Austrian Lab for AI Trust* Dossier 3

Predictive Analytics in the Allocation and Monitoring of State Transfer Payments

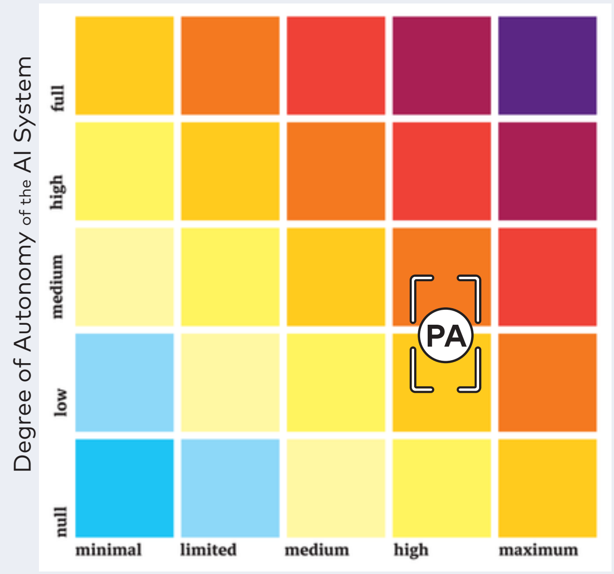
Executive Summary

The fair and transparent allocation of state transfer payments, such as family allowances, social welfare and housing benefits, presents significant challenges for public authorities and government bodies. Those affected should be helped as quickly and accurately as possible, while mistakes, inefficiency and discrimination must be avoided when processing applications. In practice, administrations often rely on predictive analytics systems to help solve these problems. As decisions supported by AI regarding the granting of state transfer payments have an immediate impact on citizens' lives, these decisions must be made with due care and caution.

In the ALAIT Risk Radar, the overall risk associated with the use of predictive analytics systems for assessing eligibility for state transfer payments is therefore classified as high to medium (dark and light orange, see illustration). This assessment is based on a high contextual risk (in accordance with the *EU AI-Act*) combined with a medium to low *degree of AI system autonomy*. In terms of autonomy, predictive analytics systems are generally not fully autonomous as clerks typically oversee the process and can intervene or make decisions at any time (Level 3, Human-on-the-Loop).

In terms of autonomy, predictive analytics systems are generally not fully autonomous as clerks typically oversee the process and can intervene or make decisions at any time (Level 3, Human-on-the-Loop).[1,8,18] The automated processing of applications leads to faster decisions and reduces the workload of clerks. However, these advantages face significant challenges: unlike humans, predictive analytics systems cannot contextualise individual decisions or consider citizens' specific life situations. This can lead to systemic unfairness If the training data is poor, AI systems introduce *systematic biases (AI Bias)* resulting in discrimination and incorrect decisions. Another problem is the lack of transparency and traceability in many AI systems. This so-called *black box problem* is particularly problematic for public authorities, who are obliged to provide reasoned decisions to their citizens and comply with legal requirements for the protection of basic rights, particularly the rights to privacy and non-discrimination.

ALAIT Risk Radar for Predictive Analytics in the Allocation and Monitoring of State Transfer Payments



AI Application in
Predictive Analytics

The ALAIT Risk Radar is a scientifically developed risk analysis tool for artificial intelligence (AI) that classifies AI applications in a context-specific manner, taking into account their degree of technical autonomy, thereby making the risk landscape visible to users at a glance. The principle is as follows: the higher the operational risk arising from the application context and the greater the degree of autonomy of the AI system with regard to decision-making, the higher the risk associated with its use. An extended bracket indicates a range within the risk classification. A low degree of autonomy in an AI system does not mean that one can simply sit back and relax. It requires a strong role on the part of the people using it. (For details on the tiered model, see p. 8ff).

Public authorities and system developers can reduce these risks by implementing the relevant risk mitigation measures listed in this article. In principle, only systems that comply with the European Convention on Human Rights should be used. Such systems must be trained using datasets that are as free from bias and discrimination as possible, and they must be state of the art. Additionally, administrative staff must ensure that enquiries and complaints from citizens are adequately addressed through human control and supervision. Clerks in government bodies must

**ALAIT is a research project initiated by the Federal Ministry for Innovation, Mobility and Infrastructure (BMIMI) aimed at developing a novel form of technology assessment for AI technologies using scientific methods.*

receive ongoing training regarding the functionality and limitations of predictive analytics systems, as well as social and legal requirements.

