Internet Engineering Task Force
Internet-Draft
Intended status: Best Current Practice
Expires: January 16, 2014

J. Pohle
HIIG

July 15, 2013

Operational Privacy for Cloud Services
draft-ietf-operational-privacy-cloud-00

Abstract

Cloud computing is a way of distributed computing over a network to use and deliver IT applications, processing capability, and storage space. Due to the cloud service providers' control over the storage and the processing of information, as well as the communication between cloud client and the cloud system, and strong tendencies of monopolization, cloud computing poses significant risks to the privacy and data protection rights of individuals.

This document describes measures by cloud service providers to support their clients to comply with privacy and data protection laws.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 16, 2014.

Copyright Notice

Table of Contents

1. Introduction

Cloud computing is a way of distributed computing over a network to use and deliver IT applications, processing capability, and storage space. Due to the cloud service providers' control over the storage and the processing of information, as well as the communication between cloud client and the cloud system, and strong tendencies of monopolization, cloud computing poses significant risks to the privacy and data protection rights of individuals.

This document intends to provide guidance for cloud service providers on how to support their clients to protect individuals and their rights to privacy and data protection. Cloud clients externalize functions that are subject to privacy and data protection laws. As they cannot externalize their legal responsibilities, they depend on a cloud service that does not hinder but supports their compliance to privacy and data protection rules.

Privacy is not a single concept, but many. It has different meanings in different disciplines, contexts, cultures, and times. Privacy theories even differ in the social relationship between entities that they describe. While some theories are based on social interactions between individuals, others focus on the relationship between individuals and organizations. On the other side, with regard to information relating to individuals and its processing, legal implementations are quite similar in their architectures: Individuals are allocated different rights that shall help them to protect themselves, and data processors are allocated obligations concerning the collection, storing, processing, and disclosure of information. The guidance in this document is generic and can be used anywhere in the world, without reference to specific legislation.

1.1. Scope

This document aims to provide guidance for cloud service providers to deliver their services in a privacy and data protection compliant manner. The scope of this guidance is restricted to privacy practices that mitigate risks for cloud clients and data subjects. Traditional security threats and risks for the cloud service provider and its services by unauthorized handling of personal data are outside the scope of this document, even if they may also affect cloud clients and data subjects.

1.2. Document Layout

The document is organized as follows. Section 2 explains the terminology used in this document. Section 3 reviews the relationship between the parties involved in providing and using cloud services. Section 4 discusses threats to privacy and data protection as they apply, and suitable mitigations, both sorted by privacy protection goals.

2. Terminology

This section defines central terms used in this document. As in [RFC4949], each entry is preceded by a dollar sign ($) and a space for automated searching.

2.1. Basic Terms

2.1.1. Entities

$ Cloud Client: An entity that uses cloud computing services.

$ Data Controller: Legal term. Any entity that determines the purposes and means of the processing of personal data and is legally responsible for compliance with privacy and data protection rules.

$ Data Processor: Legal term. Any entity that processes personal data on behalf of the data controller.

$ Data Protection Official: Legal term. Also: Privacy Official. The entity that is legally responsible for supervision of compliance with privacy and data protection rules with respect to the legal regime applicable to the data controller.

$ Data Subject: Legal term. A human being whose privacy and data protection rights are to be protected.

$ Insider: Any entity that works under control and/or on behalf of the data processor.

$ Privacy Official: Legal term. Also: Data Protection Official. The entity that is legally responsible for supervision of compliance with privacy and data protection rules with respect to the legal regime applicable to the data controller.

$ Service Provider: An entity that provides cloud computing services.

$ Third Party: Legal term. Any entity other than the data subject, the data controller, or the data processor.

2.1.2. Data and Data Processing

$ Information: The term is used in accordance with the theory of semiotics. Information has four layers or dimensions: syntax, semantics, pragmatics, sigmatics. Syntax refers to the form of the information, i.e. the data. Semantics refers to the meaning of the information. Pragmatics refers to the purpose of the information. Sigmatics refers to the object referenced by the information.

