CLOUD ACCOUNTABILITY PROJECT

D:B-2.3 Workshop 3 results (Use case domain)

Deliverable Number	D22.3
Work Package	WP 22
Version	Final
Deliverable Lead Organisation	SINTEF
Dissemination Level	PU
Contractual Date of Delivery (release)	30/04/2014
Date of Delivery	30/08/2014

Editors

Daniela Soares Cruzes (SINTEF)
Massimo Felici (HP)

Contributors

Daniela Soares Cruzes (SINTEF), Martin Gilje Jaatun (SINTEF), Børge Haugset (SINTEF), Massimo Felici (HP) , Julio Angulo (KAU), Simone Fischer-Hübner (KAU)

Reviewers

Carmen Fernandez Gago (UMA), Michela D'Errico (HP)

Table of Contents

List of tables

List of figures

Executive summary

This deliverable reports the results of the third stakeholder workshop (WS3) for the elicitation work package (WP:B-2) in the A4Cloud project. The first two elicitation workshops (WS1 and WS2) focused on stakeholder understandings of accountability (WS1) and on stakeholder perceptions of risk in the cloud (WS2). In A4Cloud, a stakeholder means a person, group or organization that affects or can be affected by the A4Cloud project results. WS3 presented stakeholders with accountability mechanisms (in particular, software tools developed by A4Cloud) in order to gather their operational experiences (or expectations) about accountability in the cloud. Due to the project decision to focus on a demonstrator to showcase all the tools developed in the project, rather than directly instantiating any one of the three identified use cases from WP:B-3, WS3 focused on A4Cloud tools rather than a specific use case domain.

In order to support focused discussions, we organised different workshops (rather than a single one) for specific cloud actors:

- Cloud subjects (WS 3.1)
- Cloud customers (WS 3.2)
- Cloud providers (WS 3.3).

These groups of cloud actor roles are aligned with the emerging cloud reference architecture (in terms of cloud roles) adopted and extended by the A4Cloud project. Each stakeholder workshop presented and used an accountability mechanism (in the specific cases, a software tool) as a means for stimulating discussions. We demonstrated software tools as a means for gathering feedback, giving stakeholders the opportunity to comment and express their accountability expectations in practice, that is, what they would like to experience (operationally) in the cloud.

In total, about 90 stakeholders (30 Cloud Subjects, 20 Cloud Customers and 40 Cloud Providers) were involved in the five workshops that comprise WS3. Thirty cloud subjects participated in the workshops (WS 3.1). They represented students in their 20s but also professional people of a higher age. In general the cloud subjects were very positive to the tool presented and the concepts of the A4Cloud project. The answers from the participants were very consistent. The cloud subjects were concerned about accountability, and happy to have tools that will help them in accomplishing it. It was clear that being “cloud” or “not cloud” was not a very clear concept for them, but after the explanation of “what the cloud is”, they understood the concept and the risks involved. The workshops were very good in the sense of creating more awareness of the cloud and also about concepts of accountability for the cloud. In total 20 cloud customers participated on our workshops (WS 3.2). They were all IT experts and most of them have worked in information security for some years. In general, the participants were very interested in the concepts around accountability and also on the tools that will be generated by the project. The workshop with cloud providers (WS 3.3) gave us the opportunity to discuss accountability from business perspectives. Most of the discussions, besides giving us some feedback on the accountability mechanism presented, highlighted accountability (and relevant supporting mechanisms) as market enabler for the cloud.

All workshops proved to be fruitful with respect to generating further insights for the tools, accountability practices (or expectations), and for the project in general. We believe that our stakeholder selection and invitation process was suitable for the A4Cloud project. When reflecting on the method for generating discussions, which led to stakeholder feedback, we argue that the method seems to be effective, and we believe that they can be reproduced in other work packages for evaluating and refining the requirements of the tools to be developed in the project.

1 Introduction

The A4Cloud project, by means of the B2 Elicitation Work Package, has engaged with a broad base of relevant stakeholders for requirement elicitation purposes in order to ensure that project activities and results reflect the needs of important stakeholder groups. Interactions with stakeholders were carried out in parallel with the conceptual developments and technical work in other work packages, to enable rapid feedback and validation of interim results. We have been following an approach based on requirements by collaboration [1]. The approach focuses on meeting two essential needs: efficiently defining user requirements while building positive, productive working relationships.

Requirements elicitation is concerned with different objectives. On the one hand, elicitation aims to understand the problem space (*how can we characterise the problem we are dealing with?*) and to identify specific requirements. Addressing this objective tends to give rise to generic requirements charactering the problem we are concerned with. On the other hand, elicitation aims also to fit specific solutions (aligned with such requirements and addressing the characterised problem) to specific user domains. Addressing such objective highlights requirements drawn from stakeholders' domains. Due to the project decision to focus on a demonstrator to showcase all the tools developed in the project, rather than directly instantiating any one of the three identified use cases from WP:B-3, WS3 focused on A4Cloud tools rather than a specific use case domain. The results of the workshops, which are described in this document, will be fed back into the requirements database and provide a reference list for the other work packages. This deliverable reports the results of WS3 that concerned with gathering requirements drawn from stakeholders' domains and their experiences.

1.1 Elicitation Workshops

The aim of involving stakeholders in workshops is to gather a broad spectrum of requirements, good practices and risks related to the cloud eco-system covering the diverse range of geographical (including legal) constraints and challenges, sector/industry-specific requirements and cloud models. As planned in the description of work (DoW) [2], four stakeholder elicitation workshops (Figure 1) are planned in the A4Cloud project, the third of which is documented in this report, named WS3.

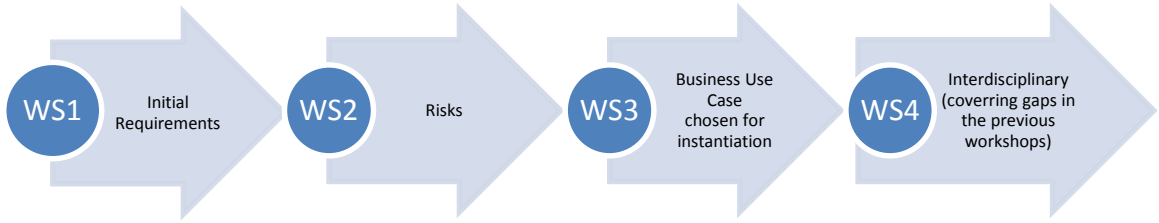

Figure 1 - Elicitation workshops and principal focus for each of them

The main goal of the first workshop (WS1) was to elicit initial accountability requirements from key stakeholders. The first workshop gathered stakeholders' feedback on accountability. It identified initial requirements in the form of accountability relationships [3]. This provided some ideas how to structure cloud ecosystems in terms of accountability relationships between actors. In the second workshop (WS2) the focus was on risks affecting cloud services, and on understanding emerging relationships between accountability, risk and trust [4].

The focus of the third workshop (WS3) was slightly changed from the description in the DoW, which states: *"WS 3 will be organised in the context of the use-case domain chosen for instantiation, MS:B-3.1 , in order to provide input for the detailed description of that use case. The use case will be presented using mock-ups and animations, with assisted walk-through of one or more worked examples. This will take place around M18. WS 3 will be led by SINTEF, with participation by HP, KAU and UMA."* As mentioned above, the project decision to shift focus from a single instantiated use case (instead targeting a demonstrator comprising all the A4Cloud tools) implied that there was no specific use case

domain to focus on. We thus decided instead to tailor the workshop organisation to specific stakeholders in the cloud ecosystem. We have organised a series of workshops (rather than a single WS3) in order to engage with different groups of cloud actors, in particular: Cloud Subjects, Cloud Customers and Cloud Providers. The main objective for the WS3 stakeholder workshops was to engage with them in order to gather experiences drawn from their own cloud ecosystems (their own experiences with cloud services). The organised WS3 workshops presented and used specific accountability tools (i.e. Data Track and the Cloud Offerings Advisory Tool, or COAT) as a means of engaging with stakeholders and gathering their feedback. This allowed us to gather feedback related directly to stakeholder experiences drawn from their own application domains (and reference cloud ecosystems). We have organised and run five different stakeholder workshops (within the WS3 umbrella) of requirements elicitation and refinement, which involved about 80 stakeholders, among cloud subjects (e.g. data subjects), cloud customers (e.g. data controllers) and cloud providers (e.g. data processors). The workshops organisation and methodological background for eliciting requirements and the results will be described in detail in Sections 2 and 3.

1.2 Relationship to Other A4Cloud Work Packages

This deliverable is the third one from WP:B-2 (Elicitation). Results from WP:B-2 will feed into a number of other work packages and deliverables in the A4Cloud project. There will be close interactions between all the WPs within stream B. In particular, WP:B-4 (Socio-economic context) and WP:B-5 (Contractual and regulatory considerations) will provide useful input which will contribute to analysis of stakeholder views, and WP:B-3 (Use-case development) will use workshop results from WP:B-2 as input to the use case descriptions. The stakeholder workshops organised within WS3 also gave us the opportunity to gather feedback on specific accountability mechanisms the project is working on. In the following we list the most important relations between this deliverable and other work packages:

- The goal of **WP:B-3 (Use-case development)** is to provide understanding of 'real-world' scenarios from three distinct user domains in the form of use-cases that inform research and development work throughout the project. The stakeholders involved in the first A4Cloud stakeholder workshop have given important input to the real world scenarios.
- The goal of **WP:C-2 (Conceptual Framework)** is to ensure a common understanding and consistent interpretation of issues relating to accountability and its contribution to trustworthy ICT. Draft content from the scoping report from WP:C-2 (MS:C-2.1) has been used when identifying the initial requirements from the first stakeholder workshop. The results from this report will be fed back into WP:C-2 and provide an initial baseline for work in other WPs in Streams C and D.
- The goal of **WP:A-3 (Dissemination)** is to ensure the proper dissemination of project results, the creation of communities of interest and the execution of training activities. WP:B-2 will rely on communication channels maintained by WP:A-3 to continue engaging with stakeholders.

1.3 Deliverable Organization

The remainder of this report is organized as follows. In Section 2, we describe the stakeholder workshops, their organisation and the elicitation methodology adopted. In Section 3 we present the individual workshops and their main results. We discuss our findings in Section 4 alongside the main concluding remarks.

2 Organisation of the Stakeholder Workshops

The focus or the workshops run in the context of WS3 were the stakeholders of the tools to be delivered by the project. This section describes the elicitation methodology adopted for guiding the organisations of the stakeholder workshops. It also explains how the stakeholder workshops are aligned with the cloud roles identified in the discussion of accountability in cloud ecosystems (within the C2 Conceptual Framework Work Package).

2.1 Elicitation Workshop

A requirements workshop is a structured meeting in which a carefully selected group of stakeholders and content experts work together to define, create, refine and reach closure on deliverables that represent user requirements [1]. Requirements workshops are based on the premise that a small group of knowledgeable, motivated people is more effective than one or two development "heroes". The benefit of the workshop process is that it nurtures team communication, decision-making, and mutual understanding. Workshops are also an effective way to bring together customers, users and software suppliers to improve the quality of software products. Requirements workshops can bridge communication gaps among project stakeholders. Co-creating models in a requirements workshop expedites mutual learning and understanding. By asking focused questions in the workshop, the workshop facilitator helps participants define requirements at different levels of specificity. Each workshop is treated as a mini-project, like any project, each workshop requires planning, role clarification and infrastructure. It has a beginning, middle and an end, as shown in Figure 2. Deliverables are defined beforehand.

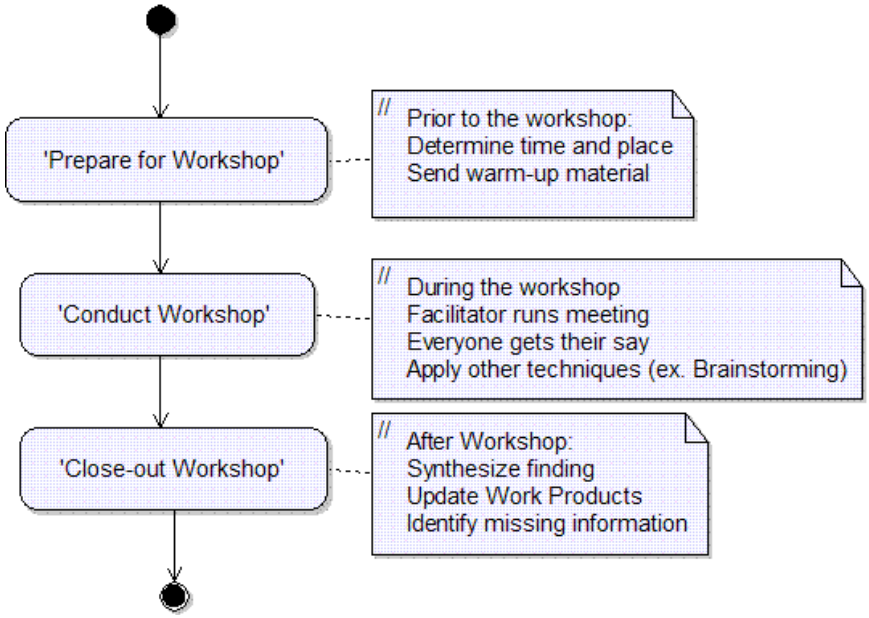

Figure 2 - Requirements Workshop Process

Face-to-face communication and interaction through active stakeholder participation is strongly encouraged when eliciting requirements. However, supporting a face-to-face process is difficult in complex situations involving multiple diverse stakeholders such as in the A4cloud project. To be able to capture the stakeholders' understanding of the concept of accountability, but without influencing them with how the challenges related to accountability are seen from the A4Cloud project, the stakeholders were exposed to software tools and previous elicited requirements from WS1 and WS2. The following motivation for the workshop was presented in the invitation letter to the stakeholders: *"We need to better understand Accountability in the cloud to create better tools and mechanisms that will allow cloud providers to be responsible stewards of customers' data"*. In order to gather stakeholder feedback drawn from their own experiences with cloud services, in each stakeholder workshop an accountability mechanism (in the specific cases, a software tool) was presented and used as a means for stimulating discussions. The first two workshops (WS1 and WS2) focused on stakeholder understandings of

accountability (WS1) and on stakeholder perceptions of risk in the cloud (WS2). The use of software tools as a means for gathering feedback gives stakeholders the opportunity to comment and express their accountability expectations in practice, that is, what they would like to experience (operationally) in the cloud.

2.2 Stakeholder Workshops by Cloud Roles

The C2 Conceptual Framework Work Package has analysed and extended the different cloud roles for the actors in a cloud ecosystem. In A4Cloud, the well-known NIST cloud supply chain taxonomy [5] was extended to create the following cloud accountability taxonomy composed of 7 main roles (the C2 deliverable D:C-2.1 provides a detailed analysis of these roles):

1. **Cloud Subject**: An entity whose data are processed by a cloud provider, either directly or indirectly. When necessary we may further distinguish:
    a. Individual Cloud Subject, when the entity refers to a person.
    b. Organisation Cloud Subject, when the entity refers to an organisation.
2. **Cloud Customer**: An entity that (a) maintains a business relationship with, and (b) uses services from a Cloud Provider. When necessary we may further distinguish:
    a. Individual Cloud Customer, when the entity refers to a person.
    b. Organisation Cloud Customer, when the entity refers to an organisation.
3. **Cloud Provider**: An entity responsible for making a cloud service available to Cloud Customers
4. **Cloud Carrier**: The intermediary entity that provides connectivity and transport of cloud services between Cloud Providers and Cloud Customers.
5. **Cloud Broker**: An entity that manages the use, performance and delivery of cloud services, and negotiates relationships between Cloud Providers and Cloud Customers.
6. **Cloud Auditor**: An entity that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation, with regards to a set of requirements, which may include security, data protection, information system management, regulations and ethics.
7. **Cloud Supervisory Authority**: An entity that oversees and enforces the application of a set of rules.

