

**USING AI TO MANAGE
MINIMUM INCOME BENEFITS
AND UNEMPLOYMENT
ASSISTANCE: OPPORTUNITIES,
RISKS AND POSSIBLE POLICY
DIRECTIONS**

**OECD ARTIFICIAL
INTELLIGENCE PAPERS**

June 2024 **No. 21**

OECD Artificial Intelligence Papers



OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to els.contact@oecd.org.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Cover image: © Kjpgargeter/Shutterstock.com

© OECD 2024

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at www.oecd.org/termsandconditions.

Using AI to manage minimum income benefits and unemployment assistance: Opportunities, risks and possible policy directions

Annelore Verhagen

While means-tested benefits such as minimum income benefits (MIB) and unemployment assistance (UA) are an essential safety net for low-income people and the unemployed, incomplete take-up is the rule rather than the exception. Building on desk research, open-ended surveys and semi-structured interviews, this paper investigates the opportunities and risks of using artificial intelligence (AI) for managing these means-tested benefits. This ranges from providing information to individuals, through determining eligibility based on pre-determined statutory criteria and identifying undue payments, to notifying individuals about their eligibility status. One of the key opportunities of using AI for these purposes is that this may improve the timeliness and take-up of MIB and UA. However, it may also lead to systematically biased eligibility assessments or increase inequalities, amongst others. Finally, the paper explores potential policy directions to help countries seize AI's opportunities while addressing its risks, when using it for MIB or UA management.

Keywords: Artificial Intelligence; Social Protection; Means-Tested Benefits; Minimum Income Benefits; Unemployment Assistance.

JEL codes: C8, H53, I3, J68, O3.

Résumé

Bien que les prestations soumises aux conditions de ressources, telles que les prestations de revenu minimum (PRM) et l'allocation chômage (AC), constituent un filet de sécurité essentiel pour les personnes à faible revenu et les chômeurs, leur utilisation incomplète est la règle plutôt que l'exception. En s'appuyant sur une revue de la littérature, des enquêtes ouvertes et des entretiens semi-structurés, ce papier étudie les opportunités et les risques liés à l'utilisation de l'intelligence artificielle (IA) pour la gestion de ces prestations soumises aux conditions des ressources. Cela couvre la fourniture d'informations aux individus, la détermination de l'éligibilité basée sur des critères statutaires prédéterminés et l'identification des paiements indus, et la notification des individus de leur statut d'éligibilité. L'une des principales opportunités de l'utilisation de l'IA à ces fins est qu'elle pourrait améliorer la rapidité et de l'utilisation des PRM et de l'AC. Cependant, cela pourrait également entraîner, entre autres, des évaluations d'éligibilité systématiquement biaisées ou accroître les inégalités. Enfin, le papier explore les politiques publiques potentielles pour aider les pays à saisir les opportunités de l'IA tout en tenant compte de ses risques, lorsqu'elle est utilisée pour la gestion des PRM ou de l'AC.

Abstract

Bedarfsabhängige Beihilfen wie Mindesteinkommensleistungen (MIL) und Arbeitslosenunterstützung (AU) bieten wichtigen sozialen Schutz für Menschen mit niedrigem Einkommen und Arbeitssuchende. Dennoch werden diese Leistungen oft nur unvollständig in Anspruch genommen. Die vorliegende Studie untersucht anhand von Sekundärliteratur, offenen Umfragen und teilstrukturierten Interviews die Chancen und Risiken des Einsatzes von künstlicher Intelligenz (KI) in der Verwaltung dieser bedarfsabhängigen Leistungen. KI könnte die rechtzeitige Inanspruchnahme von MIL und AU verbessern, etwa durch bessere Informationsvermittlung und Benachrichtigung von Einzelpersonen, einer automatisierten Prüfung der Anspruchsberechtigung, oder der Erkennung unrechtmäßiger Zahlungen. Unter den Risiken der Technologie sind potentiell verzerrte oder falsche Anspruchsbewertungen, datenschutzrechtliche Bedenken, oder eine verstärkte digitalen Kluft. Schließlich beschreibt die Studie Leitlinien, um die Möglichkeiten von KI für die Verwaltung von MIL- oder AU-Management zu nutzen und Risiken zu minimieren.

Acknowledgements

This publication contributes to the OECD's Artificial Intelligence in Work, Innovation, Productivity and Skills (AI-WIPS) programme, which provides policy makers with new evidence and analysis to keep abreast of the fast-evolving changes in AI capabilities and diffusion and their implications for the world of work. The programme aims to help ensure that adoption of AI in the world of work is effective, beneficial to all, people-centred and accepted by the population at large. AI-WIPS is supported by the German Federal Ministry of Labour and Social Affairs (BMAS) and will complement the work of the German AI Observatory in the Ministry's Policy Lab Digital, Work & Society. For more information, visit <https://oecd.ai/workinnovation-productivity-skills> and <https://denkfabrik-bmas.de/>.



This paper would not have been possible without the input from country experts from Australia, Canada, Costa Rica, the Czech Republic (Czechia), Denmark, Estonia, Finland, the Netherlands and New Zealand on the use of AI in social protection.

Special thanks go to Angelica Salvi Del Pero for her guidance and valuable feedback throughout the project. The report also benefitted from helpful comments and insights provided by colleagues from the Directorate for Employment Labour and Social Affairs (Monika Queisser, Valerie Frey, Herwig Immervol, Theodora Xenogiani, Dorothy Adams, Raphaëla Hyee, Anne Lauringson, Annikka Lemmens, Ailbhe Brioscú, and Diego Eslava), and the Public Governance Directorate (Andrea Uhrhammer, Carlos Santiso, Miguel Amaral, and Helene Wells). The work was carried out under the supervision of Stefano Scarpetta (Director of Employment, Labour and Social Affairs).

Table of contents

Résumé	4
Abstract	5
Acknowledgements	6
Executive summary	9
Synthèse	10
Zusammenfassung	12
Glossary	14
1 Introduction	17
2 Technical possibilities and use cases	20
2.1. Providing information and support to clients	21
2.2. Filling out and processing client information	23
2.3. Assessing eligibility and undue payments	25
2.4. Notifying individuals about eligibility decisions	29
3 Opportunities	30
3.1. Decreased administrative burden	30
3.2. Efficient benefit allocation and related social spending	30
3.3. Improved accuracy and fairness of decision-making	31
3.4. Improved client experience and trust	32
3.5. Increased take-up	32
4 Risks	33
4.1. Incorrect or biased eligibility assessments	33
4.2. Privacy concerns	34
4.3. Increased inequalities and decreased take-up among certain groups	35
4.4. Lack of transparency and explainability	36
4.5. Risks for accountability	37

5 Seizing opportunities, addressing risks: Possible policy directions	38
5.1. Human intervention and oversight	38
5.2. Strategic inclusiveness	39
5.3. Ensuring accountability early on	39
5.4. Transparency and explainability	40
5.5. Data quality and data governance	40
5.6. Capacity to understand, develop, and use AI	40
5.7. Proactive engagement by AI stakeholders	41
5.8. Flexibility and experimentation	41
6 Conclusions	43
References	44
Annex A. Country responses to the questionnaire	53
FIGURE	
Figure 4.1. Households with internet access at home, by income group	36
INFOGRAPHIC	
Infographic 2.1. The technical possibilities of using AI throughout the process of managing means-tested benefits	20
TABLES	
Table 1. Working-age benefits by entitlement criterion	15
Table 1.1. Working-age benefits by entitlement criterion	18
Table A A.1. Countries' use of digital tools such as AI and/or administrative data for managing minimum income benefits, unemployment benefits, or other income support benefits	53

Executive summary

Despite the vital role of means-tested benefit programmes in protecting individuals and families from economic uncertainties and poverty, many OECD countries struggle with low take-up rates, particularly for minimum income benefits (MIB) and unemployment assistance (UA). Building on desk research, open-ended surveys and semi-structured interviews, this paper explores how AI can enhance the efficiency and uptake of MIB and UA while ensuring that it is used in a trustworthy way. Key findings include:

- **AI can be used throughout the management process of MIB and UA.** Common use cases include AI-powered chatbots for information dissemination and analysing queries to improve support quality. By leveraging its capability to swiftly handle vast amounts of data and continuously adjust its outputs, AI could also simulate potential eligibility and benefit levels for individuals based on pre-defined statutory criteria. However, there are only few of this type of AI use cases, and the ones that do use AI mostly focus on detecting undue payments rather than automatically enrolling eligible individuals. Furthermore, AI's predictive capabilities extend to identifying individuals who may require benefits in the future, enabling a preventative approach to benefit provision.
- **When used for MIB or UA management, AI has the potential to automate administrative tasks, improve decision-making speed and accuracy, enhance client experiences, and – importantly – improve take-up.** It may thereby complement and improve upon existing digital tools for benefit management. AI may reduce the burden on public servants involved in managing MIB and UA, allowing them to focus on more complex cases. Moreover, AI could enable efficient benefit allocation and timely adjustments, minimising the severity and duration of financial hardships and preventing undue payments. Through data-driven eligibility assessments, AI could also promote fairness and accuracy in decision-making, and its use could improve client experiences through 24/7 support and simplified application processes. Most importantly, using AI for MIB or UA management may improve the take-up of these benefits among eligible individuals by improving information accessibility, reducing access barriers, and mitigating social stigma.
- **Poorly designed or implemented AI for MIB or UA management can lead to biased assessments, privacy concerns, exacerbate inequalities, and there are risks of a lack of transparency, explainability and accountability.** Biased assessments can result in individuals being wrongfully denied benefits or wrongly allocating benefits to those who do not need them, perpetuating inequalities and potentially violating human rights. Privacy concerns arise from the extensive personal data processed by many AI systems, increasing the risk of misuse or unauthorised access. Furthermore, difficulties accessing digital services may exacerbate inequalities and decrease take-up among vulnerable groups. Lack of transparency in AI decision-making processes makes it challenging for individuals to understand outcomes or seek redress for adverse effects. The complexity of accountability mechanisms further complicates the identification of responsible parties in case of errors or harm caused by AI systems.

Balancing the opportunities and risks of AI use for MIB and UA management requires thoughtful consideration and appropriate safeguards to ensure trustworthy use of AI. Countries must carefully consider whether AI is the optimal solution and ensure trustworthy implementation through maintaining human determination and oversight; a transparent consideration of the share of false positives and false negatives that are accepted by the system; regular risk assessments and audits; ensuring transparency and explainability and good data governance of AI systems used; and fostering capacity-building initiatives. Initiatives like the OECD AI Principles provide a global framework for trustworthy AI. Countries can take advantage of opportunities to collaborate internationally on AI approaches and standards.

Synthèse

Malgré le rôle vital des programmes de prestations sous condition de ressources pour protéger les individus et les familles des incertitudes économiques et de la pauvreté, de nombreux pays de l'OCDE sont confrontés à de faibles taux d'utilisation, en particulier pour les prestations de revenu minimum (PRM) et l'allocation chômage (AC). S'appuyant sur une revue de la littérature, des enquêtes ouvertes et des entretiens semi-structurés, ce document explore la manière dont l'IA peut améliorer l'efficacité et l'utilisation des PRM et d'AC, tout en garantissant qu'elle est utilisée de manière fiable. Les principales conclusions sont les suivantes:

- **L'IA peut être utilisée tout au long du processus de gestion des PRM et de l'AC.** Les cas d'utilisation courants incluent les chatbots alimentés par l'IA pour la diffusion d'informations et l'analyse des demandes afin d'améliorer la qualité de l'assistance. En exploitant sa capacité à traiter rapidement de grandes quantités de données et à ajuster ses résultats en permanence, l'IA pourrait également simuler l'éligibilité des individus pour des prestations sur la base de critères statutaires prédéfinis. Toutefois, ce type de cas d'utilisation de l'IA est très peu répandu, et ceux qui utilisent l'IA se concentrent principalement sur la détection des paiements indus plutôt que sur l'inscription automatique des personnes éligibles. En outre, les capacités de prédiction de l'IA s'étendent à l'identification des personnes susceptibles d'avoir besoin de prestations à l'avenir, ce qui permet une approche préventive de la distribution des prestations.
- **Lorsqu'elle est utilisée pour la gestion des PRM ou de l'AC, l'IA a le potentiel d'automatiser les tâches administratives, d'améliorer la rapidité et la précision de la prise de décision, d'améliorer l'expérience des clients et – surtout – d'augmenter le taux d'utilisation.** Elle peut ainsi compléter et améliorer les outils numériques de gestion des prestations déjà disponibles. L'IA pourrait réduire la charge de travail des fonctionnaires chargés de la gestion des PRM et de l'AC, leur permettant ainsi de se concentrer sur des cas plus complexes. En outre, l'IA pourrait permettre une attribution efficace des prestations et des ajustements en temps opportun, réduisant ainsi la gravité et la durée des difficultés financières et évitant les paiements indus. Grâce à des évaluations de l'éligibilité fondées sur des données, l'IA pourrait également promouvoir l'équité et la précision des décisions, et son utilisation pourrait améliorer l'expérience des clients grâce à une assistance 24 heures sur 24 et 7 jours sur 7 et à des processus de demande simplifiés. Avant tout, l'utilisation de l'IA pour la gestion des PRM ou de l'AC pourrait améliorer l'utilisation de ces prestations par les personnes éligibles en améliorant l'accessibilité de l'information, en réduisant les obstacles à l'accès et en atténuant la stigmatisation sociale.
- **Une IA mal conçue ou mal mise en œuvre pour la gestion des PRM ou de l'AC peut entraîner des évaluations biaisées, à des problèmes de protection de la vie privée, à l'exacerbation des inégalités, ainsi qu'à un risque de manque de transparence, d'explicabilité et de responsabilité.** Les évaluations biaisées peuvent se traduire par le refus injustifié d'accorder des prestations ou par l'attribution injustifiée de prestations à des personnes qui n'en ont pas besoin, ce qui perpétue les inégalités et peut constituer une violation des droits de l'homme. Les problèmes de protection de la vie privée proviennent des vastes données personnelles traitées par de nombreux systèmes d'IA, ce qui accroît le risque d'utilisation abusive ou d'accès non autorisé. En outre, les difficultés d'accès aux services numériques peuvent exacerber les inégalités et réduire l'utilisation de ces services par les groupes vulnérables. Le manque de transparence des processus décisionnels de l'IA fait qu'il est difficile pour les individus de comprendre les résultats ou de demander des recours en cas d'effets négatifs. La complexité des mécanismes de

responsabilité complique encore l'identification des parties responsables en cas d'erreurs ou de dommages causés par les systèmes d'IA.

Établir un équilibre entre les possibilités et les risques liés à l'utilisation de l'IA pour la gestion des PRM et de l'AC nécessite une réflexion approfondie et des protections appropriées afin de garantir une utilisation fiable de l'IA. Les pays doivent examiner attentivement si l'IA est la solution optimale et garantir une mise en œuvre fiable en préservant la détermination et la surveillance humaines, en examinant de manière transparente la part de faux positifs et de faux négatifs acceptés par le système, en procédant régulièrement à des évaluations des risques et à des audits, en garantissant la transparence et l'explicabilité ainsi qu'une bonne gouvernance des données des systèmes d'IA utilisés, et en encourageant les initiatives de renforcement des capacités. Des initiatives telles que les principes de l'IA de l'OCDE fournissent un cadre mondial pour une IA digne de confiance. Les pays peuvent saisir les occasions de collaborer au niveau international sur les approches et les normes en matière d'IA.

