Masso, A., Kasapoglu, T., Tammpuu, P., & Calzada, I. (2024). Datafied Control and Selection of Digital Identity: Estonian e-Residency as 'Citizenship by Connection', DOI: https://zenodo.org/doi/10.5281/zenodo.12666718

Datafied Control and Selection of Digital Identity: Estonian e-Residency as 'Citizenship by Connection'

Anu Masso [1] [2] [3]
Tayfun Kasapoglu [1]
Piia Tammpuu [4]
Igor Calzada [5] [6]

Abstract

This article contributes to the discussions on the datafied control, selection, and (re)construction of digital borders globally. We examine the reasons for the prevailing datafied inclusion and exclusion patterns revealed by the Estonian e-residency program. This government-supported digital identity program gives non-residents remote access to Estonian e-services and business environments. Relying on the in-depth interviews conducted among the experts responsible for planning and implementing the e-residency program (n=8) and e-residents (n=25), we examined the understandings and practices on the datafied control of the e-residency applicants and the active e-residents. The findings reveal that despite the program's global spread, the applicants' datafied control and selection may legitimise and reproduce global inequalities through (re)constructed digital borders. However, data connect individuals striving for a placeless lifestyle through their digital transactions resulting in an emerging regime that this article coins as 'citizenship by connection'. Thus, datafied control constructs not only reality and borders as connected with one's territory but also creates digital borders through new practices of such 'citizenship by connection' emerging regime.

Keywords:

E-residency, Estonia, digital identification, social datafication, borders, citizenship.

Highlights

[1] Ragnar Nurkse Department of Innovation and Governance, Tallinn University of Technology
[2] Smart City Centre of Excellence (Finest Twins), Tallinn University of Technology
[3] 2022 Global Digital Governance Fellow at Stanford University
[4] Institute of Social Studies, University of Tartu
[5] Wales Institute of Social and Economic Research and Data (WISERD), Civil Society ESRC Centre, Social Science Research Park (SPARK), Cardiff University
[6] Fulbright Scholar-In-Residence (S-I-R), US-UK Fulbright Commission, California State University

- The globally spread E-residency program reconstructs the state-individual relationship.
- Data connects individuals with a placeless lifestyle through their digital transactions.
- The desire for placeless work conflicts with the datafied digital migration control.
- The datafied selection of E-residents may legitimise the global digital borders.
- Data does not create national borders but global ‘citizenship by connection’.

1. Introduction

Societies have been experiencing an increasing application of digital platforms, enabling global access to digital identity and a location-independent and placeless lifestyle. However, there is limited knowledge that focuses on the critical aspects of these technologies and factors that hinder their effective implementation. This article aims to contribute to these discussions by examining how the data connects and separates individuals striving for a placeless lifestyle through their digital transactions and how it results in an emerging algorithmic regime that this article coins as ‘citizenship by connection’.

This article focuses on the Estonian e-Residency program as a specific and paradigmatic digital identification case study. The Estonian e-Residency, as a global digital identification system (hereinafter Digi-ID), exemplifies the evidence where the cross-border data flows are necessary to enable the practices of the e-residency through online transactions, on the one hand, and to trace and control these activities to ensure the security of the program, on the other hand. As such, the Estonian e-residency case enables us to introduce and explain the pattern in which multiple individuals desire freedom of mobility and location-independent work (Barlow, 1996) invariably conflicts with the restrictive datafied control. This pattern is especially visible in digital identification systems spread globally, where cross-border data flows are not always operationally and technically possible and not practised due to the restrictions and regulations of data exchange not directly stemming from the E-residency program but global regulations and trends. However, there are no studies about how these new datafied governance tools and approaches are seen and understood by diverse groups applying these control mechanisms or those who are the targets of these digital identity programs.

Hence, this article aims to understand how the digital identification program and its relevant data collection and analysis practices are understood from the perspective of e-residents and the experts who have designed this program. Besides, the peculiarity of e-residency (and one of the contradictions) boils down to the fact that e-residency is limited to the electronic environment and is clearly distinguished from physical migration and the provision of physical mobility opportunities. As such, it can be imagined as borderless, although at the same time, data-based bordering practices or logics/rationalities also work in it.

Therefore, it is necessary to expand the scope of empirical research and include examples where digital identification technologies are not only used to govern physical migration and physical mobility but also virtual mobility, making visible the digital borders and bordering practices that exist within the promised borderless world.

This article applies a qualitative approach using in-depth interviews with two groups: experts that develop and implement Estonia's e-residency program and e-residents from different nationalities and regions with different motives for applying for an e-resident digi-ID. This article strives to answer three research questions: (1) How and why the connections and disconnections of an Estonian e-residency program are seen and constructed through the data? (2) How are the global data flows used and seen to ensure the proper digital services and the practising and controlling the placeless lifestyle for the e-residents? (3) How is the datafied control practised and understood, based on the positions of experts and e-residents, for deciding who is 'risky' and who is ‘safe’ to allow to join the program?

2. Digital identification and Estonian E-residency

Several digital identification schemes have been developed and implemented nationally, such as Aadhaar in India and biometric ID in Tunisia. Large differences exist in the functions and implementations of these digital identification systems globally.

