A Toolchain and Interoperability Framework to enhance privacy and individual control at the Edge

Panagiotis Katrakazas
Zelus P.C.
Athens, Greece
ORCID ID: 0000-0001-7433-786X

Theodora Kallipolitou
Zelus P.C.
Athens, Greece
ORCID ID: 0000-0001-5059-4909

Stella Markopoulou
Zelus P.C.
Athens, Greece
ORCID ID: 0000-0001-9819-3143

Argyro Chronopoulou
Zelus P.C.
Athens, Greece
a.chronopoulou@zelus.gr

***Abstract*— Internet-of-Things (IoT) has the potential to create new services and markets by allowing the exploration of new, often completely different ways of doing things, based on the clustering and aggregation of data from different sources and fields of activity. As technology advances, new ethical, legal, and technological concerns arise. In this paper, we present five key pillars of innovation towards privacy-preserving edge computing, regarding smart sampling of IoT devices, anonymous authentication and consent management, dynamic data-driven pattern management, opportunistic IoT clustering, distributed IoT data governance, and resource integrity validation.**

The overall concept of this paper, is to create a comprehensive methodological framework and toolset for definition, deployment and operation of privacy-compliant IoT platforms tailored to specific use-cases. During this process we are not creating "yet another IoT platform", but rather building upon past efforts to the maximum extent possible. This approach takes into consideration existing solutions in the following areas: high-level concepts and standards; integration and interoperability frameworks; IoT platforms and infrastructural elements.

Keywords*— *consent management; opportunistic clustering; interoperability; smart behaviour; data veracity

2022 IEEE International Smart Cities Conference (ISC2) | 978-1-6654-8561-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISC255366.2022.9922309

I. Introduction

With the emergence of the Internet-of-Things (IoT) and the growing ubiquity of sensors in our day-to-day lives, there are increasing demands both from individuals and companies to improve:i) the richness of information that can be extracted (sensed), ii) the kind of actions that can be taken (actuated), and iii) the value that can be realized from available devices [1]. As IoT interoperability continues to improve, the ability for emerging IoT applications to extend beyond their originally limited "cluster" of devices to others on the same network (e.g. in-car sensors and personal wearables) becomes both a significant opportunity for disruptive innovation and a significant risk to individuals and companies alike. However, at the same time, it raises significant questions about data "ownership", ethics and liability for semi-autonomous decision making among others [2], [3].

A further shift from pure sensing to actuation is also becoming mainstream, which especially at a pivotal point of technological immaturity, underpins the importance of auditability in decision-support [4]. Given the high degree of interoperability and the creation of new cross-cluster services, the question of liability remains an open one. While policy efforts to clarify ownership of sensor data have been ongoing, liability issues can easily be envisioned across three levels: 1) the data source (e.g. IoT device) itself, 2) integration across data sources and 3) impact from a resulting actuation event which requires attention to be put on ensuring data accuracy and integrity [5]. While impact from an actuation event is perhaps the most visible output (e.g. in the form of a vehicular collision), the ability to verify why an actuation event went wrong and at what point along the data processing and value chain errors were introduced can be seen to not only making meaningful technical and research contributions to the state of the art in distributed IoT data governance but also as a key enabler and market qualifier for organisations aiming to provide value-added services that involve actuation across heterogeneous IoT devices in a dynamic regulatory environment with an unclear IoT liability regime.

At the same time, changes in the European regulatory environment, such as the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) and NIS directive. have begun to place individuals in a greater position of power over their data. This necessitates a re-think of the traditional data value chain, and facilitates a move towards more inclusive business models in which individual trust and consent must be carefully managed across the data life-cycle to realize new opportunities for personalized service delivery and value co-creation [6].

Furthermore, as the number of smart things continues to increase and the amount of data produced keeps on growing, intelligent data processing and control is no longer something that only happens on the back-end in faraway data centres through the cloud, but increasingly must be handled at the edge (fog computing) [7]. While there has been a great deal of emphasis on increasing accountability, transparency, and security for the handling of personal data in the cloud, little emphasis has been placed on what happens at the network edge (in the fog), where 'edge' and 'fog' are considered as one layer in the case of computational devices [8].

In this paper, we propose a framework to bridge this gap, on one hand by enabling intelligence to be applied at the Edge (bottom-up), while being able to benefit from the existing software ecosystem(s) and standardisation on the other (top-down).

II. Concept and Methodology

A. Concept

The overall concept of the work presented here, depicted in Figure 1, is to create a comprehensive methodological framework and toolset for definition, deployment and operation of privacy-compliant IoT platforms tailored to specific use-cases. During this process, we will not be creating "yet another IoT platform", but rather building upon past efforts to the maximum extent possible. This approach takes

