

The logo for Zendata, featuring the word "zendata" in a white, lowercase, sans-serif font. The letter "o" is replaced by a stylized circular graphic consisting of three concentric circles. The logo is centered on a blue background with a subtle pattern of thin, white, intersecting lines forming a grid-like structure.

# zendata

**Roma, 17/06/2025**

**Oggetto:** Proposta di realizzazione di una piattaforma basata su intelligenza artificiale per automatizzare il processo di gestione delle pratiche di anticipazione del Fondo Pensione BCC.

<b>1. Executive Summary.....</b>	<b>4</b>
1.1 Panoramica.....	4
1.2 Obiettivi del Sistema.....	4
1.3 Funzionamento del Sistema.....	5
<b>2. Architettura.....</b>	<b>7</b>
2.1 Panoramica dell'Architettura.....	7
2.3 Layer di Ingestion.....	7
2.4 Layer di Storage.....	7
2.5 Layer di AI.....	8
2.6 Layer Applicativo.....	8
2.7 Autenticazione e Sicurezza.....	9
<b>3. Struttura Economica.....</b>	<b>11</b>

## Chi siamo

ZenData è una startup dinamica e innovativa, specializzata nello sviluppo di soluzioni basate sull'intelligenza artificiale, per aumentare l'efficienza e stimolare la crescita.

Le nostre soluzioni abbracciano una vasta gamma di settori e reparti, spaziando dallo sviluppo di assistente virtuale per pubbliche amministrazioni e professionisti autonomi, fino all'analisi e all'ottimizzazione dei processi produttivi aziendali.

La nostra strategia è orientata alla progettazione di soluzioni che si integrano in maniera semplice e fluida nei processi operativi dei nostri clienti, garantendo un'adozione senza attriti e un impatto positivo immediato.

Siamo convinti che il nostro approccio all'innovazione, unito alla capacità di personalizzare le soluzioni in base alle specifiche esigenze dei nostri clienti, ci permetta di offrire un valore aggiunto significativo e di contribuire al successo dei nostri partner commerciali.



# 1. Executive Summary

## 1.1 Panoramica

Il presente progetto prevede la realizzazione di una Proof of Concept (PoC) **production ready** di un sistema avanzato basato su tecnologie di intelligenza artificiale, finalizzato all'automazione della pre-analisi delle pratiche di anticipazione del Fondo Pensione BCC. La PoC, della durata di **8 settimane**, ha l'obiettivo di validare la soluzione su dati reali, mantenendo un livello di rischio contenuto. In questa fase iniziale, l'infrastruttura sarà configurata con alcune **semplificazioni nei meccanismi di integrazione con i sistemi BCC**, che verranno poi sviluppati in modo completo e ad hoc nella successiva fase di messa in produzione. In caso di esito positivo, verrà inoltre condotta un'analisi dell'effort necessario per l'industrializzazione della soluzione, con particolare attenzione agli aspetti di evoluzione software, rafforzamento della sicurezza infrastrutturale **e valutazione della portabilità verso ambienti on-premise o private cloud**, qualora richiesto dal contesto operativo del cliente.

## 1.2 Obiettivi del Sistema

Il sistema deve permettere la gestione end to end delle richieste di anticipazione e riscatto:

- Deve permettere il caricamento delle richieste che verranno analizzate in diversi formati (i.e. docx, pdf, markdown, txt).
- Processa la documentazione (parsing documentale) per trasformarla in un formato gestibile dall'intelligenza artificiale.
- Estrae le informazioni rilevanti, categorizza i documenti e ne verifica la conformità alle policy e regolamenti del fondo, in maniera tale da supportare l'operatore nella gestione delle pratiche.
- Supporta, in un'interfaccia web, la gestione delle pratiche processate dall'intelligenza artificiale.

## 1.3 Funzionamento del Sistema

La soluzione proposta si articola in tre componenti principali: **Frontend web per la gestione delle pratiche**, **Backend API per la logica applicativa**, e **moduli AI asincroni per l'analisi documentale**.

Gli operatori avranno a disposizione un'interfaccia web per caricare le richieste direttamente sulla piattaforma (formati supportati: PDF, DOCX, TXT, ecc.). Una volta caricati, i documenti vengono automaticamente elaborati attraverso un workflow asincrono che include:

1. **Parsing e normalizzazione** del contenuto documentale.
2. **Estrazione automatica dei metadati** rilevanti per la pratica:
  - Dati anagrafici (nome, cognome, codice fiscale, indirizzo)
  - Dati patrimoniali (percentuali di proprietà, tipo di proprietà, intestatari)
  - Informazioni finanziarie (importi, date, causali di pagamento, IBAN, ecc.)
  - Tipologia della richiesta e documenti allegati.
3. **Categorizzazione** della pratica sulla base della tipologia di anticipazione (es. spese mediche, ristrutturazione, ulteriori esigenze, ecc.).
4. **Verifica di conformità** rispetto ai regolamenti del fondo:
  - Presenza dei documenti obbligatori previsti dall'Allegato 1 per ogni tipologia
  - Correttezza temporale (es. date dei bonifici entro 6 mesi)
  - Intestazione coerente con la proprietà dell'immobile
5. **Segnalazione automatica di anomalie** o documentazione mancante attraverso un sistema di flag intelligenti.