This study focuses on the Estonian E-residency program, a digital identification scheme and an e-services platform introduced by the Estonian government in 2014, which allows foreign nationals to apply for a state-issued digital ID to gain remote access to the digital infrastructure and e-services provided by the Estonian public and private sector (Author 4). The Estonian e-residency program, launched in 2014, provides a distinct case to explore digital identification systems as it was one of the first globally applicable digital identification schemes. Estonian e-residency has been established on Estonia’s well-functioning national digital identification system and its underlying digital infrastructure. However, unlike Estonian digi-ID, delimited to Estonian citizens and (physical) residents, e-residency targets non-residents. In this respect, Estonian e-residency is the first government-supported digital identity scheme which does not delimit its users territorially (i.e., residents of a specific country) or politically (i.e., citizens of a specific country). Also, the e-residency program is enabled through a data exchange layer called X-Road. It enables government agencies to gather citizens’ data just once and securely exchange them among agencies instead of requesting them from citizens often.

Digital identification schemes like Estonian e-Residency are advertised and presented by authorities and technology companies as global and borderless (e.g., Author 3). However, these systems may not be implemented and understood as such by the users. An empirical study (Author 4) showed that digital identification programs like Estonia's e-residency are still more accessible to applicants from countries with higher e-government and economic development levels. The digital borders, thus, could also be re-imagined, condensed, and resolved (Author 7; Author 6). Although research has emphasised (Chouliaraki & Georgiou, 2022; Amoore, 2021) that digital identification systems are one of the most eminent fields where digital data traces are produced and used, we do not know how the borders are imagined and reconstructed in the context where the practices of digital identity systems are increasingly datafied.

3. Datafied control of digital identity systems

Digital identification systems use various kinds of data in (re)constructing digital borders and resulting in datafied control (Author 5; Amoore, 2021). Datafied selection and control is the tendency to make decisions about people based on data, creating complex regimes of exclusion/inclusion with the promise or illusion of neutrality and objectivity (Author 2). A wide variety of studies have explored and critically tested the use of vast databases like the Visa Information System (including also sensitive biometric data) in Europe to control irregular migration by identifying and sorting legal and irregular migrants (Sontowski, 2018; Broeders, 2016) or implementing immediate digital controls on the border over refugee movements and identity (Latonero & Kift, 2018). Therefore, the digital infrastructure for movement is seen as a sociotechnical space of flows where regular and irregular migrants, corporations and governments interact with each other and new technologies. Still, it can also easily be leveraged for surveillance and control.

Extraterritorial geographies of data flows (Mann & Daly, 2020) have significant implications for how states exercise coercive power beyond their territorial borders. The social geopolitical hierarchies of territory and data are tightly intertwined (Couldry & Mejias, 2019) as the inevitable and inherent to the rationale of capitalism. Studies on the digitisation and datafication of migration management have mainly focused on border security, where the gathering, processing, and sharing of data facilitates the practice of traceability and rationality of mobility control. Since it restricts individuals' capacity to move and resist control, such bordering often forms the identities of the individuals they target through securitisation. Therefore, in this article, we understand the bordering through data as a geographically scaleless, technopolitical conceptual assemblage that allows the operationalisation of cross-

bordering data flows beyond digital global orders and regulations. We assume that digital identification systems and interrelated datafied practices can collapse our traditional thinking about physical borders (Yanqing, 2022) and offer new opportunities and challenges.

One of the increasing ways in which datafied selection, control and bordering are practised is predictive analytics, including both implementations of automated systems in punitive policies of an already existing physical cross-border service regime (Park & Humphry, 2019) or digital environments like predicting cybercrime through controlling transnational and transactional data flows (Kennedy, 2020). Unlike other datafied control procedures, cross-border predictive policing is often characterised by complex 'citizen scoring' (Dencik et al., 2019). On such occasions, data analytics is used by governments not only for purposes of categorisation, assessment, and prediction on the individual level and in single border crossing occasions but also on the population level for developing intersectional profiles of targeted and most (un)desired migrants. Other studies have highlighted the increasing power relations that result from social categorisation through predictive analytics and the role of active individuals in resisting these emerging social hierarchies (Couldry & Mejias, 2019). Some studies have even called this practising and participating in creating the data traces or resisting the use of data a form of 'data citizenship' (Carmi et al., 2020) or 'algorithmic citizenship' (Author 1). Whereas these concepts are introduced in the literature, the understandings and practices of the digital identity holders as datafied citizens are seldom empirically studied.

Therefore, research has highlighted the need to better understand the positions of migrant users as active and 'connected' participants or as subject to biometric datafication and surveillance (Nedelcu & Soysüren, 2020; Ponzanesi, 2019). Alternatively, studies have highlighted the significant mediating role of the labour of experts and maintainers, who, by rendering local or global information systems functional, sustain the power to control migration through digital means (Author 9; Glouftsios, 2020). Although these prior studies have suggested the importance of better understanding the citizens' positions, we still do not know how the related parties of digital identity systems understand the role of data in practising and imagining the borders constructed through the datafication of these systems.

3. Methods and data

3.1. Purposeful sample strategy

To explore the (re)construction of digital borders through data and digital identification systems, this study relies on the empirical data which was collected through semi-structured interviews conducted with two groups of interviewees – Estonian e-residents who have been

issued an e-resident digi-ID by the Estonian state and experts who have been involved in planning and implementation of the e-residency project.

