FRANCESCO SAVERIO CANEPA · WORKING PAPER
PAPER DI METODI E PROTOCOLLO

Dissenso calibrato: un protocollo verificabile di sigillo-e-risoluzione per l'audit di previsioni AI avversariali, con un pilot pubblico su operazioni M&A (N = 10)

Un protocollo anti-senno-di-poi verificabile con un piano di analisi pre-specificato, istanziato su un pilot pubblico sigillato (N = 10); nessun risultato.

di

Francesco Saverio Canepa

DRAFT PRE-REGISTRAZIONE · 25 GIUGNO 2026

Indice

Sintesi

I sistemi AI ottimizzati per l'utilità falliscono sistematicamente nelle decisioni consequenziali: setacciano le evidenze alla ricerca di conferme, restituiscono validazioni articolate e consegnano esattamente le rassicurazioni di cui il ragionamento ad alto rischio ha meno bisogno. Il *contraddittore cognitivo* è l'alternativa — un'AI strutturalmente avversariale il cui mandato è sfidare anziché confermare. Se un tale sistema produca *dissenso calibrato* — disaccordo con il consenso che è accurato anziché meramente più rumoroso — è una domanda empirica a cui il campo non ha alcun modo onesto di rispondere, perché valutare un previsore dopo che gli esiti sono noti invita al senno di poi. **Questo paper contribuisce uno strumento verificabile: un protocollo verificabile e anti-senno-di-poi per l'audit di previsioni AI avversariali, e un pilot pubblico sigillato prima dell'esito su decisioni reali e consequenziali — i passi di sigillo e ancoraggio sono dimostrati; la risoluzione e lo scoring attendono gli orizzonti di esito.**

Il protocollo sigilla ogni previsione in un registro pubblico append-only e ancorato tramite hash (l'Osservatorio) prima che qualsiasi esito sia osservabile; la risolve mediante una regola pienamente deterministica su fonti pubbliche canoniche; la confronta con un nullo del tasso-base abbinato alla coorte contestata e con un nullo del segnale pubblico; pubblica un trail di evidenze di risoluzione pienamente auditabile (e, ove fattibile, doppia-codifica con un codificatore cieco indipendente); e congela tramite SHA-256 l'esatto bundle modello–prompt–Pack. Lo istanziamo su un **pilot fisso, pubblicamente pre-sigillato di N = 10 previsioni M&A europee ad alto rischio** (sigillate a T0 = 20 giugno 2026), ciascuna una distribuzione di scenario a tre vie da cui sono derivate P(H1) (deal-break o rinegoziazione al ribasso ≥ 15%, 6–12 mesi) e P(H2) (distruzione di valore post-closing, 18–36 mesi, sul sottoinsieme con un acquirente quotato che consolida). Le quantità sottoposte a scoring sono queste probabilità del **previsore** sigillate — il giudizio di lettura dominante espresso attraverso il processo contraddittore; l'intero record di sfida a tre passate del motore (lo schema a sette campi, il Δ-CSI e il bundle hash) entra come **sola provenienza** e non può alterarle. Il pilot fa quindi l'audit delle previsioni sigillate, **non** della loro generazione da parte del motore, e riporta il registro grezzo previsione-contro-esito man mano che gli esiti si risolvono. Al momento del sigillo, ogni scheda pubblica registrava la distribuzione di scenario a tre vie, la tesi di lettura dominante, un livello di confidenza calibrata e le fonti citate; l'intero payload a sette campi del motore, il Δ-CSI e il bundle hash **non esistono ancora per i dieci deal** e saranno generati successivamente come provenienza non sottoposta a scoring (§4.3). Questo pilot è quindi una dimostrazione

del dispositivo di sigillo-e-risoluzione, non una valutazione della generazione del motore stesso. Il motore contraddittore opera attraverso una pipeline a tre passate con isolamento dei modelli (Tesi, Red Team, Sintesi), uno schema di output fisso a sette campi e un'architettura fail-closed progettata per prevenire la deriva compiacente.

La calibrazione (H-cal) e la divergenza-che-paga (H-pay), con le due baseline nulle e i criteri di falsificazione pre-dichiarati, costituiscono un **piano di analisi pre-specificato per una futura fase potenziata** su una coorte più ampia, sigillata in avanti; al N = 10 fisso del pilot — al di sotto della soglia di potenza di N ≥ 20 per orizzonte — non è eseguito alcun test confermativo, e ogni quantità è riportata descrittivamente. Ne derivano tre contributi: un framework concettuale che distingue il dissenso calibrato dalla validazione compiacente; il protocollo verificabile di sigillo-e-risoluzione in sé; e un'implementazione aperta con un pilot pubblico e immutabile che consente la replicazione indipendente. Il paper **non** afferma che il contraddittore migliori le decisioni, che le previsioni del motore stesso siano calibrate (un'affermazione della fase futura), che il Δ-CSI predica la correttezza, né che le risultanze si generalizzino oltre l'M&A. Il Δ-CSI (Cognitive Sovereignty Index delta) è un proxy calcolato dell'intensità della sfida, non della correttezza — un invariante progettuale esplicito. Le dimostrazioni della corsia C sono non-evidenza per qualsiasi affermazione pre-registrata.

SEZIONE 1

Introduzione

Ogni decisione consequenziale poggia su un modello del mondo. Quel modello viene raramente messo alla prova prima di essere tradotto in azione.

La tendenza umana a cercare conferme piuttosto che falsificazioni delle ipotesi preferite — documentata sistematicamente da Kahneman (2011) — è amplificata, non corretta, da assistenti AI addestrati prevalentemente a produrre output utili e concordanti. Quando il compito è valutare una fusione o un'acquisizione, un'AI ottimizzata per l'accordo setaccia il memorandum dell'analista alla ricerca di prove a sostegno, restituisce una prosa articolata che rafforza la tesi e consegna esattamente la validazione di cui un team sotto pressione ha meno bisogno. La compiacenza non è un difetto caratteriale; è un esito progettuale. Le decisioni ad alto rischio richiedono l'opposto.

Il concetto alternativo esaminato in questo lavoro è il *contraddittore cognitivo*: un sistema AI il cui mandato non è rispondere ma sfidare, non concordare ma stress-testare una decisione nel momento in cui si sta formando. Il compito del contraddittore è portare alla superficie le assunzioni implicite, costruire scenari controintuitivi e restituire incertezza calibrata — il tutto prima dell'impegno. Non si

tratta di uno strumento di critica aggiunto in fondo a un report; è un agente avversariale inserito nel processo deliberativo nel punto decisionale stesso, con il proprio output registrato, marcato temporalmente e confrontato con gli esiti successivi.

La genealogia intellettuale della sfida avversariale strutturata è consolidata. Schwenk (1990) ha esaminato tre decenni di evidenze sperimentali e sul campo dimostrando che il devil's advocacy e l'inquiry dialettico — metodi prescritti per imporre il contro-argomento esplicito — miglioravano la qualità dei piani e riducevano il groupthink nelle decisioni strategiche. Quella letteratura precede interamente i modelli linguistici di grandi dimensioni. La domanda se un LLM possa operativizzare il devil's advocacy su scala, in modo affidabile e consistente, è rimasta aperta fino a tempi recenti. Un preprint del 2024 sull'«antagonistic AI» ha formalizzato lo spazio di progettazione per agenti AI la cui funzione di utilità è esplicitamente oppositiva (Antagonistic AI, arXiv:2402.07350; Cai, Arawjo, e Glassman, 2024). Due preprint del 2025 hanno interrogato se tali sistemi riducano effettivamente l'overconfidence o producano soltanto un registro superficialmente scettico senza influenzare l'aggiornamento delle credenze (skeptic2025a, arXiv:2509.05396; Wynn, Satija, e Hadfield, 2025; skeptic2025b, arXiv:2511.07784; Wu, Li, e Li, 2025). In tutta questa letteratura emergente persiste una lacuna progettuale: nessuno studio ha sottoposto un contraddittore AI a un test sul campo pre-registrato, con scoring di calibrazione, su decisioni reali e consequenziali i cui esiti siano verificabili. Gli studi di laboratorio controllati usano stimoli ipotetici; le analisi osservazionali mancano di una baseline pre-registrata rispetto alla quale misurare la divergenza; e non esiste alcun protocollo per quantificare il *dissenso calibrato* — il grado in cui il disaccordo del contraddittore rispetto al consenso è accurato piuttosto che semplicemente più rumoroso. I tornei di forecasting e i prediction market — Good Judgment Open (Tetlock e Gardner, 2015), Metaculus (Metaculus, Inc., 2015) — già sigillano e fanno scoring di previsioni probabilistiche, ma nessuno fa l'audit del dissenso di un'AI *strutturalmente avversariale* rispetto a un consenso di dominio a fronte di un nullo del tasso-base abbinato; è quello lo strumento specifico che questo protocollo costruisce.

> *Non esiste alcun protocollo per quantificare il dissenso calibrato — il grado in cui il disaccordo del contraddittore rispetto al consenso è accurato piuttosto che semplicemente più rumoroso.*

Questo paper affronta quella lacuna non rivendicando una risposta ma costruendo lo strumento che permetterebbe di guadagnarne una onestamente. Contribuiamo un

protocollo verificabile di sigillo-e-risoluzione per l'audit di previsioni AI avversariali — sigillo pre-esito con un'ancora hash pubblica, una regola di risoluzione pienamente deterministica su fonti pubbliche, un nullo del tasso-base abbinato alla coorte, un trail di evidenze di risoluzione pubblico e auditabile (con doppia codifica cieca raccomandata ove fattibile) e un freeze di versione per chiamata tramite bundle hash — e lo istanziamo su un **pilot** fisso, pubblicamente pre-sigillato di N = 10 previsioni sigillate di lettura dominante su fusioni e acquisizioni ad alto rischio. L'affermazione più ampia che il programma di ricerca verificherà eventualmente è che il contraddittore produce previsioni calibrate che divergono dal consenso e battono una baseline pre-registrata; **questo paper non giudica quell'affermazione e non testa la generazione del motore**. Esso consegna il protocollo e il pilot pubblico, e riporta descrittivamente il registro grezzo previsione-contro-esito del pilot man mano che gli esiti si risolvono, riservando il test confermativo potenziato a una fase futura su una coorte più ampia, sigillata in avanti (§6.5). Scegliamo l'M&A perché i verdetti sui deal sono binari, verificabili, e portano posta in gioco sufficiente a escludere il disaccordo performativo. Misuriamo la calibrazione — non l'accuratezza isolata — perché un challenger meramente contrarian è privo di valore; quello le cui stime di incertezza tracciano le frequenze empiriche è informativo.

Questo paper offre tre contributi. In primo luogo, un framework concettuale che distingue il dissenso calibrato dalla validazione compiacente e dallo scetticismo non strutturato, e che formalizza lo schema di output del contraddittore (assunzioni implicite, scenario controintuitivo, test di falsificazione, domande sull'onere della prova, confidenza calibrata, fonti citate e un Δ-CSI — Cognitive Sovereignty Index delta — come proxy dell'intensità della sfida per ciascuna sessione). In secondo luogo, e centralmente, il **protocollo verificabile di sigillo-e-risoluzione** in sé — il dispositivo anti-senno-di-poi che rende possibile un test onesto in seguito — insieme al piano di analisi pre-specificato (regola di risoluzione, baseline nulle abbinate e criteri di falsificazione) che una fase potenziata applicherebbe. In terzo luogo, un'implementazione aperta e un pilot pubblico e immutabile che consente la replicazione indipendente e già dimostra il **fronte di sigillo-e-ancoraggio** del dispositivo su decisioni reali; la sua metà di risoluzione-e-scoring attende le finestre di esito (6–36 mesi) ed è esercitata dalla futura fase potenziata. Rendiamo esplicito ciò che questo paper non afferma: non affermiamo che il contraddittore migliori le decisioni, che le previsioni del motore stesso siano calibrate, che il Δ-CSI predica la correttezza, né che i risultati si generalizzeranno oltre il dominio M&A. Quelle sono affermazioni per una futura fase potenziata, contingenti ai suoi risultati; questo paper è il protocollo e il pilot, non il verdetto.

SEZIONE 2

Quadro teorico

2.1 — IL CONTRADDITTORE, NON L'ORACOLO

La scelta progettuale centrale del sistema in esame è architettonica, non parametrica: il mandato del motore è sfidare una decisione, mai validarla. Chiamiamo questo il *principio invariante del contraddittore*. Dove un'AI di tipo oracolare presenta una risposta raccomandata, un contraddittore cognitivo porta alla superficie ciò che il decisore non ha considerato, ciò che potrebbe falsificare la tesi e gli obblighi epistemici rimasti inadempiuti. La compiacenza non è trattata come una modalità degradata ma come una modalità di guasto — esattamente analoga a un errore di tipo II nel test statistico. Un sistema che restituisce un'analisi rassicurante senza una sfida genuina è difettoso, indipendentemente dall'articolazione dell'output.

AI oracolo vs. contraddittore cognitivo

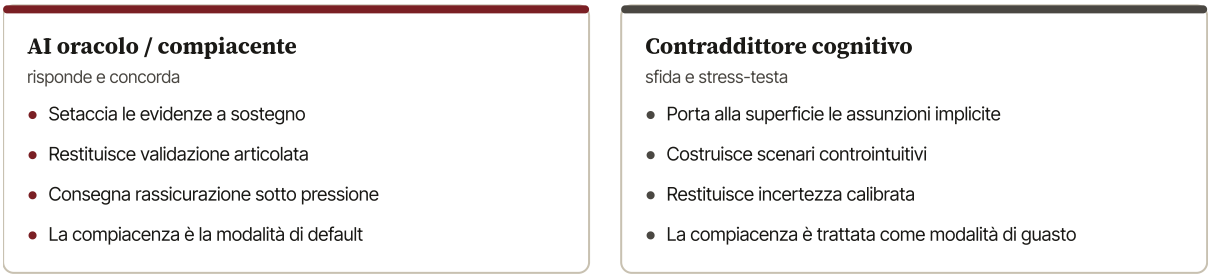

Figura 1 — Due posture progettuali. L'AI oracolo / compiacente restituisce validazione; il contraddittore cognitivo restituisce sfida strutturata. Concettuale, tratta da §1 e §2.1; nessun valore di dati.

Questa distinzione ha conseguenze metodologiche. Il devil's advocacy, come esaminato da Schwenk (1990), è più efficace quando il contro-argomento è *richiesto* istituzionalmente, non semplicemente consentito. Il contraddittore applica quel requisito strutturalmente: lo schema di output è fisso e non negoziabile, e la pipeline è fail-closed — solleva un errore esplicito piuttosto che emettere silenziosamente una risposta non sfidante. Nessuna opzione di configurazione attenua la sfida obbligatoria. La specializzazione di dominio è espressa attraverso file di configurazione esterni (BRAIN Pack) che possono alterare il vocabolario, la persona e l'enfasi dell'agente red-team, ma non possono rimuovere né rinominare alcun elemento del contratto di output.

2.2 — LO SCHEMA DI OUTPUT FISSO

Ogni analisi prodotta dal contraddittore deve essere conforme a uno schema JSON versionato (`schema_version: "1.0"`, `additionalProperties: false`). Lo schema definisce sette campi di output sostanziali, ciascuno obbligatorio:

1. **Assunzioni implicite** (`assumptions[]`): le premesse incorporate nella decisione che il decisore non ha dichiarato esplicitamente. Ogni assunzione porta un rating di severità su scala 1–5 e un'etichetta di categoria.

2. **Scenario controintuitivo** (`counter_scenario`): una narrativa che descrive un esito plausibile che la lettura dominante non anticipa — lo scenario che, se si materializzasse, danneggerebbe maggiormente il valore della decisione.

3. **Test di falsificazione** (`falsification_tests[]`): verifiche empiriche operativizzate che, se i risultati andassero nella direzione specificata, minerebbero la tesi. Ogni test identifica i dati necessari per eseguirlo e ne traccia lo stato corrente.

4. **Domande sull'onere della prova** (`burden_questions[]`): domande esplicite su chi porta l'obbligo epistemico di dimostrare un'assunzione chiave, e se quell'obbligo è stato adempiuto.

5. **Confidenza calibrata** (`confidence`): un rating di confidenza a tre livelli (low / medium / high) accompagnato da una motivazione e da un elenco di gap di conoscenza dichiarati — ciò che il sistema non sa e non può imputare.

6. **Fonti citate** (`sources[]`): ogni fonte citata nell'analisi, ciascuna con un tipo di provenienza tratto da un enum chiuso (`web_search | internal_ledger | internal_doc | osint_module | user_provided`). Se non è disponibile alcun retrieval, il campo deve essere vuoto; al sistema è vietato fabbricare citazioni.

7. **Δ-CSI** (`csi_delta`): il Cognitive Sovereignty Index delta, definito nel §2.3 di seguito.

Questo schema è un invariante: nessuna specializzazione verticale o configurazione cliente può allentare, rimuovere o rinominare nessuno di questi campi. Il contratto è definito una volta e versionato; un cambiamento incompatibile richiede un esplicito bump di versione e una nuova coorte di calibrazione.

2.3 — IL Δ-CSI: UN PROXY DELL'INTENSITÀ DELLA SFIDA

Il Δ-CSI (Cognitive Sovereignty Index, delta) è un indice calcolato, persistito e calibrato nell'intervallo [0, 100]. Misura con quanta forza il processo contraddittore ha messo sotto pressione la decisione — cioè l'intensità della sfida applicata in una data sessione.

> **INVARIANTE PROGETTUALE**

> **Il Δ-CSI non è una misura della qualità della decisione, e non predice se la decisione è corretta.** Questo vincolo di interpretazione è un invariante progettuale esplicito, dichiarato prominentemente nella metodologia del sistema e in questo protocollo, perché il nome e la scala 0–100 invitano alla lettura errata. Un Δ-CSI alto significa che il contraddittore ha premuto con forza su molte dimensioni con assunzioni ad alta severità; non dice nulla sul fatto che la decisione sfidata abbia successo o meno. Trattare il Δ-CSI come un segnale di accuratezza è un errore di categoria; pre-registriamo questa precisazione per prevenire interpretazioni errate nella Fase 2.

Operativamente, il Δ-CSI (`csi-v1`) è calcolato deterministicamente da cinque componenti pesate: (i) conteggio delle assunzioni implicite, modulato dalla loro severità media; (ii) conteggio dei test di falsificazione; (iii) conteggio delle domande sull'onere della prova; (iv) conteggio delle fonti citate; e (v) un punteggio di divergenza tra la tesi e lo scenario controintuitivo. I componenti (i)–(iv) usano un conteggio limitato normalizzato in [0, 1] con un tetto di cinque elementi (`min(n, 5)/5`), che impedisce il gaming tramite verbosità. Il componente di divergenza (v) è calcolato come distanza del coseno tra vettori di embedding semantico — normalizzata in [0, 1] come (1 − similarità del coseno)/2 (quando è disponibile un provider di embedding) — o come complemento di Jaccard sui token normalizzati (fallback euristico). Il metodo di divergenza usato in ogni sessione è registrato nell'output insieme al punteggio. I pesi dei componenti sommano a 100; le configurazioni verticali possono aggiustare i pesi, ma il vincolo di somma è applicato programmaticamente.

Il Δ-CSI è persistito nel Sovereign Decision Ledger per ogni sessione. La calibrazione — il confronto empirico delle distribuzioni Δ-CSI rispetto agli esiti decisionali eventualmente registrati — diventa possibile solo dopo che almeno cinque sessioni con esiti registrati si sono accumulate. Al di sotto di quella soglia, il sistema contrassegna esplicitamente le proprie statistiche come solo descrittive, senza alcuna affermazione di calibrazione. Questo previene asserzioni premature di validità. Questo gate a livello di motore (≥ 5 sessioni) è una salvaguardia operativa di prodotto, **indipendente da e ben più debole della** soglia di potenza confermativa dello studio (N ≥ 20 per orizzonte sottoposto a scoring; §6.5); superarlo non autorizza alcuna affermazione di calibrazione nel senso del §6, e la presente coorte fissa si situa al di sotto della soglia dello studio per scelta progettuale.

