

AI Risk Assessment - HR CV Screening System

Overview

This project presents a practical AI risk assessment applied to a CV screening system used in HR.

The focus is not on the model itself, but on how the system influences decisions and how to control that influence.

The Problem

AI systems in HR are often described as “decision support tools”; in reality, they shape who gets seen and who doesn’t.

When ranking becomes selection, the system moves from support to decision.

Without control, errors in the data become systematic and invisible.

Approach

I approached this as a system problem, not a model problem.

The assessment follows three steps:

- Identify where the system affects decisions
 - Evaluate the impact of each risk
 - Introduce controls directly in the decision flow
-

Key Insight

The system does not introduce new errors.
It makes existing ones systematic and scalable.

Main Risks Identified

- Historical data reflecting past decisions rather than actual needs
 - Implicit exclusion of candidates with non-standard profiles
 - Lack of traceability in ranking decisions
 - Misuse of personal data (GDPR exposure)
 - Misalignment between dataset and real-world requirements
-

Controls (Governance by Design)

The solution is not additional documentation, but **structural control**.

Key controls implemented:

- Separation between scoring, validation, and decision
 - No automatic exclusion based on model output
 - Mandatory human review for shortlist validation
 - Logging of input, features, and scoring logic
 - Stop rules when data quality or confidence is insufficient
-

Residual Risk

Some risks cannot be eliminated:

- Loss of context when translating human profiles into structured data
- Dependence on organizational decisions embedded in data
- Variability in human interpretation of AI outputs

The goal is not to remove all risk, but to make it visible and controllable.

Conclusion

The main risk is not technical failure, but uncontrolled influence on decisions.

By introducing architectural constraints and traceability, the system becomes verifiable and usable.

Governance is not documentation, it is a property of the system.
