Designing and Validating a Blockchain-based Architecture to Enforce Privacy in Human Robot Interaction

Viktor Vasylkovskyi
INESC-ID, Instituto Superior Técnico, University of Lisbon, Portugal
viktor.vasylkovskyi@tecnico.ulisboa.pt

Abstract

Social Robots (SR) can record large streams of raw data in the form of images, audio, Radio Frequency Identification (RFID), among other sensory sources, which could be instrumental in enforcing Human-robot interactions (HRI). However, the emotional bonds between humans and SR can raise an SR accessing/inferring profoundly private information, e.g., emotional states. Therefore, in what concerns privacy, SR may be a liability. Clarifying the ownership of data collected by robots is an issue of concern in the European Union (EU) General Data Protection Regulation (GDPR), which requires 'privacy-by-design'. The paper presents BlockRobot – a Blockchain (BC)-based software design model for developing privacy-by-design application in Human-Robot Interactions (HRI). Based on this design model, users can access the private data generated in their HRI. This paper's contributions are the design and validation of a model to develop privacy-by-design applications for HRI. This model integrates BC technology with robots to audit robotic events such as Red, Green, Blue, Depth (RGB-D) images, and RFID tags. Smart contracts and crypto-wallets accomplish identity management. To demonstrate the utility of the initial implementation of a Decentralized Application (dApp), we implement a prototype, based on the private permissioned Evolution of Scalability (EOS) Blockchain integrated with the social robot moving in a non-laboratory environment.

1. Introduction

Social robots are designed to assist humans socially. The way a person interacts with a social robot is quite different from interacting with an autonomous robot. People prefer to interact with robots with basic communication and social skills. Social interactions imply the implementation of a new vast broad of features that require multidisciplinary research [1]. Scientists need to understand the philosophical, ethical, and legal layers that motivate typical human behavior. To create social robots capable of genuine social behavior, they need to interact with humans at an emotional level [2] and act socially expected by humans. The long-term goal is to assist humans in tasks, such as education, health, entertainment, communication, and tasks requiring teamwork. On the one hand, SR can collect data from (among many other sensors) RGB and RGB-D images, record audio data, react to human voices, and detect RFID signals. RGB-D images (*cf.* Fig. 1) are depth images in which the value of each pixel in the image represents the distance to the object. They provide the ability to accurately sense and track humans and objects without the possibility of complete identification [3]. RFID signals give information regarding the position, and proximity can be inferred [4–6]. On the other hand, people tend to bond with SR that look like humans and are likely to interact with them at an emotional level [7] (in general, within the bounds of the uncanny valley paradigm [8]). Due to the variety and unprecedented volume of private data that SR can collect, they raise questions about privacy [9] and security [2], subjecting them to GDPR [10, 11] – a European Regulation that imposes legal compliance to all the public and private companies.

In Human-robot interactions (HRI), it is not easy to prove the provenance and continuous compliance with GDPR if end users cannot access the SR data. Moreover, due to the incomplete implementation of the

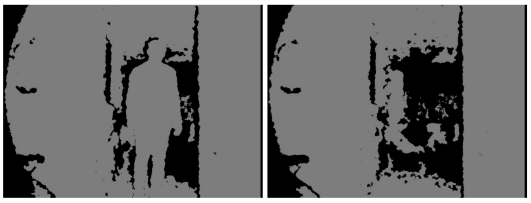

Figure 1: RGB-D Images captured by robot: Left – a person is clearly visible; Right – no person in sight.

GDPR principles, concerns on the lack of transparency may be raised. Besides the privacy enforcing techniques employed by a CA, there is growing attention to the decentralized private data management paradigm. Decentralized solutions such as Blockchain provide high-level properties of **decentralization**, **transparency**, **tamper-resistance**, and **traceability**. For that reason, Blockchain is an excellent candidate to comply with GDPR for HRI. In this paper, we show how BC, when integrated with SR, can improve users' privacy in HRI, resulting in a better experience. Using a case study, we demonstrate the utility of a BC-based architecture encompassing an implementation of a *dApp* based on a private permissioned EOS Blockchain, integrated with an SR. Our solution allows end-users (hospital staff, patients, and visitors) to control private data generated during their interaction with the SR transparently and fairly. The experiments are performed in hospital corridors and rooms. The execution data logs are computed with process mining techniques [12] to analyze the performance and bottlenecks existing in the case study.

The remaining of the paper is structured as follows. Section 2 presents the relevant background. Section 3 describes in detail the proposed architecture and its information workflows. Section 4 illustrates the working example. Section 5 provides discussion and limitations, and Section 6 concludes and points to future work.

2. Background

This section provides a brief introduction to Blockchain and exemplifies related work regarding its benefits for privacy and robotics.

2.1. Blockchain-based Technologies for Privacy

Blockchain is a distributed shared ledger replicated over the network nodes interacting between themselves based on a set of rules known as a consensus system. The primary data structure of a BC is a block encompassing a set of transactions. Each block is chained to the previous block by cryptographic hashes [13], making it virtually impossible to erase BC data. The participants can write to the BC to modify its state by creating transactions, but cannot erase it. There are three variations of BC based on the permission model: **permissionless** – any party can to read from and write to the BC; **permissioned** – access to the BC is restricted to authorized parties, and; **partially permissioned** – some actions, for example, read access, is granted to anyone, but other actions, for example, write access, may be limited to authorized parties.