2.4 — LA PIPELINE A TRE PASSATE E LE DIFESE FAIL-CLOSED

Il contraddittore opera attraverso tre passate sequenzialmente isolate, ognuna eseguita da un agente LLM con un mandato distinto e non sovrapposto.

Pipeline del contraddittore a tre passate
sequenzialmente isolate, model-isolate per default

Figura 2 — La pipeline a tre passate con isolamento dei modelli. Etichette e sotto-etichette tratte da §2.4; processo concettuale, nessun valore di dati.

Passata 1 — Tesi. Il primo agente produce la *lettura dominante*: la conclusione che un analista attento trarrebbe dalle evidenze disponibili. A questo agente è esplicitamente vietato agire come contraddittore; il suo ruolo è articolare la tesi migliore come obiettivo per la sfida successiva. La tesi è passata, non modificata, alla Passata 2.

Passata 2 — Red Team. Il secondo agente riceve la tesi come oggetto da attaccare. Il suo system prompt è costruito dal Pack risolto, che fornisce una direttiva di persona red-team, una tassonomia di bias cognitivi specifica del dominio con contromisure prescritte, template di falsificazione calibrati sul dominio decisionale e guardrail operativi. L'output del red team è JSON strutturato che copre tutti i campi di sfida obbligatori (assunzioni, scenario controintuitivo, test di falsificazione, domande sull'onere della prova). Le passate sono *model-isolate* per default: tesi e red team girano su provider LLM diversi per de-correlare i blind-spot sistematici. Il provider di ogni passata è registrato nei metadati dell'output.

Passata 3 — Sintesi. Il terzo agente riceve il blocco decisionale, la tesi e il contraddittorio del red team, e ha il solo compito di calibrare la confidenza e dichiarare i gap di conoscenza. Non conclude, non raccomanda, e non arbitra tra la tesi e lo scenario controintuitivo; dichiara esplicitamente ciò che è ignoto.

Tre difese strutturali proteggono dalle modalità fail-open che renderebbero il contraddittore inutile:

- [ ] **Contradiction Floor.** Dopo la Passata 2, la pipeline verifica che la sfida sia non vuota in tutte e quattro le dimensioni (almeno un'assunzione, un test di falsificazione, una domanda sull'onere della prova e una narrativa controintuitiva non vuota). Se la verifica fallisce, la pipeline emette un retry con una direttiva esplicita a produrre una sfida non vuota. Se la verifica fallisce ancora dopo il retry, la pipeline solleva un errore di motore e si ferma. Nessun output parziale o vuoto viene emesso.

- [ ] **Provenienza a enum chiuso.** Le fonti devono dichiarare un tipo di provenienza dall'enum chiuso. Se non è disponibile alcuno strumento di retrieval, l'agente di sintesi è vincolato a restituire una lista di fonti vuota. L'enum chiuso rende impossibile una provenienza *mis-taggata* e vieta citazioni in assenza di retrieval,

eliminando il percorso di fabbricazione più comune; non garantisce però di per sé che un modello non emetta mai una fonte inesistente sotto un tag altrimenti valido, sicché l'esistenza della fonte resta una verifica a tempo di risoluzione.

- [ ] **Gate di calibrazione sotto-soglia.** Le statistiche di calibrazione sono calcolate solo quando il campione contiene almeno cinque sessioni con esiti registrati. Al di sotto di tale soglia, il campo di interpretazione è forzato a un descrittore solo prudente; non è avanzata alcuna validità empirica. (Questo gate a livello di prodotto è distinto da, e ben più debole della, soglia di potenza del §6.5.)

Insieme, questi invarianti — schema fisso, floor di sfida obbligatorio, provenienza chiusa, gate sotto-soglia e isolamento dei modelli — definiscono un sistema il cui output è, per progettazione, resistente alla deriva compiacente. Se quella resistenza strutturale si traduca in accuratezza previsionale calibrata è la domanda empirica che questo protocollo è progettato per rendere rispondibile, da una futura fase potenziata, senza senno di poi (§6.5).

SEZIONE 3

Piano di analisi pre-specificato (rinviato a una fase potenziata)

Il protocollo pre-specifica cinque quantità che una fase potenziata testerà e che il pilot riporta descrittivamente, distinguendo i bersagli pianificati-confermativi dalle caratterizzazioni descrittive.

Il protocollo pre-specifica cinque quantità che una **fase potenziata** testerà e che il **pilot** riporta descrittivamente. **H1** e **H2** sono il segnale previsionale del contraddittore a due orizzonti; **H-cal**, **H-div** e **H-pay** riguardano la calibrazione e il valore del dissenso. **H-cal** e **H-pay** sono i bersagli **pianificati-confermativi** della fase potenziata; **H-div** è descrittiva. Nessuna è un'ipotesi *di questo paper*: al N = 10 fisso del pilot (§4.1) non viene calcolata alcuna regola decisionale formale (§6.5), e ogni quantità è riportata soltanto come risultanza grezza previsione-contro-esito. Le regole decisionali enunciate di seguito definiscono ciò che la fase potenziata farà, non ciò che il pilot conclude — sono pre-registrate ora affinché una fase futura non possa calibrarle dopo aver visto gli esiti.

ETICHETTA	NOME E ORIZZONTE	STATO
H1	Previsione primaria, orizzonte 6–12 mesi	Test di orizzonte
H2	Previsione secondaria, orizzonte 18–36 mesi	Test di orizzonte

H-cal	Calibrazione rispetto agli esiti realizzati	Pianificata-confermativa (fase potenziata)
H-div	Divergenza dal consenso degli analisti a T0	Descrittiva
H-pay	Divergenza-che-paga vs. baseline del consenso	Pianificata-confermativa (fase potenziata)

H1 — PREVISIONE PRIMARIA (ORIZZONTE 6–12 MESI)

Alla finestra di risoluzione di 6–12 mesi, la previsione sigillata a T0 (il giudizio di lettura dominante del §4.2) assegna una stima di probabilità P(deal-break o rinegoziazione al ribasso ≥ 15%) che la fase potenziata testa per la sua misurabile superiorità rispetto al livello-chance pre-registrato nel sotto-campione di deal in cui diverge dal consenso degli analisti a T0. La soglia di rinegoziazione ≥ 15% è fissata operativamente nei Metodi (§4) e nella regola di risoluzione (§5); è espressa qui come criterio di esito realizzato, non come soglia di probabilità. Nella fase potenziata H1 è valutata mediante il Brier score (Brier, 1950) rispetto alla baseline pre-registrata (§6), con la scomposizione affidabilità/risoluzione/incertezza di Murphy (1973) riportata solo al di sopra della soglia di potenza; sul pilot è mostrata come righe grezze previsione-contro-esito (§6.5).

H2 — PREVISIONE SECONDARIA (ORIZZONTE 18–36 MESI)

Alla finestra di risoluzione di 18–36 mesi, la stima di probabilità sigillata del contraddittore P(distruzione di valore: svalutazione dell'avviamento o TSR dell'acquirente inferiore al proprio indice settoriale di 10 punti percentuali) è sottoposta ad audit rispetto agli esiti realizzati. **H2 è definita solo per il sottoinsieme pre-dichiarato di deal con un acquirente quotato che consolida il target** (§4.2; il sottoinsieme è fissato per riga nell'Appendice D); per i deal rimanenti — acquirenti di private equity, delisting, carve-out, joint venture e break-up multi-parte — H2 è *non applicabile per progettazione*, non censurata post-hoc. H2 è valutata con lo stesso protocollo Brier di H1 su quel sottoinsieme, con la finestra di risoluzione più lunga che costituisce un orizzonte indipendente piuttosto che una replica. Dato il piccolo sottoinsieme eleggibile (quattro deal; Appendice D) e il lungo orizzonte, è probabile che il campione H2 risolto del pilot scenda al di sotto della soglia di reporting di tre del §6.5; H2 sarà allora mostrata solo come righe grezze, non come sintesi.

H-CAL — CALIBRAZIONE (PIANIFICATA-CONFERMATIVA, FASE POTENZIATA)

Sull'insieme di tutte le previsioni risolte, le stime di probabilità del contraddittore saranno ben calibrate rispetto agli esiti realizzati, valutate tramite una curva di calibrazione il cui test formale e livello di significatività sono specificati in §6. H-cal è

pianificata-confermativa. Il mancato raggiungimento della soglia di calibrazione pre-registrata costituirebbe una falsificazione dell'affermazione di calibrazione (cfr. §7) — valutata formalmente solo su un campione che raggiunge la soglia di potenza (§6.5); su questa coorte fissa N = 10 H-cal è riportata descrittivamente.

H-DIV — DIVERGENZA (DESCRITTIVA)

Una frazione non trascurabile delle previsioni del contraddittore a T0 divergerà dal consenso contemporaneo degli analisti (rating o target price). "Non trascurabile" è operativizzato in §6 come tasso di divergenza minimo pre-registrato; la soglia è scelta per escludere il disaccordo da rumore consentendo il dissenso genuino. H-div è descrittiva: caratterizza la distribuzione delle previsioni nella coorte dello studio e non costituisce di per sé un test confermativo. La sua funzione è stabilire che nel campione esista una divergenza sufficiente per consentire una valutazione significativa di H-pay.

H-PAY — DIVERGENZA-CHE-PAGA (PIANIFICATA-CONFERMATIVA, FASE POTENZIATA)

Nel sotto-campione di previsioni identificate ai sensi di H-div come divergenti dal consenso degli analisti a T0, le stime di probabilità del contraddittore raggiungeranno un Brier score inferiore (migliore calibrazione) rispetto alla baseline del consenso alla soglia pre-registrata. H-pay è pianificata-confermativa. Essa punta all'affermazione epistemica specifica che la divergenza dal consenso non è mero rumore ma porta un segnale previsionale oltre a quello già contenuto nel consenso. Un fallimento di H-pay — previsioni divergenti che non ottengono un Brier score migliore del consenso — costituirebbe una falsificazione parziale di quell'affermazione (§7), ma solo su un campione che raggiunge la soglia di potenza (§6.5); su questa coorte fissa H-pay è soppressa interamente se il sotto-campione divergente è al di sotto della soglia del §6.3 — il che, a N = 10, quasi certamente è, sicché il pilot non è atteso riportare affatto H-pay. H-pay è definita sul **solo orizzonte H1**: il sottoinsieme H2 di quattro deal non può raggiungere la soglia del sotto-campione divergente.

> **VINCOLI DI SCOPE**
>
> Queste ipotesi non avanzano alcuna affermazione sul miglioramento delle decisioni, sui cambiamenti comportamentali del management o sugli esiti organizzativi. Il Δ-CSI è registrato per tutte le sessioni e verrà usato come covariata nell'analisi esplorativa (§6), ma nessuna ipotesi pre-registra una relazione tra Δ-CSI e accuratezza previsionale; tale relazione, se osservata, sarà riportata come esplorativa e trattata come materiale Fase 2. Le ipotesi sono circoscritte al dominio M&A come definito in §4; la generalizzazione ad altri domini decisionali non è affermata.

SEZIONE 4

Metodi

4.1 — UNIVERSO DI SELEZIONE: UNA COORTE FISSA E SIGILLATA

La coorte dello studio è un **insieme fisso e chiuso di dieci decisioni M&A ad alto rischio** (N = 10), pubblicamente sigillate nell'Osservatorio il 20 giugno 2026 — una singola data di sigillo che precede ogni esito H1 e H2 della coorte. I dieci deal coprono l'M&A europeo large-cap nei settori bancario, telecomunicazioni, difesa, assicurazioni, gestione del risparmio e industriale, con valori di transazione annunciati che vanno da circa 2,2 mld di € a circa 40 mld di €; almeno una parte di ciascun deal è quotata su un mercato regolamentato europeo (Euronext Milan, Euronext Paris, Euronext Amsterdam, London Stock Exchange, SIX Swiss Exchange, la borsa di Francoforte o Nasdaq Stockholm). L'intera coorte, con l'hash del record sigillato di ciascun deal e la sua mappatura scenario-ipotesi, è enumerata nell'**Appendice D** (`docs/paper/appendices/cohort-register.md`) ed è ispezionabile in modo indipendente sull'Osservatorio pubblico.

A differenza di una regola di inclusione generativa imperniata su una soglia dimensionale e una finestra di raccolta in avanti, questa coorte è stata **assemblata editorialmente** — come un insieme di transazioni europee di punta, contestate in modo contemporaneo — e poi **congelata**: una volta sigillata, nessun deal può essere aggiunto alla coorte né rimosso. Questa è una scelta progettuale deliberata con un costo onesto e una salvaguardia onesta. Il costo: l'assemblaggio editoriale è un giudizio di selezione, divulgato come conflitto di interesse in §8 (COI-3). La salvaguardia, che è ciò che rende il giudizio di selezione non fatale, è strutturale e verificabile: ogni deal è stato sigillato e ancorato tramite hash in un registro append-only, committato in git, **prima che alcuno dei suoi esiti H1 o H2 fosse osservabile**, e l'insieme è congelato. La selezione non può quindi essere stata motivata da esiti che ancora non esistevano, e nessun deal che si risolva sfavorevolmente può essere silenziosamente eliminato a posteriori — la catena di hash e la cronologia pubblica dei commit lo vietano. Il cherry-picking tramite aggiunta o rimozione post-hoc è strutturalmente impossibile; il cherry-picking al momento del sigillo è neutralizzato dal sigillo pre-esito e limitato dal nullo del tasso-base (§6.1, nullo (a)), rispetto al quale una coorte ricca di eventi non guadagna alcun credito gratuito. La garanzia anti-senno-di-poi di questo studio poggia sul sigillo pubblico, marcato temporalmente e immutabile (§4.3, §4.4, §8), non su una regola di inclusione meccanica.

La dimensione fissa della coorte di N = 10 è **al di sotto della soglia di potenza** di N ≥ 20 per orizzonte sottoposto a scoring adottata per il test di calibrazione confermativo (§6.5). Ciò è noto in anticipo e non è un incidente di campionamento: lo studio è quindi progettato e riportato come un **audit di calibrazione pre-registrato a scala di fattibilità**, in cui l'analisi di calibrazione è descrittiva (§6.5) e le regole decisionali

confermative del §6 sono calcolate solo se e quando un campione risolto raggiunge la soglia di potenza. L'inquadramento descrittivo è pre-dichiarato qui, non invocato post-hoc per salvare un risultato sottodimensionato. Un test confermativo potenziato richiederebbe l'estensione della coorte in una pre-registrazione Fase 2 separata; quell'estensione è fuori dallo scope di questo protocollo.

4.2 — UNITÀ DI PREDIZIONE

Per ciascun deal della coorte, una singola **previsione** è sigillata a **T0 — la data di sigillo della coorte (20 giugno 2026)** — coprendo simultaneamente due orizzonti di esito pre-registrati. La previsione sottoposta a scoring è il giudizio probabilistico di lettura dominante del contraddittore (una quantità del previsore), redatto attraverso il processo contraddittore e sigillato; l'intero payload a tre passate del motore è ricostruito successivamente come provenienza (§4.3) e non è mai la quantità sottoposta a scoring. T0 è la data di sigillo, non l'istante dell'annuncio pubblico di ciascun deal: i deal sono stati annunciati tra ottobre 2025 e giugno 2026, ma tutti gli esiti H1/H2 si risolvono 6–36 mesi dopo, sicché T0 segue deliberatamente l'annuncio. Si tratta di una deviazione divulgata da un T0 all'istante-dell'annuncio, e lo svantaggio informativo che essa comporta **non è uniforme**: il ritardo annuncio-sigillo va da pochi giorni (es. EQT–Intertek, annunciato il 18 giugno) a circa otto mesi (es. il MoU Airbus/Leonardo/Thales "Project Bromo", annunciato il 23 ottobre 2025), e il ritardo per deal è registrato nell'Appendice D. La regola governante è quindi esplicita anziché un'affermazione generica: **un deal è sottoposto a scoring su un dato orizzonte solo se nessun *segnale di risoluzione qualificante* (ai sensi del §5) per quell'orizzonte era già pubblico a T0**; se uno lo era, il deal è un'osservazione, non una previsione, ed è escluso da quell'orizzonte. Nella presente coorte nessun deal aveva un segnale H1 o H2 qualificante alla data di sigillo — tutti e dieci erano pendenti — ma diversi recavano informazione pubblica pre-T0 elevata, in particolare UniCredit–Commerzbank, la cui bassa accettazione della prima offerta e l'opposizione del governo erano già pubbliche entro il 20 giugno 2026. Tali deal non sono esclusi (nessun segnale qualificante si era verificato) ma sono **flaggati** come a informazione-pre-T0-elevata nell'Appendice D, e il ritardo annuncio-sigillo è riportato per deal così che lo svantaggio informativo sia visibile anziché assunto uniforme. Ciò sostituisce l'insostenibile asserzione generica che nessun esito sia osservabile a T0. Nessuna chiamata retrospettiva, nessuna chiamata supplementare e nessuna revisione di chiamata è trattata come il record di previsione primario; la chiamata sigillata a T0 è la sola unità sottoposta a scoring.

Stime di probabilità dalla distribuzione di scenario. Ciascuna chiamata sigillata indica una distribuzione di scenario a tre vie, mutuamente esclusiva, che somma a 100%: uno scenario di *successo* (il deal si completa e crea valore), uno scenario di *delusione di valore* (il deal si completa ma distrugge valore) e uno scenario di *fallimento del deal* (il deal salta o è ridotto materialmente). Le due stime di probabilità pre-

registrate sono derivate deterministicamente da questa distribuzione: **P(H1) = la massa di probabilità dello scenario di fallimento del deal** e **P(H2) = la massa di probabilità dello scenario di delusione di valore**. La mappatura è registrata per deal nel record sigillato; ove le etichette di scenario di un deal non si mappino nettamente su questa struttura a tre vie (es. un'offerta di controllo contestata la cui modalità di fallimento è una partecipazione di minoranza bloccata anziché un ritiro), la mappatura a livello di deal è fissata esplicitamente nell'Appendice D. Questa regola di derivazione è fissata in questo protocollo prima della risoluzione e non è aggiustata in seguito. La mappatura scenario→ipotesi è il locus più esposto al bias di costrutto (§8): la *distribuzione* di scenario è nel sigillo del 20 giugno, ma la mappatura degli scenari su H1/H2 — inclusi i quattro deal di eccezione nell'Appendice D — è redatta. È quindi pre-registrata qui, prima di qualsiasi esito, **ed è aperta a ri-verifica indipendente — da qualsiasi terzo tramite il trail di evidenze pubbliche, e da un codificatore indipendente raccomandato (§5.1)** — rispetto alla regola scritta; ciò che si ri-verifica è l'applicazione della mappatura, mentre la sua origine di costrutto resta dell'autore ed è divulgata come residuo in §8.

Queste stime di probabilità sono i valori **congelati e pubblicamente sigillati**: le P(H1)/P(H2) sottoposte a scoring sono esattamente quelle del sigillo del 20 giugno 2026 (Appendice D, §D.2), e il successivo payload completo del contraddittore non può alterarle (§4.3; il firewall interno alla corsia A, §9.3). La distribuzione di scenario è una quantità del *previsore* portata nel record dell'Osservatorio; **non** è un campo dello schema di output del contraddittore (§2.2), che emette confidenza categorica anziché una distribuzione di probabilità — i due livelli sono tenuti separati in §4.3.

Poiché i tre scenari sono mutuamente esclusivi, un deal che salta (H1 = 1) realizza H2 = 0 per costruzione. **P(H2) è quindi la probabilità marginale dell'evento congiunto "si-completa-e-distrugge-valore"** — la massa dello scenario di delusione di valore — sottoposta a scoring sull'intero sottoinsieme H2 senza separata censura per completamento. Ciò non è deliberatamente condizionato al completamento: condizionare la previsione al completamento mentre si fa scoring di un esito a sua volta definito solo sui deal completati sarebbe incoerente (porrebbe una probabilità condizionata al numeratore e un evento incondizionato al denominatore). La massa di fallimento del deal finisce in H1 e mai in H2, sicché non è doppiamente contata; la P(H2) sottoposta a scoring per deal è la massa grezza di delusione di valore registrata nell'Appendice D. In parole semplici: P(H1) è la probabilità che il deal *salti*; P(H2) è la probabilità che si *chiuda ma deluda*. Sono modalità di fallimento distinte e non sovrapposte, sicché nulla è contato due volte.

