Published in "Cloud Computing and Services Science", CCIS Vol 581, pp 38-57, 03 February 2016
http://link.springer.com/chapter/10.1007%2F978-3-319-29582-4_3

Accountability through Transparency for Cloud Customers

Martin Gilje Jaatun[1], Daniela S. Cruzes[1],
Simone Fischer-Hübner[2], and Julio Angulo[2]

[1] Department of Software Engineering, Safety and Security
SINTEF ICT
NO-7465 Trondheim, Norway
{martin.g.jaatun,danielac}@sintef.no
http://infosec.sintef.no

[2] Karlstad University
Karlstad, Sweden
info@kau.se
http://www.kau.se/

Abstract. Public cloud providers process data on behalf of their customers in data centres that typically are physically remote from their users. This context creates a number of challenges related to data privacy and security, and may hinder the adoption of cloud technology. One of these challenges is how to maintain transparency of the processes and procedures while at the same time providing services that are secure and cost effective. This chapter presents results from an empirical study in which the cloud customers identified a number of transparency requirements to the adoption of cloud providers. We have compared our results with previous studies, and have found that in general, customers are in synchrony with research criteria for cloud service provider transparency, but there are also some extra pieces of information that customers are looking for. We further explain how A4Cloud tools contribute to addressing the customers' requirements.

Keywords: Cloud Computing, Accountability, Transparency, Privacy, Security

1 Introduction

Cloud computing, which allows for highly scalable computing and storage, is increasing in importance throughout information technology (IT). Cloud computing providers offer a variety of services to individuals, companies, and government agencies, with users employing cloud computing for storing and sharing information, database management and mining, and deploying web services, which can range from processing vast datasets for complicated scientific problems to using clouds to manage and provide access to medical records [1].

Several existing studies emphasize the way technology plays a role in the adoption of cloud services, and most of these studies conclude that the most important challenges are related to security, privacy and compliance [2–6]. Cloud service users may hand over valuable and sensitive information to cloud service providers without an awareness of what they are committing to or understanding of the risks, with no control over what the service does with the data, no knowledge of the potential consequences, or means for redress in the event of a problem. In the European A4Cloud research project[3], our focus is on accountability as the most critical prerequisite for effective governance and control of corporate and private data processed by cloud-based IT services. We want to make it possible to hold cloud service providers accountable for how they manage personal, sensitive and confidential information in the cloud, and for how they deliver services. This will be achieved by an orchestrated set of mechanisms: preventive (mitigating risk), detective (monitoring and identifying risk and policy violation) and corrective (managing incidents and providing redress). Used individually or collectively, they will make the cloud services in the short- and longer-term more transparent and trustworthy for:

- users of cloud services who are currently not convinced by the balance of risk against opportunity
- their customers, especially end-users who do not understand the need to control access to personal information
- suppliers within the cloud eco-system, who need to be able to differentiate themselves in the ultimate commodity market.

In this paper we report on the results of an elicitation activity related to transparency requirements from the perspective of cloud customers. A Cloud Customer in our context is an entity that (a) maintains a business relationship with, and (b) uses services from a Cloud Provider; correspondingly, a Cloud Provider is an entity responsible for making a [cloud] service available to Cloud Customers.

Transparency is the property of an accountable system that is capable of 'giving account' of, or providing visibility of, how it conforms to its governing rules and commitments [7]. 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 behavior of a system or process. An accountable organization is transparent in the sense that it makes the policies on 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. In an ideal scenario, the user knows the information requirements and is able to communicate that clearly to the provider, and in return, the provider is transparent and thus willing to address the regulatory and legislative obligations required with regard to the assets.

[3] `http://a4cloud.eu`

The rest of the chapter is organized as follows. Section 2 presents some background from the literature. Section 3 explains the methods that we used to elicit the views of the stakeholders. In section 4 we present the results, and in section 5 we illustrate how the tools developed by the A4Cloud project contribute to meeting the customer transparency requirements. We discuss our findings compared to related work in section 6, and draw our conclusions in section 7.

2 Related work

Transparency is closely connected to trust [8]. Onwubiko [9] affirms that trust is a major issue with cloud computing irrespective of the cloud model being deployed. He says that cloud users must be open-minded and must not wholeheartedly trust a provider just because of the written-down service offerings without carrying out appropriate due diligence on the provider; where certain policies are not explicit, users should ensure that missing policies are included in the service contract. By understanding the different trust boundaries, each cloud computing model assists users when making decision as to which cloud model they can adopt or deploy.

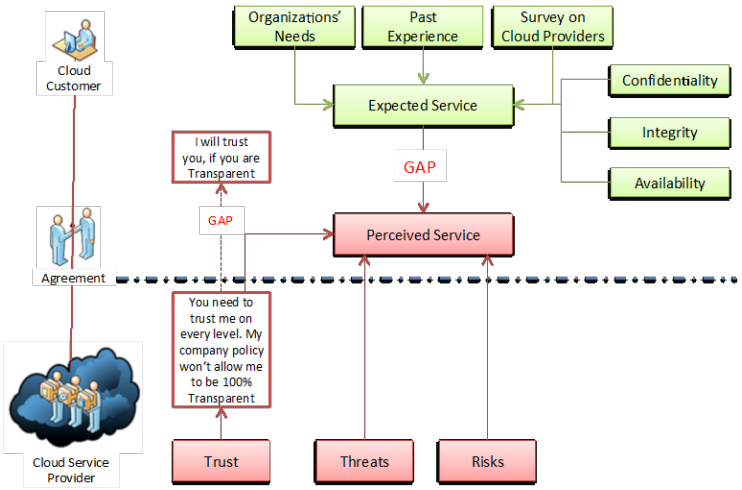

Fig. 1. Understanding Cloud Computing Gaps

Khorshed et al. [10] highlight the gaps between cloud customers' expectations and the actually delivered services, as shown in Fig. 1 (adapted from Khorshed et al. [10]). They affirm that cloud customers may form their expectations based on their past experiences and organizations' needs. They are likely to conduct some sort of survey before choosing a cloud service provider similar to what people do