BC has been implemented in different areas, including improving privacy. In the scope of access control, Zyskind et al. propose a decentralized privacy solution using a Blockchain to enforce a self-sovereign access-control management [14]. Truong et al. propose a BC-based solution to improve the integrity and transparency of data stored by centralized entities [15]. Faber et al. proposed a human-centric and GDPR-compliant personal data and identity management system based on the Blockchain technology [16].

2.2. ROS, RFID and Blockchain in Robotics

Many robots operate on ROS (Robot Operating System) – a middleware to develop robot applications. The ROS contains robotic events stored temporarily in the robot's named structures – ROS topics [17]. The use of ROS when developing robots is a natural choice. It is common for robots operating on ROS to be enabled with RFID sensors to detect RFID tags. RFID technology enables detecting contactless ID tags in scenarios where identification features with minimal false positives are required [6]. When integrated with robots and ROS, these wireless sensors provide information on RFID tags' position [6].

In the literature on HRI, some efforts demonstrate how BC can enhance privacy in HRI. Ferrer et al. [18] use BC technology to improve privacy when sharing sensitive data between robots and machine learning models. Their solution is prominent in increasing transparency in HRI. A research conducted by Degarding and Alexandre [19] shows how Blockchain can assist in auditing robotics events. Fernandes and Alexandre also proposed Blockchain's use to audit robotics events by using Tezos's technology [20]. Their research intends to explore the BC immutability properties to solve the problem of tampering the records of robotic actions in the context of factory robots. Integrating BC with robots seems to improve cooperation by providing the robots with a single immutable source of truth where they can share the knowledge base [21].

With this research in mind, BC seems like an excellent candidate in developing privacy-by-design applications to improve privacy in HRI. *Private permissioned Blockchain* is the most reasonable choice to enhance the privacy of individuals' data because it provides restricted access to the network.

3. Architecture Design

This section describes the design architecture of the *BlockRobot dApp* (Fig. 2). The system is composed of the basic interaction between a robot and a human, starting in the **HRI layer**. The robot is equipped with RGB-D camera – used to record streams of images; and an RFID sensor that can detect RFID tags from which the information about proximity can be inferred [6]. We equip individuals with RFID tags embedded into smart cards, so that robot with RFID laser can detect individuals without the line of sign or ambiguity (with anonymization techniques). After the basic interaction of a human by the robot, both streams of images and the RFID tags are published as separate ROS events. Furthermore, the BlockRobot API can access that data in **API layer** through the ROS Middleware – software that allows us to read robotic events in the *publish-subscribe pattern*. The BlockRobot API is a trusted server that handles data transformations of ROS events into private data, and operates the off-chain storage by synchronizing the on-chain transactions with off-chain database updates. Once the BlockRobot API main algorithms retrieve private data from ROS events, the data can be persisted in **Blockchain and Persistence layer**, which represents the storage of data – through BC and off-chain repository, and the access of data – managed by the smart contracts. To modify BC's state, individuals and robots interact with BC are the peers – by issuing the transactions.

3.1. Design Guidelines

In this section, we provide the main design guidelines to develop privacy-by-design applications for HRI.

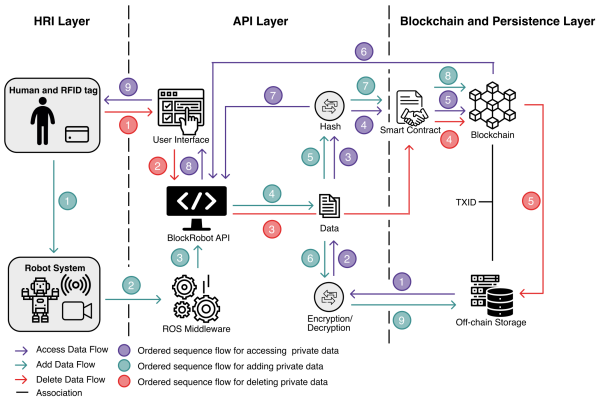

Figure 2: High-level System Architecture of BlockRobot

Correct data outline: *(i)* The communication should be achieved by subscribing to the ROS topics through ROS middleware, both to prevent additional overhead on the robot and obtain all the relevant data for private data classification algorithms, e.g., RFID tags and RGB-D images. *(ii)* The individuals' identities should be anonymized to ensure data confidentiality and security.

Data Persistence *(i)* BC should be used as a trust anchor. In other words, the data to be stored on-chain are the hashes of private data. BC is immutable storage, therefore placing hashes on BC will provide verifiability of data integrity. *(ii)* The private data should be persisted off-chain, e.g., traditional DBMS (*e.g.*, PostgreSQL, MongoDB), a cloud storage (*e.g.*, AWS or Azure), a distributed storage system (*e.g.*, IPFS, [22]). Storage of the private data off-chain allows it to be deleted, as required by GDPR' "Right to be Forgotten" [10, p. 32]. Furthermore, to provide confidentiality of private data, *(iii)* the off-chain data should be encrypted, naturally, with the data owner's keys. As an example, the possible algorithms to consider are Elliptic Curve Cryptosystem or RSA [23]. *(iv)* For faster lookup, the off-chain and on-chain data should be linked, for instance, as in the example further demonstrated in section 4 by Transaction ID (TXID). *(v)* The off-chain and on-chain should be synchronized. Once a new block is validated in the BC, the off-chain database should be triggered to align the change. Accordingly, if a participant or a robot update the BC state, then the off-chain state will update as well. *(vi)* Robots and participants should act as peers in Blockchain to provide the provenance and non-repudiation of private data.

