Proceedings of the 56th Hawaii International Conference on System Sciences | 2023

# Building Digital Trust to Protect Whistleblowers - A blockchain-based Reporting Channel

Dr. Petra Maria Asprion
University of Applied Sciences
Northwestern Switzerland (FHNW)
petra.asprion@fhnw.ch

Hermann Grieder
University of Applied Sciences
Northwestern Switzerland (FHNW)
hermann.grieder@fhnw.ch

Frank Grimberg
University of Applied Sciences
Northwestern Switzerland (FHNW)
frank.grimmberg@fhnw.ch

## Abstract

*Organizations today need internal reporting channels to report illegal/unethical misconduct. For this purpose, organizations set up one or more - often digital - internal reporting channels. Persons/Employees who want to report misconduct, so-called whistleblowers, expose themselves to reprisals and therefore need trustworthy reporting channels which ensure ´Digital Trust´. Blockchain, a technology that overcomes the need for trust due to its properties of immutability and integrity of data, could be promising as underlying technology for a digital reporting channel which is recognized as trustworthy. In our research, we explored multiple perspectives relevant to a trustworthy digital reporting system. Applying design science research, we evaluated the current state of the art of (digital) reporting channels and developed a prototypical blockchain-based reporting solution called "Integrity@Inside". The prototype is being iteratively demonstrated and pre-evaluated.*

**Keywords:** Blockchain, Compliance, Digital Trust, EU Directive 2019/1937, Reporting Channel, Whistleblowing

## 1. Introduction

More than 30 years ago, Near and Miceli (1983) defined whistleblowing as "*the disclosure by (former) organization members of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action*". This description can still be used today and has been a kind of guiding principle for our research.

The topic whistleblowing has been discussed from many perspectives for years (Olesen, 2018). Most prominently, the Edward Snowden case (e.g., Fidler & Ganguly, 2015, Salvo & Negro, 2016) was widely followed by the public. From an academic perspective, whistleblowing has been well researched, but the view of organizations has been rather neglected so far (Medojevic et al., 2020).

Often the first to know about irregularities are employees or people which are in contact with an organization. (European Parliament and Council, 2019, p. 17). People or employees that report misconduct or fraud harmful to public interest, act as 'whistleblowers' and in doing so, play a key role in exposing and preventing breaches, and in safeguarding the welfare of society. The consequences of misconduct or fraud can create distortions of competition, increase costs, undermine the interests of investors or shareholders, lower attractiveness for investment and create an uneven playing field for all businesses affecting the proper functioning of internal markets (European Parliament and Council, 2019).

Conversely, for organizations this means that if they can reduce these negative consequences, they have a considerable advantage and are therefore also interested in the information provided by whistleblowers. Thus, not only the public has an interest in uncovering illegal or unethical behaviour, but also organizations of all kinds, as the direct costs (legal costs and fines) or indirect costs (damage to reputation, loss of customers) are considerable. (OECD, 2015) Therefore, whistleblowing is gaining importance from a corporate perspective (e.g., Olesen, 2018, Skupień, 2021, Dewan, 2022), not least due to the European Directive 2019/1937 which treats on the protection of persons who report breaches of European Union (EU) law. This affects a large number of organizations (European Parliament and Council, 2019).

Not least to meet legal obligations, large organizations in particular have been establishing digital reporting solutions for some years now, which are intended to enable whistleblowers to submit information in a trustworthy manner and to communicate anonymously with the respective recipients.

Digital reporting solutions for organizations exist on the market in many forms, but not well researched for example, on effectiveness and efficiency in relation to various issues, such as how exactly trust is ensured. To explicitly represent trustworthiness, a potential solution could be the application of a blockchain based digital reporting system. Why blockchain? Because it is a relatively new technology that enables the creation of transparent, immutable transaction records, which could increase trust in the technology used, at least for informed users who are somewhat familiar with blockchain and its security properties (Lee et al., 2022).

## 2. Multiple Perspectives

The discussion about (internal) reporting and reporting channels, as well as trustworthy digital solutions (from a whistleblower perspective) in organizations can be conducted by alternating between viewpoints. In this contribution, we discuss the topic from multiple perspectives: (1) legal obligations – as a fundamental driver for organizations to act, (2) risks for whistleblowers - and their obvious need for a trustworthy reporting channel, (3) ´Digital Trust´ as a relatively new and promising flanking concept, (4) existing digital reporting solutions and their trustworthiness promise, and (5) ´Blockchain´, used as a possible approach for enabling a trustworthy – Digital Trust promising - digital reporting channel.

URI: https://hdl.handle.net/10125/103160
978-0-9981331-6-4

### 1.1 Perspective 1: Legal Obligations

The EU addressed whistleblowing and passed a corresponding law with the EU Directive 2019/1937 (European Parliament and Council, 2019), which has been in force since 17 December 2021 for all EU nations.

The directive is mandatory for organizations with more than 50 employees, whereas organizations with fewer employees are simply encouraged to implement reporting channels for whistleblowing (European Parliament and Council, 2019). A key point of the directive stipulates legal entities in the private sector to establish internal reporting channels guaranteeing a certain form of Digital Trust (which needs to be defined individually) for whistleblowers, e.g., integrity of the supported data and a diligent follow-up. Typical reporting channels include e-mail, personal contact, phone, letter/fax, hotline, web-based platforms, mobile app, and social media (Hauser et al., 2021).

### 1.2 Perspective 2: Risks for whistleblowers

There are numerous reports that employees who report observed or suspected misconduct are at great risk of retaliation, both personally and professionally (e.g., Compliance Week & SAI360, 2021; Paul & Townsend, 1996; Qusqas & Kleiner, 2001). Most retaliation is experienced by employees that are relatively vulnerable, posing the lowest risk to the company, especially when (top) management support is missing (Near & Jensen, 1983). Fears of retaliation play a key role for the individual observing misconduct in whether to speak out. For this reason, a trustworthy reporting channel is essential for whistleblowers.

### 1.3 Perspective 3: Digital Trust as flanking concept

There have been numerous attempts to define the relatively new concept of Digital Trust. One as sufficient identified definition is published by Ritter (2019); he paraphrases Digital Trust as *"the perceived confidence individuals have in the ability of people, technology, and processes to build a secure digital environment. This means organizations have demonstrated to their stakeholders that they can provide safety, privacy, security, reliability, and data ethics with, for example, their online programs or devices. Once an individual uses a company's product, they confirm their digital trust in the business"*. This description puts the human´s recognition of trustworthiness with a digital system in the center.

