Garbage in, garbage out?

Why measure and how to measure error rates as we head into AI

UX Connect, Aarhus, June 2025

Caroline Jarrett

@cjforms.bsky.social

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

https://bsky.app/profile/martinjordan.com/post/3kzmgtkdcuz2k

I got interested in better forms because of errors

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0   Understanding the costs of data capture: paper, automatic and with the internet - Effortmark

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

Write a sticky note for the work that you do

Write another sticky note for what brought you to that work

Compare your answers with your neighbour

I'll be collecting all the sticky notes later, so please keep them

2 minutes

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

FUDGE | definition in the Cambridge Learner's Dictionary

I admit that I fudged my bus pass photo

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

I chose one that doesn't comply with passport rules

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Takeaway

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Agenda

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Some errors hit the headlines

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

B B C NEWS

Home   UK   World   Business   Politics   Tech   Science   Health   Family & Education

Entertainment & Arts   Stories

Disabled woman told to repay £5,000 of benefits

YVONNE BUCKLEY

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)

38.4%

61.6%

■ Submitted   ■ Not Submitted

https://www.calsaws.org/wp-content/uploads/2025/02/BenefitsCal-Usage-Metrics-Report-Nov-Dec-2024.pdf
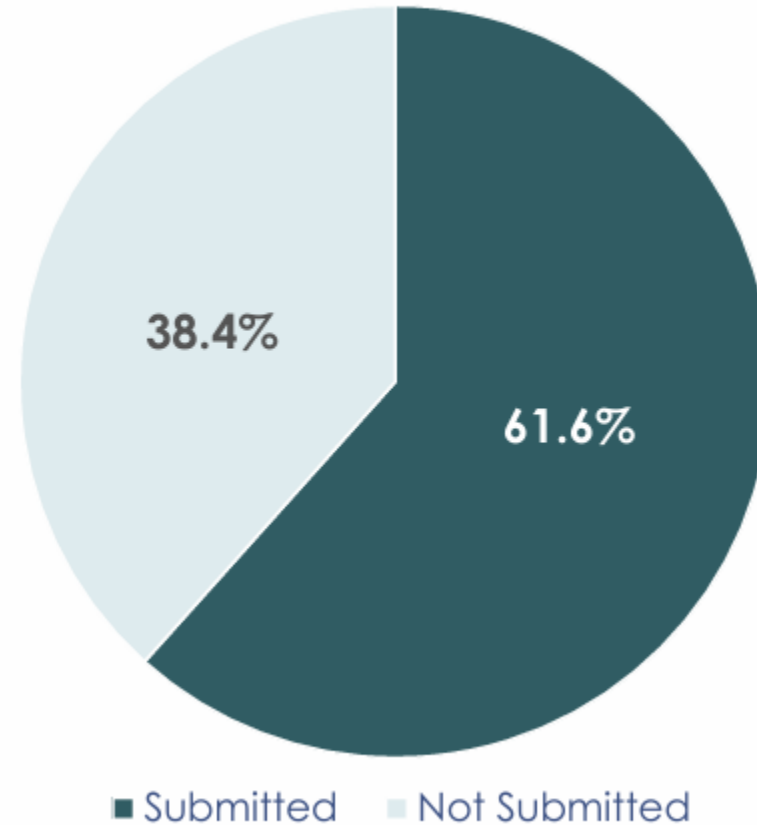
The result matters more than the journey

People rarely fill in forms for fun

We're aiming to achieve some result

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Photo by Marcos Ramírez on Unsplash

Let's learn from our experiences and ideas about errors

Write on a sticky note:
- a brief description of a service or product related to your work
- what you consider as an error in that service or product
- whether you know or don't know the error rates

If you have different types of error, please write one sticky note for each type

Then compare your note(s) with your neighbour

5 minutes

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Agenda

~~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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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	?

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0   https://bit.ly/5wayserrors

For example, "non-uptake" and "over-uptake" gets complex quickly

Week 39: OK Arrrr - Frankie Roberto

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.

We have heard some valid reasons for recording this though, including to schedule follow-up actions, to opt patients out from vaccination invitations, or for analysis.

