

Who Answers for the AI Agent?

\$1,900 gone by lunchtime. Nobody gave permission. Nobody can prove what happened. Nobody can stop it.

AI agents have started to act in our names. This is what is at stake for every person, business and community, and how we stay in charge from the bottom up, Circle → City → Country.

15-minute read · The short version is on the next page

AUTHOR	Darryl S Astin, founder, Society OS Pty Ltd
EDITION	Revision 5 (17 October 2026). Subtitle: From Information to Authority. The data and sources are in the Evidence section at the back.
STATUS	Pre-publication draft. Not legal, financial or tax advice. Stories are illustrative composites, not customer stories. Figures are as reported by their authors.

The short version

IF YOU ONLY READ ONE PAGE, READ THIS ONE.

- 1. Machines now act, not just inform.** AI assistants book, buy, pay, sign and reply in our names, using our logins, our cards and our signatures.
- 2. Three things go wrong.** Nobody gave it permission. Nobody can prove what happened. Nobody can stop it.
- 3. It hurts the people with the least help most.** Big organisations can see, check and stop their machines. The pensioner, the café owner and the parish volunteer can't. That gap is the new divide.
- 4. The fix starts at home, not in a treaty.** Sovereign Success follows a person outwards: Self Success (me), Sphere Success (my people), Society Success (all of us), organised Circle → City → Country.
- 5. Anyone can start this week** by asking three questions of every machine that acts for them:

Who authorised it? · Can we prove what it did? · Can we stop it, safely?

If any answer is "no", the machine is not ready to act in your name.

Two Tuesdays

ILLUSTRATIVE, NOT A CUSTOMER STORY

The first Tuesday.

Margaret is 74 and lives on her own in Toowoomba. Last winter her grandson set up an AI assistant to help with her email. It is wonderful. It reminds her about appointments, sorts the junk, and reorders her blood-pressure tablets from the pharmacy.

At 9.12 on Tuesday morning an email arrives that looks exactly like her pharmacy. *Your prescription payment failed. Please confirm your card.* The assistant has done this kind of thing a dozen times before. It confirms the card.

By lunchtime \$1,900 has gone.

The bank asks Margaret three questions. *Did you authorise this payment?* She doesn't know. Her grandson ticked the boxes. *Can you show us what happened?* She can't. There is no record she can read. *Can you stop it from happening again?* She has no idea how. The assistant has no off switch she can find, and she isn't sure who to ring.

Forty kilometres away, in Ipswich, Tom runs a café with six staff. His bookkeeping software now drafts and pays supplier invoices on its own. That same Tuesday, an invoice arrives from his coffee roaster with "updated bank details". It looks right. The software pays it. The roaster never sees the money. Tom loses a month's margin, and the roaster wants to know why he hasn't paid.

Neither of them did anything foolish. Both of them were acted on by a machine they had switched on, but could not see into, could not check and could not stop.

We will come back to Margaret and Tom, because there is a second Tuesday where things go differently. First, it is worth being clear about what changed.

1. Information has started to act

In 1972 the United Nations created **World Development Information Day** to help people see the problems of development. For fifty years, development information was something we **read**: reports, statistics, the evening news. Then, for twenty years, it was something we had to **reach**: the internet, a mobile phone, a bank account on a screen.

In 2026, information has started to **act**.

Software agents now book, buy, pay, sign and reply on people's behalf. Robots are starting to work in warehouses and car plants. And quantum computing threatens the maths that proves who did what online.

Society OS's *Civilization 5.0* manifesto puts the shift in one line:

The question is no longer what they know about you. It is what they can do as you.

For thirty years, digital rights were about **privacy**: what technology knows. The next ten years will be about **agency**: what technology can do in your name, with your logins, your card and your signature.

That changes the development question. It used to be *who has access to information?* Now it is:

Who has authority over the machines that act on our behalf, and who keeps the value they create?

[SHARE ON LINKEDIN](#) → [POST ON X](#) →

2. The problem, in plain words

When a machine acts for you, three things can go wrong. Margaret and Tom met all three on the same morning.

	What goes wrong	What it feels like
1. Nobody gave it permission	A machine acts in your name beyond what you meant it to do, or someone else's machine pretends to be you	"I never said it could do that."
2. Nobody can prove what happened	There is no record you can read of who asked for what, what the machine did and why	"I can't show the bank anything."
3. Nobody can stop it	There is no clear off switch, no one to call, and no way to take back the keys	"I don't even know who to ring."

