



Hiệu suất của ứng dụng và các công cụ khác



Lisboa

POWER PLATFORM
BOOTCAMP



AI-powered data extraction: Azure OpenAI vs Document Intelligence!



POWER PLATFORM
BOOTCAMP



Massimo Bonanni





UNIVERSIDADE
DE LISBOA



Lisbon School
of Economics
& Management
Universidade de Lisboa



This event is organised by the following non-profit technical communities:

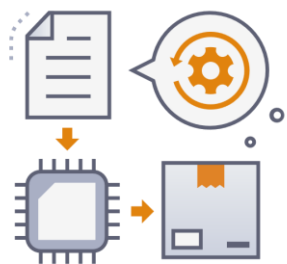


Data extraction from document

- Extracting data from documents is a common scenario in many business processes.
- Involving the conversion of unstructured, semi-structured or structured document into a structured data.
- Is crucial in industries such as finance, legal, healthcare, and many others, where decision-making often depends on the information contained within various documents such as invoices, contracts, medical records, and forms.



Technologies and Tools



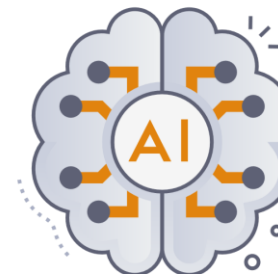
OCR Technology

For converting images of text into machine-encoded text. Advanced OCR solutions can handle complex layouts and even handwritten text.



Natural Language Processing (NLP)

For understanding the context of the extracted text, identifying entities, and classifying them appropriately.