We used the identified roles in order to organise the WS3 stakeholder workshops based on the different cloud roles. In particular, we focused on three main cloud roles, namely, Cloud Subjects (WS 3.1), Cloud Customers and Carrier (WS 3.2) and Cloud Providers (WS 3.3). Figure 3 highlights how the different WS3 stakeholder workshops cover the identified roles.

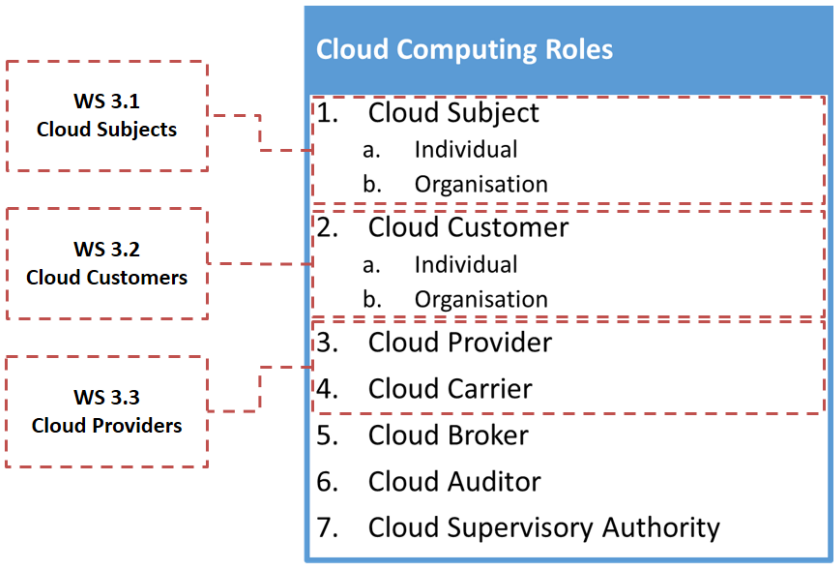

Figure 3 - Workshops' organization covering different cloud computing roles

Next section reports the main stakeholder feedback gathered for each individual workshop. We believe that our stakeholder selection and invitation process was suitable for the A4Cloud project. When

reflecting on the method for generating discussions, which led to stakeholder feedback, we argue that the method seems to be effective, and we believe that they can be reproduced in other work packages for evaluating and refining the requirements of the tools to be developed in the project.

3 Stakeholder Workshops

In the context of WS3, we have performed five different stakeholder workshops as shown in Figure 4. Our focus was to ask the stakeholders to primarily express functional requirements. We have also focused on two of the tools (i.e. the Data Track Tool and the COAT Tool) that are developed in the A4Cloud project. Transparency of the data processing and sharing practices of online services play a key role not only for endowing users with control over their own data, but also for the prosperity of democratic societies. Studies have also shown that transparency from the service provider can promote trust on that service [7]. For this reason we ran a set of interviews on transparency requirements, to both elicit requirements of transparency and to refine the requirements we have elicited in previous workshops in the project. We assume that by using the same approach, the requirements from other tools can be elicited and refined. As shown in Figure 4, we run two workshops focusing on the perspective of the cloud subjects, with a total of 30 stakeholders. One workshop and a set of interviews focusing on the perspective of the cloud customers, with a total of 19 stakeholders; and one workshop focusing on the perspective of the cloud providers, with 30 stakeholders. The results of the workshops are presented in the sections below.

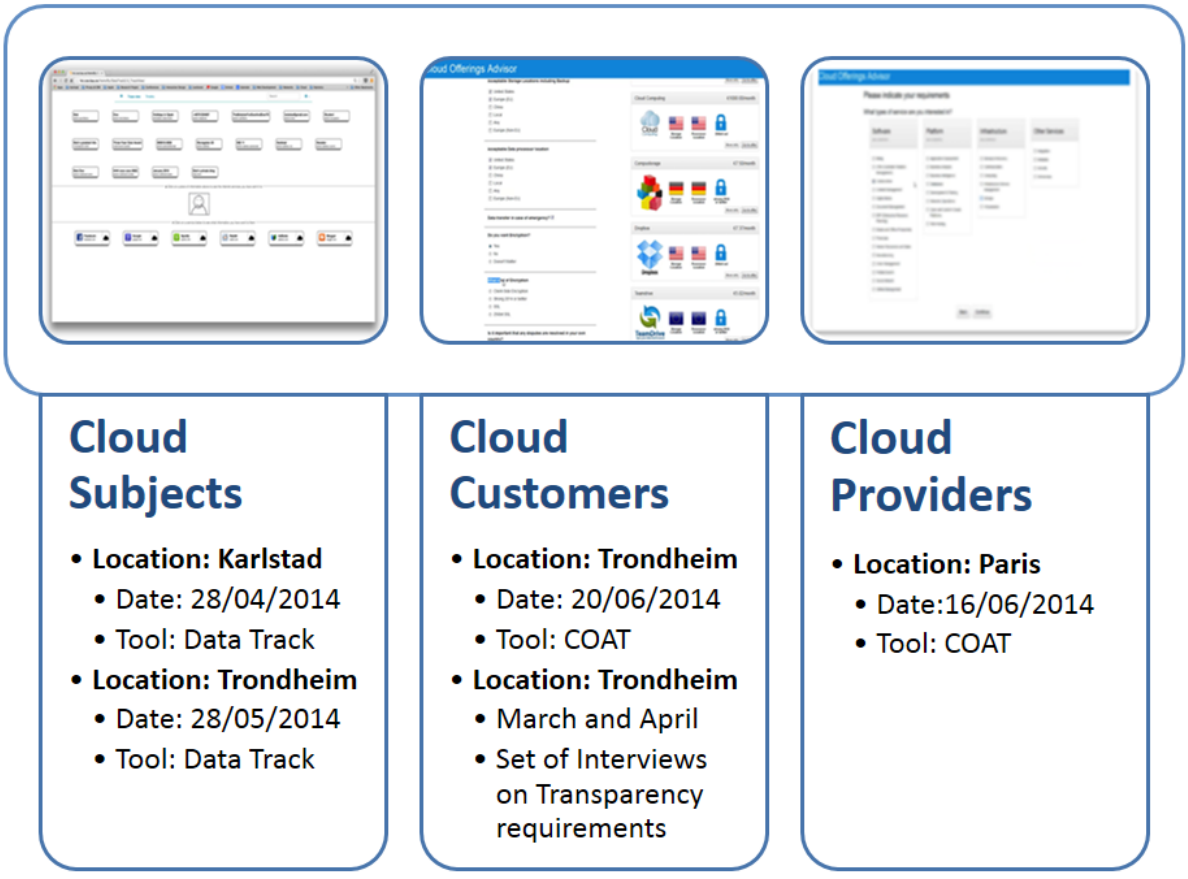

Figure 4 - WS3 stakeholder workshops

3.1 Workshop 3.1: Cloud Subjects

Cloud subjects were involved in the refinement and elicitation of requirements in two different locations. We used the Data Track tool as a way to expose them to the A4Cloud concepts and tools. The reason for choosing Data Track was that the users could better understand the concepts behind the tool once that it is more closely related to their day-to-day activities on the web. Section 3.1.1 presents the results from the workshop we ran in Karlstad with 19 participants and section 3.1.2 presents the results from the workshop we ran in Trondheim with 11 participants.

3.1.1 Karlstad Workshop

The goal of this workshop was to present the Data Track tool and collect the participants' perceptions on the tool and refine the requirements that are stated for the tool. All the material used in the workshop can be found in Appendix C.

3.1.1.1 Selection of stakeholders

Invitation was sent to students at Karlstad University (KAU) to participate in a workshop on the use and opinions about Internet services and possible solutions on how to keep track of their personal data. It was said to the students that different solutions would be discussed and demonstrated prototypes would be presented. Participation was voluntary, but participants received breakfast and two cinema tickets. The workshop took place on the 28th of April from 9:00 to 11:30 at KAU. Nineteen people answered the questionnaire and attended the workshop.

3.1.1.2 Data Collection and Analysis

As described in Appendix A, the data collection comprised a pre-questionnaire, recordings of discussions and a post-questionnaire on the perceptions of the tool. The goal of the pre-questionnaire was to understand the participants' behaviour on the cloud, to to rate their trust on some services and also rate how "sensitive" (private) the different personal data items are for the participants. This pre-questionnaire also contained a consent form for the researchers to use the data collected for research purposes. Informed consent was collected according to guidelines from computer science research and legal requirements pursuant to Art. 10 EU Directive 95/46/EC.

During the session we presented the project and the tool. After this presentation, we divided the group in two groups of 9 people and asked them to discuss the following questions for 20 minutes:

1. Do you think you will use this tool?
2. In which context do you think this tool will be interesting?
3. What did you like the most about this tool?
4. What did you like the least about this tool?
5. What would you like this tool to notify you about?

The discussions were recorded and transcribed by the two facilitators of the sessions (see transcriptions in Appendix C).

After the discussions the final post-questionnaire was handed out (see Appendix C). The post-questionnaire comprised four main sections. The first three sections asked the participants to assess their agreements with statements on:

i) Usability questions about the tool (for example: the tool is easy of use, the tool is easy to learn quickly). Options were: strongly disagree, disagree, neither, agree and strongly agree.
ii) How well the tool will help to accomplish the goals of the project (for example: the usage of the tool will enable cloud service providers to give their users appropriate control and transparency over how their data is used). Options were: strongly disagree, disagree, neither, agree and strongly agree.
iii) Functional requirements of the tool (for example: The Tool should let me know which data is collected by which services and for what purpose they are going to use my data.). Answers could be from not important (0) to very important (5).

The last section asked the participants to freely write other comments (extra requirements, improvements, suggestions, recommendations, justification of their answers) about the tool. All data from the questionnaires were tabulated and analysed quantitatively; all the qualitative data from the session were analysed and also incorporated in the results. All the details of the answers are shown in Appendix C and the main results are shown in the next section.

3.1.1.3 Results

From the 19 participants, seven are from 18 to 23 years old, and 9 from 24 to 30 years. Only three participants were over thirty years old. Figure 5 shows how people rate their trust in services, from very low trust (red) to very high trust (green). In general, people showed to trust in large scale banks, credit cards and government services. Also, they demonstrated a high trust in service providers such as Amazon, Paypal, Microsoft and Apple. Figure 6 shows how sensitive various types of data are for the participants in the workshop, from very public (green) to very private (red). Data, such as name, gender, age, and nationality are seen as "public" data. Other than these, all data are sensitive somehow to people in different levels of sensitiveness. More details on the answers are provided in the Appendix C. It is important to note that these ratings can be used for prioritizing which information should be made secure. In addition this information can be used to start creating groups of information in categories, as pointed by the participants as one nice to have feature in the tools.

On the question of how much the participants agree with the level of usability of the tool, the participants agree that:

1. The tool will probably be used frequently in the future.
2. The tool is not complex.
3. The tool is easy to use.
4. The tool is easy to learn quickly.
5. The tool's functionalities are clear and understandable
6. The tool integrates well with current practices.
7. The tool integrates well various functionalities.

The participants had different views on some other questions on usability:

8. Some believe that the tool requires user training and some said neither, but there was a big variation in the opinions.
9. The participants in general average believe that the tool does not require users to learn a lot of new concepts, but there was a big variation in the results.
10. The participants showed to neither agree nor disagree with the affirmation that: the tool lacks many useful functionalities.

On the question on how participants agree that the objectives of A4Cloud are accomplished by the Data Track tool, participants agree that the Data Track tool:

11. gives users more control and transparency over how the data is used in the cloud

In general, participants agreed, with some cases of neutral, that the Data Track tool:

12. enables cloud service providers to give their users appropriate control and transparency over how their data is used.
13. enables users to make choices about how cloud service providers may use and will protect data in the cloud.
14. will make the relationship between users and providers substantially easier because it will be easier to see who is responsible for the problems.

The participants were neutral or disagreed to the following statements on the Data Track tool:

15. will substantially increase users' trust in cloud services.
16. will substantially reduce the number of serious security problems.

On the question on how the participants rated the importance of the listed functional requirements of the tool, all functionalities were rated as important, without much variation on the opinions. The only one that was not averaged as important was requirement 30, with a great variance on the answers. Requirement 29 was also a bit controversial, so not all believe this is important. The list of requirements is as follows:

17. The Tool should let me see which of my data was sent by me and which data was collected automatically by the service
18. The Tool should let me know which data is collected by which services and for what purposes are they going to use my data.

19. The Tool should inform me whether the Cloud service stores my data in an encrypted form or if they can see the contents of my data.
20. The Tool should let me check how my data has been processed by the Cloud service and what conclusions they can draw about me based on this.
21. The Tool should allow me to find out whether my data has been used in a way that was not specified in the privacy policy when I sent my data.
22. The Tool should allow me to see how my data has been passed on to other Cloud services.
23. The Tool should allow me to correct (edit) or to delete the data that the Cloud services have about me.
24. The Tool should let me know the country in which my data is stored and the laws that apply to that country.
25. The tool should inform me when other people send data about me to a Cloud service.
26. The Tool should provide a report that tells me how risky or secure is to have my data in a Cloud service.
27. The Tool should inform me about the risks and threats associated with Cloud services.
28. The Tool should enable assessing the security level of the service providers.
29. When the Cloud service doesn't do what they promised me at the time of registration, the Tool should let me do something about it.
30. The Tool should send me many notifications per week to inform me how the Cloud service is handling my data.

Table 1 shows the list of additional requirements elicited for the data track tool, divided by the session where it was elicited.

Table 1 - List of Elicited Requirements – Data Track Tool – Cloud Subjects Karlstad

List of Elicited Requirements	
From Discussions Group 1 (Julio)	# Providers should have a list of all data they collect about users. # Data track should show a ranking of the providers (like a reputation system). # Data track should classify data and track only data that are set as sensitive to the user. # Accountability tools could have a monitoring tool for showing the status of the data. # Data track should help tracking information such as videos and photos. # Data Track database should be encrypted. # Providers should be transparent about mechanisms used to protect data.
From Discussions Group 2 (Daniela):	# Data Track tool should be more pro-active. # Data Track should help with some warnings about the security of the data. # Data Track should have a very secure system to protect all the data that is collected about services. # Data Track tool should have a strong security level on its database. # Data track should help users to delete the data they do not want to have spread on the providers (direct link, set of steps, guide, etc). # Data Track should help to find the policies from each provider. # Data Track should help users proactively by warning them about both explicit and implicit data that providers collects about users, so I can act before user submit data. # Data track should allow user to classify data in different levels of sensitiveness and also different classes of data (banking data, health data, personal information etc). # Providers should inform users if some breach happens to the data. # On Transparency, providers should give information about: where is the information, if they have connection with other companies, other countries. # For accountability, providers should have policies written in a language that is more understandable for users. # Upfront information that increases upfront trust: reputation, personal recommendations, reviews on websites. # Accountability tools should provide a plugin to tell me that something about me is being sent to somewhere, I would not open the website for trying to check which data will be spread to a server. # Data track should allow classification and prioritization of data. # Providers should give enough information to users about changes and the impact of the changes to the user. So the user can be aware of the risks that the change will imply.
From Forms	# Data Track tool should have a status of how safe a person is based on the data from providers. # Data Track should have different profiles of data information, so you do not see all at once but

	the different data from the different profiles. # Data Track should have a guide to how to delete data in different providers or a link to the where the provider explains how to delete data from the provider. # Data track should give a warning if the user is sending more data than he/she wants. # Data Track should show the level of security of a certain data type that the user is concerned . # Data Track could have a weekly, monthly or annually "warning" of what is the safe level of the data.

