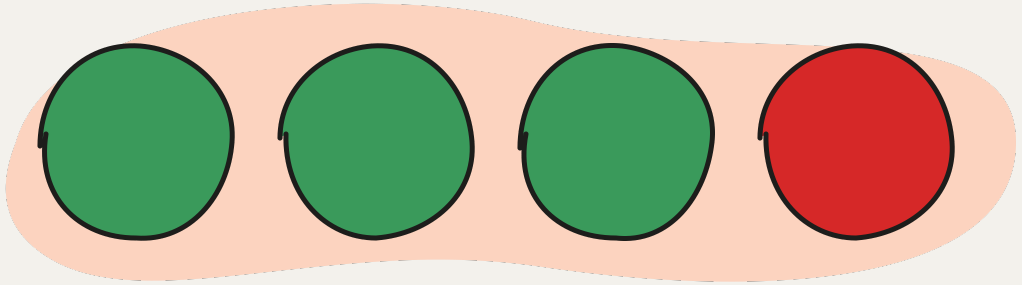
Green, green, green, red

Listening for trouble in government's biggest projects



Emily Lawson
COO, Newton

Mallory Durran
Mission Director and CTO, Nesta

The key was to find the crib

1942 Bletchley

QXRMBLTZKPAVWETTERJNDUHOYGFCSIEB

LZCNWQAPRMTDWETTEROBKVYHSXGUFIJA

HVSOJKDXEQNLWETTERMTRZPACYWGIBUF

PDGYUAKWZXRCWETTERLNSHQBOMEJVTIF

YNEOTSBHCLGUWETTERXAQPZDVKJMWRIR

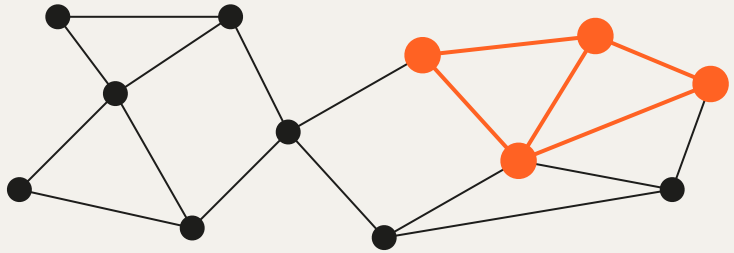
KAFJXIRPNVOZWETTERUYDCLEGTSWBHMQ

Source: The National Museum of Computing

The model was right. Nobody would cross the gap

2022 Whitehall

Dark-web network mapping, modern slavery



Institutional trust



Machine insight

99%

hit rate on historic data

Acting on machine evidence, not human review, was judged too risky

Operational action

£0

invested

Green, green, green, red

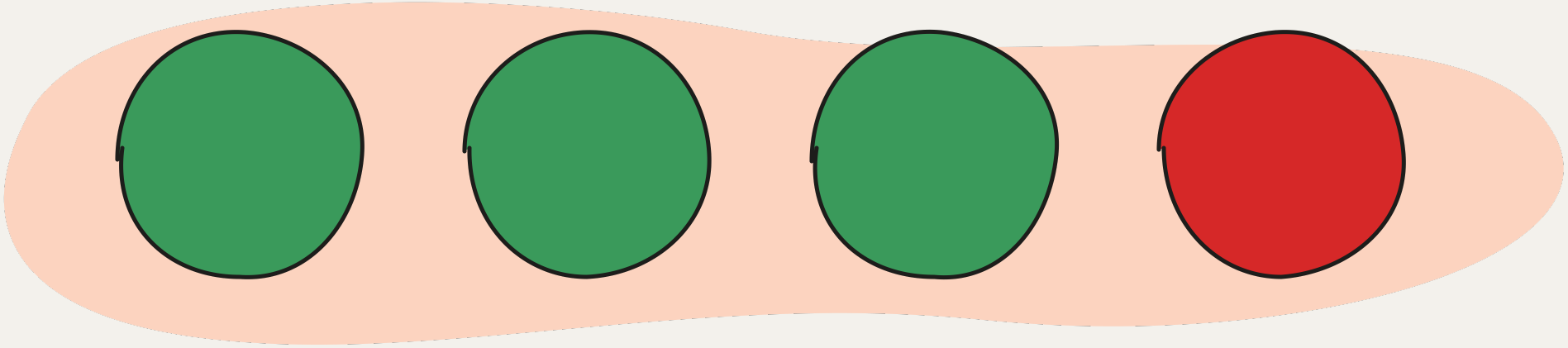
235 cross-government projects reported into the Cabinet Office each month