The World Economic Forum's Centre for Cybersecurity (WEF CfCs) (World Economic Forum, n.d.) seeks to establish a global consensus and to define what measures are useful to improve and establish the trustworthiness of digital technologies. To be considered trustworthy, the WEF CfCs states that any technology must be secure, fulfilling the CIA triad´s requirements (confidentiality, integrity & availability of data and system) and that it must be responsibly used. They identified a deficit in Digital Trust due to the lack of assurance regarding secure and responsibly used systems. Consequently, the WEF CfCs has posed a demand for an evidence-based assessment to prove digital trusted systems.

Other authors, e.g., Lee et al.(2022), Mubarak & Petraite (2020) or Shin (2019) see Digital Trust as a technical solution to the underlying problem of the role of humans in trust formation, and state that digital technologies can remove the need for trust in people through the use of automatically enforced rules and processes.

The discussions around Digital Trust bring us to blockchain and what this technology offers by default in terms of Digital Trust or security mechanisms. Blockchain has been argued to be a possible driver of Digital Trust in academic discussions, due to its unique characteristics including privacy and security, immutability and redundancy of data, and decentralization, e.g., Lee et al., 2022, Mubarak & Petraite, 2020, or Shin, 2019.

In the context of digital reporting systems, these mechanisms available in the blockchain could play a key role in gaining the trustworthiness of potential users or whistleblowers so that they use the digital reporting system precisely because it is blockchain-based.

### 1.4 Perspective 4: Existing digital reporting solutions

With focus on Digital Trust, we analyzed several existing digital solutions of reporting channels based on predefined criteria. Additionally, to the criterion of the underlying architecture (traditional centralised database vs. blockchain-based decentralised data storage), we assessed solutions with focus on the available fundamental security mechanisms.

To select and systematize the criteria, we used as benchmark digital applications with blockchain as underlying architecture developed for supporting eHealth – the use of data to support health and healthcare. Why applications from the health/healthcare sector? In eHealth, the importance of trustworthy digital applications has long been recognized and one identified promising solution is blockchain-based applications. The collected security mechanisms from eHealth applications and related publications are shown in Table 2. In addition, we investigated functions for anonymous feedback, archiving, storage, and the availability of a sufficient case management.

The analyzed solutions were ´Whistleblower-Software´, ´ithikios´, ´Falcony´, ´Whistle Willow´, ´EQS Integrity Line´, ´Whispli´, and ´AKARION Compliance Cloud´.

The results of our benchmark analysis can be summarized as follows: all analyzed solutions used a centralized classical database, none of the solutions were technologically based on blockchain. This leads to the conclusion that the providers do not consider their advantages promising enough.

Regarding the security mechanisms: All the analyzed solutions provide inherent access controls, permissions, and anonymous feedback functionality. However, all of them had only vaguely described their security mechanisms and how they (technically) implemented them was not disclosed. Furthermore, the evaluated solutions can be either operated and monitored centrally from an organizations' internal body or by external third parties.

Vendors of commercial reporting solutions sometime provide standards and certifications they achieved (e.g., ISO 22301, ISO 27001, ISO 27018, ISO 90001, ISAE 3000) to increase the credibility and to promote a special form of Digital Trust. This can help organizations as a trust-building measure to create and improve. This can be achieved by implementing and explicitly outlining and explaining measures which stand for Digital Trust existence. Such explicitly outlined measures could be dedicated functions to provide confidentiality, reported case/data integrity, the continuous availability of the solution itself (and included case management) as well as other security relevant mechanisms (see Table 2 (left column)).

According to our discussions in the demonstration phase (chapter 6), people reporting misconduct will consider using available systems instead of directly reporting the misconduct to external parties, such as the media or the government. Therefore, we claim that explicitly proven and explained (to the potential reporting persons) mechanism for Digital Trust is an important pre-condition to motivate people to report a certain misconduct or failure within an organization.

### 1.5 Perspective 5: A blockchain-based reporting channel

As introduced earlier, a solution to prove and outline Digital Trust could be the use of blockchain technology, e.g., through a web-based application for reporting misconduct. Blockchain is a peer-to-peer network built on top of the internet and can be defined as a time-stamped series of immutable records of data with various Digital Trust or security mechanisms like transparency or irreversible records that can be shared across participants.

The history of blockchain were described by Haber & Stornetta (1991). Blockchain was first mentioned in October 2008 as part of a proposal for Bitcoin (Nakamoto, n.d.), as a virtual currency system that dispensed with a central authority for issuing ´money´, transferring property and confirming transactions (Wüst & Gervais, 2018). Whereas Bitcoin itself is controversially discussed, the underlying as secure identified technology works nearly perfect and inspires researchers and practitioners to develop numerous applications (Crosby, 2016).

Blockchain provides a security mechanism to guarantee among others confidentiality, integrity, and availability and uses encryption/decryption mechanisms based on the concept of cryptographic hash algorithms such SHA256, SHA512, and Merkle tree (Obaid, 2019). In addition, there are different blockchain technologies, some of them provide data privacy related characteristics explicitly (Moriggl et al., 2019). These characteristics encouraged us to design a blockchain-based reporting solution to be able to show a prototypical application, a showcase, in order to move from the abstract discussion to a real-life case discussion.

Blockchain-based solutions can be divided into public and private (Ellervee et al., 2017). While public blockchains are decentralized peer-to-peer networks, a private blockchain is a special type of blockchain, controlled by a consortium of responsible, dedicated, trusted people. In a private blockchain, operators of the network decide who can join the network, read, and write to the blockchain and keep a record of the distributed ledger (Wüst & Gervais, 2018). Distributed ledgers can be described as a record of consensus with cryptographic maintained audit trails that are validated by nodes (Ellervee et al., 2017). Therefore, blockchain is a way to implement a distributed ledger, but not all distributed ledgers necessarily employ blockchains.

Using a private blockchain can be important for organizations that want to benefit from the fundamental security mechanisms or other characteristics of blockchains, such as ´verifiability´, visibility´ or ´transparency´ as described by Wüst & Gervais (2018), but do not want to share the content of the blockchain with the outside world.

Overall, a blockchain-based reporting solution could provide mechanisms to support Digital Trust by promoting confidentiality, integrity, and availability, as well as irrefutable records. For example, a person or organization cannot deny or contest their role in authorizing the creation or change of a record (Burns et al., 2020).

