Hup!

Inquadra il codice: sito, vetrina e documenti

Technology Assessment

Hup! · mentore AI, giudice AI a gettoni, vetrina con ricerca AI

Versione 1 · 29 novembre 2026. Valutazione della tecnologia sul modello delle valutazioni HTA: che cos'è, quanto è matura, quanto ci si può fidare, quanto costa, quali rischi porta e come la verifichiamo. Ogni dato ha la fonte tra parentesi; le etichette sono le stesse del business plan.

Conclusione in breve. La tecnologia è semplice, economica e già in parte in uso: il mentore AI è in produzione, vetrina, ricerca AI e giudice sono costruiti e provati in locale. Il punto debole non è tecnico ma di evidenza: non sappiamo ancora quanto il giudizio dell'AI coincida con quello dei giudici umani sui nostri pitch. Per questo il giudice AI allena ma non decide il livello, finché una misura su almeno 100 pitch non dice il contrario [DEDUZIONE].

1. Che cosa valutiamo

Componente	Che cosa fa	Stato al 29/11/2026
Sito e app (PWA)	Pagine statiche su Cloudflare Pages, installabili sul telefono; registrazione audio nel browser	In linea (hup-9t8.pages.dev), v12 pronta
Mentore AI (COACH-v0.7)	Chat educativa: riconosce la fase (6 fasi), cerca nei cinque volumi, poi sul web; legge PDF e audio del pitch	In produzione con il system message v0.3; v0.7 pronta
Giudice AI a gettoni (GIU-001)	Report dettagliato del pitch con verdetto; 1 gettone = 1 €, 3 di prova; registra l'allenamento senza cambiare il livello	Costruito e provato in locale
Vetrina e scheda (VET-001, VET-002)	Testo o audio → scheda scritta dall'AI → codice personale → conferma email → vetrina	Costruito e provato in locale (32 schede di esempio)
Ricerca AI (MATCH-001)	Il database propone fino a 40 candidati, l'AI sceglie e spiega i 5 migliori	Costruito e provato in locale
Livelli	Regola nel database: sale solo con «promosso» davanti a un giudice umano; il canale AI non sale mai	Provato in locale (trigger SQL)

Architettura

Pagina del sito → webhook n8n → (trascrizione audio OpenAI) → agente con modello Gemini 3.5 Flash e riserva gpt-5-mini → strumento «biblioteca» su Pinecone (indice asd, namespace hopup, embeddings Gemini a 3.072 dimensioni) e «cercatore web» → risposta alla pagina → conteggi senza testo nella tabella hup_eventi su Supabase. Schede, tentativi, gettoni e ricariche stanno su Supabase (Postgres) con sicurezza a livello di riga attiva; la chiave di servizio resta solo in n8n.

2. Maturità (scala TRL)

Componente	TRL stimato	Perché
Sito e app	7	In uso pubblico, senza incidenti noti
Mentore AI	6	In produzione, pochi utenti, nessuna misura di qualità
Giudice AI	4–5	Provato in laboratorio con dati finti
Vetrina e ricerca AI	4–5	Provata in laboratorio su 32 schede di esempio
Regola dei livelli	5	Provata su database di prova, logica semplice

Scala TRL della Commissione europea (Horizon 2020, allegato G) [FATTO]; i livelli assegnati sono una nostra stima [STIMA · peso 60%].

