Sense and Sensibility in COVID-19 medical credentials

A Value Sensitive Design perspective on the use of Self Sovereign Identity enabled access to healthcare facilities.

Authors (in alphabetical order):
Georgy Ishmaev – Delft University of Technology
Roderick Noordhoek - Rabobank
Marlies van Steenbergen – Utrecht University of Applied Sciences
Nadia Vermaes - Rabobank

Executive Summary

This white paper presents findings of the Ethics Working Group, from the conceptual phase of investigation into the ethical issues of the uNLock solution, providing identity management solutions for sharing and presentation of medical COVID-19 credentials (test results) in the context of healthcare institutions. We have provided an outline of direct and indirect stakeholders for the uNLock solution and mapped values, benefits, and harms to the respective stakeholders. The resulting conceptual framework has allowed us to lay down key norms and principles of Self Sovereign Identity (SSI) in the specific context of uNLock solution. We hope that adherence to these norms and principles could serve as a groundwork for anticipatory mitigation of moral risk and hazards stemming from the implementation of uNLock solution and similar solutions. Our findings suggest that even early stage of conceptual investigation in the framework of Value Sensitive Design (VSD), reveals numerous ethical issues. The proposed implementation of the uNLock app in the healthcare context did not proceed further than prototype stage, thus our investigation was limited to the conceptual stage, and did not involve the practical implementation of VSD method involving translation of norms and values into engineering requirements. Nevertheless, our findings suggest that the implementation of VSD method in this context is a promising approach that helps to identify moral conflicts and risks at a very early stage of technological development of SSI solutions. Furthermore, we would like to stress that in the light of our findings it became painfully obvious that hasty implementation of medical credentials system without thorough ethical assessment, risks creating more ethical issues rather than addressing existing ones.

UPDATE: on the 14th of January the Dutch Health Council published the report on “Testbewijzen voor SARS-CoV-2: ethische en juridische voorwaarden.” This report goes beyond the scope of the original uNLock solution. Therefore, the working group did not incorporate it in the white paper.

Table of Contents

This white paper is presented by the Ethics Working Group of the uNLock Consortium

1. Introduction

In April 2020, when the Netherlands was moving towards an 'intelligent lockdown' because of the COVID-19 pandemic, a group of professionals in the field of Blockchain and Self Sovereign Identity in the Netherlands gathered to work out a solution capable of alleviating at least some of the aspects of blanket lockdowns and help to 'uNLock' the Dutch society. It was decided that the professionals and organizations they represent would cooperate as a consortium to build an identity solution based on the principles of Self Sovereign Identity (SSI) (Alan, 2016), that could provide individuals with the opportunity to safely receive, store and share their COVID-19 test results. The application of the solution would be focused on the healthcare sector and its employees, in a hope that the proposed solution could help to streamline sharing of test results required for employees of healthcare institutions in a privacy preserving manner.

From the very start of the project, there was an appreciation for the fact that the implementation of such a system even in narrow application context is rife with ethical issues. The healthcare application context was also partially considered in hope that such an application embedded in existing professional norms for the healthcare employees would introduce fewer novel ethical risk. However, given the high stakes and risks, that any such system of medical credentials carries in the context of a pandemic, it was also clear that failure to address these challenges would undermine the justification for the very existence of the uNLock solution. The task force focusing on these issues was divided into an Ethics Working Group and an Ethics Committee.

2. Ethics Working Group in 'uNLock' consortium

The focus of the Ethics Working Group was the development of a conceptual framework for the anticipatory identification of values and ethical issues both for the specific technical uNLock solution and for the ecosystem in general, using methods of Value Sensitive Design. The Ethics Committee consists of ethicists with backgrounds in different fields of research, aiming to be representative of various perspectives on ethical issues pertaining to social, technological, and business aspects of the proposed solution. The main task of the external Ethics Committee is to judge any application of the solution before starting with a pilot or 'go live'. The framework developed by the Ethics Working Group also aims to assist the Ethics Committee providing one of the tools for the evaluation of the ethical desirability of the chosen designs for the uNLock system.

The Ethics Working Group supports the uNLock Consortium with the investigation aiming to provide the list of direct and indirect stakeholders of the solution, their values, benefits, and harms, and application specific conceptualizations of related ethical norms. The Value Sensitive Design (Friedman et al., 2002; van den Hoven et al. 2015) approach was chosen as a methodological basis for the research.

Up to now, the Ethics Working Group has carried out the first phase of the conceptual investigation in the VSD framework, including:

- list of the main indirect/direct stakeholders;
- respective benefits and harms;
- related stakeholder values;
- conceptualization of these values;
- application specific conceptualizations of corresponding norms.

This white paper is written by the Ethics Working Group in the hope that these findings not only identify key ethical challenges of the uNLock solution but may also provide insights to other projects aspiring to deliver SSI based solutions for medical credentials in the COVID-19 pandemic.

3. What is Self Sovereign Identity

Self Sovereign Identity (SSI) is a broad family of technological solutions for digital identity management, largely inspired by the idea that individuals should own and control their identity with minimal reliance on administrative authorities. The idea of 'self-sovereignty' in this context can be understood as the concept of individual control over identity relevant private data, the capacity to choose where such data is stored, and the ability to provide it to those who need to validate it. Build on the basis of decentralized ledger technology SSI enables issuance and sharing of verified credentials in a secure and privacy-enhancing way.

For example, if an individual Alice wants to prove her date of birth, she can share a credential cryptographically signed by a trusted issuer such as a government. Alice can let the other party verify that this credential was issued to her and that it contains cryptographically provable claim about her date of birth without asking the permission from the issuer or informing the issuer about this information exchange. This is an open-ended technological stack that can be implemented in a multitude of configurations. However, at its core, the SSI approach aims to protect digital identity owners' freedom and personal autonomy through the decentralization of key technological components. The high-level description of SSI and key standards are laid out by Christopher Alan in the paper 'The Path to Self Sovereign Identity' (2016).

