The Norwegian
Tax Administration

A Data-Driven Approach to Fraud Handling: An Overview of the Norwegian Risk Model

Sofia, 2026

Martin Iver Jacobsen

The Norwegian
Tax Administration

Agenda

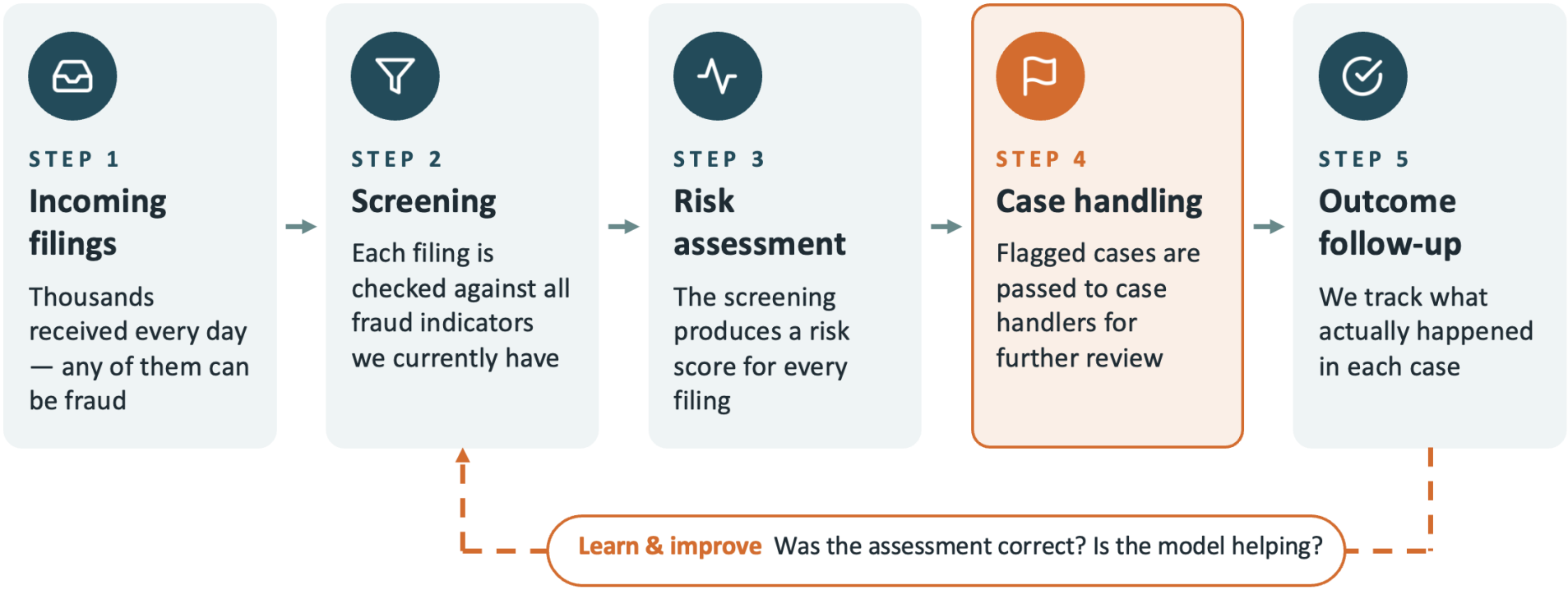
1. The Risk Model
   A framework for more data-driven fraud management

