



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze e Ingegneria

Corso di Laurea in Informatica

L'impronta della personalità nell'interazione uomo-IA: uno studio sul co-design generativo basato sul modello OCEAN

Relatore:

Chiar.mo Prof.
Fabio Vitali

Candidato:

Emanuele Di Giacomo

Correlatrice:

Dott.ssa
Fatima Rabia Yapicioglu

Sessione Marzo 2026

Anno Accademico 2024/2025

Abstract

L'utilizzo sempre più massivo di strumenti basati su *IA generativa* sta ridefinendo i flussi di lavoro nel design di interfacce grafiche, dando origine a nuovi paradigmi di interazione uomo-IA in cui la progettazione si basa su processi di *co-design* assistiti dall'intelligenza artificiale. Tuttavia, i pesi che le differenze individuali hanno sul risultato ottenuto da una *GenAI* rimangono un campo ancora poco esplorato. Questo lavoro si propone di indagare come i tratti della personalità individuale, misurati tramite modello *OCEAN*, possano influenzare il processo di co-creazione di interfacce grafiche tramite IA.

La metodologia prevede uno studio sperimentale su un campione di 21 partecipanti. L'individuo, dopo aver completato il questionario *OCEAN*, svolge un task che consiste nella creazione di una Single Page Application (SPA) per l'acquisto di biglietti ferroviari, utilizzando il modello Claude Sonnet 4.5. L'analisi dei dati si concentra su tre blocchi distinti:

1. **Blocco P**: la forma dei *prompt* durante l'interazione uomo-macchina (lunghezza, specificità);
2. **Blocco D**: le scelte di *design* riflesse nell'artefatto generato (funzionalità, colori, font), misurate attraverso metriche quantitative ben definite;
3. **Blocco I**: le dinamiche *iterative* del processo (numero di revisioni).

I risultati suggeriscono che i tratti di personalità lascino una traccia osservabile nel processo di co-design, con intensità differente a seconda del dominio considerato. Il segnale più coerente emerge nel Blocco P, dove profili caratterizzati da impulsività motivazionale e calore relazionale tendono a formulare prompt più

concisi e visivamente orientati, mentre profili ad alta Coscienziosità producono testi più articolati. Sul piano del design, la Coscienziosità si associa a interfacce strutturalmente più ricche, mentre il Nevroticismo tende a correlare con output più semplici e meno dinamici. Le ipotesi iniziali relative all'Apertura alle Esperienze come predittore principale delle scelte estetiche e al Nevroticismo come driver di un maggior numero di revisioni non trovano piena conferma nei dati, suggerendo che le relazioni tra personalità e co-design generativo siano più mediate e multidimensionali di quanto previsto. La ricerca si propone come contributo esplorativo alla comprensione del co-design generativo, avanzando l'ipotesi che l'IA possa funzionare non come strumento neutro, ma come specchio psicologico che riflette, almeno parzialmente, i tratti individuali di chi la guida.

Indice

1	Introduzione	1
2	Inquadramento Teorico e Strumenti	7
2.1	Il Modello OCEAN (Big Five)	7
2.2	La GenAI e i Large Language Models (LLM)	9
2.3	Il Paradigma del Co-Design Generativo	13
2.4	L'Utente come Variabile di Design	16
3	Metodologia	19
3.1	Il Disegno della Ricerca	19
3.2	Il Campione	21
3.3	Somministrazione del Test OCEAN	23
3.4	Il Task di Co-Design	24
3.5	La Raccolta dei Dati	25
3.6	Le Metriche di Analisi	27
3.6.1	Analisi del Prompt (P)	28
3.6.2	Analisi del Design (D)	30
3.6.3	Analisi dell'Iterazione (I)	35
4	Analisi dei Dati	37
4.1	Introduzione al Capitolo Analitico	37
4.2	Analisi Esplorativa dei Dati (EDA)	39
4.2.1	Distribuzione punteggi OCEAN nel campione	39
4.2.2	EDA - Blocco P (Prompt)	43
4.2.3	EDA - Blocco D (Design)	47

4.2.4	EDA - Blocco I (Iterazione)	51
4.2.5	Sintesi EDA e implicazioni	54
4.3	Scelta dello Strumento Correlazionale	54
4.3.1	Correlazioni Bivariate	54
4.3.2	Giustificazione del PLS Regression	55
4.4	Analisi PLS Regression	56
4.4.1	Configurazione del Modello PLSR	56
4.4.2	Risultati - Blocco P (Prompt)	57
4.4.3	Risultati - Blocco D (Design)	59
4.4.4	Risultati - Blocco I (Iterazione)	61
5	Discussione e Conclusioni	65

Capitolo 1

Introduzione

Il Cambio di Paradigma nel Design. Il panorama contemporaneo del design sta attraversando una trasformazione senza precedenti, che ridefinisce non solo i flussi di lavoro, ma l'essenza stessa dell'atto creativo. Per decenni, il rapporto tra il progettista e i propri strumenti è stato mediato dal paradigma del *Computer-Aided Design* (CAD) [1].

In questo scenario, il software agiva come un'estensione passiva della mano dell'utente: un set di strumenti deterministici destinati a tradurre istruzioni tecniche precise in output grafici. Il controllo era assoluto, ma la responsabilità della "scintilla" ideativa e della sua esecuzione formale ricadeva interamente sul singolo individuo, in un processo di creazione spesso solitario dinanzi al rigore di una tela bianca.

Oggi, l'avvento dell'intelligenza artificiale generativa (GenAI) sembra segnare il superamento di questo modello, introducendo nuove dinamiche di *Co-Design Generativo*. Non siamo più di fronte a uno strumento che si limita a eseguire, ma a un'entità partner capace di interpretare, elaborare e visualizzare stimoli e intuizioni in tempo reale [2].

La tela bianca tende a trasformarsi in uno spazio di dialogo collaborativo, dove l'interazione si sposta dal piano della mera esecuzione a quello della [3]. In questo nuovo ecosistema, rappresentato in questo studio dall'utilizzo di modelli avanzati come Claude Sonnet 4.5 e la sua funzione *Artifacts* [4], la velocità di iterazione permette un riscontro immediato tra l'input dell'utente e la risposta del sistema.

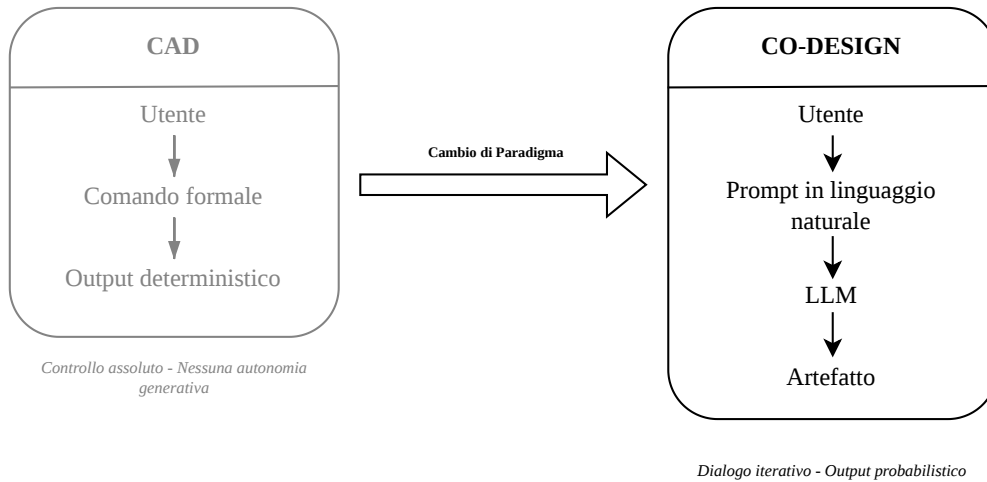


Figura 1.1: *Confronto tra il paradigma del Computer-Aided Design (CAD) e il co-design generativo mediato da LLM.*

La Dimensione Umana nell’Interazione. Nonostante la natura relazionale della GenAI, una parte significativa del dibattito attuale tende a considerare l’intelligenza artificiale principalmente come uno strumento esecutivo di alta precisione. In quest’ottica, l’attenzione della ricerca e della pratica si è spesso concentrata sull’ottimizzazione del *Prompt Engineering*: lo studio di come formulare istruzioni tecnicamente efficaci per ottenere il miglior risultato possibile dal modello [5].

Tuttavia, emerge uno spazio di indagine meno esplorato che riguarda il soggetto che opera dietro il prompt. L’atto di istruire una macchina, pur basandosi su un linguaggio logico, rimane un’attività umana mediata da inclinazioni individuali. Se due utenti diversi affrontano il medesimo compito progettuale, sia esso la creazione di un’interfaccia per l’acquisto di biglietti ferroviari, è probabile che le loro istruzioni presentino variazioni non solo tecniche, ma anche stilistiche e comportamentali.

L’ipotesi alla base di questo lavoro è che il fattore umano, inteso come l’insieme dei tratti che caratterizzano la personalità dell’utente, possa rappresentare una variabile rilevante nel processo di co-design.

L’IA come Specchio Psicologico. A fronte della complessità descritta, la presente tesi propone di superare la visione dell’intelligenza artificiale come mera

scatola nera esecutiva, ipotizzando che l'interazione con i modelli generativi possa essere letta come una proiezione delle caratteristiche individuali dell'utente. La soluzione risiede nel considerare l'IA non come uno strumento neutro, bensì come uno "specchio psicologico" capace di riflettere e rendere tangibili i tratti latenti di chi la interroga.

L'ipotesi centrale di questo lavoro suggerisce che la personalità dell'utente, mappata scientificamente attraverso il modello Big Five (OCEAN) che scompone la personalità umana in cinque grandi dimensioni [6], agisca come la variabile indipendente fondamentale nel determinare la forma, il tono e l'esito del processo di co-design. In questo paradigma, il modello generativo adatta le proprie capacità probabilistiche agli input ricevuti, i quali risultano carichi delle inclinazioni psicologiche del progettista.

Adottare il modello OCEAN permette di scomporre l'astrattezza dell'estro creativo in dimensioni misurabili e correlate. Si ipotizza, ad esempio, che un'alta Apertura possa spingere il risultato verso un'estetica più audace e non convenzionale, mentre un'elevata Coscienziosità possa tradursi in prompt più strutturati e orientati al rigore funzionale. Al contrario, livelli significativi di Nevroticismo potrebbero manifestarsi in una dinamica di revisione più frequente e cauta, rivelando un bisogno di controllo costante sull'output e una tendenza alla semplificazione del design.

In definitiva, questo studio intende suggerire che l'output visivo e funzionale di una *Single Page Application* (SPA) non è solo il prodotto di un calcolo algoritmico, ma la firma digitale di un incontro unico tra l'architettura logica della macchina e l'architettura attitudinale dell'uomo. Analizzare questo riflesso significa rendere empiricamente esplorabile il co-design generativo come un processo di espressione del sé, dove la tecnologia, nel caso specifico rappresentata dalle potenzialità di Claude Sonnet 4.5, funge da catalizzatore per la traduzione della personalità in forma visibile.

Obiettivi e Research Questions. Alla luce del mutato scenario tecnologico, il presente lavoro si pone l'obiettivo generale di indagare la natura del legame tra la struttura psichica dell'utente e l'output generato dai modelli linguistici di

grandi dimensioni (LLM). Nello specifico, si intende verificare se, e in che misura, i tratti di personalità rilevati tramite il modello Big Five (OCEAN) agiscano come variabili indipendenti capaci di modellare l'esperienza di co-design su Claude Sonnet 4.5.

L'indagine si articola attorno a quattro quesiti di ricerca fondamentali, che mirano a coprire l'intero spettro dell'interazione, dal momento dell'input testuale alla definizione dell'interfaccia visiva:

- **RQ1 (Dimensione Linguistica): In che modo i tratti del modello OCEAN influenzano la struttura e la semantica dei prompt?**

L'analisi mira a determinare se profili con alta Gradevolezza tendano a utilizzare un registro più cortese e collaborativo, o se tratti come la Coscienziosità si traducano in istruzioni caratterizzate da maggiore precisione tecnica e rigore sintattico, influenzando così la chiarezza del compito affidato alla macchina;

- **RQ2 (Dimensione Estetico-Funzionale): Esiste una correlazione significativa tra i tratti psicologici e le scelte di design riflesse nel codice generato?**

Si vuole esplorare se la firma psicologica dell'utente sia rintracciabile negli elementi formali della Single Page Application (SPA), quali la palette cromatica, la scelta dei font e la complessità dell'architettura informativa. Ad esempio, ci si interroga se l'Apertura alle Esperienze correli con soluzioni visive più audaci e sperimentali rispetto a profili più conservativi;

- **RQ3 (Dinamica Iterativa): Come influenzano i tratti della personalità il flusso di lavoro e la persistenza nel raffinamento dell'output?**

Si intende analizzare se alti livelli di Nevroticismo o di Coscienziosità portino a un numero maggiore di iterazioni e a una minore accettazione del primo output, indicando una propensione al perfezionismo o, al contrario, un'insicurezza procedurale nell'interazione con il modello;

- **RQ4 (Dimensione del Sé): Può l’output della GenAI essere considerato un *proxy comportamentale* dell’utente?**

Infine, la ricerca assume una prospettiva di sintesi, cercando di esplorare se l’artefatto finale, pur essendo prodotto da un LLM, possa essere interpretato come una proiezione digitale del sé di chi lo ha utilizzato, superando la visione dell’IA come strumento neutro per abbracciarla come specchio riflessivo.

Il carattere di questa tesi è dichiaratamente esplorativo: non si mira a stabilire leggi deterministiche universali, quanto piuttosto a mappare un territorio di confine tra psicologia della personalità e interazione uomo-macchina, fornendo basi empiriche per futuri sistemi di IA capaci di adattarsi proattivamente all’utente.

Sintesi Metodologica. Per rinsaldare le ipotesi teoriche presentate e rispondere alle domande di ricerca formulate, lo studio si avvale di un approccio empirico basato sulla triangolazione di dati psicometrici, linguistici e tecnici. La metodologia è stata strutturata per osservare l’interazione uomo-IA in un ambiente controllato ma flessibile, capace di far emergere le sfumature individuali nel processo di co-design.

Il disegno sperimentale ha coinvolto un campione di 21 partecipanti, i quali sono stati inizialmente sottoposti a un test di personalità basato sul modello dei *Big Five* (OCEAN). Questa prima fase ha permesso di delineare il profilo psicometrico di ogni utente, stabilendo le variabili indipendenti della ricerca attraverso i punteggi ottenuti nelle dimensioni di Nevroticismo, Estroversione, Apertura alle esperienze, Gradevolezza e Coscienziosità.

Successivamente, ogni partecipante è stato impegnato in un task di progettazione di una Single Page Application (SPA) dedicata all’acquisto di biglietti ferroviari. L’interazione è avvenuta attraverso il modello Claude Sonnet 4.5, sfruttando la funzionalità Artifacts per il rendering immediato del codice HTML, CSS e JavaScript. La scelta di questo specifico strumento è motivata dalla sua capacità di fornire un feedback visivo istantaneo, permettendo all’utente di percepire l’IA come un partner proattivo e di reagire in tempo reale agli output generati.

Il protocollo di ricerca è stato definito come semi-strutturato: pur imponendo requisiti funzionali minimi, ha garantito ai partecipanti piena libertà espressiva in termini di estetica, tono di voce e complessità delle funzionalità. I dati raccolti includono:

- i log completi delle conversazioni (analisi dei *prompt*);
- gli artefatti tecnici prodotti (analisi delle scelte di design tradotte in codice).

L'analisi comparativa di questi dataset mira a verificare se, e in che misura, l'output della GenAI possa essere interpretato come un riflesso delle predisposizioni psicologiche dell'utente, validando empiricamente il paradigma dell'IA come "specchio" della personalità.

Capitolo 2

Inquadramento Teorico e Strumenti

2.1 Il Modello OCEAN (Big Five)

Nel campo della psicologia della personalità, il modello dei Big Five (noto anche come modello OCEAN) rappresenta attualmente lo standard per la descrizione e la misurazione delle differenze individuali in ambito scientifico [7]. La scelta di adottare questo paradigma all'interno della presente ricerca non è casuale: a differenza di altri test psicometrici più settoriali o meno validati scientificamente (come il test di Condizionamento di Pavlov), il modello OCEAN offre una struttura robusta, replicabile e ampiamente documentata nella letteratura accademica internazionale [8].

Sviluppato e perfezionato a partire dagli anni '80, principalmente grazie ai lavori di Costa e McCrae [6, 9], il modello postula che la personalità umana possa essere sintetizzata in cinque grandi dimensioni ortogonali tra loro. Tali tratti sono considerati tendenze comportamentali relativamente stabili nel tempo, capaci di influenzare il modo in cui l'individuo percepisce l'ambiente e, di conseguenza, come interagisce con gli strumenti a sua disposizione [10].

Di seguito si delineano le cinque dimensioni, analizzandone brevemente il potenziale impatto nel contesto creativo:

- *Apertura alle esperienze (Openness to Experience)*: Misura il grado di curiosità intellettuale, sensibilità estetica e preferenza per la varietà. In un task

di co-design, ci si aspetta che soggetti con alta Apertura tendano a esplorare soluzioni visive non convenzionali e a sfruttare la GenAI come motore di divergenza creativa;

- *Coscienziosità (Conscientiousness)*: Riguarda il livello di organizzazione, persistenza e motivazione nel perseguimento degli obiettivi. Individui coscienti tendono a essere meticolosi nella definizione dei requisiti e nella struttura logica del progetto, influenzando probabilmente la precisione tecnica dei prompt inviati;
- *Estroversione (Extraversion)*: Indica l'energia, la socievolezza e la tendenza a cercare stimoli esterni. Nell'interazione uomo-macchina, questo tratto potrebbe riflettersi nella lunghezza dei prompt e in un approccio più "discorsivo" e frequente con l'LLM;
- *Gradevolezza (Agreeableness)*: Rappresenta la disponibilità verso gli altri, la fiducia e la cooperazione. Sebbene l'IA sia un'entità non umana, la ricerca suggerisce che tratti di alta gradevolezza possano tradursi in un tono più cortese (uso di formule di mitigazione) anche durante l'immissione di comandi tecnici;
- *Nevroticismo (Neuroticism)*: Valuta la stabilità emotiva rispetto all'insicurezza e all'ansia. Livelli elevati di nevroticismo potrebbero manifestarsi in una minore tolleranza verso l'errore dell'IA, portando a un numero maggiore di iterazioni e a un processo di raffinamento più frammentato o prudente.

La validità scientifica di questo modello risiede nella sua natura cross-culturale e nella sua capacità predittiva [9]. Come si osserverà nel corso della trattazione, il modello OCEAN non serve qui a etichettare l'utente, ma a fornire una lente analitica per interpretare la variabilità degli output generati da Claude. L'ipotesi di fondo è che, sebbene lo strumento tecnologico sia identico per tutti i partecipanti, il "filtro" della personalità individuale produca percorsi di progettazione unici, rendendo l'interfaccia finale una proiezione tangibile del profilo psicometrico del suo autore.

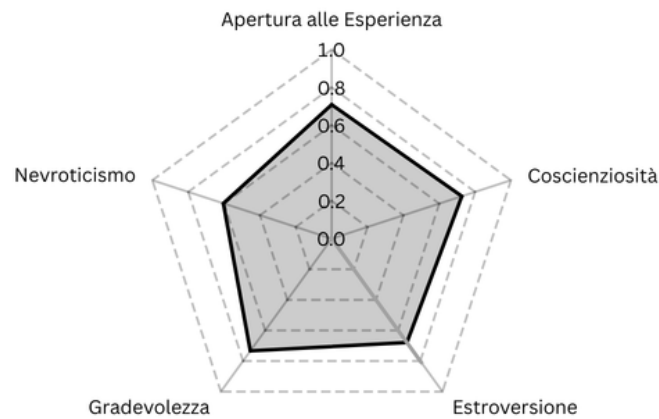


Figura 2.1: *Profilo medio del campione ($N = 21$) sul modello OCEAN (Big Five), rappresentato su scala normalizzata 0–1. I valori riportati esprimono la media aritmetica dei punteggi normalizzati per ciascuna dimensione (Costa e McCrae, 1992).*

2.2 La GenAI e i Large Language Models (LLM)

Se la sezione precedente ha definito il profilo psicologico dell’utente come variabile indipendente di questo studio, è ora necessario contestualizzare lo strumento tecnologico attraverso cui tale profilo si esprime e si rende osservabile. Comprendere la natura dei modelli linguistici di grandi dimensioni e le specificità della piattaforma adottata non è un esercizio puramente tecnico, ma una premessa indispensabile per interpretare correttamente i dati raccolti e le dinamiche di interazione che ne derivano.

