

# BANDI iNEST

Vs

Ricercatori

17 Maggio 2024



## Spoke di riferimento: 3

- **Acronimo:** E-CLIP
- **Titolo Progetto:** Piattaforma Evoluta per la gestione doCumentale Lean ed Intelligente del Patrimonio informativo aziendale
- **Key-words:** Intelligenza Artificiale, Machine Learning, Reti neurali profonde, grandi modelli di linguaggio, Analisi semantica del contesto, Analisi dei contenuti, Analisi Documenti, Sistema di gestione dei dati, Lean Intelligence, dati IoT
- **Partecipanti:** Nicola Savino, Marco Tedesco, Salvatore Miseo, Michele La Rocca, Massimo Di Pietro, Andrea Bruno
- **Durata:** 15 mesi
- **Budget totale:** 566.231 €
- **Coordinatore:** Nicola Savino

- **Abstract:**

Il progetto E-CLIP (Enhanced Content Lean Intelligence Platform) è più di un sistema di gestione documentale. E' in grado di raccogliere informazioni da più fonti (mail, minute meeting, dati dalla linea, documenti legali, fiscali etc...) e differenti dipartimenti aziendali (dal marketing, dalle vendite, amministrazione...), analizzarli ed estrarne il contenuto semantico per renderli disponibili in maniera semplice ed intuitiva.

Opera come un collettore per gestire ed accettare, in una forma di plug-in, ulteriori dati da agenti di elaborazione specifici di dominio (ad esempio dati da linee manifatturiere con relativi sensori virtuali, moduli predittivi...) e combinare questi dati con le informazioni presenti nei documenti.

L' AI (DNN e LLM, reti neurali profonde e grandi modelli di linguaggio) renderà semplice la ricerca di contenuti semantici nei documenti e disponibile il valore dell'informazione per poter creare nuovi servizi, ottimizzare la produttività, catturare insights nascosti, essere di supporto alle decisioni con un impatto su tutta la struttura aziendale e sui suoi processi che ne gioveranno in conoscenza ed efficienza operativa.

Il progetto che presentiamo rivoluzionerà le dinamiche lavorative democratizzando la tecnologia e permettendo anche alle realtà più piccole di competere a livelli più alti.

## Stato dell'arte

Le aziende sono oggi sommerse da una mole immensa di dati provenienti da più fonti come ad esempio sensori IoT da linee produttive, analisi mercato, informazioni sulla user experience che guidano le strategie per la value proposition e la pianificazione della produzione. La mole di informazioni cresce in maniera esponenzialmente e rappresenterà un patrimonio e l'asset su cui le aziende baseranno la loro competitività nei rispettivi settori.

L'enorme sforzo tecnico e scientifico, in ambito manifatturiero, è focalizzato sulla creazione di Digital Twin di prodotto, per la simulazione della produzione con l'utilizzo dell'AI per analizzare ed estrarre dati dalla sensoristica IoT delle linee di produzione e produrre insights, ad esempio, per la predictive maintenance, lo zero defect manufacturing, la just in time production, al fine di ottimizzare la qualità e la produttività. Lo stesso processo, che si potrebbe applicare al patrimonio di documentazione aziendale, non riceve la stessa attenzione.

Di fatto c'è tutta una mole di dati e documenti formali che non sono legati direttamente alla "produzione", generati per esigenze legislative o per rispondere al processo, ma sono necessari ed indispensabili per la vita dell'azienda (mail, documenti contabili, contratti, minute meeting...). La dematerializzazione e i sistemi di gestione documentale (DMS), convertendo i documenti in formato elettronico, hanno risolto il problema della gestione della documentazione aziendale efficiente, sicura, economica e conforme alle normative vigenti che comunque necessita di gestione manuale nella catalogazione ed archiviazione.

Negli ultimi anni, abbiamo assistito a una rivoluzione nel campo delle tecnologie di intelligenza artificiale (IA), in particolare nelle reti neurali profonde (DNN) e nei grandi modelli di linguaggio (LLM). Altra frontiera di ricerca, che nel campo dell'AI si sta esplorando, sono i Grandi Modelli di Linguaggio (LLM), come GPT-4 (acronimo di Generative Pretrained Transformer, implementato in CHAT-GPT), che rappresentano il vertice delle capacità attuali in termini di comprensione e generazione del linguaggio naturale e sono basati su architetture di tipo Transformer.

Con il progetto E-CLIP (Enhanced Content Lean Intelligence Platform), intendiamo aprire la strada a nuove metodologie di lavoro per fare evolvere i processi di gestione, archiviazione, consultazione, ricerca e analisi dei documenti e creare un collettore per gestire ed accettare, in una forma di plug-in, ulteriori dati e agenti di elaborazione specifici del dominio (ad esempio dati da linee manifatturiere con relativi sensori virtuali, moduli predittivi), combinandone il valore.