First, the interviews were conducted with Estonian e-residents. We compiled the sample of interviewees based on the purposeful sampling strategy (Suri, 2011). Based on the purposeful strategy, our sample included those issued an e-resident digi-ID card. To ensure the variability in the interview responses, we designed the sample to include participants from a diverse set of nationalities and geographic regions (including both EU nationals as well as third-country nationals) and with different personal motives (e.g., instrumental, business-related motives *vs* non-instrumental motives and interests) for applying for an e-resident digi-ID. The final sample included mainly younger age groups (<50) and male applicants (23 males, 2 females). The age and gender composition of the sample and the motivational and national profiles corresponded to the overall structure of e-residents (Author 4). To reach out to potential interviewees, a special call for participation was posted in a private Facebook discussion group, 'e-residents of Estonia', connecting e-residents and people interested in Estonian e-residency. We additionally used snowball sampling approach, where interviewees were asked to suggest other potential interviewees. However, in both recruitment cases, we followed the purposeful sampling strategy method to assure the heterogeneity of the responses. The interviews were carried out between March 2019 and October 2020. In total, 25 participants were recruited for the study and analysed here.

Second, the interviews were conducted with key government agencies involved in the planning, development, and implementation of the e-residency project. Based on the purposeful sampling strategy, all the experts have been related to the e-residency project before or after its launch in 2014. To assure the heterogeneity of the responses, two main groups of experts having diverse expertise regarding e-residency were selected in the study: 1. experts working in different ministries that have been involved in the policy design and strategic direction of the e-residency program; 2. experts working in government agencies responsible for the implementation of the e-residency program. The expert sample equally included males (n=4) and females (n=4). In total, eight experts were interviewed from May to June 2020.

3.2. Data collection and analysis techniques

All interviews were conducted online via audio or video-enabled channels such as Skype, Google, Facebook (video), Messenger or Facetime. The online interviewing method was used both regarding the participants' geographically dispersed places of residence (in the case of e-residents) and to allow participation in the study during a tight schedule and a Covid-19

pandemic situation (in the case of experts). The primary focus of interviews with e-residents was on individual rationales for be(com)ing an e-resident, personal experiences and evaluations of the e-residency concept regarding its advantages and limitations, understandings of one's status as an e-resident in Estonia, experiences, and imaginaries in regard of the accessibility to the project, awareness, and understandings about the data sharing, use and potential datafied control of the e-residents. The expert interviews used similar questions about the start of the project, obstacles, and challenges in the course of the project, personal experiences and understandings of the e-residency concept, the potential global meanings of the project, and the role of data used to plan, control and predict the-residency project, its applications and activities.

Each interview was approximately 1-1.5 hours long. All interviews were recorded under the consent of the interviewee. The interviews were transcribed using automated solutions; all the automatically transcribed texts were edited and revised when needed, based on the audio listening to the recorded interview. Interviews with e-residents were conducted in English. Interviews with experts were conducted in Estonian, and the interview extracts presented in the analysis were translated word by word manually.

For the analysis of the collected interview data, we used thematic analysis as the principal method (Braun & Clarke, 2006). We started our analysis with a close reading of the interview transcripts and inductive coding, focusing not only on thematic aspects of *what* participants talked about (e.g., their motives for becoming an e-resident and practices using the digital identification, their ways of positioning themselves as e-residents vis-à-vis the Estonian state, and globally etc.) but also on *how* talked about these aspects, for example, what kind of rhetorical devices they deployed to make their arguments and what type of reasoning they used in their talk. The first coding was followed by further systematisation and organisation of initial codes around emerging themes and sub-themes. In this process, initial codes were also revised and refined or merged where appropriate to better structure the identified themes. The inductive inference logic was used for organising the textual data into broader thematic categories, which were analysed in the next section. For the coding and analysis of the interview data, we combined the manual and software-aided techniques using the qualitative data analysis software MAXQDA (Kuckartz & Rädiker, 2019).

The findings presented in this article are illustrated with extracts from the interview transcripts. To ensure the anonymity and confidentiality of the study participants, only numeric acronyms (eR1-eR25, Exp1-8) and nationalities (i.e., citizenship) in the case of e-residents are presented in the analysis. To assure the anonymity of the experts, no detailed socio-

demographic or professional information about the participants is presented in the analysis because a limited number of people have been working on the e-residence program.

4. Findings

4.1. Borders in Place – E-residency for Whom

During the interviews, the experts often pointed out a shift in the marketing efforts of the e-residency program. The experts explained that the program was initially developed for and advertised to the whole world with the promise of granting remote access to digital services in Estonia. The e-residency program was portrayed as an innovative solution for eliminating the need to cross borders. It could help brand Estonia as a digital state and improve its reputation, as previously discussed by the researchers (Author 1; Author 3; Drechsler, 2018). However, despite a positive perception, both police and tax authorities raised their concerns about potential security problems. Although different kinds of risks had been mapped in policy documents before the implementation of the e-residency program in 2014, there was minimal public debate(s) about the risks and security concerns during the programme's early years in 2014-2015.