Zusammenfassung

Bedarfsabhängige Beihilfen spielen eine wichtige Rolle für den sozialen Schutz von Einzelpersonen und Familien vor wirtschaftlichen Unsicherheiten und Armut. Dennoch haben viele OECD-Länder mit niedrigen Inanspruchnahmequoten zu kämpfen, insbesondere bei Mindesteinkommensleistungen (MIL) und Arbeitslosenunterstützung (AU). Diese Studie untersucht anhand von Sekundärforschung, offenen Umfragen und teilstrukturierten Interviews, wie der Einsatz von KI auf zuverlässige Weise die Effizienz und Inanspruchnahme von MIB und UA verbessern könnte. Die wichtigsten Ergebnisse sind folgende:

- **KI kann während des gesamten Verwaltungsprozesses von MIL und AU eingesetzt werden.** Gängige Anwendungsfälle sind KI-gestützte Chatbots für die Verbreitung von Informationen und die Analyse von Anfragen zur Verbesserung der Supportqualität. KI hat die Fähigkeit, große Datenmengen schnell zu verarbeiten und Auswertungen kontinuierlich anzupassen. Die Technologie könnte dazu genutzt werden, anspruchsberechtigte Personen und deren Leistungshöhe auf der Grundlage vorgegebener gesetzlicher Kriterien automatisch zu erfassen. Es gibt allerdings bisher nur wenige Beispiele solcher KI-Anwendungen, die sich meist darauf beschränken, unrechtmäßige Zahlungen zu erkennen. KI hat das Potenzial, Personen zu identifizieren, die in der Zukunft Leistungen benötigen könnten, was einen präventiven Ansatz für Sozialleistungen ermöglichen würde.
- **In der Verwaltung von MIL und AU hat KI das Potenzial, administrative Abläufe zu automatisieren, eine schnellere und genauere Entscheidungsfindung zu unterstützen, das Kundenerlebnis zu verbessern und – besonders wichtig – die Inanspruchnahme zu erhöhen.** Auf diese Weise kann KI die bestehenden digitalen Tools für die Verwaltung von Leistungen ergänzen und verbessern. KI könnte die Arbeitsbelastung senken, so dass sich Beschäftigte in den Behörden auf komplexere Fälle konzentrieren können. Darüber hinaus könnte KI eine effiziente Leistungszuteilung und fristgerechte Anpassungen ermöglichen, um soziale Härtefälle schneller abzumildern und unrechtmäßige Zahlungen zu verhindern. Durch datengestützte Bewertungen der Anspruchsberechtigung könnte KI auch Gleichbehandlung und Genauigkeit in der Entscheidungsfindung unterstützen, und die Erfahrungen der Kunden durch einen 24/7-Support und vereinfachte Antragsverfahren verbessern. Am Wichtigsten ist jedoch, dass der Einsatz von KI in der Verwaltung von MIL oder AU die Inanspruchnahme dieser Leistungen erhöhen könnte, indem Informationen besser zugänglich gemacht, Zugangsbarrieren abgebaut und soziale Stigmatisierung gemildert werden.
- **Mangelhaft konzipierte oder umgesetzte KI-Anwendungen für die MIL- oder AU-Verwaltung bergen das Risiko voreingenommener Beurteilungen, datenschutzrechtlicher Probleme und einer verstärkten digitalen Kluft.** Bedenken bestehen zudem hinsichtlich eines Mangels an Transparenz, Erklärbarkeit und Rechenschaftspflicht. Voreingenommene Bewertungen könnten dazu führen, dass Einzelpersonen zu Unrecht Leistungen verweigert werden oder dass Leistungen falsch zugewiesen werden. Dadurch könnten Ungleichbehandlungen fortbestehen und möglicherweise die Menschenrechte verletzt werden. KI-Systeme verarbeiten oft umfangreiche personenbezogene Daten, was Bedenken hinsichtlich des Schutzes der Privatsphäre und Risiken des Missbrauchs oder unbefugten Zugriffs mit sich bringt. Ein ungleicher Zugang zu digitalen Diensten kann zudem Ungleichheiten verschärfen und die Inanspruchnahme von Leistungen durch benachteiligte Gruppen verringern. Mangelnde Transparenz von KI-unterstützten Entscheidungsprozessen macht es für Einzelne schwierig, die Ergebnisse zu verstehen oder Rechtsmittel gegen nachteilige Auswirkungen einzulegen. Komplexere Verantwortlichkeiten

würden es zusätzlich schwieriger machen, verantwortliche Parteien im Falle von Fehlern oder Schäden, die durch KI-Systeme verursacht werden, zu identifizieren.

Die Chancen und Risiken des Einsatzes von KI in der Verwaltung von bedarfsabhängigen Sozialleistungen erfordern sorgfältige Überlegungen und angemessene Schutzmaßnahmen, um einen zuverlässigen Einsatz von KI zu gewährleisten. Behörden müssen abwägen, ob KI die optimale Lösung ist, und eine vertrauenswürdige Umsetzung sicherstellen können. Dazu gehört, menschliche Aufsicht und Entscheidungsgewalt beizubehalten, die Anteile falsch positiver und falsch negativer Ergebnisse transparent zu überprüfen, regelmäßige Risikobewertungen und Kontrollen durchzuführen, die Transparenz und Erklärbarkeit sowie eine gute Datenverwaltung der verwendeten KI-Systeme zu gewährleisten, und Initiativen zum Aufbau von Kapazitäten zu fördern. Initiativen wie die OECD AI Principles bieten einen globalen Rahmen für die internationale Zusammenarbeit an Ansätzen und Standards für vertrauenswürdige und zuverlässige KI.

Glossary

Artificial Intelligence

The AI Group of Experts at the OECD has defined AI systems as “a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.”

Additionally, the Explanatory Memorandum specifies that topics typically encompassed by the term “AI” and in the definition of an AI system include categories of techniques such as machine learning and knowledge-based approaches, and application areas such as computer vision, natural language processing, speech recognition, intelligent decision support systems, intelligent robotic systems, as well as the novel application of these tools to various domains. AI technologies are developing at a rapid pace and additional techniques and applications will likely emerge in the future. The OECD definition aims to be flexible by reflecting a broad understanding of AI, and actors using this definition are encouraged to exercise judgement on its relevant scope, depending on the context it is being used in (OECD, 2024^[1]).

Outputs of AI systems generally reflect different tasks or functions performed by AI systems. They include, but are not limited to (OECD, 2022^[2]):

- recognition (identifying and categorising data, e.g. image, video, audio and text, into specific classifications as well as image segmentation and object detection);
- event detection (connecting data points to detect patterns, as well as outliers or anomalies);
- forecasting (using past and existing behaviours to predict future outcomes);
- personalisation (developing a profile of an individual and learning and adapting its output to that individual over time);
- interaction support (interpreting and creating content to power conversational and other interactions between machines and humans, possibly involving multiple media such as voice text and images);
- goal-driven optimisation (finding the optimal solution to a problem for a cost function or pre-defined goal);
- reasoning with knowledge structures (inferring new outcomes that are possible even if they are not present in existing data, through modelling and simulation).

Machine Learning

Machine learning (ML) is a set of techniques that allows machines to improve their performance and usually generate models in an automated manner through exposure to training data, which can help identify patterns and regularities rather than through explicit instructions from a human. The process of improving a system’s performance using machine learning techniques is known as “training” (OECD, 2024^[1]).

Natural Language Processing

Natural Language Processing (NLP) is a component of AI that enables machines to understand human language. By analysing the meaning of individual words, as well as the grammar that specifies the relationship between the words, NLP can extract meaning from large amounts of text and documents. When NLP is combined with ML techniques (also known as ‘statistical NLP’), it becomes possible to identify the most likely meaning of a sentence or phrase (Nadkarni, Ohno-Machado and Chapman, 2011^[3]). Common applications of NLP are automatic translations, grammar and spell check software, or automatically producing summaries of documents.

Social benefits

Social benefits to households are typically broken down into social transfers in kind and monetary social benefits. Transfers in kind are related to the provision of certain goods or services, meaning that households have no discretion over the use of these benefits. **Monetary social benefits** are typically transferred in cash and therefore allow households to use the cash indistinguishably from other income.

Means-tested monetary social benefits

Monetary social benefits may include benefits such as old-age benefits, child benefits, or other cash transfers to households to meet their financial needs in case of unexpected events, such as sickness, unemployment, housing, education, or family circumstances (OECD, 2024^[4]). Except for child benefits, which are typically universal, most monetary social benefits for working-age individuals are conditional on income (“means-tested benefits”) and/or on past contributions or past employment (see Table 1). This paper focuses on **means-tested benefits**. In particular, the focus lies on minimum income benefits and unemployment assistance.

Table 1. Working-age benefits by entitlement criterion

	Conditional on past contributions or past employment	Available irrespective of past contributions or past employment
Means-tested	E.g., Unemployment assistance (Austria)	Minimum income benefits (e.g. social assistance, housing benefit), but also some unemployment assistance programmes (Australia, Germany, UK)
Not means-tested	E.g., Unemployment insurance benefit, disability pensions, sickness benefit	Universal benefits, in practice: child benefits

Source: Based on Hye et al. (2020^[5]), How reliable are social safety nets?: Value and accessibility in situations of acute economic need, <https://doi.org/10.1787/65a269a3-en>.

Minimum income benefits

Minimum income benefits (MIB) are aimed to prevent extreme hardship and employ a low-income criterion as the central entitlement condition. These means-tested payments are typically received by households with no other income sources, although they can also top up incomes of low-paid workers and recipients of other benefits. Examples of MIB are social assistance benefits, means-tested lone-parent benefits, means-tested housing benefits, or any income support for jobseekers that is not conditional on past work contributions (Hye et al., 2020^[5]; Immervoll, 2010^[6]; OECD, 2024^[4]; OECD, 2024^[7]).

Unemployment assistance

Unemployment assistance (UA) consists of benefits to jobseekers that are usually means-tested. They thereby differ from unemployment insurance benefits, which are usually *not* means-tested. In some countries (e.g. Austria), UA is conditional on past contributions or past employment, whereas in others (e.g. Australia, Germany, United Kingdom), it is not. UA either provides a second-tier safety net for those who have exhausted their rights to unemployment insurance, or serve as a principal form of income support for jobseekers who were not entitled to unemployment insurance benefits in the first place (OECD, 2024^[7]).

Automated decision-making

Automated decision-making (ADM) can take many forms, ranging from supporting human decision-making through profiling and automatic recommendations (also known as “augmented intelligence” or “shared decision-making”) to fully automatic decision-making that does not require any human involvement (“full-ADM”) (European Commission, 2024^[8]; European Law Institute, 2022^[9]; ICO, 2024^[10]; US Congress, 2022^[11]).

1 Introduction

Despite the critical role of monetary social benefit programmes in shielding people from economic uncertainties and poverty, people do not always request or receive the benefits to which they are entitled – also known as “non-take-up” (OECD, 2024^[12]). The replacement or supplement of income provided by monetary social benefit programmes (see Box 1.1) protects individuals and families against economic and social risks such as ill health, old age, or job loss, and contributes to preventing and decreasing poverty. This makes them crucial for ensuring societal well-being and a sustainable economy (OECD, 2020^[13]).¹ The failure to provide benefits to those who are genuinely in need of and/or entitled to support increases the risk of poverty and exclusion, particularly for the poorest and most vulnerable individuals and families (Marc et al., 2022^[14]). Although non-take-up rates vary significantly across programmes and countries, incomplete take-up is the rule rather than the exception (Dubois and Ludwinek, 2015^[15]; Ko and Moffitt, 2022^[16]; Marc et al., 2022^[14]; OECD, 2018^[17]), and many individuals across OECD countries feel they cannot access public benefits easily in times of need (OECD, 2023^[18]).

Improving take-up of minimum income benefits (MIB) and unemployment assistance (UA) is of particular interest (see Box 1.1). While these means-tested benefits may not always succeed in fully lifting individuals out of poverty (Hyee et al., 2020^[5]; OECD, 2023^[19]), they are an essential safety net for low-income households and the unemployed and decrease financial strain on them (Almeida, De Poli and Hernández, 2022^[20]; OECD, 2019^[21]). Yet, take-up of MIB and UA is particularly low: on average across OECD countries, less than one-third of poor working-age households receive MIB (Hyee et al., 2020^[5]) or UA (OECD, 2018^[17]). In some countries (e.g. Latvia, Estonia, Poland), MIB take-up among the poor² is even as low as 5% (Almeida, De Poli and Hernández, 2022^[20]).

In countries with low levels of automation in the management of social benefit programmes, it can be difficult for caseworkers to assess and re-assess eligibility more than once every few months or even once per fiscal year, because eligibility criteria often require the verification of several data sources. As a result, some people continue to experience severe financial hardship while waiting for the receipt of the MIB or UA they have become eligible for, or they continue to receive undue payments – country-level data indicate substantial losses of funds due to undue payments in social benefit programmes (OECD, 2020^[13]). Adding to these challenges, social support systems tend to grow in complexity (Marc et al., 2022^[14]), increasing the workload of public servants. In light of these challenges, it is becoming imperative for countries to enhance not only the coverage but also the efficiency of their MIB and UA programmes. This paper investigates how artificial intelligence (AI – see Box 1.2) could help to improve take-up and efficiency of MIB and UA, and what are associated risks that would need to be addressed.

¹ All countries except those with very low incomes offer some kind of social benefits for lower income individuals and families (Ko and Moffitt, 2022^[16]).

² Below 40% of median equivalised disposable income (Almeida, De Poli and Hernández, 2022^[20]).

Box 1.1. Defining “minimum income benefits” and “unemployment assistance”

Social benefits to households are typically broken down into social transfers in kind and monetary social benefits. Transfers in kind are related to the provision of certain goods or services, meaning that households have no discretion over the use of these benefits. **Monetary social benefits** are typically transferred in cash and therefore allow households to use the cash indistinguishably from other income.

Monetary social benefits may include benefits such as old-age benefits, child benefits, or other cash transfers to households to meet their financial needs in case of unexpected events, such as sickness, unemployment, housing, education, or family circumstances. Except for child benefits, which are typically universal, most monetary social benefits for working-age individuals are conditional on income (means-tested benefits) and/or on past contributions or past employment (see Table 1.1). This paper focuses on **means-tested benefits**. In particular, the focus lies on minimum income benefits and unemployment assistance, which are an essential safety net for low-income households and the unemployed, but for which take-up rates are typically low.

Table 1.1. Working-age benefits by entitlement criterion

	Conditional on past contributions or past employment	Available irrespective of past contributions or past employment
Means-tested	E.g., Unemployment assistance (Austria)	Minimum income benefits (e.g. social assistance, housing benefit), but also some unemployment assistance programmes (Australia, Germany, United Kingdom)
Not means-tested	E.g., Unemployment insurance benefit, disability pensions, sickness benefit	Universal benefits, in practice: child benefits

Minimum income benefits (MIB) are aimed to prevent extreme hardship and employ a low-income criterion as the central entitlement condition. These means-tested payments are typically received by households with no other income sources, although they can also top up incomes of low-paid workers and recipients of other benefits. Examples of MIB are social assistance benefits, means-tested lone-parent benefits, means-tested housing benefits, or any income support for jobseekers that is not conditional on past work contributions.

Unemployment assistance (UA) consists of benefits to jobseekers that are usually means-tested. They thereby differ from unemployment insurance benefits, which are usually *not* means-tested. In some countries (e.g. Austria), UA is conditional on past contributions or past employment, whereas in others (e.g. Australia, Germany, United Kingdom), it is not. UA either provides a second-tier safety net for those who have exhausted their rights to unemployment insurance or serve as a principal form of income support for jobseekers who were not entitled to unemployment insurance benefits in the first place.

Source: Hye et al. (2020^[5]), How reliable are social safety nets?: Value and accessibility in situations of acute economic need, <https://doi.org/10.1787/65a269a3-en>; Immervoll (2010^[6]), Minimum-income benefits in OECD countries: Policy design, effectiveness and challenges, <https://doi.org/10.1787/218402763872>; OECD (2024^[4]), Social benefits to households, <https://doi.org/10.1787/423105c6-en>; OECD (2024^[7]), Tax-benefit web calculator, www.oecd.org/els/soc/benefits-and-wages/tax-benefit-web-calculator/#d.en.500997.

Countries are planning to increase their use of digital tools such as AI for managing means-tested benefits in the coming years (see Annex A). This is not surprising, considering that AI is recognised for its ability to handle complex tasks, which may help with determining benefit eligibility based on pre-defined statutory criteria; and AI-powered conversational agents such as chatbots, which are increasingly being used in customer services, may help to support individuals seeking information on specific means-tested benefits. However, there is a notable gap in understanding AI applications for means-tested social benefits like MIB and UA. Drawing on desk research, an open-ended survey on *Harnessing Technology and Data to Improve Social Protection Coverage and Social Assistance Delivery* among members of the OECD Working Party on Social Policy (hereafter: the “OECD Technology for Social Protection Questionnaire”) and semi-structured interviews with stakeholders engaged in AI use for MIB or UA, this study addresses this gap by first providing an exploration of the technical and practical possibilities of using AI for MIB and UA. Additionally, it seeks to assist policy makers in making informed decisions regarding AI adoption for MIB or UA, weighing the possibilities of the different types of AI tools and assessing associated opportunities and risks. Furthermore, the study aims to help policy makers establish appropriate safeguards – should AI be adopted for these means-tested benefits – to seize opportunities and address risks.