Introduction

Government transfer payments refer to financial support such as family allowances, social welfare and student grants, which are provided to eligible citizens without any direct quid pro quo. These government transfer payments are intended to provide security for citizens and promote equal opportunities.[22,23] At the same time, they must be allocated efficiently, fairly and in accordance with the law – a requirement that is becoming increasingly important in view of scarce public resources and the public debate on distributive justice. Therefore, complex eligibility assessments are increasingly supported by artificial intelligence (AI) systems such as predictive analytics and automated decision-making systems (ADM). Predictive analytics uses historical data and statistical methods to identify patterns and predict future events.[1] Thus, predictive analytics systems provide a basis for decision-making, but do not make decisions themselves. ADMs take this a step further: they use the information generated by predictive analytics or other AI systems to independently take concrete decisions based on defined criteria or objectives – for example, in the automatic approval or refusal of an application.[6]

The main advantage of these technologies lies in their ability to analyze large amounts of data in a short period of time and identify patterns or make predictions about future events.[2] Employees can use them to detect anomalies such as fraud and money laundering more quickly and accurately. In Austria, the Ministry of Finance has set up a Predictive Analytics Competence Centre (PACC) to support the authorities in controlling tax and subsidy payments to businesses.[3] Other countries already use predictive analytics and AES to distribute government transfer payments or detect social fraud.[4,5] For example, the Swedish Board of Student Finance (CSN) uses a system that, in most cases, makes fully automated decisions about the payment of financial support to students. The system combines information from the Board of Student Finance with publicly available information in Sweden, such as tax data and personal data of applicants. As agency employees only need to intervene in the event of an appeal, the system ideally offers a high quality of service and speed and, therefore, is also rated positively by applicants.[5] However, this approach shifts the burden of verifying the accuracy of decisions onto citizens.

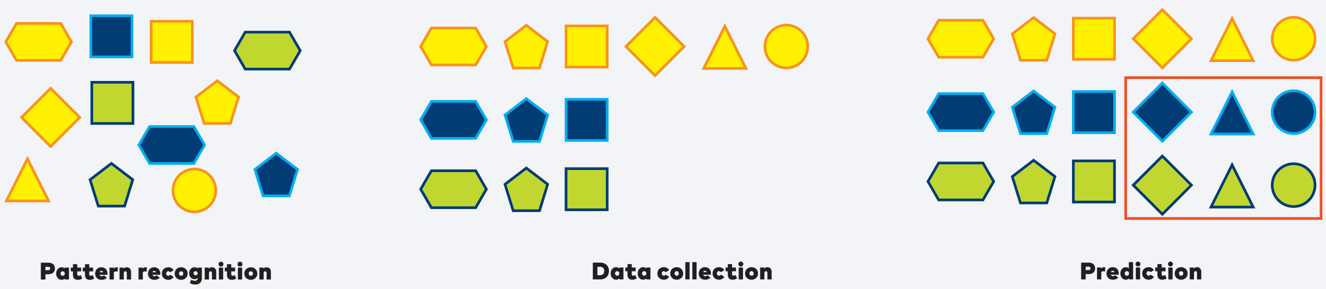
Despite the advantages of these systems, their introduction can also pose major challenges, which in any case require careful consideration of the opportunities and risks and appropriate measures.[5,9] The use of AI systems is not always an appropriate solution, especially when fundamental rights are violated or vulnerable population groups are disadvantaged. Studies show that many of these systems tend to disproportionately negatively affect socially disadvantaged population groups, as they are over- or underrepresented in the datasets and are thus systematically discriminated against.[34] A striking example is the 'SyRI' system used in the Netherlands to detect social fraud. It violated the right to privacy according to the European Convention on Human Rights and was suspended after massive criticism and a court decision. In addition, it delivered flawed results, led to a lack of transparency and fairness, and had serious consequences for thousands of affected families.[10,11,12]

Technology Description

Predictive Analytics often uses Machine Learning (ML), a method that analyses training data to predict future events and manage risks.[13,14] At its core, machine learning is based on pattern recognition and classification using labeled data (***supervised learning***) or – less commonly in the field of predictive analytics – unlabeled data (***unsupervised learning***).[8,20] Machine learning tools include:[21]

1. **Artificial neural networks**: are computer models consisting of many interconnected nodes ("neurons"). These nodes apply weighted mathematical functions to inputs, transform them using activation functions, and pass the results on to subsequent layers..
2. **Decision trees**: are tree-like structures that make step-by-step decisions based on the characteristics of the data. Each node represents a condition that leads to different branches and finally to a classification or prediction.
3. **Support vector machines (SVM)**: are supervised learning methods that assign data points to classes by finding an optimal separating line (hyperplane). One example is the distinction between spam and non-spam emails based on characteristics such as certain words, record lengths or sender information.