3.1. Relevance of the research (medical credentials)

While SSI solutions are still largely in the phase of experimental technology, some of its promised affordances seem to be a perfect fit for the requirements of medical credentials such as privacy and autonomy of credentials' holders. One of the key value conflicts in such systems concerns conflicting requirements between the capacity to share these types of credentials, with stringent requirements for (private) data protection and respect for the privacy of credentials owners. SSI based identity solutions promise to reconcile these conflicts with solutions where individuals can be in full control of their personal information and be able to retrieve, store, and share it in the form of credentials with parties of their choice, minimizing the risk of data leakage and unauthorized third party access.

At the same time, the SSI approach is not a 'silver-bullet' solution that can resolve all ethical issues in this context. This apprehension is illustrated by the idea of an 'immunity passport' that would let individuals with assumed immunity bypass quarantine measures. This solution was and is still touted by different governments in the context of the COVID-19 pandemic, often without regards to the lack of scientific knowledge on immune response, and with a superficial appreciation of perverse socio-economic incentives that such schemes introduce. Furthermore, any emergency measure for society wide access-control based on medical data risks becoming a permanent fixture of systematic discrimination and bio-surveillance. These concerns are illustrated by the 'colour code' COVID-19 apps granting individuals access to public spaces integrated with opaque infection risk assessment algorithms and 'social credit scores' (Zhong, 2020).

Thus, any solution for digital medical certificates for COVID-19 not only has to address privacy and users' autonomy issues - something that can be accomplished with the help of the SSI approach. Such a solution also needs to pass the test of efficiency, proportionality, and ethical acceptability. The latter requires not only a valid scientific basis but context-specific ethical frameworks for the assessment of these solutions, developed with the participation of all affected stakeholders.

4. About the uNLock Consortium

uNLock is an open and non-profit Dutch consortium, driven by the conviction that the setup and development of a decentralized Self Sovereign Identity network can only be done through decentralized collaboration. The uNLock use case kick-started the consortium and served as a catalyst to the motivation of the participants to find a solution with a direct, positive impact on society. See also: unlockapp.nl/#wiewezijn

uNLock is an initiative of Dutch Blockchain Coalition (DBC), Universiteit Leiden, Rabobank, TNO, Deloitte, Ledger Leopard, CMS en stichting RINIS.

5. About the uNLock application

uNLock is an application that has been built during the Covid-19 pandemic to provide a tool for the Dutch health care facilities to determine if a person is compliant with the entry requirements of the facility. In the development phase of uNLock, the Ethics Working Group has defined a set of norms and values that the uNLock application and the uNLock consortium should uphold and adhere to.

uNLock is premised on a situation in which COVID-19 tests for the healthcare sector are widely available. Once a person has been tested, she or he receives a unique credential of that test result that can be saved in the uNLock application on a smartphone. As soon as this person wants to enter a health care institution, the desk clerk requests the person to scan a unique barcode provided by the institution (from a safe distance). The visitor can then read the access policy on her/his smartphone and receive a notification whether her/his COVID-19 test results are in compliance with the access policy of this particular health care institution. After that, the visitor can decide to

show the cryptographically signed proof of compliance with the access policy to the desk clerk, whereby the digital proof's authenticity and validity are checked.*

The workflow of uNLock app is illustrated by the following high level scheme (Fig. 1):

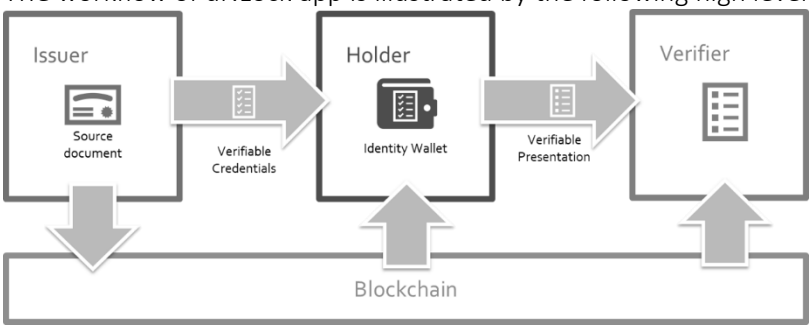

Figure 1.

The three actors in a decentralized self-sovereign identity architecture are the issuer, holder, and verifier. Issuers are entities that provide other entities with identity information. They can issue and revoke this information. Holders are often natural persons (can also be entities) that build an identity profile from their interaction with issuers and want to share that contextual identity with other parties; this other party is the verifier. Verifiers are entities who wish to provide a service to the holder but have to verify identity information about that holder first. The holder's identity information provided in a credential by the issuer is verified by the verifying entity. This identity information sharing process can be applied to any situation that requires an identity attribute.

6. Methodology of the Ethics Working Group - Value Sensitive Design

Value Sensitive Design (VSD) is "a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" (Friedman et al., 2002). VSD supports a design of technological innovations that not only takes into account instrumental aspects such as functionality, reliability, and ease of use but also the moral values of individuals and societies (Flanagan, Howe & Nissenbaum 2008). VSD defines human value as "what is important to people in their lives, with a focus on ethics and morality (Friedman & Hendry 2019, p. 4). The use of a technological artifact can both support and hinder values. Therefore, values must be considered throughout the entire design process.

VSD places much emphasis on the fact that not only the values of direct stakeholders must be considered, such as the users of technological innovation, but also the values of indirect stakeholders who may indirectly be impacted by the innovation. For example, future generations,

* The solution as described here was the initial scope of the uNLock solution on which this white paper is written. This research was confined to the specific context of uNLock solution as a prototype for the sharing of COVID-19 credentials for healthcare employees. Any further application of uNLock solution lies outside of the scope of a current report.

or individuals who cannot or will not use a service. The values of these stakeholders, as well as potential tensions between them, are investigated from a conceptual, empirical and technical perspective, and translated into design choices. At the conceptual level, the relevant stakeholders and values are identified and defined, based on existing literature and knowledge. At the empirical level the actual perception of these values by the various types of stakeholders is studied by employing methods such as interviews, focus groups or experiments, leading to further elaboration of the values into norms. At the technical level, the values and norms are translated into technical design. The three perspectives are iteratively employed.

