Proceedings of the
2nd European Conference on the Impact of Artificial Intelligence and Robotics

A Virtual Conference hosted by
Instituto Universitário de Lisboa (ISCTE-IUL)
Portugal

22–23rd October 2020

Edited by
Dr Florinda Matos

A conference managed by ACI, UK

Proceedings of the

European Conference on the Impact of Artificial Intelligence and Robotics
ECIAIR 2020

Supported By
Instituto Universitário de Lisboa (ISCTE-IUL)
Portugal

Edited by
Florinda Matos

22 – 23rd October 2020

Review Process
Papers submitted to this conference have been double-blind peer reviewed before final acceptance to the conference. Initially, abstracts were reviewed for relevance and accessibility and successful authors were invited to submit full papers. Many thanks to the reviewers who helped ensure the quality of all the submissions.

Ethics and Publication Malpractice Policy
ACIL adheres to a strict ethics and publication malpractice policy for all publications – details of which can be found here:
http://www.academic-conferences.org/policies/ethics-policy-for-publishing-in-the-conference-proceedings-of-academic-conferences-and-publishing-international-limited/

Conference Proceedings
The Conference Proceedings is a book published with an ISBN and ISSN. The proceedings have been submitted to a number of accreditation, citation and indexing bodies including Thomson ISI Web of Science and Elsevier Scopus.

Author affiliation details in these proceedings have been reproduced as supplied by the authors themselves.

The Electronic version of the Conference Proceedings is available to download from DROPBOX https://tinyurl.com/ECIAIR20 Select Download and then Direct Download to access the Pdf file. Free download is available for conference participants for a period of 2 weeks after the conference.

The Conference Proceedings for this year and previous years can be purchased from http://academic-bookshop.com

E-Book ISBN: 978-1-912764-73-0
Book version ISBN: 978-1-912764-74-7

Published by Academic Conferences International Limited
Reading
UK
www.academic-conferences.org

Contents

ECIAIR Preface

These proceedings represent the work of contributors to the 2nd European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR 2020), hosted by ACI and Instituto Universitário de Lisboa (ISCTE-IUL), Portugal on 22-23 October 2020. The Conference Chair is Dr Florinda Matos, and the Programme Chairs are Dr Ana Maria de Almeida and Prof Isabel Salavisa, all from Instituto Universitário de Lisboa (ISCTE-IUL), Portugal.

ECIAIR is now a well-established event on the academic research calendar and now, in its 2nd year, the key aim remains in the opportunity for participants to share ideas and meet people who hold them. The conference was due to be held at Instituto Universitário de Lisboa (ISCTE-IUL), Portugal, but because of the global Covid-19 pandemic, it was moved online as a virtual event. The subjects covered in the papers illustrate the wide range of topics that fall into this important and ever-growing area of research.

The opening keynote presentation is given by Prof. Mário Figueiredo, from University of Lisbon, Portugal, on the topic of "Artificial Intelligence: Historical Aspects, Modern Applications, and Implications". The second day of the conference will be open by Prof. Jean-Gabriel Ganascia, Université Pierre et Marie Curie (UPMC), France and a member of the Institut Universitaire de France, France, who will talk about "Why do we need Ethics and not just Regulations in AI and Robotics?".

With an initial submission of 60 abstracts, after the double blind, peer review process there are 25 academic research papers, 1 PhD research paper, and 1 work-in-progress paper published in these Conference Proceedings. These papers represent research from Brazil, Cuba, Denmark, Finland, Germany, Poland, Portugal, Russia, Sweden, Switzerland, UK, USA.

We hope you enjoy the conference.

Dr Florinda Matos
Instituto Universitário de Lisboa (ISCTE-IUL)
Portugal
October 2020

