



EPIC-PP-2026-015 · FLAGSHIP RESEARCH · AI GOVERNANCE

Authority Without Office: Machine Judgement and the Missing Appointment in Public Decision-Making

Why an accurate system can be an authority on a question without ever having authority over a citizen, and what governments should build to close the gap

EPIC – Emirati Perspectives and Insight Centre

Authority Without Office: Machine Judgement and the Missing Appointment in Public Decision-Making

EPIC – EMIRATI PERSPECTIVES AND INSIGHT CENTRE

EPIC-PP-2026-015 · AI Governance · Published 20 October 2026

© 2026 EPIC – Emirati Perspectives and Insight Centre. Released under Creative Commons Attribution 4.0 International (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>

EPIC is an independent research initiative in formation. It is not yet a registered or licensed entity. EPIC does not accept donations, payments or contracts.

Abstract

Public debate about artificial intelligence in government usually asks whether machine outputs are accurate, fair or explainable. This paper asks a prior question: by what right does a machine output bind anyone? Drawing on Joseph Raz's distinction between theoretical and practical authority, and on Hobbes's account of authors and actors, it argues that a system can be an authority on a question, in the sense that its outputs give good reasons for belief, without ever having authority over a person. Raz himself considers, and rejects, the objection that practical authority needs something like an appointment in addition to the service conception's conditions; this paper takes up that objection and defends a version of it for public decisions. Its claim is that, in public administration, binding practical authority belongs to offices, and offices have holders who answer, can be overruled, can be challenged and were placed there by someone entitled to place them. The paper names the failure that occurs when this structure is missing "pre-emptive drift": the gradual slide of a machine output from advice to directive, without any act of appointment. It traces that drift through four documented cases, Australia's Robodebt scheme, the Dutch childcare benefits scandal, England's 2020 A-level grading algorithm and the Wisconsin Supreme Court's decision in *State v. Loomis*, and draws on the published figures for each. It then proposes an "office test" with four conditions (answerability, revocability, contestability and standing) that any machine-assisted public decision should meet before its output may bind a citizen. Existing law, including Article 22 of the GDPR and Article 14 of the EU AI Act, addresses human oversight but leaves the question of office largely open. The paper closes with the implications for the United Arab Emirates, whose Cabinet approved an AI-driven Regulatory Intelligence Office in April 2025, and argues that the UAE's choice to house the system inside an office of the Cabinet is the right institutional instinct.

Keywords: artificial intelligence; authority; public administration; Joseph Raz; Hobbes; automated decision-making; accountability; automation bias; AI governance; United Arab Emirates

I. Introduction: the letter with no signature

Between 2015 and 2019, hundreds of thousands of Australians received letters from their government telling them they owed money. The amounts were calculated by matching annual tax-office income data against fortnightly welfare declarations and averaging the difference. The scheme raised about A\$1.7 billion in alleged debts against some 526,000 people (Sydney Morning Herald 2023). In late 2019 the Commonwealth settled a test case and accepted that key parts of the scheme were unlawful. The Royal Commission that examined it concluded that "Robodebt was a crude and cruel mechanism, neither fair nor legal, and it made many people feel like criminals. In essence, people were traumatised on the off-chance they might owe money. It was a costly failure of public administration, in both human and economic terms" (Royal Commission into the Robodebt Scheme 2023, p.xxix).

Accounts of Robodebt usually dwell on the arithmetic. Averaging annual income across fortnights is a crude way to estimate entitlement, and it produced debts that did not exist. The arithmetic was the visible flaw. The deeper one was that a calculation had been allowed to function as a decision. A number produced by matching two datasets arrived in a citizen's letterbox with the force of a demand, and the burden fell on the citizen to disprove it. Nobody in particular had decided that this person owed this sum. The system had produced a figure, and the figure had become a debt.

This paper is about that conversion. Accuracy matters and bias is real, but the question here comes before both: what has to be true before a machine output can bind a person, whether by creating an obligation, denying an entitlement, assigning a grade or informing a sentence? The answer, developed in dialogue with Joseph Raz's theory of authority and partly against it, is that binding force belongs to offices, and that an office is a specific institutional arrangement that machines cannot occupy and that governments too often forget to build around them.

The argument has a practical edge. Governments, including the UAE's, are moving quickly to put artificial intelligence at the centre of public administration. The question of office is the one they can still settle by design, before the cases arrive.



Figure 1. Frontispiece of Thomas Hobbes's *Leviathan* (1651), engraving by Abraham Bosse. Hobbes's figure of the sovereign composed of the citizens who authorise him is the classic image of authority as something conferred. British Library copy via Wikimedia Commons; public domain (CC0).

2. Two kinds of authority

2.1 Raz's service conception

The most influential modern account of authority is Joseph Raz's service conception. Its core is the normal justification thesis: a person has authority over another when the subject would better conform to the reasons that apply to them by following the authority's directives than by trying to act on those reasons directly (Raz 2006, p.1014). Raz adds an independence condition: the matters must be ones on which it is better to conform to reason than to decide for oneself (p.1014). Legitimate directives then pre-empt: they do not simply add one more reason to the balance, they replace some of the background reasons on which they were meant to be based (pp.1018–1019).

On a first reading this looks like a charter for machine authority. If a well-designed system would lead an official, or a citizen, to conform to the applicable reasons more reliably than unaided judgement, the normal justification thesis seems satisfied. Many advocates of algorithmic government make exactly this argument, often without knowing they are making it.

2.2 Being an authority and having authority

Raz himself blocks the inference. In his 2006 restatement he distinguishes theoretical from practical authority, and the distinction is decisive for machines. "In theoretical matters," he writes, "someone either is or is not an authority, but no one has authority" (p.1034). A reliable encyclopaedia, a leading physician or an excellent forecasting model can be an authority on a subject. Its statements give us reason to believe. Nothing more is needed for that kind of authority than the reliability itself.

Practical authority, the kind that gives reasons for action and can bind, raises a further question, and here Raz's text must be read with care. In Part IV of the same article he sets out what he calls the "qualification objection" to his own view: that the service conception describes when someone would be a good authority, not what it is to be one. The objection's illustration is political: "Not everyone who can be a good prime minister of a country is the prime minister of that country ... Something else has to happen to give them the task, to make it their task" (p.1032). Applied to practical authorities, "so the argument goes", they "need to be made authorities, not necessarily by being appointed to the job, but something like an appointment has to be there" (p.1035).