Inconsistently

Cabinet Office team reported the data

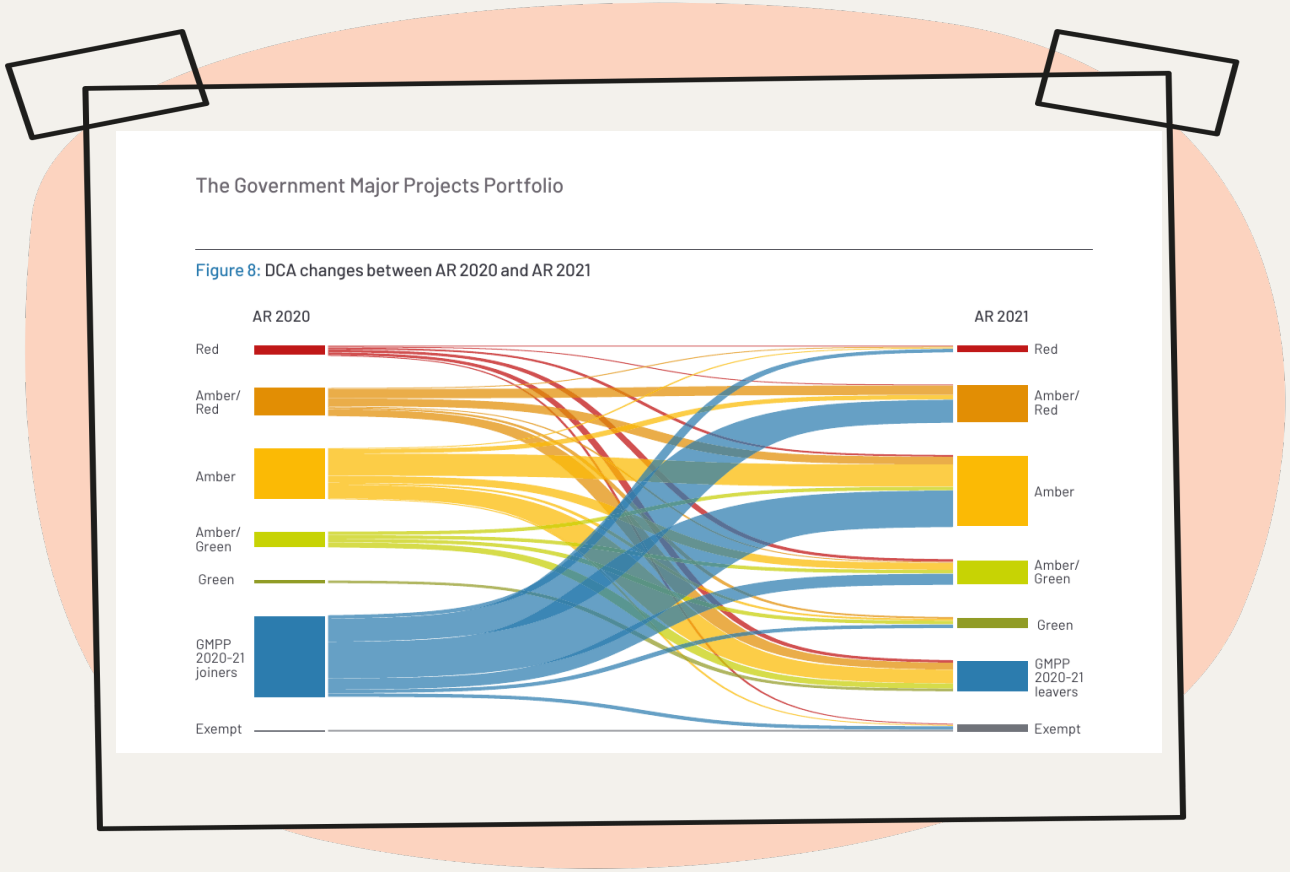
The reports were boxed

Intervention was limited



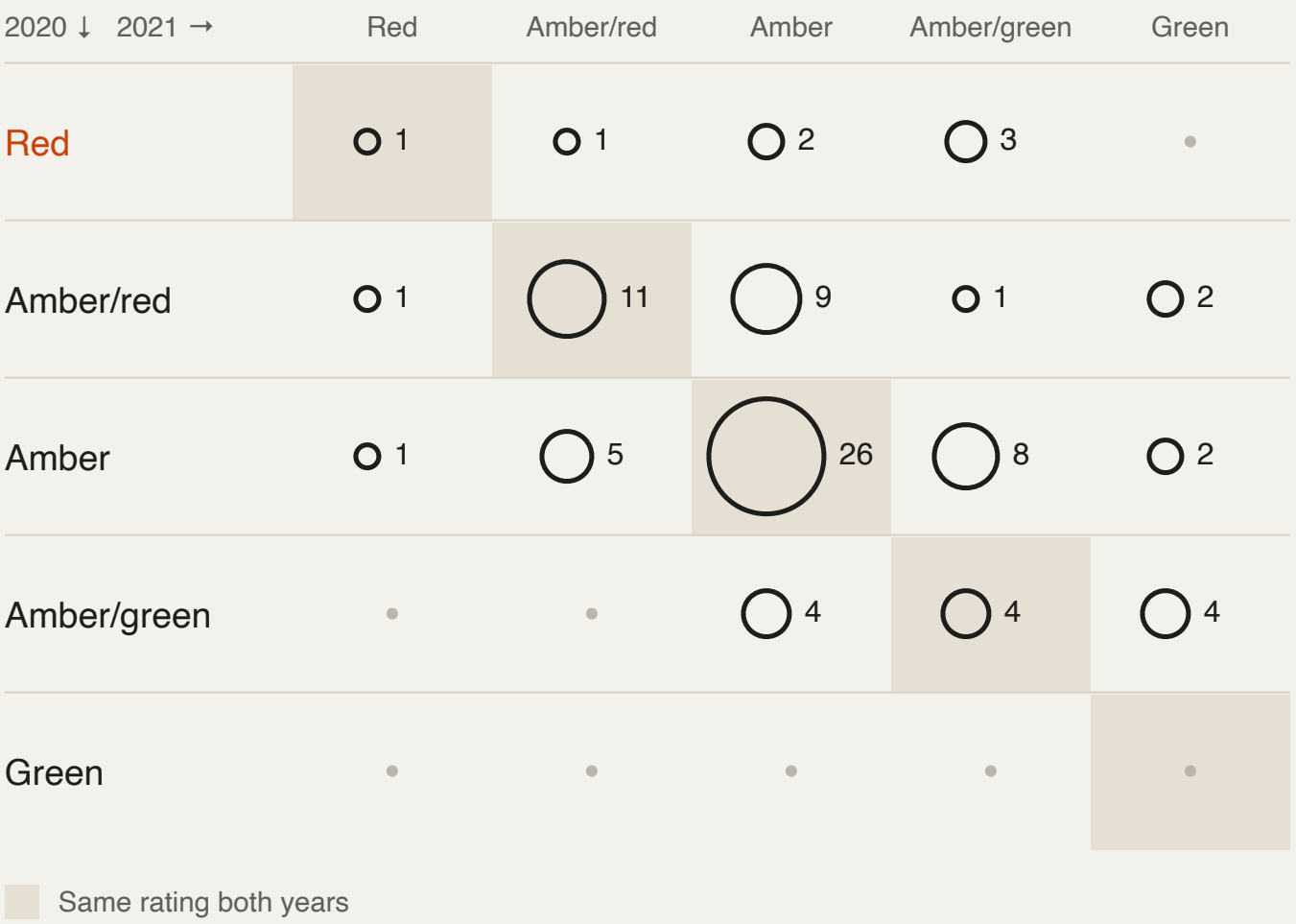