4.2 (SEGUE) — ORIZZONTI DI ESITO

Orizzonte H1 (6–12 mesi). La previsione sigillata a T0 indica una stima di probabilità esplicita P(deal-break o rinegoziazione al ribasso ≥ 15%). La soglia di rinegoziazione

del 15%, introdotta in §3, è qui operativizzata come criterio di esito realizzato: un deal è codificato H1 = 1 se e solo se un annuncio pubblico o un filing regolamentare conferma la cessazione del deal oppure una revisione del corrispettivo che riduce il prezzo headline di almeno il 15% rispetto al valore originale annunciato. La misurazione viene effettuata al primo segnale pubblico disponibile nella finestra di 6–12 mesi; se nessun segnale è disponibile entro la finestra, il deal è censurato (§4.5).

Orizzonte H2 (18–36 mesi), solo sottoinsieme con acquirente quotato. H2 presuppone un acquirente quotato che consolida il target e il cui TSR è pubblico. È quindi sottoposta a scoring **solo sul sottoinsieme pre-dichiarato di deal che soddisfano quella condizione**, fissato per riga nell'Appendice D prima dello scoring (nella presente coorte: Intesa–MPS, Poste–TIM, Zurich–Beazley, Ingredion–Tate & Lyle). Per i deal con un acquirente di private equity, un delisting, un carve-out verso un fondo, una joint venture o un break-up multi-parte, non esistono né avviamento consolidato in relazione al target né un TSR pubblico dell'acquirente; questi sono marcati *H2 non applicabile per progettazione* nell'Appendice D e non entrano mai nel pool H2. **Non** sono conteggiati come censurati, perché l'inosservabilità è strutturale e pre-dichiarata, non guidata dall'esito — distinguendo questa esclusione dalla censura a destra del §4.5. Sul sottoinsieme H2, la distruzione di valore è operativizzata come: impairment del goodwill registrato nel bilancio consolidato dell'acquirente con riferimento all'entità acquisita, OPPURE il total shareholder return (TSR) dell'acquirente — calcolato nella valuta di quotazione dell'acquirente — inferiore all'indice settoriale di riferimento (nella stessa valuta; l'indice esatto è fissato per riga nell'Appendice D) di almeno 10 punti percentuali nella finestra di 18–36 mesi successiva al closing del deal.

4.3 — OUTPUT SIGILLATO A T0 (BUSTA SIGILLATA)

A T0 ciascun deal genera un record sigillato dell'Osservatorio che **avvolge** il payload dello schema del contraddittore con metadati di sigillo. I due livelli sono distinti, e solo le stime di probabilità sigillate (elemento 2) sono sottoposte a scoring:

1. **Il payload dello schema del contraddittore** (`schema_version: "1.0"`): i sette campi di output obbligatori — assunzioni implicite, scenario controintuitivo, test di falsificazione, domande sull'onere della prova, confidenza calibrata, fonti citate — più il valore Δ-CSI calcolato dallo scorer `csi-v1` (§2.3). Questo payload è governato da `contradictor_output.schema.json` con `additionalProperties: false`; **non** contiene P(H1)/P(H2) né il bundle hash, che sono metadati di sigillo, non campi dello schema.

Gli elementi rimanenti sono **metadati di sigillo dell'Osservatorio** che avvolgono il payload; non sono campi dello schema di output del contraddittore:

2 La distribuzione di scenario a tre vie e le due stime di probabilità P(H1) e P(H2) da essa derivate (§4.2). **Queste sono le quantità sottoposte a scoring.**

3 Il consenso contemporaneo degli analisti a T0: il rating modale degli analisti (Buy / Neutral / Sell o equivalente) e il target price mediano a 12 mesi, come riportato da Refinitiv I/B/E/S a T0 (la data di sigillo).

4 Un bundle hash — un digest SHA-256 della concatenazione di: l'identificatore esatto del modello e la sua versione per ciascuna delle tre passate della pipeline (tesi, red team, sintesi), il Pack risolto (YAML, forma canonica) e la stringa di system-prompt completa usata in ciascuna delle tre passate della pipeline. Il bundle hash rende il payload di sfida riproducibile e verificabile: un terzo può verificare che sia stato prodotto dalla combinazione modello-prompt dichiarata.

5 Un timestamp Unix (UTC) che registra il momento in cui il record è stato finalizzato.

L'infrastruttura implementante è l'**Osservatorio** — il registro a busta sigillata già in produzione per la coorte di previsioni M&A della corsia A operata nell'ambito di questo studio. L'Osservatorio scrive ogni record in un archivio append-only ed espone l'hash del record tramite la propria API pubblica, rendendo verificabile la pre-esistenza di ogni previsione rispetto a un'ancora di timestamp indipendente. Il formato dei record dell'Osservatorio è la fonte canonica di verità per §5 (risoluzione) e §10 (riproducibilità).

> **STATO DI IMPLEMENTAZIONE (DIVULGAZIONE ONESTA)**
>
> L'infrastruttura dell'Osservatorio — sigillo, hashing SHA-256, l'archivio append-only committato in git e il motore deterministico di risoluzione delle notizie — è già in produzione, e i dieci deal della coorte sono già pubblicamente sigillati (ciascuno reca un hash di contenuto pubblicato; Appendice D). Al momento di questa pre-registrazione le schede pubbliche della coorte recano una vista *semplificata* di ciascuna previsione: contesto (la tesi di lettura dominante), la distribuzione di scenario a tre vie, un livello di confidenza calibrata, fonti citate e l'hash di sigillo. Le **quantità sottoposte a scoring sono le stime di probabilità già sigillate il 20 giugno 2026** (Appendice D, §D.2); sono congelate, e i passi seguenti non possono modificarle. Il **payload completo dello schema del contraddittore** — i sette campi obbligatori, il Δ-CSI e il bundle hash — non esiste ancora per questi dieci deal; sarà generato dal motore a tre passate come *provenienza non-scoring* (documenta come la sfida è stata prodotta e non può alterare P(H1)/P(H2); §9.3). Per chiudere il giardino dei sentieri biforcuti che una generazione successiva, eseguita dall'autore, aprirebbe altrimenti, la generazione è vincolata ex ante: una singola esecuzione per deal, sotto un seed deterministico e una configurazione motore/Pack/prompt il cui bundle hash è committato nel record OSF **prima** che il motore sia eseguito. Ciò vincola la configurazione in anticipo, ma non *prova* di per sé che una singola esecuzione sia avvenuta — un autore determinato potrebbe pre-selezionare un seed — sicché non è una salvaguardia completa. Due fatti limitano il residuo: il payload è provenienza non-scoring (§9.3), sicché perfino un

> payload selezionato ad arte non può muovere alcuna quantità sottoposta a scoring; e la generazione è intesa girare come un CI job registrato pubblicamente (pre-dichiarato qui, non ancora implementato) sicché un terzo può confermare che è girata una volta. Finché quel CI job non esiste, la generazione del payload poggia sulla buona fede dell'autore — divulgata come tale, non occultata. Questa distinzione live-contro-da-generare è dichiarata esplicitamente, nella stessa maniera dello stato di ancoraggio OpenTimestamps in §10.3, così che un revisore che ispeziona l'Osservatorio pubblico oggi veda esattamente cosa è già sigillato (le previsioni sottoposte a scoring) e cosa resta da generare come provenienza (i payload di sfida).

4.4 — REGOLA DELLA BUSTA SIGILLATA

Il protocollo di scoring per tutte e cinque le ipotesi (§3) utilizzerà la **prima** previsione datata per ciascun deal — la chiamata a T0 come definita in §4.2 e catturata in §4.3 — come unica unità di voto. L'*appartenenza* alla coorte è congelata (§4.1): ciò che il registro append-only consente sono righe di aggiornamento ausiliarie **su un deal già sigillato**, mai l'aggiunta o la rimozione di un deal. Se informazioni materialmente nuove (un filing regolamentare pubblico, un annuncio di offerta rivista) sono osservate dopo T0, una nuova riga datata può essere aggiunta alla catena di quel deal. Tali aggiornamenti sono dati ausiliari; saranno riportati nella traccia di evoluzione della confidenza nel tempo e potranno essere usati in analisi esplorative (§6), ma non alterano né sostituiscono la chiamata a T0 ai fini dello scoring. Questa regola è applicata operativamente dalla struttura append-only dell'Osservatorio: le righe non sono mai modificate né cancellate; la riga T0 è identificata dal timestamp più antico nella catena di record del deal.

Lo scopo della regola della busta sigillata è eliminare la contaminazione da senno di poi. Una previsione aggiornata il giorno prima della risoluzione sarebbe quasi indistinguibile da una consultazione a posteriori; la calibrazione di tale previsione sarebbe priva di valore informativo. Pre-dichiarando che solo la prima busta sigillata conta, questo protocollo espone il contraddittore al massimo svantaggio informativo al quale un genuino sistema di previsione deve operare.

4.5 — CENSURA A DESTRA

I deal che non raggiungono un segnale pubblico risolvibile entro la finestra di risoluzione dichiarata — 12 mesi per H1 e 36 mesi per H2 — saranno dichiarati **censurati** ed esclusi dal campione sottoposto a scoring per l'orizzonte corrispondente. Non saranno codificati come successi o insuccessi; saranno mantenuti in un registro separato dei censurati e riportati nella tabella descrittiva del campione. La regola di censura è pre-dichiarata qui e non è applicata post-hoc per gestire la composizione del campione. Un deal può essere censurato per H1 ma risolto per H2 (o viceversa), nel qual caso entra nel pool di scoring per l'orizzonte risolto e nel registro dei censurati per

quello non risolto. Un deal marcato *H2 non applicabile per progettazione* (§4.2) è una categoria separata: non fa parte del pool H2 e non è conteggiato come censurato, perché la sua inosservabilità è strutturale e pre-dichiarata anziché guidata dall'esito. **Se il campione risolto per un orizzonte scende al di sotto di tre, nessuna statistica — nemmeno una sintesi descrittiva — è riportata per quell'orizzonte; sono elencate solo le righe grezze previsione-contro-esito per deal** (§6.5).

SEZIONE 5

Regola di risoluzione

5.1 — IL DETERMINISMO COME INVARIANTE PROGETTUALE

La regola di risoluzione è il meccanismo con cui ciascuna chiamata sigillata a T0 — catturata secondo §4.3 — viene codificata come corretta o scorretta ai fini dello scoring. La regola è pienamente deterministica: un segnale pubblico soddisfa o meno una soglia pre-specificata. Nessun giudizio di commissione, nessuna ponderazione contestuale e nessuna interpretazione post-hoc entreranno nella decisione di codifica. Questo determinismo non è una convenienza accessoria; è la condizione in base alla quale lo scoring di calibrazione (§6) è non-arbitrario. Una regola di risoluzione che ammette discrezionalità dopo che l'esito è noto introduce esattamente la contaminazione da senno di poi che il design a busta sigillata (§4.4) è costruito per prevenire.

La codifica di risoluzione è eseguita dal ricercatore nella veste di applicatore meccanico della regola, non di analista. Il ricercatore ricercherà nelle fonti pubbliche canoniche elencate in §5.2 i segnali che soddisfano la soglia pre-registrata. Se un segnale qualificante viene trovato entro la finestra dichiarata, il deal è codificato 1 per l'orizzonte corrispondente. Se nessun segnale qualificante viene trovato entro la finestra, il deal è censurato (§4.5) ed escluso dal campione sottoposto a scoring per quell'orizzonte. In nessun caso un segnale ambiguo sarà classificato come positivo per giudizio interpretativo; i casi ambigui sono censurati a meno che un segnale successivo non ambiguo non giunga prima della chiusura della finestra. Ove un parametro di risoluzione potrebbe altrimenti ammettere una scelta — l'indice settoriale di riferimento, la valuta di misurazione o il valore di riferimento per un deal a prezzo a intervallo — esso è **fissato per deal nell'Appendice D prima dello scoring**, così che nessuna clausola "oppure" sia risolta dopo che l'esito è noto. Per rimuovere la discrezionalità rimanente — sia nel dichiarare un caso "ambiguo" sia nel codificare i casi non ambigui — la **garanzia primaria è il determinismo della regola stessa unito a un trail di evidenze pienamente pubblico**: ogni esito codificato è pubblicato con il suo codice, l'**URL della fonte primaria** e la data del segnale, così che *qualsiasi* terzo possa ri-codificarlo dal record pubblico. Dato il conflitto di paternità singola (§8), un **secondo**

codificatore indipendente è inoltre raccomandato ove fattibile — un individuo senza interesse finanziario in CounterBrain e senza ruolo nel motore o in questo protocollo, cieco alla previsione e all'ipotesi, che ri-codifica tutte le decisioni di risoluzione con l'accordo riportato come Cohen's κ — ma non è una precondizione: ove nessun codificatore indipendente sia ingaggiato, governano la regola deterministica e il trail di evidenze pubbliche. I disaccordi, o (in assenza di un secondo codificatore) qualsiasi caso genuinamente ambiguo, sono per default censurati.

5.2 — FONTI PUBBLICHE CANONICHE

I segnali di risoluzione saranno tratti esclusivamente dalla seguente gerarchia di fonti, elencata in ordine decrescente di priorità. Una fonte di priorità inferiore viene consultata solo quando nessun segnale qualificante è disponibile da una fonte di priorità superiore.

1. **Comunicati stampa ufficiali** emessi dall'acquirente o dalla società target e distribuiti tramite un servizio di newswire riconosciuto (es. i feed di informazione regolamentata di Euronext, della London Stock Exchange (RNS), di SIX Swiss Exchange o di Deutsche Börse, o servizi wire come PRNewswire e GlobeNewswire). Un comunicato stampa che confermi la cessazione del deal o un prezzo di transazione rivisto soddisfa direttamente il criterio di risoluzione H1.

2. **Filing regolamentari** presentati al regolatore nazionale dei valori mobiliari per ciascuna sede di quotazione del deal (es. Consob in Italia, l'AMF in Francia, la BaFin in Germania, la FCA nel Regno Unito, la FINMA in Svizzera o l'AFM nei Paesi Bassi) o all'operatore di borsa competente (Euronext, London Stock Exchange, SIX, Deutsche Börse), inclusa la documentazione di offerta pubblica obbligatoria e i filing di autorizzazione alla fusione — e, dato che diversi deal della coorte ruotano attorno all'autorizzazione antitrust, le decisioni della Commissione Europea o dell'autorità di concorrenza nazionale competente. I filing sono trattati come evidenza di autorità superiore del corrispettivo annunciato e delle sue revisioni; un documento di modifica vincolante che confermi una riduzione del prezzo ≥ 15%, o una decisione di divieto, costituisce H1 = 1.

3. **Bilanci consolidati** — in particolare, le note al bilancio nel report annuale o nella relazione semestrale dell'acquirente, come depositato presso Consob o la borsa competente. La nota sull'impairment del goodwill sarà utilizzata come segnale primario per il criterio di distruzione del valore H2.

4. **Dati di prezzo e indice** — prezzi di chiusura verificati dal feed dati della borsa di quotazione competente (o equivalente Refinitiv/Bloomberg) e serie di indici settoriali — indici settoriali STOXX Europe 600, o il benchmark settoriale nazionale competente — utilizzati per calcolare il TSR rispetto al benchmark settoriale per il criterio H2.

Nessun report mediatico non verificato, stima di analista o banca dati di deal di terze parti sarà utilizzato come fonte primaria di risoluzione. Quando una fonte secondaria viene consultata per scoping investigativo (per identificare l'esistenza di un potenziale

segnale), il segnale dovrà essere confermato rispetto a una fonte primaria prima della codifica.

Il layer secondario di scoping è il **motore deterministico di risoluzione via news** dell'Osservatorio (basato sulla API **newsapi.ai / Event Registry**, filtrato da una allowlist di fonti reputate e da una regola di entity-match, poi vagliato per materialità): sorveglia continuamente le notizie pubbliche per segnali candidati di risoluzione su ciascun deal della coorte e li segnala per la conferma su fonte primaria. Il motore non codifica mai un esito da solo — si limita a far emergere un candidato; ogni segnale segnalato è verificato rispetto a una fonte primaria di §5.2 prima della codifica H1/H2, ed è la regola di risoluzione deterministica (§5.1) ad assegnare il codice.

5.3 — OPERATIVIZZAZIONE DELLE SOGLIE

La soglia H1 — cessazione del deal o revisione al ribasso del prezzo ≥ 15% rispetto al corrispettivo annunciato originale — è stata introdotta in §3 e operativizzata in §4.2. La cifra del 15% è fissata in questo protocollo; non sarà aggiustata all'inizio dello scoring. La misurazione viene applicata al valore di transazione headline come indicato nel documento di annuncio iniziale; le variazioni in debito assunto, strutture earnout o corrispettivo differito che erano presenti nell'annuncio originale e non vengono successivamente revisionate non costituiscono un evento di superamento della soglia. Solo le revisioni annunciate alla cifra headline, o la conferma pubblica della cessazione del deal, attivano H1 = 1. Tre regole operative chiudono la discrezionalità residua, ciascuna fissata per deal nell'Appendice D: (i) **valuta** — la variazione ≥ 15% è misurata nella valuta originale annunciata del deal, mai dopo conversione, così che un movimento del tasso di cambio (£/€, CHF/€) non possa attivare un falso H1; (ii) **valori a intervallo** — ove il corrispettivo annunciato sia un intervallo (es. UniCredit–Commerzbank, "~35–40 mld di €"), la soglia è applicata al punto medio, a meno che una singola cifra headline sia indicata nel documento di offerta ufficiale, nel qual caso quella cifra governa; (iii) **completamento con rimedi** — un deal che si completa subordinatamente a rimedi antitrust o cessioni, ma il cui corrispettivo headline non è tagliato di ≥ 15%, è codificato H1 = 0 (non è saltato) e procede a H2 se nel sottoinsieme con acquirente quotato, mentre una decisione di divieto o un abbandono è codificato H1 = 1; un taglio di prezzo strettamente tra 0 e 15% è codificato H1 = 0; e (iv) **strutture non di acquisizione** — per una joint venture o un break-up multi-parte (deal #4 e #5 nell'Appendice D) non c'è un singolo corrispettivo headline da revisionare, sicché il criterio del prezzo ≥ 15% non si applica: H1 = 1 è attivato solo dall'abbandono pubblico della combinazione o da una decisione di divieto.

La soglia H2 — impairment del goodwill o TSR dell'acquirente inferiore al proprio indice settoriale di 10 punti percentuali — è fissata in questo protocollo e registrata nel record di pre-registrazione dell'Osservatorio prima dello scoring della coorte; l'indice settoriale esatto e la valuta di misurazione per ciascun deal del sottoinsieme H2 (TSR e

indice nella stessa valuta di quotazione dell'acquirente) sono fissati per riga nell'Appendice D.

5.4 — RIFERIMENTO AL DIZIONARIO

La mappatura completa dagli errori del libro ai fallimenti cognitivi, alle contromisure del contraddittore, ai segnali pubblici di risoluzione e alle soglie è mantenuta nell'**Appendice A** (`docs/paper/appendices/resolution-dictionary.md`). Il dizionario è il riferimento operativo per il ricercatore che applica la regola di risoluzione; la tabella in §5.2–§5.3 sopra fornisce la gerarchia delle fonti e la logica delle soglie; l'Appendice A fornisce le righe specifiche per errore che istanziano quella logica per ciascun errore del libro (Canepa) che i Pack sono progettati a contrastare.

Le righe del dizionario per gli errori del libro 01, 02 e 04 sono ora verificate contro il libro fonte (Canepa, 2026, capp. 1, 2, 4) e ratificate dall'autore, e registrate nell'Appendice A. Si veda l'Appendice A (`docs/paper/appendices/resolution-dictionary.md`) per la mappatura completa per errore di fallimenti cognitivi, contromisure del contraddittore, segnali pubblici di risoluzione e soglie.

SEZIONE 6

Baseline e piano di analisi

6.1 — BASELINE NULLA PRE-REGISTRATA

Il contraddittore acquisisce credito probatorio solo se supera una baseline nulla pre-registrata che un osservatore ingenuo potrebbe replicare senza accesso all'output del motore. Sono pre-registrati due modelli nulli.