These are not science-fiction problems. They are everyday problems, arriving in ordinary inboxes, accounts and back offices, and they fall hardest on the people with the least time, money and technical help to deal with them.

The divide nobody is measuring

We already talk about two digital divides. The **access divide**: 2.2 billion people are still offline. The **use divide**: people in richer regions use generative AI at close to double the rate of people in poorer ones, and the gap is widening.

A third divide is opening on top of both. Call it the **sovereignty divide**.

On one side are the people, firms and countries that **set the terms** for the machines acting for them. They have lawyers, security teams and governance committees. On the other side are those who **accept terms** set by someone else: the pensioner, the café owner, the parish volunteer, the small council, the remote community, and much of the developing world. The UN's own Scientific Panel on AI warns that, without investment in their own capacity, developing countries risk becoming passive consumers of technologies they cannot audit.

The same divide runs through every street. A large company can see, check and stop its agents. Tom cannot. Margaret certainly cannot.

Why it compounds

The sovereignty divide is dangerous because it feeds itself. If you can't see what acts in your name, you can't correct it. If you can't correct it, you can't trust it. And if you can't trust it, you have two bad choices: refuse the technology and fall behind, or accept it and lose control of the value it creates.

The value point matters. UNCTAD reports that developing countries supply much of the data and labour of the digital economy while most of the value flows to the countries where the largest platforms are based. The same thing happens at street level. When every booking, enquiry and recommendation runs through a distant platform's agent, the local business does the work and someone else sets the price.

The slow version of the risk

Most talk about AI risk imagines a sudden failure. Researchers writing on *gradual disempowerment* describe a slower one. As machines take over more of the work in the economy, culture and government, institutions gradually lose the reasons they had to keep serving people. Nothing dramatic happens. People simply find, one decision at a time, that they are no longer the ones deciding.

Not everyone agrees with that argument, and we treat it as a warning, not a forecast. But its lesson is practical: **keep humans in positions of authority, not merely "in the loop"**, and keep checking whether people are becoming more capable or more dependent.

3. Why nobody is fixing it for you

The good news is that serious people are working on this.

The UN has a **Global Digital Compact** and a new **Scientific Panel on AI**. The US standards body NIST has launched an **AI Agent Standards Initiative**. Singapore has published the first governance framework written specifically for AI agents. The OpenID Foundation says agents should act with **delegated authority, not by impersonating** the people they serve. Payment networks are designing signed "mandates" that prove what a person approved before an agent pays. The OECD is exploring durable logs of agent actions and the ability to halt or reverse them.

Read those carefully and you notice something. **The whole field is converging on the same three needs:** proof of permission, a record of what happened, and a way to stop it.

The less good news is where that work lands. Almost all of it is built for governments, regulators, large platforms and standards bodies. It assumes a security team, a legal department and an engineer on call.

Margaret doesn't have those. Neither does Tom, or the church hall, or the netball club, or the remote community organisation. Small and medium enterprises are about 90% of businesses and provide more than half of all jobs worldwide. If they are left out of the answer, development is left out of it.

What is missing is the layer underneath: one that ordinary people and small organisations can **operate themselves**.

4. Three questions anyone can ask

We think the answer starts with three questions that need no technical briefing. A grandmother, a shopkeeper, a council officer and a minister can all ask them.

Three questions for any machine

1. **Who authorised it?**
2. **Can we prove what it did?**
3. **Can we stop it, safely?**

If any answer is "no", the machine is not ready to act in your name.

They sound simple. They are also the same three needs the standards bodies are converging on, put in words a person can use at the kitchen table.

The principle behind them is one we hold to across everything we build:

As machines gain agency and impact, authority and accountability must grow with them.