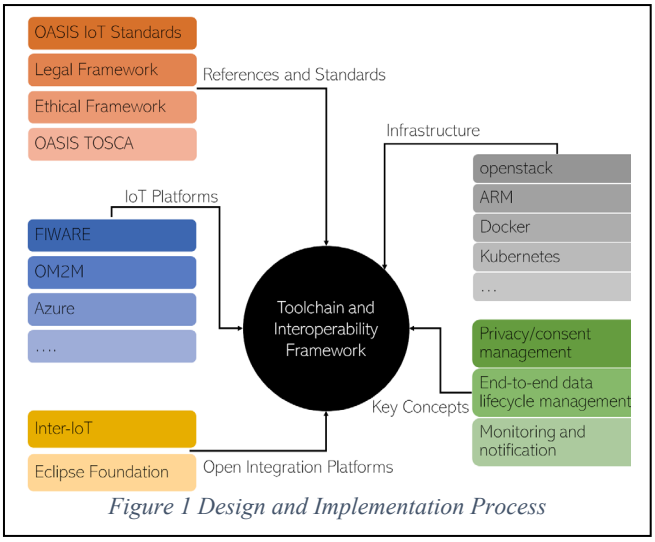

Figure 1 Design and Implementation Process

into consideration existing solutions in the following areas: high-level concepts and standards; integration and interoperability frameworks; IoT platforms and infrastructural elements.

1) References, Standards and methodologies: Approaches like those developed by OASIS (link: https://www.oasis-open.org/, accessed on: June 22, 2022), where the aim is developing software and performance engineering methods for Big Data through model transformations, can be expressed in OASIS TOSCA application blueprints. This approach will be applied to the IoT domain, and specifically to the formal definition of security and consent management policies.

2) Open Integration Platforms: Past projects like Inter-IoT (link: https://inter-iot.eu/, accessed on: June 22, 2022), as well as communities, like the Eclipse Foundation (link: https://www.eclipse.org/, accessed on June 22, 2022) and the Oniro Project (link: https://oniroproject.org/, accessed on June 22, 2022), which enable technical interoperability of IoT platforms at various levels (application, middleware, network, gateway, and device). They also provide bridges for several IoT middleware (like FIWARE, OM2M, SOFIA2), while their main benefit is an open methodology, development framework and a tool for integration of any IoT platform.

3) IoT Platforms: Many EU research initiatives, international academic research outputs and industrial developments created a highly competitive ecosystem of IoT platforms. Driving forces behind them are very different: purely academic proof-of-concept, interoperability of devices or solution of specific application domain problems. Furthermore, the "platform" concept does not have a well-accepted definition. It may refer to a cloud infrastructure, application, middleware or gateway level. It may or may not infer that the platform has been implemented with IoT specifics taken into account. In the context of our framework an "IoT platform" is a specific solution covering part of the IoT stack. The framework we suggest is not about building a new IoT platform to solve a specific problem of preserving privacy and consent management, but rather promote the usage of IoT integration platforms.

4) Infrastructure components: From an infrastructural point of view, modern IoT solutions are required to overcome the classical centralized cloud model where data collected remotely is transferred in traditional data centres in order to be elaborated: requirements related with low latency, low bandwidth usage, high responsiveness, high scalability, guaranteed quality of service (QoS), and data privacy control need the support of computational, storage and network capacity where data is produced, outside the data centres (i.e. proximity cloud). Moreover leveraging cloud computing concepts, like scalability, optimal resource allocation, allows the IoT domain to become more flexible, adopting typical paradigms of Cloud Computing like virtualization of resources and functions, elasticity, self-provisioning and DevOps. The suggested framework intends to use existing open source components wherever possible as the basis for delivering its architectures and methodology, as highlighted in (1) and (2).

5) Key Concepts: We define three key concepts of technical and ethical nature, dealing with Privacy/Consent Management, End-to-end data lifecycle management and monitoring and notification:

- **Privacy/Consent Management**: This is the main innovation in the technical sense. This component will manage and oversee privacy and consent management of pieces of data through the whole lifecycle of the data and across all layers. It will incorporate all the main concepts defined by the ethical and legislative constraints (including dynamic consent and revocation, management of privacy-preserving dynamical sampling of health signals and similar concepts).
- **End to end data lifecycle management**: With cloud services being deployed to the edge (top-down) and end-users expressing their own data usage limitations, privacy preferences, and consent (bottom-up), the data life cycle must be considered in an environment of heterogeneous data sources and a distributed data processing architecture involving multiple actors along the data value chain. This component will specifically focus on the extension of data life-cycle management across heterogenous data sources, and ways in which data governance can be applied across the overall distributed data processing architecture in the IoT.
- **Monitoring and Notification**: The monitoring and notification component provides the technological basis for monitoring sensor values and taking action (whether through reconfiguration of sensors, directly driving actuators, or simply raising an external notification to trigger other components). Monitoring and notification will similarly be exposed at all layers of the architecture, from the Things level through to the cloud, providing the basis for holistic audit and logging of data flows across the whole stack.

B. Methodology

A crucial goal of the Toolchain and Interoperability Framework (TIF) is the development of a methodology and a set of tools to assess the security level of any IoT platform component. This would enable us to compare components among different providers: "Thermometer A has got a higher security level of B", or "middleware component A has a higher security than B". Using this methodology, we will carry on an extensive assessment of existing IoT platforms and devices. The result of this effort is trifold: (1) raising user awareness about IoT security; (2) testing of the methodology and tools in practice; (3) selection of best-quality components for the reference framework implementation. We thus create a formal process to assess, select combine and reuse existing IoT platform components.