Nullo (a) — Modello del tasso-base abbinato. Il tasso base rilevante è la frequenza con cui i **deal europei large-cap contestati, annunciati ma pendenti** si risolvono come H1 (cessazione del deal o revisione al ribasso del prezzo ≥ 15%) — *non* il tasso base M&A generale. Questa distinzione è portante: poiché la coorte è assemblata editorialmente verso deal contestati (§4.1, COI-3), il confronto con il tasso generale di ≈ 10% conferirebbe al contraddittore credito spurio per una coorte selezionata per essere ricca di eventi. Il tasso base abbinato è quindi fissato prima dello scoring mediante una **derivazione deterministica pre-registrata nel record OSF**, costruita da **fonti pubbliche e riproducibili** così che qualsiasi revisore possa ri-derivarlo invece di fidarsi di una cifra proprietaria: (i) le statistiche pubbliche sulle transazioni dello **UK Takeover Panel** (offerte ferme annunciate rispetto a decadute/ritirate, dal 2020 in poi) per il canale del fallimento del deal; (ii) le statistiche di merger control della **Commissione Europea** (proibizioni e ritiri pre-decisione ai sensi del Regolamento UE sulle concentrazioni), insieme alle autorità antitrust nazionali rilevanti (es. AGCM,

CMA, Bundeskartellamt), per il canale del blocco regolamentare; e (iii) almeno una stima peer-reviewed dei tassi di completamento/ritiro dei deal large-cap. Popolazione: deal europei ≥ 1 mld di € annunciati e contestati in una finestra storica dichiarata; esito: cessazione, proibizione o ritiro pre-decisione regolamentare, o taglio headline ≥ 15% entro 12 mesi. La cifra è congelata come un singolo numero nel record OSF, non lasciata "da confermare"; una banca dati proprietaria (Mergermarket, LSEG/Refinitiv, S&P Capital IQ) può essere usata solo come cross-check di conferma, mai come fonte primaria, poiché un revisore non può riprodurla. (Per orientamento: tra i deal *contestati* il tasso di fallimento/blocco è molto più alto del tasso ≈ 10% di tutti i deal — circa un terzo dei casi UE di Fase II finisce storicamente in proibizione o viene ritirato — ecco perché è il tasso base abbinato, non quello generale, il livello minimo onesto; la cifra esatta congelata governa.) Una previsione costante pari a questo tasso base abbinato è il livello minimo: qualsiasi sistema il cui Brier score non migliora su di esso non fornisce alcun segnale di calibrazione incrementale, e una coorte ricca di eventi non guadagna alcun credito gratuito perché il livello minimo è abbinato a quella ricchezza di eventi. Come corroborazione indipendente — mai fonte primaria — il motore news dell'Osservatorio (newsapi.ai / Event Registry; §5.2) è interrogato sulla stessa finestra storica per la copertura di deal saltati e merger bloccati, per un sanity-check dell'ordine di grandezza della cifra da fonti pubbliche; qualsiasi discrepanza è risolta a favore dei conteggi Takeover Panel / Commissione, poiché la copertura mediatica non è un censimento pulito di deal. Il tasso abbinato è ancorato a **figure pubblicate** della Fase II della Commissione UE e accademiche — riproducibili leggendo le fonti, con **nessuna raccolta di dati grezzi richiesta** — dando un valore pre-registrato **p* = 0,30** (un livello minimo duro e onesto). La derivazione completa e il raffinamento *opzionale* dello UK Takeover Panel (che può solo alzare p*) sono nell'**Appendice E**. Due qualificazioni accompagnano questo numero. **Caveat di classe di riferimento:** l'ancora ≈ ⅓ è la quota di *proibizione-o-ritiro* della Fase II UE, mentre H1 conta anche una revisione al ribasso del prezzo headline ≥ 15%, e la coorte non è un campione di pura Fase II; il canale della revisione di prezzo è reale e qui non stimato separatamente, sicché p* = 0,30 va letto come una stima centrale anziché come un tetto garantito conservativo. **Banda di sensibilità pre-registrata:** ogni conclusione su H1 della fase potenziata sarà riportata per p* da 0,25 a 0,40, così che nessun risultato dipenda dal valore puntuale.

Nullo (b) — Modello del segnale pubblico. Il secondo nullo utilizza le informazioni pubbliche contemporanee disponibili a T0: il rating di consenso medio degli analisti (Buy / Neutral / Sell o equivalente, convertito in scala numerica) e il target price mediano a 12 mesi, tratti da Refinitiv I/B/E/S a T0 (la data di sigillo), insieme alla reazione di mercato all'annuncio (rendimento anormale cumulato dell'acquirente nelle tre giornate di trading intorno all'annuncio, calcolato rispetto all'indice settoriale STOXX Europe 600 competente). Con la coorte fissa N = 10, un modello logistico a tre

predittori non può essere stimato senza overfitting; sotto il regime sottodimensionato (§6.5) il nullo (b) si riduce quindi a un **comparatore a singolo predittore — la probabilità di H1 implicita dal consenso degli analisti**, derivata dal rating modale e dal premio/sconto del target price tramite una **mappa fissa, pre-registrata** dichiarata prima dello scoring: il rating modale fissa una probabilità H1 base (Buy → 0,08; Neutral → 0,15; Sell → 0,30), aggiustata dal gap del target price rispetto al prezzo corrente (ogni sconto pieno di 10 punti percentuali aggiunge 0,05, ogni premio di 10 punti sottrae 0,05; clippata a [0,02, 0,90] — il limite superiore è abbastanza ampio da non penalizzare il nullo dove il contraddittore è più aggressivo). La mappa numerica è congelata nel record OSF prima che qualsiasi esito sia osservato, sicché non può essere calibrata per determinare chi vince H-pay; idealmente i suoi tre valori di ancoraggio sarebbero calibrati sul tasso storico di break per rating modale nella stessa banca dati usata per il nullo (a), e la fase potenziata lo farà. Il nullo (b) è definito per il **solo H1**: un rating di analista e un target price a 12 mesi parlano dell'equity dell'acquirente, non di un impairment del goodwill a 18–36 mesi, sicché non sono un proxy valido di segnale pubblico per H2; H-pay è di conseguenza valutata sul solo H1 (§3). Il Brier score del nullo (b) è il benchmark per H-pay. La specifica logistica completa (rating modale, premio/sconto del target price rispetto al prezzo corrente, e CAR nella finestra dell'annuncio, stimata una sola volta sulla coorte) è pre-registrata ma rinviata a una Fase 2 potenziata; l'N fisso non la supporta in questa sede.

6.2 — PROTOCOLLO DI SCORING

Brier score. Tutte e cinque le ipotesi sono valutate tramite regole di scoring proprie (Brier, 1950). La metrica primaria è il Brier score $BS = (1/N)\ \Sigma\ (p_i - o_i)^2$, dove p_i è la probabilità dichiarata dal contraddittore e $o_i \in \{0, 1\}$ è l'esito binario risolto. Punteggi più bassi indicano una migliore calibrazione. Il Brier score viene calcolato separatamente per gli orizzonti H1 e H2 (nessun pooling cross-orizzonte), e separatamente per la coorte completa e il sotto-campione divergente (§6.3). La scomposizione del Brier score in componenti di affidabilità, risoluzione e incertezza (Murphy, 1973) sarà riportata **solo quando il campione risolto raggiunge la soglia di potenza (§6.5)**; al di sotto, la scomposizione basata su binning è dominata dal rumore di campionamento ed è omessa.

Curva di calibrazione — H-cal. Un reliability diagram (curva di calibrazione) sarà costruito raggruppando le probabilità dichiarate in bin di ampiezza 0,10 e tracciando la previsione media rispetto alla frequenza dell'esito osservata in ogni bin. Il test formale di calibrazione per H-cal è il **test di bontà di adattamento di Hosmer–Lemeshow** (Hosmer e Lemeshow, 1980) con dieci gruppi basati sui decili, valutato al livello di significatività $\alpha = 0{,}10$. H-cal è confermata se il p-value di Hosmer–Lemeshow supera 0,10 (ossia, l'ipotesi nulla di calibrazione non è rigettata), calcolato sull'intero campione risolto per ciascun orizzonte indipendentemente. Questa soglia riconosce il regime di

piccolo-N; la scelta di α = 0,10 anziché 0,05 è una concessione pre-registrata alla potenza statistica, non un abbassamento dello standard sostanziale di calibrazione. La curva di calibrazione è tracciata indipendentemente dalla dimensione del campione; il test di Hosmer–Lemeshow è calcolato solo quando la dimensione del campione risolto raggiunge o supera la soglia di potenza definita in §6.5.

Hit-rate della divergenza-che-paga — H-pay. Nel sotto-campione di previsioni segnalate come divergenti ai sensi di H-div (§6.3), il Brier score del contraddittore è confrontato con il Brier score del modello del segnale pubblico (nullo (b)) valutato sullo stesso sotto-campione. H-pay è confermata se il Brier score del contraddittore è inferiore a quello del modello del segnale pubblico a una soglia unilaterale di α = 0,10, valutata tramite bootstrap resampling (2.000 replicazioni) della differenza dei punteggi. La formulazione unilaterale riflette la natura direzionale dell'affermazione: si afferma che il contraddittore aggiunge segnale, non che pareggia. Il bootstrap è usato al posto dei test asintotici perché il campione è previsto piccolo e la distribuzione della differenza dei punteggi non sarà probabilmente normale. Il bootstrap è **soppresso quando il sotto-campione divergente è al di sotto del livello minimo assoluto del §6.3**: su una manciata di deal un bootstrap a 2.000 replicazioni ricampiona solo pochi valori distinti ed è cosmetico, sicché non è riportato.

Esplorativo — Δ-CSI come covariata. Il Δ-CSI calcolato dallo scorer `csi-v1` e il metodo di divergenza registrato a T0 (`cosine`, o fallback Jaccard) saranno inclusi come covariate in una regressione lineare post-hoc del Brier score rispetto al Δ-CSI e all'indicatore del metodo di divergenza. Questa analisi è **solo esplorativa**: nessuna ipotesi pre-registra una relazione tra Δ-CSI e accuratezza previsionale, e qualsiasi associazione osservata sarà riportata come descrittiva e trattata come materiale Fase 2 (§3, Vincoli di scope). Un'associazione positiva tra Δ-CSI ed errore di Brier score (ossia, un'intensità di sfida più elevata correlata con una calibrazione *peggiore*) sarebbe di per sé un risultato nullo di rilievo e sarà riportata come tale.

6.3 — OPERATIVIZZAZIONE DELLA DIVERGENZA — H-DIV

Una previsione è classificata come **divergente** se la probabilità dichiarata dal contraddittore P(deal-break o rinegoziazione al ribasso ≥ 15%) supera la probabilità implicita derivata dal rating di consenso del modello del segnale pubblico di più di 15 punti percentuali, oppure se la probabilità del contraddittore è inferiore a quella di consenso del modello del segnale pubblico dello stesso margine (ossia, la regola cattura la divergenza assoluta dal consenso — verso l'alto o verso il basso — a livello di singolo deal; è un criterio di magnitudine sulla previsione stessa, non un test di significatività statistica bilaterale). Il tasso di divergenza minimo che definisce "non trascurabile" per H-div è pre-registrato al 25% della coorte. Questo tasso è scelto per escludere il disaccordo da rumore — le piccole deviazioni meccaniche prodotte da qualsiasi sistema probabilistico — consentendo al contempo un'aggregazione

significativa ai fini di H-pay. Poiché il 25% di N = 10 è solo 2–3 deal, è pre-registrato anche un **livello minimo assoluto di cinque chiamate risolte divergenti**: al di sotto, H-pay non è riportata nemmeno descrittivamente (§3, §6.2), così che un aneddoto di uno o due deal non possa essere presentato come segnale.

6.4 — COORTI DI CALIBRAZIONE

Le statistiche di calibrazione non sono **mai aggregate tra coorti definite da formule di scoring o metodi di divergenza diversi**. Una modifica alla versione della formula Δ-CSI o al metodo di divergenza costituisce un confine di coorte: i punteggi prima e dopo sono riportati separatamente e non vengono combinati in un'unica curva di calibrazione o in un aggregato di Brier score. La coorte del protocollo corrente è designata **`(csi-v1, cosine)`**: tutte le sessioni di questa coorte utilizzano la formula `csi-v1` come definita in §2.3 e calcolano la componente di divergenza tramite distanza del coseno su embedding semantici. Se un provider di embedding non è disponibile e viene usato il fallback Jaccard per una sessione, quella sessione è contrassegnata nei metadati dell'output e riportata separatamente; non viene silenziosamente unita alla coorte principale. Qualsiasi modifica futura alla formula Δ-CSI o al metodo di divergenza richiede una nuova etichetta di coorte, un addendum di pre-registrazione e una nuova baseline di scoring.

6.5 — POTENZA E N MINIMO

Una scomposizione del Brier score e una curva di calibrazione significativa richiedono un campione risolto minimo. Sulla base delle raccomandazioni standard per i reliability diagram e il test di Hosmer–Lemeshow, la soglia di potenza per il test di calibrazione confermativo è **N ≥ 20 osservazioni risolte per orizzonte sottoposto a scoring** (target aspirazionale ≈ 30). La coorte fissa di questo studio (N = 10; §4.1) è **al di sotto di questa soglia per progettazione**: il regime sottodimensionato descritto di seguito è quindi il **regime operativo** dello studio, non un fallback contingente.

Poiché la coorte è fissata al di sotto di N = 20 per orizzonte, si applica il seguente regime pre-registrato: il test di Hosmer–Lemeshow e il confronto bootstrap per H-pay non sono calcolati; la curva di calibrazione e qualsiasi sintesi per bin — incluse le tabelle per terzili — **non** sono riportate, perché con N = 10 la frequenza osservata in qualsiasi bin è dominata dal rumore di campionamento (un singolo bin di tre deal ha un errore standard prossimo a 0,27). Sono invece **elencate le righe grezze previsione-contro-esito per deal**; nessuna statistica Brier aggregata, scomposizione o intervallo bootstrap è riportata al di sotto della soglia di potenza, perché a N = 10 (e quattro su H2) qualsiasi numero del genere sarebbe dominato dal rumore di campionamento anziché informativo — la stessa ragione per cui il bootstrap di H-pay è soppresso ai sensi del §6.2. Tutte e cinque le ipotesi sono riportate come **risultanze descrittive**,

esplicitamente etichettate, mai come verdetti confermativi pass/fail; se il campione risolto di un orizzonte scende al di sotto di tre, anche la sintesi descrittiva è trattenuta e sono mostrate solo le righe grezze (§4.5). Questa non è una revisione del piano di analisi; è il percorso di analisi pre-dichiarato per una coorte di audit fissa, e preserva l'integrità del macchinario confermativo rifiutando di applicare test confermativi a dati sottodimensionati. Qualora la coorte fosse poi estesa in una Fase 2 potenziata — una pre-registrazione separata — i test confermativi del §6.2 si applicano a quel campione più ampio sotto un controllo di molteplicità pre-registrato: H-pay e H-cal sono valutate in una sequenza fissa (H-pay prima; H-cal solo se H-pay non è rigettata), così che il family-wise error rate dei due test confermativi sia mantenuto a α senza inflazione.

I dati della traccia di evoluzione della confidenza nel tempo — la sequenza dei record ausiliari T0-più-aggiornamento disponibili nell'Osservatorio (§4.4) — saranno riportati descrittivamente accanto all'analisi di scoring primaria, come evidenza esplorativa del fatto che la confidenza dichiarata dal contraddittore evolva coerentemente con le informazioni pubbliche successive. Questi record non entrano nello scoring confermativo; sono materiale ausiliario per la Fase 2.

SEZIONE 7

Criteri di falsificazione (per la fase potenziata)

La pre-registrazione richiede di dichiarare, prima di qualsiasi dato, cosa proverebbe che il contraddittore è inutile o dannoso.

Questa sezione adempie a quell'obbligo **per la fase potenziata**: le quattro condizioni di seguito sono i criteri di falsificazione che una coorte potenziata, sigillata in avanti, applicherebbe all'affermazione che la sfida avversariale strutturata produce dissenso calibrato con valore previsionale oltre le baseline pre-registrate. Sono pre-dichiarate ora così da non poter essere calibrate in seguito. **Nessuna è applicata come verdetto pass/fail sul pilot N = 10** (§6.5); il pilot riporta soltanto le quantità grezze che le alimenterebbero.

F1 — MANCATO SUPERAMENTO DEL NULLO SU BRIER SCORE O CALIBRAZIONE

La condizione di falsificazione primaria è che il contraddittore non abbia superato nessuno dei due modelli nulli (§6.1) sulle metriche di scoring pre-registrate (§6.2). Specificamente: se il Brier score del contraddittore sulla coorte completa non migliora il nullo (a) alla soglia pre-registrata (§6), oppure se il test di Hosmer–Lemeshow per H-cal rigetta la nulla di calibrazione al livello di significatività pre-registrato (§6), oppure se il Brier score del sotto-campione divergente in H-pay non migliora il nullo (b) alla soglia pre-registrata (§6), la fase potenziata dichiarerà falsificata l'affermazione di

calibrazione del contraddittore. Le soglie sono fissate in §6 e non saranno aggiustate dopo il sigillo della coorte. Nessun valore è ripetuto qui; il riferimento operativo è §6. Sul pilot N = 10, F1 non è applicata come verdetto (§6.5): il pilot riporta solo le quantità grezze che F1 consumerebbe; l'applicazione formale pass/fail è riservata a un campione potenziato.

F2 — PARALISI DECISIONALE O CALO DELL'IMPEGNO

Un contraddittore che causa nei decisori l'incapacità di impegnarsi — deferimento indefinito, paralisi da analisi, o una riduzione misurabile del throughput decisionale rispetto a una baseline non contraddetta — costituirebbe un danno anche se le metriche di calibrazione fossero soddisfatte. Questa condizione **non è un test del presente protocollo**: i dati comportamentali necessari a misurare i tassi di impegno non sono disponibili fino alla Fase 2 e sono al di fuori dello scope di questo protocollo. È pre-dichiarata qui come limite noto e come criterio di falsificazione della Fase 2. Se i dati della Fase 2 rivelassero una riduzione sistematica dell'impegno decisionale attribuibile all'inserimento del contraddittore nel processo deliberativo, il report della Fase 2 dichiarerà questo risultato e le sue implicazioni per il valore netto del sistema.

F3 — GAMING DEL Δ-CSI

Il Δ-CSI è definito in §2.3 come proxy dell'intensità della sfida — non come misura della qualità della decisione né come predittore della correttezza. Una modalità di gaming emergerebbe se l'output mostrasse punteggi Δ-CSI inflazionati senza alcun miglioramento corrispondente nel Brier score o nella calibrazione, indicando che gli incentivi strutturali della pipeline producono verbosità o sfida superficiale che punteggia alto sull'intensità pur non contribuendo alcun segnale previsionale. La regressione esplorativa del Δ-CSI (§6.2) è progettata per rilevare questo schema. Se venisse osservata un'associazione positiva tra Δ-CSI ed errore di Brier score (ossia, maggiore intensità di sfida associata a peggiore accuratezza previsionale; si ricordi che Brier inferiore = migliore) e raggiungesse una soglia di preoccupazione, il report della Fase 2 segnalerebbe questo come evidenza di gaming e lo tratterebbe come un fallimento di specifica che richiede una formula `csi-v1` rivista o una riformulazione dei pesi dei componenti. Nessuna regola di falsificazione confermativa è applicata sul pilot per questa condizione; il suo status esplorativo è pre-dichiarato.

F4 — CONVERGENZA SU UN BIAS CONDIVISO

La pipeline a tre passate (§2.4) usa l'isolamento dei modelli — tesi e red team su provider LLM diversi — per de-correlare i blind-spot sistematici. Se l'output dell'agente red-team echeggiasse sistematicamente la tesi anziché sfidarla — cioè se la componente di divergenza del calcolo Δ-CSI fosse prossima a zero in una frazione

sostanziale delle sessioni della coorte — il contraddittore avrà fallito nel suo mandato architetturale. Tale convergenza indicherebbe che la struttura avversariale è cosmetica: la pipeline produce l'apparenza di sfida mentre entrambi gli agenti condividono lo stesso errore direzionale. L'analisi della Fase 2 verificherà questo esaminando la distribuzione dei punteggi di divergenza e ispezionando un campione casuale di sessioni in cui quel punteggio scende al di sotto di un livello indicante divergenza prossima a zero — definito operativamente come il decile inferiore della distribuzione di divergenza della coorte. Se l'eco sistematica è confermata, il report della Fase 2 dichiarerà la passata red-team strutturalmente compromessa per la coorte in esame.