The following diagram shows the general operating principle of predictive analytics with machine learning:



First, large amounts of heterogeneous data from various sources are collected and processed. Then, using ML algorithms, patterns, correlations, and recurring structures within this data are identified. Based on these patterns, future events can be predicted and recommendations for action can be derived.

Opportunities

Die Nutzung von Predictive Analytics in der Vergabe und The use of predictive analytics in the allocation and control of government transfer payments can bring several benefits for the respective stakeholders:

Opportunities for Citizens:

- **Quick decisions:** Unlike manual processing, machine-based decisions offer the advantage of high speed. Citizens can quickly receive information about their eligibility.[1]
- **Consistency in decision-making:** A predictive analytics system that is based on careful and realistic model assumptions, trained with high-quality data and set up correctly, ideally provides a transparent, compliant, fair and consistent basis for decision-making.[1]

Opportunities for Public Sector Employees:

- **Decision Support and Workload Reduction:** With the help of predictive analytics systems, the Predictive Analytics Competence Centre (PACC) achieved a reduction in workload for its employees of up to 40%.[20] In particular, by automating time-consuming and repetitive tasks, automated decision-making systems offer the potential to minimize routine administrative tasks for administrative staff, thereby freeing up more time for more challenging tasks.[1]

Opportunities for Government Organisations:

- **Scope and speed:** Machine learning can automate many administrative tasks, thereby significantly reducing staff requirements. Considering tight budgets and the current and future shortage of qualified personnel, the use of predictive analytics offers an way to reduce costs and speed up administrative processes by up to 31 %.[1,13,14]

- **Improved fraud control:** The key advance of predictive analytics systems is that they can analyze much larger and more complex data sets in real time. This reveals risk patterns that cannot be detected at all with conventional control mechanisms, or only with considerable effort. The risk assessment tool ARACHNE, developed by the European Commission, supports the European Social Fund (ESF) and the European Regional Development Fund (ERDF) in detecting fraud and mismanagement in the use of public funds.[2] Predictive analytics tools are also used by the Italian government, for example, to detect corruption and money laundering in public procurement, thereby helping to successfully prevent mafia activities.[2,13] In addition, the European Union is currently developing AI systems to monitor compliance with EU law. For example, the Claudette project has developed a predictive analytics system that identifies unlawful or unfair clauses in contracts and data protection policies.[35]

- **Evidence-based policy making:** Predictive analytics systems allow conclusions to be drawn about likely future developments from historical data. This information can be used in a variety of ways in policy-making (so-called 'evidence-based' policy-making):[8,10]
  1. Identification of needs and analysis of the effects of monetary programmes.
  2. Creation of prognoses and scenarios for the design of customized prevention programmes.
  3. Continuous reintegration of data enables ongoing evaluation and adaptation of programmes (so-called policy learning).
  4. Development of early warning systems for initiating government measures; for example, some municipalities in Denmark and Spain are experimenting with prediction models for the need for assistance among older citizens so that appropriate measures can be initiated in a timely manner.[10]

Challenges and Risks

Despite the advantages mentioned, the use of predictive analytics to monitor and allocate government transfer payments is also subject to criticism and challenges.

1. **"Surveying" Citizens:** A European study carried out with the Austrian Academy of Sciences shows that the use of predictive analytics can lead to citizens being "monitored" and reduced to data that can be compared with average values. Important contextual information from specific cases may be neglected. The discrepancy between algorithms and people's concrete social reality can lead to unfair results.[29,32]

2. **Increase in Inequality and *Bias (AI Bias)*:** Biased predictive analytics applications in the social welfare system often target the very people it is supposed to protect. Several projects across Europe have come under significant criticism due to problems with biased automated algorithms.[4,7,10,12] The above example of the Dutch SyRI system is the most prominent case illustrating the drastic effects of such systems on individuals and society through bias.

