate the breakdown of efforts for Adler. For Adler, 6 out of 15 person months were spent on the tasks under consideration – data flow control, data management and data protection compliance. In terms of percentage, this accounts for 40% of the total efforts as shown in Figure 4.

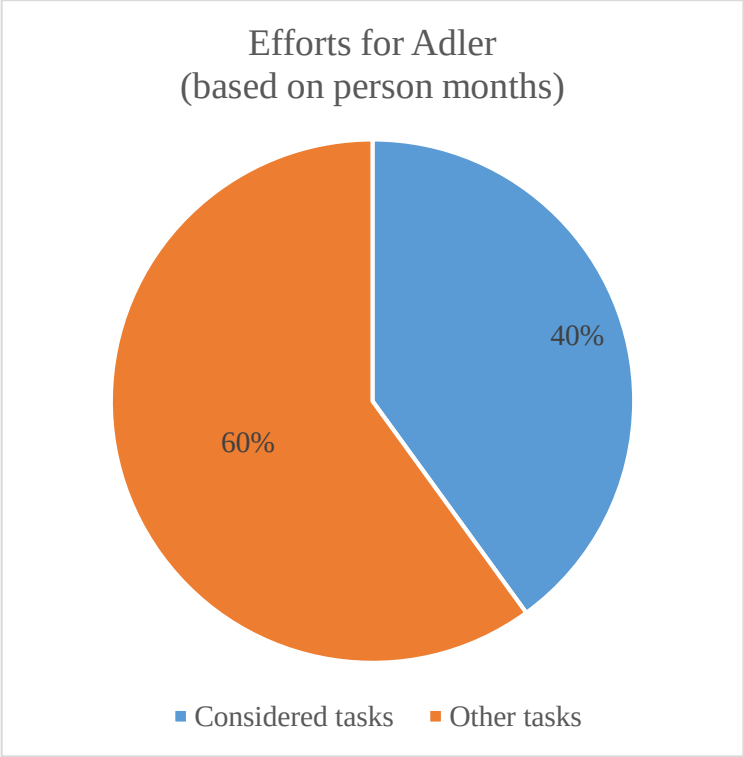

Figure 4: Breakdown of efforts for Adler

2.2.4 Projection of costs for a scenario where tasks undertaken by WUV would be done by Adler

To generalise the costs involved for data management for any company, it would be more accurate to consider a scenario where all the tasks are performed by the company itself. Thus, for the projection we use the person months spent by WUV for tasks directly related to the industry partners' use cases along with the costs associated per person month for the partners.

In total, WUV spent 18 person months for tasks directly related to both the industry partners. Roughly, the distribution was about 50-50 %. Thus, 9 out of the 18 total person months were associated with Adler and 9 were associated with DITE.

In the considered scenario, where WUV would not be involved, we argue that some person months would be saved as employees of the company would be more familiar with the IT infrastructure of the company. WUV spent some time to understand the different IT systems. Hence, it is safe to assume that if WUV's tasks were done by a company, then approx. 2 person months can be saved. This leads to a total of 7 person months for tasks that were undertaken by WUV. While, person months spent by the industry partners would not change.

For the projection of costs involved, we think that Adler's example would provide a better idea as they focussed more on customer related use cases. As discussed previously, Adler spent around 6 person months for data management related tasks. Adding another 7 for the adjusted efforts (7 out of 9) undertaken by WUV, leads to a total of 13 person months. For cost estimation, if we multiply 13 person months with the per person month costs for Adler then we end with a total of **€191,986.30**.

Thus, if a similar company like Adler would have undertaken all the considered tasks then the costs for the company would have been around €200,000.00. However, there are tasks which extend beyond the project for which we provide a short summary in the next section.

3 Tasks extending beyond the project

In this section, we discuss the complex tasks involved with implementation of privacy controls which extend beyond the project. The implementation of privacy controls has already started but more time would be required for completion. The complexity indicates a rough idea of the time and cost that would be associated with the implementation. The controls have been described in D5.3.

3.1 Adler

Adler has already started implementing privacy controls. Following is a list of controls that are in the process of implementation:

Service Description

Regarding the service description, they have prepared initial drafts covering information required based on the GDPR. The description would also cover information for the smart fitting room. However, before finalisation it would be important for Adler to ensure that the description is concise yet easy to understand (as per Article 12 of the GDPR). This might require legal consultation (particularly for terms like *conciseness, transparency, intelligible* etc.) and might consume more time for confirmation as the parameters are not clearly defined in the GDPR.

RFID Emblems and accessibility of information

Figure 5: New store notice with RFID logo at an Adler store

Adler is revamping the notices that they use in the stores such that emblems for RFID would be easily noticeable for the customers. Figure 5 shows the updated store notice which includes the RFID emblem. Currently, Adler is exploring the possibilities to include the RFID emblem

in all the marketing posters used in the store. Thus, the RFID emblems can be seen not only on the price tags but also around the stores on the posters and notices.

