THE AUTOMATION CEILING

AI could do the job tomorrow. The math still caps how much you can automate.

Five measurable constraints, not model capability, set the real ceiling on cognitive automation. Most business cases never run the numbers.

TERENCEKOK.COM

What actually caps automation

01 The corrected ceiling formula, 1/(f·v)

02 Why a second model buys less safety than you think

03 Compute scarcity reverses the cost advantage

04 Why maximum substitution isn't maximum profit

05 The five leading indicators to actually track

Everyone cites 1/f. It's wrong.

The standard claim: if a fraction **f** of workload needs human adjudication, your automation ceiling is 1/f. That's only true if reviewing a decision costs as much as producing it. It almost never does.

Max speed-up ≤ 1 / (f · v)

v = adjudication effort as a share of full production effort. At f=0.10, v=1.0 → 10x ceiling. At f=0.10, v=0.2 → 50x.

2

votes of effective independence from a 9-model LLM judge panel

Nine frontier judges, seven model families, tested against 100 human annotations per item. About three-quarters of the panel's nominal independence disappeared because the models made the same mistakes on the same items. The best single judge matched or beat the full panel every time.

66%

of all AI compute now goes to inference, up from 33% in 2023

Push automation demand up fast enough and the price of the constrained input, serving capacity, rises with it. That erodes the exact cost advantage that justified the substitution in the first place. Comparative advantage returns through the price of what's scarce.

Maximum substitution isn't maximum profit.

- Where every competitor holds equivalent AI capability at comparable cost, the capability itself generates no rent.
- Cost reduction transmits into price. The surplus lands with whoever holds a **scarce complement**: proprietary data, distribution, regulated position, institutional trust.
- PwC: AI-exposed entry roles are now 7x more likely to require the judgement and leadership skills once reserved for senior grades.

35% vs −10%

growth in "seniorised" entry roles since 2019, versus a 10% decline in the rest

AI strips out the routine work that used to double as an apprenticeship, and pulls judgement earlier into the career. That's not a shrinking labour market. It's a splitting one, and the split rewards exactly the thing that's hardest to automate.

PWC, 2026 GLOBAL AI JOBS BAROMETER

"

Nine judges from seven model families behaved like two. If your assurance plan depends on models catching each other's errors, that plan is resting on a number close to one.

ON REDUNDANCY THAT ISN'T REDUNDANT

Five leading indicators, not one coverage number

- [ ] Escalation fraction (f), is routing discipline holding or drifting
- [ ] Adjudication effort ratio (v), is review engineered cheap or redoing the work
- [ ] Detection rate on seeded defects, does review still catch what it should
- [ ] Error correlation across the model estate, real redundancy or the same mistake twice
- [ ] Queue utilisation at the 95th percentile, one incident from a backlog

READ THE FULL ANALYSIS

Size your ceiling before you size your roadmap.

The full piece walks through the formula, the correlation math, and the five indicators worth putting on your dashboard this quarter. Free on terencekok.com.

Terence Kok
AI & Innovation Strategist

SAVE · COMMENT
FOLLOW