These are not Raz's own conclusions. He replies at once that the objection's failure for theoretical authority "seems to me to establish that it fails altogether" (p.1035), conceding at most that his account may be "only a partial analysis" of practical authority (p.1035). He then considers a "more modest objection", that the service conception states a necessary but not a sufficient condition for practical authority, which he finds "more plausible" (p.1036). His answer to it is that the conditions of legitimacy already do much of the work: political authorities must in fact be followed by considerable parts of the population, so that "there cannot be an unknown political authority" (p.1036), and the service conception "recognizes and has some account of the difference between being qualified to hold authority, and having authority". Whether that account is adequate, he writes, "is still open" (p.1037).

This paper takes up the question Raz leaves open, and answers it differently. EPIC's claim is that, at least for decisions taken by public administration, the qualification objection is right: expertise, however great, does not create the authority to bind a citizen. Something must happen to place a decision in someone's hands, and in a constitutional state that something is an appointment or an institutional equivalent (an election, a statutory delegation, a commission). It creates what political language calls an office. Raz would not accept this as a general analysis of authority, and the paper does not claim his support for it. It claims only that his own concessions (that practical authority depends on being followed, known and able to change things, and that the adequacy of his account here is open) leave room for it, and it defends the requirement on its own terms in the sections that follow.

2.3 What an office is

An office is a role that persists across holders, defined by a jurisdiction (the class of cases it may decide), a source (the body that created it and appoints to it), a holder (the person currently answerable for its exercise) and a set of relations to others (who may overrule it, who may appeal against it, to whom it reports). When a tax inspector issues an assessment, the binding force does not come from the inspector's arithmetic. It comes from the office the inspector holds, which was created by statute, filled by appointment, bounded by law and subject to appeal. The arithmetic may be wrong and the assessment still valid until set aside. The arithmetic may be right and the assessment still void if the inspector lacked jurisdiction.

Machines can be extraordinarily good at the arithmetic. They can be authorities on questions in Raz's theoretical sense. What they cannot do is hold an office, because an office requires a holder who can answer for its exercise, and an answer is something only an accountable agent can give.

2.4 Why accuracy cannot close the gap

One might expect the gap between theoretical and practical authority will close as systems improve. If a model is right ninety-nine times in a hundred, why should its outputs not simply bind? There are three answers, and none depends on the model being wrong.

The first is that error is distributed. A model that is right ninety-nine times in a hundred is wrong for one person in a hundred, and that person experiences the error as a decision about them. Offices exist so that the one person has somewhere to go. Raz's normal justification thesis is stated person by person: authority is justified over a particular subject when that subject would do better by following it (Raz 2006, p.1014). A population-level accuracy rate does not establish that for any individual in particular, and certainly not for the individual the model has misjudged.

The second is that binding decisions do more than track facts. A debt notice, a grade or a sentence also allocates burdens, closes options and expresses the state's judgement of a person. These are acts, and acts need agents who can own them. Accuracy is a property of a prediction. Ownership belongs to a relation between an institution and the person it governs.

The third is that authority must be reversible to remain legitimate. An office can be held to account, reformed or abolished through political processes that citizens can reach. A system's authority, if it were allowed to rest on accuracy alone, would be answerable only to whoever controls its training and evaluation, which is to say to a technical process citizens cannot reach at all. Authority would survive, relocated to a place outside public life.

3. Authors and actors

Hobbes reached a version of the same point three centuries earlier. In chapter 16 of *Leviathan* he distinguished the actor, who performs an action, from the author, who owns it. "And as the right of possession, is called dominion; so the right of doing any action, is called 'authority.' So that by authority, is always understood a right of doing any act; and 'done by authority,' done by commission, or license from him whose right it is" (Hobbes 1651, ch.16).

Hobbes's framework fits machine decision-making with uncomfortable precision. A system that produces a debt notice or a grade is an actor. It does something. Whether it acts by authority depends on whether some author has owned the action, by commissioning it within limits. Hobbes is explicit that "things inanimate cannot be authors, nor therefore give authority to their actors" (ch.16). He adds a warning that applies directly to automated government: a person who deals with an actor "not knowing the authority he hath, doth it at his own peril", and an actor's commission extends "so far forth as is in their commission, but no further" (ch.16).

Two consequences follow. First, every binding machine output must be traceable to an author, a human office that commissioned the system for that class of case and owns its results. Second, the commission has limits. A system authorised to flag discrepancies for investigation has not been authorised to issue debts. When the output goes beyond the commission, it is, in Hobbes's language, an action without authority, however accurate it may be.

4. Pre-emptive drift

4.1 The mechanism

Machine outputs rarely start as directives. They are introduced as advice: a risk score to help a caseworker prioritise, a statistical model to help moderate grades, a matching tool to identify cases for review. Over time, through a series of small and individually reasonable steps, the advice comes to function as the decision. This paper calls that process **pre-emptive drift**, borrowing Raz's term. The output begins to pre-empt the official's own reasons, which is what directives do, without anyone having appointed the system to direct.

Four mechanisms drive the drift.

Volume. A system that produces thousands of outputs a day cannot be reviewed case by case at the same depth as a human caseload. Review becomes sampling, then confirmation, then formality.

Default. When the system's output is the default and departing from it requires justification, the burden of reasons shifts. The official must explain why they disagree with the machine; nobody must explain why they agreed.

Reversal of the burden. When the output is communicated to the citizen as a finding, the citizen must disprove it. In Robodebt, the effect was that people had to show they did not owe money the state had never proved they owed.

Opacity to the reviewer. When the official responsible for review cannot see why the system produced its output, review is impossible in substance. Amnesty International's investigation of the Dutch childcare benefits system found exactly this: when a person was flagged as a fraud risk, "a civil servant was required to conduct a manual review but was given no information as to why the system had generated a higher-risk score" (Amnesty International 2021).

4.2 Rule by Nobody

Hannah Arendt described the political condition at the end of this drift in 1969, long before machine learning. Bureaucracy, she wrote, is "the rule by an intricate system of bureaux in which no men, neither one nor the best, neither the few nor the many, can be held responsible, and which could be properly called the rule by Nobody". She drew the conclusion that matters here: "if we identify tyranny as the government that is not held to give account of itself, rule by Nobody is clearly the most tyrannical of all, since there is no one left who could even be asked to answer for what is being done" (Arendt 1969).

