Department: Internet of Things, People, and Processes
Editor: Schahram Dustdar, dustdar@dsg.tuwien.ac.at

Architectural Considerations for Privacy on the Edge

Christos Tsigkanos, Cosmin Avasalcai, and Schahram Dustdar
TU Vienna

***Abstract*—Novel pervasive systems integrate technologies and paradigms, such as mobile and cloud computing and Internet of Things, where systems are composed of heterogeneous infrastructures and services. Privacy emerges as a first-class design goal throughout such systems' development lifecycle, and suggests its management to occur architecturally at the network edge, closer to end-users as the privacy stakeholders. We discuss concerns emerging from privacy requirements and how they pertain to contemporary pervasive systems, and we distill architectural considerations highlighting privacy protection mechanisms and tactics for edge computing.**

"And about whatever I may see or hear in treatment, in the life of human beings—things that should not ever be blurted out outside—I will remain silent, holding such things to be unutterable;" Article 8 of the Hippocratic Oath provides a strong metaphor for engineering privacy-aware—by design and by default—systems.[1] Hippocrates talk about treatment of possibly sensitive medical information by a healthcare provider. Current more than ever, Hippocrates provides us the foundation of privacy-aware data management: a system may collect, use for some intended purpose, but not misuse or disclose private information. Such an ancient principle is particularly relevant in the increasingly integrated and pervasive computing environments of today, where mobile, cloud, and Internet of Things (IoT) converge inducing systems that collect, process, and disseminate information, which may be sensitive. Traditionally, organizations have been viewed as trusted custodians of information; however, data breaches, misuse, or malicious use of the sensitive information can harm privacy of the individuals.

Especially relevant in today's integrated world, comprehensive privacy mechanisms are essential

Digital Object Identifier 10.1109/MIC.2019.2935800
Date of current version 7 October 2019.

Published by the IEEE Computer Society

for the widespread uptake and acceptance of the systems we engineer, as the ever-increasing number of devices collecting (possibly sensitive) data and interacting with the physical environment poses privacy risks. Regularly acknowledged in contemporary legal and regulatory frameworks, privacy emerges as a first-class design goal throughout an application's development lifecycle. Pervasive systems are particularly sensitive to this; resource-constrained IoT devices, cloud offloading, and generally heterogeneous software components operate within diverse administrative domains.[2,3] At the same time, the volume of the data generated by devices close to end-users grows exponentially. Resource-constrained IoT devices limit the techniques that can be used to deliver efficient and effective privacy-preserving schemes.

We argue that the drive toward edge computing can help pervasive systems engineering tackle privacy threats. We elaborate the role of edge computing, by outlining concerns emerging from privacy requirements and how they pertain to contemporary pervasive systems and discuss software architecture considerations to meet privacy needs. The edge computing paradigm can help preserve privacy and protect users: first, by establishing privacy controls at a layer close to data-producing end-users and subjects, second, by minimizing the need to transmit sensitive data to the cloud for analysis, and third, by offering opportunities for stronger privacy with respect to the data collection and identification through an edge-centric anonymization. Since user-facing software components are the ones that generate and act on sensitive data, empowering the edge appears to be a reasonable decision. However, considerable engineering challenges arise to support this. We make the case that the decentralization inherent in the edge computing paradigm yields significant benefits for privacy.

In the following, we start from a societal perception of privacy—how privacy has come to be treated in legal terms[4]—and we adapt such concepts for contemporary pervasive and ubiquitous systems. Subsequently, we discuss architectural considerations that highlight privacy protection mechanisms and tactics for edge computing.

PRIVACY CONCERNS IN PERVASIVE SYSTEMS

Privacy concerns that we identify are based on a loose interpretation of the legal privacy taxonomical framework designed by Solove,[4] and adapted for contemporary ubiquitous systems; we treat them as building blocks of privacy requirements. Motivated by architectural considerations, we, however, propose a different grouping and ignore ones that are software-architecture-irrelevant (e.g., interrogation as an activity that may violate privacy). We follow the data lineage within IoT systems to discuss privacy concepts; data are collected, processed, or analyzed in edge or cloud computers and possibly disseminated, closing the loop with a possible control flow back to user-facing devices. In the following, for each privacy concern we identify, we discuss challenges, opportunities, and emerging solutions within the edge-based systems.