7. Value Sensitive Design in uNLock

7.1. Conceptual investigation

In line with the Value Sensitive Design approach (Friedman and Hendry, 2019), the Ethics Working Group started with a conceptual investigation. This entails answering questions such as: What are the values of the stakeholders? Whose values should be supported in the design process? How are values supported or diminished by particular technological designs? How should we engage in trade-offs among competing values in the design, implementation, and use of information systems (e.g., autonomy vs. security, or anonymity vs. trust)? Should moral values (e.g. a right to privacy) have greater weight, or even trump, non-moral values (e.g., aesthetic preferences)? Value Sensitive Design takes up these questions under the rubric of conceptual investigations: philosophically-informed analyses of the central constructs and issues under investigation. (Friedman et al., 2002). Conceptual investigation can be seen as thoughtful consideration of how stakeholders might be socially impacted by one's technological designs (Friedman et al., 2002).

7.2. Iterative exploration on stakeholder values

VSD methods involve different types of investigation including conceptual, empirical, and technical, that are meant to inform each other rather than be engaged as separate, strictly sequential activities. While VSD can begin with any type of investigation we follow Friedman et al. (2013) arguments that a stakeholder analysis should be taken as one of the first steps. This white paper presents the findings from the first steps of the Ethics Working Group investigation, focusing on the conceptual analysis. The task of this step is two-fold. One is a robust stakeholder analysis informed by empirical research carried out by the other members of the consortium and direct collaboration with consortium participants. The second goal of the conceptual analysis was the elucidation of key stakeholder values through the method of specification (van de Poel, 2015), as an anticipatory tool for the identification and resolution of potential value conflicts. We take a broad inclusive interpretation of values as suggested by Friedman et al. (2013) referring to what persons or groups consider important in life, circumscribed by the set of specific values with ethical import to system design. We also suggest that separation of the investigations is a conceptual tool, meant to appreciate interactional aspects of design, and different investigations should not result in separate tracks within the project. Thus, the resulting conceptual framework is a tool that not only is meant to inform the design of the system, but that should be updated and refined in an iterative manner throughout the later stages of empirical and technical investigations, including broader engagement with stakeholders and a feed-back cycle of the uNLock system technological development.

7.3. Steps of investigation.

The following steps were taken by the Ethics Working Group over the past period:

Step 1: Defining and mapping the stakeholders of the uNLock solution, their interests, benefits and harms.

In order to ensure that all ethical and social issues will be addressed in the Value Sensitive Design, detailed insights and information regarding the uNLock application and the stakeholders are required in the conceptual investigation.

It is possible to distinguish between two classes of stakeholders: direct and indirect. Direct stakeholders refer to parties – individuals or organizations – who interact directly with the solution. Indirect stakeholders refer to all other parties who are affected by the use of the solution.

Based on the information that the Ethics Working Group gathered on the application and the uNLock consortium, it worked to identify the main stakeholders of the uNLock solution. The Ethics Working Group first analyzed the main goal of the uNLock application, which was defined by the consortium as: ''to provide verified proof of Covid-19 test results''. The direct stakeholders - the group including direct users of the uNLock application – were identified. At the next step the indirect stakeholders - the group of persons and institutions that is indirectly affected by the use of the uNLock application – were identified.

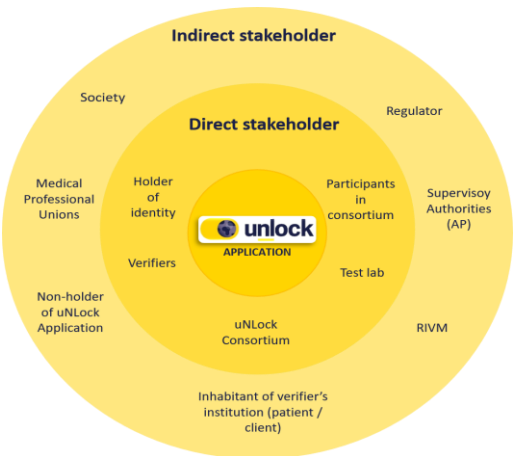

Figure 2.

Based on this distinction, the Ethics Working Group composed an overview of stakeholders (see Fig. 2), together with the benefits and harms that they could expect from the uNLock solution (Annex 1). A more detailed overview of all direct and indirect stakeholders and their interests (benefits/harms) can be found in the appendix. Example:

Holder of Identity (user of the uNLock application) The holder of identity would be able to access work by using the uNLock application to provide an irrefutable proof of health status tests. This would benefit the holder of identity by reducing the administrative burden of providing documented proof for each updated health test and the application could minimize the personal data shared by the holder of identity.
Benefits: Access to work (or people they care for); Less administrative stress; Irrefutable proof of health status; Possibility to minimize data sharing
Harms: Identity theft; Unintentionally making detrimental decisions concerning own identity; Losing identity proof, thus not being able to use it; Identity being revoked; App does not work for whatever reason; Access wrongly denied; Need to own a smart phone; Extra burden on administrative tasks; Obligation to be tested; Not being able to be tested; False idea of certainty and safety (i.e. due to quality of test or infection after test)

Step 2. Translating interests, benefits and harms into values and norms

Values and their definitions often can be vague or too abstract for the instrumentalization in the specific context, for instance formulation of design requirements. Furthermore, overly abstract definition of values often risks ignoring or obscuring relevant social, economic, and cultural differences. Thus, the second step of our conceptual investigation was to provide conceptualizations of identified values in the context of the uNLock solution and scope of its use.

Step 3: Harmonize the uNLock value set with existing research on values and norms of SSI

In order to appreciate the normative content of Self Sovereign Identity concept and elaborate on a morally conscious approach to the uNLock application, the Ethics Working Group conducted a conceptual investigation on the 10 principles of Self Sovereign Identity (2016) by Christopher Allen.

