Privacy and Trust in Cloud-Based Marketplaces for AI and Data Resources

Vida Ahmadi Mehri(✉) and Kurt Tutschku

Blekinge Institute of Technology, Karlskrona, Sweden
{vida.ahmadi.mehri,kurt.tutschku}@bth.se

Abstract. The processing of the huge amounts of information from the Internet of Things (IoT) has become challenging. Artificial Intelligence (AI) techniques have been developed to handle this task efficiently. However, they require annotated data sets for training, while manual preprocessing of the data sets is costly. The H2020 project "Bonseyes" has suggested a "Market Place for AI", where the stakeholders can engage trustfully in business around AI resources and data sets. The MP permits trading of resources that have high privacy requirements (e.g. data sets containing patient medical information) as well as ones with low requirements (e.g. fuel consumption of cars) for the sake of its generality. In this abstract we review trust and privacy definitions and provide a first requirement analysis for them with regards to Cloud-based Market Places (CMPs). The comparison of definitions and requirements allows for the identification of the research gap that will be addressed by the main authors PhD project.

Keywords: Privacy · Trust · Marketplace · IoT · Cloud · AI

1 A Market Place for Artificial Intelligence and Data

Bonseyes' Market Place (MP) for AI [1–4] aims at engaging the various stakeholders, e.g. data providers, model, or application designer, into business among AI resources, i.e. data sets, models, training facilities, etc. The business around the resources may accelerate the model design and reduces the design costs. The MP will provide functions to offer, sell, pay or use AI resources and data sets. The proposed MP will be implemented by a cloud system in order to deal with the large size of data sets and to permit elasticity for the AI resources. This led to the notion of a CMP. As any MP, a CMP requires mechanisms to enforce *privacy* and *trust*. However, the separation between resource location (e.g., storage location) and resource availability (e.g. data availability) in cloud systems makes it more challenging to implement trustful mechanisms for these features as in non-virtualised systems.

J.-P. Steghöfer and B. Esfandiari (Eds.): IFIPTM 2017, IFIP AICT 505, pp. 223–225, 2017.
DOI: 10.1007/978-3-319-59171-1

2 Trust and Privacy Definitions for Network and Clouds

Trust and Trust Dimensions: a widely agreed definition for *trust* in networks, Clouds and systems is given in IETF Internet security glossary: as "... the extent to which someone who relies on a system can have confidence that the system meets its specifications, i.e., that the system does what it claims to do and does not perform unwanted functions" [5]. The view of applications, Clouds and networks as "systems" leads to the definition of multiple *trust dimensions* [6]. These dimensions comprise (a) "device trust", i.e. the reliability of IoT devices to produce data correctly, (b) "operation trust", refers to the combination of data traceability and analytics, (c) "communication trust", builds on confidentiality, integrity, and authenticity in data transmission, (d) "infrastructure trust", which aims at the transparency and predictability of processing.

Privacy and Privacy Dimensions: the IETF glossary also provides a definition for *privacy*: "... the right of an entity to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share information about itself with others" [5]. R.S. Poore defines privacy as a required context of personal Identifiable Information (PII) which have to be under control of the individual person who is the owner of it [7].This view on privacy leads to *privacy dimensions* such as, cf. [6, 8]:

- Identity privacy: avoid the disclosure of users identity
- Location privacy: avoid the disclosure location information for specific user
- Device privacy: avoid the disclosure of device and security information
- Communication privacy: refers to encryption algorithm for confidentiality
- Access privacy: privilege levels for authorised data access
- Operation privacy: avoid the disclosure of data processing techniques
- Footprint privacy: avoid the identity disclosure by behavioural analysis
- Query privacy: avoid the identity disclosure by analysis of the origin of queries

3 Privacy and Trust Requirements for CMPs

In general, the privacy levels for an AI resource or a data set depend on the importance of the resource or of the type of PII stored in it. For example, medical records need very high levels of protection since a leakage of information may embarrass a specific person. Hence, if such data sets are traded then the specific levels of privacy needs to be maintained at the various locations where the data is accessed or processed, otherwise the users will lose their *trust* into the MP.

CMPs might host data sets with very different privacy levels at very different locations. As a result, they must enable a differentiated, transparent and even traceable handling of data. Some data sets may not be allowed to leave a certain physical premise due to regulation or provider policy, while others can do so. A CMP must support both modes of data availability at the required privacy levels for the sake of its generality. This feature is often denoted as the ability for "privacy by design" [9] of an architecture.

Since data sets are associated in a MP with a value, the infringement of this value by disclosing the data to other users needs to be avoided. Here, the privacy requirements, as seen from the data provider, turn into the problem of "Digital Right Management" that needs to be addressed by the MP architecture or its mechanisms and functions.

4 Conclusion

It is obvious that the definitions of trust and privacy do not directly address the virtualisation features of Cloud system. Particularly, the implications by separating between storage location and data availability in the Cloud are not clear yet. The privacy and trust dimensions might partly match with the application requirements of CMPs. Hence, their suitability for evaluation Cloud mechanisms needs to be investigated. These investigations as well as the design of mechanism for privacy and trust in CMPs will define the work in this PhD project.

Acknowledgment. This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 732204 (Bonseyes). This work is supported by the Swiss State Secretariat for Education Research and Innovation (SERI) under contract number 16.0159. The opinions expressed and arguments employed herein do not necessarily reflect the official views of these funding bodies.

References

1. BONSEYES - Artificial Intelligence Marketplace. https://www.bonseyes.com/
2. Bonseyes Consortium. Grant agreement number - 732204 bonseyes - annex 1 (part a): Description of action. Available on request from Bonseyes Consortium at https://www.bonseyes.com/, October 2016
3. Fricker, S., Maksimov, Y.: Pricing of data products in data marketplaces. In: 8th International Conference on Software Business (ICSOB) (2017, in submitted)
4. Llewellynn, T., Milagro, M., Deniz, O., Fricker, S., Storkey, A., Pazos, N., Velikic, G., Dahyot, R., Koller, S., Goumas, G., Leitner, P., Dasika, G., Wang, L.: Bonseyes: platform for open development of systems of artificial intelligence. In: ACM International Conference on Computing Frontiers (2017, in submitted)
5. RFC 2828. Internet security glossary, May 2000
6. Daubert, J., Wiesmaier, A., Kikiras, P.: A view on privacy and trust in iot. In: 2015 IEEE International Conference on Communication Workshop (ICCW), pp. 2665–2670, June 2015
7. Poore, R.S.: Anonymity, Privacy, and Trust. Inf. Syst. Secur. **8**(3), 16–20 (1999)
8. Cheng, Y., Naslund, M., Selander, G., Fogelström, E.: Privacy in machine-to-machine communications a state-of-the-art survey. In: 2012 IEEE International Conference on Communication Systems (ICCS), pp. 75–79, November 2012
9. Article 29 data protection working party and working party on police and justice, the future of privacy. Joint contribution to the Consultation of European Commission on the legal framework for the fundamental right to protection of personal data, WP168, December 2009