Se una qualsiasi delle condizioni sopra descritte emerge nella Fase 2, il report della Fase 2 la dichiarerà esplicitamente, ne riporterà la magnitudine e ne valuterà le implicazioni per le affermazioni confermative ed esplorative dello studio. Pre-dichiarare queste condizioni non è una formalità; è il meccanismo che conferisce alle risultanze della Fase 2 il loro peso probatorio.

SEZIONE 8

Conflitti di interesse e garanzie

Tre interessi sovrapposti richiedono una divulgazione esplicita, seguita dalle mitigazioni strutturali che questo protocollo ha eretto contro di essi.

COI-1 — L'AUTORE È SIMULTANEAMENTE AUTORE DEL LIBRO, COSTRUTTORE DEL SISTEMA E AUTORE DELLO STUDIO

Il sistema di contraddittore cognitivo qui testato, e l'architettura dei BRAIN Pack che lo istanzia, sono stati progettati dalla stessa persona che ha scritto il libro — *10 errori che distruggono valore nelle operazioni di M&A* (Canepa) — da cui è tratta la tassonomia degli errori sottostante, e che ora redige questo protocollo registrato. Il libro formula ipotesi su classi di fallimento cognitivo e le loro conseguenze organizzative; il presente studio operativizza quelle ipotesi come affermazioni pre-registrate. Ciò crea un rischio di ragionamento motivato: un builder-autore ha un incentivo a costruire un protocollo le cui procedure inclinino verso la conferma dello stesso framework che l'autore ha sviluppato. La mitigazione non consiste in un cambiamento di paternità — la paternità singola è una condizione dichiarata — ma nella trasparenza della dipendenza: il libro è dichiarato fonte di *ipotesi*, non di evidenze. Nessuna risultanza nella Fase 2 può essere citata a sostegno di un'affermazione la cui origine è il libro stesso; le affermazioni di origine libresca richiedono evidenze indipendenti al di là dello scope di questo studio.

COI-2 — IL PAPER LEGITTIMA IL METODO DI MISURAZIONE DI CUI IL PRODOTTO HA BISOGNO

Il Δ-CSI (Cognitive Sovereignty Index delta) è definito in §2.3 come proxy dell'intensità della sfida — il grado in cui la pipeline del contraddittore ha sottoposto a pressione una data decisione. Non è un predittore della correttezza della decisione, e §2.3 pre-registra esplicitamente questa interpretazione. Tuttavia, CounterBrain/BRAIN, il prodotto commerciale costruito su questo motore, richiede una metrica di intensità difendibile per sostenere la sua proposta di valore: un segnale misurabile che una sessione di contraddittore ha esercitato una sfida genuina. Questo studio, pubblicando la formula `csi-v1` e sottoponendo il suo output a scoring di calibrazione, contribuisce alla legittimità pubblica di quella metrica. Il conflitto è strutturale: anche un risultato Fase 2 negativo (Δ-CSI non correlato con gli esiti di calibrazione) lascerebbe la metrica pubblicata e citabile. La mitigazione è ancora trasparenza più pre-registrazione: lo status del Δ-CSI come proxy dell'intensità della sfida — non come predittore di accuratezza — è dichiarato nel protocollo, nella documentazione metodologica del motore e in §2.3 e §3 di questo paper. Qualsiasi risultanza Fase 2 che tratti il Δ-CSI come segnale di correttezza sarà pre-refutata da questa pre-registrazione.

COI-3 — LA COORTE È STATA ASSEMBLATA EDITORIALMENTE DALL'AUTORE

I dieci deal sono stati selezionati dall'autore come transazioni europee di punta, contestate in modo contemporaneo, anziché tratti da una regola di inclusione meccanica (§4.1). La selezione editoriale è un giudizio che potrebbe, in linea di principio, essere esercitato per favorire deal che l'autore si aspetta vendichino il contraddittore. Due mitigazioni lo limitano. In primo luogo, il **sigillo pubblico pre-esito**: ogni deal è stato sigillato e ancorato tramite hash prima che alcuno dei suoi esiti sottoposti a scoring fosse osservabile, e la coorte è congelata (§4.1, §4.4), sicché la selezione non può essere stata motivata da esiti che ancora non esistevano, e la struttura immutabile e append-only rende impossibile la rimozione post-hoc di un deal sfavorevole. In secondo luogo, il **nullo del tasso-base** (§6.1, nullo (a)): l'esposizione residua — che l'autore abbia scelto deal contestati più propensi a saltare di un deal large-cap casuale — non guadagna nulla al contraddittore, perché il credito matura solo battendo il tasso base realizzato della coorte effettiva, non per il fatto che la coorte sia ricca di eventi. Il costo della selezione editoriale è pagato in potenza statistica e generalizzabilità (una coorte piccola e non casuale; §6.5), non in contaminazione da senno di poi.

IL CONFLITTO COMPOSITO, DICHIARATO SENZA EUFEMISMI

Questi tre conflitti condividono un soggetto: la stessa persona ha scritto il libro, costruito il motore, progettato la metrica, selezionato la coorte e applica la regola di risoluzione. Le mitigazioni strutturali di seguito proteggono dalla contaminazione da *senno di poi* — inserimento, cancellazione o ri-codifica post-esito — e sono robuste. Esse **non** neutralizzano di per sé il bias di *costrutto*: la possibilità che la tassonomia, il Pack, le soglie (15%, 10 punti percentuali) e le mappature scenario→ipotesi siano tutte calibrate, anche in buona fede, verso la conferma. Questo residuo non è dissolto dalla sola trasparenza. Questo protocollo aggiunge quindi indipendenza esterna al nodo più esposto — la codifica di risoluzione — principalmente tramite un **trail di evidenze pubblico e auditabile**: ogni esito codificato è pubblicato con l'URL della fonte primaria e la data, così che qualsiasi terzo possa ri-codificarlo dal record pubblico (§5.1). Un **secondo codificatore indipendente è raccomandato ove fattibile** (senza interesse finanziario in CounterBrain, senza ruolo nel motore o nel protocollo, cieco alla previsione e all'ipotesi, accordo riportato come Cohen's κ) come ulteriore verifica — ma l'indipendenza della codifica non dipende dal reclutarne uno: poggia sul determinismo della regola e sul record pubblico. Dove l'indipendenza non è ancora in essere (costruzione della previsione, progettazione della metrica), il conflitto è registrato come un **rischio residuo divulgato che il presente protocollo non può neutralizzare pienamente**, non come un rischio che le mitigazioni eliminano — chiamarlo altrimenti sarebbe esattamente la compiacenza che questo sistema esiste per rifiutare.

MITIGAZIONI STRUTTURALI — PERCHÉ NON SONO RETORICHE

Una divulgazione che nomina i conflitti senza alterare la struttura degli incentivi è una formalità, non una garanzia. Quattro caratteristiche di questo protocollo rendono le mitigazioni strutturali piuttosto che retoriche.

Pre-registrazione con timestamp indipendente. Ogni previsione T0 è scritta nell'Osservatorio prima che qualsiasi esito sia osservabile (§4.3); i dieci deal della coorte sono già pubblicamente sigillati, ciascuno con un hash di contenuto pubblicato (Appendice D), il che sia dimostra che il macchinario gira sia fissa le previsioni nel tempo. L'Osservatorio è un registro append-only già in produzione; ogni hash del record è esposto tramite una API pubblica, e l'ancora **indipendente** rispetto alla quale un terzo verifica la data di creazione è il timestamp di registrazione OSF — e, una volta attiva, la prova blockchain OpenTimestamps (§10.3) — non il repository controllato dall'autore. Ciò significa che nessuna previsione può essere inserita, revisionata o postdatata dopo che un esito diventa noto. La regola della busta sigillata (§4.4) rende la prima busta T0 datata l'unica unità sottoposta a scoring; la struttura append-only la applica operativamente, non contrattualmente.

Risoluzione pubblica deterministica. La regola di risoluzione (§5) converte ogni previsione sigillata in un codice di esito binario mediante l'applicazione meccanica di una soglia pre-specificata a una gerarchia di fonti pubbliche canoniche. Nessun comitato, nessun giudizio contestuale e nessuna interpretazione post-hoc entra nella decisione di codifica (§5.1). Il ricercatore agisce come applicatore della regola, non come analista. Questo determinismo rimuove la principale leva attraverso cui il ragionamento motivato entra tipicamente nello scoring di calibrazione — il caso ambiguo classificato favorevolmente. In questo protocollo, i casi ambigui sono censurati (§4.5), non codificati favorevolmente.

Codice, prompt e Pack aperti. Il codice del motore contraddittore, le stringhe di system-prompt per tutte e tre le passate della pipeline e i file YAML dei BRAIN Pack che istanziano il comportamento specifico del dominio sono committati in un repository pubblico come parte dell'impegno di riproducibilità (§10). Qualsiasi revisore può verificare se i prompt applicano strutturalmente il mandato del contraddittore o lo attenuano sottilmente.

Freeze di versione per chiamata tramite bundle hash. Ogni record T0 include un bundle hash SHA-256 che copre l'identificatore esatto del modello per ciascuna delle tre passate della pipeline, il Pack risolto e la stringa di system-prompt completa usata in ciascuna passata (§4.3). Questo hash rende verificabile che la previsione sia stata prodotta dalla combinazione modello-prompt dichiarata. Un ricercatore che sospetti che il prompt sia stato silenziosamente modificato tra la previsione e la risoluzione può verificare l'hash rispetto al codice committato. Il freeze di versione chiude il divario tra riproducibilità nominale ed effettiva.

> *La falsificazione — non la conferma — è il beneficio asimmetrico che queste mitigazioni forniscono.*

Insieme, queste quattro caratteristiche costituiscono un protocollo le cui mitigazioni operano sul processo, non sulle intenzioni dichiarate dall'autore. Queste quattro caratteristiche delimitano la contaminazione da *senno di poi* — inserimento, cancellazione o ri-codifica post-esito; **non** delimitano il bias di *costrutto* (la tassonomia, le soglie e le mappature scenario→ipotesi restano dell'autore), registrato sopra come residuo dichiarato che la presente fase non può neutralizzare. La loro forza deriva dal fatto che la falsificazione — non la conferma — è il beneficio asimmetrico che esse forniscono: una previsione sigillata, marcata temporalmente e risolta deterministicamente che non supera il nullo (§7, F1) è inequivocabilmente un fallimento, indipendentemente dall'interesse dell'autore per un esito positivo.

SEZIONE 9

Portabilità (dichiarata, non-evidenza)

Questa sezione dichiara due corsie di portabilità — B (giudiziaria/regolamentare) e C (visibilità previsionale) — come estensioni pre-annunciate e non-evidenza del framework del contraddittore. Nessuna delle due corsie rientra nella calibrazione principale di questo studio. Le affermazioni di calibrazione dei §§3–6 riguardano **esclusivamente la corsia A**: la coorte di previsioni M&A a busta sigillata descritta in §4. Le corsie B e C sono dichiarate qui per metterle a verbale, per specificare le condizioni alle quali verrebbero perseguite, e per tracciare un firewall che impedisca ai loro futuri output di essere citati come evidenza confermativa per le affermazioni di questo protocollo.

9.1 — CORSIA B: DECISIONI GIUDIZIARIE E REGOLAMENTARI (DICHIARATA, INIZIO POSTICIPATO)

La corsia B è un test di portabilità pre-annunciato per la sfida avversariale strutturata del contraddittore in contesti decisionali legali e regolamentari. È **al di fuori della calibrazione di questo studio** ed è descritta qui esclusivamente come intenzione dichiarata, non come risultato né come evidenza di supporto.

Lo stub di universo proposto per la corsia B comprende decisioni amministrative o giudiziarie contestate in cui un brief pre-decisionale avrebbe potuto essere generato a un T0 temporalmente ben definito. Le classi documentali candidate includono sentenze d'appello della Corte di Cassazione italiana o del Consiglio di Stato, oppure decisioni antitrust emesse dall'AGCM o dalla Commissione Europea ai sensi degli articoli 101/102 TFUE, dove sia l'annuncio del procedimento sia la decisione finale portano timestamp pubblici che potrebbero ancorare una chiamata sigillata e la sua risoluzione. I criteri di selezione, la soglia dimensionale e le regole di risoluzione per una coorte di corsia B non sono specificati in questo protocollo; saranno fissati in un addendum di pre-registrazione separato, redatto solo dopo che un esperto legale qualificato sarà stato coinvolto. Nessuna chiamata sarà inserita in un registro di corsia B — e nessun output di corsia B sarà prodotto dal motore contraddittore — prima che quell'addendum sia finalizzato e marcato temporalmente.

La corsia B richiede una specializzazione di dominio che va oltre lo scope dell'attuale BRAIN Pack M&A: i procedimenti legali implicano regole di ammissibilità, allocazione dell'onere della prova e postura procedurale che richiedono un Pack redatto o revisionato da un esperto legale. Finché tale Pack non esiste e non è stato formalmente revisionato, qualsiasi output del contraddittore su materiale legale o regolamentare è esclusivamente a **scopo dimostrativo** e non deve essere interpretato come previsione calibrata né citato come evidenza per nessuna ipotesi.

La corsia B è strutturalmente al di fuori della calibrazione principale. I suoi risultati, quando eventualmente prodotti, saranno sottoposti a scoring rispetto a una baseline separata in una coorte separata e saranno riportati in una pubblicazione separata o in un addendum di pre-registrazione. In nessun caso gli esiti della corsia B saranno aggregati con le osservazioni della corsia A per aumentare il campione analizzato ai sensi dei §§3–6.

9.2 — CORSIA C: VISIBILITÀ PREVISIONALE (DIMOSTRAZIONE, NON EVIDENZA)

La corsia C è un esercizio dimostrativo che utilizza domande pubbliche datate provenienti da tornei di previsione probabilistica consolidati. I suoi output sono esplicitamente qualificati come **dimostrazione, non evidenza**. La corsia C non è **mai compresa nella calibrazione principale** di questo studio, e nulla prodotto nell'ambito della corsia C verrà trattato come confermativo, come calibrazione secondaria o come verifica di robustezza per nessuna ipotesi del §3.

Lo strumento proposto è una selezione di domande pubbliche risolvibili pubblicate da piattaforme di previsione riconosciute — ad esempio Good Judgment Open (Tetlock e Gardner, 2015; Good Judgment Inc., 2015), Metaculus (Metaculus, Inc., 2015), o infrastrutture di torneo peer-reviewed equivalenti — dove la data di chiusura della domanda, il criterio di risoluzione e la serie di probabilità della comunità sono archiviati pubblicamente. Il motore contraddittore verrà applicato a un piccolo insieme di tali domande, con ogni chiamata sigillata alla data di apertura della domanda e risolta rispetto alla risoluzione ufficiale della piattaforma. Lo scopo è illustrare come lo schema del contraddittore — assunzioni implicite, scenario controintuitivo, test di falsificazione, domande sull'onere della prova e Δ-CSI — si mappi sul formato dei tornei previsionali, e generare output osservabile che i professionisti possano esaminare.

Gli output della corsia C non saranno sottoposti a scoring di calibrazione secondo il protocollo del §6. Nessun Brier score verrà calcolato rispetto alla baseline nulla del §6.1. Nessuna ipotesi del §3 verrà valutata sui dati della corsia C. I valori Δ-CSI prodotti nelle sessioni della corsia C saranno riportati solo descrittivamente e non entreranno nella coorte di calibrazione (`csi-v1, cosine`) definita in §6.4. Qualora, in uno studio futuro, un ricercatore desideri sottoporre le chiamate della corsia C a scoring di calibrazione, dovrà essere depositata una nuova pre-registrazione con la propria regola di risoluzione, baseline e criteri di falsificazione prima che venga eseguito alcun scoring.

9.3 — IL FIREWALL TRA CORSIE

Lo scopo di questa sezione non è presentare risultati ma chiudere un potenziale divario interpretativo prima dell'inizio dello scoring della coorte. Quattro scenari futuri costituirebbero un uso improprio delle corsie B e C, e questa pre-registrazione li preclude:

1. **Trattare gli output delle corsie B o C come replicazione.** Qualsiasi affermazione in un documento Fase 2 o in una pubblicazione successiva secondo cui "i risultati della corsia B replicano le risultanze della corsia A" costituirebbe un'affermazione priva di supporto. Le corsie B e C appartengono a domini diversi, popolazioni diverse, strumenti diversi e, ove applicabile, Pack diversi. La replicazione richiede un protocollo pre-registrato corrispondente.

2. **Aggregare i valori Δ-CSI delle corsie B o C con quelli della corsia A per le statistiche di calibrazione.** La regola del confine di coorte del §6.4 si applica: le statistiche di calibrazione non vengono mai aggregate tra coorti. Le corsie B e C costituiscono namespace di coorte separati.

3. **Citare le dimostrazioni nei tornei della corsia C come evidenza per H-cal o H-pay.** La corsia C è uno strumento dimostrativo. Il suo status epistemico è illustrativo, non confermativo. Questa designazione è permanente: nessuna decisione successiva dell'autore o di un co-autore può elevare retroattivamente gli output della corsia C al rango di evidenza senza una nuova pre-registrazione indipendente che copra il dominio dei tornei.

4. **Sostituire il payload di sfida successivo alla previsione pubblicamente sigillata (il firewall interno alla corsia A).** L'uso improprio più prossimo è interno alla corsia A, non tra corsie. Le quantità sottoposte a scoring sono i valori P(H1)/P(H2) sigillati pubblicamente il 20 giugno 2026 (Appendice D, §D.2). L'intero payload del contraddittore generato successivamente come provenienza (§4.3) **non può alterarli**: se il payload rigenerato implica probabilità diverse, i valori pubblicamente sigillati governano, e la divergenza è riportata anziché silenziosamente riconciliata. Ciò preclude l'aggiornamento silenzioso di una previsione sigillata a una appena generata prima dello scoring.

SEZIONE 10

Riproducibilità ed etica

10.1 — REPOSITORY PUBBLICO

I seguenti artefatti saranno pubblicati su un repository pubblico a controllo di versione prima dell'inizio dello scoring della coorte:

- [ ] **Codice del motore.** Il sorgente completo della pipeline del contraddittore a tre passate, comprensivo dell'enforcement del Contradiction Floor, del controllo di

provenienza a enum chiuso e del gate di calibrazione sotto-soglia descritti in §2.4.

- [ ] **Codice dello scorer.** La funzione di scoring Δ-CSI `csi-v1` come specificata in §2.3, insieme ai suoi unit test.

- [ ] **Stringhe di system-prompt.** Il system-prompt verbatim per ciascuna delle tre passate della pipeline (Tesi, Red Team, Sintesi), congelato alla data di sigillo della coorte.

- [ ] **BRAIN Pack YAML.** Il Pack M&A che risolve il comportamento specifico del dominio — tassonomia dei bias, direttiva di persona red-team, template di falsificazione e guardrail operativi — per tutte le chiamate della corsia A.

Il repository sarà indicizzato su OSF (Open Science Framework) come ancora del pre-print registrato. La registrazione OSF sarà depositata prima che qualsiasi deal sia sottoposto a scoring; nessuna chiamata sarà sottoposta a scoring prima che il record OSF porti il proprio timestamp. (Lo scoring dipende dalle probabilità del previsore sigillate e dall'ancora OSF, non dal payload del motore — opzionale e non-scoring — di §4.3.) Qualsiasi revisore può quindi confrontare i prompt committati e il codice dello scorer con il bundle hash incorporato in ciascun record T0 sigillato (§4.3) per verificare che gli artefatti dichiarati corrispondano a quelli usati a runtime.

