Enabling trustworthy personal data protection in eHealth and well-being services through privacy-by-design

Tomás Robles, Borja Bordel, Ramón Alcarria and Diego Sánchez-de-Rivera

Abstract
Users are each day more aware of their privacy and data protection. Although this problem is transversal to every digital service, it is especially relevant when critical and personal information is managed, as in eHealth and well-being services. During the last years, many different innovative services in this area have been proposed. However, data management challenges are still in need of a solution. In general, data are directly sent to services but no trustworthy instruments to recover these data or remove them from services are available. In this scheme, services become the users’ data owners although users keep the rights to access, modify, and be forgotten. Nevertheless, the adequate implementation of these rights is not guaranteed, as services use the received data with commercial purposes. In order to address and solve this situation, we propose a new trustworthy personal data protection mechanism for well-being services, based on privacy-by-design technologies. This new mechanism is based on Blockchain networks and indirection functions and tokens. Blockchain networks execute transparent smart contracts, where users’ rights are codified, and store the users’ personal data which are never sent or given to external services. Besides, permissions and privacy restrictions designed by users to be applied to their data and services consuming them are also implemented in these smart contracts. Finally, an experimental validation is also described to evaluate the Quality of Experience (in terms of user satisfaction) and Quality of Service (in terms of processing delay) compared to traditional service provision solutions.

Keywords
Well-being services, Blockchain, privacy-by-design, data protection, trust-enhancing technologies

Date received: 3 May 2019; accepted: 29 January 2020

Handling Editor: Luis Castro

Introduction
Future pervasive sensing platforms, supported by new technological paradigms such as Internet of Things (IoT) and cyberphysical systems (CPS), will enable a new generation of eHealth and well-being services. These innovative services will improve their performance (including their personalization, cost, and processing delay) with respect to currently existing solutions. In order to do that, a large catalog of biological signals, environmental measures, and people’s behavioral information (among other data sources) must be captured and processed by well-being services. Besides, edge computing architectures must be implemented to reach that objective.

In this context, where digital services are totally integrated into citizens’ daily living, several different governmental institutions (such as the European Universidad Politécnica de Madrid, Madrid, Spain

Corresponding author:
Borja Bordel, Universidad Politécnica de Madrid, 28040 Madrid, Spain.
Email: bbordel@dit.upm.es

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Commission) have proposed regulations to protect these personal critical data from malicious uses. In particular, future eHealth and well-being services should respect the right of users with respect to their personal data: right to access, modify, and be forgotten. However, most times, well-being services cannot provide trustworthy instruments to execute these rights and, when provided, they do not guarantee a complete execution according to published regulation.

This problem is highly related to current service architectures. In this scheme, personal data about users' health or situation are acquired and directly sent to remote servers where well-being services are executed. Sometimes, these data are also stored in local repositories, but they are always sent to service providers' servers. This approach, in fact, is very useful, as remote agents (such as doctors) access these data very easily once processed by services. Nevertheless, it causes a very worrying situation: service providers replicate and store all users' personal and critical data. Obviously, these service providers (legally) cannot employ users' data with commercial or other similar objectives, and users can always execute the rights described in the corresponding regulation (General Data Protection Regulation (GDPR), in Europe, for example).

In real situations, users must trust service providers to remove (or modify) their data when requested, as no guarantee is provided. Besides, usually, this information is employed with commercial purposes without the users' permission. In a digital society where, more and more, people are concerned about their data privacy and where, more and more, digital services and flowing personal information are essential elements in the daily living, this situation must be addressed.

On the other hand, privacy options provided by well-being services are usually limited (as in Web 2.0 services), so users must manage permissions according to service providers' criteria. However, users must be enabled to create and design their own privacy restrictions to be applied to their data in a totally free manner.

Therefore, the objective of this article is to describe a new trustworthy personal data protection mechanism for eHealth and well-being services, based on privacy-by-design (PbD) technologies. Blockchain networks are employed to provide trust among agents (users and service providers) and to guarantee users' control on their personal data.

In particular, Blockchain networks use the notion of smart contract (SC), also known as a cryptocontract, as a computer program that directly controls logic execution and transfer of assets between parties in a Blockchain network. General privacy policies (contained in regulation) and user-defined privacy policies will be integrated into these SCs, enabling a PbD approach. Ambient intelligence devices and other similar components will be authorized to interact with these SCs in order to automate data capture. Personal data will be stored in a trustworthy database, managed by the Blockchain network. Well-being services will be structured following an edge computing architecture, so data will never be sent to service providers' remote servers. Indirection functions and tokens will be employed to represent personal data in well-being services, but these anonymous mechanisms will be only resolved locally in the users' infrastructure.

The rest of the article is organized as follows: section “Related work” introduces the state of the art in privacy and trust-enhancing technologies, with a special interest in well-being services; section “Trustworthy personal data protection in well-being services” describes the privacy regulation requirements and the proposed trustworthy mechanism from a mathematical and technological point of view; section “Implementation and experimental validation” describes a first real implementation of the proposed solution and the experimental validation carried out to evaluate the proposal performance; section “Results and discussion” presents the obtained results and discusses about their validity. Finally, section “Conclusion and future work” concludes and describes future developments.

Related work

Privacy- and trust-enhancing technologies are, nowadays, a very relevant research topic. Thus, many different works addressing this problem in eHealth and well-being services may be found.

Most common works in this area are position papers describing pending challenges or future research opportunities. Areas analyzed with a special interest are wellness mobile apps and mobile health services. In general, these works are focused on third parties trying to capture personal data but do not consider privacy challenges associated to service providers and their potential malicious behavior.

With a more technological view, a large catalog of mechanisms to protect users' privacy in well-being services have been reported: from protocols to be implemented into cloud infrastructures supporting health services, to solutions to hide personal information in medical records, logs, and other eHealth solutions. Solutions based on traditional cryptographic methods (including revocation procedures and other innovative functionalities) have been reported. However, in the last 2 years, different authors have discussed how Blockchain may revolutionize this area.

Although Blockchain solutions for well-being services are still sparse, some different proposals may be found. In general, in these works, Blockchain networks are employed to maintain SCs representing electronic medical records (EMRs), so data in these digital objects
are managed by Blockchain instruments. Other similar proposals described architectures that enable users to control their personal data sharing from mobile apps, but in all these approaches (finally) data are sent to service providers (a risky situation as already described). Real use cases of these technologies have also been described. Mobile eHealth services have also been studied to be integrated with Blockchain, and some tamper-resistant solutions have been reported. In this case, Blockchain networks are employed to manage the data update process, so no manipulation is allowed.

All these proposals, however, are designed following the same approach: personal data are finally sent to service providers. Access, storage, or actualization is managed through Blockchain, but (at the end) well-being services obtain a copy of all personal information. With the proposed mechanism in this article, this situation is avoided.

This new approach, based on edge computing and distributed data governance, introduces security concerns and ethical challenges which have also been analyzed especially in the context of clinical trials. Equally, the use of IoT devices and pervasive sensing platforms as data sources presents additional security challenges.

First, information transfer between the sensing platform and the Blockchain infrastructure must be secure. Lightweight protocols and security solutions specifically designed for eHealth applications have been reported. Moreover, mechanisms are required to automate the interaction between electronic devices and Blockchain networks, while addressing the challenges related to remote devices and public key infrastructure (PKIs). Some solutions based on homomorphic cryptography may be found, for ambient intelligence applications, water management, international commerce, and domestic scenarios. Other solutions based on physical unclonable functions have been reported. Besides, intelligent mechanisms to protect IoT and CPS deployments have also been studied.

Finally, PbD mechanisms have also been investigated in the context of well-being services, especially in scenarios combining health services, digital technologies, and hardware devices, such as IoT elements. Discussions about legal issues related to this approach have also been published. In any case, no real technological implementation of these proposals has been reported, a gap that is filled in this article.

**Trustworthy personal data protection in well-being services**

The concept of Blockchain has arisen as a trustworthy, authentic, shared, public (although some implementations define private networks) peer-to-peer system to manage information. A collection of remote nodes creates a unidimensional immutable sequential chain of blocks, linked to lateral blocks through hash functions. This chain describes a historical record of all transactions performed among the peers, so all involved nodes must create a consensus about what transactions are valid and accepted (in order to maintain the information and system integrity). Although different consensus mechanisms are now available and under discussion, the proof of work is the most extended one.

Blockchain presents some relevant benefits, to be considered in well-being services. It is a transparent and auditable technology, so the logic coding the users’ right is clearly exposed. Besides, as a decentralized solution, it is more resistant to faults and attacks. Furthermore, in most modern Blockchain systems (such as Ethereum), nodes maintain and execute SCs as cryptographic “boxes” containing programming logic that is executed if certain conditions are met. This enables users to design their own privacy policies and permissions in a very easy manner.