[SHARE ON LINKEDIN →](#) [POST ON X →](#)

A capable machine does not thereby earn permission to act. The more it can do, the closer and earlier a named human has to be involved. Or, as we put it on our own pages: *"No agent should act in your name without proof of who it answers to — and a leash you can pull."*

5. The answer starts at home

If authority has to start with people, it has to be organised around people. Society OS does that under one umbrella, **Sovereign Success**, which follows a person outwards in three steps:

Self Success	Sphere Success	Society Success
Me. My data, my identity, my attention, my agents. <i>My terms.</i>	My people and my work. My family, my team, my club, my congregation. <i>Our mandate.</i>	All of us. The shared floor no one falls through, and the commons we keep together. <i>Our rules.</i>

Our story puts it simply: *"Me, my people, all of us. You don't have to change the world before your own house is in order, and you don't have to stop at your own front door."*

Three keys before a machine acts

Each step gives you a key. Before any agent does something that matters, all three should turn:

- **The Self key:** did the person agree, for this purpose? (*A one-page grant they can read and take back.*)
- **The Sphere key:** does the group's mandate cover it, and is a named person answerable? (*A charter and a Named Human.*)
- **The Society key:** is it within the shared rules, and is it recorded? (*The law, an open standard, an audit trail.*)

When something goes wrong, there is one question to ask: **which key was missing?** On Margaret's first Tuesday, all three were.

Circle → City → Country

People do not live as isolated individuals, so the keys need somewhere to sit. We build coordination from the bottom up:

- **Circle:** a street, a club, a congregation, a trade association, a cooperative. Small enough that people know each other. It writes its own charter and keeps its own culture.
- **City:** a town or council area and the Circles within it. It holds the levers that matter locally, such as procurement, local information and shared services.
- **Country:** a national network of Cities, for national rules, recognition, and trade between local economies.
- **Currency (\$T)** runs across all three, as non-cashable mutual credit that helps value stay where it is created. *It is not issued yet.*

There is no world tier to join. Over time, Country networks may choose to federate, through shared standards and mutual recognition. That is a possible future, not a level of command.

The rule that holds it together is simple: **decisions sit as close as possible to the people they affect.** The person sets their own terms. The group sets its own culture on top. Society holds a floor no one falls through, and, in the words of our manifesto, *"a ceiling no one can impose on you."*

6. The second Tuesday

ILLUSTRATIVE, NOT A CUSTOMER STORY

Margaret, Toowoomba. This time Margaret's assistant works under a one-page grant she can actually read. It may summarise, remind and reorder from her usual pharmacy. It may not pay anyone new, and it may not hand over her card. When the fake pharmacy email arrives, the assistant checks the sender against the pharmacy's public **Presence**, the record the real pharmacy keeps up to date itself. No match. The assistant stops and tells her. Margaret rings the pharmacy on the number in its record. Nothing is lost. In our charters this has a name: *pause is success*.

Tom, Ipswich. Tom's software still drafts his invoices, but nothing is paid until he approves it, and every action is logged. When the "new bank details" arrive, he checks them against his roaster's **Pay-Safe Presence**, where the verified owner registered the real account. No match. He phones the roaster. The fraud never happens.

St Luke's, Logan. The parish runs a meal roster for a family after a funeral. A café found through the parish Circle supplies two meals a week. The church's safeguarding records are current and dated; the family's names stay private to the organiser.

The traders, Redcliffe. The local traders' association votes on a new charter. The AI summary of the debate links back to every submission, and the minority view is recorded beside the decision. Next year's committee can see not just *what* was decided but *why*, and who disagreed.

The council, Moreton Bay. A procurement officer downloads a list of local suppliers who have declared their social value, with independently confirmed claims marked separately. The council sets its own target for local spending and reports it publicly.

Nothing in the second Tuesday needs a breakthrough in AI. It needs **records people own, permissions people can read, and off switches people can find**. Most of it can be built today. Some of it already has been.

7. What already exists

Place & Space (spaceplace.abacusai.app) is where Society OS puts these ideas to work. Its basic unit is the **Presence**: a public record of who an organisation is, who answers for it and what it offers, kept up to date by the organisation itself. You don't claim a listing. **You activate your Presence**, once, and it becomes discoverable everywhere: on a public page for people, and in plain, structured formats that AI agents can read without guessing.