$ Personal Data: Legal term. Any information relating to an individual who can be identified, directly or indirectly.

2.2. Security

Security primarily means the protection of assets in one's own interest against generally illegitimate information interests. The attacker can be an individual, a group of individuals, or an organization.

The goal is to protect information and systems from unauthorized access, use, collection, modification, destruction, or disclosure.

Confidentiality, integrity, and availability are core principles of information security.

2.3. Privacy

Privacy and data protection primarily protect against generally legitimate information interests, especially by more powerful players. Therefore the attacker is generally an organization.

The goal is to protect individuals from legitimate, but unreasonable access, use, collection, modification, destruction, or disclosure of information relating to them, and generally unreasonable interference with their personal affairs.

The decision whether a legitimate information interest is deemed to be lawful is either based on a statutory rule, on the informed consent of the data subject, or a carefully weighting of the competing interests by the data controller. In either way, the responsibility for legal compliance is allocated by law to the data controller.

Many different sets of privacy and data protection principles have been developed over the years. Following the classic CIA triad, a goal-oriented approach is being used in this document. The privacy protection goals are confidentiality, integrity, availability, transparency, unlinkability, and intervenability ([RostPfitzmann] and

[Hansen]). As the controller is responsible for compliance with privacy and data protection rules, these goals are to be evaluated from the perspective of the controller with respect to the rights of the data subject.

The applicable law protecting the individual is to be determined by the controller. The law could be different for any or all processors.

3. Social Interaction Model

This document is focussing on organizational information processing. Both the cloud service providers and the cloud clients are considered organizations.

The service provider provides cloud computing services to the cloud client. The cloud client uses these services to store and process information related to individuals, i.e. personal data. The cloud client determines the purposes and the means of the processing of personal data and is called "data controller" or "controller" by law and allocated responsibility for legal compliance.

The relationship between service provider and cloud client is based on a contract. The service provider stores and processes the client's data on behalf of the client. By law, the service provider is called "data processor" or "processor".

The service provider may subcontract additional subcontractors. These subcontractors are processors, too.

4. Threats and Mitigations

This section is sorted by privacy protection goals: confidentiality, integrity, availability, transparency, unlinkability, and intervenability. For each privacy protection goal, possible threats are listed and described. For each threat, operational practices that mitigate it are described. Threats are drawn liberally from [Steinmueller], [FIP], [Solove], and [Art29].

4.1. Confidentiality

Confidentiality refers to preventing of information becoming known to any unauthorized entity. The confidentiality of personal data is compromised if any processor facilitates surveillance for any third party, or discloses personal data to any third party without the explicit consent of the data controller and/or the data subject.

Information may become known to any third party that is unauthorized with respect to the legal regime applicable to the controller even if this third party may be an authorized entity with respect to the legal regime applicable to one or more of the data processors. Any third party that might seem trustworthy from the processor's point of view also might not be trusted by the data controller or the data subject, respectively. The third party might apply the information to a context different from where it originated. Information therefore might change its meaning. The third party might process the information for a purpose different from that for which it was collected. The potential for loss of confidentiality can undermine the data subject's trust to the data controller and/or the data processor, and therefore reduce the willingness for exchanging information in the first place.

4.1.1. Disclosure

Disclosure is the revelation of information about an individual to any third party.

Threats: The data processor might either actively or passively disclose information to any third party. Active voluntary disclosure refers to giving away information that already is under control of the data processor. Passive voluntary disclosure refers to not preventing insiders from disclosing information to any third party.