Therefore, in this section, we describe a Blockchain-based solution to guarantee privacy according to regulation and users’ designed requirements. The proposed solution follows an edge computing architecture and employs indirection functions to operate correctly. Trustworthiness is ensured by Blockchain behavior and (mobile) prosumer environments allow the creation of PbD rules.

**Regulation requirements and the PbD paradigm**

Any solution focused on data protection must, at least, guarantee the considered requirements and users’ rights in regulation. Different regions and governmental institutions around the world have promoted different regulations, but (among them) the “General Data Protection Regulation” created by the European Commission is the most popular and extended one.

In particular, GDPR identifies some data protection features for digital services. Below, data protection requirements are briefly presented, so later the service architecture and the data protection mechanism may be designed according to the following characteristics:

**REQ#1—transparency.** All well-being services must process personal data in a transparent manner vis-à-vis data owners.

**REQ#2—purpose limitation.** Personal data can only be accessed and processed for legitimate and limiter purposes, explicitly authorized by users.

**REQ#3—data minimization.** Accessed and processed personal data must not be excessive for the well-being services’ purpose.
REQ#4—accuracy. Well-being services must guarantee that personal data are processed with accuracy and results are updated if possible.

REQ#5—storage limitation. Personal data must not be stored for longer than required to reach the service objective.

REQ#6—integrity and confidentiality. Personal data must always be securely processed, transmitted, manipulated, and stored.

REQ#7—accountability. Mechanisms enabling users to check service compliance with the above requirements must be provided.

Apart from these requirements, the proposed trustworthy data protection mechanism must be compliant with users’ rights (REQ#8), namely:

- Right to unambiguous consent;
- Right that only relevant, necessary, accurate, and legitimate data are processed in a specific, fair, and transparent manner;
- Right to access one’s own personal data;
- Right to be properly informed when personal data are processed;
- Right to rectification;
- Right to protection against the use of personal data for automated profiling;
- Right to be forgotten;
- Right to security measures.

The proposed architecture (see section “Service architecture and the data protection mechanism”) will satisfy the described requirements but following a PbD methodology. In particular, the PbD approach improves privacy-enhancing technologies (PETs), such as Blockchain, so users can avoid the zero-sum situation where either they share their personal data with a trustworthy repository. In particular, all interactions with the database are public and irrevocable if accepted (for which a majority of service users participating in the network must approve the transaction). In any case, the database configuration must be coherent with the proposed architecture. No access from any host different to the database controller in the Blockchain.

Service architecture and the data protection mechanism

Figure 1 shows the proposed service architecture for future well-being services, where data protection mechanisms are embedded into design (not included as an extra functionality). Thus, REQ#11 is directly met. Architecture presents some numbers in red to better represent architectural transactions and mentions in the following paragraphs.

The described architecture follows the edge computing scheme. Users (or prosumers, as they can also create their own privacy rules) are monitored through sensing elements which insert information coming from biomedical devices or through some health care solutions that are installed in general-purpose processing components. As currently there are plenty of legacy systems for data acquisition in health and well-being, the proposed architecture also provides a way to communicate these legacy systems to business core services (where healthcare solutions and well-being solutions co-exist). All these devices make up a sensing layer which can operate in an autonomous manner.

These data are, then, stored in a trustworthy data repository through a Blockchain network and an edge layer connected to the Blockchain using application programming interfaces (APIs). This data repository belongs to users and (by default) maintains data privacy and protection (nobody has access permission). With this feature, REQ#10 is met.

As the data repository is managed by a database controller embedded into a Blockchain network, it is a trustworthy repository. In particular, all interactions with the database are public and irrevocable if accepted (for which a majority of service users participating in the network must approve the transaction). In any case, the database configuration must be coherent with the proposed architecture. No access from any host different to the database controller in the Blockchain.
should be allowed. Traditional instruments such as firewalls or traffic filters could be employed. On the other hand, as transactions are public and edge computing devices are operating with a Blockchain network, data must be encrypted and devices must be authorized to operate with the Blockchain network. Section "Trustworthy data protection: the mathematical approach" is focused on this problem. In any case, the use of these encryption schemes from the information generation moment guarantees that REQ#6 and REQ#13 are fulfilled.

Moreover, as mentioned in the initial paragraphs of section “Trustworthy personal data protection in well-being services,” SCs in the Blockchain are transparent pieces of code, which are public and have no access restrictions. In that way, privacy-preserving mechanisms turn transparent in our proposal, and REQ#7 and REQ#14 are met.

On the other hand, the Blockchain network also executes some additional SCs located in the privacy management system. A prosumer environment where users may create the privacy rules to be applied to their data is also considered. In this environment, users employ a domain-specific language to authorize well-being services to operate with their personal data and create other privacy rules to be considered. Later, this description is translated into a deployable SC in the SC creation engine. The same operation can be performed by data managers, which are in charge of implementing GDPR users’ rights in a different SC. Both SCs manage personal data privacy. These characteristics ensure that REQ#8 and REQ#15 are fulfilled.

As well-being services need users’ personal data, they participate in the Blockchain network with an electronic identifier (in the Blockchain, a public key calculated using elliptic curves). This identifier must be explicitly included in the privacy rules defined by users to grant permissions for the well-being services (REQ#2). Then, using a new SC acting as a service linker (services hosted in remote servers or in a cloud layer) request for the needed personal data to operate. This request is then redirected to the privacy management system, where an SC focused on tokenization and indirection obtains the personal data from the trustworthy repository, but returns to the service a token calculated with an indirection function (not the raw personal data). These tokens are pointers to personal data stored in the data repository, which allow operating with data but without sharing the real information.
complexity of the server providing the well-being services. As the main disadvantage, the proposed mechanism highly increases its complexity, so more computationally powerful machines would be required, as well as more efficient software instruments (in order to ensure correct data transmission).

With these tokens, service providers may operate without revealing their core business, but may generate parametric results which are later resolved when indirections get undone by users (data owners). This approach, based on edge computing schemes, prevents possible privacy violations, even if no risk is perceived (REQ#9), and enables privacy in execution of well-being services (REQ#12).

Parametric service results are only resolved in the edge layer, where devices belong to users. However, if remote experts (such as doctors) must consume the service results, then a new layer is necessary: the fog layer. In the fog layer, a collection of nodes resolves parametric results and indirections in a transparent manner vis-à-vis data owners (REQ#1). Thus, remote experts only obtain processed data and personal information according to the privacy rules defined by users, and only the essential information (not a complete copy of users’ personal data). REQ#3 is, in that way, met. Processed results are, finally, sent from the fog layer to an edge layer where they can be manipulated, enriched, and so on by remote experts. Moreover, edge and fog layers will be configured to forget any parametric or resolved results after being consumed by users, ensuring storage limitation.

Furthermore, and finally, if necessary, remote experts may enrich personal data with new information, which is stored in the trustworthy data repository following the previously described procedure. If new information about a user is stored, that user, and all authorized agents according to the privacy rules, will be notified, so service results and other relevant information and knowledge may be updated. Thus, REQ#4 is also fulfilled. Table 1 summarizes the main architectural components and the requirements they mainly address.

### Table 1. Mapping of requirements and architectural components.

| Requirement | Architectural component | Description |
|-------------|-------------------------|-------------|
| #1          | Fog layer               | Parametric results are resolved in a flexible and transparent manner vis-à-vis data owners. |
| #2          | User privacy rules SC   | Privacy rules are defined by users to grant permissions for the well-being services. |
| #3          | Cloud and fog layers    | Parametric service results are resolved so that remote experts only obtain processed data and only essential information. |
| #4          | Notification SC         | New information is notified to authorized agents, so that relevant information and knowledge are enriched. |
| #5          | Edge and fog layers     | Layers forget any parametric or resolved results after being consumed by users, ensuring storage limitation. |
| #6          | DB controller and encryption schemes | The use of encryption schemes from the information generation moment ensures integrity and confidentiality. |
| #7          | Privacy management SC   | SCs are transparent and source code is available. Users can check service compliance with GDPR in SC specification. |
| #8          | GDPR SC                 | User rights are codified in GDPR SC, in a transparent way. Rights are publicly available, and compliance can be checked. |
| #9          | Edge layer and indirection schemes | Edge computing provides tokenization and indirection schemes, preventing privacy violations, even if no risk is perceived. |
| #10         | Data repository          | Trustworthy data repository belongs to users and, by default, maintains data privacy and protection. |
| #11         | Architectural design    | Data protection mechanisms are embedded into architectural design, being privacy aware by default. |
| #12         | Indirection in well-being services | Make possible users maintain their privacy, at the same time supporting the business of well-being service providers. |
| #13         | DB controller and encryption schemes | The use of encryption schemes from the information generation moment ensures end-to-end security. |
| #14         | Privacy management SC   | Privacy-preserving instruments and mechanism are transparent through Blockchain transparency capabilities. |
| #15         | User privacy rules SC   | Users’ rights are mainly considered in this component. Other components must comply with that. |