Per comprendere appieno le potenzialità di questi strumenti, è utile ripercorrerne brevemente la traiettoria evolutiva. I sistemi di intelligenza artificiale orientati al linguaggio naturale hanno attraversato nel corso degli ultimi decenni trasformazioni radicali sul piano architetturale e applicativo. I primi modelli di Natural Language Processing (NLP), sviluppati a partire dagli anni Novanta, erano fondamentalmente sistemi discriminativi: classificavano testi, estraevano entità, riconoscevano pattern sintattici [11]. La generazione di testo coerente e contestualmente appropriato rimaneva un obiettivo distante, ostacolato dall’incapacità di questi sistemi di modellare le dipendenze a lungo raggio tra le parole di una sequenza [12]. La svolta decisiva è avvenuta nel 2017, con la pubblicazione del

paper "Attention Is All You Need" [13] da parte di ricercatori di Google, che ha introdotto l'architettura Transformer. Questo paradigma, basato sul meccanismo di self-attention, ha consentito ai modelli di pesare dinamicamente la rilevanza di ogni elemento di una sequenza rispetto a tutti gli altri, superando i limiti delle precedenti architetture ricorrenti.

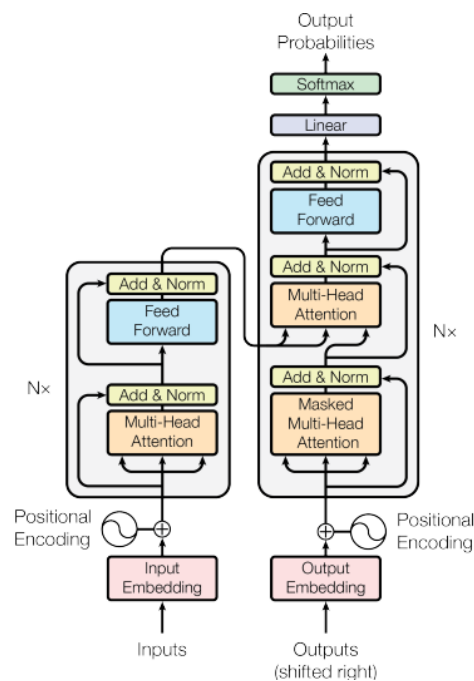


Figura 2.2: Architettura del Transformer. Da "Attention is All You Need" - Google.

Da questo punto in poi, la scala dei modelli addestrati su questa architettura è cresciuta esponenzialmente: i Large Language Models (LLM), tra cui GPT-3, PaLM e Claude, sono il prodotto diretto di questa evoluzione. Addestrati su corpus testuali di dimensioni miliardarie attraverso tecniche di apprendimento auto-supervisionato, questi modelli sviluppano rappresentazioni distribuite del linguaggio che permettono loro non solo di completare testi, ma di ragionare, sintetizzare, tradurre e, per quanto rilevante in questo studio, generare codice funzionante.

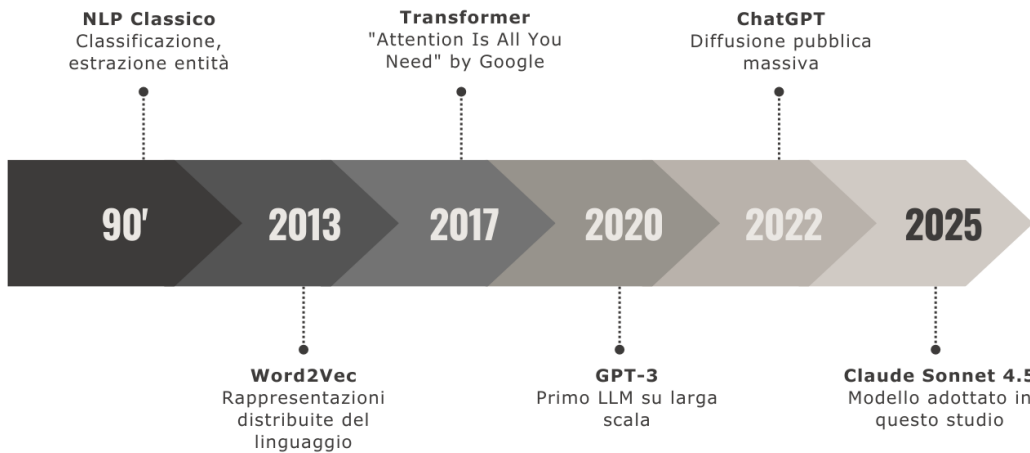


Figura 2.3: *Evoluzione dei principali paradigmi nei modelli linguistici, dall’NLP classico degli anni Novanta ai Large Language Models contemporanei. In evidenza Claude Sonnet 4.5, lo strumento adottato nel presente studio.*

Quest’ultima capacità merita un approfondimento specifico, in quanto costituisce il fondamento tecnico dell’intero disegno sperimentale. Il codice HTML, CSS e JavaScript condivide con il linguaggio naturale una struttura sequenziale regolata da regole sintattiche [14], il che lo rende accessibile alla modellazione probabilistica degli LLM. Studi recenti hanno evidenziato come modelli sufficientemente grandi, addestrati su repository di codice sorgente oltre che su testo, siano in grado di generare interfacce funzionanti a partire da descrizioni in linguaggio naturale [15, 16]. Questa capacità di traduzione da linguaggio intenzionale a linguaggio formale è centrale per questa ricerca: il partecipante esprime un’intenzione progettuale in linguaggio naturale, e il modello la traduce in un artefatto visivo e interattivo. All’interno dell’ecosistema degli LLM disponibili, la presente ricerca ha adottato Claude Sonnet 4.5, sviluppato da Anthropic, come strumento di co-design.

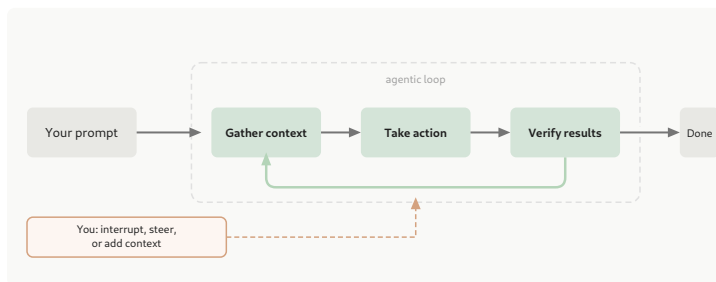


Figura 2.4: *Agentic Loop di Claude Sonnet 4.5. Dalla documentazione di Anthropic.*

Questa scelta è motivata da ragioni sia tecniche sia metodologiche. Sul piano tecnico, Claude appartiene alla famiglia di modelli che adottano il paradigma Constitutional AI [17], un approccio all’addestramento che privilegia la coerenza e la sicurezza delle risposte, riducendo la variabilità indesiderata degli output. Questa caratteristica è particolarmente rilevante in un contesto sperimentale, dove si desidera che le differenze negli artefatti prodotti siano attribuibili prevalentemente alla variabilità nei prompt degli utenti, e quindi alla loro personalità, piuttosto che a instabilità intrinseche del modello. Sul piano metodologico, l’elemento più significativo è rappresentato dalla funzione Artifacts: una modalità di interazione che consente al modello di generare e renderizzare in tempo reale blocchi di codice HTML, CSS e JavaScript direttamente nell’interfaccia di conversazione. A differenza di una chat testuale standard, in cui il codice viene restituito come testo inerte che l’utente deve copiare, compilare ed eseguire autonomamente, la funzione Artifacts collassa il ciclo di feedback in un’unica finestra visiva. Il partecipante formula un prompt, il modello genera il codice e l’interfaccia grafica appare immediatamente a fianco della conversazione, senza alcuna mediazione tecnica aggiuntiva [4]. Questa caratteristica ha implicazioni non banali per il disegno della ricerca. Il rendering immediato trasforma l’interazione da un processo astratto, in cui l’utente immagina il risultato senza vederlo, a un processo percettivo e reattivo, in cui la risposta visiva dell’artefatto stimola nuove intenzioni progettuali. È plausibile ipotizzare che questo meccanismo amplifichi le differenze individuali: un utente con alta Apertura alle esperienze potrebbe reagire all’output visivo con stimoli verso una maggiore sperimentazione, mentre un individuo con alta Coscienziosità potrebbe concentrarsi su revisioni di precisione e corre-

zione di dettagli funzionali. La piattaforma, in questo senso, non è uno sfondo neutro dell'esperimento, ma un catalizzatore delle predisposizioni individuali. È opportuno, infine, affrontare una questione metodologica che potrebbe essere sollevata rispetto all'uso di uno strumento commerciale in un contesto di ricerca: in che misura le caratteristiche intrinseche di Claude Sonnet 4.5 possono costituire una variabile confondente? Il rischio teorico è che lo stile generativo del modello, inteso come le sue preferenze implicite in termini di struttura del codice, palette cromatiche di default o layout tipografici, possa livellare le differenze tra i partecipanti, riducendo la variabilità attribuibile alla personalità. Si riconosce questa limitazione come inerente al disegno sperimentale, e la si discuterà più approfonditamente nella sezione metodologica dedicata. Tuttavia, si osserva che studi analoghi sull'interazione uomo-LLM hanno evidenziato come le variazioni nel modo in cui gli utenti formulano le proprie richieste, in particolare nella specificità, nel tono e nella struttura, producano output significativamente differenti anche a parità di modello utilizzato [18]. Questo suggerisce che la "firma" dell'utente nel prompt sia sufficientemente informativa da lasciare una traccia nell'artefatto generato [19], rendendo l'obiettivo della ricerca metodologicamente sostenibile.

2.3 Il Paradigma del Co-Design Generativo

Il concetto di co-design non nasce con l'intelligenza artificiale generativa. Le sue radici affondano in una tradizione di ricerca che, a partire dagli anni Settanta, ha progressivamente messo in discussione il modello del progettista come unico soggetto creativo, riconoscendo il valore della partecipazione e della co-creazione tra attori con competenze eterogenee [20]. In questo paradigma originario, il "co" rimandava alla collaborazione tra persone: utenti finali, committenti, esperti di dominio. L'avvento dei modelli linguistici di grandi dimensioni introduce una discontinuità radicale in questa tradizione: per la prima volta, il partner nel processo co-creativo non è un soggetto umano, ma un sistema computazionale capace di elaborare e restituire contenuti con un livello di coerenza e pertinenza contestuale sufficientemente elevato da simulare, almeno sul piano funzionale, le dinamiche di

un dialogo progettuale. Si parlerà pertanto di co-design generativo per indicare quel processo in cui la produzione di un artefatto visivo o funzionale non è il risultato di un singolo atto creativo, bensì l'esito di una conversazione iterativa tra un utente umano e un LLM [21]. Per comprendere la portata di questa transizione, è utile contestualizzarla rispetto al paradigma che l'ha preceduta. Il Computer-Aided Design (CAD), nella sua accezione classica, definisce un rapporto tra uomo e macchina fondato sul principio del controllo deterministico: l'utente impartisce istruzioni precise (comandi, coordinate, parametri) e il software le esegue con fedeltà assoluta, senza interpretare né aggiungere. In questo scenario, la variabilità del risultato finale è interamente imputabile alla variabilità dell'input umano, il che rende il sistema uno strumento trasparente, nel senso di non introdurre autonomia generativa. Il co-design mediato dalla GenAI sovverte questo equilibrio su almeno due piani. In primo luogo, l'input non è più un comando formale ma un enunciato in linguaggio naturale, soggetto alle ambiguità, alle implicature e alle sfumature pragmatiche tipiche della comunicazione umana. In secondo luogo, la risposta del modello non è una traduzione univoca dell'input, ma una generazione probabilistica che attinge a un vasto corpus di conoscenza pregressa, introducendo nel processo un contributo che non era esplicitamente contenuto nell'istruzione originale [22]. L'utente, di fronte all'output visivo generato, non valuta quindi solo se le sue istruzioni sono state eseguite correttamente, ma reagisce a una proposta estetica e funzionale che si sovrappone parzialmente alle sue intenzioni, modificando retroattivamente le sue aspettative e orientando i prompt successivi.

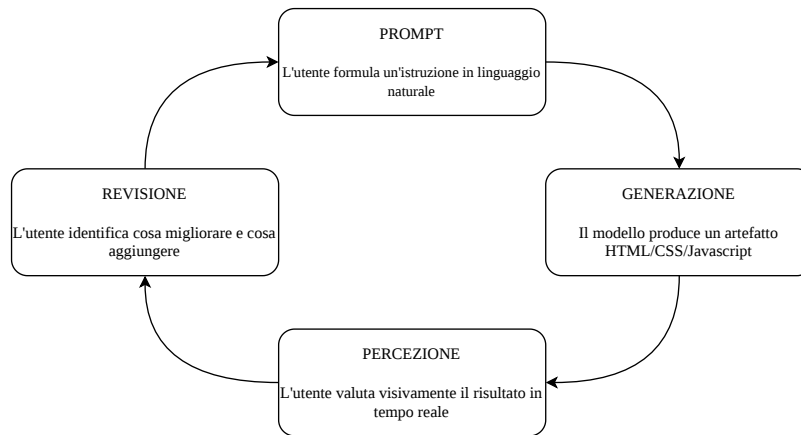


Figura 2.5: *Il ciclo iterativo del co-design generativo.*

È in questo contesto che il concetto di prompt engineering assume una valenza che va oltre la mera competenza tecnica. Nella letteratura più recente, l'atto di formulare un prompt efficace è stato analizzato prevalentemente come un problema di ottimizzazione: come strutturare una richiesta per massimizzare la qualità dell'output [5]. Questa prospettiva, tuttavia, tende a trascurare una dimensione di natura diversa. Il prompt è, prima di essere un'istruzione tecnica, un atto linguistico attraverso cui l'utente non solo descrive un obiettivo, ma esprime una postura comunicativa e una visione implicita del problema. La scelta di un imperativo diretto ("crea una navbar") rispetto a una forma mitigata ("potresti aggiungere una navbar in alto?") non è neutra: rivela un diverso rapporto con lo strumento e, secondo l'ipotesi di questo studio, un diverso profilo di personalità. Il prompt engineering, in questa lettura, non è soltanto una pratica procedurale ma anche espressiva, il cui contenuto latente risulta tanto informativo quanto quello manifesto. Il co-design generativo si configura quindi come un processo bidirezionale. L'utente guida attraverso la formulazione del prompt; il modello risponde con un artefatto che, a sua volta, ridefinisce lo spazio di possibilità percepito dall'utente, orientando le iterazioni successive. In letteratura, questo meccanismo è stato descritto come un ciclo di action-reaction in cui la persistenza dell'utente nel raffinamento dell'output e la sua capacità di reagire costruttivamente agli errori o alle divergenze del modello costituiscono variabili critiche per la qualità del risul-

tato finale [23]. È precisamente in questo ciclo iterativo che si manifesta, secondo l'ipotesi alla base della presente ricerca, l'influenza dei tratti di personalità: non come condizione iniziale statica, ma come variabile dinamica che si esprime diversamente a ogni nodo del dialogo uomo-macchina, modulando tanto la struttura linguistica dei prompt quanto la soglia di accettazione degli output intermedi.

2.4 L'Utente come Variabile di Design

Se i paragrafi precedenti hanno delineato i due pilastri concettuali di questa ricerca, ovvero il modello psicometrico OCEAN e l'architettura dei sistemi di co-design generativo, il presente paragrafo si propone di esplorare lo spazio teorico in cui questi due domini si incontrano: la relazione tra tratti di personalità e scelte estetiche e funzionali nell'interazione con sistemi digitali. Si tratta, come anticipato, del territorio teoricamente più avanzato dell'inquadramento, in quanto la letteratura specifica su questo incrocio è ancora in fase di consolidamento [24]. Tuttavia, esistono contributi consolidati che, pur non affrontando direttamente il co-design mediato da GenAI, offrono basi empiriche sufficientemente solide per formulare ipotesi plausibili e metodologicamente sostenibili. La relazione tra personalità e preferenze estetiche ha attirato l'attenzione della ricerca psicologica già a partire dagli studi sul tratto di Apertura alle Esperienze [25]. Nella tradizione inaugurata da Costa e McCrae, questo tratto è stato associato non solo alla curiosità intellettuale, ma anche a una spiccata sensibilità estetica e a una maggiore tolleranza verso la complessità visiva e concettuale. Individui con punteggi elevati in Apertura tendono a preferire stimoli più articolati, non convenzionali e ricchi di variazione, mentre soggetti con punteggi più bassi mostrano generalmente una preferenza per configurazioni più semplici, ordinate e familiari. Applicato al contesto del design digitale, questo modello suggerisce che la firma visiva di un'interfaccia possa non essere neutrale rispetto al profilo psicologico di chi l'ha commissionata o progettata, ma possa invece riflettere, in modo più o meno consapevole, le predisposizioni percettive ed estetiche del suo autore. La Coscienziosità, d'altra parte, introduce una dimensione diversa ma ugualmente rilevante. Studi condotti nell'ambito del-

l'interazione uomo-computer hanno evidenziato come individui con alti punteggi in questo tratto tendano a strutturare le proprie interazioni con i sistemi digitali in modo più sistematico, a pianificare i propri obiettivi con maggiore anticipo e a preferire interfacce che privilegino la chiarezza funzionale rispetto alla ricchezza decorativa [26]. In questo senso, si può ipotizzare che la Coscienziosità agisca come variabile moderatrice non solo nella struttura dei prompt, ma anche nella direzione estetica e nell'architettura informativa dell'artefatto finale, orientando le scelte verso una maggiore semplicità organizzativa e una gerarchia visiva più esplicita. Anche il Nevroticismo risulta pertinente in questa prospettiva, sebbene in modo meno diretto. La letteratura sull'*HCI* (*Human-Computer Interaction*) suggerisce che individui con alta instabilità emotiva tendano a manifestare una minore fiducia nei confronti delle proprie scelte, a mostrare una maggiore sensibilità agli errori e agli esiti inattesi, e a richiedere più frequentemente conferme esterne durante l'esecuzione di un compito [27]. Tradotto nel contesto del co-design generativo, questo pattern comportamentale potrebbe riflettersi non tanto in specifiche scelte cromatiche o tipografiche, quanto in una dinamica iterativa più frammentata, in cui il numero di revisioni richieste all'IA e la tendenza a semplificare o ridurre la complessità del progetto potrebbero costituire un indicatore osservabile del tratto. È in questa direzione che si orienta, parzialmente, la RQ3 formulata in questo studio. La Gradevolezza, tratto meno studiato nel contesto della HCI rispetto agli altri, è stata associata in letteratura a una maggiore propensione alla cooperazione e alla deferenza verso gli interlocutori. Sebbene l'IA non sia un soggetto umano, alcuni studi recenti hanno evidenziato come individui con alta Gradevolezza tendano ad adottare un registro comunicativo più cortese e mitigato anche nelle interazioni con sistemi automatizzati, come se il tratto si esprimesse indipendentemente dalla natura dell'interlocutore [28]. Questa evidenza, seppur ancora preliminare, sostiene l'ipotesi che il tono dei prompt possa costituire un indicatore rilevante della Gradevolezza dell'utente, aprendo una dimensione linguistica dell'analisi che si intreccia con quella estetico-funzionale. L'Estroversione, infine, è il tratto per cui le ipotesi nel contesto dell'HCI risultano forse più intuitive. La tendenza a ricercare stimoli sociali e a esprimersi in modo espansivo potrebbe

tradursi, nel dialogo con l'IA, in prompt più lunghi e discorsivi, in una maggiore frequenza di interazione e, sul piano del design, in interfacce più dense di contenuto e articolate sul piano delle funzionalità [29]. Si tratta, tuttavia, di un'ipotesi che richiede una verifica empirica attenta, in quanto la relazione tra espansività comunicativa e complessità visiva non è necessariamente lineare. A fronte di questo quadro teorico, emerge con chiarezza che la relazione tra personalità e output di design non si esprime attraverso un'unica variabile, ma attraverso un sistema di influenze intrecciate che agiscono sia sul piano linguistico dei prompt sia sul piano visivo e funzionale degli artefatti. È in questo spazio che si colloca l'ipotesi centrale della presente ricerca: che l'output generato dall'IA non sia il prodotto di un processo neutro, ma possa essere letto come una proiezione mediata dei tratti individuali di chi ha guidato l'interazione [30]. In questa prospettiva, il modello generativo potrebbe agire come un amplificatore delle intenzioni dell'utente, rendendo visibili predisposizioni che altrimenti rimarrebbero implicite nel processo creativo [31]. L'artefatto finale, in altri termini, potrebbe configurarsi come ciò che in questa ricerca si definirà, più avanti, un possibile proxy comportamentale del profilo psicologico del suo autore, una nozione su cui ci si soffermerà con maggiore rigore nella sezione metodologica dedicata.

Capitolo 3

Metodologia

3.1 Il Disegno della Ricerca

La presente ricerca si propone di indagare in che misura i tratti di personalità individuale, misurati attraverso il modello Big Five (OCEAN), influenzino il processo di co-design generativo e i suoi esiti osservabili, sia sul piano linguistico dei prompt sia su quello visivo-funzionale degli artefatti prodotti. Lo studio si configura come esplorativo, scelta motivata dalla relativa novità del fenomeno indagato: la letteratura sul comportamento dell'utente nell'interazione con sistemi di GenAI è ancora in fase di consolidamento [24], e il territorio in cui psicologia della personalità e co-design generativo si incontrano non ha ancora prodotto un corpus di ipotesi sufficientemente sedimentate da giustificare un disegno confermativo.

