

TOWARDS PARTICIPATORY DIGITAL SOCIETY USING THE DIGITAL TWIN PARADIGM: CITIZEN DATA MODEL

Aleksi Kopponen, Ministry of Finance, Finnish Government
Mariankatu 9, FI-00023, Helsinki, Finland,
Aleksi.Kopponen@vm.fi

Niko Ruostetsaari, Ministry of Finance, Finnish Government
Mariankatu 9, FI-00023, Helsinki, Finland, Niko.Ruostetsaari@vm.fi

Niko Mäkitalo, University of Helsinki, Department of Computer
Science

Pietari Kalminkatu 5, FI-00014, Helsinki, Finland,
Niko.Makitalo@helsinki.fi

Tommi Mikkonen, University of Helsinki, Department of Computer
Science

Pietari Kalmin katu 5, FI-00014, Helsinki, Finland,
Tommi.Mikkonen@helsinki.fi

Abstract

Historically, societal challenges have been handled as complicated ones by dividing responsibilities into various organizations that are then mandated to provide needed solutions and services from their own perspective. It has also defined the enterprise architecture that has had a tendency to mimic organizational structures and its responsibilities. Today, modern societies are facing complex challenges, such as climate change, that can't be solved without a paradigm shift of problem-solving. It also challenges information systems capabilities to co-operate in a cross-sectoral manner. Advances in ICT and its adaptation and penetration rate in every part of society introduce new possibilities that can only be benefitted in full by providing more integrated information systems. Moreover, the level of citizens' expectation has raised due to other systems that liberally connect data from different sources. To meet new goals, societal services should become more proactive and personalized, satisfying citizens' needs as they emerge. In this paper, we propose using a holistic model of digital twin paradigm for societal applications. The proposal builds on using a citizen 360-data model that reflects the characteristics of citizens that act as service users. *Based on the*

data model, societal information systems can propose actions and provide proactive services that are mass-tailored to meet individuals' needs.

Digital twin Public administration Human-centricity Digital society Mydata.

1. Introduction

In the coming years and decades, many societies will face major challenges such as ageing populations and high rates of structural unemployment. These challenges linked with the rapid technological development and global threats like climate change may be transforming the world faster than ever before. The so-called 'wicked', deeply intertwined problems cannot be resolved without more systematic efforts to harness the opportunities offered by technological advances and digitalisation to promote citizens' welfare.

To solve such problems, public and societal services need improved capabilities as well as new ways of thinking about citizens, their needs, and the role of society. Historically, challenges have been solved by dividing responsibilities into various organizations that are then mandated to provide needed services from their own perspective (Bourgon, 2017). For example, tax administration provides their own services to collect taxes. This paradigm has been very effective in organizing services from organizational perspective and service design has increased it further on. However, we have come to a situation where services do not form a clear picture from the citizens point of view. Also, legislation does not support co-developing seamless service paths with various organizations from public, private and third sector for the benefit of people's well-being.

To meet new goals, societal services should become more proactive and personalized, satisfying citizens' needs as they emerge. At the same time, associated information systems should be managed more effectively and proactively, which in turn offers organizations better opportunities to regenerate their operations, deliver savings and improve the availability, quality and customer experience with better matching of their services just-in-time with citizens real need.

In this paper, we propose applying the digital twin paradigm (Alam and El Saddik, 2017) for societal applications. The proposal builds on using a citizen data model that reflects the characteristics of citizens that act as service users. Based on the data model, societal information systems can propose actions and provide services that are mass-tailored to meet

individuals' needs. In addition, we introduce a case study where the approach and its core concepts have been evolving.

The rest of this paper is structured as follows. In Section 2, we provide theoretical background for this work, covering the new possibilities of increasing societal data reserves. In Section 3, which forms the core, we introduce the citizen data model. In Section 4, we present a case study that demonstrates the use of the data model. In Section 5, we provide an extended discussion regarding our observations and lessons learned. Towards the end of the paper, in Section 6, we draw some final conclusions.

