INT161

Detect and Highlight entities from unstructured text using machine learning

Komal Narsinghani, SAP

THE BEST RUN **SAP**

Disclaimer

The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. Except for your obligation to protect confidential information, this presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this presentation or any related document, or to develop or release any functionality mentioned therein.

This presentation, or any related document and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information in this presentation is not a commitment, promise or legal obligation to deliver any material, code or functionality. This presentation is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This presentation is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this presentation, except if such damages were caused by SAP's intentional or gross negligence.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.

Agenda

Overview

- What is Business Entity Recognition

Use-cases

- Explore various use-cases and applications of Business Entity Recognition

Technical Overview

- Architecture, APIs and Machine Learning process

Demo

- A demo showing Business Entity Recognition in action
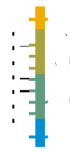
What is Business Entity Recognition?
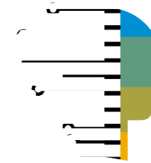
Business Entity Recognition (BER)

Detect and highlight any given type of named entity in unstructured text

Text Document for Entity Extraction > Pre trained/Custom Entity Recognition model for Text > Automatically recognize entities in text document

Powerful machine learning technology to detect and highlight any given type of named entity in unstructured text into pre-defined categories

Increase efficiency and reduce errors
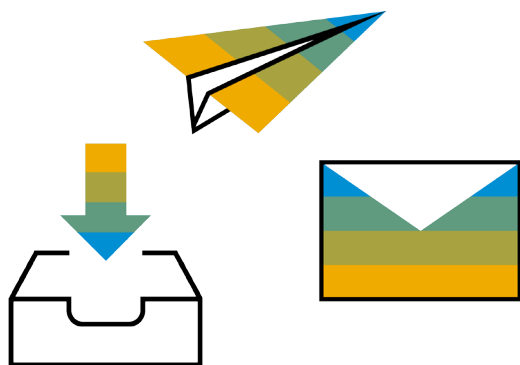
Immediately Productive with Pre-Trained models

Increase overall document processing productivity

Business Entity Recognition: Use-cases

Locating and Classifying Named Entities in Unstructured Text
Moments of truth

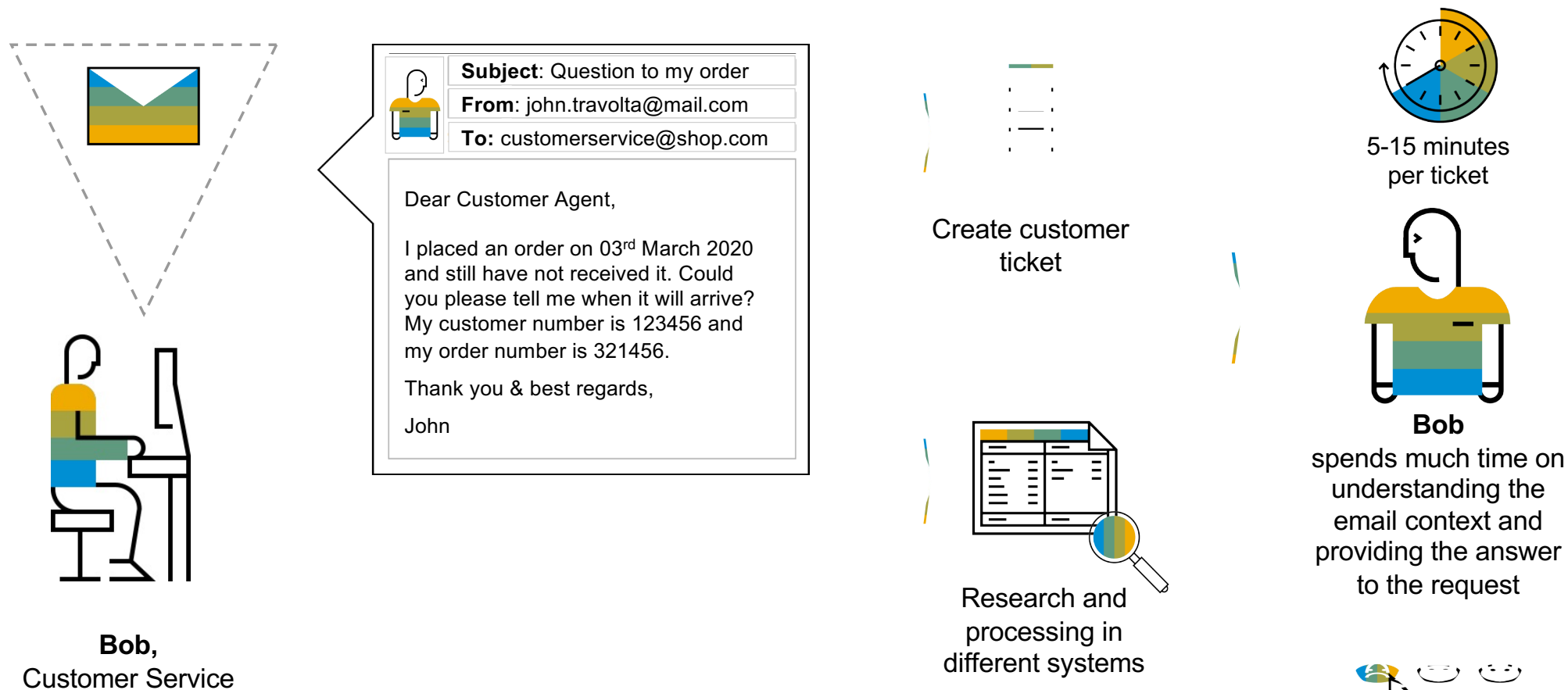The average office worker **receives 121 emails and sends about 40** each day.

