Politecnico di Torino

Responsible Artificial Intelligence

01DTEOV

A critical analysis of the ADM system of “AMS Austria” with the use of Value Sensitive Design and Critical Systems Heuristics

Candidates:
Festa Shabani (s317640)
Farisan Fekri (s308066)
Ibrahim Adouni (s332398)
Giacomo Fantino (s310624)

Referee:
Professor Antonio Vetrò

Introduction

The Austrian Public Employment Service (AMS) is a government agency responsible for implementing labor market policies in Austria. It provides services such as job placement, career counseling, and unemployment benefits [1]. In recent years, AMS introduced an automated decision-making (ADM) system designed to enhance the efficiency of these services by categorizing job seekers based on their likelihood of finding employment. This ADM system, commonly referred to as the AMS algorithm, assigns scores to individuals to prioritize resource allocation, intending to optimize employment outcomes. Depending on the score, job seekers will land in one of three groups: group A for people with high chances to find a job within half a year, group B for people with mediocre prospects on the job market who might benefit from retraining, and group C for people deemed unemployable, who will receive less help from AMS and may be discharged to other institutions.

However, the implementation of the AMS algorithm has sparked significant controversy and criticism. Critics argue that the algorithm perpetuates discrimination, particularly against women, older job seekers, and individuals with disabilities. Concerns have been raised about the transparency and fairness of the algorithmic decision-making process, with accusations that it systematically disadvantages already vulnerable groups.

In 2020, the Austrian Data Protection Authority (DPA) ruled against the AMS algorithm, citing violations of privacy laws due to its automated decision-making process lacking sufficient human oversight. This ruling led to the suspension of the algorithm's use, demanding revisions to ensure compliance with data protection regulations. Furthermore, organizations such as AlgorithmWatch have highlighted the potential for such systems to contravene EU anti-discrimination laws and emphasized the need for greater transparency and accountability in their deployment [6].

This analysis will explore the AMS ADM system through the lens of Value Sensitive Design (VSD) and Critical Systems Heuristics (CSH). These frameworks will be used to evaluate the ethical, legal, and social implications of the algorithm, assessing whether it aligns with principles of fairness, accountability, and transparency. By examining the criticisms and defenses of the AMS algorithm, this paper aims to provide a comprehensive understanding of the challenges and opportunities associated with automated decision-making in public employment services.

Driving Factors

In the last years there have been many initiatives on ethics and governance of AI:

- In April 2019 the European Commission issued the Ethics Guidelines for Trustworthy AI [11]. In this article they highlighted the ethical principles of AI, including human agency, privacy, non-discrimination and fairness.

- In the same year Australia published the Australian ethical AI principles [12], giving more emphasis on human values and human rights.
- Later the Montreal Declaration for a Responsible Development of AI [13] introduced additional aspects like intimacy, democratic participation, and solidarity principles.

- Last year the European Declaration on Digital rights [14] consolidated these principles for all digital technologies, including a chapter on the freedom of choice when interacting with algorithms and AI systems.

We hope that our work contributes to the development of ADM systems aligned with these declarations, promoting a more human-centered approach to AI.

Value Sensitive Design (VSD)

STEP A: Envision

A1. Direct Stakeholders and Their Concerns

1. **Job Seekers:** Gaining fair access to job opportunities and support resources to improve employment prospects.
2. **AMS Counselors:** Effectively assisting job seekers while managing their workload and maintaining job satisfaction.
3. **AMS Management:** Enhancing the efficiency and effectiveness of employment services to meet organizational goals and public expectations.

A2. Indirect Stakeholders and Their Concerns

1. **Families of Job Seekers:** Supporting family members to achieve stable employment and economic well-being.
2. **Employers:** Accessing a qualified and diverse pool of job candidates to meet their staffing needs efficiently.
3. **Educational and Training Providers:** Receiving accurate information on the training needs of job seekers to tailor programs effectively.

STEP B: Speculate

B1. Potentially Implicated Values

1. **Fairness**
    - Ensuring that the algorithm does not perpetuate existing biases or create new forms of discrimination based on gender, age, ethnicity, or other personal characteristics.
2. **Transparency**

- Making the workings of the algorithm clear and understandable to all stakeholders, including how decisions are made and on what data they are based.

