

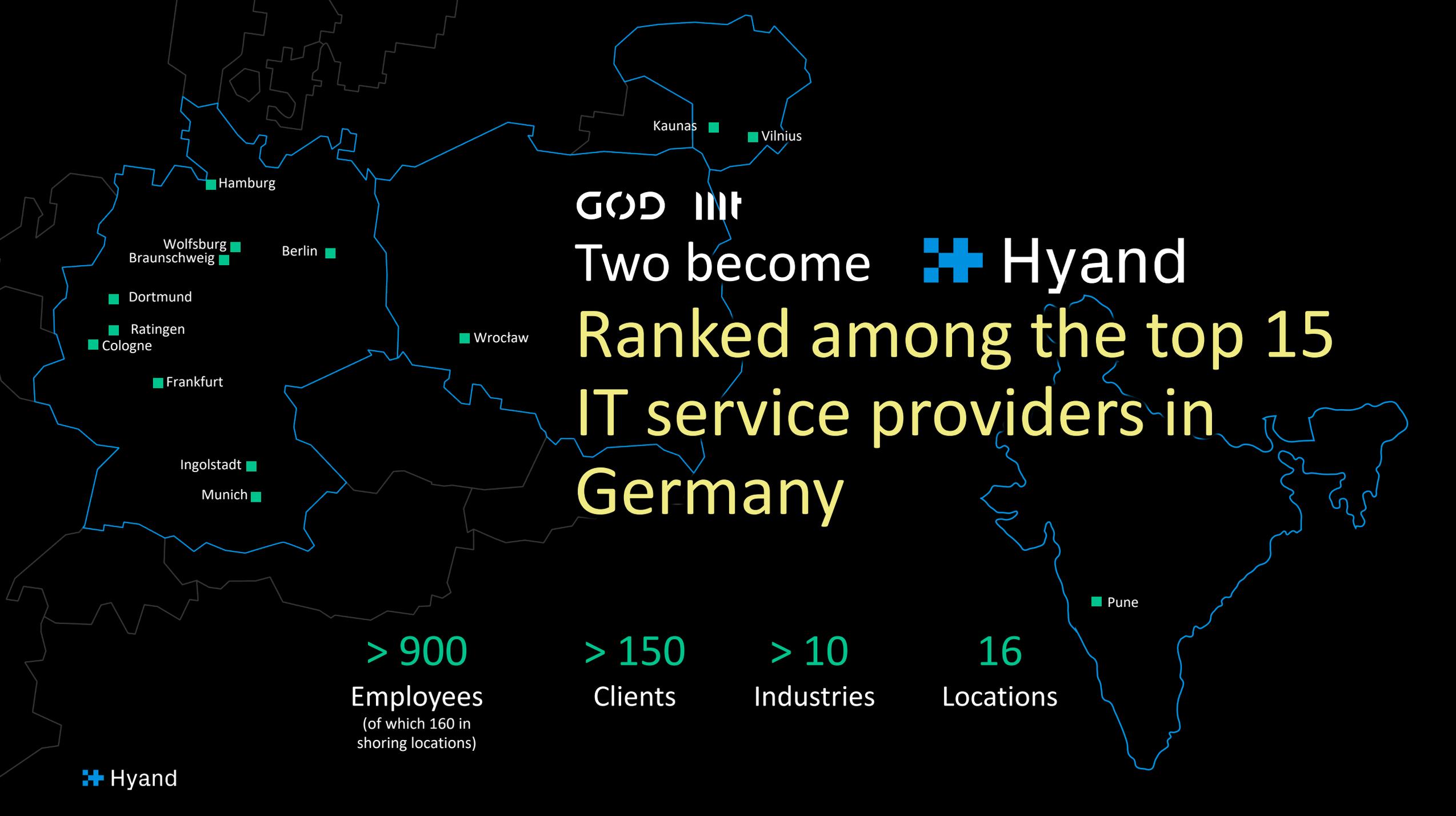
23.04.2024, Fabian Neureiter

Hands-On AI Features

An OCI Document Understanding Deep Dive

code of change

 **Hyand** by
GOD | MT



> 900

Employees
(of which 160 in
shoring locations)

> 150

Clients

> 10

Industries

16

Locations

whoami



Fabian Neureiter
Oracle APEX Developer

Part of the JavaScript Team @ Hyand

Certified Tester, specialized in automated UI Testing

Lead Developer **LCT - Low Code Testing**

Motivation for this talk

My (initial) thoughts on AI

I am a chronic late adaptor of technology

Something is trending or a lot of people are talking
about it

-> I get suspicious



First experiences with (consumer facing) AI

Github Copilot

Saw the media coverage

→ Mildly interested

Saw Youtubers using it

→ Mildly interested

Saw my former colleague Philip Hartenfeller using it

-> **Very interested!**



```
index.tsx

procedure update_step (
    pi_step_id                in lct_steps.step_id%type
    , pi_step_case_id         in lct_steps.step_case_id%type
    , pi_step_sttp_id        in lct_steps.step_sttp_id%type
    , pi_step_name           in lct_steps.step_name%type
    , pi_step_execution_sequence in lct_steps.step_execution_sequence%type
    , pi_checksum            in types_pkg.t_checksum
    , pi_user                in types_pkg.t_user default v('APP_USER')
    , pi_app_id              in apex_applications.application_id%type
    , pi_page_id             in apex_application_pages.page_id%type
    , pi_element             in types_pkg.t_vc_max
    , pi_value               in types_pkg.t_vc_max
    , pi_context             in types_pkg.t_vc_max
    , pi_page_is_modal       in number
    , pi_selector_type       in types_pkg.t_vc_max
    , pi_custom_selector     in types_pkg.t_vc_max
    , pi_custom_selector_text in types_pkg.t_vc_max
    , pi_step_timeout        in number
    , pi_force_interaction   in number
    , pi_selectOption_option in types_pkg.t_vc_max
    , pi_append_session_id   in number
    , pi_path                in types_pkg.t_vc_max
    , pi_nav_referer         in types_pkg.t_vc_max
    , pi_nav_wait_until      in types_pkg.t_vc_max
    , pi_click_btn_type      in types_pkg.t_vc_max
    , pi_click_click_count   in number
    , pi_delay               in number
    , pi_trial               in number
    , pi_presskey_modifiers  in types_pkg.t_vc_max
    , pi_presskey_key        in types_pkg.t_vc_max
    , pi_click_wait_for_navigation in number
    , pi_selectOption_select_by in types_pkg.t_vc_max
    , pi_waitMilliseconds_time in number
    , pi_remove_target       in number
    , pi_region              in types_pkg.t_vc_max
    , pi_column_id           in types_pkg.t_vc_max
)
as
    l_scope logger_logs.scope%type := gc_scope_prefix || 'update_step';
    l_params logger.tab_param;

begin
```