SC: smart contract; DB: database; GDPR: General Data Protection Regulation.
Trustworthy data protection: the mathematical approach

All users and participants (such as service providers) in a Blockchain network are provided with a key pair for authentication purposes (named user operation key, $k_U^O$). This key pair is protected by a symmetric secret user key $k_U$, so even if the operation key is stored in a host, only the owner may employ it. The operation key is automatically calculated for every user through a digital signature scheme (DSS). This DSS is described by 3-tuple (equation (1)) including a key generator $G_{sig}$, a signature function $S_{sig}$, and a verification procedure, $V_{sig}$. In our proposal, any existing DSS may be used, provided that it is homomorphic

$$G_{sig}, S_{sig}, V_{sig}$$

As all managed data by Blockchain systems are accessible by all network participants, personal data in our trustworthy data protection system must be encrypted to maintain them secured. Using the proposed DSS, well-being personal data could be encrypted by sensing elements, but some problems must be considered. First, sensing devices must operate in an autonomous manner, that is, no action from users can be requested prior to every data acquisition. However, the proposed DSS (an asymmetric encryption mechanism) only may be employed if the user secret key is provided. This key, to be secure, cannot be codified into every sensing device, but users cannot be requested to introduce the key periodically neither. Second, remote experts, nodes in the fog or edge layers, and other agents different to the data owners could potentially be authorized to access to data. If data are encrypted, all these agents should be provided with the corresponding decryption key, which is totally unsecure. Besides, all permissions should be able to be revoked dynamically. Therefore, the privacy management core and SCs in the trustworthy data protection system implemented a new and more complex cryptographic solution. Figure 2 describes the proposed cryptographic scheme.

In this figure, transactions managed by edge and sensing devices are represented by blue arrows. Transactions managed by users are represented using red arrows. Transactions managed by SCs are represented using violet arrows, and transactions managed by remote experts and services are represented using green arrows. The key and permission generation process is always started by users, which are the real participants in the Blockchain network (edge devices and

![Cryptographic scheme for the proposed trustworthy data protection system.](image-url)
Remote services are only authorized entities. Remote services and edge devices may execute also, in parallel, some initial operations, but they must wait until users finish their transactions. After users have performed all the operations, they are responsible for, in parallel, remote services and edge devices may interact with the SCs and get authorized to interact with the Blockchain network. Below, we prove a formal explanation with details.

Operation user key, as any asymmetric key for a PKI, is in fact a key pair (equation (2)): a private user operation key $k_{U-pv}$ and a public user operation key $k_{U-pb}$. This key is obtained from the key generator $G_{sig}$, using the user secret key and, optionally, an $n$-tuple of additional parameters $p_i$ increasing the generator entropy and, then, the operation key quality (equation (3))

$$k_U^* = \left\{k_{U-pb}, k_{U-pv}\right\}$$

$$G_{sig}(k_U, \{p_1, \ldots, p_n\}) = k_U^* = \left(k_{U-pv}, k_{U-pb}\right)$$

Now, it is important to remember that the required DSS is homomorphic. In homomorphic cryptography, any signed (or encrypted) information obtained from the algebraic operation of a set of clear information pieces is equivalent to operate through an algebraic function (not necessarily the same) the information piece signed (or encrypted) independently (equation (4)). Although this requirement may seem hard to fulfill, in fact many practical and common DSSs present this characteristic (such as ElGamal technique):

$$S_{sig}(m_1 \odot m_2 \odot \ldots \odot m_N) = S_{sig}(m_1) \odot S_{sig}(m_2) \odot \ldots \odot S_{sig}(m_N)$$

with $\odot$ and $\odot$ indicating the algebraic operations.

Then, to allow sensing and edge devices to interact with SC and the trustworthy data repository, a device key $k_d$ for every component is secretly obtained by users. This device key is obtained through a second generator $F_{sig}$, which presents a similar structure to the previously described generator, $G_{sig}$. In this case, however, the private operation user key $k_{upv}$ is employed. An $n$-tuple of additional parameters $r_i$ to increase the generator entropy may also be considered (equation (5)). As key generation functions are not invertible and device keys are obtained secretly, all security features are preserved. This new key generator, however, together with the corresponding device key pair, produced a piece of information $m_d$, which is unique for each key (and device), and that is named as “homomorphic constant.” As mentioned, device key is a key pair including a public device key $k_{d-pv}$ and a private device key $k_{d-pb}$

$$F_{sig}(k_{U-pv}, \{r_1, \ldots, r_n\}) = \{k_d, m_d\} = \{(k_{d-pv}, k_{d-pb}), m_d\}$$

Hereinafter, we must guarantee that the private user operation key $k_{U-pv}$ and the private device key $k_{d-pv}$ present an equivalence relation in the context of the signature function $S_{sig}(\cdot)$, and supported by the

![Figure 3. Sequence diagram for Algorithm 1.](image-url)
Algorithm 1. Authorization mechanism for edge and sensing devices.

User executes:
1 \[ (k_{U-pb}, k_{d-pb}) \leftarrow G_{pub}(\cdot) \]
2 \[ \{ (k_{d-pb}, k_{2-pb}, m_d) \leftarrow F_{sig}(\cdot) \]
3 \[ m_d^{sig} \leftarrow S_{sig}^{k_{d-pb}}(m_d) \]
4 User stores \( k_{U-pb}, m_d^{sig} \) in the smart contract
5 User sends \( k_{d-pb} \) to the edge device

Edge device executes:
6 Edge device collects information \( m \)
7 \[ s \leftarrow S_{sig}^{k_{d-pb}}(m) \]
8 Edge device sends \((m, s)\) to the smart contract

Smart contract executes:
9 auth \( \leftarrow \gamma_{sig}^{k_{U-pb}}(m_d^{sig} \otimes S_{sig}^{k_{d-pb}}(m)) \]
10 if auth then
11 \( m \) is stored
12 end if

Algorithm 2. Authorization mechanism for remote experts and services.

User executes:
1 \[ T \leftarrow G_{pub}(\cdot) \]
2 User sends \( T \) to the remote agent and to the smart contract
3 User waits for the public remote key \( k_{r-pb} \)
4 User receives the public remote key \( k_{r-pb} \)
5 User stores \( k_{r-pb} \) in the smart contract

Remote agent executes:
6 \( (k_{r-pb}, k_{r-pb}) \leftarrow G_{pub}(\cdot) \]
7 Remote agent sends \( k_{r-pb} \) to users
8 \[ T_{sig} \leftarrow S_{sig}^{k_{r-pb}}(T) \]
9 Remote agent sends \( T_{sig} \) to the smart contract

Smart contract executes:
10 auth \( \leftarrow \gamma_{sig}^{k_{r-pb}}(T_{sig}) \]
11 if auth then
12 Access is granted
13 end if

homomorphic constant (equation (6)). In this expression, the homomorphic constant \( m_d \) satisfying the equivalence relation may seem hard to obtain. Nevertheless, even if anti-collision hash functions are employed in the signature process, as private keys are different in each side of the relation (equation (6)), a mathematical solution exists for this problem, preserving the resistance to the birthday attack.\(^{42}\) Moreover, currently existing hash calculation algorithms allow obtaining results fulfilling a set of preconditions in an easy and fast manner

\[ S_{sig}^{k_{U-pb}}(m) = S_{sig}^{k_{d-pb}}(m_d \oplus m) \]

with \( \oplus \) indicating the algebraic operation.

Now, considering the previous expressions and taking into account the selected DSS in homomorphic, the signature process of any personal information \( m \) may be written as an algebraic function involving a new constant \( m_d^{sig} \) that is unique for each edge or sensing device (equation (7))

\[
S_{sig}^{k_{U-pb}}(m) = S_{sig}^{k_{d-pb}}(m_d \oplus m) = S_{sig}^{k_{d-pb}}(m_d) \otimes S_{sig}^{k_{d-pb}}(m) = m_d^{sig} \otimes S_{sig}^{k_{d-pb}}(m)
\]

with \( \otimes \) and \( \oplus \) indicating the algebraic operations.

In that way, in order to grant permission to an edge or sensing device to interact with the data protection system, the SCs, and the trustworthy data repository, the corresponding SC (named as “data integration” in Figure 1) must only store two parameters: the public user operation key \( k_{U-pb} \) and the \( m_d^{sig} \) parameter for each authorized device. As can be seen, both parameters are public and no personal information is involved or revealed. In order to complete the procedure, the authorized edge device is provided with the corresponding private device key \( k_{d-pb} \).

