

Shadow Mandate

The decision without a decider — a public-observability study of three public functions in twelve jurisdictions

*A decision has a formal holder. Between the facts and that holder, steps shape it: retrieval, analysis, recommendation, preparation. This report asks how much of that can be seen in official public **documents**.*

REPORT · V1.0 · PUBLIC RELEASE · EVIDENCE CUTOFF 2026-10-07

Summary

When a public body decides about a person, the law names who decides: a judge sets a sentence, a tax officer opens an audit, a caseworker grants a benefit. More and more often a computer system prepares that decision first. It gathers the facts, scores the case, recommends an outcome or drafts the text. This issue asks a narrow question. Looking only at official public documents, can an outsider see where the part of the system ends and the part of the official begins?

We looked at 31 systems in three public functions, justice, tax administration and social benefits, in twelve jurisdictions. For each system we recorded whether seven things are published.

What we found, in five points.

1. **The system is usually named. Its boundary is declared less often.** A public description of the system is published for 19 of the 31. A statement of who holds the decision is published for nine ¹. OBSERVED
2. **What the person affected would need is rarely published.** A notice that the system was used, an impact assessment, and a right to reasons or human review specific to automated processing are each published for two of the 31 ¹. OBSERVED
3. **Almost no one publishes how often officials depart from the system.** Such a measure is published for two systems, and both measure something adjacent to it ¹. OBSERVED
4. **Official findings exist where someone publishes them.** Seventeen systems on our list have a finding by a court, an auditor, a regulator, a parliamentary committee or a commission of inquiry ². OBSERVED
5. **The evidence does not all point one way.** Studies show officials departing from a system often, and a published divergence measure exists outside our pool. Section 8 sets this out. OBSERVED

What this is not. The matrix records what is declared in public. It does not record how any body works. A cell marked not found means not found in the public record. The issue does not order jurisdictions, and every count reflects the volume of public declaration, not the volume of use.

KEY TERMS IN PLAIN WORDS

Formal mandate. Who holds the decision by law or by design.

Shadow mandate. Where the decision is actually shaped: the steps that come before it.

Public observability. What an outsider can establish from official public documents alone, with no inside access.

The matrix. A table with one row per system and seven columns. Each column asks one question.

Published. An official document answers the question of the column. **Not found in the public record.** Our search found no such document. It may still exist.

Coder. An AI language model that fills the table by written rules and cites a document and a quotation for every cell.

Sealed protocol. The written rules, fixed and fingerprinted before coding began, so that they cannot be changed quietly afterwards.

01

Disclosure and research notice

This issue is written in a personal capacity by contributors to the Foresight Future Institute. It does not represent any contributor's employer. No case examined describes an engagement of any contributor's employer or its clients. Models from more than one developer were used for drafting, independent coding and review. All classifications follow a protocol registered before coding began.

Six amendments were added to the protocol, four of them after coding began. Appendix B lists them with their fingerprints.

RESEARCH NOTICE AND METHODOLOGY DISCLAIMER

1. **Scope.** This study is conducted in a personal capacity by contributors to the Foresight Future Institute. It is independent research into administrative transparency and algorithmic accountability. It does not reflect the views of any employer or official body.
2. **What "not found" means.** A classification of "Not found in the public record" means that the official documents named by the column were not retrieved within the search protocol used here. The gaps in that protocol are stated in section 10. The classification records public availability at the time of the study. It is not a statement that such documents do not exist, and it is not a statement about the internal workings of any public body.
3. **No ordering, no evaluation.** The report does not order jurisdictions, systems or public functions, does not score them and does not evaluate their performance. A cell records what a document says, not how a body works. Counts follow the volume of public declaration, not the volume of use.
4. **Machine-assisted coding.** Cells were coded by AI language models under a written protocol and checked as described in section 11 and Appendix A. Errors are possible. Every published cell carries a link and a quotation, so that a reader can check it.
5. **Cases and examples.** Named bodies and systems appear only where a public document or an official finding concerns them, and each statement reports what that document says. The three cases were chosen because their documents are easy to read. They are not typical and they are not a sample.
6. **Not advice, no liability.** Nothing here is legal, financial or professional advice, and the matrix is not an assessment of compliance with any law. To the extent permitted by applicable law, the authors and the Foresight Future Institute accept no liability for decisions that third parties take on the basis of this report.

7. **Sources and currency.** Links lead to documents of third parties, which can change or be withdrawn. The record closes on 7 October 2026, 23:59 (UTC+4). Later documents are listed in section 12 and are not absorbed into the matrix.

8. **Corrections.** Any body or reader may send a verifiable official document to info@foresightflow.org. If the document meets the criterion, the cell changes, the change is dated and published, and the earlier state stays visible, as set out in section 13.

02

The question

2.1 Two mandates

The formal mandate is who holds the decision by law or by design. It is visible in a statute, a procedure or a signature.

The shadow mandate is where the decision is actually shaped: in the steps below the decision itself. There are four of them. Retrieval gathers the facts of the case. Analysis scores or sorts them. Recommendation proposes an outcome. Preparation drafts the text that the official will sign.

The two mandates can coincide. They do not have to. An official can sign a decision whose facts were selected, weighed and worded by a system. The signature is then real and so is the part of the system, and only the first is on the record. The figure shows three arrangements of the same decision. In all three the official signs. They differ in how much of the shaping comes before the signature.

— FIGURE 01 · THE DECISION PIPELINE, SEEN FROM TWO SIDES

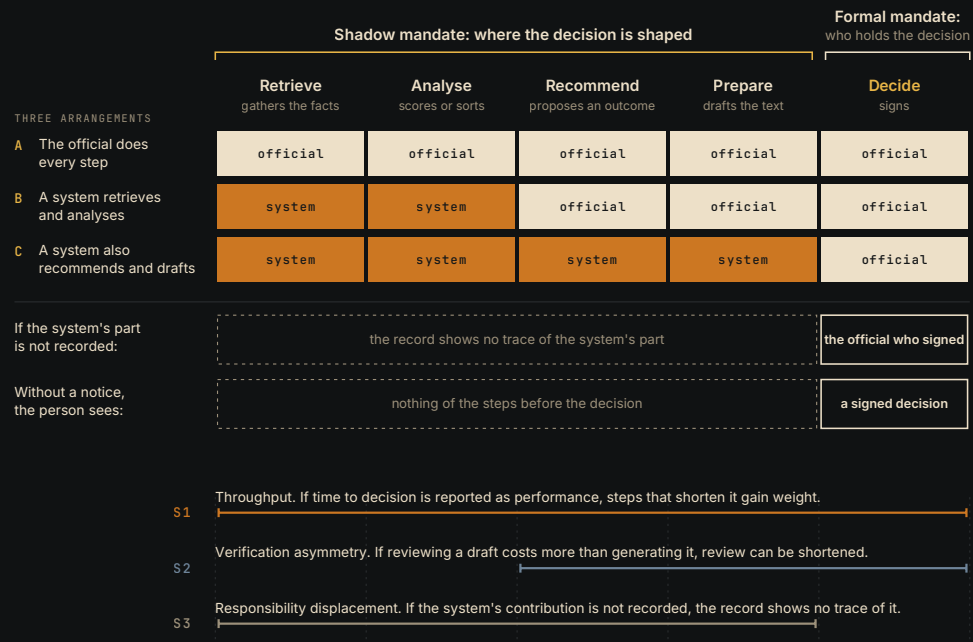


ILLUSTRATION · NO REAL BODY OR SYSTEM IS DESCRIBED · THE STRESSES ARE CONDITIONS, NOT FINDINGS

2.2 Three stresses

Digital bureaucracy can press on the line between the two mandates in three ways. Each is a mechanism. We state each as a condition, not as a finding about any body.