CodeImage



A code editor window with a dark theme. The title bar shows a tab labeled 'lct_p0029_api.pkb' with a TypeScript icon. The code contains four lines of TypeScript code that call the 'append_param' method of a 'logger' object. The code is as follows:

```
logger.append_param(l_params, 'pi_step_id', pi_step_id);  
logger.append_param(l_params, 'pi_step_case_id', pi_step_case_id);  
logger.append_param(l_params, 'pi_step_sttp_id', pi_step_sttp_id);  
logger.append_param(l_params, 'pi_step_name', pi_step_name);
```

In the bottom right corner of the editor, there is a 'CodeImage' logo consisting of three horizontal lines and the text 'CodeImage'.

```
lct_p0029_api.pkb

logger.append_param(l_params, 'pi_step_id', pi_step_id);
logger.append_param(l_params, 'pi_step_case_id', pi_step_case_id);
logger.append_param(l_params, 'pi_step_sttp_id', pi_step_sttp_id);
logger.append_param(l_params, 'pi_step_name', pi_step_name);
logger.append_param(l_params, 'pi_step_execution_sequence', pi_step_execution_sequence);
logger.append_param(l_params, 'pi_checksum', pi_checksum);
logger.append_param(l_params, 'pi_app_id', pi_app_id);
logger.append_param(l_params, 'pi_page_id', pi_page_id);
logger.append_param(l_params, 'pi_element', pi_element);
logger.append_param(l_params, 'pi_value', pi_value);
logger.append_param(l_params, 'pi_context', pi_context);
logger.append_param(l_params, 'pi_page_is_modal', pi_page_is_modal);
logger.append_param(l_params, 'pi_selector_type', pi_selector_type);
logger.append_param(l_params, 'pi_custom_selector', pi_custom_selector);
logger.append_param(l_params, 'pi_custom_selector_text', pi_custom_selector_text);
logger.append_param(l_params, 'pi_step_timeout', pi_step_timeout);
logger.append_param(l_params, 'pi_force_interaction', pi_force_interaction);
logger.append_param(l_params, 'pi_selectOption_option', pi_selectOption_option);
logger.append_param(l_params, 'pi_append_session_id', pi_append_session_id);
logger.append_param(l_params, 'pi_path', pi_path);
logger.append_param(l_params, 'pi_nav_referer', pi_nav_referer);
logger.append_param(l_params, 'pi_nav_wait_until', pi_nav_wait_until);
logger.append_param(l_params, 'pi_click_btn_type', pi_click_btn_type);
logger.append_param(l_params, 'pi_click_click_count', pi_click_click_count);
logger.append_param(l_params, 'pi_delay', pi_delay);
logger.append_param(l_params, 'pi_trial', pi_trial);
logger.append_param(l_params, 'pi_presskey_modifiers', pi_presskey_modifiers);
logger.append_param(l_params, 'pi_presskey_key', pi_presskey_key);
logger.append_param(l_params, 'pi_click_wait_for_navigation', pi_click_wait_for_navigation);
logger.append_param(l_params, 'pi_selectOption_select_by', pi_selectOption_select_by);
logger.append_param(l_params, 'pi_waitMilliseconds_time', pi_waitMilliseconds_time);
logger.append_param(l_params, 'pi_user', pi_user);
logger.append_param(l_params, 'pi_remove_target', pi_remove_target);
logger.append_param(l_params, 'pi_region', pi_region);
logger.append_param(l_params, 'pi_column_id', pi_column_id);
logger.log('START', l_scope, null, l_params);
```

CodemImage

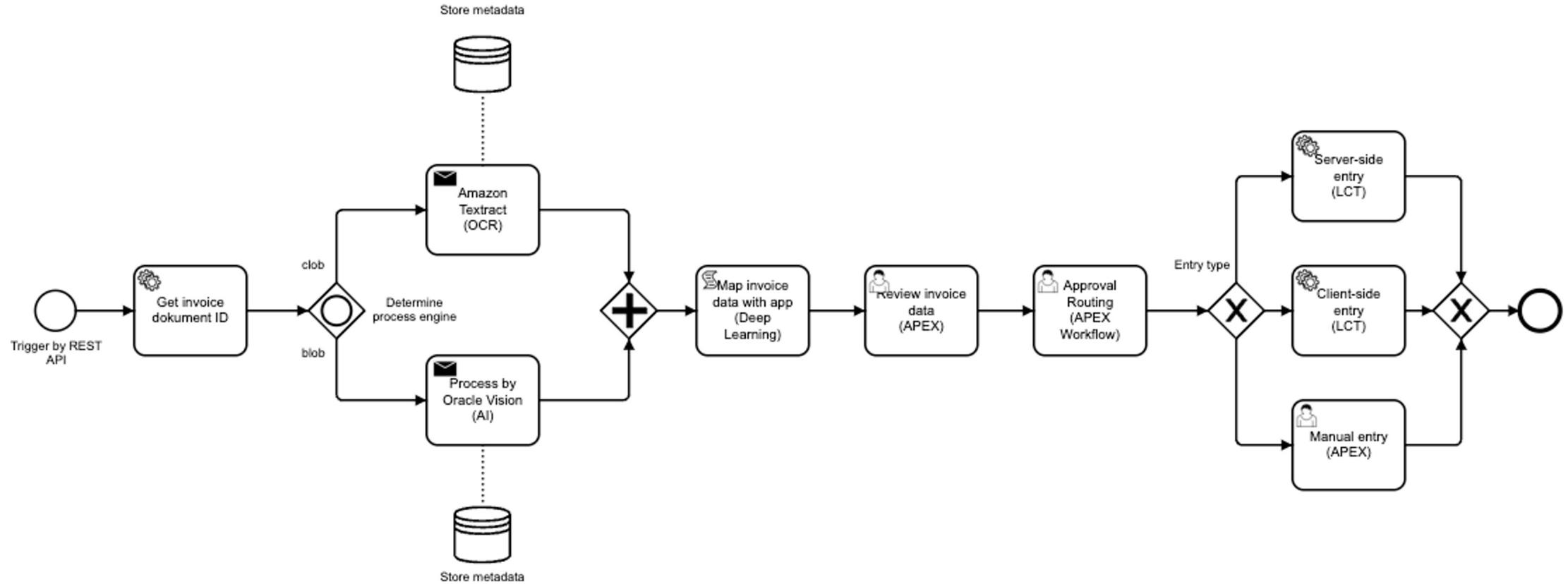
Fabian is a bit less sceptical



Realisation

< AI is first and foremost a tool like all the others and we should use it!/>

Hyperautomation



Agenda

1. Document Understanding – what is it?
2. How can we use it?
3. Live Demo – Setting up your OCI Instance
4. The Document Understanding API – A closer look
5. Live Demo – Using the API in APEX
6. (Current) Limits of the service
7. Alternatives

Document Understanding

What is it?