3. Efficacia: che cosa dice la letteratura

Domanda	Evidenza (fonte)	Forza	Che cosa ne facciamo
Allenarsi a presentare migliora il pitch?	Esperimento in 4 gare: +1/3 di deviazione standard nelle buone pratiche (Clingingsmith e Shane, Management Science, 2018)	Buona: esperimento con assegnazione casuale	Base del servizio
Il riscontro aiuta a imparare?	d = 0,48 su 435 studi, più forte su abilità cognitive e motorie (Wisniewski, Zierer e Hattie, 2020)	Buona: meta-analisi	Report con ciò che va cambiato, non solo un voto
Quanto conta la pratica?	14% della varianza (Macnamara et al. 2014); 29–61% con criteri stretti (Ericsson e Harwell 2019); 1% tra gli élite nello sport (Macnamara et al. 2016)	Discussa	Promettere miglioramento, non successo
L'AI sa valutare un pitch?	85% di accuratezza e r = 0,896 con valutazioni umane su 600 pitch (arXiv 2507.03721, 07/2025, preprint); 52% di correlazione con VC (Csaszar et al. 2024); precisione pari ai VC nel backtesting (arXiv 2603.12274, 2026)	Moderata: in parte preprint	Il giudice AI allena e fa da primo filtro
Ci sono distorsioni?	L'accuratezza apparente dipende anche da ciò che il modello sa già delle aziende (Kleinert et al., data non indicata)	Moderata	Nessuna decisione sul livello; misura sul campo

4. Sicurezza e rischi della tecnologia

Rischio	Probabilità e impatto	Contromisura già presente	Da fare
Risposte inventate (allucinazioni)	Media · medio	Libri prima del web; istruzione «non inventare»; fonti e date richieste	Controllo a campione di 30 risposte al mese
Biblioteca configurata male	Alta · alto (trovato il 29/11: la descrizione dello strumento era quella di un altro chatbot)	Descrizione corretta fornita (COACH-012)	Verifica sull'export del workflow
Contenuti dannosi	Bassa · alto	Frasi di pericolo con risposta fissa e numeri di aiuto	Filtri Gemini a «medio e oltre» invece di BLOCK_NONE (COACH-014)
Istruzioni nascoste nei testi caricati (prompt injection)	Media · medio	Testi trattati come dati; nessuna azione esterna possibile dall'agente	Prova con 10 testi trappola prima dell'apertura
Verdetto ingiusto del giudice AI	Media · medio	Non cambia il livello; gettone scalato solo se c'è il report	Misura dell'accordo con giudici umani
Attacchi ai webhook	Media · basso	Origine consentita solo dal sito; chiave anti-traffico; doppi apici nelle query	Limite di richieste per minuto in n8n o Cloudflare
Dipendenza dai fornitori	Media · medio	Modello di riserva; dati nel nostro Postgres	Piano di uscita: esportazione schede in CSV, indice ricostruibile dai PDF

5. Dati personali e regole

- **Minimizzazione.** Nella vetrina: progetto, descrizione, ruoli, area; mai nomi, email o dati particolari. Le statistiche non contengono testi; l'identificativo del browser è reso irriconoscibile (sha256 con sale) e si cancella dopo 13 mesi.
- **Consenso e conferma.** La scheda è pubblica solo con consenso esplicito e dopo il clic sulla mail; senza conferma si cancella dopo 30 giorni.
- **Trasparenza dell'AI.** Per i sistemi che interagiscono con le persone va detto che si parla con un'AI (Regolamento UE 2024/1689, art. 50) [FATTO]: lo dicono la pagina d'ingresso del mentore e la pagina del giudice.
- **Nessun dispositivo medico, nessun minore.** Servizio riservato a chi ha almeno 18 anni; non tratta dati sanitari.
- **Trasferimenti fuori dall'UE.** OpenAI e Google (modelli e embeddings), Pinecone (indice; regione dei server da verificare): basi del trasferimento da documentare fornitore per fornitore [DA VERIFICARE].
- **Precedenti del Garante.** Sanzione di 5 milioni di euro a Luka Inc. per il chatbot Replika (Garante privacy, provvedimento n. 232 del 10/04/2025, comunicato del 19/05/2025) [FATTO]: basi giuridiche non individuate, informativa inadeguata e verifica dell'età insufficiente (artt. 5, 6, 12, 13, 24, 25 GDPR) sono i punti da non ripetere.

