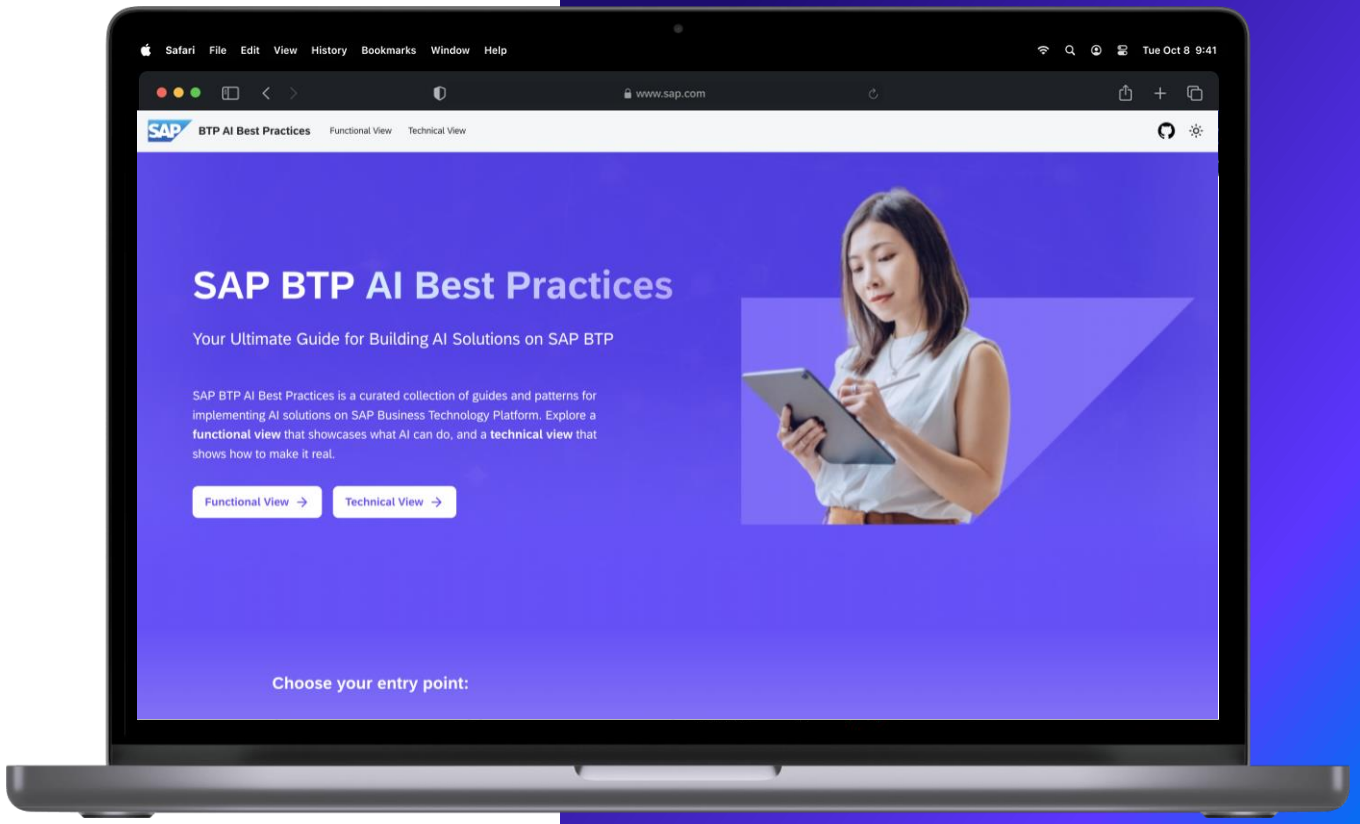
SAP BTP AI Best Practices

Document Grounding

A simple and powerful approach to improve LLM responses grounded on your enterprise data.



BTP AI Services Center of Excellence

29.10.2025

Steps

1 Overview

2 Pre-requisites

3 Key Choices and Guidelines

4 Implementation

What is Document Grounding?

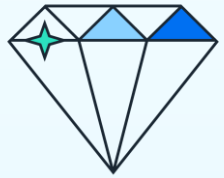
Bring enterprise knowledge into your AI workflows

Data Repository

Automatically processes documents from various sources

Retrieval

Based on input query, retrieve most similar matching documents as context



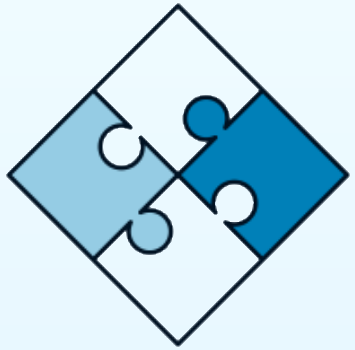
Generate

Retrieved context can be used to generate accurate and relevant answers from LLM

- Managed service under SAP BTP Generative AI Hub
- Enables Retrieval-Augmented Generation (RAG) out of the box
- Connect enterprise sources (S3, SharePoint, DMS, Work Zone, SFTP)
- Automates ingestion, embedding, and retrieval for grounded AI

Key Benefits of Document Grounding

Simplify RAG with enterprise-grade automation



Plug-and-Play Connectors

Seamlessly link enterprise data sources like S3 or SharePoint

Scalable & Efficiency

Vector databases are optimized for fast, scalable similarity search using embeddings

Extensible with Vector API

Add fine-grained control for advanced use cases

Key Concepts

- **Retrieval Augmented Generation (RAG):**

  An approach that retrieves relevant information from external knowledge sources to augment response from an LLM, ensuring accuracy and factual grounding..

