Swedish Social Insurance Agency fraud risk model

IMPLEMENTED 2013–2025 · GOVERNMENT AND PUBLIC SERVICES · SWEDEN · BAND F · REVERSED
VERSION 1.1 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

A machine-learning risk profile selected temporary parental benefit claims for checks. An inspectorate found it failed an equal treatment test, journalists reported bias, and the agency withdrew it in 2025 during a data protection inquiry.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

What went wrong: Weak oversight or controls.

AI used: Machine-learning risk profile that scores benefit claims to select them for checks.

Implemented: 2013–2025. Grade: band F, Failure, Reversed.

Headline result, as reported: Agency took its machine-learning risk profile for parental benefit checks out of use in 2025 during a data protection supervision; an inspectorate review in 2018 found it did not meet an equal treatment test

1 · The case

The facts, as reported by the 5 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Försäkringskassan (Swedish Social Insurance Agency). Sweden's national social insurance agency. The case concerns a machine-learning risk profile used to select claims for temporary parental benefit (vab) for targeted checks. Supervised by the Swedish data protection authority, IMY, in 2025. Canvas block: 8. Governance and risk. Also 3. Customers and proposition, 6. Data and technology.

Business challenge

The challenge. Not reported by the agency in the sources opened. Lighthouse Reports says claims of large-scale benefit fraud were one of the primary justifications for the system, and says its own analysis found those estimates rested on a definition of fraud that did not check whether mistakes were intentional.

How it was addressed. The agency used a machine-learning risk profile to score applications and select some for checks. Amnesty International reports that the system assigned risk scores to applicants and that those with the highest scores were automatically investigated by fraud controllers. The agency told IMY it used a machine-learning risk profile in its selection of vab claims for targeted checks. It has declined to disclose how the system works, arguing that disclosure would let fraudsters evade detection, per Lighthouse Reports. [1] [2] [3] [4] [5]

Who implemented it

Who built and ran the model is not reported in the sources opened.

Technology (SEAM pillar E)

SEAM types: T8 Predictive / ML models. Deployment: Not reported. Vendors and products: Not reported.

Where it was built in: A risk profile applied to benefit claims; claims with the highest scores were sent to the agency's fraud controllers, per Amnesty International. How it was built and hosted is not reported.

What changed

Before: claims flagged by the risk profile went to investigation, and applicants did not know they had been flagged by an algorithm, per Lighthouse Reports. After: the agency told IMY in September 2025 that the profile had been taken out of use a little over a month earlier and that it had no plans to bring it back for that handling. Measured results of the checks: Not reported in the sources opened. [1] [2] [3] [4] [5]

Cost of failure

Not reported. IMY struck out its supervision case without a finding, fine or order, and no damages or remediation cost appears in the sources opened.

What it saved

Not reported [1] [2] [3] [4] [5]

Governance and control

ISF reviewed the profile in 2018 and found it did not pass its equal treatment test; the agency rejected that conclusion. IMY opened supervision in June 2025 and closed it after the agency withdrew the profile. By Lighthouse Reports, the manager who oversees the algorithm denied wrongdoing and said a human investigator always makes the final decision. Vendor, impact assessment and monitoring inside the agency: Not reported.

S2 · Accountable owner · Evidenced

Lighthouse Reports names the manager who oversees the agency's fraud algorithm and quotes his response to its findings. [1]

A1 · Safe use · Missing (independent source shows it was absent)
ISF found that the selection did not pass its equal treatment test in its 2018 design, and Lighthouse Reports says the agency's data protection officer questioned the system's legality in a 2020 note. [3] [1]

A2 · Human in the loop · Claimed
The responsible manager argued that a human investigator always makes the final decision, per Lighthouse Reports; Amnesty International reports that the highest-scoring cases were automatically investigated by fraud controllers. [1] [2]

A3 · Monitoring and assurance · Evidenced
ISF reviewed the model's accuracy and equal treatment in 2018 and recommended an independent ethics review of selection profiles; IMY supervised the processing in 2025. [3] [5]

A4 · Vendor AI · Not reported
Not reported.

A5 · Agent controls · Not reported
Not reported.

A6 · Standards and regulation · Outside scrutiny (a regulator, court, auditor or similar body acted)
IMY opened supervision under the data protection regulation (GDPR) and closed it after the agency withdrew the profile; IMY said AI use must follow GDPR and the AI Act. [4] [5]

What worked, what did not, what is unproven

WORKED

– ISF's 2018 review found that the risk-based selection raised the accuracy of follow-up checks and was, in ISF's words (translated from Swedish), considerably more accurate than the other alternatives then available.

DID NOT

– ISF found that the selection, in its current design, did not pass ISF's test for equal treatment. ISF did not decide whether this amounted to discrimination in the sense of the Swedish Discrimination Act.

– Lighthouse Reports and Svenska Dagbladet report that their fairness tests on 2017 data found women, migrants, low-income earners and people without a university education were overselected, and that people from these groups who had done nothing wrong were more likely to be wrongly flagged. The agency's responsible manager denied wrongdoing.

– Amnesty International reports that people incorrectly flagged complained of delays and legal hurdles in accessing their entitlement.

– The agency withdrew the profile during the data protection supervision.