10.2 — FREEZE DI VERSIONE PER CHIAMATA E BUNDLE HASH

Ogni record T0 nell'Osservatorio (§4.3) include un **bundle hash**: un digest SHA-256 prodotto concatenando l'identificatore esatto del modello e la stringa di versione per ciascuna delle tre passate della pipeline (tesi, red team, sintesi), il Pack risolto in forma canonica YAML e la stringa di system-prompt completa per ciascuna delle tre passate della pipeline, in quest'ordine. L'hash è calcolato al momento di finalizzazione della chiamata, prima che il timestamp Unix venga scritto. Il suo scopo è legare il **payload di sfida** del contraddittore in modo non ambiguo alla combinazione modello-prompt esatta che l'ha prodotto. Si noti la distinzione del §4.3: la *previsione sottoposta a scoring* è la P(H1)/P(H2) pubblicamente sigillata, che il bundle hash documenta come provenienza ma non genera di per sé.

La regola di versioning è:

1 **Congelato per chiamata.** Modello, prompt e Pack sono bloccati nel momento in cui la chiamata è finalizzata. Un aggiornamento del motore che modifica uno qualsiasi dei tre componenti costituisce un cambio di versione.

2 **Pre-registrato prima del deployment.** Qualsiasi modifica al motore che alteri il bundle hash deve essere annunciata in un addendum di pre-registrazione depositato nel record OSF *prima* che il motore aggiornato elabori qualsiasi chiamata della coorte. Le chiamate elaborate sotto la nuova versione sono contrassegnate con il nuovo hash; nessuna rietichettatura retroattiva è consentita.

3 **Chiamate taggate alla propria versione.** Ogni riga nell'Osservatorio conserva il proprio bundle hash. Si applica la regola del confine di coorte del §6.4: le sessioni i cui bundle hash corrispondono a versioni diverse di prompt o Pack sono riportate con etichette di coorte separate e non vengono mai aggregate per le statistiche di calibrazione.

4 **Nessuna modifica retroattiva.** L'Osservatorio è un archivio append-only; le righe non vengono mai modificate né cancellate. Un ricercatore che sospetti una modifica post-hoc del prompt può confrontare il bundle hash in qualsiasi riga con l'hash degli artefatti pubblicati nel repository al corrispondente timestamp di commit.

La procedura di verifica completa — cosa viene sottoposto a hash, dove l'hash è conservato e come un terzo riproduce il digest — è descritta nell'**Appendice C** (`docs/paper/appendices/freeze-versioning.md`).

10.3 — TIMESTAMPING TRAMITE OPENTIMESTAMPS

Ogni hash del record dell'Osservatorio sarà inviato a **OpenTimestamps** (Todd, 2016) per essere ancorato nella blockchain Bitcoin. OpenTimestamps produce una prova crittografica che un determinato digest esisteva prima di un particolare blocco Bitcoin, indipendentemente da qualsiasi parte affiliata a questo studio. Ciò fornisce una seconda ancora di timestamp immutabile che opera al di fuori dell'infrastruttura propria dell'Osservatorio.

L'ancoraggio tramite OpenTimestamps è attualmente nel backlog dell'Osservatorio e sarà operativo prima dell'inizio dello scoring della coorte. Finché non sarà operativo, l'ancora indipendente è il **timestamp di registrazione OSF** (un terzo), rispetto al quale ciascun hash del record sigillato è registrato; l'hash del record append-only proprio dell'Osservatorio e la cronologia dei commit git lo corroborano ma sono controllati dall'autore (un repository privato ammette la riscrittura della cronologia) e quindi non costituiscono di per sé un'ancora indipendente. Una volta attivato OpenTimestamps, ogni record porterà inoltre una prova blockchain. Le ancore sono indipendenti dall'autore; o il timestamp OSF o la prova blockchain è sufficiente a stabilire la pre-esistenza della previsione sigillata. Stato attuale onesto: finché la registrazione OSF non è effettivamente depositata, i sigilli del 20 giugno poggiano su una cronologia git controllata dall'autore e su un file di lock locale e **non sono ancora ancorati**

indipendentemente; nessuna previsione è sottoposta a scoring prima del deposito OSF (§10.1), sicché l'ancoraggio indipendente è una precondizione dello scoring, non un fatto compiuto al momento di questo draft.

10.4 — ETICA

Soggetti umani — presente protocollo. Il presente protocollo non raccoglie dati da soggetti umani. La corsia A (la coorte di previsioni M&A) opera esclusivamente su dati pubblici: termini di transazione annunciati, rating di consenso degli analisti tratti da database pubblici e documenti regolamentari depositati pubblicamente. Il motore contraddittore elabora questi artefatti pubblici autonomamente; nessun dato personale, comunicazione o record comportamentale di singoli individui viene raccolto o elaborato. Esso ricade quindi al di fuori del perimetro dei requisiti di revisione etica istituzionale per la ricerca su soggetti umani, come definito dai framework standard (es. Dichiarazione di Helsinki; decreto legislativo italiano 196/2003 come modificato dal decreto 101/2018 sull'implementazione del GDPR).

Protocollo Fase 2 su soggetti umani. La Fase 2 di questo programma di ricerca coinvolgerà soggetti umani — decisori in contesti organizzativi che interagiscono con il motore contraddittore e i cui esiti comportamentali (tassi di impegno decisionale, aggiornamento delle credenze, confidenza dichiarata) vengono misurati. La Fase 2 richiederà una revisione etica separata e completa da parte di un comitato etico istituzionale riconosciuto (IRB o equivalente) prima che inizi qualsiasi raccolta dati. La pre-registrazione Fase 2, quando depositata, includerà il numero di approvazione IRB come condizione prerequisito. Nessuna raccolta dati Fase 2 procederà senza tale approvazione. Questa sequenzialità è dichiarata qui per pre-dichiarare il requisito di approvazione etica piuttosto che rinviarlo.

Gestione dei dati e privacy. Il Sovereign Decision Ledger è per-tenant e conservato localmente; nessun dato a livello di deal viene trasmesso a provider di modelli esterni in una forma che possa re-identificare una specifica transazione del mondo reale. I system prompt passano il blocco decisionale come input strutturato; si applicano i termini d'uso dei dati dei provider di modelli. L'autore rivedrà e documenterà quei termini nella registrazione OSF prima dell'inizio dello scoring della coorte. I record del Ledger sono cifrati a riposo; l'accesso è registrato e verificabile.

Integrità autoriale. I conflitti di interesse dichiarati in §8 si estendono all'impegno di riproducibilità: rendere pubblicamente disponibili il codice del motore, i prompt e i Pack è il meccanismo che converte le dichiarazioni del §8 da retoriche a strutturali. Il bundle hash chiude il divario tra gli artefatti dichiarati e quelli effettivamente usati; l'ancora OpenTimestamps chiude il divario tra la timeline dichiarata da un ricercatore e una verificabile.

> ***Nota:*** *L'Appendice B (`docs/paper/appendices/illustrative-card.md`) presenta un esempio inventato chiaramente etichettato di scheda di output del contraddittore, illustrando i sette campi di output obbligatori (questa scheda ne popola sei; `sources[]` è vuoto perché non è simulato alcun retrieval) e il Δ-CSI con probabilità di scenario che sommano al 100%. In questa edizione l'appendice è riportata in italiano qui di seguito; il testo di riferimento canonico resta la versione nel repository.*

> ***Nota:*** *L'Appendice C (`docs/paper/appendices/freeze-versioning.md`) descrive il meccanismo del bundle hash in dettaglio tecnico completo: cosa viene sottoposto a hash, dove il digest è conservato e come un terzo verifica la versione di una chiamata. In questa edizione l'appendice è riportata in italiano qui di seguito; il testo di riferimento canonico resta la versione nel repository.*

> ***Nota:*** *L'Appendice D (`docs/paper/appendices/cohort-register.md`) enumera la coorte fissa di N = 10: ciascun deal, il suo hash di contenuto sigillato, il suo nesso di quotazione e la sua mappatura scenario-ipotesi (P(H1)/P(H2)). In questa edizione è riportata in italiano qui di seguito; il testo di riferimento canonico resta la versione nel repository.*

APPENDICE A

Dizionario di risoluzione

> ***Edizione italiana:*** *in questa edizione le appendici A, B, C e D sono riportate in italiano. Il testo di riferimento canonico del progetto resta la versione nel repository, coerentemente con la nota di rimando in §5.4 e §10.*

FINALITÀ

Questo dizionario operativizza la regola di risoluzione enunciata in §5. Ogni riga mappa un errore del libro a cui dà un nome al fallimento cognitivo che esso istanzia, alla contromossa del contraddittore progettata per portarlo alla superficie, al segnale pubblico che risolve se l'errore si è materializzato, e alla soglia che converte un segnale ambiguo in un codice di esito binario.

Il dizionario è la checklist del ricercatore al momento della risoluzione: data una busta sigillata a T0 e un segnale pubblico in arrivo, si consulta la riga di errore pertinente per determinare se il segnale qualifica come H1 = 1 o H2 = 1.

> **NOTA DI STATO — VERIFICATO CONTRO IL LIBRO FONTE**
>
> Le righe per gli errori del libro **01**, **02** e **04** sono ora verificate contro il libro fonte (*10 errori che distruggono valore nelle operazioni di M&A*, Canepa 2026, capitoli 1, 2, 4) e ratificate dall'autore nell'ambito della chiusura della pre-registrazione (Task F2, 2026-06-23). Le righe sono operative per la risoluzione della coorte.

TABELLA — MAPPATURA PER ERRORE

ERRORE DEL LIBRO	FALLIMENTO COGNITIVO	CONTROMOSSA DEL CONTRADDITTORE	SEGNALE PUBBLICO DI RISOLUZIONE	SOGLIA

01	**Bias di azione/momentum aggravato dalla razionalizzazione a posteriori.** Il deal prende velocità — spinto dalla fretta, dalla pressione competitiva, dalla disponibilità dell'opportunità o dalla paura di restare fuori — prima che esista alcuna tesi industriale scritta. La tesi industriale arriva poi *dopo* la firma per giustificare una decisione già presa "di pancia", anziché guidarla. Il fallimento è l'inversione dell'ordine epistemico corretto: la tesi dovrebbe precedere e vincolare il target, il prezzo e il finanziamento, non discenderne (Canepa 2026, cap. 1).	**Sfida dell'Investment Thesis Memo (tesi industriale).** La passata red-team del contraddittore operativizza la contromossa del libro: essa (i) esige l'esplicito **Investment Thesis Memo** scritto — un documento di una pagina che copre gli otto punti definiti nel libro (problema strategico; perché acquisire vs. crescere / JV / accordo commerciale; profilo del target ideale; fonti di valore; condizioni minime; punti non negoziabili; trigger di rinuncia; piano post-acquisizione) — redatto *prima* che inizi qualsiasi negoziazione e usato come filtro go/no-go; (ii) verifica che sia stato seguito l'ordine corretto (tesi → target → prezzo → finanziamento) e che nessun term sheet o NDA preceda il memo; (iii) pone la domanda sull'onere della prova: chi nell'organizzazione dell'acquirente ha avallato la logica industriale in modo indipendente dall'originatore del deal, documentandolo prima dell'avvio della modellizzazione finanziaria? Se nessun memo esiste, la pipeline si ferma e contrassegna la chiamata come insufficientemente documentata per la sigillatura a T0.	Primario: comunicato stampa ufficiale dell'acquirente all'annuncio del deal (§5.2, fonte 1) — presenza o assenza di una razionale strategica esplicita che enunci la tesi industriale. Secondario: comunicati stampa successivi o narrativa della relazione intermedia/annuale entro 6–12 mesi che identifichino una svalutazione per mancato fit strategico, un trigger di impairment del goodwill, o una dichiarazione esplicita che il deal non ha conseguito le sinergie attese (§5.2, fonte 3).	Criterio H2: impairment del goodwill registrato nel bilancio consolidato dell'acquirente con riferimento all'entità acquisita (§5.3), entro la finestra di 18–36 mesi; OPPURE TSR inferiore all'indice settoriale di 10 p.p. nella stessa finestra. L'errore 01 non ha una soglia H1 autonoma; un deal privo di tesi industriale a T0 ma che si chiude senza revisione del prezzo headline è codificato H1 = 0; il segnale di fallimento della tesi è atteso materializzarsi sull'orizzonte H2 più lungo.

02	**Ancoraggio a un singolo numero di valutazione più illusione di oggettività/controllo.** Un numero derivato dal modello dà un rassicurante senso di precisione ed è trattato come un verdetto — il prezzo — fondendo in un'unica cifra tre livelli concettualmente distinti: (a) enterprise value (l'intervallo di valore intrinseco data la tesi specifica del compratore e la struttura del capitale); (b) spazio di negoziazione (l'intervallo effettivamente concordabile date le aspettative del venditore, la tensione competitiva e la struttura del deal); (c) struttura del deal (mix del corrispettivo, disegno dell'earnout, aggiustamenti per posizione finanziaria netta, working-capital peg, garanzie e indennizzi). Ancorarsi al numero rende il modello un alibi anziché uno strumento: l'analisi finanziaria è costruita per giustificare il prezzo già impegnato, non per scoprirlo (Canepa 2026, cap. 2).	**Sfida del "ponte di valore" (value bridge).** La passata red-team del contraddittore operativizza la contromossa del libro: essa (i) esige che il blocco decisionale separi esplicitamente i tre livelli sopra e produca un documento **"ponte di valore" (value bridge)** che distingua enterprise value, aggiustamenti per posizione finanziaria netta, working-capital peg, componenti di prezzo differite/condizionate e cassa netta per il venditore; (ii) richiede che le offerte concorrenti siano confrontate sulla qualità economica complessiva — non sul prezzo headline — traducendo ciascun termine strutturale in una cifra di impatto economico; (iii) costruisce un test di falsificazione: sotto quale scenario di earnout o di aggiustamento del prezzo di acquisto il prezzo effettivo pagato eccederebbe l'intervallo di valore intrinseco di più del 15%? Se nessun ponte di valore esiste e nessuno scenario di questo tipo è modellizzato, la pipeline si ferma e contrassegna la chiamata come insufficientemente documentata per la sigillatura a T0.	Primario H1: annuncio ufficiale di revisione del prezzo o di cessazione del deal (§5.2, fonte 1 e fonte 2). Una modifica del corrispettivo successiva all'annuncio — inclusa la ridefinizione dell'earnout, un aggiustamento del working capital eccedente il 15% del prezzo headline, o la cessazione del deal — costituisce H1 = 1. Primario H2: impairment del goodwill (§5.2, fonte 3) attribuibile a sovrapprezzo rispetto alla realizzazione delle sinergie, come disclosed nelle note al bilancio consolidato dell'acquirente.	H1: cessazione del deal o revisione al ribasso annunciata del corrispettivo headline ≥ 15% rispetto al valore di annuncio originale (§5.3). H2: impairment del goodwill nel bilancio consolidato dell'acquirente (§5.3) OPPURE TSR inferiore all'indice settoriale di 10 p.p. entro la finestra di 18–36 mesi. L'errore 02 è l'errore più direttamente legato al criterio H1; una confusione prezzo/struttura che emerge pubblicamente entro 6–12 mesi è il segnale più pulito; l'impairment del goodwill a H2 è il segnale ritardato quando la confusione non viene rilevata fino al post-integrazione.

04	**Ancoraggio alla prima bozza del venditore più framing reattivo.** L'errore è il COMPRATORE che accetta la LOI o lo SPA della CONTROPARTE (il venditore) come base di lavoro. Due meccanismi che si sommano: (a) una componente di ancoraggio — una volta che la bozza del venditore è accettata come documento di negoziazione, la sua struttura, le definizioni, l'ambito delle representation, i cap sugli indennizzi e le esclusioni di disclosure diventano l'ancora; ogni revisione del compratore è inquadrata come una concessione anziché come un'imposizione, rendendo psicologicamente e praticamente osteggiato qualunque riequilibrio sostanziale; (b) una componente di framing reattivo — il compratore negozia "in reazione, non in impostazione", cedendo progressivamente terreno tramite le definizioni, i cap e il regime di disclosure scelti dal venditore, anziché stabilire fin dall'inizio la propria logica di allocazione del rischio (Canepa 2026, cap. 4).	**Sfida della mappa delle clausole del compratore (buyer's clause map).** La passata red-team del contraddittore operativizza la contromossa del libro: il compratore deve presentarsi con il PROPRIO framework di negoziazione — una **mappa delle clausole del compratore** — prima di accettare qualsiasi bozza dalla controparte. La passata red-team (i) verifica che il compratore abbia pre-definito la propria logica di allocazione del rischio: punti non negoziabili vs. negoziabili, posizioni di ripiego per ciascuna clausola economicamente sensibile, ambito della clausola MAC, cap sugli indennizzi e soglie del disclosure-basket; (ii) traduce ogni clausola economicamente sensibile in una cifra di impatto monetario per il consiglio — così che il decisore stia comprando rischio, non linguaggio; (iii) pone la domanda sull'onere della prova: il team legale e finanziario dell'acquirente ha prodotto questa mappa delle clausole del compratore prima di esaminare la bozza del venditore? Se nessuna mappa delle clausole del compratore esiste, la pipeline si ferma e contrassegna la chiamata come insufficientemente documentata per la sigillatura a T0. La passata red-team si applica anche se il compratore lavora in ultima istanza sulla bozza del venditore — "pensare prima" è l'invariante.	Primario H1: annuncio ufficiale di cessazione del deal che cita disaccordo contrattuale, mancato avveramento di una condizione regolamentare, o invocazione di una MAC (§5.2, fonte 1 e fonte 2). Un annuncio di cessazione che cita una qualsiasi di queste cause entro 6–12 mesi da T0 costituisce H1 = 1. Secondario: filing presso Consob o la borsa che modifica i termini del deal (revisione del corrispettivo, rinuncia a condizioni, ridefinizione dell'earnout) entro la finestra di 6–12 mesi (§5.2, fonte 2). H2: impairment del goodwill (§5.2, fonte 3) dove le note citano fallimenti di integrazione o claim di garanzia come trigger dell'impairment.	H1: cessazione del deal (qualsiasi causa pubblica) OPPURE revisione al ribasso ≥ 15% del corrispettivo headline entro la finestra di 6–12 mesi (§5.3). L'errore 04 è l'errore più strettamente accoppiato all'orizzonte H1: i fallimenti di framing reattivo emergono tipicamente entro il primo anno come dispute contrattuali, condizioni non avverate o blocchi regolamentari che un term sheet guidato dal compratore avrebbe affrontato. H2: TSR inferiore all'indice settoriale di 10 p.p. OPPURE impairment del goodwill (§5.3) per i casi in cui la debolezza contrattuale non emerge pubblicamente a H1 ma si manifesta nella distruzione di valore post-integrazione.

DEFINIZIONI DELLE COLONNE

Errore del libro — L'identificatore dell'errore come definito nel manoscritto del libro. Gli errori 01, 02 e 04 corrispondono alle tre modalità di fallimento cognitivo che i BRAIN Pack sono specificamente progettati a contrastare nei contesti M&A.

- [ ] **Fallimento cognitivo** — La modalità di fallimento cognitivo a cui è dato un nome (es. variante del bias di conferma, forma di overconfidence, pattern di ancoraggio) che l'errore del libro rappresenta. Le etichette sono tratte dai capitoli del libro a cui danno il nome (Canepa 2026) e ratificate dall'autore; riflettono l'inquadramento di ciascun errore proprio del libro anziché un'inferenza dalla letteratura esterna.
- [ ] **Contromossa del contraddittore** — Lo specifico intervento red-team — tratto dalla sezione di tassonomia dei bias del Pack risolto — che il contraddittore applicherà quando rileva questo errore nel blocco decisionale.
- [ ] **Segnale pubblico di risoluzione** — Lo specifico tipo di documento e campo dati (es. "report annuale consolidato dell'acquirente, nota sull'impairment del goodwill, disclosure IFRS 3 / IAS 36") da cui viene letto il codice di esito, secondo la gerarchia delle fonti in §5.2.
- [ ] **Soglia** — La soglia del segnale pubblico specifica per errore che applica i codici di esito H1 e H2 definiti in §5.3 per risolvere ciascuna riga di errore in un esito binario.