Document Understanding – what is it?

Document Understanding is an ai-powered service of the Oracle Cloud Infrastructure (OCI) that enables its users to automatically extract information from certain kinds of Documents.

Out of the Box supported document types

- Invoices
- Receipt
- Resume
- Tax Form
- Driver's License
- Passport
- Bank Statement
- Payslip

Means to extract information

- Text Extraction
- Table Extraction
- Key value Extraction
- Document Classification

Text extraction

algorithms run as native SQL functions; not as PL/SQL scripts, call-outs or extensibility framework add-ins. Models are first class database objects that can be built, applied, shared, and audited.

In the early 2000's, starting in Oracle Data Mining Release 9.2i, Oracle's first data mining algorithms took advantage of available core Database's strengths—specifically, counting, parallelism, scalability and other database architectural underpinnings. Essentially, the first two Oracle data mining algorithms, Naive Bayes and A Priori algorithms, are based on counting principles. They count everything very quickly and then assemble conditional probability predictive models—all 100% inside the database. Neither the data, the predictive models nor the results ever leave the database.

OAA Naive Bayes algorithm can quickly build predictive models to predict e.g., "Who will churn?", "Which customers are most likely to purchase Product A?", or "What is the probability that an item will fail?". Let's take an example in a bit more detail for comprehension. Let's say we are interested in selling Product A (e.g. a motorcycle or \$500 shoes, etc.). The Oracle Advanced Analytics data mining algorithms, specifically the Naive Bayes algorithm, of all the customers who purchased Product A, it counts how many customers were male vs. female, how many rent an apartment vs. own their own home? How many have children and how many? Each of these answers involves counts that, taken together, can form a complex conditional probability model that accurately predicts whom we should target to increase our likelihood of selling more of Product A.

OAA's A Priori "market basket analysis" algorithm counts items in each customer's transactional "baskets" while looking for co-occurring items e.g. A + B appear together frequently, and then provides conditional probability AB rules. For example:

IF, "Cereal" AND "Bananas" appear in the same customer's basket

THEN, the "Milk" is also likely to appear in the basket.

WITH Confidence = 87%, and Support = 11%.

Armed with these types of new customer insights from Oracle Advanced Analytics, a store could decide to place the milk near the cereal and bananas, offer new promotional "breakfast kit" product bundles or make real-time customer specific recommendations as the customer checks-out. This is just a simple example of the types of ways that big data analytics can find "actionable insights" from data. Obviously, more data, more advanced analytics methodologies and fast enterprise wide deployment can open new doors to many new big data and analytics applications and solutions possibilities.

SQL and R Support

Where SQL is the standard language for data management and has been for 40+ years, for data analysis, various languages compete—R, SAS, Python and SQL and others. SAS, S+, SQL, SPSS and Matlab have been long time favorites, but in recent past years, open source R especially has surged to the top of the pack and Python and others have emerged. Per the KDD Nuggets data mining industry community annual poll (<http://www.kddnuggets.com/polls/>), R and SQL currently compete for #1 and #2 positions, respectively.

The good news is that Oracle Advanced Analytics supports both languages—SQL and R. There are legions of developers who know SQL for data management and Oracle provides support for data mining and advanced analytics via Oracle Advanced Analytics' SQL data mining functions and provides tight, industry leading integration with open source R statistical programming language.

Most Oracle customers are very familiar with SQL as a language for query, reporting, and analysis of structured data. It is the de facto standard for analysis and the technology that underlies most BI tools. R is a widely popular

Results i

Extraction by

Line Word

Search text

algorithms run as native SQL functions; not as PL/SQL

scripts, call-outs or extensibility framework add-ins. Models are

first class database objects that can be built, applied,

shared, and audited. In the early 2000's, starting in Oracle

Data Mining Release 9.2i,

algorithms, are based on counting principles. They count everything very quickly and then assemble conditional

probability predictive models-all 100% inside the database. Neither the data, the predictive models nor the results

ever leave the database.

OAA Naive Bayes algorithm can quickly build predictive models to predict e.g., "Who will churn?", "Which

Table extraction

Partnership Agreement Example

THIS PARTNERSHIP AGREEMENT is made this _____ day of _____, 20____, by and between the following individuals: Address: _____
City/State/ZIP: _____ Address: _____
City/State/ZIP: _____

1. Nature of Business. The partners listed above hereby agree that they shall be considered partners in business for the following purpose:

2. Name. The partnership shall be conducted under the name of _____ and shall maintain offices at [STREET ADDRESS], [CITY, STATE, ZIP]. 3. Day-To-Day Operation. The partners shall provide their full-time services and best efforts on behalf of the partnership. No partner shall receive a salary for services rendered to the partnership. Each partner shall have equal rights to manage and control the partnership and its business. Should there be differences between the partners concerning ordinary business matters, a decision shall be made by unanimous vote. It is understood that the partners may elect one of the partners to conduct the day-to-day business of the partnership; however, no partner shall be able to bind the partnership by act or contract to any liability exceeding \$ _____ without the prior written consent of each partner.

4. Capital Contribution. The capital contribution of each partner to the partnership shall consist of the following property, services, or cash which each partner agrees to contribute:

Name of Partner	Agreed Upon Cash	% Share
Sample partner	1 Million USD	20%
Another partner	1.5 Million USD	23%

The partnership shall maintain a capital account record for each partner; should any partner's capital account fall below the agreed to amount, then that partner shall (1) have his share of partnership profits then due and payable applied instead to his capital account; and (2) pay any deficiency to the partnership if his share of partnership profits is not yet due and payable or, if it is, his share is insufficient to cancel the deficiency.