Box 1.2. Defining artificial intelligence

Artificial intelligence

Simply put, AI comprises of several tools or systems that take data and, using a statistical model, generates predictions, decisions, or recommendations. AI systems can improve their predictions and recommendations over time by updating and optimising the underlying model parameters. More precisely, the OECD (2024^[1]) defines an AI system as:

“a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.”

Additionally, the Explanatory Memorandum specifies that “*Topics typically encompassed by the term “AI” and in the definition of an AI system include categories of techniques such as machine learning and knowledge-based approaches, and application areas such as computer vision, natural language processing, speech recognition, intelligent decision support systems, intelligent robotic systems, as well as the novel application of these tools to various domains*” (OECD, 2024^[1]).

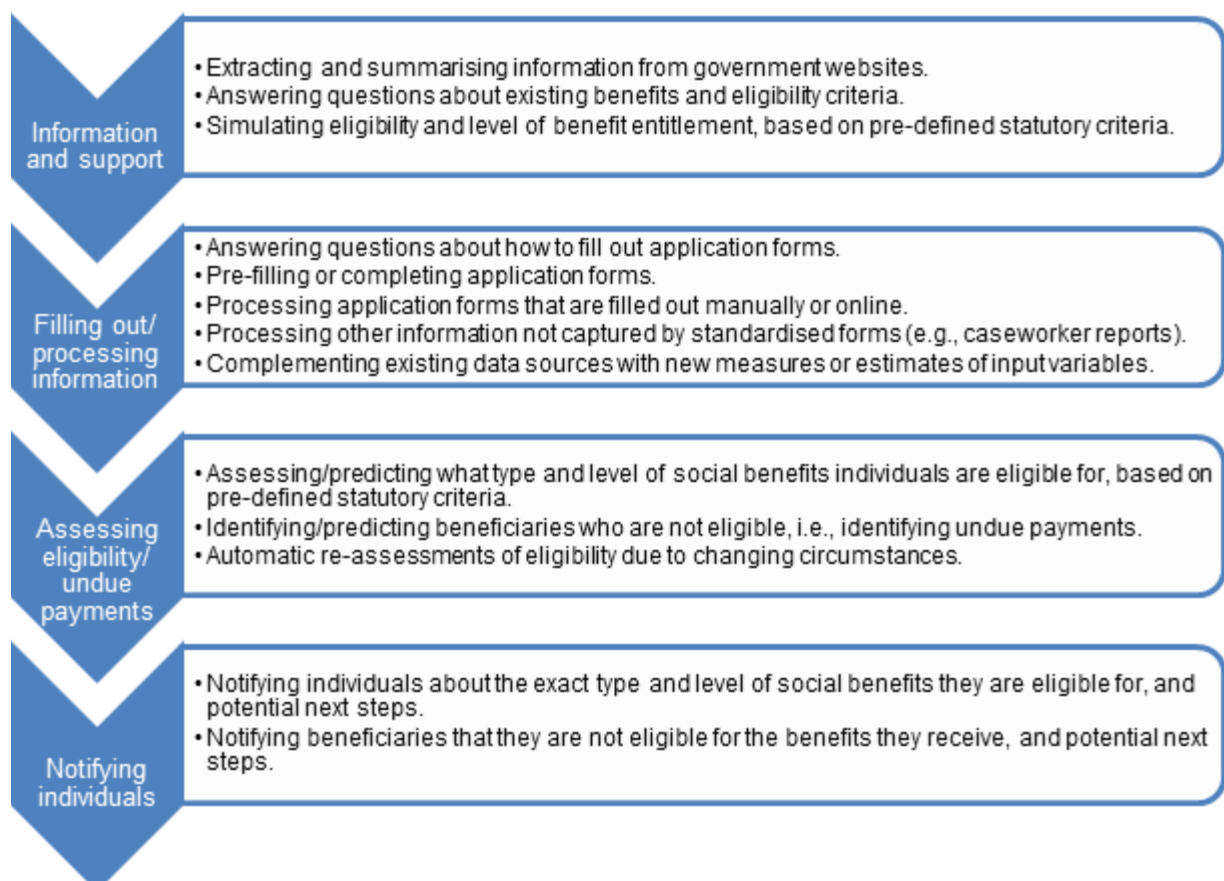
Source: OECD (2024^[1]), Explanatory memorandum on the updated OECD definition of an AI system, <https://doi.org/10.1787/623da898-en>; OECD (2023^[22]), *Future of Social Protection: Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery*, OECD Technology for Social Protection Questionnaire.

The paper continues as follows. Section 2 maps out the technical possibilities of using AI for managing MIB or UA and describes concrete examples of current or past use cases of AI for these benefits in various countries, including Australia, Austria, Belgium, Brazil, Canada, Colombia, Costa Rica, Czechia, Denmark, Estonia, Finland, Germany, the Netherlands, New Zealand, Sweden, and the United Kingdom. Section 3 analyses how these uses of AI can help to improve MIB and UA. However, using AI in this context also comes with specific risks – a topic discussed in detail in Section 4. Section 5 reflects on strategies that countries may consider to seize identified opportunities while addressing the risks, thereby ensuring that AI use for MIB and UA is trustworthy and beneficial for all. Section 6 concludes.

2 Technical possibilities and use cases

This section describes the technical possibilities of using AI to manage means-tested benefits such as minimum income benefits (MIB) and unemployment assistance (UA), and illustrates them with concrete examples of current or past use cases where possible. Infographic 2.1 shows the technical possibilities of using AI throughout the process of managing these means-tested benefits, by partially or fully automating certain tasks or services.

Infographic 2.1. The technical possibilities of using AI throughout the process of managing means-tested benefits



Source: OECD Secretariat.

This Section, which is organised according to the categories in Infographic 2.1, indeed shows that there are AI use cases throughout the benefit management process. Many AI use cases show that, when gathering information on existing social benefits and their eligibility criteria, AI can swiftly extract and summarise information from various government websites, and answer questions in any language, at any time of the day. This section also describes examples of how AI can help handle MIB and UA application forms by answering questions about the forms, pre-filling them, or processing hand-written forms and notes. Importantly, AI systems' ability to process and learn from vast amounts of data can help to determine what social benefits people are eligible for based on pre-defined statutory criteria, and swiftly allocate them, scale them up or down, or terminate them when necessary. To date, this type of AI use is mostly found to automate (parts of) the benefit termination process. When communicating benefit decisions, AI could automate the drafting of tailored letters, although there are not many examples in this regard.

While certain AI systems could manage (parts of) the process autonomously without the need for any human involvement, they are often not deployed that way, nor would that be desirable. This paper uncovers some examples where AI is used to automatically perform tasks or make decisions, but AI is more commonly used to provide recommendations to support human decision-making. Using AI in the decision-making process is also known as automated decision-making (ADM), which can take many forms, ranging from supporting human decision-making through profiling and automatic recommendations, to fully automated decision-making that does not require any human involvement ("full-ADM") (European Commission, 2024^[8]; European Law Institute, 2022^[9]; ICO, 2024^[10]; US Congress, 2022^[11]). While full-ADM (whether AI-powered or not) may significantly improve the efficiency of MIB and UA management, it raises the question whether decisions about monetary social benefits should be made without any human involvement, because these benefits have a significant impact on people's lives. Indeed, the OECD AI Principles call on AI actors to implement mechanisms and safeguards that ensure capacity for human intervention and oversight, to promote human-centred values and fairness in AI systems (OECD, 2019^[23]).

Yet, it should be noted that, while AI can demonstratively be leveraged at all stages of managing MIB and UA, this does not necessarily mean that the counterfactual is the absence of technology. In many OECD countries, microsimulation models and simple rule-based ADM systems, that some may not classify as AI, have been used for over 20 years (OECD, 2024^[12]). See Section 3 for a detailed discussion of the added value AI can bring to social protection.

2.1. Providing information and support to clients

Individuals need to be able to gather information on existing social benefits and their eligibility criteria. This enables them to assess which benefits they can apply for, and to assess whether they rightfully (do not) receive these benefits. Individuals can typically gather this information from government websites or by calling a dedicated call centre. However, information may be scattered or difficult to find and understand, for instance because of technical language used or because individuals do not master the national language. Moreover, call centres are often not available outside of business hours and may have (long) waiting times.

AI technologies could automatically summarise information and answer questions related to MIB and UA in real-time, at any time of the day. For instance, AI could extract information on social benefits from various government websites and generate summaries, which could be used to generate and automatically update explanatory notes about specific social benefits. Conversational agents such as chatbots and virtual assistants are increasingly well-known tools in government agencies, and the use of chatbots to communicate with potential programme users is currently the most common use of AI by social benefit and service agencies (OECD, 2024^[12]). For instance, 26% of respondents of the 2021 Gartner CIO survey, conducted in 166 government agencies across the world, indicate having deployed chatbots in their organisation and another 59% plan to implement them within three years (Gartner, 2021^[24]). Similarly,

Misuraca and Van Noordt (2020^[25]) find in their review of 230 AI-enabled public services across the EU that chatbots emerged as the first choice, accounting for over one-fifth of use cases.

Using a conversational agent to answer questions could free up time in call centres for questions that require more complex interactions. For instance, the Belgian National Employment Office (*Office national de l'emploi* – ONEM) set up a chatbot to ease the pressure on contact centres brought on by the unprecedented volumes following the COVID-19 crisis. The current version, called *Ori* – rolled out in December 2021 – can answer a range of questions relating to unemployment and career breaks, including where to find information on the level of unemployment benefits one may be eligible for (ISSA, 2022^[26]). Other examples of conversational agents that provide information are Digital Assistants used by Services Australia to provide real-time assistance to customers and staff with questions about online claims and social security payments (OECD, 2023^[22]), and the AI-powered chatbot OSC Caro of the Austrian social insurance agency (*Dachverband der Österreichischen Socialversicherungsträger* – SVA), which provides digital support to customers across different domains, including childcare allowances (ISSA, 2020^[27]). In 2020, the Canada Revenue Agency (CRA) started to use chatbots to provide non-confidential information on various topics including social benefits and limited programme information (see Box 2.1).

Box 2.1. AI-powered chatbots of Canada Revenue Agency

Canada Revenue Agency (CRA), the agency in charge of administering tax, benefits, and related programmes and ensuring compliance on behalf of governments across Canada, noticed that clients experienced challenges accessing information from their websites, leading them to abandon the digital portals, completing their task incorrectly, and/or resorting to telephone services to get assistance. To address these issues and optimise its public web presence, the CRA started to implement chatbots in 2020 that use Natural Language Processing (NLP) to interpret questions and information on CRA websites. These AI-powered chatbots – now available on 16 of the CRA webpages – help people navigate the complex information by redirecting them to the websites that contain the answer, or by helping them reach the right contact centre.

By design, the chatbot responds to non-account specific questions, and redirecting people to (human) contact centres remains important, because to ensure that the clients' data and privacy remain well protected, the chatbots currently do not access the clients' personal confidential data. As a result, the CRA did not experience a reduction of calls to the contact centres, but rather a change in the type of questions posed to the contact centres, with a reduction in relatively technical questions (e.g. how to find or fill out a certain form), and an increase in more complex questions that require access to the clients' personal file (e.g. which benefits they are eligible for).

Future iterations of the chatbots may include applications of generative AI, the availability of a chatbot in a secured space so that it can access the clients' confidential information, and a seamless transition between the AI-powered chatbot and the live chat with humans that is currently already available.

Source: CRA (CRA, 2024^[28]), Interview with Canada Revenue Agency, 25 January 2024; OECD (2023^[22]), Future of Social Protection: Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery, OECD Technology for Social Protection Questionnaire.

Natural Language Processing (NLP) techniques could produce easy-to-understand summaries and answers related to MIB or UA, and make them available in many languages. The Finnish Social Insurance Institution (Kela) uses the Kela-Kelpo/FPA-Folke chatbot which understands Finnish, Swedish, and English, and can reply in Finnish or Swedish. It assists clients in accessing information about various social benefits (including social assistance, parental benefits, and unemployment insurance) and helps discover, understand, and complete benefit applications – see also Section 2.2. If necessary, the bot directs

customers to other service channels, such as the call centre service. Between July and December 2022, the number of discussions increased by 67% compared to the corresponding period of the previous year, and 89% of the discussions were resolved (Kela, 2023^[29]).

AI could also monitor call volumes and analyse queries and calls to improve the availability and quality of client support and information provision about MIB and UA. For instance, based on callers' latent characteristics and previous customer behaviour, Machine Learning (ML) techniques can be used to predict future call arrival volumes and required staffing levels (Albrecht, Rausch and Derra, 2021^[30]). While potentially raising concerns regarding clients' privacy protection (see Sections 4.2 and 5.5), using AI to analyse the content of queries in conversational agents or calls can also make it possible to gather insights in recurring or trending questions, language use, or customer feedback, which can help to improve the quality of the interactions with call centres and conversational agents (Haas, McGuire Christian and Keuky, 2021^[31]). For instance, as part of its national digitisation strategy, Czechia's Ministry of Labour and Social Affairs is developing an AI system to monitor topics addressed in the call centre or other communication channels. By summarising the content of these interactions and identifying trends, training for call centre agents might be improved, amongst others (MoLSA, 2023^[32]; OECD, 2023^[22]). Finally, AI could help quickly redirect clients to the department or person best placed to answer their questions. For instance, the Austrian social insurance agency (SVA) uses an AI-powered voice recognition system that automatically forwards customer inquiries to the corresponding offices, based on the recognition of specific terms. Similarly, AI is also used to automatically dispatch emails to the corresponding departments (ISSA, 2020^[27]).

Leveraging their capability to handle large volumes of data, identify complex patterns, automatically extract data from databases, and adapt and learn from new data to improve their accuracy, AI models could be used to simulate potential eligibility and the level of MIB and UA individuals may be entitled to receive, based on pre-defined statutory eligibility criteria (see also Section 2.3). Through the analysis of personal characteristics, factors such as healthcare history, financial transactions, family composition or legal status could be taken into consideration to help people understand their potential eligibility for various benefits. In practice, however, many of these types of simulations do not rely on advanced AI technologies, but rather on more simple algorithms that may not be considered as AI. For instance, Canada's Old Age Security (OAS) Benefit Simulator asks questions about the client's age, net income, legal status, residence history, marital status, and spouse or common-law partner (if applicable) to provide estimates of the type and level of OAS benefits they are entitled to, including Old Age Security pension, Guaranteed Income Supplement and Allowance, and Allowance for the Survivor (Government of Canada, 2024^[33]). User research validated the value of the simulator that was launched in November 2022, finding an 85% success rate, in contrast to the 35% success rate for the existing OAS payment tables on Canada.ca (OECD, 2023^[22]).

2.2. Filling out and processing client information

Filling out benefit application forms can be complex and time-consuming, as can be the processing of these forms or other information needed to determine MIB or UA eligibility. For the applicant, filling out benefit application forms often requires providing detailed and comprehensive information. The documentation accompanying these forms can be complex, increasing the risk of errors and therefore unnecessary delays or even erroneous denial of MIB or UA. Processing these forms can also be time consuming and error-prone, particularly if it requires processing handwritten forms or additional open text notes provided by the applicant or in caseworker reports.

AI-powered conversational agents could provide guidance to facilitate the correct completion of social benefit application forms. Chatbots, for example, could answer questions about how to fill out specific forms and how to provide necessary documentation. Additionally, by incorporating information obtained

through these interactions or by accessing individuals' personal data such as tax administration records and employment history, the system could automatically pre-fill or complete benefit application forms (Pugh, 2023^[34]). In Finland, for example, the chatbot of the Finnish Social Insurance Institution (Kela), Kela-Kelpo/FPA-Folke, assists clients in completing benefit applications and offers personalised tips based on contextual variables during the application process (ISSA, 2022^[26]; Kela, 2023^[29]).

NLP techniques could be leveraged to extract information from unstructured data, such as notes from caseworker reports or applicant-filled forms (Fruy et al., 2022^[35]). Considering this could be done automatically, it may save significant amounts of time. For instance, Employment and Social Development Canada (ESDC) used NLP techniques to scan over 10 million open-text agent case notes within a few weeks, underscoring the scalability of this technique for handling large volumes of unstructured data (see Box 2.2).

