

Università degli Studi di Milano Bicocca

DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE

CORSO DI LAUREA IN TEORIA E TECNOLOGIA DELLA
COMUNICAZIONE



**Ottimizzazione e Scalabilità: reingegnerizzazione di
una web application per il supporto delle decisioni
in ambito ospedaliero**

Tesi di Laurea Magistrale

Relatore Prof. Federico Cabitza

Correlatore Dott. Andrea Campagner

Laureando

Giorgia Paparazzo
Matricola 888319

Giorgia Paparazzo: *Ottimizzazione e Scalabilità: reingegnerizzazione di una web application per il supporto delle decisioni in ambito ospedaliero*, Tesi di Laurea Magistrale,
© Luglio 2024.

La strada che porta alla conoscenza è una strada che passa per dei buoni incontri.

Baruch Spinoza

Ringraziamenti

Innanzitutto, vorrei esprimere la mia gratitudine al Prof. Federico Cabitza, relatore della mia tesi, per avermi permesso di partecipare a questo progetto, avermi seguita e aver avuto fiducia in me. Vorrei ringraziare il Dott. Andrea Campagner, correlatore, per l'aiuto e il sostegno fornитomi durante la stesura del lavoro.

Desidero ringraziare con affetto la mia famiglia per il sostegno e per essermi stata vicina in ogni momento durante gli anni di studio.

Ringrazio infine i miei amici nonchè colleghi di lavoro, per avermi supportata, ascoltata, istruita, per avermi sostenuta in questo percorso e per avermi fatto crescere umanamente e professionalmente.

Indice

1	Introduzione	1
1.1	Obiettivi della tesi	1
2	Incertezza nelle decisioni mediche	3
2.1	Definizione	3
2.2	Metodologia e valutazione del rischio nella pratica clinica	4
2.3	Rapporto con il paziente	6
3	Sistemi di supporto alle decisioni	9
3.1	Machine learning in ambito medico	10
3.2	Visualizzazioni vaghe	12
4	Epimetheus	17
4.1	PROM	18
5	User Experience	20
5.1	Design Centrato sull'Utente	20
5.2	Design Centrato sull'Attività	21
5.3	Il progetto in partenza	21
5.4	Modifiche apportate	21
5.5	Test	23
5.5.1	Questionario	23
5.5.2	Think aloud	26
6	Sviluppo software	30
6.1	Design dei Sistemi	30
6.2	Il progetto in partenza	30
6.3	Analisi e motivazioni	31
6.4	Sviluppi	31
6.4.1	React	32
6.4.2	Flask	32
6.4.3	Validazione	32
6.4.4	La struttura	33
7	Metodologia	35
7.1	Metodologia agile e EventStorming	35
8	Sviluppi futuri	38

9 Conclusioni	40
A Appendice	41
A.1 Guida nell'analisi dell'incertezza secondo l'EFSA	41
A.1.1 Individuazione delle incertezze	43
A.1.2 Procedure per l'individuazione delle incertezze	43
A.1.3 Quantificare l'incertezza tramite probabilità	45
A.2 Intelligenza artificiale in ambito medico: cosa ne pensano gli utenti nel 2024	46
Glossario	47
Bibliografia	49

Elenco delle figure

2.1	Natura e fonti dell'incertezza secondo Han et al.	5
2.2	Illustrazione di come si configurano le situazioni cliniche	7
2.3	Modello della decisione condivisa secondo Kon	8
3.1	Effetti applicati all'immagine per rendere il rischio del 25%, 50% e 75%. T, B e N sono, rispettivamente, l'effetto della trasparenza, sfocatura e rumore	13
3.2	Natura dell'incertezza	15
3.3	Natura dell'incertezza	16
4.1	Questionario presente nella pagina iniziale di Epimetheus	19
5.1	Pagina di risultati pre-modifiche	22
5.2	Pagina di risultati post-modifiche	22
5.3	Esempio di parametri modificabili nel caso del physical score	22
5.4	Esempio di gradiente e grafico presente in Epimetheus testati con tutti i livelli di daltonismo presenti	23
5.5	Età degli utenti che hanno risposto al questionario	23
5.6	Principali fonti da cui gli utenti consultano grafici o infografiche	24
5.7	I due gradienti messi a confronto, con focus posto sull'area centrale che rappresenta l'incertezza. Il primo gradiente è quello inizialmente presente sul progetto, il secondo è quello che si vuole adottare	24
5.8	Distribuzione delle risposte alla domanda che mette a confronto i due gradienti	24
5.9	I circulargraph sottoposti a quesito	25
5.10	Risposte al test A/B che paragona l'interfaccia iniziale con la finale	26
5.11	Boxplot del test Beta condotto dal partecipante A	27
5.12	Circulargraph del test Beta condotto dal partecipante A	27
5.13	Circulargraph del test Alfa condotto dal partecipante B	28
5.14	Circulargraph del test Alfa condotto dal partecipante C	29
6.1	Natura dell'incertezza	34
7.1	Esempio di event storming utilizzato per Epimetheus	35
A.1	Natura dell'incertezza	44

Capitolo 1

Introduzione

Al giorno d'oggi dobbiamo accettare i limiti della conoscenza e abituarci a convivere con l'incertezza, l'imprevisto e l'ignoto. Per lungo tempo, la natura e i suoi fenomeni erano avvolti da un'aura di mistero, gli eventi della vita, così come il dolore, la malattia e la morte, erano imprevedibili e incontrollabili, attribuiti agli umori e ai capricci degli dei, spesso irascibili e pernalosi. All'inizio del XX secolo, la scoperta della meccanica quantistica e lo studio dei sistemi complessi hanno messo tutto in discussione. Gli scienziati hanno compreso che molti fenomeni fisici, biologici e sociali sono governati dalle leggi della probabilità. In altre parole, gran parte di ciò che accade intorno a noi non può essere spiegato attraverso formule binarie, ma solo descritto mediante funzioni di probabilità. Questo ha segnato la fine dei dogmi, delle certezze assolute e ha sancito l'accettazione del dominio dell'incertezza nelle nostre vite.

Al giorno d'oggi questa incertezza persiste anche con l'aumento dei dati disponibili per le decisioni. Abbiamo imparato che esiste una dissociazione incolmabile tra ciò che si può prevedere disponendo di molti dati e ciò che può essere utile per decidere nelle situazioni specifiche che ci riguardano.

Con l'avvento del Machine Learning in ambito medico si è tentato di limitare l'incertezza, di fare predizioni quanto più precise possibili circa il futuro di un paziente a seguito di terapie o operazioni, ed è qui che entra in gioco Epimetheus.

1.1 Obiettivi della tesi

Epimetheus è una web application volta a fornire supporto nell'ambito delle decisioni mediche. Si tratta di un software realizzato su una base di Machine Learning che usa delle predizioni per calcolare quale potrà essere lo stato di salute del paziente 6 mesi dopo un'operazione, e lo mostra all'utente attraverso l'uso di grafici. Le operazioni prese in esame sono quelle all'anca e ginocchio. Tale software verrà utilizzato presso l'ospedale Galeazzi di Milano; l'Istituto Ortopedico Galeazzi infatti è un grande ospedale universitario dove molti interventi invasivi sono affrontati con regolarità. Tuttavia non tutti gli interventi hanno come esito un miglioramento clinico, e il 22% di questi hanno addirittura un peggioramento.

Attraverso la Data Science, quindi, Epimetheus vuole svolgere la funzione di supporto al medico affinché possa stabilire se un'operazione avrà esito positivo. Ciò avrà come conseguenza ridurre gravosi interventi a pazienti che non ne avrebbero giovamento,

ridurre le spese sostenute per trattamenti non necessari, e ottimizzazione delle risorse.

L'obiettivo di questa tesi è la trasformazione di Epimetheus da [Proof of Concept](#) a software scalabile. Per fare ciò ho assunto il ruolo di Project Manager coordinando il lavoro di altri tesisti, definendone la nuova architettura e gli sviluppi futuri; UX/UI Designer intervenendo sulla User experience del prodotto e migliorandone l'usabilità; Full Stack Developer lavorando sul codice frontend e backend.

Capitolo 2

Incertezza nelle decisioni mediche

In quest'era di medicina di precisione e di medicina predittiva, le incertezze che i medici devono affrontare sono sorprendentemente vaste, ben oltre le aspettative di una società che crede in una scienza moderna infallibile. Nonostante i progressi significativi nel campo della ricerca biomedica, i medici spesso si trovano a dover fare scelte basate su dati non completi e su una comprensione parziale di certe malattie. La medicina si trova a navigare in un mare di variabili aleatorie e incertezze, una realtà difficile da accettare per i pazienti, i quali cercano risposte chiare e sicure riguardo alla loro salute attuale e futura.

La realtà è che in sanità l'incertezza non è un'eccezione ma una costante. Questa condizione viene descritta come l'incapacità di prendere decisioni a causa di una percezione soggettiva dell'ignoranza, definita come "meta-ignoranza"¹. Spesso, l'ignoranza è considerata inaccettabile in ambito sanitario, che richiede certezze e prove inequivocabili per formulare previsioni accurate e decisioni professionali prive di dubbi. Di conseguenza, i professionisti sanitari possono sviluppare atteggiamenti di avversione o negazione dell'incertezza, cercando rifugio in false certezze.

Diviene quindi complesso per i medici comunicare queste sfumature, poiché potrebbe sembrare che ammettere la presenza di incertezza minacci la loro affidabilità. La responsabilità dei medici non è solo quella di gestire il rischio, ma anche di supportare i pazienti nel comprendere e accettare l'incertezza relativa alle loro condizioni cliniche.

2.1 Definizione

Ma cosa si intende per incertezza? L'incertezza è un termine che assume diversi significati in vari ambiti del sapere. Etimologicamente parlando, la parola certezza deriva dal verbo latino "cernere", che significa separare, distinguere. "Certum", il participio passato di "cernere", indica qualcosa di separato, distinto e chiaramente delimitato². Al contrario, le circostanze incerte sono quelle che non sono chiaramente distinguibili o separabili. Non sorprende quindi che l'incertezza venga definita come

¹Arora NK Han PK Klein WM. «Varieties of uncertainty in health care: a conceptual taxonomy». In: (2011).

²Giuseppe Parisi e Peter Konstantin Kurotschka. «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica». In: 113 (2022).

2.2. METODOLOGIA E VALUTAZIONE DEL RISCHIO NELLA PRATICA CLINICA⁴

incapacità di prendere decisioni. Questa incapacità può essere attribuita ai fatti e alla realtà o può riguardare il soggetto. Han et al. hanno definito l'incertezza come una "meta-ignoranza"³, ovvero la percezione soggettiva dell'ignoranza. In questa accezione, l'incertezza non è nella realtà, ma è un fenomeno cognitivo, emotivo e comportamentale che si manifesta nel soggetto incerto attraverso una serie di fenomeni cognitivi⁴ e affettivi, tra cui:

- Dubbio;
- Percezione di indefinito;
- Percezione di indeterminazione;
- Sensazione di non essere credibili;
- Ansia;
- Comportamenti di evitamento dell'incertezza (ambiguity aversion);

Secondo Han sono almeno tre livelli di incertezza (2.1). Il primo tipo di incertezza è quello scientifico, che deriva dalle informazioni con cui il professionista si confronta per prendere decisioni. L'incertezza è legata all'indeterminatezza degli eventi futuri, spesso espressa in termini percentuali di probabilità. Questa incertezza è misurabile e fornisce una guida utile per il decisore. Il grado di incertezza aumenta se le informazioni disponibili sono ambigue, ovvero non univoche, inaffidabili o difficili da interpretare⁵. Cresce ulteriormente quando le informazioni sono complesse, con molteplici fattori causali, stati di eventi, esiti o interpretazioni possibili, e quando i confini tra eventi sono vaghi o difficili da descrivere.

Un secondo tipo di incertezza è di natura pratico-organizzativa⁶, poiché qualsiasi intervento diagnostico o terapeutico dipende dalla qualità del sistema sanitario in cui è inserito.

Infine, l'incertezza può derivare dalle informazioni fornite dal paziente stesso sulla propria salute e sulla sua rete di assistenza. Questo è particolarmente rilevante nelle cure primarie, dove quasi tutti gli interventi iniziano dalla soggettività del paziente.

Le situazioni cliniche possono quindi essere descritte come costituite da un nucleo di rischio, con probabilità di esito noto, circondato da una "nuvola" di incertezza, che diventa più ampia quanto più vaga e imprecisa è l'informazione clinica disponibile⁷. Questa nuvola di incertezza deve essere sempre considerata, poiché influisce significativamente sulle probabilità degli esiti.

2.2 Metodologia e valutazione del rischio nella pratica clinica

Il ragionamento clinico è fondamentale in medicina per affrontare qualsiasi situazione. Oggi, si riconosce che l'approccio del medico alla situazione clinica non si basa solo

³Han PK, «Varieties of uncertainty in health care: a conceptual taxonomy».

⁴Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

⁵Ibid.

⁶Ibid.

⁷Luca Pasolli Giuseppe Parisi. *Clinica, relazione, decisione*. Franco Angeli, 2016.

2.2. METODOLOGIA E VALUTAZIONE DEL RISCHIO NELLA PRATICA CLINICA5

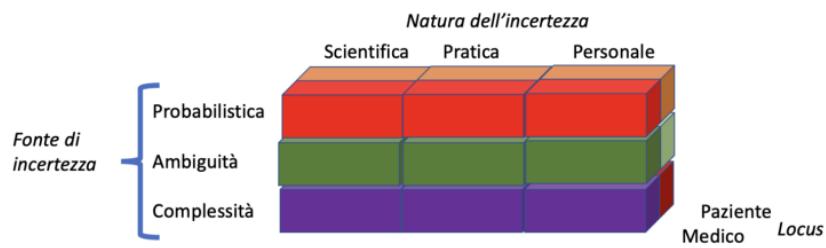


Figura 2.1: Natura e fonti dell'incertezza secondo Han et al.

su modelli analitici, consci e controllati dal ragionamento, ma anche su modelli non analitici, inconsci e intuitivi. L'insegnamento del metodo clinico raccomandato, quindi, include entrambe queste tipologie. Quando non è possibile affidarsi a un dato numerico o statistico, i medici si affidano a quella che viene chiamata *gut feeling*⁸, oppure all'euristica⁹.

In queste occasioni, è la sensazione del medico a indicare qual è la decisione giusta da prendere; c'è una sorta di consapevolezza fisica su quale sarebbe la decisione giusta, senza ragionarci troppo sopra.

Questa sensazione in realtà non è solo un'intuizione, è soprattutto un'attivazione del sistema nervoso enterico, che è composto da milioni di neuroni collegati al nostro intestino. L'euristica somiglia al gut feeling perché non è sempre razionale. Si basa sull'identificazione di poche informazioni importanti, vengono valutate e si interrompe la ricerca quando si ritiene che queste informazioni siano sufficienti per una decisione adeguata.

La conclusione è che il percorso diagnostico del medico si alterna tra euristiche, ragionamento deduttivo, e confronto con altri¹⁰. Tuttavia non sono le uniche opzioni disponibili per i medici; tra le strategie decisionali rapide ma basate su evidenze vi sono le cosiddette *Clinical decision rules* o *Clinical prediction rules*, in cui pochi segni e sintomi, insieme a esami *Point-of-care*¹¹, possono essere utilizzati per prendere rapidamente una decisione clinica.

Gigerenzer dimostra¹² come nelle situazioni di rischio conosciuto, i sistemi di decisione basati sull'analisi delle probabilità siano i migliori, mentre nelle situazioni di incertezza, le strategie euristiche rapide dei professionisti abbiano maggiore successo. Questo approccio si basa sull'idea che, in molte situazioni della vita reale, le risorse cognitive (come il tempo e l'informazione completa) sono limitate, e quindi le persone devono fare affidamento su metodi di decisione che siano efficienti in termini di risorse. In medicina, quindi, in molti casi, le strategie rapide possono essere più efficaci di quelle basate esclusivamente su metodi analitici. Utilizzare strategie euristiche evita anche di aumentare le indagini diagnostiche in situazioni di incertezza.

