Chapter 2

The Use of Digital Technology in the Fight Against Welfare Fraud: Comparative Analysis of Selected National Experiences

Łukasz Jurek

Assistant Professor, Wroclaw University of Economics and Business
ORCID: 0000-0002-0078-471X

Quote as: Jurek, Ł. (2024). The Use of Digital Technology in the Fight Against Welfare Fraud: Comparative Analysis of Selected National Experiences. In P. Luty, N. Versal, & P. Semerád (Eds.), *Knowledge and Digitalisation Against Corruption and Fraud* (pp. 23-35). Publishing House of Wroclaw University of Economics and Business.

DOI: 10.15611/2024.96.3.02

2.1. Introduction

The term ‘welfare state’ is ambiguous. It was coined in England during the Second World War in contrast to another term that was common at that time, ‘warfare state’. The replacement of ‘warfare’ by ‘welfare’ was intended to be a symbolic shift from an economy focused on pro-arms production, which provided citizens with military security, to pro-social production supposed to provide citizens with social security.

The welfare state was developed for both equity and efficiency reasons. Thus, according to Barr (2020), it can be thought of as (1) a set of institutions that provide poverty relief, redistribute income and wealth, and seek to reduce social exclusion (the ‘Robin Hood’ function), and (2) a set of institutions that provide insurance and a mechanism for redistribution over the life cycle (the ‘piggy bank’ function).

The design of the welfare state varies from country to country (Esping-Andersen, 1990). Despite some differences, the overall size (tasks, funds, administrators) of the welfare state is systematically increasing in Western countries (Quadagno, 1987), due to the growth of traditional issues caused by demographic change and shifts in industrial relations. There is also the emergence of new risks related to, for example, epidemiological issues. All this increases the pressure on public authorities to extend a safety net for citizens.

The development of the welfare state has enabled to overcome crucial social problems, however it also raised a number of concerns about operating costs and administrative efficiency. Excessive bureaucracy hampers social policy programmes, especially at the stage of benefit distribution. The multiplicity of procedures and ineffective control make the welfare system open to various types of abuse and fraud. Public funds that were to serve good causes

are being extorted by unauthorised persons. This problem, although for a long time overlooked and/or neglected, is now becoming a major challenge for theorists and policy-makers. The use of modern technology seems to be the only reasonable solution in this situation. Mass data collection, automation, and artificial intelligence can essentially increase efficiency in detecting and combating welfare fraud. It is crucial therefore, as stated by Henman (2022, p. 536), "to bring the 'digital' into 'social policy'".

So far, only a few countries have digitised their welfare infrastructure. Societies where local governments have developed AI-driven technologies to eliminate non-compliance have been involved in a massive experiment. Their experience has shown that while technology reduces old problems such as low operational efficiency and high bureaucratic costs, it also creates a new set of problems and concerns, with the emerging critical issues of data protection, privacy and e-exclusion. In retrospect, the overall outcome is mixed and to some extent controversial. For this reason, the general narratives in the literature range from the techno--optimistic to the techno-pessimistic (Fugletveit & Sørhaug, 2023).

This chapter describes the experiences of two countries (the Netherlands and Australia) in using digital technologies to detect and combat welfare fraud, and also attempts to assess the strengths and weaknesses of such a digital transformation. It is both analytical and descriptive, and based on analysis of literature and official documents.

2.2. Digital Welfare State: Theoretical Approach

The world is changing fast, and the technological revolution we are witnessing today touches all areas of society. Every aspect of human existence is more or less affected by modernity, with technologies penetrating every aspect of life. However, the pace of this diffusion varies in different areas. Public administration has traditionally been very conservative and resistant to change, hence innovations do not affect it as quickly as, for example, the corporate world. Sometimes it is necessary to stimulate this development, in order to encourage the adaptation of new tools based on technology.

In general, digital transformation may be described as the transition from 'analogue' to 'digital'. This shift has also affected the welfare state. In order to emphasise the scale and quality of change, various load-bearing labels are used in the literature, such as: digital social policy, social policy 4.0, and digital welfare state (Szatur-Jaworska, 2023).

In 2014, on the initiative of the OECD Council, member countries adopted Recommendation of the Council on Digital Government Strategies (OECD, 2014). The document called on governments to develop and implement digital strategies in order to achieve digital transformation. Technologies were to be a strategic driver to create an open, participatory, and trustworthy public sector, to improve social inclusiveness and government accountability, and to bring together government and non-government actors and develop innovative approaches to contribute to national development and long-term sustainable growth.

Digital transformation in the public sector responds to the need to modernise public services. This change seems inevitable in the face of new and growing expectations of the welfare state and the necessity of state administration to deal with increasingly complex issues. As such, it has become a political imperative to improve the efficiency, effectiveness and governance of public services by designing and implementing innovative technological solutions (OECD, 2016).