3. **Accountability**
    - Holding the AMS and its developers responsible for the outcomes and impacts of the algorithm, ensuring they can answer for any negative consequences.
4. **Privacy**
    - Protecting the personal data of job seekers, ensuring that their information is used ethically and securely.
5. **Efficiency**
    - Improving the effectiveness, better allocating the resources needed to help job seekers.
6. **Trust**
    - Building and maintaining public confidence in the AMS algorithm and the broader employment services it supports.
7. **Autonomy**
    - Respecting the individual choices and self-determination of job seekers in their career paths and the support they receive.
8. **Human Dignity**
    - Recognizing and upholding the intrinsic worth of all job seekers, treating them with respect and consideration in the algorithm's operations.
9. **Inclusivity**
    - Designing the algorithm to consider and accommodate the diverse backgrounds and needs of all job seekers.
10. **Responsibility**
    - Acknowledging the duty of the AMS and its developers to consider the broader societal impacts of the algorithm and to act in the public interest.
11. **Reliability**
    - Ensuring that the algorithm consistently provides trustworthy support for decision-making processes within the AMS.
    - Ensuring the algorithm provides precise and correct predictions about job seekers' employment prospects to avoid misallocation of resources.

B2. Investigate a Value

Value: Fairness

Definition: Fairness in the context of the AMS algorithm refers to the unbiased and equal treatment of all job seekers, ensuring that the algorithmic decisions do not favor or disadvantage any individual or group based on personal characteristics such as age, gender, ethnicity, or health status.

Differences in Team Members' Perceptions:

- Some team members believe that fairness means ensuring that the algorithm treats all job seekers equally without bias or favoritism, providing equal opportunities regardless of personal characteristics.

- Other team members argue that fairness involves allocating resources and support measures in a way that addresses the unique needs of different job seekers, particularly those from disadvantaged backgrounds.

From this discussion, the concept of fairness was further clarified and divided into two distinct sub-values: **equality**, which emphasizes uniform treatment and equal access to resources, and **equity**, which focuses on providing additional support to those who are historically disadvantaged to ensure truly fair opportunities for all.

STEP C: Explore

C1. Main Primary Values the System Supports

1. **Accountability**

   While the system supports accountability to some extent by allowing decisions to be overruled, the presence of biased human oversight can partially undermine this value as mentioned by [5].

2. **Efficiency**

   The system is designed to allocate resources and make decisions quickly, enhancing the overall efficiency of employment services.

3. **Reliability**

   The developer provided an evaluation of the model that uses the precision metric. Precision, while useful in many contexts, is not always a reliable measure when evaluating the primary values a system supports in terms of reliability. This is because precision measures the consistency of results rather than their correctness. A system can produce highly precise but incorrect results, thereby failing to be reliable. Additionally, precision does not account for systematic errors or biases which can undermine the reliability of the system's outputs [4].

C2. Value Tensions and Design Features

1. **Fairness vs. Efficiency**
   - **Tension:** Striving for fairness might require complex adjustments and individualized assessments, which can reduce the system's overall efficiency.
   - **Design Feature Favoring Efficiency:** By categorizing job seekers based on broader characteristics and predefined statistical models rather than individual nuances, the system can make decisions more quickly and allocate resources more efficiently, although potentially at the expense of fairness to some individuals.

2. **Fairness vs. Reliability**
    - **Tension:** Emphasizing the algorithm's ability to make correct predictions and classifications with high precision and reliability.
    - **Design Feature Favoring Reliability:** By prioritizing the optimization of performance metrics such as precision during the model development process to maximize the algorithm's predictive accuracy, we may end up with an optimized model that performs well technically. However, this approach can come at the expense of fairness, potentially leading to discrimination against certain groups[4].

STEP D: Adapt

D1. Mitigating Value Tensions

1. **Value Tension: Fairness vs. Efficiency**

    Implement a dynamic resource allocation system that balances efficiency with fairness.

    Use a two-phase approach where the initial classification is broad and efficient, followed by targeted interventions for those identified as potentially disadvantaged. This allows the system to maintain overall efficiency while ensuring that additional support is available for those who need it.

2. **Value Tension: Fairness vs. Reliability**

    Implement a Transparency-Driven Approach.