⁸Daniele Coen. *L'arte della probabilità. Certezze e incertezze della medicina*. Leanpub, 2021.

⁹Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

¹⁰Coen, *L'arte della probabilità. Certezze e incertezze della medicina*.

¹¹Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

¹²Gigerenzer G. «Fast and frugal heuristics: the tools of bounded rationality. Blackwell handbook of judgment and decision making». In: (2004).

Nella pratica clinica, i medici si confrontano quotidianamente con informazioni probabilistiche per prendere decisioni¹³. Esempi includono stime di sopravvivenza alla malattia in ambito oncologico, valutazione del rischio cardiovascolare, e la probabilità di eziologia batterica o virale in infezioni respiratorie. L'incertezza che ne deriva è legata all'indeterminatezza degli eventi futuri, spesso espressa in percentuali di probabilità¹⁴. Questa incertezza è misurabile e fornisce una guida utile al decisore, ma non è eliminabile, poiché rimane una previsione probabilistica e non deterministica. Nelle situazioni più ambigue e con esiti incerti, è cruciale valutare attentamente il ruolo della consulenza specialistica e la gestione condivisa del caso, che possono ridurre l'incertezza a livelli accettabili per il paziente. Tuttavia, è importante riconoscere che un certo grado di incertezza sarà sempre presente e dovrà essere gestito attraverso una comunicazione chiara e trasparente con il paziente.

2.3 Rapporto con il paziente

Comunicare le incertezze al paziente è fondamentale per costruire un percorso decisionale consapevole. Nonostante la medicina moderna si basi su numeri e analisi statistiche di esperienze pregresse, presupponendo che ogni opzione diagnostica o terapeutica sia la più appropriata, la comunicazione del medico trasmette non solo logica, ma anche significato, influenzando la percezione del paziente e le sue aspettative. È essenziale che la comunicazione delle incertezze non confonda ulteriormente il paziente, ma che si adatti alle sue caratteristiche individuali (età, cultura, fattori psicologici) sia nella complessità dell'informazione che nel linguaggio utilizzato.

Le situazioni cliniche possono essere rappresentate come formate da un corpo centrale di rischio conosciuto, affrontabile con la conoscenza probabilistica, circondato da una nube di incertezza(2.2). La gestione dell'incertezza in medicina non dovrebbe limitarsi a una semplice ricognizione anamnestica, ma deve essere un aspetto chiave del metodo clinico, incorporato in tutti gli aspetti del processo decisionale dei professionisti sanitari. Questo è essenziale per attuare azioni appropriate.

Nel processo decisionale clinico, l'incertezza influisce sulla probabilità che un evento si verifichi o meno. Riconoscere l'incertezza consente di comunicarla, prendere decisioni più consapevoli e praticare una medicina più prudente e meno onnipotente. In base alle preferenze espresse dal paziente, è possibile costruire una rete protettiva (safety netting) adottando una strategia di attesa. Vengono descritti al paziente una serie di sintomi di allarme (red flags)¹⁵ potenzialmente riscontrabili e i tempi e le modalità per cercare aiuto medico se questi sintomi si manifestassero o se non ci fossero miglioramenti in un periodo prestabilito. Idealmente, queste informazioni dovrebbero essere fornite per iscritto, dopo un processo decisionale condiviso.

Esistono situazioni in cui la diagnosi è incerta e l'alternativa potrebbe essere una malattia grave a progressione rapida. In questi casi, si può utilizzare il tempo come

¹³Coen, *L'arte della probabilità. Certezze e incertezze della medicina*.

¹⁴Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

¹⁵Ibid.

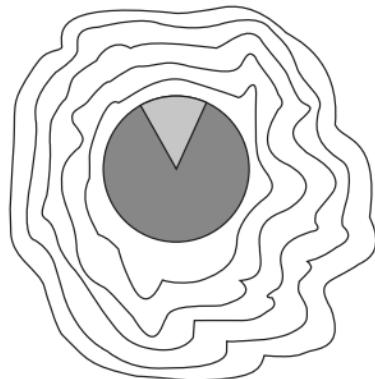


Figura 2.2: Illustrazione di come si configurano le situazioni cliniche

La parte in grigio scuro rappresenta il rischio in percentuale che un evento si verifichi; il grigio chiaro rappresenta il rischio in percentuale che un evento non si verifichi; le linee circolari intorno rappresentano la nube di incertezza

forma di test, simile a un esame ematico o strumentale¹⁶¹⁷. In questo caso l'aspetto comunicativo della strategia è cruciale, perché il medico deve creare una buona relazione di fiducia e collaborazione con il paziente, affinché quest'ultimo possa comprendere chiaramente le istruzioni. Il paziente, inoltre, deve essere consapevole dell'incertezza diagnostica, dei segni e sintomi a cui prestare attenzione e del periodo di osservazione, nonché sapere come cercare aiuto¹⁸.

È fondamentale anche l'aspetto organizzativo: il medico e il suo team devono avere accesso alle informazioni sulla decisione presa con il paziente in ogni momento, e deve esistere un sistema di ricezione delle chiamate che preveda una risposta immediata.

In generale comunque i risultati non sono sempre chiari, di fatto lo strumento più potente per aiutare il medico a decidere nell'incertezza è la condivisione del processo decisionale con il paziente o con chi se ne prende cura.

Al giorno d'oggi il paziente può influire sul processo decisionale in medicina. Kon sostiene che la decisione condivisa esiste su un continuum con il modello paternalistico (guidato dal medico) a un'estremità e il modello informativo (guidato dal paziente) all'altra, mentre al centro c'è la *decisione condivisa*¹⁹ in cui medico e paziente contribuiscono equamente, ciascuno al 50%.

Occorre considerare quale situazione clinica richiede un maggiore coinvolgimento del paziente. Questa è spesso una situazione di incertezza, contrapposta a una in cui esiste una chiara opzione migliore con un rapporto costi/benefici ben definito. In contesti incerti, le preferenze e i valori del paziente diventano cruciali poiché non ci sono rapporti chiari tra costi e benefici. Più è grande l'incertezza, più è necessario

¹⁶Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

¹⁷Heneghan C, Glasziou P e Thompson M et al. «Diagnostic strategies used in primary care». In: (2009).

¹⁸Coen, *L'arte della probabilità. Certezze e incertezze della medicina*.

¹⁹MD Alexander A. Kon. «The Shared Decision-Making Continuum». In: (2010).

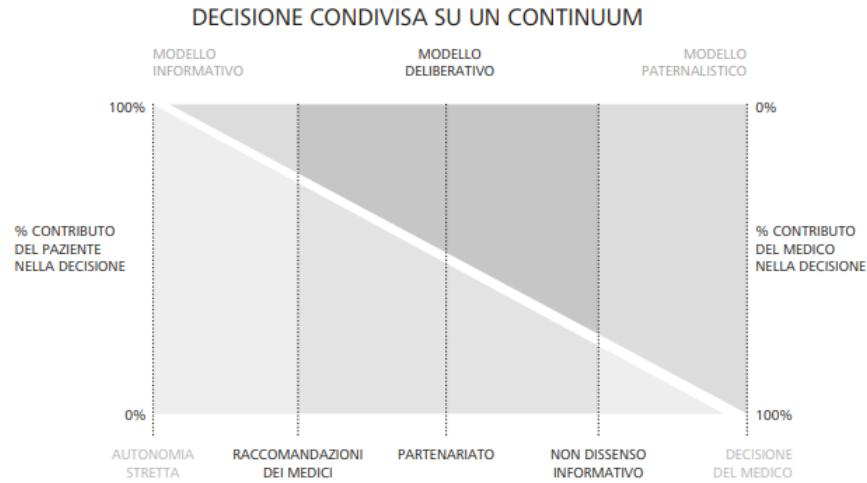


Figura 2.3: Modello della decisione condivisa secondo Kon

coinvolgere il paziente, valutando le sue conoscenze e capacità. Questo non significa solo considerare le preferenze, ma anche integrare il contributo del paziente nella decisione.

È importante distinguere la decisione condivisa dal semplice consenso informato²⁰: quest'ultimo implica un flusso informativo unidirezionale dal medico al paziente, che è eticamente e legalmente obbligatorio ma non necessariamente parte di una decisione già presa dal medico. La vera decisione condivisa implica invece uno scambio di idee tra medico e paziente, una collaborazione che porta a una decisione comune. In questo processo, il paziente agisce come consulente, aiutando il medico a prendere una decisione, il che è molto diverso da un semplice processo informativo.

²⁰Parisi e Kurotschka, «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica».

Capitolo 3

Sistemi di supporto alle decisioni

I Sistemi di Supporto alle Decisioni (DSS) sono strumenti informatici progettati per assistere gli utenti nel processo decisionale, utilizzando dati e modelli per risolvere problemi complessi e non strutturati. Questi sistemi combinano l'uso di dati, modelli analitici e strumenti di supporto interattivi per aiutare a prendere decisioni più informate. I DSS possono essere applicati in vari settori, tra cui il business, la sanità, la finanza e la gestione delle risorse umane.

Si suddividono in tre categorie principali:

- DSS orientati ai dati: si concentrano sulla raccolta e l'analisi dei dati. Utilizzano database e data warehouse per memorizzare grandi quantità di dati e strumenti di data mining per estrarre informazioni utili.
- DSS orientati ai modelli: utilizzano modelli matematici e statistici per analizzare i dati e supportare il processo decisionale. Questi modelli possono includere modelli di ottimizzazione, simulazione, previsione e analisi delle serie temporali.
- DSS orientati alla conoscenza: incorporano conoscenze specifiche del dominio, regole e logiche di business per supportare decisioni complesse. Spesso utilizzano sistemi esperti e intelligenza artificiale per rappresentare e utilizzare la conoscenza.

I DSS hanno avuto una significativa evoluzione nel corso degli anni: originariamente erano semplici strumenti di analisi statistica utilizzati per supportare decisioni aziendali specifiche. Con l'avvento della tecnologia dell'informazione e delle comunicazioni sono diventati più complessi e potenti, integrando nuove tecnologie come il data warehousing, il data mining e, più recentemente, il machine learning e l'intelligenza artificiale (AI). Negli ultimi anni, il machine learning è diventato una componente essenziale dei DSS moderni, infatti consente ai sistemi di apprendere dai dati e migliorare le loro prestazioni senza essere esplicitamente programmati. Questo è particolarmente utile nei casi in cui la capacità di analizzare grandi volumi di dati e identificare pattern nascosti può migliorare significativamente il processo decisionale.

È essenziale che i DSS moderni comunichino l'incertezza nei dati e nei risultati delle analisi. La trasparenza aiuta i decisorи a comprendere i limiti delle previsioni e a prendere decisioni più informate. L'incertezza può derivare da vari fattori, tra cui la qualità dei dati, l'accuratezza dei modelli utilizzati e l'aleatorietà intrinseca dei fenomeni osservati¹.

¹S. K. Card et al. «Understanding the Role of Visualizations on Decision Making: A Study on Working Memory». In: (2020).

La comunicazione dell'incertezza è particolarmente importante in settori dove le decisioni hanno un impatto significativo sulla vita delle persone, come la sanità. Ad esempio, un DSS utilizzato per diagnosticare una malattia deve comunicare chiaramente l'incertezza associata alla diagnosi, in modo che i medici possano considerare tutte le possibili opzioni di trattamento.

3.1 Machine learning in ambito medico

Uno degli ambiti più promettenti per l'applicazione del machine learning nei DSS è la medicina.

Negli ultimi anni, l'IA ha conosciuto uno sviluppo significativo, rendendo possibile immaginare un futuro in cui diagnosi errate e trattamenti sintomatici saranno superati. La gestione e l'analisi delle enormi quantità di dati generate dalle immagini mediche e dai test diagnostici consentono all'IA di sviluppare applicazioni sofisticate, inaugurando un'era di medicina elettronica.

Nonostante alcuni algoritmi siano in grado di eguagliare e talvolta superare i clinici in varie attività, l'IA non è ancora integrata completamente nella pratica medica quotidiana.

Molti algoritmi di IA possono apprendere dai dati, sia che si tratti di immagini mediche (come le scansioni MRI) sia di dati numerici (come la pressione sanguigna). Dopo aver elaborato questi dati, gli algoritmi sono in grado di fornire risultati probabilistici o classificazioni, come identificare un campione di tessuto come canceroso o stimare la probabilità di un coagulo arterioso. Le prestazioni degli algoritmi vengono poi confrontate con quelle dei medici per determinare se la diagnosi dell'IA è accurata e clinicamente rilevante.

Un esempio significativo di IA in ambito medico è l'algoritmo DLAD (Rilevamento Automatico basato su Apprendimento Profondo)², sviluppato presso l'Università Nazionale di Seoul. Questo algoritmo analizza le radiografie del torace per rilevare crescite cellulari anomale, come i tumori. Nei test comparativi, DLAD ha dimostrato di superare 17 su 18 medici nella capacità di rilevamento delle anomalie.

Un altro esempio significativo di algoritmo IA nel settore medico è stato sviluppato nell'autunno del 2018 dai ricercatori di Google AI Healthcare³. Questo algoritmo, denominato LYNA (Assistente dei Linfonodi), è stato progettato per identificare i tumori metastatici del cancro al seno dalle biopsie dei linfonodi. LYNA rappresenta un progresso significativo poiché è in grado di individuare aree sospette nei campioni di biopsia, una capacità che va oltre la percezione visiva umana. Nei test condotti su due diversi database, LYNA ha dimostrato un'accuratezza del 99% nel classificare i campioni come cancerosi o non cancerosi. Inoltre, ha ridotto della metà il tempo medio di revisione delle diapositive quando utilizzato dai medici come supporto alla loro analisi tradizionale.

Altri algoritmi basati su immagini hanno mostrato abilità simili nel migliorare l'accuratezza diagnostica dei medici.

Esempi come DLAD e LYNA dimostrano come gli algoritmi possano supportare i medici nella classificazione di campioni patologici, evidenziando caratteristiche delle

²Artificial Intelligence in medicine: applications, implications, and limitations. URL: <https://www.hunimed.eu/news/artificial-intelligence-in-medicine-applications-implications-and-limitations/>.

³Ibid.

immagini che necessitano di un'analisi più approfondita.

Recentemente è stato condotto un nuovo studio⁴ che esplora l'applicazione delle tecniche di machine learning (ML) in ambito ortopedico; il suo focus è esaminare articoli pubblicati negli ultimi vent'anni e farne una revisione.

In questo studio il machine learning è definito come lo studio di come gli algoritmi possono "imparare" relazioni complesse dai dati empirici, producendo modelli matematici che collegano numerose variabili a una variabile target di interesse. In medicina, questo significa poter prevedere, data una serie di immagini radiologiche, risultati di laboratorio o dati estratti da registri elettronici, etichette diagnostiche, livelli di risultato, valori di esami o opzioni di trattamento per aiutare i medici a prendere decisioni più accurate ed efficienti.

La revisione della letteratura è stata condotta eseguendo una ricerca sui database Medline e Scopus, includendo articoli che utilizzano tecniche di ML per il sistema muscoloscheletrico umano. Sono stati selezionati sei settori anatomici principali: colonna vertebrale, anca, ginocchio, caviglia, mano e piede, oltre a una procedura generale, l'artroplastica. Sono stati selezionati 70 articoli per una revisione approfondita del loro contenuto e codifica.