2. Theoretical background

The use of data and novel technological solutions provide new tools for building a human-centric and participatory society by respecting people's own basic rights to their own data. Especially use of the digital twin paradigm can function as a major transformative force in order to allow citizens to utilize their personal data in a way that enables tailor-made provision of different services. Services that are based on personal data provide users with better understanding and control over their welfare and the ability to activate services in real time. The key feature is that the user is able to manage their own data and edit them into situational, temporary profiles in order to access a personalised service offering. That is to say, the key feature is that the user is able to manage their own data and edit them into situational, temporary profiles in order to access a personalised service offering in a cross-sectoral manner.

Digital Twin is originally a concept where the aim is to create an "exact" cyber copy of a system to the Digital World (Alam and El Saddik, 2017). While initially, Digital Twin concept was developed for improving manufacturing systems, today, Digital Twins are being re-defined as digital copies of living and non-living entities (El Saddik, 2018). Creating such copies, however, requires much information from the entity and its environment. Such demand is typically addressed with Cyber-Physical Systems (CPS) and Cyber-Physical-Social Systems (CPSS). Cyber-Physical Computing leverages data from physical "real" world and cyber world processes, and send this to the Digital Twin for computation (Alam and El Saddik, 2017). Physical-Cyber-Social (PCS) Computing extends these systems by adding a third dimension to this equation that is inherent for humans: the social world. PCS Computing considers the social world

processes and relationships while encompassing data, information, and knowledge coming from these three worlds, and then to integrate, correlate, interpret, and provide human-understandable abstractions that are contextually meaningful (Sheth, Anantharam, and Henson, 2013). According to the authors, such a process can be used for improving the human experience in computing (Sheth, 2009), which is also a goal of the Digital Twins. Sheth, Anantharam, and Henson initially presented PCS Computing (2013), but others have presented similar ideas that are often referred as Cyber-Physical-Social Systems (CPSS) as shown by the survey from Zeng et al. (2016). Some of the oldest CPSS references are from 2010 to 2011 (Wang, 2010; Liu et al., 2011).

With a proper Citizen Data Model, societies can enormously increase their understanding of people's real situation and what services are really needed and by whom. It can lead to a societies where services can be proactively orchestrated around people's real need. Citizen Data Model also helps organisations to decrease failure demand of their services by providing services to citizens only when they really bring value (Kjøller and Rasmussen, 2017).

However, implementing a whole and holistic Citizen Data Model might be too difficult to be done at once. Also, data challenge shouldn't be considered as an easy one to tackle. Even though one can identify the needed elements for data, numerous challenges occur when building a model in practice. Many of these challenges are related to data usability as seen in the studies done by Lacity and Willcocks (2018). Also, the use of data is restricted by legislation such as General Data Protection Regulation (GDPR). The other challenges are related to leadership, investments, know-how of organisations and citizens, technological capabilities, digital trust and information policies in both public and private sectors.

3. Digital Twins for Societal Use: Citizen Data Model

In a human-centric society, Digital Twin concept needs to be applied with the citizens, as well. For example, in respect to the new demands the technological development poses to the working life, people continuously need new skills during their careers and to stay employed. Despite the fact that there is a range of digital learning and employment services available, it is usually difficult for citizen themselves to identify suitable services from

the vast selection. In addition, advice on continuous learning and personal career paths is limited and usually too expensive for the society. Smart and co-operative digital services and related service chains can help users to analyse their competencies and possibilities for competence development and employment.

As a part of the preliminary study of the National AuroraAI Programme in Finland (<https://vm.fi/en/auroraai-en>), a trial was concluded in which information was collected and produced about various alternatives for using AI in the context of lifelong learning and employment, taking into account responsibility, data privacy and ethical questions. The solution produced in the trial maps the user's competencies and compares them with existing educational and job offerings. Thus, Digital Twin can be then utilised to improve the effectiveness of the services and decrease utilising services that really doesn't add value for the citizen.

The Digital Twin concept can be developed in cases where the connection between the real-world person and their digital persona needs to be made invisible. The user collates their personal data to produce a compilation or summary that can be processed by the network without connecting it to the user's source data. The development and controlled testing of such a concept can be seen as important methods in the transition towards a human-centric society in which users can trust data privacy.