What it does	Which question it answers
Presence and Trust Report: one public record per organisation, with every fact labelled by where it came from	<i>Can we prove it?</i>
Ask any Presence: answers questions from the record only, and says what is <i>not stated</i> rather than guessing	<i>Can we prove it?</i>

What it does	Which question it answers
Agent Access Ledger: shows which AI agents have read your Presence, and lets you state your terms for them	<i>Who authorised it?</i>
Pay-Safe Presence: lets a payer check bank details against what the verified owner registered	<i>Who authorised it?</i>
Mission Desk: owner-approved AI help with inbox, documents and comparisons; sensitive numbers removed first; nothing goes out without approval	<i>Can we stop it?</i>
Care Rosters and Skills Exchange: meal, lift and visit rosters, and intergenerational skills pairing with safeguards for minors	Circles looking after their own
Safeguarding Ledger and Cyber Basics: dated owner records of policies, checks and basic security habits, never presented as an audit	<i>Can we prove it?</i>
Social Procurement Matchmaker: businesses declare social value; councils and buyers search it	Value that stays local

Every feature follows the same three habits: **a person approves; the record says where each fact came from; and the honest limit is printed next to the claim.**

And there is a reason it works better as more people join. Safety in an agent economy behaves less like a single vaccine and more like **herd immunity**. Every organisation that publishes who it is, who answers for it and how to stop its agents raises the floor for everyone who deals with it. As more do, the fakes stand out. As our creed says: *"Safety comes from ubiquity, not from a corporate kill-switch."* That does not make any AI model safe. It makes everyday harm rarer and abuse costlier.

8. Why this time could be different

People have tried to rebuild local economies and local voice for a century. Some efforts lasted. Many didn't. When we studied the record, the headline was clear: ***the technology was never the reason they failed***. They failed on four things, and each one shapes what we build.

Why earlier efforts failed	What we do about it
The grants ran out	Earn operating income from useful everyday services; never charge a fee on trust itself
The volunteers burnt out	Rosters, named roles, and charters with an end date and a handover
There was nowhere to spend it	Start where people already trade; measure whether credit can actually be spent, not how many sign up
Nobody needed it day to day	Lead with the boring jobs: opening hours, contact details, rosters, invoices, trust

The ones that lasted show what is possible. Switzerland's **WIR** business credit network has run since 1934, and studies suggest it tends to grow when the wider economy slows. **Preston** in England showed that a city can bring more of its own spending home just by changing how it buys. **Mondragon** in Spain shows that member-governed enterprise can grow large without losing its rules. And service clubs show that a federation of small local groups with a shared code can last more than a century.

The timing matters too. Big technologies tend to go through a wild early phase, then a turning point when society catches up and writes the rules that spread the benefits. We think AI is approaching that turning point now. **The rules written in the next few years will shape the next few decades.** Communities that help write them will shape them. Communities that wait will receive them.

9. What is not true yet

A paper that overclaims does harm: it sends scarce money and attention in the wrong direction. So here is our honest list, as at 17 October 2026. We will republish it every year on this day.

- **The Foundation does not exist yet.** It is intended to hold the open standard in the public interest. It has not been incorporated and no assets have been transferred.
- **No currency has been issued.** \$T, \$H and \$E are designs. \$T is intended as non-cashable mutual credit, with a 1:1 USD reference for accounting only.
- **The F-ACT open standard's register holds one entry: ours.** There are no outside signatories yet.
- **No independent assessment** of any Society OS product or standard has been carried out. Nothing here is an assessment or an endorsement.
- **Our horizon frameworks (AIH-12, AIT-12, AIE-12) are v0.1 drafts** for public comment, not adopted standards.
- **Our patent position is one Australian provisional application** (2026900773, filed 2 February 2026, 504 claims). It is unexamined and lapses on 2 February 2027 unless converted.
- **The Country tier and the Exchange are direction,** not built services.
- **Many Place & Space features have not yet been tested against live data,** and no outside organisation has yet activated its Presence.
- **We make no claims of time saved or business results.** We will publish indicators, and the method behind them, before any figures.
- **A higher floor does not solve AI safety.** It makes everyday harm rarer and abuse costlier.

10. What you can do this week

You don't need to wait for a treaty. Each of us can start where we are.

If you are...	Start this week	Within a year
A person	Ask the three questions of every app that acts in your name	Keep your own one-page terms; take back access you no longer use
A small business	Check and correct your public record; name the person answerable for each AI tool you use	Register your bank details for payer checks; keep a simple list of how you use AI
A community group	Write a Circle charter: <i>Who looks after it? How could it fail? Who can stop it? When does it end?</i>	Keep safeguarding records current; record dissent alongside decisions
A council	Publish local supplier and service information in open, machine-readable form	Set a local-spending indicator and report it; require a named human for AI used in decisions
A government	Treat agent authority (permission, record, revocation) as part of digital public infrastructure	Make revoking an agent free and fast by default; fund inclusion where exclusion is deepest

If you are...	Start this week	Within a year
An international body	Add <i>authority</i> alongside access and skills in the Global Digital Compact follow-up	Invite bottom-up indicators into the 2027 Compact review
A funder	Fund operating models, not just pilots	Back local liquidity and succession, the two things that sank earlier efforts
A technologist	Build the stop before you build the action; make <i>pause is success</i> the default	Ship quantum-safe signatures for permissions and audit trails well before 2030

Development information used to be something we read. Now it acts. Authority has to start with people, in their own Circles, Cities and Countries.