User Interface and BC Transactions *(i)* Private data generated in HRI should be easily retrievable and presented in a human-readable format through the user interface. *(ii)* Once presented with the data, users should be able to access and erase their private data by UI. *(iii)* Users should create a new transaction to the BC per each access or deletion of data. By being an immutable truth source, all the actions can be audited on the BC, thus providing accountability. *(iv)* To prove the user's identity securely, each individual should have a crypto wallet keeping the private and public keys. The crypto wallet should be integrated with the user interface and request a digital signature on each transaction. The digital signature is proof of the individual's identity that needs to be provided on each data's access and erasure.

3.2. Data Subject Classification

During the interactions between a robot and a human, the robot may record large streams of images.

To illustrate, we present the interaction design in Figure 3, where the robot records RGB-D images while a person walks by the robot. A person is equipped with a smart card with an embedded RFID tag. As the interaction goes, and streams of RGB-D images are published on the RGB-D Topic, the RFID laser embedded in the robot detects an RFID tag from the person's smart card. It is worth to note that the RFID tag detection and RGB-D image are asynchronous events because the RFID laser and the kinetic camera are separate devices. Therefore, the data from both events require additional coordination.

RFID tags provide information regarding the proxemics and time relating to the RFID laser. Previous research [6] points out that RFID tags are subject to background noises, depending on the environmental conditions, which may reduce the precision of the time and distance of an individual carrying tags. Still, they provide knowledge that some individuals are in the robot's vicinity, which is valuable for identification purposes. The RFID signal is composed of the *UID* – unique identifier, and the timestamp when the RFID signal was detected. Furthermore, the UID gets anonymized by hashing it, with the intuit to difficult the individual's possible re-identification.

In real-world scenarios, many people may walk by the robot. Therefore, many RFID signals will be identified simultâneously. For that reason, all the anonymized UIDs and timestamp retrieved from each RFID tag are temporarily stored in the off-chain repository. The storage of the UIDs will allow us to consult the information about all the individuals detected in HRI, and consequently, who is present in the RGB-D image recorded at a specific time.

On the other hand, to identify an individual in the network, we persist his/her anonymized UID on the BC (Further explained). Besides being immutable, placing

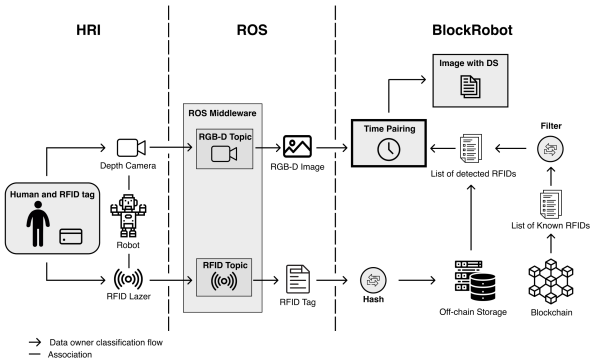

Figure 3: ROS middleware and BlockRobot interactions – Classification of data owner model

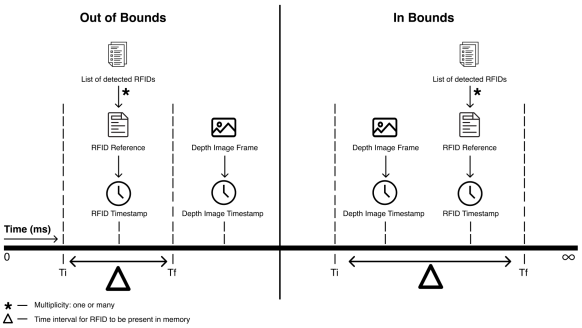

Figure 4: Time Pairing model: Left – Image detected out of bounds of RFID detection; Right – Image detected in bounds of RFID detection

RFID tags on BC provides an indirect relation between known participants and their RFID tags. Finally, the algorithm has enough data to know *(i)* who are the known members of the network – UIDs stored on-chain, and *(ii)* who are the individuals detected in a specific time interval – UIDs and timestamps temporarily stored off-chain. Then two lists of UIDs are reduced into a list of known members detected in a given interval. For each RGB-D image, this list will be used to understand whether an individual in a list is present in the RGB-D image. If yes, then we can create a unique data structure containing the RGB-D image and individuals identified in the image. The details are discussed in the *Time-Pairing Algorithm*.

3.3. Time-Pairing Algorithm

The Time-pairing algorithm consists of examining each RGB-D image and the list of RFID tags, and give as output whether that RGB-D image is a private data or not. If it is a private data, meaning there are individuals identified in the image, it also says who the individuals are. Each RFID detected is a single event in the time-pairing algorithm. Therefore if there are many participants near the robot, many RFID events will be published. One RGB-D image is also an event. Different inputs, such as a list of RFIDs, one RFID, or zero RFIDs collected from events during the HRI, will trigger different outcome scenarios. Furthermore, for different RFIDs detected, based on the timestamps of detection, we estimate how long the participant identified shall be considered to remain close to the robot. Δ **denominates this estimate**.