As with many things in the NHS, the question becomes more about where is the best place to record the information, and what happens with the data?

Hopefully our design sketches can at least help to prompt some of this wider conversation.

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Did you find any examples of these errors?

Add a sticky note that says whether
the errors that you noted
are in any of these categories:

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

3 minutes

https://bit.ly/5wayserrors

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

What about error rates?

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

UK government services must publish completion rates

GOV.UK ⌄ Menu 🔍

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').

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Measuring completion rate - Service Manual - GOV.UK

Completion rate is a relatively simple calculation

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Error rate might also be simple

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

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

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0   Thanks to Whitney Quesenbery for this example

In the error rate, what do we divide by?

For the service or product you are thinking about, what do you divide by?

If there's more than one sensible thing to measure or divide by, write extra sticky notes

3 minutes

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Agenda

~~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

Data quality can deteriorate because of errors

- **Inefficient administrative actions.** The way the process is managed across administrative and legal staff resulted in multiple errors, duplication or unnecessary effort, and difficulty in making simple fixes. Our analysis shows that around 32% of cases had at least one hearing cancelled before the hearing took place (vacated), which then had to be rearranged.

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

National Audit Office: Improving family court services for children

But some problems occur purely because of delay

15   **The government, via the FJB, does not have an overall assessment of the main drivers of delays or the capacity required to manage the system efficiently and reduce delay.** MoJ, DfE and others have carried out several reviews of different sizes and scope to identify causes of delay in family justice, identifying more than 25 different contributing factors. There is a consensus among FJB members that the system needs to focus on delays, but most of the reviews could not quantify the scale and impact of each issue on overall performance, or the resources required to deal with these causes efficiently, due to data limitations described above. The causes of delays can vary significantly by local area. For example, there has been an acute shortage of judges in London, unlike some other areas. London also has a low number of cases resolved (disposed) per available sitting day (0.6), while Essex & Suffolk, which also has long delays, disposes 1.0 case per sitting day (paragraphs 2.7 to 2.8, 3.2 to 3.4, 3.12, Figure 6 and Figure 10).

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

National Audit Office: Improving family court services for children

Delay adds to errors in Denmark, too

The Agency of Family Law and the family courts are not ensuring that processing times across the system of family law are as short as possible

In 2023, the average processing time for cases that originate in the Agency of Family Law and are settled in family court was more than 12 months and thus 4 weeks longer than in 2020. This development is caused by increased processing times in the family courts, including increased waiting time for the first court session. Due to the longer processing time, the family courts increasingly re-examine cases and conduct new interviews with the children, for instance. This practice does not support the efficiency of the system of family law and entails a risk of straining the children.

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Rigsrevisionen's report on case processing in the system of family law 2024

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

Photo by Michal Balog on Unsplash

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

There is a UK Government Data Quality Framework

Guidance

The Government Data Quality Framework

Published 3 December 2020

Contents

Foreword

Acknowledgments

Why do we need a data quality framework?

The strategic context

How to use the data

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

The Government Data Quality Framework - GOV.UK

I asked folks at another workshop to have a look

Guidance

The Government Data Quality Framework

Published 3 December 2020

Contents

Foreword

Acknowledgments

Why do we need a data quality framework?

The strategic context

How to use the data

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

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

The Government Data Quality Framework - GOV.UK

Takeaway

Data quality is important but not easy

Agenda

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

Caroline Jarrett   Effortmark.co.uk @cjforms (CC) BY SA-4.0

Reflect on what we've shared - EasyRetro

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

5 minutes

The EasyRetro we used is now closed. Contact Caroline for feedback.