Gli articoli sono stati divisi in due categorie principali: tecniche di ML convenzionali e deep learning. Tecniche di ML convenzionali:

- Decision Trees e Random Forests: Queste tecniche sono state utilizzate per classificare i soggetti con osteoartrite e per fornire interpretabilità clinica dei risultati.
- Nearest Neighbors (NN): Utilizzate per varie applicazioni, tra cui la segmentazione delle immagini.
- Regressione Lineare e Altre Tecniche Simili: Impiegate per modellare relazioni lineari tra variabili.
- Support Vector Machines (SVM): Ampiamente usate per la classificazione e il rilevamento delle caratteristiche.
- K-means Clustering e Altre Tecniche Simili: Utilizzate per raggruppare i dati in base a caratteristiche simili.
- Altre Tecniche Discriminative: Come gradient boosting machines e LDA.
- Tecniche Generative: Come i modelli probabilistici, utilizzati per prevedere la progressione della scoliosi.

Gli autori hanno concluso che, sebbene il ML abbia dimostrato risultati promettenti in vari campi ortopedici, è necessaria una valutazione rigorosa e una validazione in contesti reali prima che possa essere ampiamente adottato nella pratica clinica. Attualmente, l'adozione del ML in ortopedia è ancora in una fase preliminare, con necessità di ulteriori studi e ricerche per consolidarne l'applicabilità e l'efficacia.

⁴Giuseppe Banfi Federico Cabitza Angela Locoro. «Machine Learning in Orthopedics: A Literature Review». In: (2018).

3.2 Visualizzazioni vaghe

La comunicazione dell'incertezza è particolarmente importante in settori dove le decisioni hanno un impatto significativo sulla vita delle persone, come la sanità. Ad esempio, un DSS utilizzato per diagnosticare una malattia deve comunicare chiaramente l'incertezza associata alla diagnosi, in modo che i medici possano considerare tutte le possibili opzioni di trattamento.

Le visualizzazioni vaghe sono un concetto importante in data science che aiuta a gestire l'incertezza e la variabilità nei dati. Quando si tratta di analisi dei dati, spesso ci troviamo a dover affrontare informazioni che contengono errori, rumore o incertezze intrinseche. Le visualizzazioni vaghe sono progettate per rappresentare queste incertezze in modo che gli utenti possano avere una comprensione più completa e affidabile delle informazioni presentate.

Uno dei metodi più comuni utilizzati nelle visualizzazioni vaghe è l'impiego degli intervalli di confidenza: questi intervalli indicano la gamma di valori entro cui si prevede che un parametro si trovi con una certa probabilità. Ad esempio, in un grafico a barre, possiamo vedere delle linee verticali che rappresentano l'intervallo di confidenza per ciascuna barra, fornendo così un'indicazione visiva dell'incertezza associata a ciascun dato.

Nei grafici a linee, le bande di incertezza sono spesso utilizzate per mostrare la variabilità intorno a una linea di tendenza. Queste bande, che possono essere ombreggiate, offrono una rappresentazione visiva chiara dell'incertezza, aiutando a comprendere meglio quanto ci si può fidare di una previsione o di una tendenza osservata. Anche le mappe di calore (heatmaps), sono strumenti efficaci in questo contesto, poiché possono rappresentare dati spaziali o temporali con variazioni di colore o intensità per indicare incertezze.

I grafici di tipo violin plot e box plot sono altre tecniche utili, poiché permettono di visualizzare la distribuzione dei dati insieme alle indicazioni di variabilità e densità. In sostanza la visualizzazione vaga si propone di rappresentare dei risultati non in formato numerico o simbolico, ma attraverso immagini pittoriche in cui prevale un'incertezza visiva. Questo tipo di rappresentazione visiva è legata a tre fattori principali: vaghezza, indistintezza e sfocatura.

Attualmente, nelle scienze dure, come la matematica, la logica e le scienze naturali (biologia, chimica, fisica), l'incertezza viene rappresentata in termini di probabilità, punteggi di confidenza o percentuali. In questo approccio non è sempre chiaro se l'incertezza venga realmente compresa dai medici, con conseguente rischio di sopravvalutare le informazioni, un fenomeno noto come bias della quantificazione⁵.

Secondo lo studio "Vague Visualizations to Reduce quantification bias in shared medical decision making", dalle immagini è possibile trarre valori numerici in modo "immediato" attraverso diverse modalità, come la posizione su scale graduate, segmenti su un piano cartesiano, angoli o sfumature di colore.

Lo scopo delle visualizzazioni vaghe è quindi sfruttare il gut feeling degli utenti, ovvero lasciare che sia la loro percezione a guidarli piuttosto che la razionalità.

L'incertezza può essere rappresentata con varie tecniche e modalità. Non è stata ancora definita la "forma" migliore da utilizzare, ma quello che sembra ormai certo è che colore, tonalità, saturazione del colore, forma e la trasparenza siano i mezzi più

⁵Silvia Bordogna Michela Assale e Federico Cabitza. «Vague Visualizations to Reduce Quantification Bias in Shared Medical Decision Making». In: (2020).

efficaci. In questo studio comparativo si è evidenziato che la sfocatura, la posizione e la trasparenza favoriscono la percezione e l'intuizione da parte del fruitore finale, mentre la saturazione risulta meno utile. Relativamente alla tecnica da utilizzare, come ad esempio i glifi di linea, alcuni ricercatori hanno evidenziato che la tecnica stessa dipende anche dalla capacità di "traduzione" o decodifica dell'utente finale.

In questo studio si è voluto valutare l'efficacia rappresentativa di tre effetti visivi, precisamente: la sfocatura, la trasparenza e il rumore, nel comunicare una probabilità di rischio. Gli autori intendono per efficacia rappresentativa la rappresentazione, e quindi la soluzione grafica, che permette valutazioni coerenti, senza indurre l'utente a sovrastimare o a sottostimare il valore della probabilità. È evidente che le VV non devono ostacolare la comprensione o fuorviare i medici nelle loro valutazioni e scelte diagnostiche e terapeutiche.

Lo studio si è posto due domande:

- la VV ostacola o favorisce la stima delle probabilità?
- In caso di risposta affermativa, c'è qualche effetto tra quelli utilizzati che è più efficace o meno fuorviante degli altri?

Per testare le domande è stato creato un software web based che accetta qualsiasi immagine raster (formata da pixel) e una percentuale di probabilità come input, e restituisce in output la medesima immagine influenzata però dagli effetti visivi precedentemente specificati. Il 100% di probabilità coincide con l'immagine originale, quindi purezza al 100 per cento; il valore 0% corrisponde alla massima distorsione e viene rappresentata la massima incertezza. Nello studio sono state create 6 visualizzazioni con le seguenti percentuali: 10%, 25%, 40%, 60%, 75% e 90%; queste percentuali corrispondono a diversi quartili.



Figura 3.1: Effetti applicati all'immagine per rendere il rischio del 25%, 50% e 75%. T, B e N sono, rispettivamente, l'effetto della trasparenza, sfocatura e rumore

È stato strutturato un questionario online proposto poi a un certo numero di studenti universitari e conoscenti. I partecipanti avevano come compito quello di associare due delle 6 VV (2VV scelte casualmente per ogni effetto visivo) in due compiti di difficoltà crescente. Per tutti e due i compiti, è stato somministrato un set di riferimento a tre VV che indicavano un valore del 0%, 50% e 100%, rispettivamente.

In questo questionario il primo compito riguardava l'accuratezza visiva (RA), il secondo compito l'accuratezza assoluta. Nel presentare le visualizzazioni su cui si basava il questionario si ipotizzavano due possibili fonti di bias che avrebbero potuto influenzare l'analisi:

- l'ordine delle domande (bias di ordine);
- il valore della percentuale mostrata (ovvero, bias di campionamento).

Per ridurre il primo tipo di bias, il questionario online è stato implementato in modo che la presentazione dei 3 effetti diversi ai partecipanti fosse in ordine casuale. Le VV sono risultate utili nel comunicare probabilità sotto e sopra il 50%, con una tendenza a sovrastimare i valori bassi e a sottostimare quelli alti. Non sono emerse differenze significative tra sfocatura, trasparenza e rumore, eccetto la trasparenza, più efficace al 40%.

In particolare le visualizzazioni vaghe sono uno strumento valido per comunicare valori intermedi, mentre è stata registrata una regressione alla media quando vengono mostrati valori estremi.

Infine, non è stato individuato un metodo migliore di altri. Gli studiosi contano di effettuare altri esperimenti in ambienti controllati e reali per verificare se le decisioni prese dai medici sono diverse quando la previsione del rischio è rappresentata con quantità evidenti, o mediante una visualizzazione vaga. Si conta anche di misurare la soddisfazione dei medici.

Un altro studio interessante che utilizza le visualizzazioni vaghe è lo studio "Comparative Assessment of Two Data Visualizations to Communicate Medical Test Results Online"⁶, dove si prendono in esame test diagnostici basati su biomarcatori, cioè indicatori di una specifica condizione, nella fattispecie l'infezione da SARS-CoV-2, causa di COVID-19.

Ogni test diagnostico, sia esso basato su imaging, carica virale o presenza di antigeni, è associato a un certo margine di errore, ma razionalmente tendiamo a considerare l'esito del test solo in termini di sì/no, evitando di valutare la probabilità di avere o non avere una specifica condizione. Utilizzando gli strumenti di supporto decisionale basati su ML, l'elemento probabilistico può essere reso visibile e quindi esplicitato, ad esempio riportando i punteggi di probabilità o rappresentandoli in un qualche modo: in questo studio i ricercatori ritengono che questo possa aiutare gli utenti ad interpretare al meglio l'output della macchina. In pratica l'incertezza intrinseca di questi modelli può costituire un plus nel processo decisionale, ad esempio per scegliere se sottoporsi ad un ulteriore esame o per prediligere una terapia.

Tale studio ha quindi lo scopo di scegliere la migliore visualizzazione dei dati da presentare agli utenti relativamente ad un test ematologico per rilevare le infezioni da COVID-19, mediante il Conteggio Ematologico Completo, utilizzando un modello di ML. Questo modello è stato convalidato nella letteratura di riferimento e poi incorporato in uno strumento basato sul Web.

⁶Andrea Campagner Federico Cabitza e Enrico Conte. «Comparative Assessment of Two Data Visualizations to Communicate Medical Test Results Online». In: (2022).

Sono state confrontate due visualizzazioni progettate secondo i principi delle visualizzazioni vaghe: le stime incerte sono rappresentate evitando volutamente l'utilizzo dei simboli, cioè numeri e rappresentazioni metriche, come estensioni della lunghezza e angoli. Secondo le caratteristiche delle visualizzazioni vaghe, le quantità probabilistiche sono presentate come indizi visivi, i quali sono difficili da interpretare in termini razionali, cioè sono difficili da associare a valori numerici. Si utilizzano specificamente tonalità di colore o gradienti di saturazione e luminosità. Questa modalità è voluta e mira a comunicare ai lettori un senso di incertezza e vaghezza per far sì che i lettori comprendano effettivamente le stime visualizzate, come rischi, probabilità, dispersione. È evidente che le visualizzazioni vaghe richiedono una attenzione aggiuntiva.

Le due visualizzazioni citate sono state ideate durante due sessioni di progettazione in cui sono stati coinvolti sia gli autori di questo articolo sia i clinici che hanno sviluppato il modello statistico. I clinici sono stati adeguatamente edotti circa le caratteristiche delle visualizzazioni vaghe e sono stati invitati a co-progettare due visualizzazioni: una visualizzazione destinata a colleghi esperti nell'interpretare i test di laboratorio, e una più semplice che potesse essere più familiare ai pazienti testati.

Le visualizzazioni risultanti si basavano su metafore diverse: la prima visualizzazione si basava sul comune [test del tornasole](#) e sulla metafora della bolla livello; la seconda visualizzazione dei dati adottava la metafora del bastoncino di test, utilizzata, ad esempio, nei test di gravidanza, e quindi familiare al pubblico in generale. Lo studio è stato concepito per capire:

1. se la metafora del bastoncino di test fosse adeguata in caso di una risposta sensibile e delicata come quella relativa alla positività al COVID-19, o, come osservato in alcuni studi, finisse per confondere troppo le persone comuni
2. se una visualizzazione dei dati più tecnica, quella progettata per gli operatori sanitari, potesse essere comprensibile anche al pubblico comune.

Nella visualizzazione a bolla livello, il risultato del test è presentato attraverso la posizione di una bolla circolare all'interno di una barra a tre colori (simile al tornasole). Si valuta quindi la sua maggiore o minore vicinanza a uno degli estremi della barra per indicare una condizione COVID-19-positiva o negativa (rispettivamente all'estremo rosso più a sinistra e all'estremo blu più a destra). I risultati incerti, cioè quelli a bassa affidabilità, sono caratterizzati da una sostanziale equidistanza della bolla dagli estremi, corrispondente ad un posizionamento nella zona grigia centrale della barra tornasole. L'incertezza è anche resa evidente dalla dimensione della bolla: più grande è la bolla, maggiore è l'intervallo di confidenza della stima della probabilità.



Figura 3.2: Natura dell'incertezza

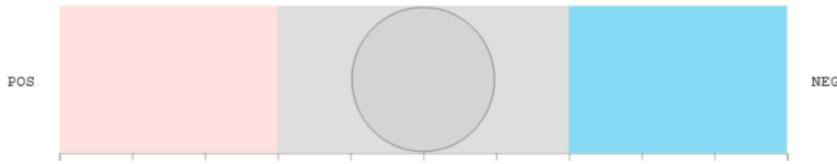


Figura 3.3: Natura dell'incertezza

La visualizzazione del bastoncino di test fornisce le medesime informazioni visualizzate tramite la bolla livello, ma attraverso affordance (in senso generale, favorire l'utilizzo) e segnali visivi diversi.

Nel dettaglio si vedono due fasce rosse: una per indicare l'affidabilità della risposta e indicata con una C maiuscola ("controllo"), e una che indica il risultato del test, indicata con un segno più singolo (+). In pratica questa visualizzazione restituisce l'output del modello in termini di opacità della barra: più trasparenti (e meno visibili) sono la fascia + e la fascia C, minore è la probabilità che il test sia associato a una condizione positiva e che il test sia affidabile. Un test quasi certamente negativo è quindi reso da un bastoncino in cui è chiaramente visibile solo la barra C, mentre un test non valido è rappresentato da un bastoncino in cui nessuna fascia rossa è visibile.

In conseguenza dei risultati positivi riscontrati nei due studi sopra citati, è nata l'idea di utilizzare lo stesso tipo di comunicazione dell'incertezza nel progetto Epimetheus.

Capitolo 4

Epimetheus

Il progetto Epimetheus nasce con l'obiettivo di migliorare le previsioni sullo stato di salute dei pazienti sottoposti a interventi chirurgici all'anca o al ginocchio. Questo prodotto si pone nel contesto degli strumenti di supporto alle decisioni mediche, utilizzando un complesso sistema di machine learning per elaborare predizioni a partire dai dati PROM (Patient-Reported Outcome Measures).

L'adozione di Epimetheus è prevista presso l'IRCCS Istituto Ortopedico Galeazzi (IOG) di Milano, un'importante struttura sanitaria universitaria specializzata nella diagnosi e nel trattamento dei disturbi muscolo-scheletrici. Presso lo IOG vengono eseguiti annualmente quasi 5000 interventi chirurgici, con una prevalenza di artroplastiche dell'anca e del ginocchio, oltre a numerose procedure relative alla colonna vertebrale. Epimetheus mira a rispondere a una necessità cruciale: migliorare la qualità delle previsioni cliniche post-operatorie. Attualmente, circa il 39% degli interventi chirurgici alla colonna vertebrale presso lo IOG non raggiunge un miglioramento clinicamente significativo, e il 22% di questi è associato a esiti negativi. Queste cifre sono dovute, in parte, al fatto che l'istituto Galeazzi è una struttura terziaria che accoglie i casi più complessi provenienti da un vasto territorio. Inoltre, le problematiche vertebrali, specialmente in una popolazione che invecchia, rappresentano sfide terapeutiche considerevoli.