At normative level, a major contribution to the digital transformation of the welfare state has been made by the United Nations. The notable report on extreme poverty and human rights (Alston, 2019) highlighted the strategic importance of the digital transformation of the welfare state as a transformation of the relation between the state and citizens, from analogue to digital, to improve governance through efficiency, integrity, and transparency. The document established the term 'digital welfare state', in which "systems of social protection and assistance are increasingly driven by digital data and technologies that are used to automate, predict, identify, survey, detect, target and punish" (Alston, 2019, p. 4). Moreover, it was stressed that

> new forms of governance are emerging which rely significantly on the processing of vast quantities of digital data from all available sources, use predictive analytics to foresee risk, automate decision-making and remove discretion from human decision makers […]. In such a world, citizens become ever more visible to their Governments, but not the other way around. (Alston, 2019, p. 4)

The concept of the digital welfare state has been subject to much development and modification since then. In their extensive study on datafication in the context of welfare state, Dencik and Kaun (2020, p. 2) explored a "shift toward a new regime in public services and welfare provision intricately linked to digital infrastructures that results in new forms of control and support". The importance of this technical-driven transformation of the welfare state was also emphasised by van Gerven (2022):

> The welfare state and its management of social risks is clearly affected by the technological transformations of the late twentieth and early twenty-first centuries. It creates a need to reorganize and recalibrate welfare state structures and systems to extend social risk protection towards a diverse set of risks, including existing (e.g. changing family structures and labour market participation patterns) and 'new' digitally driven risks (e.g. labour market insecurities induced by platform work and automation). (p. 254)

According to Henman (2022), the justifications for the digitalisation of the welfare state can be divided into two groups: 'traditional' and 'new'. The 'traditional' justifications have been behind the transformation from the beginning and have not changed over time. These are: efficiency, cost reduction, staff savings, consistency of decisions and reduction of errors. The 'new' justifications, on the other hand, have emerged more recently. These are: policy responsiveness and agility, customer service and service innovation, personalisation, overpayment and fraud detection, improved governance and enhanced accountability and democracy.

Technology offers a wide range of solutions to improve the structure and functioning of the welfare state. The above-mentioned United Nations report (Alston, 2019) identified six areas as being particularly open to the use of modern technology.

1. Identity verification: a verifiable identity is essential for claiming benefits, establishing entitlements, receiving benefits and appealing against benefit denials. Modern identification systems contain both demographic and biometric information on all residents, including an iris scan, photograph and fingerprints.

2. Eligibility assessment: IT systems support decision-making processes and increase the efficiency of analysing large databases. They relieve caseworkers of clerical tasks, and make decisions independent of subjective opinion, ensuring transparency and objectivity. They also allow for continuous monitoring of the situation of beneficiaries in terms of compliance with entitlements.

3. Welfare benefit calculation and payments: using spreadsheets and algorithms to automate the calculation of benefits, and paying them out using digitised financial services such as bank transfers and electronic payment cards.
4. Fraud prevention and detection: informatic systems allowing to match data from different sources in order to expose deception and irregularities on the part of welfare applicants.
5. Risk scoring and need classification: algorithm-based techniques to determine whether intervention is required and, if so, at what level.
6. Communication between welfare authorities and beneficiaries: traditional forms of communication (face-to-face, telephone, letter) are being replaced by online applications and interactions (e.g. chatbots).

Generally, these areas of digitalisation can be divided into two basic domains: rule-based systems and predictive systems. A rule-based system helps to verify eligibility for benefits and, if applicable, to calculate the amount of benefits. Predictive systems, on the other hand, are risk-profiling tools that sort welfare claimants into different levels of intervention. The risk assessment is based on indicators that are identified by research and/or mandated by policy. On this basis a statistical model is formed that gives a probabilistic score for each individual. As noted by Well et al. (2023, p. 45), "companies, professionals and sometimes academics develop these assessment tools in relation to a historic comparative population and past experiences of service provision, often using stakeholder consultation, trials, validity and usability testing and factor analysis".

Today, the issue of the digitalisation of the welfare state is becoming a priority for both social policy theorists and practitioners. Many research and implementation projects in this area are being undertaken in many countries. Such initiatives are now also taking place at transnational level. One example is AUTO-WELF, the first project to provide an analysis of automated welfare provision across different European welfare regimes, which examines the implementation of automated decision-making (ADM) in the welfare sector across Europe.[1]

2.3. Welfare Fraud and Welfare Surveillance

Welfare fraud is a complex and multidimensional issue (Jurek, 2024). This term is often used as synonymous with 'welfare abuse' or 'welfare crime', although these are not entirely clear-cut. Their common feature is non-compliance with the welfare rules, resulting in incorrect payments, i.e. payments made for the wrong reason or for the wrong amount. However, not every incorrect payment is a case of welfare fraud. Firstly, the non-compliance may involve both an over- and understatement of the welfare benefit. Second, the overconsumption may be the result of unintentional error or intentional behaviour. Third, intentional overconsumption may be irrational but legal (*moral hazard*) or illegal (fraud). Fourthly, fraud can be of different 'degrees of seriousness', it can be a minor offence (the so-called crimes of everyday life) or a serious crime committed on a large scale by organised criminal groups.