Then, in order to insert new personal information \( m \) in the repository, the sensing device must send to the SC (together with the information) the corresponding signature \( S_{sig}^{k_{d-pb}}(m) \). In the SC, the signature is validated using only public information (already stored in the Blockchain), thanks to specific properties of homomorphic cryptography (equation (8)). Only if a valid and authorized device private key \( k_{d-pb} \) is employed, signature may be validated through the constant value \( m_d^{sig} \). After signature validation, data repository accepts the received information

\[
\gamma_{sig}^{k_{U-pb}}(S_{sig}^{k_{d-pb}}(m)) = \gamma_{sig}^{k_{U-pb}}(m_d^{sig} \otimes S_{sig}^{k_{d-pb}}(m))
\]

On the other hand, personal information \( m \) may be encrypted using the public user operation key \( k_{U-pb} \) (that is public). In order to decrypt that information in edge, fog, and sensing devices, an identical deduction to the previous one, based on homomorphic cryptography, may be done. Authorized devices to decrypt personal information must include in the SC the corresponding homomorphic constant (different to the constant for the signature process), which is a public datum with no personal content.

The proposed scheme presents some relevant advantages. First, the SC does not maintain any private key, which could be publicly accessed. Besides, as keys are automatically generated, they can be as long as necessary to reach the desired security level. Moreover, keys may be periodically updated very easily, so cyberattacks based on statistical learning are very difficult to perform.
In order to summarize the described procedure, Algorithm 1 shows the proposed authorization mechanism. To make easier to understand the proposed solution, Figure 3 shows the same algorithm in a sequence diagram.

Now, in order to enable fog nodes and edge devices of remote experts, and remote well-being services, to access personal data in the repository, a new, additional and different authorization mechanism is required. The same strategy described previously for sensing devices could be employed; nevertheless, as experts and well-being services are remote, key distribution is a relevant problem and making users calculate all keys (as previously) is not recommended.

Therefore, in this case, we are proposing a token-based solution. In this scheme, users calculate a token $T$ for each remote agent using a specific token generator $G_{tok}$. This generator, as previous key generators, receives different parameters: in this case, the public user operation key $k_{r-pb}$ and an $n$-tuple of additional parameters $q_i$ to increase the generator entropy (equation (9)). The obtained token is sent to the authorized remote agent and to the data protection system to be stored in the corresponding SC

$$G_{tok}(k_{r-pb}, \{q_1, \ldots, q_n\}) = T$$

(9)

At the same time, the authorized remote agent, using the previously described key generator (equation (3)), generates a new remote key $k_r$ (which is also composed by a key pair, including a private and a public key). In this case, however, only additional parameters $s_i$ to increase the entropy during the generation process are injected in the generator (equation (10))

$$G_{sig}(\{s_1, \ldots, s_n\}) = k_r = (k_{r-pv}, k_{r-pb})$$

(10)

The remote agent, then, sends the public remote key $k_{r-pb}$ to the user, who sends it to the data protection system to be stored in the corresponding SC. As this information is public, no risk is caused. At this point, in order to access personal data protected by our solution, remote services or experts should send (together with their request) the received authorization token, signed with the private remote key $k_{r-pv}$. If the signature is validated, access is granted (equation (11))

$$V_{sig}(k_{r-pv}, \text{sig}(T))$$

(11)

Algorithm 2 shows the proposed authorization mechanism for remote agents. To make it easier to understand the proposed solution, Figure 4 shows the same algorithm in a sequence diagram.

It is important to note that, due to data encryption, there is a delay in the process of personal information. In that way, although access permissions and data processing may be modified dynamically, they cannot be managed as fast as some situations may require. For example, for emergency access to personal data, remote agents (such as doctors) may be affected. This delay may be controlled through the Blockchain configuration and other system settings, but a certain delay is always present caused by the proposed architecture itself.

With the proposed mechanism, access permissions are dynamically and securely managed. However, without any additional feature, remote services and experts would obtain their own copy of the users’ personal data. This situation, as mentioned previously, should
also be solved with the proposed data protection system. For this, an edge computing approach and tokenization solution based on indirection functions is proposed.

Using the proposed architecture (see Figure 1), a remote well-being service $W$ is provided from the cloud. This service is represented as a function acting over a set of sequences of personal data $d_i[n]$ (equation (12))

$$W(d_1[n], d_2[n], \ldots, d_M[n])$$

This service function is understood as the composition of four functions (equation (13)): $C(\cdot)$, $F(\cdot)$, $E(\cdot)$, and $S(\cdot)$. $C(\cdot)$ represents the service functionalities to be executed in the cloud, $F(\cdot)$ the service functionalities to be executed in the fog layer, $E(\cdot)$ the service functionalities to be executed in the edge layer, and $S(\cdot)$ indicates the service functionalities to be executed in the sensing platform. If the service under study has no functionalities to be deployed in a certain layer, the corresponding function is replaced by the identity function $Id(\cdot)$

$$W(d_1[n], d_2[n], \ldots, d_M[n]) = S \cdot E \cdot F \cdot C(d_1[n], d_2[n], \ldots, d_M[n])$$

On the other hand, tokenization SCs codify an indirection function $\Pi(\cdot)$, so for each sequence or piece of personal information an anonymous pointer or token $\tau$ (equation (14)) is generated. In order to locate samples or specific data within the sequences, tokens may be complemented with an offset value $\phi_n$. Token calculation is, in our case, based on anti-collision hash functions, the outputs of which are composed by a large number of bits (256 or even a higher number)

$$\{\tau_i, \phi_n\} = \Pi(d_i[n])$$

Then, when remote well-being services request for personal data, they obtain an anonymous token representing those data. In the cloud, service providers perform those operations which are secret as they belong to their business core. As tokens are anonymous parameters representing real data, the result of these initial operations is a parametric service result (equation (15))

$$C(\{\tau_1, \phi_1^1\}, \{\tau_2, \phi_2^2\}, \ldots, \{\tau_M, \phi_M^M\})$$

This result is, later, received by the edge or fog layer (depending on the situation). Edge devices belong to users, and fog nodes (see Figure 1) are part of the trustworthy data protection system, so components in both layers may get access to clear personal data (if permissions are granted by owners) without any relevant privacy risk and, overall, no external agent gets a copy of the personal data being processed. Then, in those layers (edge or fog), tokens are resolved through an inverse indirection function $\Pi_{inv}(\cdot)$, also codified in SCs, and after exchanging data and resolution requests (see Figure 1) with the Blockchain network in the data protection system. Besides, functionalities deployed in the fog, edge, and sensing layers are executed to generate the final service result.

Figure 5 presents the proposed service execution scheme, including tokens and indirection functions. All arrows indicate the transmission of messages, requests, or replays. Green arrows indicate transactions with the data repository. Purple arrows indicate transactions with the Blockchain network and orange arrows indicate transactions between end devices. Transactions are numbered to show the natural sequence of interactions. Furthermore, two different flows are also represented.
(labeled with letters “a” and “b”): the first one (flow “a”) representing a service provided to users or prosumers and the second one (flow “b”) representing a service provided to remote experts.

Implementation and experimental validation

The validation of this article aims to discuss whether the new proposed PbD architecture and well-being service provision scheme improves the Quality of Experience (QoE; including an improved user satisfaction) and Quality of Service (QoS; represented by the main QoS indicator, the processing delay) compared to traditional service provision solutions. In order to evaluate the performance of the proposed technology, some experiments with users were proposed. In these experiments, users interacted with well-being services provided through the new proposed mechanisms. Technical measures were collected, and (at the end) users fulfilled a survey about their experience. In this section, we describe the validation process considering the context, plan, and data gathered from the experiments.

The first experiment

Two experiments were conducted employing as main material the described system in this article. In this section, we are describing all information about the first one.

Experimental setup. The final objective of this experiment was to answer some questions regarding the effectiveness of the proposed technological solution, in terms of the reached improvement in QoE. The following research question was formulated:

Q1. Is there any significant difference (improvement) in user satisfaction reported by surveys when using the proposed scheme and when using traditional cloud-based solutions?

Method and participants. The first experiment was planned, guided, monitored, and evaluated by its authors (hereafter experts), who have more than 5 years of experience in service provision systems, cybersecurity, privacy management, and data analysis.

In order to evaluate the performance of the proposed solution, a validation scenario was built. In this scenario, a collection of well-being services was deployed following the proposed architecture and using a virtual representation of Blockchain networks (Ganache).

Using the scenario, an experiment consisting of surveys about the QoE (user experience and the perceived privacy level) was carried out. During this experiment, two well-being services collecting personal data about a community of 12 users were deployed in a home-like environment. Considered services included monitoring of users’ motivation and a pattern recognition service for activities of daily living (ADL). Involved users were adults aged from 30 to 60. The same number of women and men participated (see Table 2). Human rights and ethical issues were respected. All participants were provided with a document describing their rights and the ethical issues related to the experiment. Participants signed the document and kept a copy.