Sul piano logico, la ricerca adotta uno schema variabile indipendente–variabile dipendente. I punteggi OCEAN di ciascun partecipante rappresentano le variabili indipendenti: i tratti psicometrici stabili che si ipotizza modulino il processo di interazione con il modello generativo. Le variabili dipendenti sono di natura duplice: il comportamento linguistico nei prompt, analizzato in termini di lunghezza, tono e specificità tecnica, e le scelte di design riflesse nell'artefatto HTML finale, operazionalizzate attraverso le metriche descritte nel paragrafo 3.6. Questa struttura bipartita permette di affrontare research question che operano su piani analitici distinti ma complementari, dalla dimensione testuale dell'interazione alla sua traduzione visivo-funzionale.

Il disegno sperimentale si è articolato in tre fasi sequenziali e distinte. Nella prima fase, di natura psicometrica, ciascun partecipante ha completato in modo autonomo un questionario basato sul modello Big Five, ottenendo al termine un codice identificativo univoco e anonimizzato. Nella seconda fase, il task di co-design, i partecipanti sono stati invitati a progettare una Single Page Application (SPA) per l'acquisto di biglietti ferroviari, interagendo liberamente con il modello Claude Sonnet 4.5 attraverso la funzione Artifacts. Nella terza fase, il ricercatore ha proceduto alla raccolta sistematica dei due oggetti di analisi per ciascun partecipante: il log completo della conversazione e il codice HTML dell'artefatto finale. La sequenzialità delle fasi non è casuale, ma risponde all'esigenza metodologica di isolare la variabile psicometrica prima che l'individuo entri in contatto con il task, evitando così che l'esperienza di progettazione possa retroattivamente condizionare la risposta al questionario.

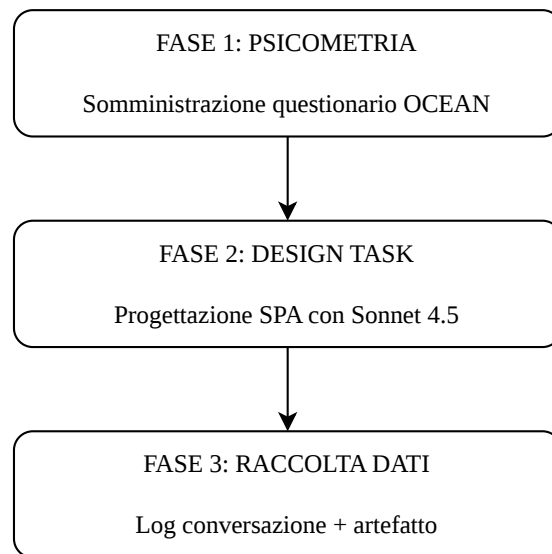


Figura 3.1: *Rappresentazione schematica del disegno della ricerca.*

Le quattro research question che strutturano l'intera indagine trovano nelle fasi descritte la propria base empirica. La RQ1 e la RQ3, relative rispettivamente alla dimensione linguistica dei prompt e alle dinamiche iterative del processo, si fondano principalmente sull'analisi dei log conversazionali prodotti nella seconda fase. La RQ2, che indaga la correlazione tra tratti psicologici e scelte di design, si basa

invece sull'analisi degli artefatti HTML raccolti nella terza fase. La RQ4, infine, che si propone di valutare l'output generativo come possibile proxy comportamentale del profilo psicologico dell'utente, rappresenta una sintesi interpretativa che attraversa tutti e tre gli strati di dati, integrando le evidenze linguistiche, visive e psicometriche in una prospettiva unitaria.

3.2 Il Campione

Lo studio ha coinvolto un campione di 21 partecipanti, reclutati secondo un criterio di campionamento per convenienza. Si tratta di un approccio non probabilistico, in cui i partecipanti vengono selezionati in base alla loro accessibilità e disponibilità piuttosto che attraverso una procedura di estrazione casuale dalla popolazione. Questa scelta, pur non consentendo una generalizzazione statistica dei risultati, è considerata metodologicamente adeguata alla natura della ricerca: l'obiettivo primario non è la produzione di stime rappresentative, ma l'identificazione di pattern e tendenze che possano costituire una base empirica per indagini future condotte su campioni più ampi e controllati [32].

Sul piano etico, la partecipazione allo studio è avvenuta su base volontaria e informata. A ciascun partecipante è stata garantita la piena anonimizzazione dei dati raccolti. A tal fine, al termine della compilazione del questionario OCEAN, il sistema ha generato automaticamente un codice identificativo univoco e non nominativo, assegnato individualmente a ciascun soggetto. Tale codice è stato successivamente inserito dall'utente in un modulo di raccolta dati in cui venivano richieste alcune informazioni demografiche di base: età, genere e occupazione. Queste variabili non costituiscono oggetto di analisi primaria, ma forniscono una descrizione più articolata della composizione del campione e potranno essere considerate, in sede di interpretazione, come possibili variabili di controllo rispetto ai risultati osservati. Il codice identificativo ha inoltre assolto una funzione di chiave di collegamento tra i tre oggetti di analisi prodotti da ogni partecipante: il profilo psicometrico, il log della conversazione con il modello generativo e l'artefatto HTML finale. In nessuna fase dell'analisi i dati sono stati trattati in forma nomi-

nativa, né è stato possibile risalire all'identità del singolo partecipante a partire dai materiali raccolti.

È opportuno riconoscere esplicitamente i limiti che il campionamento per convenienza introduce sul piano della validità esterna. La composizione del campione non è stata controllata rispetto a variabili potenzialmente confondenti, quali il livello di familiarità con strumenti di intelligenza artificiale generativa, il background formativo in ambito digitale o creativo, o la precedente esperienza con attività di web design. Tali variabili potrebbero influenzare tanto la struttura dei prompt quanto le scelte di design, introducendo una fonte di variabilità non interamente attribuibile ai tratti di personalità. Si riconosce questa limitazione come inerente al disegno sperimentale adottato, e la si terrà in considerazione nella fase di interpretazione dei risultati.

Variabile	Valore
Numerosità	$N = 21$
Età	Media = 27 anni ($DS = 7.8$; range: 20–55)
Genere	Maschile: 11 (52.4%) Femminile: 10 (47.6%)
Occupazione	Studente: 11 (52.4%); Sviluppatore SW: 2 (9.5%); Ristorazione: 2 (9.5%); Altre professioni: 6 (28.6%)

Tabella 3.1: *Caratteristiche demografiche del campione ($N = 21$).*

3.3 Somministrazione del Test OCEAN

La prima fase del protocollo sperimentale ha previsto la somministrazione di un questionario psicometrico basato sul modello Big Five, con l'obiettivo di rilevare il profilo di personalità di ciascun partecipante prima dell'avvio del task di co-design. La sequenza è intenzionale: raccogliere i dati psicometrici in una fase preliminare e indipendente rispetto all'esperienza di progettazione garantisce che i punteggi OCEAN riflettano tratti stabili dell'individuo, e non stati emotivi o disposizioni contingenti indotte dall'interazione con il modello generativo.

Lo strumento adottato è un questionario online, accessibile tramite link condiviso dal ricercatore [33], che i partecipanti hanno compilato in modo autonomo e in un momento a propria scelta, senza vincoli temporali imposti. Questa modalità di somministrazione auto-amministrata, pur rinunciando al controllo diretto delle condizioni ambientali di compilazione, è stata ritenuta preferibile in ragione della sua praticabilità logistica e della sua coerenza con il principio di partecipazione volontaria che orienta l'intero disegno. Al termine della compilazione, il sistema ha restituito automaticamente a ciascun partecipante il proprio profilo sintetico sui cinque tratti e un codice identificativo univoco, generato in modo non nominativo, che ha costituito da quel momento in poi la chiave di anonimizzazione dell'intero dataset.

L'output psicometrico grezzo è rappresentato da punteggi su cinque dimensioni principali, Nevroticismo, Estroversione, Apertura alle Esperienze, Gradevolezza e Coscienziosità, ciascuna delle quali è articolata in sei faccette specifiche, per un totale di trenta sotto-dimensioni. I punteggi grezzi, espressi su scale ordinali di ampiezza variabile, sono stati successivamente normalizzati su un intervallo comune compreso tra 0 e 1, al fine di rendere comparabili le dimensioni tra partecipanti diversi e di facilitare l'analisi correlazionale con le variabili dipendenti. I punteggi grezzi originali sono stati conservati a scopo di verifica e trasparenza metodologica.

Vale la pena soffermarsi brevemente sulla scelta di includere nell'analisi anche le faccette, e non soltanto i cinque punteggi dimensionali aggregati. Le dimensioni OCEAN, nella loro forma sintetica, catturano tendenze comportamentali ampie e

relativamente stabili, ma possono oscurare sfumature interne rilevanti. Ad esempio, due partecipanti con un punteggio identico di Nevroticismo potrebbero differire significativamente nella composizione di tale tratto: uno potrebbe presentare punteggi elevati prevalentemente nella faccetta dell'Ansia, l'altro in quella della Vulnerabilità. Includere le faccette nella struttura del dataset non implica necessariamente che tutte vengano impiegate in ogni analisi, ma preserva la granularità del profilo psicometrico, lasciando aperta la possibilità di esplorazioni più fini in sede interpretativa.

3.4 Il Task di Co-Design

La seconda fase del protocollo sperimentale ha previsto lo svolgimento di un task di progettazione assistita da intelligenza artificiale generativa. A ciascun partecipante è stato assegnato il medesimo compito: progettare un'interfaccia web per l'acquisto di biglietti ferroviari, interagendo liberamente con il modello Claude Sonnet 4.5. La scelta di un task comune e uniforme per tutti i partecipanti risponde a un'esigenza di comparabilità: affinché le differenze negli output possano essere ricondotte alla variabilità psicologica individuale, è necessario che il punto di partenza sia identico per ciascun soggetto [34].

La piattaforma adottata, Claude Sonnet 4.5, appartiene alla famiglia di modelli sviluppati da Anthropic e rappresenta, al momento della raccolta dei dati, una delle soluzioni più avanzate disponibili per la generazione di codice a partire da descrizioni in linguaggio naturale. L'elemento distintivo che ha motivato la scelta di questa piattaforma in un contesto sperimentale è la funzione Artifacts, descritta ampiamente nella sezione 2.2 dell'elaborato.

La procedura operativa si è articolata nei seguenti passaggi. Ciascun partecipante ha avuto accesso a un account Claude e ha aperto una nuova sessione di chat. In apertura della sessione, è stato fornito un prompt iniziale standardizzato, identico per tutti i partecipanti e predisposto dal ricercatore. Il testo del prompt recitava:

"Sto conducendo un esperimento di tesi. Il partecipante dovrà progettare l'inter-

faccia di acquisto di biglietti del treno. Il partecipante ha piena libertà su tutto quello che riguarda la progettazione dell'interfaccia. L'output deve essere un'interfaccia web. Non basarti su chat precedenti. Non influenzare in alcun modo le scelte dell'utente. Lascio il controllo al partecipante."

A partire da questo punto, il partecipante ha interagito liberamente con il modello fino al completamento soggettivo del task. Al termine della sessione, la chat è stata rinominata con il codice identificativo univoco OCEAN del partecipante, al fine di consentire il collegamento con il profilo psicometrico in fase di analisi.

La scelta di un protocollo deliberatamente aperto e non direttivo merita una riflessione esplicita, in quanto rappresenta una delle decisioni metodologiche più rilevanti dell'intero disegno. Non imporre requisiti funzionali stringenti, né vincoli estetici o strutturali, risponde all'obiettivo di massimizzare lo spazio espressivo a disposizione di ciascun partecipante. Un protocollo più vincolante avrebbe ridotto la variabilità degli output, limitando la possibilità di osservare come le predisposizioni individuali si manifestino nelle scelte progettuali. Al contrario, lasciare piena libertà all'utente, sia sul piano estetico sia su quello funzionale, consente di interpretare ogni scelta come un atto potenzialmente informativo del profilo psicologico di chi l'ha compiuta. Il prompt iniziale è stato inoltre formulato in modo da non suggerire alcuna direzione progettuale al modello, istruendo esplicitamente Claude a non influenzare le scelte del partecipante, così da preservare il più possibile la spontaneità dell'interazione [35].

3.5 La Raccolta dei Dati

Al termine dell'intero ciclo sperimentale, una volta concluse le sessioni di tutti i partecipanti, il ricercatore ha proceduto alla raccolta sistematica di due oggetti di analisi distinti per ciascun soggetto. La scelta di posticipare questa fase al completamento di tutte le sessioni non è casuale: raccogliere i materiali in modo aggregato e differito ha consentito di preservare l'integrità del protocollo di anonimizzazione, impedendo che il ricercatore potesse entrare in contatto con i dati di un partecipante prima che l'intera raccolta fosse conclusa. Entrambi gli oggetti

sono collegati al partecipante attraverso il codice identificativo univoco generato nella fase psicométrica. La dualità di questi oggetti non è accessoria, ma strutturalmente coerente con l'impianto della ricerca: i due livelli di analisi, linguistico e visivo-funzionale, richiedono materiali empirici di natura diversa, che solo congiuntamente permettono di rispondere all'intero spettro delle research question formulate.

Il primo oggetto di analisi è il log completo della conversazione, ovvero la trascrizione integrale dell'interazione tra il partecipante e il modello generativo nel corso dell'intera sessione. Il log è stato esportato al termine di ciascuna sessione e strutturato in un formato markdown standardizzato, organizzato per unità di scambio progressivamente numerate. Per ciascuna interazione, il documento distingue esplicitamente il contributo dell'utente dal contributo del modello, rispettivamente contrassegnati dai marcatori **UTENTE:** e **CLAUDE:**, all'interno di una sezione identificata da un'intestazione numerata del tipo **INTERAZIONE N**. Questa struttura garantisce che ogni prompt dell'utente sia isolabile e analizzabile in modo indipendente, rendendo possibile l'applicazione delle metriche linguistiche descritte nel paragrafo 3.6 senza ambiguità di attribuzione.

Il secondo oggetto di analisi è l'artefatto HTML finale, vale a dire il codice sorgente, comprensivo delle componenti HTML, CSS e JavaScript, dell'ultima versione dell'interfaccia prodotta durante la sessione. Questo oggetto rappresenta il punto di arrivo del processo di co-design e costituisce il materiale empirico su cui si fondano le analisi relative alla dimensione estetico-funzionale della ricerca. La raccolta è avvenuta mediante estrazione diretta dell'ultimo Artifact generato nella chat, identificato come la versione definitiva al termine dell'interazione. La scelta di adottare esclusivamente l'artefatto conclusivo risponde a una precisa logica metodologica: l'interesse analitico non è tracciare la traiettoria evolutiva del codice generato, ma osservare il risultato finale come esito complessivo delle scelte intenzionali accumulate nel corso del processo. L'artefatto conclusivo rappresenta, in questo senso, la sintesi più informativa del profilo progettuale del partecipante.

La connessione tra i due oggetti di analisi e il profilo psicométrico di ciascun partecipante è garantita dal codice identificativo univoco, che funge da chiave di

collegamento tra tre dataset distinti: il file dei punteggi OCEAN normalizzati, il file markdown del log conversazionale e il file HTML dell'artefatto finale. Questa architettura di collegamento consente di condurre analisi di correlazione tra variabili psicometriche e metriche derivate dai due oggetti, mantenendo al contempo l'anonimizzazione completa dei dati.

3.6 Le Metriche di Analisi

La sfida centrale di questa fase del protocollo risiede nell'operazionalizzazione: come rendere empiricamente osservabile e quantificabile una variabile di natura latente come un tratto di personalità? I punteggi OCEAN descrivono tendenze comportamentali stabili, ma non si traducono direttamente in unità misurabili. È necessario un passaggio intermedio che identifichi, nell'output osservabile del processo di co-design, indicatori plausibilmente connessi a tali tendenze. A questo scopo è stato definito un sistema strutturato di metriche di analisi, organizzate in tre blocchi funzionalmente distinti che corrispondono ai tre livelli empirici della ricerca: il comportamento linguistico nel prompt, le scelte di design riflesse nell'artefatto finale e le dinamiche iterative del processo. La logica che guida il sistema complessivo non è quella della misurazione esaustiva, ma quella della triangolazione: ciascun indicatore, considerato singolarmente, cattura solo una proiezione parziale del profilo psicologico dell'utente; è il loro insieme, letto in relazione ai punteggi OCEAN, a rendere possibile un'interpretazione più articolata.

Vale la pena chiarire, in apertura, il perimetro analitico entro cui si collocano le metriche relative all'artefatto HTML. Il codice sorgente dell'interfaccia è generato dal modello e non dal partecipante, e non costituisce pertanto, in quanto tale, un oggetto di valutazione tecnica. Ciò che si intende misurare non è la qualità o la sofisticazione del codice, ma le scelte progettuali che il partecipante ha guidato attraverso i propri prompt e che il codice rende osservabili e quantificabili. In questa prospettiva, l'artefatto HTML funge da materializzazione delle intenzioni dell'utente: ogni suo elemento formale è interpretabile come il riflesso di una preferenza estetica o di un'aspettativa funzionale che il partecipante ha,

esplicitamente o implicitamente, comunicato al modello nel corso della sessione.

3.6.1 Analisi del Prompt (P)

Il primo blocco raccoglie gli indicatori che operazionalizzano il comportamento linguistico del partecipante durante l'interazione. Tutte le metriche di questo blocco sono calcolate esclusivamente sui turni di testo prodotti dall'utente, escludendo sistematicamente le risposte del modello. Le metriche del Blocco P costituiscono il materiale empirico primario per la RQ1, che indaga la dimensione linguistica dell'interazione.

Il primo indicatore, denominato **Parole per Prompt (P1)**, misura la lunghezza media dei prompt prodotti dal partecipante nell'arco dell'intera sessione, calcolata come rapporto tra il numero totale di parole digitate e il numero totale di turni inviati. La scelta di questo indicatore è motivata dalla sua capacità di catturare, in forma proxy, la propensione del partecipante a sviluppare le proprie istruzioni in modo espansivo o, al contrario, a formularle in modo conciso e diretto. Dal punto di vista teorico, si ipotizza che questo indicatore possa correlare positivamente con l'Estroversione, tratto associato in letteratura a uno stile comunicativo più discorsivo, e con la Coscienziosità, che potrebbe indurre prompt più articolati nella specificazione dei requisiti [36]. L'estrazione è avvenuta mediante un parser che opera sui file markdown dei log conversazionali, conteggiando le parole presenti in ciascuna sezione marcata come turno dell'utente e calcolando la media sull'intero dataset di sessione.

Il secondo indicatore, **Technical Terms (P2)**, misura la densità di terminologia specialistica appartenente al dominio del web design e della programmazione nei prompt dell'utente. Operativamente, si è costruito a priori un dizionario di termini tecnici predefiniti, comprendente voci come layout, padding, navbar, grid, font-size, viewport, responsive, CSS, HTML, flex e termini affini, ed è stata calcolata la loro frequenza relativa sul totale delle parole digitate, espressa come occorrenze per cento parole. La scelta di ancorare il calcolo a un vocabolario definito prima dell'analisi risponde a un'esigenza di riproducibilità: un approccio lessicometrico senza dizionario fisso rischierebbe di produrre misure non comparabili tra

sessioni. Sul piano teorico, si ipotizza una correlazione positiva con la Coscienza, tratto associato alla tendenza e alla precisione tecnica nella formulazione delle istruzioni [26].

Il terzo indicatore del blocco, **Utilizzo di Immagini (P3)**, misura la frequenza con cui il partecipante ha richiesto esplicitamente elementi visivi nei propri prompt. Si è scelto di posizionare questo indicatore al livello del testo dei prompt anziché al livello dell'artefatto per una ragione metodologica precisa: rilevare la presenza di elementi immagine nel codice HTML finale non consentirebbe di distinguere tra immagini inserite su richiesta esplicita del partecipante e immagini introdotte autonomamente dal modello come scelta compositiva. Spostando l'analisi al livello del prompt, l'indicatore diventa direttamente attribuibile alle intenzioni del partecipante. Operativamente, si è costruito un vocabolario di riferimenti visivi espliciti, comprendente termini come immagine, foto, icona, sfondo, galleria, banner e i corrispondenti equivalenti inglesi, e se ne è calcolata la frequenza relativa per cento parole. Sul piano teorico, si ipotizza una correlazione con l'Apertura alle esperienze, in ragione della sensibilità estetica e della preferenza per interfacce visivamente ricche che caratterizzano questo tratto, e con l'Estroversione, associata alla ricerca di stimolazione percettiva [37].

KPI	Descrizione
P1 - Parole per Prompt	Lunghezza media dei prompt dell'utente.
P2 - Termini Tecnici	Frequenza relativa di termini appartenenti a un vocabolario tecnico predefinito.
P3 - Utilizzo di Immagini	Frequenza relativa di richieste esplicite di elementi visivi nei prompt del partecipante.

Tabella 3.2: *Riepilogo degli indicatori del Blocco P - Analisi del Prompt.*

3.6.2 Analisi del Design (D)

Il secondo blocco raccoglie gli indicatori che operationalizzano le scelte di design dell'utente così come si riflettono nell'artefatto HTML prodotto al termine della sessione. Come anticipato, l'oggetto di analisi non è il codice generato in quanto produzione tecnica del modello, ma il complesso di preferenze progettuali che il partecipante ha orientato attraverso l'interazione e che il codice rende misurabili. Le metriche di questo blocco costituiscono il materiale empirico primario per la RQ2 e, in modo integrativo, per la RQ4. L'estrazione delle metriche è avvenuta mediante parser automatici scritti in Python, applicati al codice sorgente degli artefatti finali.

