

AI should make us better. Is it? Find out.

A short paper for senior leaders on a question most organisations cannot currently answer - and the case for looking beyond AI performance to what AI is making the organisation become.

CONTENTS

1	Three questions	03
2	What you already measure	04
3	Why this is difficult to see	05
4	Three things worth knowing	06
5	Two situations worth recognising	07
6	What the assessment does	08
7	What it is not	09
8	Why now	10
9	What taking part involves	11
10	Where this stands, and the invitation	12
A1	AI value: what gets missed in the business case?	13
A2	Beyond the business case: what happens to capability?	14

Three questions

Think about your most significant AI deployment. Then answer three questions about it.

What did it make possible?

What did it make unnecessary?

What did it make harder to do without?

Most organisations can answer the first. Many can answer the second - it is usually in the business case.

Very few can answer the third.

That third question is the subject of this paper. Not because the answer is alarming, but because it is rarely asked - and the answer tends to matter later.

What you already measure

By now your organisation can likely tell you, for any significant AI deployment: what it cost, what it saved, whether it is faster, whether quality held up, whether customers noticed, whether people are using it, and whether it passed governance.

That is a reasonable set of measures, and a competent transformation team will have them.

What those measures were not designed to tell you

WHAT YOU CAN ANSWER	WHAT IS HARDER TO ANSWER
Did productivity improve?	What did we build with the time it released?
Did we reduce cost?	What did we give up, or become dependent on?
Is a human reviewing the output?	Is that person actually exercising judgement?
Did people get trained?	Are they becoming more capable through the work itself?
Did governance approve it?	Has governance changed since?
Can we scale it?	What new dependency does scaling create?

Nothing in the left-hand column is wrong. The point is that the right-hand column is not a harder version of the same question. It is a different question, and no measure in the left column answers it.

AI adoption is not the same as AI impact.
And AI productivity is not the same as human capability development.

Why this is difficult to see

There are two separate reasons, and they need different responses.

Some of it nobody measures

No organisation tracks whether a review step produces real scrutiny or simply records an approval. None tracks whether the expertise needed for difficult cases is still being built. These are not numbers filed in the wrong place. They do not exist.

And some of it is measured, but in a way that hides the connection

Consider how the two halves of a single decision travel through an organisation.

- **The saving** appears this year, in a named budget line, attributed to the AI initiative. It is clear, early and easy to report.
- **The consequence** appears later - as an incident, a recruitment cost, a slower response on an unusual case, a business continuity problem. Each is recorded. Each is attributed to the incident, the labour market, the team, the IT estate.

Both halves are visible. They are just never visible *together*. By the time the second half arrives, nobody connects it back to the decision that produced it.

The savings are legible and the consequences are not, which means the trade is being made without being seen.

This is not a failure of competence. It is how organisations are structured. Finance sees one part, Operations another, HR another, Technology another. The connection is nobody's job.

Three things worth knowing

These come from published research. Sources are available in the accompanying evidence paper.

1. A human in the loop is not the same as human judgement

In a study of UK general practitioners using clinical decision support, the system improved decision accuracy in 13.1% of prescribing cases. There was a net improvement in performance.

But in 5.2% of cases, clinicians switched from a **correct** prescription to an incorrect one after the system gave them poor advice. The same tool that helped more often than it harmed also caused some good decisions to become bad ones.

The study also found that **less experienced clinicians switched their decisions more often**. That matters here for a particular reason: if AI simultaneously removes the routine work through which people become experienced, the two effects compound.

2. The same tool can build capability or consume it

A growing body of research distinguishes two ways people use AI: one in which the person delegates the thinking, and one in which the person does the thinking and uses AI as support. These patterns appear to have very different effects on what people retain and develop.

What that means for leadership: it may be the way AI is used that determines the outcome, not how much it is used. Two organisations with identical adoption rates and identical productivity gains could be moving in opposite directions - and no adoption metric would tell them apart.

3. People are poor judges of their own AI productivity

In a controlled study, experienced software developers using AI tools took **longer** to complete tasks than those without. They expected to be substantially faster. Afterwards, they still believed they had been faster.

Perceived productivity and measured productivity are not the same thing - and where they diverge, the gap tends to flatter. Where your evidence that AI is working relies heavily on what people report, it is worth testing those reports against something measured.

Two situations worth recognising

You automate your customer service front line

Cost per interaction falls. Resolution times improve. Satisfaction holds. Every measure says this worked.

Meanwhile, the straightforward cases stop reaching your people. Only the difficult ones do. Your experienced staff are fine - they learned on the easy cases years ago. Your newer staff never get them. In three years, when those experienced people have moved on, the question is who handles the difficult cases, and where they learned to.

Nothing here is an argument against the automation. The question is whether anyone decided this would happen, or whether it simply did.

If that capability has to be rebuilt later, the organisation pays for it - through recruitment, training, slower handling, or some combination of the three.