ECIAIR Conference Committee

Dr Kareem Kamal A.Ghany, Beni-Suef University , Egypt; Prof Azween Abdullah, Taylors University, Malaysia; Dr Kinaz Al Aytouni, Arab International University, Syria; Dr Hanadi AL-Mubaraki, Kuwait University, Kuwait; Prof Hamid Alasadi, Iraq University college, Iraq; Prof Laurice Alexandre, Sorbonne Paris Cité , France; Dr José Álvarez-García, University of Extremadura, Spain; Dr Xiomi An, Renmin University of China, China; Prof Antonios Andreatos, Hellenic Air Force Academy, Greece; Prof Oscar Arias Londono, Institucion Universitaria de Envigado, Colombia; Dr Gil Ariely, Interdisciplinary Center, Herzliya, Israel; Dr Sotiris Avgousti, Cyprus University Of Technology, Cyprus; Prof Rosalina Babo, Polytechnic of Porto, Porto Accounting and Business School, Portugal; Prof Liz Bacon, Abertay University, UK; MSc Jordi Bieger, TU Delft, Netherlands; Prof Elias Carayannis, George Washington University, USA; Prof Davide Carneiro, Polytechnic Institute of Porto, Portugal; Prof Karen Cham, University of Brighton, UK; Prof Jim Chen, U.S. National Defense University, U.S.A.; Dr Pericles Cheng, European University Cyprus, Cyprus; Prof Koteshwar Chirumalla, Malardalen University, Sweden; Mr. David Comiskey, Ulster University, UK; Dr Leonardo Costa, Universidade Católica Portuguesa - Católica Porto Business School, Portugal; Prof Carmen-Eugenia Costea, The Bucharest University of Economic Studies, Romania; Prof Larry Crockett, Augsburg University, USA; Dr Marija Cubric, University of Hertfordshire, UK; Francesca Dal Mas, Università di Udine/Università la Sapienza, Italy; Assc Ben Daniel, University of Otago, New Zealand; Prof Justine Daramola, Cape Peninsula University of Technology, South Africa; Geoffrey Darnton, WMG, University of Warwick, UK; Mr Martin De Bonis, Alma Mater Studiorum, Italy; Prof Armando Carlos de Pina Filho, Federal University of Rio de Janeiro (UFRJ), Brazil; Dr Martin De Saulles, University of Brighton, UK; Dr María de la Cruz Del Río-Rama, University of Vigo, Spain; Dr Souâd Demigha, CRI Univ of Paris 1 La Sorbonne, France; Paolo Di Muro, Politecnico di Milano School of Management, Italy; Dr Mihaela Diaconu, ”Gheorghe Asachi” University of Iasi, Romania; Inês Domingues, DEIS-ISEC, Portugal; Dr Patricio Domingues, Polytechnic Institute of Leiria, Portugal; Prof Yanqing Duan, University of Bedfordshire, UK; Prof. John Edwards, Aston Business School, UK; Dr Kelechi Ekuma, The Global Development Institute, University of Manchester, UK; Dr Scott Erickson, Ithaca College- School of Business, USA; Dr José Esteves, IE Business School, Spain; PhD Fernanda Faini, CIRSFID - University of Bologna, Italy; Dr Georgios Fessakis, University of the Aegean, Greece; Prof Eric Filiol, ENSIBS, Vannes, France & CNAM, Paris, France; Dr Panagiotis Fotaris, University of Brighton, UK; Prof Andreas Giannakoulopoulos, Ionian University, Greece; Dr Valerie Priscilla Goby, Zayed University, United Arab Emirates; Dr Amol Gore, UAE Government HCT, UAE; Dr Paul Griffiths, Ecole de Management de Normandie, Oxford, UK; Dr Hossein hakimpour, IAU , Iran; Prof William Halal, George Washington University, USA; Prof Ali Hessami, City University London, UK; Dr Grant Howard, University of South Africa (Unisa), South Africa; Prof Ulrike Hugl, Innsbruck University, Faculty of Business and Management, Department of Accounting, Auditing and Taxation, Austria; Prof Hamid Jahankhani, Northumbria University London, UK; Dr Aman Jatain, Amity University, India; Dr Runa Jesmin, University of Roehampton, UK; Dr Jari Jussila, Häme University of Applied Sciences, Finland; Dr Selvi Kannan, Victoria University, Australia; Prof Ergina Kavallieratou, University of the Aegean, Greece; Dr Harri Ketamo, Headai ltd, Finland; Dr Nasrullah Khilji, University of Bedfordshire, UK; Prof Tatiana Khvatova, Peter the Great St. Petersburg Polytechnic Universtity, Russia; Prof Jesuk Ko, Universidad Mayor de San Andres, Bolivia; Prof Michael Kohlegger, Institute for Web Technologies & Applications, Austria; Prof Renata Korsakiene, Vilnius Gediminas Technical University, Lithuania; Prof Ibrahim Krasniqi, University of Peja, Kosovo; Prof Konstadinos Kutsikos, Business School, University of the Aegean, Greece; Dr Jean Lai, Hong Kong Baptist University, Hong Kong; Dr Isah Abdullahi Lawal, Noroff University College, Norway; Dr Efstratios Livanis, University of Macedonia, Greece; Prof Eurico Lopes, Instituto Politécnico de Castelo Branco, Portugal; Dr Martin Magdin, Constantine the Philosopher University in Nitra, Faculty of Natural Sciences, Department of Informatics , Slovakia; Dr Paolo Magrassi, Alephuture, Switzerland; Dr Hossein Malekinezhad, Islamic Azad University, Naragh Branch, Iran; Prof António Martins, Universidade Aberta, Portugal; Prof Maurizio Massaro, Ca' Foscari University of Venice, Italy; Dr Nuno Melão, Polytechnic Institute of Viseu, Portugal; Prof Anabela

