

Garbage in, garbage out?

Measuring error rates and data quality to get ready for AI

#SDinGOV

Caroline Jarrett
@cjforms.bsky.social

Sharon Dale
@pixlz.com

We miss Vicky Teinaki

Events (a weeknote, starting 16 September 2024)

5 min read · Sep 23, 2024



Vicky Teinaki

Follow

Martin Jordan announced error rate as a metric



Martin Jordan

@martinjordan.com

Non-fun fact:

There are error rates for various government services and related forms far beyond 90%.

Government does not ask people to submit the necessary information and documents and then cannot make a decision.

Error rate is now a central metric tracked in various of our work streams.

August 13, 2024 at 5:35 PM  Everybody can reply

I got interested in better forms because of errors

The image shows three examples of form input errors for an 'Amount' field. Each example consists of the label 'Amount' followed by a series of input boxes. In the first example, the first four boxes are empty, the fifth contains a minus sign, and the next two contain '1' and '0'. In the second example, the boxes contain the digits '1', '2', '5', '7', '8', '1', '6', and '3' in sequence. In the third example, the first three boxes are empty, and the next four contain the letters 'N', 'O', 'N', and 'E' in sequence.

Amount - 1 0

Amount 1 2 5 7 8 1 6 3

Amount N O N E

For “AI”, data quality matters



Dare Obasanjo

@carnage4life.bsky.social

Harsh truths of machine learning

<https://bsky.app/profile/carnage4life.bsky.social/post/3liouuj3phs2z>

The “garbage in, garbage out” principle also applies to

- Large language models (LLMs)
- Image recognition
- And everything else really



Let's meet

At your table, split yourselves up into pairs (or threes)

Compare:

- What brought you to this session today?
- What do you hope to gain and to contribute?

2 minutes



Collect reflections as we go

- Current practice Share your best practice, tip, or suggestion for others to try
- Next step Anything that you yourself plan to find out or to do differently
- Something else Any other observations or thoughts you'd like to share



bit.ly/gigo25



Agenda

Think about fudging forms

Share what we know about our error rates (if anything)

Try using “Six ways to think about errors”

Meet the data quality framework

Think about what we might do differently

I've observed that people “fudge” through forms

To fudge a form:
answer one or more questions
in a way that's not entirely truthful



I admit that I fudged my bus pass photo

The application form said that I had to submit a passport photo



I chose one that doesn't comply with passport rules



If you've fudged a form, share your story

Write a sticky note with an example of fudging a form

Share with your neighbour

2 minutes



Takeaway

There are many ways in which the data we collect isn't entirely accurate

Agenda

~~Think about fudging forms~~

Share what we know about our error rates (if anything)

Try using “Six ways to think about errors”

Meet the data quality framework

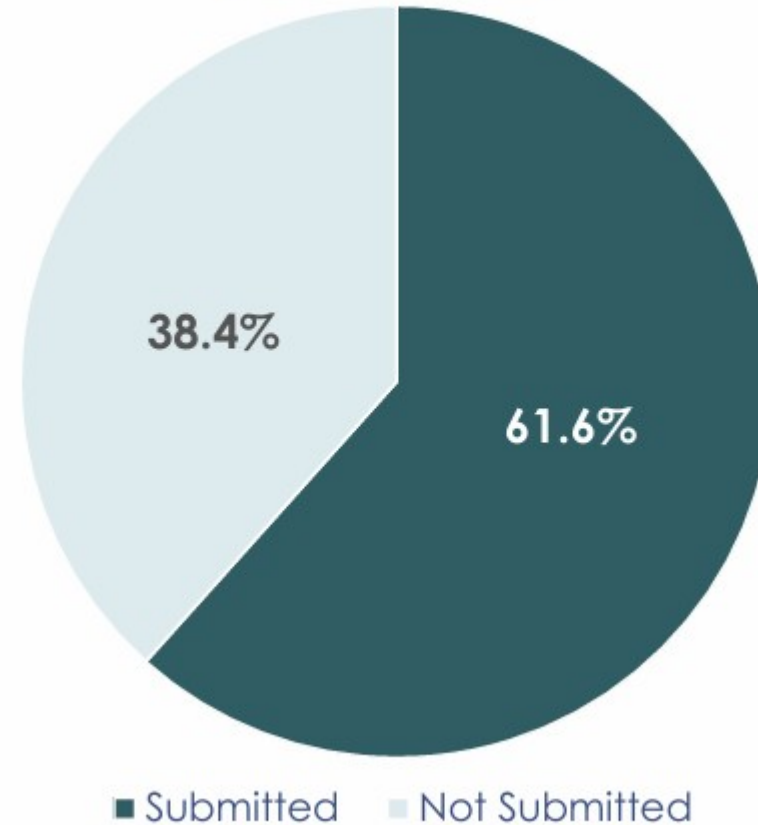
Think about what we might do differently

Share what we know about
our error rates (if anything)

Some people don't get to the end of their application

People on low incomes or no income in California can apply for government money to buy fresh food (CalFresh).