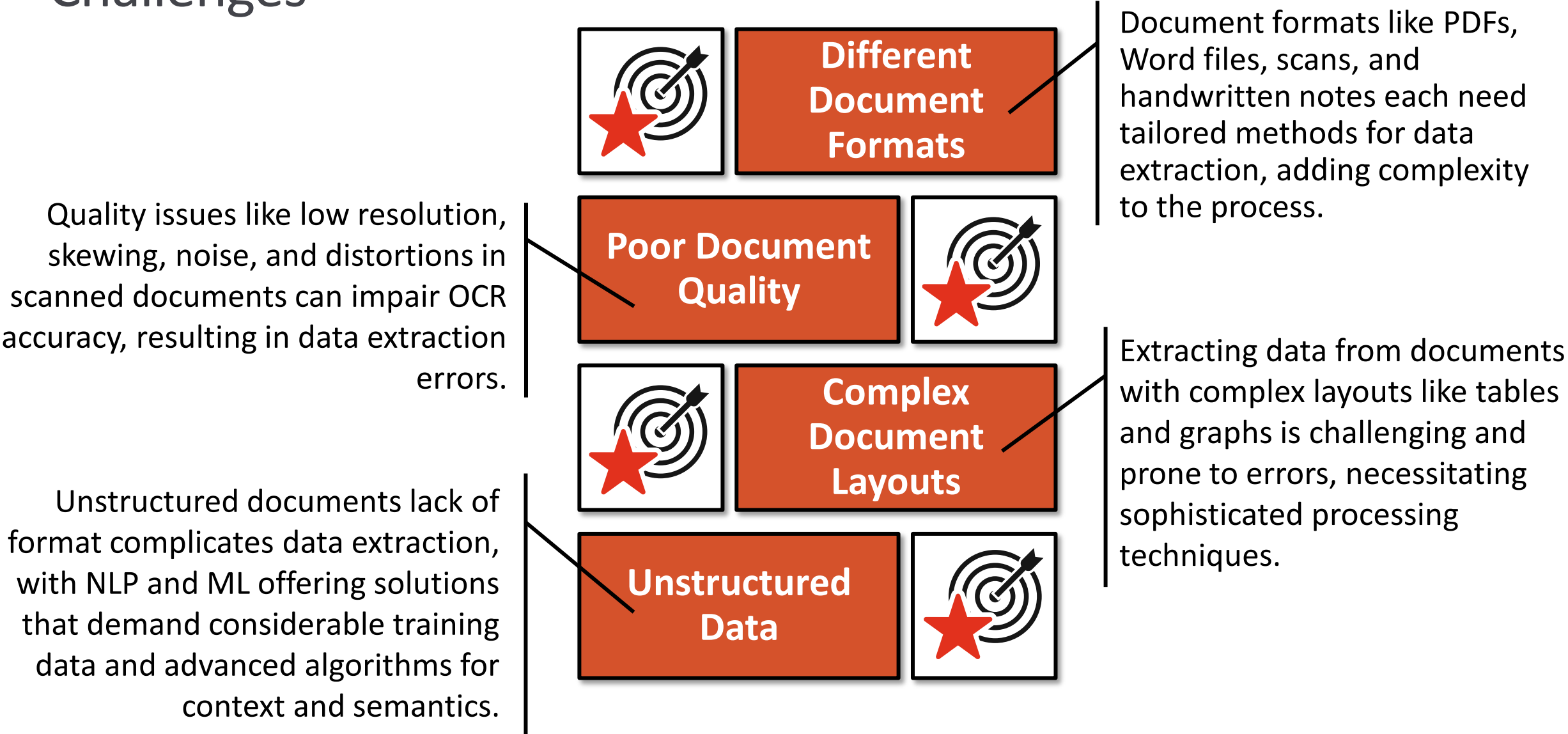


Machine Learning and Artificial Intelligence

For improving the accuracy of data extraction over time, learning from corrections and adjustments made during the post-processing phase.

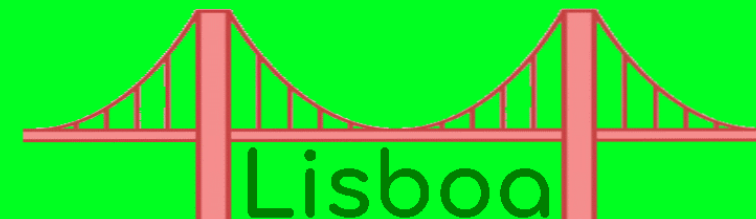


Challenges

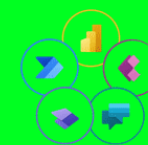


Document Intelligence

Custom Model



+



Google AI LISA GRRBLISA Google AI GRRB...

What is Document Intelligence?

Azure Document Intelligence is a cloud-based service (part of the Azure AI Services portfolio) that uses machine learning to automate the extraction of text and structure from documents.

Azure Document Intelligence can be deployed in the cloud or at the edge, and integrated with other Azure services for data processing, search, and analytics.

It is a powerful tool for enhancing data-driven strategies and enriching document search capabilities.



Document Intelligence Key Features



It can handle various types of documents, such as Pdf, Images or Office documents (only few models) composed by multiple pages



It allows users to customize the extraction process by training on their own data with minimal effort



It offers prebuilt models for common scenarios (IDs, receipts, invoices, contracts) as well as custom models for specific document types



It supports both printed and handwritten text, and provides built-in security and privacy for data and models

Models Overview

Pretrained document-analysis models

Read OCR model

Layout analysis model

Layout analysis model with KeyValuePairs option

Pretrained scenario-specific models

Pretrained models for specific scenario like US tax document, contract, invoice and so on

Custom models

Extraction Models

Classification Models

Add-on capabilities

- Use the add-on features to extend the results to include more features extracted from your documents.
- Some add-on features incur an extra cost.
- To enable a feature, add the associated feature name to the features query string property.
- You can enable more than one add-on feature on a request by providing a comma-separated list of features.