The 10 principles lay out the necessary attributes that a Self-Sovereign Identity system must have in order to uphold and protect human rights and freedom. According to Christopher Allen, ''an identity system must balance transparency, fairness, and support of the commons with protection for the individual'' in order to ensure that the user's control is at the heart of Self Sovereign Identity. Following an extensive discussion on the 10 principles, the Ethics Working Group developed conceptualizations of values and corresponding principles based on the context of uNLock. The main goal of these conceptualizations, was the translation of values into solution-specific norms, to be used in the step 2 of a Value Sensitive Design process. An overview of this assessment can be found in the appendix (Annex 3).

To conclude the conceptual investigation, the Ethics Working Group worked to harmonize conceptualizations of identified values with binding and non-binding legal sources. We have substantiated the identified values in the following legal sources: United Nations Declaration of Human Rights, Treaty on the European Union (TEU), Treaty on Functioning of the European Union (TFEU), European Charter of Human Rights (ECHR), GDPR and eIDAS.

7.4. Overview of values

The defined values were identified and conceptualized. The full list of values, conceptualizations and derived norms can be found in the appendix (Annex 2).
Example:

Value: Autonomy of identity	
Conceptualization	**Norms**
Appreciation and respect for the capacity to reflectively endorse (or not be alienated from) aspects of oneself.	Holder can solely retrieve and send his/her own credentials.
	Holder has no restrictions on to whom he/she sends credentials.
	App users have an accessible and easy way to contest injustices caused by the App

8. Study limitations

The presented conceptual stage of the VSD investigation does not claim to be a systematic and comprehensive identification of all affected stakeholders. The identification of stakeholders has to be a continuous process during the whole lifecycle of the solution to avoid exclusion and blind spots. Secondly, this conceptual mode of VSD investigation does not obviate the need for deliberative methods and tools to promote joint reflection on values during the design process – in particular, reflection by stakeholders on their own values, value tensions, and implications for design, as participants in the design process. Thus, this stage of investigation is not aiming to provide a final list of value conceptualization and norms for the design of a system. Rather it aims to provide a heuristic tool for the anticipatory identification of ethically desirable features of a resulting technical system and possible value conflicts.

9. Further investigation

The study presented in this paper represents the first step in applying VSD, identifying and conceptualizing values from the conceptual perspective.

In the different perspectives (conceptual, empirical and technical research) we have and will keep traceability in our research by applying the ethical matrix instrument: a matrix with stakeholders on one dimension and values on the other (Van der Stappen & Van Steenbergen, 2020). In the cells, we put the impact of the uNLock solution on the values of the stakeholders. We will do this for the solution as a whole, as well as for more detailed design alternatives to be considered. This will allow us to morally explain any design choices that will be made.

Looking at the set of relevant values resulting from the conceptual investigation, it is interesting to notice that some of them appeared in other digitalization studies as well. Comparing the list of values with for instance values found in comparable research in an educational context (i.e. implementation of online proctoring and design of an app for students) shows that all three applications may impact the values of autonomy, privacy, distributive justice, wellbeing and trust. In a discussion of six dominant technologies and the ethical issues that may arise from them, Royakkers et al. (2018) also refer to the values of autonomy, privacy and distributive justice, among others. Four sources are not enough to be able to draw any conclusions, of course, but it might be worthwhile to investigate whether there is a set of values that are at the core of current digitalization efforts.

One interesting avenue of further research may be to see whether it is possible to compose a validated common list of values for use as a starting point in the conceptual investigation by developers of digital solutions. Such lists have been composed before, for instance by Friedman et al. (2017), but with ongoing technological innovation, over the years other values may have come into play. To this end, we might collect and analyze cases of VSD applications in practice, and compare the outcomes with validated value lists from the academic literature. The intent of such a study, however, should not be to prevent developers from doing their own investigation into relevant values. To avoid such unintended outcomes, we may apply VSD principles to this investigation as well.

10. Insights

After performing the conceptual investigation we discussed what possible value conflicts we might encounter and have outlined three examples. However, creating an extensive list of value conflicts is not part of the scope of this white paper and should be a part of further investigation.

Scope creep vs Scope change

Controlling access to health care institutions based on COVID-19 test results is only one possible application of a credentials app. Scope creep may occur by extending the kind of credentials and/or by extending the kind of use contexts. Examples are the use of proof of negative testing to be allowed access to other public or private spaces, such as bars, restaurants, planes, trains, shops, events, or using the app to proof being vaccinated as a condition for access. Scope creep may occur without explicit ethical deliberation and may lead to undesirable consequences. On the other hand, careful ethical deliberation may inform desirability of a scope change, i.e. the well-considered decision to use the credentials app in a new situation, after careful weighing of benefits and harms for all direct and indirect stakeholders concerned.

Institutional autonomy vs Technological standardization

The availability of the app may put pressure on health care institutions to implement the app, regardless of their own view on the matter. This may negatively impact the autonomy of institutions to formulate and install their own policies. For government, it may become easier to enforce a common policy aimed at increasing the well-being of health care institution inhabitants, even further decreasing the autonomy of the institutions. These blanket regulations risk ignoring local needs and local knowledge, as well as hamper policy feed-back cycles.

Control vs Trust

In the first interviews that the uNLock team did with the tartget group ‘verifiers’ (health care institutions) the outcome was that these organizations highly rely on protocols/procedures that are shared with ‘holders’ (medical staff) and compliance with these protocols is entrusted to the ‘holders’. In other words institutions trust the medical staff to abide by the protocol which makes a ‘control’ like the uNLock solution superfluous, albeit that the solution is technically more trustworthy. The issue here is that in any case where trust is considered to be sufficient for compliance, an additional effective mechanism for the control over compliance makes trust redundant, thus undermining established social ties. Furthermore, this mechanism risks eroding an intrinsic moral value of the credential’s holder (employee) to be and be treated as trustworthy, when collaborating in an organization.