The issue of welfare non-compliance has become an important issue in European Union policy. This is linked to a number of irregularities that have arisen with the increasing coordination of national social security systems. In response to these problems, it was decided (Decision No H5 of 18 March 2010…) to call on the Member States to take appropriate remedial action in

[1] Project website: https://blogg.sh.se/digitalwelfare

this area. EU documents distinguish two types of irregularities: fraud and error (Jorens et al., 2019). A fraud is defined as any act or omission contrary to national legislation, either in order to obtain benefits from the social security system or to evade the obligation to pay public contributions to maintain this system. An error, on the other hand, is considered to be an unintentional mistake or oversight on the part of officials or citizens.

European countries are developing their own definitions of welfare abuse as well as methods for the prevention, detection and deterrence. These measures are particularly advanced in Sweden, where in 2021 the country's government adopted a special ordinance that regulates initiatives to ensure proper payments from the welfare system (Sveriges Riksdag, 2021). The coordination of these activities has been entrusted to the National Institution for Financial Management (ESV). According to its guidelines (Modin & Lindblom, 2021), a payment from the social security system is incorrect if it is made even though the conditions for receiving the benefit have not been met. The benefit may be too high, too low, or completely undue based on the applicable regulations. Incorrect payments range from unintentional errors to large--scale, systematic and organised welfare crime. To illustrate this diversity, the welfare compliance pyramid was used (Fig. 1), which contains four types of incorrect payments that differ greatly in terms of motive, severity, and structure.

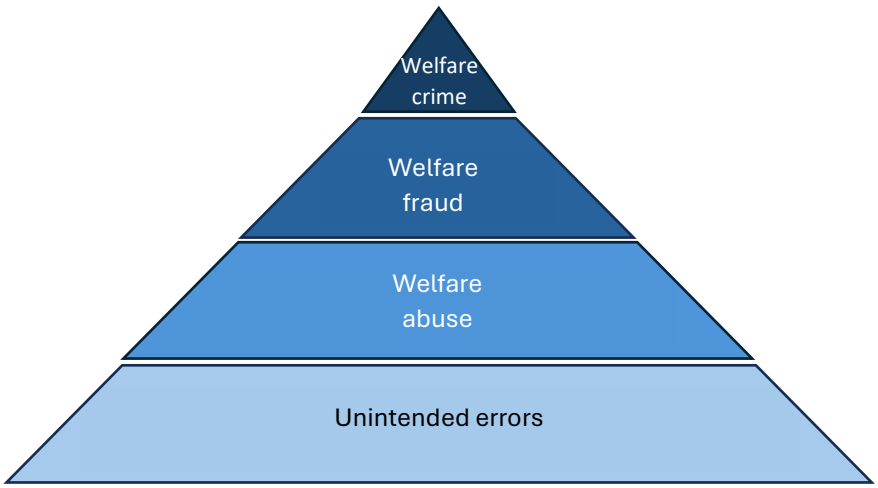

Fig. 2.1. Welfare compliance pyramid

Source: (Modin & Lindblom, 2021, p. 11).

Unintended errors are the result of mistakes that are unconscious and unintentional. Welfare abuse is the conscious provision of misinformation in order to misuse welfare entitlements. Welfare fraud is the practice of extortion and deception (e.g. bribery, forgery, cheating and lying) to obtain undue benefits, whereas welfare crime is an organised activity carried out on a large scale against welfare institutions.

Similar legislation exists in the Netherlands (Eerste Kamer, 2012). The Dutch legislation distinguishes between 'improper use of benefits' and 'benefit fraud'. Improper use is defined as "using the entitlements in a way that is allowed by law but not in the spirit of the law". Fraud, on the other hand, is defined as "the use in a manner not permitted by law".

Kukuła (2016) defined welfare fraud as

> actions of natural persons aimed at persuading a social institution to dispose of its property in a certain way and obtain material benefits not due to them, which fulfil the characteristics of a crime, consisting in deliberately misleading or exploiting the error of an institution dealing with the distribution of budget funds, aimed exclusively at providing assistance or material support to people, both those who have a source of income and those who are temporarily or permanently deprived of it. (p. 28)

It is, therefore, a manifestation of conscious and deliberate pathologies committed by private individuals in the use of welfare benefits.

By its very nature, welfare abuse is difficult to detect. The main reason is the blurred line between justified and unjustified benefit use. As described by McKeever (2012, p. 472), this boundary lies between need and greed. Usually, it is very difficult to assess whether the use of benefits is really necessary because of a difficult life situation or whether it is rather a matter of cheating. In the case of welfare fraud, however, the situation is clear, as it is a clear violation of the law.