Add-on Capability	Add-On/Free
Font property extraction	Add-On
Formula extraction	Add-On
High resolution extraction	Add-On
Barcode extraction	Free
Language detection	Free
Key value pairs	Free
Query fields	Add-On*

Document Intelligence Studio

- Online tool for Document Intelligence service
- Learn, test, experiment, train, and integrate models and features
- Use your own or sample documents
- Train custom models for classification and extraction
- Get sample code for different languages

Extract text, tables, structure, key-value pairs, and named entities from documents.



Read

Extract printed and handwritten text from images and documents.

[Try it out](#)



Layout

Extract tables, check boxes, and text from forms and documents.

[Try it out](#)



General documents

Extract labels, values and entities such as names, dates, and amounts from documents.

[Try it out](#)



General documents with private preview

Extract labels, values and entities such as names, dates, and amounts from documents.

[Private preview – sign up.](#)

Prebuilt models

Extract data from unique document types using the following prebuilt models.



Invoices

Extract invoice ID, customer details, vendor details, ship to, bill to, total tax, subtotal, line items and more.

[Try it out](#)



Receipts

Extract time and date of the transaction, merchant information, amounts of taxes, totals and more.

[Try it out](#)



Business cards

Extract person name, job title, address, email, company, and phone numbers from business cards.

[Try it out](#)



Identity documents

Extract name, expiration date, readable zone, and more from ID cards.

[Try it out](#)



Health insurance cards

Extract insurer, member, prescription, group number and more information from US health insurance cards.

[Try it out](#)



W-2

Extract employee, employer, wage information, etc. from US W-2 Tax Form.

[Try it out](#)



1098

Extract mortgage interest information from US 1098 Tax Form.

[Private preview – sign up.](#)



1098-E

Extract student loan information from US 1098-E Tax Form.

[Private preview – sign up.](#)

ĞłôçăłAÍLîşçôă ĞRRBLîşçôă ĞłôçăłAÍ ĞRRB, . . . –

Pricing

Instance	Document type	Price
Free - Web/Container ¹	All	0 - 500 pages free per month
S0 - Web/Container	Read	0-1M pages - €1.390 per 1,000 pages 1M+ pages - €0.556 per 1,000 pages
S0 - Web/Container	All Prebuilt Models: Document, Layout, Receipt, Invoice, ID, W-2, 1098 Tax forms, Health insurance card, Contract.	€9.261 per 1,000 pages
S0 - Web/Container	Custom classification	€2.779 per 1,000 pages
S0 - Web/Container	Custom extraction	€46.303 per 1,000 pages
S0 - Web/Container	Add-On ²	€5.557 per 1,000 pages
S0 - Web/Container	Query Fields	€9.261 per 1,000 pages



Demo

Session Feedback Form with Document Intelligence

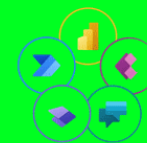
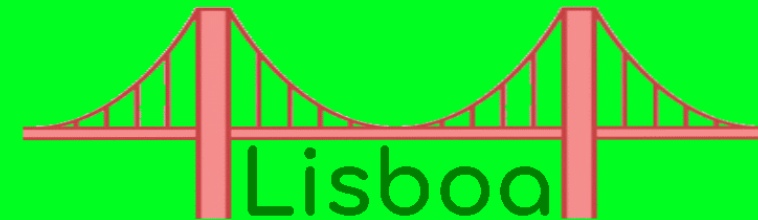


[massimobonanni/FeedbackExtractor](https://github.com/massimobonanni/FeedbackExtractor)

ĞłôčăłAÍLîşçôă ĞRRBLîşçôă ĞłôčăłAÍ ĞRRB,...

Azure OpenAI

GPT-4 Turbo with Vision



Đỗ Hải Anh Lê Thị Ngọc Hà Đỗ Hải Anh Lê Thị Ngọc Hà Đỗ Hải Anh Lê Thị Ngọc Hà Đỗ Hải Anh Lê Thị Ngọc Hà

What is Azure OpenAI?

Azure OpenAI is a cloud-based platform that offers access to various OpenAI models.