### Obiettivi:

- Miglioramento di produttività ed efficienza del processo di ricerca documentale con il nuovo approccio basato sulla descrizione dei contenuti.
- Tipo e numero di funzionalità correlate all'estrazione del contenuto semantico dei documenti.
- Miglioramento del processo di ricerca semantica mediante descrizione dell'esigenza.
- Verifica dei suggerimenti basati su eventi, identificazione dell'utilizzo, fattori sociali con proposta di informazioni "Right First Time".
- Ottimizzazione dei costi di creazione di documenti di sintesi su richieste dell'utente.

### KPI

- Tempo di Ricerca: Riduzione percentuale del tempo medio impiegato per trovare un documento specifico.
- Accuratezza della Categorizzazione: Percentuale di documenti correttamente categorizzati automaticamente rispetto al totale.
- Tasso di Errori: Riduzione del tasso di errori di richieste che portano a risultati non pertinenti o errati.
- Feedback Utente: Percentuale di feedback positivi sull'interfaccia e sulla facilità d'uso.
- Anticipazione delle Esigenze: Percentuale di suggerimenti automatici accettati o utilizzati dagli utenti.

## Stato dell'arte

Il progetto mira a incorporare tecnologie avanzate di intelligenza artificiale come deep learning e elaborazione del linguaggio naturale per realizzare E-CLIP (Content Lean Intelligence Platform). E-CLIP semplificherà i processi per aumentare l'efficienza nel lavoro quotidiano degli utenti, riducendo i tempi di ricerca e aumentando l'efficienza complessiva.

- **FASE 1 - Definizione dei Requisiti:** Questa fase si concentrerà sulla definizione dei requisiti iniziali del progetto. Saranno condotti studi di fattibilità e analisi per identificare il caso d'uso più adeguato e valutare l'introduzione di nuove opportunità in collaborazione con clienti fidati. Questo step include la raccolta e l'analisi delle esigenze reali basate sul know-how aziendale e sulle aspettative dei clienti, la definizione di requisiti di sicurezza e l'identificazione dei Key Performance Indicators (KPI) attesi.
- **FASE 2 - Studio di modelli di Machine Learning:** Questa fase si concentrerà sulla definizione del problema e l'implementazione di uno o più modelli di machine learning. L'obiettivo principale è stabilire e affinare il processo di classificazione dei documenti aziendali mediante tecniche di machine learning, prevedendo le categorie di appartenenza dei documenti.
- **FASE 3 - Sperimentazione e Selezione di Modelli di Deep Learning NLP e LLM:** Questa fase si concentrerà sullo studio e la sperimentazione e l'addestramento di modelli di Deep Learning, Natural Language Processing (NLP) e LLM (Large Language Models) per raggiungere gli obiettivi di progetto.
- **FASE 3 – Sviluppo e prototipazione:** Questa fase si concentrerà sullo sviluppo sperimentale, la prototipazione di E-CLIP per realizzare un prototipo che funzioni su dati reali attraverso cicli rapidi di sviluppo, test e revisione in un'ottica di sviluppo agile.

L'obiettivo finale di CLIP è quello di creare un sistema d'intelligenza artificiale aziendale che funzioni come un "cervello" distribuito, migliorando l'interazione e la comprensione dei documenti a livello semantico.

## Impatti attesi

Il progetto E-Clip ha effetti significativi e multidimensionali sull'economia, sulla tecnologia e sui processi aziendali. L'impatto economico, tecnologico e di processo è su più livelli:

- **Efficienza Aziendale:** L'intelligenza artificiale (AI) migliora l'efficienza aziendale analizzando e categorizzando i documenti rapidamente. Questo consente un facile accesso alle informazioni, riduce i tempi di ricerca e permette ai dipendenti di dedicarsi a compiti più strategici.
- **Produttività aziendale:** L'AI automatizza la gestione dei documenti, velocizzando l'elaborazione e migliorando la qualità dei dati e diminuendo il rischio operativo.
- **Esperienza del Cliente:** Una gestione documentale efficace migliora le interazioni con i clienti, offrendo risposte più veloci e accurate, aumentando la soddisfazione e la fedeltà del cliente.
- **Filiere Produttive:** Il progetto stimola l'innovazione nelle filiere produttive attraverso l'ammodernamento di macchinari, software avanzati, nuove linee produttive.
- **Innovazione nei Macchinari e Robotica:** Il progetto promuove la crescita nel settore dei macchinari intelligenti e delle interfacce uomo-macchina avanzate, aprendo nuove opportunità per l'industria manifatturiera.