The intelligent and dynamically evolving service chain mines and matches profile data, future data, foresight data and information about educational and job market offerings to help user to analyse competencies and job opportunities. In the future, the solution should lead to getting information about competence development needs and suitable opportunities in order to address skill gaps. Expert advice and guidance would be available whenever needed.

In the long term, the solution and the personalised service offering it facilitates could help resolve the skills-matching problem as competence needs to evolve. The greatest benefit would come from continuous use, which would give the tool continuous feedback about the user's choices and personal development, and enable it to learn and produce increasingly accurate suggestions about suitable job openings and competence development opportunities. This however implies that there is personal presence online, even when the citizens themselves are offline.

In addition, efficient use of Citizen Data Model requires a better understanding of who our citizens really are and what they really need in order to better focus the public impact on people's well-being. Understanding the well-being of the citizens as a multidimensional, emergent issue is overlooked by conventional income measurements. This is why the report by the Stiglitz Commission on the Measurement of Economic Performance and Social Progress highlighted the importance of going beyond GDP or production and tackling the more difficult task of measuring people's well-being (Stiglitz et al., 2009). Figure 1 addresses the multidimensional definition of people's well-being that is based on initiatives around the world and academic research. The Stiglitz Commission identified eight key dimensions listed on the Figure 1 that should be taken into account simultaneously when measuring people's well-being. In addition, well-being indicators to indicate well-being at a societal level have been introduced by OECD (2011). These together form a baseline to build a Digital Twin of a citizen as multidimensional taking indicators of sustainability into notice, too. As stated in the Stiglitz Commission report, statistical indicators are important for designing and assessing policies since what societies measure affect what societies really do. Hence, the paradigm shift from measuring production to measuring the real multidimensional well-being is paramount for advancing the progress of society.

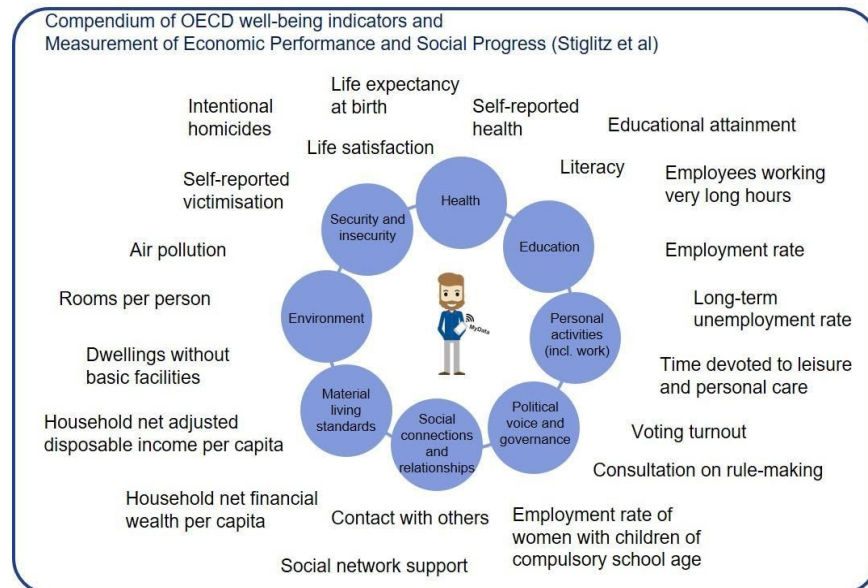


Figure 1. Citizen Data Model is based on various indicators of citizen multidimensional well-being proposed in compendium of OECD well-being indicators and Measurement of Economic Performance and Social Progress (Stiglitz et al., 2009).

To summarize, Citizen Data Model should be built through chosen business cases. From the citizen's point of view, business case is considered as a life event, where numerous service needs occurs at the same event. In order to build a proper Citizen Data Model for chosen life event, design approach needs to be applied and work should be started with building a situational analysis of people's well-being in chosen life event.