Employees spend **20 mins** on average to complete one expense report.

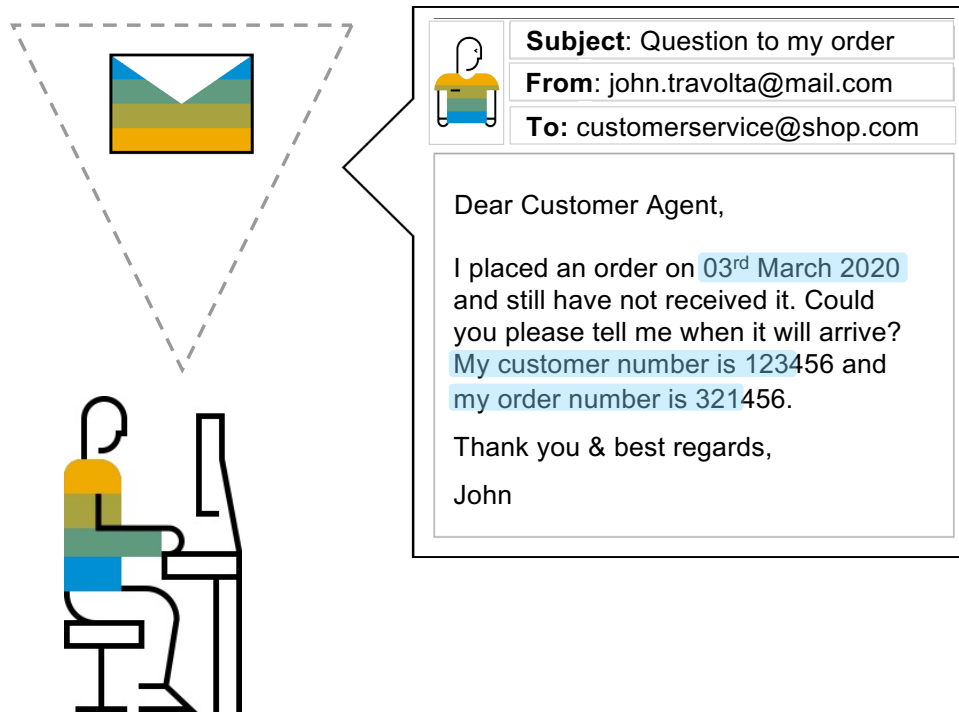# Locating and Classifying Named Entities in Unstructured Text
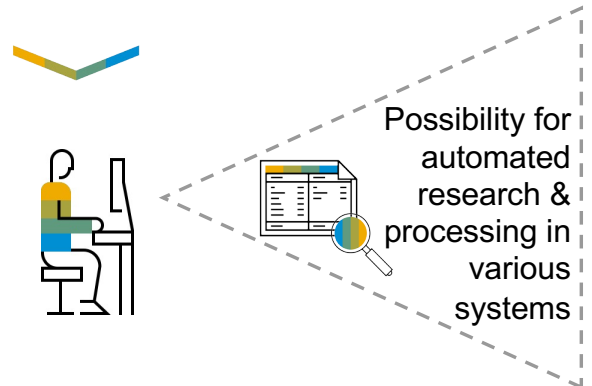## Problem: long time to understand the context of emails and processing them