Introducing the Concept of Digital-Agent Signatures for Human-Robot-Robot-Human Interaction

Alexander Pfeiffer[1.2,3], Alesja Serada[7], Mark Bugeja[3], Stephen Bezzina[6], Thomas Wernbacher[2] and Simone Kriglstein[4 5]
[1]Massachusetts Institute of Technology (MIT), Cambridge, MA, USA
[2]Center for Applied Game Studies, Donau-Universität Krems (DUK), Austria
[3]University of Malta (UoM), Msida, Malta
[4]Austrian Institute of Technology GmbH (AIT), Vienna, Austria
[5]University of Vienna, Vienna, Austria
[6]Ministry for Education and Employment, Floriana, Malta
[7]University of Vaasa, Finland
alex_pf@mit.edu
alesja.serada@uwasa.fi
mark.bugeja@um.edu.mt
mail@stephenbezzina.com
Thomas.wernbacher@donau-uni.ac.at
simone.kriglstein@ait.ac.at
DOI: 10.34190/EAIR.20.004

Abstract: Digital/electronic identities are essential components of collaborative robots/robots and human-robot/robot-human interactions. Through such identities, digital agents (AI powered software or robots/bots) are entrusted with tasks in the name of certain individuals/companies. Digital identities can come from various sources; these can be assigned by an employer, through a service provided by a government entity or an external company specializing in the creation of such signatures or generated through an interface like Facebook Connect. All these different sources offer a range of varying levels of trust, both within the institution where the signature is principally used, but especially when interacting with third parties. Ultimately, this level of trust or its valuation is a determining factor in how far the authorization of the respective digital/electronic signature goes. The authors describe the application of digital/electronic signatures from human employees or legal entities which, simultaneously with the main task, generate sub-signatures for the respective digital agent.The topic is presented from a technical perspective as well as from a social science point of view.

Keywords: Digital Agents, Digital Identity, Self-Sovereign Identity, Blockchain, AI

1. Introduction

A new high-class gaming PC bought by an 11-year-old via Alexa voice command; a reservation in a luxury restaurant, unintentionally made by google-assistant; or a 1000 Euro tax overcharge, due to an error in the AI-assisted accounting software; cases, which are solved nowadays due to well-written terms and conditions, a helpful service hotline or in the worst case in court.

Now imagine these kinds of problems on a larger scale. An unauthorized person orders production machines via an AI-assisted purchasing software; a digital assistant books an unplanned business trip without knowing who gave the order initially; a minor software error that miscalculates the tax by a "0", or production machines that have not been operated in accordance with their intended use.

Most of these cases are unlikely to be solved on a goodwill basis but will have more significant consequences. Often, however, one will have to ask first who should be held responsible at all. The ever-faster technological progress is therefore not only a blessing (e.g. through cost savings due to efficient production), but also brings problems with it. Problems which from a social, ethical, political, technological or legal we have not yet managed to solve.

Triggered by the current Covid-19 situation, automation and production using machines with AI elements are being discussed even more intensively and pursued more rapidly. (BBC Article 2020)[1]

[1] https://www.bbc.com/news/technology-52340651 (

To put the problem into perspective: We currently lack a trustworthy solution to show orders given to "AI-assisted Software Solutions", "AI-driven language assistants", to AI-assisted robots/bots" between all parties involved. The problem becomes more prominent, the more partners are involved and whether they go beyond companies, municipalities, countries or even continents.

Trust is described by Jøsang (2016) as a subjective belief in the reliability, honesty and security of an entity on which we depend for our welfare, and these entities contain software, hardware, data, people and organizations. Two components might be needed to create this trust in the digital space: digital/electronic identities and storage of data and processes on data hubs that are accepted by all parties involved.

2. Research Questions

The authors pursue the following research questions:

- How can secured digital identities be transferred to AI agents?
- What role can Blockchain technologies in connection with different forms of ID-verification play?
- Which aspects must be considered in the ethics debate, especially when AI and blockchain takes over our activities in a more complex (though still human-defined) framework.

3. Aim and Methodology

We reviewed the literature on existing solutions and discussions on the topic of digital agents, virtual environments, self-sovereign identity, qualified digital signatures, ethical discussions related to virtual world/agents and blockchain. In our research, we stress the importance of performative, multiple and adjustable digital identities that can be constructed (or generated as sets of signatures and sub-signatures) and controlled in a way similar to avatars in a virtual environment.

4. Related Research

Goodell & Aste (2019) suggest that potential users of digital identification systems should be free to operate several instances of identities, each suited for a specific aim. The Authors provided a general blueprint for 'trustless' interactions with multiple identifiers but did not extend their concept to specific use cases. In this paper, this idea is taken one step further by providing an actual technical embodiment of a similar idea.