The blockchain characteristics also promote the visibility of transactions and availability of information to support the auditability of information transacted on the blockchain (Burns et al., 2020). For example, once a person reporting misconduct files a report, the data entered cannot be deleted or changed due to the inherent immutability of data, and the information can be reviewed by those having access to the blockchain, including the reporting person.

To take advantage of the security mechanisms of blockchain-based solutions, we designed and developed a prototypical solution named Integrity@Inside (I@I) as a novel, web-based reporting platform backed by a blockchain that aims to make organizations compliant with EU law, as well as ensuring Digital Trust for the reporting person based on identified requirements (chapter 4). For the first prototype, we have implemented Digital Trust mechanisms that are shown in Table 2 (last column on the right titled with "I@I") (those with "R" have already been implemented in the prototype, those with "F" are planned for a future version. In addition, we decided to go - for the first prototype - with a private blockchain, first to enhance the public blockchain perspective which has been investigated by Habbabeh (2020). Second, because we have found from our expert interviews that many stakeholders in organizations do not feel comfortable having internal misconducts stored in a public ledger and therefore a private blockchain solves this issue.

The remainder of this paper is structured as follows. The next chapter presents the multiple perspectives from which whistleblowing can be examined; in this chapter, we did not present all conceivable perspectives, but limited ourselves to those that seemed to make sense in answering our research question. Chapter 3 outlines the research design and chapter 4 the requirements collected. In chapter 5, we present I@I – our prototypical blockchain-based reporting solution, followed by chapter 6, the demonstration and evaluation part described. Chapter 7 closes with a conclusion and outlook.

## 3. Research Design

The starting point for our research was a compliance-related research question: How can whistleblowers or persons who want to report misconduct be (better) protected? And how do they know that the reporting solution they want/need to use meets Digital Trust?

As the main outcome of our research, we wanted to develop an artifact - a prototype for a blockchain-based reporting solution. Therefore, we chose ´Design Science Research´ (DSR) from Hevner & Chatterjee (2010) as leading method with a "*design & development centered approach*" as proposed by Peffers et al. (2020). The course of our research included six steps of the DSR process (Peffers et al., 2020):

1. **Problem Identification and motivation.** This included the relevance of our research and the fact that IS research has not yet been focused on the application of blockchain in the context of trustworthy reporting systems.
2. **Objectives of a solution.** The targeted solution should be consistent in functionality with existing solutions and examine the effects on Digital Trust and security mechanisms by use of a blockchain as the underlying technology.

3. **Design and development.** Based on raised requirements from multiple sources like the EU Directive 2019/1937, existing reporting solutions (as benchmark), eHealth blockchain-based solutions (also as benchmark), as well as several qualitative interviews with experts from different fields. For the development of the prototype, we selected the open source blockchain implementation Hyperledger Fabric by the Linux Foundation (2022) as the underlying blockchain technology .
4. **Demonstration.** Our prototype should be presented to different target groups (e.g., subject matter experts (technical and compliance related), whistleblowers, journalists) continually during its development, including to students who will be involved in the development and testing. Furthermore, intermediate solutions should be presented to the advisory board that is specially established for this purpose. In addition, a final presentation event is foreseen to demonstrate the prototype to a group of subject matter experts with the intend to encourage participants to focus on the value of the idea and collect feedback.
5. **Evaluation.** Intention is to examine how effective the developed solution is and what potential for improvement there is. For this process, different target groups (e.g., subject matter experts (technical and compliance related), whistleblowers, journalists) should be involved, to gather expectations from different stakeholder groups.
6. **Communication.** Different channels should be used: One is the research community, through peer-reviewed papers and presentations at conferences. We also want intensive exchanges with practitioners, subject matter experts, but also the users of reporting platforms. The necessary groups and persons will be acquired during the project.

# 4. Requirements

To conceptualize our prototype, we collected and evaluated technical and non-technical requirements, with a special focus on Digital Trust and security mechanisms from complementary sources. To follow a structured approach, we categorized the requirements by four areas: Guiding source and first category was the legal perspective – the EU Directive 2019/1937. For category 2 we supplemented requirements with recommendations from literature and from functional descriptions of existing reporting systems. As third source (category 3) we leaned on requirements from the eHealth because data privacy and not at least cybersecurity controls are essential and need to be strongly considered to fulfill data protection requirements (Moriggl et al., 2019). As the fourth source (category 4), we collected expectations from industry subject matter experts on different levels. As a result, we identified the following in total 21 requirements for a trustworthy digital reporting solution applicable within organizations.

## 4.1 Category 1: The EU Directive 2019/1937

From the EU Directive 2019/1937 we derived requirements based on relevant paragraphs (para):

1. **Channel.** The EU Directive requires reporting channels to ensure confidentiality and access controls according to para. 9 (1a) "*[…] channels are designed, established, and operated in a secure manner that ensures that the confidentiality of the identity of the reporting person and any third party mentioned in the report is protected, and prevents access by non-authorized staff members*".
2. **Issue Receipt.** The EU Directive requires that a receipt needs to be provided according to para. 9 (1b) "*Acknowledgement of receipt of the report to the reporting person within seven days of that receipt*".
3. **Reporting Mode:** The EU Directive requires that the report should be allowed to be written according to para. 9 (2) "*The channels provided for input […] shall enable reporting in writing or orally, or both […]*".
4. **Record Keeping.** The EU Directive requires that reports need to be stored for a certain time according to para. 18 (1) "*[…] shall ensure that legal entities in the private and public sector […] keep records of every report received […]*"

## 4.2 Category 2: Existing Reporting Systems

From the literature (Medojevic et al., 2020; Tur, 2018), we derived criteria focusing on security mechanisms as well as criteria for usability and acceptance:

5. **Anonymity.** The reporting system should credibly assure that a reporting persons' anonymity is guaranteed. This will increase trust resp. Digital Trust in the reporting system, which is also important for the credibility of the compliance department of an organization. In case of loss of trust, regaining the trust of employees, customers and suppliers may turn out to be costly in the long run. Moreover, there is no guarantee that Digital Trust will be successfully restored. However, the system should allow the reporting persons, if they decide to do so, to leave their name.
6. **Accessibility.** The reporting system should guarantee accessibility in terms of time and location. The ´window of opportunity´ for reporting misconduct is often very short. If a person has decided to relay information, poor accessibility or limited availability of the system creates barriers and the whistleblower may decide not to report after all.
7. **Secure Dialog Capability.** The reporting system should provide an option for a technically secure communication channel, in which the reporting person can receive feedback, and in which the person investigating can ask questions.
8. **Multilingual user interface.** The reporting system should provide a multilingual user interface. For example, international organizations with dependencies in different countries should take care that the system is offered in all relevant languages. It should be possible to report a case in the mother tongue without any language barriers.
9. **Thematic Limitation.** There is a chance that people reporting with malintent use the system to report denunciations about their co-workers or superiors in an anonymous manner. Therefore, the system should, at the very least, present a limited pool of topics one can file a report under.