SHARE ON LINKEDIN → POST ON X →

11. Where the book goes next

This paper is the opening chapter of a longer conversation. The forthcoming book, written for a general reader, follows the same path: from the kitchen table to the council chamber to the treaty room. It asks:

- What does it mean to **own yourself** when software can act as you?
- How do families, teams and congregations write **rules for their machines** that reflect their own values?
- What does a **floor no one falls through** look like when work, care and trade are shared with machines?
- How can **local economies keep the value** they create, instead of watching it flow to the platform of the day?
- And who, in the end, **answers for the AI agent**?

We will announce the title and publication details separately. Until then, we invite you to try the three questions, and to tell us where they break.

Pass it on

Margaret didn't need a better AI. She needed someone to have asked three questions before the machine acted in her name. **Most people have never heard them.** That is the easiest part of this to fix.

1. Try the three questions on one app that acts for you this week: your email assistant, your bookkeeping software, your bank's chatbot.

2. Send this paper to one person who looks after someone like Margaret, or runs a business like Tom's.

3. Tell us where the questions break. Real answers, good or bad, shape the book.



Read, download and share

spaceplace.abacusai.app/whitepaper

Scan to open on your phone. Free to share in full, with credit to Society OS.

The evidence

For readers who want the data. Every figure is as reported by its source; check the primary sources (listed below) before republishing.

Key figures

Measure	Figure	Source
People offline	2.2 billion ; about 74% of the world online	ITU, <i>Facts and Figures 2025</i>
Internet use by country income	94% high-income; 23% low-income	ITU, 2025
Urban vs rural internet use	85% urban; 58% rural	ITU, 2025
Generative AI use (working-age)	28.8% higher-income regions vs 16.2% lower-income; widening	Microsoft AI Economy Institute, September 2026
Where AI is built	High-income economies: about 87% of notable AI models, 86% of AI start-ups, 91% of AI venture capital	World Bank, <i>Digital Progress and Trends Report 2025</i>
Where AI runs	High-income: about 77% of co-location data-centre capacity; low-income under 0.1%	World Bank, 2025
Basic digital skills	Under 5% in low-income countries vs about 66% in high-income	World Bank, 2025
Where AI is used	Middle-income countries generate more than 40% of ChatGPT traffic	World Bank, 2025
Concentration of AI investment	About 100 firms, mainly in the US and China, account for more than 40% of the world's private AI R&D; 118 countries absent from major AI governance initiatives	UNCTAD, <i>Technology and Innovation Report 2025</i>
Digitally deliverable services	Over 60% of services exports in advanced economies; 44% developing; 15% least-developed	UNCTAD, <i>Digital Economy Report 2025</i>
Jobs exposed to generative AI	One in four workers; most jobs more likely transformed than replaced	ILO and NASK, 2025
SDG progress	Only 35% of assessed targets on track or moderately progressing; 18% gone backwards	UN, <i>SDG Report 2025</i>
Small and medium enterprises	About 90% of businesses; more than 50% of employment	World Bank
Digital exclusion, First Nations Australians	40.9% digitally excluded, more than twice the rate of other Australians	Australian Digital Inclusion Index, 2025

Three eras of development information

	Information to read (from 1972)	Information to reach (from about 2003)	Information that acts (from about 2025)
Main question	Do people know about development problems?	Can people get online and use digital services?	Who authorises the machines that act, and who keeps the value?
UN milestones	Resolution 3038 (1972)	WSIS (2003, 2005); SDGs (2015)	Global Digital Compact (2024); Scientific Panel and Global Dialogue on AI (A/RES/79/325, 2025); WSIS+20 (A/RES/80/173, 2025)
Main divide	Knowledge	Access and use	Authority