Box 2.2. AI use to analyse free text caseworker reports in Canada

In Canada, recipients of the Old Age Security pension (OAS) can receive a Guaranteed Income Supplement (GIS) – a cash benefit targeting low-income old-age persons to ensure they reach a minimum income level after retirement. In January 2017, changes to the eligibility criteria for the GIS resulted in the loss of these benefits for certain individuals. Six months later, this policy change was reverted, and people who lost their GIS could request to receive these benefits again. Employment and Social Development Canada (ESDC) – the agency responsible for GIS administration – wanted to use a proactive approach to reach out to the affected individuals. However, the absence of records detailing the GIS benefits that were terminated required a meticulous review of over 10 million open-text agent case notes. Using a sample of manually assessed case notes, the ESDC leveraged NLP and ML techniques to identify the individuals that were eligible to receive the GIS again. To minimise the possibility that eligible individuals were overlooked, the system was intentionally designed to allow a higher number of false positives. Once the system identified all relevant cases (around 4 000), the list of affected clients was provided to Service Canada caseworkers for verification, leading to the reinstated benefits of about 2 000 clients that had been affected. The whole process of proactively identifying and reinstating GIS benefits was successfully completed within a few weeks.

Source: ESDC (2023^[36]), Interview with Employment and Social Development Canada, 24 October 2023; ISSA (2017^[37]), Using artificial intelligence (AI) to identify vulnerable Canadians, www.issa.int/gp/198044; ISSA (2020^[27]), Artificial Intelligence in Social Security: Background and Experiences, www.issa.int/analysis/artificial-intelligence-social-security-background-and-experiences; OECD (2023^[22]), Future of Social Protection: Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery, OECD Technology for Social Protection Questionnaire.

Using large amounts of data, ML techniques could generate estimates of input variables necessary for determining benefit eligibility or identifying undue payments, potentially complementing or augmenting existing data sources. This could be particularly useful in cases where existing data are of poor quality. For instance, using tax records as an input variable may not be reliable in countries where the poor population predominantly works in the informal sector. While non-AI technologies can already provide estimates of variables such as income based on other available data (proxy-means testing being a popular technique), this typically relies on pre-determined and static models that are not typically considered as AI. AI, and ML in particular, facilitates a more flexible data-driven approach and allows for the parameters to be automatically updated as new information becomes available.

AI systems could automatically cross-reference benefit application forms with various data sources, helping identify anomalies or irregularities in benefit applications. However, AI is again not the only technology that can perform this task. For instance, Estonia's Unemployment Insurance Fund (*Töötukassa*) uses automated decision-making to automatically verify in different databases whether an

applicant's filled out information is correct (Algorithm Watch, 2024^[38]; Esko, 2020^[39]; Nortal, 2022^[40]; Raudla, 2020^[41]; Töötukassa, 2024^[42]). Approximately 50% of all decisions regarding unemployment insurance benefits are fully automated using this system, which does not typically fall under the definition of AI, because the decision-making process is static and rule-based.³ Using AI could enable the use of more extensive information and inference for flagging cases that warrant further inspection. A notable example is found in Denmark, where the Data Mining Unit in the Public Benefits Administration (*ATP-Udbetaling Danmark*) employs ML and data analytics to detect cases suspected of social benefits fraud, that need verification by a human caseworker, who then reach out to the individuals, enabling them to rectify any inaccuracies (see also Box 2.4). Similarly, Germany's Federal Employment Agency is using ML to recognise and extract information from proof-of-study documents submitted by citizens as proof of ongoing eligibility for non-employment related child benefits (*Kindergeld*), and for which eligibility continues up to the age of 25, provided the child is undertaking education or training (Brioscú et al., 2024^[43]).

2.3. Assessing eligibility and undue payments

Determining who is eligible for MIB and UA according to pre-defined statutory criteria, and which beneficiaries may not be eligible and receive undue payments, is a crucial and time sensitive but often complex task. Any delay in the decision-making process implies that eligible individuals continue to experience severe financial hardship,⁴ and that ineligible beneficiaries continue to receive undue payments and accumulate debts. Additionally, people's personal situation and the eligibility criteria may change over time, requiring periodic re-assessments of eligibility. Yet, in countries with low levels of automation of benefit processing, determining MIB and UA eligibility could be a time-consuming task because eligibility criteria often depend on many variables, and application forms often need to be cross-referenced with various sources such as tax administration, employment records and housing registries. Even if each decision to allocate, adjust, or terminate MIB or UA would only take the caseworker a few minutes to complete, the cumulative workload could result in a significant administrative burden. Assessments or re-assessments of eligibility may therefore only be performed every few months or once per fiscal year.

While traditional statistical models could already help to swiftly determine social benefit eligibility or undue payments based on pre-defined rules and relatively simple variables, some AI systems offer enhanced capabilities in identifying complex patterns within large datasets. AI's ability to adapt may also improve the precision of (in)eligibility assessments (see Section 3.3). Using a representative sample of previous decisions regarding benefit reciprocity, combined with personal data from both recipients and non-recipients, AI systems may be able to predict the likelihood of eligibility for specific social benefits with greater accuracy and nuance. For instance, the Colombian Identification and Classification System of Potential Beneficiaries for Social Programmes (*Sisbén*) trained ML models based on data from individual surveys on living conditions (e.g. income information, or access to public services) to learn patterns and relationships between socio-economic factors and "prosperity" scores on a scale from 0 to 100. By applying these learned assessments to new applicants, public entities determine whether a person can access social benefits (OECD/CAF, 2022^[44]). The Canadian ESDC used ML models to predict the most probable outcome for reviews of employment insurance claims (increase or decrease benefit rate, or no change). Due to unprecedented volumes of employment insurance claims during the COVID-19 pandemic, a backlog of claims needing review arose. Using a training set of caseworker reports, the AI system was

³ The system checks the correctness of the data in the online benefit application form in various databases and decides if and for how long the person is entitled to compensation, and in what amount. After the decision is made, the system informs the applicant about it, accompanied by a notice stating that it is a decision based on automated processing of the request and the data in the files, and that they have the right to obtain explanations and to lodge a challenge.

⁴ Yet, while MIB and UA may decrease the severity or duration of financial hardship, they often do not succeed in fully lifting people out of poverty (Hyee et al., 2020^[5]; OECD, 2023^[19]).

able to swiftly identify claims requiring review by an officer (ESDC, 2023^[36]; ESDC, 2024^[45]; OECD, 2023^[22]).

In its efforts to increase efficiency and combat social benefits fraud, the United Kingdom Department for Work and Pensions, responsible for administering the State Pension and a range of working age, disability and ill-health benefits, developed an AI system to help process benefit claims and cross-reference them with data such as those gathered from credit-checking companies, the police, the Land Registry, or social media.⁵ The algorithms look for patterns in claims such as applications written in the same style. Once a claim is flagged as suspicious, a human investigator takes over to determine if the claim is in fact fraudulent (Digital Watch, 2019^[46]; DWP, 2018^[47]; Marr, 2018^[48]; Robinson, 2019^[49]). Another example is found in Brazil, where the National Social Security Institute (*Instituto Nacional do Seguro Social – INSS*) is using AI to speed up beneficiary death detection, amongst others by using data from digital death certificates that are now issued in real-time. The INSS turned to AI because it often erroneously continued payments to deceased beneficiaries, mainly because the civil registry officers were not able to report beneficiary deaths to the INSS in a timely manner. This led to significant financial losses, amounting to billions of Brazilian real (ISSA, 2020^[27]; ISSA, 2021^[50]). The Swedish Public Employment Services has developed a largescale fraud detection system, rolling out in early 2024. This system combines several techniques (machine learning, deep learning, social network analysis and knowledge graphs) and encompasses a wide pool of data, including information on jobseekers, employees and employers (i.e. labour market contracts data), and suppliers (partners or service providers of the Swedish PES) (Brioscú et al., 2024^[43]).

⁵ Using social media data may raise concerns regarding individuals privacy protection: see Sections 4.2 and 5.5.

Box 2.3. Algorithms to assess benefit eligibility using an inter-institutional information system in Costa Rica

In 2013, in accordance with Law 9137, Costa Rica developed the National Information System for the Single Registry of State Beneficiaries (SINIRUBE), a centralised inter-institutional information system that includes beneficiaries of all social programmes financed by the State. This system has an algorithm developed together with researchers from MIT that takes advantage of SINIRUBE to identify people and households in poverty along with the different types of social benefit programmes they are entitled to.

This is a significant gain in efficiency compared to the previous situation where individuals or households received visits from multiple social workers, one for each type of social assistance for which they were eligible. This represented a significant increase in efficiency and a reduction in the risk of asymmetries, as the assessment was based solely on the perception of each social worker. The system is also designed to improve ease of access for individuals or households requiring several types of assistance. This is because they now only have to upload their information into the system once, rather than for each type of assistance separately. For instance, using SINIRUBE, an active search process was conducted in 2022 and 2023, to identify possible beneficiaries for a temporary benefit for inflation, without anyone having to apply.

The Costa Rican Government is investigating and evaluating possibilities for further improving and updating the algorithm, and to increase the use of SINIRUBE for other purposes such as labour market matching of unemployed individuals.

Source: United Nations (2024^[51]), Comprehensive co-operation to Costa Rica's National Information System and Unique Registry of Beneficiaries (SINIRUBE), <https://sdgs.un.org/partnerships/comprehensive-co-operation-costa-ricas-national-information-system-and-unique-registry>; SCIJ (2013^[52]), Reglamento a la Ley N.º 9 137, Creación del Sistema Nacional de Información y Registro Único de Beneficiarios del Estado, www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=75585&nValor3=0&strTipM=TC; Instituto Mixto de Ayuda Social (2023^[53]), Informe final Beneficio Temporal por Inflación año 2022 y 2023, www.imas.go.cr/sites/default/files/docs/Informe_Final_Beneficio_Temporal_por_Inflacion_2022_y_2023.pdf.

As changes unfold in one of the data sources, AI systems are capable of swiftly re-evaluating eligibility or undue payments. If data are updated in real-time, AI systems could process this information directly and – importantly – recalibrate predictive models as soon as the change is registered in the data. For instance, AI systems could (re)assess eligibility for UA or adjustments in family benefits as soon as employment status or family composition are updated in the system. For instance, the New Zealand Government uses automated decision-making (not AI-powered) to support delivery of social welfare benefits, for example by adjusting how child support income affects benefit payments payable to clients (OECD, 2023^[22]). A Danish system to predict and prevent social benefits fraud is updated daily (see Box 2.4).

Box 2.4. Machine learning and data mining to predict social benefits fraud in Denmark

Until 2013, public servants in Denmark needed to log on to various government websites and computer systems to collect all the data needed to determine eligibility of benefit applicants. Additionally, they needed to re-assess all benefit recipients every three months to verify which citizens were still entitled to the benefits they receive. Between 2007 and 2013, the Danish welfare system gradually underwent a reorganisation, including the centralisation of the administration of social benefits into *ATP-Udbetaling Danmark*, and the digitisation and linking of various data registries. In 2014, the Danish Parliament passed a law granting ATP-Udbetaling Danmark the authority to consolidate the many agency data sources for monitoring purposes. This enabled the use of ML and data mining techniques (something many municipalities would not be able to invest in) and, consequently, the creation of the Data Mining Unit.

This Unit uses non-sensitive personal data from Danish authorities for their models, including data on people's personal situation (e.g. who they are married to, who they live with, number of children), housing situation (address, apartment size, number of rooms), tax data (whether they own a car, bank account details, assets, taxes), and employment data (including pay and hours worked). Using experience-based selection criteria and – specifically for family benefits – unsupervised ML for outlier detection, suspicious cases are identified and prioritised.

Notably, the system scans the data on a daily basis, so that any changes in the data can immediately be detected and re-evaluated. A digital robot subsequently generates a document with all the relevant information of the suspected individual and uploads it on a platform accessible by municipality caseworkers. They are obliged to reach out to individuals within 10 days to check if the suspicion may be due to a data error and allow the citizen to correct it or otherwise help the authorities correct their suspicion of fraud before the final decision is made. This way, suspicious cases cannot be determined based on data alone: they must be investigated in more detail by human caseworkers, and the citizen is always heard. Moreover, the extent of data that can be shared is assessed in relation to whether the data processing is proportional.

Source: Jacobsen (2020^[54]), Town Hall Meeting on Citizen Data in Focus: Machine Learning and Data Analytics, www.youtube.com/watch?v=ZZ-xK7Xp2k; Jacobsen (2020^[55]), Data-driven Fraud Detection: The Data Mining Unit (Den Fælles Dataenhed), <https://lisboncouncil.net/wp-content/uploads/2020/10/Annika-Jacobsen.pdf>; The Lisbon Council (2020^[56]), Data-Driven Public Service: The Road to Efficient and Trustworthy Government, <https://lisboncouncil.net/newsandevents/data-analytics-for-citizen-support-and-government-savings/>.

AI's predictive capabilities extend to identifying individuals who may require benefits in the future. AI could analyse patterns in vast datasets such as economic trends, employment patterns, and demographic shifts, to identify individuals who may require benefits in the future. Taking such a preventative approach may help to reduce (the level of) benefits needed by the individual. An illustrative case is the Amsterdam municipality in the Netherlands, which uses an algorithm to identify individuals at risk of poverty (see Box 2.5). Similarly, New Zealand's Youth Service has a risk-scoring algorithm that aims to predict which school leavers are at high risk of becoming long-term benefit recipients. The algorithm uses data to analyse factors such as how well the former student did at school, whether their parents received social benefits and if they were in contact with child protective services. Service providers then approach those deemed most at risk to offer a service (OECD, 2023^[22]).

Box 2.5. The use of algorithms to identify people at risk of poverty in Amsterdam

Under the Early Intervention Programme (*Vroeg eropaf*), the municipality of Amsterdam in the Netherlands uses an algorithm for early identification of people at risk of getting into debts. The idea is that, by intervening early, advice or small interventions can be enough to prevent people from getting into serious debt or being evicted from their homes. To do this, the municipality leverages RIS Matching data – an online registration and information system where creditors register payment arrears and other client information (e.g. address, gender, date of birth). Municipal civil servants can initiate the automated matching of these data with the municipality's records. Subsequently, the system automatically generates a report and assigns the individual to a specific care team (e.g. homeless, youth, entrepreneurs, back-to-work). A social worker in this team tries to reach out to the concerned citizen within 14 days to gather insight into their specific situation. Provided the citizen accepts assistance, the social worker arranges the necessary support and helps them formulate a detailed action plan. They register any agreements (or refusals to accept help) in the RIS Matching system, so that the creditors are informed and may decide to temporarily suspend their automatic debits. In order to ensure transparency, the information on the Early Intervention Programme is published on the websites of the municipality of Amsterdam and the government's Algorithm Register (see also Section 5.4).

Source: Gemeente Amsterdam (2020^[57]), Vroeg Eropaf (early intervention), <https://algoritmeregister.amsterdam.nl/ai-system/vroeg-eropaf/314/>; Inforing (Inforing, 2024^[58]), Regiesoftware voor vroegsignalering van financiële problemen (Direction software for early detection of financial problems), www.vroegsignalering.nl/nl/home; Overheid.nl – The Algorithm Register (2024^[59]), Early on, <https://algoritmes.overheid.nl/en/algoritme/66453169>.

2.4. Notifying individuals about eligibility decisions

Once it is determined who should and should not receive specific benefits or benefit amounts, this decision needs to be communicated to the relevant individuals. This could be a notification about the person's eligibility, advising them to apply for specific benefits themselves, or a notification that specific benefits will automatically be granted to them. For ineligible beneficiaries, it could be a notification that they are suspected of receiving undue payments, and/or the decision to terminate specific benefits, potentially with an explanation how to appeal the decision.

AI could automatically draft notifications that are tailored to the individual's case. The task of sending out standard notifications could easily be automated using traditional technologies, but AI could automate the drafting of notifications that include information about the main or determinant factors influencing the decision, or provide information about what would happen in a counterfactual (Doshi-Velez and Kortz, 2017^[60]). This means that, if an applicant is denied certain social benefits based on an AI system's recommendation, the AI-generated notification could clarify what factors affected the decision, whether they affect it positively or negatively and what their respective weights are (see also Section 5.4). However, in practice, AI does not appear to be used for this purpose (yet).