The design of digital identities in virtual worlds provides another fruitful perspective. Full anonymity and inconsistency of a player's identity between play sessions, provoked online abuse, even in the earliest virtual worlds (McDonough 1999). However, as McDonough (1999) shows, making all interactions between players open and public lead to players' discomfort and protective behaviors to restore at least some level of privacy. Thus, a stable identity in a digital world is required to create trust, but a person should also be able to project different facets of this identity to different actors, just as we take up different roles in social interactions. Qualitative studies of online identities and privacy management in social networks produce the same conclusion. In general, digital ethnographers have successfully demonstrated that it would be a mistake to reduce relationships between real-world identities and online personas to direct, one-to-one connection (Marwick and Boyd 2014, Bancroft and Scott Reid 2017).

5. Blockchain

Blockchain can play a key role in the non-manipulable and trusted storage and application of digital identities, their transfer to digital agents and the recording of tasks performed by those. Blockchain Systems as we know them today are based on the white paper “Bitcoin: A Peer-to-Peer Electronic Cash System”, by the anonymous author Satoshi Nakamoto (2008). Blockchain technologies belong to the Distributed Ledger Systems or DLTs in short. DLTs work through different computers that store information of the same type. The ledger is therefore divided into different locations, operated by different persons or companies, none of these people or companies has to know or personally agree with each other when using a public Blockchain. Blockchains are unique due to the way they operate, which is based upon a set of rules. These rules vary slightly depending on the Blockchain system used. Transactions are then combined in a block and stored in encrypted form. This process is intended to ensure that the same information is actually stored on the distributed systems and that there is no file or text information among them that may have the same file name and size as all the others but does not contain the correct information. The storage process of a Blockchain is, therefore based on the fact that new data blocks are continuously generated. Each of these new entries (blocks) increases the size of the Blockchain.

Blockchain systems can operate in three different ways:

- Private Blockchain: is a closed system and is operated exclusively within organizations, companies or government structures. No information is passed on to the outside world unless there is evidence that a transaction has taken place.
- Blockchain operated by a consortium: serves connected parties who have a common goal. Consortium partners may join the Blockchain, based on joint agreements.
- Public Blockchain: has no restrictions on joining or leaving the Blockchain. All information is public, although it is possible to store some information in encrypted form.
- Private and consortium Blockchains can also store information on a public Blockchain, for example, the hash value of all transactions within 24 hours. This keeps the data content itself private but ensures that no data manipulation takes place retroactively. Not block by block, but still, as in the example above, for all data older than 24 hours

6. Digital Identities

Digital/electronic identities are essential components of collaborative robots/robots and human-robot/robot-human interactions. Through such identities, digital agents (AI powered software or robots/bots) are entrusted with tasks in the name of certain individuals/companies. Digital identities can come from various sources; these can be assigned by an employer, through a service provided by a government entity (For example signatures that comply with the EIDAS regulation[2]) an external company specializing in the creation of such signatures, the self-sovereign identity (SSI) movement (Sovrin Foundation) or generated through an interface like Facebook Connect. All these different sources offer a range of varying levels of trust, both within the institution where the signature is principally used, but especially when interacting with third parties. Ultimately, this level of trust or its valuation is a determining factor in how far the authorization of the respective digital/electronic signature goes.

The first state-supported pilot project for a digital identity on blockchain in the EU was launched in Zug, Switzerland, in September 2017 (Blockchain-Identität für alle Einwohner 2017). It is based on the Ethereum blockchain. In June 2018 these blockchain identities were officially used for voting (Eixelsberger et al. 2019, 514).

Another type of projects can be seen as "data cooperatives" described by Giannopoulou (2020): they approach "data as a common value" and create tools for its collective regulation. However, community standards for data management in such projects remain opaque. If closed ecosystems of data emerge as a result, abuse and exploitation within them are technically viable. A non-authoritarian way to manage digital identities is to provide as many opportunities for integration as possible.

7. Digital Interaction & the role of digital agents

We can distinguish between three different types of interaction:

1. Human with computer interaction: The average person logs in 7-25 times per day (Greene). In the simplest form of a login system used. This can either be assigned by a system administrator (human or software) or by the user himself.
2. Computer to computer interaction: (digital-agent with digital-agent, digital-agent with software, software with digital-agent). In this case, too, a digital identity in the sense of proof of entitlement must be provided, in the best-case scenario, this process can track right back to the original source. It is essential, however, that the instruction to the software or digital-agent is guaranteed by the most secure authentication possible.
3. Computer to human interaction: This takes place when a digital-agent, approaches a human to enter further data, to perform a production step or to mark work as completed. Like the previous case, it is also vital that the system can trace from which source, or from which sources, the initial order originated.

Wooldridge points out that there is no generally agreed definition of agents. In 2000 he proposed the following definition, which reflects a revision of his thoughts from 1995. Wooldridge:

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2014.257.01.0073.01.ENG

"An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives."