The issue of preventing potential misuse of the e-residency scheme gained public attention only in the following years, as indicated in the expert interviews. As a result, rather than providing services to everyone across the globe, the e-residency program started to target specific countries in its marketing efforts, especially after publishing the White Paper on e-Residency, a strategic policy document setting further goals for the program, in 2018. This document called for a shift 'from quantity to quality' regarding potential e-residents and applicants of the program. The shift towards higher selectivity of (potential) e-residents was also an aspect that the experts unanimously emphasised and commented on in the interviews. Initially, the e-residency platform was implemented by the so-called trial-and-error experimental approach, where the goal was to clarify the potential interest and demand for such an e-service and to create a so-called primary customer base that could be used to start offering and developing additional e-services. Since the beginning, the e-residency program was also designed as a marketplace, creating a market and customers for various private sector e-service providers (e.g. accounting services, tax consulting etc.). In the second stage of development, as formulated in the aforementioned policy document in 2018, the program was given a more specific focus, and potential e-residents and applicants were differentiated into various risk categories based on their citizenship/country of origin and also prioritised based on the expected 'profitability'.

The experts often positively perceived the new, more target-oriented approach combined with datafied control principles. They underlined the importance of risk mitigation and the potential benefits/costs of the e-residency program for Estonia. While the e-residency program is still open to everyone, there is a targeted group of people who are preferred as potential e-residents as indicated by the experts, and this preference is usually based on the home country and possible activities of the applicants and their calculated and expected benefits/costs to Estonia. However, the benefits and costs were not understood exclusively in financial terms. The experts also emphasised the role of e-residents acting as 'global advocates' of the e-residency programme and contributing to the reputation of Estonia as a digitally advanced and trustworthy state. The possibility of using Estonia's e-residency programme and its services was a primary concern which could tarnish Estonia's reputation, as indicated by the experts. One of the experts explained the approach by underlining the importance of security and Estonia's reputation:

> *After all, the conceptual idea was that the state provides an e-tool that is reliable and secure and that others can trust - that the person has been identified, that the person has been verified and that we do not give access to e-solutions for criminal purposes*. (Exp 1)

As the expert pointed out, the precondition to providing a secure tool is very much related to providing it to people who have 'good' intentions and who are correctly verified against their existing identification documents (e.g., such as their national passport) and data-based background checks concerning overall 'trustworthiness'. However, as the experts pointed out, it is challenging to verify people's identities and business activities due to the lack of access to cross-border data in many cases. In this aspect, the home country and country of business operation proved to be an essential criterion in determining who is more likely to be preferred e-residency applicants, as emphasised during the interviews. A problem also voiced by experts on tax policy concerned the lack of information on how e-residency may be used for business, especially for e-residents' companies not registered in Estonia, as there is no reporting duty in such cases.

Many experts voiced concern regarding the lack of reporting and access to cross-border data, which was why providing e-residency only to certain countries would be more effective and secure. While this raises questions about data inequality, the statements of the experts indicate that starting from the initial stage of the e-residency process – even before the actual application, people from certain countries are perceived/portrayed as less desirable for the e-

residency program – program that has promised to eliminate borders, at least digitally - or virtually (Author 3).

Similarly to experts, the e-residents also underlined the importance of their home country as a factor that could affect their e-residency program application. Their reflections indicated that they are often expected to experience more positive or negative experiences in the application process based on their national citizenship. The arguments proposed by the experts were confirmed as the e-residents from both in and outside the EU expressed that they would expect the process of obtaining e-residency to be easier for EU citizens, while citizens of certain countries are considered to be potentially problematic as e-residents due to the political regime or reputation of their country of origin, which, unfortunately, is beyond the control of its citizens. The e-residents stated:

> *I am French. So probably, my access [to e-residency] is easier than if I were Pakistani or Afghan.* (eR6, French)
>
> *So, the only practical way for people from politically disadvantaged countries to access more opportunities is to relocate to other parts of the world. Seriously, there is not much else they can do. They can't change their regime. They can't influence global politics in any reasonable way. So, if you're [from a] wrong [country], I don't know from Saudi Arabia or Somalia right, you will have a tough time doing anything. So, if you are inclined to do global business, maybe it's a good idea to move somewhere else like, Estonia, for instance, or any other part of the world that doesn't carry this political stigma.* (eR12, Russian)

The statements indicate that e-residents perceive citizens from certain parts of the world to be securitised and consequently isolated from relevant business opportunities offered by Estonian e-residency. European and non-European e-residents expected the application process to go more accessible for European, especially EU citizens, as these countries do not have a political stigma and are not associated with money laundering or terrorism financing. Some e-residents openly mentioned that 'being white and male' is a privilege, and the e-residency program is no exception. The interviewees from non-EU countries often argued that the approach was somewhat justified. It was also mentioned that if someone wants to do a global business, and they are from a country with an authoritarian or totalitarian regime, they have very limited/almost no chance of doing global business successfully[71]. Thus, the limitation of

7

people's agency depending on their country of citizenship and origin was underlined. Some non-EU interviewees were happy not to be excluded from the program.

> *At least I know they [the Estonian authorities] are not excluding them [Africans] [from e-residency], which is good.* (eR23, Rwandan)

The statement of the Rwandan e-resident indicates how the perspectives of people are shaped by the relative opportunities they have regarding international services. Many international services, from money transfers (Koker, 2006) to investment opportunities, are difficult to access for many African countries due to potential risks and strict due diligence checks. As a result, the lack of exclusion from Estonia's e-residency program is perceived positively by some of the e-residents, even if European applicants are thought to be favoured overall.

