SHADOW AI: IL RISCHIO SILENZIOSO E L'OPPORTUNITÀ NASCOSTA

Strategie per governare l'innovazione non autorizzata nell'era dell'IA Generativa.

Un framework per CIO e Business Leaders.

LA REALTÀ DEI DATI: L'ADOZIONE È PIÙ VELOCE DELLA GOVERNANCE

96%

Dipendenti che utilizzano strumenti di GenAI.
(Fonte: IBM)

78%

Lavoratori che portano i propri strumenti AI (BYOAI) eludendo l'IT.
(Fonte: Microsoft/LinkedIn)

34%

Lavoratori uraiili che usano la GenAI nami'a.
(Fonte: IBM)

68%

Lavoratori italiani che usano la GenAI settimanalmente.
(Fonte: BCG/Corriere)

38%

Dipendenti che ammettono di condividere dati confidenziali senza autorizzazione.
(Fonte: Aidia/Yellow Tech)

ATTENZIONE:
La maggior parte di queste attività avviene all'insaputa dei dipartimenti IT e Sicurezza.

OLTRE LO SHADOW IT: LA DIFFERENZA È NEI DATI

SHADOW IT (Statico)

- **Esempio:** Scaricare un editor PDF non approvato.
- **Rischio:** Vulnerabilità software locale.

SHADOW AI (Apprendimento Attivo)

- **Esempio:** Inserire contratti sensibili in un LLM pubblico.
- **Rischio:** "Data Exfiltration" immediata. Il modello apprende dai tuoi segreti industriali.

INSIGHT: Non è solo un software non autorizzato. È intelligenza non autorizzata che elabora proprietà intellettuale.

PERCHÉ ACCADE? LA RICERCA DELL'EFFICIENZA

"Le persone non aggirano il sistema per malizia, ma perché il sistema ufficiale è troppo lento."

- **53%** dei dipendenti dichiara che la produttività calerebbe senza l'IA.
- **66%** paga di tasca propria gli strumenti pur di lavorare meglio. (Fonte: Forbes)

ANATOMIA DI UNA FUGA DI DATI

Da un prompt a una violazione globale.

L'AZIONE

Un ingegnere incolla codice proprietario per il debug.

IL PROCESSO

Il modello pubblico assimila il codice nel suo training set.

IL RISULTATO

Il codice segreto diventa conoscenza pubblica.

CASE STUDY: SAMSUNG (2023). Ingegneri hanno caricato codice sorgente confidenziale su ChatGPT.
Risultato: Divieto totale e sviluppo soluzioni interne.

Una volta inviato il prompt, il dato non è più sotto il vostro controllo.

I QUATTRO QUADRANTI DEL RISCHIO AZIENDALE

PRIVACY E IP

Esposizione di segreti commerciali, codice sorgente e PII. Violazione GDPR.

COMPLIANCE

Violazione dell'EU AI Act. Perdita del privilegio legale sui documenti.