Building on top of existing standards, integration and IoT platforms, the framework proposed here is not intended to provide a new standard, nor a new platform. Its purpose would be to provide a reference IoT platform architecture with an emphasis on security and user consent. This reference platform deployment will consist of selected modules from already existing platforms. The outcome is a handbook and a set of tools to integrate systems.

III. Conceptual Architecture

The application of the TIF concept, defined in the previous section, will result in the specification and deployment of specific reference platforms tailored to a particular application domain. This physical architecture, shown in Figure 2, consists of three or more horizontal layers, traditionally defined as cloud, edge, gateways and things. In addition to horizontal layers that are responsible for processing data at pre-designated abstraction and complexity levels, we define three vertical pillars, resulting from the key concepts defined in the previous section, responsible for Privacy/Consent Management, End to end data lifecycle management and Monitoring and Notification. In the following sections, we elaborate the provisional architecture of all layers and pillars defined above.

A. Cloud to fog management and opportunistic clustering

The physical architecture of the TIF framework is based on a distributed and heterogeneous cloud computing environment inspired by the fog computing approach where fog nodes are geographically distributed at the edge of the network (i.e. where a swarm of devices and, more generally, things are located). Figure 3 gives a representation of the physical architecture.

- Cloud tier: this is the traditional Cloud computing environment and can be implemented with the usage of a public cloud offer (e.g. AWS, Azure, GCP) or with the setup of a private cloud (e.g. based on OpenStack). The IaaS layer is complemented by containerization technologies (e.g. Docker and Kubernetes).
- Edge tier:
  - Cloudlets tier: this is a sort of mini-cloud or Cloudlet (e.g. based on x86-64 processors), located in the proximity of the IoT devices, offering computational and storage capacity, offering a plethora of benefits [9]. It can be implemented using an Infrastructure-as-a-Service (IaaS) layer above the bare metal (e.g. OpenStack) together with containerization and orchestration technologies (e.g. Docker and Kubernetes);
  - Gateways tier: this is the set of IoT gateways (e.g. based on ARM processors in case of Raspberry Pi) directly connected with the devices/sensors. It is implemented using containerization and orchestration