You introduce AI into a professional decision process

Decisions get faster and more consistent. A human reviews each one before it goes out - a proper control, documented, auditable.

The reviewer has thirty seconds. They have a throughput target. The output is confident and usually right. They approve 99% of what they see.

Ask whether that step is oversight or a record of approval. Then ask how you would know the difference.

If the control fails, the cost is not the cost of the review step. It is the cost of the decision the review did not challenge.

What the assessment does

It asks a structured set of questions across five connected areas, and produces a report.

1 · PURPOSEFUL TRANSFORMATION

Do we know why we are using AI, and what for?

2 · HUMAN CAPABILITY

Are our people becoming more capable, or more dependent?

3 · COLLECTIVE INTELLIGENCE

Are we thinking and working better as an organisation?

4 · RESPONSIBLE STEWARDSHIP

Can we govern this as it keeps changing?

5 · LEARNING & ADAPTATION

Can we tell what is working, and change course?

What you get back

Where the organisation stands on each area, and what the assessment found - with the answers that triggered each finding, so you can check it or disagree with it.

Each finding points at a choice rather than a verdict: where automation should stop, which capabilities need to stay alive, where released capacity should go, how much dependency is acceptable, what needs measuring before the next investment decision. Each also shows what the situation looks like if left as it is, and who would need to approve doing something about it.

Where you supply financial figures, the report sets out which potentially material costs and benefits sit outside them - so the reported return can be read in its wider context, and so you can decide whether any are worth putting a number to.

If several people take it

Invite four or more colleagues at different levels and we compare the answers. Where you agree, you have something firmer than any one person's view. Where you disagree - particularly between those who make decisions and those closest to the work - the gap itself becomes the useful thing. Both accounts can be accurate from where each person sits. The assessment makes those differences visible rather than averaging them away.

The core assessment takes 15-20 minutes. If you also complete the optional financial section - the fuller version - allow 30-45 minutes. **Answer for what actually happens, not what is intended.**

What it is not

Being clear about this matters more than the pitch.

- **It does not replace your AI strategy, your financial analysis, your risk review or your governance.** Those remain necessary. This sits alongside them.
- **It does not assume your transformation team has missed something obvious.** They are probably already doing much of this work. What is hard is connecting it across functions.
- **It does not put a pound value on human capability.** It identifies where capability may be changing. What that is worth is a question for your figures, not ours.
- **It does not tell you AI is bad for your organisation.** Capability can go up as easily as down. The assessment is built to find either.
- **It does not produce a single score.** A number would hide exactly what you need to see.
- **It is not an evaluation of individual employees.** The focus is the organisation's experience with AI.

It does not tell you what to decide. It gives you a clearer view of what is at stake when you decide.

Why now

You are currently making decisions about what to automate, what people should continue to do, where judgement should stay human, which skills to build, what to do with the savings, and how much dependency is acceptable.

Those decisions are easier to shape now than to unwind later.

The assessment is equally useful where AI is already running, because you can examine what actually happened rather than what the business case predicted.

The underlying idea

AI capability is developing extremely quickly. Organisational capability does not automatically develop at the same speed.

The organisations that benefit most from AI may not simply be those that deploy the most technology. They may be those that learn fastest about what the technology is actually doing to the organisation - the value it creates, the costs it introduces and the opportunities it reveals - and respond intelligently to what they find.

That is not a philosophical question. It is a question about the organisation you are building.

If your AI systems became twice as capable tomorrow, would your people become more capable - or simply less necessary?

What taking part involves

Deliberately simple. There is nothing to install, nothing to prepare, and no need to assemble data in advance.

1 · You complete the assessment

Online, in one sitting, answering from your own experience of AI in the organisation. The core assessment takes 15-20 minutes. If you also complete the optional financial section - the fuller version - allow 30-45 minutes. Answer for what actually happens, not what is intended.

2 · You get your report straight away

On screen as soon as you finish, ready to save or print. Nothing to wait for.

3 · Colleagues can take it too

Where four or more people across different levels complete it, we compare the responses and produce a second report showing where views align, where they diverge, and what that difference may be telling you.

4 · We go through the findings together, if useful

A conversation rather than a presentation. The questions worth asking are usually: are the expected benefits being realised, what costs are visible and which may not be, where is time being saved and where is work being created, are different teams taking overlapping or fragmented approaches, and what should be measured before the next investment decision.

5 · You tell us what worked

What was useful, what was not, and what was missing. That is the part we most need.

Two things worth saying plainly. There is no requirement to disclose commercially sensitive information - the financial section is optional and uses whatever figures you are comfortable with. And the assessment is not an evaluation of individual employees. The focus is the organisation's experience with AI, not anyone's performance within it.

Where this stands, and the invitation

Being straight with you

This assessment is being tested. It is built on a documented framework and the evidence base behind the questions is real, but the assessment itself has not yet been validated against the thing that matters most - whether it changes decisions. That is what the current trial is for.