We define Δ as a certain interval for which the person carrying a tag should be considered in the robot vicinity. For simplicity, here and after, we will refer to the RFID tag detection as the same as the individual

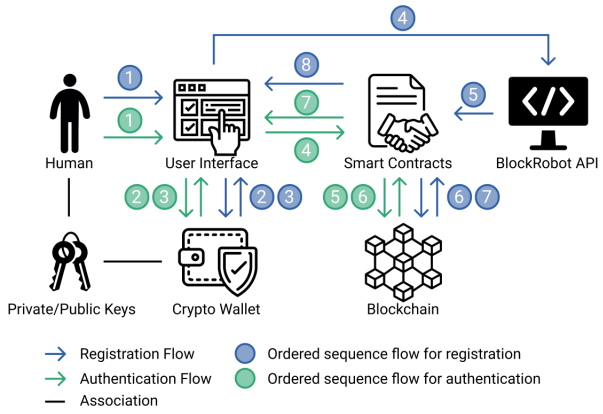

Figure 5: Identity Management Model – Registration and Authentication

being detected. For each individual detected by the robot, we take this event's timestamp and enclose it with Δ, which will give us the interval of time when the person remains near the robot. *Two* possible outcomes may happen (illustrated in Figure 4): *(i)* **The out of bound time-pairing** (Fig. 4 - Left) – In this outcome, the estimated interval of the person's presence near the robot does not intersect with the time when the image is recorded. In this case, the algorithm detects no participant in the RGB-D image, and therefore this is not a private data. *(ii)* **The inbound time-pairing** (Fig. 4 - Right) – Outcome when the estimated time when the robot sees the participant intersects with the timestamp of the RGB-D image recording. As a result, the algorithm outputs that image as a private data belonging to the person identified in the image.

If more that one person is identified during HRI, meaning that some RFID intervals overlap, then more than one participant is in the image. In such a case, the picture is the private data belonging to all these persons, and anonymization algorithms can be applied so that individuals cannot recognize other people in the image. Further, a person can request the deletion of the link between him and the image. Given the sufficient anonymity level, if the individual is no longer recognizable on the image, that image is no longer considered that participant's private data. Once the last participant requests deletion, then the image can be deleted.

3.4. Identity Management with Smart Contracts and UID

Identity management comprises two phases – proof of ownership of the UID (retrieved from RFID tag) and proof of identity by a digital signature. Once proven the identity, a user may access all the images with his/her associated UID. We illustrate the *Registration* and the *Authentication* on Fig. 5.

Registration: to interact with a private permissioned BC user must belong to the network. In our architecture, the proof of membership to the BC is managed by the smart contracts. As a first step, a user should provide his UID and a public key through a user interface and issue a new transaction to the BC with his UID (hashed) and public key as a payload. Further, once the transaction is initiated, the smart contract will require the user to sign the transaction with his/her private key. Once signed, the smart contract performs a public key recovery to prove that the digital signature belongs to the user's public key. Given the cryptographic relation between private and public keys, and a digital signature inferred from the private key, it is possible to recover a public key from the digitally signed transaction. Once the public key is confirmed to belong to the user, and the transaction is valid, the UID with the user's public key will be installed on the Blockchain, and the user can begin the process of authentication for future access the data under that same UID.

Authentication: To access private data, the user needs to authenticate to the network by providing his previously registered UID. Once provided, the UID is hashed, and the smart contract search for the UID on the BC. If any hash of UID installed on BC matches the UID that an individual claims to own, then the smart contract performs a public key recovery algorithm (explained above) to recover a public key from the digitally signed transaction. Once the public key matches the public key installed on the BC under the provided UID, the user is the UID owner. Therefore, the proof of UID ownership is successful, and the user can access his private data.

4. Implementation and case study

As a proof of concept, we established an initial implementation of a BC-based *dApp* based on the design guidelines presented in section 3, integrated with a social robot (Fig. 6) moving in hospital corridors and rooms. Previous experience pointed out that the robotic events happen on the high frequency with potentially thousands of images per minute. EOS Blockchain seemed to be an ideal solution to guarantee the quality of the data and avoid performance bottlenecks due to its Delegated Proof of Stake (DPoS) consensus that supports high transaction speed [24]. As far as we know, we are the first to integrate EOS Blockchain to audit robotic events.

The following experiences intend to test the architecture and to answer the question: *"Is it possible*

for individuals to access their private data generated during their interactions with the robot without any security breach for other participants?" In other words, we want to know whether users can control **only** their data, and personal data is correctly outlined. For the sake of our research, participants should wear wireless RFID tags embedded in smart cards to be correctly recognized by a social robot with an RFID reader. As an experimental protocol, we consider the following three stages. Pre-experiment: *(i)* configuration of the two users smart cards identified by RFID tags, *(ii)* preparation of one social robot with its dedicated camera for image recording and RFID sensor used for detecting and positioning of RFID tags, *(iii)* reset of the Blockchain and database, and create 1 (or 2) wallet key pair(s) (one for each user), *(iv)* register each user in the network with his/her RFID and key pair, *(v)* deploy the BlockRobot API and a Mongo database (off-chain repository). Experiment execution: *(i)* the individuals and the robot act as peers and are connected to the local network, *(ii)* users walk near the robot, and the robot detects a person in less than 5 seconds (in the range of 3 meters). Post-experiment: *(i)* extraction of data logs and process mining techniques are performed to evaluate the system's performance.