Source: GSMO data published on gov.uk

The IPA's view was that improvement was happening



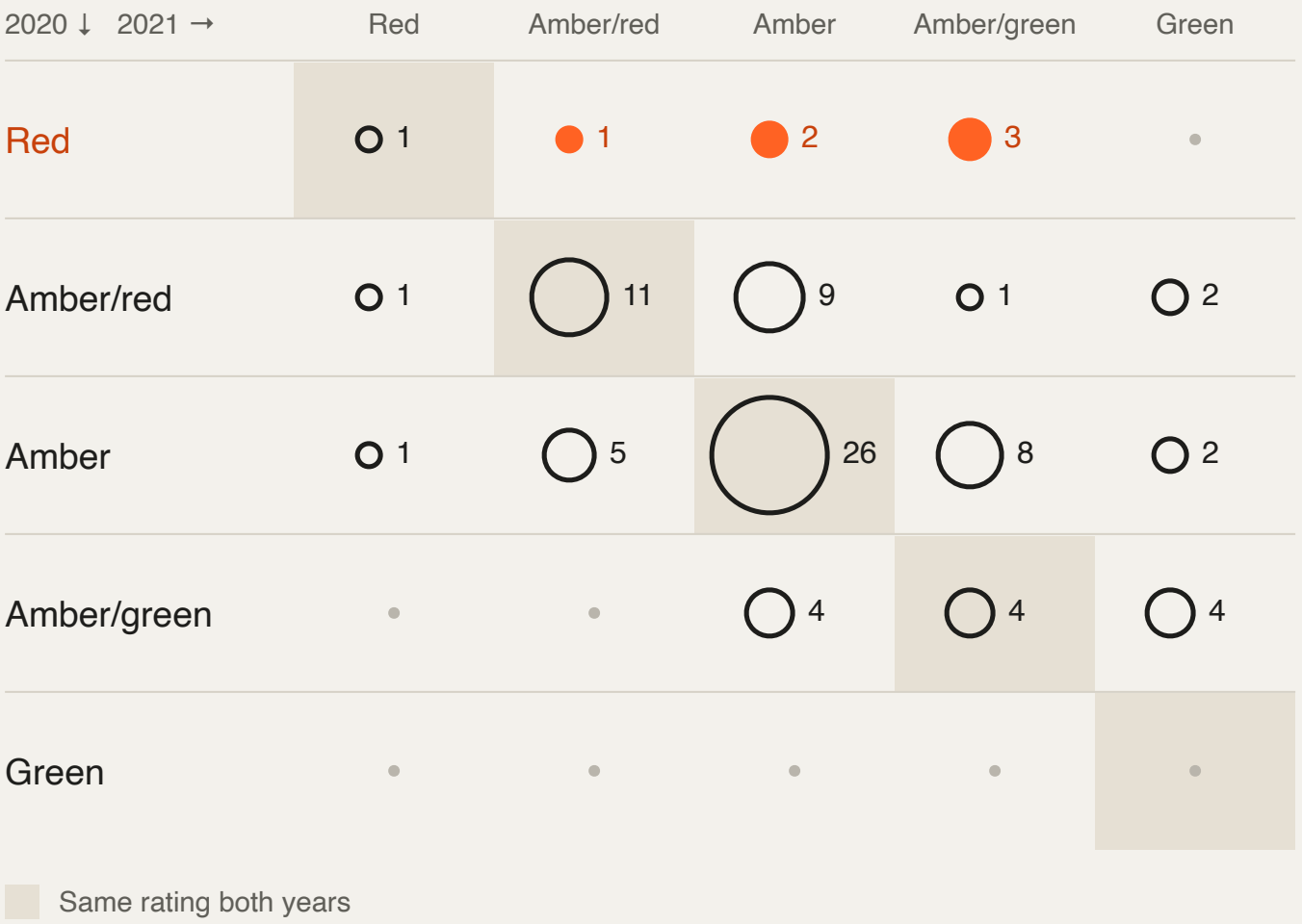
Source: Infrastructure and Projects Authority Annual Report 2021