The service supports the latest OpenAI models such as GPT-4, **GPT-4 Turbo with Vision**, GPT-3.5-Turbo, and Embeddings

The service can be accessed through REST APIs, Python/C# SDK, or the web-based interface in the Azure OpenAI Studio



GPT-4 Turbo with Vision



It can process images and generate text-based answers to queries about them



It combines language and vision skills.



The model can respond to basic questions about the contents of the images



It can also handle video input with the Vision enhancement feature

Extract Data with Prompt Engineering

Prompt Engineering is a technique that leverages natural language processing and artificial intelligence to create effective prompts for various tasks and domains.

A **prompt is a text input that guides an AI model** to produce a desired output, such as a summary, a classification, a translation, or a generation.

One of the applications of Prompt Engineering is to **extract data from texts and documents**, such as names, dates, locations, keywords, sentiments, etc.

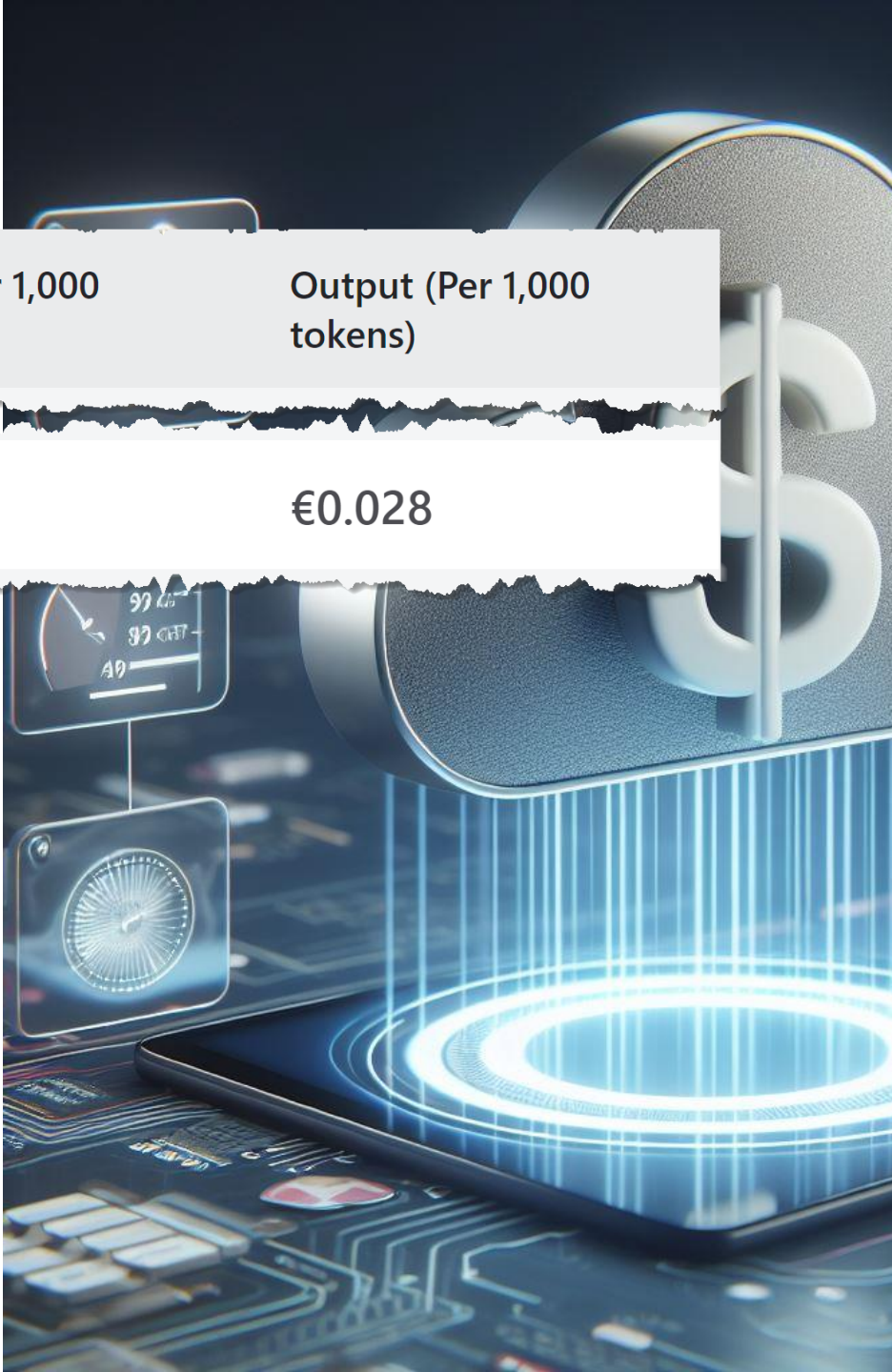


Pricing

Models	Context	Input (Per 1,000 tokens)	Output (Per 1,000 tokens)
GPT-4-Turbo-Vision	128K	€0.010	€0.028

The **detail** parameter affects the number of tokens used for the source images.

- The **low** option uses a lower resolution and saves time and tokens.
- The **high** option uses a higher resolution and consumes more time and tokens.
- The **auto** option chooses between low or high based on the image size.



Pricing

- Low detail allows the API to return faster responses and consume fewer input tokens for use cases that don't require high detail.
- These images cost **85 tokens** each, regardless of the image size.

Low Resolution



- High detail lets the API see the image in more details.
- The token cost is calculated by a series of scaling steps:
 1. The image is first scaled to fit within a 2048 x 2048 square while maintaining its aspect ratio.
 2. The image is then scaled down so that the shortest side is 768 pixels long.
 3. The image is divided into 512-pixel square tiles, and the number of these tiles (rounding up for partial tiles) determines the final cost. Each tile costs **170 tokens**.
 4. An additional **85 tokens** are added to the total cost.

High Resolution



Pricing

Session Feedback Form

Event Name: GLOBAL AI BOOTCAMP

Session Code: Session 1

How useful did you think this event was for you?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

How would you rate the session?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Do you agree that the speaker was prepared about the topic of the session?