All'atto pratico il progetto è composto principalmente da tre grafici, o visualizzazioni, che adottano il colore come strumento di comunicazione. I grafici sono:

- Boxplot: si tratta di una visualizzazione formata da due assi, l'asse X rappresenta il physical score pre-operatorio, l'asse Y rappresenta lo stesso score a 6 mesi dall'operazione. Il grafico è diviso in tre regioni principali che rappresentano lo stato di salute del paziente; in base a dove si colloca il punto del grafico si può intuire se il paziente avrà un miglioramento, peggioramento, o un esito in certo a seguito dell'operazione.
- Violinplot: questa visualizzazione compara lo stato del paziente con quello degli altri pazienti con PROM simili, mostra la distribuzione dei pazienti al pre-operatorio e al post-operatorio.
- La terza visualizzazione è un boxplot strutturato in maniera simile al test del tornasole valutato nello studio "Comparative Assessment of Two Data Visuali-

zations to Communicate Medical Test Results Online"¹ e spiegato nella sezione 3.2, e un grafico circolare dove all'esterno vi è la legenda, cioè il gradiente che rappresenta lo stato di salute del paziente, e all'interno un cerchio riempito del colore che più rappresenta quello che sarà lo stato di salute nel post-operatorio, quindi se migliorerà o peggiorerà.

4.1 PROM

Quando si parla di PROM² si parla di misure che valutano lo stato di salute percepito direttamente dal paziente. I PROM indagano vari aspetti dello stato di salute del paziente, come la percezione dei sintomi, il dolore, l'ansia, la depressione e il grado di affaticamento. Questi strumenti aiutano a comprendere come i pazienti percepiscono la loro salute e il loro livello di disabilità, oltre a misurare la qualità della vita correlata alla salute. I PROM sono essenziali per raccogliere dati sulle condizioni di salute dei pazienti, che possono includere sintomi fisici e mentali, nonché il loro impatto sulla vita quotidiana.

I PREM (Patient-Reported Experience Measures) invece sono questionari che misurano la percezione dei pazienti riguardo alla loro esperienza durante le cure ricevute. Valutano aspetti come la qualità della comunicazione tra il paziente e il personale sanitario, il supporto ottenuto per la gestione di condizioni a lungo termine, il tempo trascorso in attesa di ricevere assistenza e la facilità di accesso alle cure. I PREM forniscono un feedback cruciale sull'esperienza complessiva del paziente nel sistema sanitario, permettendo di identificare aree di miglioramento nei servizi offerti.

L'adozione di PROM e PREM nei PSP offre diversi vantaggi significativi:

- Valutazione del valore del PSP: questi strumenti permettono di dimostrare il valore del PSP in modo oggettivo agli specialisti sanitari. Monitorare l'andamento del trattamento attraverso PROM e PREM fornisce dati concreti che possono essere utilizzati per valutare l'efficacia del programma e migliorare la qualità delle cure fornite.
- Ottimizzazione dei servizi: PROM e PREM aiutano a ottimizzare i servizi offerti, consentendo di ridisegnare il percorso del paziente in base alle sue aspettative e necessità. Ad esempio, se i pazienti segnalano difficoltà nell'accesso alle cure o nella comunicazione con il personale sanitario, questi feedback possono essere utilizzati per apportare modifiche che migliorano l'esperienza del paziente.
- Personalizzazione dell'Assistenza: l'utilizzo di scale come la Morisky o l'ARMS permette di valutare l'aderenza del paziente al trattamento in modo oggettivo. Altre scale, come la PHE (Patient Health Engagement), utilizzate nei PSP di training, consentono di personalizzare l'assistenza in base al coinvolgimento del paziente nella gestione della propria terapia. Questo approccio centrato sul paziente assicura che le cure siano adattate alle specifiche esigenze e preferenze del paziente.

Nel software Epimetheus il PROM usato è il Short Form 12 (SF-12). È una versione ridotta del più completo SF-36 Health Survey e comprende 12 domande che

¹Federico Cabitza e Conte, «Comparative Assessment of Two Data Visualizations to Communicate Medical Test Results Online».

²L'utilizzo di PROM e PREM nei patient support program. URL: <https://phdlifescience.eu/utilizzo-di-prom-e-prem-nei-patient-support-program/>.

valutano diverse dimensioni della salute fisica e mentale. Le domande del SF-12 sono progettate per raccogliere informazioni su due dimensioni principali: la salute fisica e la salute mentale. La salute fisica viene valutata attraverso domande che esplorano la funzionalità fisica, il dolore e le limitazioni dovute a problemi di salute fisica. La salute mentale, invece, viene esaminata tramite domande sul benessere emotivo, le limitazioni dovute a problemi emotivi, il funzionamento sociale e l'energia o la vitalità. Le risposte si basano su una scala Likert, dove i partecipanti valutano la loro salute in vari aspetti. Il processo di scoring del SF-12 aggrega le risposte per produrre due punteggi principali: il Physical Component Summary (PCS) e il Mental Component Summary (MCS). Uno dei principali vantaggi del SF-12 è la sua brevità e facilità di somministrazione. Rispetto al SF-36, il SF-12 richiede meno tempo per essere completato, rendendolo pratico sia per studi su larga scala che per l'uso clinico quotidiano.

In Epimetheus tra le domande che si possono trovare vi sono per esempio la capacità di salire le scale, se il paziente si sente limitato nelle sue attività quotidiane, se il dolore ha interferito con le sue attività, in che misura il paziente si è sentito triste.

The form consists of several input fields and dropdown menus:

- Gender:** Radio buttons for Male and Female.
- Age patient:** Text input field.
- ASA Class:** Dropdown menu with "select ASA class" placeholder.
- Height Pre operation (cm):** Text input field.
- Weight Pre operation (kg):** Text input field.
- VAS Total pre operation:** Text input field.
- SF12 PhysicalScore Pre operation:** Text input field.
- SF12 MentalScore Pre operation:** Text input field.
- SF12 - Health self-assessment:** Text input field.
- SF12 - Ability to climb stairs:** Text input field.
- SF12 - Did the patient accomplished less than they would like due to physical problems?** Text input field.
- SF12 - Did the patient felt limited in their activities?** Text input field.
- SF12 - Did the patient accomplished less than they would like due to emotional problems?** Text input field.
- SF12 - How much pain interfered with the patient's normal work?** Text input field.
- SF12 - How much did the patient felt calm and peaceful?** Text input field.
- SF12 - How much did the patient felt having a lot of energy?** Text input field.
- SF12 - How much did the patient felt downhearted and blue?** Text input field.
- SF12 - Degree to which the patient's problems interfered with social activities:** Text input field.

At the bottom left is a **Reset** button, and at the bottom right is a **Submit** button.

Figura 4.1: Questionario presente nella pagina iniziale di Epimetheus

Capitolo 5

User Experience

Il redesign dell’interfaccia di un prodotto digitale richiede un’approfondita comprensione delle esigenze degli utenti, delle attività che devono svolgere e delle interazioni tra le varie componenti del sistema. Per raggiungere questo obiettivo, ho adottato un approccio integrato che combina il Design Centrato sull’Utente, il Design Centrato sull’Attività e il Design dei Sistemi, di cui parlerò nel capitolo 6. Questo metodo mi ha permesso di disegnare e successivamente sviluppare un’interfaccia intuitiva, efficiente e capace di rispondere alle aspettative degli utenti finali.

5.1 Design Centrato sull’Utente

Il Design Centrato sull’Utente (User-Centred Design, UCD) è stato il punto di partenza fondamentale per il mio lavoro. La comprensione del target di utenti ha guidato ogni decisione progettuale, assicurando che l’interfaccia fosse realmente adatta a chi l’avrebbe utilizzata.

Il processo di User Experience è partito con una intervista ai vari stakeholder: qui ho compreso quali fossero gli obiettivi del prodotto iniziale e quali potessero essere le necessità evolutive.

Successivamente mi sono informata per comprendere chi fossero gli utenti del mio prodotto e capire meglio le loro necessità e aspettative. Le personas hanno rivelato informazioni cruciali:

- necessità di accesso rapido alle informazioni: gli utenti hanno bisogno di trovare le informazioni immediatamente e in modo chiaro, senza dover navigare attraverso complicati menu.
- preferenza per la semplicità: il target desidera un’interfaccia pulita e semplice, priva di elementi superflui che potessero distrarre o confondere.
- efficacia ed efficienza: gli utenti sono persone spesso di fretta e non vogliono perdere tempo. Era essenziale che l’interfaccia consentisse di completare le operazioni rapidamente e senza intoppi.

Con queste informazioni ridisegnato l’interfaccia con un focus sulla chiarezza e l’usabilità, avvalendomi dell’utilizzo di mockup e wireframe. Per usabilità s’intende il grado con cui un prodotto può essere usato da determinati utenti per raggiungere

determinati obiettivi con efficacia, efficienza, soddisfazione in un determinato contesto d'uso. Per efficacia s'intende il grado di correttezza, accuratezza, completezza in cui l'utente raggiunge i suoi obiettivi; per efficienza s'intende il rapporto tra la capacità di raggiungere gli obiettivi e la quantità di risorse impiegate per raggiungerlo; per soddisfazione s'intende rendere piacevole l'esperienza. Infine ho condotto test di usabilità per raccogliere feedback dagli utenti finali, assicurandomi che ogni modifica migliorasse realmente l'esperienza utente e rispettasse i requisiti richiesti.

5.2 Design Centrato sull'Attività

Il Design Centrato sull'Attività (Activity-Centered Design) mi ha permesso di concentrarmi sulle specifiche attività che gli utenti devono svolgere per raggiungere i loro obiettivi. Questo approccio ha integrato e completato il Design Centrato sull'Utente, fornendo una visione più operativa e pratica delle esigenze degli utenti.

A partire dal POC esistente ho compreso quali sono le attività svolte dagli utenti, identificando i compiti principali e i flussi di lavoro tipici. Questo processo ha incluso l'osservazione diretta degli utenti mentre interagivano con il prodotto e la mappatura dei loro percorsi operativi.

5.3 Il progetto in partenza

Inizialmente, l'applicazione era composta da due sole pagine: una per il wizard in cui l'utente compilava il form di dettaglio del paziente, e una seconda per la visualizzazione dei risultati, dove i grafici erano disposti in una griglia che occupava tutto lo spazio disponibile. Una delle principali carenze del POC iniziale era la mancanza di contesto nei grafici. Questo rendeva difficile la comprensione per utenti senza un background in data science, poiché non potevano interpretare correttamente i dati senza una spiegazione preliminare (5.1).

Un ulteriore problema riscontrato riguardava la lunghezza del wizard, che inizialmente consisteva di quattro step. Successivamente, un passo è stato eliminato e accorpato al primo per migliorare l'usabilità.

5.4 Modifiche apportate

Nella transizione da POC a software completo, il numero di pagine è aumentato: una per il wizard, una per i risultati, una per la valutazione dell'utente circa la piattaforma, e una per il tutorial. Il tutorial è stato posizionato nell'header per garantire un facile accesso da parte dell'utente in qualsiasi momento. Inoltre, il wizard è stato semplificato, eliminando uno step.

La pagina dei risultati ha subito le modifiche più significative. È stato inserito un tabbar che permette all'utente di cambiare le visualizzazioni. L'ordine del tabbar è stato progettato in modo tale che l'utente possa vedere prima le singole visualizzazioni e poi tutte insieme in una vista complessiva (5.2). Quando si clicca sul tab "All graphs", appare un modale bloccante, che l'utente non può chiudere senza prima compilare il form contenuto al suo interno. Questo form deve essere compilato solo una volta ed è cruciale per raccogliere dati sull'utilità percepita dal medico quando visualizza i grafici. È possibile accedere a un altro form tramite il pulsante "Evaluate platform" posizionato

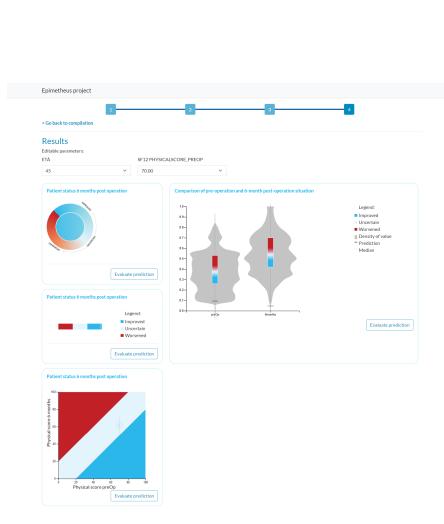


Figura 5.1: Pagina di risultati pre-modifiche

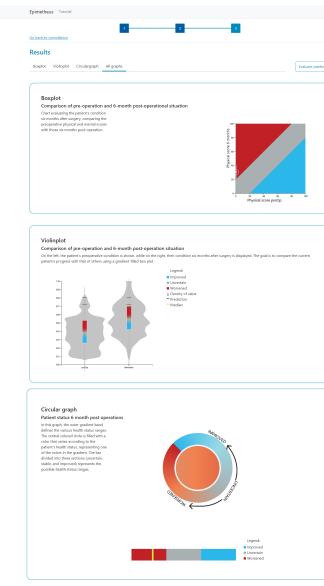


Figura 5.2: Pagina di risultati post-modifiche

in linea con il tabbar. Questo form raccoglie feedback sull'utilità percepita del software nel suo complesso.

I grafici sono interattivi, permettendo all'utente di modificare parametri specifici, come il valore del peso, con conseguente ricalcolo del grafico (5.3).

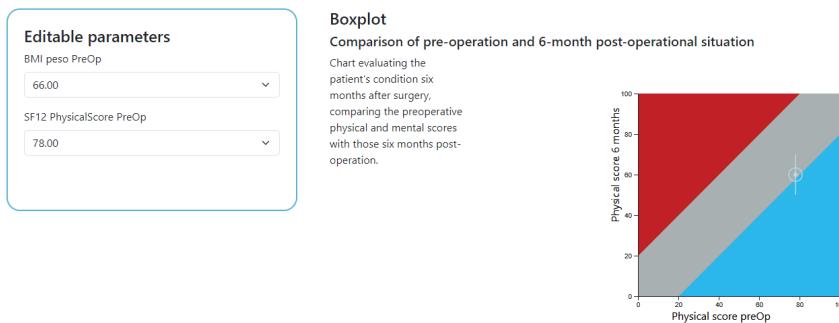


Figura 5.3: Esempio di parametri modificabili nel caso del physical score

Per quanto riguarda il Visual Design del prodotto ho seguito lo stile iniziale, mantenendo come colore primario il celeste, e uno stile complessivamente fresco e pulito grazie all'adozione del Material Design. Ho usato strategicamente le common region per dividere i vari elementi, soprattutto nella pagina di risultati dove sono visibili tutti i grafici. Sono state inoltre apportate modifiche ai colori dei grafici: inizialmente, il colore dell'incertezza era il celeste, ma poiché questo è il colore primario del brand

identity del prodotto, ho proposto di sostituirlo con il grigio. Il grigio, per definizione, si presta meglio a rappresentare il dubbio e l'incertezza. A supporto della mia tesi, test condotti su un campione demograficamente variegato hanno confermato che il grigio è percepito come colore "incerto" (5.8).

Non è mancata l'attenzione verso l'accessibilità del prodotto: il nuovo gradiente è stato testato su tutti i livelli di daltonismo rivelando una buona lettura per quasi tutti i livelli di daltonismo (5.4).

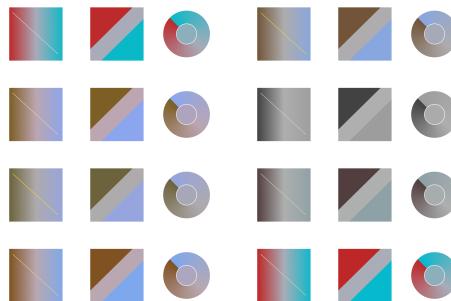


Figura 5.4: Esempio di gradiente e grafico presente in Epimetheus testati con tutti i livelli di daltonismo presenti

5.5 Test

Sono stati condotti due tipi di test, un questionario volto a fornire dati quantitativi e qualitativi, e dei think aloud volti a fornire dati qualitativi.