The proposed technology was deployed in a laboratory of the Technical University of Madrid, where a house-like scenario is created. The created scenario represented a house-like scenario, where ADL can be performed (see Figure 6). In particular, in order to guarantee that we are testing a scenario which was validated previously, we selected an application and collection of tasks to be performed which have been extensively employed in other experiments related to enhanced living environments and other similar technologies. Participants could perform in this scenario a large list of activities. Some of the most relevant and common were going to toilet, bathing, transferring objects, dressing, and feeding. In particular, all users

| Total number of participants | Mean age | Standard deviation in age | Women percentage |
|-----------------------------|----------|---------------------------|-----------------|
| 33                          | 49       | 6.3                       | 48.5            |
perform (at least) three times the following activities: (1) transfer (moves in and out of bed and/or in and out of chair); (2) feeding (gets food from plait or equivalent into mouth); (3) going to toilet (gets on and off toilet and arranging clothes); (4) bathing (sponge bath, tub bath, or shower); and (5) dressing (gets clothes from closets and drawers and puts on clothes). Users had no limit time for any activity.

As the control group, we selected a new set of 12 users which were immersed in the same environment, but, in this case, well-being services were supported by a cloud infrastructure as in traditional architectures. The same survey about their experience and perceived privacy were performed in both groups, and the obtained results were compared using statistical tools.

In this initial experiment, for developing SCs associated to the trustworthy data protection system, a Blockchain simulation environment, Ganache (Ganache homepage: https://www.trufflesuite.com/ganache) was used. Ganache delivers a personal Blockchain for Ethereum development that facilitates contract deployment, app development, and testing. The employment of Ganache is combined with the usage of the Truffle framework, a development environment and testing framework run on top of the Ethereum Virtual Machine (EVM). Truffle facilitates the SC compilation, linking, deployment, and binary management, as well as direct contract communication to test API definitions.

Privacy and tokenization contracts (shown in Figure 7) are implemented, in Solidity programming language, as described in the system’s architecture. Tokenization in this example is defined with a new token called HealthToken. Tokenization SC codifies the indirection function \( \Pi(\cdot) \), so for each sequence or piece of personal information an anonymous pointer to an indirection result is generated (as indicated in the HealthToken’s constructor). Tokens are generated (function \textit{generate}), complemented with an offset value \( \phi_n \), and resolved (function \textit{resolve}) through an inverse indirection function \( \Pi_{\text{inv}}(\cdot) \), also codified in the SC, and after exchanging data and resolution requests with the Blockchain network.

The privacy contract implements the authorization mechanisms explained in Algorithm 1. The \( k_{U-pb} \) and \( m_{\text{d}}^{\text{des}} \) parameters for each authorized device are stored in the SC through the \textit{setAuthorized} function. Fog nodes, edge devices of remote experts, and remote well-being services ask for data access or data write permission through the \textit{permission} function and signature validation is then performed through the \textit{onlyAuthorized}
requirements. After signature validation, data repository accepts the received information.

Also, GDPR is supported by CRAB (Create, Retrieve, Append, Burn) operations. In Figure 8, a Truffle management code shows the Append and Burn functions in JavaScript programming language. Append considers that an update operation is represented by appending a new version of the data. This way, the latest version represents the current state of the data and the entire version history is maintained in the Blockchain through transaction history. The Burn operation aims at deleting an asset by transferring it to an unspendable public key. An artificial public key is generated, not knowing the private key, so, it is infeasible to redeem the asset.

Figure 8(a) presents the Truffle’s management operations for creating an asset of “owner” by the append operation. updatedAsset contains the last (unspent) state of the asset so any actions need to be done to updatedAsset. Figure 8(b) presents the Truffle’s management operations for deleting a created asset by the burn operation. For the burning process, the asset is transferred to the burn address. Since the private key is not known, it is not feasible to recover the asset.

Participants used the deployed well-being services for 10 days. Basically, during this time, participants were assisted by well-being services in order to improve their motivation and the efficiency of their actions. To do that, well-being services monitored both elements (participants’ motivation and actions) and modified the environmental conditions and actuators to improve the life conditions.

After 10 days using the deployed well-being services and being informed the control and pilot groups about the behavior and functioning of both architectures (the proposed and the traditional one), all of them answered the same survey.

With respect to the validity threats of the experiment, the following observations must be made.

The most relevant threat to the internal validity is the great variability of human behavior. All participants were volunteers so the obtained results are not affected by differences in motivation, although they may be more open to the benefits of technology than standard users. All participants finished the experiment and all of them have the same educational level, so internal validity threats caused by differences in the technical abilities of participants are reduced as much as possible.

The most important external validity threats are caused, on one hand, by organic factors and, on the other hand, by situational and novel factors. As the main organic factor to be considered, all participants had a technical degree, so non-expert people may present a slightly different performance. Besides, the experiment was developed in a laboratory, so people can behave in a different manner than in real daily life. Finally, novel factors may threaten the external validity as it was the first time participants used the proposed system, and they did it for a limited time. Long-term results could be different.

Contract validity is threatened by the impossibility of obtaining or measuring the real sensations or feeling of participants in a quantitative manner. In this case, we are using surveys to sense the participants’ opinions. As people are more biased if they are aware of the expected outcome, the goal of the experiment was kept hidden from the participants.

Data gathering: test variables. In this first experiment, a survey about the proposed technology was answered by each participant in the pilot group, and the same survey
(but referred to traditional service provision mechanisms) was answered by each participant in the control group. The survey consisted of 25 questions about the satisfaction and privacy level perceived by users when using the corresponding well-being services. In general, the objective is to evaluate the QoE of users in both groups (pilot and control groups). Existing mechanisms to evaluate the relation between technology and people, such as TAM (technology acceptance model) questionnaires, are not adequate for this experiment as they do not evaluate the variables we are trying to learn about: on one hand, the service performance in terms of QoE and, on the other hand, the perceived security level by users with respect to their personal data. Thus, a new and specific questionnaire is proposed for this article.

In order to guarantee that the employed evaluation mechanism is comparable to those proposed in previous works, some questions in this survey were extracted from previous papers, and others were similar to works in the state of the art. In any case, in this article, a statistical test is also provided to guarantee the reliability of the proposed evaluation mechanism.

**The second experiment**

The second experiment is employing, as well, the described system in this article as the main material. In this section, we present all information about this second and last experiment.

**Experimental setup.** The final objective of this experiment was to answer some questions regarding the effectiveness of the proposed technological solution, in terms of the reached improvement in QoS. One research question was formulated:

\[ Q2. \text{ Is the processing delay of the proposed mechanism reduced enough to allow the proposed solution to react before the user needs change and get unsatisfied?} \]

**Method and participants.** In order to evaluate the performance of the proposed solution in terms of QoS, a new validation scenario was built. In this second scenario, well-being services, Blockchain networks, fog nodes, and edge devices were implemented into low-cost general-purpose physical devices such as Raspberry Pi computers.

In this second experiment, once management operations and SCs were tested, the infrastructure was implemented in a real Ethereum environment. We used the Ethereum implementation provided by the Geth (Geth homepage: https://geth.ethereum.org/) client, version v1.8.13, for Linux/Arm and installed it in 4 Raspberry Pi 3 Model B. In order to validate the proposed contributions, a prototype was built consisting of four Raspberry Pi 3 M006Fdel B, connected to a server through wired local area network (LAN) connection. The hardware prototype is shown in Figure 9.

APIs communicating legacy, new and remote systems to the trustworthy data protection system were implemented in Node.js, employing the web3.js (Web3.js homepage: https://github.com/ethereum/web3.js/) Ethereum JavaScript API, a collection of libraries allowing the interaction with a local or remote Ethereum node, using an HTTP, WebSocket, or IPC (interprocess communication) connection. Figure 10 presents an implementation diagram of the proposed solution.

Each API acts as a Blockchain account, which receives data from HTTP/HTTPS connections and executes SC functions or listens to events emitted by SCs. Every transaction or SC method call which involves a change in the Blockchain state will return a JavaScript promise, which must be managed accordingly. Also, a callerID (identification of the user invoking the API) is mapped to each variable for session recognition, to avoid the well-known SC’s concurrency problem.

Using the system implementation shown in Figure 10 and mentioned technologies and approaches, the processing delay in services implementing the proposed data protection mechanism was evaluated and compared to the values obtained from services following the traditional approach. In order to evaluate the processing delay, a specific software module, executed in Java SE Runtime Environment 1.9, was deployed in final nodes (biomedical devices and other similar instruments). This module monitored the time between the initial request for a service interaction or transaction and the reception of the final response. For each of those events, a timestamp was stored in a log file. Thus, we guarantee that the calculated processing delay includes all processing tasks at all layers. Finally, the log file was analyzed using the MATLAB software.

The validation described in this second experiment was planned, guided, monitored, and evaluated by its
The experiment was repeated for different block generation rates in the Ethereum network. In order to remove exogenous effects, for each block generation rate the experiment was repeated five times. To evaluate each case, operation data were captured for 24 hours. Then, the final result for the processing delay for each block generation rate is obtained as the mean value of all the obtained measures during those five experiments. Equally, the entire experiment was repeated for different numbers of users in the evaluation environment.