Throughput. Public bodies report how fast they work: time to decision, backlog, cases closed. These are legitimate measures and the public asks for them. If time to decision is what is reported, a step that shortens it gains weight.

INFERRED A draft produced in seconds changes the arithmetic of a working day in a way that a careful second reading does not.

The trace this would leave in public documents is specific. A body would publish how much the system processes or how much time it saves, and would not publish what the saved time displaced. This is why the matrix puts a column on throughput (C7) beside a column on divergence (C6). **INFERRED**

Verification asymmetry. Producing a recommendation costs a system almost nothing. Checking it costs an official time: the file has to be read again and the reasoning rebuilt. If reviewing a draft costs more than generating it, and nothing rewards the review, review can become the step that is shortened. **INFERRED**

This is not inevitable, and the evidence is mixed. In one study of bail decisions, judges overrode the recommendation of an algorithm often, and most of them did worse than the algorithm when they did ³. **OBSERVED** So the open question is not whether officials can depart from a system. It is whether anyone outside can see how often they do, and with what result. **INFERRED**

Responsibility displacement. A decision record names the official who signed. If the contribution of a system is not recorded, the record shows a decision with a decider and no trace of what shaped it. INFERRED That contribution is concrete: which facts the system retrieved, what it recommended, and whether the official changed it.

The consequence falls on two readers of the record. The person affected receives the reasons of the official and cannot contest a step that they do not know took place. An auditor or a court that reads the file later sees a human decision. Columns C4 and C5, on notice and on reasons and review, record whether a public document closes this gap. INFERRED

These are statements about a mechanism. They describe no body and no practice.

2.3 Two viewpoints

The administration sees who acts at each step: the top rows of the figure. The person affected sees only what a notice or the reasons of the official show: the bottom rows. The matrix asks the same question from both sides: is the boundary between the two mandates visible in a public document?

2.4 What the law already says

The concern is not new to the law. The EU AI Act names awareness of automation bias among the capacities that a human overseer of a high-risk system must have ⁴. OBSERVED A statute in force in the United Kingdom since 5 February 2026 defines a decision as solely automated when there is no meaningful human involvement, and tells the person assessing this to consider the extent of profiling ⁵. OBSERVED These provisions make observability, not recognition, the open question. INFERRED

2.5 What this issue does not claim

It does not claim that any body acts in a particular way. It names a failure only where an official inquiry, commission, court or auditor has established it, and it says which. It attributes no motive and no intent. It does not assess the practice of a named body.

03

How to read the matrix

3.1 Seven columns

Each row of the matrix is one system. Each column asks whether one thing is published.

COLUMN	SHORT NAME	WHAT THE CODER LOOKED FOR	A DOCUMENT THAT WOULD COUNT
C1	Who decides	A statement of who holds decision authority over the output of the system	An adopted policy or decree saying that the official, not the system, decides
C2	What the system is	A document that identifies the system, states its purpose and gives its data, its logic or the decision it informs	A register entry or an official description
C3			

COLUMN	SHORT NAME	WHAT THE CODER LOOKED FOR	A DOCUMENT THAT WOULD COUNT
	Impact assessment	A published assessment of the impact on the people affected, in full or as an official summary	A published data-protection or algorithmic impact assessment
C4	Notice to the person	A rule or notice that tells the person that the system was used in their case	A notice that names the processing, or a rule that requires notice
C5	Reasons and review	A right to reasons or to human review that is specific to automated processing in this function	A statutory provision or an official procedure
C6	Divergence rate	A number: how often officials depart from or override the output of the system	An annual report with an override rate
C7	Throughput	A number: how much the system processes, or how fast	A report with volumes or times

A cell can be in one of five states.

STATE	MEANING
Published	An admissible official source meets the criterion, and its quotation was found on the cited page
Not found in the public record	Our search found no such source. This is a statement about our search, not about the body
Not applicable	Used in column C6 only, where an official source declares that the system decides without human review
Unverified	A coder cited a source, and the page could not be retrieved to check the quotation
Disputed	The two coders disagree, and the source of the published side could not be verified

The three functions are justice (courts and tribunals, including decision support to judges), tax administration (assessment, risk selection, audit and enforcement targeting) and social benefits (eligibility, entitlement calculation, fraud detection and debt recovery).

3.2 A worked example

Take a hypothetical system. A benefits agency uses a model, System X, that flags claims for a caseworker to review. No real body or system is described here.

- **C1, who decides.** A ministerial order says that the caseworker decides every flagged claim. The cell is published.
- **C2, what the system is.** A register entry names the system, its purpose and the data it uses. The cell is published.
- **C3, impact assessment.** An annual report says that an impact assessment exists. It is not published, so the cell is not found in the public record.
- **C4, notice.** The agency has a general privacy notice. It does not name this processing, so the cell is not found in the public record.
- **C5, reasons and review.** A statute gives a right to human review of automated decisions in this function. The cell is published.
- **C6, divergence.** A report gives the number of claims flagged. It does not say how often caseworkers disagreed with the flag, so the cell is not found in the public record.
- **C7, throughput.** The annual report gives the number of claims processed each month. The cell is published.

— FIGURE 02 • A WORKED EXAMPLE: ONE HYPOTHETICAL SYSTEM, SEVEN CELLS

C1 Who decides	C2 What the system is	C3 Impact assessment	C4 Notice to the person	C5 Reasons and review	C6 Divergence rate	C7 Throughput
 PUBLISHED	 PUBLISHED	 NOT FOUND	 NOT FOUND	 PUBLISHED	 NOT FOUND	 PUBLISHED
A ministerial order says the caseworker decides every flagged claim.	A register entry names the system, its purpose and the data it uses.	A report says an impact assessment exists. It is not published. in the public record	Only a general privacy notice. It does not name this processing. in the public record	A statute gives a right to human review of automated decisions.	A report counts flagged claims, not how often caseworkers disagreed. in the public record	The annual report gives the number of claims processed each month.

HYPOTHETICAL • NO REAL BODY OR SYSTEM IS DESCRIBED

The row says what an outsider can read. It does not say whether caseworkers follow the flag, whether the order is observed, or whether the model is accurate. A row with seven published cells would describe a system whose boundary is fully declared. It would not describe a system that works well.

04

What we found

KEY TAKEAWAYS

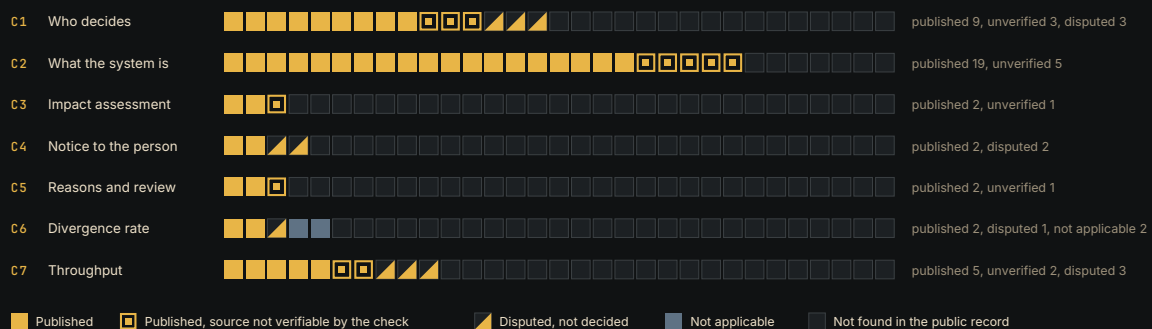
The system is described in public for 19 of the 31 systems. Who decides is declared for nine.