Data Collection and Identification

Devices that collect, store, and send information from various surrounding sources are ubiquitous in modern environments and an integral part of the IoT paradigm. Information collection—even if no information is revealed publicly—can be problematic with respect to privacy laws and regulations. Requirements may require not only data within an application to remain locally close to where it is sourced, but also for all mechanisms managing it to respect different legal or administrative frameworks (e.g., the EU General Data Protection Regulation versus the California Consumer Privacy Act) and user preferences. A further challenge is that the data often traverses through computational resources of diverse domains; enabling applications to operate across them points to another facet where data producers within an edge-based system require control over data exchange. Furthermore, collected data may be used to identify individuals. To this end, anonymization methods have been developed to counteract the possibility of linking attacks and achieving *k*-anonymity over data.[5] While such processes are well understood for the static datasets, they become challenging on dynamic and streaming data typically found in contemporary ubiquitous and IoT applications.

We advocate that because the edge is closer to data sources and users, there is not only obvious latency advantages, but also an opportunity for stronger privacy with respect to data collection and identification. This includes anonymization as well as operation within administrative domains; both suggest building appropriate data handling logic inside the edge-based software components. To counter identification threats while collecting data, anonymization can occur at the edge before transmission to the cloud. Anonymization facilities[6] need to be developed for streaming data and perhaps able to be deployed on resource-constrained edge devices, e.g., within a user's home. Finally, the risk of (re-)identification of anonymized data with machine learning (ML) methods spotting patterns needs to be acknowledged; possible mitigations can include use of diversification techniques or feeding of fake transactions.

Aggregation and Inference

Information processing in the context of privacy, involves various ways of combining data together and linking it to individual people to whom it relates. Vast amounts of personal data about individuals stored at different commercial vendors and organizations is the norm, which in combination can pose privacy risks. Furthermore, we identify information aggregation and inference as particularly relevant privacy aspects due to the widespread adoption of data-producing IoT devices and the increasing technological advances in (and need for) inference using artificial intelligence (AI) and ML methods. We treat aggregation and inference activities as similar from a technological perspective, as they similarly concern multidimensional data. For example, patterns within smart meter readings processed by an energy operator for analytics or energy efficiency purposes can reveal the occupancy of a residence.

Complex processing and inference is typically performed on the cloud and is dominated by training deep ML models requiring heavy compute capacity. The recent trend is on moving the inference part of the AI workflow close to end-devices. This may be desirable for nonfunctional requirements like security, cost or latency, but can have a positive impact on privacy as well, as user data are kept constrained to an edge device. Novel approaches, such as federated learning[7]—where user-facing devices learn a shared prediction model in a collaborative manner while keeping all the training data on the device—can hinder aggregation and inference attacks that presuppose centralized, organizationally-curated data repositories usually on the cloud.

Secondary Use, Insecurity and Exclusion

Following the data lineage within IoT systems, after data have been sourced from mobile devices, sensors, or users within some organization, privacy issues arise with use, storage, and manipulation of collected data. This privacy aspect concerns issues arising from an organization's maintenance and use of collected data. Insecurity—from a legal perspective—involves carelessness in protecting stored information from leaks and improper access. Secondary use concerns information collected for one purpose, used for a different purpose without the data subject's consent. This reflects a common principle in information privacy, where for all data collected and processed, there should be a stated purpose; usage of data for another purpose than the one it was intended for must be prevented. Exclusion concerns the failure to allow the data subject to know about the data that others have about her and participate in its handling and use. Such aspects have been reflected also in recent regulatory frameworks, such as EU/GDPR (Art. 6 – Lawfulness of processing), making compliance mandatory. The typical example lies within the healthcare domain, where there is an organic abundance of sensitive data collection for diagnostic or other medical purposes. First, data may be improperly stored leading to data leaks. Secondly, they may be shared with third parties for some other originally unintended use (such as insurance companies or research institutions). Finally, fulfilling requests by a subject (e.g., a concerned patient) for what data have been collected and with whom it was shared may be difficult due to improper data handling processes.