before choosing an Internet Service Provider (ISP). Customers are expected to also establish to what extent providers satisfy confidentiality, integrity and availability requirements. On the other hand, cloud service providers may promise a lot to entice a customer to sign a deal, but harsh reality is frequently accompanied by insurmountable barriers to keeping some of their promises. Many potential cloud customers are well aware of this, and are consequentially still sitting on the sidelines. They will not venture into cloud computing unless they get a clear indication that all gaps are within acceptable limits.

Durkee [11] says that transparency is one of the first steps to developing trust in a relationship, and that the end customer must have a quantitative model of the cloud's behavior. The cloud provider must provide details, under NDA if necessary, of the inner workings of their cloud architecture as part of developing a closer relationship with the customer. Durkee also says that this transparency can only be achieved if the billing models for the cloud clearly communicate the value (and avoided costs) of using the service. To achieve such clarity, the cloud vendor has to be able to measure the true cost of computing operations that the customer executes and bill for them.

Pauley [12] proposed an instrument for evaluating the transparency of a cloud provider. It is the only empirical evaluation that we found that focuses on transparency in the cloud as a subject of study. The study aims to help businesses assess the transparency of a cloud provider's security, privacy, auditability, and service-level agreements via self-service Web portals and publications. Pauley designed a scorecard (Table 1) to cover the assessment areas frequently raised in his research, and to begin to establish high-level criteria for assessing provider transparency. He concludes that further research is needed to determine the standard for measuring provider transparency. In our research we used a different strategy than Pauley; we have interviewed customers of cloud services to see what kind of information they would like to get from the cloud providers.

3 Method

As part of the project, we were responsible for running a set of stakeholder workshops for eliciting requirements for accountability tools. In total, our elicitation effort has involved more than 300 stakeholders, resulting in 149 stakeholder requirements. The first workshop dealt with eliciting initial accountability requirements, serving as a reality-check on the three selected business use cases we had constructed [13]. The second workshop dealt with risk perception. The aim was to focus on the notion of risk and trust assessment of cloud services, future Internet services and dynamic combinations of such services (mashups). After the first two workshops, we decided to organize multiple smaller, local workshops on each theme to ease participation of cloud customers and end users. The third set of workshops presented stakeholders with accountability mechanisms to gather their operational experiences and expectations about accountability in the cloud.

Of particular importance to this study was the risk workshop, where 15 tentative requirements related to transparency where identified. This workshop

Table 1. Pauley's Cloud Provider Transparency Scorecard

Aspect	Criteria	Mentioned in Interviews?
Business factors	1. Length in years in business > 5?	No
	2. Published security or privacy breaches?	Yes
	3. Published outages?	Yes
	4. Published data loss?	Yes
	5. Similar customers?	Yes
	6. Member of ENISA, CSA, CloudAudit, OCCI, or other cloud standards groups? No	
	7. Profitable or public?	No
Security	8. Portal area for security information?	Yes
	9. Published security policy?	Yes
	10. White paper on security standards?	Yes
	11. Does the policy specifically address multi-tenancy issues?	Yes
	12. Email or online chat for questions?	No
	13. ISO/IEC 27000 certified?	Partially
	14. COBIT, NIST SP800-53 security certified?	Partially
	15. Offer security professional services (assessment)?	No
	16. Employees CISSP, CISM, or other security certified?	Partially
Privacy	17. Portal area for privacy information?	Yes
	18. Published privacy policy?	Yes
	19. White paper on privacy standards?	Yes
	20. Email or online chat for questions?	No
	21. Offer privacy professional services (assessment)?	No
	22. Employees CIPP or other privacy certified?	Partially
External audits or certifications	23. SAS 70 Type II	No
	24. PCI-DSS	No
	25. SOX	No
	26. HIPAA	No
Service-level agreements	27. Does it offer an SLA?	Yes
	28. Does the SLA apply to all services	No
	29. ITIL-certified employees?	No
	30. Publish outage and remediation?	Yes

comprised 20 international stakeholders from the manufacturing industry, telecom, service providers, banking industry and academia, and the tentative transparency requirements were subsequently presented to our interviewees as a starting point for the discussion.

In addition to the stakeholder requirements, we have devised a set of high-level requirements which, from an organizational perspective, set out what it takes to be an accountable cloud provider [14]. These requirements intend to supplement the requirements elicitation process by providing a set of high-level "guiding light" requirements, formulated as requirements that accountable organizations should meet. In short, these requirements state that an accountable organization that processes personal and/or business confidential data must 1) demonstrate willingness and capacity to be responsible and answerable for its data practices 2) define policies regarding their data practices, 3) monitor their data practices, 4) correct policy violations, and 5) demonstrate policy compliance.

From these activities we have created a repository with requirements from all elicitation workshops, the guiding lights requirements as well as a number of more technical requirements that have originating from the conceptual work and technical packages in the project. These have been classified in terms of whether they are functional requirements, which are directly related to the actors involved in the cloud service delivery chain, or requirements for accountability mechanisms, which are related to the tools and technologies that are being developed in the project.