Impact assessment, notice, and reasons and review are each published for two.

The same pattern appears in all three functions. The share of published cells is lowest in tax administration.

For about one cell in ten, a coder cited a source that could not be checked, or the two coders disagree. Those cells are shown as they are.

— FIGURE 03 • EACH COLUMN ACROSS THE 31 CODED SYSTEMS



Squares are not jurisdictions in a fixed order. The figure does not order jurisdictions and does not combine columns into a single figure.

4.1 Column by column

What the system is (C2). A public description is published for 19 of the 31 systems, and for five more it is published but its source could not be verified ¹. OBSERVED This is the column with the most published cells, which is expected: a system enters the list because an official source mentions it.

Who decides (C1). A statement of who holds decision authority is published for nine systems, with three more unverified and three disputed ¹. OBSERVED Three examples show what such a statement looks like. In the United Kingdom, guidance for judicial office holders says that they are personally responsible for material produced in their name ⁶. OBSERVED In the United States, an audit report on a model that classifies tax returns says that the examiner can disregard the issues the model recommends ⁷. OBSERVED A transparency record of the United Kingdom for a benefits fraud model says that a human always intervenes and decides ⁸. OBSERVED

Impact assessment (C3). Published for two systems, with one more unverified ¹. OBSERVED One transparency record of the United Kingdom, for a tax debt model, states that the full data-protection impact assessment has been published ⁹. OBSERVED

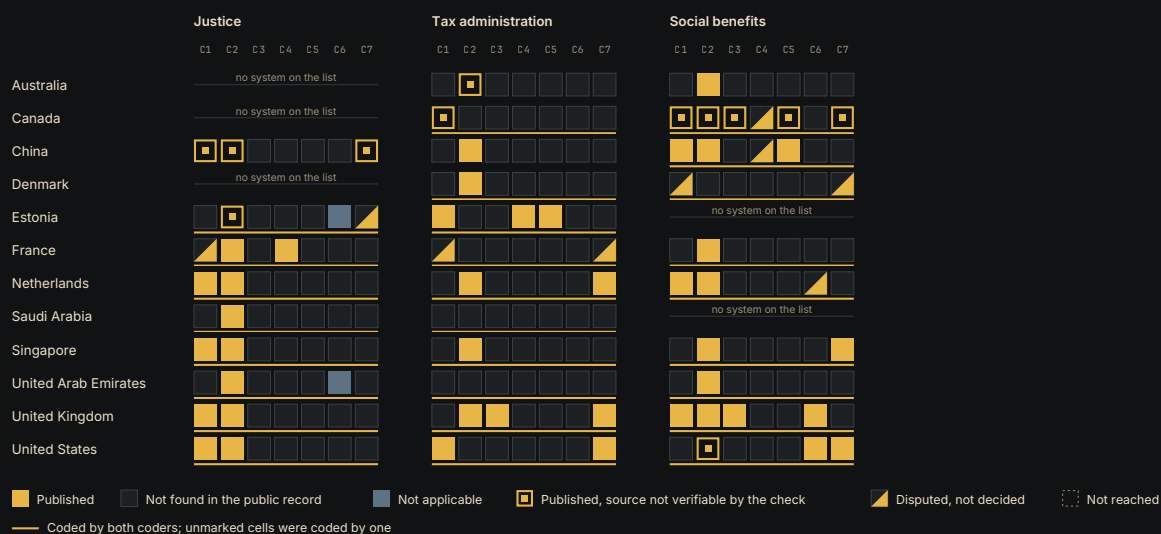
Notice to the person (C4). Published for two systems, and disputed for two more ¹. OBSERVED In France, a notice of the justice ministry names the processing called DataJust ¹⁰. OBSERVED

Reasons and review (C5). Published for two systems, with one more unverified ¹. OBSERVED In Estonia, a page of the tax authority on automated decisions says that the person may state a view, obtain an explanation and contest the automated decision ¹¹. OBSERVED

Throughput (C7). Published for five systems, with two unverified and three disputed; it is reported separately ¹. OBSERVED The published figures are of different kinds. A transparency record of the United Kingdom says that a model scores about four million customers each month ⁹. OBSERVED A parliamentary document in the Netherlands says that the signals in a register concerned 9,084 citizens ¹². OBSERVED The two coders agreed least on this column, as section 11 shows.

Divergence (C6) is the subject of section 6.

— FIGURE 04 · THE PUBLIC-OBSERVABILITY MATRIX



4.2 Function by function

Justice (nine systems). A public description of the system (C2) is published for seven, with two more unverified. A statement of who holds decision authority (C1) is published for four, unverified for one and disputed for one. No impact assessment (C3) and no statement of a right to reasons or human review (C5) is published for any of the nine, and a notice to the person affected (C4) for one ¹. **OBSERVED** Column C6 is not applicable for two systems, for which official sources declare that the system decides without human review, and not found in the public record for seven.

OBSERVED

Tax administration (twelve systems). The description of the system (C2) is published for five, unverified for one and not found in the public record for six. Authority (C1) is published for two, unverified for one and disputed for one. An impact assessment (C3), a notice (C4) and a right to reasons or review (C5) are each published for one ¹. **OBSERVED** Column C6 is not found in the public record for any of the twelve. A throughput measure (C7) is published for three and disputed for one. **OBSERVED**

OBSERVED

Social benefits (ten systems). The description of the system (C2) is published for seven, unverified for two and not found in the public record for one. Authority (C1) is published for three, unverified for one and disputed for one. An impact assessment (C3) and a right to reasons or review (C5) are each published for one and unverified for one more. A notice (C4) is published for none and disputed for two ¹. **OBSERVED** A divergence rate (C6) is published for two and disputed for one. **OBSERVED** A throughput measure (C7) is published for two, unverified for one and disputed for one.

OBSERVED

4.3 How much the counts depend on the checks

The counts depend on rulings that make the reading of three criteria, C1, C2 and C7, stricter. Six of the seven rulings are contested ¹³. INFERRED At the ruling check, published cells fell in C1 from 23 to 12, in C2 from 31 to 26 and in C7 from 6 to 4 ¹. OBSERVED The source check then moved cells whose quotation was absent from the cited page to not found in the public record, and marked cells whose page could not be retrieved as unverified ¹⁴. OBSERVED A reader who reads a criterion more loosely will count more published cells. INFERRED

05

Three cases, cell by cell

KEY TAKEAWAYS

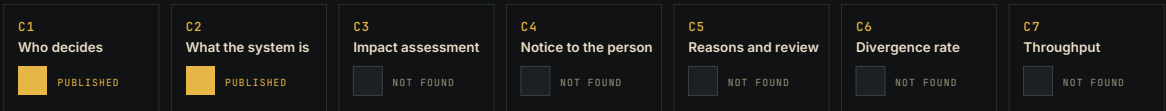
- One system per function, read cell by cell, with the public document behind each published cell.
- The cases were chosen because their documents are easy to read. They are not typical, and they are not a sample.
- The three rows differ. One has the description and who decides, and nothing else. One publishes notice and re-view. One publishes the divergence column.

The matrix is easier to trust when a row can be read against its documents. This section does that for three systems, one per function ¹. A case describes public documents. It does not assess how the body works.

5.1 Justice: a writing aid for the reasons of a sentence

In the Netherlands, a district court tried an AI tool as a writing aid for a draft of the reasons for a criminal sentence ¹⁵. OBSERVED The step concerned is preparation: the drafting of a text that a judge signs.