5.5.1 Questionario

Il questionario è stato somministrato a 45 utenti di diversa età ed estrazione sociale (Figura 5.5). Tra le prime domande che sono state poste vi è la domanda se sono soliti ad consultare grafici o infografiche nell'ambito del loro lavoro o per interesse personale; circa il 42% di loro ha risposto di sì, il 53% di loro ha risposto di no. Successivamente si è indagato se gli utenti si ritengono esperti in grafici, il 47% di loro ha risposto "abbastanza", il 36% di loro ha risposto "non molto". La maggior parte degli utenti sceglie come fonte da cui visualizzare grafici siti web e libri di testo (Figura 5.6).

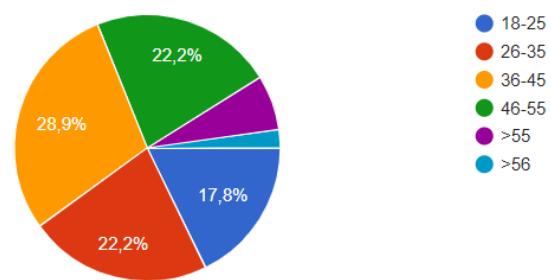


Figura 5.5: Età degli utenti che hanno risposto al questionario

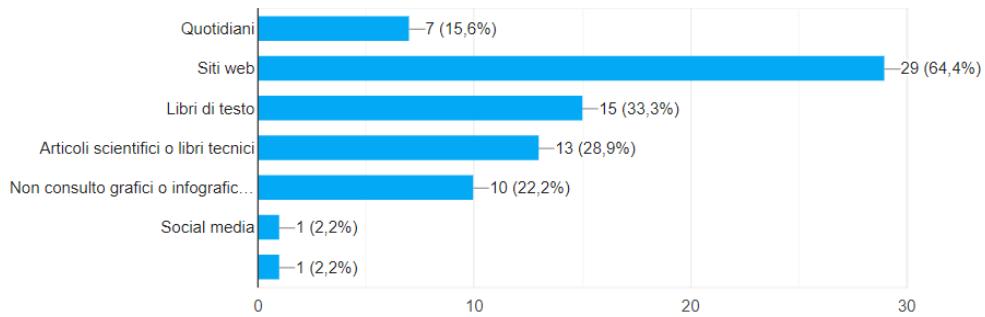


Figura 5.6: Principali fonti da cui gli utenti consultano grafici o infografiche

Successivamente il questionario si è articolato su domande relative ai colori e ai grafici. Qui si è indagato su come gli utenti percepiscono l'incertezza, quindi è stata posta una prima domanda volta a capire se la scelta del grigio invece dell'azzurro comunicasse meglio il concetto di incertezza (Figura 5.7).



Figura 5.7: I due gradienti messi a confronto, con focus posto sull'area centrale che rappresenta l'incertezza. Il primo gradiente è quello inizialmente presente sul progetto, il secondo è quello che si vuole adottare

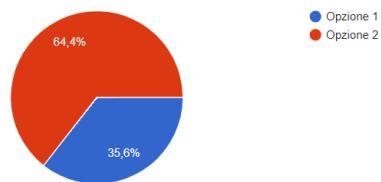


Figura 5.8: Distribuzione delle risposte alla domanda che mette a confronto i due gradienti

Ne è emerso che l'intuizione era corretta, infatti il 64% ha affermato di percepire maggiore incertezza utilizzando il colore grigio invece che azzurro (Figura 5.8).

In un'altra sezione si è indagato sulla percezione dell'incertezza nell'ambito del grafico "Circulargraph". In particolare si voleva capire se effettivamente gli utenti percepissero correttamente lo stato di salute che si voleva comunicare (Figura 5.9). Per condurre questo test sono state realizzate alcune copie del grafico, il colore del cerchio centrale è stato prelevato a partire da vari punti del grafico esterno. Successivamente

all'interno del questionario sono stati disposti i grafici in maniera casuale, in modo tale che l'utente non avesse un bias di ordine.



Figura 5.9: I circulargraph sottoposti a quesito

I risultati sono stati i seguenti:

- Nel grafico 1 il 44% degli intervistati ritiene che il paziente avrà un discreto miglioramento; il 44% ritiene che avrà un netto miglioramento;
- Nel grafico 2 il 77% degli intervistati ritiene che il paziente avrà un discreto peggioramento;
- Nel grafico 3 il 66% degli intervistati ritiene che il paziente avrà un discreto miglioramento, il 22% ritiene che avrà un esito incerto;
- Nel grafico 4 l'84% degli intervistati ritiene che il paziente avrà un esito incerto, non si può dire se migliorerà o peggiorerà;
- Nel grafico 5 il 49% degli intervistati ritiene che il paziente avrà un netto peggioramento, mentre il 38% ritiene che avrà un discreto peggioramento.

Le conclusioni che si possono trarre da questi dati sono che il colore centrale effettivamente comunica correttamente quale sarà lo stato di salute del paziente. Inserire una legenda che permetta di capire quale è il colore del miglioramento e quale è il colore del peggioramento fornisce ulteriori indicazioni all'utente per orientarsi.

Ultima domanda ha riguardato un test A/B, in particolare si è messo a paragone le due pagine di risultati pre (opzione 1) e post (opzione 2) modifiche, e si è chiesto agli utenti quali delle due opzioni preferissero e perché. Il 93% dei partecipanti ha affermato di preferire la nuova interfaccia di risultati (Figura 5.10); i motivi riportati sono stati vari, tra i principali si può citare come motivazione una maggiore pulizia dell'interfaccia, l'interfaccia è meno stancante alla vista, è più ordinata, il commento che accompagna i grafici ne favorisce la comprensione. La parola più frequente è stata "lineare".

Tra questi vi è un commento in particolare che vale la pena citare:

Se l'intento è quello di avere un quadro chiaro e completo del paziente la prima opzione è la migliore perché sono visualizzate più informazioni in poco spazio, dando un quadro generale dello stato di salute; si perde però il dettaglio o delle note che potrebbero risultare comunque interessanti in caso di pazienti/operazioni particolari che invece nella seconda opzione potrebbero essere più facili da individuare.

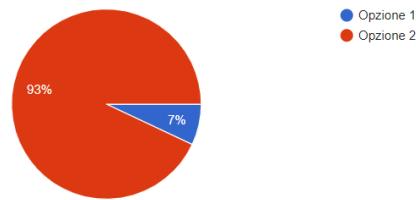


Figura 5.10: Risposte al test A/B che paragona l’interfaccia iniziale con la finale

Questo commento lascia intendere che in futuro si potrebbe aggiungere un bottone che modifichi il layout dell’interfaccia, in modo tale da avere una visualizzazione raggruppata oppure una visualizzazione a lista.

5.5.2 Think aloud

Sono stati condotti dei test di usabilità con utenti appartenenti all’ambito medico. I task richiesti agli intervistati sono stati:

- compilare il form contenente i dati e accedere alla pagina di risultati
- modificare le dimensioni su cui indagare
- accedere alla pagina di valutazione piattaforma
- tornare alla pagina del wizard e compilare un nuovo questionario

Per tutti i partecipanti sono stati condotti due test, uno con score physical (che chiameremo test Alfa) e uno con score mental (che chiameremo test Beta). I dati utilizzati per i partecipanti sono stati gli stessi, in particolare il test Alfa indagava lo stato di salute di un paziente con un physical score pre operation alto e un mental score pre operation discreto, il test Beta indagava lo stato di salute di un paziente con physical score e mental score pre operation molto bassi. Entrambi davano esiti incerti più o meno definiti, quindi si è chiesto ai partecipanti se avrebbero operato o no il paziente. È stata lasciata libera scelta sul tipo di operazione da esplorare.

Partecipante A

Il partecipante A è un medico ortopedico specializzato nelle protesi all’anca e al ginocchio. Durante l’utilizzo del software l’utente si è mosso agilmente all’interno delle pagine. Le funzionalità sono apparse chiare e intuitive, non ha riscontrato difficoltà di sorta nel portare a termine i suoi task.

Dopo un contesto iniziale sulla prima visualizzazione, ha ritenuto il grafico chiaro, è stato in grado di interagire con i parametri modificabili per verificare come lo stato di salute del paziente potesse cambiare nel post operatorio. La seconda visualizzazione (il violinplot) è stata ritenuta molto buona. È stata valutata come comprensibile, chiara, dava precise indicazioni sullo stato di salute del paziente nel post operatorio. La terza visualizzazione (il circulargraph e il boxplot) è stata valutata abbastanza positivamente. Il boxplot è stato ritenuto molto chiaro ed esplicito, il circulargraph invece è stato ritenuto meno ovvio da comprendere.

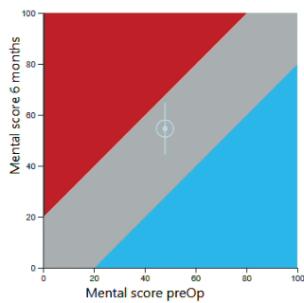


Figura 5.11: Boxplot del test Beta condotto dal partecipante A

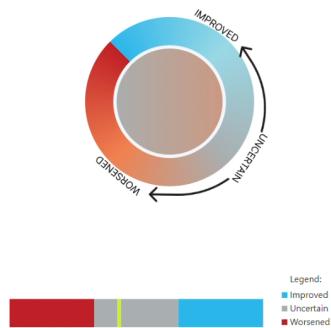


Figura 5.12: Circulargraph del test Beta condotto dal partecipante A

In questo test si è indagato l’ambito dell’operazione all’anca

L’opinione generale nei confronti dell’interfaccia è stata molto positiva, l’utente l’ha definita facilissima da comprendere. Tuttavia, quando si è chiesto di tornare alla pagina iniziale per tentare una nuova indagine, l’utente ha riscontrato difficoltà nell’individuare il link che portasse alla prima pagina.

Al secondo questionario (test alfa) compilato si è andato ad indagare l’opinione del medico nei confronti dei grafici. Nonostante i grafici dessero un esito incerto ((Figura 5.11), (Figura 5.12)) il medico ha affermato che opererebbe il paziente in quanto terrebbe conto dello stato mentale del paziente in relazione al suo dolore fisico. Quindi di fatto la sua opinione non è stata influenzata dal grafico. Secondo il partecipante, infatti, un paziente può avere un mental score negativo perché sta male, ma se a seguito di un’operazione il suo dolore scomparisse allora il suo mental score migliorerebbe. Il partecipante ha riferito che se il fattore predittivo tenesse conto dell’interazione tra la VAS e lo stato mentale del paziente il grafico sarebbe molto più affidabile.

L’utente ha sollevato inoltre che sarebbe opportuno mettere dei target del dolore percepito nel post operatorio, così da sfruttarli nell’ambito delle predizioni. Ha fatto notare inoltre che nell’ambito dello stato mentale del paziente è importante indagare anche su due fattori: il primo è il tempo intercorso tra la prima comparsa del dolore fisico ad oggi, la seconda di come si sente a seguito della lista d’attesa per l’operazione. Infatti pazienti che stanno male e che sono posti di fronte ad una lista d’attesa lunga hanno un peggioramento del loro mental score.

Una lode fatta dal medico riguarda il fatto che tra i parametri modificabili dei grafici figurano parametri concreti su cui effettivamente si può intervenire. Una mia domanda è stata se ritiene che le dimensioni disponibili su cui indagare siano appropriate a formare un’opinione circa lo stato di salute a 6 mesi dall’operazione. La risposta è stata sì, e che varrebbe la pena aggiungere l’HOOS (Hip disability and Osteoarthritis Outcome Score), il KOOS (Knee Disability and Osteoarthritis Outcome Score), e il VAS. Attualmente HOOS e KOOS non esistono nel calcolo predittivo operato dal backend, quindi questo suggerimento si configura come uno sviluppo futuro.

Un’altra domanda che è stata fatta riguarda la modalità di inserimento dei dati: è stato chiesto se l’utente ritiene più facile avere questi valori numerici da copiare nell’apposito input text, oppure se sia meglio presentare il PROM vero e proprio e lasciare che

l'utente lo compili così come è stato compilato dal paziente. L'utente ha risposto che l'attuale modalità di inserimento dei dati è preferibile. Sarebbe ancora più preferibile avere accesso al registro dei dati che ha a disposizione l'istituto Galeazzi.

Partecipante B

Il partecipante B è un tecnico audioprotesista; sebbene non sia un medico ortopedico, questo utente si è reso interessante perché nel suo lavoro affronta tematiche molto simili agli utenti target di Epimetheus, adottando anche metodologie di lavoro simili. Dopo aver spiegato il contesto di utilizzo del software, a partire dall'inserimento dei dati il partecipante è riuscito a portare a termine il wizard senza difficoltà, approdando nella pagina dei risultati. Qui ha interagito con i vari grafici osservandone i cambiamenti ove presenti. Il partecipante ha trovato molto utile la spiegazione contestuale a ciascun grafico, in quanto, soprattutto al primo utilizzo, l'utente inesperto necessita di essere inserito nel contesto applicativo. Ha commentato che i dati sono chiari, le previsioni a 6 mesi dall'operazione si rendono utili a capire come effettivamente cambierà lo stato di salute del paziente anche in relazione al cambiamento dei parametri modificabili.

Quando è stato chiesto di tornare alla pagina iniziale per condurre un altro test, l'utente ha individuato subito il link per tornare al wizard. Nell'eseguire il nuovo test si è notato come il partecipante aveva già assunto confidenza col prodotto, andando spontaneamente a modificare i parametri necessari per indagare come potesse cambiare lo stato di salute del paziente. Nel secondo test ho chiesto al partecipante se opererebbe il paziente. La risposta è stata di sì, perché se vi è una variazione del peso del paziente il grafico si sposta verso un miglioramento (Figura 5.13).

Il partecipante ha commentato che l'interfaccia è molto rapida da utilizzare e comprendere, non ci sono fronzoli, le informazioni sono tutte disponibili.

Dopo il test sul prodotto, il partecipante ha spiegato vari aspetti del suo lavoro. Come tecnico audioprotesista il suo lavoro prevede l'utilizzo di questionari, chiamati COSI (Client Oriented Scale of Improvement), attraverso i quali si va ad indagare quanto il paziente sente, qual è la sua percezione del rumore, se è in grado di distinguere determinati suoni in determinate situazioni, quanto ha disagio in situazioni rumorose (come in mezzo alla folla). Questo test viene effettuato due volte, una volta nel preprotesizzazione, una seconda nel postprotesizzazione, e viene associato a dei test strumentali che forniscono dati numerici. In questa intervista il partecipante ha spiegato che trova difficoltà a comunicare al paziente informazioni concrete e comprensibili circa il suo stato di salute, perché gli strumenti a disposizione sono pochi e forniscono per lo più dati numerici. Il partecipante ha trovato particolarmente comprensibile Epimetheus, poiché ha ritenuto

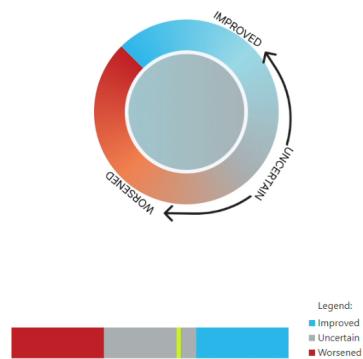


Figura 5.13: Circulargraph del test Alfa condotto dal partecipante B

In questo test si è indagato l'ambito dell'operazione all'anca

che il colore fosse molto più comunicativo rispetto ai numeri. Ha inoltre sottolineato che sarebbe estremamente utile nel suo lavoro avere dei grafici che possano confrontare il pre e post protesizzazione, in quanto avrebbe modo di capire e spiegare al paziente quale sarebbe il margine di miglioramento.

Infine ha affermato che utilizzerebbe il prodotto con il paziente.

Partecipante C

Il partecipante C è uno studente di medicina e chirurgia.

Dopo la spiegazione del contesto iniziale il partecipante si è approcciato con il prodotto in modo naturale.