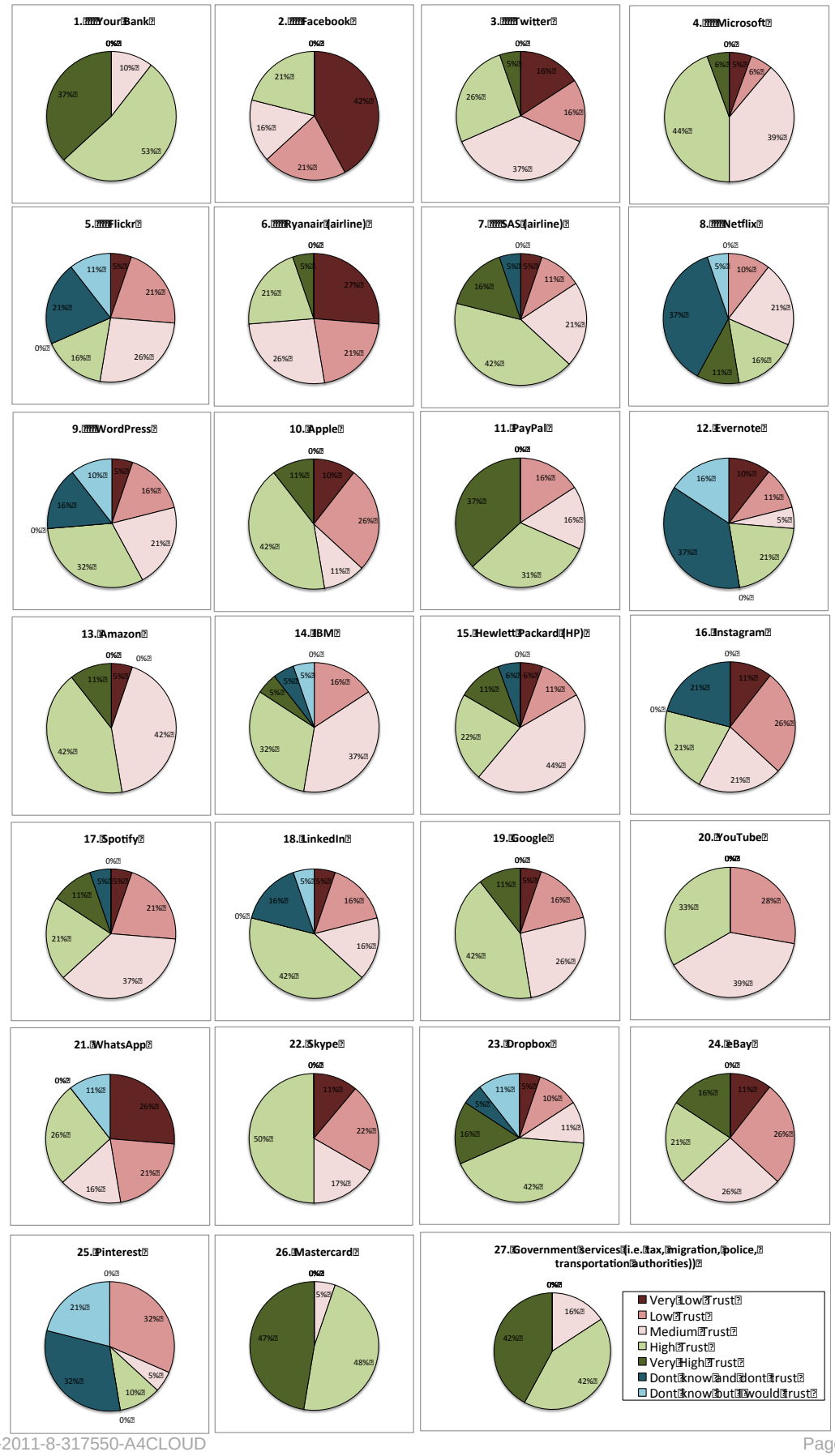
1. Your Bank
0%
10%
37%
53%
2. Facebook
0%
21%
42%
16%
21%
3. Twitter
0%
5%
16%
26%
16%
37%
4. Microsoft
0%
6%
5%
6%
44%
39%
5. Flickr
5%
11%
21%
21%
0%
16%
26%
6. Ryanair (airline)
0%
5%
27%
21%
21%
26%
21%
7. SAS (airline)
0%
5%
5%
11%
16%
21%
42%
8. Netflix
0%
5%
10%
21%
37%
16%
11%
9. WordPress
5%
10%
16%
16%
0%
21%
32%
10. Apple
0%
11%
10%
26%
42%
11%
11. PayPal
0%
16%
37%
16%
31%
12. Evernote
16%
10%
11%
5%
37%
21%
0%
13. Amazon
0%
0%
11%
5%
42%
42%
14. IBM
0%
5%
5%
5%
16%
5%
32%
37%
15. Hewlett Packard (HP)
0%
6%
6%
11%
11%
22%
44%
16. Instagram
0%
11%
21%
0%
26%
21%
21%
17. Spotify
0%
5%
5%
11%
21%
21%
37%
18. LinkedIn
5%
5%
16%
0%
16%
16%
42%
19. Google
0%
5%
11%
16%
42%
26%
20. YouTube
0%
28%
33%
39%
21. WhatsApp
0%
11%
26%
26%
21%
16%
22. Skype
0%
11%
22%
50%
17%
23. Dropbox
11%
5%
10%
5%
11%
16%
42%
24. eBay
0%
16%
11%
21%
26%
26%
25. Pinterest
0%
21%
32%
5%
32%
10%
0%
26. Mastercard
0%
5%
47%
48%
27. Government services (i.e. tax, migration, police, transportation authorities))
0%
16%
42%
42%
Very Low Trust
Low Trust
Medium Trust
High Trust
Very High Trust
Dont know and dont trust
Dont know but would trust

Figure 5 - Rates on trust that the participants have in different services / companies / websites

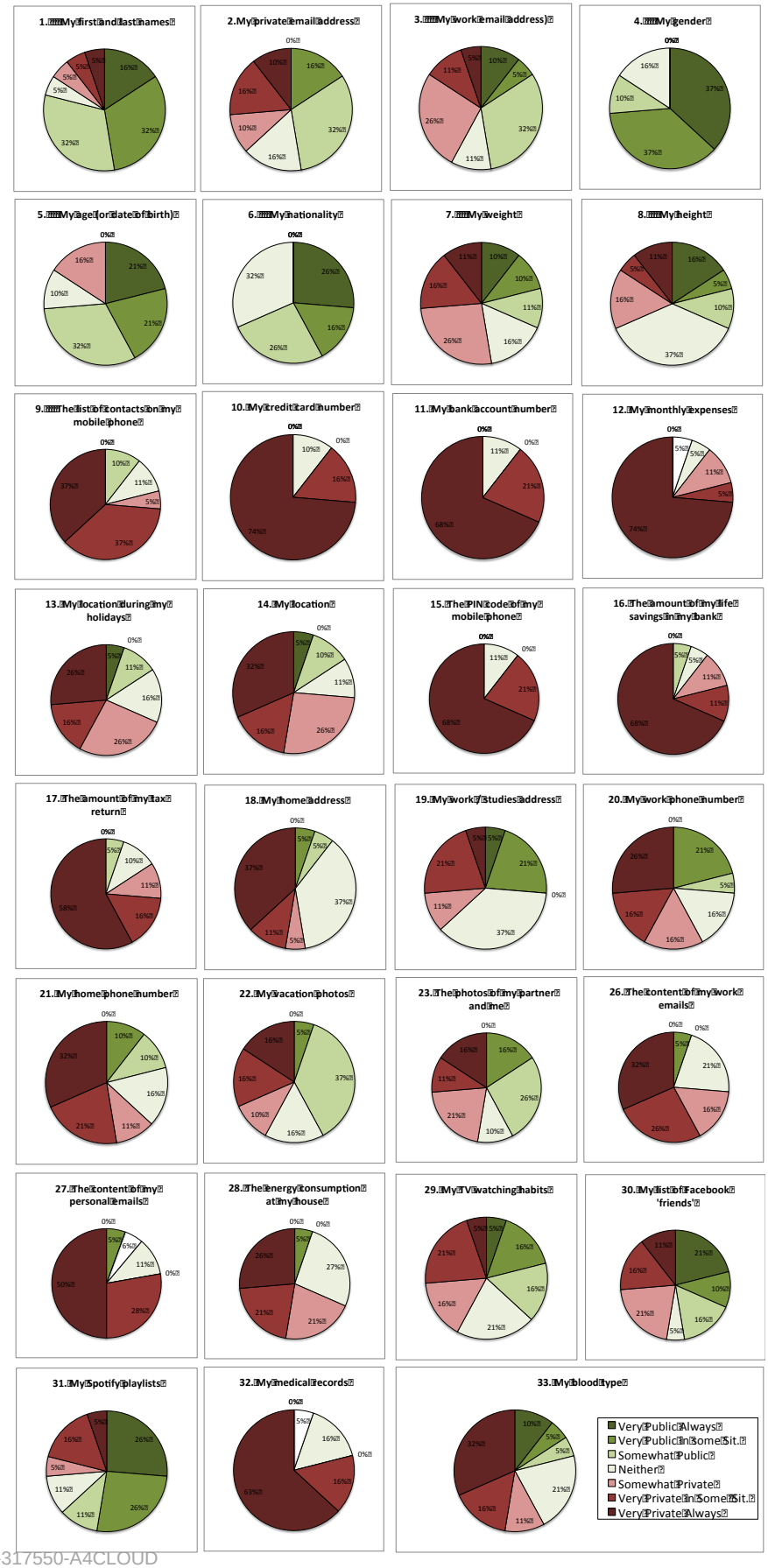
1. My first and last names
2. My private email address
3. My work email address)
4. My gender
5. My age (or date of birth)
6. My nationality
7. My weight
8. My height
9. The list of contacts on my mobile phone
10. My credit card number
11. My bank account number
12. My monthly expenses
13. My location during my holidays
14. My location
15. The PIN code of my mobile phone
16. The amount of my life savings in my bank
17. The amount of my tax return
18. My home address
19. My work/studies address
20. My work phone number
21. My home phone number
22. My vacation photos
23. The photos of my partner and me
26. The content of my work emails
27. The content of my personal emails
28. The energy consumption at my house
29. My TV watching habits
30. My list of Facebook 'friends'
31. My Spotify playlists
32. My medical records
33. My blood type
Very Public Always
Very Public In Some Sit.
Somewhat Public
Neither
Somewhat Private
Very Private In Some Sit.
Very Private Always

Figure 6 - Rates how sensitive personal data items are perceived.

3.1.2 Trondheim Workshop

The goal of this workshop was to present the Data Track tool (Appendix A) and collect the participant's perceptions on the tool and refine the requirements that are stated for the tool. All the material used in the workshop can be found in Appendix C.

3.1.2.1 Selection of stakeholders

Invitation was sent to employees at the administration in SINTEF and also students at the master degree in biology at NTNU (Norwegian University of Science and Technology) to participate in the workshop on the use and opinions about Internet services and possible solutions on how to keep track of their personal data. Participation was voluntary, but participants received two cinema tickets as compensation. The workshop took place on the 23rd of May from 9:00 to 10:30 at SINTEF. Eleven people answered the questionnaire and attended the workshop.

3.1.2.2 Data Collection and Analysis

As described in appendix C, the data collection comprised a pre-questionnaire, recordings of discussions and a post-questionnaire on the perceptions of the tool.

The goal of the pre-questionnaire was to understand the participants' behaviour on the cloud and also to rate their trust on some services and also rate how "sensitive" (private) the various pieces of information are for the participants. This pre-questionnaire also contained a consent form for the researchers to use the data collected for research purposes. Informed consent was collected according to guidelines from computer science research.

During the session we presented the project and the tool. After this presentation, we divided the group in two groups and asked them to discuss the following questions (note that they are the same as in the Karlstad workshop) for 20 minutes:

1. Do you think you will use this tool?
2. In which context do you think this tool will be interesting?
3. What did you like the most about this tool?
4. What did you like the least about this tool?
5. What would you like this tool to notify you about?

The discussions were recorded and transcribed by the two facilitators of the sessions (see transcriptions in Appendix C).

After the discussions the final post-questionnaire was handed in (see Appendix C). The post-questionnaire comprised four main sections. The first three sections asked the participants to assess their agreements with statements on:

iv) Usability questions about the tool (for example: the tool is easy of use, the tool is easy to learn quickly)
v) How well the tool will help to accomplish the goals of the project (for example: the usage of the tool will enable cloud service providers to give their users appropriate control and transparency over how their data is used).
vi) Functional requirements of the tool (for example: The Tool should let me know which data is collected by which services and for what purposes are they going to use my data.)

The last section asked the participants to freely write other comments (extra requirements, improvements, suggestions, recommendations, justification of their answers) about the tool. All data from the questionnaires were tabulated and analysed quantitatively; all the qualitative data from the session was also analysed and incorporated in the results. All the details of the answers are shown in Appendix C, and the main results are shown in the next section.

3.1.2.3 Results

The participants at the workshop organized by SINTEF were in general older than the ones from Karlstad, three were under 30. Figure 7 shows how people rate their trust in services, from very low trust (red) to very high trust (green). In general, people demonstrated trust in large scale in banks, credit cards, Paypal and government services. In general the participants do not trust the other services. Figure 8 shows how sensitive different types of data are considered to be for the participants in the workshop, from very public (green) to very private (red). Data, such as name, gender, age, and nationality are seen as "public" data. They were also more public about their home and work address than the students from Karlstad. Other than these, all data are sensitive somehow to people in different levels of sensitiveness. More details on the answers are provided in Appendix C. It is important to note that these ratings can be used for prioritizing which information should be made secure. In addition this information can be used to start creating groups of information in categories, as pointed by the participants as one nice to have feature in the tools.

On the question of how much the participants agree with the level of usability of the tool. The participants agree that:

1. The tool will probably be used frequently in the future.
2. The tool is not complex.
3. The tool is easy to use.
4. The tool is easy to learn quickly.
5. The tool's functionalities are clear and understandable
6. The tool integrates well with current practices.
7. The tool integrates well various functionalities.

The participants had different views to some other questions on usability:

8. Some believe that the tool requires user training and some said neither, but there was a big variation on the opinions.
9. The participants in general average believe that the tool does not require users to learn a lot of new concepts. But there was a big variation on the results.
10. The participants showed to neither agree nor disagree with the affirmation that: the tool lacks many useful functionalities.

On the question on whether the objectives of A4Cloud are accomplished by the Data Track tool, the participants agree that the Data Track tool:

11. gives users more control and transparency over how the data is used in the cloud
12. enables cloud service providers to give their users appropriate control and transparency over how their data is used.

In general, participants agreed, with some cases of neutral, that the Data Track tool:

13. enables users to make choices about how cloud service providers may use and will protect data in the cloud.
14. will make the relationship between users and providers substantially easier because it will be easier to see who is responsible for the problems.