A red rating told you little about next year



Source: IPA annual report 2021 Figure 8, redrawn for projects rated in both years

A red rating told you little about next year



Source: IPA annual report 2021 Figure 8, redrawn for projects rated in both years

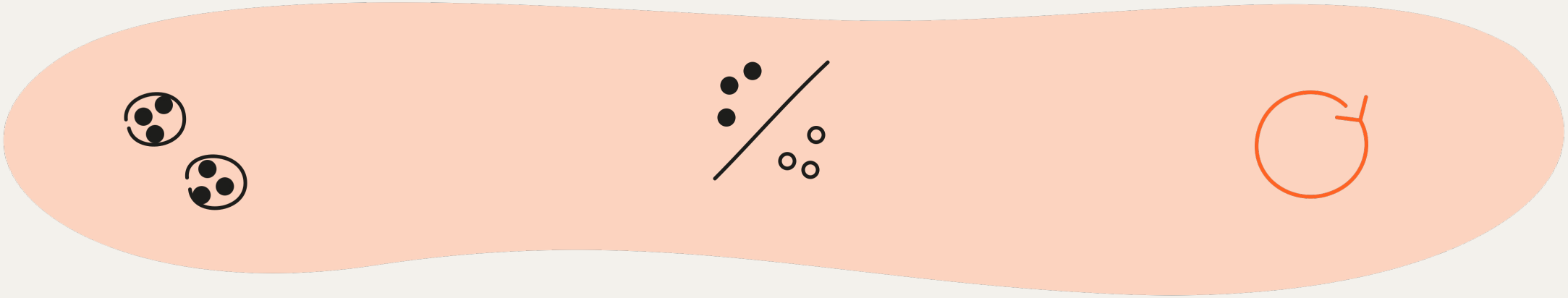
Our crib: the director’s own words

What was their ‘Wetter’?

What combination of words or phrases were the ‘tell’ that said, this project is about to go wrong?



We taught 2022 NLP what ‘nervous’ sounded like



Find patterns nobody had labelled

(Unsupervised)

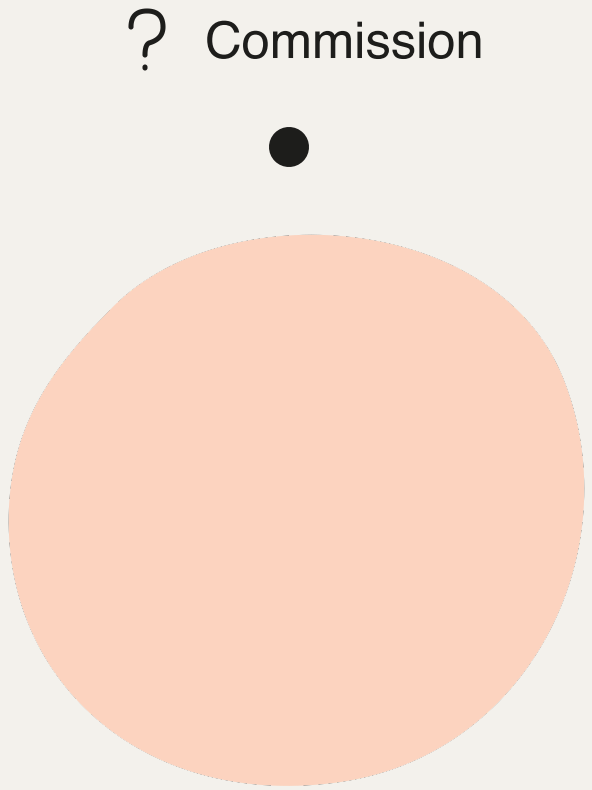
Learn from what the team had flagged

(Supervised)