3. **Data Protection:** By collecting data from income and employment records, travel histories, social media activities, etc., governments could lay the basis for extensive surveillance of their citizens..[4,13] Even if the data collection were covered by law, it would usually take place without the knowledge or explicit consent of the individuals concerned.[4] To make matters worse, anonymised data from separate databases can be de-anonymised by combining it with other data.[8]

4. **Data Quality and Misjudgements:** To ensure ***accurate*** predictions, the collected data must be organised in a specific way. However, data sets from public authorities are often confusing, outdated, or scattered across different departments. The European Parliament has found that inconsistent data (e.g. in the subsidy records of EU Member States) makes it extremely difficult to apply predictive analytics tools uniformly.[2] Incorrect pre-processing of data can lead to systematic misjudgements, e.g. denying support to those who are eligible (false negatives) or granting support to those who are not (false positives), which can have significant consequences for applicants. For instance, the Copenhagen City Council's AI system for predicting

post-hospital care needs has only been shown to be accurate four times out of five.[17] The potential efficiency gains for public authorities must not justify the serious consequences for affected citizens resulting from a misjudgement by an AI system.[10]

5. **Transparency:** The transparency of predictive analytics systems should ensure that employees can understand decisions made by predictive analytics systems to identify errors and discrepancies and be able to take responsibility for these decisions. However, from a technical point of view, the problem remains that most modern AI systems are not transparent and traceable due to their complexity – even for experts – a phenomenon also known as „***black box***".[8,16] Furthermore, the use of AI tools requires appropriate expertise and training, which must be ensured by the authorities and companies that seek to use such tools (see Article 4, ***EU AI Act***).[25]

6. **Overreliance on Technology:** Several studies have shown that people tend to trust AI systems more than humans. This phenomenon is called "automation bias".[5] This applies both to administrative staff who work with AI tools and to citizens who use governmental social benefits and services. There is a risk that people will no longer adequately question the suggestions or decisions made by predictive analytics systems, even when they have conflicting information and knowledge about a specific case.[8] This means that support by AI systems does not automatically lead to better decisions, as the current overview study by the Massachusetts Institute of Technology (MIT) from 2024 shows.[19]

Recommendations for Practical Use

For Public Authorities (System Users):

1. **Premise Human Oversight:** Human supervision and final decision-making at critical points in the decision-making process take advantage of predictive analytics systems for the allocation and control of government transfer payments and help to mitigate existing risks. This "human-in-the-loop" approach emphasises the supportive role of predictive analytics in public administration, rather than replacing humans with machines.[1] Furthermore, the focus should be on collaboration between humans and machines. In future, a new collaborative design involving multiple stakeholders will be required to enhance, rather than undermine, socio-technical synergies.[29] Explainability and ***transparency*** should also be prerequisites for supervision. Human supervision forms the basis of an ethical AI approach, enabling contextualised decisions that protect citizens' rights and prevent algorithmic errors.[16]

2. **Building and strengthening AI knowledge:** The role of the state and the authorities in dealing with and utilising AI involves setting a strong example. This means that alongside the (increasing) use of AI within the administrative apparatus, the knowledge and competences of administrative staff must be strengthened as well. Article 4 of the ***EU AI Act*** states: "Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used."[25]

3. **Precise and high-quality data:** The reliability of predictive analytics systems depends largely on the quality of the datasets used to train the system or supply it during ongoing operations. Beyond that, the datasets should be diversified to improve accuracy. For the use of data in the field of predictive analytics, the following principles must also be followed:[11,13,24]

    a. **Data minimisation:** It is recommended that only the minimum amount of personal data necessary to fulfil the task be used.

a. **Purpose limitation:** Data collected for the provision of public services should not be used arbitrarily for other purposes

4. **Alignment with guidelines and implementation of risk assessment tools prior to introduction:** All predictive analytics systems must be assessed to ensure that they comply with the guidelines of the General Data Protection Regulation (GDPR) and the ***EU AI Act***. The European AI regulation supplements the provisions of the GDPR and requires public authorities to assess the risks of the system in terms of violations of fundamental rights and to take appropriate measures prior to the implementation of predictive analytics tools.[1,8,16,29] Furthermore, bodies such as the European Office for Artificial Intelligence ('AI Office') and the European Board for Artificial Intelligence ('AI Board') monitor the implementation and application of the ***EU AI Act*** at Union level.[33]

5. **Low-threshold compliant channels:** It must be made clear that a person has the right and the option to appeal against a decision made by an automated system (Articles 13 and 52 of the ***EU AI Act***).