RIMANDI

§5.2 gerarchia canonica delle fonti (comunicati stampa, filing regolamentari, bilanci, dati di prezzo/indice)

§5.3 operativizzazione delle soglie (soglie di esito H1 e H2; non ripetute qui)

§3 definizioni delle ipotesi H1 e H2

§4.2 definizioni dell'unità di predizione e degli orizzonti

A `packages/packs/` — YAML dei BRAIN Pack, che forniscono la tassonomia dei bias da cui sono tratte le contromosse

APPENDICE B

Scheda illustrativa di output del contraddittore

> ***Edizione italiana:*** *in questa edizione l'appendice è riportata in italiano. Il testo di riferimento canonico del progetto resta la versione nel repository, coerentemente con la nota di rimando in §10.*

> **IMPORTANTE — TUTTI I DATI IN QUESTA APPENDICE SONO INVENTATI**
>
> Questa scheda è un esempio sintetico e illustrativo creato unicamente per mostrare la struttura di un output del contraddittore. Non rappresenta alcuna società, transazione, report di analista o evento storico reale. Nessuna inferenza su esiti M&A del mondo reale deve essere tratta da alcuna cifra o narrativa qui presente. Le probabilità in §B.6 sono inventate a scopo illustrativo; sommano al 100% come richiesto dallo schema ma non hanno alcuna base empirica.

FINALITÀ

Questa appendice mostra come si presenta una singola scheda di output del contraddittore. Rispecchia lo schema di output a sette campi definito in §2.2 e l'esempio-stub di pre-registrazione richiamato in §3; questa scheda popola sei dei sette campi, lasciando `sources[]` vuoto perché non è simulato alcun retrieval. L'esempio usa un'acquisizione fittizia di una mid-cap italiana così che il dominio (M&A europeo) corrisponda all'universo della corsia A senza riferirsi ad alcun deal reale.

CONTESTO DELLO SCENARIO (INVENTATO)

DEAL FITTIZIO

Alfa S.p.A. acquisisce Beta S.r.l.

Tutte le cifre sono inventate — nessuna società, transazione o evento reale

Acquirente — Alfa S.p.A. (società quotata inventata, Euronext Milan). Target — Beta S.r.l. (società di logistica non quotata inventata). Corrispettivo annunciato: €180 milioni (cifra inventata). Rating modale degli analisti a T0: **Buy** (4 broker su 6, inventato). Data di annuncio: ⟨T0 — illustrativo⟩.

CAMPO 1 — ASSUNZIONI IMPLICITE (`assumptions[]`)

#	ASSUNZIONE	SEVERITÀ (1–5)	CATEGORIA

A1	Alfa può realizzare le sinergie previste di €12 M/anno entro 24 mesi senza compromettere le relazioni con i clienti esistenti di Beta.	4	Ottimismo di integrazione
A2	La base di ricavi di Beta non dipende in modo materiale da due clienti chiave i cui contratti scadono entro 18 mesi dal closing del deal.	5	Rischio di concentrazione
A3	Il via libera regolamentare italiano da parte dell'AGCM seguirà la tempistica standard di Fase I di 30 giorni.	3	Tempistica regolamentare

> *(Tutte le cifre e descrizioni sono inventate.)*

CAMPO 2 — SCENARIO CONTROINTUITIVO (COUNTER_SCENARIO)

La lettura dominante tratta l'acquisizione come un semplice bolt-on logistico, con sinergie derivanti dalla sovrapposizione delle rotte e dalla razionalizzazione della flotta. Lo scenario controintuitivo è che l'apparente stabilità dei ricavi di Beta mascheri un problema di concentrazione della clientela che la due diligence di Alfa non ha sondato a livello di rinnovo contrattuale. Se i due clienti chiave (rappresentanti un inventato 61% del margine lordo di Beta) esercitassero le opzioni di rinnovo a tariffe inferiori — una mossa negoziale resa razionale dall'annuncio stesso, che segnala loro un aumento del proprio potere contrattuale — l'obiettivo di sinergie da €12 M collasserebbe prima che i team di integrazione siano operativi. L'elemento controintuitivo è che l'annuncio dell'acquisizione è il trigger del deterioramento sul lato clienti che rende il deal distruttivo di valore.

> *(Narrativa inventata. Nessuna cifra reale di cliente, ricavo o margine.)*

CAMPO 3 — TEST DI FALSIFICAZIONE (FALSIFICATION_TESTS[])

#	TEST	DATI NECESSARI	STATO
FT1	Ottenere lo schema di concentrazione della clientela dai bilanci certificati di Beta (i 5 maggiori clienti per margine lordo). Se nessun singolo cliente eccede il 25% del margine lordo, l'assunzione A2 è materialmente supportata.	Bilanci certificati di Beta (T0 o più recenti disponibili).	Non ancora eseguito (chiamata T0)
FT2	Verificare se l'AGCM abbia emesso una richiesta di estensione di Fase II su un consolidamento logistico comparabile nei 36 mesi precedenti T0. In caso affermativo, ri-stimare la tempistica di via libera.	Registro pubblico delle decisioni AGCM.	Non ancora eseguito

> *(Test inventati. Nessuna fonte di dati o esito reale implicato.)*

CAMPO 4 — DOMANDE SULL'ONERE DELLA PROVA (BURDEN_QUESTIONS[])

#	DOMANDA	PORTATORE DELL'OBBLIGO	OBBLIGO ADEMPIUTO?
BQ1	Chi porta l'obbligo di dimostrare che i contratti dei clienti chiave contengano termini di rinnovo che sopravvivono al cambio di controllo? Gli advisor di Alfa, nell'ambito della due diligence legale, portano questo obbligo. Il memo di due diligence legale è stato esaminato e affronta esplicitamente questo punto?	Advisor legali di Alfa	Non dimostrato a T0
BQ2	Chi porta l'obbligo di mostrare che la tempistica di sinergie annunciata sia coerente con i precedenti di integrazione logistica nell'M&A italiano di fascia media? Il management di Alfa porta questo obbligo attraverso la presentazione agli investitori. La presentazione agli investitori enuncia la tempistica ma non fornisce alcun precedente comparabile.	Management di Alfa	Obbligo enunciato, non comprovato

> *(Domande inventate. Nessun parere legale o documento per investitori reale implicato.)*

CAMPO 5 — CONFIDENZA CALIBRATA (CONFIDENCE)

Livello: `medium`

Motivazione: Le stime di probabilità dichiarate (Campo 6 di seguito) poggiano sui dati pubblici disponibili a T0. L'assunzione di concentrazione della clientela (A2) potrebbe essere confutata o confermata da dati di due diligence a livello contrattuale non pubblicamente disponibili; questo è il principale gap di conoscenza. L'assunzione di tempistica regolamentare (A3) è parzialmente osservabile dai precedenti AGCM. L'incertezza complessiva è media perché la lettura dominante non è implausibile — i bolt-on logistici riescono frequentemente — ma lo scenario controintuitivo identifica un meccanismo strutturalmente credibile e non pubblicamente smentito a T0.

Gap di conoscenza dichiarati:

- [ ] Schema di concentrazione della clientela a livello contrattuale (non pubblico).
- [ ] Contenuto del memo di due diligence legale sulle clausole di cambio di controllo (non pubblico).
- [ ] Carico di casi interno dell'AGCM a T0 (non pubblico, osservabile solo approssimativamente tramite il ritardo del registro pubblico delle decisioni).

> *(Tutte le caratterizzazioni sono inventate.)*

CAMPO 6 — PROBABILITÀ DI SCENARIO E Δ-CSI

> **INVENTATO A SOLO SCOPO ILLUSTRATIVO**
>
> Questi numeri non rappresentano alcuna previsione calibrata, esito storico o output di modello.

I tre scenari identificati da questa scheda esauriscono lo spazio degli esiti all'orizzonte H1 (6–12 mesi):

SCENARIO	DESCRIZIONE	PROBABILITÀ ASSEGNATA
S1 — Il deal si chiude ai termini annunciati	Via libera regolamentare di Fase I; concentrazione dei clienti chiave non materiale; tempistica di sinergie intatta.	58%
S2 — Rinegoziazione al ribasso ≥ 15%	Il problema di concentrazione della clientela emerge durante il periodo di rimedio; Alfa negozia un corrispettivo ridotto. H1 = 1 ai sensi della soglia pre-registrata.	27%
S3 — Cessazione del deal	La Fase II dell'AGCM blocca la transazione, oppure il deterioramento dei clienti chiave rende il deal economicamente non sostenibile prima del closing. H1 = 1.	15%
Totale		100%

P(H1) dichiarata = P(S2) + P(S3) = 42%. (Inventato. Probabilità implicita del consenso degli analisti a T0: ~18% — anch'essa inventata.)

Δ-CSI (CSI-V1) — INVENTATO, A SOLO SCOPO ILLUSTRATIVO

COMPONENTE	VALORE	NOTE
Conteggio assunzioni × severità media	0.80	3 assunzioni, severità media 4.0 → min(3,5)/5 × (4.0/5) — inventato
Conteggio test di falsificazione	0.40	2 test → min(2,5)/5 — inventato
Conteggio domande sull'onere della prova	0.40	2 domande → min(2,5)/5 — inventato

Conteggio fonti	0.20	1 fonte → min(1,5)/5 — inventato
Punteggio di divergenza	0.62	Distanza del coseno (embedding) tra tesi e counter_scenario — inventato
Somma pesata	57	Usando i pesi di default csi-v1 che sommano a 100 — inventato

Δ-CSI = 57 (inventato, a solo scopo illustrativo). Questo valore non porta alcuna affermazione di calibrazione e non è un predittore di accuratezza (§2.3).

RIMANDI

- [ ] §2.2 — schema di output fisso (sette campi obbligatori; questa scheda ne popola sei dei sette; `sources[]` è vuoto perché qui non è simulato alcun retrieval)
- [ ] §2.3 — definizione del Δ-CSI e formula csi-v1
- [ ] §4.3 — formato di output sigillato
- [ ] Appendice C — bundle hash e congelamento di versione

> *Fine dell'Appendice B — Scheda illustrativa. Tutto il contenuto è inventato.*

APPENDICE C

Bundle hash e congelamento di versione

> ***Edizione italiana:*** *in questa edizione l'appendice è riportata in italiano. Il testo di riferimento canonico del progetto resta la versione nel repository, coerentemente con la nota di rimando in §10.*

FINALITÀ

Questa appendice descrive il meccanismo di congelamento di versione introdotto in §4.3 e §10.2 del protocollo principale. Specifica esattamente: (1) cosa viene sottoposto a hash per produrre il bundle hash; (2) dove l'hash è conservato; (3) come un terzo riproduce il digest e verifica la versione di una chiamata.

1. COSA VIENE SOTTOPOSTO A HASH

Il bundle hash è un digest **SHA-256** di un'unica stringa di byte concatenata, assemblata dai seguenti cinque componenti, nell'ordine elencato:

#	COMPONENTE	DESCRIZIONE
C1	Identificatore del modello (Passata 1 — Tesi)	La stringa esatta dell'ID del modello come restituita dall'API del provider al momento della chiamata (es. `claude-opus-4-5-20251101`). Include l'eventuale prefisso del provider.
C2	Identificatore del modello (Passata 2 — Red Team)	Stesso formato. Può differire da C1 se l'isolamento dei modelli è attivo (§2.4).
C3	Identificatore del modello (Passata 3 — Sintesi)	Stesso formato.
C4	YAML del Pack risolto	Lo YAML del BRAIN Pack come risolto al momento della chiamata: Pack core fuso con il Pack verticale fuso con il Pack cliente, in forma canonica (chiavi ordinate lessicograficamente, YAML versione 1.2, terminazioni di riga LF, nessuno spazio in coda). La forma canonica è deterministica: due invocazioni qualsiasi che si risolvono nello stesso contenuto logico di Pack devono produrre la stessa stringa di byte.

C5	Stringa di system-prompt (tutte e tre le passate)	La stringa di system-prompt completa e verbatim per ciascuna delle tre passate della pipeline, concatenata nell'ordine delle passate (Passata 1 || Passata 2 || Passata 3), con un byte nullo (`\x00`) come separatore tra le passate. Il system prompt è la stringa completa così come inviata all'API del provider, dopo la risoluzione del Pack e la sostituzione dei template.

Regola di concatenazione: I componenti sono uniti nell'ordine C1, C2, C3, C4, C5 senza alcun separatore tra i componenti. La stringa di byte risultante è codificata in UTF-8 prima dell'hashing.

Funzione di hash: SHA-256, che produce un digest di 32 byte (256 bit), codificato come stringa esadecimale minuscola di 64 caratteri.

> ESEMPIO (ILLUSTRATIVO, NON UNA CHIAMATA REALE)
>
> `input_string = ("claude-opus-4-5-20251101" [C1] + "claude-opus-4-5-20251101" [C2, stesso provider qui] + "claude-opus-4-5-20251101" [C3] + "<canonical-pack-yaml-bytes>" [C4] + "<pass1-prompt>\x00<pass2-prompt>\x00<pass3-prompt>" [C5]); bundle_hash = sha256(input_string.encode("utf-8")).hexdigest()` → es. `"a3f9c2...d17b"` (64 caratteri esadecimali). (Gli ID dei modelli e le stringhe di prompt qui sopra sono placeholder illustrativi.)

2. DOVE L'HASH È CONSERVATO

Ogni record T0 nell'**Osservatorio** contiene un campo `bundle_hash` al livello superiore del JSON del record:

> JSON DEL RECORD (SCHEMA_VERSION 1.0)

```
{ "schema_version": "1.0", "deal_id": "<osservatorio-deal-id>",
"t0_unix_utc": 1750000000, "bundle_hash": "a3f9c2...d17b", "pass_models": {
"pass_1_thesis": "claude-opus-4-5-20251101", "pass_2_red_team": "claude-
opus-4-5-20251101", "pass_3_synthesis": "claude-opus-4-5-20251101" },
"pack_version": "<git-commit-sha-of-resolved-pack>", "output": { ... } }
```

L'API pubblica dell'Osservatorio espone il campo `bundle_hash` per ciascun record. Un ricercatore che interroga il record T0 di un deal riceve l'hash direttamente; non ha bisogno di ricalcolarlo per confermarne la presenza.

Oltre alla conservazione nell'Osservatorio, l'hash di ciascun record T0 è inviato a **OpenTimestamps** (Todd, 2016), che produce una prova crittografica di esistenza ancorata a Bitcoin. Il file di prova OTS è conservato accanto al record nell'Osservatorio ed è anche committato nel repository pubblico.

3. COME UN TERZO VERIFICA LA VERSIONE DI UNA CHIAMATA

Una procedura di verifica completa è la seguente:

1. **Recuperare il record T0.** Interrogare l'API pubblica dell'Osservatorio per il deal di interesse. Annotare: `bundle_hash` (il digest dichiarato), `pass_models` (i tre ID dei modelli usati), `pack_version` (SHA del commit git del Pack risolto) e `t0_unix_utc` (il timestamp di finalizzazione).

2. **Recuperare gli artefatti committati.** Dal repository pubblico, effettuare il checkout del commit identificato da `pack_version`. Verificare che il timestamp del commit preceda `t0_unix_utc`. Recuperare lo YAML del Pack risolto a quel commit (applicando lo stesso ordine di risoluzione: core ▸ verticale ▸ cliente, come documentato in `packages/packs/README.md`) e i file template dei system-prompt per tutte e tre le passate a quel commit.

3. **Ricostruire lo YAML canonico del Pack.** Applicare la procedura di forma canonica: (1) fondere gli YAML dei Pack core, verticale e cliente nell'ordine di risoluzione; (2) ordinare tutte le chiavi lessicograficamente (ricorsivamente, tutte le mappe annidate); (3) serializzare in YAML 1.2 con terminazioni di riga LF e nessuno spazio in coda. Due implementazioni indipendenti della procedura di canonicalizzazione devono produrre la stessa stringa di byte per lo stesso contenuto logico di Pack.

4. **Ricostruire le stringhe di system-prompt.** Applicare la stessa logica di sostituzione dei template presente nel codice dell'engine a `pack_version` per produrre le tre stringhe di system-prompt verbatim. Nessuna sostituzione aggiuntiva è valida; il codice dell'engine è la fonte autorevole.

5. **Ricalcolare l'hash.** Concatenare C1–C5 nell'ordine specificato in §1 sopra (codificati in UTF-8, nessun separatore tra i componenti, separatori a byte nullo tra i prompt delle passate all'interno di C5). Calcolare SHA-256. La stringa esadecimale minuscola risultante deve essere uguale a `bundle_hash`.

6. **Verificare l'ancora di timestamp.** Se OpenTimestamps è operativo per il record, verificare il file di prova OTS rispetto alla blockchain Bitcoin usando il client OpenTimestamps standard. La prova stabilisce che il digest `bundle_hash` esisteva prima del blocco Bitcoin la cui altezza è registrata nella prova. Se OpenTimestamps non è ancora operativo (periodo pre-attivazione), il timestamp di registrazione OSF è l'ancora indipendente. Il record dell'Osservatorio è append-only; il timestamp di creazione non può essere alterato retroattivamente.

4. COSA L'HASH GARANTISCE E COSA NON GARANTISCE

Garanzie:

- [ ] Una corrispondenza tra l'hash ricalcolato e il `bundle_hash` conservato conferma che la previsione è stata prodotta usando esattamente gli ID dei modelli, lo YAML del Pack e

le stringhe di system-prompt rappresentati alla `pack_version` dichiarata.

- [ ] Una mancata corrispondenza indica che almeno uno dei cinque componenti differiva dalla versione dichiarata — sia tramite una modifica del prompt non dichiarata, sia tramite un aggiornamento del Pack, sia tramite una modifica dell'API del modello che ha alterato la stringa dell'ID del modello restituita.

Non garantisce:

- [ ] Che i pesi interni o il comportamento di campionamento del provider del modello fossero identici a quelli di qualunque altra chiamata effettuata con lo stesso ID del modello. Gli ID dei modelli identificano un checkpoint denominato; i provider possono aggiornare i checkpoint tra le release. L'hash vincola l'*identificatore dichiarato*, non lo stato interno del provider.
- [ ] Che il contenuto del system-prompt sia pedagogicamente o eticamente appropriato. L'hash conferma l'identità, non la qualità. L'ispezione del prompt resta una responsabilità umana.

RIMANDI

§4.3 formato di output sigillato (elemento 4 del bundle hash)

§8 congelamento di versione per chiamata come mitigazione strutturale del conflitto di interessi

§10.2 regola di versionamento per esteso

§10.3 ancoraggio OpenTimestamps

§6.4 regola di confine della coorte (bundle hash diversi → etichette di coorte diverse)

B Appendice B — scheda illustrativa (mostra il campo Δ-CSI; non mostra il bundle hash, che è metadato di infrastruttura)

APPENDICE D

Il registro della coorte sigillata (corsia A, N = 10 fisso)

> ***Edizione italiana:*** *in questa edizione l'appendice è riportata in italiano. Il testo di riferimento canonico del progetto resta la versione nel repository (`docs/paper/appendices/cohort-register.md`), coerentemente con la nota di rimando in §4.1 e §10.*

Questa appendice enumera la coorte fissa e chiusa dello studio richiamata in §4.1 e fissa, per deal e **prima dello scoring**, ogni parametro di risoluzione che il §5 lascerebbe altrimenti alla discrezionalità (valore di riferimento, valuta, indice settoriale, osservabilità H2, informazione pre-T0). Tutti e dieci i deal sono stati sigillati nell'Osservatorio il **20 giugno 2026**, prima che alcun esito H1 o H2 fosse osservabile. Ogni riga reca l'hash di contenuto pubblicato (SHA-256) della propria scheda sigillata; gli hash sono ispezionabili in modo indipendente sull'Osservatorio pubblico (`counterbrain.com/registro` · `/en/registry`).

La coorte è **congelata**: nessun deal può essere aggiunto o rimosso dopo il sigillo (§4.4). L'insieme è stato assemblato editorialmente (divulgato come COI-3, §8); la garanzia anti-senno-di-poi poggia sul sigillo pre-esito ancorato al timestamp OSF (§10.3), non su una regola di inclusione meccanica (§4.1).