The extent of welfare fraud is difficult to measure. It is obviously inappropriate to draw conclusions from the results of inspections carried out by control bodies, because of the potentially very large underestimation bias. This type of data does not show the actual number of frauds committed, but only the number of cases detected. The relation between these two figures, i.e. the number of frauds committed and the number of frauds detected, is unknown and depends on many factors, the most important of which is the frequency and quality of controls.

The basic problem is that the general structure of the social system was not designed for monitoring. Administrators generally see the essence of their job as providing benefits, not exercising control. The effectiveness of their work is measured by the number of payments made, not by the number of wrongly withheld payments. In some cases, the rules leave administrators a lot of room for subjective judgement, which makes it difficult to identify incorrect payments. In addition, controls tend to be carried out in those areas that are easy to perform and "produce results" (Korsell et al., 2008).

It can be assumed that the extent of welfare fraud varies from country to country. This is partly due to institutional and legal differences (control measures and severity of sanctions), but above all to socio-cultural conditions. This concerns the level of so-called 'benefit morality', i.e. the individual reluctance to exploit the welfare state through fraud (Halla et al., 2010; Heinemann, 2008).

The genesis and development of the fight against welfare fraud, although taking place in different countries and at different times, are to some extent similar. The general trajectory of events is largely convergent, forming an evolutionary process consisting of four main stages.

The first stage may be called 'welfare euphoria'. This is a period of dynamic development of the welfare state that goes hand in hand with rapid economic growth. Implementation of welfare programmes is based on trust and high moral standards in society. The control system is practically non-existent.

The second stage may be called a 'grace'. This period usually coincides with an economic crisis and fiscal restrictions. The problem of welfare abuse is then noticed. The policy-makers implement a pioneer control system. First inspections show that the problem of abuse concerns the most vulnerable. The dominant rhetoric is that fraud is forced by a difficult life situation and is an act

of higher necessity in order to satisfy basic existential needs. Control measures have a negative social reception, being identified with criticism of the welfare state and the punishment of poverty. In addition, the low cost-effectiveness of control activities is emphasised, i.e. the potential gains from the detection of abuse are usually lower than the costs of the administrative measures taken. Moreover, the effectiveness of the control system is limited by poor institutional coordination and lack of access to information.

The third stage may be called 'consternation'. This is when reports appear from public institutions and/or the media about various cases of benefit abuse. These are often examples of arrogant and perfidious exploitation of the welfare system. The issue is publicised in the mass media and becomes a popular political and journalistic topic. Attention is drawn to organised criminal activity against the system. Fraud is equated with theft, and priority is given to preventing and combating this problem. Politicians make honest attempts to estimate the extent of abuse and its costs (social and financial). The infrastructure for monitoring and combating social security fraud is being put in place: staff are being recruited and given powers, special monitoring and control institutions are being set up, legislation is being adapted, and sanctions are being made more severe. However, many of these are 'shotgun' measures, as they are taken blindly and intuitively, without a proper understanding of the nature of the problem.

The fourth (last) stage can be described as the 'surveillance welfare state'. In this case, monitoring and control activities are established as a natural public task. A systematic and organised diagnosis of the problem is carried out (a reliable diagnosis using scientific methods and techniques). Supervision and control activities are digitised, automated and computerised. Modern technologies increase the cost-effectiveness of control activities. The integration and coordination of activities within different institutions takes place in order to effectively combat fraud.

2.4. Digital Welfare Surveillance in Selected Countries

The idea of digital surveillance is based on the logic of risk in detecting non-compliances. The point is that in welfare systems (especially social assistance) it is necessary to monitor the situation of beneficiaries on an ongoing basis. At the very beginning, when people apply for a benefit, their identity, social situation and 'means' (income, assets) need to be verified. Then, after the benefit has been granted, continuous monitoring must be maintained in order to detect changes that may affect the eligibility status of beneficiaries. Naturally, there is no justification for checking everyone who uses welfare benefits as it would require a huge administrative effort. Besides, a significant proportion of beneficiaries do not commit fraud. Controlling them would be a simple waste, and instead it is advisable to select the cases that require special attention. To achieve this goal, policy makers are developing tools that calculate the risk of fraud. Such measures are based on complex statistical models that make it possible to identify (flag) those with an above-average risk and refer them for inspection. The computer programmes employ large databases containing information on various aspects of life. This data is analysed using algorithms and (to some extent) machine learning (artificial intelligence). On this basis, surveillance targets people with certain characteristics that are predicted to become involved in incorrect payments (Henman & Marston, 2008).

So far, several countries have implemented such risk-based system to detect welfare fraud, with others already working at an advanced stage on such solutions. One of these countries is Slovenia, where the SURVEILWEL project is being implemented with EU funding, which

> will employ qualitative, multi-sited ethnographic fieldwork to scrutinise the realm of digital welfare surveillance within Slovenian institutions […]. The project explores how digital welfare surveillance tools impact welfare eligibility assessments and the investigation of suspected fraudulent activities. Moreover, its objective encompasses the development of efficient interventions for social and material welfare. (Horizon Europe, 2033)

Similar activities are also taking place in Sweden. In 2024, the country's government set up a special commission to study the possibilities of increasing digitalisation in the area of social security, primarily to prevent fraud, but also to ensure a more personalised service and streamline the processing of benefit claims (Sveriges Riksdag, 2024).