For System Developers:

1. **Transparency:** Besides creating knowledge, transparency is seen as a key factor in building trust in the use of AI. It is crucial to ensure that the use of AI in public administration, where it has a direct impact on those affected – namely citizens – adheres to the maximum transparency criteria. ***Explainable AI (XAI)***, which is designed to make the decisions and predictions of AI systems comprehensible and traceable to humans, can be valuable in this context.[36] Such a system can provide explanations that are understandable to humans by indicating which input factors had the greatest influence on the system's output (e.g. income level or housing cost burden). In this way, case handlers can check, validate and override automated suggestions, ensuring that the final decision remains under human control.[36]

2. **Avoidance of specific personal characteristics:** Determining exactly which characteristics of an individual are processed in algorithmic systems and contribute to decision-making is not optional, but mandatory.[29] Particular care must be taken to ensure that the use of measurable and non-measurable characteristics (proxy variables), such as nationality or ethnic origin, is not uncritically embedded in AI tools used for the allocation and monitoring of public transfer benefits. AI tools in general, but particularly in this area of application, are subject to a special obligation not to reproduce or even reinforce existing social prejudices and stereotypes.[11]

For Citizens:

1. **Building and strengthening AI literacy:** Knowledge of AI technologies, development of key skills, and awareness of rights are essential.[8] In Austria, the Rundfunk- und Telekom Regulierungs-GmbH (RTR) serves as an information and contact point for the general public. The AI Service Centre of RTR assists citizens with questions regarding transparency, rights relating to the use of AI systems, or the review of AI decisions.[33]

2. **Right to *transparency*:** Citizens have the right to be informed about how and when their personal data is used and into which systems it is transferred (Art. 15 GDPR).[13,31] The contact point in cases of suspected violations of data protection rights is the Austrian Data Protection Authority (DSB).[27,28]

3. **Possibility to review an AI decision:** Where a decision has been made with the help of automated systems and produces legal or similarly significant effects, such as approval, reduction, or recovery of benefits, citizens have the right not to be subject to a decision based solely on automated processing, including profiling, but also to obtain human review (Art. 22 GDPR).[26] In case of a complaint against a public authority, the complaint may be submitted to the Austrian Ombudsman Board.[30]

Important Terms

EU AI Act (European Union Act on Artificial Intelligence): The European regulation on artificial intelligence (AI) – the first ever comprehensive regulation on AI issued by a major regulatory authority. It focuses in particular on high-risk AI systems.

AI autonomy: The ability of an AI system to achieve a set of objectives, amidst a range of uncertainties in its environment, independently and without external intervention.

AI accuracy: Refers to an AI system's ability to make correct predictions or decisions. It is an important measure of its performance and is crucial for determining its effectiveness and reliability.

Systematic bias: Bias is the systematic difference in the treatment of certain objects, individuals or groups compared with others. Treatment refers to any kind of action, including perception, observation, representation, prediction or decision-making.

Transparency: This means that the functionality, decision-making processes and areas of application of an AI system are comprehensible, explainable and openly accessible – for developers, users and other stakeholders.

Supervised learning: A machine learning method in which the training data is pre-labeled and the system learns the pattern between the image and the label. The task of such an AI system is to find a relationship that maps each input of the training set (the data) to an output (the label).

Unsupervised learning: A subfield of machine learning in which an algorithm recognizes patterns, structures, or relationships in unlabeled data without being told what is right or wrong.

Black-box systems: This term refers to AI systems whose internal decision-making processes are neither transparent nor comprehensible to humans. This means that only inputs and outputs can be observed, without understanding the precise processing that occurs in between.

Data privacy in AI: The entirety of practices and concerns related to the ethical collection, storage, and use of personal data by artificial intelligence systems.

Explainable AI (XAI): Refers to methods and techniques that make it possible to better understand and account for the decisions and outcomes of AI systems.

Explanation of the ALAIT Risk Radar’s tiered model

The ALAIT AI Risk Radar illustrates the relationship between application risk and the autonomy of an AI system. The risk levels are based on the EU AI Act, in particular Article 6 and Annex III, which deal with high-risk areas of AI application. Lower levels of system autonomy and application risks are represented by cooler colours (blue), whilst higher levels of system autonomy and application risks are represented by warmer colours (red).