4.2. Better Data for Better Services?

As indicated earlier, the datafied control of e-residency program was applied since the start of the program 2014. However, since the emergence of the White Paper in 2018, the more focused targets of the program, and the datafied tools to implement this selection, were formulated. In this document, the target marketing countries were more specifically formulated, and the differentiation of the applications, based on the potential risks for the program and Estonian country, were specified. Therefore, also the datafied control and selection turned more thorough and elaborated. As indicated in this document and also emphasised by the experts, the datafied control measures are primarily aimed to assure the program's changed economic and political goals. Therefore, the datafied control of the applications and the activities of the e-residents were assumed to support minimising of the risks, while maximising the revenues.

The main concern raised about using data-based control procedures was the lack of reliable data that could be used to analyse and make effective decisions to ensure the e-residency program is used only for its intended purpose. To that end, the experts often discussed the importance of data and its quality. A primary concern was the limited capacity, especially when verifying background data and (identity) documents in the case of applicants coming from countries that do not share their official data with Estonian government authorities. In other words, there is a lack of access to official administrative data concerning citizens of certain countries. Official requests to verify certain information an applicant presents cannot always be submitted to the countries' respective government agencies. Because of that, Estonian government authorities have to rely on various data sources in their decision-making processes, including semi-official databases formed based on 'data scratch' from the Internet.

Potential data manipulation or incorrect information provided by the applicants was considered a severe problem. Some of the solutions suggested by the experts focused on using more advanced (and often invasive) technologies such as biometrics or accepting applications only from countries that Estonian authorities already cooperate by data collection and sharing (such as the Schengen Visa Information system). As one of the experts indicated:

> *If we talk about information and the data presented in the application, the problem is that we don't see these persons. They apply through the e-application environment and fill in all these fields, but we don't know if they did it themselves or did anyone do it for them. After all, we do not have any video identification behind it, and even if we had [such video identification], it could be manipulated very easily today.* (Exp 2)

The experts also wished for further data exchange and cooperation between countries to minimise the risk of international financial crimes. As such, it seems unlikely given that the right to realise the value of their data is at stake due to the algorithmic rivalry between the US and China/Russia and the way data ecosystems are quickly being shaped and re-arranged at the global level. One of the experts stated:

> *We don't have any kind of super database or network that we can see whether a person is involved in tax fraud, for example, or has been banned from doing business. And since we do not have this information and access, there is a high risk that a person with a business ban in their home country may become an e-resident.* (Exp 3)

As both previously mentioned statements confirm, data reliability and access are critical aspects that require cooperation between the states (Shachar, 2018). The experts often imply that 'more' data and 'better' technology are potential solutions to ensure the e-residency program aligns with its intended purposes. This approach demonstrates that data inequalities between states also affect citizens of these states who want to utilise the e-residency program. The capacity of the states to collect, store, analyse, and, when necessary, share data with other relevant parties is not the same.

Moreover, states often have different jurisdictions regarding which data to collect, how, and with whom to share it. European Union (EU) countries cooperate in multiple ways regarding their cross-border data; this networked data sharing enables these countries to access better/more data about their citizens (Mushak & Zaporozhets, 2020). Considering e-residency is a program that includes people who strive to have lifestyles and operate businesses free from a specific physical location, the risks they may pose are also networked; it requires networked data cooperation to control potential risks and risky e-residents. The lack of such networked data cooperation leads to inequalities between EU citizens and countries with less developed

data capacities/networks. The citizens that need digital solutions the most are subject to further/stricter scrutiny through datafied control, as this analysis of the e-residency system has indicated. The citizens from countries with advanced data capabilities and networks supposedly need such services relatively less, yet their engagement with such programs is further encouraged.

Similarly to the experts, interviewed e-residents also expressed that access to data about EU citizens is easier, and therefore, Estonian authorities can provide the services more efficiently. However, will this make cross-border data controls vanish and allow 'citizenship by connection' among European citizens toward pan-European data flows? This idea was supported by examples presented by informants, where applicants from the EU are granted e-residency faster than non-EU applicants. One e-resident stated:

> *If you are an EU member, they [Estonian authorities] have more control over your activities. If there is a problem, they can have access to you. But for someone in Asia or Africa, it's more problematic because they don't have control over you. So, you can [do] fraud sometimes. Actually, the main problem is money laundering. /---/ I said Africans and Asians because, from an Estonian perspective, it's hard to control and to have trust [in them]*. (eR17, Algerian)

The e-resident's statement equates data with control and, similar to expert views, also argues that a lack of data on people from Asia and Africa would mean less control for Estonia and make it more difficult for the Estonian authorities to trust these people. The power of data and the importance of sharing it with the relevant authorities have been mentioned throughout the interviews. At this point, there emerges another level of inequality: the countries with developed data capacities can run programs, collect data from their citizens and non-citizens, and improve their data capabilities even further. In cases where the data is not collected or is considered insufficient, potential opportunities are denied to people, especially those from countries with less developed data infrastructure and networks.