11. Open research questions

When dealing with emerging technologies one needs to consider the level of maturity of the solution. It is also an ethical dilemma to decide at which state certain ‘open questions’ need to be answered (i.e. before or after entering the market, providing the solution to the direct

stakeholders). To this end we have summed up some of the key issue that should be addressed before any large scale deployment.

Private/Publicly managed centralized vs Decentralized solution

What is the role of public and private organizations in a decentralized solution and in what way are the interests of these organization in conflict with the solution? Is there a risk of vendor lock in? Should public authorities manage the SSI ecosystem, or should public authorities only partake in the SSI ecosystem as an issuer and verifier?

Harmonization of legal framework

Is there a need to harmonize a legal framework for Self Sovereign Identity, or should requirements become in place as part of a decentralized discussion? Do existing laws and regulations suffice when working on SSI? Should the solution be deployed in the absence of technology specific regulation?

Open vs Closed Source

Should SSI solutions be fully open source, and publicly available? How are changes to core codebase managed? Do all contributors have an even say in any changes made to the solution?

Identity & Fraud

Does identity fraud pose a risk to the decentralized SSI ecosystem? What if there is no more single original source of truth? Can you lose your identity when you have lost your digital identity? Can a digital identity uphold in court?

Entry barriers to ecosystem (verify the verifier, or not?)

May all organizations that have a verifier role ask for a verification of all digital credentials of individuals? Who governs the obligation of the verifier towards the autonomy of the credential holder? see also (Van Deventer, 2020).

12. Conclusion

This white paper presents findings of the Ethics Working Group, from the conceptual phase of investigation into the ethical issues of the uNLock solution, providing identity management solutions for sharing and presentation of medical COVID-19 credentials (test results) in the context of healthcare institutions. We have provided an outline of direct and indirect stakeholders for the uNLock solution and mapped values, benefits, and harms to the respective stakeholders. The resulting conceptual framework has allowed us to lay down key norms and principles of SSI in the specific context of uNLock solution. We hope that adherence to these norms and principles could serve as a groundwork for anticipatory mitigation of moral risk and hazards stemming from the implementation of uNLock solution and similar solutions. Our findings suggest that even early stage of conceptual investigation in the framework of Value Sensitive Design, reveals numerous ethical

issues. The proposed implementation of the uNLock app in the healthcare context did not proceed further than prototype stage, thus our investigation was limited to the conceptual stage, and did not involve the practical implementation of VSD method involving translation of norms and values into engineering requirements. Nevertheless, our findings suggest that the implementation of VSD method in this context is a promising approach that helps to identify moral conflicts and risks at a very early stage of technological development of SSI solutions. Furthermore, we would like to stress that in the light of our findings it became painfully obvious that hasty implementation of medical credentials system without thorough ethical assessment, risks creating more ethical issues rather than addressing existing ones.

What can we do with the outcome?

We have gained insights with regard to possible value conflicts that need to be resolved for the purpose of Value Sensitive Design. We hope related SSI/Covid-19 solutions can benefit from this research and get in contact to share insights. We would like to stress the fact that this research was in light of the scope of the uNLock solution based and (to be) deployed in healthcare facilities in the Netherlands. Country and context/sector specific traits might create a complete different ethical perspective on this solution.

About the authors

Georgy Ishmaev (PhD), is a postdoctoral researcher at Distributed Systems section (EEMCS/ST) of Delft University of Technology. His research is focused on the ethical issues of blockchain technology applications and decentralization. Email: g.ishmaev@tudelft.nl

Roderick Noordhoek Msc, works at Rabobank as Compliance Advisor and Product Owner of the CLRS & Tech. Squad en Guilds. He also teaches at the Nederlands Compliance Instituut and is a secretary of the Ethics Committee of uNLock. Email: roderick.noordhoek@rabobank.nl

Dr. ir. Marlies van Steenbergen is professor Digital Ethics at Utrecht University of Applied Sciences, focusing on Value Sensitive Design of data-driven applications, and principal consultant enterprise architecture at Sogeti Netherlands BV. Email: marlies.vansteenbergen@hu.nl.

Nadia Vermaes (LL.B Candidate), works at Rabobank as Compliance Advisor and Lead of the CLRS & Self Sovereign Identity Guild. Email: nadia.vermaes@rabobank.nl

Reflections of the authors

Georgy: I am thankful for the opportunity to take part in such an ambitious and engaging project. Most valuable finding to me from our research work within the Ethics Working Group was an empirical observation that temptations to address emergency problems with technological solutions, should always be tempered with a holistic multidisciplinary investigation to avoid moral pitfalls and moral regrets.

Roderick: Working on this white paper inspired me and makes me realize again that ethics is an iterative process and that technology is not to be seen as value neutral. I dearly value working together with the Ethics Working Group that gave me both personal and professional insights.

Marlies: I am grateful for having been able to participate in this relevant Ethics Working Group and the worthwhile and honest discussions we had. The entire project confirmed for me that 1) digital innovations must be regarded and designed as socio-technical systems, not as purely technical systems, and 2) a valuable dialogue about the potential impact of a digital innovation on personal and societal values requires diversity in participants, and consequently a thorough empirical investigation.

Nadia: Our continuous investigations and discussions over the timespan of the project provided me with the insight that a comprehensive ethical framework can positively impact the potential and development of the technology.

Disclaimer

Up until the current moment of writing of this white paper all the work by the Members of Ethics Working Group and Ethics Committee was carried out on a voluntary basis without financial reimbursements or contractual obligations. This research was confined to the specific context of uNLock solution as a prototype for the sharing of COVID-19 credentials for healthcare employees. Any further application of uNLock solution lies outside of the scope of this report.

References

Alan, C. “The Path to Self-Sovereign Identity“, http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html, last visited 2020-08-28

Friedman, Batya, Peter Kahn, and Alan Borning. "Value sensitive design: Theory and methods." University of Washington technical report 2-12 (2002).