La prima visualizzazione è stata ritenuta chiara a seguito di spiegazioni circa il suo significato, la descrizione circostante ha aiutato la comprensione. La seconda descrizione è apparsa più comprensibile soprattutto a seguito di interazioni con i parametri modificabili. La terza visualizzazione è stata quella più chiara ed esplicita ed è stata quella che ha permesso più facilmente di prendere decisioni in merito allo stato di salute del paziente (Figura 5.14).

In questo contesto si è chiesto al partecipante se opererebbe il paziente e la risposta è stata che dipende. Secondo l'utente, dal momento che il paziente presenta un mental score basso sarebbe opportuno prima chiedere supporto psicologico per aiutare il paziente a migliorare il proprio stato mentale, e successivamente si potrebbe valutare una operazione al ginocchio. Il partecipante C ritiene molto importante avvalersi del supporto di una equipe medica che possa supportare e accompagnare il paziente nel suo percorso di cura. Di fronte l'indecisione la prima scelta sarebbe chiedere un confronto con colleghi.

Alla domanda se userebbe il software con il paziente ha risposto di sì.

Quando è stato chiesto di tornare alla pagina iniziale per condurre una nuova analisi l'utente è stato in grado di navigare all'interno del prodotto; il nuovo tentativo è stato subito naturale e spontaneo, sia con il wizard sia con i parametri modificabili dei grafici. In definitiva il circulagraph sembra abbia fatto da ago della bilancia nella scelta se operare o no un paziente.

Il partecipante ha commentato l'interfaccia definendola chiara, precisa, facilmente leggibile. Il contesto che contorna i grafici li ha resi più comprensibili anche agli utenti inesperti. Ha affermato inoltre che l'utilizzo del colore nel grafico permette un maggiore impatto visivo.

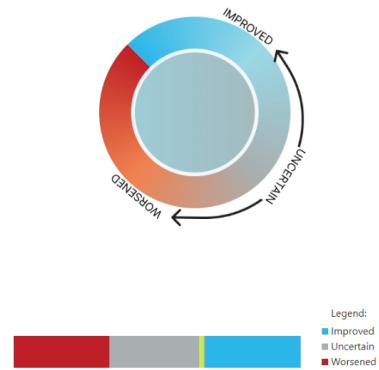


Figura 5.14: Circulargraph del test Alfa condotto dal partecipante C

In questo test si è indagato l'ambito dell'operazione al ginocchio

Capitolo 6

Sviluppo software

6.1 Design dei Sistemi

Nell’ambito dello sviluppo software ho implementato quello che si chiama *Design di Sistemi*, il quale si è reso essenziale per comprendere e gestire le interazioni tra le varie componenti del sistema. Questo approccio mi ha permesso di assicurare che tutte le parti dell’interfaccia lavorassero insieme in modo coerente ed efficiente.

Ho analizzato l’architettura del sistema, identificando tutte le componenti principali e le loro interazioni. Questo includeva il backend, il frontend, e le API utilizzate per la comunicazione tra le diverse parti del sistema.

L’obiettivo era garantire che il frontend fosse perfettamente allineato con le funzionalità offerte dal backend, facendo da filtro iniziale per avere una buona tolleranza all’errore dell’utente ed evitare il rischio di navigare in pagine di errore di difficile gestione o comprensione per l’utente inesperto.

Per implementare il Design dei Sistemi, ho utilizzato un approccio modulare, suddividendo l’interfaccia in componenti riutilizzabili. Ogni componente è stato progettato per essere indipendente e facilmente integrabile con gli altri, seguendo i principi della progettazione a strati. Ho separato il frontend (React) dal backend (Flask), garantendo che ciascuno potesse essere sviluppato, testato e aggiornato indipendentemente. Questa separazione ha facilitato non solo la manutenzione del software, ma in futuro permetterà l’implementazione di nuove funzionalità, garantendo un’evoluzione più rapida del prodotto.

6.2 Il progetto in partenza

Il software nasceva come POC, era un monolite unico dove backend e frontend risiedevano nella stessa directory.

Il backend, scritto in python, si occupava di:

- esporre un endpoint per il frontend
- realizzare calcoli di machine learning per generare le visualizzazioni

Il frontend invece era composto da due file html, uno per il form e uno per la visualizzazione dei grafici, due file javascript, uno per la validazione del form e uno per la realizzazione dei grafici, e un file css che dava lo stile a tutto il sito.

6.3 Analisi e motivazioni

Dopo una prima analisi del software e un confronto con gli stakeholder ho proposto di operare una divisione più netta di frontend e backend avvalendomi dell'ausilio di un framework per il frontend. Attraverso questa scelta si è seguito il principio di progettazione architettonica noto come *Architettura a Strati* o *Architettura a Livelli* (Layered Architecture).

Nell'architettura a strati, l'applicazione è suddivisa in diversi strati logici o livelli, ognuno dei quali ha una responsabilità specifica. Nel caso di Epimetheus:

1. Backend (Strato logico):
 - Responsabilità: Gestisce la logica di presentazione, in particolare la generazione delle risposte API per il frontend.
 - Tecnologie utilizzate: Flask, Python, API REST.
2. Frontend (Strato di Presentazione):
 - Responsabilità: Interagisce con l'utente e invia richieste API al backend per ottenere e visualizzare i dati.
 - Tecnologie utilizzate: React, JavaScript, HTML, CSS.
3. Strato di Persistenza: gestisce il salvataggio e il recupero dei dati.
 - Gestione dei dati nel frontend: quando l'utente inserisce i dati, questi vengono salvati nel local storage del browser, un tipo di storage lato client che permette di mantenere i dati localmente e persistenti anche dopo la chiusura del browser.
 - Gestione dei dati nel backend: i dati raccolti dai form di valutazione del prodotto vengono salvati in un file CSV locale. Questo approccio, noto come persistenza su file locale, permette di conservare i dati di valutazione per eventuali analisi future senza dover implementare subito un database più complesso

Questa separazione ha permesso di ottenere la divisione delle responsabilità, facilitando la manutenzione e l'aggiornamento dei singoli componenti senza influenzare gli altri. La modularità introdotta permette inoltre di scalare l'applicazione con maggiore facilità, ad esempio nel caso in futuro si volesse applicare moduli di registrazione e login, quindi avere cura della sicurezza, oppure ampliare il numero di grafici da mostrare all'utente. Il mio focus era principalmente la possibilità di permettere in futuro di ampliare il software operando un limitato numero di modifiche, quindi era importante portare fin da ora il frontend ad una discreta maturità.

6.4 Sviluppi

Gli sviluppi sono proseguiti in parallelo tra me e i colleghi che ho coordinato. Io mi sono occupata principalmente del frontend, sono partita dall'interfaccia del wizard

mentre ho indirizzato e supportato i colleghi nella realizzazione di endpoint, correzione del backend e creazione del salvataggio dei questionari compilati dagli utenti sulla piattaforma.

Dapprima sono stati realizzati endpoint separati per la validazione e il submit del form relativo ai dati del paziente. Successivamente, su mia indicazione, le logiche sono state accorpate in un unico endpoint che rispondesse con la visualizzazione oppure elencasse i campi non validi. La validazione è stata implementata sia lato frontend che lato backend per garantire maggiore robustezza.

6.4.1 React

Il frontend è stato sviluppato in React, le principali librerie usate sono state:

- **Bootstrap**: per lo stile e alcuni componenti preconfezionati.
- **Axios**: per le chiamate API.
- **Toastify**: per mostrare notifiche (toast) in fase di validazione del form del paziente.
- **D3.js**: per la creazione dei grafici a partire dai dati pervenuti dal backend

Queste librerie sono state installate attraverso il package manager npm, facilitando la gestione delle dipendenze del progetto. Il vantaggio di npm è l'accesso a migliaia di pacchetti open source, rendendo più semplice l'aggiornamento e la gestione delle librerie utilizzate, e la messa a disposizione di script di build per effettuare il build e deploy del progetto.

6.4.2 Flask

Flask è un microframework per Python, è stato scelto per la sua semplicità e flessibilità in quanto consente di costruire applicazioni web leggere e modulari. Ho operato modifiche anche al backend, in particolare nella risposta inviata al frontend per calcolare le visualizzazioni. In origine, tra i campi modificabili vi era l'età, ma a seguito di analisi si è concluso che non era una dimensione di interesse per il medico-utente, pertanto è stata sostituita con il peso.

Ho inoltre predisposto la POST degli endpoint che ricevono i form di valutazione dell'utente, in modo tale che il collega maggiormente incentrato sul backend potesse successivamente sviluppare le feature per il salvataggio a CSV del form compilato.

La struttura del backend è così composta:

- **flaskapp.py**: è il punto d'ingresso principale per lanciare il progetto e contiene le rotte e le logiche di gestione delle richieste.
- **utils.py**: è il file dove sono presenti tutte le funzioni che si occupano di realizzare le visualizzazioni da inviare al frontend.

6.4.3 Validazione

Il software opera una doppia validazione:

- una lato frontend quando l’utente compila i campi, in modo tale da avere subito feedback se qualcosa non è corretto; questa viene fatta tramite chiamata a funzioni di validazioni implementate inizialmente dalla collega e successivamente evolute da me affinchè funzionassero in ambiente React.
- una lato backend, il quale controlla il form pervenuto e verifica se sia tutto conforme.

Se lato frontend non dovessero esserci segnalazioni di errori, ma al contrario il backend dovesse rispondere con un errore, questo viene notificato al frontend e l’utente può vedere a video una pagina di errore con la descrizione di cosa è andato storto. Nel caso in cui il backend dovesse rispondere errore 500 l’utente vedrà una pagina di errore con descrizione generica.

6.4.4 La struttura

Lato frontend le pagine principali sono le seguenti:

- **Wizard:** è presente il wizard iniziale dove l’utente compila i dati del paziente; si compone di 2 step: il primo è la scelta della zona da operare, il secondo è l’effettivo inserimento manuale dei dati del paziente.
- **Results:** è la pagina di risultati dove l’utente visualizza i grafici e interagisce con gli stessi.
- **Evaluation:** è la pagina di valutazione dell’applicativo, dove è presente un form che l’utente può compilare per darci il suo feedback.
- **Tutorial:** è una pagina accessibile dall’ header che descrive l’applicativo e funge appunto da tutorial per l’utente.

All’interno di queste pagine è possibile trovare diversi componenti volti a suddividere ulteriormente la complessità delle pagine e delle logiche.

Lato frontend mi sono avvalsa inoltre di alcuni pattern, tra i quali lo *State Management Patterns* utilizzando il *global context*, ovvero un contesto globale che permette di condividere informazioni tra i diversi componenti senza dover sottostare alla relazione parent-child. All’interno del global context possiamo trovare la gestione degli step, gran parte del form iniziale compilato dall’utente in quanto molte informazioni sono necessarie in più pagine e componenti, i dati fetchati in quanto servono alla pagina Results per calcolare le visualizzazioni. Il global context si occupa anche di controllare anche se parte di questi dati sono già presenti nel local storage del browser, quindi di inizializzare le rispettive strutture dati.

Il fetch delle API relative al submit del wizard avviene tramite custom hooks; questa scelta è stata fatta perché in base a quale operazione l’utente sceglie di esplorare, vengono chiamati endpoint differenti. Il custom hooks ha permesso quindi di articolare queste logiche più strutturate su un file a parte dividendo opportunamente le responsabilità dei vari componenti. Utilizzare custom hooks assieme al global context ha permesso inoltre di semplificare la gestione delle informazioni condivise all’interno dell’applicativo.

La costruzione del wizard compilato dall'utente avviene in maniera dinamica, questo vuol dire che in base alla prima preferenza espressa dall'utente (spine o hip/knee) i campi da compilare cambieranno. Questa scelta è stata fatta per più motivi: inizialmente le operazioni da esplorare erano tre, ovvero hip, knee, spine; in base alla scelta dell'utente i campi del form erano effettivamente diversi, ed era ridondante utilizzare componenti react differenti in base al form necessario. Per implementare questa feature mi sono avvalsa dell'uso di un JSON di dati, dove viene inserita la lista di campi che devono comparire, le loro caratteristiche, ovvero con che nome devono essere inviati a backend e che tipo sono (number, text, select, radio), e quale sarà la funzione che si occuperà della validazione. Questa implementazione permette di poter ampliare il prodotto aggiungendo altre preferenze in modo facile e rapido, limitando quindi le modifiche da apportare ad due soli file per il frontend (il JSON con la lista dei campi e la lista di funzioni di validazione).

```
const formDataManually = {
  Default: [
    {
      type: "radio",
      label: "Gender",
      name: "sesto",
      options: ["Male", "Female"],
    },
    {
      type: "number",
      label: "Age patient",
      name: "anni_ricovero",
      validation: "validateAge",
    },
    {
      type: "select",
      label: "ASA Class",
      name: "classe_asa",
      options: ["select ASA class", "1", "2", "3", "4"],
    },
  ],
}
```

Figura 6.1: Natura dell'incertezza

Capitolo 7

Metodologia

7.1 Metodologia agile e EventStorming

Gli sviluppi si sono svolti seguendo la metodologia agile¹: sono state schedulate due riunioni mensili allo scopo di tenersi aggiornati sugli sviluppi in corso, in assenza di riunioni vi è stato scambio di messaggi per mantenere presente e continua la comunicazione.

Il principio che si è seguito è quello dello sviluppo iterativo, ovvero tenendo presente l'obiettivo finale del software si è proceduto per step implementando prima soluzioni più semplici ed immediate e poi si è proseguito nell'evoluzione di alcune funzionalità per implementare soluzioni più performanti e scalabili.

Le comunicazioni all'interno del team si sono basate su riunioni online, di persona e condivisione di materiale informativo per accordarci sul lavoro da svolgere. Il materiale usato si è basato sulla condivisione di documenti drive e di file Figma.

Parte integrante dell'attività di comunicazione è stata l'uso dell' EventStorming².

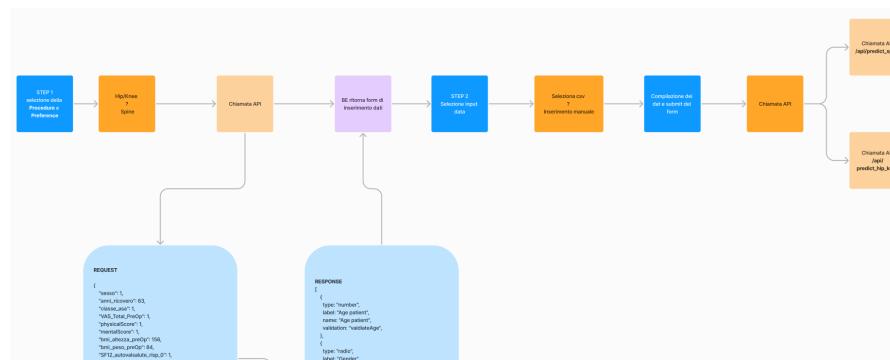


Figura 7.1: Esempio di event storming utilizzato per Epimetheus

L'EventStorming è una tecnica collaborativa utilizzata principalmente per esplorare e modellare processi complessi e sistemi software attraverso la narrazione di eventi che

¹Manifesto agile. URL: <https://agilemanifesto.org/iso/it/manifesto>.

²Alberto Brandolini. *Introducing EventStorming*. Scienza e idee, 2021.

accadono all'interno di un sistema. È stata sviluppata da Alberto Brandolini nel 2013 ed è particolarmente utile nel contesto del [Domain-Driven Design](#) (DDD). L'obiettivo principale di EventStorming è di ottenere una comprensione condivisa del dominio tra tutti i partecipanti, che può includere sviluppatori, stakeholder, esperti di dominio e altri membri del team.