But he also says that this definition does not yet reflect the degree of autonomy of an agent nor the space in which it is located.

Monostori et al. describe the role of agents along with other factors as follows. Digital agents should:

- Have a purpose to fulfill,
- Perform autonomous behavior and control both of their actions within the environment,
- Perform real-time information processing and adapt themselves to new situations,
- Prioritize events in accordance with their preferences,
- Exhibit intelligence, to some degree, from applying fixed rules to reasoning, planning, and learning capabilities,
- Interact with their environment in which they are operating, including the interaction with other agents,
- Be adaptive, that is, capable of tailoring their behavior to the changes of the environment without the intervention of their designer,
- Work as genuinely and transparently as possible, and
- Be credible and trustworthy in providing information to others.

Obviously, spaces can exist where more than one agent exists. Huhns and Stephens describe the conditions of such environments:

- Multiagent environments provide an infrastructure specifying communication and interaction protocols.
- Multiagent environments are typically open and have no centralized designer.
- Multiagent environments contain agents that are autonomous and distributed, and may be self-interested or cooperative.

Burden and Savin-Baden (2019) define four different types of "AI-Systems", which can be well adopted for the thoughts about digital agents:

- Simple Algorithms – probably 99% of most computer programs, even complex Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) systems since they are highly linear and predictable.
- Complex Algorithms – programs such as, but not limited to, machine learning, deep learning, neural networks, Bayesian networks and fuzzy logic where the complexity of the inner code starts to move beyond simple linear relationships. Many systems currently referred to as AI sit here.
- Artificial General Intelligence – closer to what the public image of AI is, a system that can be applied to a wide range of problems and solve them to a better or similar level as a human.
- Artificial Sentience – beloved of science-fiction, code which 'thinks' and is 'self-aware'.

8. Proposal of the E-ID Wallet concept and Digital Agent Signature

How can the various concepts of digital identities described so far be applied to digital agents? And what key role can Blockchain technologies play? To answer the first research question, the authors would now like to discuss the concepts of the E-ID wallet and the digital agent signature.

An E-ID wallet - as perceived by the authors - is a wallet for blockchain-tokens, which is linked to one or more digital signatures of the owner(s). Valid signatures include government-issued signatures, any signature from a self-sovereign identity app, a signature issued by an educational institution, signatures issued by an identity verification company or a connection to a social media account. There are different levels of trust in the digital signature to be considered.

Depending on the selected signature type, the proof of the signature transaction is stored and displayed differently. The signature hash value can be published, for example, on a protocol page of the respective trust center, on the blockchain used by the SSI app, in which case, a token (including the private data as encrypted message) is sent from the (signed E-ID wallet) of the SSI app provider to the newly signed E-ID wallet of the user.

The signature chain can be retrieved for each E-ID wallet. For example, if an E-ID wallet is dedicated to a department of a company, the user will see that the primary account of the company and the person responsible for the department have signed as well as the user who originally signed the main account of the company and how the primary person responsible for the department got this status. Whether the private data is publicly accessible or encrypted is, of course, always subject to the person or institution and their needs. In other words, whether it is essential that everyone can see whom the wallet is assigned to, or if only persons, company-partners, or other departments of a company who gain access to this information should know the ownership.

However, the distinctive feature of the E-ID Wallet is that, in addition to digital identities, it can and should also hold blockchain-based tokens and can, therefore, be used for utility tokens linked to an identity on the one hand, but also for cryptocurrencies as a form of payment with proof of identity (to counteract money laundering and other similar problems) on the other.

Now it is a matter of connecting digital signatures and E-ID wallets with digital agents so that their distribution of tasks and progress is stored with the highest possible security and allows the digital agents to interact with third parties.

It should also be noted that some blockchain-token wallets already offer the possibility to name the wallet publicly (for example original Ardor wallet, but here it is a pure self-authentication). One way of proving the identity of a blockchain wallet would be to have a digitally signed PDF that specifies the blockchain address and is digitally signed by the user, with a transaction taking place from the blockchain address that references the PDF and its hash value together with possible location.

The basic idea of the digital agent signature is to connect the (digital) identity of the user of digital-agent with the digital agent itself. The user uses a signature that is available in his SSI app. This allows the user to select the appropriate signature for the different applications. When the digital-agent is instructed to carry out an administrative task, such as monitoring and paying tax returns, a government-issued signature will be used. If the digital-agent is instructed to search for and purchase the best possible car insurance, a signature issued by an identity verification company is used, a digital-agent is instructed to compile and enroll in the best schedule for study. That signature issued by the university is then used, and if the digital-agent is acting on behalf of the user on a love mediation platform, an even lower level of verification may be sufficient.