(please check the appropriate number 1=disagree and 5=completely agree)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Comment:
Amazing Session

Session Feedback Form

Event Name: GLOBAL AI BOOTCAMP

Session Code: Session 1

How useful did you think this event was for you?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

How would you rate the session?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Do you agree that the speaker was prepared about the topic of the session?
(please check the appropriate number 1=disagree and 5=completely agree)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Comment:
Amazing Session

Scaling to fit
2048x2048



Session Feedback Form

Event Name: GLOBAL AI BOOTCAMP

Session Code: Session 1

How useful did you think this event was for you?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

How would you rate the session?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

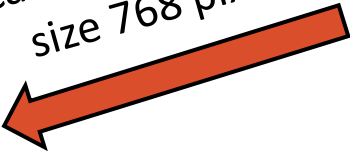
1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Do you agree that the speaker was prepared about the topic of the session?
(please check the appropriate number 1=disagree and 5=completely agree)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Comment:
Amazing Session

Scaling to minimum
size 768 pixels



Divided into 512-
pixel square tiles



Session Feedback Form

Event Name: GLOBAL AI BOOTCAMP

Session Code: Session 1

How useful did you think this event was for you?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

How would you rate the session?
(please check the appropriate number where 1=not at all useful and 5=extremely useful)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Do you agree that the speaker was prepared about the topic of the session?
(please check the appropriate number 1=disagree and 5=completely agree)

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Comment:
Amazing Session

Image Tokens:
 $4 \times 170 + 85 = 765$

Prompt Tokens:
System → 18
User → 465

Input Tokens:
 $1248 \leq 0.020\text{€}$

GLOBAL AI BOOTCAMP GLOBAL AI BOOTCAMP GLOBAL AI BOOTCAMP

Demo

Session Feedback Form with GTP-4 Vision and Prompt Engineering



[massimobonanni/FeedbackExtractor](https://github.com/massimobonanni/FeedbackExtractor)

Takeaways



Document Intelligence supports PDF and DOCX files up to 2000 pages and run on containers

GPT 4.0 – Vision can extract data from unstructured documents

You can orchestrate both calling Document Intelligence first and GPT 4.0-Vision after to refine data.