UNPROVEN

– Whether the profile breached data protection law: IMY closed the case without examining the merits.

– Whether the withdrawn profile has been replaced by another model: Not reported in the sources opened.

– Whether the 2017 results hold for the model as it ran in later years: the analysis used a 2017 dataset.

– The agency's reasons for withdrawing the profile: Not reported in the sources opened.

Timeline

2013 Amnesty International says the agency has used the machine learning system since at least this year [2]

2018 Inspectorate for Social Insurance (ISF) reports that the risk-based selection raises accuracy but, in its current design, does not pass the ISF equal treatment test; the agency rejects the conclusion [3]

2020 A redacted note from the agency's data protection officer questions the legality of the system, per Lighthouse Reports [1]

Oct 2021 Lighthouse Reports sends its first freedom-of-information request to the agency; it is rejected [1]

Nov 2024 Lighthouse Reports and Svenska Dagbladet publish their investigation; Amnesty International calls for the system to be discontinued [2]

Jun 2025 The data protection authority (IMY) opens supervision of the agency's risk-based selection for temporary parental benefit checks [4]

Sept 2025 The agency tells IMY (reply received 5 September) that it used a machine-learning risk profile, took it out of use a little over a month earlier and has no plans to reinstate it [5]

Nov 2025 IMY closes the case on 18 November because the profile is no longer in use [5]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: Not applicable. Grade: F, Failure. Proof: Reversed. Level of change: Not graded.

No level assessed: band F, a deployment withdrawn after an inspectorate finding and a regulator's supervision, with reported unequal selection. The regulator made no finding and imposed no sanction.

Graded F. A state inspectorate found the risk profile failed its equal treatment test, and the agency took the profile out of use while the data protection authority was supervising it. The regulator made no finding and imposed no sanction, so the grade rests on the withdrawal and on the reported unequal selection and delays.

The SEAM view and lessons learnt

The model was reviewed from outside twice, in 2018 by an inspectorate and in 2025 by a regulator, and both turned on one question: did the selection treat groups equally? The agency also kept the workings confidential, which left applicants and reviewers unable to test it. A mid-sized or larger organisation using scores to decide who is checked needs a fairness test it can run itself, a policy on what it will disclose, and a plan for what happens when a reviewer's finding goes against it.

1. Test any selection model for unequal outcomes by group before and after launch.
2. Decide in advance what you can disclose to reviewers and to the people affected.
3. Answer an inspector's finding with evidence, and record why you accept or reject it.

Evidence

Score 3 of 4. Independent: a state inspectorate review (2018), a data protection authority decision (2025) and a press analysis supported by Amnesty International. Limits: IMY made no finding on the merits; the press analysis used 2017 data; the agency disputed the 2018 analysis.

Evaluation method. Regulator and inspectorate documents plus investigative journalism: ISF report 2018:5 (summary), IMY supervision decision IMY-2025-11828, and statistical fairness tests by Lighthouse Reports and Svenska Dagbladet.

Sources disagree. Amnesty International describes the system as scoring social security applicants generally; ISF, IMY and the Lighthouse dataset concern temporary parental benefit, and this synopsis limits its findings to that benefit. IMY's decision records that the agency took the profile out of use during supervision and that IMY struck out the case; it records no order to stop.

Status. Closed: IMY decision of 18 November 2025. The agency says it has no plans to reinstate the profile for that handling.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/forsakringskassan-welfare-fraud-risk-model

[1] Lighthouse Reports: Sweden's Suspicion Machine
https://www.lighthousereports.com/investigation/swedens-suspicion-machine/

[2] Amnesty International: Sweden: Authorities must discontinue discriminatory AI systems used by welfare agency
https://www.amnesty.org/en/latest/news/2024/11/sweden-authorities-must-discontinue-discriminatory-ai-systems-used-by-welfare-agency/

[3] Inspektionen för socialförsäkringen (ISF): Sammanfattning av rapport 2018:5, Profilering som urvalsmetod för riktade kontroller (Swedish)
https://isf.se/download/18.6e75aae16a591304891f1e8/1565330423683/Profilering%20som%20urvalsmetod%20f%C3%B6r%20riktade%20kontroller-ISF-sammanfattning-2018-05.pdf

[4] IMY (Integritetsskyddsmyndigheten): Avslutad tillsyn efter att Försäkringskassan tagit AI-system ur bruk (Swedish)
https://www.imy.se/nyheter/avslutad-tillsyn-efter-att-forsakringskassan-tagit-ai-system-ur-bruk/

[5] IMY: Beslut efter tillsyn enligt dataskyddsförordningen, Försäkringskassan, IMY-2025-11828 (Swedish, PDF)
https://www.imy.se/globalassets/dokument/beslut/2025/beslut-tillsyn-forsakringskassan.pdf

Försäkringskassan (Swedish Social Insurance Agency) website: https://www.forsakringskassan.se

Version 1.1 · Published 15 November 2026 · Updated 15 November 2026 · Facts checked against sources 15 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/forsakringskassan-welfare-fraud-risk-model

VERSION HISTORY

v1.1 15 November 2026 Method update: standards and regulation (A6) relabelled "outside scrutiny" because the evidence is a regulator, court, auditor or similar body acting, not the organisation's own compliance.
v1.0 15 November 2026 First published