If the signature is used in a work context, a custom SSI-app provided by the company with authorized signatures, or an SSI-app that the user typically uses is used and where there are one or more signatures to choose from, is able to verify the identity of the user, the user’s position in the company, the user’s rights within the company and also the authenticity of the company for which the user is operating can be utilized.
The particular feature of the digital-agent signature is that a token transfer is triggered during the signature process. From the E-ID wallet of the person or company for whom this person works to the digital-agent's E-ID wallet. The data can be controlled via a system of shared keys in various levels, which are again connected to an identity management system.

9. Ethical debate

The discussion on identities for digital agents and their authorizations must not only be conducted from a technical perspective, but also from an ethical perspective. The authors would now like to address the different perspectives of the discussion and thus answer the second research question. Identity management in digital spaces acquired new meanings after the introduction of blockchain technologies. The possibility to create an indestructible and automatically verifiable personal record gave a new meaning to individual autonomy online, but it also raised several ethical concerns. Among others, binding a personal identity to a single non-destructible digital record violates the 'right to be forgotten', which is also a part of the European General Data Protection Regulation. It contradicts the principle of 'purpose limitation', which states that personal data should be kept as long as it is required by the purpose of collecting it, but no longer (GDPR).

Privacy, in general, is the recurrent topic in ethical debates on data subjects and digital identities. Earlier implementations of blockchain such as Bitcoin and later many other cryptocurrencies sought to resolve the issue of privacy by anonymity. However, such cryptocurrencies as Bitcoin offer pseudonymity, at best, and their lack of identity protection has been uncovered in a rather short time (Reid and Harrigan 2012). This concern leads to the development of more privacy-oriented cryptocurrencies such as Monero and Zcash as secure alternatives to Bitcoin and Ether. However, despite sophisticated cryptography, some rather intuitive methods to identify owners of financial assets have been used by Wall Street traders "for decades if not centuries" (Yermack 2017, 18), and they remain relevant in case of blockchain, as the discovery of the Bitcoin wallet that allegedly belongs to Satoshi Nakamoto suggests (Voell 2020). Blockchain Developer Wikis, started to publish tutorials on how to secure sensitive data using a hybrid solution that stores the sensitive data in a centralized database and places a unique proof of all operations on the public blockchain. (see, e.g. Ardordocs[3])

The promise of anonymity backfired by the association of cryptocurrencies with criminal activities such as drug trade and money laundering (see Latimer and Duffy 2019 for current evaluation of financial risks related to cryptocurrencies). However, even communities that put anonymity first tend to operate under somehow authentic social personas, as the study of the actual 'darknet' bitcoin users by Bancroft and Scott Reid has shown. Sellers of illicit goods maintain stable pseudonymous identities that function in the same way as brands, in order to manage their reputation among buyers (Bancroft and Scott Reid 2017). Once again, privacy is challenged, re-constructed and re-negotiated, and new digital aspects of old identities are generated, confirmed and managed to enable social and financial interactions online.

The next concern, related to our proposal, is the question of consent. Human consent in information systems becomes a means "to mediate the expression of autonomy through technological applications" (Giannopoulou 2020) by well informed and self-determining subjects. However, retaining agency through consent in a technological society becomes dubious. How informed is informed consent in global information systems that are too complex to understand? Truly informed consent presumes the amount of responsibility and a cognitive load probably unbearable for a human being who is not a security engineer by occupation. As Langdon Winner summarizes in his writings on autonomous technology: "With the overload of information so monumental, possibilities once crucial to citizenship are neutralized" (Winner 2001, 296).

Furthermore, all data subjects involved in electronic communications leave 'digital traces' that can be scrapped without consent. In the best-case scenario, this data can be used to enable more comfortable coexistence of humans and non-humans in responsive smart environments, which should be the goal of digital identity projects. In the worst case, prevented from "forming or formulating a desire" (Rouvroy et al. 2013), a human agent is deprived of choice, purpose and opportunities for self-actualization, much like in a science fiction film The Matrix (1999). In reality, scrapping seemingly non-private data to use it for algorithmic decisions on personal safety have created controversies around AirBnB, among others, for banning marginalized but otherwise law-abiding users from the service (Dickson 2020; see also Jhaver et al. 2018 on coping behaviors of AirBnB hosts when 'negotiating' with opaque algorithms).

Another concern, especially in communication between human and non-human agents, arises when a digital identity is prioritized over the real human being when making an important decision. Consequences can be grave in case of an algorithmic decision about human matters. In critical information studies, a concept of a "digital" or a "statistical double" has been introduced, and the potentially repressive rule of algorithms has been described as "algorithmic dominance" (Giannopoulou 2020) or "algorithmic governmentality" (Rouvroy et al. 2013).