Subject: Question to my order	
From: john.travolta@mail.com	
To: customerservice@shop.com	

Dear Customer Agent,

I placed an order on 03rd March 2020 and still have not received it. Could you please tell me when it will arrive? My customer number is 123456 and my order number is 321456.

Thank you & best regards,

John

Bob,
Customer Service

Create customer ticket

Research and processing in different systems

5-15 minutes per ticket

Bob
spends much time on understanding the email context and providing the answer to the request

Locating and Classifying Named Entities in Unstructured Text
Solution: Business Entity Recognition

Subject: Question to my order
From: john.travolta@mail.com
To: customerservice@shop.com

Dear Customer Agent,

I placed an order on 03rd March 2020 and still have not received it. Could you please tell me when it will arrive? My customer number is 123456 and my order number is 321456.

Thank you & best regards,

John

Business Entity Recognition

Customer #	123456
Order #	321456
Date	03.03.2020

Create customer ticket

Possibility for automated research & processing in various systems
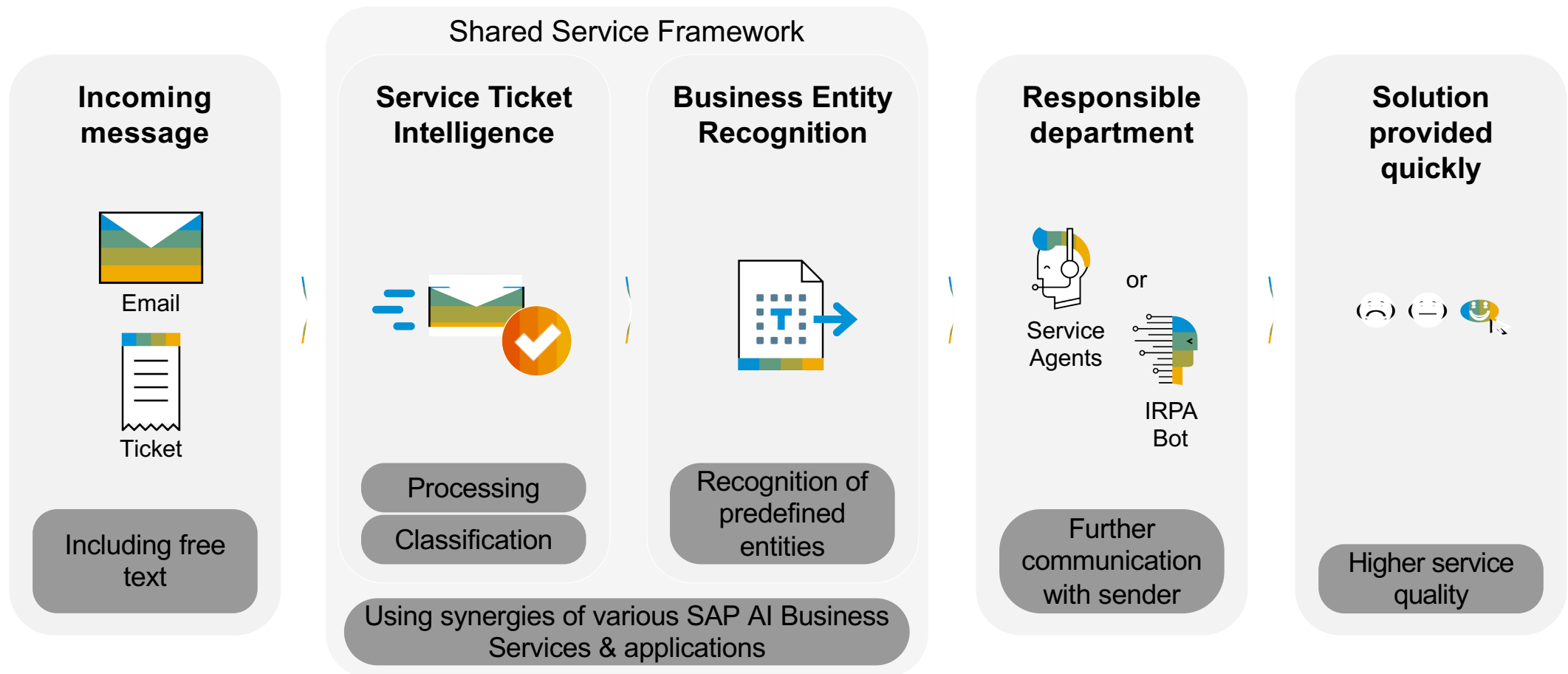
Bob,
Customer Service