For refining and confirming the elicited requirements of transparency, we have performed an interview study with eight interviewees, followed by an in-depth analysis of the collected information.

Invitations were sent to our list of contacts in Norwegian software companies. Participation was voluntary. Eight people accepted to participate in the interviews. The participants were all IT security experts working with cloud related projects. The participants represented six different organizations: a consultancy, 2 cloud service providers (1 public, 1 private), an application service provider, a distribution service provider, and a tertiary education institution.

The interviews were performed on Skype and lasted about one hour. The main questions of 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? (Fig. 2)
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? (Fig. 3)
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? (Fig. 4)

The eight interviews for this study were transcribed into text documents based on the audio recordings. For further analysis of the transcription, we followed the Thematic Synthesis recommended steps proposed by Cruzes and Dybå

[15]. Thematic synthesis is a method for identifying, analyzing, and reporting patterns (themes) within data. 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. Five steps were performed in this research: initial reading of data/text (extraction), identification of specific segments of text, labeling of segments of text (coding), translation of codes into themes, creation of the model and assessment of the trustworthiness of the model.

4 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 (Fig. 2):

1. clear statements of what is possible to do with the data,
2. conformance to data agreements,
3. how the provider handles data,
4. location,
5. who else other than the provider is participant of the value chain,
6. multi-tenant situations,
7. what the provider does with the data,
8. procedures to leave the service
9. assurance that the user still owns the right to the data.

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 [web] page where they could tell me about security mechanisms, for example, 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 related information about how the providers handle the data. On location, the respondents are concerned about where the data is physically stored, and the legal jurisdiction of the services. Another important piece of information is about sub-providers, if there are any; where they are located and whether they meet legal requirements of the customer's location. Multi-tenant situations are a concern of the customers, and they would like to have this information transparent. Also, information on how the providers ensure that data from one customer will not be accessed by another customer.

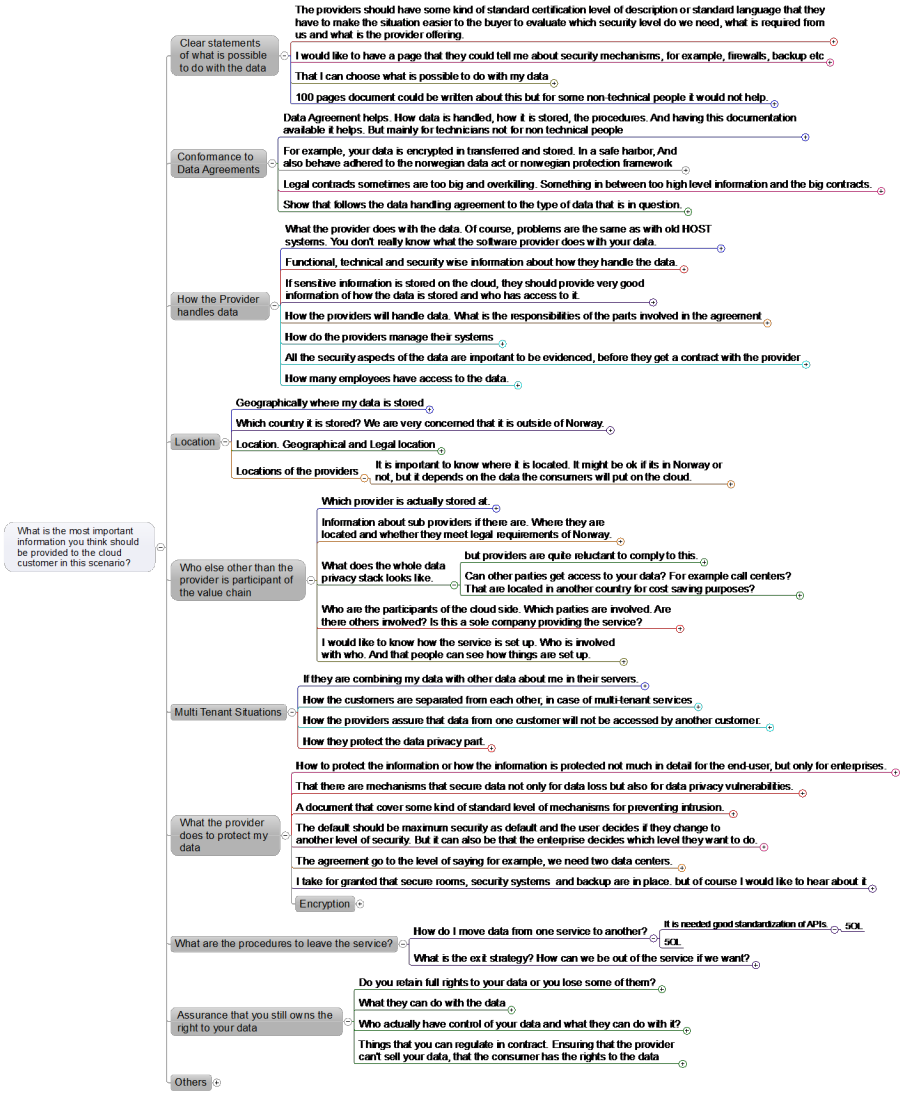

Fig. 2. Important Upfront Information for Transparent Services

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 that they would like to have the procedures to leave the service and on how to move data from one service to another transparent. Besides, they would like to have the assurance that they still own the rights 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: 1) upfront transparency; 2) community discussions, 3) customer awareness; 4) way out; 5) reputation; 6) encryption; 7) data processor agreements; and 8) location.