While great importance was attributed to digital data and its accessibility as an effective way to control (and prevent) potential criminal activities, some of the e-residents raised concerns about data, potential bias within the data and ethics. As the e-residents stated:

> *As far as I know, this is still an unsolved problem with implementing ethical algorithms and ensuring that your learning data is unbiased*. (eR1, Russian)
>
> *No algorithms and no data are unbiased*. (eR4, Hungary)
>
> *I don't think there is a problem in using data if the data owner has been informed*. (eR15, Kenya)

The informants expected Estonia to consider the well-being of its (prospective) data citizens by ethically handling their data with no bias. As the statements indicate, e-residents' concerns draw attention to potential downsides of technological solutions and require further analysis focusing on sensitive issues such as data ethics, bias, informed consent, fair/non-discriminatory data, and technologies. While e-residents largely supported data collection efforts as a part of the e-residency program, the concerns they discussed raised questions about another layer of inequality that data citizens have very little control over. Thus, ensuring the right to use and realise their data and protecting their digital rights revealed crucial for e-residents. In addition, because the data is mainly cross-border, it can be assumed that it is even more challenging to detect and eliminate potential biases and ensure data quality and its proper application. In cases where the services are vital to people, potential downsides of technology must be considered from the perspective of authorities and experts and the data citizens that are subject to these technologies.

4.3. Data for Deciding Who is Risky and Who is Safe

The informants discussed risks and risk mitigation through data and more advanced technology. There were different approaches to how to accept people for the e-residency program. While some experts argued for strict controls before accepting the applicants, others argued that the risks could never be entirely eliminated but instead minimised. Therefore, they emphasised the importance of the benefits the e-residency program may bring to Estonia and that more people should be admitted into the program. As one expert indicated (Exp 4), from the state's point of view, it is a matter of finding a balance between expected/desired economic income (which can be maximised by granting e-residency to the maximum possible number of people) and possible risks (potential misuse of e-resident's digital-ID and its prevention through comprehensive control mechanisms). In other words, what is the balance point between the country's economic interests and the country's security, both being public interests? Therefore, one of the experts also emphasised that to earn (more) income, the state must be ready to take (more) risks.

Accordingly, some experts supported the idea of accepting reasonably risky e-residents due to the potential financial benefits the program may have. A solution proposed to overcome risks was to monitor the activities of e-residents after their application is accepted. This way, the experts can create individual risk profiles based on new data they can collect and the data they collected during the application process. This will ensure that the program is being used per its intended purposes and that predictions about the e-residents are correct. An expert stated:

> *Today is that we have finally come so far in our technical development that we will then move from the emphasis on ex-ante control to ex-post control. If so far, we had done this in the form of ex-ante controls within our capacity and that of the state in general, i.e. before we granted or refused the request, we did these background inquiries and checks, now - as we see that technology in the follow-up view: that if an e-resident already has an e-residence and uses e-services, look more closely, more systematically, based on the risk profile, of what he or she is doing with the card* (Exp 5).

As understood from the statements, the selectivity applied by the experts for the e-residency program is based on potential security risks an applicant may pose. As a solution, the applicant's data is analysed, and the possible risk probability is predicted based on the available data. In cases where the application is accepted, they are still subject to further monitoring to ensure that how they use their e-resident digi-ID is in accordance with the purposes of the programme.

There is an ongoing effort to improve the data and prediction capacities of the program. However, the data-based control mechanisms for e-residency depend significantly on data exchange between countries. Besides, the state in its administrative processes, i.e. in these same data-based control mechanisms, is forced to use many different data sources, including the so-called Google 2 databases, i.e. data scraped from internet resources, which one of the experts referred to, the quality of which (integrity, validity, reliability, so on.) are often doubtful. In essence, the state uses, among other things, data created by private platforms (especially if it does not have access to the so-called national databases of other/non-EU countries) to make administrative decisions - i.e. whether and to whom to issue its digital ID.

Once the application process is complete and e-residency is granted, the e-residents do not focus on ongoing monitoring processes. They express an overall willingness to share relevant data, but not with any country, however, to the countries such as Estonia, whose digital state and the democratic political regime are trusted. At this point, the e-residents consider themselves digital citizens of Estonia whose data has been reviewed, verified, and confirmed by the e-residency program. They have passed any possible risk analysis and proved to be trusted digital citizens of Estonia who are ready to use the services provided by the e-residency program. As a result, they expect to have further benefits from this selection process. One of the informants stated:

> *This card needs to tell the consulate or embassy of other country that this African guy is already ahead of the curve because being the [e-]citizen of Estonia, and Estonia is the most advanced digital nation in the world. So, if Estonia trusts this citizen, it means that*

the EU can trust this citizen, and the embassy [of another country] also can trust. (eR13, Senegal)

The e-resident wants the information produced through data and predictive analytics about themselves as trusted e-residents to be acknowledged by other EU countries. Seen as a trustworthy e-citizen of Estonia, the e-resident wants the benefits of applying to his activities within other EU countries as well.

Contrary to positive expectations that e-residency provides extra benefits, many e-residents discussed issues preventing them from using e-residency services. For example, an African e-resident stated:

> *The purpose of the e-residency is to help Africans to overcome the barrier with the African passport and to get e-passports, to do business like the French, the Belgian and the Estonians. But in practice, no, it is the same constraints, the same thing for me – many banks refused, and some banks accepted me. But it is something that is very hard, really hard*. (eR22, Tunisia)

The statements regarding difficulties in accessing banking services despite being an e-resident were abundant throughout the interviews. Even if the Estonian government grants them e-residency, many African informants stated that accessing essential banking services (as private corporate services) in Estonia was difficult for them because of their African passports. Thus, the e-residency granted to these informants did not fully facilitate their participation in the financial sphere in Estonia, and that is not because of the e-residency program and its datafied control but due to strict precautions banks take for citizens of certain countries. This shows that the e-residency program and the opportunities it provides operate, even if partially, within a larger system (global financial system) of control and inequalities that consider specific geographies riskier and consequently deny the people some services creating further disadvantages for people who are already isolated due to their home countries. While global networks create, use, store, share, and analyse data, the countries that often fall out of these networks and lack the relevant data capacities are perceived as risky, leading to severe consequences for the citizens of these countries.