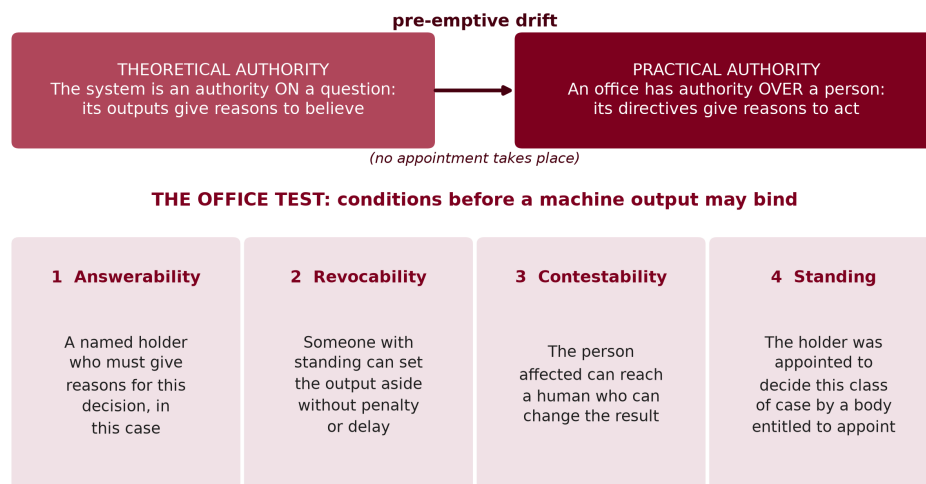
Automated decision systems make rule by Nobody easier to achieve and harder to see. A bureaucracy at least has desks. When a decision emerges from a data match, a model and a mail merge, there may be no desk at which it was taken. The official who signs the policy did not decide the case. The official who handles the complaint did not make the decision being complained about. The system did not decide either, because systems cannot decide in the sense that matters. The decision has been taken by Nobody.

4.3 How drift differs from error

Drift should be distinguished from malfunction. A system can be accurate and still drift, and a drifted system can be fixed technically and still lack authority. The 2020 A-level algorithm in England was not a coding error. It did what it was designed to do. Its outputs were wrong for many individuals because the design gave a statistical estimate the force of a grade. The failure was institutional before it was technical.

4.4 Drift in generative systems

The cases in this paper involve scoring and matching systems, but the same drift is now possible with generative models. A caseworker who uses a language model to draft the reasons for a decision letter has not delegated the decision. A caseworker who signs, un-read or barely edited, a letter the model drafted from a case file has delegated something closer to it. The model has answered a question and has also written the state's reasons. When those reasons are challenged, the official must defend text they did not compose, on the basis of a reading of the file they may not have made. Each of the four mechanisms reappears: volume (more letters per day), default (the draft is the starting point), reversal of the burden (the citizen must find the error in fluent prose) and opacity to the reviewer (the official cannot see why the model emphasised one fact over another). The office test applies without modification. What matters is whether a named holder made the decision and can give reasons for it that are their own, whichever tool produced the words.



Source: EPIC conceptual framework. The theoretical/practical distinction is from J. Raz, 'The Problem of Authority: Revisiting the Service Conception' (2006); the appointment (standing) condition is EPIC's, taken from the 'qualification objection' that Raz sets out and rejects (pp.1032–1037).

EPIC

Figure 2. Theoretical authority, practical authority and the office test. Pre-emptive drift occurs when outputs move from the first box to the second without any act of appointment. EPIC conceptual framework, drawing on Raz's distinction between theoretical and practical authority (2006, pp.1032–1037) and on the qualification objection he discusses there, which EPIC adopts and Raz rejects.

5. Four documented cases

The cases below were chosen because each has an authoritative public record: a royal commission, a parliamentary inquiry, regulatory statistics or a supreme court judgment.

Each case was also chosen to show the drift at a different point.

- **Robodebt** is the case where commission was missing from the start: the need for legislative change dropped out of the record before the system was built, so the failure is one of standing (l.118).
- **The Dutch childcare benefits scandal** keeps a human reviewer in the loop and still fails, because the reviewer was given no reasons, which makes it the test case for the view that oversight alone is the remedy (l.124).
- **England's 2020 A-levels** involve a model that did what it was designed to do, so the case separates drift from error and shows how quickly offices can take a decision back (l.91, l.109).
- **State v. Loomis** is the partial success, where a court let a score inform a decision while keeping the judge answerable for it, and it shows the office test can be met (l.152–154).

Taken together they run from a scheme with no office behind it to a court that kept its office intact.

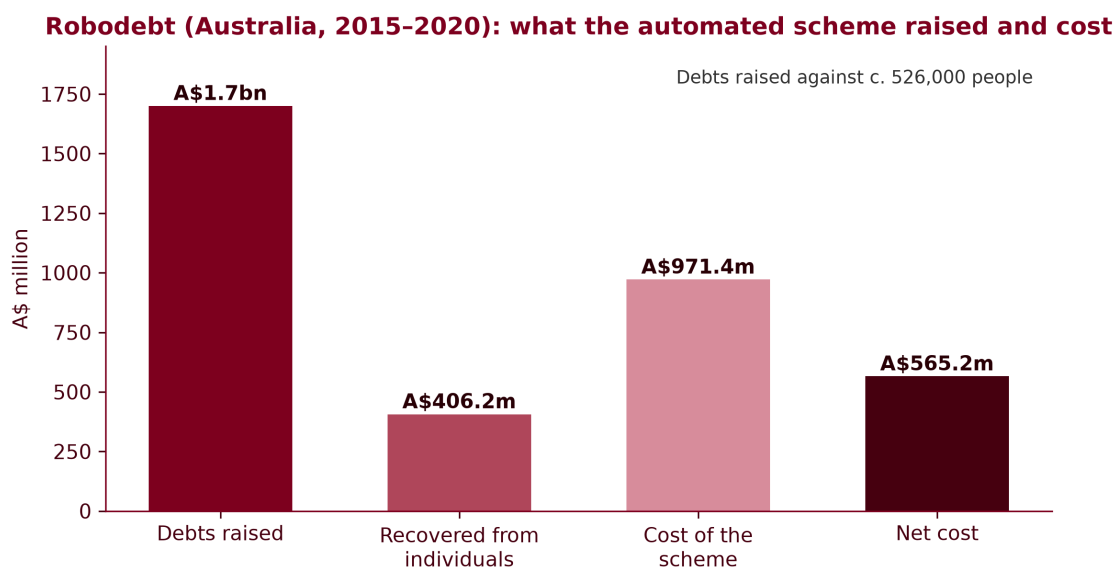
Table 1 summarises the documented facts.