— FIGURE 05 · JUSTICE: A WRITING AID FOR THE REASONS OF A CRIMINAL SENTENCE



COLUMN	STATE	WHAT THE PUBLIC DOCUMENT SAYS	SOURCE
C1	Published	The court describes the tool as a supporting language aid in drafting a criminal judgment.	15
C2	Published	A news item of the court says that judges used AI as a writing aid for a draft of the reasons for the sentence.	15
C3	Not found in the public record	No impact assessment was found in the public record.	

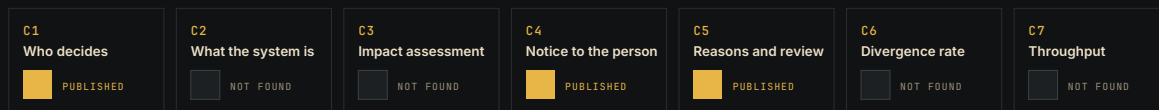
COLUMN	STATE	WHAT THE PUBLIC DOCUMENT SAYS	SOURCE
C4	Not found in the public record	No notice to the defendant that the aid was used was found in the public record.	
C5	Not found in the public record	No right to reasons or review specific to this use was found in the public record.	
C6	Not found in the public record	No figure on how often judges changed the draft was found in the public record.	
C7	Not found in the public record	No throughput figure was found in the public record.	

Both published cells rest on one news item of the court. OBSERVED

5.2 Tax administration: automatic decisions on paying a tax debt in instalments

In Estonia, a taxpayer can ask the tax authority to pay a tax debt in instalments, and where set conditions are met the decision is prepared automatically ¹⁶. OBSERVED It is the only coded system in tax administration for which both a notice and a right to review are published ¹. OBSERVED

— FIGURE 06 · TAX ADMINISTRATION: AUTOMATIC DECISIONS ON PAYING A TAX DEBT IN INSTALMENTS



COLUMN	STATE	WHAT THE PUBLIC DOCUMENT SAYS	SOURCE
C1	Published	The authority's page says that an application that does not meet the conditions for an automatic decision is reviewed by an official.	¹⁶
C2	Not found in the public record	A coder cited the page for this column. The third model ruled the description too general to identify one system, so the cell is not found in the public record. The ruling is contested.	
C3	Not found in the public record	No impact assessment was found in the public record.	
C4	Published	The authority's page on personal data says that an automatically prepared decision carries a note saying so.	¹⁷
C5	Published	The same page says that the person may state a view, obtain an explanation and contest the automatic decision.	¹⁷
C6	Not found in the public record	First coded not applicable. A check found that an official takes over when an application is not granted automatically, so the system is not declared to decide without human review. No divergence figure was found in the public record.	
C7	Not found in the public record	No throughput figure was found in the public record.	

Two cells of this row were changed by checks: C2 by a contested ruling, and C6 by the check of the cells first coded not applicable ¹⁸. OBSERVED The row shows how much a single reading of a criterion can move a cell.

5.3 Social benefits: a model that flags requests for a benefit advance

In the United Kingdom, a machine-learning model scores requests for an advance on a benefit and refers some of them to a human decision maker ⁸. OBSERVED Four of its seven cells are published, including the divergence column ¹.

OBSERVED



COLUMN	STATE	WHAT THE PUBLIC DOCUMENT SAYS	SOURCE
C1	Published	The transparency record says that a human always intervenes and decides, and that the model makes no auto-mated decision.	8
C2	Published	The record describes a supervised machine-learning classifier that risk-assesses requests for advances.	8
C3	Published	An effectiveness assessment for April 2025 to March 2026 is published. It states that concerns about effectiveness, unfair treatment and impact on legitimate claimants are minimal.	19
C4	Not found in the public re-cord	The department's privacy notice covers fraud detection and, according to the effectiveness assessment, names artificial intelligence as a method of processing. The third model decided that a general notice does not tell the person that the system was used in their case, so the cell is not found in the public record.	
C5	Not found in the public re-cord	A coder cited a general statement of data-protection rights. The third model decided that it is not specific to this function.	
C6	Published	The assessment defines a true discovery rate: the share of referrals later denied payment by a human decision maker. It reports that rate as relative likelihoods between groups of claimants and gives no overall figure.	19
C7	Not found in the public re-cord	A coder cited a figure on payment delay. Under the ruling on throughput it does not qualify.	

An audit body has also reported on this model; it is one of the 17 systems of section 7 ². OBSERVED

5.4 What the three rows show together

The cells do not move as a block. A body can publish who decides and nothing for the person affected, or the reverse. This is why the matrix keeps seven columns apart and does not add them up. INFERRED

06

The empty column

KEY TAKEAWAYS

Before coding began we named column C6, the divergence rate, a candidate empty column.

It is not empty. It is published for two of the 31 systems.

Both relate the selection of a system to a later human decision. Neither is a count of overrides kept for that purpose.

Outside our pool, one sentencing commission publishes such a figure every year.

Column C6 asks for a published number: how often human officials depart from or override the output of the system. Of the seven columns it is the only one that speaks directly to the second and third stresses of section 2.2.

It is published for two of the 31 coded systems ¹. **OBSERVED** In the United Kingdom, an official effectiveness assessment of a model that flags potentially fraudulent advance claims, published in July 2026, defines a true discovery rate: the share of referrals later denied payment by a human decision maker ¹⁹. It reports that rate as relative likelihoods between groups of claimants and gives no overall figure ¹⁹. **OBSERVED** In the United States, an audit report of 2009 states that, in a sample of 575 cases that a model had selected for a quick determination, 38, about 7 percent, were not allowed ²⁰. **OBSERVED** Both relate the selection of a system to a later human decision. Neither is a count of overrides kept for that purpose, and the first gives only relative figures. **INFERRED**

In one more cell the coders disagree and the source of the published side could not be verified, so the cell is disputed and not decided ¹. **OBSERVED** In two systems an official source declares that the system decides without human review, so the column is not applicable. **OBSERVED** Four cells were first coded not applicable. A check of them against a further ruling moved two to not found in the public record, because the sources showed direct judicial supervision or review by an official. The sources of one of the two that remain could not be retrieved, and its value stays as coded ¹⁸.

OBSERVED In the other 26 systems no such measure was found in the public record. **OBSERVED**

— FIGURE 08 · COLUMN C6, PUBLISHED DIVERGENCE RATE, BY FUNCTION

Justice



Tax administration



Social benefits



Published Not found in the public record Not applicable: declared to decide without human review Source not verifiable by the check Disputed

Outside the pool, not coded: one US state sentencing commission publishes each year how often judges agree with its risk assessment. So a published divergence measure exists in the world. The claim here is limited to the cells in the pool, as coded.

The column is therefore not empty. It is nearly empty, and the two filled cells report something adjacent to what the column asks. The claim is limited to the cells in the pool, as coded: one US sentencing commission outside the pool publishes each year how often judges agree with its risk assessment ²¹. OBSERVED

07

Official findings

KEY TAKEAWAYS

Seventeen systems on the list have an official finding about them: three in justice, seven in tax administration, seven in social benefits.