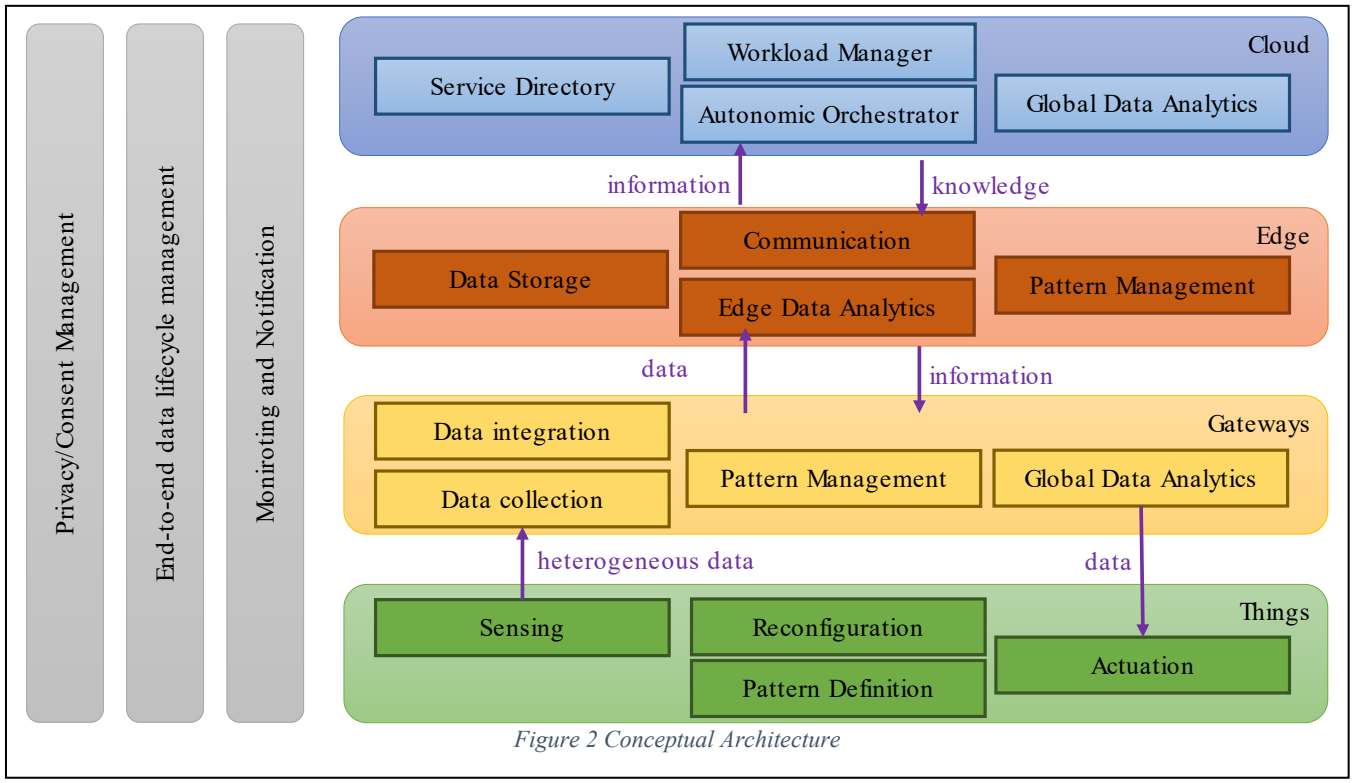

Figure 2 Conceptual Architecture

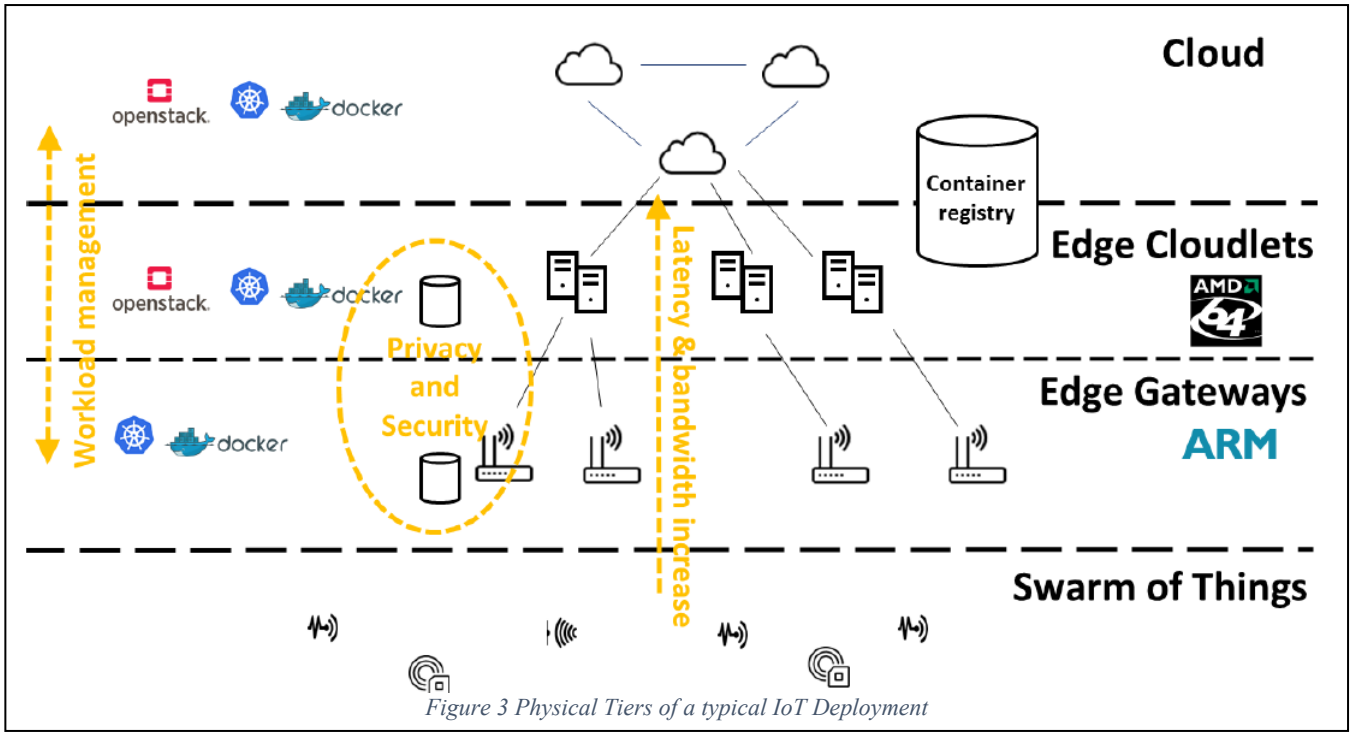

Figure 3 Physical Tiers of a typical IoT Deployment

technologies (e.g. Docker and Kubernetes). In specific cases, when the computational and storage capacity offered by the sole Gateways tier is enough for supporting the load of the services requested by a specific instantiation of the TIF reference implementation, this tier can be merged with the Cloudlets tier.

- Things tier: this tier represents the swarm of devices, sensors, and actuators that collect data and execute actions on the real environment. They are directly connected to the Edge tier (Gateways tier).

Disparate services and functionalities are offered by the different tiers as shown in Figure 3. Each of them is implemented following the paradigm of microservices leveraging separation of concerns and using the containerization as the enabling technology. Containerization, or operating system virtualization, provides a number of advantages with respect to traditional virtualization offered by Virtual Machines (VM) (i.e. containers are lightweight, faster, more efficient and highly portable with respect to virtual machines), is ideal to be exploited in a microservices architectures and offers the possibility to materialize DevOps targeting innovative approaches as the ones foreseen in Immutable Infrastructure Model [10], Infrastructure as Code [11] and Pets vs Cattle metaphors (link: https://traefik.io/blog/pets-vs-cattle-the-future-of-kubernetes-in-2022/, accessed on: 22 June 2022) that are keys in contexts where the fleet management of thousands of devices must be addressed.