La tecnica si basa su diverse componenti e fasi. Gli eventi rappresentano i cambiamenti di stato significativi che accadono all'interno del sistema. Vengono descritti in modo semplice, solitamente come frasi al passato, e organizzati cronologicamente lungo una linea temporale. Questo aiuta a visualizzare il flusso dei processi e a identificare eventuali lacune o incongruenze. Gli attori, che sono le persone o i sistemi che causano gli eventi, vengono collegati agli eventi che generano. I comandi sono azioni o intenzioni che provocano eventi e vengono posizionati prima degli eventi che innescano. Gli aggregati sono cluster di eventi e comandi che rappresentano unità logiche all'interno del dominio, identificati in una fase successiva per semplificare e strutturare il modello. Infine, le politiche e le regole di business influenzano come e quando gli eventi accadono. Esistono diverse varianti di EventStorming, ognuna con un focus specifico. Il Big Picture EventStorming viene utilizzato per avere una visione d'insieme del dominio, coinvolgendo un gran numero di partecipanti per identificare tutti gli eventi rilevanti, i punti di interesse e le problematiche principali. Il Process Modeling EventStorming si concentra su specifici processi o flussi all'interno del dominio, modellando con precisione come funzionano. Il Design-Level EventStorming è impiegato per progettare soluzioni tecniche specifiche, coinvolgendo principalmente sviluppatori e concentrandosi su come implementare gli eventi e i comandi nel codice.

L'EventStorming offre numerosi benefici. Favorisce la collaborazione tra tutti i partecipanti, aiutando a costruire una comprensione comune del dominio, permette di visualizzare facilmente le aree del sistema che potrebbero essere poco chiare o incomplete, coinvolgendo direttamente gli esperti di dominio per garantire che il sistema rifletta accuratamente la realtà del business, inoltre, la natura visiva e interattiva dell'EventStorming lo rende un metodo rapido e flessibile per esplorare e modellare domini complessi rispetto ai metodi tradizionali di documentazione.

La metodologia nello sviluppo software ha visto l'utilizzo di un approccio *top-down*, ovvero sono partiti prima da una visione d'insieme delle interazioni tra le parti, e successivamente si sono concentrati sui dettagli. Questo approccio implica partire dal livello più alto di astrazione per poi suddividere il sistema in componenti sempre più piccoli e specifici.

Le caratteristiche dell'approccio top-down prevedono una definizione chiara degli obiettivi e delle funzionalità globali del sistema, dove viene creata una visione d'insieme che descrive il sistema nel suo complesso, includendo le sue principali componenti e interazioni. Successivamente vi è la decomposizione, ovvero il sistema viene suddiviso in sottoinsiemi più piccoli dove ogni modulo rappresenta una parte specifica della funzionalità complessiva del sistema. Ogni modulo viene quindi progressivamente raffinato, accompagnandolo da ulteriore documentazione e/o pianificazione.

I vantaggi dell'approccio top-down sono chiarezza e coerenza, migliore gestione della complessità in quanto suddividere il sistema in componenti più piccole permette una migliore gestione del progetto e della sua complessità, ed infine maggiore facilità di manutenzione. Lo svantaggio è che richiede maggiori tempi di sviluppo in quanto parte del tempo è impiegato per la progettazione, e se dovesse esserci un errore nella visione globale questo si ripercuoterebbe su tutto il progetto. In questo caso avere costanti interazioni con i vari soggetti coinvolti nello sviluppo di Epimetheus ha permesso di

limitare gli errori di visione.

Per quanto riguarda la collaborazione tra sviluppatori ci siamo avvalsi dell'uso di **Git**, un sistema di controllo di versione che ci ha permesso di tracciare le modifiche e di lavorare in contemporanea sul progetto. La separazione del frontend dal backend è stata operata su un branch separato rispetto a quello iniziale, mentre l'utilizzo di questo sistema di versionamento ci ha permesso di tornare indietro quando necessario.

Capitolo 8

Sviluppi futuri

Poichè che l'approccio adottato è di tipo iterativo, sono state discusse delle linee guida per le possibili iterazioni future. Queste iterazioni mirano a migliorare ulteriormente l'efficacia e l'efficienza del sistema, sia dal lato backend che dal lato frontend.

Un importante passo avanti nel backend sarebbe la transizione dal salvataggio dei dati su file CSV a un sistema di database. L'attuale implementazione con file CSV è stata scelta per la sua rapidità di sviluppo e la sua capacità di fornire risultati immediati e sicuri. I file CSV, infatti, permettono di tenere traccia dei dati in modo semplice e immediato. Tuttavia, l'adozione di un database porterebbe a una gestione più strutturata e scalabile dei dati, facilitando operazioni di ricerca, aggiornamento e analisi.

Inoltre, tenere traccia della versione del software attualmente in uso consentirà di confrontare, in futuro, come i pareri dei medici riguardo alla piattaforma evolveranno in base agli avanzamenti del prodotto. Questo monitoraggio sarà essenziale per valutare l'impatto delle modifiche e per guidare lo sviluppo di nuove funzionalità.

Un altro sviluppo futuro riguarda il ripristino delle visualizzazioni relative alla spina dorsale. Attualmente, questo aspetto è limitato dall'obsolescenza delle librerie utilizzate, che sono state deprecated. Sarà necessario identificare nuove librerie che possano sostituirle, mantenendo o migliorando le funzionalità precedenti.

Un'evoluzione significativa è l'implementazione di un sistema di login per gli utenti. Questo permetterebbe di salvare i dati della sessione utente, migliorando l'esperienza e la personalizzazione del servizio. La gestione delle sessioni utente rappresenta un passo fondamentale verso un sistema più sicuro e strutturato.

Una richiesta iniziale dello stakeholder principale era quella di spostare maggiormente le logiche del wizard lato server. In particolare, è stato suggerito di ottenere il form relativo al paziente tramite una POST, dove il frontend comunica al backend le preferenze impostate dall'utente, e il backend risponde con l'intero form compilare. Sebbene questa specifica non sia stata implementata per motivi di tempo, la modularità del frontend rende questa integrazione facilmente realizzabile in futuro.

Un altro miglioramento auspicabile riguarda l'ampliamento delle visualizzazioni disponibili. Aumentare il numero e la varietà di visualizzazioni permetterebbe di avere un quadro sempre più chiaro dello stato di salute del paziente a seguito delle varie

operazioni, migliorando la capacità dei medici di prendere decisioni informate.

Un’ulteriore ottimizzazione potrebbe essere l’introduzione di una doppia visualizzazione nella pagina dei risultati: una a griglia, simile a quella iniziale, e l’attuale a lista. La visualizzazione a griglia è più compatta e permette di vedere tutti i grafici con un colpo d’occhio, migliorando l’efficienza dell’interfaccia.

Molto interessante è stato il test condotto con il partecipante B ([5.5.2](#)) in quanto è emerso l’interesse per questo prodotto anche nell’ambito delle protesi audiometriche. Una implementazione di questo tipo richiederebbe sicuramente una base di machine learning differente per il backend, lato frontend invece sarebbe sufficiente inserire una pagina iniziale dove si potrebbe differenziare il contesto di utilizzo, e mantenere all’interno dell’applicativo la stessa struttura dell’interfaccia.

Capitolo 9

Conclusioni

Il risultato di tale tesi è stato la realizzazione di un prodotto scalabile che si presterà ad essere modulare, espandibile e strutturato. È stato molto formativo per me avere questa opportunità come Project Manager in quanto ho avuto modo di mettere in pratica tutte le nozioni apprese nel mondo del lavoro. Grazie ad Epimetheus ho potuto affinare le mie capacità di progettazione, ho avuto modo di testare e migliorare le mie abilità in ambito di design. I test di usabilità sono stati molto importanti non solo perché hanno confermato le ipotesi preliminari ma anche perché hanno svelato nuovi ambiti di interesse per un prodotto come Epimetheus.

Sebbene l'utilizzo del colore per comunicare incertezza sia poco diffuso i test hanno dimostrato come questo tipo di grafici si presta molto bene ad essere compreso sia da parte di utenti esperti che inesperti. C'è quindi terreno fertile per portare avanti la ricerca nell'ambito di questo tipo di visualizzazione vaghe.

Appendice A

Appendice

A.1 Guida nell’analisi dell’incertezza secondo l’EFSA

Il comitato scientifico dell’Autorità europea per la sicurezza alimentare (EFSA) ha sviluppato una guida dettagliata¹ sull’analisi dell’incertezza nelle valutazioni scientifiche, con l’obiettivo di fornire una base solida per il processo decisionale in situazioni di incertezza.

Tale comitato è composto da eminenti scienziati di tutta Europa, ciascuno con una comprovata eccellenza scientifica e una vasta esperienza in varie discipline rilevanti per il mandato dell’EFSA. Tra le loro competenze si annoverano la valutazione dei rischi per la salute umana, l’epidemiologia, la microbiologia, la nutrizione umana, la chimica, la biologia e molte altre aree scientifiche. Questo gruppo multidisciplinare lavora per sviluppare metodologie armonizzate di valutazione del rischio e per garantire l’omogeneità dei pareri scientifici.

Le questioni affrontate dal comitato sono spesso di natura multidisciplinare, richiedendo approcci innovativi e integrati per la valutazione del rischio. Tra queste, vi sono l’armonizzazione dell’uso di ipotesi predefinite in assenza di dati reali.

Secondo tale comitato, l’analisi dell’incertezza nelle valutazioni scientifiche varia a seconda della natura della valutazione stessa. È essenziale identificare il tipo di valutazione in corso per poi seguire le linee guida specifiche.

Tipologie di Valutazione:

- Valutazioni standardizzate: Queste seguono procedure prestabilite e dettagliate, spesso utilizzate per prodotti regolamentati come le autorizzazioni di preimmissione sul mercato. Queste procedure richiedono dati da studi specifici e includono criteri e calcoli standardizzati.
- Valutazioni specifiche per il caso in esame: Necessarie quando non esistono procedure standardizzate. I valutatori devono sviluppare un piano di valutazione personalizzato, utilizzando elementi standardizzati dove possibile ma adottando approcci specifici per altre parti.

¹Comitato scientifico dell’EFSA. «Guida sull’analisi dell’incertezza nelle valutazioni scientifiche». In: (2018).

- Elaborazione o revisione di documenti orientativi: Coinvolge la creazione o l'aggiornamento di procedure standardizzate.
- Valutazioni urgenti: Richiedono completamento rapido con risorse limitate e implicano approcci semplificati.

In alcune aree dell'EFSA, una valutazione standardizzata può evidenziare la necessità di un'analisi più approfondita. In questi casi, si passa a una valutazione specifica per il caso, sostituendo elementi standardizzati con approcci su misura.

Le valutazioni possono essere quantitative (stima di una quantità) o qualitative (risposta verbale a una domanda). Anche se una valutazione è qualitativa, le conclusioni devono essere ben definite. L'incertezza può essere espressa quantitativamente tramite probabilità, pertanto, è raccomandato che i valutatori cerchino di quantificare l'incertezza anche quando utilizzano metodi qualitativi, poiché i metodi qualitativi hanno comunque un ruolo importante nell'analisi dell'incertezza.

Quando si analizza l'incertezza in una valutazione standardizzata, è cruciale distinguere tra incertezze standard e non standard. Prima di tutto, bisogna determinare se esistono incertezze non standard. Se non ci sono, basta riportare che la valutazione non presenta incertezze non standard. In caso contrario, bisogna valutare l'impatto di queste incertezze sul risultato della procedura standard. Se l'incertezza non standard è significativa, potrebbe essere necessario passare a una valutazione specifica per il caso in esame.

Per le valutazioni specifiche per il caso, i valutatori devono seguire istruzioni pertinenti e, se necessario, consultare pareri scientifici di supporto. Dopo aver pianificato la valutazione e identificato le incertezze, bisogna decidere se suddividere l'analisi dell'incertezza in parti. Questa scelta dipende dalle domande e quantità rilevanti per la valutazione.

Valutare le incertezze collettivamente è il metodo più semplice, dove tutte le incertezze vengono valutate insieme nell'ambito della valutazione specifica.

Se l'analisi richiede risposte a domande sì/no, i valutatori possono esprimere le incertezze utilizzando probabilità e combinarle tramite calcoli. Questo richiede che ogni parte dell'analisi risponda a domande sì/no e che il ragionamento sia espresso in un modello logico formale. In alternativa, le incertezze possono essere valutate separatamente, sia qualitativamente che quantitativamente, e poi combinate secondo giudizio esperto. Anche se alcune parti vengono valutate qualitativamente, è consigliabile esprimere un giudizio di probabilità per la valutazione complessiva.

In caso di valutazioni scientifiche urgenti, i valutatori devono adottare un approccio rapido che si adatti ai tempi e alle risorse disponibili. Anche in situazioni di emergenza, l'analisi dell'incertezza è essenziale. Pertanto, si consiglia di valutare tutte le incertezze in un unico passaggio. Questo metodo, sebbene meno preciso rispetto ad altri, offre una base ragionevole per una consulenza preliminare, a condizione che l'incertezza aggiuntiva della valutazione semplificata sia chiaramente indicata nelle conclusioni.

Durante la creazione o la revisione di documenti di orientamento contenenti procedure standardizzate, è spesso necessario eseguire un'analisi dell'incertezza. Questa analisi verifica se la procedura proposta è sufficientemente prudenziale e se soddisfa gli obiettivi di gestione per quella specifica classe di prodotti o problemi. Una procedura ben calibrata può essere applicata ripetutamente senza dover valutare le incertezze

standard ogni volta.

Se un'analisi dell'incertezza è richiesta per altri motivi durante lo sviluppo di un documento di orientamento, essa deve essere trattata come una valutazione specifica per il caso in esame. Quando una procedura viene utilizzata in più aree di lavoro dell'EFSA, la sua calibrazione e revisione devono essere svolte congiuntamente da tutti i soggetti coinvolti. Inoltre, se una procedura standardizzata fa parte di un protocollo internazionale, eventuali modifiche devono essere apportate in consultazione con i partner internazionali e la comunità scientifica.

A.1.1 Individuazione delle incertezze

Incertezze standard e non standard

Incertezze standard: Sono gestite da procedure standardizzate o da elementi standardizzati della valutazione. Ad esempio, nella valutazione del rischio chimico, un fattore di default pari a 100 può gestire le incertezze relative alla tossicità. In studi sperimentali, l'incertezza di misurazione è considerata standard se le linee guida sono state seguite senza deviazioni. Le incertezze standard non devono essere riconsiderate in ogni nuova valutazione che segue la procedura standard, poiché sono già state valutate durante la definizione della procedura stessa.

Incertezze non standard: Includono qualsiasi deviazione da una procedura standard o elementi non coperti dalla stessa. Ad esempio, studi che non seguono le linee guida standard o che presentano carenze nella reportistica, oppure l'uso di studi non standard o di "livello superiore". Queste incertezze devono essere valutate caso per caso, poiché non sono coperte dal margine di manovra delle incertezze standard.

Proporzione delle incertezze: In ogni tipo di valutazione possono esserci sia incertezze standard che non standard, ma in proporzioni variabili. Le valutazioni standardizzate tendono ad avere meno incertezze non standard, mentre altre valutazioni ne presentano di più.

A.1.2 Procedure per l'individuazione delle incertezze

Fonti di incertezza: Ogni valutazione deve indicare le fonti di incertezza individuate, preferibilmente sotto forma di elenco o tabella, per trasparenza.

Verifica sistematica: I valutatori devono verificare sistematicamente le incertezze in ogni parte della valutazione, includendo sia gli input (dati, stime, evidenze) che i metodi utilizzati (metodi, calcoli, modelli statistici, ragionamento, giudizio esperto), per ridurre al minimo il rischio di tralasciare incertezze importanti. Nelle valutazioni standardizzate, devono essere individuate solo le incertezze non standard.

Incertezze negli input: Queste vengono individuate durante la valutazione delle evidenze dalla letteratura o dai database esistenti. Se sono disponibili approcci strutturati per la valutazione delle evidenze, dovrebbero essere utilizzati. Devono anche prestare attenzione a eventuali incertezze aggiuntive non elencate.