In order to simplify the logistics associated to this experiment, users in this second experiment were simulated. These users were technologically configured as autonomous agents. To support this scheme, in a central server, a collection of resource-constrained virtual machines was deployed to generate the personal data to be protected by our solution and consumed by well-being services. Each machine was provided with only 50 MB RAM memory. In order to configure and manage in an easy manner the large number of virtual machines to be employed, the LXC (Linux Containers) technology and the libvirt interface were used. In that way, a simple script may be used to run all the desired instances and connect them to the real Blockchain network and data protection system, as well as to the deployed well-being services.

The simulated scenario represented a house-like scenario, where ADL can be performed. All configurations and decisions about the scenario in this second experiment were equivalent to those taken in the first experiment. Three different tests were conducted in this second experiment. In the first test, 10 different users (virtual machines) were deployed and run. These users simulated ten family members living in the same house. In the second test, 40 people were simulated (and the same number of machines deployed). This second scenario represented a medium-sized office, which is another relevant scenario for individual well-being services. The last and final test represented a large industrial area, where up to 100 people are around. Thus, 100 virtual machines were deployed. In the largest scenarios, several servers and service instances are usually deployed.

Finally, some considerations about the validity threats of the experiment must be made. First, internal validity is mainly threatened by changes in the operating system, which tends to be slower as the time passes, and the numerical precision, which may affect mathematical calculations (such as block mining). In order to reduce this risk, simulations were repeated several times, and mean values were obtained. And, second,
with respect to external validity, the obtained results are clearly dependent on the underlying hardware (situational factors). In the proposed experiment, a standard hardware architecture was employed, but (in the future) advanced architecture may generate slightly different results.

Data gathering: test variables. In the second experiment, the simulated users (agents) in the service provision platforms were provided with a new and embedded module being able to calculate the processing delay for each well-being service request. These agents and modules were connected to a virtual central server, where data were received, stored, and processed. Basically, each service request generated a timestamp indicating the beginning of the processing tasks. Together with this timestamp, the identifier of the service request was stored. The same process was repeated for each service reply. Thus, through simple mathematical operations, different measures for the well-being service processing delay were obtained. As said before, the final results were obtained as the mean value of all the obtained measures.

Results and discussion

In this section, we present and discuss the obtained results from the experiments described in the previous section. Besides, we elaborate on the reasons and implications of the experimental results, from the point of view of well-being services in an IoT environment, comparing the obtained results to other studies in the state of the art.

Results

The proposed survey to be answered by users included 25 questions. Fifteen questions were about the user experience with the deployed well-being services, and 10 additional questions were about the perceived privacy level.

All questions were answered with a number, according to a Likert-type scale with five elements: representing the minimum mark (one) “I totally disagree,” and the maximum mark (five) “I totally agree.” The complete scale is presented as follows:

1. I totally disagree;
2. I disagree;
3. I neither disagree nor agree;
4. I agree;
5. I totally agree.

All the participants were treated anonymously by experts. No personal data related to the participants’ identification were stored or transmitted outside the research team. All the experiments involving users were performed under the conditions of respect for individual rights and ethical principles that govern research involving humans.

A description of all the obtained answers is presented in Figure 11. In order to make it easier to understand the reported answers, mean values and variances have been rounded with a precision of 0.5 units. Using also these answers, and before going any further, we evaluate the reliability of the employed mechanism.

In order to evaluate the reliability, we use Cronbach’s alpha test. As two different variables are being evaluated (the service performance in terms of QoE and the perceived security level by users), two
different tests must be conducted. In this case, both tests were calculated from variances in data. For this purpose, reports from the pilot and control group were aggregated. Table 3 shows the obtained results for both tests. As can be seen, both values are near 0.9, which indicates a high level of reliability and consistency. For these initial experiments, we consider these values enough (although future and advanced works should try to increase these values for the alpha parameter).

As the captured data from surveys in the control and pilot groups present overlapped statistical distributions (see Figure 11), it is not possible to scientifically and directly determine if there is a global and relevant improvement in user experience using the proposed architecture. A statistical test is, then, needed. While parametric tests require guaranteeing some conditions about the probability distribution of the relevant variables in the population under study, in this case, we cannot reach that level of knowledge and measurement. We have no studies or information to fix coherent and reasonable assumptions about the mean, variance, and so on of the QoE or security level perceived by users in the context of well-being services. In this context, non-parametric tests are the most adequate instrument. Therefore, we employ the Mann–Whitney U test to evaluate this improvement. The p value indicates the significance level of the Mann–Whitney U test. Different tests for different significance levels (alpha parameter) were conducted. Significance levels have been selected to be the most usual and standard in the state of the art (among works related to similar topics and experiments). Table 4 shows the results of this initial experiment.

As can be seen, most answers (up to 11 from a total of 15) in questions about the well-being service quality (QoE or user satisfaction) do not present a statistically relevant difference. In other words, although the service architecture has been totally modified and new layers (such as the edge or fog layer) have been included, the users’ experience (QoE) has not been affected. Even, in some particular areas, such as the reaction time (Q3 and Q12), the proposed new architecture based on edge computing generates a better experience than previous and traditional approaches. Moreover, in questions related to unobtrusiveness (Q7) and discretion (Q15), users have reported also an improvement when well-being services are supported by edge computing technologies and the proposed data protection system.

However, questions about privacy and data protection present a totally different behavior. In all of them, users in the pilot group (employing well-being services supported by edge computing technologies and the proposed data protection system) have reported a much greater privacy and security feeling (regarding their personal data) in well-being services implementing the proposed data protection system. Considering the proposed architecture, this higher perceived security level (regarding personal data) may be explained by the transparent logic and permission that control access to data; by the ability of users to monitor and control transactions involving their data; and by the possibility of every participant to access the data being collected. Although the deployed well-being services (for human motivation and activity monitoring) collect relevant personal data (biological and social), the proposed mechanism helps users feel secure, thanks to the trust provision solutions supported by Blockchain networks.

Differences are especially statistically relevant in questions about access control to the personal data (by third parties, remote experts, etc.) and with respect to the user rights described in regulation (Q17). As a conclusion, although the user experience is not affected by the employment of the innovative proposed service architecture, the perceived privacy and data security levels are much higher.

Once the improvement in the perceived privacy level by users was proved, a more technological evaluation was carried out using a real Blockchain network, as mentioned in section “Implementation and experimental validation.” To obtain relevant results, the processing delay of the proposed technology is compared to those in traditional well-being service provision schemes. In particular, in these traditional schemes, all data are sent directly to a cloud server where they are processed, and the corresponding reply or action is calculated and sent back (contrary to the proposed edge computing approach where data are processed step by step at different levels). As large amounts of raw personal data may be required by cloud servers in this approach, a bottleneck may be originated in the access point to the well-being service, if appropriate configuration and traffic engineering policies are not implemented.

Figure 12 shows the results of the second experiment, for processing delays in well-being services consumed by 10 people.

As can be seen, when services are being consumed by a reduced number of people, traditional architectures present a lower processing delay. In fact, although edge computing schemes improve the processing delay in comparison to traditional architectures, the block generation rate in standard Blockchain networks (what determines how fast transactions are made through these networks) is not lower than 1 second, so the global delay...
In the proposed well-being service architecture is higher. Besides, for small amounts of data, it is still more efficient to send all raw data directly to the cloud server than processing the information at different steps. In fact, although the block generation could be reduced as much as desired, in order to preserve the properties of Blockchain systems, it is necessary to introduce a difficulty parameter in the consensus algorithm which forces miners to repeat the same calculation several times until it fulfills the requirements to be accepted. This time, in any case, is acceptable for general-purpose and common scenarios such as smart homes or enhanced living environments. However, other applications, especially in critical infrastructures or situations, may benefit in a limited manner of this scheme. Furthermore, real-time applications are