The participants were neutral or disagreed to the following statements on the Data Track tool:

15. will substantially increase users trust in cloud services.
16. will substantially reduce the number of serious security problems.

On the question on how the participants rated the importance of the listed functional requirements of the tool, all functionalities were rated as important, without much variation in the opinions. The only one that was not averaged as important was requirement 30, with a great variance on the answers. The requirement 29 was also a bit controversial, so it is not all that thinks this is important, but still it gets a high score. The list of requirements is as follows:

17. The Tool should let me see which of my data was sent by me and which data was collected automatically by the service

18. The Tool should let me know which data is collected by which services and for what purposes are they going to use my data.
19. The Tool should inform me whether the Cloud service stores my data in an encrypted form or if they can see the contents of my data.
20. The Tool should let me check how my data has been processed by the Cloud service and what conclusions they can draw about me based on this.
21. The Tool should allow me to find out whether my data has been used in a way that was not specified in the privacy policy when I sent my data.
22. The Tool should allow me to see how my data has been passed on to other Cloud services.
23. The Tool should allow me to correct (edit) or to delete the data that the Cloud services have about me.
24. The Tool should let me know the country in which my data is stored and the laws that apply to that country.
25. The tool should inform me when other people send data about me to a Cloud service.
26. The Tool should provide a report that tells me how risky or secure is to have my data in a Cloud service.
27. The Tool should inform me about the risks and threats associated with Cloud services.
28. The Tool should enable assessing the security level of the service providers.
29. When the Cloud service doesn't do what they promised me at the time of registration, the Tool should let me do something about it.
30. The Tool should send me many notifications per week to inform me how the Cloud service is handling my data.

Table 2 shows the list of additional requirements elicited for the data track tool, divided by the session where it was elicited.

Table 2 - List of Elicited Requirements Data Track tool – Cloud Subjects - Trondheim

List of Elicited Requirements	
From Discussions Group 1 (Daniela):	# In case of deletion of information, the tool could show if the information is already completely deleted from servers or not or just unavailable, but still stored on the servers. # The tool has to have a maximum security level, because it is very dangerous to have all data in one specific tool. # The tool should give a warning on when privacy rules change on the websites. And which data is affected by the change of the rules. # The tool could have an option of not only seeing which information is everywhere, but also if I want to change an information, then change in all services (for example, I changed my home address and now I want to update this information in all sites I have this information). # The tool could have an option of SHARE information with other sites or other people. And track who I shared this information. # The tool could provide a "save" button in which you could save some information and have it locally. # The tool should be tested on elderly people.
From Discussions Group 2 (Martin):	# The tool should be tested in different contexts of information (for example: health data, personal financial data, personal data) # The tool has to be a locally installed tool, not a cloud or internet service. #In case of deletion of information, the tool could show if the information is already completely deleted from servers or not or just unavailable, but still stored on the servers. # The tool should show explicitly which country the information is stored or can be stored. # The tool should show explicitly which information access has the country government of where the information is stored. #The tool should show explicitly which information third parties companies has access to.

Requirements From Questionnaires	# The tool should be usable by people that are not used to PCs. # The tool should be usable by older people. # The tool should be usable by young people. # The tool should inform about privacy changes or any other changes on the provider.

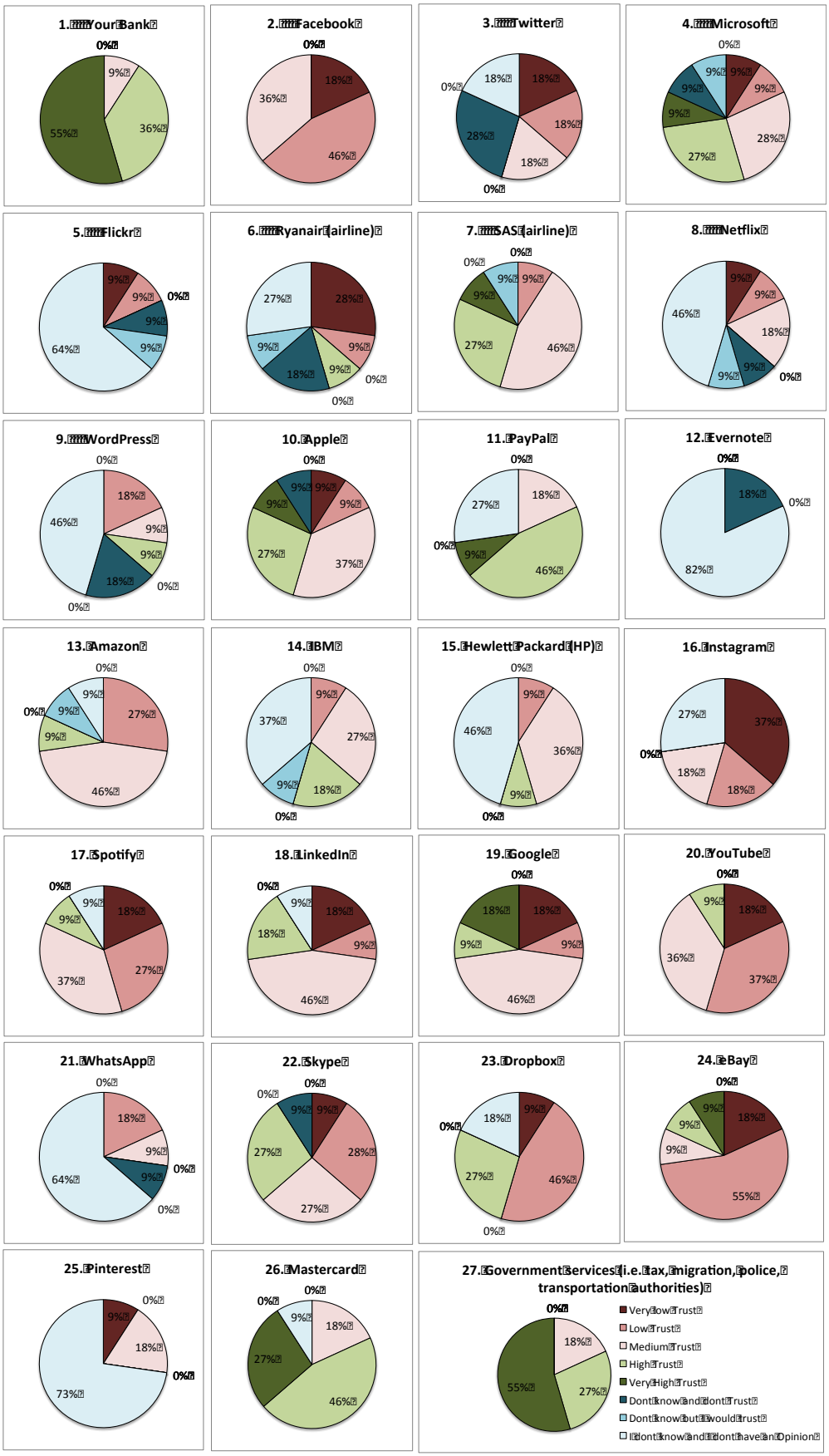
1. Your Bank
2. Facebook
3. Twitter
4. Microsoft
5. Flickr
6. Ryanair (airline)
7. SAS (airline)
8. Netflix
9. WordPress
10. Apple
11. PayPal
12. Evernote
13. Amazon
14. IBM
15. Hewlett Packard (HP)
16. Instagram
17. Spotify
18. LinkedIn
19. Google
20. YouTube
21. WhatsApp
22. Skype
23. Dropbox
24. eBay
25. Pinterest
26. Mastercard
27. Government services (i.e. tax, migration, police, transportation authorities)
Very Low Trust
Low Trust
Medium Trust
High Trust
Very High Trust
Dont know and dont Trust
Dont know but would Trust
I dont know and I dont have an Opinion

Figure 7 - Rates on trust on different services / companies / websites

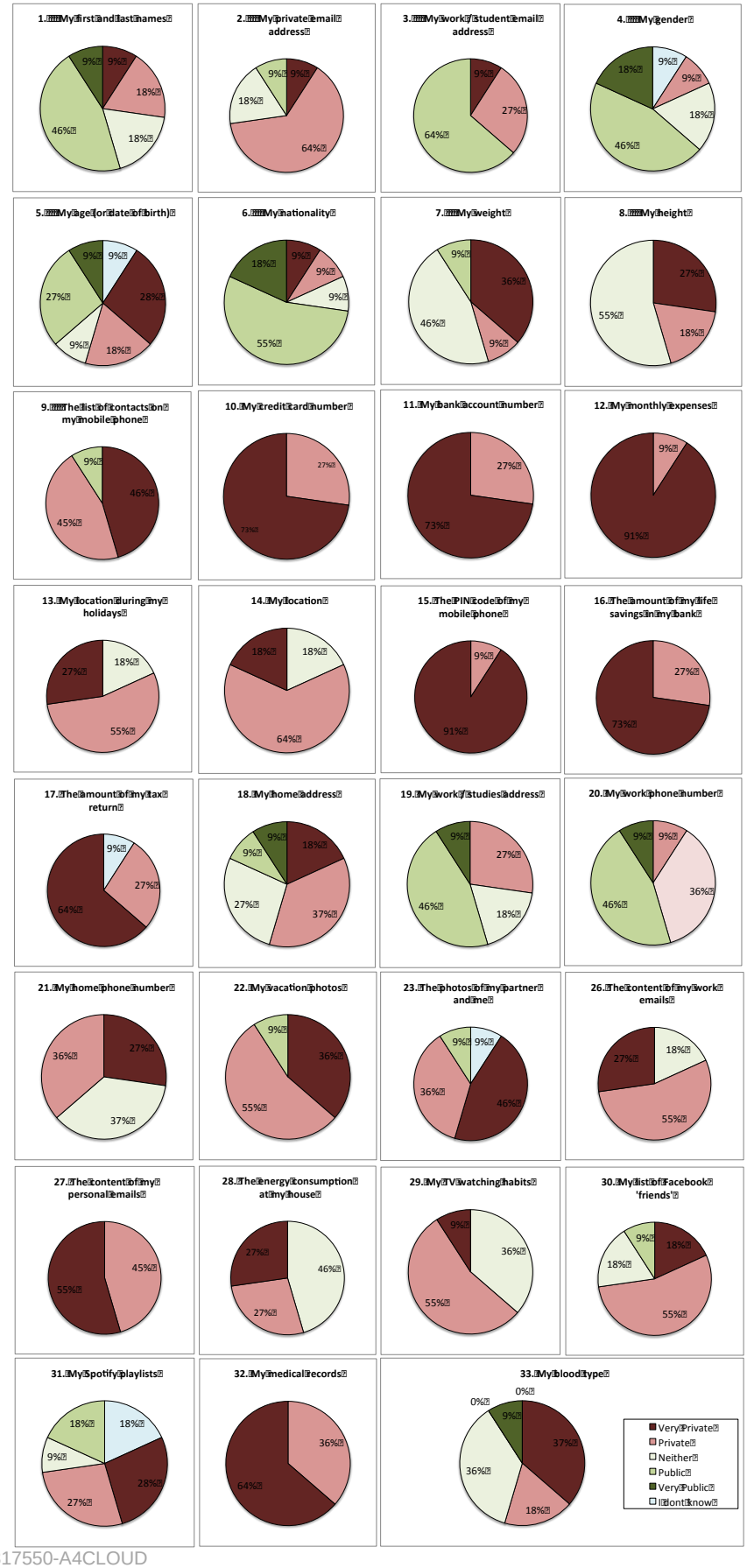
1. My first and last names
9% 9% 18% 18% 46%
2. My private email address
9% 9% 18% 64%
3. My work/student email address
9% 27% 64%
4. My gender
9% 18% 9% 18% 46%
5. My age (or date of birth)
9% 9% 27% 28% 9% 18%
6. My nationality
9% 18% 9% 9% 55%
7. My weight
9% 36% 46% 9%
8. My height
27% 55% 18%
9. The list of contacts on my mobile phone
9% 46% 45%
10. My credit card number
27% 73%
11. My bank account number
27% 73%
12. My monthly expenses
9% 91%
13. My location during my holidays
27% 18% 55%
14. My location
18% 18% 64%
15. The PIN code of my mobile phone
9% 91%
16. The amount of my life savings in my bank
27% 73%
17. The amount of my tax return
9% 27% 64%
18. My home address
9% 9% 18% 27% 37%
19. My work/studies address
9% 27% 46% 18%
20. My work phone number
9% 9% 36% 46%
21. My home phone number
27% 36% 37%
22. My vacation photos
9% 36% 55%
23. The photos of my partner and me
9% 9% 36% 46%
26. The content of my work emails
18% 27% 55%
27. The content of my personal emails
45% 55%
28. The energy consumption at my house
27% 46% 27%
29. My TV watching habits
9% 36% 55%
30. My list of Facebook 'friends'
9% 18% 18% 55%
31. My Spotify playlists
18% 18% 9% 28% 27%
32. My medical records
36% 64%
33. My blood type
0% 0% 9% 37% 36% 18%
Very Private
Private
Neither
Public
Very Public
I don't know

Figure 8 - Rates on sensitiveness (private) the pieces of information.

3.1.3 Concluding Remarks for Cloud Subjects

In general the Cloud Subjects were very receptive to the concepts exposed in the workshops. They seem to have gotten a good understanding of what the project A4Cloud is about and what are the goals we would like to achieve. In this sense, these workshops showed to be a very good strategy for that. It was clear during the workshops that the data subjects are not completely aware of what "being on the cloud" means, for them is just another "web service" or "online". Therefore they were also happy to understand more about the "cloud as a service".

From the discussions on the sessions, we noticed that there is not much hope that there will be remediation in case of data breaches. They seem to be longing for tools such as the ones proposed in A4Cloud for feeling more "secure" on using online services. They were all glad to know that tools such as Data Track will be available soon.

We also notice, that when exposing the data subjects to the concepts of the A4Cloud, one should be aware of the age of the data subjects. It was clear to us that for example, for young data subjects, examples such as health care use cases, do not convey the message or help them to relate to the concepts. On the other hand, the use case works very well with older data subjects, from 40s for example. Future work can be to have a session on teenagers, and the comments from parents are that they are completely unaware of any security issues, therefore, it is probable that there would be a need for yet another type of "use case" to convey a message to them.

3.2 Workshop 3.2: Cloud Customers

Cloud customers were involved in the refinement and elicitation of requirements in two different types of elicitation workshops. We used the COAT tool as a way to expose them to the A4Cloud concepts and tools as well as the transparency requirements that were elicited in WS1 and WS2. The reason for choosing COAT was that the customers could better understand the concepts behind the tool once that it is more closely related to their cloud related activities in their organizations. Section 3.2.1 presents the results from the workshop we ran in Trondheim with 11 participants and section 3.1.2 presents the results from the transparency requirements interviews that we performed through Skype in Trondheim with 8 participants.

3.2.1 COAT Trondheim Workshop

The goal of this workshop was to present the project, the transparency interview results and the COAT tool (Appendix B) and collect the participant's perceptions on the tool. All the material used in the workshop can be found in Appendix C.

3.2.1.1 Selection of stakeholders

Invitation was sent to our list of contacts in software companies in Trondheim. Participation was voluntary, but at the end of the workshop we offered the participants a lunch. The workshop took place on the 20th of June from 10:00 to 12:00 at SINTEF. Eleven people attended the workshop.