### 4.3 Category 3: eHealth Solutions as Benchmark

When it comes to secure data and the necessary security mechanisms, an analogy can be made with electronic health data. In this area, it is essential that patients' data meet high security requirements. For this purpose, we have examined blockchain-based solutions in the eHealth environment which deal with privacy-sensitive data.

We do not intend to describe the analyzed solutions in detail, but we have investigated the security and privacy requirements of eHealth data to derive requirements for our reporting solution. As mentioned earlier, eHealth stakeholders are extremely concerned about the security and privacy of their eHealth data and need to fulfill certain regulatory requirements to support data privacy. This privacy concern also applies to reporting persons, which is why we are applying the requirements of eHealth to our reporting solution. Therefore, from literature regarding eHealth systems, we derived further security and Digital Trust mechanisms also of relevance in today's cybersecurity criteria, which are compiled in Table 2: the most recommended security mechanism can be concluded with the following criteria (marked in grey in Table 2):

11. **Access Control.** These include ways for dealing with digital access rights, data availability, and (faster) access to records.
12. **Authentication**. These include various ways to prove user's identity. Commonly, users prove their identity by providing credentials, i.e., an agreed piece of information shared between the user and the system.
13. **Confidentiality**. These include ways to limit information access and disclosure to authorized users and preventing access or disclosure to unauthorized ones.
14. **Encryption/Decryption.** Encryption includes ways to convert a readable message to an unreadable form to prevent unauthorized parties from reading it. Decryption (the opposite) – includes ways to convert encrypted message back to its original (readable) format.
15. **Privacy Preservation.** An important concept for the management of sensitive data. When the data is transferred or communicated between different parties then it is compulsory to provide security so that other parties do not know what data is communicated. The method requires evaluating the data set's usefulness for the user. The sensitive information is eliminated, twisted, or modified to achieve confidentiality.

**Table 2. Blockchain-based Security Mechanism in eHealth Solutions**

| References<br>Security Mechanism | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | **I@I** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Access Control | X |  |  | X |  | X | X | X | X | X | X | **R** |
| Authentication |  |  |  | X | X |  |  |  |  |  |  | **R** |
| Availability |  |  |  |  | X |  |  |  |  |  |  | **R** |
| Confidentiality |  |  |  | X | X |  |  |  |  |  |  | **R** |
| Consistency |  |  | X |  |  |  |  |  |  |  |  | **R** |
| Data Integrity |  | X |  | X | X |  |  |  |  |  |  | **R** |
| Encryption. Decryption | X | X | X |  |  |  | X |  |  |  |  | **F** |
| Forward and Backward Secrecy |  | X |  |  |  |  |  |  |  |  |  | **R** |
| Non-Repudiation |  | X |  |  |  |  |  |  |  |  |  | **R** |
| Privacy Preservation |  |  | X |  |  | X | X | X | X | X | X | **R** |
| Resisting Replay Attacks |  | X |  |  |  |  |  |  |  |  |  | **F** |
| User Untraceability |  | X |  |  |  |  |  |  |  |  |  | **R** |
| Smart Contracts (time based) |  |  |  |  |  | X |  |  |  |  |  | **F** |

**Abbreviations**
I@I (right column) – mechanism in I@I prototype
R - realized in the current version of I@I
F - will be considered in the follow-up version of I@I
X - was mentioned in the analyzed publication

**References**
[1] Shahnaz et al., 2019
[2] Chen et al., 2020
[3] Daraghmi et al., 2019
[4] Kaur et al., 2021
[5] Moriggl et al., 2019
[6] Xia et al., 2017
[7] Azaria et al., 2016
[8] Amofa et al., 2018
[9] Dagher et al., 2018
[10] Zhang et al., 2018
[11] Shen et al., 2019

### 4.4 Category 4: Expertise from Subject Matter Experts

From qualitative interviews with experts and stakeholders of reporting solutions, experts of blockchain-based solutions with focus on security, and from discussions with our project advisory board members we collected the following criteria:

16. **Guided questionnaire vs simple form**. Some experts, especially, when managing many cases, prefer to initially have as much information as possible as to increase efficiency by limiting the workload due to back-and-forth communication with the person reporting. This can be achieved by having a structured questionnaire guiding the reporting person through the reporting process. However, for whistleblowers this could be counterproductive as it may lead to a feeling of being invalidated by the types of questions. In

addition, a guided questionnaire could be counter-productive in terms of a character limit in the case's description to protect the reporting person.

17. **Ease of Use**. Some expert stressed their dislike of the complexity of the current reporting system within their organization. Here, a less complex solution could help to lower the barrier for people to make the decision to file a report.
18. **Limited input.** The reporting system should be providing only limited space for writing, primarily to protect the reporting person. In case she/he later regrets her/his statement (in a formulation arising from an affect). However, this requirement contradicts somehow with #14 and needs to be decided on a case by case basis to find a balanced way to deal with reporting persons' potential intention to change the content later again and to provide enough space to reproduce an observation as accurately as possible.
19. **Campaigns and external hosting**. While a system on its own can resolve technical challenges and some security issues, trust in the system must be cultivated through diverse channels and campaigns and not at least through an organization´s culture. The experts we surveyed see an increase of reported cases whenever they launch awareness campaigns and trainings for their employees. One argument that increases Digital Trust for potential whistleblowers is that the solution is hosted and operated by an independent third party outside the organization.
20. **Deletion and Archiving**. From practice, we know that cases are deleted after a certain retention period mandated by applicable laws (e.g., 10 years for Swiss Law). However, cases that are deemed unsubstantiated are not deleted but anonymized and archived. Likewise important is the consideration in the design of any solution, how data can be transferred and deleted, especially also when opting for an outsourced setup (transfer & deletion at the end of a contractual period).
21. **Storage**. From practice, many organizations do not feel ´comfortable´ having their internal misconducts stored in a public ledger or other potentially publicly available solutions (e.g., a cloud-based solution was controversially discussed).

# 5. Integrity@Inside – THE SOLUTION

In this chapter, we describe I@I – our solution: first the applied methodology we used, second the conceptual, and third the architectural design.