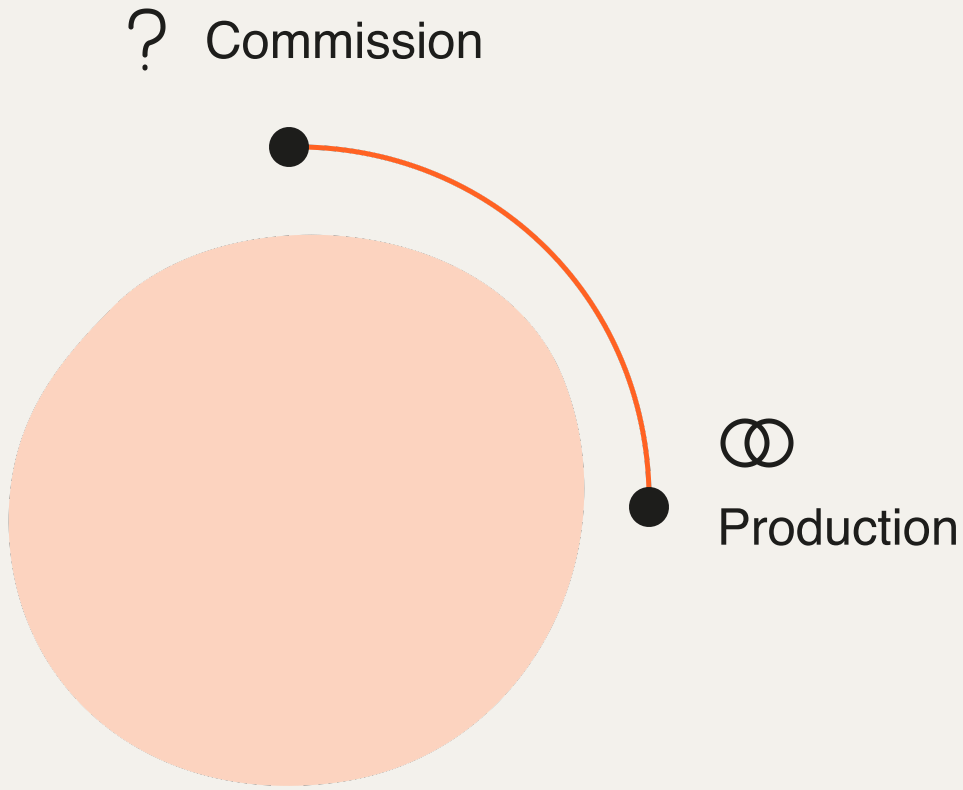
Check every signal by hand, feed it back

(Manual feedback loop)