Table 1. Four cases of machine judgement in public decisions: documented facts

Case	Period	Function of the system	Documented scale or outcome	Source
Robodebt (Australia)	2015–2020	Income averaging to raise welfare overpayment debts	A\$1.7bn raised against c.526,000 people; A\$406.2m recovered; A\$971.4m cost; net cost A\$565.2m	Royal Commission (2023), via SMH (2023)
Childcare benefits (Netherlands)	From 2013	Risk-classification model flagging benefit claims as potentially fraudulent	Inquiry found "fundamental principles of the rule of law" violated; tax office admitted 11,000 people received extra scrutiny because of dual nationality; cabinet resigned 15 January 2021	Tweede Kamer (2020); BBC (2021); Amnesty (2021)
A-level grading (England)	August 2020	Statistical model moderating teachers' assessed grades	39.1% of grades lowered from teachers' assessments; reversed within four days of results	Ofqual data via Schools Week (2020)
COMPAS in sentencing (Wisconsin, US)	2016 judgment	Proprietary risk assessment in pre-sentence reports	Court ruled risk scores may not determine incarceration or severity of sentence; mandatory written warnings	<i>State v. Loomis</i> , 2016 WI 68

5.1 Robodebt: a calculation that became a debt

Robodebt was designed to identify possible overpayments by comparing annual income reported to the tax office with income declared fortnightly to the welfare agency. Used as an investigative lead, such a comparison has value. Used as the basis of a debt, it does not, because annual averages say nothing reliable about fortnightly income for people with irregular work. The Royal Commission's report, delivered on 7 July 2023 after nine weeks of public hearings, 115 witnesses and 1,092 submissions, ran to 990 pages and made 57 recommendations (Sydney Morning Herald 2023).



Source: Royal Commission into the Robodebt Scheme, Report (7 July 2023), figures as reported by the Sydney Morning Herald (7 July 2023).

EPIC

Figure 3. Robodebt: debts raised, amounts recovered and the cost of the scheme (A\$ million). Source: Royal Commission into the Robodebt Scheme (2023), figures as reported by the Sydney Morning Herald, 7 July 2023.

The finding most relevant to the office question concerns legal authority itself. The Commission found that a stated need for legislative change to implement the scheme disappeared from a policy document taken to Cabinet in 2015 (Sydney Morning Herald 2023). In the terms of this paper, the question of commission, whether any author had the right to authorise this action, was dropped from the record before the system was built. Everything that followed was the work of an actor without an author.

5.2 The Dutch childcare benefits scandal: a reviewer without reasons

The Dutch case shows drift inside a system that nominally kept a human in the loop. From 2013 the tax administration used a risk-classification model to select childcare benefit applications for scrutiny. Amnesty International found that the model used Dutch nationality as a risk factor, so that non-Dutch nationals received higher risk scores, and that a self-learning element reproduced the bias over time "with no meaningful human oversight" (Amnesty International 2021). The parliamentary inquiry committee's report of

December 2020, *Ongekend onrecht* ("Unprecedented injustice"), concluded that fundamental principles of the rule of law had been violated (Tweede Kamer 2020). The tax administration admitted that 11,000 people had been subjected to extra scrutiny simply because they had dual nationality, and the Rutte government resigned on 15 January 2021. The Prime Minister said that "innocent people have been criminalised and their lives ruined" and that responsibility lay with the cabinet: "The buck stops here" (BBC News 2021). In December 2021 the Dutch Data Protection Authority fined the tax administration €2.75 million for unlawful and discriminatory processing of applicants' (dual) nationality (Autoriteit Persoonsgegevens 2021).

The most important sentence in the Amnesty report concerns the reviewer. A civil servant was required to review each flagged case, but was given no information about why the system had flagged it. The human in the loop held a position without the means to exercise it. That is the office emptied of content: a holder who must sign, who cannot see, and who cannot therefore answer.

The Dutch case also shows what office looks like when it reasserts itself. The cabinet's resignation was a political act of answerability. It came years too late for the families concerned, but it located responsibility in an office whose holders could be named.

5.3 England's A-levels, 2020: a statistical estimate given the force of a grade

When the pandemic cancelled examinations in 2020, England's exam regulator, Ofqual, asked schools for centre-assessed grades and then moderated them with a statistical model designed to keep national results in line with previous years. When results were published on 13 August 2020, 39.1 per cent of grades had been lowered from the teachers' assessments: 35.6 per cent by one grade, 3.3 per cent by two and 0.2 per cent by three. Grades were unchanged in 58.7 per cent of cases and raised in about 2.3 per cent (Schools Week 2020; The Guardian 2020). On 17 August the government reversed course and allowed students the higher of the two grades (Ofqual 2020).

Table 2. England, August 2020: calculated A-level grades compared with centre-assessed grades

Change from centre-assessed grade	Share of grades
Three grades lower	0.2%
Two grades lower	3.3%
One grade lower	35.6%
No change	58.7%
One grade higher	2.2%
Two grades higher	0.05%

Change from centre-assessed grade	Share of grades
Three grades higher	0.01%

Source: Ofqual data published 13 August 2020, as reported by Schools Week (2020).