Il primo indicatore del blocco, **Complessità Layout (D1)**, misura il numero di sezioni funzionali distinte e identificabili nell'artefatto, intese come aree dell'interfaccia dedicate a un compito specifico dell'utente, quali la ricerca di tratte, la visualizzazione dei risultati, la selezione dei biglietti, il pagamento o la conferma. Il parser HTML rileva i tag semantici principali e i contenitori con attributi identificativi semanticamente significativi, che vengono quindi mappati manualmente a una tassonomia funzionale predefinita. Il valore finale corrisponde al numero intero di categorie funzionali distinte presenti nell'artefatto. Si ipotizza una correlazione positiva con la Coscienziosità, tratto associato alla tendenza a strutturare l'interfaccia in aree ben delimitate e funzionalmente coerenti, e con l'Apertura alle esperienze, che potrebbe esprimersi in una maggiore articolazione delle funzionalità offerte [37].

Il secondo indicatore, **Ricchezza Funzionale (D2)**, misura il numero di elementi interattivi distinti presenti nell'artefatto, come indicatore della profondità operativa dell'interfaccia al di là della sua articolazione strutturale in sezioni. Il parser conta le occorrenze di elementi appartenenti a categorie predefinite: campi di input, form, bottoni e componenti con listener JavaScript espliciti. Il valore finale è un punteggio composto ottenuto sommando il numero di elementi interattivi. L'inclusione di questo indicatore risponde all'esigenza di distinguere due dimensioni concettualmente separate: un'interfaccia può essere composta da poche sezioni ma presentare un ricco apparato interattivo, o viceversa. Si è scelto

di misurare entrambe le dimensioni autonomamente per preservare questa distinzione analitica. Si ipotizza una correlazione positiva con la Coscienziosità, in ragione dell'attenzione ai requisiti funzionali, e con l'Apertura alle esperienze per la possibile inclusione di funzionalità non standard [37].

Il terzo indicatore, **Divergenza Creativa (D3)**, è l'unico del sistema a essere assegnato mediante valutazione ordinale da parte del ricercatore, su una scala a quattro livelli che va da un valore 0, corrispondente a un'interfaccia del tutto convenzionale, a un valore 3, corrispondente a un concept radicalmente divergente dalla norma attesa per un sito ferroviario, con l'introduzione di elementi come gamification, narrative design o metafore visive alternative. La valutazione è guidata da una rubrica predefinita e assistita da indicatori automatici estratti dal codice: la presenza di sezioni non mappabili alla tassonomia funzionale di base, la frequenza di keyword creative nel codice sorgente e il numero di blocchi di animazione nel CSS. In caso di ambiguità, si è scelto di assegnare il punteggio per difetto, privilegiando la prudenza valutativa. Questo indicatore presenta il collegamento teorico più diretto con l'Apertura alle esperienze, tratto stabilmente associato alla preferenza per la novità e alla sensibilità estetica. La scelta di includerlo nonostante la sua natura ordinale e parzialmente soggettiva è motivata dalla considerazione che alcune dimensioni della creatività progettuale non sono interamente catturabili da metriche puramente quantitative [37].

Il quarto indicatore, **Indice di Animazione (D4)**, misura la presenza di elementi dinamici e transizioni visive nell'artefatto, come indicatore dell'intenzione del partecipante di creare un'esperienza percettivamente ricca. Il valore è calcolato come somma ponderata di due componenti estratte via parser: il numero di blocchi @keyframes dichiarati nel CSS, che indicano animazioni complete e strutturate, e il numero di proprietà transition o animation dichiarate in selettori distinti, pesate con un coefficiente riduttivo di 0.5 per riflettere la loro minore complessità rispetto alle animazioni vere e proprie. Si ipotizza una correlazione con l'Estroversione, tratto associato alla ricerca di stimoli e alla vivacità espressiva, e con l'Apertura alle esperienze, in ragione della propensione alla sperimentazione estetica.

Il quinto e il sesto indicatore, **Varietà Cromatica (D5)** e Saturazione Media

(D6), si collocano entrambi nell'ambito dell'analisi della palette cromatica adottata nell'interfaccia, ma ne misurano dimensioni distinte. La Varietà Cromatica è calcolata estraendo via espressione regolare tutti i valori cromatici esadecimali dichiarati nel CSS, convertendoli nello spazio percettivo HSL e raggruppandoli in cluster sulla base della loro prossimità. Il valore finale corrisponde al numero di cluster cromatici percettivamente distinti. La Saturazione Media è invece calcolata come media aritmetica della componente di saturazione HSL di tutti i colori estratti, ed esprime la vivacità complessiva della palette, distinguendo tra toni puri e accentuati e toni neutri o pastello. Per entrambi gli indicatori si ipotizza una correlazione positiva con l'Apertura alle esperienze e con l'Estroversione, mentre una correlazione negativa potrebbe emergere con il Nevroticismo, nella misura in cui la preferenza per palette conservative e poco sature possa riflettere una tendenza alla riduzione della complessità percettiva come strategia di controllo.

Il settimo indicatore, **Gerarchia Visiva (D7)**, misura l'ampiezza della scala dimensionale tra gli elementi tipografici dell'interfaccia, intesa come indicatore della chiarezza dell'ordine informativo imposto dal partecipante. Il valore è calcolato come coefficiente di variazione dei valori font-size dichiarati nel CSS, dopo normalizzazione di tutte le unità di misura in pixel equivalenti. Un coefficiente di variazione elevato indica una gerarchia visiva marcata, in cui titoli, sottotitoli e testo corrente sono chiaramente differenziati; un valore prossimo a zero indica uniformità tipografica. Si ipotizza una correlazione positiva con la Coscienziosità, tratto associato all'importanza attribuita alla strutturazione chiara e prioritizzata delle informazioni [26].

L'ottavo indicatore, **Varietà Tipografica (D8)**, misura la ricchezza espressiva dell'apparato tipografico dell'artefatto attraverso un punteggio composito che combina il numero di famiglie di font distinte dichiarate e il numero di combinazioni uniche di peso e stile tipografico, queste ultime pesate con un coefficiente riduttivo di 0.5. La distinzione tra numero di famiglie e numero di varianti è analitica: l'uso di una sola famiglia con molte varianti stilistiche rappresenta una scelta tipografica diversa rispetto all'uso di più famiglie con varianti ridotte, e il punteggio composito tenta di catturare entrambe le dimensioni. Si ipotizza una

correlazione positiva con l'Apertura alle esperienze e con l'Estroversione [37].

Il nono indicatore, **Equilibrio Visivo (D9)**, misura la proporzione tra le dichiarazioni CSS relative alla spaziatura interna ed esterna (padding e margin) e il totale delle dichiarazioni dimensionali presenti nel foglio di stile, come proxy della densità percettiva del layout. Valori elevati indicano un layout arioso e generoso nella gestione dello spazio bianco; valori ridotti indicano una maggiore densità di contenuto. Questo indicatore è incluso con una cautela interpretativa esplicita: un layout arioso può riflettere sia la preferenza per la chiarezza e la riduzione del carico cognitivo, associabile alla Coscienziosità, sia, per ragioni di natura diversa, la tendenza del Nevroticismo a semplificare e ridurre la complessità visiva. Si è scelto di interpretarlo sempre in combinazione con le metriche di Layout Complexity e Ricchezza Funzionale, al fine di disambiguare il suo significato contestuale [26].

KPI	Descrizione
D1 - Complessità Layout	Numero di sezioni funzionali distinte identificabili nell'artefatto.
D2 - Ricchezza Funzionale	Numero di elementi interattivi distinti presenti nell'artefatto.
D3 - Divergenza Creativa	Grado di allontanamento dall'interfaccia ferroviaria convenzionale verso soluzioni originali sul piano estetico o concettuale.
D4 - Indice di Animazione	Misura della presenza di elementi dinamici e transizioni visive dichiarati nel CSS dell'artefatto.
D5 - Varietà Cromatica	Numero di colori percettivamente distinti dichiarati nel CSS dell'artefatto.
D6 - Saturazione Media	Vivacità complessiva della palette cromatica adottata.
D7 - Gerarchia Visiva	Ampiezza della scala dimensionale tra gli elementi tipografici.
D8 - Varietà Tipografica	Ricchezza espressiva dell'apparato tipografico.
D9 - Equilibrio Visivo	Proporzione tra dichiarazioni di spaziatura e dichiarazioni dimensionali complessive nel CSS.

Tabella 3.3: *Riepilogo degli indicatori del Blocco D - Analisi del Design.*

3.6.3 Analisi dell'Iterazione (I)

Il terzo blocco raccoglie gli indicatori che operazionalizzano le dinamiche del processo iterativo, ossia il modo in cui il partecipante ha gestito il flusso di revisione e raffinamento dell'artefatto nel corso della sessione. Questi indicatori si fondano sull'analisi strutturale dei log conversazionali e costituiscono il materiale empirico primario per la RQ3, che indaga come i tratti di personalità influenzino la persistenza nel raffinamento dell'output.

Il primo indicatore, **Iterazioni Totali (I1)**, è il più diretto del blocco: corrisponde al numero totale di turni di dialogo inviati dall'utente nel corso dell'intera sessione, computati come il conteggio assoluto delle sezioni marcate come turno dell'utente nel log strutturato. Si tratta di un indicatore grezzo, non normalizzato rispetto alla complessità dell'artefatto prodotto, poiché si è ritenuto che la durata assoluta del processo iterativo sia in sé una variabile informativa, indipendentemente dal risultato finale ottenuto. Si ipotizza una correlazione positiva con il Nevroticismo, tratto associato a una minore tolleranza verso l'output inatteso e a un maggiore bisogno di controllo iterativo, e con la Coscienziosità, in ragione del perfezionismo procedurale. Una correlazione negativa potrebbe emergere con l'Estroversione, nella misura in cui un approccio comunicativo più espansivo nei prompt iniziali possa ridurre la necessità di iterazioni correttive successive [38].

Il secondo indicatore, **Architettura del Prompt (I2)**, è una variabile categoriale binaria che descrive la modalità con cui il partecipante ha strutturato il processo costruttivo dell'artefatto. Si distingue tra un approccio top-down, in cui il primo prompt contiene una descrizione complessiva e strutturata dell'interfaccia attesa, seguita da raffinamenti progressivi, e un approccio bottom-up, in cui i prompt iniziali specificano un sottoinsieme limitato di funzionalità, aggiungendo componenti in modo incrementale nel corso della sessione. La classificazione è stata assegnata manualmente dal ricercatore sulla base dell'analisi dei primi tre prompt di ciascuna sessione. In presenza di un pattern misto o non chiaramente classificabile, si è attribuita la categoria prevalente; in caso di ambiguità persistente, si è applicato un criterio conservativo. Si ipotizza un'associazione tra approccio top-down e Coscienziosità, tratto associato alla pianificazione anticipata e alla vi-

sione strutturata del progetto, e tra approccio bottom-up e Nevroticismo o bassa Coscienziosità, come espressione di una gestione più contingente e frammentata del processo [26].

Il terzo indicatore, **Attenzione ai Dettagli (I3)**, misura la quota di prompt della sessione dedicati alla correzione o al raffinamento di elementi microscopici dell'interfaccia, quali dimensioni di un singolo elemento, spaziature, colori di specifici componenti o variazioni tipografiche puntuali, in rapporto al totale dei prompt inviati. La classificazione dei singoli prompt come orientati al dettaglio o orientati a modifiche strutturali e funzionali è avvenuta manualmente, sulla base di una rubrica che distingue tra interventi locali, circoscritti a un singolo elemento visivo, e interventi globali, che modificano la struttura o l'architettura dell'interfaccia nel suo insieme. Il valore finale, espresso come rapporto nell'intervallo $[0, 1]$, indica la proporzione di prompt classificati come micro-revisioni sul totale della sessione. Si ipotizza una correlazione positiva con la Coscienziosità, in ragione del perfezionismo e dell'attenzione alla cura formale, e con il Nevroticismo, che potrebbe manifestarsi in un'insoddisfazione ricorrente verso l'output e in un ricorso frequente a piccole correzioni come forma di controllo sull'ambiente generativo [38, 26].

KPI	Descrizione
I1 - Iterazioni Totali	Numero assoluto di turni di dialogo inviati dal partecipante nel corso dell'intera sessione.
I2 - Architettura del Prompt	Modalità di costruzione dell'artefatto: approccio <i>top-down</i> oppure <i>bottom-up</i> .
I3 - Attenzione ai Dettagli	Quota di prompt dedicati a modifiche locali e microscopiche dell'interfaccia, in rapporto al totale dei turni della sessione.

Tabella 3.4: *Riepilogo degli indicatori del Blocco I - Analisi dell'Iterazione.*

Capitolo 4

Analisi dei Dati

4.1 Introduzione al Capitolo Analitico

Il presente capitolo costituisce il nucleo empirico della ricerca e si propone di tradurre le ipotesi teoriche formulate nei capitoli precedenti in un'analisi sistematica dei dati raccolti nel corso del disegno sperimentale. Prima di entrare nel merito delle analisi, è opportuno orientare il lettore sulla logica complessiva che struttura il capitolo, sull'architettura dei dati su cui si fondano le analisi e sui limiti interpretativi entro cui i risultati devono essere letti.

Le quattro research question alla base di questa ricerca operano su piani analitici distinti e trovano riscontro in fasi dell'analisi differenti. La RQ1, relativa alla dimensione linguistica dei prompt, e la RQ3, che indaga le dinamiche iterative del processo, si fondano sull'analisi dei log conversazionali e dei KPI estratti rispettivamente dai Blocchi P e I. Entrambe saranno affrontate a partire dall'Analisi Esplorativa dei Dati (EDA), che ne descrive la distribuzione, e approfondite attraverso l'analisi PLS Regression (PLSR), che ne esplora le relazioni con i punteggi OCEAN. La RQ2, che indaga la correlazione tra tratti psicologici e scelte di design, si basa invece sui KPI del Blocco D, estratti dall'artefatto HTML finale. La RQ4, infine, rappresenta una domanda di sintesi interpretativa che non si presta a essere trattata come oggetto di un'analisi separata: attraversa l'intero capitolo trovando risposta nella lettura convergente dei risultati dei tre blocchi, e la sua trattazione esplicita è riservata al Capitolo 5, sezione 5.5.

Il materiale empirico su cui si fonda l'analisi è organizzato in tre dataset distinti, ciascuno corrispondente a uno dei livelli di osservazione previsti dal disegno sperimentale. Il primo dataset, relativo al Blocco P, raccoglie i KPI estratti dai log conversazionali di ciascun partecipante: la lunghezza media dei prompt (P1), la densità di terminologia tecnica (P2) e la frequenza di richieste esplicite di elementi visivi (P3). Il secondo dataset, relativo al Blocco D, raccoglie i KPI estratti dall'artefatto HTML finale: la complessità del layout (D1), la ricchezza funzionale (D2), la divergenza creativa (D3), l'indice di animazione (D4), la varietà cromatica (D5), la saturazione media (D6), la gerarchia visiva (D7), la varietà tipografica (D8) e l'equilibrio visivo (D9). Il terzo dataset, relativo al Blocco I, raccoglie gli indicatori delle dinamiche iterative: il numero totale di prompt inviati (I1), l'architettura del primo prompt (I2, variabile categoriale binaria top-down o bottom-up) e la quota di prompt orientati al raffinamento di dettaglio (I3). I tre dataset sono collegati ai profili psicometrici dei partecipanti attraverso il codice identificativo univoco generato nella fase psicometrica, che funge da chiave di congiunzione tra le variabili indipendenti, ovvero i punteggi OCEAN normalizzati, e le variabili dipendenti operazionalizzate nei tre blocchi.

Il capitolo si articola in due blocchi analitici sequenziali e funzionalmente complementari. Il primo, corrispondente alla sezione 4.2, è dedicato all'Analisi Esplorativa dei Dati: si descrivono le distribuzioni delle variabili, si identificano eventuali anomalie o effetti di pavimento e soffitto, e si valuta la variabilità effettiva del campione come condizione necessaria per le analisi successive. In un campione di dimensioni ridotte come quello adottato in questo studio, l'EDA costituisce una fase importante che determina la sostenibilità delle analisi correlazionali: una distribuzione eccessivamente omogenea su una o più dimensioni OCEAN ridurrebbe il potere discriminativo dell'analisi multivariata [39]. Il secondo blocco, comprendente le sezioni 4.3 e 4.4, è dedicato all'analisi correlazionale. Come si argomenterà nella sezione 4.3, la scelta del metodo PLS Regression in luogo delle più tradizionali correlazioni bivariate è motivata dalla dimensione ridotta del campione e dalla necessità di analizzare simultaneamente le relazioni tra tutti i tratti OCEAN e i KPI dei tre blocchi.

È necessario, in chiusura di questa introduzione, esplicitare l'avvertenza interpretativa che governa l'intero capitolo. Data la natura esplorativa dello studio e la dimensione del campione ($n=21$), le analisi presentate non hanno l'obiettivo di produrre generalizzazioni statisticamente rappresentative, né di stabilire relazioni causali tra tratti di personalità e comportamenti di co-design. L'obiettivo è più circoscritto: identificare pattern e tendenze plausibili nel campione osservato, valutarne la coerenza rispetto alle ipotesi teoriche formulate nel Capitolo 2 e fornire una base empirica preliminare per ricerche future condotte con campioni più ampi e disegni sperimentali più controllati. I risultati vanno interpretati con la cautela metodologica appropriata a un disegno di questo tipo, tenendo presente che l'assenza di significatività statistica non implica necessariamente l'assenza di relazione tra le variabili, quanto piuttosto i limiti intrinseci di un campione di questa dimensione [32].

4.2 Analisi Esplorativa dei Dati (EDA)

4.2.1 Distribuzione punteggi OCEAN nel campione

Prima di procedere all'analisi della distribuzione psicometrica è opportuno descrivere brevemente la composizione demografica del campione, le cui caratteristiche costituiscono un elemento di contesto rilevante per l'interpretazione dei profili OCEAN osservati. Il campione è composto da 21 partecipanti, con un'età media di 27 anni e una mediana di 24 anni, a indicare una distribuzione asimmetrica positiva: la maggior parte dei soggetti si colloca nella fascia compresa tra i 20 e i 24 anni, con alcuni partecipanti più adulti che spostano la media verso l'alto, su un intervallo che va da un minimo di 20 a un massimo di 55 anni. La distribuzione di genere risulta sostanzialmente bilanciata, con 11 partecipanti di genere maschile e 10 di genere femminile. Sul piano occupazionale, la categoria più rappresentata è quella degli studenti, che costituisce il 52% del campione. Le restanti occupazioni comprendono, tra le altre, due sviluppatori software, un analista di dati e uno sviluppatore di interfacce utente, vale a dire profili con una familiarità presumibilmente più elevata rispetto alla media con gli strumenti digitali e con

il web design. Questa eterogeneità occupazionale non è stata controllata come variabile sperimentale, ma viene riconosciuta come potenziale fonte di variabilità non interamente imputabile ai tratti di personalità.

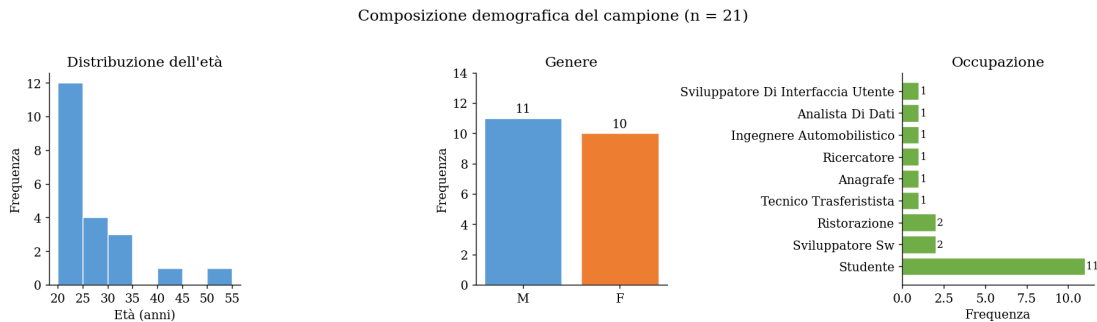


Figura 4.1: *Statistiche riguardanti il campione di ricerca.*

Completata la descrizione demografica, si passa all'analisi della distribuzione psicometrica. I punteggi OCEAN di ciascun partecipante, espressi su una scala normalizzata nell'intervallo compreso tra 0 e 1, sono stati analizzati in termini di tendenza centrale e dispersione al fine di valutare la variabilità effettiva del campione come condizione preliminare necessaria per l'analisi correlazionale. In un campione di dimensioni ridotte, una distribuzione eccessivamente omogenea su una o più dimensioni della personalità ridurrebbe il potere discriminativo dell'analisi multivariata, limitando la possibilità di osservare relazioni tra i tratti e i KPI di comportamento [39]. La tabella seguente riporta le statistiche descrittive calcolate per ciascuno dei cinque tratti.