The change of color conveys the increased risk of AI applications. This color scale can be used to identify the overall risk: violet and dark red – very high, red and dark orange – high, light orange and yellow tones – medium, blue tones – low overall risk. Ideally, high application risks and system autonomy should be avoided or only be used after very careful consideration.

Level of Autonomy of the AI System

Level 1: No Autonomy

AI is a passive tool; it is people who make all the decisions and take action.

Example: Diagnostic systems that display raw medical data or analyse the data (without providing recommendations!).

Recommended use cases: Scenarios involving high risks or in which ethical decisions are of crucial importance (e.g. medical diagnostics, the justice system).

Level 2: Low Degree of Autonomy (Human-in-the-Loop)

The AI provides recommendations or options, but users remain responsible for selecting and approving actions.

Example: AI suggests optimal routes for logistics, or recommendation systems in e-commerce.

Recommended use cases: Tasks of medium complexity involving moderate risks (e.g. supply chain optimisation).

Level 3: Medium Degree of Autonomy (Human-on-the-Loop)

The AI carries out certain tasks autonomously, with humans intervening in exceptional cases.

Example: AI-supported manufacturing processes where the system controls the machines, but users intervene in the event of anomalies.

Recommended use cases: Scenarios in which continuous human involvement is not required, but critical risks necessitate human oversight (e.g. industrial automation, monitoring of financial transactions).

Level 4: High Degree of Autonomy (Human in Control)

The AI system operates largely autonomously, but allows users to override it themselves in order to avoid undesirable results.

Example: Autonomous vehicles.

Recommended use cases: Low- to medium-risk environments (e.g. logistics, basic traffic management).

Level 5: Full Autonomy with Minimal Supervision

The AI system operates independently and requires minimal or no human intervention.
Human involvement is limited to long-term audits.

Example: Autonomous agricultural machinery, AI for electricity distribution, underground railways, airport rail links.

Recommended use cases: Environments with low safety or ethical risks and high reliability of the AI system (e.g. repetitive tasks in controlled environments).

Scope of Application Risk

Stage 1: Minimal Risk

The AI system has no impact on the user or the decision-making process.

Examples: Filters, NPCs in computer games, recommendationalgorithms without serious consequences (DeepL, other translation tools).

Criteria: No direct impact on the high-risk areas of the EU AI Act.

Stage 2: Limited Risk

AI systems that interact with users but do not make high-risk decisions. The risk increases when there is a lack of transparency regarding the involvement of AI.

Examples: AI systems that interact with users but do not make high-risk decisions. The risk increases when there is a lack of transparency regarding the use of AI.

Criteria: Areas not included in the 'high-risk' list set out in the EU AI Act.

Stage 3: Medium Risk

Artificial intelligence systems do not have any particular impact on individuals; rather, their effects are felt at a collective or societal level.

Examples: Generative AI such as ChatGPT and other systems that can indirectly influence the environment in which they are used and may ultimately lead to major changes in society and people's lives.

Criteria: AI systems that are available for public use and have the potential to influence and, in the long term, change existing practices.

Stage 4: High Risk

Any algorithm used in what the EU AI Act defines as 'high-risk areas' or which has a direct impact on the life and limb of individuals.

Examples: Medicine, biometrics, critical infrastructure, education and vocational training, employment, access to services, law enforcement, migration.

Criteria: Classification as a 'high-risk sector' under the EU AI Act is only applicable if the rules on transparency and data quality are complied with.

Stage 5: Extreme Risk

Any algorithm used in what the EU AI Act defines as 'high-risk areas'.

Examples: Medicine, biometrics, critical infrastructure, education and vocational training, employment, access to services, law enforcement, migration.

Criteria: Classification as a 'high-risk area' if the rules on transparency and data quality are NOT complied with.

Sources

1 OECD. (2024). Using AI to manage minimum income benefits and unemployment assistance: Opportunities, risks and possible policy directions (OECD Artificial Intelligence Papers No. 21; OECD Artificial Intelligence Papers, Vol. 21). https://www.oecd.org/en/publications/using-ai-to-manage-minimum-income-benefits-and-unemployment-assistance_718c93a1-en.html

2 Schwarscz, A. (2021). Use of big data and AI in fighting corruption and misuse of public funds—Good practice, ways forward and how to integrate new technology into contemporary control framework.