Some answers were overlapping towards the answers from the first question: upfront transparency, location and conformance to data processor agreement. Interesting answers for this question were related to community discussions, customer awareness and reputation. The respondents said that it increases their trust in 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 also mentioned "Way out", meaning that they would like to have webpages telling them what to do to remove the data from the service provider.

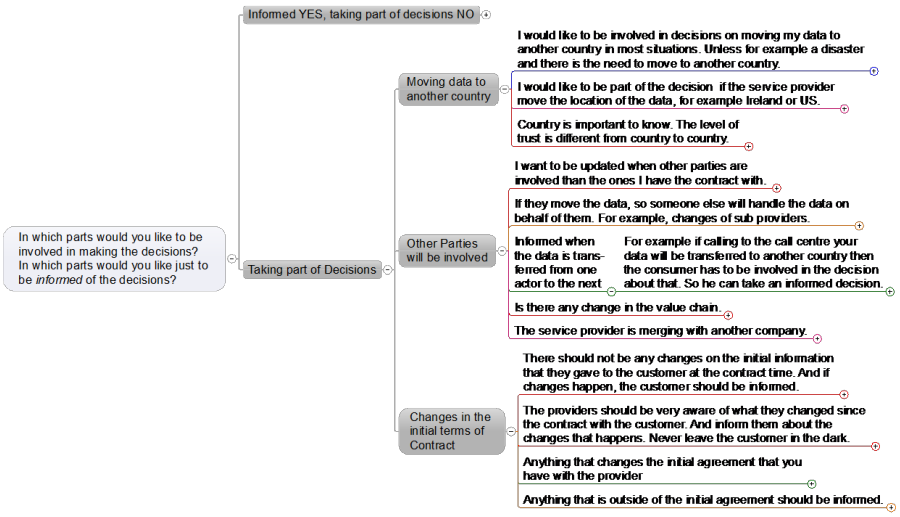

Fig. 3. Involvement on making Decisions

On the questions: "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 (Figure 4); the exceptions were when the provider was moving data to another country, other parties are introduced in the service provider 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 in the decisions", and another one said: "there are some decisions that we don't need to explicitly know about, but it has to be regulated by some other agreement about the responsibility of each one towards the data". One respondent also 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 center 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 [was signed], and inform them about the changes that happen. Never leave the customer in the dark."

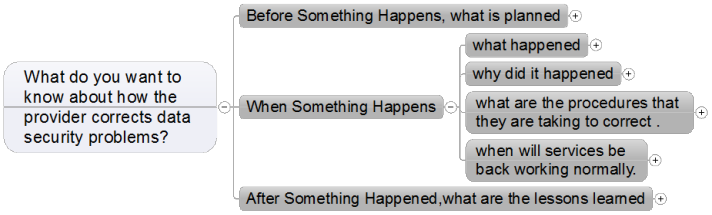

Fig. 4. Transparency on Correction of Data Security Problems

When asked on what they would want to know about how the provider corrects data security problems, it was again surprising to learn that the participants have not thought much on what they could expect from the providers if some security issue happens. Most of the respondents needed further elaboration of the question before they would start saying something. Then, the participants stated that they would like to know what is planned before something happens; when something happens they want to know how the providers are handling the situation, why the problem happened, and when will the services be back online. 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. These responses are collated in the taxonomy shown in Fig. 4.

5 Transparency Tools

Many of the transparency mechanisms that customers expressed a desire for are actually being developed by the A4Cloud project [14]. Furthermore, a central theme of A4Cloud is the development of the Accountability PrimeLife Policy Language (A-PPL), which allows end users to specify a privacy policy that also covers accountability requirements, including transparency [16]. A4Cloud is developing an A-PPL Engine which will serve as a Policy Decision Point for the associated policies at each cloud provider. Other tools developed by A4Cloud include the Cloud Offerings Advisory Tool (COAT), which allow cloud customers to select an appropriate cloud provider based on relevant accountability requirements, including transparency [17]. This will eventually allow transparency requirements to be built into standard cloud service level agreements (SLAs), where transparency is just one of several security attributes [18].

In the following subsections, we will show in more detail how the A4Cloud DataTrack tool enhances transparency for end users by allowing users to visualize the personal data that have been disclosed to different online services.

5.1 The Data Track tool

The Data Track transparency-enhancing tool was initially developed as part of the European FP6 and FP7 research projects PRIME[4] and PrimeLife[5]. Initially, the Data Track consisted of a history function for keeping a log of each transaction in which a user discloses personal data. The log contained a record for the user on which personal data were disclosed to whom, for which purposes, which credentials and/or pseudonyms have been used in this context as well as the details of the agreed-upon privacy policy. These transaction records were stored at the user side in a secure manner. During the PrimeLife project and in the A4Cloud project, the Data Tack tool has been extended with online access functions, conceptually allowing users to exercise their data subjects' rights to access their data at the remote services sides and to request correction or deletion of their data (as far as this is permitted by the service side).