Questões?

Questions?

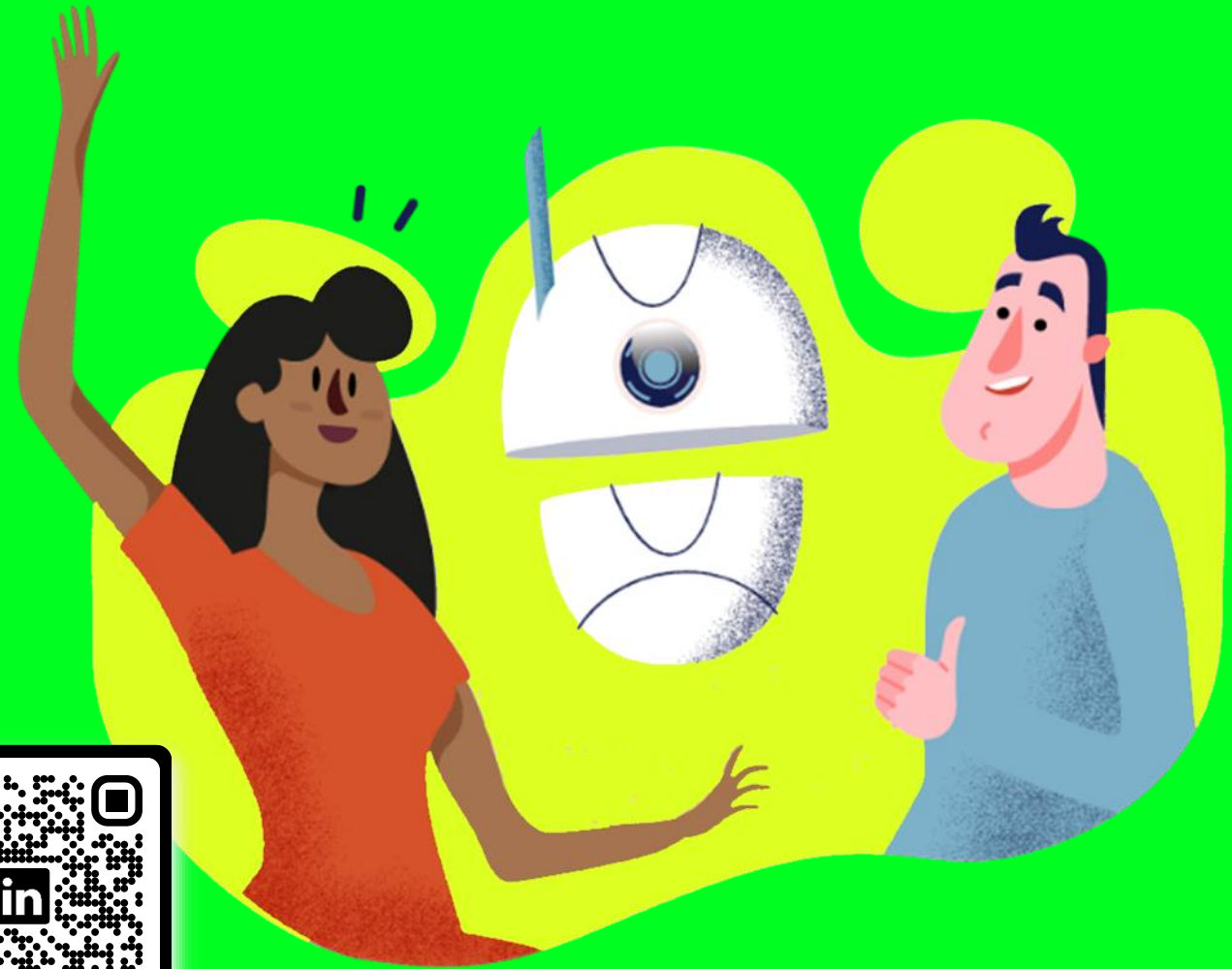


Massimo Bonanni

Microsoft Technical Trainer

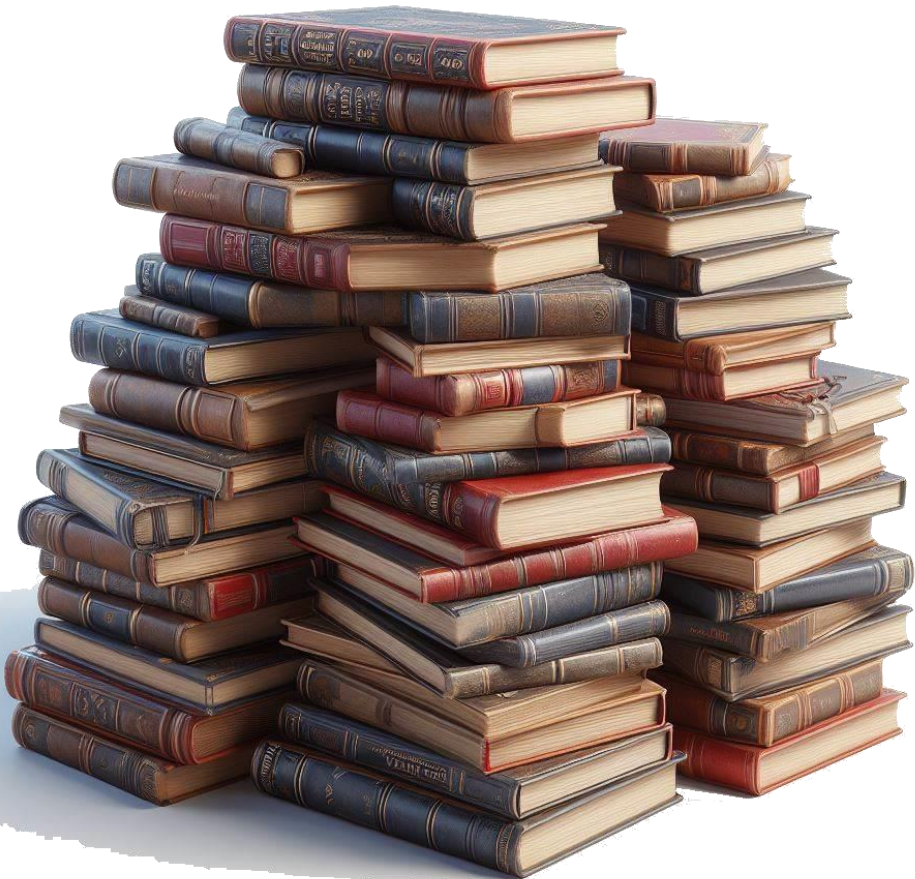
massimo.bonanni@microsoft.com

@massimobonanni



Global AI Lisboa GRRB Lisboa Global AI GRRB...

References



[Azure AI Document Intelligence](#)

[Document Intelligence documentation - Quickstarts, Tutorials, API Reference](#)

[Document Intelligence Studio](#)

[Pricing - Azure AI Document Intelligence](#)

[Azure OpenAI Service – Advanced Language Models](#)

[How to use the GPT-4 Turbo with Vision model - Azure OpenAI Service](#)

[Azure OpenAI Studio](#)

[Azure OpenAI Service - Pricing](#)



Hiệu quả của việc sử dụng công nghệ trong việc xây dựng và vận hành doanh nghiệp