Information privacy research has long developed privacy models and frameworks to ensure compliance. P-RBAC[8] is able to capture roles and permissions, actions on data, conditions, and obligations that arise in privacy requirements, whereas Contextual Integrity's model of

communicating agents[9] shows suitability for streaming data. However, integration in engineering processes and architectures within privacy-compliant ubiquitous systems have to be investigated. In edge computing architectures, with user-facing devices handling (possibly sensitive) data and interacting with the physical environment, the edge device is by definition located within the administrative domain of its local IoT devices—one can take that as the devices being in the same privacy scope. Thus, the edge can be treated as a first-class entity regarding privacy and data management, and can ensure that the data flows between the edge and other external components (i.e., other edge nodes, the cloud etc) always respect defined privacy policies in the system.[10]

Decision and Boundaries

Privacy does not always involve information. Harms may come from invasive acts that disturb an individual's personal boundaries and tranquility. A manifestation of this are the legal protections of the privacy of the home, protecting it from trespass and external nuisances (for hundreds of years).[4] Invasive acts may harm privacy, and—to borrow from cyber-physical systems terminology—are a form *actuation*. The interplay between computational and physical aspects must be additionally considered.[11] For example, a virtual assistant making actuation decisions, such as enabling indoor surveillance cameras when occupants are home, can violate privacy boundaries of subjects.

The drive to decentralization points to putting trust and security in the hands of users. Edge computers operated by and within reach of users (e.g., virtual assistants in their home), and potentially invasive (e.g., in charge of controlling window blinds or cameras) should operate transparently and in an accountable manner. Putting security in the hands of users can be a double-edged sword, but empowering users to make their own decisions about control actions, devices and networks they own is desirable from a privacy perspective.

Appropriation and Distortion

Traditionally, organizations have been viewed as trusted custodians of information; however, data breaches or malicious use of sensitive information can harm privacy of individuals. Generally, appropriation involves the use of the data subject's identity to serve the aims and interests of another, such as the deliberate use of someone else's personal data in context of identity theft. Distortion consists of the inappropriate dissemination of false or misleading information about individuals, for example misuse of personal data acquired by a corporation in a marketing campaign.[4]

Failures of the various data custodians, the increased prevalence of sensitive data nowadays, and the mistrust placed on institutional organizations to manage them calls for new paradigms. Security concepts, such as identity management and digital signatures can be leveraged in a decentralized manner for sensitive data security and management at the edge of networks. Personal data for instance, can be signed before reaching data stores, rendering source verification, and lineage tracking possible. Solutions made possible with blockchain technology can be further used, with validation that occurs in edge nodes outside of control of organizations, but within the control—and trust sphere—of users.

EDGE ARCHITECTURES FOR PRIVACY-PROTECTION

In edge-based systems, the edge software component is by definition (by virtue of deployment) located close or within the administrative domain of the local end-user. Software components operated by the end-user and hosted on, e.g., mobile or IoT devices interact with the edge. As such, one can take that the edge and user components are architecturally in the same privacy scope, and trust exists between them. Sensitive data flows between the edge and external components (i.e. other edge nodes, or the cloud) should always respect privacy policies, over which the user should have complete control. As such, edge components must enforce privacy policies when data are shared, processed, and collected.

To distill the previous into an architecture, we loosely follow the conventional description structure of ISO/IEC/IEEE 42010—a software architecture view, architecture description and privacy as a cross-cutting perspective. We take the concerns of the previous section to be the functional viewpoint of a privacy-protecting edge architecture,