In its backend the architecture of the Data Track consists of four high-level components. First, the *user interface* component, which displays different visualizations of the data disclosures provided by the Data Track's *core*. Second, the *core* component is a backend to the UI with local encrypted storage. Through a RESTful API, the core is able to provide a uniform view to the UI of all users' data obtained from a service provider via so called *plugins*. Third, the *plugin* component provides the means for acquiring data disclosures from a given source (e.g., a service provider's database) and parsing them into the internal format readable by the core. Fourth, the Data Track specifies a generic *API* component that enables a service provider to support the Data Track by providing remote access, correction, and deletion of personal data. Based on solutions

[4] EU FP6 project PRIME, `https://www.prime-project.eu/`
[5] EU FP7 project PrimeLife `http://primelife.ercim.eu/`

proposed by Pulls *et al.* [21], the transfer of data through a service's API can be done in a secure and privacy-friendly manner. By retrieving data from different services through their provided APIs users would be able to import their data immediately into the Data Track and visualize it in different ways, thus providing immediate value for end-users.

Detailed descriptions of the initial Data Track's proof-of-concept, user interfaces and results of its usability evaluations are given by Fischer-Hübner et al. [22], and further design process is described by Angulo et al. [20]. The security and privacy mechanisms of its software implementation have been documented by Hedbom, Pulls et al. [23–25].

5.2 Visualizing data disclosures.

The design of the Data Track's UI considers different methods for visualizing a user's data disclosures in a way that is connected to this user's momentary intentions. Based on the ideas from previous studies suggesting ways to display data disclosures [26, 27] and the creation of meaningful visualizations for large data sets [28–30], we have designed and prototyped two main visualizations for the Data Track as part of the A4Cloud project, we refer to them as the *trace view* and the *timeline view*

The main *trace view* interface, shown in Fig. 5, is separated into three main panels. The services to which the user has released information appear in the bottom panel and the information attributes that have been released by the user to these different services appear in the top panel. The user is represented by the panel in the middle, with the intention of giving users the feeling that this interface is a *place* that focuses on them (i.e., data about them and services that they have contacted). When the user clicks on one (or many) service(s) from the bottom panel, a trace is shown to the personal attributes (represented with graphical icons) that have been disclosed to the selected service(s). Similarly, if the user selects a personal attribute from the top panel, a trace is shown to the service(s) to which the selected attribute has been disclosed at some point in time. By its design, the trace view lets users answer the question of "*what information about me have I sent to which online services?*"

In order to cater for users perceptual capabilities and considering the screen real state, filtering mechanisms are put in place that would allow users to filter for information that is relevant to what they want to find out. In the trace view, users can search using free-text (i.e., by typing the name of a company, like Flickr or Spotify, or the name of a personal attribute, like 'credit card' or 'heart rate'), they can also select categories of data or individual pieces of data, as well as the number of entities to be displayed on the screen.

The other visualization presents each disclosure in chronological order, thus name the *timeline view*. In this view, shown in Fig. 6, each circle along the vertical line represents the service to which personal data has been disclosed at a specific point in time. Each box besides a circle contains the personal attributes that were sent with that particular disclosure. In order to keep the size of the boxes consistent and to not overwhelm users with visual information, the boxes

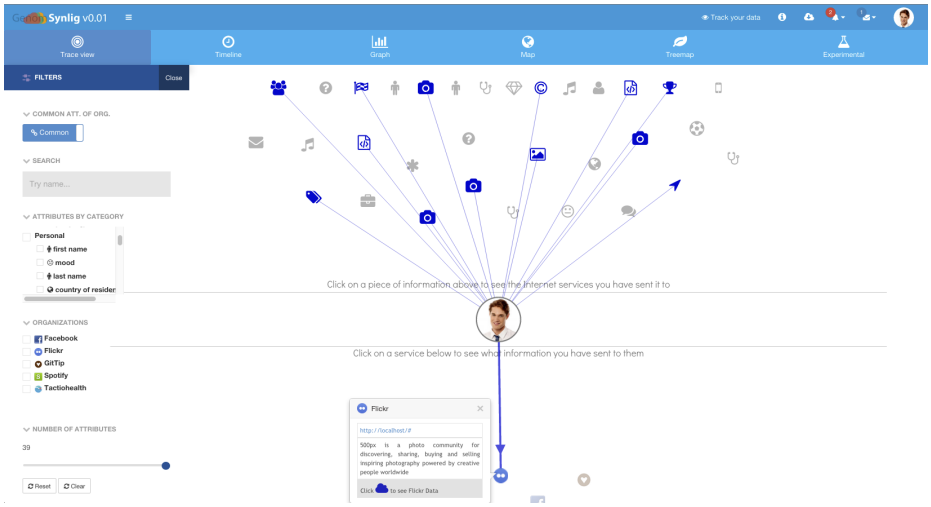

Fig. 5. The prototype of the *trace view* interface of the Data Track tool

only show four attributes initially, and users have the option to look at the rest of the attributes in that particular disclosure by clicking in the "Show more" button. Users can scroll vertically indefinitely, thus unveiling the disclosures of data that they have made over time, and allowing them to answer the question "*what information about me have I sent to which online services at a particular point in time?*"

Filters have also been considered for the timeline view, allowing users to search, for instance, for all disclosures made in a specified time interval, or all disclosures made to a particular service.

Thanks to the envisioned architecture in the A4Cloud project, which considers the use of the A-PPL Engine mentioned earlier, the Data Track would allow its end-users to access personal data about them that is located in a service's side (i.e., stored in the service's databases). In both, the trace view and the timeline view, a button (in shape of a cloud) located besides a service providers logo, opens up a dialog showing users the data about them that is located on the services' side. This dialog, shown in Fig. 7, presents not only the personal attributes that have been explicitly collected by the service provider, but also data about the user that has been derived from analysis. Through this dialog users would also be able to request correction or deletion of personal attributes, thus being able to exercise their data access rights.