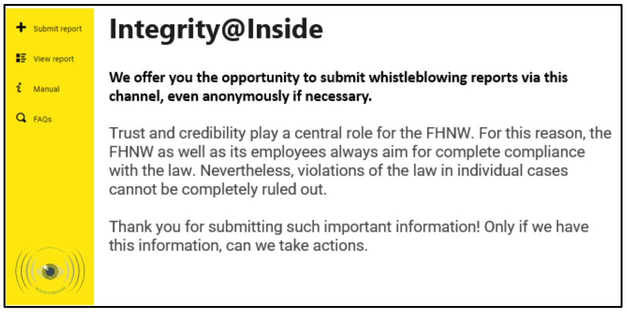
## 5.1 Applied Methodology

As a foundation, we relied on a previous study (Habbabeh et al., 2020) wherein a prototypical whistleblowing solution in form of a ´marketplace´ was conceptualized. Based on this previous study, we designed and developed I@I, as an intra-organizational blockchain-based application by applying DSR (Hevner & Chatterjee, 2010) and following a process-oriented setup (Peffers et al., 2020) as described in chapter 3. Our main objective was to design and develop a prototypical blockchain-based application of a reporting solution that would satisfy most of the requirements outlined in chapter 4.

## 5.2 Conceptual Design

I@I consists of two views, one for the reporting person (Figure 1) and one for the organization's authorized personnel or commissioned third party (Figure 2).

**5.2.1 Reporting Person View.** The web application guides the reporting person through the process of filling out a report to alleviate unnecessary barriers and allow for straightforward reporting. The website features a user manual and a section for ´Frequently Asked Questions´ (FAQ) to address any question or concern a reporting person might have ahead of reporting a case. Further, once a case has been filed, I@I allows the reporting person to monitor the report's status and chat anonymously with a so-called ´case worker´ (i.e., authorized personnel from the related organization or a delegated trusted third party).



Integrity@Inside

**We offer you the opportunity to submit whistleblowing reports via this channel, even anonymously if necessary.**

Trust and credibility play a central role for the FHNW. For this reason, the FHNW as well as its employees always aim for complete compliance with the law. Nevertheless, violations of the law in individual cases cannot be completely ruled out.

Thank you for submitting such important information! Only if we have this information, can we take actions.

**Figure 1. I@I - excerpt: reporting person landing page**

The ´end-to-end´ reporting process consists of four steps:

(1) The reporting person receives a security notice that highlights all the safeguards that have been implemented and the measures that should be taken to ensure identity protection. I@I does not allow users/reporting persons to continue without confirmation of having read the security notes.

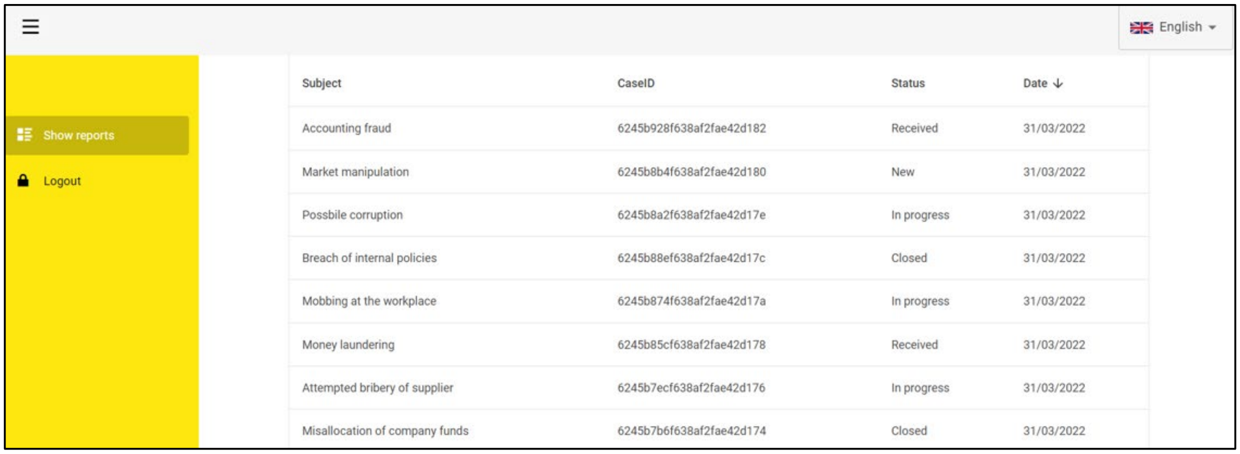
(2) The reporting person is asked to enter the details of the case through a form that is structured similarly to an email – easy to use - with a subject line and a field to describe the case in detail without any pre-structured entry fields.

(3) There is the possibility to include the reporting person's name (optional) and also optional to upload files/documents that potentially serve as evidence to the case.

(4) The system returns a case number and a random password to the reporting person so they can access the case later. The use of a system-generated password is an additional safeguard to ensure that passwords will not allow the case to be traced back to a reporting person's identity. The reporting person is advised to note down the password and case number as they are only shown this information once and it cannot be retrieved later, which is a compromise between convenience and guaranteeing the anonymity of the reporting person.

**5.2.2 Organization's authorized personnel.** Persons that report misconduct can use the system without the need for creating an account or registration or any login. Yet, from an organizational view user access management for the case workers is needed as only authorized personnel or third party workers are allowed to have access to the cases and process them. Case workers can view the list of all cases (Figure 2) and process individual cases.

Organization's authorized personnel can change the status of the case, from new to processing and closed depending on the state, access uploaded files, and leave messages for the reporting person. Because there is no way to inform a reporting person about new messages, reporting persons are advised to check the case status regularly.



| Subject | CaseID | Status | Date ↓ |
|---|---|---|---|
| Accounting fraud | 6245b928f638af2fae42d182 | Received | 31/03/2022 |
| Market manipulation | 6245b8b4f638af2fae42d180 | New | 31/03/2022 |
| Possbile corruption | 6245b8a2f638af2fae42d17e | In progress | 31/03/2022 |
| Breach of internal policies | 6245b88ef638af2fae42d17c | Closed | 31/03/2022 |
| Mobbing at the workplace | 6245b874f638af2fae42d17a | In progress | 31/03/2022 |
| Money laundering | 6245b85cf638af2fae42d178 | Received | 31/03/2022 |
| Attempted bribery of supplier | 6245b7ecf638af2fae42d176 | In progress | 31/03/2022 |
| Misallocation of company funds | 6245b7b6f638af2fae42d174 | Closed | 31/03/2022 |

**Figure 2. I@I - excerpt: organization's authorized personnel case overview**

**5.2.3 Supplementary Design Choices.** Derived from the requirements outlined in chapter 2, we took the following decisions:

(1) To protect a reporting person's identity, we advise them not to use I@I in an organizational network with internal devices (e.g., laptops, workstations) which could possibly be used to trace them, but rather use a remote location with personally obtained devices.