5. Profits and Losses. The profits and losses of the partnership shall be divided by the partners according to a mutually agreeable schedule and at the end of each calendar year according to the proportions listed above. 6. Term/Termination. The term of this Agreement shall be for a period of ____ years, unless the partners mutually agree in writing to a shorter period. Should the partnership be terminated by unanimous vote, the assets and cash of the partnership shall be used to pay all creditors, with the

Results ⓘ

Table results: (1)

Table 1

Search text

Name of Partner	Agreed Upon Cash	% Share
Sample partner	1 Million USD	20%
Another partner	1.5 Million USD	23%

Document classification

1 THIS FIRST BANK

FIRST CHOICE ACCOUNT

MANDARIN BRANCH
4444 THIS STREET
ANYTOWN, STATE 00000-0000

CUSTOMER SERVICE 24 HOURS A DAY, 888-000-0000

JOHN Q. CUSTOMER
1234 MAIN STREET
ANYTOWN, STATE 00000-0000

ACCOUNT
12345-678910
STATEMENT PERIOD
1-1-2000 TO 2-1-2000

THANK YOU FOR BANKING WITH THISFIRST

SUMMARY OF YOUR ACCOUNTS

CHECKING	SAVINGS
BEGINNING BALANCE	500.000
DEPOSITS	538.68
WITHDRAWALS	629.73
SERVICE CHARGES/FEEES	6.00
ENDING BALANCE	794.57
MIN BAL ON 2-1-00	54.87

CHECKING ACTIVITY

DEPOSITS POSTED	AMOUNT	DESCRIPTION
1-05	230.66	DEPOSIT
1-15	765.50	DEPOSIT

WITHDRAWALS CKNO	PAID	AMOUNT
106	1-16	632.75
107	1-24	35.98
108	1-26	72.43

POSTED

AMOUNT	DESCRIPTION	
1-21	40.00	ATM WITHDRAWAL
1-22	20.00	ATM WITHDRAWAL

CHECKING SERVICES CHARGE AND FEE SUMMARY

AMOUNT	DESCRIPTION
6.00	MONTHLY SERVICE CHARGE

Results ⓘ

Label	Confidence
BANK STATEMENT	100.00%
OTHERS	0.00%
INVOICE	0.00%
PAYSLIP	0.00%
RESUME	0.00%

- ▶ Request
- ▶ Response

How does it work?

How do we know that this is an invoice?

- By looking at the documents **structure**
- By looking at the **text** within the document

GASQ Service GmbH • Rothenburger Str. 11 • 90443 Nuremberg

Mr. Fabian Neureiter
Wilhelm-Stumpf-Straße 84
44789 Bochum

GASQ Service GmbH
Rothenburger Str. 11
90443 Nuremberg
Germany
Tel. +49 911 990078 0
Fax +49 911 990078 99

info@gasq.org

Nuremberg, 21.06.2022

Invoice No. 67508

Reference: [REDACTED]

Dear Mr. Fabian Neureiter

We are invoicing you according to your order as follows:

Pos	Description	Amount	Unit Price	Net Price	VAT (19%)	Gross Price
1	ISTQB CTFL E-Exam Exam Date: 07/25/2022 Exam Place: Düsseldorf Participant: Fabian Neureiter	1	215,00 €	215,00 €	40,85 €	255,85 €
Total				215,00 €	40,85 €	255,85 €

The exam fee has been successfully charged to your credit card.

If you should have any further questions, please do not hesitate to contact us.

Sincerely,
Oliver Braun

ML Technologies working in the Background

Document Understanding

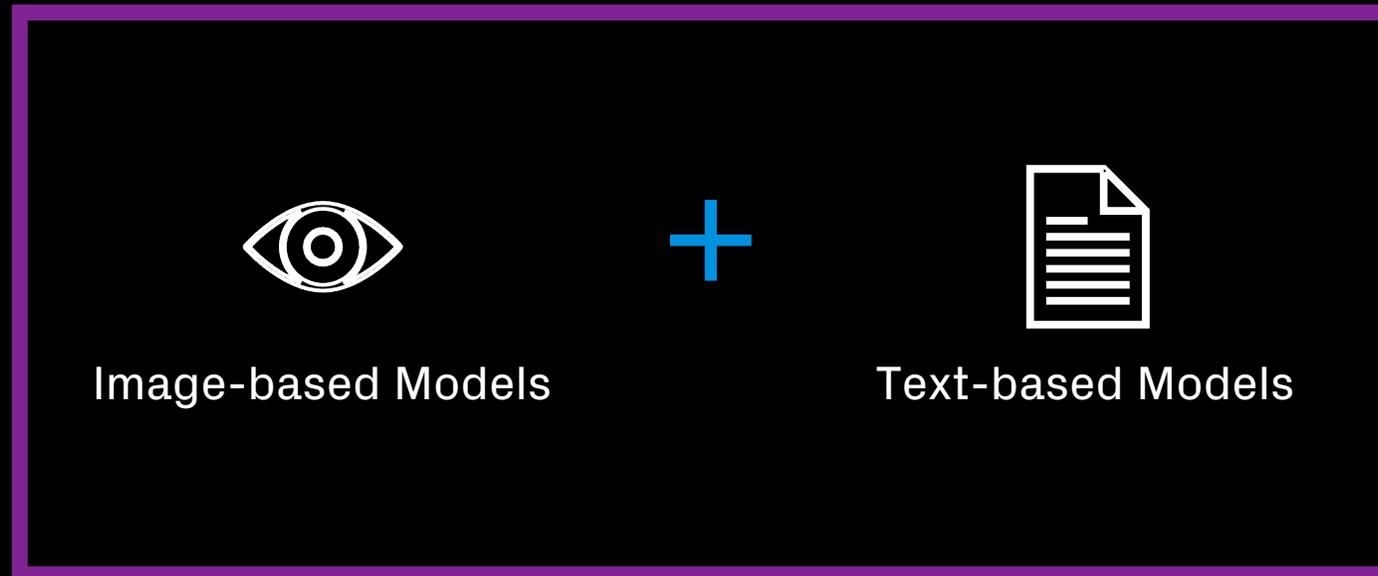


Image-based models

- Have learned an **abstract representation** of a document type
- Classify based on the learned **abstraction**

GASQ Service GmbH • Rothenburger Str. 11 • 90443 Nuremberg

Mr. Fabian Neureiter
Wilhelm-Stumpf-Straße 84
44789 Bochum

GASQ Service GmbH
Rothenburger Str. 11
90443 Nuremberg
Germany
Tel. +49 911 990078 0
Fax +49 911 990078 99

info@gasq.org

Nuremberg, 21.06.2022

Invoice No. 67508

Reference: [REDACTED]

Dear Mr. Fabian Neureiter

We are invoicing you according to your order as follows:

Pos	Description	Amount	Unit Price	Net Price	VAT (19%)	Gross Price
1	ISTQB CTFL E-Exam Exam Date: 07/25/2022 Exam Place: Düsseldorf Participant: Fabian Neureiter	1	215,00 €	215,00 €	40,85 €	255,85 €
Total				215,00 €	40,85 €	255,85 €

The exam fee has been successfully charged to your credit card.

If you should have any further questions, please do not hesitate to contact us.