TIF can use Docker-based technologies for handling containers together with a compatible container orchestrator for managing containers in distributed, heterogeneous and multi-clustered environment. The main goal of a container orchestrator is to take care of the deployment, placement, scaling, and health check of application containers, governing their life cycle, offering services for covering the main infrastructural functionalities (e.g. load balancer, reverse proxy, Domain Name System (DNS), persistent storage and container to container communications) and finally easing operations of clusters hosting a high number of containers. Among the available container orchestrators (e.g. Mesos, Docker Swarm, Kubernetes), selection will be based on the one that best fits with the objectives and customization needs.

At the moment, given the type of workload foreseen in relevant use cases, the relevant adoption and the vibrant community behind it, Kubernetes appears to be the preferred choice [12]. Deploying Kubernetes (but also other container orchestrators) conceived for traditional data centres, one of the issues is how to deploy it in a distributed and decentralised environment. Essentially Kubernetes is composed of a master node and many worker nodes, each managed by an agent called kubelet; the deployment and distribution of Kubernetes master and its relation with Kubernetes workers depends on the architecture (centralized vs decentralized) adopted (link: https://kubernetes.io/docs/concepts/overview/components/, accessed on: June 22, 2022). Traditionally Kubernetes is deployed in a centralized way inside a data centre but also a decentralized deployment (with many masters, each one controlling its own worker nodes) is possible even though not yet mainstream (e.g. Kubernetes Cluster Federation in [13]).

Such a decentralised deployment is indeed requested for TIF in order to make the whole system fault tolerant and resilient. To make the TIF infrastructural easy to deploy, maintain, evolve, monitor, and to lower the complexity of operating such a framework, a set of tools will be developed for the automatic deployment of the infrastructural components, for adding new edge nodes, for re-configuring existing ones and for monitor the whole system:

- Modelling tools like juju (link: https://jaas.ai/, accessed on: June 22,2022) or automated configuration tools like Ansible (link: https://www.ansible.com/, accessed on: June 22,2022) , Chef Infra (link: https://www.chef.io/products/chef-infra, accessed on: June 22,2022) or Puppet (link:

https://puppet.com/, accessed on: June 22,2022) can help to simplify the deployment of the infrastructural cluster and to keep it up to date.

- Monitoring tools like Elastic Stack (link: https://www.elastic.co/products/, accessed on: June 22,2022) , OpenStack Telemetry (link: https://wiki.openstack.org/wiki/Telemetry, accessed on: June 22,2022) or Zenoss (link: https://www.zenoss.com/, accessed on: June 22,2022) can be used or customised to keep the system healthy.

IV. Discussion

A. Progress in Cloud/Edge Computing applied to IoT

Cloud Computing and IoT are two well-known paradigms apparently belonging in two different architectural models: one highly centralised and another highly distributed. Nevertheless the massive diffusion of devices and components (e.g. sensors, wearable devices), that can generate a huge amount of data, is incentivizing their integration in order to take advantage of the two different yet complementary approaches inspiring the respective technologies: on one side IoT can benefit from the ideally infinite resources offered by the cloud, its scalability, flexibility and elasticity; on the other side cloud computing can offer its services outside traditional data centres, spreading the real word and opening new scenarios able to exploit its resources [14].

However, traditional cloud computing, where resources are stored in few and big data centres, is not enough for satisfying the requirements of the modern IoT domain where a huge amount of data is generated from many disparate sources, often geographically disperse [15]. All this demands a distributed and/or decentralised cloud architecture where location-aware services are placed on the network edge so as to overcome issues related to real-time responses, low latency, low bandwidth availability and utilisation while ensuring resilience and tolerance with respect to network faults [16].

Following this trend, different initiatives have been proposed from both industrial and academic stakeholders: Fog Computing, promoted by CISCO, Mobile Edge Computing (MEC), supported by ETSI and Cloudlets, launched by CMU, are all aiming to offer a distributed architecture for enabling scenarios where pervasive and widespread cloud services are requested, namely automotive, industry 4.0, content and data management, smart cities, tactile internet [17]. Nowadays many open communities and initiatives have been created in order to govern and evolve such Edge/Fog technologies (e.g. OpenFog Consortium, OpenEdgeComputing). A similar evolution has been followed also by the big providers of the public Cloud (namely Amazon, Google and Microsoft) that recently launched proprietary proximity Cloud solutions for supporting IoT needs (e.g. AWS Greengrass).

What is proposed in our research, aims at leveraging the distributed Edge/Fog architecture advancing the state of the art in terms of applying cloud concepts like (function) virtualization, scalability, flexibility, self-healingcapabilities and DevOps patterns so as to enable Immutable Infrastructure Model and Pets vs Cattle like approaches, typical of traditional data centres, also at the edge of the network. In this line, a distributed, decentralized, heterogeneous and autonomic architecture is proposed to be conceived and deployed, able to offer not only computational and storage capacity close to the sources of the data but also to provide instruments for automatic deployment, fleet management, resource optimization, workload placement and support for multi-tenancy and privacy at the edge of the network.

B. Progress in Monitoring and Notification Frameworks

IoT monitoring and notification frameworks have primarily been used for triggering simple actions based on certain thresholds being reached by sensors. This may include a range of actions, raising alerts/alarms, sending email/push notifications to an administrator, or as a means of triggering the next step in process automation [18].