SICUREZZA

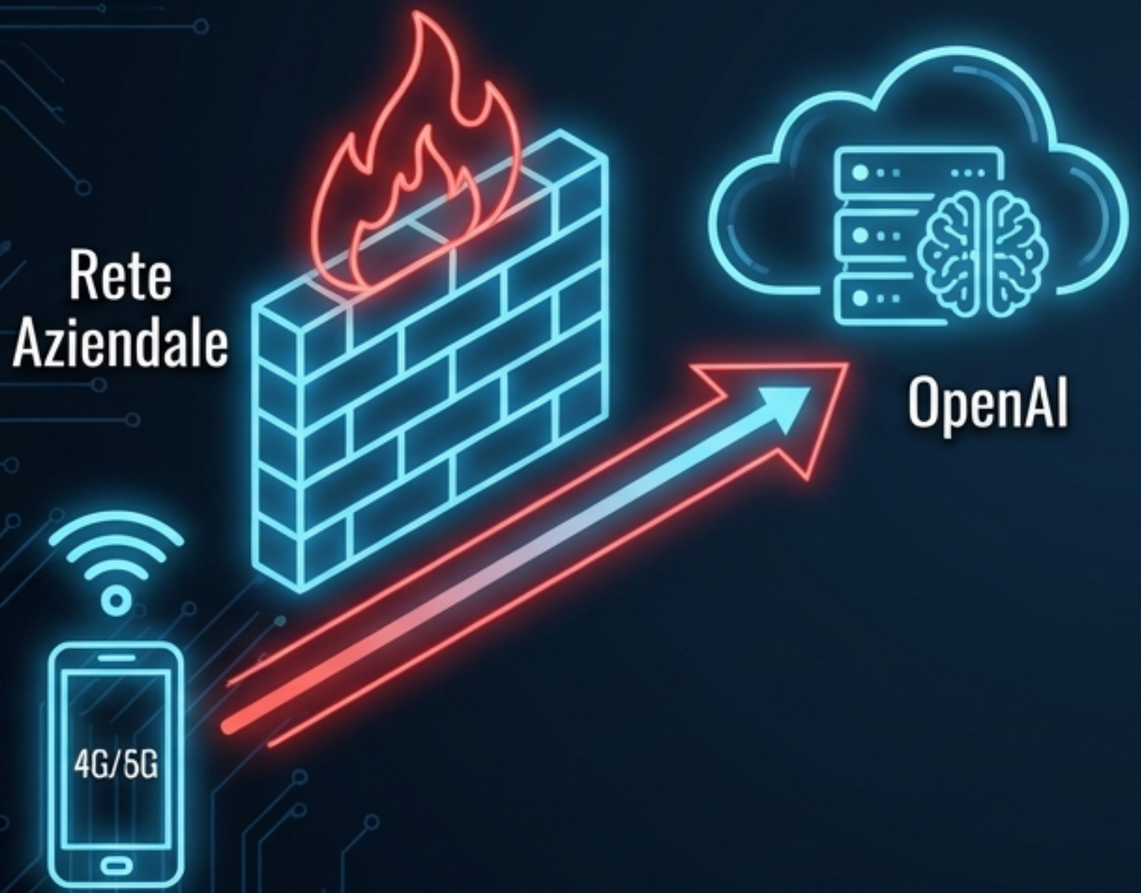
Prompt injection, malware tramite estensioni browser, superfici d'attacco espanse.

AFFIDABILITÀ

Decisioni basate su 'allucinazioni', dati obsoleti o bias algoritmici.

L'ILLUSIONE DEL CONTROLLO

Perché il blocco totale fallisce.