3 Opportunities

Modernising social protection through digital tools may bring many opportunities (OECD, 2024^[12]), and using AI for managing means-tested benefits such as minimum income benefits (MIB) and unemployment assistance (UA) has the potential to improve upon these advances. It may alleviate administrative burden through task automation, provide caseworkers with valuable insights for improved decision-making, minimise delays between benefit eligibility and receipt or debt accumulation, improve client interactions with social services, and – most importantly – improve the take-up of MIB and UA among eligible individuals. Yet, the extent to which these opportunities of AI are realised in practice remains uncertain, due to a lack of evidence and because the realisation of the opportunities critically hinges on AI being trustworthy and beneficial for individuals, which requires certain safeguards to be put in place (see Section 5). This section presents the potential opportunities; Section 4 will discuss the risks that systems used for MIB or UA can pose if they are not used in a trustworthy way.

3.1. Decreased administrative burden

By (further) automating tedious and repetitive tasks, AI systems offer the potential to decrease the administrative burden on public servants engaged in MIB or UA management. Non-AI technologies and -automation could also help with this, but Section 2 shows that, with the help of AI, caseworkers may additionally no longer need to process hand-written forms and open-text caseworker reports, or cross-reference benefit application forms with an array of data sources. Instead, they can receive individuals' full case reports with the click of a button, including recommendations about their (in)eligibility for specific benefits.⁶ AI can also answer questions 24/7.

Caseworkers and customer service employees could use the time freed-up by AI to focus on more complex cases. For instance, (Milanez, 2023^[61]) shows that, in the finance and manufacturing sectors, the automation of tedious tasks through AI allowed workers to spend more time supporting customers and colleagues across the firm. For caseworkers, this could mean that they can spend more time on individuals that require client consultations, home visits, or compassionate decision-making. For customer service employees, the experience of the Canadian Revenue Agency of using a chatbot to help people find the answer to their questions was that, while the overall volume of calls remained the same, the type of questions asked to call centre agents changed towards questions about their individual cases (CRA, 2024^[28]).

3.2. Efficient benefit allocation and related social spending

AI's ability to quickly process and analyse large volumes of data enable it to quickly assess and reassess MIB and UA eligibility in view of changing circumstances. Depending on how frequently data are updated in the system (see Section 5.5), this could help reduce the intensity and duration of financial hardship as

⁶ Evidence from other sectors also shows that AI tends to automate more repetitive tasks than it creates (Lane, Williams and Broecke, 2023^[112]).

compared to “traditional”, non-AI powered systems, in which benefit eligibility may be updated every few months or once per fiscal year. The ESDC use case in Canada (Box 2.2) shows that AI can re-assess social benefit eligibility of over 10 million individuals within a few weeks – a task that would likely take a human team of caseworkers several months or even years due to the required open-text analysis. With AI, caseworkers no longer needed to perform the time-intensive task of reading through vast amounts of open text, and can instead focus on verifying the decisions made by the system, thereby significantly speeding up the process.

In addition to identifying eligible individuals, AI could also improve the speed and efficiency in identifying individuals who are at high risk of financial or situational hardships. While human caseworkers and non-AI technologies can perform this task as well – and indeed, they often do – they would not be able to discern patterns in as little time and using as many variables as an AI system can. AI’s predictive power and speed enables social protection agencies to implement measures such as targeted interventions or financial counselling early on, thereby mitigating potential challenges before they escalate.

Similarly, AI tools could help scale down benefits in a timelier way and communicating decisions more promptly. As a result, in addition to making the targeting of benefits more effective, AI could minimise or even prevent undue payments. For instance, the Brazilian use case described in Section 2.3 highlights the use of AI to speed up beneficiary death detection. Using AI could also help prevent undue payments by flagging anomalies in benefit application forms, as exemplified by the Danish use cases described in Box 2.4.

Overall, using AI to swiftly allocate, scale up or down, or terminate MIB and UA may improve efficiency of social spending, helping to ensure that public resources are allocated where they are needed most and when they are needed. AI would also allow social protection agencies to have more accurate and up-to-date estimations of current needs for MIB and UA. Combined with AI-powered identification of patterns and trends, this could allow policy makers to better anticipate shifts in demand for MIB and UA, and proactively plan for upcoming challenges. For instance, when demand for MIB or UA is expected to increase, additional resources can be earmarked for future benefit recipients, investments in human resources can be made to efficiently manage the increased demand, and measures to prevent hardship can be implemented.

3.3. Improved accuracy and fairness of decision-making

The use of trustworthy AI may promote data driven, objective, and consistent benefit eligibility assessments. While there are undoubtedly risks of biases in AI systems (see Section 4.1 for a detailed discussion), humans are unfortunately not infallible either. If well designed and trained on unbiased data, AI systems used for MIB or UA eligibility assessments⁷ may be more accurate and data driven, because AI enables the use of a richer variety of data as inputs than previously possible, such as unstructured data (e.g. images and open text), and because it can harness information generated by digitised service delivery. It therefore creates opportunities for improved problem definition and policy framing, and allows for a quicker, deeper, and more precise understanding of citizen preferences and context (Berryhill et al., 2019^[62]). AI-driven eligibility assessments may also improve fairness by eliminating variations across caseworkers or regions. Yet, realising this opportunity critically hinges on the AI system making use of high quantity and quality data and algorithms (see Section 5.5). Considering this can be challenging, there is a risk of incorrect or biased outcomes when using AI for MIB or UA eligibility assessments: see Section 4.1.

⁷ Including the use of AI to improve the accuracy and timeliness of the variables used as input for defining benefit eligibility.

3.4. Improved client experience and trust

AI may improve interactions with social protection agencies. For instance, AI-powered conversational agents can provide plain-language information about MIB and UA and how to apply to them, with information tailored to the individual's case, and do so 24/7. Moreover, AI's potential to free up time for caseworkers and call centre agents could make it easier to find someone available to provide information and support, thereby facilitating human-to-human interactions. Additionally, individuals who prefer *not* to talk to a person can still receive the necessary information and support, for instance by interacting with AI-powered conversational agents. The automated nature of AI interactions can create a less intimidating environment, potentially encouraging more individuals to explore and access available benefits.

AI may simplify the application process. AI's speed in processing structured and unstructured data can help social protection agencies implement proactive MIB and UA systems in which individuals receive a notification about their eligibility for – or even enrolment in – specific benefit programmes, if they wish to do so. This could significantly decrease the necessary effort for individuals to enrol in MIB or UA programmes, thereby improving take-up (see Section 3.5). Even systems for MIB or UA where people need to apply for benefit themselves, the application process could be simplified by using AI, because of its potential to improve information provision about available benefits and how to apply to them, and to pre-fill benefit application forms.

AI's potential to provide more accurate and timely decisions about benefit eligibility may also enhance trust in social protection agencies. This is particularly likely if using AI reduces the number of inaccurate rejections, downward adjustments, or terminations of MIB or UA (i.e. decreasing the probability of false negatives in determining benefit eligibility), or if it speeds up enrolment and upward adjustments. Yet, this opportunity requires there to be public trust in AI, which critically hinges on AI being trustworthy and beneficial for individuals. This, in turn, requires certain safeguards to be put in place (see Section 5).

3.5. Increased take-up

One of the key opportunities of using AI for managing MIB or UA is its potential to increase the take-up of these benefits among eligible individuals. Indeed, most AI-use cases for MIB and UA analysed in this paper have as one of their primary aims improving take-up. There are many reasons for non-take-up, including lack of information, complexity of access, and social barriers such as stigma (Dubois and Ludwinek, 2015^[15]; OECD, 2023^[18]; OECD, 2024^[12]; de Schutter, 2022^[63]). As discussed above, AI may decrease non-take-up through each of these channels: it could improve the quality and availability of information, facilitate access to MIB and UA by automatically enrolling individuals into programmes or pre-filling forms, and decrease feelings of stigmatisation or shame by providing the option to interact with a machine rather than a person. Moreover, with more time to focus on complex cases, caseworkers can proactively identify eligible individuals who might not have sought assistance due to the complexity of access or lack of information. Yet, if not designed or implemented well, AI could also *decrease* take-up: see Section 4.3.

4 Risks

The integration of AI into the realm of means-tested benefits such as for managing minimum income benefits (MIB) and unemployment assistance (UA) is not without risks – see also (OECD, 2024^[12]). Evidence from various incidents⁸ shows that, if not designed or implemented well, AI systems used to manage MIB or UA can generate incorrect or biased eligibility assessments, lead to privacy infringements, decrease take-up among certain groups, and lack the transparency, explainability and accountability needed for people to exercise their rights (e.g. right to due process, or providing consent to automated decision-making) and identify risks, which may ultimately cause harm and decrease trust in public institutions. Greater awareness of these risks – and appropriate safeguards to address them (see Section 5) – may empower actors, including policy makers, caseworkers, and citizens, to use AI systems for MIB and UA with caution.

4.1. Incorrect or biased eligibility assessments

Incorrect or biased MIB or UA eligibility assessments mean that individuals in precarious situations do not, or no longer receive the benefits they are entitled to (also known as “false negatives” in benefit eligibility assessments), or they lead to the wrongful allocation of MIB or UA to those who do not actually need them (also known as “false positives”). Biased assessments could lead to discrimination, in violation of the United Nations’ Universal Declaration of Human Rights (United Nations, 1948^[64]) and anti-discrimination laws. False positives or -negatives in benefit eligibility assessments typically lead to repayment demands, which can be substantial and often accompanied by fines or interest charges, and it could lead to wrongfully labelling individuals as engaging in fraudulent activity. This may not only increase financial hardship of already-poor individuals: it may also have severe negative consequences on their mental health and well-being, and people’s overall trust in public institutions. Additionally, for governments, the costs for compensation of financial, emotional, and other damage caused to victims of erroneous benefit denials or terminations may be substantial.

The use of AI and automated systems can perpetuate inaccuracies in MIB and UA calculations. For example, the process known as “income averaging” used by the algorithm of the former debt recovery scheme in Australia to assess income and benefit entitlement (Robodebt) produced inaccurate results and did not comply with the income calculation provisions of the Social Security Act 1991. As a result, debt notices were wrongfully issued to affected welfare recipients who would have to prove they did not owe a debt that could be many years old (Meers et al., 2017^[65]; Murray, Cheong and Paterson, 2023^[66]; OECD, 2024^[12]; Royal Commission into the Robodebt Scheme, 2023^[67]).

AI systems can carry over and systematise biases inherent in the data they are trained on, potentially leading to discrimination. While bias can also occur in decisions made by human agents, the use of AI can amplify the scale of these incidents. Yet, it can be difficult to get the data right – see Section 5. For instance,

⁸ The OECD AI Incidents Monitor (AIM) documents AI incidents to help policy makers, AI practitioners, and all stakeholders worldwide gain valuable insights into the incidents and hazards that concretise AI risks (OECD.AI, 2024^[110]). Over time, the AIM will help to show patterns and establish a collective understanding of AI incidents and their multifaceted nature and serve as an important tool for trustworthy AI.

biased investigation or reporting of undue payments can cause AI to systematise these biases and systematically assign higher risk scores to specific demographic groups. A recent study has shown that the algorithms of the United States' Internal Revenue Service (IRS) disproportionately select Black taxpayers for audits, even though the tax collection agency does not collect information on race. The authors argue that the bias stems from the fact that Black taxpayers are more likely to make certain mistakes in claiming tax credits and file the types of tax returns the IRS typically targets. For instance, people filing for the Earned Income Tax Credit are more likely to be selected for audits (Elzayn et al., 2023^[68]). The IRS has identified this problem in the algorithm and is making changes to how people are selected for audit (OECD, 2024^[12]).

Bias can also be introduced during the process of training or curating the data or the algorithm, or in the selection of biased proxies. For instance, a Dutch algorithm to identify fraudulent applications for childcare benefits was suspended in 2019, after the Dutch Data Protection Authority found it had illegally used nationality as a variable, leading to over 34 000 wrongly identified cases of fraud affecting over 94 000 children (Dutch Ministry of Finance, 2024^[69]), at least 1 000 of whom were placed into custody as a result of the accusations (European Parliament, 2022^[70]) – later known as the “childcare benefits scandal”. According to Amnesty International (2021^[71]), “the use of nationality in the risk classification model reveals the assumptions held by the designer, developer and/or user of the system that people of certain nationalities would be more likely to commit fraud or crime than people of other nationalities”.

Yet, while AI could scale up and systematise incorrect and biased benefit eligibility assessments, this would not necessarily lead to large-scale incidents, if there are (effective) mechanisms in place to regularly check the system's output for bias or discrimination and/or that enable making exceptions for extreme cases, and if these mechanisms are well enforced (see also Section 5). For instance, one of the key issues regarding the childcare benefits scandal in the Netherlands, was that legislation offered insufficient options to intervene or deviate in the event of unforeseen and unreasonably harmful consequences in individual situations, such a hardship clause or discretionary powers (Childcare Allowance Parliamentary Inquiry Committee, 2020^[72]).

Vulnerabilities in AI systems can also increase benefit fraud, undermining the integrity of social protection and exacerbating inequalities. By analysing the patterns and criteria used by the AI system to make eligibility determinations, malicious actors could reverse-engineer the AI models and strategically tailor their fraudulent applications to mimic legitimate cases and evade detection. This highlights that there is a tension between transparency and risk of abuse of the system (see Section 4.4) and fraud, that introduces vulnerabilities into the system. Additionally, hackers could exploit vulnerabilities in the software or network infrastructure supporting AI systems, to manipulate outcomes in their favour (Comiter, 2019^[73]). They could also exploit these vulnerabilities to steal personal data: see Section 4.2.

4.2. Privacy concerns

AI systems for managing MIB or UA often access and combine large amounts of sensitive personal data, posing a risk of privacy breaches if the data governance is inadequate. For instance, data could be misused, used without the needed consent, or inadequately protected (GPAI, 2020^[74]; OECD, 2022^[75]). Although the risk of a privacy breach for digital technologies is not limited to using AI, the personal data collected or processed by AI systems are often more extensive than data collected or processed by humans or through other technologies, thereby increasing the potential harm if something goes wrong. For instance, AI systems for MIB or UA management often integrate data from various agencies, increasing the risk of privacy violations through unauthorised access or unintended disclosures. Additionally, obtaining meaningful consent for data use by public institutions can be problematic, due to the power imbalance between citizens and the government, and because the opaque nature of AI decision-making processes can make it difficult to understand how personal data are being used or processed (see Section 4.4).

Moreover, AI systems are capable of detecting subtle patterns and correlations within large datasets, allowing them to make predictions or classifications about individuals that extend far beyond the information explicitly provided. For instance, AI systems can infer sensitive information of individuals (e.g. religion, sexual orientation, or political affiliations) based on non-sensitive data (Wachter and Mittelstadt, 2019^[76]). Privacy breaches are a violation of people's fundamental rights as well as data- and privacy protection legislation – in place in all OECD member countries (UNCTAD, 2024^[77]).

4.3. Increased inequalities and decreased take-up among certain groups

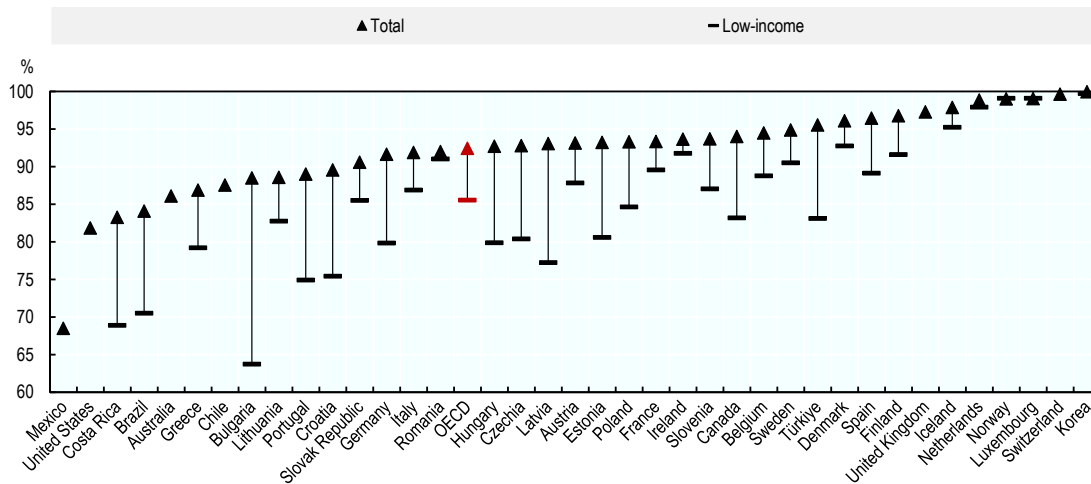
People with difficulties accessing or engaging with digital services may not be able to benefit from AI as much as others. Certain AI systems are not (easily) accessible to people with certain disabilities, with low levels of digital skills, and/or without internet access at home. For instance, on average across OECD countries, more than a third of adults have low digital problem-solving skills or have no experience with a computer, and in some countries (e.g. Republic of Türkiye, Mexico, Chile), this share is higher than 60% (Verhagen, 2021^[78]). At best, this decreases people's ability to benefit from AI. For instance, caseworkers with low levels of digital skills may experience difficulties using AI systems that decrease their administrative burden, potentially exposing them to the risk of automation due to declining productivity compared to those who use AI.