Situational analysis process consists three phases, as shown in Figure 2:

1. Choosing a life event and identifying its key stakeholders that have valuable information about people in a chosen life event
2. Collecting information and creating a Citizen Data Model in chosen life event in a cross-sectoral co-operation. Citizen Data Model requires collecting information around all well-being indicators at a local and a national level.
3. Co-analysing results and start strengthening well-being of people in chosen clusters

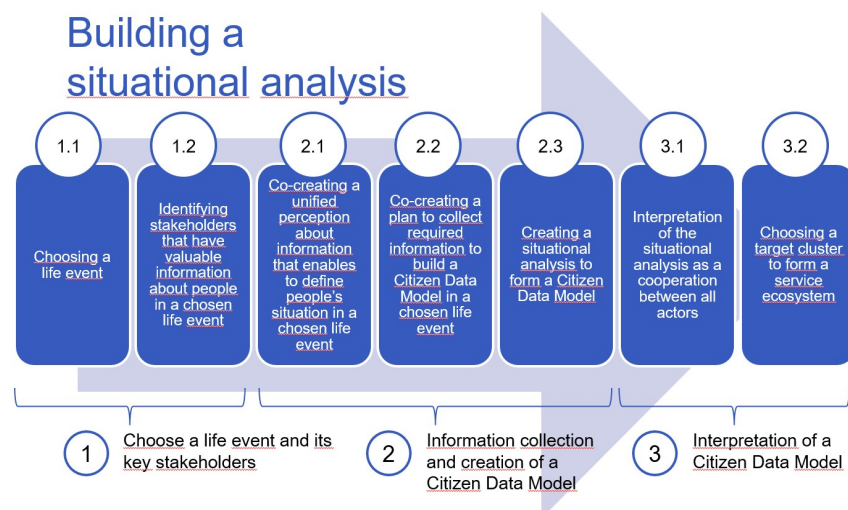


Figure 2. Building a Citizen Data Model through situational analysis.

In order to succeed building a situational analysis, it is paramount to gather experts around every organisation that are involved with chosen life event into a public-private-people-partnership co-operation. Participants together form a shared understanding of what really happens in chosen life event and what data is needed to understand it better. Data can also be gathered by interviewing people experiencing chosen life event. As a result, situational analysis visualises the differences of people's multidimensional well-being in chosen life event, as illustrated in Figure 3, which presents an example of a situational analysis in the City of Turku in Finland and cluster analysis was done by GoforeOyj. As a part of the preliminary study on the AuroraAI national artificial programme, City of Turku investigated the well-being of students of universities and universities of applied sciences that were starting their studies in the City area. In order to measure the multidimensional well-being using the framework of Stiglitz Commission report, the City concluded a comprehensive survey on the target group focusing on the key dimensions of well-being addressed by the Commission report. Based on a cluster analysis implemented on the survey data, City of Turku created the situational analysis on the well-being of the target group which enables a better understanding of who the City residents really are and what they really need - in an anonymous manner.

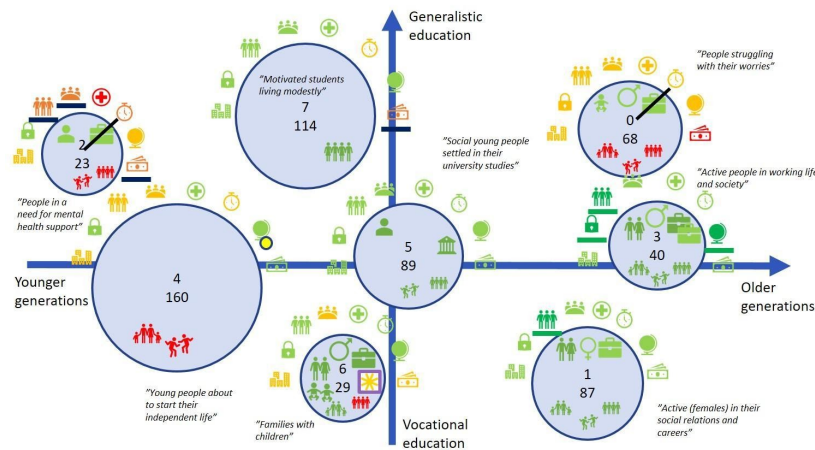


Figure 3. Machine learning can be applied to deepen the understanding of people's well-being in chosen life events. Cluster analysis can be done by utilising Citizen Data Model (Cluster analysis was done by GoforeOyj).