Sincerely,
Oliver Braun

Example: Document Classification

Hybrid models

Optical Character Recognition (OCR)

- Recognize and extract **words** and/or **lines**
- Works on digitally created and handwritten documents

Nuremberg GASQ Service GmbH Mr. Fabian Neureiter

Rothenburger Str. 11 Wilhelm-Stumpf-Strae 84 90443 Nuremberg

44789 Bochum Germany Tel. +49 911 990078 0 Fax +49

911 990078 99 Oliver Braun info@gasq.org

Nuremberg, 21.06.2022 Invoice No. 67508 Reference: SPR1882

Dear Mr. Fabian Neureiter We are invoicing you according to

your order as follows:

Pos	Description	Amount	Unit	Price
Net	Price	VAT (19%)	Gross	Price
1	ISTQB CTFL E-Exam	1		
1	215,00 €	215,00 €	40,85 €	255,85 €
	Exam	Date:		
07/25/2022	Exam	Place: Dsseldorf	Participant: Fabian Neureiter	
Total	215,00 €	40,85 €	255,85 €	The exam fee has
been	successfully	charged	to	your credit card. If you should
have	any	further	questions,	please do not hesitate to
contact	us.	Sincerely,	Oliver Braun	GASQ Service GmbH

Example: Document Classification

Text-based models

- Detect **contextual meanings** of extracted words
- Classify based on the learned **context** of the words used in an invoice

Nuremberg GASQ Service GmbH Mr. Fabian Neureiter
Rothenburger Str. 11 Wilhelm-Stumpf-Strae 84 90443 Nuremberg
44789 Bochum Germany Tel. +49 911 990078 0 Fax +49
911 990078 99 Oliver Braun info@gasq.org
Nuremberg, 21.06.2022 Invoice No. 67508 Reference: SPR1882
Dear Mr. Fabian Neureiter We are invoicing you according to
your order as follows: Pos Description Amount Unit Price
Net Price VAT (19%) Gross Price 1 ISTQB CTFL E-Exam
1 215,00 € 215,00 € 40,85 € 255,85 € Exam Date:
07/25/2022 Exam Place: Dsseldorf Participant: Fabian Neureiter
Total 215,00 € 40,85 € 255,85 € The exam fee has
been successfully charged to your credit card. If you should
have any further questions, please do not hesitate to
contact us. Sincerely, Oliver Braun GASQ Service GmbH

How can we use it

OCI Set Up

How can we use it? - OCI Set Up

OCI Instance Set Up

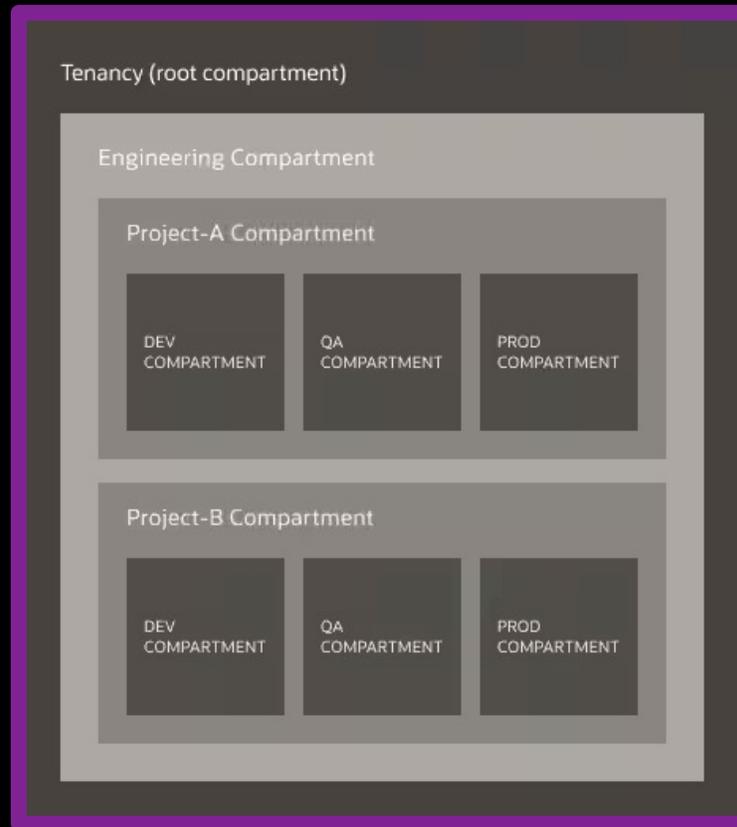
1. Upgrade to paid instance
2. Create a compartment
3. Create a group
4. Create Buckets in Object Storage
5. Create Policies

Upgrade to paid instance – Pricing (21.04.2024)

- Pay-as-you-go structure
- Amount is calculated on a **transaction** basis
 - Transaction = Amount of operations per page of the document
 - Billing is per 1000 Transactions
 - 5000 Transactions free
- Difference between individual and corporate accounts