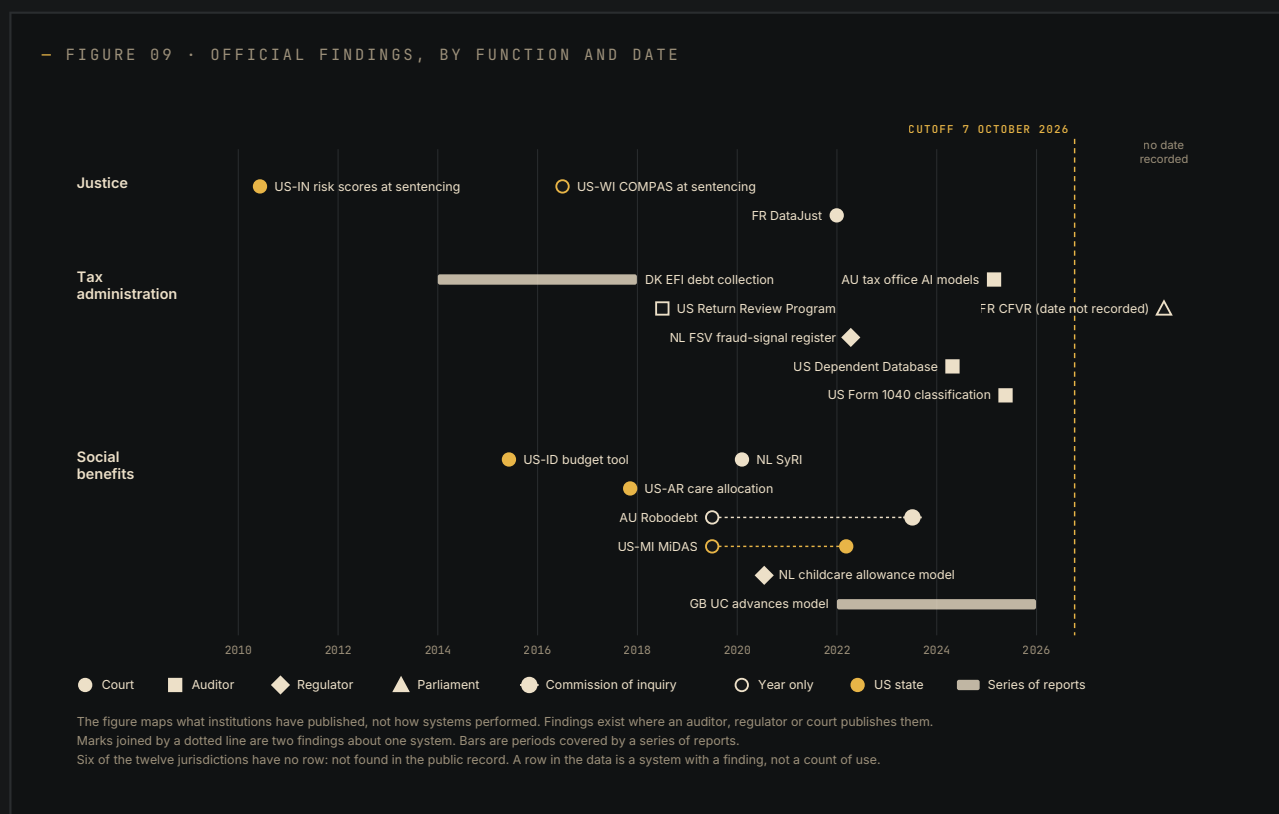
A finding is not always adverse. In all three justice rows a court allowed the use or rejected the challenge.

Four of the seven social-benefit systems are retired, against two of seven in tax administration. The order of finding and retirement is shown row by row, and no cause is claimed.

For six of the twelve jurisdictions no official finding was found in the public record.

Findings exist where an auditor, regulator or court publishes them. This section covers the 17 systems on the list for which such a body has made findings about the system itself ². INFERRED

Three examples show the range. In the Netherlands, a court ruled in February 2020 that the legislation for a system known as SyRI was in breach of Article 8 of the European Convention on Human Rights ². OBSERVED In Australia, a commission of inquiry reported in July 2023 on the scheme known as Robodebt ². OBSERVED In the United States, an audit report of May 2025 made recommendations on a model that classifies tax returns for examination ². OBSERVED The first two systems are retired. The third is live.



Courts account for seven of the findings, auditors for six, data-protection regulators for two, a parliamentary committee for one, and a court together with a commission of inquiry for one ². OBSERVED Ten of the systems were live at the cutoff, six were retired and one was a pilot with no live effect ². OBSERVED

A finding is not always adverse. In all three justice rows a court allowed the use or rejected the challenge, and two of them state limits on how a score may be used ². OBSERVED In tax administration five of the seven findings come from auditors working to their own programme, with one regulator and one parliamentary report ². OBSERVED In social benefits four of the seven come from courts, and the others from a regulator, an auditor and a court together with a commission of inquiry ². OBSERVED

Four of the seven social-benefit systems are retired, against two of seven in tax administration. No causal claim follows ². OBSERVED

The date of first use is verified for too few of the 17 systems to state how long a system ran before its first finding. This issue does not state that interval. OPEN

Eight of the 17 rows are in the United States: three federal rows from two audit bodies and five state court rows, marked separately in the figure ². OBSERVED For six of the twelve jurisdictions no official finding was found in the public record. That says nothing about the absence of anything else. OBSERVED

The timeline maps the reach of the institutions that publish. It does not show where systems have worked less well.

INFERRED

08

Evidence against the thesis

Without this section the issue would be one-sided. These are the strongest rows we found against the thesis and against our reading of the matrix. Most come from studies of bail decisions and child-welfare screening, not from tax or benefits.

People do depart from the system. In a study of bail decisions, judges overrode the algorithm's recommendation for 12% of low-risk defendants and for 54% of high-risk ones. In the overrides of 90% of the judges, the outcome was worse than the algorithm's ³. OBSERVED Departure was frequent, so this is not a picture of rubber-stamping. It was mostly not better, so it is not evidence that review works.

A system can raise accuracy and consistency. A machine-learning model applied to bail data simulated crime reductions of up to 24.7% at unchanged jailing, with lower racial disparity ²². OBSERVED This is a simulation on past data, not an evaluation of a deployed tool. In a large online experiment on price estimates, algorithmic assistance raised both accuracy and the consistency of estimates across respondents ²³. OBSERVED That is a laboratory result, not a government setting.

Workers corrected a tool's errors. In child-maltreatment hotline screening, workers were less likely to follow a tool's recommendation when the score shown to them was misestimated ²⁴. OBSERVED This is outside the three functions.

A divergence measure can be published. A US state sentencing commission publishes every year a chapter on how often judges agree with its risk assessment; its overall concurrence rate was above 75% in the 2024 report ²¹.

OBSERVED This sits outside the pool and is not coded.

The law already names the concern. The provisions in section 2.4 show that the boundary between support and decision is part of the vocabulary of the law. Observability, not recognition, is the open question. INFERRED

What is not established. No study read here concerns tax administration or social benefits. Evidence that automated tools raised consistency in a government setting is not established. OPEN Whether declared human review is effective review is not established. OPEN

09

What would make the shadow mandate observable

This section contains no finding about any body. It reads the matrix the other way round: for each column, what a public document would have to say for the cell to read published. A legislator, a regulator, an auditor, or a body that wants its own boundary to be visible could use it as a checklist. INFERRED

COLUMN WHAT WOULD MAKE THE CELL READ PUBLISHED

- | | |
|----|---|
| C1 | An adopted instrument that says who decides in relation to the output of the system, and that the official can depart from it |
| C2 | A public entry that names the system and gives its purpose and its data, its logic or the decision it informs |

COLUMN WHAT WOULD MAKE THE CELL READ PUBLISHED

C3	The impact assessment itself, or an official summary of it; a statement that one exists is not enough
C4	A notice to the person that names the system, or a rule that requires such a notice
C5	A right to reasons or to human review that is specific to automated processing in the function
C6	A record of how often officials accept, change or reject the output of the system, published at least once a year with the number of cases
C7	The throughput figures that are published already, placed beside the figures for C6