(2) Certain information about the reporting person could be inferred by their use of such a digital internal reporting solution. Examples are the time of day, when a person files a report or when a person repeatedly provides additional information through the chat function. To this end, from a technical viewpoint, we do not allow for additional information outside the HTTP request payload to be stored or logged at any time during the process. As stakeholders of I@I need to know when a report was filed, the date of filing is recorded and shown, however the time of filing is intentionally omitted.

(3) The prototypical solution of I@I supports two languages to address the (minimum) requirements of a multilingual system. However, offering more languages is easily achieved by adding the corresponding translations.

(4) To be compliant with the EU Directive 2019/1937 paragraph 18 (1), reports and records received are stored in a secure environment.

### 5.3 Architecture Design

The architecture design of I@I follows classical a three-tier client/server architecture (Aarsten & Brugali, 1996), consisting of a client ´Frontend Application´ (left box), a ´Server Application´ (center box) and a ´Private Blockchain´ (right box) using ´Hyperledger Fabric Network´ (Linux Foundation, 2022), an open source, permissioned blockchain framework .

**5.3.1 Frontend Application.** The web-based frontend application provides an easy-to-use interface for both the reporting person and the case worker(s). Reporting persons can create a report. Case workers can access all reports and update the report status. Further, both parties can send and receive messages through the application.

The web-application combines advantages of systems to provide functionality with advantages of web pages to be easily accessible via a browser. The frontend is implemented using the APS.NET Core Blazor Framework (Microsoft, 2022) to create interactive dynamic client-side web pages. HTTPS requests to the server application are standardized using an openAPI specification (*OpenAPI Specification*, 2022). For packaging, executing, deploying the application into different environments, Docker (2021) was selected.

**5.2.2 Server Application.** The server application bridges the communication between the client-side web application and the private blockchain. HTTPS requests, as defined in an openAPI specification (*OpenAPI Specification*, 2022), are sent from the client application to the corresponding service on the server. The service connects to and invokes chain code of the private blockchain through the gateway provided by the Hyperledger Fabric Software Development Kit (SDK). The server application was built using the Node.js runtime, the Express web application framework, and the Hyperledger Fabric SDK. Again, Docker was used to package, execute, and manage the application as a single, immutable object.

**5.3.3 Private Blockchain.** Currently there are various blockchain platforms available provided as open source (Analytics Insight, 2022). We decided to use the open-source blockchain Hyperledger Fabric (Linux Foundation, 2022). Hyperledger Fabric is a modular permissioned decentralized ledger technology (DLT) platform for developing applications aimed for use within private enterprises (Linux Foundation, 2022). The use of a blockchain has some advantages over traditional databases. Due to the decentralization of the system, an organization can increase the Digital Trust in the system by setting up nodes (Organization 1 and 2 in Figure 3) for different stakeholders, all with their respective copy of the ledger.

Especially in a private permissioned blockchain, one can envision a setup where the organization, an employee representative body, and possibly an external independent party each run a separate node. This would allow for independent inspection of the ledger's current state and transactions. Further, the risk of malicious manipulation of the ledger, such as deleting or altering reports, can be alleviated, as each node is audited by different parties. Therefore, we designed a network structure that consists of two organizations, emblematic of e.g., the company and an employee representative body.

Each organization has one peer node that holds a copy of the ledger of the blockchain. The peer nodes communicate through a dedicated channel to update the ledger after the respective *´chain code´* (Hyperledger Fabric's version of a smart contract) has been invoked. Due to the modular design of Hyperledger Fabric, the network can be adapted to accommodate for more participants if needed.

of whistleblowing came up. Here, all participants agreed that any kind of financial incentive would be unadvisable as this might motivate people to file unjustified and illegitimate claims.

Furthermore, we collected many hints that were relevant for the evaluation phase, e.g., how to convince a company to trust the system and test it, or how to convince future users of the security and trustworthiness of the system.

Furthermore, we presented the proposed solution to an expert panel for evaluation in Lausanne, Switzerland, consisting of lawyers, researchers, and practitioners. We learned that there is some opposition to the thought of immutable data, as this could be a problem for organizations who do not want total transparency. This could be an indication that exactly these characteristics of blockchain, such as immutability and redundancy could be a driver for Digital Trust.

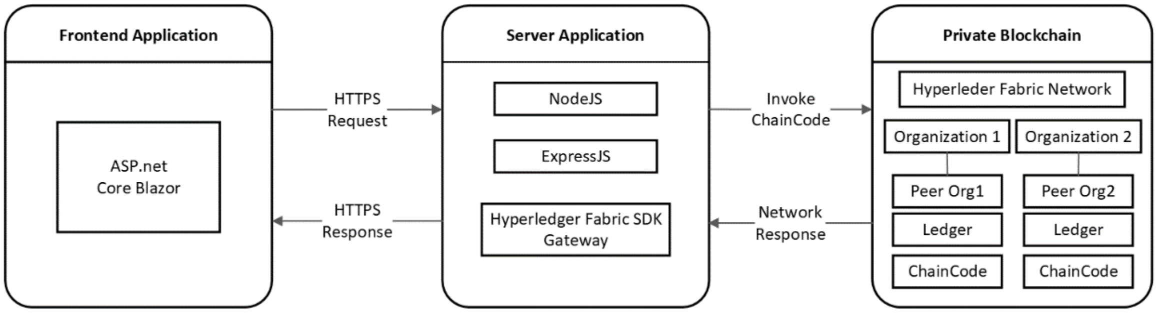


**Figure 3 I@I Architecture Design**

## 6. Demonstration and Evaluation

According to Peffers et al. (2020) DSR process flow, the design phase is followed by the demonstration - "*use the artefact to solve problem*", and the evaluation phase - "*iterate back to design*". Our prototype was demonstrated continually during its development: First with students who were partly involved in the development and who tested the application in the role of a whistleblower and eagerly gave feedback, which in turn was incorporated into the I@I solution. Furthermore, we demonstrated intermediate solutions to various stakeholders, such as our funding organisation and to the advisory board that was specially established for this purpose.