PROPORTION OF APPLICATIONS SUBMITTED
(NOV – DEC)



<https://www.calsaws.org/wp-content/uploads/2025/02/BenefitsCal-Usage-Metrics-Report-Nov-Dec-2024.pdf>

Some errors hit the headlines

[Disabled Ipswich woman told to repay £5,000 of universal credit - BBC News](#)



The result matters more than the journey

People rarely fill in forms for fun

We're aiming to achieve some result



Let's learn from our experiences with errors

Write on a sticky note:

- a brief description of a service that your team is responsible for
- an example of something that you consider as an error in that service

Please write one sticky note for each error

2 minutes



Now compare your stickies with your neighbour

Do you have similar or different errors?

Maybe write more sticky notes for any extra errors that you now think about

3 minutes



Agenda

~~Think about fudging forms~~

~~Share what we know about our error rates (if anything)~~

Try using “Six ways to think about errors”

Meet the data quality framework

Think about what we might do differently

I've been thinking about ways to think about errors

Thanks to the team at HMRC who inspired this

Errors	Examples
Problems along the way	Tried to buy 2Kg carrots, ordered 2 carrots
Wrong result	Ordered pork pie, got lettuce
Unnecessary action	No delivery arrives, called to find out why
Delayed-impact problem	Credit card OK when placed order, failed later
Non-uptake	Decide to use a different supermarket
Over-uptake	Accidentally placed order twice
Technology issue	Supermarket goes offline due to cyber-attack
Something else	?

Non-uptake and over-uptake can get complex

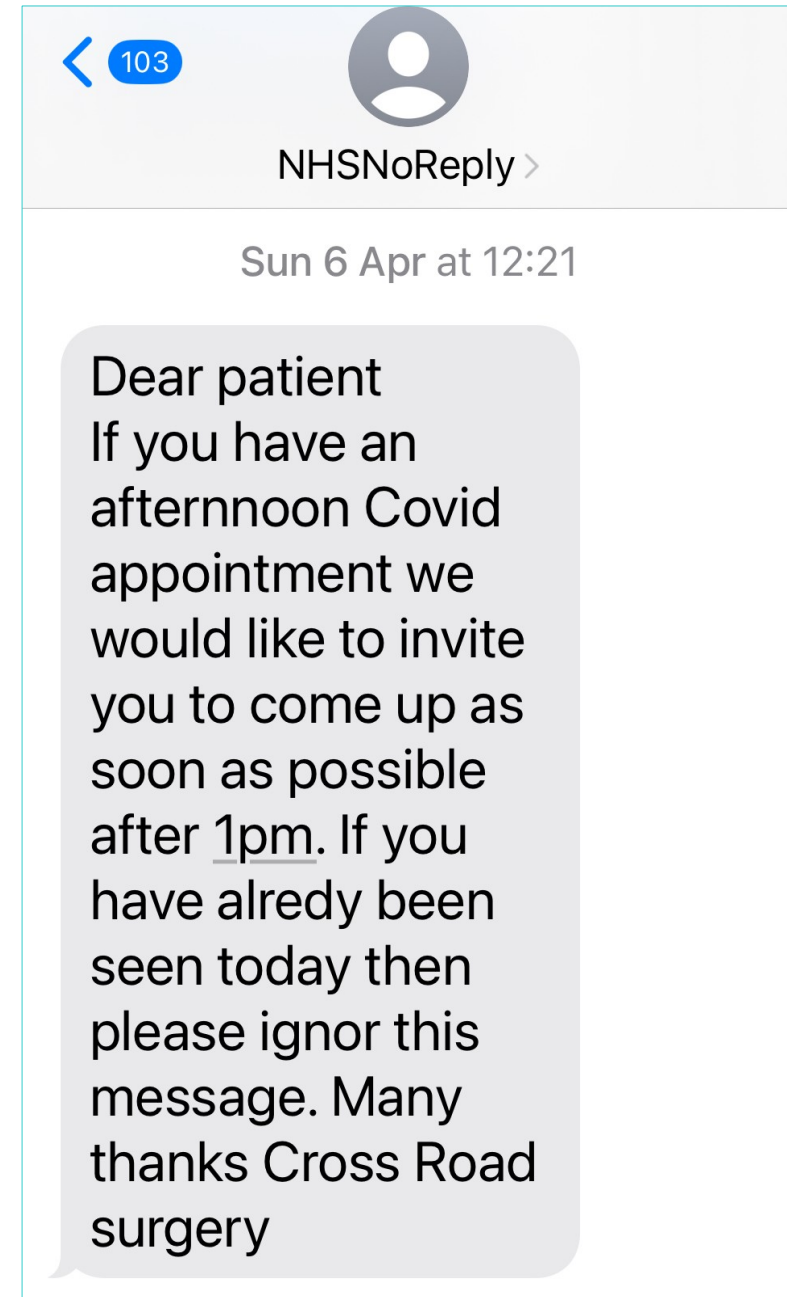
Not recording a vaccination

We did some timeboxed design exploration this week about how we could better let users record when someone is not given a vaccination.