Un'osservazione trasversale ai cinque tratti rivela una tendenza comune: tutti i punteggi si concentrano nella metà superiore dell'intervallo normalizzato, con medie comprese tra 0.60 per il Nevroticismo e 0.73 per Gradevolezza e Coscienziosità. Questo schema riflette con ogni probabilità le caratteristiche demografiche del campione, composto prevalentemente da giovani adulti in parte provenienti da ambienti accademici, per i quali è plausibile attendersi profili psicometrici orientati verso valori medio-alti nelle dimensioni adattive della personalità. Tale tendenza non inficia la validità delle analisi successive, ma costituisce un vincolo interpretativo che sarà mantenuto in considerazione nel corso del capitolo.

Sul piano della variabilità, il tratto che offre il maggiore potere discriminativo per

Tratto	Media	Mediana	DS	Min	Max	Range	CV (%)
Nevroticismo	0.60	0.60	0.11	0.42	0.78	0.36	17.7
Estroversione	0.68	0.68	0.09	0.46	0.89	0.43	12.7
Apertura alle esperienze	0.71	0.70	0.08	0.57	0.92	0.35	10.8
Gradevolezza	0.73	0.72	0.08	0.59	0.90	0.31	11.5
Coscienziosità	0.73	0.71	0.09	0.56	0.92	0.36	12.3

Tabella 4.1: *Statistiche descrittive dei punteggi OCEAN nel campione ($n = 21$), calcolate sui valori normalizzati nell'intervallo $[0, 1]$. CV = coefficiente di variazione ($DS / Media$).*

l'analisi correlazionale è il Nevroticismo, che presenta il coefficiente di variazione più elevato del campione ($CV = 17.7\%$) e la deviazione standard più ampia (0.107), distribuita su un range di 0.36 punti tra il minimo di 0.42 e il massimo di 0.78. Si tratta dell'unico tratto in cui il valore minimo si colloca con una certa distanza dal centro della scala, lasciando aperta una finestra di variabilità apprezzabile anche nel quadrante inferiore, condizione favorevole alla rilevazione di differenze individuali nelle analisi successive. L'Apertura alle esperienze mostra invece la variabilità più contenuta del gruppo, con un coefficiente di variazione del 10.8%, una deviazione standard di 0.077 e un valore minimo di 0.57 che implica l'assenza nel campione di partecipanti con bassa apertura. Questo addensamento nella metà superiore della scala configura un potenziale effetto di restrizione della varianza su questo tratto, che è al contempo, come anticipato nel Capitolo 2, uno dei predittori teoricamente più rilevanti per le variabili di design. Tale condizione non esclude la possibilità di osservare relazioni con i KPI del Blocco D, ma ne riduce il potere discriminativo e richiede cautela nell'interpretazione di eventuali

coefficienti che coinvolgono questa dimensione.

Estroversione e Coscienziosità presentano profili di variabilità analoghi, con coefficienti di variazione rispettivamente del 12.7% e del 12.3% e range di 0.43 e 0.36 punti. Entrambi i tratti mostrano una distribuzione che copre un segmento utile della scala, rendendo i due indicatori teoricamente adeguati per l'analisi correlazionale. La Gradevolezza è, tra i cinque tratti, quello con il range più ristretto (0.31 punti) e il valore minimo più elevato (0.59), aspetto che, unitamente al CV dell'11.5%, segnala una distribuzione compatta nella parte alta della scala. Questo sarà rilevante nella sezione 4.4, dove si dovrà valutare con attenzione la capacità esplicativa di questo tratto rispetto ai KPI del Blocco P.

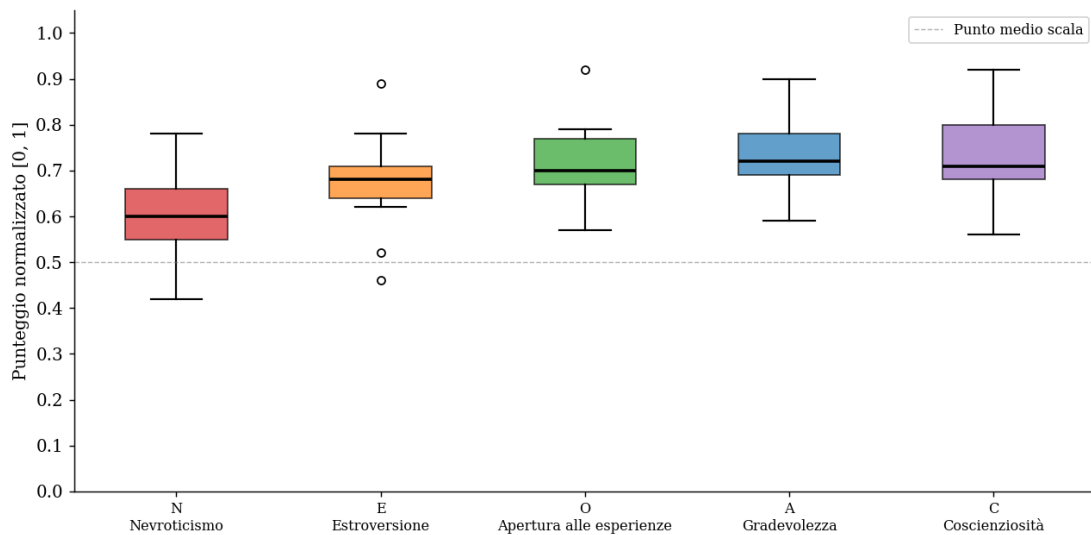


Figura 4.2: *Boxplot rappresentante la distribuzione dei cinque tratti OCEAN.*

In sintesi, la distribuzione dei punteggi OCEAN nel campione presenta una variabilità complessivamente moderata, con il Neuroticismo come tratto più disperso e l'Apertura alle esperienze come tratto più omogeneo. Tutti i punteggi si collocano nella metà superiore della scala normalizzata, delineando un campione psicometricamente non rappresentativo della popolazione generale, bensì orientato verso profili adattivi. Questa condizione, pur costituendo un limite intrinseco del campionamento per convenienza adottato, non pregiudica la possibilità di osservare pattern correlazionali all'interno del campione, a condizione che le relazioni siano interpretate come tendenze locali e non come stime della popolazione di

riferimento.

4.2.2 EDA - Blocco P (Prompt)

La presente sezione analizza la distribuzione dei tre indicatori estratti dai log conversazionali del Blocco P: la lunghezza media dei prompt (P1), la densità di terminologia tecnica (P2) e la frequenza di richieste esplicite di elementi visivi (P3). L'obiettivo è duplice: valutare la variabilità inter-individuale di ciascun indicatore come condizione necessaria per le analisi correlazionali successive, e identificare eventuali anomalie distributive, quali asimmetrie marcate, effetti di pavimento o valori estremi, che richiedano una gestione metodologica esplicita. La tabella seguente riporta le statistiche descrittive calcolate per ciascuno dei tre KPI.

Variabile	Media	Mediana	DS	Min	Max	CV (%)
P1 - Parole per Prompt	84.75	26.50	184.16	5.83	789.00	217.3
P2 - Termini Tecnici / 100 parole	4.93	3.68	3.67	0.00	12.74	74.5
P3 - Menzioni Immagini / 100 parole	1.47	0.95	2.09	0.00	7.55	144.3

Tabella 4.2: *Statistiche descrittive dei tre indicatori di interazione ($n = 21$). CV = coefficiente di variazione ($DS/Media$).*

- **P1 - Parole per Prompt:** presenta la distribuzione più eterogenea dell'intero blocco, con una media di 84.75 parole per prompt e una mediana di 26.50, a segnalare una forte asimmetria positiva. Il divario tra le due misure di tendenza centrale è interamente imputabile alla presenza di due

partecipanti che hanno condotto l'intera sessione sperimentale attraverso un unico prompt di grande estensione, raggiungendo rispettivamente 789 e 426 parole. Per questi due casi, il calcolo di una media per prompt perde di fatto il suo significato operativo: trattandosi di un singolo messaggio, il valore di P1 coincide con l'estensione complessiva della sessione. Questa scelta comportamentale, condensare l'intera richiesta in un unico testo articolato anziché distribuirla in più iterazioni, è in sé un dato interpretativo rilevante, potenzialmente associabile a tratti di personalità come la Coscienziosità elevata o una bassa propensione all'interazione dialogica, e sarà ripresa nella sezione 4.4. Al netto di questi due casi, la distribuzione di P1 mostra una variabilità comunque sostenuta, con valori che spaziano da un minimo di 5.83 parole per prompt a un massimo di 71.78 tra i restanti 19 partecipanti. Il coefficiente di variazione complessivo risulta pari al 217.3%, un valore che, pur tecnicamente indicativo di un'elevata dispersione, riflette in larga misura la distorsione prodotta dai due casi estremi piuttosto che una distribuzione genuinamente differenziata;

- **P2 - Termini Tecnici per 100 Parole:** presenta la distribuzione più regolare del blocco, con una media di 4.93 termini tecnici ogni 100 parole e una mediana di 3.68, valori sufficientemente prossimi da indicare un'asimmetria contenuta. La deviazione standard è di 3.67, su un range compreso tra 0.00, per un partecipante che non ha impiegato alcun termine tecnico riconosciuto nel vocabolario di riferimento, e 12.74. Il coefficiente di variazione del 74.5% attesta una buona variabilità inter-individuale, che si traduce in un potere discriminativo adeguato per l'analisi correlazionale: il campione copre un ventaglio di stili linguistici che va da un registro sostanzialmente colloquiale a uno spiccatamente tecnico. Si segnala che un partecipante presenta un valore di P2 pari a 12.74, che supera la soglia convenzionale di $\mu + 2\sigma$ e configura pertanto un caso outlier su questo indicatore. Si tratta di un profilo con un numero contenuto di prompt ($n = 5$) e un'elevata concentrazione di lessicospecialistico, condizione che amplifica l'indice normalizzato per 100 parole in virtù del denominatore ridotto. Questo caso sarà tenuto

in considerazione nell'interpretazione dei risultati PLSR relativi a P2;

- P3 - Menzione di Elementi Visivi per 100 Parole:** presenta la distribuzione più asimmetrica del blocco, caratterizzata da un effetto di pavimento pronunciato: 8 partecipanti su 21, pari al 38% del campione, non hanno mai fatto riferimento esplicito a elementi visivi nel corso dell'intera sessione, registrando un valore di P3 pari a zero. Per i restanti 13 partecipanti, il valore medio si attesta a 1.45 menzioni ogni 100 parole, con una mediana complessiva di 0.95. La distribuzione presenta due casi con valori elevati, rispettivamente 7.55 e 5.84, che superano la soglia di $\mu + 2\sigma$ e sono identificati come outlier: si tratta di partecipanti con un numero molto elevato di prompt ($n = 108$ e $n = 57$), nei quali le richieste di immagini sono state frequenti e distribuite lungo l'intero arco della sessione. Il coefficiente di variazione del 144.3% segnala nominalmente un'elevata dispersione, ma questo valore è in larga misura artefatto dall'asimmetria della distribuzione: la presenza di molti zeri e di pochi valori alti produce una dispersione relativa elevata che non corrisponde a una distribuzione ben differenziata lungo tutta la scala.

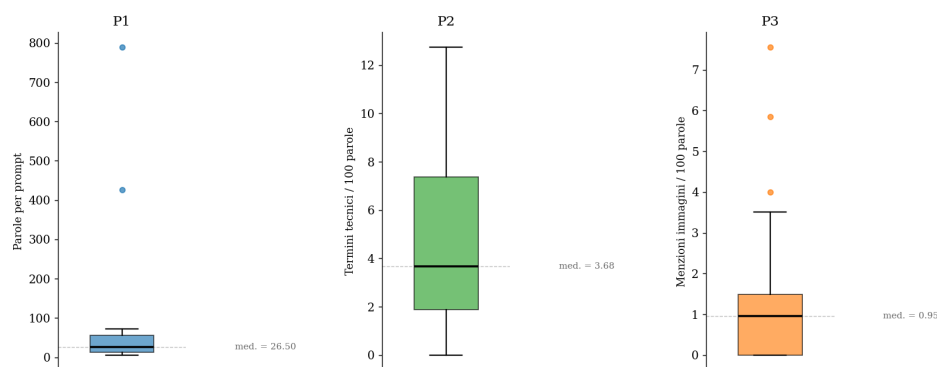


Figura 4.3: *Boxplot rappresentante la distribuzione dei KPIs del blocco P.*

In sintesi, i KPI del Blocco P mostrano una variabilità inter-individuale complessivamente buona, con P2 come indicatore distribuito in modo più regolare e dunque più direttamente adatto alle analisi correlazionali. P1 e P3 presentano distribuzioni fortemente asimmetriche, la prima per la presenza di due casi con

strategie di interazione radicalmente diverse dalla norma, la seconda per l'effetto di pavimento prodotto da un'ampia quota di valori nulli. Questa condizione non preclude le analisi successive, ma ne circoscrive l'interpretabilità e richiede l'adozione di metodi robusti alla non-normalità, quale il PLSR discusso nella sezione 4.3.

4.2.3 EDA - Blocco D (Design)

La presente sezione analizza la distribuzione dei nove indicatori estratti dall'artefatto HTML finale prodotto da ciascun partecipante. Ai fini dell'analisi, i KPI del Blocco D sono organizzati in tre famiglie funzionalmente distinte. La prima comprende le **variabili strutturali**: la complessità del layout (D1), la ricchezza funzionale (D2), la gerarchia visiva (D7) e l'equilibrio visivo (D9). La seconda raggruppa le **variabili estetiche**: l'indice di animazione (D4), la varietà cromatica (D5), la saturazione media (D6) e la varietà tipografica (D8). La terza categoria è rappresentata dalla variabile ordinale della **divergenza creativa** (D3), di natura non continua e trattata separatamente sul piano metodologico.

Variabile	Media	Mediana	DS	Min	Max	CV (%)
D1 - Complessità Layout	7.05	6.00	3.32	2.00	15.00	47.16
D2 - Ricchezza Funzionale	28.52	25.00	21.30	4.00	76.00	74.68
D4 - Indice Animazione	5.62	4.50	4.84	0.00	19.50	86.18
D5 - Varietà Cromatica	10.19	10.00	5.13	4.00	26.00	50.29
D6 - Saturazione Media	43.64	45.37	20.21	3.88	87.07	46.30
D7 - Gerarchia Visiva	0.41	0.38	0.14	0.25	0.82	32.95
D8 - Varietà Tipografica	2.69	2.50	0.97	1.50	5.00	35.98
D9 - Equilibrio Visivo	0.62	0.62	0.16	0.17	0.94	25.81

Tabella 4.3: *Statistiche descrittive degli indicatori continui di design (n = 21). CV = coefficiente di variazione (DS/Media).*

- **D1 - Complessità Layout:** presenta una media di 7.76 sezioni o elementi strutturali e una mediana di 6, con un range compreso tra 2 e 15. Il coefficiente di variazione si attesta intorno al 44%, valore che attesta una

buona variabilità inter-individuale: i partecipanti adottano strategie di organizzazione dello spazio piuttosto differenziate, da layout lineari a strutture articolate in molte sezioni. Un partecipante con $D1 = 15$ supera la soglia $\mu + 2\sigma$ e costituisce un caso outlier su questo indicatore.

- **D2 - Ricchezza Funzionale:** è l'indicatore con la dispersione assoluta più elevata dell'intero blocco: la media è pari a 28.52 funzionalità, la mediana a 25, su un range che si estende da un minimo di 4 a un massimo di 76. Il coefficiente di variazione supera il 74%, riflettendo profonde differenze nelle scelte funzionali dei partecipanti. Il partecipante con $D2 = 76$ (quasi il doppio del secondo valore più alto) costituisce un outlier manifesto e corrisponde a un profilo con un numero estremamente elevato di funzionalità implementate, riconducibile a una strategia progettuale orientata alla massima completezza dell'interfaccia.
- **D7 - Gerarchia Visiva:** presenta una media di 0.41 e una mediana di 0.38, su un range da 0.25 a 0.82. La distribuzione è relativamente compatta, con un coefficiente di variazione attorno al 30%, il più contenuto tra i KPI continui del blocco. Il partecipante con $D7 = 0.82$ supera la soglia $\mu + 2\sigma$ e si distingue per una struttura gerarchica visiva particolarmente marcata, coerente con un approccio progettuale orientato all'ordine e alla differenziazione dei livelli informativi.
- **D9 - Equilibrio Visivo:** presenta una media di 0.62 e una mediana di 0.62, su un range da 0.17 a 0.94. La distribuzione è generalmente concentrata nella fascia superiore della scala, ad indicare che la maggior parte dei partecipanti produce layout con una spaziatura adeguata. Il partecipante con $D9 = 0.17$ costituisce un outlier basso manifesto: il valore corrisponde a un layout molto denso, quasi privo di margini e spaziatura, che riflette una scelta progettuale orientata alla massima densità informativa.

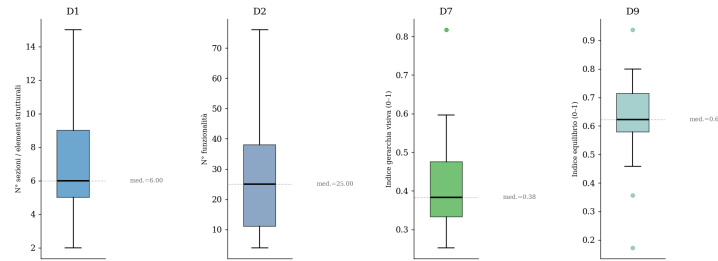


Figura 4.4: *Boxplot* rappresentante la distribuzione dei KPIs strutturali del blocco D.

- D4 - Indice di Animazione:** presenta una media di 5.62 classi CSS animate e una mediana di 4.5, con un range da 0 a 19.5. Due partecipanti non presentano alcuna animazione ($D4 = 0$), indicando scelte progettuali orientate alla semplicità. Il coefficiente di variazione elevato segnala che l'uso dell'animazione è tra le variabili di design più differenziate nel campione. Il partecipante con $D4 = 19.5$ supera la soglia $\mu + 2\sigma$, confermandosi come outlier anche su questa dimensione estetica.
- D5 - Varietà Cromatica:** registra una media di 10.19 colori unici e una mediana di 10, con un range da 4 a 26. La distribuzione è relativamente regolare rispetto ad altri KPI del blocco, con un coefficiente di variazione attorno al 50%. Un partecipante con $D5 = 26$ supera la soglia $\mu + 2\sigma$, distinguendosi per un utilizzo cromatico considerevolmente più ampio rispetto alla norma del campione.
- D6 - Saturazione Media:** presenta il range più ampio dell'intero blocco, compreso tra 3.88 e 87.07, con una media di 43.64 e un coefficiente di variazione superiore al 45%. Questa variabilità riflette scelte cromatiche polarizzate: alcuni partecipanti hanno adottato palette desaturate o quasi monocromatiche, altri palette a elevata intensità cromatica. Il partecipante con $D6 = 87.07$ costituisce un outlier manifesto su questo indicatore. Vale la pena osservare che varietà cromatica (D5) e saturazione media (D6) non mostrano una relazione lineare netta: palette più ampie non risultano sistematicamente associate a maggiore saturazione, suggerendo che le due

dimensioni catturano aspetti parzialmente indipendenti del comportamento cromatico.

- **D8 - Varietà Tipografica:** presenta una media di 2.69 font-family distinti e una mediana di 2.5, con un range da 1.5 a 5. La distribuzione è la più omogenea del blocco, con un coefficiente di variazione del 36%. La maggior parte dei partecipanti impiega tra 2 e 3.5 famiglie tipografiche, riflettendo un comportamento relativamente convergente nelle scelte di stile testuale. Un partecipante con $D8 = 5$ supera la soglia $\mu + 2\sigma$.

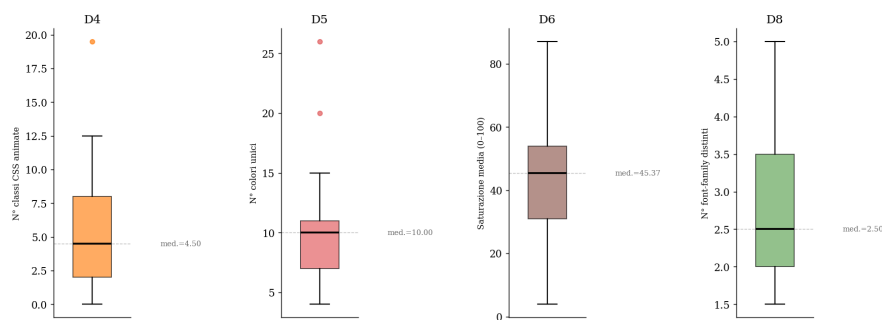


Figura 4.5: *Boxplot rappresentante la distribuzione dei KPIs estetici del blocco D.*

- **D3 - Divergenza Creativa:** è una variabile ordinale con scala da 0 (design convenzionale) a 3 (design molto creativo). La distribuzione osservata nel campione è fortemente concentrata sul valore 1, che raccoglie 10 partecipanti su 21 (48%), seguita dal valore 2 con 6 partecipanti (29%), dal valore 0 con 4 partecipanti (19%) e dal valore 3 con un solo partecipante (5%). Questo effetto di centralità attorno al livello 1 segnala che la scala è utilizzata, ma che la divergenza creativa elevata rimane un comportamento raro nel campione. La scarsa presenza ai livelli estremi della scala riduce il potere discriminativo dell'indicatore e suggerisce cautela nell'interpretazione di eventuali relazioni con i tratti OCEAN.