Semantics and timeliness of information

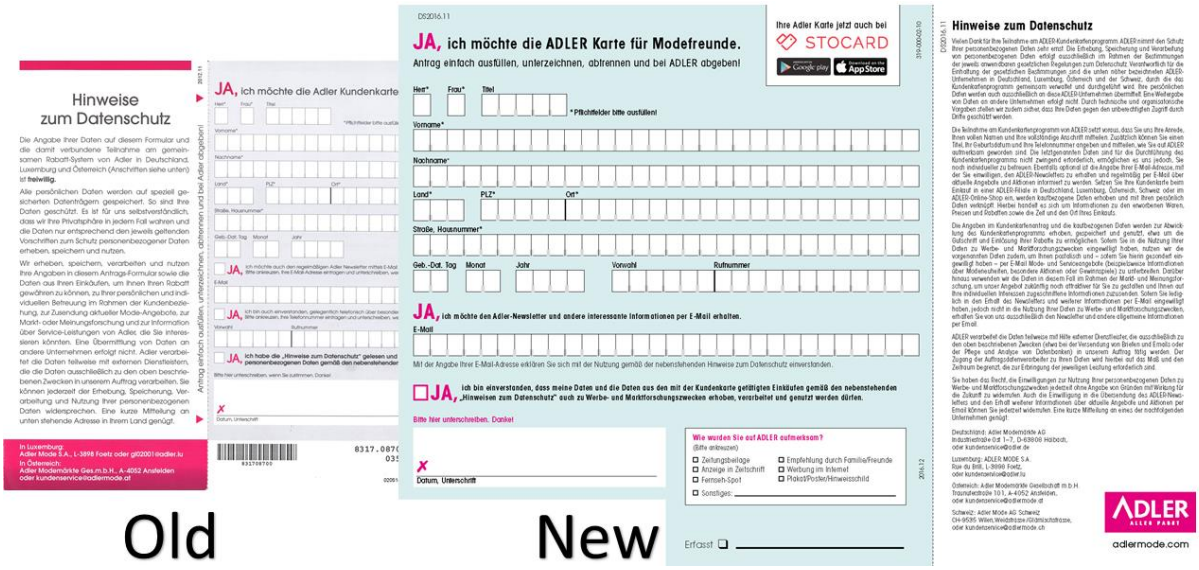

Figure 6: Side by side comparison of the old and new form for loyalty card registration

Adler is also updating the privacy policy both for online shop and for loyalty card used in the stores. Updating the information would ensure the timeliness of the information. Figure 6 compares the new form (not finalised, initial draft) and the old form used for the registration for the loyalty card. The new form differentiates the optional fields better as compared to the old form. Also, the privacy policy is more elaborate in the new draft which also contains contact information for privacy related grievances.

Ensuring purpose related processing & anonymisation

Currently, as RFID data is well separated from the customer data, it is not personally identifiable. However, to confirm the non-identifiability of RFID data with the addition of smart services it is important to audit the systems regularly. Thus, Adler is exploring the possibilities to have periodic audits done by a 3rd party to affirm the non-identifiability of RFID data. This would help them in demonstrating that RFID data is maintained separately and not combined with personal data.

Thus, Adler is working on ensuring that all the information provided and the communication done with their customers related to data processing is done in a "*concise, transparent, intelligible and easily accessible form, using clear and plain language*" (Article 12, para 1).

Due to the unclear description of these terms, we believe that it would be both time and cost intensive to ensure the compliance with respect to information and communication.

3.2 Diffusione Tessile (DITE)

DITE is currently in the process of revamping/upgrading their fidelity card program. Therefore, the controls would be implemented along with the implementation of the new program. Tasks related to providing information and for all communication with customers for data processing would be similar to the tasks discussed for Adler. We would thus discuss other time intensive tasks.

As discussed in D5.3 currently there are 6 different stakeholders for the data. Also, third parties are involved with the process of disseminating the newsletters to the customers. As the GDPR requires the companies to provide rights to customers like the right to be forgotten, it would be important for DITE to ensure proper communication channels among the different systems to handle such rights related requests. In other words, if a customer requests the erasure of data then DITE must delete it from all their systems i.e. from store system, main server, DITE headquarters, third party systems and all other systems/databases storing any personal data from the customer. Currently, the requests are addressed as well but are handled manually. After the GDPR comes into force, we believe that the number of requests might increase making it crucial to automate the processes for erasure which would take substantial efforts to execute in our opinion.