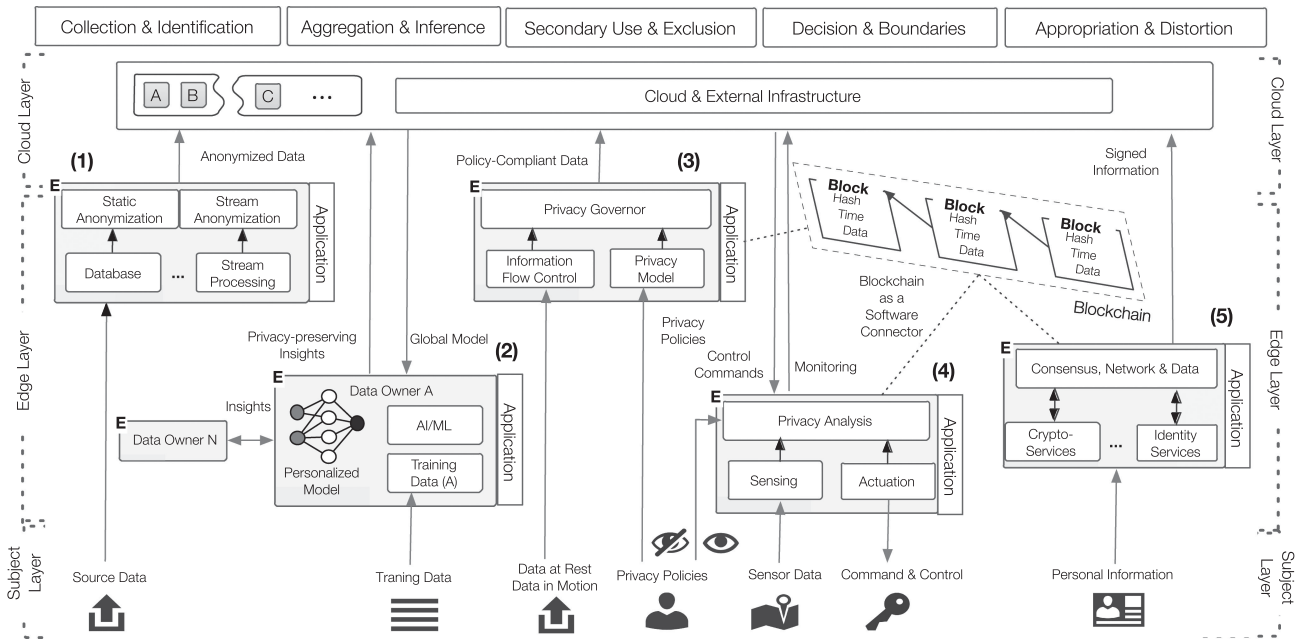

Figure 1. Combined dataflow and deployment architectural diagram showing different architectural manifestations for the edge as a first-class entity for privacy protection.

showing scenarios describing how various architectural elements and views should address privacy.

Architecture Description

A privacy-protecting platform must accommodate the decentralized nature of today's infrastructures, where sensitive data are distributed between multiple devices in the network, and apply privacy protection close to where data are collected. Such a platform may be composed of three entities: first, the subject or user that has total control of his/her private data, second, the privacy-supporting software components representing facilities used to enforce privacy measures, and third, the organizational infrastructure, typically centrally located in the cloud (see Figure 1).

The cloud layer consists of multiple data servers, located at remote (to the data sources) locations, where organizations collect data for further processing. The cloud may be in a different administrative domain compared to the origin of data—privacy controls should be enforced before data are sent to the cloud. We advocate that the role of this architectural layer is overall *privacy compliance* with requirements, laws and regulations—the burdens of specification and implementation move to lower architecture layers. The edge layer is populated with geolocated distributed edge devices that may or may not belong on the same administrative domain themselves. Such computing devices have different characteristics—from resource-constrained energy-optimizing devices to powerful gateways and edge servers where computation can be offloaded. The role of the edge layer is *privacy implementation*—measures and techniques are deployed in edge nodes as privacy-protecting software components, and act as intermediary monitoring and enforcement facilities. Finally, the subject layer concerns data sources and *privacy specification*. Data are generated by different user-facing devices as primary sources. Specification of privacy requirements and control of the edge facilities employed to satisfy them lies similarly within control of the user.

Software Architecture View

The view of edge entity internals consists of two layers: first, the application layer, fulfilling user goals and second, the edge support layer, facilitating privacy governance. The former implements business logic and is application specific. The latter has a supportive role to applications; this includes provisioning, configuration

management, data storage, and event processing, but also privacy governance facilities: anonymization, policy enforcement, and control. Such capabilities may be exposed to applications.

In Figure 1, edge components (middle layer) act as intermediaries between privacy subjects and primary data sources (lower layer), and the organizational, external infrastructure (top layer). Privacy concepts are shown within a horizontal division, with the architecture outline of the respective edge component. Within each, dataflow and typical deployment of software components is illustrated, highlighting various capabilities. Such supportive privacy components can be accessed and used by end-user applications.

Privacy Perspective

We treat the privacy perspective as a crosscutting concern, i.e., capabilities that cut across architectural layers (from privacy subjects to organizations and cloud infrastructures), and outline how privacy protection mechanisms can be employed at the edge.