Bob
can now provide a faster response and use his time for value-creating tasks

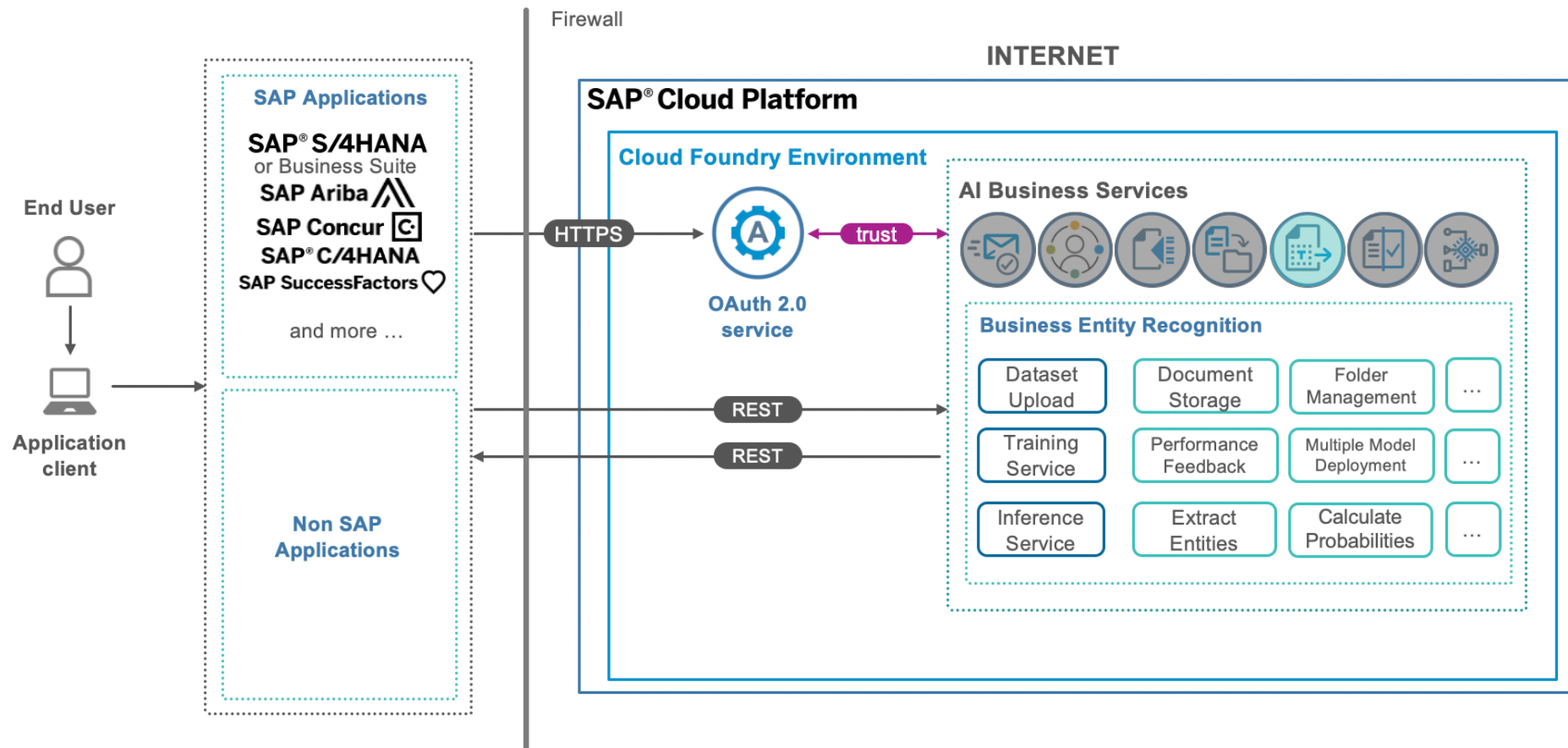# Locating and Classifying Named Entities in Unstructured Text
Shared Services Framework