To date, the Netherlands had the broadest experience in using digital technology to combat welfare fraud. Since the 2000s, data from a variety of sources have been used to identify potential abuses. One example was the ADM system, which used information from households on registered water consumption to compare it with the number of declared residents. In the following years, fraud investigation became a major application of 'public sector data analytics' (Zajko, 2023). As a result, three algorithmic fraud risk detection systems were successively implemented, first 'Waterproof' (2004-2007), then 'Black Box' (2008-2014) and most recently 'SyRI' (since 2015).

SyRI (for 'System Risk Indication') was described by Van Bekkum and Zuiderveen Borgesius (2021, p. 325) as 'a socio-technical infrastructure'. It was designed to identify potential welfare fraud by generating risk notifications and sending alerts to administrative bodies. Individuals are shortlisted and flagged as those 'at high risk of fraudulent behaviour' by identifying discrepancies in their personal data. The system cross-referenced almost all the information the government has on its citizens, relating to: employment, penalties and convictions, taxes, assets, denial of benefits, residence, identity, integration, compliance with the law, education, pension, reintegration, debts, welfare benefits, permits and exemptions, and health insurance (Wieringa, 2023).

SyRI operates on the basis of the 'black box method'. The system downloads some information about welfare beneficiaries and then analysis it using a certain algorithm. However, neither the sort of data used, nor the risk assessment model remained unknown for a general public (Van Bekkum & Borgesius, 2021).

From the outset, SyRI has been the subject of much controversy. In particular, it has been accused of a lack of transparency in the flow of information used. Claimants were not informed what kind of data, even sensitive data, was being used. The system of risk reporting was also unknown. Decisions were made automatically on the basis of a blind verification mechanism (using pseudonymised data). However, claimants were not warned that they had been flagged as a fraud risk and were not told why they had been flagged. There was also criticism that SyRI was only being used in selected communities, known as 'problem neighbourhoods', i.e. those with high rates of poverty, crime and unemployment. It was perceived by the general public as an element of victimisation and discrimination that should not be sanctioned by the law. Based on these arguments, in 2018 several civil rights organisations filed a lawsuit against the Dutch state to stop the use of SyRI. In 2020, the court ruled that neither the legislation on SyRI nor its use met the requirements of Article 8 of the European Convention on Human Rights. A fair balance had not been struck between the public interest in detecting fraud on the one hand and the human right to privacy on the other. As a result, SyRI was found to be unlawful (Rachovitsa & Johann, 2022).

Another country with a relatively rich experience of digital surveillance is Australia, where welfare governance has been transformed into a paternalistic concept of welfare for several decades; the 2007 reform introduced a pioneering income management system called BasicCard. It allows welfare recipients to buy only 'essential' items (food, fuel, clothing, rent payments, etc.), and avoid purchasing prohibited products (alcohol, tobacco, pornography, etc.) (Dee, 2013). Moreover, the Australian government has a 'zero tolerance' policy towards welfare non-compliance. Preventive measures were targeted in particular at 'welfare overpayments'. The point is that in Australia, as in many other countries, eligibility for welfare support is linked to financial situation. An increase in income may reduce eligibility for benefit. However, information about earnings does not flow to the welfare agency in real time and it is the responsibility of the individual to report any changes in this regard. If they fail to do so (for whatever reason), they may receive more support than they are entitled to and become debtors.

To eliminate this problem, in July 2016 the Department of Human Services (currently known as Services Australia) implemented an Online Compliance Intervention (OCI) programme, commonly known as Robodebt (welfare debt recovery system). It was intended to verify eligibility to welfare benefits without an excessive human-resources burden. The system drew data from two different sources: social welfare data on benefit payments (from Centrelink), and tax data on earnings reports (from the Australian Taxation Office). Information on earnings was matched with information on received welfare benefits. The system automatically detected discrepancies between the two. On this basis, an algorithm identified 'suspected' welfare recipients and issued them with debt notices (Rinta-Kahila et al., 2022).

Robodebt has been criticised for inaccurate assessments, illegality, shifting the burden of proof of debt onto welfare recipients, poor support and communication, and coercive debt collection (Braithwaite, 2020). Thus, in 2017, within a year of its implementation, it came under scrutiny by public bodies, while at the same time, activists from non-governmental organisations began to raise awareness of the system's shortcomings. The entrenched problems with the scheme continued until a legal challenge led the government to suspend the system in 2019 (Rinta-Kahila et al., 2024).