3.2.1.2 Data Collection and Analysis

As described in Appendix C, the data collection comprised recordings of discussions and a post-questionnaire on the perceptions of the tool. Informed consent was collected according to guidelines from computer science research. During the session we presented the project and the tool. After this presentation, we asked them to discuss the following questions for 20 minutes:

1. Do you have any feedback on the concept of the tool?
2. Do you have any feedback on the implementation of the tool?
3. Do you have any suggestion on what else the tool can cover?

The discussions were recorded and transcribed by the two facilitators of the sessions (see transcriptions in Appendix C).

After the discussions the final post-questionnaire was handed in (see Appendix C). The post-questionnaire was comprised of three main sections. The first two sections asked the participants to assess their agreement with statements on:

i) Usability questions about the tool (for example: the tool is easy of use, the tool is easy to learn quickly)
ii) How well the tool will help to accomplish the goals of the project (for example: the usage of the tool will enable cloud service providers to give their users appropriate control and transparency over how their data is used).

The last section asked the participants to freely write other comments (extra requirements, improvements, suggestions, recommendations, justification of their answers) about the tool. All data from the questionnaires were tabulated and analysed quantitatively; all the qualitative data from the session was analysed and incorporated in the results. All the details of the answers are shown in Appendix C and the main results are shown in the next section.

3.2.1.3 Results

On the question of how much the participants agree with the level of usability of the COAT tool. The participants agree that:

1. The tool is not complex.
2. The tool is easy to use.
3. The tool is easy to learn quickly.
4. The tool's functionalities are clear and understandable
5. The participants in general average believe that the tool does require users to learn a lot of new concepts.

The participants had different views to some other questions on usability:

6. Only four participants believe that the tool will probably be used frequently in the future.
7. Some believe that the tool requires user training and some said neither, but there was a big variation on the opinions.
8. The participants showed to neither agree nor disagree with the affirmation that: The tool integrates well with current practices and the tool integrates well various functionalities.
9. The participants showed to neither agree nor disagree with the affirmation that: the tool lacks many useful functionalities.

On the question on how participants agree that the objectives of A4Cloud are accomplished by the COAT tool. Participants agree that the COAT tool:

1. enable cloud service providers to give their users appropriate control and transparency over how their data is used.
2. enable users to make choices about how cloud service providers may use and will protect data in the cloud.

The participants neither agreed nor disagreed the following statements on the COAT tool:

3. will substantially increase users trust in cloud services.
   - Most are between neutral and agree, tending more to agree.
4. will give users more control and transparency over how the data is used in the cloud
   - Most are between neutral and agree, tending more to agree.
5. will substantially reduce the number of serious security problems.
   - Most are neither. Three people said disagree.
6. will make it substantially easier the relationship between users and providers because it will be easier to see who is responsible for the problems.
   - Most are between neutral and agree. Basically half and half.

Table 3 shows the list of additional requirements elicited for the COAT tool, divided by the session where it was elicited.

Table 3 - List of Elicited Requirements COAT tool – Cloud Customers - Trondheim

List of Elicited Requirements	
From Discussions	# The tool should provide a validation of seriousness, economy and so on, of the providers. # The tool needs to be independent and with no hidden criteria for showing one provider on top of the list, other than the explicit ones. # The tool should allow the user to select criteria of what he/she is going to buy, without asking for an actual offer. # There should be a verification process behind the process for adding a provider to the list of providers of the tool. One possibility is some crowd sourcing in that people who trust a certain provider can add response in some way. # There should be a process of updating prices of the providers, to assure that the prices showed on the tool for each provider are the actual prices. # The architecture for how to find the criteria you seek, needs to be simple. Everyone should be able to add, recommend criteria, but it needs to be based on a sound architecture. # The criteria should be created based on what the users are asking for an not just on what is offered by the providers. # You could also, instead of a having a library view of it all, use the concept of tags. You have some choices, but afterwards you use some tags that are added, that converge on something that is a trend of needs that people ask for. And perhaps you can collect them in sets because they use different terms for the same concept. Later on, take all free text tags, and turn them into choices. Based on the data, this is what comes up, this is what users believe is most relevant. And then add, many people add requirements, possibly in a tag format, and then you will start generating new tags that can trend high up. # The tool should have an option to get ratings to the providers based on customers opinions. # The tool should provide a history of the provider, incidents, how they were solved etc.
Requirements From Questionnaires	# The tool should consider needs of large corporations and organizations # The tool needs to address changes in laws and regulations and best practice quickly it will need to be dynamic in nature instead of a static library of option and "correct answers" # The tool should show the history of the provider (incidents, etc) # The tool should make data available as a data service

3.2.2 Transparency Requirements Interviews

Transparency is the property of an accountable system that it is capable of **'giving account**' of, or **providing visibility** of, how it conforms to its **governing rules and commitments**. Transparency involves operating in such a way as to maximize the amount of and **ease-of-access to information** which may be obtained about the structure and behaviour of a system or process. An accountable organization is transparent in the sense that it makes the policies defined about treatment of personal and confidential data known to relevant stakeholders, can demonstrate how these are implemented, provides appropriate notifications in case of policy violation, and responds adequately to data subject access requests.

The goal of this workshop was to present the project, then expose the customers to the transparency requirements we elicited in previous workshops in the project, refine them, and elicit more details on requirements for transparency. All the material used in the interviews can be found in Appendix C.

3.2.2.1 Selection of stakeholders

Invitations were sent to our list of contacts in software companies in Trondheim. Participation was voluntary. Eight people accepted to participate on the interviews. The participants were all IT security experts working with cloud related projects. The participants represented six different organizations.

3.2.2.2 Data Collection and Analysis

All the material used in the interviews can be found in Appendix C. The interviews were performed on SKYPE and lasted about one hour. The main questions on the interview were:

1. What is the most important information you think should be provided to the cloud customer when buying services from cloud service providers?
2. In which parts would you like to be involved in making the decisions? In which parts would you like just to be *informed* of the decisions?
3. What would increase your trust that the data is secure in this scenario?
4. What do you want to know about how the provider corrects data security problems?
5. From the following list of requirements of transparency that have been elicited in the project (Table 4):
    a. What is your opinion about them?
    b. Which extra information should be added?

Table 4 - List of Transparency Requirements for the Interview Guide

#	Requirement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	The cloud provider is responsible to the cloud consumer for the provision of evidence of data segregation.					
2	The cloud provider is responsible to the cloud auditors, Regulators and Data Protection Authorities (DPAs) for the provision of evidence of compliance of data segregation with respect to legislative regimes.					
3	The cloud provider is responsible to the cloud consumer for the implementation of different policies tailored to the nature of data, privacy laws and needs of the cloud consumer.					
4	The cloud provider is responsible to the cloud consumer that data are used for the intended purposes.					
5	The cloud provider is responsible to the cloud consumer for the provision of rights management on data.					
6	The cloud provider is responsible to the cloud consumer for asking the explicit consent for any operation on data.					
7	The cloud provider is responsible to the cloud consumer for revoking data consent if requested.					
8	The cloud provider is responsible to the cloud consumer for asking the explicit consent every time any operation is performed on data.					
9	The cloud provider is responsible to the cloud consumer for the provision of data classification mechanisms supporting different data security levels (e.g. confidential or non-confidential).					
10	The cloud provider is responsible to the cloud consumer for the provision of custom-made data security levels.					
11	The cloud provider is responsible to the cloud consumer for the provision of the highest data security level as default.					
12	The cloud provider is responsible to the cloud consumer for allowing the use of data encryption.					
13	The cloud broker is responsible to the cloud consumer for the provision of evidence of non-data aggregation (or effective data segregation).					
14	The cloud provider is responsible to the cloud consumer for the provision of evidence of data collection practices.					
15	The cloud provider is responsible to the cloud consumer for the provision of evidence of data gathered, inferred or aggregated.					

The eight interviews for this study were transcribed into text documents based on the audio recordings. For further analysis of the transcription, recommended steps proposed by Cruzes and Dybå [6] were followed. Five steps were performed (as described in Figure 9): initial reading of data/text (extraction), identification of specific segments of text, labelling of segments of text (coding), translation of codes into themes, creation of the model and assessment of the trustworthiness of the model. Thematic synthesis is a method for identifying, analysing, and reporting patterns (themes) within data. It is one of the most common methods for synthesis of evidence in SE [6]. Thematic synthesis resembles some of the characteristics of grounded theory analysis, in that the themes emerge from (are grounded in) the primary data. It minimally organizes and describes the data set in rich detail and frequently interprets various aspects of the research topic. It comprises the identification of the main, recurrent or most important (based on the specific question being answered or the theoretical position of the reviewer) issues or themes arising from a body of evidence. The level of sophistication achieved by this method can vary; ranging from simple description of all the themes identified, through to analyses of how the different themes relate to one another in a conceptual map. The advantage of thematic synthesis is that it provides a means of organizing and combining the findings from a large, diverse body of research. It can handle qualitative and quantitative findings, and it can be a deductive, theoretically driven approach or an inductive one, in which themes 'emerge' from the process of synthesis.

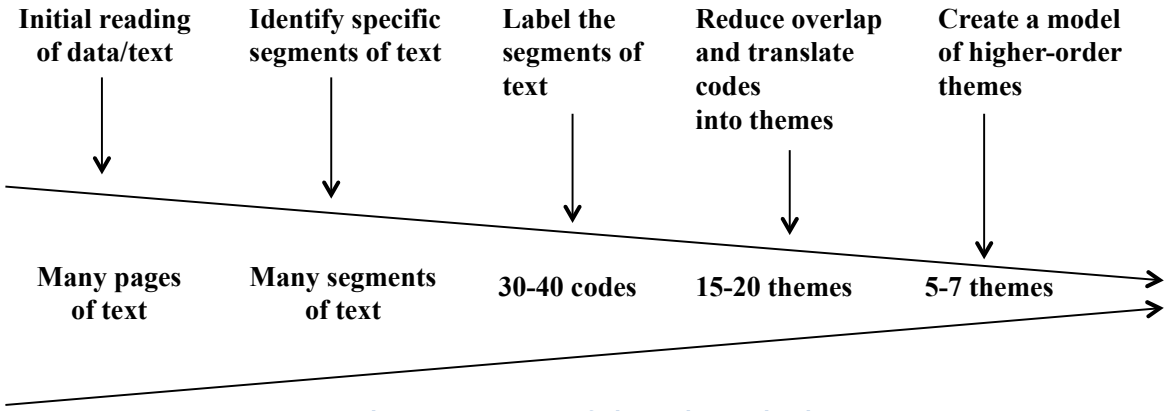

Figure 9 - Process of Thematic Synthesis

3.2.2.3 Results

For the question "What is the most important information you think should be provided to the cloud customer in this scenario?" the participants talked mostly about nine themes: clear statements of what is possible to do with the data, conformance to data agreements, information on how the provider handles data, location, who else other than the provider is participant of the value chain, multi-tenant situations, what the provider does with the data, procedures to leave the service and assurance that the user still owns the right to the data. As shown in Figure 10, the colour represents the different respondents, and their opinions to the question.

One respondent commented that even though he would like to have clear statements of what is possible to do with the data: "100 pages document could be written about this but for some non-technical people it would not help at all". Another one said: "I would like to have a page that they could tell me about security mechanisms" (e.g. firewalls, backup, etc.)

On the conformance to data agreements, the respondents agree that having Data Agreements helps. But it is mainly for technicians not for non-technical people. On how the provider handles data the respondents said that they would like to have functional, technical and security wise information about how the providers handle the data. On location, it is about geographically where the data is stored and what the legal location of the services is. Another important information is about sub providers if there are any. Where they are located and whether they meet legal requirements of the customers' location. Multi-tenant situations are a concern of the customers and they would like to have this information

transparent. And also, information on how the providers assure that data from one customer will not be accessed by another customer.

It is also important for transparency to know what the provider does to protect customers' data. One respondent said that he would like to have information on: "How to protect the information or how the information is protected not much in detail for the end-user, but only for enterprises."

It was also highlighted by the respondents that they would like to have transparent the procedures to leave the service and on how to move data from one service to another. Besides they would like to have the assurance that they still own the right to their data.

On the question "What would increase your trust that the data is secure in this scenario?" the participants mentioned eight different themes: upfront transparency, community discussions, costumer awareness, way out, reputation, encryption, data processor agreements and location. Some were overlapping towards the answers from the first question: upfront transparency, location and conformance to data processor agreement. Interesting answers for this question were answers related to community discussions, customer awareness and reputation. The respondents said that it increases their trust to a cloud provider if they know that the provider has an active security research team, or participates in security communities. The respondents also said that for security: "Customers should be proactive and make sure that all the documentation is there". And another one commented on the importance of having webpages telling what customers could do to keep the data safe. Two participants mentioned also "Way out" meaning that they would like to have webpages telling what to do to remove the data from the service provider.

On the question: In which parts would you like to be involved in making the decisions? In which parts would you like just to be informed of the decisions? It was surprising that the participants mostly answered that they would like to be informed but not really taking part of every decision (See Figure 11), the exceptions were when the provider was moving data to another country, other parties will be involved in the value chain or there are significant changes in the initial terms of contract. One participant said: "Some customers sometimes have some requests. But in general they do not care about taking part of the decisions" and another ones said: "there are some decisions that we don't need to explicitly know about it, but it has to be regulated by some other agreement about the responsibility of each one towards the data". On moving data to another country, one respondent said: "I would like to be involved in decisions on moving my data to another country in most situations. Unless for example a disaster and there is the need to move to another country." Some respondents said that they would like to be informed when the data is transferred from one actor to the next, one of them added: "For example if calling to the call centre your data will be transferred to another country then the customers has to be involved in the decision about that. So he can take an informed decision." On changes in the initial terms of Contract, one respondent said: the providers should be very aware of what they changed since the contract with the customer. And inform them about the changes that happen. Never leave the customer in the dark."

When asked on what would they want to know about how the provider corrects data security problems. It was surprising to perceive that the participants have not thought much on what they could expect from the providers if some security issue happens. Most needed that we elaborated more on the question so they would start saying something. And then it was possible to get to the taxonomy as shown in Figure 12, in which the participants stated that they would like to know before something happens, what is planned, when something happens, how the providers are handling the situation and also be informed of the reasons why the problem happened and it was very important to know when will the services be back. Interesting was also the fact that the participants wanted to know how the providers are improving their services after something happens, based on lessons learned.

After analysing all the collected information we compiled a list of requirements elicited in the interviews, as shown in Table 5. The main "topics" mentioned by the respondents were related to what is possible to do with the data, conformance to data agreements, data handling, value chain, multi-tenant situations, and protection of the data, decisions and corrections of the data.