Shared Service Framework

Incoming message

Email

Ticket

Including free text

Service Ticket Intelligence

Processing

Classification

Business Entity Recognition

Recognition of predefined entities

Using synergies of various SAP AI Business Services & applications

Responsible department

Service Agents

or

IRPA Bot

Further communication with sender

Solution provided quickly

Higher service quality

Business Entity Recognition: Technical Overview

Business Entity Recognition Architecture

Firewall

INTERNET

SAP® Cloud Platform

Cloud Foundry Environment

SAP Applications

SAP® S/4HANA
or Business Suite
SAP Ariba
SAP Concur
SAP® C/4HANA
SAP SuccessFactors

and more …

End User

Application client

Non SAP Applications

HTTPS

OAuth 2.0 service

trust

AI Business Services

Business Entity Recognition

Dataset Upload	Document Storage	Folder Management	…
Training Service	Performance Feedback	Multiple Model Deployment	…
Inference Service	Extract Entities	Calculate Probabilities	…

REST

REST

Business Entity Recognition: APIs

BER has a set of APIs to perform the following operations:

- Create dataset and upload documents

- Trigger training to get a custom model based on your dataset

- Deploy the custom model to use it for inference

- Perform inference with the deployed model to extract entities

API: Trigger training to create custom models

Datasets Training Data APIs : Upload, delete and get training data

| GET | /datasets | Retrieve all datasets |

| POST | /datasets | Creates a new training dataset entry |

| DELETE | /datasets/{datasetId} | Deleting a dataset and its contents |

| GET | /datasets/{datasetId} | Get specific dataset details |

| GET | /datasets/{datasetId}/documents | Get documents present in a dataset |

| POST | /datasets/{datasetId}/documents | Upload a training document |

| DELETE | /datasets/{datasetId}/documents/{documentId} | Deletes specific training document of the given dataset ID |

| GET | /datasets/{datasetId}/documents/{documentId} | Retrieve specific training document detail of the given dataset ID |