We are working with a small number of organisations already using or introducing AI, who are willing to take it, say what worked and what did not, and tell us whether it showed them anything.

The test is simple

- Did it show you something about your AI transformation that you were not already seeing?
- And did it change, or sharpen, a decision?

If the answer to both is no, we would rather know.

It is **free during the trial**. In return we ask only for an honest view of whether it showed you anything useful.

Interested in taking part?

Take a look at the assessment: capacity2grow.com/human-flourishing-ai-assessment

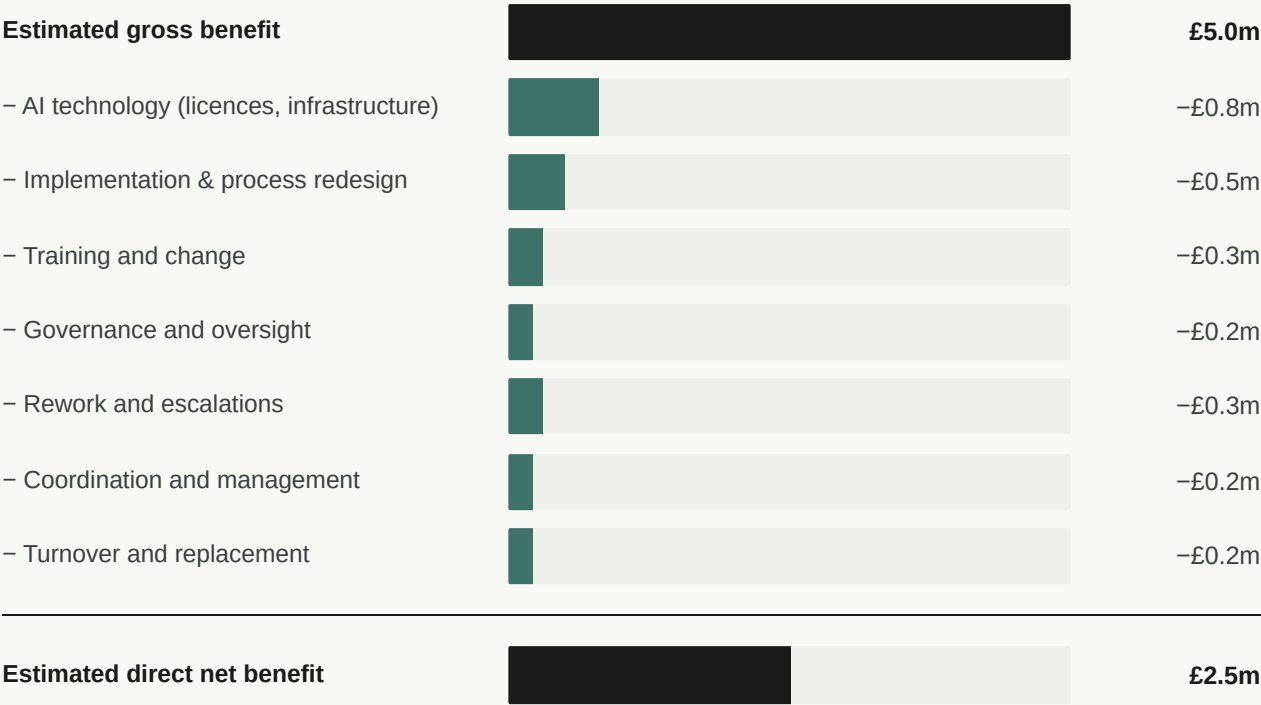
I would be very happy to explain the approach, answer questions, and discuss whether a trial would be useful for your organisation.

Firooz Firoozmand

Capacity2Grow

AI value: what gets missed in the business case?

Suppose an organisation estimates that an AI programme will create **£5m of annual gross benefit**. That might look like a very successful transformation. But let us make some of the costs visible.



Still a substantial benefit. But notice what has not been priced - and that is where the business case gets more interesting.

The figures above are illustrative assumptions, not industry benchmarks. In a real assessment, the organisation's own financial and operational data should be used.

Beyond the business case: what happens to capability?

The £2.5m calculation is useful. But it still does not tell the whole story. What happens to the things that are difficult to put a reliable value on?

EXPERTISE

What expertise is lost when experienced people no longer do the work?

JUDGEMENT

What judgement stops developing because people no longer get the experience?

LEARNING

What learning opportunities disappear?

RELATIONSHIPS

What happens to trust, relationships and collaboration?

RESILIENCE

What happens to organisational resilience if fewer people understand how the underlying work actually works?

These should not simply be given invented pound values. They are **potential value at risk** - things that may matter to the organisation but are harder to quantify reliably.

If people stop developing certain capabilities today, what happens to the organisation's capability in year three or year five? The effect could compound over time.

So the question is not simply "**Should we automate?**" It is:

"What should AI automate, what should it augment, and what human capability do we still need to deliberately develop?"

Value created. Value transferred. Value potentially lost.

AI should make the organisation better - not simply cheaper.