3 Bundesministerium Finanzen. (n.d.). Predictive Analytics Competence Center. Retrieved 15 July 2025, from https://www.bmf.gv.at/themen/betrugsbekaempfung/einheiten-betrugsbekaempfung/Predictive-Analytics-Competence-Center.html

4 Amnesty International. (2024, November 12). Denmark: AI-powered welfare system fuels mass surveillance and risks discriminating against marginalized groups – report. Amnesty International. https://www.amnesty.org/en/latest/news/2024/11/denmark-ai-powered-welfare-system-fuels-mass-surveillance-and-risks-discriminating-against-marginalized-groups-report/

5 Wihlborg, E., Larsson, H., & Hedström, K. (2016). 'The Computer Says No!' – A Case Study on Automated Decision-Making in Public Authorities. 2016 49th Hawaii International Conference on System Sciences (HICSS), 2903–2912. https://doi.org/10.1109/HICSS.2016.364

6 Turdibayeva, K. (2025, February 7). What is Automated Decision-Making? ProcessMaker. https://www.processmaker.com/blog/what-is-automated-decision-making/

7 Allhutter, D., Cech, F., Fischer, F., Grill, G., & Mager, A. (2020). Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective. Frontiers in Big Data, 3. https://doi.org/10.3389/fdata.2020.00005

8 Thapa, B. (2019). Predictive Analytics and AI in Governance: Data-driven government in a free society. The European Liberal Forum. https://liberalforum.eu/publication/predictive-analytics-and-ai-in-governance-data-driven-government-in-a-free-society

9 Hoffman, B. (2024, March 10). Automation Bias: What It Is And How To Overcome It. Forbes. https://www.forbes.com/sites/brycehoffman/2024/03/10/automation-bias-what-it-is-and-how-to-overcome-it/

10 AlgorithmWatch. (2019a). How Dutch activists got an invasive fraud detection algorithm banned. AlgorithmWatch. https://algorithmwatch.org/en/syri-netherlands-algorithm/

11 Dutch scandal serves as a warning for Europe over risks of using algorithms. (2022, March 29). POLITICO. https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/

12 Henley, J., & Booth, R. (2020, February 5). Welfare surveillance system violates human rights, Dutch court rules. The Guardian. https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules

13 Bhupatiraju, S., Chen, D., Jankin, S., Kim, G., Kupi, M., & Maqueda, M. R. (2023). Government Analytics Using Machine Learning.

14 Boulieris, P., Pavlopoulos, J., Xenos, A., & Vassalos, V. (2024). Fraud detection with natural language processing. Machine Learning, 113(8), 5087–5108. https://doi.org/10.1007/s10994-023-06354-5

15 Joshi, R., Nair, V., Singh, M., & Nair, A. (2021). Leveraging Natural Language Processing and Predictive Analytics for Enhanced AI-Driven Lead Nurturing and Engagement.

16 Santiso, C. (n.d.). Public Governance in the Age of Artificial Intelligence. Retrieved 1 July 2025, from https://www.chandlerinstitute.org/governancematters/public-governance-in-the-age-of-artificial-intelligence

17 AlgorithmWatch. (n.d.). DENMARK. AlgorithmWatch. Retrieved 1 September 2025, from https://algorithmwatch.org/en/automating-society-2019/denmark/

18 AlgorithmWatch. (2019b). NETHERLANDS. AlgorithmWatch. https://algorithmwatch.org/en/automatingsociety-2019/netherlands/

19 Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human Behaviour, 8(12), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

20 Kölem, Ö. Interview über Predictive Analytics in PACC. Interview by Ilya Faynleyb, 19. Aug. 2025. Personal interview.

21 Kumar, V., & L., M. (2018). Predictive Analytics: A Review of Trends and Techniques. International Journal of Computer Applications, 182(1), 31–37. https://doi.org/10.5120/ijca2018917434

22 Stadt Wien. (n.d.). Förderhandbuch. Retrieved 27 August 2025, from https://www.wien.gv.at/spezial/foerderhandbuch/allgemeines-zu-forderungen/begriff-der-forderung-was-ist-eine-forderung/

23 Bundeszentrale für politische Bildung. (n.d.). Transferzahlungen. bpb.de. Retrieved 27 August 2025, from https://www.bpb.de/kurz-knapp/lexika/lexikon-der-wirtschaft/20866/transferzahlungen/