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory (pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4

Friedman, B.,Hendry, D., and Borning, A. (2017), "A Survey of Value Sensitive Design Methods", Foundations and Trends® in Human–Computer Interaction: Vol. 11: No. 2, pp 63-125.

Friedman, B., & Hendry, D. (2019). *Value sensitive design: Shaping technology with moral imagination*. The MIT Press.

Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. Ethics and Information Technology, 20(2), 127-142.

uNLock website: https://www.uNLockapp.nl/

Van Deventer, O., 'VERIFY THE VERIFIER - ANTI-COERCION BY DESIGN', *TNO*, https://blockchain.tno.nl/blog/verify-the-verifier-anti-coercion-by-design

Van den Hoven, J., Vermaas, P. E., & Van de Poel, I. (2015). Handbook of ethics, values, and technological design: Sources, theory, values and application domains. Springer.

Van de Poel, I. (2015). Conflicting Values Value conflict in Design for Values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), Handbook of Ethics, Values, and Technological Design (pp. 89–116). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_5

Zhong, R. (2020, May 26). China’s Virus Apps May Outlast the Outbreak, Stirring Privacy Fears. The New York Times. https://www.nytimes.com/2020/05/26/technology/china-coronavirus-surveillance.html

Annex 1: Overview of stakeholders and their benefits/Harms

Direct stakeholders

Holder of Identity (user of the uNLock application) The holder of identity would be able to access work by using the uNLock application to provide an irrefutable proof of health status tests. This would benefit the holder of identity by reducing the administrative burden of providing documented proof for each updated health test and the application could minimize the personal data shared by the holder of identity.
Benefits: Access to work (or people they care for); Less administrative stress; Irrefutable proof of health status; Possibility to minimize data sharing
Harms: Identity theft; Unintentionally making detrimental decisions concerning own identity; Losing identity proof, thus not being able to use it; Identity being revoked; App does not work for whatever reason; Access wrongly denied; Need to own a smart phone; Extra burden on administrative tasks; Obligation to be tested; Not being able to be tested; False idea of certainty and safety (i.e. due to quality of test or infection after test)
Issuers (test laboratories) Covid-19 test laboratories would be able to provide verified proofs of test results that can be communicated quickly to the holder of identity through the uNLock application. Therewith, can the test laboratories update the health status of the identity holders be updated easily after a new covid-19 test.
Benefits: Faster communication of test results; Provide verified proofs of test results; Easily update status of test
Harms: Increased security risks (i.e. through introducing more complexity to IT system); Only persons who have app can get tested (depending on scope); more fraud (because of power of solution)
Verifiers (hospitals, care homes etc.) The verifiers are the institutions that could use the uNLock application to control the access management of their facilities. Currently, the uNLock application is intended for providing controlled access to the verifiers facility for employees. Which means that the uNLock application is to be used in an employee and employer relationship.
Benefits: Give controlled access to visitors (employees/volunteers/medical researchers); Safety of personnel and hospital; Reduced stress caused by uncertainty; Increased speed of test results; Effective implementation of protocol leading to less risk
Harms: Unfair distribution of access priviledge; False idea of certainty and safety (i.e. due to quality of test or infection after test); Extra burden on administration of verifier; Increased security risks
uNLock Consortium The uNLock Consortium would be able to contribute to society by providing an application that could benefit the healthcare sector in enforcing its Covid-19 policy with regard to the entrance to the health care facility. Therewith, would the uNLock application benefit the creation of an Self Sovereign Identity digital infrastructure in the Netherlands.
Benefits: Contribution to society (in the context of healthcare); Create SSI infrastructure for the Netherlands
Harms: Delivering harmful product; Failure of Consortium
Participants in the uNLock Consortium The participants in the uNLock Consortium would be able to build upon their knowledge and experience. Therewith, would they be able to promote their services to society.
Benefits: Knowledge building; Promote their services
Harms: Reputational damage; Loss of public trust; Liability for harm caused; Conflict/disagreements between consortium partners

Indirect stakeholders

Society
Benefits: UNLocking workplaces for vital functions of health professionals; Health; Economy through increased health of society; Unburdening health sector
Harms: Exclusion/discrimination (no test/proof/app, no access); People intentionally attracting COVID-19 infection to get access (future risk); Misuse in other contexts; Function creep; Fraudulent tests (if system is more effective, more reason for testing, more fraud); Power misuse; More infections because of false sense of safety; Loss of public trust in experts.
Medial professional unions Protecting the rights of medical professionals that who can be part of the solution playing the role of Holder
Benefits: UNLocking society (in the context of healthcare); Complaints from professionals
Harms: Need for unavailable expertise; Less power
Non-holder of the uNLock application
Benefits: UNLocking society (in the context of healthcare)
Harms: Exclusion from access/work; Being marginalized; Peer pressure; Fear of job loss
Inhabitant of the facility of the verifier (hospitals, care homes etc.)
Benefits: Health; Visitors allowed
Harms: Less personnel; False idea of certainty and safety (i.e. due to quality of test or infection after test)
RIVM
Benefits: Insight in test results; Better access control
Harms: Doing business through non-public entity; More infections because of false sense of safety; Reputational damage; Loss of public trust
Supervisory authorities (AP)
Benefits: Privacy-proof app
Harms: Reputational damage; Loss of public trust
Regulator
Benefits: UNLocking society (in the context of healthcare); Health; Available research; on use of SSI/increased legal clarity on application of SSI
Harms: Lack of regulation; Claims of discrimination; Misuse in other contexts; Fraudulent tests; Power misuse; Reputational damage; Loss of public trust