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

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

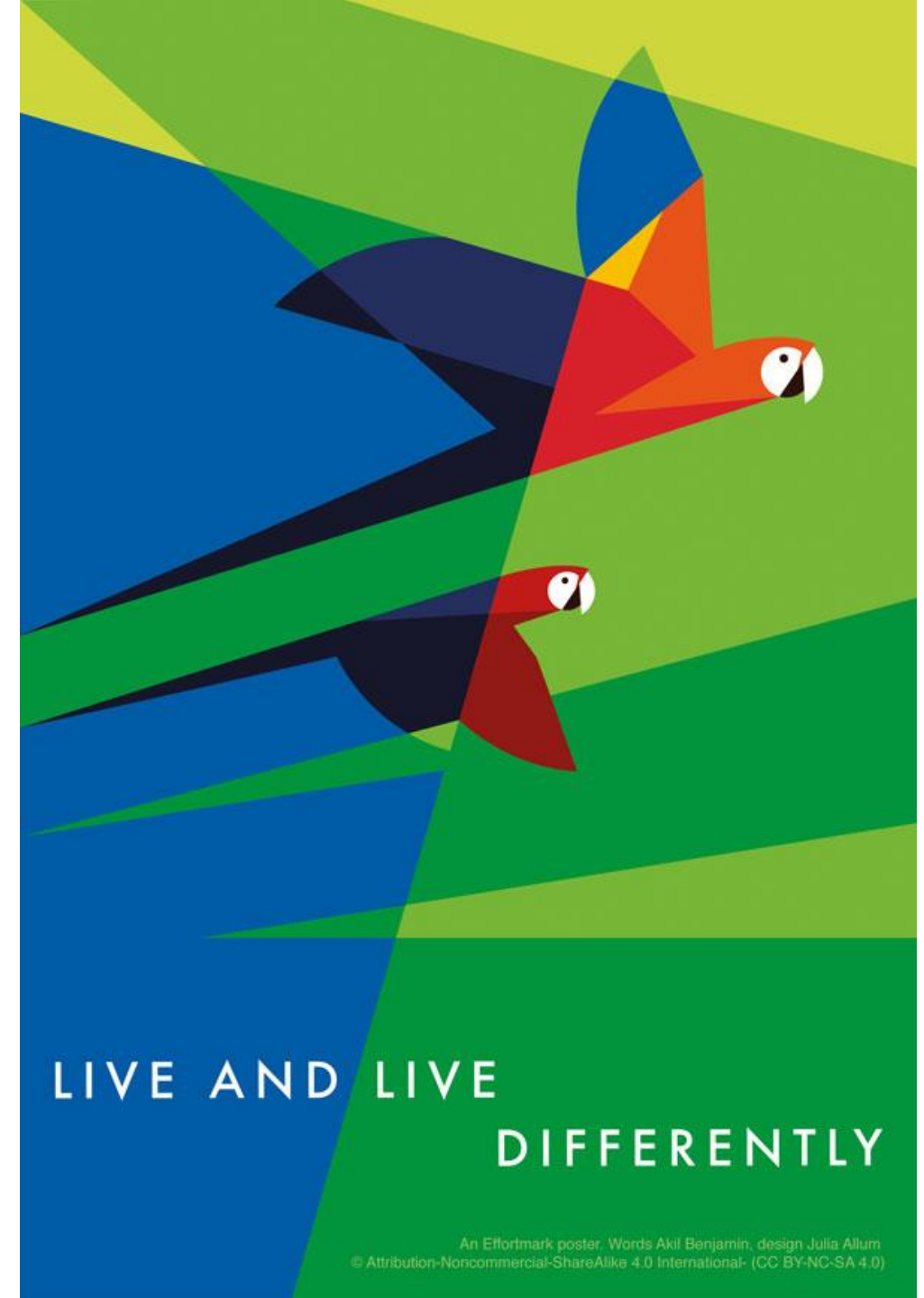Disability News   DWP helped cause mental distress of claimant who took her own life

Academic research

The King's Fund   Lost In The System: The Need For Better NHS Admin

You need to see me in person for a sticker

There is a poster for each sticker on my website

[Effortmark posters - Effortmark](#)

Caroline Jarrett Effortmark.co.uk @cjforms (CC) BY SA-4.0

LIVE AND LIVE
DIFFERENTLY

An Effortmark poster. Words Akil Benjamin, design Julia Allum
© Attribution-Noncommercial-ShareAlike 4.0 International- (CC BY-NC-SA 4.0)

Keep up the conversation: Caroline Jarrett

BlueSky @cjforms.bsky.social

caroline.jarrett@effortmark.co.uk

www.effortmark.co.uk