4.1. Experiments

We performed many different experiments and summarised them in the four main scenarios. *(i)* One person walking by the robot carrying the RFID card. The robot is recording streams of RGB-D images and detects an RFID tag. At the end of the scenario, the person can access his/her data and visualize the images recorded by the robot. *(ii)* The same scenario as before, with one person and one robot recording RGB-D images and detecting the person with its RFID laser. However, in this scenario, the data is corrupted. An example of

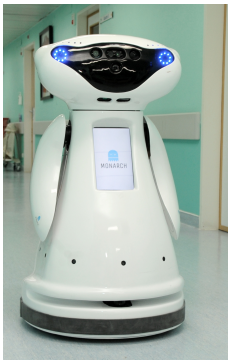

Figure 6: MOnarCH Robot – the social robot used in the experiments

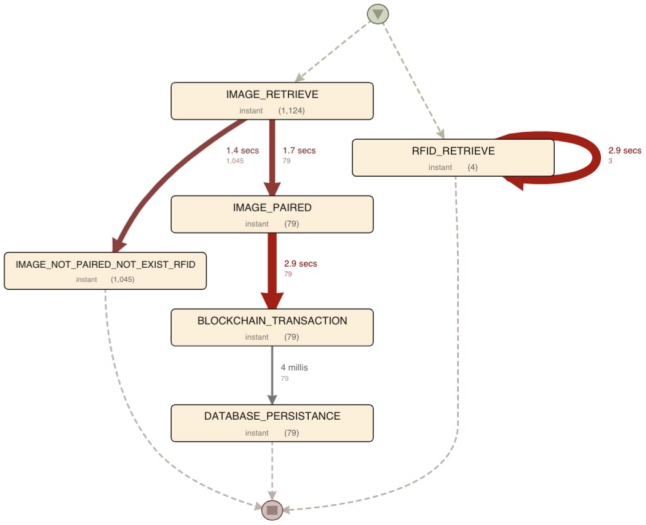

Figure 7: Graph showing the execution of system events during experiment 1.

corrupted data may result from a malicious party gaining access to the off-chain repository and modifying the recorded data. This configuration shows the *Integrity* property of the personal data assured by the proposed framework because of the use of Blockchain as a trust anchor. If some data is corrupted, then the integrity test will fail. Still, the user can access his/her personal data and have the information that the data was corrupted. The sensors' corruption that could occur in the external environment is out of this project's scope. *(iii)* A third scenario manifest slightly different configuration. Here, two users walk by the robot, one at the time, without intersecting their images. This experiment demonstrates the feasibility of the BlockRobot algorithms to outline the private data to the participants correctly. As a result of the experiment, two individuals could access their private data without causing any security breach to the other participant. *(iv)* In the last experiment, two individuals walk by the robot at the same time. In the real-world scenario, this is possibly the most common situation because it is common to have many people walking in the hospital corridors simultaneously. In this scenario, both individuals were identified by the robot, and their RFID intervals intersected, meaning that two participants are identified in the same image. Both can access the images because it is their private data. However, the other's individuals shapes are blurred as an attempt to anonymize them in the image. In such a case, the participants' confidentiality is guaranteed because it is not possible to detect them; hence their identities are covered by applying the anonymization techniques,

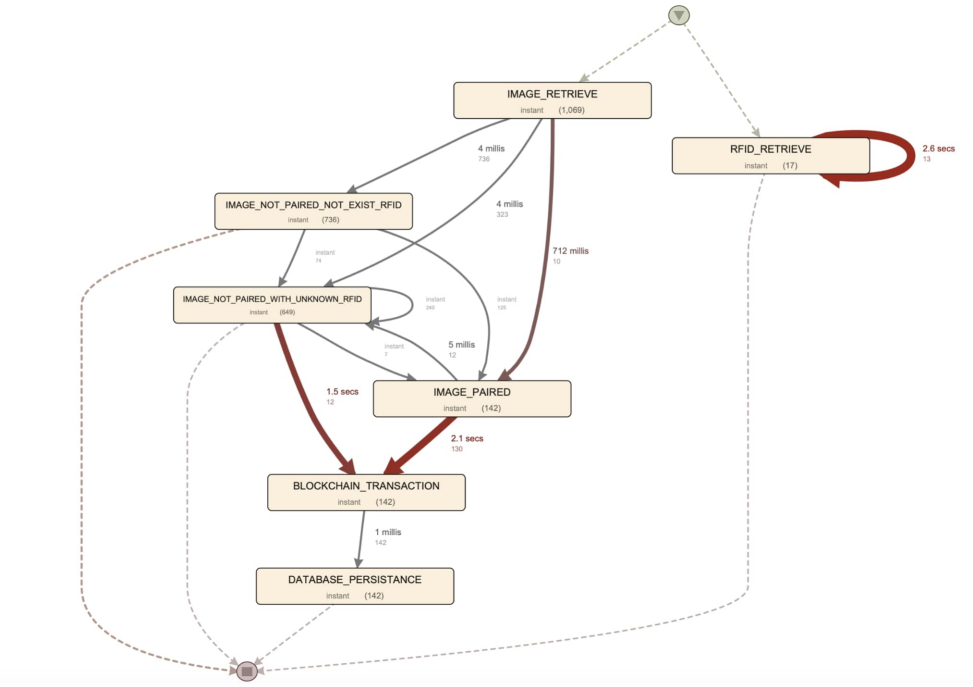

Figure 8: Graph showing the execution of system events during experiment 2.

e.g., blurring faces. Therefore the user can access data or revoke access (delete data) without breaching of data privacy of other users as no information is given about other data subjects' identities in the recording.