Mitigations: To assure confidentiality, insiders are to be prevented from disclosing personal data to any third party outside of the data processor's control. If any information disclosure is required by law, it is done only in accordance with written instructions that define responsibilities, necessary requirements, procedures, and logging. If not explicitly forbidden by law, the data controller is to be informed about the disclosure in advance, otherwise afterwards. To enable a data controller's reasonable decision about the legal compliance, the data processor shares all necessary information about the disclosure with the data controller, especially the identity of the third party, the purpose and the extent of the disclosure. If the data processor is forbidden by law to inform the data controller, the data processors informs the appropriate data protection or privacy official. If the data processor is explicitly forbidden by law to inform the data subject, the data controller, or the data protection or privacy official about a particular disclosure, the data processor informs the data controller in advance about all circumstances under which it will disclose personal data to any third party. If the data processor is explicitly forbidden by law to disclose any information about the disclosure of personal data to one or more third parties, the data processor MUST NOT declare that it is providing a secure or privacy compliant or data protection compliant service.

4.1.2. Surveillance

Surveillance is the observation or monitoring of an individual's behavior, activities, communications, or other changing information. While disclosure generally refers to a single act, surveillance generally implies continuity over a period of time.

Threats: The effects of surveillance on the individual can range from discomfort to angst and behavioral changes such as inhibition and self-censorship. Surveillance may also harm autonomy and self-determination. The privacy of the individual may be harmed even if the individual is not aware of the surveillance because the probable possibility of surveillance may be enough to change the individual's behavior.

The data processor might surveil the data subject by collecting more information about the individual's behavior, activities, communications, or other changing information than necessary for providing the service as expected by the data controller. The data processor might enable or facilitate eavesdropping by intentionally not using proper communication security measures or none at all. The data processor might enable or facilitate man-in-the-middle attacks by re-routing communications before they are encrypted.

Mitigations: To assure confidentiality, the data processor implements and uses proper security measures to prevent surveillance, especially by encrypting communications between nodes under control of the data subject, the data controller, and the data processor respectively. Information not collected need not to be protected against surveillance. Security measures such as encryption are to be selected based on a privacy and data protection analysis that extends a routine security analysis with respect to organizations as possible attackers and their abilities. If the data processor is explicitly required by law to enable or facilitate surveillance, the data processor informs the data controller in advance about all circumstances under which it may not prevent surveillance. If the data processor is explicitly forbidden by law to disclose any information about the surveillance by one or more third parties, the data processor MUST NOT declare that it is providing a secure or privacy compliant or data protection compliant service.

4.2. Integrity

Integrity refers to maintaining and assuring the accuracy and consistency of information over its entire life-cycle.

If decisions are based on inaccurate or inconsistent information, they may be wrong and/or otherwise adversely affecting the data subject.

4.2.1. Misattribution

Misattribution occurs when information or communications related to one entity are attributed to another.

Threats: The data subject may either be linked to information that is not related to him or her, or not linked to information that is or should be related to him or her. Misattribution might result in wrong decisions about the individual. The data processor might disclose wrong personal data to a third party, or withhold personal data that should be disclosed. The data processor might also delete wrong personal data, or preserve personal data that should be deleted. Misattribution might also lead to the data subject's exclusion from accessing its personal data, or from using services that he or she should be able to use.

Mitigations: To prevent misattribution, the data processor implements and uses proper forms of identification or authentication. To mitigate the consequences of misattribution, the data processor enables the data subject and/or the data controller to challenge the attribution or non-attribution of information. The data processor therefore enables the data subject and/or the data controller to

learn which information is attributed to them. Information that is neither collected nor stored cannot be misattributed.

4.2.2. Inaccuracy

Inaccuracy occurs when information are not correct, not up-to-date, or in a wrong context.

Threats: The data subject may be adversely affected by wrong decisions based on incorrect, outdated, or out-of-context information. This may be especially severe if this information is disclosed to any third party, as it often cannot review the accuracy of the information, or it has no particular interest to do so.

Mitigations: To assure accuracy, the data processor enables the data subject as well as the data controller to learn which information related to them is stored, and to change it. In particular, the data processor provides an easy way for data subjects and data controllers to correct and update stored personal data. The data processor also enables data subjects and data controllers to specify, change, and delete the context in which they perceive stored personal data as correct.