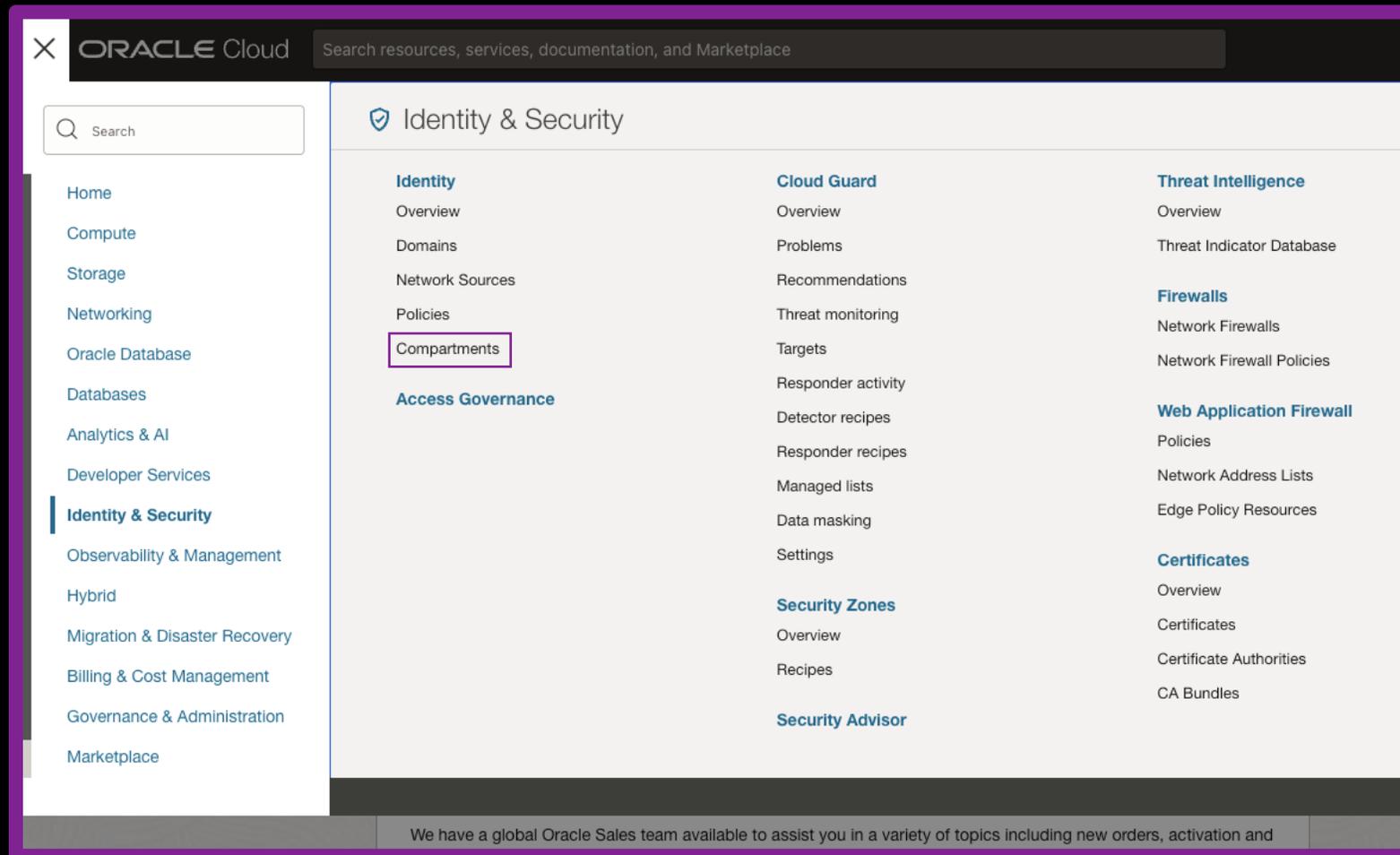
How can we use it? - OCI Set Up

Creating a compartment – what are compartments



How can we use it?- OCI Set Up

Creating a compartment - How to



How can we use it?- OCI Set Up

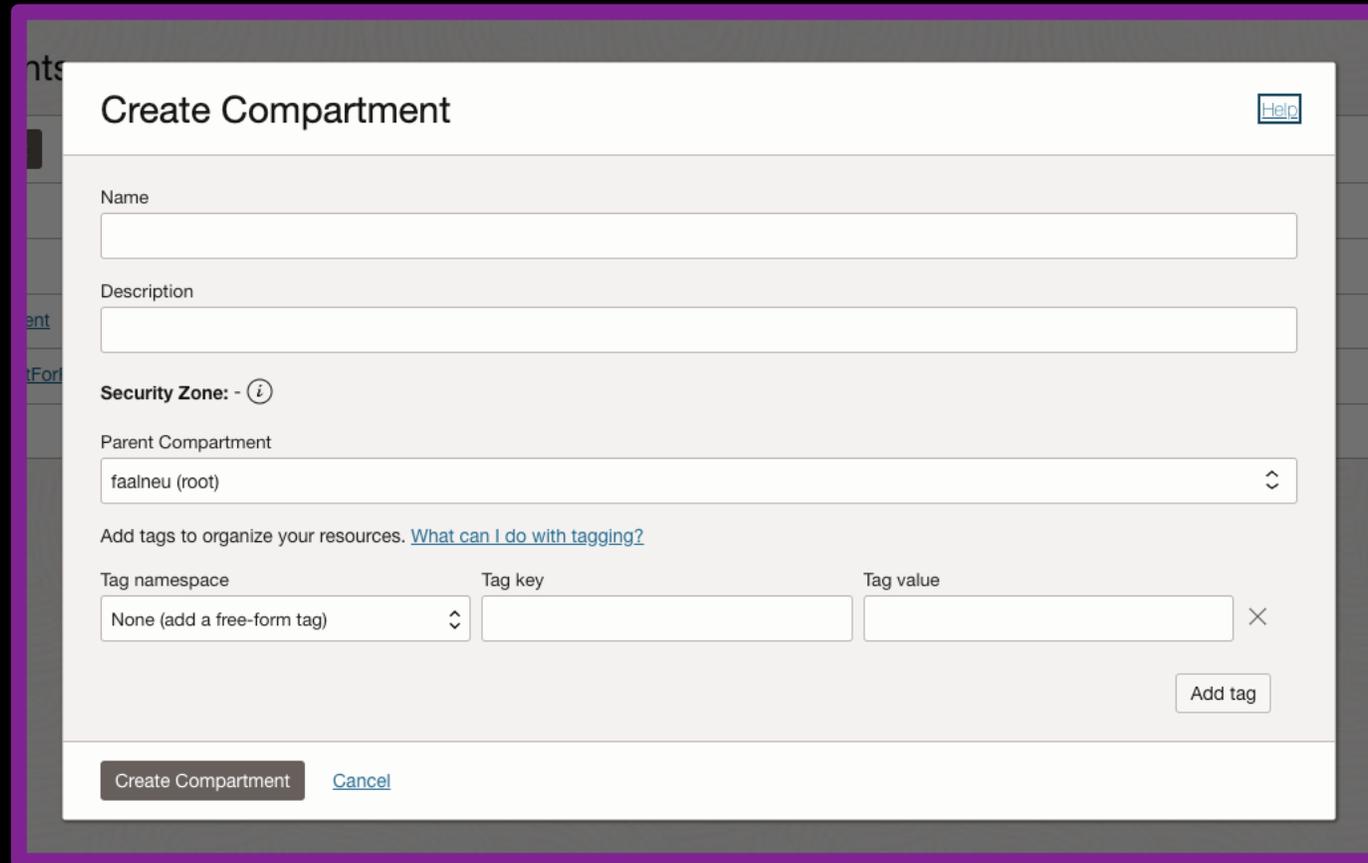
Creating a compartment - How to

The screenshot shows the Oracle Cloud console interface for managing compartments. On the left, there is a navigation menu with options like Overview, Domains, Network Sources, Policies, and Compartments (which is currently selected). Below the menu, there are filter options for State (set to 'Active | Deleting') and Tag filters (with 'add | clear' links). The main content area is titled 'Compartments' and features a 'Create Compartment' button at the top left, which is highlighted with a red box. Below the button is a table listing existing compartments.