4.2. Results

We analyzed the performance of the system using process mining techniques [12]. from the data resulting from two experiments: *(i)* one person and one robot (Fig. 7) and *(ii)* two persons interacting with one robot simultaneously (Fig. 8). Further data is presented in statistical tables in GitHub links: statistical data with one person [1] and with two persons [2]

Each operation corresponds to the following system's event. IMAGE_RETRIEVE – the generation of a new image recorded by the robot; RFID_RETRIEVE – the generation of a new RFID tag recorded by the robot; IMAGE_NOT_PAIRED_NOT_EXIST_RFID – image not associated because no RFID time intersects with image time; IMAGE_NOT_PAIRED_WITH_UNKNOWN_RFID – image not associated with the RFID because it is unknown RFID (non-registered); IMAGE_PAIRED – an image associated with RFID by time intersection; BLOCKCHAIN_TRANSACTION – Blockchain transaction validated and the hash of image is stored in Blockchain; DATABASE_PERSISTANCE – the encrypted image is stored on BC.

By analyzing the logs in both experiments, the most time-consuming event occurs when IMAGE_PAIRED execute successfully, and BLOCKCHAIN_TRANSACTION is triggered – showing that Blockchain transaction plays a cost-effective part of the system. Moreover, all the IMAGE_PAIRED instances have a correspondingly BLOCKCHAIN_TRANSACTION instance; therefore, no transaction has been considered corrupted. In the first experiment, the difference from the activity frequency of IMAGE_RETRIEVE to the frequency of DATABASE_PERSISTANCE shows that only 7% of the images are being paired and stored encrypted. With the increase of RFIDs, in the second experiment, this value increases to 13%. In the second experiment (Figure 8), a more complex network of events shows that in some instances are identified some other RFIDS that were not involved directly in this experiment (*e.g.*, other health care personnel). In those situations, only the identified RFIDs are considered, and a

[1] https://github.com/vvasylkovskyi/eos-web/blob/master/experiments/experiment-1-person.pdf

[2] https://github.com/vvasylkovskyi/eos-web/blob/master/experiments/experiment-2-persons.pdf

BLOCKCHAIN_TRANSACTION triggered.

All the code relative to the current implementation is open-source and available on Github [3] altogether with the datasets from discussed experiences [4], and the detailed diagrams [5].

5. Discussion and Limitations

This section discusses the proposed solution, including the privacy validation in HRI, the GDPR compliance, and the Blockchain benefits evaluation and possible security and privacy threats.

5.1. Privacy Validation in HRI

According to *Art. 4* of the GDPR, *"personal data means any information relating to an identified or identifiable natural person (data subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors [...]."*, even when data are encrypted or hashed, it qualifies as personal data under EU law [11].

We analyzed the GDPR legal documents [10,11] and identified the main privacy conditions. We illustrate in Figure 9 the mapping between the GDPR privacy requirements and the BlockRobot design choices. From the related work, it seems clear that currently, SRs support basic privacy requirements such as **principle of purpose limitation**, **data minimization**. The data that the robot collect has the sole purpose of providing suitable interactions. Moreover, the robot knows the minimum amount of information, inferred from its embedded sensors, for the HRI. Another GDPR privacy rule is the **Storage Limitation** – enforced by destroying the data that is no longer necessary [25].

In Social Robotics, from both legal and ethical perspectives, we foresee that users should have *information transparency and fairness*. Fairness is defined by M. Fink in GDPR [10, p. 64] as clarity regarding the purpose of data, meaning that individuals understand the purpose of data collection. Regarding transparency in privacy, the **accountability** and **provenance** of data is required. Individuals should be notified of all data processing activities in the past and the future. The provenance means the non-repudiation of the data – the source of the data should be clear and unquestionable. In the context of social robots, this means that individuals should be assured in some way that the data they have access to is the data resulting from their interaction with that robot.

Results of the experiments in the previous section validate that individuals are given *privacy by opacity* – no individual can access the data that does not belong to him. Moreover, the off-chain private data is encrypted, thus providing the **confidentiality** of the data. Besides, off-chain storage enables individuals with the possibility to **erase the data or revoke access**. Furthermore, by using BC as a trust anchor, it is possible to audit the history of data processing activities in retrospective and grant the participants with data **accountability**, and the possibility to **verify the data integrity** by hashing data. Finally, if robots interact in the BC as peers and issue transactions, the data's **provenance** is guaranteed by design.