In sintesi, i KPI del Blocco D mostrano un grado di variabilità inter-individuale complessivamente buono per le variabili continue, con D2 e D6 come indicatori più dispersi e D8 come il più omogeneo. La variabile non continua D3 presenta una distribuzione concentrata che ne limita il potere discriminativo per effetto di

centralità. Questa condizione orienta la selezione delle variabili da includere nel modello PLSR, come si discuterà nella sezione 4.3.

4.2.4 EDA - Blocco I (Iterazione)

La presente sezione analizza la distribuzione dei tre indicatori relativi alle dinamiche iterative del processo di co-design. Il Blocco I comprende variabili di natura eterogenea: I1 (numero totale di prompt inviati, variabile continua) misura l'intensità del processo iterativo; I2 (architettura del primo prompt, variabile binaria: 0 = bottom-up, 1 = top-down) descrive la strategia di apertura dell'interazione; I3 (rapporto tra prompt di dettaglio e prompt totali, variabile continua nell'intervallo $[0, 1]$) quantifica la proporzione di iterazioni orientate al raffinamento piuttosto che all'estensione funzionale.

Variabile	Media	Mediana	DS	Min	Max	CV (%)
I1 - N° Iterazioni	16.86	9.00	26.21	1.00	108.00	155.48
I3 - Rapporto Dettaglio	0.40	0.45	0.34	0.00	1.00	82.91

Tabella 4.4: *Statistiche descrittive degli indicatori continui di interazione ($n = 21$). DS = deviazione standard; CV = coefficiente di variazione ($DS/Media$).*

- **I1 - Numero di Interazioni:** è l'indicatore con la variabilità più elevata dell'intero studio. La mediana si attesta a 9 prompt, valore che ben rappresenta il comportamento tipico del campione, mentre la media è pari a 16.86, fortemente influenzata dalla presenza di un partecipante con 108 prompt che costituisce un outlier manifesto secondo la soglia $\mu + 2\sigma$. Al netto di questo caso estremo, i valori si distribuiscono in un range da 1 a 57, con una concentrazione prevalente nella fascia bassa (1–11 prompt). Il coefficiente di variazione supera il 150%, valore che riflette in larga misura

la distorsione prodotta dall'outlier piuttosto che una distribuzione genuinamente bimodale. Ciononostante, I1 rimane l'indicatore con il maggiore potere discriminativo inter-individuale del blocco: le differenze nel numero di iterazioni tra i partecipanti sono reali e sostanziali, e rappresentano un segnale informativo per l'analisi correlazionale con i tratti OCEAN.

- **I3 - Rapporto Prompt di Dettaglio:** presenta una media di 0.40 e una mediana di 0.45, su un range da 0.0 a 1.0. La distribuzione presenta due caratteristiche degne di nota. In primo luogo, 5 partecipanti registrano un valore pari a 0.0, indicando sessioni in cui nessun prompt è stato classificato come orientato al dettaglio: si tratta di profili con un numero contenuto di iterazioni (2–5 prompt), nei quali l'intera interazione è risultata strutturata attorno all'estensione funzionale piuttosto che al raffinamento. In secondo luogo, due partecipanti presentano un valore pari a 1.0, configurando un potenziale effetto soffitto puntuale: si tratta di sessioni con un unico prompt totale ($n = 1$), nel quale l'unica iterazione è stata classificata come prompt di dettaglio, rendendo il rapporto necessariamente uguale a 1 per effetto del denominatore unitario. Questa condizione è un artefatto della misura e non un dato comportamentale interpretabile. Il coefficiente di variazione si attesta intorno all'83%, che attesta una buona variabilità nel campione, pur in presenza delle concentrazioni agli estremi della scala.

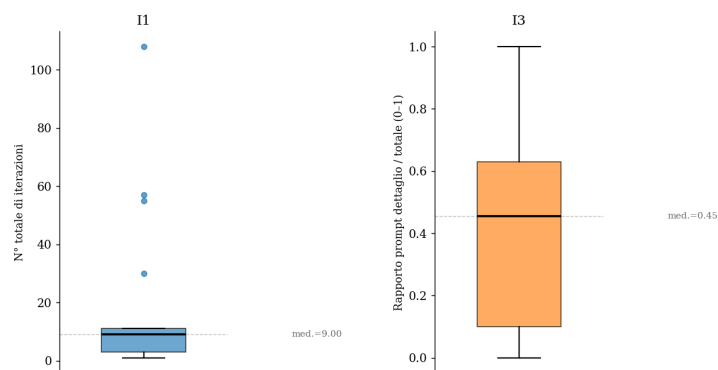


Figura 4.6: *Boxplot rappresentante la distribuzione dei KPIs I1 e I3.*

- **I2 - Architettura del Processo:** è una variabile binaria che classifica

l'approccio con cui il partecipante ha strutturato il primo prompt: bottom-up (0), ovvero partendo da elementi specifici o dalla componente funzionale, oppure top-down (1), ovvero a partire da una descrizione generale dell'interfaccia. Nel campione si osserva una lieve prevalenza dell'approccio top-down, adottato da 12 partecipanti su 21 (57%) contro 9 che hanno optato per un approccio bottom-up (43%). La distribuzione è sufficientemente bilanciata da non configurare un effetto pavimento o soffitto, ma il basso potere discriminativo intrinseco di una variabile dicotomica in un campione di questa dimensione ne sconsiglia l'inclusione diretta nel modello PLSR. Come si argomenterà nella sezione 4.4.4.

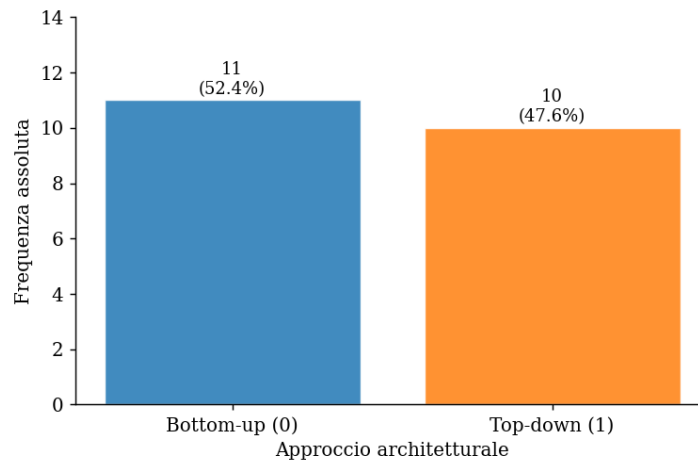


Figura 4.7: Rappresentazione della distribuzione del KPI I2.

In sintesi, i KPI del Blocco I offrono un quadro variegato: I1 è l'indicatore con il maggiore potere discriminativo, pur richiedendo attenzione alla presenza dell'outlier estremo; I3 mostra una distribuzione con caratteristiche bimodali che impone cautela nell'interpretazione; I2, di natura binaria, sarà trattata con cautela. Complessivamente, il Blocco I fornisce una descrizione quantitativa delle strategie iterative adottate dai partecipanti che, integrata con i profili psicometrici, potrà contribuire a identificare pattern di interazione associati ai tratti di personalità.

4.2.5 Sintesi EDA e implicazioni

La presente sezione offre un quadro di raccordo tra i risultati delle tre analisi esplorative, con l'obiettivo di orientare la selezione delle variabili per il modello PLSR presentato nella sezione 4.3 e di segnalare le condizioni distributive che richiedono una gestione metodologica esplicita nelle analisi correlazionali.

Sul piano della variabilità inter-individuale, i KPI che presentano le condizioni più favorevoli per rilevare pattern correlazionali con i tratti OCEAN sono, nei rispettivi blocchi: P2 (densità di terminologia tecnica, $CV = 74.5\%$) e P1 (lunghezza media dei prompt, fortemente dispersa sebbene con distribuzione asimmetrica); D2 (ricchezza funzionale, $CV > 74\%$) e D6 (saturazione media, range estremamente ampio); I1 (numero di iterazioni, $CV > 150\%$, con outlier manifesto a 108). Questi indicatori combinano una varianza sufficiente a potenzialmente differenziare i partecipanti con una copertura dell'intera gamma della scala, condizione necessaria per l'analisi PLSR.

Complessivamente, l'analisi evidenzia un campione con variabilità inter-individuale sufficiente su gran parte delle dimensioni misurate, pur con alcune aree di omogeneità o distribuzione anomala che richiedono cautela interpretativa. Il passaggio alle analisi correlazionali della sezione 4.3 avviene dunque con una mappa delle condizioni distributive del campione che permette di contestualizzare i coefficienti PLSR e di valutarne la plausibilità alla luce della struttura dei dati.

4.3 Scelta dello Strumento Correlazionale

4.3.1 Correlazioni Bivariate

Come primo passo esplorativo, sono state calcolate le correlazioni bivariate tra ciascuno dei cinque punteggi OCEAN e ciascun KPI dei tre blocchi, applicando sia il coefficiente di Pearson sia quello di Spearman per tenere conto della non-normalità di alcune distribuzioni osservata nella sezione precedente. L'analisi non ha prodotto correlazioni statisticamente significative.

Questo esito è metodologicamente atteso e non implica l'assenza di relazione tra

le variabili. Con un campione di ventuno partecipanti, la potenza statistica dei test bivariate è intrinsecamente limitata: la probabilità di rilevare effetti di entità piccola o media come statisticamente significativi rimane bassa, anche in condizioni distributive favorevoli [40]. Vi è inoltre una ragione concettuale: le relazioni tra tratti di personalità e comportamenti di co-design difficilmente si esprimono attraverso coppie di variabili isolate. È più plausibile che si manifestino come pattern multivariati, in cui più tratti agiscono congiuntamente su più KPI, e che un approccio che analizza una coppia alla volta non sia in grado di catturare. Queste considerazioni motivano il ricorso al metodo descritto nella sezione successiva.

4.3.2 Giustificazione del PLS Regression

L'assenza di correlazioni bivariate significative dopo correzione per test multipli non chiude l'indagine, ma ne dirige il percorso analitico verso un approccio più adeguato alla struttura dei dati disponibili. Le relazioni tra tratti di personalità e comportamenti di co-design difficilmente si manifestano come associazioni dirette e isolabili tra una singola dimensione psicometrica e un singolo KPI: è più plausibile che esse emergano come pattern distribuiti su più variabili simultaneamente, in cui la struttura latente condivisa tra i predittori e tra le variabili risposta porta una quota di informazione che le correlazioni bivariate, per costruzione, non sono in grado di rilevare. Questa considerazione motiva il ricorso alla PLS Regression (PLSR), un metodo di regressione multivariata particolarmente adatto alle condizioni sperimentali del presente studio [41].

Il principio fondamentale su cui si basa il PLSR consiste nella decomposizione simultanea di entrambi i blocchi di variabili (predittori e variabili risposta) in componenti latenti che massimizzano la covarianza tra i due spazi. Il modello non cerca di predire ciascun KPI indipendentemente dagli altri, bensì identifica le direzioni di massima varianza condivisa tra il profilo psicometrico dei partecipanti e l'insieme delle metriche osservate. Le componenti latenti che ne risultano non sono costruiti direttamente osservabili, ma dimensioni sintetiche interpretabili sulla base dei pesi con cui ciascuna variabile originale vi contribuisce. Questo approccio si rivela particolarmente prezioso in contesti in cui i predittori presentano tra loro

un grado non trascurabile di multicollinearità: i cinque tratti OCEAN e le loro trenta sfaccettature, pur essendo teoricamente organizzati in una struttura gerarchica, mostrano in pratica correlazioni interne che possono destabilizzare le stime ottenute con una regressione ordinaria. Il PLSR gestisce strutturalmente questo problema, poiché opera su rappresentazioni compresse dei dati originali piuttosto che sui valori grezzi delle singole variabili.

Un secondo vantaggio, di particolare rilevanza nel contesto del presente studio, riguarda il comportamento del metodo in presenza di campioni di dimensioni ridotte rispetto al numero di variabili in gioco. La letteratura metodologica ha documentato l'adeguatezza del PLS per configurazioni in cui n è inferiore alla soglia di 30-50 unità e il rapporto osservazioni-variabili è sfavorevole, condizioni che caratterizzano esattamente il disegno adottato in questa ricerca, con 21 partecipanti e un blocco predittivo di 35 variabili psicometriche. In tale configurazione, approcci alternativi come la regressione lineare multipla andrebbero incontro a stime instabili e a un rischio di overfitting difficilmente controllabile [42].

4.4 Analisi PLS Regression

4.4.1 Configurazione del Modello PLSR

Il modello PLSR è stato applicato separatamente ai tre blocchi di KPI, mantenendo invariato il blocco predittivo X in tutti e tre i run: le 35 variabili psicometriche ottenute combinando i cinque punteggi macro-dimensionali OCEAN con le trenta sfaccettature di secondo livello, entrambi normalizzati su scala $[0, 1]$. I tre blocchi risposta Y corrispondono rispettivamente ai KPI del Blocco P $\{P1, P2, P3\}$, ai KPI del Blocco D $\{D1, D2, D3, D4, D5, D6, D7, D8, D9\}$ e ai KPI continui del Blocco I $\{I1, I3\}$. La variabile binaria $I2$ è trattata separatamente, come si descriverà nella sezione 4.4.4.

L'implementazione è stata realizzata in Python mediante la classe `PLSRegression` della libreria `sklearn.cross_decomposition`, con parametro `n_components=1`. La scelta di limitare il modello a una sola componente latente (PLS1) è motivata da ragioni metodologiche esplicite. Con 21 osservazioni e 35 predittori, l'estrazio-

ne di componenti aggiuntive aumenterebbe il rischio di overfitting senza apportare un guadagno interpretativo giustificabile: ogni componente successiva alla prima cattura quote decrescenti di covarianza condivisa tra i due blocchi, e in un campione di questa dimensione il segnale di tali quote tende a diventare indistinguibile dal rumore campionario. PLS1 identifica la direzione di massima covarianza tra il profilo psicometrico dei partecipanti e l'insieme dei KPI osservati, ed è sufficiente a descrivere il pattern principale di associazione che questa ricerca intende esplorare [43]. Tutte le variabili sono state standardizzate mediante `StandardScaler` prima dell'applicazione del modello, operazione necessaria per rendere confrontabili i contributi di variabili espresse su scale diverse.

I risultati sono presentati per blocco, riportando per ciascuna analisi: (a) la correlazione tra i punteggi latenti del blocco X e del blocco Y su PLS1, come misura della covarianza catturata dalla componente; (b) i pesi delle sfaccettature OCEAN (`x_weights_`), dei quali si commentano i dieci con valore assoluto più elevato; (c) i pesi delle variabili risposta (`y_weights_`), indicativi della direzione e dell'entità con cui ciascun KPI contribuisce alla componente latente; (d) le correlazioni di Spearman bivariate tra le dieci sfaccettature con peso assoluto maggiore e i KPI del blocco, come supporto visivo e di triangolazione dell'interpretazione PLSR. Si precisa che i coefficienti qui discussi non sono stime robuste di relazioni nella popolazione generale, ma indicatori di tendenza osservabili nel campione, da interpretare con la cautela metodologica propria di un disegno esplorativo con $n = 21$.

4.4.2 Risultati - Blocco P (Prompt)

La correlazione tra i punteggi latenti del blocco X e del blocco Y su PLS1 risulta pari a $r = 0.823$, il valore più elevato tra i tre run del modello. Questo indice segnala che la componente latente estratta cattura una quota sostanziale di covarianza condivisa tra il profilo psicometrico dei partecipanti e il loro comportamento linguistico nei prompt. Con $n = 21$, un valore di questa entità non può essere interpretato come una stima robusta della relazione nella popolazione, ma indica che esiste nel campione una struttura di associazione tra personalità

e stile di interazione sufficientemente coerente da emergere con chiarezza anche dalla sola prima componente del modello.

I dieci predittori con peso assoluto più elevato su PLS1, in ordine decrescente, sono: Assenza di moderazione (+0.369), Allegria (+0.316), Auto-efficacia (-0.307), Ordinatezza (-0.265), Assertività (-0.263), Zelo (+0.261), Fiducia (+0.241), Amichevolezza (+0.220), Interesse artistico (-0.204), Moralità (+0.194). La struttura dei segni delinea due poli interpretativamente distinti. Il polo positivo è dominato da sfaccettature riferibili a impulsività motivazionale e calore relazionale: Assenza di moderazione (faccetta del Nevroticismo che cattura la tendenza a cedere agli stimoli), Allegria ed Entusiasmo (faccette dell'Estroversione), Zelo (faccetta della Coscienziosità orientata alla motivazione intrinseca), Fiducia e Amichevolezza (faccette della Gradevolezza). Il polo negativo è caratterizzato invece da sfaccettature connesse all'autoregolazione e alla competenza procedurale: Auto-efficacia (fiducia nelle proprie capacità), Ordinatezza e Assertività, che nella loro combinazione descrivono un profilo orientato al controllo strutturato dell'interazione.

Sul lato della variabile risposta, P1 (Parole per Prompt) mostra il peso assoluto più elevato con segno negativo (-0.367), seguita da P3 (Menzione di Elementi Visivi, +0.250) e da P2 (Termini Tecnici per 100 Parole, +0.085). La polarità dei pesi implica che il polo positivo della componente, associato a impulsività e calore relazionale, si manifesta in prompt mediamente più brevi e con una maggiore presenza di richieste esplicite di elementi visivi. Il polo negativo, associato ad autoregolazione e competenza, corrisponde invece a prompt più estesi ma con una minore enfasi sugli elementi grafici espliciti. P2 si rivela sostanzialmente ortogonale alla componente estratta, il che suggerisce che la densità di terminologia tecnica non faccia parte del pattern principale di variazione linguistica nel campione e non si articoli lungo l'asse individuato da PLS1. La componente latente sembra pertanto identificare una dimensione interpretabile come **stile comunicativo espressivo-visivo**: i partecipanti con profili più impulsivi, entusiasti e orientati alle relazioni tendono a formulare istruzioni concise, dirette e ricche di riferimenti visivi, mentre i partecipanti con profili più autoregolati e strutturati producono testi più articolati, ma con minore esplicitazione delle preferenze gra-

fiche. Questa lettura è coerente con le ipotesi formulate nel Capitolo 2 riguardo all'Estroversione e alla Gradevolezza come predittori di uno stile comunicativo espansivo e relazionale, sebbene il pattern emerso sia più sfumato di quanto atteso: le sfaccettature di livello superiore agiscono in combinazione trasversale ai tratti macro, e la direzione di alcuni contributi si discosta dalle ipotesi iniziali.

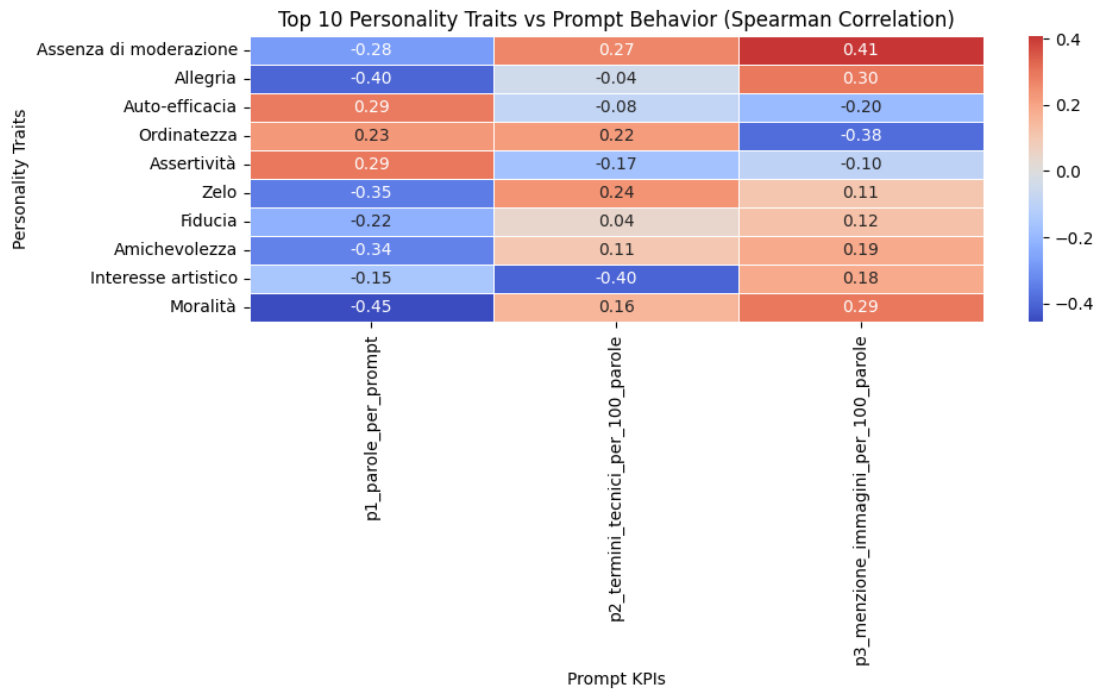


Figura 4.8: Heatmap di Correlazione di Spearman a supporto visivo.