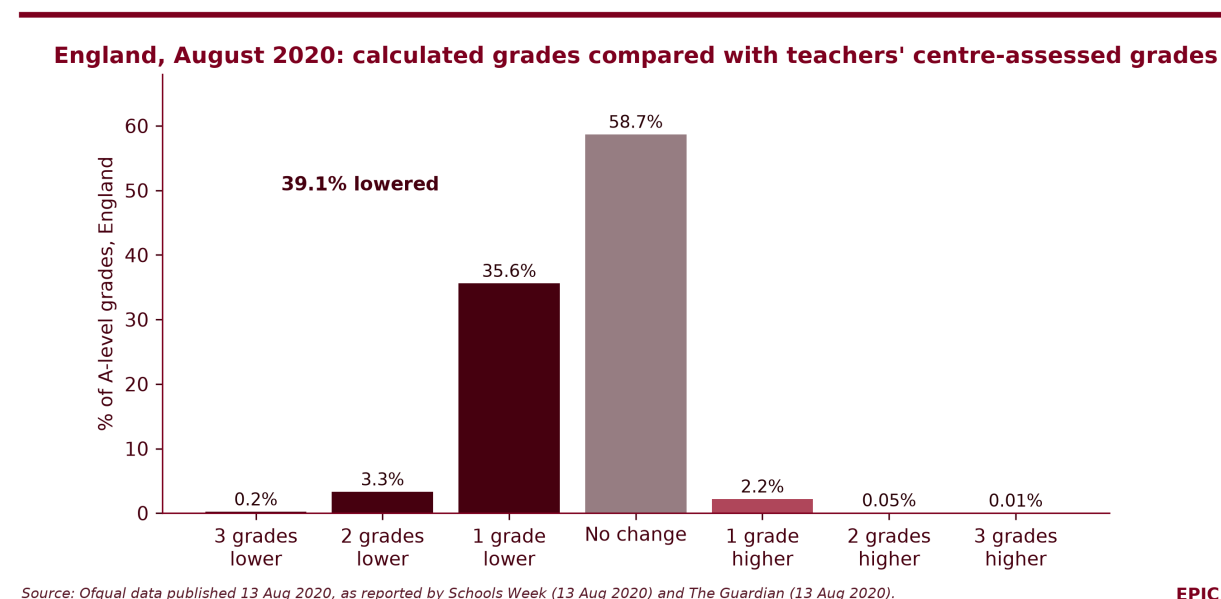


Figure 4. England, August 2020: calculated A-level grades compared with teachers' centre-assessed grades.
Source: Ofqual data as reported by Schools Week and The Guardian, 13 August 2020.

The model gave a defensible answer to a statistical question: what distribution of grades would this school have achieved, given its history? The failure was to treat the answer to that question as the answer to a different one: what grade has this student earned? A grade is a judgement about a person, issued by an office (the examining body) on the basis of that person's performance. The model produced an estimate about a population and assigned it to individuals. It was an authority on the first question and had been given practical force over the second. The four-day reversal was the system of offices reclaiming a decision it should never have delegated.

5.4 State v. Loomis: an office that held

The Wisconsin case is included because it is the partial success. Eric Loomis challenged the use of COMPAS, a proprietary risk assessment, in his sentencing. In July 2016 the Wisconsin Supreme Court upheld the sentence but set limits. Risk scores may not be used "to determine whether an offender is incarcerated" or "to determine the severity of the sentence", and may not be the determinative factor in deciding whether an offender can be supervised safely in the community (*State v. Loomis* 2016, ¶198). Pre-sentence reports that include COMPAS must carry a written advisement listing the tool's limitations, among them that its proprietary nature has prevented disclosure of how scores are determined and that, because the scores are based on group data, "they are able to identify

groups of high-risk offenders" and not a particular high-risk individual (§100). The sentencing court must explain the factors, other than the score, that support its sentence (§99).

Critics have argued that these safeguards are thin, and the critique has merit: a warning label does not make a proprietary model contestable. What *Loomis* did preserve was the office. The judge remains the decision-maker, must give reasons that stand independently of the score, and can be appealed. The score is permitted to be an authority on a question (the statistical risk profile of offenders like this one) and forbidden to have authority over the defendant. That is the Razian distinction written into a judgment.

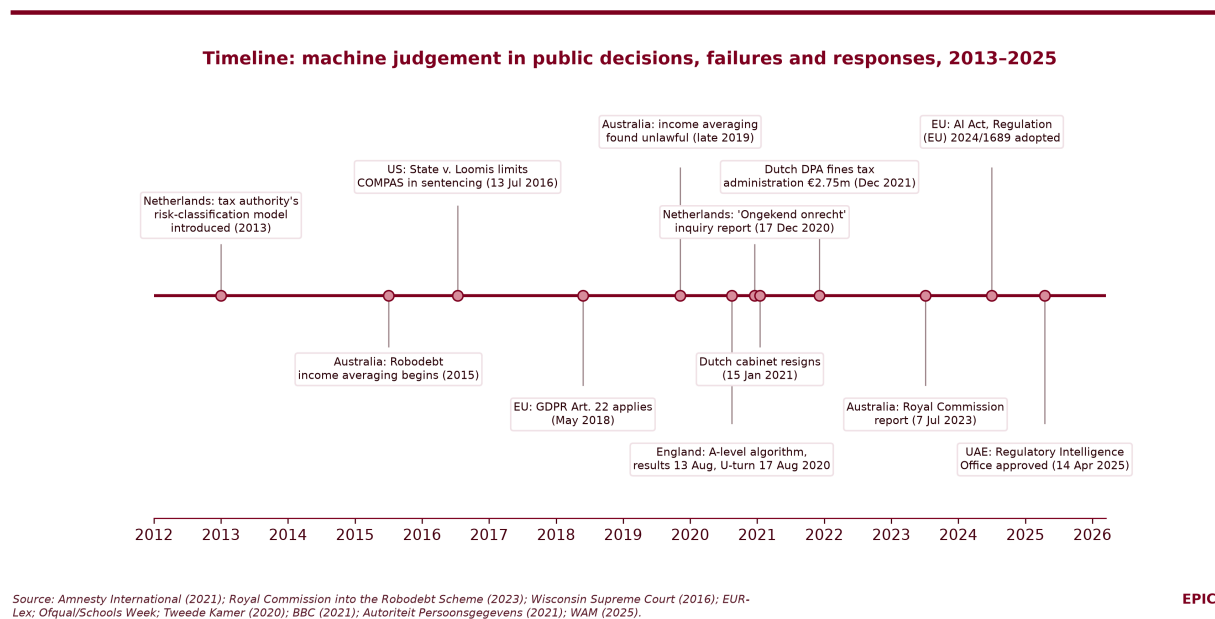


Figure 5. Timeline of documented failures and responses involving machine judgement in public decisions, 2013–2025. Sources: Amnesty International (2021); Royal Commission (2023); State v. Loomis (2016); EUR-Lex; Ofqual/Schools Week (2020); Tweede Kamer (2020); BBC (2021); Autoriteit Persoonsgegevens (2021); WAM (2025).