6. Costi unitari e scalabilità

Voce	Costo (fonte)
Trascrizione di 1 minuto di pitch	0,003 $ con gpt-4o-mini-transcribe (listino OpenAI, 29/11/2026) [FATTO]
Modello per un report (circa 20.000 token letti e 1.500 scritti)	circa 0,008 $ con gpt-5-mini a 0,25 $/M e 2 $/M (listino OpenAI, 29/11/2026) [DEDUZIONE]
Costo prudenziale per prova del giudice	0,05 € (modello v3, con ricerca e memoria) [DEDUZIONE]
Pagamento con Stripe	1,5% + 0,25 €: 26,5% su 1 €, 6,5% su 5 € (stripe.com/it/pricing, 29/11/2026) [FATTO]
Tecnologia fissa nel 2027	3.000 € l'anno nello scenario base (modello v3) [IPOTESI]

Scalabilità. I costi variabili crescono con le prove, non con le persone: il margine per prova resta intorno a 0,89 € con ricariche da 5 gettoni (modello v3) [DEDUZIONE]. Il collo di bottiglia è umano: i giudici delle tappe e delle sessioni online.

7. Piano di verifica

Che cosa misuriamo	Come	Soglia per andare avanti
Accordo tra giudice AI e giudici umani	100 pitch giudicati da entrambi alle tappe di dicembre e gennaio; correlazione e kappa di Cohen sul verdetto	r ≥ 0,6 e kappa ≥ 0,4 per usare l'AI come primo filtro [IPOTESI]
Utilità del report	Domanda dopo il report: «Sai che cosa cambiare?»; seconda prova entro 7 giorni	Almeno 60% sì; punteggio medio in salita alla seconda prova [IPOTESI]
Disponibilità a pagare	Quante schede comprano gettoni dopo i 3 di prova	Almeno 10% entro marzo 2027 [IPOTESI]
Qualità della ricerca AI	Su 20 richieste di prova: quante delle 5 proposte sono giudicate pertinenti	Almeno 3 su 5 [IPOTESI]
Sicurezza	10 testi trappola e 10 frasi di pericolo	Zero istruzioni eseguite; risposta di aiuto nel 100% dei casi

8. Azioni aperte

Codice	Azione	Stato
COACH-012	Descrizione corretta dello strumento biblioteca nel workflow in produzione	Da fare (testo fornito)
COACH-013	Query delle statistiche in modalità espressione	Da fare
COACH-014	Filtri di sicurezza Gemini a «medio e oltre»	Da fare
COACH-015	Caricare tutti e cinque i volumi con i metadati (carica libri v0.5)	Da verificare il numero di vettori
VET-001, VET-002, MATCH-001, GIU-001	Importare, collegare le credenziali, attivare e mettere gli indirizzi in config.js	Da fare

Codice	Azione	Stato
SQL v0.4	Lanciare su Supabase: vetrina, tentativi, livelli, gettoni, ricariche	Da fare

Fonti

Fonte	Documento e data
Clingingsmith e Shane	Management Science 64(11), 2018
Wisniewski, Zierer e Hattie	Frontiers in Psychology, 22/01/2020
Macnamara, Hambrick e Oswald	Psychological Science, 2014
Macnamara, Moreau e Hambrick	Perspectives on Psychological Science, 2016
Ericsson e Harwell	Frontiers in Psychology, 25/10/2019
arXiv 2507.03721	preprint, 04/07/2025
arXiv 2602.01684	preprint, 2026 (cita Csaszar et al. 2024)
Bae et al.	DIALECTIC, arXiv 2603.12274, 2026
Kleinert et al.	valutazione di nuove imprese con LLM, data non indicata
Commissione europea	Horizon 2020, allegato G, livelli di maturità tecnologica
Regolamento UE 2024/1689	AI Act, art. 50
Garante privacy	provvedimento n. 232 del 10/04/2025 su Luka Inc. (Replika), comunicato del 19/05/2025
OpenAI	developers.openai.com/api/docs/pricing, 29/11/2026
Stripe	stripe.com/it/pricing, 29/11/2026
Hup!	workflow n8n, SQL v0.4 e prove locali del 29/11/2026