- **Embedding:**

  Numerical representations of text segments (chunks) that capture semantic meaning and are stored for efficient similarity-based retrieval.

- **Chunking:**

  Chunking in Retrieval-Augmented Generation (RAG) refers to the process of dividing large text corpora into smaller, manageable pieces known as chunks. Please refer Chunking strategies for more details.

- **Managed Connectors:**

  Prebuilt connectors provided by SAP to integrate data sources such as S3, SharePoint, or DMS without custom development.

Common Use cases

Grounded AI across enterprise workflows

- **Invoice Validation & Processing:**

  Validate extracted invoice data against purchase orders, goods receipts, or contracts to ensure compliance.

- **Risk and Compliance Summarization:**

  Analyse lengthy policies or audit reports to extract key compliance points and obligations

- **Contract Clause Extraction & Verification:**

  Identify, compare, and verify specific contractual clauses (e.g., payment terms, liability limits) across documents.

- **Contextual Assistance in SAP Business Processes:**

  Power intelligent copilots that provide policy lookups or contextual answers within SAP applications.

- **Knowledge Retrieval from Internal Repositories:**

  Let employees query internal handbooks or SOPs in natural language and get grounded, context-aware answers.

Security and Compliance

Built for enterprise trust and control

- **Enterprise-Grade Protection:**

  All processing stays securely within SAP-managed environments.

- **Data Isolation and Control:**

  Each tenant's data and embeddings remain fully isolated.

- **Secret Management:**

  Credentials stored as encrypted Generic Secrets in AI Launchpad.

- **Content Security and Masking:**

  Sensitive or PII data protected through masking and filtering.

- **Compliance and Data Residency:**

  Meets GDPR and ISO 27001; data stays in your regional data center.

Pre-requisites

Business

- SAP AI Core with the “Extended” tier on SAP BTP (Pricing Information)
- SAP AI Launchpad (Pricing Information)

Technical

- SAP Business Technology Platform subaccount (Setup Guide)
- SAP AI Core (Setup Guide)
- SAP AI Launchpad (Setup Guide)
- Credentials for your document source (Setup Guide)

SAP Business Technology Platform (SAP BTP)

- SAP Business Technology Platform (BTP) is an integrated suite of cloud services, databases, AI, and development tools that enable businesses to build, extend, and integrate SAP and non-SAP applications efficiently.