5.2. Security and Privacy Threats

Identity management based on RFID and BC is an innovative approach. In BlockRobot, identity management is based on placing the RFID tags on the BC, which any network member can visualize. One of the privacy-sensitive topics to discuss is the degree to which the placing of RFID tags on the Blockchain is secure. If the owner of the RFID tag happens to be discovered, then everybody in the network knows who the individual that has been "seen" by the robot is. However, it is worth noting that they will not access the individuals' images even in his/her identity disclosure. For that reason, it is crucial to anonymize the RFID tags before placing them in BC. In the definition of the GDPR [10], the anonymity is the degree to which it is impossible to re-identify a data subject from the data. We attempt to anonymize that data by hashing the UIDs extracted from the tags; however, a sufficiently motivated malicious entity may re-identify the hash's

Privacy Requirements	Design Guidelines
Purpose limitation	SR moving in environment
Data minimization	Recording of the Depth images
Storage limitation	Destruction of data in SR
Integrity	Hashing of data on-chain
Confidentiality	Encryption of data off-chain
Provenance	SR interacting as a peer in BC
Accountability	Audit history on-chain
Accuracy	Identity Pseudonimyzation/RFID
Access	Identity Management/Crypto Wallet
Withdraw the concent or deletion	Storage of data off-chain

Figure 9: Privacy Validation – GDPR Requirements and BlockRobot Design Guidelines mapping

[3] https://github.com/vvasylkovskyi/eos-web

[4] https://github.com/vvasylkovskyi/eos-web/tree/master/experiments

[5] https://github.com/vvasylkovskyi/eos-web/tree/master/documentation

actual UID. The platform might be prone to pattern analysis despite the use of hashed UIDs. Whether the storage of hashed RFID tags on Blockchain provides the irreversibility in identification is the specific question to each system/business involved, in our case, HRI.

In theory, the same problem may apply to the storage of the hashes of depth images on BC. However, in the case of depth images, it seems highly unlikely that some intruder may reconstruct the depth image by knowing only the hash. Therefore, we assume that the hashes of depth images on BC are a secure option.

Another relevant topic for discussion is the security of using the BlockRobot API as a trusted intermediary. BlockRobot API access to the robotic events transferred through the network is subject to a *Man-in-the-middle* attack. Such danger can be mitigated by delegating the BlockRobot API's computations to the robot wherein the smart contracts execute all the data transformations when social robots issue transactions. Hence, due to the simplification of the experimental environment, BlockRobot API is used as an intermediary to demonstrate the concept's feasibility.

5.3. GDPR and the Right to be Forgotten

GDPR, in its literature [11], explicitly states the obligation of any entity with access to individuals data to provide them with the **"Right to be Forgotten"** – the ability to erase the data. In this sense, as one of the leading design guidelines, we state that the private data should not be stored on the Blockchain due to its immutability, which invalidates the data's erasure. As a solution, the privacy-by-design applications for HRI should have off-chain storage where the data can be erased [26]. Still, an open question remains whether the corresponding hash of private data, which remains on the BC forever, is considered private data, subject to particular environments and implementations, and related to the hashes' degree of anonymity (the sufficiently anonymous data is no longer private data [10]). In the experience performed in section 4, it seems highly unlikely that some malicious entity may re-identify the owner of an arbitrary RGB-D image by knowing only the hash stored on BC. Therefore, hashes of binary files, such as RGB-D images, when stored on BC, at first glance, may seem unlinkable to any outer information.

Additionally, a loose term of **erasure** leaves discussion and flexibility as to when some data shall be considered deleted. A possible solution could be to destroy the private keys in the BC environment that enable access to that data. Hence, by being inaccessible or unreadable, in some interpretations, that private data may be considered erased.

5.4. EOS Blockchain Evaluation

Robotic events happen at a very high frequency with potentially thousands of events per minute. The experience with BC applications pointed out to use EOS Blockchain [24] because of its promising results regarding the transaction throughput and potential of scalability. Further, the results of our performance evaluation indicate that there is a place to improve performance.

Moreover, the current state of the art shows other implementations that attempt to integrate robotic events with the BC.Iin the literature, the use of *Tezos* [18, 20] or *Aitheon* [27] BC was found. Besides, a new research area is emerging – the Blockchain interoperability [28] – the ability to connect two or more BC cooperating. Given the benefits of BC, we foresee the integration of many Blockchains with robots in the future. Therefore, the BlockRobot could enhance the privacy in HRI and cooperate with some other BC that offers some other solutions.

6. Conclusions and Future work

SR bears potential in personal assistant robots, healthcare, manufacturing industry, education assistants, defense agents, public space hosting, and many more. In particular, people tend to engage with robots that look more like humans on an emotional level. Individuals have limited knowledge about the robots' inference abilities and what information they may happen to know during their HRI. We need to provide them with adequate privacy-by-design features to comply with GDPR requirements, both legally and ethically. BlockRobot aims at being a step in the right direction while stimulating the discussion on privacy-by-design in SR.

As future work, performing demonstrations with real stakeholders to evaluate the perceived **fairness** and **transparency** by individuals during HRI is a natural step. Moreover, the images are being presented in the RAW format, and using a stream processing could increase the user experience and the quality of visualization of the pictures. Another task would deem to remove the trusted API for data transformations and delegate the responsibility to the robots, wherein the business logic is written by smart contracts and robots perform transactions on their own. Additionally, performing the evaluations regarding the hashes' anonymity stored on-chain, as these have special attention under GDPR scope [10, p. 31]. The anonymity improvements can be analyzed if using salted

of peppered hash [10, p. 31], wherein the risks can be measured by pattern analyzes or collision resistance attack. Moreover, other sensors may be more invasive to users' privacy, e.g., audio and emotional state. Future studies on the inclusion of these privacy-sensitive data shall be considered. We believe that the matter of privacy in HRI is a complex issue and needs much work to be done and that our research will serve as a ground point to further investigation.