Therefore, the analysis results revealed that the datafication of the e-residency program includes both external and internal processes. Externally, the datafied control of digital identification is seen as the desired target, i.e., datafied control is publicly and strategically understood as an innovative way of implementing novel data-driven tools. However, there are several obstacles in implementing the datafied control when the e-residency as a digital system crosses the state borders, which assumes data transfers across societies, which needs

negotiations for data sharing. Internally, the participants of the e-residency program often tend to internalise their datafied transactions and meanings of these transactions, that in turn tends to modify their feelings and attachment to the global affiliation with this digital identity system.

5. Conclusion

This article used Estonia's e-residency program as a proxy for digital identification systems assumed to provide global access to public and private e-services in Estonia. We aimed to explore how experts and data subjects experience, imagine and understand the control of the e-residency program through data resulting in an emerging regime that we coin as 'citizenship by connection'.

This article reveals that the datafication of digital migration through digital identification systems entails controlling the applications, access and practices that do not only lead to the control but also the reconstruction of digital borders through social selection and sorting. Our research indicates that the e-residency program aims to include a diverse set of people, and practices of exclusion and securitisation are present at various stages starting from the initial marketing efforts to the ex-post control through data. Both the experts and e-residents acknowledge the differences in approaches towards EU and non-EU applicants; however, this is not discussed in terms of discrimination but often with a focus on security/risk and the capacity of relevant authorities to verify data and (identity) documents. However, many of the e-residents and experts also underlined Estonia's responsibility to deal with the data in a fair, unbiased, and ethical way. The expectation from Estonia regarding an ethical approach to data indicates that new kinds of data and data relations between people and authorities require people to participate in policy debates as data citizens, as previously stated in literature (Gregory & Bowker, 2016). Rather than solely focusing on technological solutions, this study confirms the importance of addressing social, political, and ethical dimensions globally (Tavmen, 2020). Simply put, e-residents as data citizens should be informed about relevant data practices and have enough voice to argue for their interests and digital rights.

Therefore, this article reveals that data connecting individuals striving for a placeless lifestyle through their digital transactions do not create reality as connected with one's land or national connections, but in the form of digital borders through new forms of '*jus nexum',* 'citizenship by connection'. These new emerging digital practices, often named in the literature as datafied or algorithmic citizenship, are characterised by connection processes through global datafied transaction practices and a state of scaleless connectedness, as this study revealed. This lifestyle connected through data does not allude to normative concepts but strives to 'decode'

the social space as a multifaceted place-centred form of diversity (Kitchin & Dodge, 2014; Sachs, 2009). Yet, the promises of innovative technology programs are not fully realised for everyone, and global inequalities are often difficult to eradicate, even if technological remedies are being developed and applied.

The study utilised a qualitative approach and explored newly emerging relations where de-territorialized data that connect people to endless global and local points result in the exclusion of some groups based on the Estonian e-residency case. The interviews were conducted with a small number of e-residents and experts. Further research that explores multiple cases with larger samples in different fields would provide an essential contribution to literature in terms of understanding the emergence of new data relations at and across the borders around digital citizenship, even beyond nation-states.

Acknowledgements

We want to thank all the interviewees who took part in our study. We are also very thankful to our colleagues Tam Abaku, Mergime Ibrahimi, Rahi Patra and Tsagana sBadmaeva for their support in collecting the interview data used in this study. We also thank Prof. Ines Mergel and Prof. Rainer Kattel for their valuable discussions on the e-residency program. This research was financed by the Development Program ASTRA of Tallinn University of Technology (2014-2020.4.01.16-0032), NordForsk and Estonian Research Council project 'Critical understanding of predictive policing' (ETAG20083), and ESRC Programme ES/S012435/1.

References

Amoore, L. (2021). The deep border. *Political Geography*. https://doi:10.1016/j.polgeo.2021.102547.

Anwar, M.A., & Graham, M. (2022). *The Digital Continent: Placing Africa in Planetary Networks of Work*. Oxford: Oxford University Press.

Barlow, J. P. (1996). *A declaration of the independence of cyberspace*. https://vimeo.com/111576518?ref=tw-v-share [Google Scholar]

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Broeders, D. (2016). The New Digital Borders of Europe: EU Databases and the Surveillance of Irregular Migrants. *International Sociology*. https://doi.org/10.1177/0268580907070126

Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review, 9*(2). https://policyreview.info/articles/analysis/data-citizenship-rethinking-data-literacy-age-disinformation-misinformation-and

Chouliaraki, L., & Georgiou, M. (2022). *The Digital Border: Migration, Technology, Power*. NYC Press: NYC.

Couldry, N., & Mejias, U. A. (2019). Making data colonialism liveable: How might data's social order be regulated? *Internet Policy Review*, *8*(2). https://policyreview.info/articles/analysis/making-data-colonialism-liveable-how-might-datas-social-order-be-regulated

Drechsler, W. (2018). Pathfinder: e-Estonia as the β-version. *eJournal of eDemocracy & Open Government, 10*(2), 1-22.