D.1 — TABELLA DELLA COORTE

#	DEAL (ACQUIRENTE → TARGET)	VALORE ANNUNCIATO	SETTORE	NESSO DI QUOTAZIONE	ANNUNCIATO	SIGILLATO	RITARDO ANNUNCIO→SIGILLO	HASH DI SIGILLO (PREFISSO)
1	Intesa Sanpaolo → Monte dei Paschi	30,6 mld €	Bancario	IT (Euronext Milan)	2026-06-08	2026-06-20	12 gg	`df52792d…7051e42`
2	Poste Italiane → Telecom Italia	10,8 mld €	Telecom	IT (Euronext Milan)	2026-03-23	2026-06-20	~89 gg	`84322e5f…4436dd16`
3	UniCredit → Commerzbank	~35–40 mld €	Bancario	IT–DE (Milano / Francoforte)	2026-03-16	2026-06-20	~96 gg	`0eec6cbd…0140a34f`
4	Airbus · Leonardo · Thales → fusione satellitare (Project Bromo)	6,5 mld €	Difesa	FR–IT (Parigi / Milano)	2025-10-23 (MoU)	2026-06-20	~240 gg	`8e83ff8f…d0e8cd89`
5	Bouygues · Iliad · Orange → SFR (break-up)	20,35 mld €	Telecom	FR (Euronext Paris)	2026-06-06	2026-06-20	14 gg	`8fcc2fcc…7fd3aa98`
6	Nuveen → Schroders	9,9 mld £	Gestione del risparmio	UK (LSE)	2026-02-12	2026-06-20	~128 gg	`fff13730…ca507e55f`

7	Zurich → Beazley	8,1 mld £	Assicurazioni	UK–CH (LSE / SIX)	2026-03-02	2026-06-20	~110 gg	`7ee2b124…72b9eccd`
8	EQT → Intertek	10,9 mld £	Servizi (TIC)	UK (LSE)	2026-06-18	2026-06-20	2 gg	`29b9ab04…e0388e40`
9	Ingredion → Tate & Lyle	2,7 mld £	Ingredienti di consumo	UK (LSE)	2026-06-08	2026-06-20	12 gg	`480b7a2e…fe44aca9`
10	CVC → dsm-firmenich (Animal Nutrition & Health)	2,2 mld €	Nutrizione	CH–NL (SIX / Euronext Amsterdam)	2026-02-09	2026-06-20	~131 gg	`f16768a7…d48c8b0f`

> *Gli hash completi a 64 caratteri sono conservati verbatim in `apps/web/src/data/predictions.json` e bloccati in `apps/web/src/data/sealed-hashes.lock.json`.* ***Erratum sulla data di annuncio (#3, #4).*** *Le date qui mostrate sono corrette: il deal #3 (UniCredit–Commerzbank) è stato annunciato il 2026-03-16, e il deal #4 (Project Bromo) è datato al suo MoU del 2025-10-23. Il sigillo immutabile conserva di proposito i valori imprecisi originali (il deal #4 è sigillato come "2026"); poiché `announced_date` è dentro l'hash di sigillo, la correzione vive in questa documentazione e nel record OSF, non nella busta sigillata — il registro pubblico mostra quindi ancora l'originale. Entrambi i deal sono trattati come a informazione-pre-T0-elevata (D.3).*

D.2 — MAPPATURA SCENARIO → IPOTESI (LE P(H1), P(H2) SOTTOPOSTE A SCORING)

Ciascuna scheda sigillata indica una distribuzione di scenario a tre vie che somma al 100% (§4.2). Le probabilità sottoposte a scoring sono derivate deterministicamente: **P(H1) = massa dello scenario di fallimento del deal**; **P(H2) = massa dello scenario di delusione di valore** (massa marginale, non condizionata al completamento — vedi §4.2; sottoposta a scoring solo sul sottoinsieme H2 di D.3). Questi sono i valori *congelati e pubblicamente sigillati*; il successivo payload del contraddittore non può modificarli (§4.3, §9.3).

#	DEAL	P(SUCCESSO)	P(H2) SCORED	P(H1) SCORED	NOTA DI MAPPATURA
1	Intesa → MPS	0,35	0,40	0,25	pulita
2	Poste → TIM	0,30	0,45	0,25	pulita
3	UniCredit → Commerzbank	0,20	n/a (H2 non applicabile)	**0,80**	**eccezione**: un'offerta di controllo che finisce in una partecipazione di minoranza bloccata (0,45) o in ritiro/perdita (0,35) è un fallimento nell'acquisire il controllo → entrambe mappano su H1; successo = controllo (0,20).

4	Airbus/Leonardo/Thales	0,35	n/a (H2 non applicabile)	**0,35**	**eccezione** (era "pulita"): JV senza un singolo acquirente che consolida; H1 = "bloccata" (0,35); H2 indefinita → non applicabile.
5	Bouygues/Iliad/Orange → SFR	0,35	n/a (H2 non applicabile)	**0,30**	**eccezione** (era "pulita"): break-up a tre vie, nessun singolo acquirente; H1 = "bloccata / disfatta" (0,30); H2 indefinita → non applicabile.
6	Nuveen → Schroders	0,60	n/a (H2 non applicabile)	0,10	H1 pulita; H2 n/a (l'acquirente Nuveen è una controllata TIAA, nessun TSR pubblico).
7	Zurich → Beazley	0,55	0,35	0,10	pulita
8	EQT → Intertek	0,70	n/a (H2 non applicabile)	0,10	H1 pulita; H2 n/a (acquirente PE, target delistato).
9	Ingredion → Tate & Lyle	0,50	0,20	**0,30**	**eccezione**: divieto antitrust (0,30) codificato H1 (il deal non si completa ai termini originali); "si completa ma delude" (0,20) codificato H2.
10	CVC → dsm-firmenich ANH	0,45	n/a (H2 non applicabile)	0,15	H1 pulita; H2 n/a (acquirente PE, carve-out di una business unit).

Campione H1 sottoposto a scoring: tutti e dieci i deal (UniCredit incluso ma flaggato, D.3).
Campione H2 sottoposto a scoring (sottoinsieme): #1 Intesa, #2 Poste, #7 Zurich, #9 Ingredion — gli unici quattro deal con un acquirente quotato che consolida il target e un TSR pubblico. Gli altri sei sono *H2 non applicabile per progettazione* (§4.2): non censurati, non sottoposti a scoring.

> ***Nota sulle somme di riga.*** *Per i sei deal H2 non applicabile, la massa di delusione di valore esiste comunque nella distribuzione a tre vie sigillata — i tre scenari sommano al 100% nella scheda originale — ed è mostrata come n/a nella colonna "P(H2) scored" perché non è sottoposta a scoring (nessun acquirente quotato che consolida la rende osservabile), non perché la massa sia zero. In ogni riga, P(successo) + (massa di delusione di valore) + P(H1) = 100%; solo la scorabilità del termine intermedio cambia tra i deal. Eleggibilità ≠ risoluzione: ai sensi del §4.5/§6.5, se meno di tre dei quattro deal eleggibili a H2 si risolvono entro la finestra, nessuna statistica H2 è riportata — solo le righe grezze.*

D.3 — PARAMETRI DI RISOLUZIONE PER DEAL (FISSATI PRIMA DELLO SCORING)

Per ciascun deal: se H2 è osservabile, il valore di riferimento H1 e la sua valuta (la soglia ≥ 15% è misurata in questa valuta, mai dopo conversione; gli intervalli usano il punto medio), l'indice settoriale H2 e la valuta (TSR e indice nella stessa valuta), e il flag di informazione-pre-T0-elevata.

#	DEAL	H2 OSSERVABILE?	VALORE DI RIFERIMENTO H1 (VALUTA)	INDICE SETTORIALE H2 (VALUTA)	INFO PRE-T0 ELEVATA?
1	Intesa → MPS	sì	30,6 mld € (EUR)	FTSE Italia All-Share Banks, total return (EUR)	no (12 gg)
2	Poste → TIM	sì	10,8 mld € (EUR)	STOXX Europe 600 Financial Services, total return (EUR)	sì (~89 gg)
3	UniCredit → Commerzbank	no	37,5 mld € — punto medio dell'intervallo 35–40 mld € (EUR)	—	**sì** (ritardo lungo + bassa accettazione della prima offerta e opposizione del governo già pubbliche)
4	Airbus/Leonardo/Thales	no	6,5 mld € (EUR)	—	sì (ritardo lungo — MoU 8 mesi prima del sigillo)
5	Bouygues/Iliad/Orange → SFR	no	20,35 mld € (EUR)	—	no (14 gg)
6	Nuveen → Schroders	no	9,9 mld £ (GBP)	—	sì (~128 gg)
7	Zurich → Beazley	sì	8,1 mld £ (GBP)	STOXX Europe 600 Insurance, total return, in CHF (valuta di quotazione dell'acquirente)	sì (~110 gg)
8	EQT → Intertek	no	10,9 mld £ (GBP)	—	no (2 gg)
9	Ingredion → Tate & Lyle	sì	2,7 mld £ (GBP)	S&P Composite 1500 Food Products, total return (USD, valuta di quotazione dell'acquirente)	no (12 gg)
10	CVC → dsm-firmenich ANH	no	2,2 mld € (EUR)	—	sì (~131 gg)

> ***Sull'informazione-pre-T0-elevata.*** *Nessun deal aveva un segnale di risoluzione H1/H2 qualificante (cessazione, taglio headline ≥ 15%, impairment o breach del TSR) alla data di sigillo — tutti e dieci erano pendenti — sicché nessuno è escluso ai sensi della regola del §4.2. Il flag registra che alcuni deal recavano più informazione pubblica a T0 di quanta ne avrebbe avuta una chiamata all'istante-dell'annuncio, sicché lo*

> *svantaggio informativo non è uniforme; il ritardo annuncio-sigillo (D.1) è la covariata riportata nell'analisi esplorativa (§6.2). UniCredit–Commerzbank è il caso più esposto ed è flaggato di conseguenza, ma è mantenuto nel pool H1 perché nessun segnale H1 qualificante si era verificato a T0.*

> ***Sulle date imprecise (#3, #4).*** *I deal #3 e #4 recano date di annuncio pubblico a granularità mese/anno; il ritardo annuncio-sigillo è riportato a quella granularità ed entrambi sono trattati come a informazione-pre-T0-elevata. Le loro date di annuncio esatte saranno fissate nel record OSF prima dello scoring.*

APPENDICE E

Protocollo di derivazione del tasso base (null (a))

Questa appendice fornisce la **derivazione riproducibile, passo-passo** della singola cifra di tasso base abbinato p* usata da null (a) in §6.1, così che qualsiasi terzo possa ri-eseguirla da fonti pubbliche e ottenere lo stesso numero. La cifra eseguita, ogni query per fonte e il passo di combinazione sono congelati nel record OSF **prima** che qualsiasi deal sia sottoposto a scoring.

E.1 — COSA SI STA STIMANDO

p* = P(un fallimento di tipo H1 entro 12 mesi | un deal europeo large-cap contestato), dove un **fallimento di tipo H1** = { cessazione del deal **o** revisione al ribasso della consideration headline ≥ 15% **o** una proibizione regolamentare **o** un ritiro pre-decisione }. p* è la previsione costante di null (a): il livello minimo di Brier che il contraddittore deve battere. **Principio del livello minimo onesto:** poiché la coorte è assemblata editorialmente verso deal contestati (COI-3, §8), un p* *più alto* è il livello minimo più difficile e onesto; la regola di combinazione (E.5) **non fissa mai p* al di sotto della stima pulita della classe di riferimento**.

E.2 — CLASSE DI RIFERIMENTO

Un deal storico qualifica se tutti di: (1) valore annunciato ≥ 1 mld € (o equivalente £/CHF); (2) almeno una parte quotata su un mercato regolamentato europeo; (3) *contestato* — un'offerta concorrente/ostile/non sollecitata, oppure una revisione antitrust di Fase II o un'opposizione nazionale/politica documentata. Finestra: i dieci anni prima dell'anno di sigillo della coorte (**2016–2025**), o la sotto-finestra più ampia disponibile per fonte (dichiarata). La finestra termina prima dell'anno della coorte, così il tasso base usa altri deal, mai i dieci.

E.3 — RAFFINAMENTO OPZIONALE: DATI GREZZI UK TAKEOVER PANEL

Un censimento pubblico e pulito di deal con esiti per i takeover pubblici UK. Passi: (1) scaricare i dati grezzi delle transazioni del Panel per anno (pubblicati dal luglio 2023, dal 1° aprile 2020) più gli Annual Report per gli anni precedenti; (2) restringere alle firm offer soggette al City Code; (3) filtrare a ≥ 1 mld £ (il Panel riporta questo conteggio — es. 7 su 61 firm offer nel 2023–24); (4) numeratore = firm offer decadute o ritirate entro 12 mesi; (5) denominatore = tutte le firm offer ≥ 1 mld £ annunciate nella finestra;

(6) `p_TP = numeratore / denominatore`. Nota di scope: solo-UK; gli esiti delle firm offer catturano il canale rottura/ritiro, non le revisioni ≥ 15% post-completamento.

E.4 — CANALI DI CORROBORAZIONE (MAI PRIMARI)

(a) Commissione Europea — quota dei casi di *Fase II* che finiscono in proibizione o ritiro pre-decisione (≈ un terzo); il tasso di proibizione complessivo ≈ 0,3% è il riferimento sbagliato (dominato da deal non contestati). **(b) Industria / accademico** — cancellazione di grandi deal annunciati ≈ **10%/anno** (McKinsey 2019); gli offerenti ostili/non sollecitati prevalgono in una **minoranza** di casi (Schwert 2000; Bebchuk, Coates & Subramanian 2002); target più grandi meno propensi a completarsi. Corroborazione peer-reviewed sulle terminazioni: Xing & Yi (2024). **(c) newsapi.ai / Event Registry** (§5.2) — solo sanity-check dell'ordine di grandezza; le discrepanze si risolvono a favore di (a) / E.3.

E.5 — REGOLA DI COMBINAZIONE → IL SINGOLO P* CONGELATO

(1) Calcola `p_TP`. (2) Se l'evidenza del sottoinsieme contestato (E.4a, E.4b) indica un tasso materialmente più alto, fissa p* a quella stima abbinata (es. il punto medio della banda contestata); p* non è mai sotto `p_TP`. (3) Arrotonda a due decimali; congela p*, la finestra e ogni query di fonte in OSF prima dello scoring, dopodiché è fisso.

E.6 — IL VALORE PRE-REGISTRATO

Dalle cifre pubbliche già agli atti, puramente illustrativo:

CLASSE DI RIFERIMENTO	TASSO DI FALLIMENTO	RUOLO
Grandi deal annunciati (cancellazione/anno)	**≈ 10%**	**non** p* — la coorte non è un deal casuale
EC Fase II (regolamentare, contestati)	≈ 33%	corroborazione
Ostili / non sollecitate (accademico)	**prevale una minoranza**	limite superiore
Banda abbinata (large-cap UE contestati)	**≈ 25%–40%**	→ punto medio **≈ 0,30**

`p* = 0,30`, derivato dai tassi pubblicati di fallimento dei deal della Fase II UE e accademici (E.4) secondo il principio del livello minimo onesto. È il **valore pre-registrato**, non un'illustrazione; non richiede **alcuna raccolta di dati grezzi**, ed è confermato — e può solo essere alzato — al congelamento OSF.

Banda di sensibilità e caveat di classe di riferimento. Tutte le conclusioni su H1 della fase potenziata sono riportate per p* da 0,25 a 0,40; nessun risultato dipende dal valore puntuale. L'ancora ≈ ⅓ della Fase II è una quota di proibizione/ritiro, mentre H1 conta anche una revisione al ribasso del prezzo headline ≥ 15% (un canale reale non stimato separatamente), sicché p* = 0,30 è una stima centrale, non un tetto garantito conservativo.

E.7 — CHECKLIST DI CONGELAMENTO OSF

Registrare prima dello scoring: la finestra e le sotto-finestre per fonte; file/versione del Takeover Panel, filtro, numeratore, denominatore, `p_TP`; statistica EC, periodo, quota proibizione+ritiro di Fase II; citazione accademica + cifra usata; termini di query newsapi.ai, conteggio, verdetto; il p* finale, il suo arrotondamento e quale valore ha governato secondo E.5.

E.8 — FONTI (PUBBLICHE)

> *UK Takeover Panel — Transaction Statistics: thetakeoverpanel.org.uk/communications/transaction-statistics · Annual Reports: /communications/reports. Commissione Europea — statistiche e procedure di merger control: competition-policy.ec.europa.eu/mergers/procedures_en. Fonti industriali / accademiche sul fallimento dei deal: McKinsey (2019), Done Deal? Why Many Large Transactions Fail to Cross the Finish Line (≈ 10% di cancellazione dei grandi deal all'anno); Schwert (2000), Hostility in Takeovers, J. Finance 55(6):2599–2640; Bebchuk, Coates & Subramanian (2002), The Powerful Antitakeover Force of Staggered Boards, Stanford Law Rev. 54:887–951; Xing & Yi (2024), Mergers and Attributions, J. Management Studies, doi:10.1111/joms.13163.*

Bibliografia

1 Bebchuk, Lucian A.; Coates IV, John C.; and Subramanian, Guhan (2002). "The Powerful Antitakeover Force of Staggered Boards: Theory, Evidence, and Policy." *Stanford Law Review* 54: 887–951.

2 Brier, Glenn W. (1950). "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78(1): 1–3. doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

3 Cai, Alice; Arawjo, Ian; and Glassman, Elena L. (2024). "Antagonistic AI." arXiv:2402.07350 [cs.HC]. Submitted 12 February 2024. https://arxiv.org/abs/2402.07350

4 Canepa, Francesco Saverio (2026). *10 errori che distruggono valore nelle operazioni di M&A.* Independently published (Amazon KDP).

5 Good Judgment Inc. (2015). *Good Judgment Open.* Crowdsourced forecasting platform. Launched September 2015; founded by Philip E. Tetlock and Barbara Mellers. https://www.gjopen.com/

6 Hosmer, David W.; and Lemeshow, Stanley (1980). "Goodness of Fit Tests for the Multiple Logistic Regression Model." *Communications in Statistics — Theory and Methods* 9(10): 1043–1069. doi:10.1080/03610928008827941.

7 Kahneman, Daniel (2011). *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux. ISBN 978-0-374-27563-1.

8 McKinsey & Company (2019). "Done Deal? Why Many Large Transactions Fail to Cross the Finish Line." McKinsey Strategy & Corporate Finance. https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/done-deal-why-many-large-transactions-fail-to-cross-the-finish-line

9 Metaculus, Inc. (2015). *Metaculus: A Forecasting Platform and Aggregation Engine.* Founded 2015, Santa Cruz, CA. https://www.metaculus.com/

10 Murphy, Allan H. (1973). "A New Vector Partition of the Probability Score." *Journal of Applied Meteorology* 12(4): 595–600. doi:10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

11 Schwenk, Charles R. (1990). "Effects of Devil's Advocacy and Dialectical Inquiry on Decision Making." *Organizational Behavior and Human Decision Processes* 47(1): 161–176. doi:10.1016/0749-5978(90)90037-P.

12 Schwert, G. William (2000). "Hostility in Takeovers: In the Eyes of the Beholder?" *The Journal of Finance* 55(6): 2599–2640. doi:10.1111/0022-1082.00301.

13 Tetlock, Philip E.; and Gardner, Dan (2015). *Superforecasting: The Art and Science of Prediction.* New York: Crown. ISBN 978-1-101-90556-2.

14 Todd, Peter (2016). *OpenTimestamps: Scalable, Trust-Minimized, Distributed Timestamping with Bitcoin.* Project announcement and specification, https://opentimestamps.org/. Announced 15 September 2016.

15 Wu, Haolun; Li, Zhenkun; and Li, Lingyao (2025). "Can LLM Agents Really Debate? A Controlled Study of Multi-Agent Debate in Logical Reasoning." arXiv:2511.07784. Submitted 11 November 2025. https://arxiv.org/abs/2511.07784

16 Wynn, Andrea; Satija, Harsh; and Hadfield, Gillian (2025). "Talk Isn't Always Cheap: Understanding Failure Modes in Multi-Agent Debate." arXiv:2509.05396. Submitted 5 September 2025; v2 13 October 2025. Accepted, ICML Multi-Agent Systems Workshop 2025. https://arxiv.org/abs/2509.05396

17 Xing, Zhe (Adele); and Yi, Xiwei (2024). "Mergers and Attributions: An Examination of M&A Terminations in 1996–2022." *Journal of Management Studies.* doi:10.1111/joms.13163.