Incertezze nei metodi: Queste non sono generalmente gestite dagli schemi di valutazione delle evidenze esistenti. I valutatori dovrebbero utilizzare la colonna destra della Tabella 1 per individuare le incertezze nei metodi applicati alla valutazione. Anche qui, devono essere considerati eventuali tipi aggiuntivi di incertezza.

Uso delle tabelle: Le tabelle e gli elenchi devono facilitare l'individuazione delle incertezze, non la loro classificazione. I valutatori non dovrebbero impiegare troppo tempo a far corrispondere le incertezze alle tipologie elencate.

Classificazione delle incertezze: I valutatori devono determinare quali incertezze sono standard e quali non lo sono, poiché questo influenzera come verranno gestite nei passaggi successivi dell'analisi dell'incertezza.

Incertezze associate agli input della valutazione	Incertezze associate alla metodologia della valutazione
1) Ambiguità 2) Accuratezza e precisione delle misure 3) Incertezza relativa al campionamento 4) Dati mancanti negli studi 5) Studi mancanti 6) Ipotesi relative agli input 7) Stime statistiche 8) Incertezza relativa all'estrapolazione (ovvero limiti in termini di validità esterna) 9) Altre incertezze	1) Ambiguità 2) Fattori esclusi 3) Ipotesi distribuzionali 4) Utilizzo di valori fissi 5) Relazione tra le parti della valutazione 6) Evidenze relative alla struttura della valutazione 7) Evidenze relative al processo di gestione delle evidenze tratte dalla letteratura 8) Giudizio esperto 9) Calibrazione o validazione con dati indipendenti 10) Dipendenza tra fonti di incertezza 11) Altre incertezze

Figura A.1: Natura dell'incertezza

Definizione delle priorità delle incertezze

Definire le priorità delle fonti di incertezza è utile in diverse fasi della valutazione. All'inizio, aiuta a selezionare le incertezze più importanti per un'analisi approfondita. Durante la valutazione, individua le aree in cui cercare più dati o coinvolgere ulteriori esperti. Alla fine, identifica potenziali aree di ricerca futura. Le priorità dovrebbero basarsi sul contributo di ciascuna fonte di incertezza all'incertezza complessiva della valutazione, tenendo conto della dimensione dell'incertezza e della sua influenza sul risultato. L'influenza delle diverse incertezze può essere valutata utilizzando giudizi esperti e scale ordinali. Si possono utilizzare metodi come le “tabelle d’incertezza” o l’approccio NUSAP.

Quando si usano modelli quantitativi, l'analisi di sensibilità può valutare l'influenza delle incertezze sugli input del modello. Le scelte riguardanti la struttura del modello possono essere verificate ripetendo la valutazione con alternative.

Per affrontare l'incertezza in modo efficace, è fondamentale che la valutazione sia suddivisa in parti principali, come esposizione e pericoli in una valutazione del rischio chimico, e in parti minori, come singoli parametri o studi. Questa suddivisione consente di eseguire l'analisi dell'incertezza a diversi livelli: valutare tutte le incertezze collettivamente, suddividerle in parti principali e combinarle successivamente, o suddividerle ulteriormente in parti più piccole e combinarle gradualmente. È essenziale combinare le parti dell'analisi per caratterizzare l'incertezza complessiva, definendo in anticipo come queste parti verranno integrate. Utilizzare un diagramma di modello concettuale può migliorare la trasparenza e il rigore del processo.

L'efficienza e l'affidabilità della valutazione dipendono dal livello di suddivisione scelto. Trattare separatamente le parti con incertezze rilevanti è generalmente più

affidabile, mentre valutare tutte le incertezze collettivamente può essere più rapido ma meno preciso. Alcune incertezze minori possono essere considerate in una fase successiva della caratterizzazione complessiva, mentre quelle maggiori dovrebbero essere combinate tramite calcoli per garantire maggiore affidabilità. Quando si utilizzano modelli quantitativi, è utile quantificare l'incertezza per ciascun parametro e considerare anche le incertezze non quantificate, come quelle relative alla struttura del modello.

Per valutare correttamente l'incertezza, è essenziale che le domande e le quantità d'interesse siano chiaramente definite. Qualsiasi ambiguità aggiunge ulteriore incertezza e complica la valutazione. Se una domanda o quantità non è già ben definita, i valutatori devono provvedere a farlo per l'analisi dell'incertezza. Una quantità o domanda è ben definita se i valutatori possono concordare sulla risposta. Per ottenere ciò, si può specificare un esperimento, uno studio o una procedura che determinerebbe la risposta. Ad esempio, una misura ben definita per una quantità d'interesse dovrebbe includere specifiche di tempo, popolazione, luogo e condizioni. Allo stesso modo, la presenza o assenza di condizioni o meccanismi specifici deve essere dettagliata, così come i risultati di uno studio scientifico o un calcolo rilevante per la valutazione. Durante la formulazione delle domande, è importante identificare e sostituire termini ambigui o soggetti a giudizio di gestione del rischio con termini più chiari o numeri. Se i termini di riferimento sono aperti, è necessario che le conclusioni siano riferite a quantità ben definite o contengano affermazioni chiare e ben definite, necessarie per valutare ed esprimere l'incertezza associata.

L'espressione qualitativa dell'incertezza utilizza parole o categorie ordinali, senza quantificare le possibili risposte o probabilità. Questo metodo è utile in varie situazioni, come la definizione delle priorità delle incertezze e la descrizione di singole fonti di incertezza durante l'analisi. Le espressioni qualitative sono utili per definire le priorità, descrivere singole fonti di incertezza come passaggio preliminare per la quantificazione complessiva, comunicare i risultati usando una scala di probabilità approssimata con descrittori qualitativi, descrivere incertezze non incluse nella valutazione quantitativa e per la reportistica quando richiesto da decisori o normative.

Per descrivere singole fonti di incertezza, si raccomanda l'uso di scale ordinali per migliorare coerenza e trasparenza. Queste scale dovrebbero essere parte della pianificazione della valutazione e possono variare in formalità a seconda delle necessità. Le espressioni qualitative dovrebbero essere usate per elaborare un giudizio quantitativo sull'impatto combinato delle incertezze, basandosi su una logica chiara e non su calcoli arbitrari.

A.1.3 Quantificare l'incertezza tramite probabilità

Il Comitato scientifico raccomanda di esprimere quantitativamente l'impatto combinato delle incertezze. Le probabilità, specificate completamente o parzialmente, possono essere derivate da dati o giudizio esperto.

Probabilità: La probabilità varia da 0 a 1, espressa come percentuale da 0% a 100%. Per una domanda sì/no, 0% significa che la risposta è sicuramente negativa, 100% che è sicuramente positiva, e valori intermedi rappresentano vari gradi di certezza.

A.2. INTELLIGENZA ARTIFICIALE IN AMBITO MEDICO: COSA NE PENSANO GLI UTENTI NEL 2024

Distribuzioni di probabilità: L'incertezza relativa a una quantità non variabile può essere espressa tramite una distribuzione di probabilità, mostrando la probabilità relativa di diversi valori. Un'espressione parziale dell'incertezza può indicare la probabilità che un intervallo di valori includa il valore reale.

Uso dei dati: Quando disponibili, i dati dovrebbero essere utilizzati tramite analisi statistica, anche se il giudizio esperto interviene sempre, ad esempio nella scelta del modello statistico. L'incertezza può essere quantificata direttamente o tramite calcoli dopo aver quantificato singole fonti di incertezza.

A.2 Intelligenza artificiale in ambito medico: cosa ne pensano gli utenti nel 2024

Secondo una recente indagine² dell'EngageMinds HUB, il Centro di ricerca dell'Università Cattolica, gli italiani ad oggi segnalano fiducia e timori verso l'uso dell'intelligenza artificiale in ambito medico.

Dal loro studio emerge che 6 italiani su 10 sono favorevoli all'uso dell'Intelligenza artificiale in ambito sanitario, di questi, l'88% la userebbe per semplificare il linguaggio dei referti, l'86% come supporto al medico per effettuare una diagnosi e l'80% come aiuto per stabilire una terapia farmacologica adeguata, mentre quasi 6 italiani su 10 la utilizzerebbero come strumento per un'autoanalisi.

Di opinione meno positiva sono 7 italiani su 10 secondo i quali l'AI potrà causare una perdita della relazione e del contatto diretto con il medico.

Tra le principali opportunità che l'uso delle tecnologie digitali potranno portare, poco meno di 8 italiani su 10 (78%) riferiscono che esse porteranno ad una maggiore accessibilità nell'accesso e nell'uso dei servizi, una riduzione dello spreco di carta e un maggior coinvolgimento del paziente grazie ad una maggiore accessibilità al proprio fascicolo sanitario. Il 74% crede che le AI potranno ridurre i costi a lungo termine; poco meno di 7 su 10 ritengono che possa esserci un miglioramento nei monitoraggi tramite devices (68%), mentre poco più di 6 su 10 si aspetta che le AI possano migliorare le diagnosi (63%). Il 68% degli italiani ritiene che l'uso di tecnologie digitali possano migliorare il monitoraggio da remoto.

Un ulteriore rischio che gli italiani percepiscono è legato ai dati sensibili: per il 63% l'uso dell'Intelligenza Artificiale potrà causare delle problematiche legate alla gestione della privacy, mentre per il 60% legate alla diffusione di dati sensibili.

²EngageMinds Hub e Consumer Food Health Engagement Research Center. «Intelligenza artificiale e medicina». In: (2023).

Glossario

clinical decision rules Clinical decision rules (CDRs) sono strumenti utilizzati dai medici per migliorare e standardizzare il processo decisionale clinico. Queste regole sono derivate da studi di ricerca clinica e utilizzano informazioni cliniche specifiche per fornire raccomandazioni su come gestire un paziente. Le CDRs combinano diverse variabili cliniche per prevedere la probabilità di una determinata malattia o esito e forniscono una guida su quali azioni intraprendere, come ulteriori test diagnostici o trattamenti. Esempi noti di CDRs includono le regole di Ottawa per determinare se un paziente con un trauma alla caviglia necessita di radiografie..
[5](#), [47](#)

clinical prediction rules Clinical prediction rules (CPRs) sono simili alle CDRs, ma sono specificamente focalizzate sulla previsione di esiti clinici. Le CPRs utilizzano una combinazione di segni, sintomi e risultati di test per calcolare la probabilità che un paziente sviluppi un determinato esito clinico, come la probabilità di una diagnosi, il rischio di complicazioni, o la risposta a un trattamento. Le CPRs sono strumenti quantitativi che possono aiutare i medici a stratificare i pazienti in base al rischio e a prendere decisioni più informate. Un esempio è il punteggio Wells per la previsione della probabilità di trombosi venosa profonda.. [5](#), [47](#)

domain driven design Il Domain-Driven Design è un approccio allo sviluppo del software che si concentra sulla comprensione e sulla modellazione del dominio del problema, ovvero il settore di attività a cui il software si applica.. [42](#), [47](#)

POC La Proof of Concept (POC), in italiano "prova di concetto", è una realizzazione preliminare di un'idea o di una teoria per dimostrarne la fattibilità. È utilizzata in diversi campi, tra cui tecnologia, ingegneria, ricerca scientifica e sviluppo di prodotti, per verificare se un concetto o una proposta ha il potenziale per essere sviluppato ulteriormente e diventare un prodotto o un servizio funzionante.. [2](#), [47](#)

point-of-care Point-of-care (POC) si riferisce a qualsiasi servizio medico o test diagnostico effettuato al momento e nel luogo del trattamento del paziente. In altre parole, i test POC sono eseguiti vicino al paziente, spesso durante la visita medica, piuttosto che in un laboratorio remoto. Questo approccio consente di ottenere risultati rapidi e di prendere decisioni terapeutiche immediate. Esempi comuni di test POC includono test di gravidanza, glucometri per il controllo del glucosio nel sangue, e test rapidi per l'influenza o il COVID-19. L'obiettivo del POC è migliorare l'efficienza del trattamento, ridurre i tempi di attesa per i risultati e migliorare la qualità delle cure fornite. [5](#), [47](#)

quartili Una volta ordinati i dati, i quartili sono i tre valori che dividono l'insieme dei dati in quattro intervalli di uguale numerosità. Il secondo quartile coincide con la mediana della distribuzione. Per estensione di significato, si dice quartile anche ognuno dei quattro intervalli così determinati.. [19](#), [48](#)

test del tornasole Scala di Colore: Le cartine tornasole sono dotate di una scala di colori che varia da 1 a 14. Questo permette di determinare con precisione il livello di pH di una sostanza. La scala si suddivide in: 1-3 per indicare forte acidità, 4-6 per acidità, 7 per neutralità, 8-11 per alcalinità e 12-14 per forte alcalinità.. [21](#), [48](#)

Bibliografia

Libri di riferimento

- Brandolini, Alberto. *Introducing EventStorming*. Scienza e idee, 2021 (cit. a p. 35).
- Coen, Daniele. *L'arte della probabilità. Certezze e incertezze della medicina*. Leanpub, 2021 (cit. alle pp. 5–7).
- Giuseppe Parisi, Luca Pasolli. *Clinica, relazione, decisione*. Franco Angeli, 2016 (cit. a p. 4).
- Womack, James P. e Daniel T. Jones. *Lean Thinking, Second Edition*. Simon & Schuster, 2010.

Articoli di riferimento

- al, S. K. Card et. «Understanding the Role of Visualizations on Decision Making: A Study on Working Memory». In: (2020) (cit. a p. 9).
- Alexander A. Kon, MD. «The Shared Decision-Making Continuum». In: (2010) (cit. a p. 7).
- C, Heneghan, Glasziou P e Thompson M et al. «Diagnostic strategies used in primary care». In: (2009) (cit. a p. 7).
- dell'EFSA, Comitato scientifico. «Guida sull'analisi dell'incertezza nelle valutazioni scientifiche». In: (2018) (cit. a p. 41).
- Federico Cabitzza, Andrea Campagner e Enrico Conte. «Comparative Assessment of Two Data Visualizations to Communicate Medical Test Results Online». In: (2022) (cit. alle pp. 14, 18).
- Federico Cabitzza Angela Locoro, Giuseppe Banfi. «Machine Learning in Orthopedics: A Literature Review». In: (2018) (cit. a p. 11).
- G, Gigerenzer. «Fast and frugal heuristics: the tools of bounded rationality. Blackwell handbook of judgment and decision making». In: (2004) (cit. a p. 5).
- Han PK Klein WM, Arora NK. «Varieties of uncertainty in health care: a conceptual taxonomy». In: (2011) (cit. alle pp. 3, 4).
- Hillen, Marij A. «Tolerance of uncertainty: Conceptual analysis, integrative model, and implications for healthcare». In: (2017).

- Hub, EngageMinds e Consumer Food Health Engagement Research Center. «Intelligenza artificiale e medicina». In: (2023) (cit. a p. 46).
- Lucia Craxì, Vito Di Marco. «Come gestire l'incertezza nella pratica clinica?» In: (2018).
- Michela Assale, Silvia Bordogna e Federico Cabitza. «Vague Visualizations to Reduce Quantification Bias in Shared Medical Decision Making». In: (2020) (cit. a p. 12).
- Parisi, Giuseppe e Peter Konstantin Kurotschka. «Dalla medicina basata sulle evidenze alla decisione condivisa con il paziente: come gestire l'incertezza nella pratica clinica». In: 113 (2022) (cit. alle pp. 3–8).

Siti web consultati

- Artificial Intelligence in medicine: applications, implications, and limitations.* URL: <https://www.hunimed.eu/news/artificial-intelligence-in-medicine-applications-implications-and-limitations/> (cit. a p. 10).
- L'utilizzo di PROM e PREM nei patient support program.* URL: <https://phdlifescience.eu/lutilizzo-di-prom-e-prem-nei-patient-support-program/> (cit. a p. 18).
- Manifesto agile.* URL: <https://agilemanifesto.org/iso/it/manifesto> (cit. a p. 35).