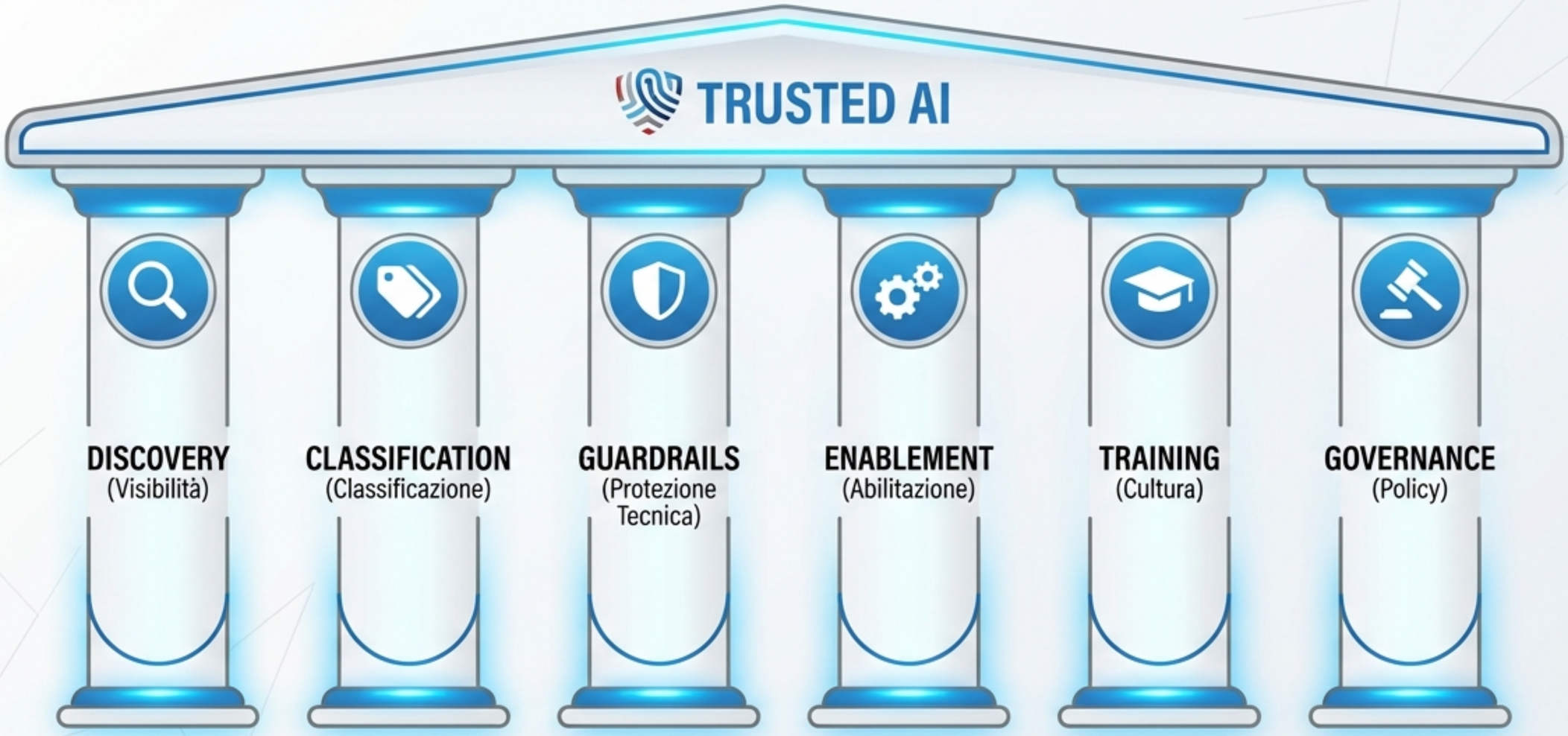


IL PARADOSSO DEL DIVIETO

- Se l'azienda blocca ChatGPT sulla rete, i dipendenti passano ai dispositivi personali (BYOD).
- Conseguenza: Perdita totale di visibilità (Zero Logs).
- Dato: 43% dei dipendenti usa app IA su dispositivi personali per lavorare." (Fonte: CIO/1Password)

Non si può governare ciò che non si vede.
Il blocco sposta il problema, non lo risolve.

DAL BUIO ALLA LUCE: UNA STRATEGIA IN 6 PILASTRI



BigID

VISIBILITÀ: VEDERE L'INVISIBILE

Tecniche per scoprire lo Shadow AI nella rete aziendale.

1. **Analisi Traffico Web (CASB/SWG):** Monitorare connessioni verso domini AI noti (OpenAI, Hugging Face).
2. **Scansione Estensioni Browser:** Rilevare plugin che leggono il contenuto dello schermo.
3. **Analisi Log SSO:** Verificare accessi tramite credenziali aziendali su piattaforme terze.
4. **Sondaggi Anonimi:** Chiedere ai dipendenti cosa usano (approccio non punitivo).

Tech Stack: Strumenti di 'SaaS Discovery' e 'AI-SPM' sono essenziali.

CLASSIFICAZIONE: IL SISTEMA A SEMAFORO

VERDE - APPROVATO

Istanze Enterprise (es. Copilot protetto, Sandbox interna). Sicuro per dati interni.

GIALLO - LIMITATO

Tool pubblici permessi solo per dati non sensibili (es. brainstorming). Vietato l'uso di PII o Codice.

ROSSO - VIETATO

Strumenti non verificati, malware noti, o piattaforme che rivendicano la proprietà dell'input. Blocco totale.