4.3. Availability

Availability refers to assuring timely and reliable access to information and services.

If information related to them are not available, data subjects might be hampered in excercising their rights to know what information related to them is stored and processed, to amend or correct inaccurate information, or to prove or disprove information vis-a-vis third parties. Without access to the information about them, data subjects may also be hindered to intervene in its storage, processing, and use. Without access to personal data they are responsible for, and information about how it has been stored, processed, used and/or deleted, data controllers may not be able to prove compliance with privacy and data protection rules.

4.3.1. Exclusion

Exclusion is the failure to allow individuals to access information related to them, or to use services that they should be able to use. It also refers to the failure to allow data controllers to access information they are legally responsible for.

Threats: Exclusion of data subjects and data controllers reduces accountability on the part of the data processor and any third party

that received information about the individual. Neither the data subject nor the data controller may be able to build up trust in the data processor if personal data is not available to be reviewed. If data subjects or data controllers are excluded from using services without reasonable grounds, it also may undermine trust in the data processor, especially if the data subject and/or the data controller had used services provided by the data processor in the past.

Mitigations: To assure availability, the data processor implements and uses proper access control mechanisms to give authorized entities access to information while preventing unautorized ones. If an entity could be successfully identified, access to all information related to that entity will be granted. If any exclusion is required by law, the data controller is to be informed in advance. If the data processor is not explicitly forbidden by law, the data processor hands over all information related to the excluded entity to the data controller or the data protection or privacy official, respectively.

4.4. Transparency

Transparency refers to assuring that systems, processes, and information could be comprehended, verified, and evaluated.

Most people do not or not fully understand how modern computing, especially cloud computing, really works, and what risks are associated with its use. To use cloud computing services despite their lack of knowledge about it, trust is necessary. The lack of transparency can undermine the data subject's trust to the data controller and/or the data processor, and therefore reduce the willingness for exchanging information in the first place.

Without transparency, an individual may be treatened as an object, not as a subject. Such treatment would threaten human dignity, and undermine personal autonomy and individual self-determination. As a result, most privacy and data protection laws include the individual's right to be informed and corresponding provisions to require data controllers informing individuals about the collection and use of personal data, especially its purposes.

4.4.1. Opacity

Opacity occurs when the data controller and/or the data subject are not reasonably informed about the entities contractually or otherwise involved, the entities authorized for accessing stored information or communications, the systems being used, the means of information processing being used, the locations of the systems, or the information stored on the data controller's behalf.

Threats: The consequence is that neither the data controller nor the data subject are able to adequately weight the risks of using the cloud services for compliance with privacy or data protection rules.

Mitigations: The data processor informs the data controller about the entities contractually or otherwise involved, the entities authorized for accessing stored personal data, and the conditions under which these entities may access stored personal data. The data processor also specifies the systems being used, their locations, and which legal regimes apply. It also describes the processes that are used for collecting, storing, processing, communicating, and deleting personal data in a comprehensible manner for the data controller to review the compliance with privacy and data protection rules. For all personal data collected and stored, the data processor defines the purposes of its processing and use in advance, and informs the data controller adequately. The data processor also reveals which information with respect to the individual's behavior, activities, or communications are collected while providing the cloud services.

4.4.2. Repudiation

Repudiation occurs when an entity is able to successfully challenge any of its past actions.

Threats: If the data processor is able to repudiate its past actions or that of entities acting on behalf of the data processor, it reduces accountability on the part of the data processor.

Mitigations: The data processor logs all of its actions concerning personal data and that of entities acting on behalf of the data processor. The logs are recorded in a manner that they are usable as evidence in court. If the data processor is required by law to not log particular actions, the processor informs the data controller in advance in an adequate manner to enable a data controller's reasonable decision about the legal compliance. If the data processor is forbidden by law to inform the data controller, the data processor informs the appropriate data protection or privacy official. If the data processor is explicitly forbidden by law to inform the data subject, the data controller, or the data protection or privacy official about actions which are not being logged, the data processor MUST NOT declare that it is providing a secure or privacy compliant or data protection compliant service.