2.5. Conclusions

The digital transformation of the welfare state is a dynamic process taking place in various areas of infrastructure, management, contact with citizens, service delivery, and control and monitoring. However, the experience gained so far in implementing digital surveillance does not inspire optimism. The examples of the Netherlands and Australia, where the process is most advanced, revealed a number of critical problems. This is not to say that the projects have not worked at all and should be entirely abandoned. At this stage, the problems must be treated as challenges. In order to overcome these problems and advance the digitalisation process, the systems need to be adapted to the surrounding conditions (legislation, infrastructure, social attitudes) and, conversely, the surrounding conditions (legislation, infrastructure, social attitudes) also need to be changed and (to some extent) adapted to the system.

The first crucial problem is transparency. Digital surveillance is accused of lacking clear information about the construction of the algorithm, i.e. the method of data analysis. The system operates as a 'black box'. The general public does not know what data is collected, how it is analysed and what factors are crucial for identifying non-compliance.

This problem should be addressed as a question of dominant values in the public sphere – what is more important, transparency or effectiveness? Clear criteria for decision-making appear to be at the heart of government accountability and administrative justice, however the clarity of assessment rules raises concerns about feedback reactions on behalf of those subject to control. The situation can be compared to the fight against traffic offences. Effective surveillance requires a conspiratorial approach so that drivers do not know when or where they will be checked, otherwise they will temporarily change their behaviour and drive according to the rules for a while, only to commit offences at a later stage. One can be sure that the same is true for welfare fraud. Making the control criteria known to the public triggers a behavioural reaction, i.e. adjustment to these criteria. Potential offenders may mimic certain features just to cheat the algorithm and reduce the risk of being 'flagged' as a fraudster.

Freedom of information is at odds with the effectiveness of the surveillance system. It is therefore necessary to reach a consensus on how and to what extent control is to be carried out. Auditing should ensure clear principles of risk calculation on the one hand and efficiency of control on the other.

The second crucial problem is operational performance. To date, the predictive validity of risk assessment tools has been very low. Failure to detect welfare fraud has led to various negative consequences, such as human tragedies (deprivation, aggression, divorce), as well as a decline in public trust in the institutions of the state. However, it should be clearly stated that the fault for these errors is allegedly not with the system but with the human who developed it. The algorithms were merely 'decision trees' that estimated the risk based on the data provided and the criteria adopted. A satirical scene from the Polish comedy *Miś* [Bear] comes to mind here, where one of the characters (a collection agent) comments on the use of a computer: "it will always make a mistake when adding, sir. Not a month goes by without it making an error". This is, of course, a humorous approach to the subject of digitalisation. In reality, the problem is not the technology *per se*, but the 'input', i.e. the quality of the statistical model and the assumptions on which the algorithm is based. The human factor still plays a dominant role here.

Risk assessment tools are already used in various areas of the welfare state. Abuse detection is no exception. Algorithms and artificial intelligence will play an increasingly important role, whether we are in favour of it or not, therefore an appropriate attitude should be adopted. The failures made so far should not be seen as a 'warning' about the use of digital technology in government, but as an important lesson to be learnt. Systems, by definition, require continuous progress by eliminating errors and adapting to changing external realities. Initial failures cannot invalidate the whole solution, but rather highlight the need for modification and improvement.

According to Zajko (2023), digital transformation only replaces human bureaucrats by autonomous machinery. Digital technologies are designed to replicate the decisions made by people in the analogue world. They automate the decision-making process based on parameters and constraints set by humans. Wrong decisions are therefore the result of wrong assumptions. In addition, the role of the algorithm is to calculate an individual's risk score based on an assessment of a dataset of people claiming benefits, in order to identify those who are most likely to be fraudsters. The problem is that this is a very complex issue. The propensity to commit welfare fraud is conditioned by personal and contextual factors. Many of these are not obvious and are still poorly recognised. For this reason, it is extremely difficult to create an algorithm that effectively predicts human behaviour.

Therefore, at the current stage of development, digital technology should only be treated as a tool to support the decision-making process. Algorithmic governance should take a hybrid form, with humans making the final decision on cases that have been algorithmically flagged for scrutiny.

The third crucial problem is the issue of personal data protection. To be effective, digital surveillance systems need to access various databases, including sensitive ones, without the knowledge or consent of welfare recipients. It is claimed that such scrutiny violates the right to privacy. However, as noted by Henman and Martson (2008), such arguments are based on ideas of personhood and social governance inherited from the early days of western modernity. They presented three assumptions for data-based surveillance. Firstly, concerns about privacy are predicated on a liberal ideal of personal property and personhood – an independent, rational, self-updating human being. Yet in a world of growing interdependence and collective destiny, the right to be left alone is outdated. Secondly, the notion of privacy is based on a dichotomy of public and private responsibility, which is problematic from the perspective of a welfare state where the boundary between the two is unclear. Thirdly, the image of surveillance is changing, from an omnipresent panopticon-style vision to the tracking of a person's data ('dataveillance').