Artefacts always have ethical values encoded in them (Winner 1980). 'Platform ethics' of blockchains, enforced by 'smart contracts', can potentially magnify existing biases and power disbalances in electronic systems. Use of blockchain in the capacity of "a trust machine" does not guarantee fairness. In his discussion of potential applications of blockchains in corporate governance, David Yermack notes that "the regulations embedded in a blockchain's software code could favor some participating companies at the expense of others" (Yermack 2015, 27) and stresses the importance of possible human intervention. As an example of such intervention, he reminds of the DAO hack of the Ethereum platform in 2016: after the hack, 85% of Ethereum miners agreed to 'hard fork' the compromised platform and negated the consequences of illicit behavior.

[3] Ardordocs: https://ardordocs.jelurida.com/Securing_sensitive_data_with_the_blockchain

This example also shows that human judgement should always be prioritized, even – and especially – if not all human actors agree about the same priorities. The fully automated algorithmic consensus is the scenario that theoretically leads to 'domination' of robots over humans. If digital identities are prioritized over natural persons in the system, the consensus in it will likely be in favor of artificial intelligence, and it will probably exclude specifically human interests from consideration. Biometrics combined with blockchain and AI is yet another development of this same potentially harmful scenario: while blockchain is immutable and tamper-proof, a human body is not. A reasonable level of doubt should be guaranteed in every system that combines indestructible records with potentially flaccid biometric data.

In general, how concerned should we be with artificial agents? Ethical concerns about machine intelligence are often magnified with the existential fear of achieving 'singularity', the future event that will herald the total superiority of machine superintelligence over human capacities (Bostrom 2002). However, another vision appears to be more realistic - a cooperative vision of human and non-human entities who cooperate to reach common goals set by human actors and system designers (machines ultimately lack goal-setting abilities). In his review of original ideas of artificial intelligence by Alan Turing and J.C.R. Licklider, Oscar Schwartz shows how Turing's 'automotive vision' feeds the anxiety of "computers automating and replacing humans". In contrast, Licklider's "hybrid vision of AI" relies on human-machine collaboration that harnesses the power of machine intelligence (Schwartz 2018).

10. Conclusion

From a technical perspective: The digital agent signature combines self-sovereign identity (e.g.: digital qualified signatures) with verified blockchain wallets (E-ID wallets) and non-tradable utility tokens as a carrier medium for data and authorization to operate. This not only provides a complete record of interactions with and between digital-agents and their tasks, but it also ensures that you can see who has given the orders for the actions. All this information can be kept private, either in whole or in part, with the ability to assign shared-keys for access rights. Only the hash value that a transaction has taken place should be public or at least shared between the consortium.

From the ethnic perspective: Virtual worlds have taught us that ‘social types’ in the material worlds can be compared to “real-life avatars” that interact with other human and non-human actors. A game educator James Paul Gee has argued that a gamer’s self is a unified “sum and intersection” of online and offline identities and experiences (Gee 2015, 100). This understanding invites us to consider digital identities that are unified and plural at the same time, as a better fit for realistic, social and respectful implementations of identity management technologies. A cybernetic model of privacy in electronic networks should be re-evaluated to correspond to the social, intrinsically contextual way of practicing privacy in interaction with human and non-human agents. While technical limitations of ‘restricted access constitute the cybernetic model’, the social model is about ‘shared access’, a dynamic way to establish and negotiate boundaries and connections in virtual environments.

The authors suggest that understanding identity management for digital agents is similar to how we perceive avatars. As identities which are stable and consistent; they are attached to a single human person or a legal entity that can be identified on request. At the same time, this identity solution allows human agents to control which accounts and personal records to provide in interaction with non-human entities to achieve their goals in a private and secure, and yet transparent manner.

11. Further Research

The authors propose to consider both aspects together in future research projects. Technical progress should always be accompanied by a socio-political perspective. The authors would like to pursue this goal in further research projects. The proposals from this "Vision-Paper" should be put into practice and a prototype of an E-ID wallet should be created and tested and discussed in different use cases. Different interaction variations (man to man, man to machine, machine to machine) will be considered and the focus will be to discover possible fields of problems.

References

Cryptomarket Users. Information, Communication & Society 20, no. 4 (April 3): 497–512. https://doi.org/10.1080/1369118X.2016.1187643.

Gee, James Paul. (2014). Unified Discourse Analysis: Language, Reality, Virtual Worlds and Video Games. 1 edition. London ; New York: Routledge,

Bancroft, A. and Reid P. S. (2017). Challenging the Techno-Politics of Anonymity: The Case of

Blockchain-Identität für alle Einwohner (2017). Stadt Zug. https://www.stadtzug.ch/newsarchiv/383355

Bostrom, N. (2002). Existential Risks. Analyzing Human Extinction Scenarios and Related Hazards. Journal of Evolution and Technology, Vol. 9, No. 1.