Two fundamental limitations are identified with these kinds of approaches – (1) There is a general assumption that an event should always be triggered when a sensing threshold is met, with the event handler needing to determine whether to filter out the event or take action, and (2) A largely consistent quality is assumed based on a device being able to provide a specific type of reading of interest to the application developer (e.g. a cardiograph from a heart rate monitor) [19].

While advances in semantic interoperability and device discovery (e.g. via the Web of Things) have opened up access to a wider range of sensors that may be capable of providing a requested reading, the individual sensors may provide widely different quality guarantees [20].

Monitoring and notification services for IoT appear in many commercial Cloud-based IoT platform offerings, such as Microsoft's Azure Logic Apps, Bug Labs Dweet.io, and ptc's Axeda Machine Cloud. Open source initiatives, such as the Open Connectivity Foundation's IoTivity are also increasingly providing these capabilities.

Considerations such as the privacy preferences and individual consent should be factored in and will provide limits for the notification mechanism, allowing for opt-in/out provisions to be made where appropriate. Additional technical considerations, such as the quality (and integrity) of the data source will further be considered as limiting the kind of actions that could be taken (e.g. a Fitbit detecting a flat heart rate is unlikely to be sufficient cause for notifying emergency services, while a patient under continuous health monitoring may have more precise and reliable equipment from which notification to emergency services could be considered as an option).

C. Progress in Cloud Data Lifecycle Management

The data lifecycle represents the process from creation of the data to its destruction, including the storage, update, use, transfer, sharing, and archival of the data. Sticky policies are policies that attach conditions to data to define how this data is allowed to be used during its lifecycle. Many surveys analysed the major issues pertaining to data security in the cloud computing environment, and many of them propose security solutions for these issues [21]. A major part of these solutions only considers one side of security or a particular phase of data lifecycle by focusing on data privacy and confidentiality, data access control, data integrity or data availability. Other works proposed that the cloud data security must be considered from the data life cycle [22].

The works related to data lifecycle security management are limited to data life cycle security analysis and a set of

recommendations. In addition, existing approaches based on sticky policies paradigm has no consideration of the data categories or the distinction of access privilege, some of them describe access policies with access structure, which does not fit to express the complex privacy aware access policies.

Our objective is to apply these recommendations and go further by layering the data lifecycle with governance policies, managing machine readable data sticky policies and enforcing them by key standard technologies of data security in the cloud such as identity and access management and other technologies issued from this project such as the anonymous credentials and monitoring and notification framework. We also aim to decentralize data lifecycle management for enforcing end-user privacy policies in dynamic and decentralized computational environments by implementing a framework for managing and enforcing sticky policies amongst a multitude of data centers and third-party cloud and IoT service providers. The TIF framework will give the end-user a central role in the specification and observation of his privacy changes.

D. Progress in Consent Management

Per Gartner definition: "*Consent management is a system, process or set of policies for allowing consumers to determine what information they are willing to permit their various providers to access. It enables consumers to affirm their participation in initiatives and to establish privacy preferences to determine who will have access to their personal information, for what purpose and under what circumstances. Consent management supports the dynamic creation, management and enforcement of consumer, organizational and jurisdictional privacy directives*" (link: https://www.gartner.com/en/information-technology/glossary/consent-management, accessed on June, 22 2022). The development of the European Commission's General Data Protection Regulation (GDPR) has increased the focus on consent management as the regulation contains requirements for explicit and unambiguous consent as one of the allowable mechanisms for the processing of personal data. Critically, however, consent management is primarily used to increase the level of trust between the data provider and the data consumer, which is anticipated to lead to better outcomes for both.

The management of consent via traditional means is administratively time and resource consuming. Current jurisdictions that required consent have experienced a great increase in paperwork and storage for the signed documents. Traditional consent capture has to be completed each time consent is required. There is no interoperability or exchange of consent between services.

The advent of GDPR and a desire from service providers to find better mechanisms for improving consumer trust has led to the development of personal data management systems. These systems allow the individual to manage their digital identity and include consent management so the individual can control how their data is used. Standards are emerging to ensure that the personal data management systems are interoperable. The Kantara initiative has a proposed standard for the consent receipt.

Our framework aims to make a first-of-a-kind deployment of consent management for IoT devices at the edge. Similar attempts seem to gain traction as well (e.g., [23], [24]) thus highlighting the need for such an approach. This will ensure that the digital rights and privacy of those interacting with the devices will be treated as seriously as network security. It will enable the management of consent preferences so that interactions with the devices will be unobtrusive to the data subject whilst maintaining the quality and integrity of the data.

E. Advantages and Disadvantages

The infrastructural architecture described so far has a number of advantages but also some disadvantages with respect to a traditional IoT architecture where cloud computing services are offered centrally. The following tables summarise the main pitfalls and pearls of our approach.