4. Case Study

In today's labor market, people continuously need new skills during their careers. To meet such demand, there is a range of digital learning and employment services available. However, for an individual, it can be difficult to identify suitable services from the vast selection. Furthermore, in many cases advice on continuous learning and personal career paths is limited. To understand the challenges of the field, we have carried out a trial on gaining a foothold in working life through competence development.

In the trial, information was collected and produced about various alternatives for using AI in the context of lifelong learning and employment, taking into account responsibility and ethical questions. The trial was participated by numerous partners, including representatives from the National Agency for Education, the Ministry of Education and Culture, the Ministry of Employment and the Economy, and the following enterprises: Osaamisbotti, Headai, Accenture, Fjord and SuomenTilaajavastuu Oy (Kopponen and Ruostetsaari, 2019, pp. 38-42). In addition, development was assisted by students, learning and employment experts from various sectors, and the AuroraAI developer community. On average, 50 people participated

in the workshops and development work. An average of 15 test students were involved in the development process at all times.

Test setup was technically made in co-operation of three private organization. These organizations were SuomenTilaajavastuu Oy, who's responsibilities involved authorizing use of mydata. Osaamisbotti Oy was responsible of skill mapping and continuous support. The third organization was Headai that was responsible for analysing skills and making job suggestions for citizens. The pilot used OsaamisbottiOy's general mapping tool, which was supplemented by a few more sector specific mapping. Here, the selected sample areas were construction and digital skills.

SuomenTilaajavastuuOy'sMyData Wallet enables to utilise citizen's MyData to personalise and optimise services based on people's real profiles and needs. With MyData Wallet, citizens are able to control their own data flows and decide in which services MyData is utilized. In this test setup, citizen wants to map his/her skills (Figure 4) with labour markets by giving permission for Headai service to utilise MyData from his/her MyData Wallet. To achieve this, a citizen triggers an information request where Headai requests to use her/his skill map found from MyData Wallet. Citizen is able to see the exact purpose of utilisation of MyData and where it is going to be utilised. In order to proceed, citizen gives the permission to utilise MyData. Citizen is able to remove permission anytime she/he wants. If so, Headai can't access her/his skill map anymore. MyData Wallet also gives an overview for citizens to understand all the given permissions for every services and purposes her/his MyData is utilised.