Attraverso la dashboard, l'operatore può:

- Consultare lo **stato delle pratiche** (nuova, in lavorazione, completata).
- Visualizzare i **metadati estratti automaticamente** e modificarli se necessario.
- Verificare i **flag e le criticità rilevate** dal sistema AI.

Il sistema riproduce le logiche dell'attuale applicativo istruttorio, proponendo un **flusso di domande guidate** che riflette il processo di verifica manuale:

- “L'intestatario della fattura corrisponde all'iscritto?”
- “La fattura è già stata rimborsata in precedenza?”
- “Il bonifico è associato correttamente alla spesa?”

Queste domande vengono **pre-compilate** sulla base dell'analisi automatica dei documenti, facilitando il lavoro dell'operatore.

1. Caricamento documenti sulla piattaforma.
2. Estrazione automatica di metadati e informazioni chiave.
3. Verifica AI della completezza e conformità documentale.
4. Presentazione della scheda pratica con flag ed evidenze.
5. Revisione e validazione da parte dell'operatore.

## 2. Architettura

---

L'infrastruttura sarà ospitata su Microsoft Azure, sfruttando servizi cloud per garantire prestazioni ottimali, sicurezza avanzata e possibilità di evoluzione continua senza impatti sull'operatività.

### 2.1 Panoramica dell'Architettura

L'architettura si compone di quattro layer funzionali:

- **Layer di Ingestion:** Gestisce l'upload e la pre-elaborazione della documentazione da analizzare tramite OCR.
- **Layer di Storage:** Include database e storage per dati strutturati e file.
- **Layer di AI:** Utilizza modelli di intelligenza artificiale per estrazione delle informazioni rilevanti dai documenti in perimetro.
- **Layer Applicativo:** Comprende il frontend per gli utenti e il backend per la logica di business.

### 2.3 Layer di Ingestion

Il layer di ingestion sarà responsabile dell'acquisizione e pre-elaborazione delle schede tecniche:

- **Acquisizione:** Gli utenti caricheranno le pratiche di anticipazione del Fondo Pensione tramite un'interfaccia Azure Static Web App.
- **Pre-elaborazione:**
  - Un Azure Function App utilizzerà **Azure Cognitive Services (OCR API)** per estrarre testo da documenti.
  - I dati estratti saranno salvati in **Azure Cosmos DB for MongoDB**.

### 2.4 Layer di Storage

Il layer di storage si compone di:

- **Azure Cosmos DB for MongoDB:**
  - Database scalabile per gestione utenti e ruoli contestuali alle label e tag estratti dalle pratiche di anticipazione.
  - Schema flessibile NoSQL adatto a possibili cambiamenti sulle tipologie di dati da estrarre dai documenti.
  - Accesso sicuro tramite Private Endpoint nella VNet.
- **Azure Blob Storage:**

- Archiviazione dei documenti caricati sulla dashboard amministrativa

## 2.5 Layer di AI

Questo layer utilizza tecnologie AI per le funzionalità chiave dell'applicazione:

- **Large Language Model (LLM):** engine generativo utilizzato per estrarre le informazioni di interesse dai documenti
- **Function asincrone** per la gestione scalabile del processing documentale
- **Estrazione Dati:** L'obiettivo è realizzare un workflow di AI capace di suddividere il task di labeling dei documenti in differenti step che rendano l'estrazione delle informazioni dai documenti il più accurata possibile.

## 2.6 Layer Applicativo

Il layer applicativo sarà composto da un backend sviluppato in **Python** tramite il framework **FastAPI**, distribuito in modalità container Docker su **Azure App Service**, per garantire isolamento, scalabilità e semplicità nella gestione delle risorse cloud.

Il backend esporrà API REST per due principali funzionalità:

- **Gestione documentale:** caricamento, rimozione e aggiornamento dei documenti caricati dagli utenti;
- **Analisi AI:** accesso ai risultati generati dal motore di intelligenza artificiale sui documenti elaborati, inclusi i dati estratti e le eventuali valutazioni automatizzate.

Il frontend dell'applicazione sarà sviluppato in **Next.js** e distribuito anch'esso su **Azure Static Web App o App Service**, consentendo un'integrazione rapida e flessibile con i sistemi esistenti, oltre a facilitare la fase di validazione e adozione da parte degli utenti finali.

Le comunicazioni tra i layer avverranno attraverso una **Virtual Network** con configurazione di **Private Endpoint**, in modo da garantire la massima sicurezza nelle connessioni tra backend, storage e moduli AI.

Tutti i **segreti applicativi** e le **chiavi di accesso** saranno gestiti tramite **Azure Key Vault**, utilizzando **Managed Identity** per eliminare la necessità di credenziali esplicite e assicurare un livello elevato di protezione.



Infine, il monitoraggio del sistema e la raccolta delle metriche operative saranno affidati a **Azure Application Insights**, abilitando un'osservabilità avanzata e il troubleshooting in tempo reale.