5.5 A rising number of incidents

These cases are not isolated. The Stanford AI Index reports that the AI Incident Database recorded 362 AI-related incidents in 2025, up from 233 in 2024 (Stanford HAI 2026, ch.3), an increase of about 55 per cent (EPIC calculation). The database counts incidents reported in public sources, its totals are revised over time, and many entries concern private systems, so the figure measures visibility as much as frequency. It is nonetheless a fair indication of direction: more systems, in more settings, producing more outputs that someone has allowed to bind.

6. The office test

If binding force belongs to offices, the practical question is how to tell whether a machine-assisted decision sits inside one. The paper proposes four conditions, each derived from the account of office in section 2.3. All four must be met before a machine output may bind a citizen.

Answerability. There is a named office whose holder must give reasons for this decision, in this case, in terms that do not reduce to "the system said so". The reasons must be available to the holder, which means the holder must be able to see why the system produced its output. The Dutch reviewer who was given no reasons fails this condition by design.

Revocability. Someone with standing can set the output aside, in the ordinary course of business, without penalty, special justification or delay. If departing from the system requires more paperwork than following it, the output has acquired a directive force no one conferred.

Contestability. The person affected can reach a human decision-maker who has the power to change the result, before the decision takes effect where the stakes allow it. The burden of proof stays where the law places it. A system that turns a data mismatch into a debt the citizen must disprove fails this condition, as Robodebt did.

Standing. The office was created to decide this class of case, by a body entitled to create it, and the system was commissioned within that jurisdiction. This is Hobbes's condition of commission, and the "something like an appointment" that the qualification objection discussed by Raz demands (Raz 2006, p.1035; Raz himself rejects that objection, as section 2.2 explains). Robodebt's dropped legislative step is the clearest failure of standing in the record.

The test is deliberately institutional. It asks nothing about the system's accuracy, which is a separate requirement. A highly accurate system can fail the office test, and a mediocre one can pass it if its outputs are treated as advice by an office that answers for the decision. That is the point. Accuracy earns a system theoretical authority. Only the office test can justify practical force.

6.1 A worked example

Consider a hypothetical municipal housing authority that introduces a model to prioritise applications for emergency housing. The example is illustrative and describes no real system.

In the first design, the model ranks applications, caseworkers receive the ranked list each morning and allocate units from the top, and applicants receive a letter stating their priority band. Departing from the ranking requires a written justification approved by a su-

pervisor. Applicants who disagree may ask for a review, which is carried out by the same team using the same list. No document records which office decided to use the model for allocation, as distinct from triage.

Applied to this design, the office test fails on every condition. Answerability fails because no caseworker decides an individual allocation; they process a list. Revocability fails because departure costs more than compliance. Contestability fails because review returns the applicant to the system that produced the result. Standing fails because the commission was never recorded.

In the second design, the model's ranking is shown to caseworkers alongside the case file and the main factors that drove the score. Each allocation is recorded as the decision of a named caseworker, with reasons that refer to the applicant's circumstances. Caseworkers depart from the ranking whenever their reasons warrant it, and the rate of departures is reported monthly as a sign of healthy use. Applicants may ask a senior officer, who has not seen the model's score, to reconsider. The authority's board has adopted a resolution authorising the model for prioritisation, setting its limits and naming the post responsible.

The model is identical in both designs. Its accuracy is identical. Only the second design gives it a legitimate place in the exercise of public authority, because only the second places it inside an office.

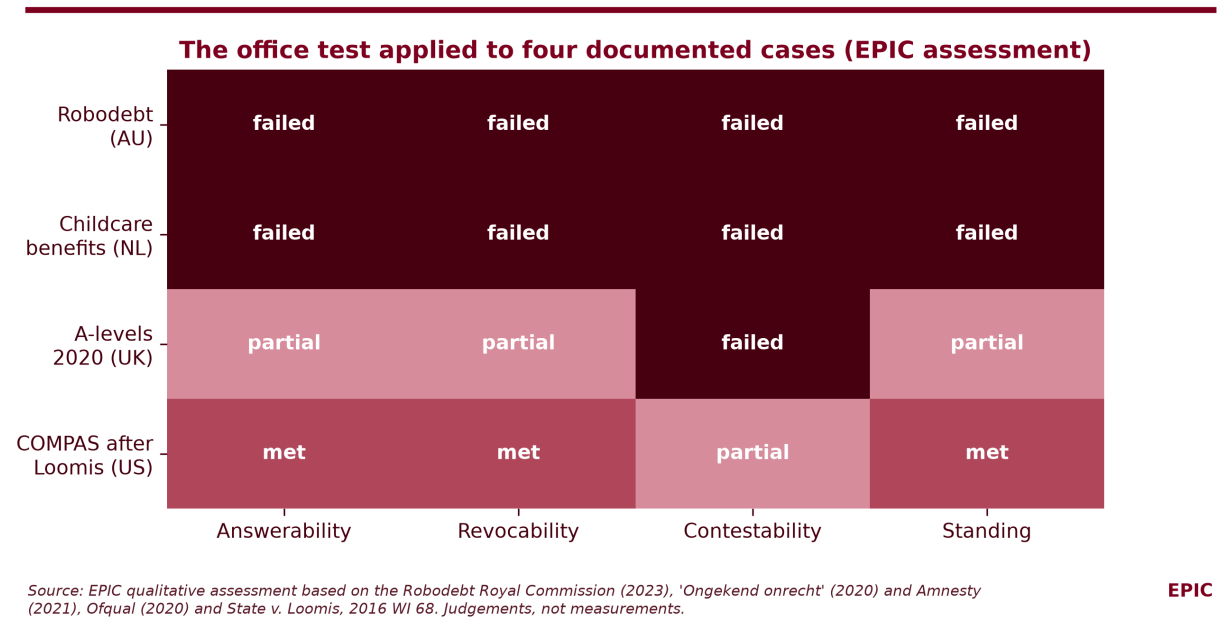


Figure 6. The office test applied to the four cases. EPIC qualitative assessment, not a measurement, based on the sources listed for each case.