Burden, D.; Savin-Baden, M. (2019) Virtual Humans, Today and Tomorrow, part of: Chapman & Hall/CRC, Artificial Intelligence and Robotics Series, CRC Press, Taylor & Francis Group, Boca Raton, USA

Dickson, EJ. (2020). Who's Allowed to Use Airbnb? Rolling Stone (January 8) https://www.rollingstone.com/culture/culture-news/airbnb-sex-worker-discrimination-935048/

Eixelsberger, W., Wundara, M. und Huemer, W. (2019). Blockchain in der Verwaltung. Handbuch E-Government. Springer Fachmedien Wiesbaden, pp. 506-518. https://doi.org/10.1007/978-3-658-21402-9_43

GDPR Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) http://data.europa.eu/eli/reg/2016/679/2016-05-04

Giannopoulou, A. (2020). Algorithmic Systems: The Consent Is in the Detail? Internet Policy Review 9, no. 1 (March 23). https://policyreview.info/articles/analysis/algorithmic-systems-consent-detail.

Goodell, G., and Tomaso A. (2019). A Decentralized Digital Identity Architecture. Frontiers in Blockchain. doi: https://doi.org/10.3389/fbloc.2019.00017.

Greene, K.K., Kelsey, J., Frankli, J.M. (2016) Measuring the usability and security of permuted passwords on mobile platforms. Technical report NISTIR 8040, National Institute of Standards and Technology, Information Access Division, Information Technology Laboratory, 100 Bureau Drive (Mail Stop 8940) Gaithersburg, MD 20899-8940

Huhns, M. N., Stephens L. M. (2000) Multiagent Systems and Societies of Agents in Gerhard Weiss, Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, MIT Press, USA

Jhaver, S., Karpfen, Y., & Antin, J. (2018). Algorithmic Anxiety and Coping Strategies of Airbnb Hosts. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 1–12. https://doi.org/10.1145/3173574.3173995

Jøsang, A. (2014). Identity management and trusted interaction in Internet and mobile computing. IET Information Security. 8. 67-79. 10.1049/iet-ifs.2012.0133.

Latimer, P. and Duffy, M.. (2019). Deconstructing Digital Currency and Its Risks: Why ASIC Must Rise to the Regulatory Challenge. Federal Law Review 47, no. 1 (March), pp. 121–50. https://doi.org/10.1177/0067205X18816237.

Marwick, A. E., and Boyd, D. (2014). Networked Privacy: How Teenagers Negotiate Context in Social Media. New Media & Society 16, no. 7 (November 1), pp. 1051–67. https://doi.org/10.1177/1461444814543995.

McDonough, J. P. (1999). Designer Selves: Construction of Technologically Mediated Identity within Graphical, Multiuser Virtual Environments. Journal of the American Society for Information Science, vol. 50, no. 10, pp. 855–69.

Monostori L., Váncza J., Kumara S.R.T., 2006, Agent-based systems for manufacturing, Annals of the CIRP, Vol. 55/2.

Nakamoto, S. (2008). Bitcoin: a peer-to-peer electronic cash system. Available at: https://bitcoin.org/bitcoin.pdf

Reid, Fergal, and Martin Harrigan (2012). An Analysis of Anonymity in the Bitcoin System. ArXiv:1107.4524 [Physics], May 7. http://arxiv.org/abs/1107.4524.

Rouvroy, A., Berns, T. and Libbrecht, E.. (2013). "Algorithmic Governmentality and Prospects of Emancipation." Réseaux No 177, no. 1 (October 14), pp. 163–96.

Oscar, S. "Competing Visions for AI." Digital Culture & Society 4, no. 1 (March 1, 2018): 87–106. https://doi.org/10.14361/dcs-2018-0107.

Tobin, A., and Reed, D. (2016). The Inevitable Rise of Self-Sovereign Identity. The Sovrin Foundation.

Voell, Z. (2020). 50 BTC Just Moved for First Time Since 2009 – But It Doesn't Look Like Satoshi. Coindesk (May 20). https://www.coindesk.com/50-btc-just-moved-for-first-time-since-2009-but-it-doesnt-look-like-satoshi

Winner, Langdon. (1978). Autonomous Technology. Technics-out-of-Control as a Theme in Political Thought. The MIT Press.

Winner, L. (1980). Do Artifacts Have Politics? Daedalus, Vol. 109, No. 1, Modern Technology: Problem or Opportunity? (Winter, 1980), pp. 121-136.

Wooldridge, M. (2000) Intelligent Agents in Gerhard Weiss, Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, MIT Press, USA

Wooldridge, M., Jennings, N. R. (1995) Intelligent agents: Theory and practice. The Knowledge Engineering Review, 10(2)

Yermack, D. (2017) Corporate Governance and Blockchains. Review of Finance 21, no. 1 (March 1), pp. 7–31. https://doi.org/10.1093/rof/rfw074.