| Indicator | ID | Question | p value | Significance $\alpha = 0.05$ | Significance $\alpha = 0.005$ | Significance $\alpha = 0.001$ |
|-----------|----|----------|---------|-------------------------------|-------------------------------|-------------------------------|
| Satisfaction | Q1 | My experience with well-being services was good. | 0.0814 | n.s. | n.s. | n.s. |
| | Q2 | I felt my well-being increases with these services. | 0.0905 | n.s. | n.s. | n.s. |
| | Q3 | The system reacted quickly and adequately to my needs. | 0.0127 | * | n.s. | n.s. |
| | Q4 | My daily living was not disturbed by well-being services. | 0.0913 | n.s. | n.s. | n.s. |
| | Q5 | I would like to repeat the experience. | 0.0632 | n.s. | n.s. | n.s. |
| | Q6 | I will recommend the experience to my colleagues. | 0.0975 | n.s. | n.s. | n.s. |
| | Q7 | Well-being services and sensing devices were unobtrusive. | 0.0027 | * | n.s. | n.s. |
| | Q8 | The way in which well-being services addressed my problems was adequate. | 0.0546 | n.s. | n.s. | n.s. |
| | Q9 | I felt well-being services were a professional and commercial product. | 0.0957 | n.s. | n.s. | n.s. |
| | Q10 | Provided services, in general, had a good quality. | 0.0964 | n.s. | n.s. | n.s. |
| | Q11 | Compared to other well-being services, in this experiment services showed, at least, the same performance. | 0.1576 | n.s. | n.s. | n.s. |
| | Q12 | Compared to other well-being services, services in this experiment reacted, at least, as fast as the other services. | 0.0097 | * | n.s. | n.s. |
| | Q13 | In this experiment, I do not miss any benefit provided by other well-being services. | 0.0957 | n.s. | n.s. | n.s. |
| | Q14 | In its current form, the well-being deployment is adequate. | 0.0585 | n.s. | n.s. | n.s. |
| Security | Q15 | Well-being services are discreet in the manner they help users. | 0.0080 | * | n.s. | n.s. |
| | Q16 | I trust my personal data were private and secure. | 0.0014 | * | * | n.s. |
| | Q17 | I have guarantees about the respect to my rights as a user in relation to my private data. | 0.0004 | * | * | * |
| | Q18 | I am totally sure my data are never sent to third parties. | 0.0009 | * | * | * |
| | Q19 | I could verify how my data are protected. | 0.0049 | * | * | n.s. |
| | Q20 | I could apply to my personal data all privacy restrictions or permission I desired. | 0.0035 | * | * | n.s. |
| | Q21 | I worried about the use other people could make with my personal data after the experiment. | 0.0026 | * | * | n.s. |
| | Q22 | I am satisfied about where my data are processed and analyzed in well-being services. | 0.0035 | * | * | n.s. |
| | Q23 | Compared to other well-being services, in this experiment I felt my data are more private and protected. | 0.0049 | * | * | n.s. |
| | Q24 | I would like the well-being services to implement additional data protection policies. | 0.0046 | * | * | n.s. |
| | Q25 | In this experiment, I felt more comfortable about remote experts accessing my personal information. | 0.0006 | * | * | * |

n.s.: not significant.
* Significant.
feasible, but the maximum tolerated delay must be considered before implementing the proposed solution.

On the contrary, as more people are using the deployed services, for example, 40 users (see Figure 13), an improvement caused by the local computation of service results in edge computing schemes is more relevant. Potential bottlenecks caused by a large amount of raw data being concentrated in the same point in the cloud server (as described before) tend to disappear. In this case, two regions may be distinguished. For block generation rates below 5 seconds, the proposed architecture presents a lower processing delay. However, once this threshold is overcome, the delay caused by the block generation rate in the Blockchain network dominates the system behavior and the processing delay grows up linearly with that rate.

Usually, public Blockchain networks present a block generation rate around 14 seconds. For this value, the proposed architecture presents a globally lower processing delay if, at least, 100 people consume the deployed well-being services (see Figure 14). In this case, the aggregated processing delay for all people is around 26 seconds for traditional architectures, while, if block generation rates are slightly reduced, this amount may be decreased using the proposed edge computing architecture.

Finally, as a summary for all tests in the second experiment, we describe the growth of processing delay. Regarding time complexity, traditional architectures are independent of the block generation rate, as the “block” concept is not being used. Thus, they show an order of $O(1)$ (in Big O notation (http://web.mit.edu/16.070/www/lecture/big_o.pdf)), implying that the time to complete the experiment is unrelated to the problem size. On the other hand, we calculate a delay in the proposed scheme of $O(\log n)$. This result is coherent with current Blockchain proposals\(^49\) where full nodes only need to perform $O(\log n)$ operations per block, with the help of data availability proofs.

**Discussion**

The proposed solution requires a total change in the well-being service provision architectures. In particular, data analysis tasks are distributed among all layers in the system, and parametric service results are sent from the cloud servers to the user devices to be resolved using the corresponding personal data. This new approach requires all components in the system to improve their
complexity and computational power: the edge and sensing elements because they must, now, perform some data processing activities; and the cloud servers because the creation of service parametric results is a complicated mission.

Besides, in order to allow a transparent and auditable privacy management, the proposed well-being service architecture includes a Blockchain network, as other recent proposals.\textsuperscript{30,31} This introduces a problem for standard users, as they cannot manage that kind of network. In that way, a new role focused on managing the Blockchain system and providing privacy is introduced. Furthermore, this scheme may prevent the proposed scheme to be employed in critical scenarios or real-time applications (because of the mining block delay). This situation could affect the implantation of IoT technologies in certain scenarios, such as critical medical alerts. In previous proposals,\textsuperscript{8} these networks are maintained by the community of agents taking benefits of them, contrary to the current work where service providers must enhance their services with these new functionalities.

Selected experiments to evaluate the performance of the proposed solution are focused on domestic scenarios,\textsuperscript{8} where well-being services are an additional value but not an essential element. In this situation, studies have shown that users do not perceive any difference, in terms of satisfaction, between well-being services provided through traditional architectures and services provided through the proposed new scheme. Nevertheless, the perceived privacy level is much higher in solutions based on the proposed mechanism. This first experiment shows that the proposed scheme is an adequate approach, at least in common domestic scenarios. The same proposal may be employed in other relevant scenarios, reported in the state of the art, related to important personal data, such as eHealth.\textsuperscript{21} Additional experiments might be required to evaluate if users feel satisfied in the same manner in other scenarios where privacy is not a critical issue for them (e.g. in critical situations).

In any case, the proposed mechanism in this article is the first that allows a totally personal management of privacy. Other works on privacy preservation for IoT applications\textsuperscript{50} understood privacy as anonymous data, and their proposal is focused on trustworthy algorithms for data anonymization. However, this approach is not completely adequate for well-being services, where responses are personal and depend on the individual being sensed. Besides, sometimes, specific hardware devices are required\textsuperscript{32} which in our case are not needed.

On the other hand, the Blockchain network, as mentioned, introduces a minimum processing delay which may prevent well-being services to work in strict real time. Nevertheless, most of the security problems identified in other relevant applications, for example, those based on data mining,\textsuperscript{51} are totally addressed and solved using the proposed mechanism. In standard domestic scenarios, moreover, the introduced processing delay by Blockchain is negligible, although in other situations (medical emergencies) it may be relevant.

Finally, additional experiments would be required to evaluate the influence of the community in the perceived privacy level, and the privacy management performed by users. Some works focused on social networks and privacy preserving\textsuperscript{52} have proved that close friends and social communities affect these indicators, so future experiments should evaluate the proposed solution in established social groups, instead of in groups of random people. Current experiments cannot show these effects.

**Conclusion and future work**

In this article, a new trustworthy personal data protection mechanism is proposed for well-being services supported by edge computing and IoT devices, based on PbD technologies. This new mechanism is based on Blockchain networks and indirection functions and tokens. Blockchain networks execute transparent SCs, where users’ rights are codified, and store the users’ personal data which are never sent or given to external services. Besides, permissions and privacy restrictions designed by users to be applied to their data and services consuming them are also implemented in these SCs.

A new architecture following the PbD principles and edge computing paradigm is proposed for well-being services. Moreover, a cryptographic framework to manage authorizations and the parametric resolution of services is also presented.

The provided experimental validation shows that user experience is not affected by the use of this new architecture and even sometimes is improved. However, users feel a much greater privacy level using the proposed data protection system. Technically, the proposed architecture may present a much lower processing delay for large communities of users.

Future work will consider Blockchain networks implementing other consensus mechanisms (such as proof of stake) in order to reduce the processing delay of the proposed data protection mechanism for small communities of users.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research leading to these results has received
funding from the Ministry of Science, Innovation and Universities through the COGNOS project, and is also based on VACADENA (No. RTC-2017-6031-2) project ideas.