    Develop a transparency-driven approach to address the tension between fairness and reliability in the decision-making process:

    - **Algorithmic Transparency**: Ensure that the algorithms used for classification and resource allocation are transparent and explainable. Provide detailed documentation on the data sources, feature engineering, and model architecture to allow for external auditing and understanding of the decision-making process.
    - **Continuous Monitoring and Feedback**: Continuously monitor the system's performance, including metrics like precision, recall, and demographic parity, to identify and address any biases or inconsistencies. Regularly review and update the algorithms based on the feedback and performance monitoring to improve fairness and reliability over time.
    - **Human Oversight and Intervention**: Incorporate human experts, such as career counselors, to periodically review the algorithmic decisions and provide oversight.

The inspiration for the above changes came from the recently published guidelines, like [13] and [14].

STEP E: A LOOK INTO THE FUTURE

E1. Envisioning Activity: Work of the Future

Ways the System May Change the Nature of Work:
For the better:

- **Enhanced Matching:** The system could improve the efficiency of job matching by analyzing job seeker profiles and employer requirements more effectively, leading to better job placements and reduced unemployment rates.
- **Skill Development:** By identifying skill gaps and recommending targeted training programs, the system could facilitate continuous learning and skill development among job seekers, fostering a more adaptable and competitive workforce.
- **Remote Work Facilitation:** With advancements in technology, the system might facilitate remote work opportunities, enabling greater flexibility and work-life balance for employees.

For the worse:

- **Algorithmic Bias:** On the downside, if the system's algorithms are biased or discriminatory, it could perpetuate existing inequalities in the labor market and disadvantage certain groups.
- **Job Insecurity:** Automation and algorithmic decision-making may lead to job displacement and increased job insecurity for certain occupations, particularly those susceptible to automation.

E2. Reflection on Values

After considering the potential impacts of the system on the future of work, there might be changes in how we perceive and prioritize certain values:

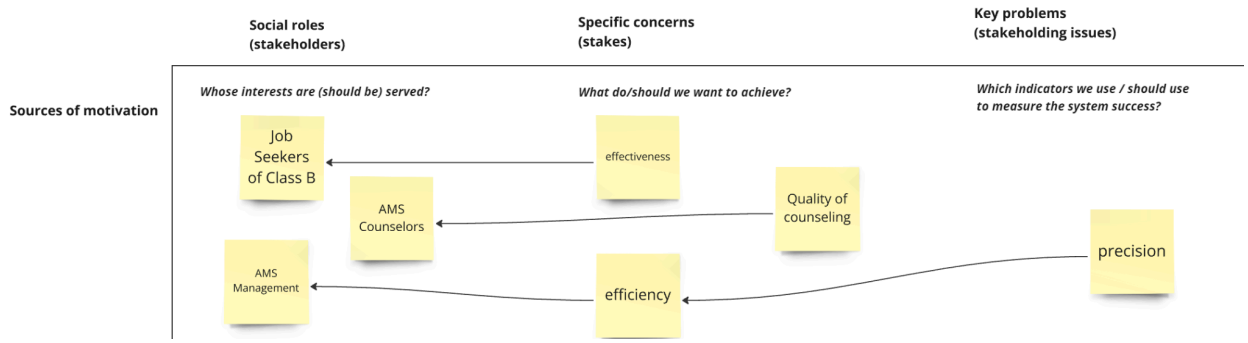
- **Fairness:** Given the potential for algorithmic bias and job insecurity, ensuring fairness in job matching and algorithmic decision-making becomes even more crucial. The AMS system should not enforce existing historical bias, otherwise in the future the nature of work will be changed for the worse. In this setting adding both definitions of equity and equality (as defined in B2) is fundamental.
- We also decided to introduce a new value: **Flexibility,** as a value to ensure the system can adapt to changing labor market conditions and stakeholder needs. This means the algorithm should remain adaptable to evolving labor market conditions, technological

advancements, and feedback from stakeholders, continuously improving its effectiveness and fairness.