Data collection and identification threats concern sensitive data occurring at the subject layer, which are used to uniquely pinpoint a subject. Anonymization techniques[5] can be used to counter privacy threats from multidimensional data, performed at the edge before data reaches an organizational cloud infrastructure [see diagram (1) of Figure 1]. *k*-Anonymity and *k*-indistinguishability have been historically employed for data-at-rest—contemporary ubiquitous systems, however, are often characterized by streaming data. We advocate development and deployment (and oversight) of such techniques at (or by) the edge node.[6]

Aggregation and inference threats are about using diverse data and advanced methods to identify subjects by combining information, typically with AI techniques. Recent developments on federated learning have shown that ML applications can be engineered in a decentralized manner, with training data not leaving the privacy sphere of the subject.[7] The evocative term *edge AI* has emerged, where deployment of AI/ML techniques occurs at edge nodes. In such a case [see diagram (2) of Figure 1], personalized models reside and are trained by subject data at nodes, whereas insights are shared with other data owners. The key idea is that insights shared can preserve privacy, as personal data do not leave the scope of the node.

Secondary use, insecurity, and exclusion threats can be tackled by using privacy models to control how, what and for which purpose data are collected, shared, and processed. Given specification of privacy policies capturing subject preferences (or privacy laws and regulations), a privacy governor can operate on incoming subject data, ensuring compliance [see diagram (3) of Figure 1]. Furthermore, if the subject has control of the edge node, privacy policies can by-design dictate how data leaves the scope.

Decision and boundaries threats can emerge by control actions that come from outside a subject's privacy scope. Control actions are usually coupled with sensing (e.g., a virtual assistant opens the window blinds when daylight is detected). Privacy analysis wrapping such actions can occur at the edge [e.g., at the subject's home network, see diagram (4) of Figure 1], forbidding or allowing command, and control based on subject's explicit preferences.

Appropriation and distortion refers to deceptive use of a subject's personal information. Cryptographic advances have enabled widespread schemes to counter such trust issues. Digital signatures can verify authenticity or source of messages or documents, and symmetrically, can also provide nonrepudiation. Such measures can be employed to validate that the sensitive information has not been altered, and that it belongs to the subject that generated it. Moreover, recent developments have shown the benefits of blockchains for trust management without central authorities or servers. A blockchain deployment in an edge-to-edge network where nodes collectively adhere to a protocol for communication and validation, can support such functions in a decentralized manner and ensure that personal information is not appropriated by others.[12] Functionality essentially entails recording transactions between subject parties or organizational entities in a verifiable and permanent manner. We suggest that the edge can host such supportive functions [see diagram (5) of Figure 1] and can make them available to end-user applications, by considering a blockchain as a software connector for identity, trust and verification facilities.

EMERGING RESEARCH AGENDA

Contemporary pervasive systems integrate multiple technologies and paradigms in systems that are composed of heterogeneous infrastructures and services. Privacy emerges as a first-class design goal throughout such systems' application development lifecycle, and suggests its management to occur architecturally at the network edge, closer to the end-devices. We discussed aspects emerging from privacy requirements and how software architecture considerations pertain to edge-enabled systems, and highlighted privacy protection mechanisms and tactics for edge computing.

As future work, we identify providing a complete reference architecture that engineers and organizations can use for documenting viewpoints as per ISO/IEC/IEEE 42010. Qualitative aspects and other nonfunctional requirements, such as performance often are in conflict with the privacy protection, as it adds processing overhead. Such design tradeoffs need to be carefully considered, as, e.g., timeliness of data or events can be critical in edge-based systems. We ignored network and communication issues, which must be treated as well in operationalizations of the privacy-protecting architectures discussed. Finally, identification and employment of specific technological options for the refinement of the architectural components discussed is highly desired.

ACKNOWLEDGMENTS

This work was supported in part by the Research Cluster "Smart Communities and Technologies (Smart CT)" at TU Vienna, and in part by the EU H2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 764785, FORA–Fog Computing for Robotics and Industrial Automation.