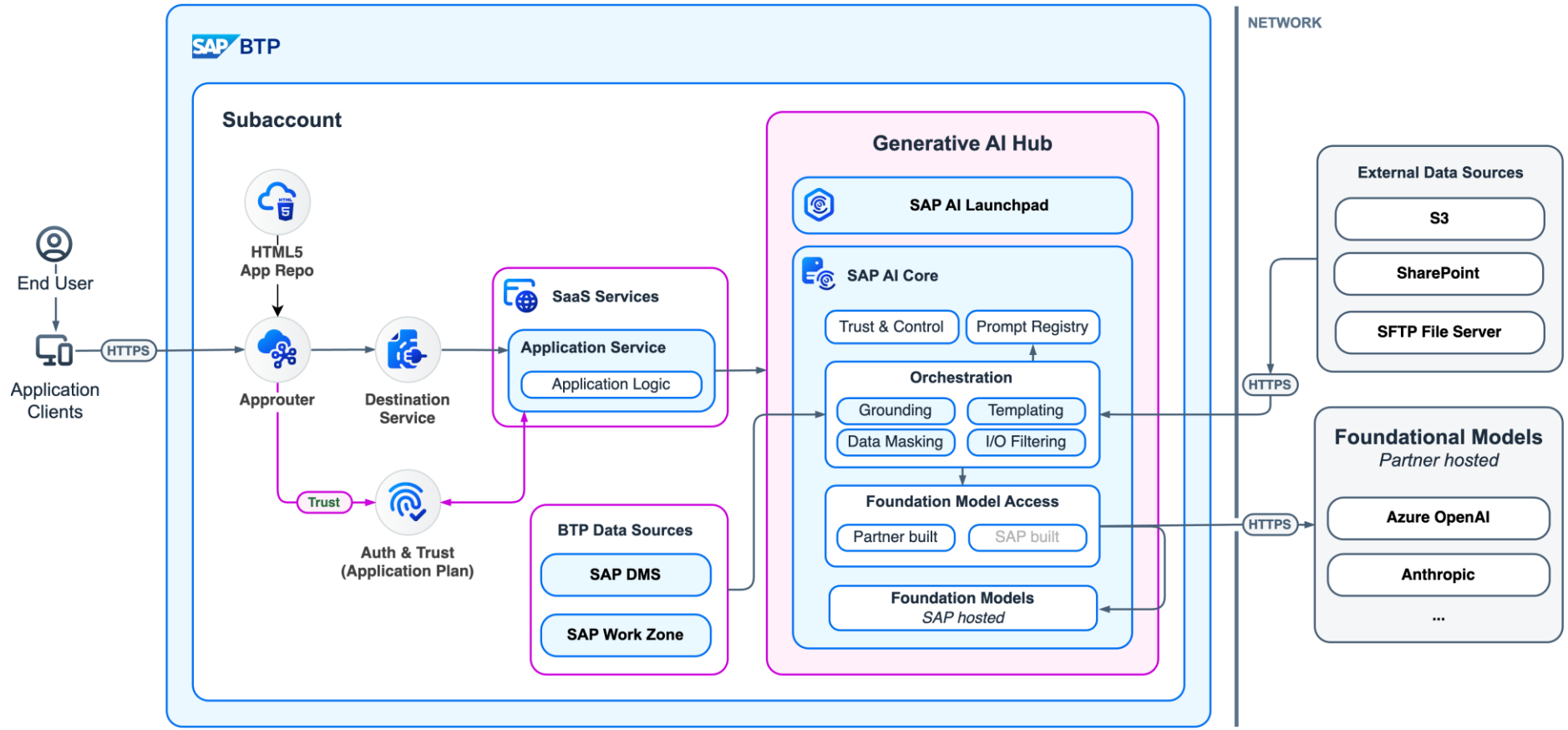
SAP AI Core

- SAP AI Core is a managed AI runtime that enables scalable execution of AI models and pipelines, integrating seamlessly with SAP applications and data on SAP BTP that supports full lifecycle management of AI scenarios.

SAP AI Launchpad

- SAP AI Launchpad is a multitenant SaaS application in SAP BTP. Customers can use SAP AI Launchpad to manage AI use cases (scenarios) across multiple instances of AI runtimes.

High-level reference architecture



Key Choices and Guidelines

Optimizing your grounding setup

Key Choices and Guidelines

1

Decisions that impact the quality and performance of your Vector Embeddings

Data Repository Creation

- A Data Repository (Collection) represents your enterprise knowledge base
- Create via AI Launchpad, Pipeline API, or SAP Cloud AI SDK
- Use clear naming for easy identification and management
- For S3, avoid specifying sub-directory paths if using metadata
- Use Vector API for direct uploads of chunked data and their corresponding metadata
- Refresh repositories to validate ingestion and updates

Key Choices and Guidelines

Decisions that impact the quality and performance of your Vector Embeddings

Metadata Management

- Add metadata per chunk using Vector API
- For S3/SharePoint/SFTP, host metadata server separately
- SAP DMS manages metadata internally

Grounding Configuration

- Define repository IDs and retrieval parameters
- Tune Top-K and chunk count for accuracy vs performance
- Use metadata filters for focused retrieval

Orchestration

- Combine multiple repositories, including SAP Help
- Use Orchestration Service for dynamic LLM access
- Apply masking and filtering for data protection

Implementation

Programming Model reference to implement Document Grounding

Python

SDK

- SAP Cloud SDK for AI (Python) (For building apps)
- SAP AI Core SDK and AI API Client SDK (AI Core lifecycle)

Reference Code

- Document Grounding with S3 docs using API
- Document Grounding with S3 docs using Python SDK
- Custom repository with Vector API

Learning Journeys

- Predictive AI with SAP AI Core
- SAP Generative AI Hub

Java

SDK

- SAP Cloud SDK for AI (Java)

Reference Code

- SAP Cloud SDK for AI - Tutorial

Learning Journeys

- There are currently no learning journeys using the official SDK.

JavaScript

SDK

- SAP Cloud SDK for AI (JavaScript)

Reference Code

- SAP Cloud SDK for AI - Tutorial

Learning Journeys

- There are currently no learning journeys using the official SDK.

Reference Code and Setup Guidelines

Python

- **Code Samples (Python Notebooks)**
  - **S3 Integration:** Jupyter Notebooks Using SDK and API
  - **Vector Data Ingestion:** Jupyter Notebook using Vector API
- **Setup Guides**
  - **SAP DMS:** Connect DMS with Document Grounding
  - **SharePoint:** Integration setup with Document Grounding
  - **SAP Work Zone:** Linking SAP Workzone content
  - **SFTP Server:** Secure file transfer configuration
- **API Collection**
  - **Document Grounding APIs:** Manage repositories, pipelines, and queries

Contributors

Behera, Bhagabat Prasad

Rzhaksynskyi, Andrii

Thank you