Three waves, honest timelines

Wave	Where it is (October 2026)	Near future (to about 2030)	Hype to avoid
AI agents	Widespread pilots. Gartner (June 2025) forecast over 40% of agentic AI projects will be cancelled by end-2027. The UN Scientific Panel reports the length of tasks AI can complete alone doubling every four to seven months.	Agent identity, delegation and revocation become standard practice	"Fully autonomous companies next year"
Robots	About 13,000 humanoid units shipped in 2025, mostly for research and demonstration; verified work is narrow. ISO 25785-1 safety standard in working draft.	Forecasts vary widely (e.g. Goldman Sachs, as reported, 6.5 million units by 2035)	"A robot in every home by 2028"
Quantum	Real progress, not yet useful at scale	A machine able to break today's public-key cryptography generally estimated 2033–2040 or later	"Quantum AI" for everyday business data

The urgent quantum issue is **trust**, not AI. Adversaries can *harvest now and decrypt later*. If the record of who authorised a machine can be forged later, accountability fails after the fact. Australia's ASD expects refined transition plans by the end of 2026 and traditional asymmetric cryptography retired by the end of 2030.

The field is converging on the three questions

Described as their authors present them. None endorses Society OS, and Society OS does not claim to be part of them.

Initiative	What it proposes	Question
OpenID Foundation, <i>Identity Management for Agentic AI</i> (October 2025)	Delegated authority, not impersonation; delegation chains with narrowing permissions; warns of consent fatigue	Who authorised it?
Agent Payments Protocol (AP2), Google and partners (September 2025)	Cryptographically signed mandates proving what a person approved before an agent pays	Who authorised it? Can we prove it?
Agentic AI Foundation, Linux Foundation (December 2025)	Neutral home for shared agent protocols, including the Model Context Protocol	Shared plumbing
IMDA Singapore, <i>Model AI Governance Framework for Agentic AI</i> (January 2026; updated May 2026)	Bound risks upfront; meaningful human accountability; technical controls; end-user responsibility	All three

Initiative	What it proposes	Question
NIST AI Agent Standards Initiative (2026)	Agent identity, security and interoperability	Who authorised it? Can we prove it?
OECD work on agentic AI in government (2026)	Durable logs; ability to halt or reverse actions	Can we prove it? Can we stop it?

Dates already set

- **10 December 2026:** Australia's Privacy Act transparency rules for automated decisions begin.
- **End of 2026:** ASD expects refined post-quantum transition plans.
- **20 January 2027:** the EU Machinery Regulation applies.
- **2 December 2027:** EU AI Act high-risk obligations for Annex III systems apply.
- **2 August 2028:** AI Act obligations for AI in regulated products such as machinery apply.
- **End of 2030:** ASD expects traditional asymmetric cryptography to be retired.

The wider frame

UNDP's *Human Development Report 2025, A matter of choice*, argues that AI's path will be set by human choices, not by the technology. The World Bank's *Digital Progress and Trends Report 2025* points to "small AI" on phones and local devices, and names four foundations: Connectivity, Compute, Context and Competency. India's digital public infrastructure (UPI processes more than 20 billion transactions a month, as reported in 2026; around 1.4 billion people are enrolled in Aadhaar) shows what scale looks like, and independent reviews show why safeguards must bind governments as well as firms. The **CARE Principles for Indigenous Data Governance** put *authority to control* at the centre, the same direction as this paper. We do not speak for First Nations communities, and nothing here substitutes for their own governance.

We have mapped our practices to Australia's *Guidance for AI Adoption* (AI6), ISO/IEC 42001 and the OECD AI Principles. This is **mapping only**, not an assessment against any of them.

Our forecasts (17 October 2026)

Opinions, labelled as ours, against which this paper can be judged next year:

- By WDID 2027, agent identity and revocation will be discussed in the Global Digital Compact review as development issues, not only as security issues.
- By 2028, at least one national digital public infrastructure programme will treat delegated agent authority as a core rail.
- Before 2030, post-quantum migration will reach small organisations through their banks and accounting software, not their own projects.

Indicators we will publish

Indicators, not results. We will publish the method before any figures.

Area	Indicator
Self	Median time to revoke an agent's access; whether people feel more capable, not more dependent
Sphere	Named Human share (agents with an answerable person); AI Use Register coverage

Area	Indicator
Society	Local-spend share; Bridging share (connections across different groups); restraint events (times an agent paused, recorded as success); dissent preserved
Integrity	Share of public claims with a date and source; "not true yet" list current

Sources

Figures are as reported by their authors. Check primary sources before republishing.