Annex 2: uNLock Values, Conceptualizations and Norms

Value: Autonomy of identity	
Conceptualization	**Norms**
Appreciation and respect for the capacity to reflectively endorse (or not be alienated from) aspects of oneself.	Holder can solely retrieve and send his/her own credentials. Holder has no restrictions on to whom he/she sends credentials. App users have an accessible and easy way to contest injustices caused by the App
Value: Honour of the individual	
Conceptualization	**Norms**
Value of recognition and approval that links reputation with conduct and helps sustain existing patterns of social ties. It is intrinsically tied to respect and the worthiness to be respected.	Holder can retrieve and share credentials without involvement of credentials issuer System does not allow the use of medical credentials for profiling or creation of reputation scores System prevents the collection of credentials or other private data by the third parties
Value: Dignity	
Conceptualization	**Norms**
Recognition of treating humans as self-governing persons and respect for the inherent capacity for upholding one's principles.	Installation and use of app is always optional and voluntary Holder can choose to install App. Holder gives Informed and independent consent on Terms and Conditions of use.
Value: Individual agency	
Conceptualization	**Norms**
Individual agency - ability for an individual to act in accordance with a goal the agent has adopted on the basis of an overall practical assessment of her options and opportunities.	Holder gives informed and independent consent on Terms and Conditions of use. Holder is properly informed of how to exercise their right to contest Terms and Conditions are minimized to what is required for the goal of the solution
Value: Transparency	
Conceptualization	**Norms**
Availability and integrity of information, the conditions of its accessibility including considerations on how this information may pragmatically or epistemically support the user's decision-making process.	Terms and Conditions for use of the App are published on the website on A2 language level. Code of solution is open-source. The information on functioning of a system and data flows is provided to users in understandable form Terms and Conditions are formulated for maximum transparency. Interests of the different Partners of the Consortium are are publicly declared Partners comply to CoC.

Value: Privacy	
Conceptualization	**Norms**
Right of an individual to determine what information about himself or herself can be communicated to others.	Lawfulness, fairness and transparency - Personal data must be processed lawfully, fairly and in a transparent manner; Purpose limitation - Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes (with exceptions for public interest, scientific, historical or statistical purposes); Data minimisation - Personal data must be adequate, relevant and limited to what is necessary in relation to purposes for which they are processed Accuracy - Personal data must be accurate and, where necessary, kept up to date. Inaccurate personal data should be corrected or deleted; Retention - Personal data should be kept in an identifiable format for no longer than is necessary (with exceptions for public interest, scientific, historical or statistical purposes); Integrity and confidentiality - Personal data should be kept secure. Accountabilty - One must not only comply with the six general principles, but also be able to demonstrate compliance with them.
Value: Non-discrimination	
Conceptualization	**Norms**
Guarantee that human rights are exercised without discrimination of any arbitrary kind.	Covid-19 test result in the App cannot be sole decisive aspect for employment. Safeguards are implemented to prevent the use of the app in other contexts. The absence of the app does not affect rights of a credentials holder in any way.
Value: Organisational repsonsibility	
Conceptualization	**Norms**
The possibility of a moral hazard. A person is able to make decisions and/or take actions on behalf of another person or entity, where the person is motivated to act in their own best interest, which are contrary to those of the person or entity they are working on behalf of. Embracing distributed responsibility and anticipatory approach towards risks and harms	Any (potential) conflict of interest is avoided for which safeguards are in place (I.e. Code of Conduct). Employees of the solution must be able to exercise their tasks independently and in the best interest of the solution and its users formalized by contracts between the Consortium, its partners and third parties. Responsibilities of the entities involved in the consortium are defined and documented. The information on the identities and responsibilities of parties involved in the consortium is publicly available The members of consortium are always able to express concerns regarding possible risks and harms to users and non-users of the app

Value: Security	
Conceptualization	**Norms**
Ensure the presence of peace, safety and the protection of human rights or absence of crisis or threat to human dignity.	All critical elements of the solution is subject to independent security audits
	The solution is assessed by independent, external stakeholders (Technical and Ethical board)
Value: Freedom of movement	
Conceptualization	**Norms**
Freedom of movement of individuals throughout the world in public spaces and spaces where they are entitled to.	The non-use of an app, can not be a sole basis for formal or de-facto restrictions on the freedom of access for individuals to public spaces or other spaces where individuals are entitled to be Safeguards are implemented to prevent the use of app as access management tool in other contexts.
	Verify the verifier and issuer, credentials can only be sent and checked by verified issuers and verifiers.
Value: Well being	
Conceptualization	**Norms**
The state of being comfortable, healthy and happy. The well-being of society, medical professionals, clients/patients and non-holders should be sustained and protected.	The solution cannot do physical or psychological harm. This should be assessed through pilots and interviews of Holders, Issuers and Verifiers.
Value: Welfare (sub value of well being)	
Conceptualization	**Norms**
The state of physical and material well-being. People being able to do their job and make a living.	The solution and its use may not be an incentive for employers to terminate the contract of an employee or discriminate toward potential employees, this is described in Terms and Conditions.
Value: Health (connected to well being)	
Conceptualization	**Norms**
Six main dimensions of health are distinguished: bodily functions, mental functions & perception, spiritual/existential dimension, quality of life, social & societal participation, and daily functioning. The ability to adapt and to self-manage.	The solution may not be an argument to force individuals to work when the individual feels unhealthy this is described in Terms and Conditions.
	Storing of the Holder's credentials is only allowed if Holder provides them him-/herself for storage. Storing of the Holder's credentials is only allowed through a secure, compliant system and can only be viewed by entitled medical staff.