At worst, difficulties accessing or engaging with digital services, such as those that are AI-powered, may decrease MIB or UA take-up of certain groups, especially the most vulnerable. In many OECD countries, low-income households – who are more likely to need MIB and UA – are less likely to have internet access at home (see Figure 4.1). This would not necessarily decrease take-up if traditional off-line or non-digital solutions continued to exist. However, since many AI systems need as much digitised data as possible to function optimally, introducing AI for MIB and UA management may be accompanied by an overall digitisation of public services and a discontinuation or discouragement of the possibility to submit paper-based benefit applications, for example.⁹ Using AI for MIB and UA management may therefore unintentionally decrease take-up among certain groups. For instance, while 98% of households who make a claim for Universal Credit in the United Kingdom does so successfully online, a small number of people with complex needs or without access to the internet are not able to use the online process. Indeed, Human Rights Watch has found that those who do not regularly use computers or the internet experience difficulties accessing their welfare benefits: a problem that is exacerbated by decreasing availability of computers in libraries and community centres due to budget cuts (Human Rights Watch, 2020^[79]). In response, the DWP provides a range of support to make the service more accessible (see Section 5).

⁹ Yet, AI may also enable the processing of hand-written forms – see Section 2.2.

Figure 4.1. Households with internet access at home, by income group

Percentage of households



Note: Reference year is 2023, except for Austria, Canada, Costa Rica, Ireland, Korea and Mexico, for which the reference year is 2022, the United States (2021), the United Kingdom (2020), Chile and Iceland (2017), and Australia (2016). Low-income households are in the bottom quintile of the national income distribution.

Source: OECD (2023^[80]), ICT Access and Usage by Households data, <https://data-viewer.oecd.org/?chartId=0144c9cd-4669-476f-a762-3a4affc8317f>.

4.4. Lack of transparency and explainability

Individuals are not always aware of the scope, extent or even existence of AI use to manage MIB or UA. AI use can be difficult to detect without explicit disclosure, and even if individuals are aware of their own or their caseworkers' interactions with AI, obtaining insights into AI's decision-making process can be difficult, for instance due to the complexity of the system or developers' reluctance to disclose information. This lack of transparency poses important limits to people's ability to exercise specific rights (e.g. the right not to be subject to automated decision-making), detect risks (e.g. risks of bias or for their privacy), and/or effectively question outcomes. In the United Kingdom, for instance, the Information Commissioner has raised serious concerns that the DWP rejected freedom of information requests and blocked inquiries from MPs about their AI system for fraud detection in welfare claims. While the DWP argues that transparency could enable fraudsters (see Section 4.1), the Information Commissioner is worried about the lack of transparency (Booth, 2023^[81]).

AI systems, particularly those using complex technologies like deep neural networks, yield outcomes that can be difficult or even impossible to explain, even by the designers of the systems. This lack of explainability means that it can be extremely challenging for individuals to understand why they were denied MIB or UA or had their benefits adjusted, potentially leading to mistrust of the system even if it made the correct decision. A lack of explainability also makes it difficult for individuals to provide informed consent to the use of such systems, or to identify and seek redress for adverse effects caused by AI systems for MIB or UA. For instance, since Serbia's introduction of an algorithm to determine people's eligibility for social benefits, over 22 000 people have lost their benefits without knowing why (Caruso, 2022^[82]). In the Netherlands, one of the problems in the childcare benefits scandal was that the tax authorities were unable to explain their decision to identify parents as fraudsters, including what the parents had done wrong and how they could correct their mistake (Amnesty International, 2021^[71]; Kuźniacki, 2023^[83]). Similarly, the Los Angeles Office of Child Protection terminated its AI-based project, citing the "black-box" nature of the algorithm (Lokshin and Umapathi, 2022^[84]).

4.5. Risks for accountability

It is not always clear which actor linked to the AI system is responsible for preventing that something can go wrong, or who is accountable if something does go wrong: the developer, the provider, or the user of the system. This is related to the fact that, unlike traditional goods and services, some AI systems can change as they are used, by learning from new data. Using “black box” AI – including pre-trained models that are not reviewed or validated, and AI systems that are licensed from third party vendors – poses particular due process concerns (Campolo et al., 2017^[85]). In recent years, legislators have made efforts to promote accountability mechanisms, such as impact assessments and/or audits to provide evidence and assurance that they are trustworthy and safe to use (see Section 5). However, enforcing these policies can be difficult because developers, providers and users of AI systems do not necessarily reside in the same jurisdictions. For instance, Amnesty International argues that public and private sector players may circumvent regulatory obligations by exporting the same AI systems to other countries (Nolan, Maryam and Kleinman, 2024^[86]).

Additionally, individuals eligible for MIB or UA may not always possess the skills or resources to verify whether they have been treated unjustly and how to contest erroneous decisions based on AI. Consequently, they may encounter significant barriers in seeking redress when their benefit applications are unjustly rejected or when their existing benefits are erroneously adjusted or terminated (Barca and Chirchir, 2019^[87]; Lokshin and Umapathi, 2022^[84]). This is even more challenging considering that understanding the workings of an AI system can be a difficult task for anyone (see Section 4.4).

5 Seizing opportunities, addressing risks: Possible policy directions

As the previous sections have shown, AI presents both significant opportunities and risks when used for managing means-tested benefits such as minimum income benefits (MIB) and unemployment assistance (UA). This section explores possible policy directions for trustworthy AI, that is policies that seize the opportunities and address the risks. The findings of this section are based on insights from the survey “*Harnessing Technology and Data to Improve Social Protection Coverage and Social Assistance Delivery*” which was distributed among members of the OECD Working Party on Social Policy in the Summer of 2023 (OECD, 2023^[22]), semi-structured interviews with country experts on the use of AI for social benefits management, as well as the OECD AI Principles (OECD, 2019^[23]). The section also builds on previous OECD work on use of AI in the public sector (Berryhill et al., 2019^[62]; OECD, 2023^[88]; Ubaldi et al., 2019^[89]), and the recent OECD report *Modernising access to social protection: Strategies, technologies and data advances in OECD countries* (OECD, 2024^[12]).

5.1. Human intervention and oversight

The OECD AI Principles – which set standards for trustworthy AI – emphasise the importance of mechanisms that ensure human intervention and oversight in AI systems, to uphold human-centred values and fairness (OECD, 2019^[23]). Human intervention and oversight are especially important when AI is used to support decisions that have direct consequences on people’s rights and well-being, such as decisions regarding MIB or UA eligibility. For such decisions, many governments require having humans “in the loop” (e.g. they may have to approve a decision) or “on the loop” (e.g. they are able to view and check the decisions being made): for most examples of AI use for MIB or UA management discussed in this paper, the decision-making power for assigning or terminating benefits remains with human caseworkers. For European countries, the EU AI Act stipulates additional legal requirements, including human oversight, for most AI systems used in MIB or UA management, by categorising them as high-risk AI systems (Council of the European Union, 2024^[90]).¹⁰ However, uncertainties persist regarding the legal role and accountability of humans involved in the decision-making process (Enarsson, Enqvist and Naarttijärvi, 2021^[91]), underscoring the need for clarity about public officials’ role in the decision-making process, and ensuring that they have the required competence to monitor the AI system effectively and spot anomalies.

Human intervention and oversight in the management of MIB and UA also enable deviation from (or termination of) AI-enabled decision-making, for instance when the decisions appear to be biased, discriminatory or inaccurate, or when they have disproportionate consequences on individuals or households. This may promote inclusivity and accessibility of the AI-powered tools for MIB and UA management. Human-to-human interaction may also help foster transparency and explainability of the

¹⁰ Specifically, the EU AI Act identifies the following as “high risk” AI systems: “AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services” (Council of the European Union, 2024^[90]).

decision-making process, which are beneficial for accountability and trust. For instance, the United Kingdom's DWP provides additional support for people with complex needs or those who are not able to use online processes, for example with face-to-face support in local Jobcentres. Similarly, in Canada, individuals can access assistance through dedicated helplines staffed by trained personnel that provide guidance and support. Moreover, some provinces offer in-person assistance through local government offices or community centres, enabling individuals to directly interact with staff members who can offer personalised assistance tailored to their needs.

5.2. Strategic inclusiveness

Using AI for determining eligibility or identifying undue payments of MIB or UA requires specific considerations to strike a balance between maximising costs at the risk of denying some eligible individuals the benefits they are entitled to (“false negatives”) and maximising take-up even if that means assigning benefits to some people who are not eligible (“false positives”). Determining the threshold for false positives and false negatives in AI systems used for MIB or UA management is crucial yet complex, requiring careful considerations of technical capabilities and ethical implications. For instance, in the use case of the Canadian Guaranteed Income Supplement (GIS) programme (see Box 2.2), experts made a strategic decision to design the model to err on the side of inclusion to mitigate the risk of leaving eligible individuals without much-needed support, even if it means accepting some level of error which resulted in more work for Service Canada caseworkers (ISSA, 2020^[27]; ESDC, 2023^[36]).¹¹ However, while this type of strategy ensures that more eligible individuals receive social benefits, it also raises concerns about the allocation of limited resources and the potential for misuse or abuse of the system. Achieving strategic inclusiveness requires a nuanced approach that considers both the immediate needs of vulnerable populations and the long-term sustainability of social protection.

5.3. Ensuring accountability early on

Establishing clear accountability mechanisms for AI use in MIB and UA management – meaning that the systems follow the law and/or principles of trustworthiness – is crucial for the use of trustworthy AI and the enforcement of regulations (OECD, 2023^[88]). Conducting impact- and risk assessments before and during the implementation of AI systems can identify risks to individuals’ safety and rights, as well as opportunities early on. For example, the Dutch Fundamental Rights and Algorithms Impact Assessment (FRAIA) helps to map the risks to human rights in the use of algorithms by government organisations (Government of the Netherlands, 2024^[92]), although it is not always clear how addressing the risks identified in the FRAIA can be enforced. Quality labels and certifications for trustworthy AI systems can also help. An increasingly popular tool to assess AI systems after implementation is “AI auditing” or “algorithmic auditing”. Generally speaking, in an algorithmic audit, a third-party assesses to what extent and why an algorithm, AI system and/or the context of their use aligns with ethical principles or regulation.

However, measures such as impact- and risk assessments will only promote accountability if they are properly enforced and communicated to the public in a clear and simple manner (see Section 5.4), with explanations on decisions can be reviewed and appealed. This, in turn, calls for adequately trained staff (see Section 5.6) and an adequate complaint processes in place, with measures to ensure follow-up. Regulatory impact assessments (RIAs) may help to identify and quantify the benefits and costs likely to flow from regulatory or non-regulatory options to ensure accountability of AI use for MIB and UA (OECD, 2021^[93]).

¹¹ Note that, even with strategic inclusion in place, human caseworkers still needed to verify the decisions made by the system.

5.4. Transparency and explainability

Transparency and explainability mechanisms enable individuals to understand decisions, detect risks, effectively question outcomes of AI systems, provide informed consent to the use of such systems, and exercise their rights, amongst others. Transparency consists of disclosing to affected parties the fact that AI is being used to make predictions, recommendations or decisions that concern them, which helps build trust in AI systems and the governments that use them. For instance, the Dutch Government's public Algorithm Register provides information on algorithms used by government institutions, including their purpose, impact, level of human determination and oversight, and data used (Overheid.nl, 2024^[94]). In Estonia, applicants for unemployment benefits are informed by the Unemployment Insurance Fund's use of ADM when submitting an application, and the use of ADM is displayed on the decision document (Algorithm Watch, 2024^[38]).

Explainable AI (also known as "XAI") consists of being able to provide information how AI systems reach their predictions, recommendations, or decisions. One approach for achieving XAI is to provide post-hoc explanations of how the system reached its outcome, which can be particularly useful for complex systems such as neural networks or deep learning algorithms that cannot be easily explained (also known as "black box" AI). Post-hoc explanations could for instance include information about the main or determinant factors influencing the decision, or provide information about what would happen in a counterfactual (Doshi-Velez and Kortz, 2017^[60]). However, considering that some experts argue for an overall ban on "black box" AI for core public agencies in high stakes domains such as welfare (Campolo et al., 2017^[85]), building interpretability into the design of the system may be preferable when AI is used for MIB or UA. This type of XAI is also known as "self-interpretable" or "white box" AI, and should be so self-explanatory that it in many cases allows for a visualisation of a decision tree of each decision the system made to reach its outcome (EDSP, 2023^[95]).

5.5. Data quality and data governance

Data are the fundamental building block of AI. For instance, Section 3 highlighted the significant opportunity of improving efficiency in allocation of MIB or UA critically hinges of the availability of data that are updated in near real-time. Yet, obtaining high-quality (e.g. timely and unbiased) data is costly and administratively complex, and generating a stable consensus across society on the trade-offs, for instance between privacy, transparency and service quality, can be challenging (Janssen and van den Hoven, 2015^[96]). Effective data governance requires arrangements across technical, policy, regulatory or institutional domains that affect data and their cycle – including creation, collection, storage, use, protection, access, sharing and deletion of data (OECD, 2024^[12]). The OECD provides guidelines and frameworks to develop data governance policies in the public sector and beyond (OECD, 2019^[97]; OECD, 2022^[98]), emphasising the necessity of a comprehensive whole-of-government approach to data governance.

5.6. Capacity to understand, develop, and use AI

To promote a trustworthy use of AI systems for MIB or UA management, managers in the public sector need the skills not only to assess whether AI is a viable solution but also to explain the benefits and risks associated with AI use in this context to affected caseworkers. Caseworkers, in turn, not only need to be able to understand the opportunities and risks of AI, but also need the skills to work alongside, interpret, or complement the technology. Government IT staff will need to have the skills to develop, implement and/or maintain the AI system. This requires investing in AI literacy training across the organisation and establishing a clear narrative explaining the opportunities of AI for MIB and UA management. For instance,

Ireland developed a national certification programme to upskill civil servants on the ethical application of AI in government (OPSI, 2022^[99]). Training programmes for elementary AI literacy often provide a basic understanding of what AI is and what it can and cannot do. Most courses review machine learning and deep learning, as they form the basis for most AI applications today. Some courses also discuss ethical issues in AI, addressing algorithmic bias or the black box nature of AI (Lassébie, 2023^[100]). AI literacy training for public servants may not only ensure trustworthy use of AI for MIB and UA management by enabling them to assess its opportunities and risks and how to address them; it may also reduce workplace anxiety about the implications for staff. Guidelines are also needed, to support caseworkers and their managers on compliance with accountability measures applicable to their use of AI for MIB or UA management, including national or international auditing standards (e.g. on auditor independence, representative analysis, access to data, code and models, and consideration of adversarial actions).

Countries are developing training programmes to promote digital skills and AI literacy in the general population as well (Lassébie, 2023^[100]), which will help people understand how AI use for MIB or UA may affect them, and help them interact with these systems if needed. This is not only necessary for accountability and transparency, but it may also help to build trust and improve take-up of trustworthy AI-powered benefits. Importantly, access to the internet and relevant devices, such as a computer or mobile phone, will be critical to benefiting from new technologies such as AI. In the United States, for example, the *Lifeline* programme helps ensure that low-income consumers can afford 21st century connectivity services, by providing a discount on telephone service and/or broadband internet (FCC, 2024^[101]).