Table 5 - List of Requirements from Transparency interviews

List of Elicited Requirements	
What is possible to do with the data	# The provider should inform the cloud customer with clear statements of what is possible to do with the data # The provider should allow the cloud customer to choose what is possible to do with his/data data # The provider should have a page to inform the cloud customer about security mechanisms, for example, firewalls, backup etc. # The provider should have some kind of standard certification level of description or standard language that they have to make the situation easier to the buyer to evaluate which security level does he need, what is required from the buyer and what is the provider offering. # The provider should have a document explaining what are the procedures to leave the service and take the data out of their servers. # The provider should have a document in which they describe the ownership of the data.
Conformance to Data Agreements	# The provider should make available the technical documentation on how data is handled, how it is stored, the procedures. And having this documentation available it helps. # There should be documentation of procedures in different levels of abstraction, for example for technical staff or for cloud subjects # The provider should show that they follow the data handling agreement to the type of data that is in question. # The provider should provide geographical information of where the data is stored.
Data Handling	# The provider should provide functional, technical and security wise information about how they handle the data. # The provider should provide very good information of how the data is stored and who has access to it.
Value chain	# In case of using services from other parties, the provider should inform cloud customers on what are the responsibilities of the parts involved in the agreement. # In case of using services from other parties, the provider should inform about the existence of sub providers, where they are located and whether they meet legal requirements of the country of the cloud customer.
Multi Tenant Services	# The provider should inform the cloud customers on cases of multi-tenant services. # In case of multi-tenant services, the provider should inform how the customers are separated from each other. # In case of multi-tenant services, the provider should inform how they assure that data from one customer will not be accessed by another customer.
Protection of the data	# The provider should inform the cloud customer on how to protect the information or how the information is protected not much in detail for the end-user, but only for enterprises. # The provider should have a document describing the mechanisms that secure data not only for data loss but also for data privacy vulnerabilities.
Decisions and Information	# The cloud providers should get the consent of the cloud customer before moving the data to another country, in cases where new parties will be involved in the value chain and on changes on the initial terms of contract.
Correction of the data	# The cloud provider should have a document stating what are the procedures and mechanisms planned for cases of security breaches on customers data. # In case of security breaches, the cloud provider should inform the cloud customers on what happened, why did it happen, what are the procedures they are taking to correct the problem and when will services be normalized.

Figure 10 - Important Upfront Information for Transparent Services.

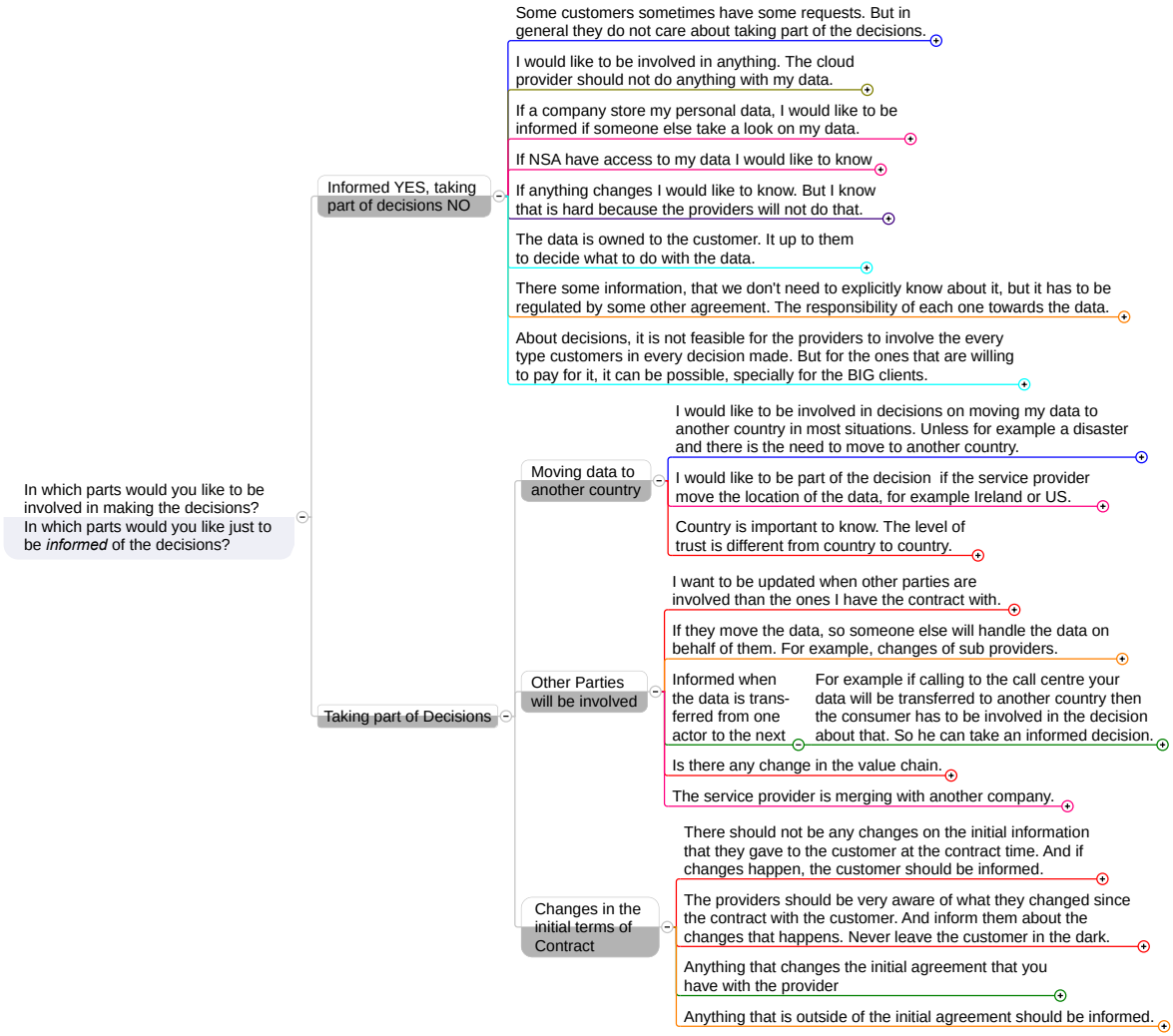

Figure 11 - Involvement on making Decisions

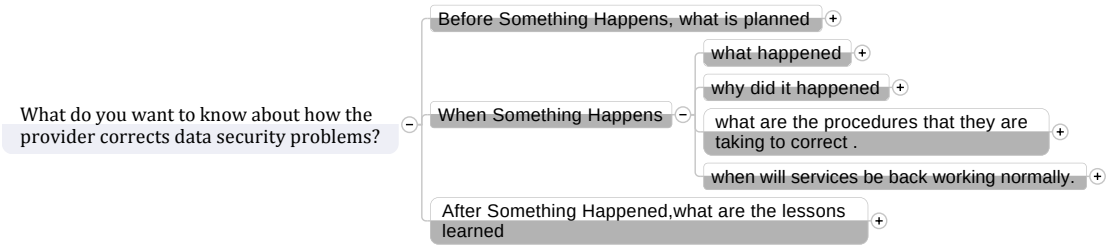

Figure 12 - Transparency on Correction of Data Security Problems.

Finally we asked the participants to rate the table of previously elicited requirements, as shown in (Table 4). We also asked which extra information should be added to the requirements. As shown in Table 6 - Agreement on the List of Transparency Requirements for the Interview Guide, the colours highlights the tendencies to accept the requirement (green) and red colours the tendencies to disagree with the requirements. As it can be seen in the table, the participants have in general positive reaction to requirements 1, 2, 4, 5, 6, 7, 13, 14, 15.

Table 6 - Agreement on the List of Transparency Requirements for the Interview Guide

#	Requirement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	The cloud provider is responsible to the cloud consumer for the provision of evidence of data segregation.	0	1	0	4	1
2	The cloud provider is responsible to the cloud auditors, Regulators and Data Protection Authorities (DPAs) for the provision of evidence of compliance of data segregation with respect to legislative regimes.	0	1	0	2	3
3	The cloud provider is responsible to the cloud consumer for the implementation of different policies tailored to the nature of data, privacy laws and needs of the cloud consumer.	0	1	2	3	1
4	The cloud provider is responsible to the cloud consumer that data are used for the intended purposes.	0	1	0	3	2
5	The cloud provider is responsible to the cloud consumer for the provision of rights management on data.	0	1	0	2	3
6	The cloud provider is responsible to the cloud consumer for asking the explicit consent for any operation on data.	0	1	1	4	0
7	The cloud provider is responsible to the cloud consumer for revoking data consent if requested.	0	1	0	3	2
8	The cloud provider is responsible to the cloud consumer for asking the explicit consent every time any operation is performed on data.	0	3	0	3	0
9	The cloud provider is responsible to the cloud consumer for the provision of data classification mechanisms supporting different data security levels (e.g. confidential or non-confidential).	0	1	1	2	2
10	The cloud provider is responsible to the cloud consumer for the provision of custom-made data security levels.	0	4	0	2	0
11	The cloud provider is responsible to the cloud consumer for the provision of the highest data security level as default.	0	3	1	1	1
12	The cloud provider is responsible to the cloud consumer for allowing the use of data encryption.	0	0	1	3	1
13	The cloud broker is responsible to the cloud consumer for the provision of evidence of non-data aggregation (or effective data segregation).	0	0	0	3	2
14	The cloud provider is responsible to the cloud consumer for the provision of evidence of data collection practices.	0	1	0	4	1
15	The cloud provider is responsible to the cloud consumer for the provision of evidence of data gathered, inferred or aggregated.	0	1	0	3	2

On requirements 3, 9 and 12, the participants also had in general a positive reaction but with some neutrals that make them weak agreements. On requirement 3 (The cloud provider is responsible to the cloud consumer for the implementation of different policies tailored to the nature of data, privacy laws and needs of the cloud consumer), a participant said that the reason for neutral is that he would think that it would be yes for some systems and no for others, and he was not sure if a customer could come

up with new policies. Another participant said: *"I am hesitant with the word tailored. We cannot go to every customer and tailor specific needs, in general it will be groups of customers. Groups are formed based on size and complexity. But more possible that bigger customers are willing to pay more. For example, we could go from multi-tenant solutions to more private solutions".* On requirement 9 (The cloud provider is responsible to the cloud consumer for the provision of data classification mechanisms supporting different data security levels, e.g. confidential or non-confidential), one participant said that the customer has their share of responsibility, "*It is more a must for the consumer, that they have knowledge about data classification and that they have knowledge about the usage of it. The provider should be able to have solutions that support the classifications*". Another one said that he believes this is hard to implement. On requirement 12 (The cloud provider is responsible to the cloud consumer for allowing the use of data encryption), one participant commented that: *"This requirement should be rewritten, the provider does not allow anything! Just provide the functionality. If the customers uses or not then it is their responsibility. The customer will never complain to the availability of it"*.

On requirement 8 (The cloud provider is responsible to the cloud consumer for asking the explicit consent every time any operation is performed on data), one participant that agreed said: "*Any is too strong. For example, if providers are changing servers from one place to another they don't need to ask*". Another one that agreed said also: "*This is hard to implement, any is too strong, some operation I would agree. Some would be, moving data to another country. Maybe that is the most important thing*". The participants that disagreed added that: "*The customer should be responsible, customers should push the things and not only the providers*", another participant commented that: "in most cases this is not practical, this is very difficult to get it through".

On requirement 10 (The cloud provider is responsible to the cloud consumer for the provision of custom-made data security levels), one of the participants that disagreed said: *"In this case it is dependent on what kind of system is this. If I am a customer and I will use a completely different type of encryption protocol for example, it is unfair that I must require the provider to comply with that. So in most cases it is up to the cloud providers to provide the service of encryption and the customer should learn how to use.".* Another participant said: "*Not sure about custom made! There will be different levels, but providers can't go to each customer when they want to do that*". Finally, one of the participants that agreed said: "I *am not sure about this one, it would be nice, but it is hard to implement*."

On requirement 11 (The cloud provider is responsible to the cloud consumer for the provision of the highest data security level as default). For this requirement one participant said: *"That should be defined by each customer. And this costs! So it depends on if the customer pays for the extra service or not. Now in some services for example the customer wants to set the lower security by default"*.

3.2.3 Concluding Remarks for Cloud Customers

In general, the Cloud Customers were very receptive to the concepts exposed in the workshops. They seem to have gotten a good understanding of what the project A4Cloud is about and what are the goals we would like to achieve. In this sense, these workshops showed to be a very good strategy for that.

The COAT tool did not seem to have a "wow" factor from the cloud customers. There are similar tools in the market and they did not seem to believe that the tool will be largely used in the future.

When asked about transparency, there is a longing for having more explicit criteria to judge service providers and to rank them accordingly. They all seem positive to the results from the interviews and all want to see more on it.

We also notice, that when exposing the cloud customers to the concepts of the A4Cloud, one should be aware of the size of organizations and domain of the business.
Some specific points on the transparency results we would like to highlight are:

- Explicit consent for data operations is seen as overkill by some
- Custom-made security levels are a "nice to have feature", but they understand that it costs and that not all providers will offer that.

- Many don't want highest security as default; this may be a reflection on a "you get what you pay for" attitude, and preferring the cheapest version as default. A question is: Should we still advocate highest security, to force customers to actively downgrade?

All the requirements elicited from the cloud customers and cloud subjects were consolidated in the requirements repository (see Table 7).

Table 7 - Consolidated List of Requirements from Customers and Data Subjects

List of Elicited Requirements	
What is possible to do with the data	# The provider should show clear statements of what is possible to do with the data # The provider should allow the cloud customer to choose what is possible to do with the data # The provider should have a page that they could tell the cloud customer about security mechanisms, for example, firewalls, backup etc. # The provider should have some kind of standard certification level of description or standard language that they have to make the situation easier to the buyer to evaluate which security level do we need, what is required from us and what is the provider offering. # The provider should have a document explaining what are the procedures to leave the service and take the data out of their servers. # The provider should have a document in which they describe the ownership of the data. # Providers should have policies written in a language that is more understandable for users.
Protection of the data	# The provider should inform the cloud customer on how to protect the information or how the information is protected not much in detail for the end-user, but only for enterprises. # The provider should have a document describing the mechanisms that secure data not only for data loss but also for data privacy vulnerabilities.
Correction of the data	# The cloud provider should have a document stating what are the procedures and mechanisms planned for cases of security breaches on customers data. # In case of security breaches, the cloud provider should inform the cloud customers on what happened, why did it happen, what are the procedures they are taking to correct the problem and when will services be normalized.
Data Handling	# The provider should provide functional, technical and security wise information about how they handle the data. # The provider should provide very good information of how the data is stored and who has access to it.
Conformance to Data Agreements	# The provider should make available the technical documentation on how data is handled, how it is stored, and the procedures. And having this documentation available it helps. # There should be documentation of procedures in different levels of abstraction, for example for technical staff or for cloud subjects # The provider should show that they follow the data handling agreement to the type of data that is in question. # The provider should provide geographical information of where the data is stored. # Accountability tools could have a monitoring tool for showing the status of the data.
Value chain	# In case of using services from other parties, the provider should inform cloud customers on what are the responsibilities of the parts involved in the agreement. # In case of using services from other parties, the provider should inform about the existence of sub providers, where they are located and whether they meet legal requirements of the country of the cloud customer.
Multi Tenant Services	# The provider should inform the cloud customers on cases of multi-tenant services. # In case of multi-tenant services, the provider should inform how the customers are separated from each other. # In case of multi-tenant services, the provider should inform how they assure that data from one customer will not be accessed by another customer.
Decisions and Information	# The cloud providers should get the consent of the cloud customer before moving the data to another country, in cases where new parties will be involved in the value chain and on changes on the initial terms of contract.