Value: Solidarity	
Conceptualization	**Norms**
Willingness to give psychological and/or material support when another person is in a difficult position or needs affection. Unity or agreement of feeling or action, especially among individuals with a common interest; mutual support within a group. Desire to unburden medical professionals.	The solution should benefit all users and not harm non-users, more than it does harm. I.e. if an employee has to wait before he/she can start working this can not be seen as free time(?).
Value: Stakeholder power (relevance)	
Conceptualization	**Norms**
The capacity or ability to directly influence the behaviour of others regarding one's own rights and legitimate interests. Knowledge about other people can provide power over them or others.	Stakeholders of the solution may not exercise its power by any means to pursue their own interests
Value: Inclusion	
Conceptualization	**Norms**
Ensure that the needs of disadvantage (social) groups such as those without access to mobile internet, illiterate, disabled people, are considered so that no one is left behind.	The solution shall include citizens as best to its abilities and shall not restrict disadvantaged (social) groups in using the solution.
Value: Distributive justice (fairness)	
Conceptualization	**Norms**
To treat all individuals with equal characteristics in the context of data that is processed and interaction to the system equally. And with the same respect in terms of use.	The solution must treat users equally and in alignment with the terms of use of the solution. The solution should not put unjust burden on its users. The solution should not put unjust burden on non-users of app or other affected stakeholders
Value: Autonomy of employee	
Conceptualization	**Norms**
Everyone has the right to work, to free choice of employment, to just and favourable conditions of work. Within the relationship between employer/employee, the employee is only bound by law and contracts, the latter can only be signed by informed consent and without any form of pressure inflicting on the autonomy of the individual.	Employees must be able to determine for themselves whether they want to use the solution and must be able to give informed consent. A suitable alternative is offered.

Value: Trust	
Conceptualization	**Norms**
The justified believe and comfort of an individual or entity in the reliability of the system and/or all related entities, and competence and benevolence of individuals, that have a direct impact on the system. A psychological state compromising the intention to accept vulnerability based on positive expectations of the intentions and behaviours of another.	The solution must safeguard the justified belief and comfort of the user by being truthful and transparent about the uNLock system and the uNLock consortium.
Value: Psychological attitude (sub-value trust)	
Conceptualization	**Norms**
Justified belief that entity (individual, organisation) trusted to act on your behalf, will reliably and completely act in your best interests.	The solution (and the uNLock consortium) must operate transparently and in the best interests of its direct and indirect stakeholders.
Value: Acting on trust (sub-value trust)	
Conceptualization	**Norms**
Delegation of power to other party to act in our best interests at the expense of increased vulnerability from that party.	Terms and conditions are written in the best interest of stakeholders and actively accepted by stakeholders.
Value: Value of justified trust (sub-value trust)	
Conceptualization	**Norms**
Capacity to delegate complex or costly actions on your behalf to other party without exessive risks.	Terms and conditions are written in the best interest of stakeholders and actively accepted by stakeholders.
Value: Trust capital (sub-value trust)	
Conceptualization	**Norms**
Sustained justified belief in a given society that benefits of cooperation based on trust overweight possible risks and vulnerabilities.	Solution is supported and maintained by trustworthy parties in a transparent way
Value: Institutional Reputation	
Conceptualization	**Norms**
The generalized beliefs and/or opinions of the public in the system or the entities that have a direct impact on the system, which can be found on any public source/forum and is translated by i.e. journalist.	The uNLock Consortium must contribute as effectively as possible to maintain the generalized beliefs and/or opinions of the public about the participants participating in the solution and the healthcare system in general.

Value: Individual's Reputation	
Conceptualization	**Norms**
Value of recognition and approval that links reputation with conduct and helps sustatin existing patterns of social ties. It is intrinsically tied to respect and the worthiness to be respected.	The solution must put up safeguard to protect the reputation and dignity of its users. System does not allow the use of medical credentials for profiling or creation of reputation scores Holders always have independent access to the latest up-to-date status of their credentials System does prevents the sharing collection of credentials or other private data with third parties
Value: Accessibility	
Conceptualization	**Norms**
Having the opportunity and capacity to access the system for authorized parties.	The solution must be accessible and up-to date for all authorized parties and on-boarded users at all times.
Value: Efficiency	
Conceptualization	**Norms**
Sustainable and optimized use of resources and time.	The solution supports efficiency goals of its users.
Value: Autonomy of policy	
Conceptualization	**Norms**
Every organization has the right to draft and effectuate policy that requires adherence by stakeholders that have a contractual relationship to the organization. Provided, that these policies do not contradict stablished ethical and legal standards.	Verifiers must be able to uphold their own policies.
Value: Right to complain	
Conceptualization	**Norms**
Practically feasible opportunity for all individuals that are direct or indirect stakeholders to object, to make suggestions, to be heard and taken seriously.	uNLock Consortium is responsible and accountable for facilitation and protection of any direct stakeholder's right to file a complaint. Furthermore, it may not restraint direct stakeholders to file a complaint at the relevant authority.

Annex 3: Matching SSI principles to uNLock Values

SSI Principles by Christopher Allen (2016)	uNLock VSD conceptualization	uNLock VSD value
Existence: user must have an independent existence.	Autonomy of choice to use the provided ID scheme or alternative method (e.g. a paper credential).	Autonomy of identity
Control: user must control their identities.	Information is only accessed with consent.	Dignity
Access: users must have access to their own data.	User must be able to access their data and any associated claims without the interference of gate keepers or intermediaries. The individual should only be granted access to his/her/its own identity and not those of others.	Individual agency
Transparency: systems and algorithms must be transparent.	The system design is open-source. The system must operate in an intelligible and easily accessible format, using ''clear and plain language''. The implications of the use of the system must be explained to the user.	Transparency and trust
Persistence: identities must be long-lived	Temporary identifiers necessary for privacy	Privacy
Portability: information and services about identity must be transposable.	Transferability of this ID scheme in other contexts has limited desirability.	Privacy and dignity
Interoperability: identities should be as widely usable as possible.	The user should be able to provide cross international border identification, without losing control of what information is shared. Users should be able to maintain their identities across platforms and geographical locations.	Non-discrimination / fairness
Consent: user must agree to the use of their identity.	Each data transaction must be authorized (and only executed when) by user's consent.	Trust and agency
Minimization: disclosure of claims must be minimized.	Data shared in the credentials should be as minimal as possible. Developers and system administrators should consider data minimization techniques.	Privacy and security
Protection: the rights of users must be protected.	Users data and any associated claims must be as protected as possible and data may not be re-used in other contexts.	Security