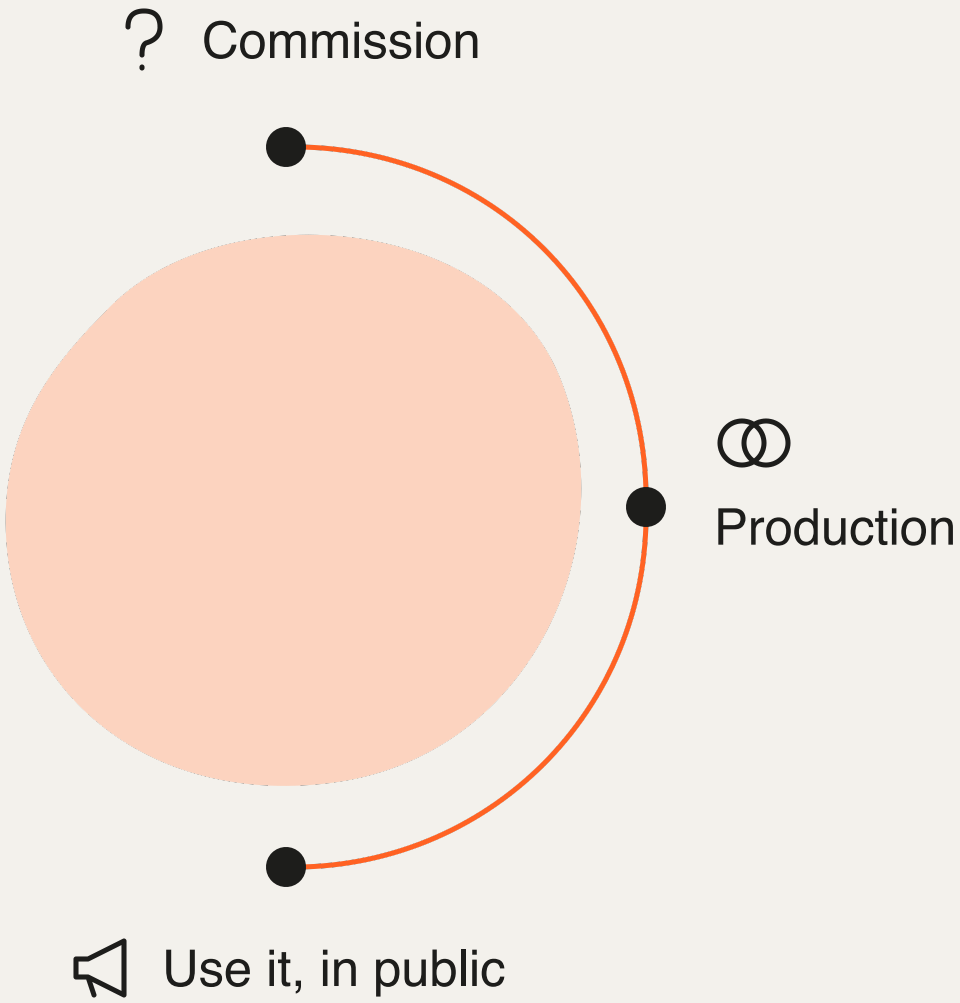
We started with the right question



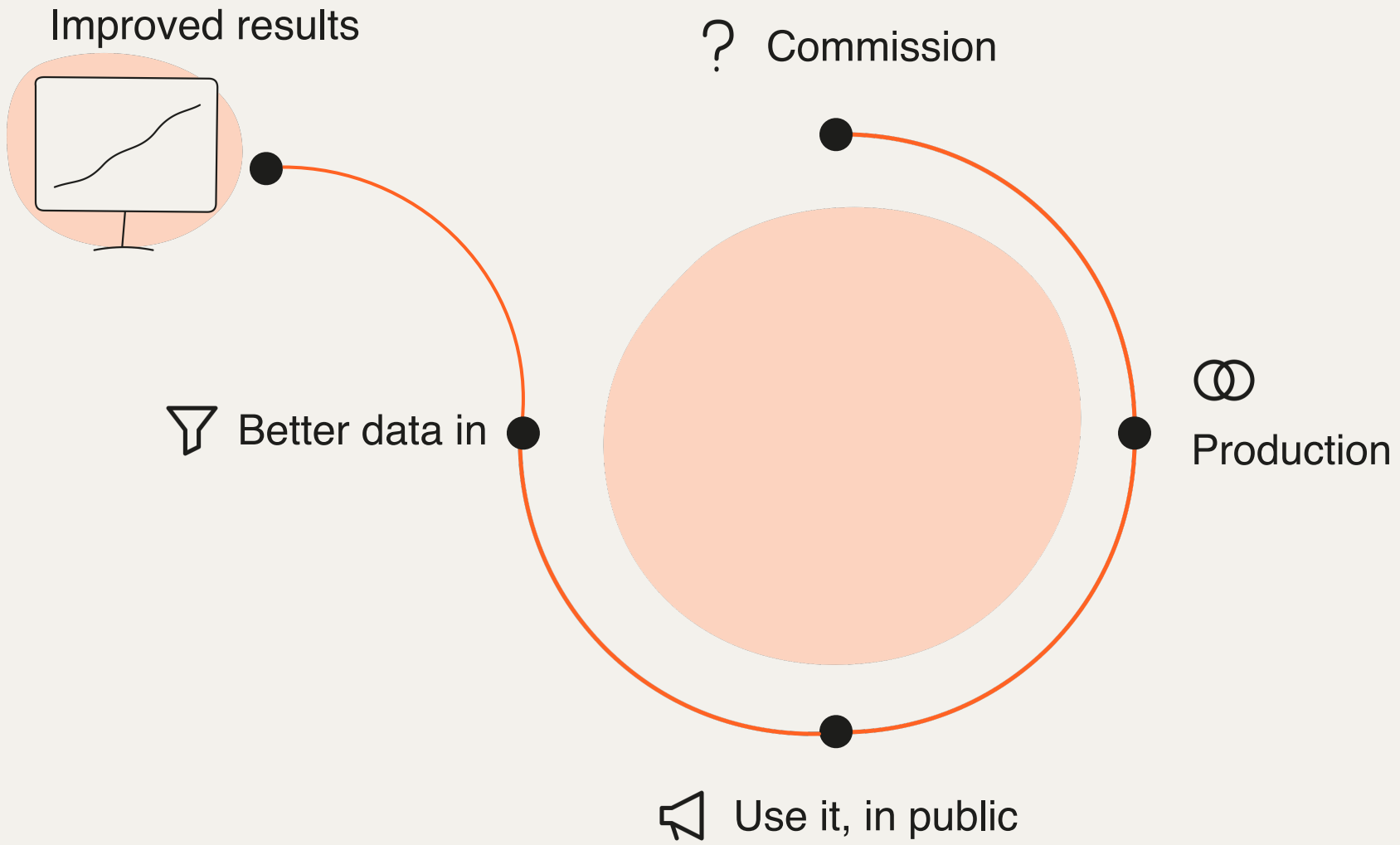
We trusted the team and the tech



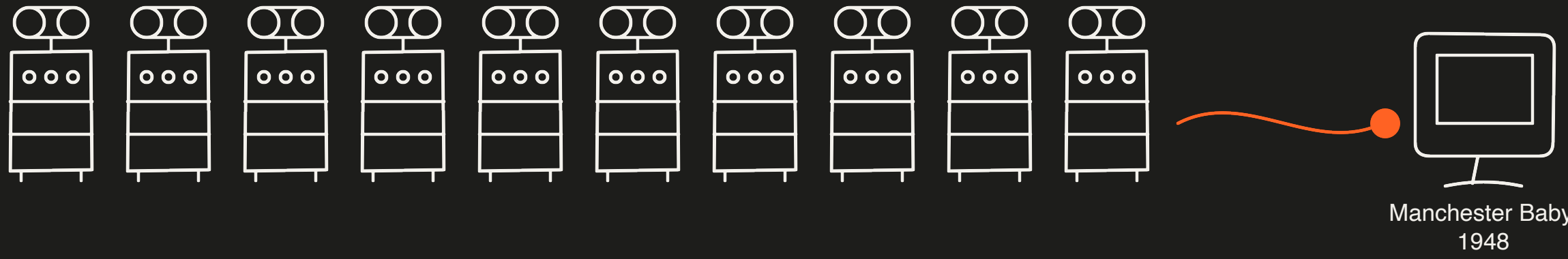
We used the data to intervene



Candour improved data and results

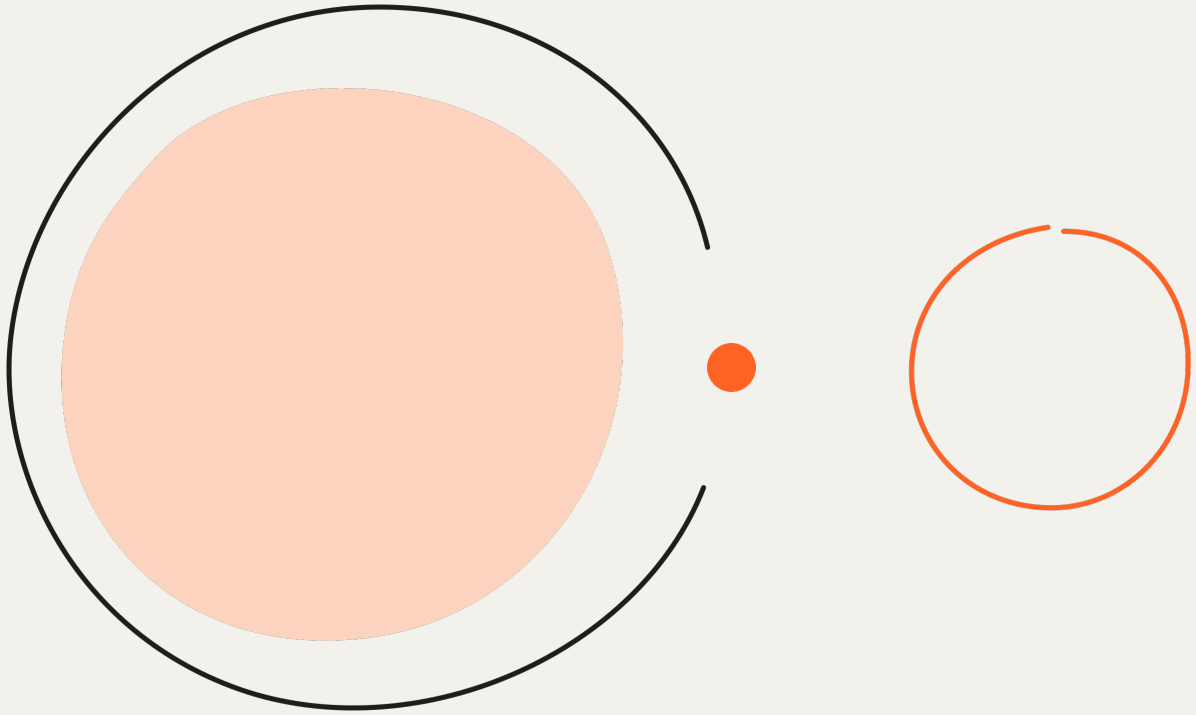


8 of 10 Colossus machines were dismantled after the war



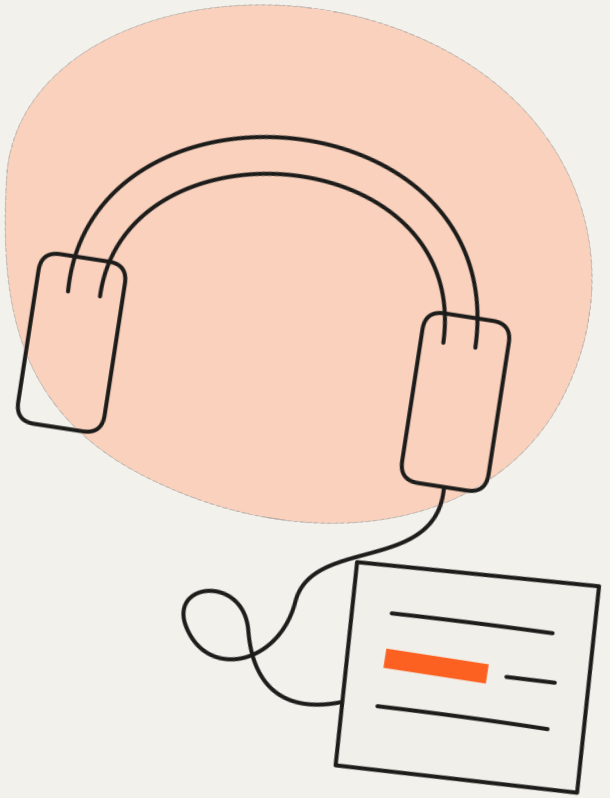
Source: GCHQ, Colossus at 80 (2024)

Our approach was dismantled…but our people and better tools live on

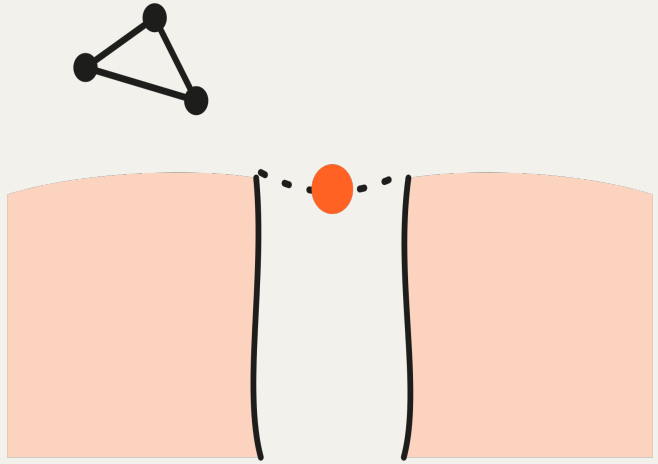