Following the rhetoric of today's researchers and practitioners, one gets the impression that too much attention is being paid to the problems and limitations of digital transformation and not enough to its potential benefits. Technology not only facilitates existing activities, it opens up new possibilities. The digitalisation of the welfare state allows for more individualised and differentiated decisions, thus opening up a wider path for conditionality in social policy (Szatur-Jaworska, 2023). This process rests (at least in principle) on three fundamental values: efficiency, integrity and transparency (Alston, 2019).

It may seem that digitalisation is an inevitable path. Welfare states must undergo this transformation, otherwise they will not exist. The question, then, is why does the development of digital surveillance meet with such resistance in society? It is associated with 'disciplining', 'punishing' or 'criminalising' (Fenger & Simonse, 2024), however welfare surveillance is in fact nothing new as controls have been carried out for a long time. Henman and Martson (2008) argued that digitalisation re-configures the nature of surveillance. It concentrates the power and capacity of authorities to assert norms, monitor behaviour and enforce compliance. It means that the form of the relation between the citizen and the welfare state is changing, but not its content.

There are many other serious concerns about digital transformation, such as high costs, underdeveloped infrastructure and e-exclusion. From a historical perspective, however, these problems are only technical details. First of all, digital transformation should not be seen as a cost but as an investment that will bring tangible benefits in the future. As for e-exclusion, this is a general problem that needs to be solved, not just in the welfare sector. In addition, techno-pessimists often stress the problem of the 'dehumanisation' of social policy. Such an argument of 'digital rigidity' concerns the lack of empathy between the state and the citizen (Ranchordás, 2022). Paradoxically, this feature is seen by techno-optimists as the main advantage of digital surveillance. Computer algorithms and big data are culturally constructed as accurate, objective and true, and there is no room for subjective discretion. Simple compliance with the law, and no exceptions to the rules. If someone does not meet the eligibility criteria, the computer always says "No" (Henman, 2022).

References

Alston, P. (2019). Digital Technology, Social Protection and Human Rights: Report. In *GA /74/493* (Vol. 17564, Issue October). http://statements.unmeetings.org/media2/21999189/sr-extreme-poverty-ga-3rd-cttee-statement-f.pdf

Barr, N. (2020). *The Economics of the Welfare State* (6th ed.). Oxford University Press.

Braithwaite, V. (2020). Beyond the Bubble that is Robodebt: How Governments That Lose Integrity Threaten Democracy. *Australian Journal of Social Issues*, *55*(3), 242-259. https://doi.org/10.1002/ajs4.122

Dee, M. (2013). Welfare Surveillance, Income Management and New Paternalism in Australia. *Surveillance and Society*, *11*(3), 272-286. https://doi.org/10.24908/ss.v11i3.4540

Decision No H5 of 18 March 2010 concerning cooperation on combating fraud and error within the framework of Council Regulation (EC) No 883/2004 and Regulation (EC) No 987/2009 of the European Parliament and of the Council on the coordination of social security systems (*OJ* C 149/5).

Dencik, L., & Kaun, A. (2020). Datafication and the Welfare State. *Global Perspectives*, *1*(1), article 12912. https://doi.org/10.1525/gp.2020.12912

Eerste Kamer. (2012). Wet aanscherping handhaving en sanctiebeleid SZW-wetgeving (33 207). Eerste Kamer der Staten-Generaal. https://www.eerstekamer.nl/wetsvoorstel/33207_wet_aanscherping_handhaving

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. Polity Press.

Fenger, M., & Simonse, R. (2024). The Implosion of the Dutch Surveillance Welfare State. *Social Policy and Administration*, *58*(2), 264-276. https://doi.org/10.1111/spol.12998

Fugletveit, R., & Sørhaug, C. (2023). Lost in Digital Translations: Studies of Digital Resistance and Accommodation to the Welfare State in Practice. In R. Fugletveit & C. Sørhaug (Eds.), *Lost in Digital Translations: Studies of Digital Resistance and Accommodation to the Welfare State in Practice* (pp. 7-9). Cappelen Damm Akademisk. https://doi.org/https://doi.org/10.23865/noasp.196.ch0

Halla, M., Lackner, M., & Schneider, F. G. (2010). An Empirical Analysis of the Dynamics of the Welfare State: The Case of Benefit Morale. *Kyklos*, *63*(1), 55-74. https://doi.org/10.1111/j.1467-6435.2010.00460.x

Heinemann, F. (2008). Is the Welfare State Self-Destructive? A Study of Government Benefit Morale. *Kyklos*, *61*(2), 237-257. https://doi.org/10.1111/j.1467-6435.2008.00400.x

Henman, P. W. F. (2022). Digital Social Policy: Past, Present, Future. *Journal of Social Policy*, *51*(3), 535-550. https://doi.org/10.1017/S0047279422000162

Henman, P., & Marston, G. (2008). The Social Division of Welfare Surveillance. *Journal of Social Policy*, *37*(2), 187-205. https://doi.org/10.1017/S0047279407001705

Horizon Europe. (2033). *Welfare Surveillance: Digital Seams in the Social Safety Net. Project Description*. https://cordis.europa.eu/project/id/101130820

Jorens, Y., Coninck, M. De, Wispelaere, F. De, Smedt, L. De, & Pacolet, J. (2019). *Fraud and Error in the Field of EU social Security Coordination*. Publications Office of the European Union.