Acknowledgements

The first and second authors state that this work was supported by the European Commission program H2020 under the grant agreement 822404 (project QualiChain) and by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UIDB/50021/2020 (INESC-ID). The third author state that this work was supported by LARSyS – FCT Plurianual funding 2020-2023.

References

[1] E. Cucco, M. Fisher, L. Dennis, C. Dixon, M. Webster, B. Broecker, R. Williams, J. Collenette, K. Atkinson, and K. Tuyls, "Towards robots for social engagement," 2017. [Department of Computer Science, University of Liverpool, UK].

[2] C. Lutz and A. Tamò, "Privacy and healthcare robots–an ant analysis," in *We Robot 2016: the Fifth Annual Conference on Legal and Policy Issues relating to Robotics*, 2016.

[3] M. Dziergwa, P. Kaczmarek, and J. Kedzierski, "Rgb-d sensors in social robotics," *Journal of Automation Mobile Robotics and Intelligent Systems*, vol. 9, 2015.

[4] F. Hegel, M. Lohse, A. Swadzba, S. Wachsmuth, K. Rohlfing, and B. Wrede, "Classes of applications for social robots: A user study," in *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 938–943, IEEE, 2007.

[5] F. Alonso-Martín and M. A. Salichs, "Integration of a voice recognition system in a social robot," *Cybernetics and Systems: An International Journal*, vol. 42, no. 4, pp. 215–245, 2011.

[6] J. Sequeira and D. Gameiro, "A probabilistic approach to RFID-based localization for human-robot interaction in social robotics," *Electronics*, vol. 6, no. 2, p. 32, 2017.

[7] J. Vitale, M. Tonkin, S. Herse, S. Ojha, J. Clark, M.-A. Williams, X. Wang, and W. Judge, "Be more transparent and users will like you:a robot privacy and user experience design experiment," 2018.

[8] A. Richert, S. Müller, S. Schröder, and S. Jeschke, "Anthropomorphism in social robotics: empirical results on human–robot interaction in hybrid production workplaces," *AI & SOCIETY*, vol. 33, no. 3, pp. 413–424, 2018.

[9] M. Finck, "Blockchains and data protection in the european union," *Eur. Data Prot. L. Rev.*, vol. 4, p. 17, 2018.

[10] M. Finck, *Blockchain and the General Data Protection Regulation: Can Distributed Ledgers be Squared with European Data Protection Law?: Study*. European Parliament, 2019.

[11] R. Jay, W. Malcolm, E. Parry, L. Townsend, and A. Bapat, *Guide to the General Data Protection Regulation (GDPR)*. Sweet & Maxwell, 2017.

[12] W. Van Der Aalst, "Data science in action," in *Process mining*, pp. 3–23, Springer, 2016.

[13] T. Laurence, *Blockchain for dummies*. John Wiley & Sons, 2019.

[14] G. Zyskind, O. Nathan, *et al.*, "Decentralizing privacy: Using blockchain to protect personal data," in *2015 IEEE Security and Privacy Workshops*, pp. 180–184, IEEE, 2015.

[15] N. B. Truong, K. Sun, G. M. Lee, and Y. Guo, "Gdpr-compliant personal data management: A blockchain-based solution," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1746–1761, 2019.

[16] B. Faber, G. C. Michelet, N. Weidmann, R. R. Mukkamala, and R. Vatrapu, "Bpdims: a blockchain-based personal data and identity management system," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[17] J. M. O'Kane, "A gentle introduction to ros," 2014.

[18] E. C. Ferrer, O. Rudovic, T. Hardjono, and A. Pentland, "Robochain: A secure data-sharing framework for human-robot interaction," *arXiv preprint arXiv:1802.04480*, 2018.

[19] B. Degardin, "Blockchain for robotic event recognition," *Covilhã: Universidade da Beira Interior*, 2018.

[20] M. Fernandes and L. A. Alexandre, "Robotchain: Using tezos technology for robot event management," *Ledger*, 2019.

[21] A. Brown, "Rise of the machines? amazon's army of more than 100,000 warehouse robots still can't replace humans because they lack 'common sense'," *Daily Mail (accessed 9 March 2019)*, 2018.

[22] J. Benet, "Ipfs-content addressed," *Versioned, P2P File System*, vol. 2, 2014.

[23] J. Izaguirre, J. Furgeson, and Q. Ma, "Proceedings of cse 331, data structures, department of computer science and engineering, university of notre dame, notre dame, indiana," pp. 63–67, 2000.

[24] I. Grigg, "Eos-an introduction," *White paper. https://whitepaperdatabase.com/eos-whitepaper*, 2017.

[25] J. Sequeira, P. Lima, A. Saffiotti, V. Gonzalez-Pacheco, and M. A. Salichs, "Monarch: Multi-robot cognitive systems operating in hospitals," in *ICRA 2013 workshop on many robot systems*, 2013.

[26] I. S. Michael Ault, "Why new off-chain storage is required for blockchains," *NIST National Institute of Standards and Technology*, 2019.

[27] Aitheon, "Aitheon whitepaper," 2018. [Online].

[28] R. Belchior, A. Vasconcelos, S. Guerreiro, and M. Correia, "A survey on blockchain interoperability: Past, present, and future trends," *arXiv preprint arXiv:2005.14282*, 2020.