What did we learn?

Leaders don't need to do the coding….	…they need to know how good software is structured so that they know what to ask for
You need to know what you are listening for…	…so you can commission a question that needs answering instead of an overly-specific tool
We need to make it safe to be heard…	…and reward team candour with leaders speaking up in public
And build it to outlast us…	…invest in the plumbing – data, process, people



What help do we need?



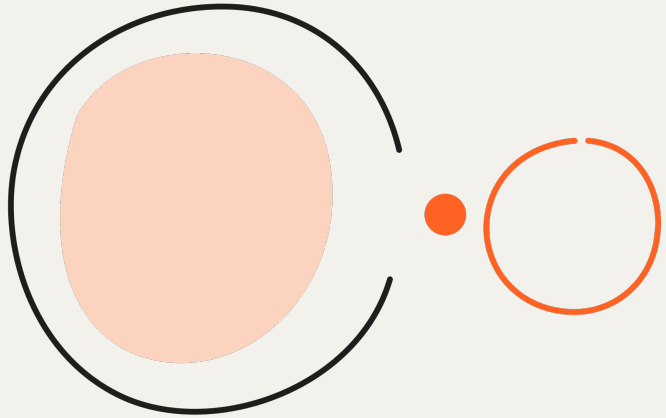
Stories of crossing the gap