BigID

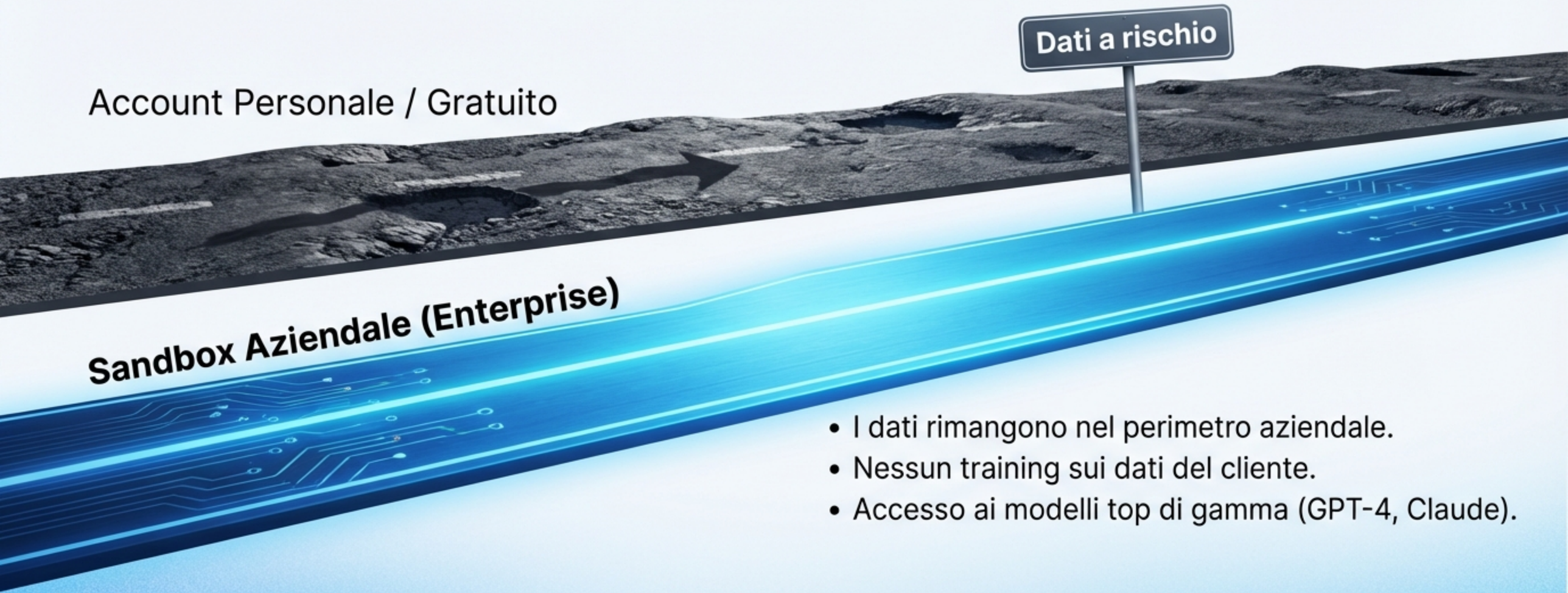
GUARDRAIL TECNICI: AUTOMAZIONE DELLA SICUREZZA

- **DLP (Data Loss Prevention):** Blocca carte di credito e codice sorgente.
- **Browser Isolation:** Isola le sessioni web rischiose.

- **Oscuramento Dati:** Anonimizzazione automatica.
- **Controlli API:** Monitoraggio integrazioni codice.

ENABLEMENT: L'ALTERNATIVA SICURA

La miglior difesa è fornire strumenti migliori.



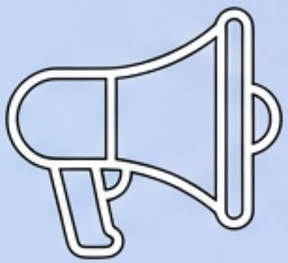
- I dati rimangono nel perimetro aziendale.
- Nessun training sui dati del cliente.
- Accesso ai modelli top di gamma (GPT-4, Claude).

I dipendenti scelgono la via di minor resistenza. Fornite un'alternativa sicura.

IL FATTORE UMANO: CULTURA E FORMAZIONE

Formazione Ruolo-Specifica

Gli sviluppatori hanno rischi diversi (codice) rispetto all'HR (dati personali). Formazione mirata.



Politica di Amnistia

Incoraggiare la segnalazione di tool in uso senza timore di punizioni. Mappare il reale utilizzo.

AI Champions

Nominare responsabili in ogni dipartimento per guidare l'adozione responsabile.

Policy Update: Aggiornare il Codice Etico e la AUP (Acceptable Use Policy).

L'ORIZZONTE NORMATIVO: GDPR E AI ACT

GDPR:
L'uso di Shadow AI impedisce la gestione dei diritti (es. diritto all'oblio). Se non sai dove sono i dati, non puoi cancellarli.

EU AI ACT:
Obblighi di trasparenza e gestione del rischio. La Shadow AI rende impossibile la conformità.

IMPATTO ECONOMICO:
I costi di un breach da Shadow AI sono significativamente più alti (**+670k USD** in media).

TRASFORMARE L'OMBRA IN VANTAGGIO COMPETITIVO

Raccomandazioni:

1. **Non ignorare:** Lo Shadow AI è già qui.
2. **Non solo bloccare:** Abilita alternative sicure.
3. **Governare:** Visibilità continua e policy dinamiche.

Vincenzo Calabro'
Information Security Officer & Digital Forensics Analyst | Researcher | Trainer

La sicurezza non deve frenare l'innovazione, deve renderla sostenibile.