AuroraAI Osaamisbotti

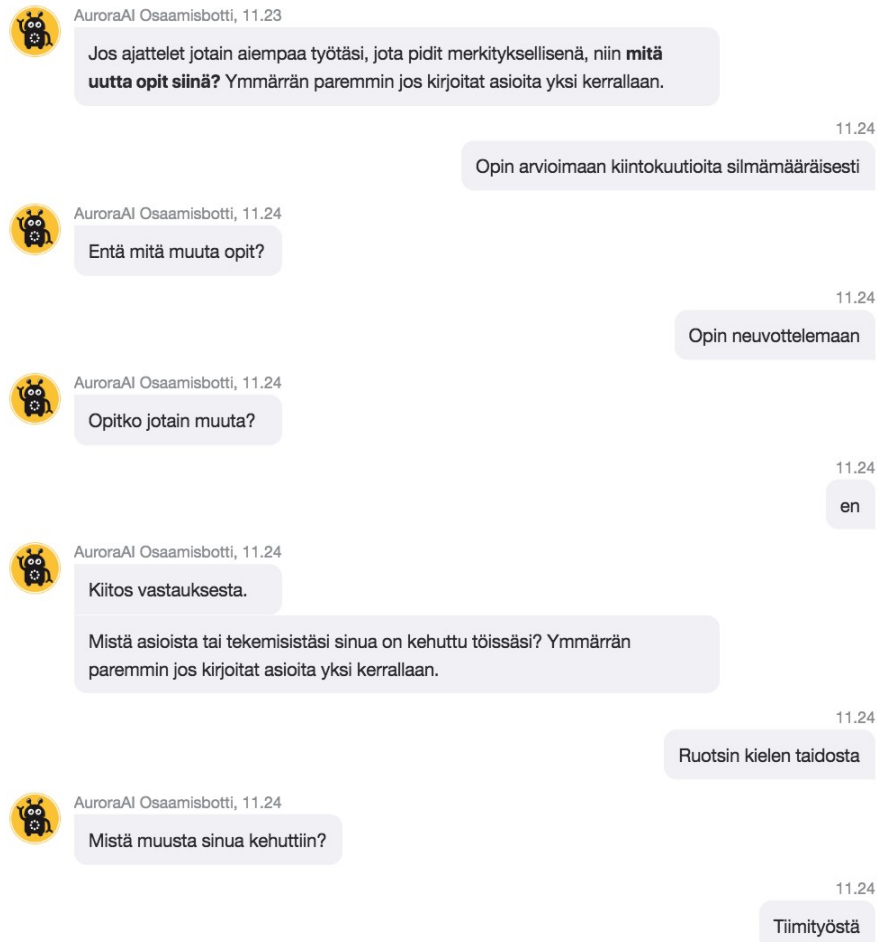


Figure 4. The survey asked about skills, interests, goals and need for support (available in Finnish only). Dialogue is characterized by dynamic interaction, i.e. the number of questions and answers depends on how the discussion progresses. Chatbot was made by Osaamisbotti Oy.

After permission in gutilisation of MyData, citizen skill profile is to be done. That rises two challenges (Kopponen and Ruostetsaari, 2019, pp. 38-42): Firstly, the expression of knowledge/skills is very difficult, and, each citizen

has a unique skill profile. In other words; there are no two identical citizen skill profiles. Secondly, young people especially can find it difficult to articulate their skills. To tackle these issues, a conversational chatbot technology was utilised as a personal interface to map citizen skills, described in detail in (Lacity and Willcocks, 2018). In test setup, students were discussing 10 minutes with a chatbot, that built a customised skill profile based on discussions for each students, who were then able to utilize it in services found from AuroraAI network. One key finding was to see that citizen's should not be left alone just with a chatbot. It is important to keep conversation open also with real student counselor or career coach. This key finding is supported with other studies of utilising robotic process automation (RPA) and cognitive automation (CA) tools. Those studies suggest to invite customers to try the automation, but keep other channels open, even though major part of discussions are with chatbot.

Rakennusala		tammikuu 15. 2019
Kysymys	Käyttäjän vastaus	
Hei Kartoitamme osaamistasi ja kiinnostusta rakennusalalla. Kerrotko aluksi kuinka monta vuotta sinulla on alan työkokemusta?	4	
Mikä rakentamisen osa-alue kuvaa parhaiten osaamistasi tai kiinnostusta?	Pintakäsittely	
Onko sinulla jotakin muuta rakennusalan osaamista?	Maalaus	
Onko sinulla jokin näistä korteista tai sertifikaateista?	Tulityökortti	
Onko sinulla voimassaolevaa ensiapukoulutusta?	Ei ensiapukoulutusta	
Millaisista töistä olet kiinnostunut?	Viimeistelytyöt	

Figure 5. Customised tool for mapping know-how in construction industry (available in Finnish only).

Completed skill profile includes hidden skills, interests, motivation, and even her/his dreams. A citizen can utilise her/his profile as a DigiMe in

AuroraAI network. In this test setup, DigiMewas transferred to Headai service to make natural language based analysis to understand what skill related words is found from it. Headai also creates a digital twin from labor market. Thus, Headai service links DigiMe with digital twin of labour market to show the interdependencies of citizen skills with real needs in the labor market and what kind of vacancies and opportunities one can find for her/himself. Interdependencies are done with analysing meanings of discussions with a citizen, not by just finding and matching keywords. It is also important to visualise matching skills for citizens, whether or not they are hidden or proven, since citizen her/himself might not actually understand her/his skills that might be relevant in the labour market. Also, any lack of skills or potential to improve on some areas should be visualised in order to empower people to improve their skill set.