Name	Status	OCID	Authorized	Security Zone ⓘ	Subcompartments
faalneu	● Active	...lkjxja	Yes	-	2
ai_services_compartment	● Active	...bpxixq	Yes	-	0
ManagedCompartmentForPaaS	● Active	...n4zvfq	Yes	-	0

How can we use it?- OCI Set Up

Creating a compartment - How to

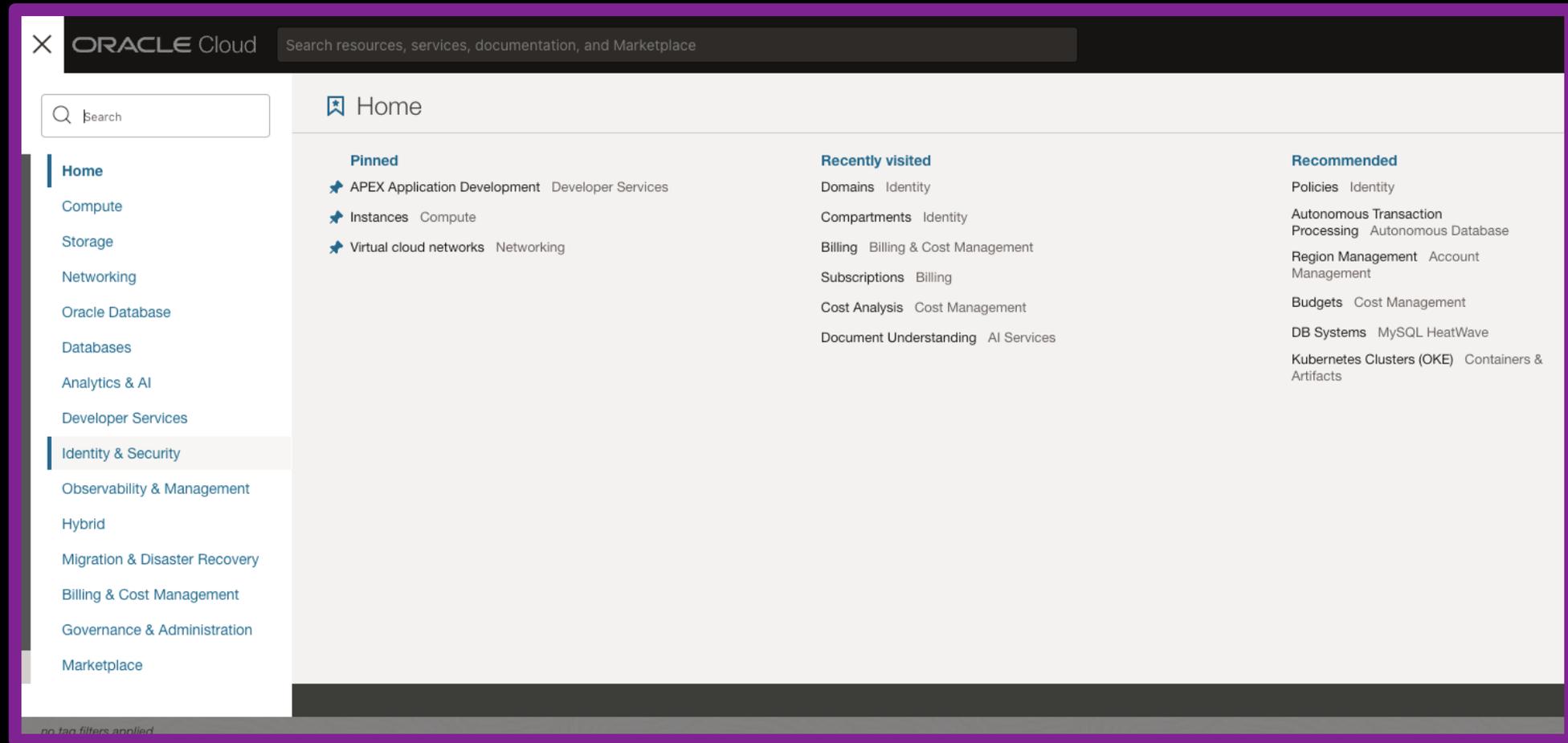


The screenshot shows the 'Create Compartment' dialog box with the following fields and options:

- Name:** An empty text input field.
- Description:** An empty text input field.
- Security Zone:** A dropdown menu with a help icon (i) and the text 'Security Zone: - (i)'. It is currently set to 'None (add a free-form tag)'.
- Parent Compartment:** A dropdown menu showing 'faalneu (root)' with a double arrow icon.
- Tagging:** A section with the text 'Add tags to organize your resources. [What can I do with tagging?](#)' and three input fields: 'Tag namespace' (set to 'None (add a free-form tag)'), 'Tag key' (empty), and 'Tag value' (empty). An 'Add tag' button is located to the right of these fields.
- Buttons:** At the bottom, there are two buttons: 'Create Compartment' and 'Cancel'.

How can we use it?- OCI Set Up

Creating a group



How can we use it?- OCI Set Up

Creating a group

The screenshot shows the Oracle Cloud Identity console interface. The breadcrumb navigation is 'Identity > Domains > Default domain > Groups'. The main heading is 'Groups in Default Domain'. A search bar is present with the placeholder text 'Search by group name or description.'. Below the search bar, there is a 'Create group' button and a 'More actions' dropdown menu. A table lists the existing groups:

<input type="checkbox"/>	Name	Description	Created
<input type="checkbox"/>	ai_services_users	-	Fri, Apr 19, 2024, 20:00:39 UTC
<input type="checkbox"/>	All Domain Users	A group representing all users.	Fri, Nov 17, 2023, 12:58:51 UTC
<input type="checkbox"/>	Administrators	Administrators	Fri, Nov 17, 2023, 12:58:46 UTC

At the bottom of the table, it says '0 selected'. On the left sidebar, the 'Groups' menu item is highlighted. At the bottom of the page, there is a 'Tag filters' section with 'add | clear' links and the text 'no tag filters applied'.

How can we use it?- OCI Set Up

Creating a group

The screenshot shows the Oracle Cloud console interface for creating a group. The breadcrumb navigation is Identity > Domains > Default domain > Groups. The left sidebar shows the navigation menu with 'Groups' selected. The main content area is titled 'Create group' and contains the following form fields:

- Name:** A required text input field.
- Description:** A text input field.
- User can request access:** A checkbox.
- Users:** A section titled 'Optional' with the instruction 'Select users to assign this group.' It includes a search bar and a table of users.

<input type="checkbox"/>	First name	Last name	Email
<input type="checkbox"/>	Fabian	Neureiter	fa.al.neu@gmail.com

At the bottom of the form, there are 'Create' and 'Cancel' buttons.