The assessment in Figure 6 is a judgement and should be read as such. Robodebt and the Dutch system fail all four conditions on the documented record. The A-level model meets some conditions partially: Ofqual was a body with standing, ministers did answer for the

outcome and the decision was revoked, but students had no effective route to contest an individual grade before it affected university places. *Loomis* comes closest to passing, with contestability only partial because the model itself remains proprietary.

7. What the law already says, and what it leaves open

European law has engaged with these questions more directly than most. Article 22(1) of the General Data Protection Regulation gives a person "the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her" (Regulation (EU) 2016/679). The provision captures the problem of decisions without a human decision-maker, but it is triggered only by decisions based *solely* on automated processing. Pre-emptive drift lives in the gap: a human signature on a decision the human did not in substance make takes the case outside the literal scope of the rule.

The EU AI Act goes further. Article 14 requires that high-risk systems be designed so that they "can be effectively overseen by natural persons". The people assigned to oversight must be enabled "to remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (automation bias)", and "to decide, in any particular situation, not to use the high-risk AI system or to otherwise disregard, override or reverse the output" (Regulation (EU) 2024/1689, Art.14(1), (4)(b), (4)(d)). For remote biometric identification, Article 14(5) adds a stricter rule: identifications must be separately verified by at least two natural persons "with the necessary competence, training and authority". The Act also includes, in Article 86, a right to an explanation of individual decision-making.

The word "authority" in Article 14(5) is revealing. The drafters understood that oversight without authority is empty. Even so, the Act's architecture is built around the oversight *function*, a set of capabilities that a person must have. It says much less about the *office*: who appoints the overseer, to whom the overseer answers, what happens when they override, and how the affected citizen reaches them. Article 14(4)(d) in particular secures the legal possibility of override, the revocability condition, but a legal possibility is not the same thing as an ordinary practice. The office test is meant to fill that space.

Table 3. How existing provisions map onto the office test

Provision	Answerability	Revocability	Contestability	Standing
GDPR Art.22	Indirect (requires human involvement)	Not addressed	Partial (rights to human intervention and to contest, under Art.22(3))	Not addressed

Provision	Answerability	Revocability	Contestability	Standing
AI Act Art.14	Partial (oversight persons must understand outputs)	Addressed (power to disregard, override or reverse)	Not addressed directly	Partial (Art.14(5): "competence, training and authority")
AI Act Art.86	Partial (explanation of individual decisions)	Not addressed	Partial	Not addressed
<i>State v. Loomis</i>	Addressed (court must give independent reasons)	Addressed	Partial	Addressed (judge retains decision)

Source: EPIC analysis of the texts cited. "Partial" indicates that the provision addresses part of the condition.

8. Objections

8.1 "If the machine is more accurate, Raz's own theory says we should defer"

The normal justification thesis does say that a person should defer to a source that would help them conform better to their reasons. Applied to machines, it justifies something real: officials should take reliable systems seriously as theoretical authorities and should often follow their advice. It does not, on this paper's argument, make the system a practical authority over the citizen. Here EPIC departs from Raz. Raz rejects the qualification objection, the view that practical authority requires "something like an appointment" beyond the service conception's conditions (2006, p.1035), and a strict Razian could reply that a sufficiently reliable system, consistently followed, would meet those conditions. The paper's answer has three parts. First, Raz himself treats the sufficiency of his conditions for practical authority as an open question (p.1037). Second, the normal justification thesis is assessed subject by subject (p.1014), and a population-level accuracy rate does not show that a given citizen, least of all a misjudged one, would do better by deferring (section 2.4). Third, the cases in section 5 show that where no office stood behind the output, nobody could be asked to answer for it, and that this, rather than inaccuracy alone, was what made the outcomes unacceptable. The appointment requirement is therefore EPIC's position, defended on institutional grounds, not a doctrine attributed to Raz. Deference by an official to good advice is how a well-run office uses expertise. The office remains the authority. The advice remains advice.

8.2 "Human officials are biased and inconsistent too"

They are. The office test makes no claim that human judgement beats machine judgement case by case. Its claim concerns who can answer. A biased official can be named, questioned, overruled and removed. The office survives the holder's failings because it has the machinery for correcting them. A biased system inside a structure that lacks an office has no such machinery, which is why the Dutch and Australian errors ran for years.

8.3 "Keeping a human in the loop solves the problem"

A human in the loop is necessary and insufficient. The Dutch reviewer was in the loop. What the reviewer lacked was the substance of an office: the information to give reasons, the practical freedom to depart and the standing to change the outcome. A loop describes a workflow. An office describes a relation of authority. Governments should require the second and will find that it entails the first.

9. Implications for the United Arab Emirates

The UAE is among the governments moving fastest to place artificial intelligence inside the machinery of state. On 14 April 2025 the Cabinet, chaired by His Highness Sheikh Mohammed bin Rashid Al Maktoum, approved the launch of what it described as the first integrated regulatory intelligence ecosystem in the UAE Government and the establishment of a Regulatory Intelligence Office under the General Secretariat of the Cabinet. The system is designed to connect federal and local legislation with judicial rulings, executive procedures and public services, to track the impact of laws using large-scale data, to "regularly suggest updates to our legislation", and to accelerate the legislative process by up to 70 per cent (WAM 2025). The Cabinet described the initiative as reflecting a commitment to "a new generation of governance that blends human insight with AI efficiency".

Seen through the office test, the design choice that matters most is the one least discussed: the system was placed inside an office. It suggests; the Cabinet and the country's constitutional legislative process decide. The Regulatory Intelligence Office has a defined source (the Cabinet), a location (the General Secretariat) and a jurisdiction (the design, implementation and coordination of the system with federal and local authorities). That is precisely the structure that Robodebt lacked. A system that proposes legislative updates to an office that must decide on them is an authority on the state of the law and its effects. It should never become an authority over the citizen.

The risk to watch is drift. As the system becomes trusted and its suggestions numerous, the default could shift from "the office considers the suggestion" to "the office must justify declining it". The safeguards that keep the UAE's design on the right side of the line are

institutional and can be built now:

A register of binding uses. Each government entity should publish the classes of decision in which machine outputs contribute to a binding outcome for an individual or business, and the office that owns each class.