MyData Wallet

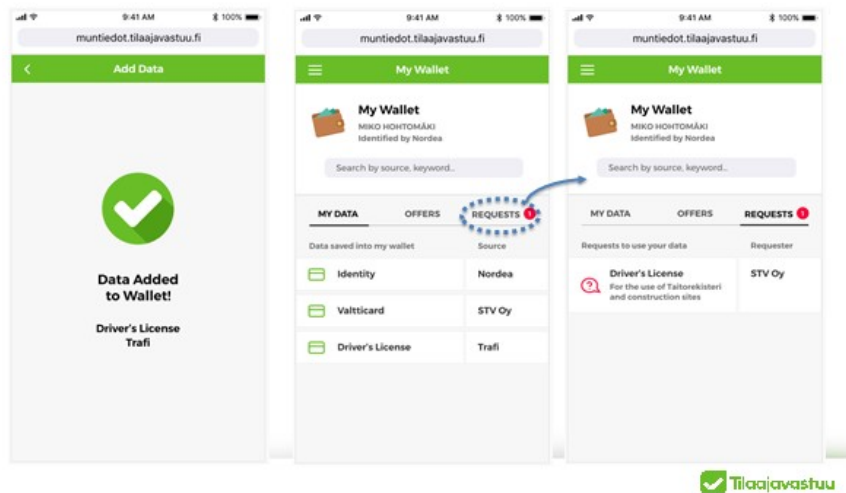


Figure 6. Pilot carried out an experiment with DigiMe - identity management and authentication (with personal consent to use data). MyData wallet was provided by SuomenTilaajavastuu Oy.

The differences between a keyword search and using machine learning in reasoning is enormous. A keyword search do not account any smart understanding of what a citizen really means or needs. A machine learning in

reasoning, on the other hand, enables to build an understanding of intentions and who citizens are and what they really need. In this set up, Headai used AI to create a maximum labour market value for citizens instead of just matching open vacancies with skills. In order to do that, labour market data is deeply analysed to create an understanding of what skills are really needed, not just by looking after job titles. By comparing previous year with this year and how skill searches have changed, and combining that with educational information and curriculum, Headai provides an analysis of what we really are training in our society and does it fit with real need in the labour markets. It also reveals hot spots of lacking skills in the labour market. Visualising mismatches also enables decision makers build better understanding of what skills citizens should build for themselves. Furthermore, citizens different situations can be analysed through segmentation or cluster analysis.

5. Discussion

Obviously, there are numerous technical and operational issues that should also be considered. Any large operation that involves numerous stakeholders can create complications regarding the stakeholders' views and responsibilities, and our needs are by no means different. The stance taken by us in the trial is that by committing to the platform's fundamental compatibility requirements, it is possible to integrate just about any system in the system.

The societal dimension of the work calls for certain special characteristics, such as accountability, which imply that the requirements also include policy and operations related issues. To govern these, a trust council has been established, and several contractual aspects are considered. Furthermore, all the associated services are subject to direct monitoring as well as to applications such as Saidot.ai that offer transparency and other desirable functions for various stakeholders, including also individuals.

A further technical and operational dimension to be considered is Intellectual Property Rights (IPR) and associated responsibilities and opportunities for further monetization of existing implementations in other contexts. With the increased interest in platform economy, a key challenge is which parts of the system being designed are scalable enough for generalization towards society as a service, and how should such a system be governed. Obviously,

this may also mean reconsidering the role of the government in general when considering its role as a societal stakeholder in ICT.

6. Conclusions

In our quest to build a human-centric society, where information technology is used proactively to help citizens' daily activities, we have only taken very first steps. In the trial presented in this paper, we have so far demonstrated the technical capabilities as well as the feasibility of public-private-people-partnership as the underlying model for the development. Furthermore, it has also become apparent that while politics change, information technology is so deeply embedded in the Finnish society that initiatives do not directly follow political guidance.