4 Future development processes

In this section, we discuss some tentative plans for extending the customer related RFID use case scenarios which may or may not, depending of the exact implementation plans, require comprehensive privacy impact assessment. Thus, these could be the tasks demanding more time for ensuring that the privacy of customers is not compromised.

4.1 Sewn-in tags

Adler currently integrates the RFID tag with the price label. However, quite frequently the tags are detached or damaged which leads to inefficiency. If tags were to be sewn-in the garments then these sort of issues would not arise. Thus, Adler is exploring the possibilities for using sewn-in tags. Though, it should be noted that there are no concrete plans for its implementation and the idea can be scraped in the future.

Sewn-in tags were discussed in D5.4 where it was concluded that sewn-in tags could potentially make the RFID data personally identifiable and hence could lead to customer profiling. This issue can be solved by either taking a consent from the customers or just deleting/anonymising the RFID data. However, there is one more issue which is difficult to control. Currently, the RFID tags used are not encrypted and can be read by any adversary who has the right RFID reader. Also, as RFID tag contains a unique ID, it would not be difficult to track the customers with the sewn-in RFID tag using this ID. Thus, if sewn-in tags are not implemented properly then they expose the customers to adversaries that might be interested to track them.

To avoid the tracking, Adler has three main options [1, 2]– 1) destroy/deactivate the tag at checkout, 2) Reduce the reading range of the tag at checkout, 3) encrypt the tags. Every option has it pro and cons which need to be analysed. Destroying the tags may reduce the utility, in case of returns, RFID tags with removable antennas are not easily available and encryption in the tags would increase the cost substantially.

Thus, if sewn-in tags are implemented in the future then Adler would have to control all the possible privacy threats such that customers do not face any privacy implecation.

4.2 Indoor location estimation

DITE has already implemented RFID RTLS, which is used for estimation and prediction of location of the garments on the shop floor. As discussed in D5.3, the readers flush out the data after every set of readings. This ensures that the RTLS data cannot be correlated to the customers. However, in future, different techniques could be used for the collection which

might enable retention of RTLS data for a longer time. This could potentially lead to profiling of customers' location in the stores.

Similarly, as smartphones are becoming more popular, companies are now investing in WiFi based location tracking for analysing the customers' movement pattern in the stores. In the past, such techniques have been highly criticised for the underlying privacy implications. An example of such a case is the use of smart WiFi enabled trash cans in London[1], which was later scraped due to the lack of privacy protection.

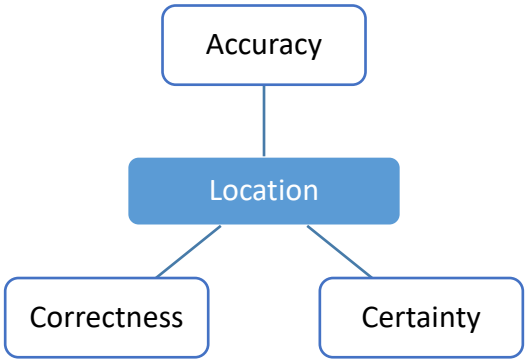

Figure 7: The three dimensions of location accuracy

To mitigate the privacy threats associated with location estimation, companies should either seek consent for processing such data or should anonymise the data such that customers are not identifiable. In case anonymisation is selected then, either the dimensions of accuracy, certainty or correctness (illustrated in Figure 7) for the location data should be adjusted such that no location trace is uniquely identifiable [3]. Otherwise, other anonymisation techniques must be used [4] to strip down the parts which make the customers directly or indirectly identifiable. However, it is important to note that anonymisation should be done using the best practices and analysed periodically to ensure its effectiveness.

Thus, we believe that for anonymisation of location data would require substantial efforts to ensure that it customers are not uniquely identifiable using that data.

[1] https://qz.com/112873/this-recycling-bin-is-following-you/

5 Conclusions

In this deliverable, we discussed all the major privacy related tasks undertaken by the industry partners during the project, tasks that extend beyond the project as well as the probable tasks that might be required in case some discussed ideas materialise in the future. The collected list of tasks gives an idea about which tasks related to privacy could end up being time and cost intensive for the companies.

6 References

[1] G. Karjoth and P. A. Moskowitz, Eds., *Disabling RFID tags with visible confirmation: Clipped tags are silenced*: ACM, 2005.

[2] M. David and N. R. Prasad, "Providing Strong Security and High Privacy in Low-Cost RFID Networks," in *Security and Privacy in Mobile Information and Communication Systems: First International ICST Conference, MobiSec 2009, Turin, Italy, June 3-5, 2009, Revised Selected Papers*, A. U. Schmidt and S. Lian, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 172–179.

[3] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, Eds., *Quantifying location privacy*: IEEE, 2011.

[4] K. El Emam and C. Alvarez, "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques," *International Data Privacy Law*, vol. 5, no. 1, pp. 73–87, 2015.