Critical Systems Heuristics (CSH)

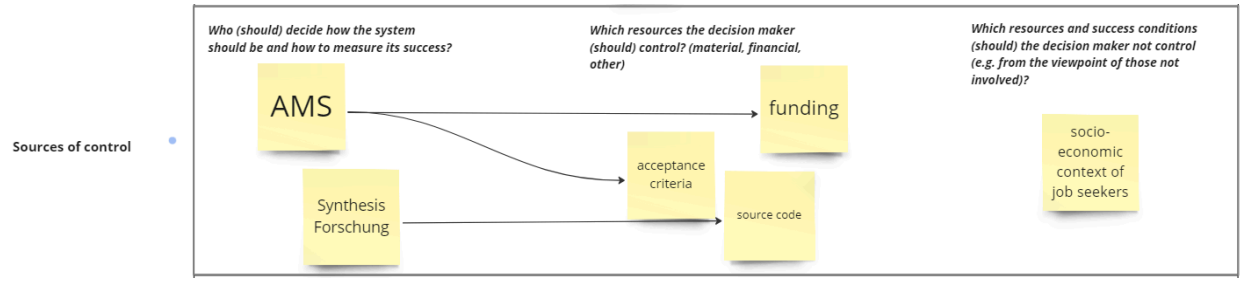
For better visualization, you can refer to the [Miro Link]. However, for a comprehensive understanding, we will elaborate on each row in this chapter.

AS IS Map:



Sources of Motivation:

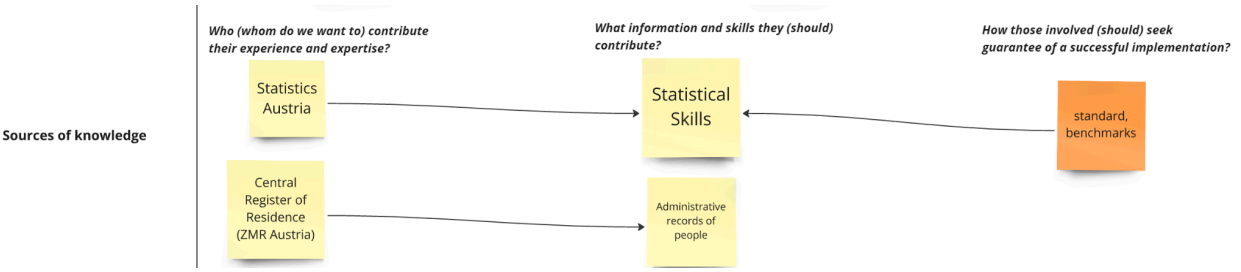
1. **Whose interests are served?**
   The primary beneficiaries are the job seekers classified as Class B, who receive resources from the agency. Additionally, AMS counselors and management are also served by the system.
2. **What do we want to achieve?**
   The aim is to effectively allocate resources to assist job seekers in finding employment. AMS management seeks to optimize resource allocation to maximize employment while minimizing resource use. Counselors aim to enhance their counseling effectiveness and provide high-quality guidance to job seekers.
3. **Which indicators do we use to measure the system's success?**
   The current measure of success is precision [4], which indicates how people are allocated to different groups. However, this metric may not fully capture fairness in resource allocation.



Sources of Control:

1. **Who decides how the system should be and how to measure its success? Which resources the decision maker controls?**
   The Austrian Public Employment Service (AMS) defines the system requirements, acceptance criteria, and provides funding and oversight. AMS has the authority over the broader strategic and operational decisions. However, the development and implementation of the algorithm are handled by Synthesis Forschung [8], whose control is limited to the technical aspects of the system.

2. **Which resources and success conditions the decision maker does not control?**
   The decision-makers (AMS and Synthesis Forschung) cannot control the socio-economic context of job seekers, which significantly impacts job placement success.



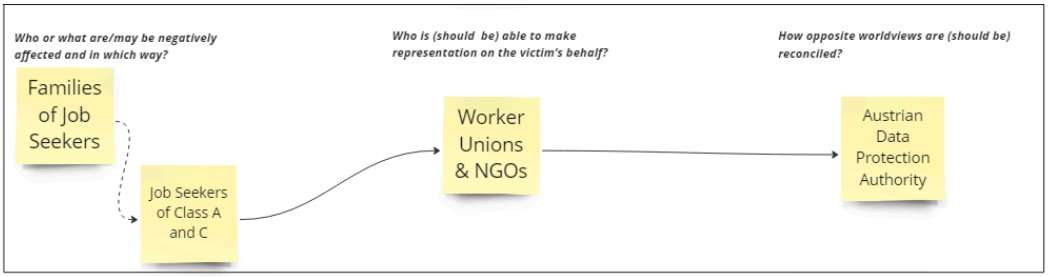
Sources of Knowledge:

1. **Who contributes with their experience and expertise and what information and skills do they contribute?**
   Statistics Austria [9] is an independent federal institution responsible for performing scientific services in the field of federal statistics. It provides objective and impartial data on various aspects of Austrian society, including economic, demographic, social, ecological, and cultural conditions. The actual data is provided by the Central Register of Residents (ZMR) [10], the Central Register, which collects mandatory administrative records of austrian people.