TABLE I. ADVANTAGES IDENTIFIED

Advantages
Ability to adapt to the distributed and decentralised architectures typical of new emerging technologies like 5G and IoT: cloud computing needs to escape the data centres and be offered as an enabler for intrinsically distributed technologies
Lower the latency and improve the response time (values around 10 ms or even less are possible with respect to hundreds of ms needed in case of usage of the public cloud): this is particularly important when actuation takes place.
Lower the bandwidth usage thanks to the elaboration of data at the edge of the network: raw data or streaming video can be pre-elaborated by the nodes at the edge (one network hop away from the things) and then stored locally; only the results of such an elaboration is sent to the cloud
Make the system more resilient to network faults: in the case of faults in the network connectivity between the edge and the cloud, the functionalities offered by the edge nodes continue to work without any service disruption. However, a reconciliation will be necessary once the connectivity is restored
Enable data privacy control and preservation keeping sensitive data on the user's premises and under the user management: user can decide which kind of data can be moved to the cloud and which should be kept locally

TABLE II. DISADVANTAGES IDENTIFIED

Disadvantages
Higher complexity in managing and operating a distributed, heterogeneous and decentralised architecture
More difficult to exploit economies of scale property of centralised cloud computing environments
The potential involvement of different types of stakeholders that make the management of the whole infrastructure more difficult: cloud providers, telco providers, edge providers etc

While the first disadvantage can be mitigated by the adoption of automatic deployment and operation tools, the second and third ones are intrinsic to the decision to opt for a model different from the one offered by the public cloud. Nevertheless, this can be seen as an opportunity to overcome the traditional idiosyncrasies of public cloud model (e.g. vendor lock-in) in favour of a multi-provider one.

In summary, the work foreseen in order to materialise the infrastructural architecture depicted above mainly targets the following aspects, advancing the state of the art in each of them: (a) Automatic deployment, fleet management and operational models in a highly distributed, decentralised and heterogeneous context involving swarm of devices; and (b) Multi-tenancy and privacy support at the edge of the network considering tenant isolation and privacy policy enforcement; (c) Decentralization of IaaS (e.g. OpenStack) and container orchestrator (e.g., Kubernetes) leveraging current federation initiatives; (d) Workload placement and management in a distributed and heterogeneous environment presenting different constraints and requirements on the different tiers; and (e) Monitoring and self-healing from the edge to the cloud.

V. Conclusion

The work presented here intends to facilitate data security protection across all stages of data lifecycle. The data lifecycle represents the process from creation of the data to its destruction, including the storage, update, use, transfer, sharing, and archival of the data. To achieve this, machine-readable policies are stuck to the data at the data creation phase. Current work mainly focus on privacy and confidentiality, data access control, data integrity or data availability.

The suggested framework intends to enhance existing work for all activities with data during its lifecycle as proposed by the Cloud Security Alliance. Adopting the UniversAAL open-source middleware (link: https://www.universaal.info/, accessed on: June 22, 2022), the TIF will ensure high level of interoperability with the latest "semantic web" standards for IoT. Using standardised protocols, we will develop new ontologies especially in the domain of ambient sensing and affective computing for the understanding of the emotional cues. The increasing usage and expected exponential growth of sensor networks creates the need of managing this growing complexity. UniversAAL creates a future-proof environment by avoiding domain-specific API's and vendor lock in. On a similar note we intend to capitalize on UniversAAL and FIWARE to seamlessly integrate multi-dimensional health monitoring technologies into a single tool, to transmit collected data using the latest IoT protocols. Additionally, logical structure of the infrastructure components in the TIF stack and relationships between the functional components and how they interact with one another to achieve their respective goals are described. This will lead to a framework that combines its elements seamless into an integrated solution that solves all suggested goals. In the meantime, it will be ensured that the results will have an impact on interested parties.

Acknowledgment

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957337, project MARVEL.

References

[1] A. Sestino, M. I. Prete, L. Piper, and G. Guido, 'Internet of Things and Big Data as enablers for business digitalization strategies', *Technovation*, vol. 98, p. 102173, Dec. 2020, doi: 10.1016/j.technovation.2020.102173.

[2] A. Reyna, C. Martín, J. Chen, E. Soler, and M. Díaz, 'On blockchain and its integration with IoT. Challenges and opportunities', *Future Generation Computer Systems*, vol. 88, pp. 173–190, Nov. 2018, doi: 10.1016/j.future.2018.05.046.

[3] B. Nagajayanthi, 'Decades of Internet of Things Towards Twenty-first Century: A Research-Based Introspective', *Wireless Pers Commun*, vol. 123, no. 4, pp. 3661–3697, Apr. 2022, doi: 10.1007/s11277-021-09308-z.

[4] A.-C. Karlsen and M. Wallberg, *The effects of digitalization on auditors' tools and working methods : A study of the audit profession*. 2017. Accessed: Jun. 22, 2022. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:hig:diva-24578

[5] P. Brous, M. Janssen, and P. Herder, 'The dual effects of the Internet of Things (IoT): A systematic review of the benefits and risks of IoT adoption by organizations', *International Journal of Information Management*, vol. 51, p. 101952, Apr. 2020, doi: 10.1016/j.ijinfomgt.2019.05.008.

[6] M. D. Bormida, 'The Big Data World: Benefits, Threats and Ethical Challenges', in *Ethical Issues in Covert, Security and Surveillance Research*, vol. 8, R. Iphofen and D. O'Mathúna, Eds. Emerald Publishing Limited, 2021, pp. 71–91. doi: 10.1108/S2398-601820210000008007.