REFERENCES

1. K. Barker *et al.*, "A data privacy taxonomy," in *Proc. Brit. Nat. Conf. Databases*, 2009, pp. 42–54.
2. J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the internet of things: Threats and challenges," *Secur. Commun. Netw.*, vol. 7, no. 12, pp. 2728–2742, 2014.
3. C. Tsigkanos, S. Nastic, and S. Dustdar, "Towards resilient internet of things: Vision, challenges, and research roadmap," in *Proc. 39th IEEE Int. Conf. Distrib. Comput. Syst.*, Jul. 2019, pp. 1754–1764.
4. D. J. Solove, "A taxonomy of privacy," *Univ. Pennsylvania Law Rev.*, vol. 154, no. 3, 2006, Art. no. 477.
5. C. C. Aggarwal, "On *k*-anonymity and the curse of dimensionality," in *Proc. 31st Int. Conf. Very Large Databases*, 2005, pp. 901–909.
6. J. Cao, B. Carminati, E. Ferrari, and K.-L. Tan, "CASTLE: Continuously anonymizing data streams," *IEEE Trans. Dependable Secure Comput.*, vol. 8, no. 3, pp. 337–352, May/Jun. 2011.
7. K. Bonawitz *et al.*, "Towards federated learning at scale: System design," 2019, *arXiv:1902.01046*.
8. Q. Ni *et al.*, "Privacy-aware role-based access control," *ACM Trans. Inf. Syst. Secur.*, vol. 13, no. 3, 2010, Art. no. 24.
9. A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum, "Privacy and contextual integrity: Framework and applications," in *Proc. IEEE Symp. Secur. Privacy*, 2006, Art. no. 15.
10. N. Li, C. Tsigkanos, Z. Jin, S. Dustdar, Z. Hu, and C. Ghezzi, "POET: Privacy on the edge with bidirectional data transformations," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, Mar. 2019, pp. 1–10.
11. C. Tsigkanos, L. Pasquale, C. Ghezzi, and B. Nuseibeh, "On the interplay between cyber and physical spaces for adaptive security," *IEEE Trans. Dependable Sec. Comput.*, vol. 15, no. 3, pp. 466–480, May/Jun. 2018.
12. M. A. Khan and K. Salah, "Iot security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, 2018.

Christos Tsigkanos is a Researcher with the Distibuted Systems Group at TU Vienna. Previously, he was a Postdoctoral Researcher at Politecnico di Milano, where he received the Ph.D. defending a thesis entitled "Modeling and Verification of Evolving Cyber-Physical Spaces." His advisor was Prof. Carlo Ghezzi. His research interests lie in the intersection of dependable systems and formal aspects of software engineering, and include security and privacy in distributed, self-adaptive and cyber-physical systems, requirements engineering and formal verification. He received the B.Sc. degree in computer science from the University of Athens, Greece, and the M.Sc. degree in software engineering from the University of Amsterdam. Contact him at: christos.tsigkanos@tuwien.ac.at.

Cosmin Avasalcai is a Research Scientist and a Ph.D. student working under the supervision of Prof. Schahram Dustdar with the Distibuted Systems Group, TU Vienna. Before joining the Distributed Systems Group, he worked for two and a half years as a research assistant with the Technical University of Denmark (DTU), developing modeling and decision making tools for the automotive industry for mixed-criticality applications in dynamic and changeable real-time environments and the oil and gas industry for structural monitoring. His current activities include quality of service-aware resource provisioning in edge computing within the Fog Computing for Robotics and Industrial Automation European Training Network project. After finishing the bachelor studies from Technical University "Gheorghe Asachi" Iasi, Romania, in 2015, he received the master's degree in computer science and engineering from DTU in 2017. Contact him at: c.avasalcai@dsg.tuwien.ac.at.

Schahram Dustdar is a Professor of Computer Science with the Distributed Systems Group, TU Vienna, Austria, and an IEEE Fellow. He was an Honorary Professor of Information Systems with the University of Groningen, The Netherlands from 2004–2010, a Visiting Professor with the University of Sevilla, Spain from 2016–2017, and a Visiting Professor with the University of California at Berkeley, USA in 2017. He is an elected member of the Academia Europaea, where he is the Chairman of the Informatics Section. He was the recipient of the ACM Distinguished Scientist Award in 2009, the IBM Faculty Award in 2012, and the IEEE TCSVC Outstanding Leadership Award for outstanding leadership in services computing in 2018. He is the Co-Editor-in-Chief for the ACM *Transactions on Internet of Things* and the Editor-in-Chief for *Computing* (Springer). He is also an Associate Editor for the IEEE Transactions on Services Computing, the IEEE Transactions on Cloud Computing, the ACM *Transactions on the Web*, and the ACM *Transactions on Internet Technology*. He is on the Editorial Board of *IEEE Internet Computing* and the *Computer Magazine*. Contact him at: dustdar@dsg.tuwien.ac.at.