A named holder for every class. Each register entry should name the office, and the post within it, that answers for decisions in that class.

Reasons that stand without the score. Where an output contributes to a binding decision, the decision should record reasons that would stand if the output were removed, as *Loomis* requires of sentencing courts.

Override as ordinary practice. Departing from a system's output should carry no more administrative burden than following it, and the rate of overrides should be monitored. A rate near zero is a warning sign.

A human route for the citizen. Every person affected by a machine-assisted decision should be able to reach an official with power to change the result before the decision takes effect, wherever the stakes allow.

Commission before deployment. Before a system is used to produce outputs that bind, the competent authority should record the legal basis and the limits of the commission. Robodebt shows what happens when that step disappears.

Procurement that preserves the office. Contracts for systems used in binding decisions should guarantee the government access to the information its officials need to give reasons: the factors that drove an output, logs of system behaviour and the right to audit. The *Loomis* court had to warn judges that a proprietary tool's workings were undisclosed. A government that buys a system on those terms has bought an actor it cannot fully answer for.

These measures are modest, which is a point in their favour. They do not slow the adoption of AI. They make clear, at the moment of adoption, where the authority sits.

10. Conclusion

The question of whether machines can have authority usually generates either enthusiasm or alarm. Raz's distinction allows a calmer answer. A good system can be an authority on many questions, and governments are right to use it. It should not have authority over anyone, because in public decisions binding authority ought to be conferred, and what is conferred is an office with a holder who answers. That is a requirement this paper defends, going beyond what Raz himself accepts. When governments forget this, ma-

chine outputs drift from advice to command without any appointment taking place, and the result is Arendt's rule by Nobody: decisions that bind citizens and for which nobody can be asked to account.

Robodebt and the Dutch childcare system ran for years as actors without authors, while the A-level reversal and *Loomis* show offices reclaiming a decision or keeping hold of it. The office test asks the same of every binding decision, however capable the system behind it: a holder who can be asked, in Arendt's phrase, to answer for what is being done.

References

Amnesty International (2021) 'Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms', 25 October 2021, introducing the report *Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal* (EUR 35/4686/2021). Available at: <https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/> ; report: <https://www.amnesty.org/en/documents/eur35/4686/2021/en/>

Arendt, H. (1969) 'A Special Supplement: Reflections on Violence', *The New York Review of Books*, 27 February 1969. Available at: <https://www.nybooks.com/articles/1969/02/27/a-special-supplement-reflections-on-violence/>

Autoriteit Persoonsgegevens (2021) Fine of €2.75 million imposed on the Tax and Customs Administration for discriminatory and unlawful processing of the (dual) nationality of childcare benefit applicants, 7 December 2021.

BBC News (2021) 'Dutch Rutte government resigns over child welfare fraud scandal', 15 January 2021. Available at: <https://www.bbc.com/news/world-europe-55674146>

European Union (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), Article 22. *Official Journal of the European Union*, L 119.

European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Articles 14 and 86. Consolidated text available at: <https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-14>

Hobbes, T. (1651) *Leviathan*, chapter 16, 'Of Persons, Authors, and Things Personated'. Text available at: <https://standardebooks.org/ebooks/thomas-hobbes/leviathan/text/chapter-16>

Prime Minister of Australia (2023) 'Final report of the Royal Commission into the Robodebt Scheme', media release, 7 July 2023. Available at: <https://www.pm.gov.au/media/final-report-royal-commission-robodebt-scheme>

Raz, J. (1986) *The Morality of Freedom*. Oxford: Clarendon Press.

Raz, J. (2006) 'The Problem of Authority: Revisiting the Service Conception', *Minnesota Law Review*, 90, pp.1003–1044. doi:10.24926/265535.919

Royal Commission into the Robodebt Scheme (2023) *Report*. Canberra: Commonwealth of Australia, 7 July 2023 (Commissioner Catherine Holmes AC SC). Available at: <https://robodebt.royalcommission.gov.au/system/files/2023-09/rrc-accessible-full-report.PDF>

Ofqual (2020) 'Statement from Roger Taylor, Chair, Ofqual', 17 August 2020. London: GOV.UK. Available at: <https://www.gov.uk/government/news/statement-from-roger-taylor-chair-ofqual>

Schools Week (2020) 'A-level results 2020: 8 key trends in England's data', 13 August 2020. Available at: <https://schoolsweek.co.uk/a-level-results-2020-8-key-trends-in-englands-data/>

Stanford HAI (2026) *The AI Index 2026 Annual Report*, chapter 3, 'Responsible AI'. Stanford, CA: Stanford Institute for Human-Centered Artificial Intelligence, April 2026. Available at: https://hai.stanford.edu/assets/files/ai_index_report_2026_chapter_3_responsible_ai.pdf

State v. Loomis, 2016 WI 68, 371 Wis. 2d 235, 881 N.W.2d 749 (Wisconsin Supreme Court, 13 July 2016). Available at: <https://www.courtlistener.com/opinion/4237997/state-v-eric-l-loomis/>

Sydney Morning Herald (2023) Clun, R. and Whitbourn, M., 'What you need to know about the robo-debt royal commission', 7 July 2023. Available at: <https://www.smh.com.au/politics/federal/what-you-need-to-know-about-the-robo-debt-royal-commission-20230707-p5dmkb.html>

The Guardian (2020) Hill, A. and Davies, C., 'A-level results day 2020 live: 39.1% of pupils' grades in England downgraded', 13 August 2020. Available at: <https://www.theguardian.com/education/live/2020/aug/13/a-level-results-day-2020-live-students-teachers-government-ucas-mock-exams-triple-lock-nick-gibb>

Tweede Kamer (2020) *Ongekend onrecht*, report of the Parliamentary Interrogation Committee on Childcare Benefits (chair: Chris van Dam), 17 December 2020. The Hague: Tweede Kamer der Staten-Generaal.

WAM (Emirates News Agency) (2025) 'UAE Cabinet, chaired by Mohammed bin Rashid, approves launch of first integrated regulatory intelligence ecosystem in UAE Government', 14 April 2025. Available at: <https://www.wam.ae/en/article/bj6abdw-uae-cabinet-chaired-mohammed-bin-rashid-approves>