4.4.3 Risultati - Blocco D (Design)

La correlazione latente su PLS1 tra profilo psicometrico e KPI del design risulta pari a $r = 0.738$, valore inferiore a quello del Blocco P ma comunque indicativo di una struttura di covarianza rilevabile nel campione. La maggiore dimensionalità del blocco risposta (nove KPI anziché tre) rende il compito della componente latente intrinsecamente più complesso: PLS1 deve sintetizzare in un'unica direzione le relazioni tra 35 predittori e nove misure di design eterogenee per dominio e scala di misura.

I dieci predittori con peso assoluto più elevato sono: Coscienziosità (+0.314), Livello di attività (+0.304), Auto-efficacia (+0.294), Vulnerabilità (-0.286), An-

sia (-0.285), Nevroticismo (-0.271), Disciplina (+0.209), Conseguitore di risultati (+0.204), Zelo (+0.202), Ordinatezza (+0.201). Il polo positivo è dominato da un cluster compatto di sfaccettature della Coscienziosità (Auto-efficacia, Disciplina, Conseguitore di risultati, Zelo, Ordinatezza) e dalla macro-dimensione stessa, affiancate da Livello di attività (faccetta dell'Estroversione). Il polo negativo è caratterizzato quasi esclusivamente da componenti del Nevroticismo: Vulnerabilità, Ansia e il tratto macro, insieme alla faccetta della Depressione, che non compare tra i dieci principali ma è presente con peso -0.199.

D1 (Complessità Layout, +0.277) e D9 (Equilibrio Visivo, -0.250) sono i KPI con peso assoluto più elevato, seguiti da D4 (Indice di Animazione, +0.193) e D3 (Divergenza Creativa, +0.169). I quattro KPI restanti (D5, D7, D8, D2, D6) mostrano pesi molto contenuti (da +0.052 a -0.013), indicando che la componente estratta cattura prevalentemente le dimensioni strutturali e dinamiche del design, mentre la complessità funzionale (D2), la saturazione cromatica (D6) e le scelte tipografiche (D7, D8) non partecipano in misura rilevante al pattern principale. Il segno negativo di D9 merita attenzione: il polo positivo della componente (Coscienziosità, attività) si associa a un minore equilibrio visivo, ovvero a layout meno ariosi e più densi di contenuto. Questa evidenza può sembrare controintuitiva, ci si aspetterebbe che un profilo coscienzioso prediligesse layout ordinati e spaziosi, ma diventa più interpretabile se letta congiuntamente al peso positivo di D1: i partecipanti con alta Coscienziosità tendono a strutturare interfacce più articolate, con un numero maggiore di sezioni funzionali, che occupano maggiore spazio visivo riducendo il rapporto di spaziatura catturato da D9. La componente latente sembra identificare una dimensione interpretabile come **complessità deliberata**: i partecipanti con profili psicometrici orientati alla Coscienziosità e all'attività producono interfacce strutturalmente più ricche, più animate e più creative, mentre i partecipanti con profili nevrotici, in particolare ansiosi e vulnerabili, tendono verso output visivamente più semplici, meno articolati e meno dinamici.



Figura 4.9: Heatmap di Correlazione di Spearman a supporto visivo.

4.4.4 Risultati - Blocco I (Iterazione)

La correlazione latente su PLS1 tra profilo psicometrico e KPI iterativi risulta pari a $r = 0.690$, il valore più contenuto tra i tre run. Il blocco risposta comprende due sole variabili continue, I1 e I3, esclusa I2 che, in ragione della sua natura binaria, è analizzata separatamente attraverso un test di Mann-Whitney descritto nell'ultima parte della sezione.

I dieci predittori con peso assoluto più elevato sono: Allegria (+0.376), Depressione (-0.343), Altruismo (+0.296), Ordinatezza (-0.302), Moralità (+0.275), Spirito d'avventura (-0.259), Gradevolezza (+0.230), Emotività (+0.208), Disciplina (+0.183), Fiducia (+0.192). Il polo positivo aggrega sfaccettature di natura relazionale e prosociale: Allegria (faccetta dell'Estroversione), Altruismo, Moralità, Fiducia ed Emotività (faccette della Gradevolezza), insieme al tratto macro Gradevolezza, componendo un profilo caldo, empatico e orientato alla collaborazione. Il polo negativo presenta una struttura meno omogenea: Depressione (faccetta

del Nevroticismo), Ordinatezza (faccetta della Coscienziosità) e Spirito d'avventura (faccetta dell'Apertura alle esperienze) convergono nel descrivere un profilo caratterizzato da ridotta vitalità motivazionale e tendenza alla sistematicità procedurale.

I1 (Iterazioni Totali, +0.345) è nettamente dominante rispetto a I3 (Rapporto Prompt di Dettaglio, +0.105): la componente cattura prevalentemente la variazione nell'intensità iterativa, ovvero nel numero di scambi prodotti durante la sessione, e solo marginalmente la qualità dell'iterazione espressa dal rapporto tra prompt di dettaglio e prompt totali. Questo risultato suggerisce che la dinamica principale differenziabile dalla personalità in questo blocco riguarda la quantità di interazione con il modello, non la strategia di raffinamento adottata.

La componente latente sembra identificare una dimensione interpretabile come **coinvolgimento dialogico con il modello**: i partecipanti con profili prosociali ed entusiasti tendono a produrre sessioni più estese, con un numero maggiore di scambi, come se trattassero il modello generativo come un interlocutore con cui mantenere un dialogo prolungato. I partecipanti con profili orientati alla sistematicità o caratterizzati da bassa vitalità motivazionale tendono invece a sessioni più brevi.

Un risultato di particolare interesse riguarda il Nevroticismo macro e la sua faccetta principale Ansia, entrambi con peso quasi nullo su PLS1 (rispettivamente -0.044 e +0.019). Questo dato si discosta dall'ipotesi formulata nel Capitolo 2, secondo cui livelli elevati di Nevroticismo avrebbero dovuto associarsi a un maggior numero di revisioni per effetto di insicurezza e bassa tolleranza all'output inatteso. L'unica faccetta del Nevroticismo che mostra un segnale rilevante è la Depressione (-0.343), con segno opposto a quanto atteso: i partecipanti con punteggi più elevati su questa faccetta tendono a iterare di meno, plausibilmente per un effetto di riduzione della motivazione e della persistenza nel compito piuttosto che per un meccanismo di controllo ansioso. Analogamente, la Coscienziosità macro mostra un peso quasi nullo (+0.004), mentre le sue sfaccettature si distribuiscono in direzioni opposte: Ordinatezza ha peso negativo (-0.302), associandosi a minori iterazioni, mentre Zelo ha peso positivo (+0.256), associandosi a maggiori itera-

zioni. Questo pattern suggerisce che le sfaccettature di secondo livello agiscano in modo divergente rispetto al tratto aggregato, catturando aspetti comportamentali distinti che una lettura basata sui soli cinque tratti macro non sarebbe in grado di discriminare.

La variabile I2, che classifica l'approccio con cui il partecipante ha strutturato il processo costruttivo dell'artefatto (top-down, n=12, versus bottom-up, n=9), è stata analizzata mediante test di Mann-Whitney applicato separatamente a ciascuno dei 35 predittori OCEAN. Si segnala che, in assenza di correzione per confronti multipli, il livello di significatività $\alpha=0.05$ su 35 test classifica la variabile I2 come non rappresentativa di alcuna correlazione. Pertanto, non verrà considerata nel capitolo conclusivo,

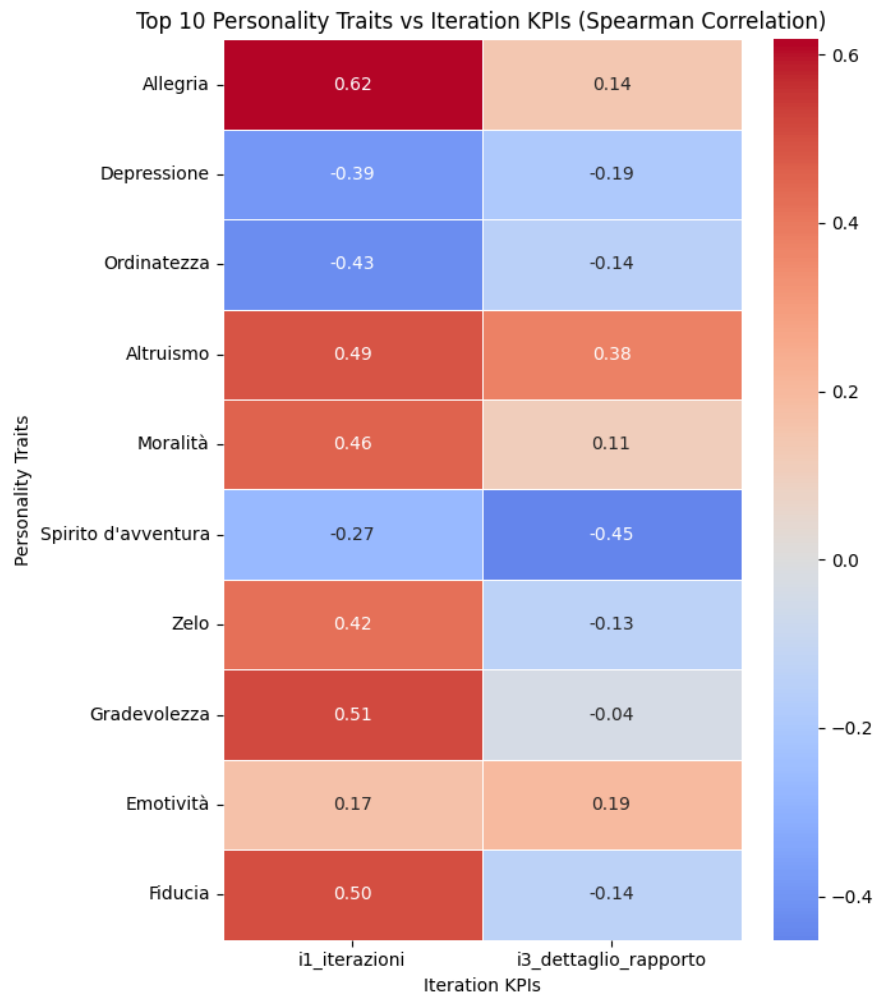


Figura 4.10: Heatmap di Correlazione di Spearman a supporto visivo.

Capitolo 5

Discussione e Conclusioni

Interpretazione in Chiave OCEAN e Research Questions. Prima di procedere alla lettura interpretativa dei risultati emersi dall'analisi PLSR, è necessario ribadire con chiarezza la cornice analitica entro cui tale lettura si colloca. I pattern individuati nelle sezioni precedenti rappresentano tendenze osservabili nel campione, non relazioni causali accertate né stime robuste della popolazione di riferimento. L'interpretazione in chiave OCEAN è guidata dalle ipotesi teoriche formulate nel Capitolo 2 e deve essere letta come un esercizio di plausibilità teorica: si valuta se i coefficienti emersi dall'analisi siano coerenti con le previsioni della letteratura, riconoscendo che la dimensione del campione non consente di trarre conclusioni definitive. Con questa avvertenza metodologica esplicitata, è possibile procedere a una lettura convergente dei tre run del modello PLSR alla luce delle quattro research question che strutturano l'intera indagine.

RQ1 - Dimensione Linguistica: *In che modo i tratti del modello OCEAN influenzano la struttura e la semantica dei prompt?*

I risultati del Blocco P offrono il segnale più forte tra i tre run del modello, con una correlazione latente di $r = 0.823$ che indica una struttura di covarianza marcata tra il profilo psicometrico dei partecipanti e il loro comportamento linguistico. Tuttavia, la lettura di questo segnale richiede una revisione parziale delle ipotesi iniziali, che si rivela essa stessa un contributo interpretativo di rilievo.

L'ipotesi formulata nel Capitolo 2 prevedeva che l'Estroversione e la Coscienziosità

fossero i predittori primari della lunghezza e dell'articolazione dei prompt, il primo per la propensione a uno stile comunicativo espansivo, il secondo per la tendenza a specificare i requisiti in modo dettagliato. I risultati dell'analisi PLSR suggeriscono invece che la dimensione latente principale non si articola lungo questi due tratti macro, bensì lungo un asse che aggregano sfaccettature di secondo livello trasversali a più tratti: il polo positivo è dominato da Assenza di moderazione, Allegria, Zelo, Fiducia e Amichevolezza, mentre il polo negativo è caratterizzato da Auto-efficacia, Ordinatezza e Assertività. Si tratta di una configurazione che descrive, più che un contrasto tra Estroversione e Coscienziosità, un'opposizione tra impulsività motivazionale e calore relazionale da un lato, e autoregolazione strutturata dall'altro.

La direzione dei pesi sulle variabili risposta complica ulteriormente la lettura. P1, ovvero la lunghezza media dei prompt, mostra il peso assoluto più elevato con segno negativo: questo implica che il polo positivo della componente, associato a impulsività e calore, si manifesta in prompt mediamente più brevi, non più lunghi come l'ipotesi sull'Estroversione avrebbe suggerito. La spiegazione più plausibile è che i partecipanti con profili impulsivi e relazionali tendano a interagire con il modello in modo diretto e iterativo, formulando istruzioni concise e affidandosi alle iterazioni successive per costruire progressivamente l'artefatto. Al contrario, i partecipanti con profili autoregolati, caratterizzati da auto-efficacia e ordinatezza, producono prompt più estesi che includono già al loro interno una specificazione più completa delle aspettative.

Il peso positivo di P3 (menzione di elementi visivi, +0.250) al polo impulsivo-relazionale trova invece una coerenza più diretta con le ipotesi iniziali: si osserva che i partecipanti caratterizzati da entusiasmo e calore tendono a esplicitare le proprie preferenze estetiche in modo più spontaneo e frequente, chiedendo al modello immagini, icone o sfondi. Questo comportamento è interpretabile come un'espressione di sensibilità visiva comunicata in forma diretta, coerente con la letteratura che associa la vivacità emotiva a una maggiore attenzione per la dimensione percettiva degli stimoli.

L'aspetto più interessante, e metodologicamente significativo, è la sostanziale ir-

rilevanza di P2 (densità di terminologia tecnica) rispetto alla componente latente estratta: il peso pressoché nullo di questo indicatore segnala che la densità tecnica del linguaggio non fa parte del pattern principale di variazione linguistica nel campione. In altri termini, la specificità del vocabolario impiegato non sembra essere una dimensione psicologicamente saliente nel differenziare i partecipanti: essa può dipendere in misura più rilevante dal background formativo e dalla familiarità con il web design che dai tratti di personalità, e la sua scarsa associazione con la struttura latente individuata dal PLSR suggerisce che, in ricerche future con campioni più ampi, questo indicatore meriterebbe di essere affiancato da misure più raffinate del background tecnico individuale, da impiegare come variabile di controllo.

RQ2 - Dimensione Estetico-Funzionale: *Esiste una correlazione tra i tratti psicologici e le scelte di design riflesse nel codice generato?*

Il run del Blocco D restituisce una correlazione latente di $r = 0.738$, inferiore a quella del Blocco P ma comunque indicativa di una struttura di covarianza osservabile, pur tenendo conto della maggiore complessità del blocco risposta, che comprende nove KPI eterogenei per scala e dominio. La struttura dei predittori su PLS1 rivela un contrasto netto tra due configurazioni psicometriche. Il polo positivo è dominato dalla Coscienziosità, sia come macro-dimensione sia attraverso le sue sfaccettature di secondo livello (Auto-efficacia, Disciplina, Conseguitore di risultati, Zelo, Ordinatezza). Il polo negativo è invece caratterizzato quasi esclusivamente da componenti del Nevroticismo: Vulnerabilità, Ansia e il tratto macro. Questo contrasto delinea un'opposizione interpretabile come intenzionalità deliberata contro inibizione ansiosa nel processo di design. I partecipanti con profili coscienziosi e attivi sembrano produrre interfacce strutturalmente più ricche: D1 (complessità del layout, +0.277) è il KPI con il peso positivo più elevato, indicando che questi profili tendono a organizzare l'interfaccia in un maggior numero di sezioni funzionali. L'associazione positiva con D4 (indice di animazione, +0.193) e D3 (divergenza creativa, +0.169) suggerisce che la stessa configurazione psicometrica si accompagna a una maggiore elaborazione dinamica e a soluzioni progettuali me-

no convenzionali, anche se quest'ultimo risultato va letto con la cautela imposta dalla scala ordinale e dalla distribuzione concentrata di D3 descritta nella sezione 4.2.3.

Un elemento apparentemente controintuitivo merita una riflessione esplicita: il segno negativo di D9 (equilibrio visivo) al polo positivo della componente implica che i partecipanti con alta Coscienziosità tendono a produrre layout meno ariosi, ovvero più densi di contenuto. Ci si sarebbe potuti aspettare il contrario, poiché l'ordine e la sistematicità associati a questo tratto fanno presupporre una preferenza per interfacce chiare e ben spaziose. La chiave interpretativa risiede nella lettura congiunta di D9 e D1: un'interfaccia articolata in molte sezioni funzionali occupa necessariamente maggiore spazio visivo, riducendo il rapporto tra spazio vuoto e spazio complessivo. In questo senso, la densità del layout non segnala disordine, ma l'ambizione a coprire un maggior numero di funzionalità e aree tematiche, che è essa stessa un'espressione del perfezionismo procedurale associato alla Coscienziosità.

Il polo negativo, dominato da Nevroticismo, Vulnerabilità e Ansia, si associa a interfacce meno articolate e meno dinamiche. Questo risultato è coerente con le ipotesi formulate nel Capitolo 2, secondo cui l'instabilità emotiva potrebbe manifestarsi in una tendenza a semplificare e ridurre la complessità visiva come strategia di controllo dell'incertezza. Si osserva tuttavia che i KPI relativi alle scelte cromatiche (D5, D6) e tipografiche (D7, D8) mostrano pesi molto contenuti su PLS1, suggerendo che la palette e la tipografia non partecipino in misura rilevante al pattern principale di variazione nel campione. Questo dato apre uno spazio per ricerche future che indagino le dimensioni estetiche con strumenti di analisi più raffinati, eventualmente capaci di cogliere pattern cromatici non rilevabili da misure aggregate come la saturazione media o il numero di cluster.

RQ3 - Dinamica Iterativa: *Come i tratti di personalità influenzano il flusso di lavoro e la persistenza nel raffinamento dell'output?*

Il Blocco I restituisce la correlazione latente più contenuta tra i tre run, $r = 0.690$, che rimane comunque indicativa di una struttura di covarianza rilevabile

nel campione. La configurazione dei predittori su PLS1 è quella che più si discosta dalle ipotesi iniziali, offrendo spunti interpretativi di particolare interesse.

Il polo positivo è dominato da sfaccettature di natura relazionale e prosociale (Allegria, Altruismo, Moralità, Fiducia, Emotività) convergenti intorno alla macro-dimensione della Gradevolezza. Il polo negativo aggrega invece Depressione (faccetta del Nevroticismo), Ordinatezza (faccetta della Coscienziosità) e Spirito d'avventura (faccetta dell'Apertura alle esperienze), componendo un profilo che si potrebbe descrivere come orientato alla sistematicità e caratterizzato da ridotta vitalità motivazionale.

La componente identifica così una dimensione che si è definita nella sezione 4.4.4 come coinvolgimento dialogico: i partecipanti con profili prosociali ed entusiasti tendono a produrre sessioni più estese, con un numero maggiore di scambi con il modello, come se trattassero il modello generativo come un interlocutore con cui mantenere un dialogo prolungato e iterativo. Questa lettura apre una prospettiva: nella relazione con uno strumento di natura non umana, il tratto della Gradevolezza, concettualmente associato alla cooperazione, sembra esprimersi in una propensione a iterare l'interazione, amplificando la sua presenza nel processo di co-design.

L'aspetto che più si discosta dalle ipotesi teoriche è la sostanziale irrilevanza del Nevroticismo macro su questa componente: il suo peso è quasi nullo (-0.044), e l'unica faccetta che mostra un segnale interpretabile è la Depressione (-0.343), con un segno che non è quello atteso. Secondo le ipotesi formulate nel Capitolo 2, livelli elevati di Nevroticismo avrebbero dovuto associarsi a un maggior numero di revisioni, per effetto di insicurezza e bassa tolleranza verso l'output inatteso. I dati suggeriscono invece che la faccetta della Depressione si associ a una minore propensione all'iterazione, plausibilmente per un effetto di riduzione della motivazione e della persistenza nel compito, piuttosto che per un meccanismo di controllo ansioso. Il risultato suggerisce che la relazione tra instabilità emotiva e dinamiche iterative possa essere mediata da dimensioni interne al Nevroticismo, come la vitalità motivazionale e il senso di efficacia, che i soli punteggi macro non sono in grado di discriminare.

Analogamente, la Coscienziosità macro mostra un peso quasi nullo su PLS1 (+0.004), mentre le sue sfaccettature si distribuiscono in direzioni opposte: Ordinatezza contribuisce negativamente (-0.302), mentre Zelo ha un contributo positivo (+0.256). Questo pattern suggerisce che la Coscienziosità non agisca come un predittore unitario delle dinamiche iterative, ma che le sue componenti interne abbiano effetti divergenti: l'ordine e la sistematicità tendono a concentrare l'interazione in meno scambi, mentre la motivazione intrinseca associata allo Zelo spinge verso sessioni più ricche di iterazioni. La capacità del PLSR di far emergere questa eterogeneità interna al tratto costituisce, in sé, un contributo metodologico rilevante: il ricorso alle trenta sfaccettature in luogo dei cinque soli punteggi macro si rivela una scelta giustificata, capace di rivelare una granularità comportamentale che una lettura di primo livello non sarebbe in grado di cogliere.