24 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance), 119 OJ L (2016). http://data.europa.eu/eli/reg/2016/679/oj/eng

25 Article 4: AI literacy | EU Artificial Intelligence Act. (n.d.). Retrieved 5 September 2025, from https://artificialintelligenceact.eu/article/4/

26 Art. 22 GDPR – Automated individual decision-making, including profiling. (n.d.). General Data Protection Regulation (GDPR). Retrieved 5 September 2025, from https://gdpr-info.eu/art-22-gdpr/

27 Unternehmensberatung, A. (n.d.). § 77 DSGVO (Datenschutz-Grundverordnung)—JUSLINE Österreich. Retrieved 5 September 2025, from https://www.jusline.at/gesetz/dsgvo/paragraf/77

28 Datenschutzbehörde, Ö. (n.d.-b). Österreichische Datenschutzbehörde. Österreichische Datenschutzbehörde. Retrieved 5 September 2025, from https://dsb.gv.at/

29 Allhutter, D., Alushi, A., Cavalcanti De Alcântara, R., Männiste, M., Pentzold, C., & Sosnowski, S. (2024). Public value in the making of automated and datafied welfare futures. Internet Policy Review, 13(3). https://doi.org/10.14763/2024.3.1803

30 Volksanwaltschaft—Beschwerdewegweiser. (n.d.). Retrieved 7 September 2025, from https://volksanwaltschaft.gv.at/fuer-probleme-mit-behoerden/beschwerdewegweiser/

31 Art. 15 DSGVO – Auskunftsrecht der betroffenen Person. (n.d.). Datenschutz-Grundverordnung (DSGVO). Retrieved 7 September 2025, from https://dsgvo-gesetz.de/art-15-dsgvo/

32 Macht Digitalisierung den Sozialstaat gerechter? (2024, May 6). scilog - Das Magazin des Wissenschaftsfonds FWF. https://scilog.fwf.ac.at/magazin/macht-digitalisierung-den-sozialstaat-gerechter

33 KI-Servicestelle der RTR. (n.d.). RTR. Retrieved 16 September 2025, from https://www.rtr.at/rtr/service/ki-servicestelle/ki-servicestelle.de.html

34 Petheram, A. (2019, September 19). Predictive Analytics, Public Services and Poverty. Oxford Insights. https://oxfordinsights.com/insights/predictive-analytics-public-services-and-poverty/

35 Sartor, G. (2025). Artificial Intelligence for Monitoring the Application of EU Law

36 Mehdiyev, N., Houy, C., Gutermuth, O., Mayer, L., & Fettke, P. (2021). Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes – On the Potential of XAI in Tax Audit Processes. In F. Ahlemann, R. Schütte, & S. Stieglitz (Eds), Innovation Through Information Systems (pp. 413–428). Springer International Publishing. https://doi.org/10.1007/978-3-030-86790-4_28

About the ALAIT Project

The Austrian Lab for AI Trust (ALAIT) is a research and development project initiated by the Austrian Federal Ministry for Innovation, Mobility and Infrastructure (BMIMI) to build trust through knowledge in the field of artificial intelligence (AI). The ALAIT project aims to empower interested individuals and key societal groups to use AI technologies responsibly and to establish ethical and high-quality standards for the application of AI.

The project is led by **winnovation** (Gertraud Leimüller and Lena Müller-Kress) and implemented by a consortium including **leiwand.ai** (Rania Wazir and Silvia Wasserbacher-Schwarzer), **TU Vienna** (Sabine Köszegi and Ilya Faynleyb) and **Austria Press Agency – APA** (Verena Krawarik and Sophia Marecek).

The ALAIT dossiers are available on the project homepage: https://science.apa.at/project/alait/

The contents of this dossier reflect the current state of the art and have been carefully compiled according to scientific criteria. However, they do not constitute legally binding information or advice.

Imprint

Media owner and publisher:
winnovation consulting gmbh
Linke Wienzeile 124/5
1060 Vienna

Published in 2025, funded by the Austrian Federal Ministry for Innovation, Mobility and Infrastructure.

Bundesministerium Innovation, Mobilität und Infrastruktur

Acknowledgments:

We would like to thank the following expert for their helpful feedback on earlier versions of this dossier: Doris Allhuter