5.3 User evaluations of the Data Track's UI.

Throughout the A4Cloud project, the user interface of the Data Track has gone through several iterative rounds of design and user evaluations. The evaluations had the purpose of testing the level of understanding of the interface, but also

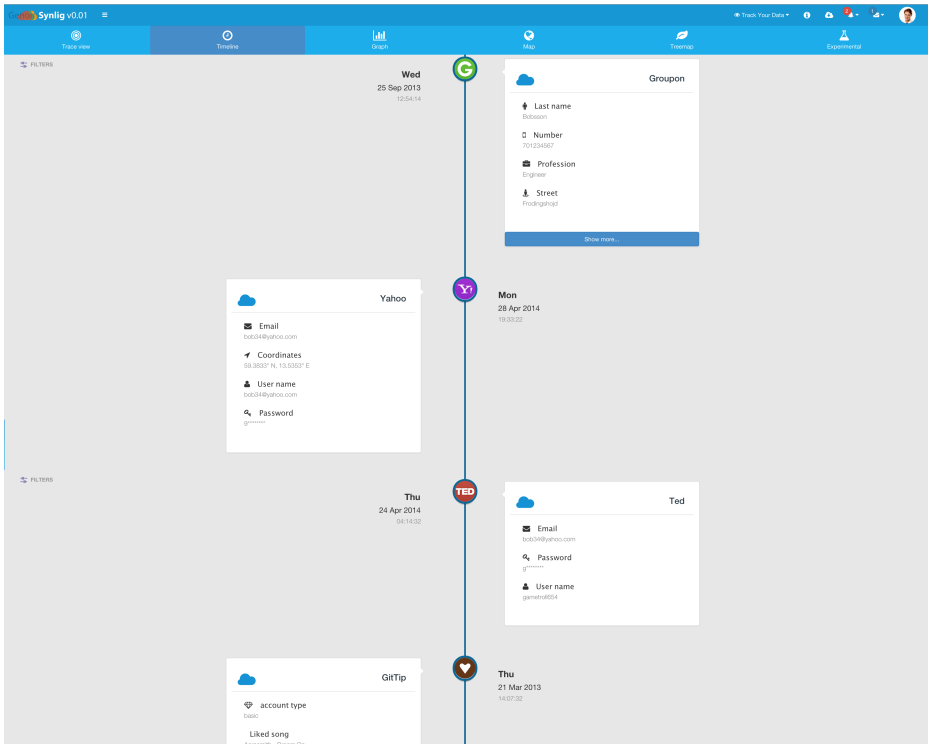

Fig. 6. The prototype of the *timeline view* interface of the Data Track tool

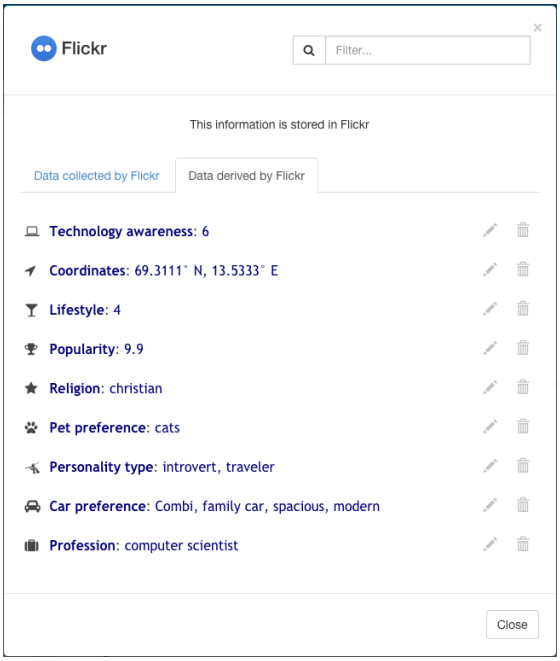

Fig. 7. The pop-up dialog showing the explicitly sent and derived data stored at the service's side

as a method for gathering end-user requirements on the needs and expectations that such a tool should provide to its users.

Usability testing of earlier designs of the Data Track revealed that lay users expressed feelings of surprise and discomfort with the knowledge that service providers analyze their disclosed data in order to derive additional insights about them, like their music preferences or religion. In general, evaluations have also shown that participants understand the purpose of the tool and ways to interact with it, identifying correctly the data that has been send to particular service providers, and using the filtering functions to answer questions about their disclosed personal data. The tests also revealed users' difficulties when differentiating between data that is locally stored under their control in their computers and data that is accessed on the services' side (and shown through the pop-up dialog), as well as skepticism of the level of security of the data stored locally.

During an evaluation workshop, attendees discussed the advantages and possible risks of using such a tool, as well as the requirements to make such a tool not only user-friendly but also adopted in their routinary Internet activities. One participant, for instance, commented that the transparency that the Data Track provides, would encourage service providers to comply with their policies and be responsible stewards of their customers data, "*it would keep me informed and hold big companies in line.*". Another participant mentioned the benefit of be-

coming more aware of disclosures made to service providers, "*makes you aware of what information you put on the Internet, you probably would be more careful.*" On the other hand, a participant commented on the risk of accumulating large amounts of personal data in a single place, "if there is one tool collecting all the data, then it is a single point of failure...".

6 Discussion

After analyzing all the collected information we compiled a list of requirements elicited in the interviews, as shown in Appendix A. 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, protection of the data, decisions and corrections of the data.