Emmer, M., & Kunst, M. (2018). "Digital Citizenship" Revisited: The Impact of ICTs on Citizens' Political Communication Beyond the Western State. *International Journal of Communication, 12*, 2191–2211.

E-residency 2.0 White Paper. (2018). https://s3.eu-central-1.amazonaws.com/ereswhitepaper/e-Residency+2.0+white+paper+English.pdf

Faure, L., Vendramin, P., & Schurmans, D. (2020). A situated approach to digital exclusion based on life courses. *Internet Policy Review*, *9*(2). https://policyreview.info/articles/analysis/situated-approach-digital-exclusion-based-life-courses

Goggin, G., Hollier, S., & Hawkins, W. (2017). Internet accessibility and disability policy: Lessons for digital inclusion and equality from Australia. *Internet Policy Review*, *6*(1). https://policyreview.info/articles/analysis/internet-accessibility-and-disability-policy-lessons-digital-inclusion-and

Graham, M., De Sabbata, S., & Zook, M. A. (2015). Towards a study of information geographies: (Im)mutable augmentations and a mapping of the geographies of information. *Geo: Geography and Environment*, *2*(1), 88–105. https://doi.org/10.1002/geo2.8

Gregory, J., & Bowker, G. C. (2016). The data citizen, the quantified self, and personal genomics. In D. Nafus (Ed.), *Quantified: Biosensing Technologies in Everyday Life*, 211.

Kitchin, R., & Dodge, M. (2014). *Code/Space: Software and Everyday Life*. MIT Press.

Koker, L. (2006). Money laundering control and suppression of financing of terrorism: Some thoughts on the impact of customer due diligence measures on financial exclusion. *Journal of Financial Crime*, *13*(1), 26–50. https://doi.org/10.1108/13590790610641206

Kuckartz, U., & Rädiker, S. (2019). Analyzing qualitative data with MAXQDA (pp. 1-290). Basel, Switzerland: Springer International Publishing.

Latonero, M., & Kift, P. (2018). On Digital Passages and Borders: Refugees and the New Infrastructure for Movement and Control. *Social Media + Society*, *4*(1), 2056305118764432. https://doi.org/10.1177/2056305118764432

Leese, M. (2020). Fixing State Vision: Interoperability, Biometrics, and Identity Management in the EU. *Geopolitics*, *0*(0), 1–21. https://doi.org/10.1080/14650045.2020.1830764

Mann, M., & Daly, A. (2020). Geopolitics, jurisdiction and surveillance. *Internet Policy Review*, *9*(3). https://policyreview.info/geopolitics-jurisdiction-surveillance

Mathew, A. J. (2016). The myth of the decentralised internet. *Internet Policy Review*, *5*(3). https://policyreview.info/articles/analysis/myth-decentralised-internet

Mukherjee, R. (2019). Jio sparks Disruption 2.0: Infrastructural imaginaries and platform ecosystems in 'Digital India.' *Media, Culture & Society*, *41*(2), 175–195. https://doi.org/10.1177/0163443718818383

Mushak, N., & Zaporozhets, A. (2020). Law Enforcement Cooperation of the Member States of the European Union. *Law Review of Kyiv University of Law*, 2020(3), 338-342.

Oreglia, E., & Ling, R. (2018). Popular Digital Imagination: Grass-Root Conceptualization of the Mobile Phone in the Global South. *Journal of Communication*, *68*(3), 570–589. https://doi.org/10.1093/joc/jqy013

Sachs, W. (2009). *The Development Dictionary: A Guide to Knowledge as Power*. Bloomsbury Academic.

Sander, I. (2020). What is critical big data literacy and how can it be implemented? *Internet Policy Review*, *9*(2). https://policyreview.info/articles/analysis/what-critical-big-data-literacy-and-how-can-it-be-implemented

Schia, N. N. (2018). The cyber frontier and digital pitfalls in the Global South. *Third World Quarterly*, *39*(5), 821–837. https://doi.org/10.1080/01436597.2017.1408403

Shachar, A. (2018). The marketization of citizenship in an age of restrictionism. *Ethics & International Affairs*, *32*(1): 3-13.

Sontowski, S. (2018). Speed, timing and duration: Contested temporalities, techno-political controversies and the emergence of the EU's smart border. *Journal of Ethnic and Migration Studies*, *44*(16), 2730–2746. https://doi.org/10.1080/1369183X.2017.1401512

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative research journal.*

Tavmen, G. (2020). Data/infrastructure in the smart city: Understanding the infrastructural power of Citymapper app through technicity of data. *Big Data & Society*, *7*(2), 2053951720965618.

Toots, M. (2019). Why E-participation systems fail: The case of Estonia's Osale.ee. *Government Information Quarterly*, *36*(3), 546–559. https://doi.org/10.1016/j.giq.2019.02.002.

Vassil, K., Solvak, M., Vinkel, P., Trechsel, A.H., & Alvarez, R.M. (2016). The diffusion of internet voting. Usage patterns of internet voting in Estonia between 2005 and 2015. *Government Information Quarterly*, *33*(3), 453–459. https://doi.org/10.1016/j.giq.2016.06.007.

[1]For these reasons, Estonia has also stopped to accept submissions and stopped issuing applications submitted by Russian and Byelorussian citizens due to the war in Ukraine. https://learn.e-resident.gov.ee/hc/en-us/articles/4575271559441-Restrictions-on-Russia-and-Belarus