## 2.7 Autenticazione e Sicurezza

Per garantire la protezione dei dati e il controllo degli accessi, il sistema adotterà sistema di autenticazione ed autorizzazione basato su token JWT.

L'uso delle **Managed Identities** permette ai servizi Azure di autenticarsi in modo sicuro senza necessità di gestire manualmente credenziali statiche. In particolare, l'**App Service** utilizzerà una **System-Assigned Managed Identity** per ottenere token da Azure AD e accedere in modo sicuro a **Key Vault**, **Azure Cosmos ed Azure blob storage**, sfruttando il **Role-Based Access Control (RBAC)** e politiche di accesso dedicate.

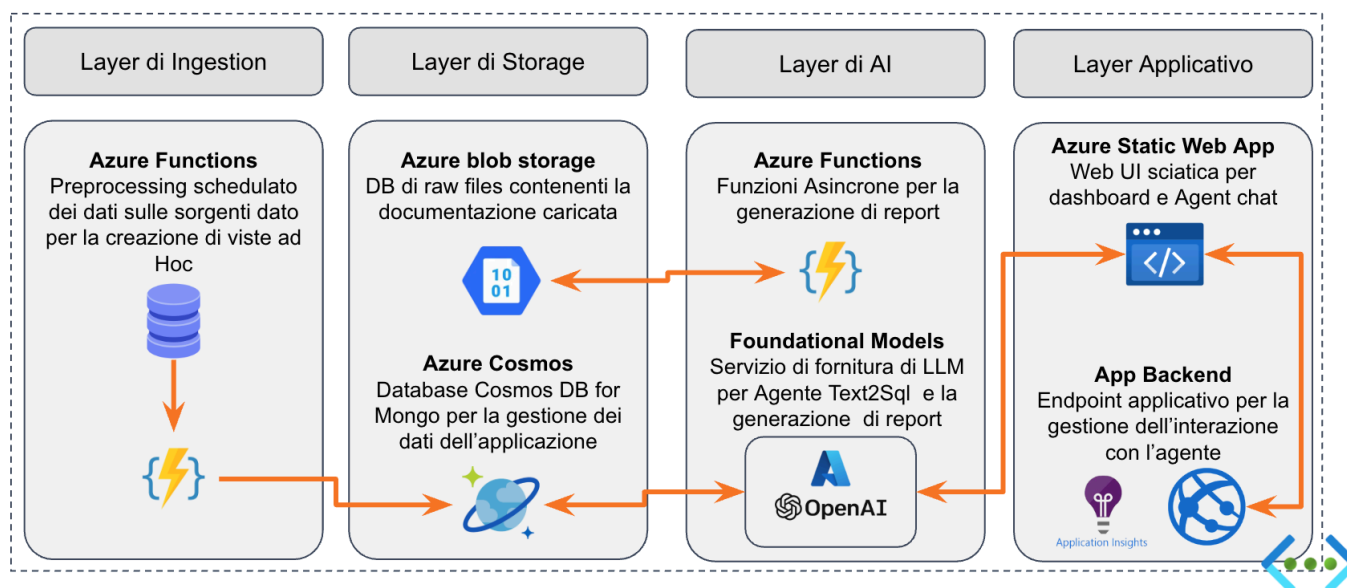
Il sistema sarà isolato all'interno di una **Virtual Network (VNet) dedicata**. Questa configurazione permette di limitare l'accesso ai soli sistemi autorizzati e instradare il traffico attraverso connessioni private, evitando esposizioni pubbliche. Sarà prevista l'integrazione con la stessa Vnet per gli assistenti virtuali e la dashboard.

L'architettura prevede:

- **VNet dedicata**
- **Integrazione dell'App Service con la VNet**, per connettersi internamente a **Azure Cosmos Db for Mongo**, **Blob storage** e **Key Vault** tramite IP privati.
- **Private endpoint per il database NoSQL e Blob storage**, assegnati a una subnet dedicata per garantire che il traffico resti confinato all'interno della VNet, senza esposizione pubblica.
- **Private endpoint per Key Vault**, che limita l'accesso solo ai servizi interni alla rete privata.

Questa strategia garantisce un'elevata sicurezza dell'infrastruttura, proteggendo i dati e minimizzando i rischi legati agli accessi non autorizzati.

Di seguito si riporta l'architettura ad alto livello precedentemente descritta:



### 3. Struttura Economica

---

Di seguito si riporta il dettaglio dei costi complessivi di progetto relativi alle macro attività di progetto:

Descrizione	Effort (gg)	Costo (€)
Analisi Funzionale	10	3.500
Sviluppo Interfaccia Web	20	7.000
Sviluppo app backend e Ai workflows	30	13.500
<b>Totale</b>		<b><del>24.000€</del></b>
<b>Totale Scontato</b>		<b>21.500</b>

Tutti i prezzi sono da considerarsi IVA esclusa.

La seguente offerta è **valida per 30 giorni** dalla data di emissione

# zendata



[info@zendata.it](mailto:info@zendata.it)

[www.zendata.it](http://www.zendata.it)