Training Training Job APIs : Trigger training job and get status

| POST | /training/jobs | Submit training job |

| DELETE | /training/jobs/{jobId} | delete training job |

| GET | /training/jobs/{jobId} | Get training status |

Deployments Model Deployment APIs : Create, delete and get model deployments

| GET | /deployments | Get all deployment details |

| POST | /deployments | Deploy a model |

| DELETE | /deployments/{deploymentId} | Undeploy a successfully deployed model |

| GET | /deployments/{deploymentId} | Get information of a deployed model |

API: Inference

Inference Inference Job APIs : Trigger inference job and get extracted entities

| POST | `/inference/jobs` | Submit inference job |

| GET | `/inference/jobs/{jobId}` | Get inference status and extracted entities |

Summary

| Upload Data | Train Model | Deploy Model | Prediction |

Upload Data
- Use Pre-trained models for specific business problems
- Upload own entities with annotations to train a custom specific model

Train Model
- Train new entity models
- Create model version
- Get feedback on model performance

Deploy Model
- Activate trained model version
- Deactivate model version when necessary

Prediction
- Provide data for entity recognition
- Get machine learning results + confidence scores

Business Entity Recognition: Data Management

Training Data

- Stored in Minio object storage

- The data is isolated for each tenant

- The service user has the full control of the data- upload and delete APIs are provided

- It is recommended to use anonymization to restrict processing of PIIs by the service

Inference Data

- Texts provided for inference are temporarily stored in the database (PostgreSQL)

- A scheduled job runs every 24 hours to delete the texts which are older than 12 hours

- Hence, no inference text would be stored for more than 24 hours

Business Entity Recognition: Machine Learning Process

| Data pre-processing | Model Training | Model Selection |

Data pre-processing
- Validate training data format
- Validate annotated labels
- Split the dataset to train and test dataset
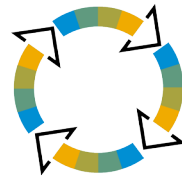
Model Training
- Deep Learning applied to Natural Language Processing
- Basic components: convolutional and recurrent neural networks
- Transfer learning capability:
  - Pretrained language models (Denoising Auto-Encoder)
  - Word embeddings
- Support extension to multilingual data (through language specific embedding)

Model Selection
- Multiple models generated
- Each model validated on the test dataset
- Select model with best f1-score averaging across label classes

Demo

Locating and Classifying Named Entities in Unstructured Text

More information

Related SAP TechEd sessions

- INT110– Learn how to Consume AI Business Services for Business Document Processing
- INT201– Infuse Intelligence into Business Processes with SAP AI Business Services
- INT820– Road Map: SAP AI Business Services
- INT937– Expert Q&A

Public SAP Web sites

- SAP Community: https://community.sap.com/topics/artificial-intelligence
- SAP Help Page: https://help.sap.com/viewer/product/Business_Entity_Recognition

Continue your learning experience from SAP TechEd in 2020

Your exclusive path to build and maintain SAP solution skills anytime, any place

Get empowered with access to relevant, up-to-date digital learning for SAP TechEd participants through a complete enablement solution that drives adoption and innovation.

Deepen your learning experience from SAP TechEd

[Activate your free access](#) to SAP Learning Hub, event edition, for:

- **Learning Journey** illustrations to guide you through **complementary** self-paced learning content
- **Content specific to SAP TechEd** in the online **SAP Learning Room for SAP TechEd**
- Access to SAP experts in **special live sessions**

Deepen and validate your SAP solution skills

[Subscribe](#) to SAP Learning Hub, solution editions, for:

- **Solution-specific Learning Journey guides, content, collaborative learning,** and **hands-on practice** for your role and goals
- Drive performance and business success with validated solution expertise from the **SAP Global Certification** program

Your benefits

- Gain insight into the latest innovations, and master software proficiency
- Keep skills up-to-date, and enable performance and business success with help from SAP solution experts
- Achieve competitive advantages and digital transformation success with trusted certifications

500,000+

Learners in SAP Learning Hub

100+

Experts getting certified per day

150+

SAP Global Certifications

Thanks for attending this session.

Contact for further topic inquiries

Komal Narsinghani
komal.narsinghani@sap.com

THE BEST RUN **SAP**

Follow us

THE BEST RUN **SAP**