A read-aloud for kids, parents and teachers

Read It Together

Sure Is Not the Same as Right

Why the machine sounds certain when it is wrong

Last time we learned the secret. The machine picks the word that usually comes next.

Today a puzzle. If it only guesses, why does it sound so sure? And what happens when the usual word is the wrong one?

Forty minutes. A grown-up, a child or a class, and one machine at the front.

Step 1 · The starter · 5 min

What is the capital of Australia?

Ask it out loud and take the answers. Most rooms say "Sydney!" quickly and happily.

It is Canberra. Nobody was lying. Nobody was careless. Sydney just felt right.

Here is the point, before any machine appears. Being sure and being right are two different things, and from the inside they feel exactly the same.

Do this together

If someone did say Canberra, ask the room how sure they were of Sydney a moment ago. The lesson works best when the confident wrong answer came from people first.

Why it happens

The usual word, when the usual word is wrong

The machine does not check. It cannot. It picks the word that usually comes next.

Most of the time the usual word is also the true one, and the answer is fine.

But when the true answer is rare and a nearby answer is common, it picks the common one. With the same steady voice.

It was not trying to trick you. It was doing exactly what it always does.

Step 2 · Make it wrong · 20 min

The pair challenge

In pairs, get the machine to say something confidently wrong. Write down what you asked and what came back.

Questions that work: how many days February had in a certain year. Whether a made-up book was written by a real author. The score of a match that never happened. Something that happened last week. Something about you.

Then two sentences. Sentence one: what it said. Easy.

Sentence two: why that particular wrong answer, and not some other wrong answer. That is the whole lesson.

Do this together

"It picked the thing that usually comes next" is the destination, in whatever words you have. "It saw lots of writing about leap years near the word February" counts. "It was broken" does not. Ask what it was doing instead.

Watch your words

Saying it the true way

When it gets something wrong, it is tempting to say it lied, or it got confused. Neither is true. Try these instead.

> Someone might say: It lied to me.
>
> **It produced the words that usually follow my question, and this time they were not true.**

> Someone might say: It got confused.
>
> **The true answer was rare and a nearby answer was common, so it picked the common one.**

> Someone might say: AI is rubbish.
>
> **It is unreliable on rare facts, recent things, numbers, and anything about me.**

Step 3 · The plenary · 10 min

Which one was hardest to spot?

Put three of your examples where everyone can see them.

Ask which of the three would have been hardest to notice, and why.

It is usually the one where the wrong answer sounded sensible rather than silly.

That is the thought to leave with. The dangerous mistakes are not the ridiculous ones.

Next time · Plate 3

So how would I know?

You have watched a confident answer be wrong. The next question is the obvious one.

Plate 3: what a source is, what only looks like one, and the most useful sentence in research.

For the grown-up

Why this plate matters

The failure to guard against is a child concluding "AI is rubbish". It is as useless as "AI is magic" and lasts about a week.

The outcome is not distrust. It is a working model of when to expect an error: rare facts, recent events, anything with a number in it, anything about themselves.

Children will want to make it say something rude. Anyone can make a machine say a bad word, and nobody learns anything. Making it say something false and explaining why is the harder trick, and the one that transfers.

If a pair cannot get a wrong answer out of it, that is a good problem. Have them write down what they tried. Showing where it is reliable is real work too.