RQ4 - IA come Specchio: *Può l'output della GenAI essere considerato un proxy comportamentale del profilo psicologico dell'utente?*

La quarta research question è la più ambiziosa dell'intero studio e, come anticipato nell'introduzione al capitolo, la più vulnerabile alla dimensione ridotta del campione. La sua trattazione non può avvalersi di un run del modello dedicato, ma deve emergere come lettura convergente dei risultati dei tre blocchi, cogliendo le coerenze e le tensioni tra i pattern identificati nelle sezioni precedenti.

Ciò che i dati sembrano suggerire è che sia plausibile, entro i limiti di uno studio esplorativo con ventuno partecipanti, interpretare l'artefatto finale come portatore di una traccia del profilo psicologico di chi ha guidato la sua generazione. Questa traccia non è tuttavia né universale né uniforme: si manifesta in modo differenziato a seconda del dominio considerato. Nel Blocco P, il segnale è più forte e più coerente: la lunghezza e il contenuto implicito dei prompt dell'utente sembrano riflettere dimensioni della personalità osservabili e interpretabili in chiave psicométrica. Nel Blocco D, la traccia è leggibile soprattutto nelle scelte strutturali e dinamiche dell'interfaccia, mentre le dimensioni estetiche mostrano un segnale più debole. Nel Blocco I, la personalità sembra associarsi prevalentemente all'intensità del dialogo con il modello, ovvero alla quantità di interazioni

prodotte, più che alla qualità della strategia iterativa.

Questa eterogeneità non indebolisce la tesi centrale del lavoro, bensì la arricchisce di sfumatura. L'ipotesi che l'IA funzioni da specchio psicologico non implica necessariamente che ogni dimensione dell'artefatto finale rifletta in eguale misura tutti i tratti di personalità: è più verosimile che lo specchio sia selettivo, ovvero che alcuni tratti lascino una traccia più visibile in certi domini del processo di co-design e meno in altri. I risultati qui presentati suggeriscono che le differenze psicologiche individuali si manifestino in modo più nitido nelle dimensioni processuali, come il modo in cui l'utente costruisce le proprie istruzioni e gestisce il dialogo con il modello, che in quelle puramente estetiche, che potrebbero essere in parte livellate dalla propensione generativa di Claude a produrre soluzioni visivamente coerenti con gli standard del web design.

Va riconosciuta esplicitamente una tensione teorica nei risultati emersi: alcune ipotesi formulate nel Capitolo 2 non trovano conferma nei dati, o trovano conferma parziale attraverso sfaccettature diverse da quelle attese. L'Apertura alle esperienze, teoricamente il predittore più forte delle scelte estetiche, non emerge come voce dominante in nessuno dei tre run del modello PLSR, in parte a causa della restrizione di varianza descritta nella sezione 4.2.1. Il Nevroticismo non si associa a un maggior numero di iterazioni come ipotizzato, bensì, nella sua faccetta della Depressione, a una minore persistenza nel processo. Queste discordanze non devono essere lette come fallimenti della cornice teorica, ma come contributi all'affinamento delle ipotesi per ricerche future: indicano dove le previsioni basate sulla letteratura richiedono una specificazione più fine, e suggeriscono che la relazione tra personalità e co-design generativo sia più mediata e multidimensionale di quanto un modello diretto avrebbe previsto.

In definitiva, i risultati complessivi dell'analisi PLSR offrono una risposta parzialmente affermativa alla RQ4: è possibile rintracciare nell'artefatto finale, e nel processo che lo ha generato, segnali che riflettono la struttura psicometrica di chi ha guidato l'interazione. Questi segnali non sono univoci, non sono presenti in eguale misura in tutti i domini osservati, e non raggiungono la soglia della significatività statistica in un campione di questa dimensione. Tuttavia, la loro coerenza

interna rispetto al quadro teorico di riferimento e la loro osservabilità attraverso metodi multivariati che valorizzano la struttura latente dei dati suggeriscono che il paradigma dell'IA come specchio psicologico non sia una metafora vuota, ma un'ipotesi empiricamente esplorabile e metodologicamente sostenibile, che uno studio futuro con campione ampliato e variabili di controllo più rigorose potrebbe sottoporre a una verifica più stringente.

Implicazioni e Direzioni Future. I risultati dell'analisi PLSR non si esauriscono nella risposta alle quattro research question: per loro stessa natura esplorativa, aprono indicazioni metodologiche e spunti tematici che orientano la ricerca futura. La presente sezione le articola su due piani: quello empirico, dove si identificano le condizioni per verifiche più robuste, e quello applicativo, dove si discute il potenziale del paradigma dell'IA come specchio psicologico per sistemi di co-design adattativi.

Sul piano della ricerca empirica, la prima e più urgente direzione riguarda l'ampliamento del campione. Con ventuno partecipanti, la potenza statistica dell'analisi rimane intrinsecamente limitata, e i pattern emersi dal PLSR non possono essere trattati come stime robuste della popolazione di riferimento. Un campione di cinquanta o più partecipanti, reclutato secondo un criterio di stratificazione intenzionale rispetto ai profili OCEAN consentirebbe di superare il principale vincolo interpretativo di questo studio: la restrizione di varianza osservata sull'Apertura alle esperienze, che ha limitato il potere discriminativo proprio del tratto teoricamente più rilevante per le scelte estetiche. In un campione più ampio e stratificato, sarebbe inoltre possibile isolare sottogruppi psicometricamente omogenei e confrontarli in modo diretto [40].

Una seconda direzione di sviluppo riguarda il controllo delle variabili confondenti, che nel presente studio sono state riconosciute ma non gestite analiticamente. La familiarità con strumenti di intelligenza artificiale generativa, il background formativo in ambito digitale o creativo e la precedente esperienza con attività di web design costituiscono fonti di variabilità non interamente imputabili ai tratti di personalità. Studi futuri dovrebbero misurare queste variabili in modo sistematico, attraverso scale validate di competenza digitale percepita e questionari di

esperienza pregressa, includendole nel modello come covariate e permettendo così di isolare il contributo netto della personalità al netto dell'expertise individuale. Una terza direzione riguarda l'estensione del protocollo a contesti e strumenti diversi. Il presente studio ha adottato un unico task, ovvero la progettazione di un'interfaccia per l'acquisto di biglietti ferroviari, e un unico modello generativo, Claude Sonnet 4.5 con la funzione Artifacts. Entrambe le scelte, pur giustificate dalla necessità di controllo sperimentale, introducono un doppio vincolo di generalizzabilità: non è noto se i pattern osservati si mantengano con task di natura diversa, né se siano specifici di Claude o condivisi da altri modelli generativi come GPT-5 o Gemini. Studi comparativi che varino sistematicamente il task e il modello adottato potrebbero chiarire in quale misura la firma psicologica dell'utente dipenda dal tipo di problema progettuale e in quale misura sia invece modulata dalle caratteristiche generative del modello specifico, contribuendo così a separare il contributo dell'utente da quello dello strumento nella produzione dell'artefatto finale.

Sul piano applicativo, se la personalità dell'utente lascia una traccia osservabile nel processo di co-design, nella struttura dei prompt, nelle scelte di design e nell'intensità iterativa, allora è teoricamente plausibile progettare sistemi generativi capaci di rilevare queste tracce in tempo reale e di adattare il proprio comportamento in risposta ad esse. Un modello generativo che riconoscesse segnali di alta Coscienziosità nei prompt del suo interlocutore, attraverso indicatori come la lunghezza, la specificità tecnica e la struttura top-down delle istruzioni, potrebbe calibrare le proprie proposte verso un maggiore rigore funzionale e una presentazione più ordinata delle opzioni. Analogamente, segnali di impulsività motivazionale o di calore relazionale potrebbero orientare il sistema verso suggerimenti più visivamente vivaci e verso una gestione più dialogica del processo iterativo.

Si tratta, è opportuno precisarlo, di un'implicazione di carattere speculativo, che richiederebbe una base empirica considerevolmente più robusta di quella qui presentata prima di poter essere tradotta in scelte di design. Essa indica tuttavia una direzione di ricerca promettente che si inserisce nel dibattito più ampio sulla personalizzazione adattiva dei sistemi di intelligenza artificiale: un dibattito in

cui la comprensione del fattore umano come variabile strutturale, e non come rumore da minimizzare, potrebbe rivelarsi un contributo teorico di prima rilevanza. In questo senso, il presente studio si propone non come una risposta definitiva, ma come un passo esplorativo in una direzione che la crescente pervasività degli strumenti di GenAI nel lavoro creativo rende sempre più urgente approfondire.

Bibliografia

- [1] Jianfeng Wu, Yuting Cai, Tingyu Sun, Keer Ma, and Chunfu Lu. Integrating aigc with design: dependence, application, and evolution – a systematic literature review. *Journal of Engineering Design*, 36(5-6):758–796, 2024. Fornisce una mappatura completa della trasformazione del design landscape, analizzando l'evoluzione dai paradigmi tradizionali (CAD) verso l'integrazione sistematica della GenAI.
- [2] Andrew Peckham et al. Generative ai meets cad: enhancing engineering design to manufacturing processes with large language models. *The International Journal of Advanced Manufacturing Technology*, 138(4):1120–1135, 2025. Analizza la transizione tecnologica dai sistemi deterministici CAD verso il Generative AI-assisted Design (GAD).
- [3] Z. Liu. Human-ai co-creation: A framework for collaborative design in intelligent systems. *arXiv preprint arXiv:2507.17774*, 2025. Propone un framework teorico per il co-design collaborativo, definendo i ruoli di partner tra uomo e sistemi intelligenti.
- [4] Anthropic. Introducing claude sonnet 4.5: Toward an ai that works, not just chats. *Anthropic News*, 2025. Documentazione ufficiale sul modello Claude 4.5, con focus sulle capacità di rendering immediato (Artifacts) e iterazione proattiva.
- [5] Kiran Ramnath, Kang Zhou, Sheng Guan, Soumya Smruti Mishra, Xuan Qi, Zhengyuan Shen, Shuai Wang, Sangmin Woo, Sullam Jeoung, Yawei Wang, Haozhu Wang, Han Ding, Yuzhe Lu, Zhichao Xu, Yun Zhou, Balasubrama-

- niam Srinivasan, Qiaojing Yan, Yueyan Chen, Haibo Ding, Panpan Xu, and Lin Lee Cheong. A systematic survey of automatic prompt optimization techniques. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 33078–33110, Suzhou, China, November 2025. Association for Computational Linguistics.
- [6] Paul T. Costa and Robert R. McCrae. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources, 1992. Testo fondamentale per la definizione del modello OCEAN citato nell'inquadramento teorico.
- [7] Oliver P. John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research*, volume 2, pages 102–138. Guilford Press, 1999. Rassegna critica essenziale che documenta come il modello Big Five sia diventato la tassonomia standard per la ricerca scientifica sulla personalità.
- [8] Florencio Jr. Kabigting. The discovery and evolution of the big five of personality traits: A historical review. *GNOSI: An Interdisciplinary Journal of Human Theory and Praxis*, 4(3):119–132, 2021. Fornisce una prospettiva storica aggiornata, definendo il Big Five come il costrutto di personalità più ricercato e utilizzato al mondo ad oggi.
- [9] Robert R. McCrae and Antonio Terracciano. Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3):547–561, 2005. Supporta la validità cross-culturale del modello Big Five.
- [10] Ethan Zell and Tara L. Lesick. Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, 90(4):559–573, 2022. Sintesi quantitativa recente che conferma la robustezza e la pervasività del modello OCEAN come gold standard per prevedere il comportamento e le performance umane.

- [11] Dan Jurafsky and James H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Pearson, 3rd edition, 2024. L'opera di riferimento per l'NLP. Documenta l'evoluzione dai sistemi a stati finiti degli anni '90 ai moderni modelli di linguaggio generativi.
- [12] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994. Articolo seminale che ha definito formalmente l'ostacolo tecnico citato nel testo: l'incapacità dei modelli sequenziali di gestire informazioni distanti nella sequenza.
- [13] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30, 2017. Pietra miliare per l'architettura Transformer alla base degli LLM.
- [14] Miltiadis Allamanis, Earl T. Barr, Premkumar Devanbu, and Charles Sutton. A survey of machine learning for big code and naturalness. *ACM Computing Surveys (CSUR)*, 51(4):1–37, 2018. Testo che formalizza l'ipotesi della "naturalness": il codice sorgente ha proprietà statistiche simili al linguaggio naturale, permettendo l'uso di modelli probabilistici.
- [15] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique P.O. Pinto, Jared Kaplan, et al. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*, 2021. Documento tecnico fondamentale su OpenAI Codex. Valida l'efficacia del pre-training su repository pubblici (GitHub) per la generazione di codice funzionante.
- [16] Y. Pan et al. Webgen-bench: Evaluating llms on generating interactive and functional websites from scratch. *arXiv preprint arXiv:2505.03733*, 2025. Ricerca recentissima che valuta la capacità di modelli come Claude 3.5 Sonnet

di creare codebase multi-file (HTML/CSS/JS) coerenti e funzionali partendo da zero.

- [17] Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, et al. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073*, 2022. Documento tecnico fondamentale per la filosofia Constitutional AI di Claude/Anthropic.
- [18] J. D. Zamfirescu-Pereira et al. Why johnny can't prompt: How non-ai-experts try (and fail) to design llm prompts. *Proceedings off the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2023. Studio fondamentale del CHI che evidenzia come le differenze nei modelli mentali degli utenti producano strategie di prompting e risultati di design estremamente variegati.
- [19] L. Maman et al. Prompt variability effects on llm code generation. *ResearchGate - Preprint Series*, 2025. Studio che dimostra come il background dell'utente e il modo di decomporre un problema portino a variazioni significative nel codice generato (HTML/CSS), validando la sensibilità dell'IA alla "firma" dell'utente.
- [20] Elizabeth B.-N. Sanders and Pieter Jan Stappers. Co-creation and the new landscapes of design. *CoDesign*, 4(1):5–18, 2008. Definisce il passaggio dal CAD al design partecipativo e collaborativo.
- [21] Ben Shneiderman. Human-centered ai: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6):495–504, 2020. Analizza il passaggio dell'IA da strumento a partner, fondamentale per definire la "discontinuità radicale".
- [22] JuHyun Lee, Ning Gu, and Anthony P. Williams. Parametric design strategies for the generation of creative designs. *International Journal of Architectural Computing*, 12(3):263–282, 2014. Esplora le strategie di progettazione parametrica, distinguendo tra approcci "problem-forwarding" e "solution-reflecting" per favorire la creatività nel settore AEC.

- [23] A. Yuan, A. Coenen, E. Reif, and D. Ippolito. Wordcraft: a human-ai collaborative editor for creative writing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2022. Descrive empiricamente il ciclo di iterazione e raffinamento (action-reaction) tra utente e LLM.
- [24] J. Rezwana and M. L. Maher. Human-centered ai communication in co-creativity: An initial framework and insights. *arXiv preprint arXiv:2505.18385*, 2025. Indica esplicitamente che la ricerca sull’interazione uomo-AI nei domini co-creativi è in una fase iniziale, supportando la tesi della letteratura in consolidamento.
- [25] Alexandre N. Tuch, Javier A. Bargas-Avila, Klaus Opwis, and Frank H. Wilhelm. Visual complexity and prototypicality as determinants of the first impression of websites. *International Journal of Human-Computer Studies*, 70(11):794–811, 2012. Supporta la relazione tra tratti individuali e percezione della complessità estetica.
- [26] T. Alves, J. Natálio, J. Henriques-Calado, and S. Gama. Incorporating personality in user interface design: A review. *Personality and Individual Differences*, 155:109709, 2020. Analisi sistematica che conferma come la Coscienza influenzi la preferenza per interfacce strutturate, pulite e con gerarchie visive esplicite.
- [27] Roslyn Harold, Kaylin E. Hill, Roma Kamat, Greg Perlman, Roman Kotov, Camilo J. Ruggero, Douglas B. Samuel, and Dan Foti. Error-related brain activity shapes the association between trait neuroticism and internalizing symptomatology in two tasks. *International Journal of Psychophysiology*, 204:112404, 2024. Fornisce la base neurologica per il monitoraggio dell’errore e la sensibilità agli esiti inattesi nel tratto del Nevroticismo.
- [28] Clifford Nass and Youngme Moon. Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1):81–103, 2000. Pietra miliare

del paradigma CASA. Spiega perché applichiamo regole sociali (cortesia) alle macchine.

- [29] Jon Oberlander and Alastair J. Gill. Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes*, 42(3):239–270, 2006. Supporta l'ipotesi che l'Estroversione corredi con messaggi più lunghi e un registro più "espansivo".
- [30] T. Pascu. Personalizing llm-powered conversational agents for promoting mental well-being. In *Thesis - Politecnico di Milano*, 2025. Indaga l'orchestrazione di architetture LLM e Big-Five personality prompting per la creazione di interfacce calibrate.
- [31] Yuchang Liu, Yongzhong Yang, and Haoran Xu. From humans to ai: understanding why ai is perceived as the preferred co-creation partner. *Frontiers in Psychology*, 16(1695532):1–15, 2025. Esamina i fattori psicologici che portano a preferire l'IA come partner creativo rispetto agli esseri umani nei processi di co-creazione.
- [32] J. Lazar, J. H. Feng, and H. Hochheiser. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann, 2nd edition, 2017. Valida l'uso di campioni ridotti (n circa 20) negli studi HCI di natura esplorativa, dove l'identificazione di comportamenti e tendenze prevale sulla significatività statistica universale.
- [33] Big Five Test. Test della personalità basato sul modello dei big five, 2026. Strumento open-source impiegato per la misurazione dei cinque principali tratti della personalità (OCEAN).
- [34] W. R. Shadish, T. D. Cook, and D. T. Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin, 2002. Testo fondamentale sulla validità interna: spiega come mantenere costanti le condizioni ambientali (il task) sia essenziale per attribuire le variazioni osservate alle differenze tra i soggetti.

- [35] M. T. Orne. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11):776–783, 1962. Fonte classica per giustificare l'approccio non direttivo: spiega come le istruzioni del ricercatore possano involontariamente suggerire il comportamento atteso (demand characteristics), inquinando la spontaneità dei partecipanti.
- [36] J. W. Pennebaker and L. A. King. Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6):1296–1312, 1999. Studio seminale che stabilisce come la quantitativo di parole (word count) sia un indicatore stabile della personalità, associando l'Estroversione a una maggiore verbosità e produzione testuale.
- [37] K. Fayn, P. J. Silvia, Y. Erbas, H. Tiliouine, and P. Kuppens. Aesthetic emotions and aesthetic people: Openness predicts sensitivity to novelty in the experiences of interest and pleasure. *Frontiers in Psychology*, 6:1877, 2015. Dimostra empiricamente che l'Apertura predice la sensibilità estetica e l'interesse per stimoli visivi nuovi e complessi, supportando il legame tra tratto e preferenza per interfacce ricche.
- [38] J. L. Szalma and G. S. Taylor. Individual differences in response to automation: The five factor model of personality. *Journal of Individual Differences*, 32(2):70–83, 2011. Fonte chiave per il Nevroticismo: associa il tratto a un aumento dello stress e del carico di lavoro percepito in compiti automatizzati, supportando l'idea di un maggior bisogno di controllo e verifica (iterazione).
- [39] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson. *Multivariate Data Analysis*. Cengage Learning, 8th edition, 2018. Sottolinea come l'esame della distribuzione e della varianza sia un "prerequisito assoluto" prima di applicare tecniche multivariate, specialmente per valutare il potere discriminativo delle variabili.
- [40] F. D. Schönbrodt and M. Perugini. At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5):609–612, 2013. Dimostra che

le correlazioni bivariate sono estremamente instabili e oscillanti in campioni inferiori a $n=150$, rendendo la significatività bivariata un indicatore poco affidabile in studi con $n=21$.

- [41] H. Abdi. Partial least squares regression and projection on latent structure methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(1):97–106, 2010. Fornisce la giustificazione tecnica per l'uso delle componenti latenti nel massimizzare la covarianza e spiega come il PLSR risolva strutturalmente il problema della multicollinearità tra i predittori.
- [42] M. Haenlein and A. M. Kaplan. A beginner's guide to partial least squares analysis. *Understanding Statistics*, 3(4):283–297, 2004. Fonte cruciale per studi con campioni ridotti; documenta come il PLS non richieda le assunzioni distributive della regressione OLS e sia efficace con campioni inferiori a 50 unità e un alto numero di variabili.
- [43] S. Wold, M. Sjöström, and L. Eriksson. Pls-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2):109–130, 2001. Testo fondamentale che spiega come l'obiettivo del PLS sia estrarre solo le componenti con reale potere predittivo, avvertendo che l'aggiunta di componenti oltre la prima, in presenza di campioni piccoli, spesso modella solo il rumore (noise) invece del segnale.