[7] T. Alam, 'Cloud-Based IoT Applications and Their Roles in Smart Cities', *Smart Cities*, vol. 4, no. 3, Art. no. 3, Sep. 2021, doi: 10.3390/smartcities4030064.

[8] A. A.-N. Patwary *et al.*, 'Towards Secure Fog Computing: A Survey on Trust Management, Privacy, Authentication, Threats and Access Control', *Electronics*, vol. 10, no. 10, Art. no. 10, Jan. 2021, doi: 10.3390/electronics10101171.

[9] S. Shahhosseini *et al.*, 'Exploring computation offloading in IoT systems', *Information Systems*, vol. 107, p. 101860, Jul. 2022, doi: 10.1016/j.is.2021.101860.

[10] A. Mikkelsen, T.-M. Grønli, and R. Kazman, *Immutable Infrastructure Calls for Immutable Architecture*. 2019. Accessed: Jun. 22, 2022. [Online]. Available: http://hdl.handle.net/10125/60142

[11] A. Rahman, R. Mahdavi-Hezaveh, and L. Williams, 'A systematic mapping study of infrastructure as code research', *Information and Software Technology*, vol. 108, pp. 65–77, Apr. 2019, doi: 10.1016/j.infsof.2018.12.004.

[12] A. Malviya and R. K. Dwivedi, 'A Comparative Analysis of Container Orchestration Tools in Cloud Computing', in *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2022, pp. 698–703. doi: 10.23919/INDIACom54597.2022.9763171.

[13] F. Faticanti, D. Santoro, S. Cretti, and D. Siracusa, 'An Application of Kubernetes Cluster Federation in Fog Computing', in *2021 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, Mar. 2021, pp. 89–91. doi: 10.1109/ICIN51074.2021.9385548.

[14] H. F. Atlam, A. Alenezi, A. Alharthi, R. J. Walters, and G. B. Wills, 'Integration of Cloud Computing with Internet of Things: Challenges and Open Issues', in *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Jun. 2017, pp. 670–675. doi: 10.1109/iThings-GreenCom-CPSCom-SmartData.2017.105.

[15] A. Yousefpour *et al.*, 'All one needs to know about fog computing and related edge computing paradigms: A complete survey', *Journal of Systems Architecture*, vol. 98, pp. 289–330, Sep. 2019, doi: 10.1016/j.sysarc.2019.02.009.

[16] E. Ahvar, S. Ahvar, S. M. Raza, J. Manuel Sanchez Vilchez, and G. M. Lee, 'Next Generation of SDN in Cloud-Fog for 5G and Beyond-Enabled Applications: Opportunities and Challenges', *Network*, vol. 1, no. 1, Art. no. 1, Jun. 2021, doi: 10.3390/network1010004.

[17] M. Sheikh Sofla, M. Haghi Kashani, E. Mahdipour, and R. Faghih Mirzaee, 'Towards effective offloading mechanisms in fog computing', *Multimed Tools Appl*, vol. 81, no. 2, pp. 1997–2042, Jan. 2022, doi: 10.1007/s11042-021-11423-9.

[18] A. I. Paganelli *et al.*, 'A conceptual IoT-based early-warning architecture for remote monitoring of COVID-19 patients in wards and at home', *Internet of Things*, vol. 18, p. 100399, May 2022, doi: 10.1016/j.iot.2021.100399.

[19] N. Tax, N. Sidorova, and W. M. P. van der Aalst, 'Discovering more precise process models from event logs by filtering out chaotic activities', *J Intell Inf Syst*, vol. 52, no. 1, pp. 107–139, Feb. 2019, doi: 10.1007/s10844-018-0507-6.

[20] J. Lanza, L. Sánchez, D. Gómez, J. R. Santana, and P. Sotres, 'A Semantic-Enabled Platform for Realizing an Interoperable Web of Things', *Sensors*, vol. 19, no. 4, Art. no. 4, Jan. 2019, doi: 10.3390/s19040869.

[21] A. Ometov, O. L. Molua, M. Komarov, and J. Nurmi, 'A Survey of Security in Cloud, Edge, and Fog Computing', *Sensors*, vol. 22, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/s22030927.

[22] J. Koo, G. Kang, and Y.-G. Kim, 'Security and Privacy in Big Data Life Cycle: A Survey and Open Challenges', *Sustainability*, vol. 12, no. 24, Art. no. 24, Jan. 2020, doi: 10.3390/su122410571.

[23] H. Song, R. Dautov, N. Ferry, A. Solberg, and F. Fleurey, 'Model-based fleet deployment in the IoT–edge–cloud continuum', *Softw Syst Model*, May 2022, doi: 10.1007/s10270-022-01006-z.

[24] K. Rantos, G. Drosatos, A. Kritsas, C. Ilioudis, A. Papanikolaou, and A. P. Filippidis, 'A Blockchain-Based Platform for Consent Management of Personal Data Processing in the IoT Ecosystem', *Security and Communication Networks*, vol. 2019, p. 1431578, Oct. 2019, doi: 10.1155/2019/1431578.