The prototype sparked a lot of discussions around various topics, including technical, organizational, and legal aspects. The head of compliance of a multinational organization found the solution "*innovative and that it could potentially lead to more people speaking up*" in her organization but remarked that the benefits of the blockchain as the underlying technology might need preceding communication and education efforts, to generate enough trust for the reporting person.

Rejection of the idea of a blockchain-based solution came from a corporate lawyer who was concerned that through the immutability of data, organizations might face long lasting consequences if they cannot entirely control how the filed reports are stored. Further, the concept of having to share filed reports in distributed ledgers in a consortium-like environment caused unease. During the discussions, the topic of financial incentives through cryptocurrency in the context

Therefore, it is important to have a neutral, decentralized solution where opposing parties with decision power cannot force through their wishes to possibly/potentially redact or hide whistleblowing cases.

The artefact demonstrated largely met expectations and the blockchain-based technical details were also intensively challenged and discussed in terms of additional benefits to increase Digital Trust.

An important finding was that knowing that I@I is a blockchain-based solution increased the willingness of some people to whom we had demonstrated I@I to provide information - as whistleblowers. Especially for the technically informed, the blockchain was an additional motivation to trust the system and therefore to use it. There was also the opposite reaction, namely that some people were suspicious of a blockchain-based solution and therefore did not trust it.

We completed the demonstration phase in July 2022. In the context of this phase, two publications - as dissemination engagement - have been submitted but not published yet.

The evaluation following the demonstration phase is currently being planned: We are preparing to test the prototype at a university - as a web-based reporting system for students. However, this phase has not yet started, as it requires extensive stakeholder management to convince a university that a reporting system is beneficial and needs to be supported. A lot of educational work still needs to be done here. The latter is currently being planned as a separate research project namely to answer the question of what an education and communication concept could look like.

## 7. Conclusion & Outlook

In this contribution, we described the overall relevance of digital reporting solutions and tried to find answers for our two research questions: First, how whistleblowers can be (better) protected? Second, how do they know that the reporting solution they want/need to use is trustworthy enough?

In chapter 2, based on multiple perspectives, we outlined initially the need for reporting solutions, not least because of legal obligations. We described some risks which need to be mitigated when we talk about a digital (web-based) reporting solution. We discussed the challenges of the concept of Digital Trust and provided the idea to increase Digital Trust by providing a blockchain-based digital reporting system. Further, we analyzed some existing digital reporting systems – to disclose their security mechanisms and their overall functionalities – not at least as benchmark for our I@I prototype which should use blockchain as the underlying technology.

In chapter 3, we described our research design which we aligned with the DSR process flow from Peffers et al. (2020).

In chapter 4, we raised and categorized the requirements which are necessary to address the various challenges of an internal reporting system.

In chapter 5, we presented the conceptual and architectural designs of I@I, which can be adapted easily because only open-source products were used.

Chapter 6 focused on the demonstration and evaluation of our research. The feedback and discussions confirmed the finding from Lee & Fargher (2018), namely that the presences of internal whistleblowing systems reduce the likelihood of whistleblowing to outside third parties, which is beneficial to the organization, as they can correct the issue before they become public. The demonstration phase further revealed that the rollout needs to be carefully prepared and that stakeholders needs to be involved in all stages.

In our future work, first and foremost, the piloting of the current prototype is planned. This will be focused on the useability (e.g., ease of use as requirement) of the technical solution. In addition, we intend to analyze and investigate the topic of Digital Trust in more depth, in line with the initiative of the WEF CfCs (World Economic Forum, n.d.), that claims that an evidence-based assessment of what drives Digital Trust (e.g., between citizens and tech, between governments and other organizations, among private sector actors) is important also in relation to the ability to measure improvements (or erosion) against generally accepted Digital Trust metrics. Furthermore, we know from interviews and discussions, that an adequate ´marketing´ of a tool like I@I is important for its success. Hence, we plan to design and assess a campaign to advertise our solution.

### Funding & Acknowledgement

We would like to express our gratitude to the KBA-NotaSys Integrity Fund (https://banknote-solutions.koenig-bauer.com/en/integrity-fund) for funding our research. Further, we would like to thank the subject matter experts and the project’s advisory board members for providing valuable insights, recommendations, and guidance as well as being available for the related discussions.

## 8. References

Aarsten, A., & Brugali, D. (1996). *Patterns for Three-Tier Client/Server Applications*.

Amofa, S., Sifah, E. B., Obour Agyekum, K. O.-B., Abla, S., Xia, Q., Gee, J. C., & Gao, J. (2018). A Blockchain-based Architecture Framework for Secure Sharing of Personal Health Data. *IEEE 20th International Conference on E-Health Networking, Applications and Services (Healthcom)*, 1–6. https://doi.org/10.1109/HealthCom.2018.8531160

Analytics Insight. (2022, July 27). *Top 10 Open-Source Blockchain Platforms to Explore in 2022*. https://www.analyticsinsight.net/top-10-open-source-blockchain-platforms-to-explore-in-2022/

Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). *MedRec: Using Blockchain for Medical Data Access and Permission Management*. 6. https://doi.org/10.1109/OBD.2016.11

Burns, J., Steele, A., & Cohen, E. E. (2020). *Blockchain and internal control*. 40.

Chen, C.-L., Deng, Y.-Y., Weng, W., Sun, H., & Zhou, M. (2020). *A Blockchain-Based Secure Inter-Hospital EMR Sharing System*. https://doi.org/10.3390/app10144958

Compliance Week, & SAI360. (2021). *What whistleblowers want compliance officers to know* (p. 31). www.complianceweek.com/e-books/e-book-what-whistleblowers-want-compliance-officers-to-know/30983.article

Crosby, M. (2016). *BlockChain Technology: Beyond Bitcoin*. *2*, 16.

Dagher, G. G., Mohler, J., Milojkovic, M., & Marella, P. B. (2018). Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustainable Cities and Society*, *39*, 283–297. https://doi.org/10.1016/j.scs.2018.02.014

Daraghmi, E.-Y., Daraghmi, Y.-A., & Yuan, S.-M. (2019). MedChain: A Design of Blockchain-Based System for Medical Records Access and Permissions Management. *IEEE Access*, *7*, 164595–164613. https://doi.org/10.1109/ACCESS.2019.2952942

Dewan, K. (2022). Tatiana Bazzichelli (ed), Whistleblowing for Change: Exposing Systems of Power and Injustice. *Journal of Conflict and Security Law*, krac025. https://doi.org/10.1093/jcsl/krac025

Docker Inc. (2021, October 6). *Docker*. https://www.docker.com/products/docker-desktop/

Ellervee, A., Matuleviˇcius, R., & Mayer, N. (2017). *A Comprehensive Reference Model for Blockchain-based Distributed Ledger Technology*. 14.