Of these, the sixth is the one the matrix found least often. It is also the cheapest to produce where a system already logs its own recommendations, because the comparison with the final decision is then a count. INFERRED

Two cautions apply. Publication is not performance: in the bail study of section 8, officials departed from a system often and mostly did worse ³. OBSERVED A published divergence rate would show how the boundary is used, not whether it is used well. And the evidence here is limited: 31 systems, read by two AI coders whose agreement on one column is below the threshold set in advance (section 11). INFERRED

10

Limits

- **The matrix records declaration.** It does not record how a body works. A cell marked published says that a document exists. It does not say that the document describes what happens.
- **Counts follow public declaration.** A jurisdiction that publishes more produces more rows. Counts must not be read as the volume of use ²⁵.
- **No ordering.** The issue does not order jurisdictions and does not combine cells into a single figure.
- **One system per cell.** The protocol allowed three systems per jurisdiction and function. For lack of time one was coded: the first by the rule of the protocol. In the cells where more systems qualified, the coded system is not a sample of the others.
- **The coders are language models.** They can share patterns that come from training, and independence between them is therefore limited. INFERRED We reduced the effect with different developers, a sealed protocol, blind coding and a third model for disputes. We did not remove it.
- **Four systems were coded by one coder only.** They are marked in the matrix.
- **A share of the sources could not be checked.** For about three in ten cells that a coder marked published, the cited page could not be retrieved, in part for technical reasons on our side ²⁶. OBSERVED Those cells are shown as unverified, not as published.
- **The checks work in one direction.** The ruling check and the source check can move a cell from published to not found in the public record, and never the other way. The counts before and after are in section 4.3.
- **Not-found cells rest on our searches.** They were closed with a general search engine, and some language queries were not run. A document we did not find may exist.
- **Agreement on throughput is weak.** On column C7 the two coders agreed less than the threshold set in advance.

- **Screening of states is partial.** The matrix is a matrix of countries.
- **Evidence in section 8 is uneven.** The studies available to us concern bail decisions and child-welfare screening, not tax or benefits.
- **Copies of judgments.** Some court decisions are cited from a court or repository copy and not from the issuing court's own site.
- **Findings reflect publication.** The official findings in section 7 exist where an auditor, regulator or court publishes them. They show the reach of those institutions. INFERRED
- **Declared review is not effective review.** Column C5 records whether human review is declared, not whether it works.

11

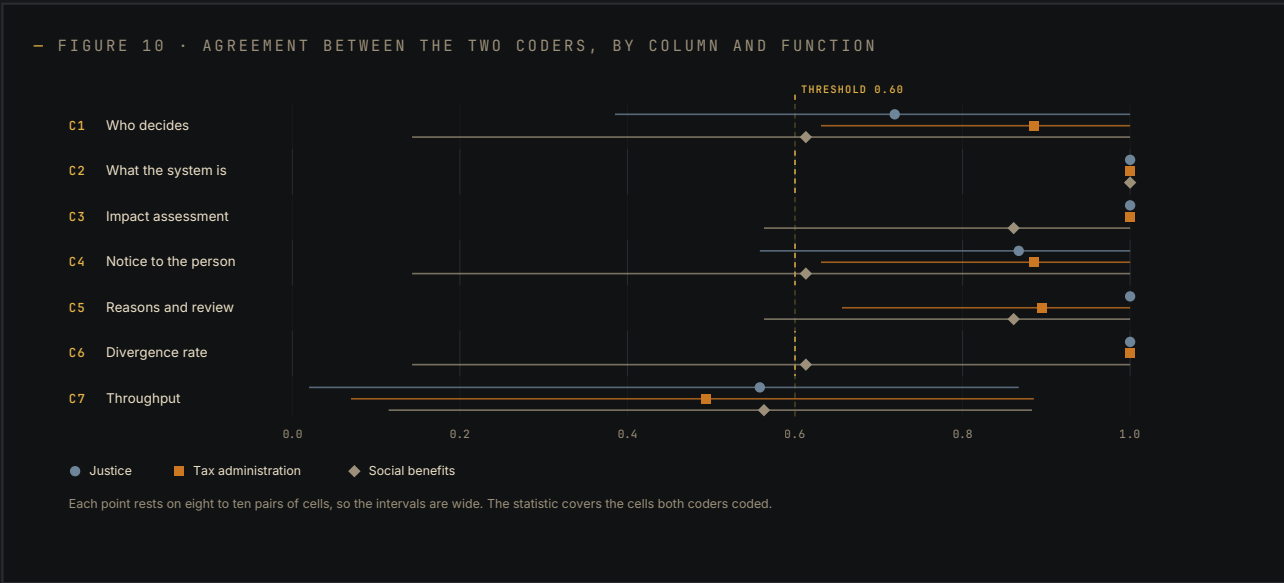
How we did it, in brief

Rules first. The rules for filling the matrix were written down and fingerprinted before any cell was coded. Six amendments followed, four of them after coding began; none changed a criterion, and two record rulings on how a criterion reads. Appendix A gives the method in full and Appendix B the fingerprints.

Why AI coders. Each cell needs a search of official sources in the languages of twelve jurisdictions, and a decision by a written rule. Two AI language models from different developers did this separately. Neither saw the other's results, the thesis of this issue or the rules for reading the results. Every published cell carries a link and a quotation, so a reader can check it without trusting the coder. AI coders are not free of bias, and they can be wrong in the same way at the same time; the limits in section 10 say so. INFERRED

Three checks. First, a third model ruled on seven questions about how a criterion reads, and the cells those rulings touched were checked against them. Second, the page behind every published cell was fetched and the quotation had to be on it. Third, where the two coders disagreed, the third model decided from the retrieved passage; where the passage could not be retrieved, the cell stays disputed ¹⁴. OBSERVED

How far the coders agreed. Both coders finished 27 of the 31 systems ¹. OBSERVED Across all columns their agreement before any check was 0.86 in justice, 0.87 in tax administration and 0.70 in social benefits, on a scale where 1 is full agreement and the threshold set in advance is 0.60 ²⁷. OBSERVED On throughput (C7) it was below that threshold in all three functions. OBSERVED Each figure rests on eight to ten pairs of cells, so the intervals are wide.



12

After the cutoff

Documents published after 7 October 2026 are outside the record. None is listed at the date of publication. Later documents are added here, with their date, rather than absorbed into the matrix.

13

Correction policy

The matrix can be corrected. Any body, or any reader, can point to an official document that we missed. If the document meets the criterion, the cell changes, the change is dated and published, and the earlier state stays visible. We did not contact any body before publication.

APPENDIX A

Method in detail

A.1 The instrument

The matrix has seven columns. A cell is published only if an official source meets the column criterion. Otherwise it is not found in the public record.

COLUMN CRITERION

- C1 An official source states who holds decision authority in relation to the system's output
- C2 An official document identifies the system, states its purpose and says at least one of: data sources, logic, or the decision it informs

COLUMN CRITERION

C3	A pre-deployment or periodic impact assessment is public, in full or as an official summary
C4	An official rule or notice provides that the person is told that the system was used in their case
C5	A legal provision or official procedure gives a right to reasons or to human review specifically for automated or AI-assisted processing in this function
C6	An official publication gives a quantitative measure of how often human officials depart from or override the system's output
C7	An official publication gives a throughput measure; reported separately, outside the matrix block

Column C6 was a candidate empty column by design. It is not applicable only where an official source declares that the system decides without human review.

The matrix structure is adapted from the author's book *AI Agents at the Border* (ERA Society) ²⁸. Every criterion in C1 to C5 is drawn from an external instrument, listed in A.2.