2. Measuring Effectiveness
   How do we know whether our efforts are working?

The Norwegian Tax Administration

The Risk Model: From Filing to Learning

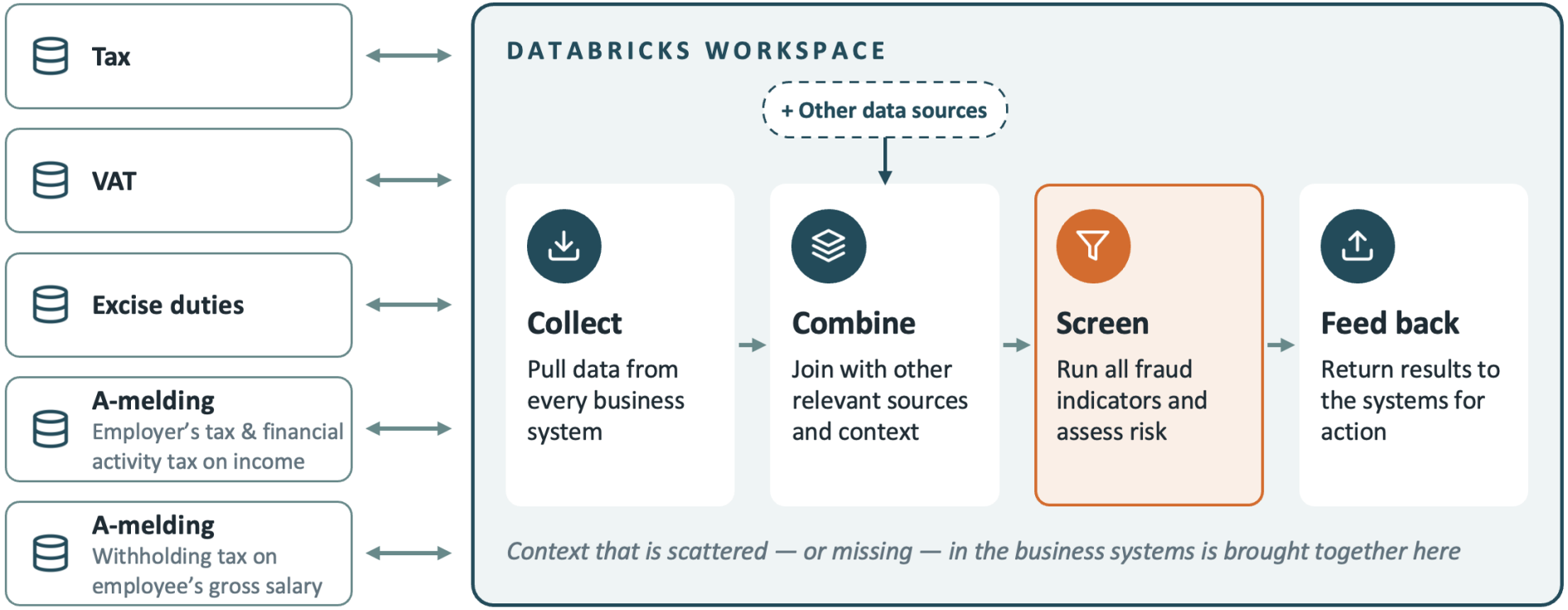
Thousands of filings a day — screened, assessed, reviewed and checked against the real outcome



The Norwegian Tax Administration

The Risk Model: The Reality

Different units, systems, history and legislation — combined, screened and fed back from one workspace



The Norwegian
Tax Administration

From Data to Detection

- New possibilities enabled by Databricks
- Expert knowledge + data = fraud indicators
- Expanding access to high-value data sources (Improving fraud detection requires better data, not just better models)

The Norwegian Tax Administration

A Theory of Indicators

Example from VAT: following a flagged return through the process



THE KEY QUESTION

How much evidence does a particular indicator provide that something is actually wrong?

The Norwegian Tax Administration

A Theory of Indicators

Two kinds of indicators with different roles

MODES

A particular way in which something worth investigating tends to occur — a pattern or mechanism

- ○ Mode A
- ◉ **Mode B**
- ○ Mode C
- ○ None

Mutually exclusive: at most one mode per filing, or none

SIGNALS

An indicator in its own right that can strengthen the evidence provided by a mode

- ☑ **Signal 1**
- ☐ Signal 2
- ☑ **Signal 3**
- ☐ Signal 4

Can co-occur: several signals can be present at once

WHAT WE ARE AFTER

Y = 1 means fraud. It is rare, with base rate π, and can occur even when every Z and A is 0.



The Norwegian Tax Administration

A Theory of Indicators

Putting it together: Bayes in log-odds form

$$L = \ln\frac{\pi}{1-\pi} + \ln\frac{P(Z_i=1|Y=1)}{P(Z_i=1|Y=0)} + \sum_{j=1}^{m}\left[A_j \ln\frac{a_j}{g_j} + (1-A_j)\ln\frac{1-a_j}{1-g_j}\right]$$

PRIOR

Base rate — low, since Y = 1 is rare

MODE

How much more typical mode i is when Y = 1 than when Y = 0

SIGNALS

Each signal shifts the score up when present

FROM SCORE TO PROBABILITY

$$p = \frac{1}{1+e^{-L}}$$

The logistic function turns the log-odds score L back into a probability between 0 and 1



The Norwegian Tax Administration

A Theory of Indicators

Operationalizing the model



BEFORE DAY 1

Set the prior

Experts set the base rate π from threat assessments

$$\pi$$

If Σs_i is close to 1, π may be set too low

DAYS 1–30

Review everything

- N all filings in the period
- N_i hits on mode i
- B_i of those, marked for "audit"
- C_{ij} marked for "audit" with both i and signal j

$$s_i = \frac{B_i}{\pi N}$$

$$a_j = \frac{\sum_{i \in \mathcal{M}_j} C_{i,j}}{\sum_{i \in \mathcal{M}_j} B_i}$$

← added signal strength

DAY 30

Initial scores

π cancels out, and each combination k of mode and signals gets a starting score

$$L_{k0} = \ln \frac{B_i}{N_i - B_i} + \Sigma \text{ signals}$$

DAY 31 ONWARDS

Learn from feedback

b_k of n_k flagged messages were worth a look

$$p_k^{\text{new}} = \frac{q\, p_{k0} + b_k}{q + n_k}$$

q = weight on the starting score: small reacts fast, large stays stable

- Updates always restart from L_{k0} and use all feedback since launch
- If behaviour shifts, count n_k and b_k over the last X days only
- Combinations never flagged have $n_k = 0$ and keep L_{k0}

The Norwegian
Tax Administration

Thank you!