Pauley [12] designed a scorecard reproduced in Table 1 to cover the assessment areas frequently raised in the research, and to begin to establish high-level criteria for assessing provider transparency. When comparing our list of elicited requirements (see Appendix A) to Pauley's scorecard, we can see some slight differences in the criteria that Pauley described as information that should be provided by the cloud providers and the information that the customers are looking for. In the criteria about the business factors, the customers did not mention being concerned about the number of years in business, nor about membership of CSA, CloudAudit, OCCI, or other cloud standards groups, or if the providers are profitable or public. There is a possibility that the respondents did not mention these criteria because (a) companies in Norway are usually stable, and (b) membership of a group or association does not in itself guarantee good performance or compliance, even if the group or association promotes a certain standard.

On the security and privacy aspects, the customers mentioned all the criteria, but they did not mention directly the standards/certifying bodies, such as ISO/IEC 27000, COBIT and NIST, but they mentioned that it would be nice to know if the provider was certified somehow, based on some criteria. The customers also did not mention the need to know about "external" audits. One of the reasons for not mentioning security standards and certification bodies may be that companies that we have investigated are predominantly private companies in Norway, where there are not strong requirements from the certification bodies yet.

One important aspect not very much explored in Pauley's scorecard is that customers would like providers to be transparent about what is possible to do with the data. In addition, customers were quite concerned about transparency on exit procedures ("way out") and ownership of the data. The concern over data ownership is interesting seen in the light of Hon et al. [31], who found no evidence of cloud contracts leading to loss of Intellectual Property Rights.

Another aspect further mentioned by the customers is on the decisions made on "ongoing" services, where the customers would like that: "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."

Physical location and legal jurisdiction, as well as specific information on the value chain was a very important aspect to be transparent about for the cloud customers, and it was not explicitly mentioned in Pauley's scorecard.

The interviewees did not show a desire for the kind of detailed information Durkee [11] deems necessary (the inner workings of their cloud architecture as part of developing a closer relationship with the customer), and as also pointed out by Durkee, some respondents were also aware that the costs of such clarity may be prohibitive, and we might add that this level of disclosure seems highly unlikely for ordinary customers of commodity cloud services.

The Data Track tool that we have described in Section 5 focuses more on end users (*data subjects*) than professional cloud users, but is clearly relevant for the customers of the cloud users. However, the tool can also be used to follow up on what a provider claims to be able to do with the data (A.1). It can be used to follow up on the geographical location of the customer's data (A.2), and can also help illustrating the existence of services from other parties (A.4).

7 Conclusions

Cloud computing has been receiving a great deal of attention, not only in the academic field, but also amongst the users and providers of IT services, regulators and government agencies. The results from our study focus on an important aspect of accountability of the cloud services to customers: transparency.

The customers made explicit all the information that they would like the providers to be transparent about. Much of this information can be easily provided at a provider's website. Our contention is that being transparent can be a business advantage, and that cloud customers who are concerned with, e.g., privacy of the data they put into the cloud, will choose providers who can demonstrate transparency over providers who cannot.

Our study increases the body of knowledge on the criteria needed for more accountable and transparent cloud services, and confirms the results from previous studies on these criteria. The list of requirements in Appendix A complements, in part, the existing criteria.

An area for future research is to further evaluate how cloud providers currently make the information required by cloud customers available. In addition, what are the effects of having transparent services in terms of costs and benefits to cloud customers and providers. Besides, we plan to increase the number of participants responding to our interview guide and adding strength to the evidence provided in this paper. Another aspect we would like to investigate, is if the results will be different for users of the different types of services (e.g., SaaS vs IaaS).

Acknowledgements

This paper is based on joint research in the EU FP7 A4CLOUD project, grant agreement no: 317550.

References

1. Paquette, S., Jaeger, P.T., Wilson, S.C.: Identifying the security risks associated with governmental use of cloud computing. Government Information Quarterly **27**(3) (2010) 245 – 253
2. Kuo, A.M.: Opportunities and challenges of cloud computing to improve health care services. J. Med. Internet Res. **13**(3) (2011) e67
3. Gavrilov, G., V., T.: Security and privacy issues and requirements for healthcare cloud computing. In: Proceedings of ICT Innovations. (2012)
4. AbuKhousa, E., Mohamed, N., Al-Jaroodi, J.: e-health cloud: Opportunities and challenges. Future Internet **4**(3) (2012) 621
5. Rodrigues, J.J., de la Torre, I., Fernandez, G., Lopez-Coronado, M.: Analysis of the security and privacy requirements of cloud-based electronic health records systems. J. Med. Internet Res. **15**(8) (2013) e186
6. Ahuja, S.P., Mani, S., Zambrano, J.: A Survey of the State of Cloud Computing in Healthcare. Network and Communication Technologies **1**(2) (2012) 12–19
7. Felici, M., Koulouris, T., Pearson, S.: Accountability for data governance in cloud ecosystems. In: Cloud Computing Technology and Science (CloudCom), 2013 IEEE 5th International Conference on. Volume 2. (Dec 2013) 327–332
8. Yang, H., Tate, M.: A descriptive literature review and classification of cloud computing research. Communications of the Association for Information Systems **31** (2012)
9. Onwubiko, C.: Security issues to cloud computing. In Antonopoulos, N., Gillam, L., eds.: Cloud Computing. Computer Communications and Networks. Springer London (2010) 271–288
10. Khorshed, M.T., Ali, A.S., Wasimi, S.A.: A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. Future Generation Computer Systems **28**(6) (2012) 833 – 851 Including Special sections SS: Volunteer Computing and Desktop Grids and SS: Mobile Ubiquitous Computing.
11. Durkee, D.: Why cloud computing will never be free. Commun. ACM **53**(5) (May 2010) 62–69
12. Pauley, W.: Cloud provider transparency: An empirical evaluation. Security Privacy, IEEE **8**(6) (Nov 2010) 32–39
13. Bernsmed, K., Tountopoulos, V., Brigden, P., R'ubsamen, T., Felici, M., Wainwright, N., Santana De Oliveira, A., Sendor, J., Sellami, M., , Royer, J.C.: Consolidated use case report. A4Cloud Deliverable D23.2 (2014)
14. Jaatun, M.G., Pearson, S., Gittler, F., Leenes, R.: Towards strong accountability for cloud service providers. In: Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on. (Dec 2014) 1001–1006
15. Cruzes, D.S., Dybå, T.: Recommended steps for thematic synthesis in software engineering. In: Proceedings of ESEM 2011. (2011) 275–284