5.7. Proactive engagement by AI stakeholders

To build trust and support in AI use for MIB or UA management and maximise their potential and address their risks, policy makers should engage with stakeholders and individuals who may be directly or indirectly affected by these systems, including caseworkers or their representatives, and low-income individuals. Stakeholders can be involved in designing, piloting, and scaling AI-powered services, fostering innovation and user-centred approaches (ISSA, 2023^[102]). For instance, workshops with caseworkers and frontline staff could help to understand the challenges they face in managing MIB or UA and how AI tools could support (or are supporting) their work. Pro-active public services, in particular, that anticipate and handle user needs before they have to take action (e.g. completing a benefit application form) cannot be achieved without a very thorough understanding of user needs, the key to which is user engagement (Berryhill et al., 2019^[62]). For instance, Ireland is moving to systematic involvement of citizens in the creation of digital and other solutions to help combat non-take-up of benefits and services, for instance by testing prototype solutions with customers (Grace, 2018^[103]; OECD, 2024^[12]). AI can also help facilitate and process stakeholder engagement and provide personalised feedback to participants (Bono Rossello, Clarinval and Simonofski, 2023^[104]).

5.8. Flexibility and experimentation

To harness the benefits of AI use for MIB or UA management and to minimise the risk of adopting dangerous, unethical, or sub-optimal AI in this context, social protection agencies need the freedom to try new ways of managing these means-tested benefits. Regulatory experimentation, which can be particularly useful in addressing innovation-induced disruptions and the resulting uncertainty,. However, adopting an experimental approach to AI use may counteract efforts to put in place robust systems and consistent processes across government. For instance, the EU General Data and Protection Regulation (GDPR) stipulates that people have a right to meaningful human involvement in decisions that have a significant

impact on their lives (Official Journal of the European Union, 2016^[105]), which could limit using AI-powered automated decision-making for MIB and UA management in EU countries.¹²

One way to strike a balance the needed controls and embracing experimentation is through regulatory experimentation. If used appropriately and in combination with other relevant approaches and regulatory co-operation, regulatory experimentation can improve adaptive learning, policy coherence, and the evidence base for regulatory design, delivery and adaptation, resulting in more effective and efficient public policies (OECD, 2021^[93]; OECD, 2024^[106]). A common form of regulatory experimentation is the creation of regulatory “sandboxes”, which offer a safe environment to test new AI technologies without exposing users or society to undue risk, fostering innovation while providing regulators with valuable insights to adapt regulatory frameworks. While often geared towards the private sector, sandboxes are increasingly being considered for public sector AI (Appaya, Gradstein and Haji Kanz, 2020^[107]; Attrey, Leshner and Lomax, 2020^[108]; Madiaga and Van De Pol, 2022^[109]).

Moreover, to ensure that AI solutions for MIB and UA management remain trustworthy post-implementation, it is crucial to embed flexibility in the decision-making processes, for instance by ensuring human intervention and oversight (see Section 5.1). Post-implementation flexibility also requires regularly monitoring, reviewing and updating the AI systems to adapt to new data and evolving contexts, ensuring ongoing compliance with existing regulations. Continuous adaptation and rigorous oversight can ensure that AI technologies remain accountable, transparent, and aligned with ethical standards, thus preventing issues from becoming entrenched.

¹² Yet, exceptions to the rule about human involvement in automated decisions exist under EU law. If suitable measures are in place to protect people’s rights, freedoms and legitimate interests, automated decisions may be allowed. For example, in Estonia, the Unemployment Insurance Act allows fully automated decision-making to attribute or reject UB to applicants, by informing applicants that the decision was automated, that they have a right to be heard and to submit a request for internal review (Barros Vale and Zanfir-Fortuna, 2022^[111]).

6 Conclusions

This paper shows that AI can be used throughout the process of managing means-tested benefits such as minimum income benefits (MIB) and unemployment assistance (UA), from providing information to potential beneficiaries to determining eligibility according to pre-determined statutory criteria. This brings significant opportunities as well as risks for social protection. For instance, using AI for MIB or UA management may improve take-up by automatically enrolling individuals into programmes or pre-filling forms. Yet, if not designed or implemented well, the opposite effect may be achieved with risks that some groups are left behind. For example, groups with low digital skills or groups with limited access to digital devices or internet at home. Similarly, while AI's data-driven decision-making may improve accuracy and fairness of eligibility assessment based on pre-defined statutory criteria, it may also systematise inaccuracies and bias. This is particularly problematic for MIB or UA, that often specifically serve vulnerable groups in the population.

Understanding the balance between the opportunities of AI adoption for MIB and UA management and the associated risks is therefore crucial, and careful consideration must be given to its implementation. The decision-making process regarding AI adoption for MIB and UA management must consider whether AI is the optimal solution for the problem at hand and ensure that AI systems used for MIB or UA management are trustworthy. This includes maintaining human intervention and oversight, transparency and explainability, including in how false positives and false negatives are treated, and good data governance of AI systems used. It also includes regular risk assessments and audits, and fostering capacity-building initiatives to equip public servants and citizens alike with the necessary skills and knowledge to understand, develop or use AI. This requires establishing robust controls and accountability mechanisms, while also embracing flexibility and experimentation to adapt to evolving challenges and opportunities. Further research and stakeholder consultation will be necessary to develop detailed implementation strategies and address specific regulatory and ethical considerations.

Countries can take advantage of opportunities to collaborate internationally on AI approaches and standards, so that they do not need to handle every aspect of developing robust agendas and ecosystems for AI use for MIB and UA management. Innovative policy efforts in these areas are often scattered and lack coherence, limiting the potential for collective learning and the scaling up of good practices (OECD, 2023^[88]). Collaborating internationally on AI approaches and standards offers countries opportunities to address common challenges and explore collaborative approaches. Initiatives like the OECD AI Principles provide a global framework for trustworthy AI.

References

- Albrecht, T., T. Rausch and N. Derra (2021), “Call me maybe: Methods and practical implementation of artificial intelligence in call center arrivals’ forecasting”, *Journal of Business Research*, Vol. 123, pp. 267-278, <https://doi.org/10.1016/j.jbusres.2020.09.033>. [30]
- Algorithm Watch (2024), *Automating Society Report 2020: Estonia*, <https://automatingsociety.algorithmwatch.org/report2020/estonia/> (accessed on 23 April 2024). [38]
- Almeida, V., S. De Poli and A. Hernández (2022), “The effectiveness of Minimum Income schemes in the EU”, *JRC Working Papers on Taxation and Structural Reforms*, No. 09/2022, https://joint-research-centre.ec.europa.eu/system/files/2022-09/JRC130676_the_effectiveness_of_minimum_income_schemes_in_the_eu_final_version.pdf (accessed on 23 April 2024). [20]
- Amnesty International (2021), *Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal*, <https://www.amnesty.org/en/documents/eur35/4686/2021/en/> (accessed on 23 April 2024). [71]
- Appaya, M., H. Gradstein and M. Haji Kanz (2020), *Global Experiences from Regulatory Sandboxes*, The World Bank Group. [107]
- Attrey, A., M. Leshner and C. Lomax (2020), “The role of sandboxes in promoting flexibility and innovation in the digital age”, *Going Digital Toolkit Note*, No. 2, <http://www.oecd.org/going-digital>. (accessed on 26 April 2023). [108]
- Barca, V. and R. Chirchir (2019), *Building an integrated and digital social protection information system*, <https://www.giz.de/en/downloads/giz2019-en-integrated-digital-social-protection-information-system.pdf> (accessed on 23 April 2024). [87]
- Barros Vale, S. and G. Zanfir-Fortuna (2022), *Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities*, The Future of Privacy Forum, <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> (accessed on 23 April 2024). [111]
- Berryhill, J. et al. (2019), “Hello, World: Artificial intelligence and its use in the public sector”, *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [62]

- Bono Rossello, N., A. Clarinval and A. Simonofski (2023), “BeCoDigital – Digital Co-Creation of Public Services with Citizens: Understanding Pre-Conditions, Technologies, and Outcomes”, *The 17th International Conference on Research Challenges in Information Science*, <https://ceur-ws.org/Vol-3601/short3.pdf> (accessed on 23 April 2024). [104]
- Booth, R. (2023), *UK warned over lack of transparency on use of AI to vet welfare claims*, <https://www.theguardian.com/politics/2023/sep/03/uk-warned-over-lack-transparency-use-ai-vet-welfare-claims> (accessed on 23 April 2024). [81]
- Brioscú, A. et al. (2024), “A new dawn for public employment services: Service delivery in the age of artificial intelligence”, *OECD Artificial Intelligence Papers*, No. 19, OECD Publishing, Paris, <https://doi.org/10.1787/5dc3eb8e-en>. [43]
- Campolo, A. et al. (2017), *AI Now 2017 Report*, https://assets.ctfassets.net/8wprhvnpc0/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb14f2b74b2be64c3ce0c78/AI_Now_Institute_2017_Report_.pdf (accessed on 23 April 2024). [85]
- Caruso, F. (2022), *Serbia, algorithmic discrimination rehearsals*, https://www.europeandatajournalism.eu/cp_data_news/serbia-algorithmic-discrimination-rehearsals/ (accessed on 23 April 2024). [82]
- Childcare Allowance Parliamentary Inquiry Committee (2020), *Letter from the Parliamentary Inquiry Committee*, https://www.houseofrepresentatives.nl/sites/default/files/atoms/files/verslag_pok_definitief-en-gb.docx.pdf. [72]
- Comiter, M. (2019), “Attacking Artificial Intelligence AI’s Security Vulnerability and What Policymakers Can Do About It”, *Belfer Center Paper*, <https://www.belfercenter.org/sites/default/files/2019-08/AttackingAI/AttackingAI.pdf> (accessed on 23 April 2024). [73]
- Council of the European Union (2024), *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement*, <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> (accessed on 23 April 2024). [90]
- CRA (2024), *Interview with Canada Revenue Agency, 25 January 2024*. [28]
- de Schutter, O. (2022), *A/HRC/50/38: Non-take-up of rights in the context of social protection - Report of the Special Rapporteur on extreme poverty and human rights*, <https://www.ohchr.org/en/documents/thematic-reports/ahrc5038-non-take-rights-context-social-protection-report-special> (accessed on 23 April 2024). [63]
- Digital Watch (2019), *UK Department for Work and Pension to use AI for benefit claims*, https://dig_watch/updates/uk-department-work-and-pension-use-ai-benefit-claims (accessed on 23 April 2024). [46]
- Doshi-Velez, F. and M. Kortz (2017), “Accountability of AI Under the Law: The Role of Explanation”, *Ethics and Governance of Artificial Intelligence Initiative*, https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf (accessed on 23 April 2024). [60]

- Dubois, H. and A. Ludwinek (2015), *Access to social benefits: Reducing non-take-up*, Eurofound, [15]
<https://www.eurofound.europa.eu/en/publications/2015/access-social-benefits-reducing-non-take> (accessed on 23 April 2024).
- Dutch Ministry of Finance (2024), *Feiten en cijfers (Facts and numbers)*, [69]
<https://herstel.toeslagen.nl/dashboard-kinderopvangtoeslag/> (accessed on 23 April 2024).
- DWP (2018), *Department for Work and Pensions Annual Report and Accounts 2017-18 for the year ended 31 March 2018*, [47]
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721224/dwp-annual-report-and-accounts-2017-2018.pdf (accessed on 23 April 2024).
- EDSP (2023), *TechDispatch #2/2023 - Explainable Artificial Intelligence*, [95]
https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2023-11-16-techdispatch-22023-explainable-artificial-intelligence_en.
- Elzayn, H. et al. (2023), “Measuring and Mitigating Racial Disparities in Tax Audits”, *Stanford Institute for Economic Policy Research (SIEPR)*, [68]
<https://siepr.stanford.edu/publications/working-paper/measuring-and-mitigating-racial-disparities-tax-audits> (accessed on 23 April 2024).
- Enarsson, T., L. Enqvist and M. Naarttijärvi (2021), “Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts”, [91]
<https://doi.org/10.1080/13600834.2021.1958860>,
<https://doi.org/10.1080/13600834.2021.1958860>.
- ESDC (2024), *Reducing Employment Insurance Backlog: A Machine Learning Approach*, [45]
<https://open.canada.ca/data/en/dataset/24d2cab2-6a0d-4234-9239-b6ce102ebabd> (accessed on 23 April 2024).
- ESDC (2023), *Interview with Employment and Social Development Canada, 24 October 2023*. [36]
- Esko, L. (2020), *2015. aastal käivitunud partnerluse jooksul viis Finestmedia senise moraalselt vananenud töötuskindlustuse infosüsteemi üle kaasaegsele lihtsalt hallatavale platvormile, muutes Eesti Töötukassa töötus-kindlustushüvitise menetluse täisautomaatseks (During the partnership that started in 2015, Finestmedia transferred the morally outdated unemployment insurance information system to a modern, easy-to-manage platform, making the unemployment insurance benefit procedure of the EUIF fully automatic)*, [39]
<https://finestmedia.ee/projekt/eesti-tootukassa-tootuskindlustuse-infosusteemi-uuendamise-ja-automatiseerimine/> (accessed on 23 April 2024).
- European Commission (2024), *Can I be subject to automated individual decision-making, including profiling?*, [8]
https://commission.europa.eu/law/law-topic/data-protection/reform/rights-citizens/my-rights/can-i-be-subject-automated-individual-decision-making-including-profiling_en (accessed on 22 April 2024).
- European Law Institute (2022), *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*, [9]
https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf (accessed on 22 April 2024).

- European Parliament (2022), *Parliamentary question - O-000028/2022: The Dutch childcare benefit scandal, institutional racism and algorithms*, [70]
https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html (accessed on 23 April 2024).
- FCC (2024), *Lifeline Program for Low-Income Consumers*, [101]
<https://www.fcc.gov/general/lifeline-program-low-income-consumers> (accessed on 23 April 2024).
- Fruy, G. et al. (2022), *Artificial Intelligence for human flourishing: the case of AI-driven social welfare systems*, [35]
<https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/AI-social-welfare> (accessed on 23 April 2024).
- Gartner (2021), *Gartner Says Government Organizations Are Increasing Investment in AI, But Their Workforce Remains Apprehensive*, [24]
<https://www.gartner.com/en/newsroom/press-releases/2021-10-05-gartner-says-government-organizations-are-increasing-> (accessed on 23 April 2024).
- Gemeente Amsterdam (2020), *Vroeg Eropaf (early intervention)*, [57]
<https://algoritmeregister.amsterdam.nl/ai-system/vroeg-eropaf/314/> (accessed on 23 April 2024).
- Government of Canada (2024), *Old Age Security Benefits Estimator*, [33]
<https://estimateursv-oasestimator.service.canada.ca/en> (accessed on 23 April 2024).
- Government of the Netherlands (2024), *Fundamental Rights and Algorithms Impact Assessment (FRAIA)*, [92]
<https://www.government.nl/documents/reports/2021/07/31/impact-assessment-fundamental-rights-and-algorithms> (accessed on 23 April 2024).
- GPAI (2020), *Data Governance Working Group: A Framework Paper for GPAI's work on Data Governance*, [74]
<https://gpai.ai/projects/data-governance/gpai-data-governance-work-framework-paper.pdf> (accessed on 14 February 2023).
- Grace, E. (2018), *Enabling Citizens: A Two-Way Street...*, [103]
<https://www.wheel.ie/sites/default/files/media/file-uploads/2018-08/Two-Way%20Street.pdf> (accessed on 23 April 2024).
- Haas, A., M. McGuire Christian and R. Keuky (2021), *Supercharging customer service with AI: Reimagining AI strategies to unlock a true end-to-end customer experience*, [31]
<https://www2.deloitte.com/us/en/insights/topics/marketing-and-sales-operations/global-marketing-trends/2022/end-to-end-customer-experience-ai.html> (accessed on 23 April 2024).
- Human Rights Watch (2020), *Automated Hardship: How the Tech-Driven Overhaul of the UK's Social Security System Worsens Poverty*, [79]
https://www.hrw.org/sites/default/files/media_2020/09/uk0920_web_0.pdf (accessed on 23 April 2024).
- Hyee, R. et al. (2020), "How reliable are social safety nets?: Value and accessibility in situations of acute economic need", *OECD Social, Employment and Migration Working Papers*, No. 252, OECD Publishing, Paris, <https://doi.org/10.1787/65a269a3-en>. [5]
- ICO (2024), *What is automated individual decision-making and profiling?*, [10]
<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/> (accessed on 22 April 2024).