Are you acting on machine evidence before human confirmation, and, if so, what makes that feel safe?



Robust early-warning tools

Particularly in government, if you're building something like this we'd like to compare notes



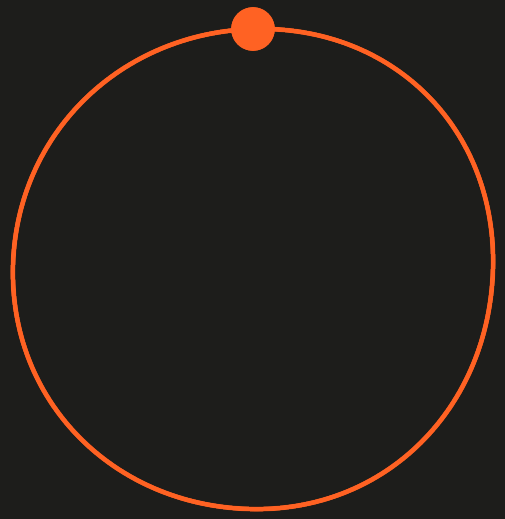
Life beyond the sponsor

What have you done to keep an approach live through a change of leadership

The machine is the easy part

The value is in knowing what to listen for, and being willing to act on what you hear.





Emily Lawson
COO, Newton

Mallory Durran
Mission Director and CTO, Nesta

Newton

nesta