	# Providers should give enough information to users about changes and the impact of the changes to the user. So the user can be aware of the risks that the change will imply. # Providers should give information about: where is the information, if they have connection with other companies, other countries. # Providers should have a list of all data they collect about users. # Providers should make it available information on: their reputation, personal recommendations, reviews on websites.
Data Track	# Data Track should have a very secure system to protect all the data that is collected about services (for example, data base should be encrypted). # Data track should help users to delete the data they do not want to have spread on the providers (direct link, set of steps, guide, etc). # Data Track should help to find the policies from each provider. # Data Track should help users proactively by warning them about both explicit and implicit data that providers collects about users, so I can act before user submit data. # Data track should allow user to classify data in different levels of sensitiveness and also different classes of data (banking data, health data, personal information etc). # Data track should allow classification and prioritization of data. # The tool could have an option of not only seeing which information is everywhere, but also if I want to change an information, then change in all services (for example, I changed my home address and now I want to update this information in all sites I have this information). # The tool could have an option of SHARE information with other sites or other people. And track who I shared this information. # Data Track tool should have a status of how safe a person is based on the data from providers. # Data Track should have different profiles of data information, so you do not see all at once but the different data from the different profiles. # Data Track should have a guide to how to delete data in different providers or a link to the where the provider explains how to delete data from the provider. # In case of deletion of information, the tool could show if the information is already completely deleted from servers or not or just unavailable, but still stored on the servers. # Data Track should show the level of security of a certain data type that the user is concerned. # Data Track could have a weekly, monthly or annually ""warning"" of what is the safe level of the data. # The tool should give a warning on when privacy rules changes on the websites. And which data is affected by the change of the rules. # The tool should inform about privacy changes or any other changes on the provider. # Data track should help tracking information such as videos and photos. # The tool has to be a locally installed tool, not a cloud or internet service. # The tool should show explicitly which country the information is stored or can be stored. # The tool should show explicitly which information access has the country government of where the information is stored. # The tool should be tested in different contexts of information (for example: health data, personal financial data, personal data) # The tool should be usable by people that are not used to PCs. # The tool should be usable by older people as well as young people.
COAT	# The tool should consider needs of large corporations and organizations # The tool needs to address changes in requirements law and best practice quickly it will need to be dynamic in nature instead of a static library of option and "correct answers"# The tool should provide a validation of seriousness, economy and so on, of the providers. # The tool needs to be independent and with no hidden criteria for showing one provider on top of the list, other than the explicit ones. # The tool should allow the user to select criteria of what he/she is going to buy, without asking for an actual offer. # There should be a verification process behind the process for adding a provider to the list of providers of the tool. One possibility is some crowd sourcing in that people who trust a certain provider can add response in some way. # There should be a process of updating prices of the providers, to assure that the prices showed on the tool for each provider are the actual prices. # The architecture for how to find the criteria you seek, needs to be simple. Everyone should be able to add, recommend criteria, but it needs to be based on a sound architecture. # The criteria should be created based on what the users are asking for and not just on what is offered by the providers.

	# Providers can adjust to new trends, existing providers can then possibly after adjusting get that button about providers that match your needs today. And people who have used a provider can provide ratings, who you can trust companies. Trend analysis is a classical thing to trying to understand the market. # The tool should provide a history of the provider, incidents, how they were solved etc. # The target for such a tool would be ones who seek providers, but are uncertain about what to look for. How do we handle our data etc. So what I miss from this demonstration… tick off that you should store personal data etc… so, what are the regulations concerning this sector or type of information you store. So you can add this to the rest of the development process when you make something. So you are for instance going to store a credit card number because you think it is important to store it locally. What does this mean, which rules and regulations are there concerning this? Or birth number, whatever you can come up with.

3.3 Workshop 3.3: Cloud Providers

This section highlights the main remarks from the first stakeholder meeting of the HP's Cloud28+, Cloud of Clouds, Made in Europe – Secured Locally initiative. The objective of the meeting was to keep the Cloud 28+ community informed. In addition to updates about the initiative, the meeting involved presentations and discussions on a topic at the heart of the European strategy on cloud computing. During the meeting, the Cloud28+ management team gave an update about the initiative and related activities. Researchers from HP Labs (Bristol) who lead the Cloud Accountability Project (A4Cloud) shared some results on accountability, risk and trust, governance and control of corporate and private data processed by cloud based IT services. The meeting was an opportunity to discuss with stakeholders on relevant issues affecting cloud adoption and trust in the cloud. It was also an opportunity for gathering together the Cloud28+ community.

3.3.1 Workshop Stakeholders

The Cloud Providers' workshop was attended by 41 participants representing 22 different stakeholders among service providers, independent software/system vendors, university/research and public administration/government. The stakeholders attending the workshop were mostly, service providers but also with participants from other relevant groups. Figure 13 shows the stakeholders' distribution.

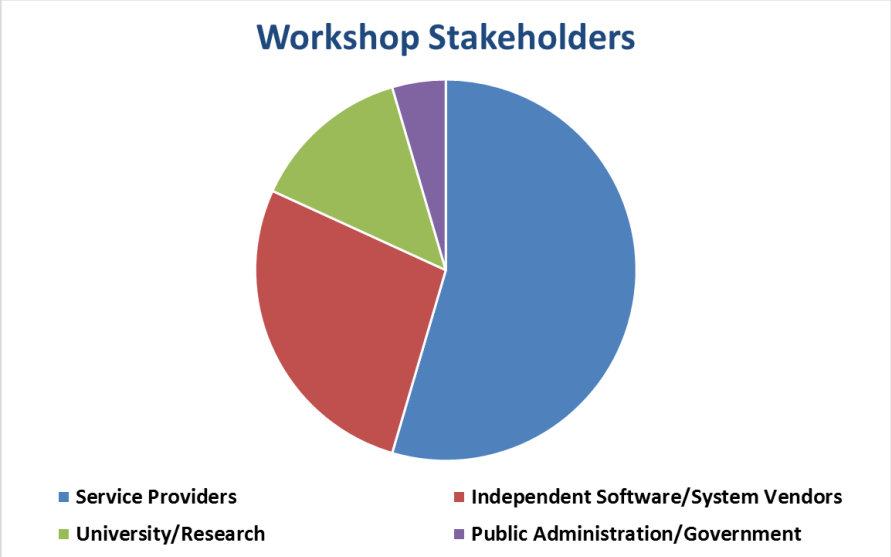

Figure 13 - Workshop Stakeholders

The workshop agenda (Table 8) included different presentations by HP. The talks pointed out on-going industrial cloud initiatives as well as research topics addressed within the A4Cloud project. Last talk focused intentionally on the presentation of one of the tools (namely, Cloud Offerings Advisor or Cloud Offerings Advisory Tool – see Appendix B for a description of the main tool functionalities) developed

within A4Clolud, followed by round table discussions within focused groups. Small groups of up to 6 people (per table) were moderated by HP in order to gather stakeholder feedback.

3.3.2 Presentations' Abstracts

The following paragraphs describe the different presentations at the stakeholder workshop.

- **Cloud28+ vision, Xavier Poisson:** This presentation was concerned with the current challenges and opportunities for developing a European cloud strategy. Cloud services need to address several challenges (e.g. vendor lock-in, security, compliance, etc.) in order to increase the trust in the cloud. These challenges affect cloud adoptions. Cloud providers addressing such challenges will lead the way for supporting the envisaged cloud strategy for Europe. Moreover, they will be part of the potential developments due to the adoption of the cloud.
- **HP Helion Network, Colin l'Anson:** This presentation was concerned with the new HP Helion Network offerings. It highlights the main cloud offering features, which are based on OpenStack technology. Relying on OpenStack addresses some of the challenges faced by cloud services. The second part of the presentation discussed how HP Helion Network is aligned with the objectives of the EU cloud strategy.
- **Cloud technology – a view from HP Labs, Julio Guijarro:** This presentation highlighted current and future issues concerned with cloud, security, big data and mobility. It gave an overview of on-going research at the Cloud and Security Lab.
- **Cloud risk, trust and accountability, Siani Pearson:** This presentation was concerned with emerging cloud threats affecting trust in the cloud. It gave an overview of the complexity of on-going legislative activities constraining cloud governance. The presentation then motivated a rationale supporting the need for accountability in the cloud.
- **Cloud Accountability Project, Nick Wainwright:** This presentation gave a brief overview of the EU Cloud Accountability Project (A4Cloud) funded by the European Commission's seventh framework programme (FP7).
- **Tools for accountability, Massimo Felici:** This presentation was concerned with one of the tools currently developed by the Cloud. Accountability Project. The Cloud Offerings Advisor, the tool presented, is aligned with the Cloud28+ vision and the development of a European cloud strategy. The demo video was followed by round table discussions in order to gather stakeholder feedback.

Table 8 - Workshop Agenda

Session	Agenda Item	Speaker
Morning Sessions		
10:00-10:45	**Welcome and Opening**	
	Cloud28+ vision	Xavier Poisson HP Converged Cloud
	HP Helion network	Colin l'Anson HP Converged Cloud
10:45-11:00	**Coffee break**	
11:00-12:30	**Session 1: Cloud Technology and Cloud Trust**	
	Cloud technology – a view from HP Labs	Julio Guijarro HP Labs
	Cloud risk, trust and accountability	Siani Pearson HP Labs
12:30-14:00	**Lunch**	
Afternoon Sessions		

14:00-16:00	**Session 2: Cloud Accountability Project**	
	Introduction	Nick Wainwright HP Labs
	Tools for accountability	Massimo Felici HP Labs
	Round table discussions and feedback	
16:00-16:15	**Closing remarks**	Colin I'Anson HP Converged Cloud

3.3.3 Comments and Questions

This section lists the gathered questions and comments grouped by the main areas addressed in the technical presentations. The highlighted points informed the final workshop remarks.

3.3.3.1 Cloud Risk, Trust and Accountability

- Promise of cloud is 'something magic' (that is, 'not transparent').
- Enterprises will give different answer (about accountability) from what the 'customer' will say. Definition of accountability will be different depending on whom you ask. Legal people will give different answers too.
- SAP Hana System has a very strong optimisation for particular architectures. Performance varies across processor architectures. It is necessary to expose performance of underlying hardware (i.e. different performances across different platforms). Governance would require assuring that the underlying technology is approved for the specific operational context.
- Open solutions that can be assessed/audited (e.g. OpenStack)
- Although infrastructures can be certified, how to deal with dynamic ecosystems? It is necessary to identify (monitoring) mechanisms once an infrastructure has been certified.
- For UK government, can HP Helion be approved for use by UK government? And, is HP going to do that?
- Unclear understanding of cloud service contracts; cloud offerings tailored to different types of customers.
- How does 'account' (and accountably) work with standards? For instance, using standard frameworks and ITIL processes.
- Does account include information about failures from supply chain, e.g. availability in a geographic zone; this would be useful.
- Clarifying the multi-disciplinary (cultural) perspectives by technology independent models (e.g. like ontologies bridging different stakeholders).

3.3.3.2 Cloud Accountability Project: Tools for accountability

- Is requirement for storage and processing different?
  This leads to some confusion in the way the tool present storage/processing requirements and relate them to the roles of data controller and data processor.
- Could geographic requirements be 'more granular' for customers who want to identify a specific country?
  This points to the identification of services and processers being located in a specific region. Currently, it is possible to specify broad continental areas like EU or large states like US and China.
- The tool (Cloud Offerings Advisor) is a bit too 'supplier focussed'. Customers have (need) to have answered certain questions in advance. Could the tool's requirements not be more 'open' questions? (Maybe supporting subsequent refinements)

- How much information should a user need to know before using the tool?
- Does the tool gather such (contractual, cloud offer) information automatically (online)?
  This point is questioning how to populate the offers available in the tool. The tool support matching offers rather than searching offers. Therefore, a critical functionality is concerned with populating the offers in the tool. This could be done directly by providers (or automatically by the tool). It is currently assumed that providers are willing to provide their offers via the tool. The tool configuration (in terms of offers available) would need to take into account also offers by other providers too.
- How much of the contract is revealed to the users?
  This relates strongly to the “More Info” and “Show legal terms” functionalities (buttons) currently supported by the tool.

- There is a need for a revision management – every time there is a new agreement there is need to highlight this – what about detection of the change to alert the customer that there is a change to the agreement?
  This highlights the necessity to combine the offering tool with other tools concerned with monitoring of cloud services and notification to customers.
- It is necessary that there be an alliance ‘of the open’ (referring to federated cloud services)
- Who is the person who is actually using the tool? Unclear purpose and business model for the tool.
- There is a mixture of terms including contract terms and technical terms?
- All the tools that you make have to work together (interoperability across different tools as well as providers).
- Tool “feels a bit like a ‘car configurator’ and there should be something at a higher level more of the business level”.
- Looks like just a broker, if they won’t give you the data, or if the data isn’t available, won’t work.
- We have to provide comparable measure for performance
- What happens next? Is it automated to purchase?
- How do you track the success of this? How does the provider track whether the customer came to the provider as a result of this tool? Commercial questions – how do you list them in the tool if all-same price?
- Difficult to have a tool capturing all offerings; it is very difficult to collect detailed offers. Moreover, the tool would be difficult to use if there are too many offers (how such offers would be listed? By price?). At the EU level one strategy could be to identify a minimum level of requirements (to comply with)
- The presented tool should probably focus on fewer key selection criteria for EU services like data location and security (simplified selection criteria).
- A common language for expressing legal (contract) terms seems to be necessary.
- The tool should be tailored to different expertise. The tool seems supporting more customers rather than providers.
- Would the tool become ‘intelligent’ in filtering/ordering specific offerings based on previous contextualised selected criteria?

3.3.3.3 Comments from Round Table Discussions

- Accountability comes “towards the end of the process” (I think this means once something is being used/is operational) whereas you have put it at the start
- Infrastructure services are not where the interest is – this should be about a marketplace of business processes which is where the real customer interest is
- Accountability should be about business processes, but if you get into this there are many other rules to take account of such as financial sector rules or healthcare rules
- SaaS customers seek outsourcing of business processes – there should be templates for business processes
- Contract flow is simpler at lower levels in the stack

- Who carries responsibility for loss of data?
- It is difficult [impossible] to make a full catalogue of services
- The presented tool would not be of interest for large providers (such as telecom companies) who are not interested in providing their services in such manner (personalised cloud service arrangements rather than open offers).
- Tool should focus on specific areas – locale, data protection, and security. It should not try to address all non-functional or functional requirements.
- What about more granular services (e.g. services like returning location, or the value of a currency, etc. that other services use)
- Specialist SLAs are the differentiator for some service, lack of transparency may be a business tactic
- A cloud "broker" should have open APIs
- Difficult to get a tool that have the full set of services; start from Key European criteria (huge differentiator) that we have to search: Data location, Security levels, then move to offering, get a standard pricing.
- Generate a request for proposal (RFP) out of the tool. Way to define the initial criteria of selection. First filtering. All criteria in a shot is impossible. Layers of criteria (from more general to most individual). Thousands of services would be a nightmare.
- G-Cloud approach in the UK; the government has the power to ask every provider willing to enter it to sign security agreements
- Have an European "CE" stamp (minimum level of service)
- Common language and contracts description
- Risk and trust. Verify that levels of security and risks are ok. Audit moving the 'speed' of cloud
- Have a sales team or support team to help the customer.
- Ensure the quality of sources. Policy of certification to ensure customers the services will be correctly executed.

3.3.4 Main Remarks

The discussions with stakeholders and the feedback received during the workshop highlighted different interesting comments concerning the specific tool presented (i.e. Cloud Offerings Advisor). We summarise such comments and feedback in some remarks that may apply to other accountability mechanisms (tools in particular) too.