A.2 Where the criteria come from

Columns C1 to C5 are drawn from named instruments. Columns C6 and C7 are observational and are not anchored in any norm. We searched the sealed list of instruments for a provision that requires publication of a divergence rate and found none ²⁹. OBSERVED

COLUMN INSTRUMENTS

C1	EU AI Act Arts 14 and 26 ⁴ ; GDPR Art 22 ³⁰
C2	EU AI Act Arts 49 and 71 ³¹ ; the algorithmic transparency recording standard of the United Kingdom ³² ; the national algorithm register of the Netherlands ³³
C3	EU AI Act Art 27 ³⁴ ; the algorithmic impact assessment of Canada ³⁵
C4	EU AI Act Art 26(11) ³⁶ ; GDPR Arts 13 to 15 ³⁷
C5	EU AI Act Art 86 ³⁸ ; UK GDPR Arts 22A and 22C ⁵ ; the directive on automated decision-making of Canada ³⁹

The instruments are references for the criteria. They are not tests applied to any body, and they bind different jurisdictions in different ways. At the cutoff, the AI Act obligations for the Annex III high-risk systems were scheduled to apply from 2 December 2027 under Regulation (EU) 2026/1744 ⁴⁰. OBSERVED

A.3 Three functions

Each system is assigned to one function. L1 is justice: courts and tribunals, including decision support to judges and tribunal members. L2 is tax administration: assessment, risk selection, audit and enforcement targeting. L3 is social benefits: eligibility, entitlement calculation, fraud detection and debt recovery. Courts and tribunals go to L1; administrative decisions go to L2 or L3 ⁴¹. Customs control and customs duty assessment are outside every function ⁴².

A.4 Which systems enter

A system enters the matrix if an admissible official source declares or describes automated or AI-assisted processing used in decisions in the function (basis a), or if an official inquiry, commission, ombudsman, auditor, parliamentary committee, court or tribunal has made findings about it (basis b). Rule-based automation, machine learning, generative AI and agentic systems are in scope. Systems that only file, schedule, process payments or store documents are out. Pilots declared to have no live effect are out unless basis (b) applies. The time window runs from 1 January 2010 to the cutoff ⁴³.

A.5 Admissible sources

Six source types are admissible: statute and regulation (S1); official policy, guidance and register entries (S2); official reports of the body (S3); reports of an official inquiry, commission, ombudsman, auditor-general or parliamentary committee (S4); judgments and decisions of courts and tribunals (S5); and official statements by the body or by an official state news agency (S6). Press reporting, vendor material, academic papers, NGO reports and social media are not admissible. They can lead us to an admissible source and never change a cell ⁴⁴.

A review written by officials of the operating body or of its parent department counts as S3, not S4, unless the reviewer is independent by statute ⁴⁵. A copy of a judgment is admissible when it sits on a court or government site or on a case-law repository and the text itself shows the court, the case number and the date ⁴⁶.

A.6 The pool and the list

The pool has twelve jurisdictions, shown in alphabetical order in the matrix figure. Each jurisdiction is a cell at the national level. The US federal level and each US state that enters are separate cells ⁴⁷. A US state enters when a basis (b) system is found in it, and is then to be screened in all three functions ⁴⁸. Our screening of the states was partial. Where state rows are shown, they are marked separately from the national rows.

The union of the two coders' system lists after dispute rulings held 127 systems. OBSERVED Where more than three systems qualified in one jurisdiction and function, three were listed: basis (b) systems first, then those with the most recent qualifying official source ⁴⁹. The list holds 62 systems; 65 rows beyond the cap are published in the data release ⁴⁹.

OBSERVED

The cap interacts with the volume of public declaration. In the Netherlands, the tax-administration cell holds 64 systems; most are entries in the national algorithm register ⁵⁰. OBSERVED Three were listed under the cap rule. For register entries, the date used is the date of last modification. The count of 64 records the volume of declaration. It does not record the volume of use. INFERRED

Rows that describe a framework, strategy or programme without identifying a system were listed and not coded ⁵¹.

A.7 Coding and agreement

Of the 62 systems on the list, the 31 that rank first in their cell at the national level, including the US federal level, were coded ⁵². Amendment 3, issued after coding began, set this scope. It changed no criterion ⁵². Two coders, both language models from different developers, coded independently and neither saw the other's results, the interpretation lock or

any material about motives. Both coders finished 27 of the 31 systems; the other four were finished by one coder only ¹.

OBSERVED A system coded by one coder only is marked as single-coded in the matrix ⁵². A third model resolved disputes.

Every published cell was also checked mechanically: the cited page was fetched and the quotation had to appear in it ⁵². A cell whose quotation was absent from the retrieved page became not found in the public record unless the other coder's source for it passed the check. A cell whose page could not be retrieved stays published and is marked unverified. No cell became not found in the public record because a page could not be retrieved ¹⁴. **OBSERVED**

Seven questions about how a criterion reads arose during coding. The third model ruled on them, and six of the seven rulings are marked contested ¹³. The rulings concern the column on declared authority (an agency-wide statement counts only if it is an adopted binding policy), the column on throughput (a comparison in words without figures does not qualify), the column on public description (a function without a name counts only if it identifies a single system), mirror copies of official records, the status of systems said not to be in production, and two cases of newspaper material on official sites ¹³. Five mechanical triggers flagged cells for a check against these rulings, in two rounds. In the first round 39 cells were flagged: 22 became not found in the public record and 17 stayed as coded ¹³. **OBSERVED** The check can change a cell to not found in the public record and never to published. A later check of the cells coded not applicable in column C6 worked the same way: it moved two cells to not found in the public record and none the other way ¹⁸. **OBSERVED**

Of the 27 cells on which the coders disagreed, 13 were decided by the third model from a verified passage and 9 remain disputed because the source could not be verified; the rest were settled by the source check ¹. **OBSERVED**

Agreement as returned, before any reconciliation, on the cells both coders coded, as Gwet's AC1 ²⁷ {obs}:

COLUMN	JUSTICE	TAX ADMINISTRATION	SOCIAL BENEFITS
C1	0.72	0.89	0.61
C2	1.00	1.00	1.00
C3	1.00	1.00	0.86
C4	0.87	0.89	0.61
C5	1.00	0.90	0.86
C6	1.00	1.00	0.61
C7	0.56	0.49	0.56
All columns	0.86	0.87	0.70

Column C7 is below the threshold of 0.60 in all three functions. **OBSERVED** Each cell of the table rests on eight to ten pairs, so the intervals are wide; for C7 in tax administration the interval runs from 0.07 to 0.89. **OBSERVED** In social benefits, columns C1, C4 and C6 are only just above the threshold. **OBSERVED** The statistic describes the cells both coders coded, which is a subset, and it is reported before reconciliation. **INFERRED**

A.8 Screening dates and limits

Screening closed on 7 October 2026 at 23:59 (UTC+4), and this is the cutoff⁵³. A completion sweep covered cells that were still open and leads recorded in the notes. It did not repeat the earlier passes. Cells that were open at closing remain marked as incomplete in the data.

APPENDIX B

Sealed documents

Two amendments were added during screening and sealed before coding began. Four more were issued after coding began: Amendment 3 narrowed the coded set and fixed the file format; Amendment 4 recorded rulings of the third model on seven interpretation questions, six of them marked contested; Amendment 5 set how a source check that could not be completed is treated; and Amendment 6 recorded three further rulings.