2. **How do those involved seek guarantees of a successful implementation?**

No information was found regarding this part, but we can safely assume that Statistics Austria likely provided benchmarking and validation to ensure the data used in the system meets high standards of accuracy and relevance. This might have included comparing the algorithm's performance against established benchmarks and using methods such as K-anonymity to protect privacy.

Sources of legitimacy



Sources of Legitimacy:

1. **Who or what are/may be negatively affected and in which way?**
   If job seekers are misclassified by the algorithm (e.g., classified into groups that do not accurately reflect their needs), they and their families may suffer from inadequate support. This could lead to prolonged unemployment and financial stress. Some job seekers are more vulnerable than others, including women, older job seekers, individuals with disabilities, and ethnic minorities.

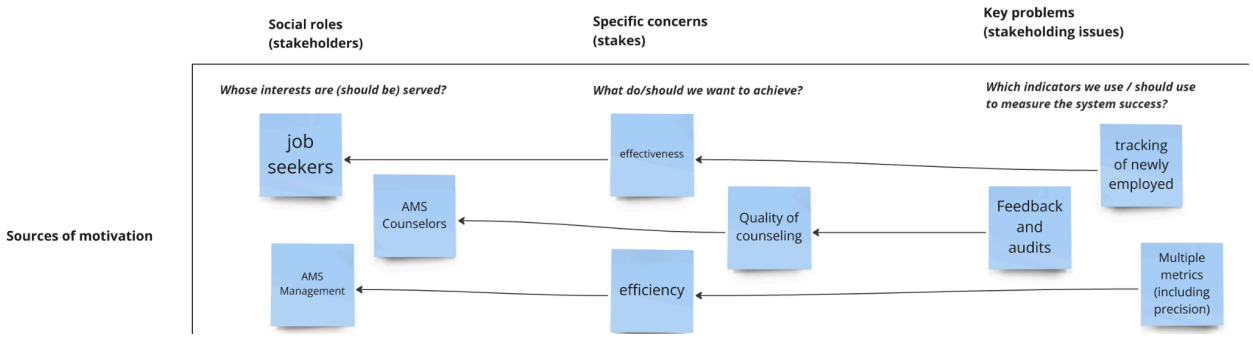
2. **Who is able to make representation on the victim's behalf?**
   Currently, worker unions and non-governmental organizations (NGOs) play a significant role in representing job seekers. These organizations advocate on behalf of those affected by the system, using available documents and data to challenge misclassifications and seek justice for disadvantaged groups.

3. **How opposite worldviews are reconciled?**
   The model lacks mechanisms for reconciling opposing worldviews. However, Austrian Data Protection Authority (DPA) has intervened in the past, such as halting the AMS algorithm due to privacy law violations. They can act on behalf of job seekers to ensure compliance with data protection regulations and fair treatment.

SHOULD BE Map:

To avoid the unfairness of the division of workers in 3 groups (that doesn't have enough granularity to understand the needs of a person) and keeping the efficiency of the resource allocation we decided to change the ADM: instead of assigning a group to each person, it will assign the number of allocated resources to a person considering its background, needs and

possibilities. We also introduced in different parts of the system external actors, like union workers in the source of control and third party evaluators in the source of legitimacy.