Whenever we come to this topic it opens up a bit of an existential can of worms, as our service is called [Record a vaccination](#), and the scope of recording non-vaccinations could be quite wide.

Non-vaccinations covers anything including non-attendance at a booked appointment, a walk-in patient who has to be turned away for some reason, or someone changing their mind at the last minute.

Non-uptake and over-uptake happens in real life



Did you find any examples of these errors?

Try allocating your errors to the chart

- Problems along the way
- Wrong result
- Unnecessary action
- Delayed-impact problem
- Non-uptake / over-uptake
- Technology problem
- Something else?

3 minutes



What about error rates?

And now, what about error rates?

If you know the error rate for any of the errors, add a star
(plus, optional: add a sticky note with a comment)

1 minute



UK government services must publish completion rates

 **GOV.UK**

Menu

Search

Beta [Contact the Service Manual team](#) if you have feedback, questions or suggestions.

[Service manual](#) > [Measuring success](#)

Measuring success

Measuring completion rate

[Give feedback about this page](#)

From: [Performance analysis community](#)

Contents

- [Calculating completion rate](#)
- [Configure your service to measure completion rate](#)
- [Counting users who are ineligible for a service](#)

Your service's completion rate is the number of digital transactions that your users complete as a percentage of all digital transactions that your users start.

This includes transactions where the user receives support from someone to use the digital service (called '[assisted digital support](#)').

Completion rate is a relatively simple calculation

$$\text{Completion rate} = \frac{\text{Number of people who complete}}{\text{Number of people who start}}$$

Error rate might also be simple

$$\text{Error rate} = \frac{\text{Number of errors}}{\text{Number of people who start}}$$

Other error rate calculations are available

$$\text{Error rate} = \frac{\text{Number of attempts with an error}}{\text{Number of people who start complete are eligible other?}}$$

Example: there are at least 3 definitions of turnout

In an election, a failure to vote is an error

$$\text{Turnout} = \frac{\text{Number of votes counted}}{\text{Number of people who}}$$

1. Send in postal votes or go in person
2. Are on the electoral roll
3. Are in the voting-eligible population

In the error rate, what do we divide by?

For the service you are thinking about, what would you divide by?

Discuss as a table

3 minutes



How can we solve the missing stars?

Some of the errors do not have a star for a known error rate

Any ideas for changing that, maybe by doing some measuring?

3 minutes



Agenda

~~Think about fudging forms~~

~~Share what we know about our error rates (if anything)~~

~~Try using “Six ways to think about errors”~~

Meet the data quality framework

Think about what we might do differently

Previous takeaway

There are many ways in which
the data we collect isn't
entirely accurate

Data quality can deteriorate or change all by itself

People change

- Move house
- Change jobs
- Lose their phones
- And many others

Organisations change

- Add a new product
- Delete an old one
- Are affected by energy costs
- And many others



Missing data can have serious effects

The Home Office destroyed the documents that migrants later needed to prove their right to remain



'Windrush' migrants facing deportation threat

Michael Braithwaite was let go after his employer ruled him an illegal immigrant.

Like many other members of the "Windrush generation" of Caribbean migrants who moved to the UK before the 1971 Immigration Act, Michael had the right to remain but no documents to prove it.

'Windrush generation' deportation threat

UK · 13 April 2018

['Windrush' migrants facing deportation threat - BBC News](#)

Previous takeaway

There are many ways in which
the data we collect isn't
entirely accurate

Reframed

There are many ways in which
the data we use isn't
entirely accurate

What is the longer-term life of our data?

Are there data quality issues that might arise over time?

Write one sticky note for each data quality issue.

Compare with your neighbour.

3 minutes



There is a UK Government Data Quality Framework



I asked folks at another workshop to have a look



It looks a bit like WCAG – I think there are probably useful ideas in it, but right now I can't see how to use it myself

Takeaway

Data quality is important
but not easy

Agenda

~~Think about fudging forms~~

~~Share what we know about our error rates (if anything)~~

~~Try using “Six ways to think about errors”~~

~~Meet the data quality framework~~

Think about what we might do differently

Reflect on what we've shared – now with voting

- Current practice Share your best practice, tip, or suggestion for others to try
- Next step Anything that you yourself plan to find out or to do differently
- Something else Any other observations or thoughts you'd like to share



bit.ly/gigo25

5 minutes



Takeaway

AI may help to start useful conversations about error rates and data quality

I'm starting to collect some resources

General advice on measurement

Steve Messer [Metrics, measures and indicators: a few things to bear in mind](#)

Official / government guidance

ONS [Quality in official statistics - Office for National Statistics](#)

Headline stories

The Guardian [Windrush victims could have compensation reconsidered](#)

Disability News [DWP helped cause mental distress of claimant who took her own life](#)

Academic research

The King's Fund [Lost In The System: The Need For Better NHS Admin](#)

Keep up the conversation: Caroline Jarrett

BlueSky @cjforms.bsky.social

caroline.jarrett@effortmark.co.uk

www.effortmark.co.uk