4.5. Unlinkability

Unlinkability refers to assuring that an entity is not able to distinguish whether two pieces of information are related or not, or whether an information is related to a particular individual or not. It also refers to assuring that an entity is not able to use information for any non-authorized purpose.

4.5.1. Linkage

Linkage occurs when an entity is able to combine different formerly separated pieces of information related to the same event, behavior, act, individual, or group of individuals.

Threats: While individuals might be aware of the different pieces of information that an organization knows about them, through linkage organizations may infer new information about them. Linkage may thus disregard individuals' expectations of the limits of what organizations know about them. As a consequence, it undermines individuals' autonomy in their decision whether to share personal data with organizations or to withhold. Linking personal data that originate from different contexts, organizations disrespect individuals' decisions about the separation of different social roles and a socially adequate role-playing.

Mitigations: To assure unlinkability, the data processor collects only the amount of personal data that is necessary for providing the service as expected by the data controller. Information that is neither collected nor stored cannot be linked. Information without identifiers is harder to link, especially without additional knowledge. The data processor therefore removes identifiers at the earliest possible moment. If personal data is collected for different purposes, the data processor separates the data, so that entities performing tasks for one purpose may not access information stored for a different puspose. For this purpose, role-based access control mechanisms provide a widely accepted solution.

4.5.2. Identification

Identification is a special case of linkage. It occurs when an entity is able to link a piece of information to a particular individual.

Threats: There are situations where it is socially accepted or legally required to be identifiable. As many activities and communications occur outside of these controlled situations, the risk to be identifiable might deter individuals from free and self-determined activities and communications, and from excercising their

rights like free speech or participating in political, religious, or other legally protected activities. Entities able to identify individuals may excercise control over them, or discriminate them based on their activities or communications.

Mitigations: To prevent identification of individuals, the data processor provides services that can be used anonymously. Information that is neither collected nor stored cannot be related to an individual. If recognition of individuals is necessary, the data processor enables the use of pseudonyms. Pseudonyms should be under the individuals' control, contextually limited, and easily revocable. Often, transaction pseudonyms are sufficient. Information that can be used to uncover pseudonyms are stored separately. Identifiers are removed at the earliest possible moment.

4.5.3. Secondary Use

Secondary use occurs when information collected for one purpose is being used for a different purpose.

Threats: In modern, functionally differentiated societies, individuals play different roles in different social contexts. To decide on how to play these roles and what information to share in different contexts and for what purpose is part of the autonomy of the individual that is threatened when personal data collected for one purpose is being used for a different purpose without the individual's explicit consent. Organizations may opt for secondary use to earn additional income, or to pursue other organizational goals. The risk of unauthorized secondary use may discourage individuals to share information in the first place.

Mitigations: Purposes and acceptable uses are defined by contract between data processor and data controller. The data controller is responsible for obtaining individuals's consent. The data processor only uses personal data in accordance with the contract with the data controller.

There exist generally accepted cases of secondary use like data backups or information security but only if it is done in accordance with written instructions that define responsibilities, procedures, and logging. The data controller is informed about the kind of secondary use and associated protection measures in advance.

4.6. Intervenability

Intervenability refers to the ability of an entity to intervene in the storing and processing of personal data to protect its own rights or the rights of the data subject. Intervenability includes explicitly the ability of the individual to intervene in legitimate cases of storing and processing of personal data.

4.6.1. Provider Lock-in

Provider lock-in occurs when an entity is not able to access all its information, to take it out, or to take it out in a usable format.