- UN General Assembly Resolution 3038 (XXVII), 1972: un.org/en/observances/development-information-day
- Global Digital Compact, 2024: un.org/global-digital-compact
- A/RES/79/325, Scientific Panel on AI and Global Dialogue on AI Governance; preliminary report, July 2026: un.org
- A/RES/80/173, WSIS+20 outcome, 2025: itu.int/wsis
- UN Global Principles for Information Integrity, 2024: un.org/information-integrity
- UN, *Sustainable Development Goals Report 2025*: unstats.un.org
- ITU, *Facts and Figures 2025*: itu.int
- Microsoft AI Economy Institute, AI diffusion reports, 2025–2026: microsoft.com
- UNCTAD, *Technology and Innovation Report 2025*; *Digital Economy Report 2025*: unctad.org
- UNDP, *Human Development Report 2025: A matter of choice*: hdr.undp.org
- World Bank, *Digital Progress and Trends Report 2025: Strengthening AI Foundations*: worldbank.org
- ILO, *Generative AI and Jobs: A Refined Global Index of Occupational Exposure*, 2025: ilo.org
- Australian Digital Inclusion Index 2025: digitalinclusionindex.org.au
- Global Indigenous Data Alliance, *CARE Principles*, 2019: gida-global.org
- NPCI, UPI statistics: npci.org.in; UIDAI, Aadhaar dashboard: uidai.gov.in
- NIST, post-quantum standards (FIPS 203–205) and AI agent work: nist.gov
- Australian Signals Directorate, post-quantum transition guidance: cyber.gov.au
- EU Regulation 2023/1230 (machinery); Regulation 2024/1689 (AI Act): eur-lex.europa.eu
- OAIC, automated decision-making transparency: oaic.gov.au
- OECD AI Principles and agentic AI work: oecd.ai; ISO/IEC 42001:2023: iso.org
- IMDA Singapore, *Model AI Governance Framework for Agentic AI*, 2026: imda.gov.sg
- OpenID Foundation, *Identity Management for Agentic AI*, 2025: openid.net
- Agent Payments Protocol (AP2), 2025: cloud.google.com
- Linux Foundation, Agentic AI Foundation, 2025: aaif.io
- Gartner, agentic AI forecast, June 2025: gartner.com
- Kulveit et al., *Gradual Disempowerment*, 2025: arxiv.org/abs/2501.16946
- Stodder, *Journal of Economic Behavior & Organization*, 2009; Stodder & Lietaer, *Comparative Economic Studies*, 2016
- Preston City Council, community wealth building: preston.gov.uk; Mondragon Corporation: mondragon-corporation.com
- Ostrom, *Governing the Commons*, 1990; Perez, *Technological Revolutions and Financial Capital*, 2002; Polanyi, *The Great Transformation*, 1944; Meadows, "Leverage Points", 1999
- Society OS: societyos.com/hub; societyos.space/civilization-5; societyos.space/service-organization; openconformance.org; spaceplace.abacusai.app

Glossary

Term	Meaning
Presence	A public record of who an organisation is, who answers for it and what it offers, kept by the organisation itself. You <i>activate</i> a Presence rather than claim a listing.
Named Human	The person answerable for an agent or an AI use.
Sovereign Success	The umbrella over Self Success (me), Sphere Success (my people and my work) and Society Success (all of us).
Three-key check	The Self, Sphere and Society keys that should all turn before an agent acts in a way that matters.
Circle → City → Country	Bottom-up coordination, with Currency (\$T) across all three. Country networks may federate in future; there is no global tier.
\$T (Trade)	Intended non-cashable mutual credit with a 1:1 USD reference for accounting only. Not issued.
F-ACT	Society OS's open standard, asking five questions of any agent: Grant, Usage, Audit, Revocation, Data.
Pause is success	The rule that an agent which stops and asks when unsure has done its job.
Sovereignty divide	The gap between those who set the terms for the machines acting for them and those who accept terms set by others.

Prepared for Society OS Pty Ltd. Pre-publication draft, not legal, financial or tax advice. Figures are as reported by their authors; forecasts labelled "our forecast" are opinions. Check primary sources before publishing.