DOCUMENT	SHA-256	SEALED
Coding protocol v1.0	7ec38368ff0023bc1d7ce02335ae209fb5fa9e078311e510e5abc179e1de7674	6 October 2026, 12:13 (UTC+4)
Interpretation lock v1.0	19745fee1bc41b7a43efbe1e89560897e18b34422e72b156484da8aed8bb89e0	6 October 2026, 12:13 (UTC+4)
Amendment 1	fe31f35cf69171e46d17b28637f8ecddcb636da0df8558f9d9781ef3e7674ef8	6 October 2026, 12:48 (UTC+4)
Amendment 2	d9e0c357287e9f8668a4cb1f6b2e72722b34c689e6591acb6b29e0ac1bfec8b2	6 October 2026, 18:44 (UTC+4)
Amendment 3	77ab2463db418cf069f58d149171b6d217168fca438bec9ed46f7f250c59e5ce	6 October 2026, 23:45 (UTC+4)
Amendment 4	e59cf3eaa559ccfad179700039687f4f88a7247c6685e26171729f8e08e45f2e	7 October 2026, 08:03 (UTC+4)
Amendment 5	9a5144b72b0daa8e2cb658a6131bc3cd3c2e2cf884dd8cada335d51dcbdf626	7 October 2026
Amendment 6	85cf942f8afddbe13d783900ad0504f5c2a448fd3ebf31ef0c6aaefca6942834	7 October 2026
System list v1	5a8b1a01193e3fb2eeb03c1e36459368bd310d3060b497594f386653cc0ba099	6 October 2026, 21:59 (UTC+4)

The headings of the sealed amendments read "draft, unsealed" because each text was sealed as written, at the time recorded here. Nothing in them is provisional. The SHA-256 of the file is the seal.

How to read claims

- OBSERVED

 recorded in a primary document
- ALLEGED · CONTESTED

 asserted by a party, disputed
- INFERRED

 the author's reading, marked as the author's
- OPEN

 no means of establishing it was found here

Source register

1	Reconciled coding of the 31 systems in scope: counts after the ruling check, the source check and the decisions of coder C (Amendments 4 to 6); file reconciled/reconciled_coding_wide.csv.
2	Ledger block E, v0.2: 17 systems with an official finding about the system; dates are those of the finding document; first-use dates are not verified.
3	Angelova, Dobbie and Yang, Review of Economic Studies (2025), DOI 10.1093/restud/rdaf084; NBER working paper 31747 and NBER Digest, January 2024 (abstract level).
4	EU AI Act, Arts 14 and 26: human oversight and deployer obligations (official text checked 6 October 2026).
5	Data (Use and Access) Act 2025, s. 80: UK GDPR Arts 22A to 22D; in force 5 February 2026 (S.I. 2026/82).
6	Judiciary of England and Wales, AI guidance for judicial office holders (refreshed version, April 2025).
7	Treasury Inspector General for Tax Administration, report 2025-308-022 on the use of AI in selecting returns for examination (May 2025).
8	GOV.UK algorithmic transparency record: DWP Universal Credit advances model.
9	GOV.UK algorithmic transparency record: HMRC Business Propensity to Pay.
10	French ministry of justice, personal-data notice on the processing called DataJust (justice.fr).
11	Estonian Tax and Customs Board, pages on the processing of personal data and on the scheduling of tax debt (emta.ee).
12	Dutch parliamentary paper, annex blg-1014238 (officielebekendmakingen.nl), on a register of fraud signals.
13	Amendment 4: rulings of coder C on seven interpretation questions (six marked contested) and the ruling check of flagged cells.
14	Amendment 5: handling of source checks that could not be completed; states unverified and disputed.
15	Rotterdam District Court, news item on a trial of AI as a writing aid in a criminal judgment (rechtspraak.nl).
16	Estonian Tax and Customs Board, page on the scheduling of tax debt (emta.ee).
17	Estonian Tax and Customs Board, page on the processing of personal data (emta.ee).
18	Amendment 6: three further rulings of coder C and the way decisions are applied.
19	UK Department for Work and Pensions, Effectiveness Assessment of the Universal Credit Advances Model, 1 April 2025 to 31 March 2026, published 9 July 2026 (GOV.UK).
20	US Social Security Administration, Office of the Inspector General, National Rollout of Quick Disability Determinations, audit report A-01-09-19030 (May 2009).
21	Virginia Criminal Sentencing Commission, Annual Report 2024, chapter on concurrence with the nonviolent offender risk assessment (contents and summary text; rate tables not retrieved).
22	Kleinberg, Lakkaraju, Leskovec, Ludwig and Mullainathan, Quarterly Journal of Economics 133(1), 2018, DOI 10.1093/qje/qjx032 (abstract level; a simulation).
23	

Grgic-Hlaca, Ali, Gummadi and Vaughan, CSCW 2024 (arXiv 2407.11225) (abstract level; laboratory experiment).

24 De-Arteaga, Fogliato and Chouldechova, CHI 2020 (arXiv 2002.08035) and the working-paper version on misestimated scores (abstract level; outside the three functions).

25 Interpretation rule R1: the number of systems reflects the volume of public declaration, not the volume of use.

26 Source check under Amendments 3 and 5: results per coder for the published cells of the 31 systems.

27 Specification of agreement and reconciliation (document 06).

28 The author's book AI Agents at the Border (ERA Society), the source of the matrix structure.

29 Ledger block A: no provision in the sealed list of instruments requires publication of a human divergence rate.

30 GDPR Art 22(1) and (2)(b): automated individual decisions.

31 EU AI Act Arts 49 and 71 and Annex VIII: registration and the EU database.

32 Algorithmic transparency recording standard of the United Kingdom (GOV.UK hub page).

33 National algorithm register of the Netherlands (algoritmes.overheid.nl).

34 EU AI Act Art 27: fundamental-rights impact assessment.

35 Directive on Automated Decision-Making of Canada, s. 6.1: algorithmic impact assessment.

36 EU AI Act Art 26(11): information to persons subject to a high-risk system.

37 GDPR Arts 13(2)(f), 14(2)(g) and 15(1)(h): information on automated decisions.

38 EU AI Act Art 86: right to explanation of individual decisions.

39 Directive on Automated Decision-Making of Canada, ss. 6.3 and 6.4: human intervention and recourse.

40 Regulation (EU) 2026/1744, Official Journal 24 July 2026, in force 27 July 2026: Annex III application date 2 December 2027.

41 Amendment 1, rule A6: courts and tribunals go to L1; administrative decisions go to L2 or L3.

42 Amendment 2, rule B9: customs outside all functions.

43 Coding protocol v1.0, section 4: inclusion rule and time window.

44 Coding protocol v1.0, section 5: admissible sources S1 to S6.

45 Amendment 2, rule B7: reviews and source type S4.

46 Amendment 2, rule B2: court decisions and hosting.

47 Amendment 1, rule A3: country cells; US federal level and each eligible state are separate cells.

48 Amendment 1, rule A4: a US state with a basis (b) system is screened in all three functions.

49 System list v1, sealed: 62 systems, from a union of 127 before the cap; 65 rows beyond the cap.

- | | |
|----|--|
| 50 | Algorithm register of the Netherlands, entries of the tax administration. |
| 51 | Amendment 1, rule A8: definition of findings. |
| 52 | Amendment 3: coded scope, empty cells, file format, single- and double-coded systems, mechanical source check; issued after coding began; it changes no criterion. |
| 53 | Amendment 2, rule B1: closing date and cutoff, 7 October 2026, 23:59 (UTC+4). |

A stress test, not a prediction; not an allegation against any named body. Nothing here constitutes legal advice. Statements of fact are limited to matters recorded in the public sources cited and are reports of what those sources state.