16. Azraoui, M., Elkhiyaoui, K., ʹOnen, M., Bernsmed, K., Santana De Oliveira, A., Sendor, J.: A-ppl: An accountability policy language. In Garcia-Alfaro, J., Herrera-Joancomartí, J., Lupu, E., Posegga, J., Aldini, A., Martinelli, F., Suri, N., eds.: Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance. Volume 8872 of Lecture Notes in Computer Science. Springer International Publishing (2015) 319–326
17. Alnemr, R., Pearson, S., Leenes, R., Mhungu, R.: Coat: Cloud offerings advisory tool. In: Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on. (Dec 2014) 95–100
18. Jaatun, M.G., Bernsmed, K., Undheim, A.: Security slas – an idea whose time has come? In Quirchmayr, G., Basl, J., You, I., Xu, L., Weippl, E., eds.: Multidisciplinary Research and Practice for Information Systems. Volume 7465 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2012) 123–130
19. Fischer-Hübner, S., Angulo, J., Pulls, T.: How can cloud users be supported in deciding on, tracking and controlling how their data are used? In Hansen, M., Hoepman, J.H., Leenes, R., Whitehouse, D., eds.: Privacy and Identity Management for Emerging Services and Technologies. Volume 421 of IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg (2014) 77–92
20. Angulo, J., Fischer-Hübner, S., Pulls, T., Wästlund, E.: Usable transparency with the data track: A tool for visualizing data disclosures. In: Extended Abstracts in the Proceedings of the Conference on Human Factors in Computing Systems. CHI '15, Seoul, Republic of Korea, ACM (2015) 1803–1808
21. Pulls, T.: Preserving privacy in transparency logging. PhD thesis, Karlstad University Studies. 2015:28 (June 2015)
22. Fischer-Hübner, S., Hedbom, H., Wästlund, E.: 13. PrimeLife - Privacy and Identity Management for Life in Europe. In: Trust and Assurance HCI. Springer (June 2011) 261
23. Hedbom, H., Pulls, T., Hjärtquist, P., Laven, A.: Adding secure transparency logging to the prime core. In: Post-Proceedings of the Fifth International Summer School: Privacy and Identity Management for Life, Nice, France, 7th - 11th September 2009, in press (2010)
24. Hedbom, H.: A survey on transparency tools for enhancing privacy. In: The future of identity in the information society. Springer (2009) 67–82
25. Pulls, T., Peeters, R., Wouters, K.: Distributed privacy-preserving transparency logging. In: Workshop on Privacy in the Electronic Society. WPES '13, Berlin, Heidelberg, Germany (November 2013) 83–94
26. Kani-Zabihi, E., Helmhout, M.: Increasing service users' privacy awareness by introducing on-line interactive privacy features. In: Information Security Technology for Applications. Springer (2012) 131–148
27. Kolter, J., Netter, M., Pernul, G.: Visualizing past personal data disclosures. In: Availability, Reliability, and Security, 2010. ARES'10 International Conference on, IEEE (2010) 131–139
28. Becker, H., Naaman, M., Gravano, L.: Beyond trending topics: Real-world event identification on twitter. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM'11). (2011)
29. Freeman, L.C.: Visualizing social networks. Journal of social structure **1**(1) (2000) 4
30. Kairam, S., MacLean, D., Savva, M., Heer, J.: Graphprism: compact visualization of network structure. In: Proceedings of the International Working Conference on Advanced Visual Interfaces, ACM (2012) 498–505

31. Hon, W., Millard, C., Walden, I.: Negotiating Cloud Contracts – Looking at Clouds from Both Sides Now. STAN. TECH. L. REV. **81** (2012) Queen Mary School of Law Legal Studies Research Paper No. 117/2012.

A List of Requirements from Transparency Interviews

A.1 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 his/data data
- The provider should have a page that they could tell the cloud customer about security mechanisms, e.g., 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.

A.2 Conformance to Data Agreement

- The provider should make available the technical documentation on how data is handled, how it is stored, and the procedures.
- 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.

A.3 Data handling

- The provider should provide functional, technical and security-related information about how they handle the data.
- The provider should provide very good information on how the data is stored and who has access to it.

A.4 Value chain

- In case of using services from other parties, the provider should inform cloud customers on what the responsibilities of the parties involved in the agreement are.

– 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.

A.5 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.

A.6 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.

A.7 Decisions

– 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.

A.8 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.