Jurek, Ł. (2024). Nadużycia socjalne w Polsce: perspektywa badawcza i implikacje praktyczne. *Polityka Społeczna*, *3*, 29-37. https://doi.org/10.5604/01.3001.0054.6277

Korsell, L., Hagstedt, J., & Skinnari, J. (2008). Från kelgrisar till styvbarn – Fusket med välfärdssystemen. *Nordisk Tidsskrift for Kriminalvidenskab*, *95*(1), 21-38. https://doi.org/10.7146/ntfk.v95i1.71704

Kukuła, Z. (2016). *Przestępczość socjalna z perspektywy prawa karnego i kryminologii*. Difin.

McKeever, G. (2012). Social Citizenship and Social Security Fraud in the UK and Australia, *Social Policy and Administration*, *46*(4), 465-482. https://doi.org/10.1111/j.1467-9515.2012.00844.x

Modin, M., & Lindblom, E. (2021). *Rapport: Underlag inför 2022 års omfattningsstudier*. Ekonomistyrningsverket (ESV). https://www.esv.se/contentassets/f88f15625bad43ed92227a65d3c651d0/2021-37-underlag-infor-2022-ars-omfattningsstudier.pdf

OECD. (2014). Recommendation of the Council on Digital Government Strategies. OECD/LEGAL/0406. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406

OECD. (2016). Digital Government Strategies for Transforming Public Services in the Welfare Areas. In *OECD Comparative Study*. http://www.oecd.org/gov/digital-government/Digital-Government-Strategies-Welfare-Service.pdf

Quadagno, J. (1987). Theories of the Welfare State. *Annual Review of Sociology*, *13*(1), 109-128. https://doi.org/10.1146/annurev.so.13.080187.000545

Rachovitsa, A., & Johann, N. (2022). The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case. *Human Rights Law Review*, *22*(2), 1-15. https://doi.org/10.1093/hrlr/ngac010

Ranchordás, S. (2022). Empathy in the Digital Administrative State. *Duke Law Journal*, *71*(6), 1341-1389. https://doi.org/10.2139/ssrn.3946487

Rinta-Kahila, T., Someh, I., Gillespie, N., Indulska, M., & Gregor, S. (2022). Algorithmic decision-making and system destructiveness: A case of automatic debt recovery. *European Journal of Information Systems*, *31*(3), 313-338. https://doi.org/10.1080/0960085X.2021.1960905

Rinta-Kahila, T., Someh, I., Gillespie, N., Indulska, M., & Gregor, S. (2024). Managing Unintended Consequences of Algorithmic Decision-Making: The Case of Robodebt. *Journal of Information Technology Teaching Cases*, *14*(1), 165-171. https://doi.org/10.1177/20438869231165538

Sveriges Riksdag. (2021). Förordning (2021:663) om arbetet med att säkerställa korrekta utbetalningar från välfärdssystemen. https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/forordning-2021663-om-arbetet-med-att_sfs-2021-663/

Sveriges Riksdag. (2024). Kommittédirektiv: Ökad digitalisering på socialförsäkringsområdet för att motverka bidragsbrott och förebygga felaktiga utbetalningar. https://www.riksdagen.se/sv/dokument-och-lagar/dokument/kommittedirektiv/okad-digitalisering-pa-socialforsakringsomradet_hcb159/

Szatur-Jaworska, B. (2023). Polityka społeczna a cyfryzacja. *Polityka Społeczna*, *592*(8), 1-8.

Van Bekkum, M., & Borgesius, F. Z. (2021). Digital Welfare Fraud Detection and the Dutch SyRI Judgment. *European Journal of Social Security*, *23*(4), 323-340. https://doi.org/10.1177/13882627211031257

Van Gerven, M. (2022). Studying Social Policy in the Digital Age. In K. Nelson, R. Nieuwenhuis, M. Yerkes (Eds.), *Social Policy in Changing European Societies: Research Agendas for the 21st Century* (pp. 251-264). https://doi.org/10.4337/9781802201710.00025

Well, L., Currie, M., & Stewart, J. (2023). Surveillance, Discretion and Governance in Automated Welfare: the Case of the German ALLEGRO System. *Science & Technology Studies*, *36*(1), 42-58. https://doi.org/10.23987/sts.100490

Wieringa, M. (2023). Hey SyRI, Tell Me About Algorithmic Accountability: Lessons from a Landmark Case. *Data and Policy*, *5*, e2. https://doi.org/10.1017/dap.2022.39

Zajko, M. (2023). Automated Government Benefits and Welfare Surveillance. *Surveillance and Society*, *21*(3), 246-258. https://doi.org/10.24908/ss.v21i3.16107