Threats: The data processor might refrain from using standard data formats and service interfaces with the intent to bind the data controller to the cloud service. Without access to all personal data stored on behalf of the data controller, the data controller may not be able to migrate from one cloud provider to another, even if the data controller would be required by law to do so. If the data processor is able to prevent the data controller from migration to another cloud service provider, it reduces accountability on the part of the data processor. Even if the data controller might be able to access all personal data stored on its behalf, the lack of interoperability might hinder the transfer of personal data to a new cloud service provider.

To assure intervenability, the data processor uses standard data formats, open service interfaces, and other means facilitating interoperability.

4.6.2. Denial of User Intervention

Denial of user intervention occurs when the data subject is restrained from excercising its rights of access, correction, objection, deletion, or blocking.

Threats: The data processor might impose technical and/or organizational obstacles to restrain data subjects from excercising their rights. Technical obstacles include lack of access to all information related to the individual, excessive retention periods, or incomprehensible technical systems to be used by the data subject to excercise its rights. Organizational obstacles include the necessity to use manual measures like handwritten letters or to follow laborious procedures to access, correct, object, delete, or block personal data.

Mitigations: To assure intervenability, the data processor provides open service interfaces to be used by data subject to excercise their

rights of access, correction, objection, deletion, or blocking. If the data processor is not able to provide such interfaces, it supports the data controller to provide such interfaces. The data processor also provides a single point of contact for data subjects to excercise their rights, or supports the data controller to provide such a contact. The data processor informs the data controller about its technical and organizational measures to assure or support intervenability. The deletion of personal data SHOULD NOT be more complicated for the data subjects and the data controller than the provision of personal data.

4.6.3. Retention

Retention occurs when information is stored longer than necessary, or after the data subject or the data controller has deleted it. Retention also occurs when original information is stored after it was changed on behalf of the data subject or the data controller.

Threats: To retain personal data longer than necessary or authorized increases the risks of security breaches or violations of privacy and data protection rights of individuals. The data processor might retain personal data for economic reasons and need-to-know concerns. Information stored on backups might not be deleted to save costs.

Mitigations: To assure intervenability, the data processor ensures that no information is retained after the data subject or the data controller has deleted it, and that no original information is retained after it was changed on behalf of the data subject or the data controller. If the data processor is not able to ensure this, the data processor informs the data controller about the maximum retention period and associated protection measures in advance.

5. Acknowledgements

Here Be Dragons.

6. IANA Considerations

This document does not require actions by IANA.

7. Security Considerations

This document describes privacy and data protection measures that cloud service providers should consider and provide in addition to regular security measures. If privacy and data protection requirements conflict with security requirements, cloud service providers will have to weight the competing interests and make a balanced decision. Both the weighting and the decision MUST be recorded in writing in a form to be usable in court.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[Art29] Article 29 Data Protection Working Party, "Opinion 05/2012 on Cloud Computing", 2012.

[FIP] U.S. Department of Health, Education, and Welfare, "Records, Computers, and the Rights of Citizens", 1973.

[Hansen] Hansen, M., "Top 10 Mistakes in System Design from a Privacy Perspective and Privacy Protection Goals", Privacy and Identity Management for Life Vol. 375, pp. 14-31, 2012.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", RFC 4949, August 2007.

[RostPfitzmann] Rost, M. and A. Pfitzmann, "Datenschutz-Schutzziele -- revisited", Datenschutz und Datensicherheit Vol. 33, No. 6, pp. 353-358, 2009.

[Solove] Solove, D., "A Taxonomy of Privacy", University of Pennsylvania Law Review Vol. 154, No. 3, pp. 477-560, 2006.

[Steinmueller] Steinmueller, W., Lutterbeck, B., Mallmann, C., Harbort, U., Kolb, G., and J. Schneider, "Grundfragen des Datenschutzes", BT-Drs. VI/3826 Appendix 1, 1971.

Author's Address

Joerg Pohle
Humboldt Institut fuer Internet und Gesellschaft
Berlin
Germany

Phone: +49 30 2093 3490
Email: joerg.pohle@hiig.de