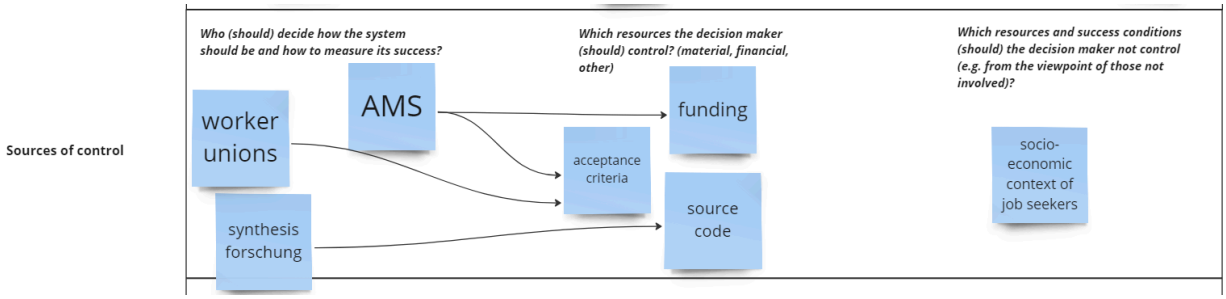


Sources of Motivation:

1. **Whose interests should be served?**
   The system should serve the interests of all job seekers, ensuring that resources are allocated based on individual needs rather than group classifications. This includes providing adequate support to those who may be at a disadvantage as well, such as women, older job seekers, individuals with disabilities, and ethnic minorities.

2. **What should we want to achieve?**
   We do not propose any changes to the current objectives shown in AS IS map.

3. **Which indicators should we use to measure the system's success?**
   Collecting feedback from job seekers on their experience, perceived fairness, and satisfaction with outcomes is essential. Additionally, accuracy and reliability metrics using precision, recall, and other statistical measures should be employed to evaluate the algorithm's accuracy. Long-term impact metrics, such as career progression, employment stability, and economic improvements, are crucial for monitoring the sustained impact of the system. Finally, gathering feedback from AMS counselors on system usability and its impact on their ability to support job seekers effectively will help ensure the system meets operational needs.

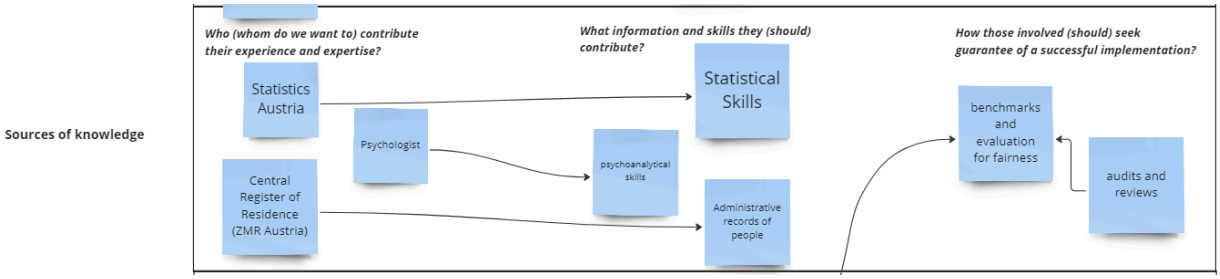


Sources of Control:

1. **Who should decide how the system should be and how to measure its success? Which resources should the decision maker control?**

   The decision on how the system should be and how to measure its success should be made collaboratively by the Austrian Public Employment Service (AMS) and worker unions. AMS, representing management, can ensure that the system aligns with organizational goals and public expectations. Worker unions can provide insights into the needs and challenges faced by job seekers. AMS should control resources such as funding. Meanwhile, Synthesis Forschung should be in charge of the technical aspects, including the development and maintenance of the algorithm's code.

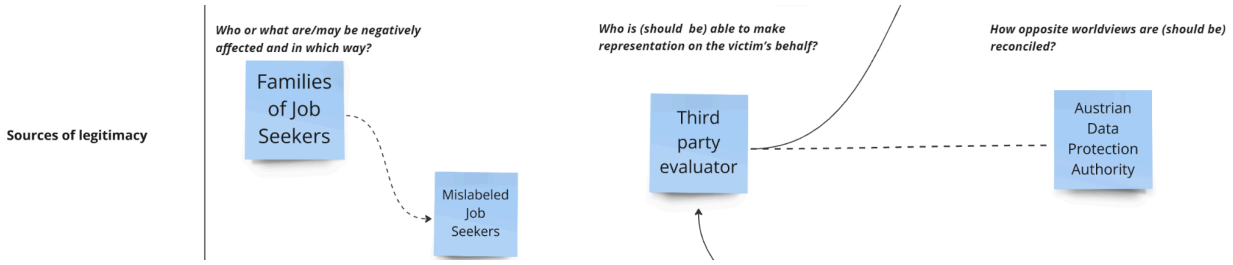
2. **Which resources and success conditions should the decision maker not control?**
   We did not propose any changes to this part with respect to the AS IS map.