- Immervoll, H. (2010), “Minimum-Income Benefits in OECD Countries: Policy Design, Effectiveness and Challenges”, *OECD Social, Employment and Migration Working Papers*, No. 100, OECD Publishing, Paris, <https://doi.org/10.1787/218402763872>. [6]
- Inforing (2024), *Regiesoftware voor vroegsignalering van financiële problemen (Direction software for early detection of financial problems)*, <https://www.vroegsignalering.nl/nl/home> (accessed on 23 April 2024). [58]
- Instituto Mixto de Ayuda Social (2023), *Informe final Beneficio Temporal por Inflación año 2022 y 2023*, [https://www.imas.go.cr/sites/default/files/docs/Informe Final Beneficio Temporal por Inflación 2022 y 2023.pdf](https://www.imas.go.cr/sites/default/files/docs/Informe%20Final%20Beneficio%20Temporal%20por%20Inflaci%C3%B3n%202022%20y%202023.pdf). [53]
- ISSA (2023), *Towards customer-centric design and agile methodologies in social security institutions*, <https://www.issa.int/analysis/towards-customer-centric-design-and-agile-methodologies-social-security-institutions> (accessed on 23 April 2024). [102]
- ISSA (2022), *Artificial intelligence in social security institutions: The case of intelligent chatbots*, <https://www.issa.int/analysis/artificial-intelligence-social-security-institutions-case-intelligent-chatbots> (accessed on 23 April 2024). [26]
- ISSA (2021), *Priorities for social security: Trends, challenges and solutions*, <https://www.issa.int/sites/default/files/documents/2021-11/2-Americas%20priority%20report%20WEB-rev.pdf> (accessed on 23 April 2024). [50]
- ISSA (2020), *Artificial Intelligence in Social Security: Background and Experiences*, <https://www.issa.int/analysis/artificial-intelligence-social-security-background-and-experiences> (accessed on 23 April 2024). [27]
- ISSA (2017), *Using artificial intelligence (AI) to identify vulnerable Canadians*, <https://www.issa.int/gp/198044> (accessed on 23 April 2024). [37]
- Jacobsen, A. (2020), *Data-driven Fraud Detection: The Data Mining Unit (Den Fælles Dataenhed)*, <https://lisboncouncil.net/wp-content/uploads/2020/10/Annika-Jacobsen.pdf> (accessed on 23 April 2024). [55]
- Jacobsen, A. (2020), “Town Hall Meeting on Citizen Data in Focus: Machine Learning and Data Analytics”, https://www.youtube.com/watch?v=_ZZ-xK7Xp2k (accessed on 23 April 2024). [54]
- Janssen, M. and J. van den Hoven (2015), “Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy?”, *Government Information Quarterly*, Vol. 32/4, pp. 363-368, <https://doi.org/10.1016/j.giq.2015.11.007>. [96]
- Kela (2023), *Kelan chattirobotti vastaa nyt myös työttömyysturvaa koskeviin kysymyksiin (Kela’s chatbot now also answers questions about unemployment insurance)*, <https://www.kela.fi/ajankohtaista/5167813/kelan-chattirobotti-vastaa-nyt-myos-tyottomyysturvaa-koskeviin-kysymyksiin> (accessed on 23 April 2024). [29]
- Ko, W. and R. Moffitt (2022), “Take-up of Social Benefits”, *IZA Discussion Papers*, No. 15351, <https://docs.iza.org/dp15351.pdf> (accessed on 23 April 2024). [16]
- Kuźniacki, B. (2023), *The Dutch childcare benefit scandal shows that we need explainable AI rules*, <https://www.uva.nl/en/shared-content/faculteiten/en/faculteit-der-rechtsgeleerdheid/news/2023/02/childcare-benefit-scandal-transparency.html?cb&cb> (accessed on 23 April 2024). [83]

- Lane, M., M. Williams and S. Broecke (2023), “The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers”, *OECD Social, Employment and Migration Working Papers*, No. 288, OECD Publishing, Paris, <https://doi.org/10.1787/ea0a0fe1-en>. [112]
- Lassébie, J. (2023), “Skill needs and policies in the age of artificial intelligence”, in *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, <https://doi.org/10.1787/638df49a-en>. [100]
- Lokshin, M. and N. Umapathi (2022), *AI for social protection: Mind the people*, <https://www.brookings.edu/articles/ai-for-social-protection-mind-the-people/> (accessed on 23 April 2024). [84]
- Madiega, T. and A. Van De Pol (2022), *Artificial intelligence act and regulatory sandboxes*, European Parliamentary Research Service, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI\(2022\)733544_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf) (accessed on 26 April 2023). [109]
- Marc, C. et al. (2022), “Non-take-up of minimum social benefits: quantification in Europe A salient phenomenon still not making public policy headlines”, *DREES Reports*, No. 94, https://drees.solidarites-sante.gouv.fr/sites/default/files/2022-07/Non-take-up%20of%20minimum%20social%20benefits%20quantification%20in%20Europe.%20A%20salient%20phenomenon%20still%20not%20making%20public%20policy%20headlines_0.pdf (accessed on 23 April 2024). [14]
- Marr, B. (2018), *How The UK Government Uses Artificial Intelligence To Identify Welfare And State Benefits Fraud*, <https://www.forbes.com/sites/bernardmarr/2018/10/29/how-the-uk-government-uses-artificial-intelligence-to-identify-welfare-and-state-benefits-fraud/?sh=71b036840cb9> (accessed on 23 April 2024). [48]
- Meers, A. et al. (2017), *Lessons learnt about digital transformation and public administration: Centrelink’s online compliance intervention*, Commonwealth Ombudsman, https://www.ombudsman.gov.au/_data/assets/pdf_file/0024/48813/AIAL-OCI-Speech-and-Paper.pdf (accessed on 23 April 2024). [65]
- Milanez, A. (2023), “The impact of AI on the workplace: Evidence from OECD case studies of AI implementation”, *OECD Social, Employment and Migration Working Papers*, No. 289, OECD Publishing, Paris, <https://doi.org/10.1787/2247ce58-en>. [61]
- Misuraca, G. and C. Van Noordt (2020), *AI Watch - Artificial Intelligence in public services*, JRC, <https://doi.org/10.2760/039619> (accessed on 23 April 2024). [25]
- MoLSA (2023), *Interview with Czechia’s Ministry of Labour and Social Affairs, 15 November 2023*. [32]
- Murray, T., M. Cheong and J. Paterson (2023), *The flawed algorithm at the heart of Robodebt*, <https://findanexpert.unimelb.edu.au/news/62817-the-flawed-algorithm-at-the-heart-of-robodebt#:~:text=So%2C%20underpinning%20the%20Robodebt%20scheme,lives%20of%20the%20most%20disadvantaged.> (accessed on 23 April 2024). [66]
- Nadkarni, P., L. Ohno-Machado and W. Chapman (2011), “Natural language processing: an introduction”, *Journal of the American Medical Informatics Association*, Vol. 18/5, pp. 544-551, <https://doi.org/10.1136/AMIAJNL-2011-000464>. [3]

- Nolan, D., H. Maryam and M. Kleinman (2024), *The Urgent but Difficult Task of Regulating Artificial Intelligence*, <https://www.amnesty.org/en/latest/campaigns/2024/01/the-urgent-but-difficult-task-of-regulating-artificial-intelligence/> (accessed on 23 April 2024). [86]
- Nortal (2022), *Estonian Unemployment Insurance Fund prevents unemployment with Artificial Intelligence*, <https://nortal.com/insights/estonian-unemployment-insurance-fund-prevents-unemployment-with-artificial-intelligence/> (accessed on 23 April 2024). [40]
- OECD (2024), “Explanatory memorandum on the updated OECD definition of an AI system”, *OECD Artificial Intelligence Papers*, No. 8, OECD Publishing, Paris, <https://doi.org/10.1787/623da898-en>. [1]
- OECD (2024), *Modernising Access to Social Protection: Strategies, Technologies and Data Advances in OECD Countries*, OECD Publishing, Paris, <https://doi.org/10.1787/af31746d-en>. [12]
- OECD (2024), “Regulatory experimentation: Moving ahead on the agile regulatory governance agenda”, *OECD Public Governance Policy Papers*, No. 47, OECD Publishing, Paris, <https://doi.org/10.1787/f193910c-en>. [106]
- OECD (2024), *Social benefits to households* (indicator), <https://doi.org/10.1787/423105c6-en> (accessed on 22 April 2024). [4]
- OECD (2024), *Tax-benefit web calculator*, <https://www.oecd.org/els/soc/benefits-and-wages/tax-benefit-web-calculator/#d.en.500997> (accessed on 22 April 2024). [7]
- OECD (2023), *Evaluation of Belgium’s COVID-19 Responses: Fostering Trust for a More Resilient Society*, OECD Publishing, Paris, <https://doi.org/10.1787/990b14aa-en>. [19]
- OECD (2023), *Future of Social Protection: Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery*, *OECD Technology for Social Protection Questionnaire*. [22]
- OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/0655b570-en>. [88]
- OECD (2023), *Main Findings from the 2022 OECD Risks that Matter Survey*, OECD Publishing, Paris, <https://doi.org/10.1787/70aea928-en>. [18]
- OECD (2023), *OECD ICT Access and Usage by Households data*, <https://data-viewer.oecd.org/?chartId=0144c9cd-4669-476f-a762-3a4affc8317f>. [80]
- OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris, <https://doi.org/10.1787/40d53904-en>. [98]
- OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/e3d783b0-en>. [75]
- OECD (2022), “OECD Framework for the Classification of AI systems”, *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, <https://doi.org/10.1787/cb6d9eca-en>. [2]
- OECD (2021), *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation (OECD/LEGAL/0464)*, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464>. [93]

- OECD (2020), *Countering Fraud in Social Benefit Programmes: Taking Stock of Current Measures and Future Directions*, OECD Publishing, Paris, <https://doi.org/10.1787/71df2657-en>. [13]
- OECD (2019), *Recommendation of the Council on Artificial Intelligence*, OECD, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (accessed on 31 August 2021). [23]
- OECD (2019), *Society at a Glance 2019: OECD Social Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/soc_glance-2019-en. [21]
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [97]
- OECD (2018), “Unemployment-benefit coverage: Recent trends and their drivers”, in *OECD Employment Outlook 2018*, OECD Publishing, Paris, https://doi.org/10.1787/empl_outlook-2018-9-en. [17]
- OECD.AI (2024), *OECD AI Incidents Monitor (AIM)*, [https://oecd.ai/en/incidents?search terms=%5B%5D&and condition=false&from date=2014-01-01&to date=2024-02-27&properties config=%7B%22principles%22:%5B%5D,%22industries%22:%5B%5D,%22harm types%22:%5B%5D,%22harm levels%22:%5B%5D,%22harmed entities%22:%5B%5D%7D&only threats=false&order by=date&num results=20](https://oecd.ai/en/incidents?search%20terms=%5B%5D&and%20condition=false&from%20date=2014-01-01&to%20date=2024-02-27&properties%20config=%7B%22principles%22:%5B%5D,%22industries%22:%5B%5D,%22harm%20types%22:%5B%5D,%22harm%20levels%22:%5B%5D,%22harmed%20entities%22:%5B%5D%7D&only%20threats=false&order%20by=date&num%20results=20) (accessed on 23 April 2024). [110]
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/1f334543-en>. [44]
- Official Journal of the European Union (2016), *Regulation (EU) 2016/679 of the European Parliament and of the Council*, <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (accessed on 23 November 2021). [105]
- OPSI (2022), *Driving Artificial Intelligence in Public Services through accredited upskilling and project identification*, <https://oecd-opsi.org/innovations/ai-certification-ireland/> (accessed on 23 April 2024). [99]
- Overheid.nl (2024), *The Algorithm Register of the Dutch government*, <https://algoritmes.overheid.nl/en> (accessed on 23 April 2024). [94]
- Overheid.nl - The Algorithm Register (2024), *Early On*, <https://algoritmes.overheid.nl/en/algoritme/66453169> (accessed on 23 April 2024). [59]
- Pugh, J. (2023), *AI Could Be a Godsend for Accessing Public Benefits—Or a Nightmare*, <https://inthesetimes.com/article/ai-artificial-intelligence-food-stamps-snap-public-benefits> (accessed on 23 April 2024). [34]
- Raudla, M. (2020), *Töötukassa automatiseeritud infosüsteem otsustab raha jagamise sekunditega (The unemployment fund’s automated information system decides the distribution of money in seconds)*, <https://virumaateataja.postimees.ee/6904529/tootukassa-automatiseeritud-infosusteem-otsustab-raha-jagamise-sekunditega> (accessed on 23 April 2024). [41]

- Robinson, M. (2019), *Rise of the 'welfare robots': DWP tests artificial intelligence system to judge whether benefit claims are TRUE - as part of £8m-a-year plan to create 'virtual workforce' with world's first 'bot billionaire'*, <https://www.dailymail.co.uk/news/article-7574415/DWP-tests-artificial-intelligence-judge-benefit-claims-TRUE.html> (accessed on 23 April 2024). [49]
- Royal Commission into the Robodebt Scheme (2023), *Royal Commission into the Robodebt Scheme: Report*, <https://robodebt.royalcommission.gov.au/publications/report> (accessed on 23 April 2024). [67]
- SCIJ (2013), *Reglamento a la Ley N.º 9137, Creación del Sistema Nacional de Información y Registro Único de Beneficiarios del Estado*, http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=75585&nValor3=0&strTipM=TC. [52]
- The Lisbon Council (2020), *Data-Driven Public Service: The Road to Efficient and Trustworthy Government*, <https://lisboncouncil.net/newsandevents/data-analytics-for-citizen-support-and-government-savings/> (accessed on 23 April 2024). [56]
- Töötukassa (2024), *Data Protection Terms: When does the Unemployment Insurance Fund make an automatic decision regarding you*, <https://www.tootukassa.ee/en/data-protection-terms/data-protection-estonian-unemployment-insurance-fund/when-does-unemployment> (accessed on 23 April 2024). [42]
- Ubaldi, B. et al. (2019), "State of the art in the use of emerging technologies in the public sector", *OECD Working Papers on Public Governance*, No. 31, OECD Publishing, Paris, <https://doi.org/10.1787/932780bc-en>. [89]
- UNCTAD (2024), *Data Protection and Privacy Legislation Worldwide*, <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide> (accessed on 11 January 2023). [77]
- United Nations (2024), *Comprehensive cooperation to Costa Rica's National Information System and Unique Registry of Beneficiaries (SINIRUBE)*, <https://sdgs.un.org/partnerships/comprehensive-cooperation-costa-ricas-national-information-system-and-unique-registry> (accessed on 23 April 2024). [51]
- United Nations (1948), "Universal declaration of human rights", <https://www.un.org/en/about-us/universal-declaration-of-human-rights> (accessed on 20 October 2021). [64]
- US Congress (2022), *H.R.6580 - Algorithmic Accountability Act of 2022*, <http://www.congress.gov/> (accessed on 20 January 2023). [11]
- Verhagen, A. (2021), "Opportunities and drawbacks of using artificial intelligence for training", | *OECD Social, Employment and Migration Working Papers*, No. 266, <https://doi.org/10.1787/22729bd6-en> (accessed on 30 June 2022). [78]
- Wachter, S. and B. Mittelstadt (2019), "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", *Columbia Business Law Review*, Vol. 2, <https://papers.ssrn.com/abstract=3248829> (accessed on 26 April 2023). [76]

Annex A. Country responses to the questionnaire

Table A A.1. Countries' use of digital tools such as AI and/or administrative data for managing minimum income benefits, unemployment benefits, or other income support benefits

	Yes	No, but planning to	No, and not planning to	[No response on this question]	Other
Austria				X	
Australia				X	
Canada	X				
Colombia				X	
Czechia		X			
Lithuania	X ¹	X			
Norway			X ²	X	This has not been evaluated yet but will be part of the strategy implementation process
New Zealand				X	Some ADM use for social benefits
Slovak Republic		X			
United States	X				

Note: Country response to the question: "Does your government use digital tools, such as AI, and/or administrative data to predict entitlement to the following benefits, or to adjust benefit amounts depending on changes in personal circumstances? (Yes, No but planning to, No and not planning to)": Social assistance benefits, Other minimum income benefits, Unemployment benefits, Other income support benefits.

¹ Only for "other income support benefits".

² Only for social assistance. Added note: "This benefit is not suited for automatic evaluation, as it is mostly discretionary".

Source: OECD (2023^[22]), Future of Social Protection: Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery, OECD Technology for Social Protection Questionnaire.