- **Business Model**: Stakeholders discussed the business model of the Cloud Offerings Advisor. In particular, from a business perspective, they envisioned different roles of the tool. One possible user group for the tool will be cloud brokers. Stakeholders also provided further feedback how to populate the Cloud Offerings Advisor with information about providers' offerings. The tool is aligned with common business terms adopted by service providers, and can be extended to clarify service offerings to cloud customers. Clarifications of the different aspects of cloud services help cloud customers to use the Cloud Offerings Advisor effectively and to make the most of its functionalities. The usability of the tool (also in terms of guidance to customers) is among the critical non-functional requirements. The main functionality of the tool is matching rather than searching for cloud offerings.
- **Non-Functional Requirements:** The list of non-functional requirements captured by the selection criteria (for matching cloud services) can be tailored to cover specific areas of interest (e.g. data protection, location). The identification of such criteria can benefit from experiences of other types of services (e.g. in the telecom domain, prices of offerings have no direct meaningful relation to the specific services and their quality for customers).
- **Legal and Contractual Terms:** The analysis of legal and contractual terms is a challenging one. Most legal and contractual terms seem to be to a certain extent comparable, although it can be difficult to assess them on a more technical (and operational) level. However, the usage of the tool will ease the comparison of legal and contractual across different cloud services. The current version of the Cloud Offerings Advisor supports the gathering of cloud offerings by a standardised templates. One possible extension of the current implementation involves the automatic gathering of such information.

- **Standards and Certifications:** One important perspective is concerned with the standards and certifications adopted by a group of federated service providers. This suggests moving towards a common auditing/certification scheme adopted by different providers. At the technical level, this requires a level of interoperability across different services. At the governance level, this requires a corresponding certification/auditing scheme for cloud services.

4 Discussion and Concluding Remarks

In total, about 90 stakeholders were involved in the workshops that comprise WS3. All workshops proved to be fruitful with respect to generating new requirements for the tools and for the project in general. The requirements elicited in WS3 were not only directly related to the tools explored in the workshops, but also general requirements that stakeholders expect for accountable cloud providers. The main themes were related to what is possible to do with the data, conformance to data agreement, protection correction and data handling mechanisms, value chain and multi-tenant information, and other information they would like to receive from the providers.

It is important to highlight that in general the cloud subjects were very positive to the tool presented and the concepts of the A4Cloud project. The answers from the participants were very consistent. The cloud subjects showed that they were concerned with accountability, and happy to have tools that will help them in accomplishing it. It was clear that being “cloud” or “not cloud” was not a very clear concept for them, but after the explanation of “what is the cloud”, they understood the concept and the risks involved. The cloud subjects also voiced concern with respect to possible limitations of accountability of cloud providers, and were happy to have tools that will help them in accomplishing accountability and transparency. The workshops were very good in the sense of creating more awareness of the cloud and also about concepts of accountability for the cloud.

For demonstrating the project, we found out that the “health use case” does not work in all cases, the younger people said that anything related to health doesn’t have an impact on them and that they can’t relate to it, for them it seems too far from their reality. On the other hand, it works better for people that are over 40. The video showed in the session was very good to convey the message of the data track tool. As one lesson learned we believe this is a good medium for getting the message of the project across to cloud subjects. In the video we focused on showing the main functionalities of the tool and also the goals of the project. The questionnaire was also a very good complement to consolidate the message to the participants of the workshop.

Cloud customers involved IT experts and most of them have worked with security for some years. In general, the participants were very interested in the concepts around accountability and also on the tools that will be generated by the project. Some concepts, such as transparency, create a high level of interest in cloud customers. In regard to the COAT tool, they were concerned if the tool would be useful for large corporations. The process used in the workshop seems to work well with cloud customers, but we need to be clearer about how the users of the tools will benefit from them.

The workshop with cloud customers was a real opportunity to present on-going work within the A4Cloud project (in particular, how specific accountability mechanisms would support a European cloud market) to business stakeholders, who share an interest in adopting technological innovations as market enablers for the cloud. The stakeholders feedback gathered gave us some insights about the need to clarify the business models of accountability mechanisms (in particular, software tools) in order to facilitate their deployments in operational environments. Moreover, the stakeholder feedback pointed out the critical role of governance in the cloud. On the one hand, the adoption of accountability mechanisms would push towards a standardization of cloud offerings. This would enable comparisons across different cloud providers and ease the adoption from cloud customers. This is the reason why accountability is perceived as a potential market enabler for the cloud. On the other hand, emerging cloud standards and third party certifications (assessments) have a critical role in shaping future technological developments for the cloud.

References

1. Ellen Gottesdiener, Requirements by Collaboration: Workshops for Defining Needs, Addison Wesley, 2002.
2. A4CLOUD, Accountability For Cloud and Future Internet Services, Annex I, Description of Work, Grant agreement 317550, Version date 2012-09-13.
3. Nils Brede Moe (Ed.), Stakeholder Workshop 1 Results (Initial Requirements), A4Cloud, Deliverable D2.1, Version 1.0, March 2013.
4. Erdal Cayirci (Ed.), Risk Modelling for Cloud Services Workshop Results, A4Cloud, Deliverable D 2.2, Version 1.0, November 2013.
5. F. Liu et al., NIST Cloud Computing Reference Architecture, NIST Special Publication 500-292, September 2011.
6. Daniela S. Cruzes, Tore Dybå: Recommended Steps for Thematic Synthesis in Software Engineering. ESEM 2011: 275-284, 2011.
7. Yang, Haibo and Tate, Mary. "A Descriptive Literature Review and Classification of Cloud Computing Research," Communications of the Association for Information Systems: Vol. 31, Article 2, 2012.

Appendixes

A. Data Track

The Data Track is a user-side transparency-enhancing tool that provides users with a history function documenting what personal data the user has revealed to which services sides under which conditions. Also, the Data Track provides functions that allow users to access their personal data at the remote services side online.

Earlier versions of the Data Track tools were conceptualized and developed under the European FP7 projects PRIME and PrimeLife. The program aims at giving users more transparency and control over the personal information that they distribute over different online services at different times. Its goal is to let users know in an understandable and easily manageable way which personal information has been given to which internet and cloud service providers and to let users control this information in different ways, such as revoking it, correcting it, or exercising their right to be forgotten.

For A4Cloud project, a novel graphical user interface for visualizing the users' information in the Data Track tool has been implemented, as shown in Figure 14. More detailed descriptions of the A4Cloud Data Track user interface can be found in the A4Cloud Deliverable D:C-7.1. This way of showing the tracking of the users' data has been called the "trace view", presenting an overview of which personal data item have been sent to service providers as well as which service providers might have similar personal data items of a user. The idea is that users should be able to see all the information (displayed in the top of the UI) that they have submitted to services on the Internet (these Internet services are shown in the bottom panel of the interface). If the user clicks on one of the Internet services in the bottom panel she will be shown arrows pointing to the personal data items possessed by that service. Similarly, if the user clicks on a personal data items in the top panel, arrows will indicate which Internet services obtained these data.

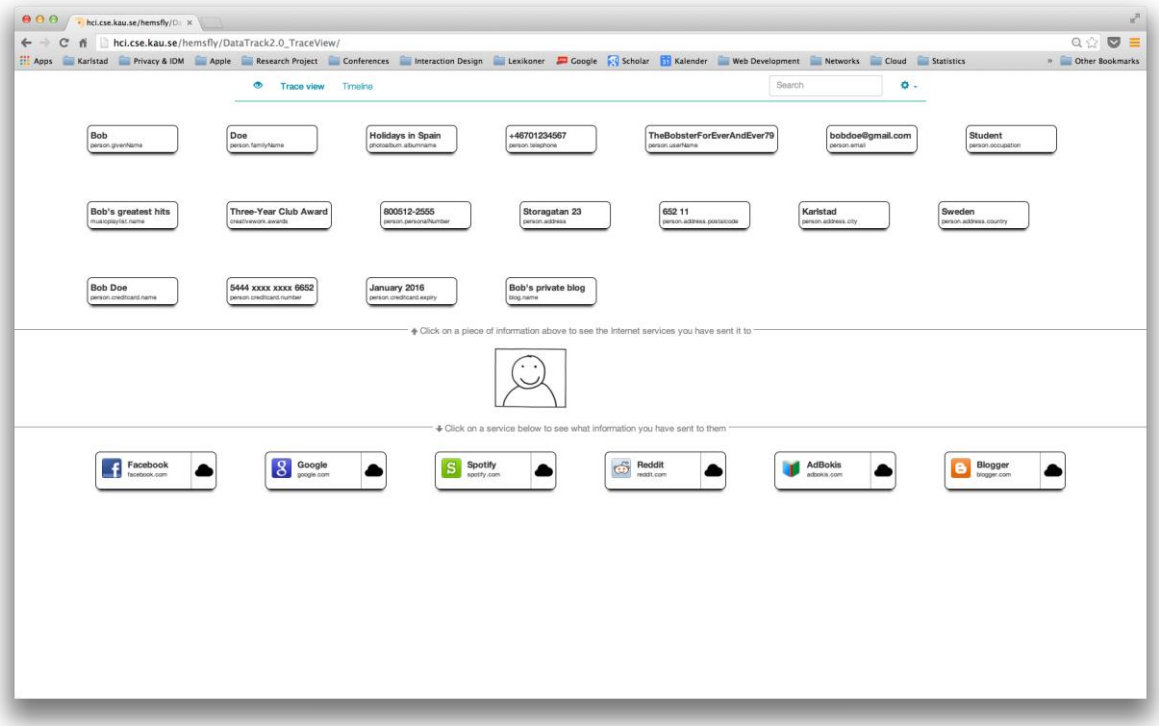

Figure 14 - Data Track - Trace View

For example, if a user wants to buy a book at Amazon.com, she has to fill out and send over the Internet some personal information to Amazon in order to get the book delivered to her home. She must, for instance, provide her name, her address, an email, a telephone, a password, and her credit card information, among other things. In addition, Amazon could obtain extra implicitly revealed or derived information without the user being aware of it, for example, the time that she bought the book, the type of books she might be interested in, the location she was at when she bought the book, whether she is a reliable customer, and so on..

The Data Track program helps users to visualize which personal information was explicitly disclosed by them to Amazon at what point in time. Also, users can see which implicitly disclosed and which derived information was collected by the online service. Users are then able to correct or remove their information from the Amazon servers if the Amazon service allows it. For instance, the user might want to see which home address she submitted to Amazon the first time when she became a member, and if she has moved since then she might want to replace the old address information with her new address. She can also see the information that Amazon stores about her, such as whether she is a reliable customer, the location from where she bought the book, and so on.

B. Cloud Offerings Advisory Tool (COAT)

This Section summarises the main features of the Cloud Offerings Advisory Tool (COAT) or Cloud Offerings Advisor.

Tool description
This tool will be used to provide information/guidance to potential cloud consumers (SMEs and data subjects) on: how to understand and assess what a cloud service provider is offering from a privacy and security perspective, how to compare offerings (from a data protection compliance and provider accountability point of view), and offer guidance on the meaning of the comparison attributes. The output is a guided comparison of the service offers along with an explanation of potential risks. The tool also logs the offered advice and the user's decision for accountability purposes. Figure 15 shows a screenshot of the Cloud Offerings Advisory Tool (in particular, the matching of different cloud offers).

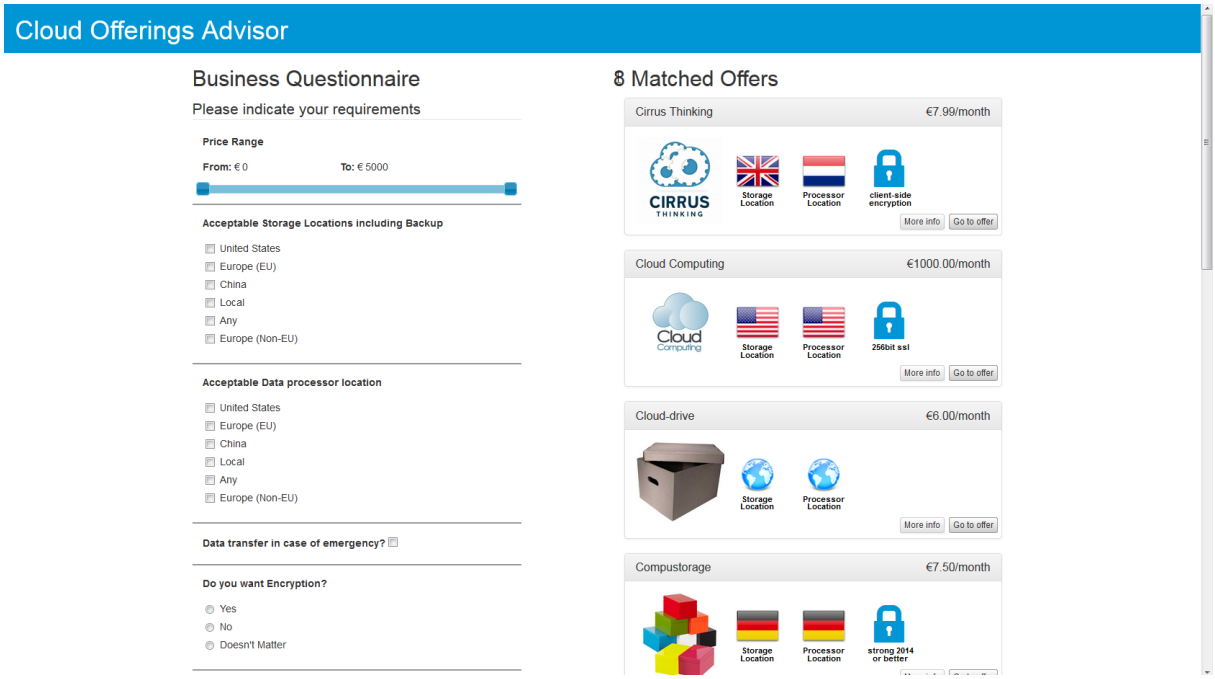

Figure 15 - Cloud Offerings Advisory Tool (COAT)

Inputs
- User info. (location, Roles, contact details)
- Context (contextual info)
- What are the needs and requirements?
- structured service offerings
- a model of cloud contracts and points of attention
- Reputation info
- Knowledge-Base of threats

Processes
The processes inside the tool will include:
- User Requirements Questionnaire
- Logging component
- Information and explanation Component
- Comparison Component
- Matchmaking Component

Outputs
- Guidance on things to pay attention to when exploring and comparing the terms of service offerings
- Overview of comparable service offerings
- a requirement list to give to the CSP
- SME guidance
- User Interfaces
- Input: Questionnaire asking for certain input
- Output: a report with a comparison between different service offerings
- Users

The intended user of the system is SMEs and data subjects (end-users)

Main features
- Checking user requirements
- Checking Offers by cloud service providers
- Comparing offers by cloud service providers
- Explaining the terms of offerings
- Suggesting best offerings that match the user requirement
- Give general guidance to users on service offerings

Use case
The tool will be useful in:
- Taking decisions about service offerings
- Understanding the contract terms of the service offerings

Benefits
- Cloud Service Providers
    - Decrease complexity for the customers to pick a cloud provider (like you!)
    - Highlight the unique criteria in your offer easily (what distinguishes your contract offer from the others)
    - Increase market exposure for cloud providers! (growing market opportunities)
    - Match cloud demands with offerings
- Cloud Customers
    - Ease the comparison of alternative cloud offerings (increasing transparency by clarifying contractual terms)
    - Ease the public concerns about the cloud (thorough transparent contract terms, guidance, etc.)
    - Increase trust in cloud offerings

C. List of Workshop Materials and Raw Data

This Section describes the main material used in all workshops run in WS2. The material is organized as follows:

C.1 Cloud Subjects Workshop – Data Track Tool – Karlstad, Sweden – April 2014
C.2 Cloud Subjects Workshop – Data Track Tool – Trondheim, Norway – May 2014
C.3 Cloud Customers Workshop – Coat Tool – Trondheim,Norway – June 2014
C.4 Cloud Customers Interviews – Transparency Requirements – Norway – 2014