Specific conclusion from the participants of the test setup is that the emergence of the ecosystem cannot be built without cross-sectoral co-operation. Ecosystem must be developed and led through dialogue, and solutions must be formulated together. Also, the key to solving difficult challenges is to identify opportunities to work together instead of concentrating on obstacles, and to provide developers the needed ability, motivation and enthusiasm to succeed in given tasks. Co-development and acquisition models need to be reformed to allow for designing solutions and leveraging the best capabilities from various organisations. All in all, test setup has shown that seamless service chains implemented in a co-operation by various digital service providers bring the greatest added value to the individual. What is also needed is a collaborative architecture that benefits both the individual and other ecosystem actors, as well as authorising public sector data for the DigiMe.

Despite the fact that the trial version has been successfully deployed, there are numerous identified challenges that will need special attention. While it is almost universally agreed that the state and municipalities should support its citizens, aiming to implement this ideal with computers and software provokes reactions. While building understanding regarding citizens' needs, there is a fine line between empowering and controlling the individuals, and with the trial version, we have only scratched the surface of what is acceptable and what is not. Furthermore, this line is different for different individuals, not to mention different nations. Additional challenges emerge when considering regulation and ethical issues in more general as they evolve. What is now possible may not be acceptable in the future, and vice versa. Hence, although individual services, key design decisions, and

learnings of AuroraAI are reusable in other countries, detailed data sets or behavioral patterns are probably not.

7. References

1. Alam, K. M., & El Saddik, A. (2017). C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access*, 5, pp. 2050-2062.
2. Ala-Pietilä, P., Bauer, W., Bergmann, U., Bielikova, M. et al (2019). Policy and investment recommendations for trustworthy Artificial Intelligence. Independent High-Level Expert Group on Artificial Intelligence, set up by the European Commission.
3. Bourgon, J. (2017). The New Synthesis of Public Administration Fieldbook. Dansk psykologiskforlag A/S. pp. 31-32
4. El Saddik, A. (2018). Digital twins: The convergence of multimedia technologies. *IEEE MultiMedia*, 25(2), pp. 87-92.
6. Kjøller, A. and Rasmussen, N. (2017). The hidden potential of eliminating failure demand. Implement Consulting Group.
7. Kopponen, A. and Ruostetsaari, N., (Ed.) (2019). AuroraAI – Towards a people-oriented society, Development and implementation plan 2019–2023 based on the preliminary study on the Aurora national artificial intelligence programme. Ministry of Finance: <https://vm.fi/en/auroraai-en> (23.12.2019)
8. Lacity, M. and Willcocks, L., (2018). Robotic Process and Cognitive Automation: The Next Phase. SB Publishing.
9. Liu, Z., Yang, D. S., Wen, D., Zhang, W. M., & Mao, W. (2011). Cyber-physical-social systems for command and control. *IEEE Intelligent Systems*, 26(4), pp. 92-96.
10. OECD (2011), Compendium of OECD well-being indicators. OECD: www.oecd.org (23.12.2019)
11. Sheth, A., Anantharam, P., and Henson, C. (2013). Physical-cyber-social computing: An early 21st century approach. *IEEE Intelligent Systems*, (1), pp. 78-82.
12. Sheth, A. (2009). Computing for human experience: Semantics-empowered sensors, services, and social computing on the ubiquitous web. *IEEE Internet Computing*, 14(1), pp. 88-91.
13. Stiglitz, J. E., Sen, A., and Fitoussi, J. P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress.

14. Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, 94(9-12), pp. 3563-3576.
15. Zeng, J., Yang, L. T., Lin, M., Ning, H., & Ma, J. (2016). A survey: Cyber-physical-social systems and their system-level design methodology. *Future Generation Computer Systems*.
16. Wang, F. Y. (2010). The emergence of intelligent enterprises: From CPS to CPSS. *IEEE Intelligent Systems*, 25(4), pp. 85-88.