**ORCID iDs**

Borja Bordel https://orcid.org/0000-0001-7815-5924
Ramón Alcarria https://orcid.org/0000-0002-1183-9579

**References**

1. Cui X. The Internet of Things. In: Moran S (ed.) *Ethical ripples of creativity and innovation*. London: Palgrave Macmillan, 2016, pp.61–68.
2. Bordel B, Alcarria R, Robles T, et al. Cyber-physical systems: extending pervasive sensing from control theory to the Internet of Things. *Pervas Mobile Comput* 2017; 40: 156–184.
3. Bordel B, Iturrioz T, Alcarria R, et al. Provision of next-generation personalized cyber-physical services. In: *Proceedings of the 2018 13th Iberian conference on information systems and technologies (CISTI)*, Caceres, 13–16 June 2018, pp.1–6. New York: IEEE.
4. Satyanarayanan M. The emergence of edge computing. *Comput* 2017; 50(1): 30–39.
5. Bordel B and Alcarria R. Assessment of human motivation through analysis of physiological and emotional signals in Industry 4.0 scenarios. *J Ambient Intel Humanized Comput* 2017; 29: 1–21.
6. Voigt P and Von dem Bussche A. The EU General Data Protection Regulation (GDPR): a practical guide. 1st ed. Cham: Springer, 2017.
7. O’Connor Y, Rowan W, Lynch L, et al. Privacy by design: informed consent and internet of things for smart health. *Proc Comput Sci* 2017; 113: 653–658.
8. Alcarria R, Bordel B, Robles T, et al. A blockchain-based authorization system for trustworthy resource monitoring and trading in smart communities. *Sensors* 2018; 18(10): 3561.
9. Robles T, Bordel B, Alcarria R, et al. Blockchain technologies for private data management in AmI environments. *Multidiscip Digital Publish Inst Proceed* 2018; 2(19): 1230.
10. Yüksel B, Küpçi A and ÖzkasapÖ. Research issues for privacy and security of electronic health services. *Future Gener Comput Syst* 2017; 68: 1–13.
11. Huckvale K, Prieto JT, Tilney M, et al. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med* 2015; 13(1): 214.
12. Kotz D, Gunter CA, Kumar S, et al. Privacy and security in mobile health: a research agenda. *Computer* 2016; 49(6): 22–30.
13. Jiang Q, Khan MK, Lu X, et al. Privacy preserving three-factor authentication protocol for e-Health clouds. *J Supercompu* 2016; 72(10): 3826–3849.
14. Ben-Assuli O. Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments. *Health Policy* 2015; 119(3): 287–297.
15. Pika A, Wynn MT, Budiono S, et al. Privacy-preserving process mining in healthcare. *Int J Environ Res Public Health* 2020; 17(5): 1612.
16. Zhang Y, Zheng D and Deng RH. Security and privacy in smart health: efficient policy-hiding attribute-based access control. *IEEE Int Thing J* 2018; 5(3): 2130–2145.
17. Qian H, Li J, Zhang Y, et al. Privacy-preserving personal health record using multi-authority attribute-based encryption with revocation. *Int J Inform Secur* 2015; 14(6): 487–497.
18. Esposito C, De Santis A, Tortora G, et al. Blockchain: a panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Comput* 2018; 5(1): 31–37.
19. Azaria A, Eklaw A, Vieira T, et al. MedRec: using blockchain for medical data access and permission management. In: *Proceedings of the 2016 2nd international conference on open and big data (OBD)*, Vienna, 22–24 August 2016, pp.25–30. New York: IEEE.
20. Yue X, Wang H, Jin D, et al. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. *J Med Syst* 2016; 40(10): 218.
21. Eklaw A, Azaria A, Halamka JD, et al. A case study for blockchain in healthcare: “MedRec” prototype for electronic health records and medical research data. In: *Proceedings of the 2016 2nd international conference on open and big data (OBD)*, Vienna, 22–24 August 2016, p.13. New York: IEEE.
22. Bravo J, Hervás R, Fontecha J, et al. m-Health: lessons learned by m-Experiences. *Sensors* 2018; 18(5): 1569.
23. Ichikawa D, Kashiyama M and Ueno T. Tamper-resistant mobile health using blockchain technology. *JMIR Mhealth Uhealth* 2017; 5(7): e111.
24. Patel V. A framework for secure and decentralized sharing of medical imaging data via blockchain consensus. *Health Inform J* 2019; 25: 1398–1411.
25. Ribitzky R, Clair JS, Houlding DI, et al. Pragmatic, interdisciplinary perspectives on blockchain and distributed ledger technology: paving the future for healthcare. *Blockchain Healthc Today* 2018; 1: v1.24.
26. Hoy MB. An introduction to the blockchain and its implications for libraries and medicine. *Med Ref Serv Quart* 2017; 36(3): 273–279.
27. Nugent T, Upton D and Cimpoesu M. Improving data transparency in clinical trials using blockchain smart contracts. *F1000Res* 2016; 5: 2541.
28. Wang Z, Ding H, Han J, et al. Secure and efficient control transfer for IoT devices. *Int J Distrib Sens Netw* 2013; 9(11): 503404.
29. Kadiyala ST, Sowdhari A, Valina S, et al. Security/privacy in health care monitoring using wireless sensor networks. *UB ScholarWorks*, 2017, https://scholarworks.bridgeport.edu/xmlui/handle/123456789/1881
30. Bordel B, Lebigot P, Alcarria R, et al. Digital food product traceability: using blockchain in the international commerce. In: *Proceedings of the 2018 international conference on digital science.* Budva, 19–21 October 2018, pp.224–231. Cham: Springer.
31. Bordel B, Martin D, Alcarria R, et al. A blockchain-based water control system for the automatic
management of irrigation communities. In: Proceedings of the 2019 IEEE international conference on consumer electronics (ICCE), Las Vegas, NV, 11–13 January 2019, pp.1–2. New York: IEEE.

32. Pérez-Jiménez M, Sánchez BB, Migliorini A, et al. Protecting private communications in cyber-physical systems through physical unclonable functions. *Electronics* 2019; 8(4): 390.

33. Bordel B, Alcarria R, Sánchez-de-Rivera D, et al. Protecting industry 4.0 systems against the malicious effects of cyber-physical attacks. In: *Proceedings of the international conference on ubiquitous computing and ambient intelligence*, Philadelphia, PA, 7–10 November 2017, pp.161–171. Cham: Springer.

34. Bordel B, Alcarria R, Robles T, et al. Stochastic and information theory techniques to reduce large datasets and detect cyberattacks in ambient intelligence environments. *IEEE Access* 2018; 6: 34896–34910.

35. Bender JL, Cyr AB, Arbuckle L, et al. Ethics and privacy implications of using the internet and social media to recruit participants for health research: a privacy-by-design framework for online recruitment. *J Med Intern Res* 2017; 19(4): e104.

36. Becker J, Breuker D, Heide T, et al. Can we afford integrity by proof-of-work? Scenarios inspired by the Bitcoin currency. In: Böhme R (ed.) *The economics of information security and privacy*. Berlin: Springer, 2013, pp.135–156.

37. Christidis K and Devetsikiotis M. Blockchains and smart contracts for the Internet of Things. *IEEE Access* 2016; 4: 2292–2303.

38. Cavourian A. *Privacy by design: the 7 foundational principles: information and privacy*. Toronto, ON, Canada: Commissioner of Ontario, 2009, p.5.

39. Bordel B, Alcarria R, de Rivera DS, et al. Process execution in cyber-physical systems using cloud and cyber-physical internet services. *J Supercomput* 2018; 74(8): 4127–4169.

40. Bos JW, Costello C, Longa P, et al. Selecting elliptic curves for cryptography: an efficiency and security analysis. *J Cryptograph Eng* 2016; 6(4): 259–286.

41. Schnorr CP and Jakobsson M. Security of signed ElGamal encryption. In: *Proceedings of the international conference on the theory and application of cryptography and information security*, Kyoto, Japan, 3–7 December 2000, pp.73–89. Berlin: Springer.

42. Bellare M and Kohno T. Hash function balance and its impact on birthday attacks. In: *Proceedings of the international conference on the theory and applications of cryptographic techniques*, 2–6 May 2004, pp.401–418. Berlin: Springer.

43. Bordel Sánchez B, Pérez Jiménez M and Sánchez de Rivera D. Recognition of activities of daily living in enhanced living environments. *IT Converg Pract* 2017; 4(4): 18–31.

44. Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 914–919.

45. Liu Y, Gummadi KP, Krishnamurthy B, et al. Analyzing Facebook privacy settings: user expectations vs. reality. In: *Proceedings of the 2011 ACM SIGCOMM conference on internet measurement conference*, November 2011, pp.61–70. New York: ACM.

46. Yang Z, Cai S, Zhou Z, et al. Development and validation of an instrument to measure user perceived service quality of information presenting web portals. *Inform Manag* 2005; 42(4): 575–589.

47. Nemoto T and Beglar D. Likert-scale questionnaires. In: *Proceedings of the JALT 2013 conference proceedings*, Kobe, Hyogo, Japan, 25–28 October 2013, pp.1–8. Japan: JALT Publications.

48. Cowles M and Davis C. On the origins of the .05 level of statistical significance. *Am Psycholog* 1982; 37(5): 553–558.

49. Yu M, Sahraei S, Li S, et al. Coded Merkle tree: solving data availability attacks in blockchains. Cryptology Eprint Archive, Report 2019/1139, https://eprint.iacr.org/2019/1139

50. Xu L, Jiang C, Wang J, et al. Information security in big data: privacy and data mining. *IEEE Access* 2014; 2: 1149–1176.

51. Du J, Jiang C, Chen KC, et al. Community-structured evolutionary game for privacy protection in social networks. *IEEE T Inform Forensic Secur* 2017; 13(3): 574–589.