Sources of Knowledge:

1. **Whom do we want to contribute with their experience and expertise and what information and skills they should contribute?**
   We include psychologists in this part to provide psychoanalytic insights that help understand the backgrounds and potential of job seekers. Their expertise ensures that the algorithm evaluates capabilities without introducing unfairness, considering physical and mental health, as well as social factors that affect employability.

2. **How should those involved should seek guarantees of a successful implementation?**
   Benchmarking and evaluation should explicitly include fairness metrics, provided by third-party evaluators who are impartial and objective. Regular audits and reviews by these independent evaluators will help maintain the integrity and effectiveness of the system.



Source of Legitimacy:

1. **Who or what should be negatively affected and in which way?**
   Ideally, no one should be negatively affected by the system.

2. **Who should be able to make representation on the victim's behalf?**
   Except for worker unions and non-governmental organizations (NGOs) that currently represent the interests of job seekers, an independent ombudsman or third-party evaluators, such as the European Centre for Algorithmic Transparency, should have the authority to review and contest the algorithm's decisions.

3. **How opposite worldviews should be reconciled?**
   We decided to keep the Austrian Data Protection Authority (DPA). However, it should be made clearer how the DPA can interact with AMS and the involved third-party evaluators to reconcile differing perspectives.

With these last two changes an unemployed person will have more trust in the system, knowing that it always has a third party evaluator that can analyze his demands and a DPA that can create a dialogue between AMS and the evaluator.

Conclusion

In this paper, we conducted a critical analysis of the AMS algorithm using Value Sensitive Design (VSD) and Critical Systems Heuristics (CSH) frameworks. Our examination revealed significant ethical, legal, and social implications associated with the algorithm's deployment, particularly regarding fairness, transparency, and accountability. We identified sources of discrimination within the system and proposed changes to enhance its fairness and effectiveness.Furthermore, we stressed the necessity of maintaining transparency, providing adequate support to vulnerable groups, and ensuring that no one is negatively affected by the algorithm's decisions. By implementing these recommendations, we aim to create a more human-centered and equitable automated decision-making system. This work serves as a foundation for further analysis and improvements, promoting principles of fairness, accountability, and transparency in the broader field of automated decision-making.

Bibliography

[1] I. Nagl. *Labour Market Policy Austria – Overview. Reporting Year 2022*. Federal Ministry of Labour and Economy, Vienna, 2023. Translation by E. Holzmair-Ronge.

[2] Austrian Public Employment Service (AMS). *Labour Market Data*. Department of Labour Market Research and Career Information, Vienna, 2024.

[3] APA. "Further Wait for Ruling on Controversial AMS Algorithm's Legitimacy." *The Austrian Press Agency*, July 21, 2023.

[4] 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*, Article 5. doi: 10.3389/fdata.2020.00005

[5] Cech, F., Fischer, F., Grill, G., Human, S., Lopez, P., & Wagner, B. (2019, October 3). The AMS algorithm lacks an instruction leaflet. *Futurezone*.

[6] Kayser-Bril, N. (2019, October 6). Austria's employment agency rolls out discriminatory algorithm, sees no problem. *AlgorithmWatch*.

[7] Arbeitsmarktservice Österreich. (2024). Registered unemployed persons by gender, duration of registration and age groups - outflow from unemployment and average duration of terminated unemployment (days). AMS Austria. June 3, 2024.

[8] IAB-Colloquium. (2022, April 25). *Insights into the Labor Market Opportunities Assistance System (AMAS) of the Austrian Public Employment Service (AMS).*

[9] Statistics Austria. Responsibilities and principles of Statistics Austria. Federal Statistical Office of Austria. Accessed June 29, 2024.

[10] Zentrales Melderegister (ZMR). Austrian Federal Ministry of the Interior (BMI). Accessed June 29, 2024.

[11] European Commission. High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. April 8, 2019.

[12] Australian Government. (n.d.). *Australia's AI Ethics Principles.*

[13] Montréal Declaration for a Responsible Development of Artificial Intelligence. (2018)

[14] European Union. (n.d.). *European Declaration on Digital Rights and Principles*.