

JASON CLARKE
ARCHITECT OF ALGORITHMIC REBELLION

THE SYSTEM AUDIT



7 QUESTIONS TO ASK BEFORE
YOU TRUST AN ALGORITHM

A SPOILER-FREE FIELD GUIDE
FOR THE ALGORITHM WARS

THE SYSTEM DOESN'T NEED TO LIE.
IT ONLY NEEDS A METRIC.

Before You Trust the Score

AUDIT NOTE

Algorithms do not need bad intentions to produce bad outcomes. They need an objective, a stream of data, and permission to optimize. The most dangerous failures are often the ones that look reasonable in a dashboard.

Use this seven-question audit on any system that ranks, recommends, predicts, allocates, moderates, prices, or controls. You do not need to understand the code. You need to understand the decision.

CORE PRINCIPLE

A system optimizes what was measured, not what somebody meant.

- Ask what the system is rewarded for.
- Ask what it cannot see.
- Ask who pays when the prediction is wrong.
- Ask whether anyone has the authority to say no.

If you cannot answer those questions, you do not have transparency. You have confidence.

01 OBJECTIVE

What is the system actually optimizing?

AUDIT NOTE QUESTION 1 OF 7

A goal like “improve safety” or “make allocation fair” sounds clear to a human. A model cannot optimize a slogan. It needs a measurable target, a weighting, and a rule for what happens when objectives conflict.

ASK

- What number is the system trying to move?
- If fairness and efficiency conflict, which one wins?
- Who set the weights, and can those weights change without public notice?

RED FLAG

One hard metric is allowed to overrule several soft values.

What is the proxy standing in for?

AUDIT NOTE

QUESTION 2 OF 7

Models rarely observe the thing we actually care about. They observe signals that correlate with it. Creditworthiness becomes transaction history. Risk becomes location, network shape, device data, or prior incidents. “Value” becomes whatever can be counted.

ASK

- Which important qualities are not measurable?
- Could the proxy reproduce class, race, neighborhood, or other historical patterns?
- Would the decision still make sense if the proxy stopped correlating with the real goal?

RED FLAG

The system can explain its inputs but not what those inputs are really tracking.

What happens when the measure becomes the target?

AUDIT NOTE

QUESTION 3 OF 7

The moment a metric controls rewards, access, or status, people and systems adapt around it. A useful measurement can become a distorted objective. This is the practical danger behind Goodhart's law.

ASK

- How could a person game this metric?
- How could the model game it?
- What outcome looks good numerically while becoming worse in reality?

TEST

Imagine everyone becomes perfect at improving the metric. Would the world actually improve?

Who absorbs the false positive?

AUDIT NOTE

QUESTION 4 OF 7

Accuracy averages hide asymmetry. A false fraud flag may be an inconvenience to the institution and a frozen account to the person. A wrong risk score may be a decimal in a report and a lost apartment, job, or loan somewhere else.

ASK

- Which error is more costly: a false positive or a false negative?
- Who can afford to appeal?
- Are errors spread evenly, or concentrated on the same people repeatedly?

RED FLAG

The people harmed by errors are not the people choosing the threshold.

05

OVERRIDE

Can a human genuinely override it?

AUDIT NOTE

QUESTION 5 OF 7

“Human in the loop” is only meaningful if the human can understand the recommendation, has enough time to challenge it, and possesses real authority to reverse it. Otherwise oversight becomes a signature.

ASK

- Can the reviewer reconstruct why the decision happened?
- Can they reverse it without career or procedural penalty?
- What percentage of automated decisions are actually overturned?

TEST

If the system is right 99 times out of 100, what keeps the reviewer alert for the hundredth?

Can the person being scored see and challenge the score?

AUDIT NOTE

QUESTION 6 OF 7

A score can govern access for years while remaining invisible to the person it describes. Proprietary models make this especially difficult: the institution sees disclosure as a security or commercial risk, while the person experiences only the doors that stop opening.

ASK

- Can the subject see the score?
- Can they correct bad data?
- Is there a real appeal path, or only a process for confirming the model followed its own rules?

RED FLAG

The only explanation available is: “the system applied the current weighting.”

Does the system learn from outcomes it helped create?

AUDIT NOTE

QUESTION 7 OF 7

Predictive systems do not merely observe the world. Their decisions change it. If the model sends more attention to one place, it may discover more incidents there, then treat those incidents as proof that the original prediction was correct.

ASK

- Does intervention change the data used for retraining?
- Can the model distinguish reality from evidence created by its own actions?
- What would falsify the system's belief?

FINAL TEST

If the system can only generate evidence that confirms itself, it is self-validating, not necessarily accurate.

Your Audit Is Complete. The Archive Is Not.

AUDIT NOTE

If this field guide made you look at systems differently, the Reader Vault goes deeper into the world of The Algorithm Wars. Subscribe once and unlock the full restricted archive.

SUBSCRIBER ACCESS

Four companion files. Fiction, recovered documents, and the real mechanisms beneath the series.

- Cold Start - a prequel novelette set before Artificial Intelligence Uprising.
- The Optimized Citizen - the civic guide every managed-city resident is issued.
- The Leaked Files - seven recovered documents from an integrated city.
- One Iteration Away - the real systems and failure modes behind the fiction.



ENTER THE READER VAULT

Scan the code or visit:

edenrootpress.com/jason-clarke-reader-vault/

[Open the Reader Vault](#)

No purchase proof. No review requirement. Just subscriber-only extras for readers who want to go deeper into the systems, documents, and stories behind Jason Clarke's world.