European Parliament and Council. (2019). *Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law*. 40.

Fidler, D. P., & Ganguly, S. (Eds.). (2015). *The Snowden reader*. Indiana University Press.

Habbabeh, A., Asprion, P. M., & Schneider, B. (2020). *Mitigating the Risks of Whistleblowing An Approach Using Distributed System Technologies*. 12.

Haber, S., & Stornetta, W. S. (1991). How to time-stamp a digital document. *Journal of Cryptology*, *3*(2), 99–111. https://doi.org/10.1007/BF00196791

Hevner, A., & Chatterjee, S. (2010). *Design Research in Information Systems* (Vol. 22). Springer US. https://doi.org/10.1007/978-1-4419-5653-8

Kaur, J., Rani, R., & Kalra, N. (2021). Blockchain-based framework for secured storage, sharing, and querying of electronic healthcare records. *Concurrency and Computation: Practice and Experience*, *33*(20), e6369. https://doi.org/10.1002/cpe.6369

Lee, N. M., Varshney, L. R., Michelson, H. C., Goldsmith, P., & Davis, A. (2022). Digital trust substitution technologies to support smallholder livelihoods in Sub-Saharan Africa. *Global Food Security*, *32*, 100604. https://doi.org/10.1016/j.gfs.2021.100604

Lennane, K. J. (1993). "Whistleblowing": A health issue. *BMJ*, *307*(6905), 667–670. https://doi.org/10.1136/bmj.307.6905.667

Linux Foundation. (2022). Hyperledger Fabric. *Hyperledger Foundation*. https://www.hyperledger.org/use/fabric

Medojevic, B., Milojkovic, E., & Brink, J. (2020). *Establishing Guidelines for Internal Whistleblowing Systems*. 76.

Microsoft. (2022). *Blazor*. Microsoft. https://dotnet.microsoft.com/en-us/apps/aspnet/web-apps/blazor

Moriggl, P., Asprion, P. M., & Kramer, F. (2019). *Blockchain as an Enabler for Cybersecurity Use Case: Electronic Health Records in Switzerland*. 12.

Mubarak, M. F., & Petraite, M. (2020). Industry 4.0 technologies, digital trust and technological orientation: What matters in open innovation? *Technological Forecasting and Social Change*, *161*, 120332. https://doi.org/10.1016/j.techfore.2020.120332

Nakamoto, S. (n.d.). *Bitcoin: A Peer-to-Peer Electronic Cash System*. 9.

Near, J., & Jensen, T. (1983). The Whistleblowing Process: Retaliation and Perceived Effectiveness. *Work and Occupations*, *10*(1), 3–28. https://doi.org/10.1177/0730888483010001001

Obaid. (2019, September 30). *Understanding the basics of blockchain—Nourish the roots of technology*. DataFlair. https://data-flair.training/blogs/basics-of-blockchain-technology/

OECD. (2015). *Corporate Governance and Business Integrity: A Stocktaking of Corporate Practices* (p. 120). OECD. http://www.oecd.org/corruption/corporate-governance-business-integrity-stocktaking-corporate-practices.htm

Olesen, T. (2018). The democratic drama of whistleblowing. *European Journal of Social Theory*, *21*(4), 508–525. https://doi.org/10.1177/1368431017751546

*OpenAPI Specification*. (2022). https://spec.openapis.org/oas/latest.html

Paul, R. J., & Townsend, J. B. (1996). Don't kill the messenger! Whistleblowing in America - A review with recommendations. *Employee Responsibilities and Rights Journal*, *9*(2), 149–161. https://doi.org/10.1007/BF02622256

Peffers, K., Tuunanen, T., Gengler, C., Rossi, M., Hui, W., Virtanen, V., & Bragge, J. (2020). *Design Science Research Process: A Model for Producing and Presenting Information Systems Research*.

Qusqas, F., & Kleiner, B. H. (2001). The difficulties of whistleblowers finding employment. *Management Research News*, *24*(3/4), 97–100. https://doi.org/10.1108/01409170110782702

Ritter, J. (2019). *What is digital trust?* WhatIs.Com. www.techtarget.com/whatis/definition/digital-trust

Salvo, P. D., & Negro, G. (2016). *Framing Edward Snowden: A comparative analysis of four newspapers in China, United Kingdom and United States*. https://doi.org/10.1177/1464884915595472

Shahnaz, A., Qamar, U., & Khalid, A. (2019). Using Blockchain for Electronic Health Records. *IEEE Access*, *7*, 147782–147795. https://doi.org/10.1109/ACCESS.2019.2946373

Shen, B., Guo, J., & Yang, Y. (2019). MedChain: Efficient Healthcare Data Sharing via Blockchain. *Applied Sciences*, *9*(6), 1207. https://doi.org/10.3390/app9061207

Shin, D. D. H. (2019). Blockchain: The emerging technology of digital trust. *Telematics and Informatics*, *45*, 101278. https://doi.org/10.1016/j.tele.2019.101278

Skupień, D. (2021). *Towards a Better Protection of Workplace Whistleblowers in the Visegrad Countries, France and Slovenia*. Wydawnictwo Uniwersytetu Ł&#243;dzkiego. https://www.ceeol.com/search/book-detail?id=1007429

Tur, K. (2018). Hinweisgebersysteme. In A. Kleinfeld & A. Martens (Eds.), *CSR und Compliance: Synergien nutzen durch ein integriertes Management* (pp. 277–290). Springer. https://doi.org/10.1007/978-3-662-56214-7_18

World Economic Forum. (n.d.). *Digital Trust*. World Economic Forum. Retrieved April 28, 2022, from https://www.weforum.org/projects/digital-trust/

Wüst, K., & Gervais, A. (2018). Do you Need a Blockchain? *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, 45–54. https://doi.org/10.1109/CVCBT.2018.00011

Xia, Q., Sifah, E. B., Asamoah, K. O., Gao, J., Du, X., & Guizani, M. (2017). MeDShare: Trust-Less Medical Data Sharing Among Cloud Service Providers via Blockchain. *IEEE Access*, *5*, 14757–14767. https://doi.org/10.1109/ACCESS.2017.2730843

Zhang, P., White, J., Schmidt, D. C., Lenz, G., & Rosenbloom, S. T. (2018). FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data. *Computational and Structural Biotechnology Journal*, *16*, 267–278. https://doi.org/10.1016/j.csbj.2018.07.004