Using Domains– Good Practice

- Learn from my mistake, **don't** create the group in the default domain
- Create a separate domain for everything not administration related
- For more information:

<https://blogs.oracle.com/ateam/post/oci-iam-identity-domains-best-practices>

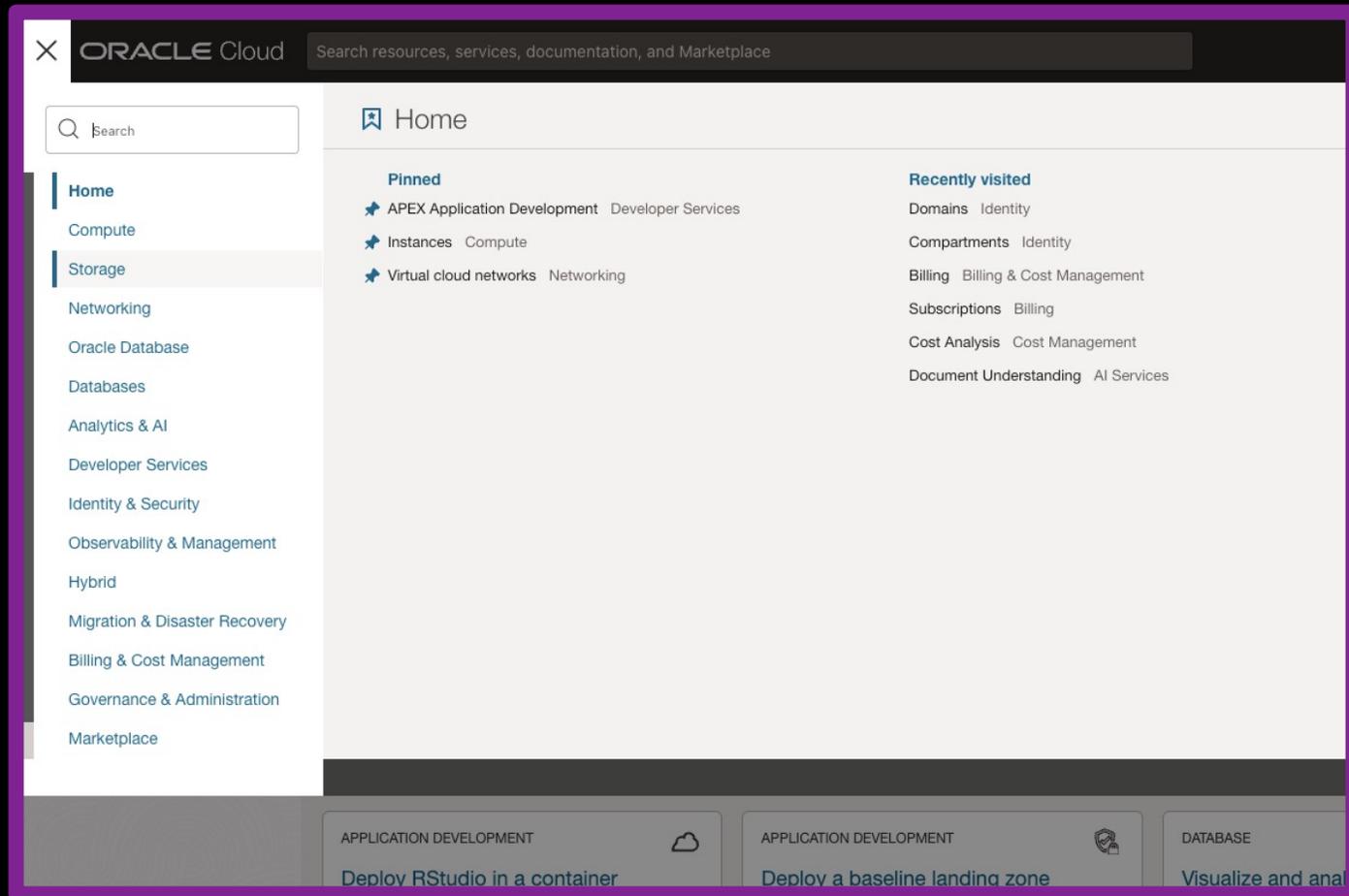
How can we use it?- OCI Set Up

Creating Buckets in Object Storage

- Used for:
 - Storing documents we want to process
 - The output of document analysis

How can we use it?- OCI Set Up

Creating Buckets in Object Storage



How can we use it?- OCI Set Up

Creating Buckets in Object Storage

The screenshot displays the Oracle Cloud console interface. On the left, the 'Object Storage & Archive' section is visible, showing a list of buckets under the 'ai_services_comp' compartment. The 'Create Bucket' dialog box is the central focus, with the following configuration:

- Bucket Name:** bucket-20240421-1310
- Default Storage Tier:** Standard (selected), Archive (unselected)
- Encryption:** Encrypt using Oracle managed keys (selected), Encrypt using customer-managed keys (unselected)
- Resource logging:** Disabled (toggle switch)

At the bottom of the dialog, there are 'Create' and 'Cancel' buttons. The background page shows a list of buckets with names like 'document_analysis_output' and 'documents_input'.

How can we use it?- OCI Set Up

Set up policies

The screenshot displays the Oracle Cloud console interface. At the top, there is a search bar with the text "Search resources, services, documentation, and Marketplace". Below this, a navigation sidebar on the left lists various categories: Home, Compute, Storage, Networking, Oracle Database, Databases, Analytics & AI, Developer Services, Identity & Security (highlighted), Observability & Management, Hybrid, Migration & Disaster Recovery, Billing & Cost Management, Governance & Administration, and Marketplace. The main content area is titled "Home" and is divided into three columns: "Pinned", "Recently visited", and "Recommended".

- Pinned:**
 - APEX Application Development Developer Services
 - Instances Compute
 - Virtual cloud networks Networking
- Recently visited:**
 - Domains Identity
 - Compartments Identity
 - Billing Billing & Cost Management
 - Subscriptions Billing
 - Cost Analysis Cost Management
 - Document Understanding AI Services
- Recommended:**
 - Policies Identity
 - Autonomous Transaction Processing Autonomous Database
 - Region Management Account Management
 - Budgets Cost Management
 - DB Systems MySQL HeatWave
 - Kubernetes Clusters (OKE) Containers & Artifacts

At the bottom left of the console, a small text indicator reads "no tag filters applied".

How can we use it?- OCI Set Up

Set up policies

The screenshot displays the Oracle Cloud console interface for a specific policy. At the top, the Oracle Cloud logo and a search bar are visible. The breadcrumb navigation shows 'Identity > Policies > Policy Details'. The main content area features a large green circle with a white 'P' representing the policy, with the status 'ACTIVE' below it. To the right of the circle are buttons for 'Edit Policy', 'Add tags', and 'Delete'. Below these buttons are two tabs: 'Policy Information' (selected) and 'Tags'. The 'Policy Information' tab contains the following details: OCID: ...msp5ctda (with 'Show' and 'Copy' links), Compartment: faalneu (root), Description: policy for things related to document understanding, and Created: Fri, Apr 19, 2024, 20:15:46 UTC. On the left side, there is a 'Resources' section with a sub-section for 'Statements'. The 'Statements' section contains an 'Edit Policy Statements' button and a list of three policy statements: 'allow group ai_services_users to use ai-service-document-family in tenancy', 'allow group ai_services_users to use object-family in compartment ai_services_compartment', and 'allow group ai_services_users to manage object-family in compartment ai_services_compartment'.

How can we use it

Ways of using the document understanding service

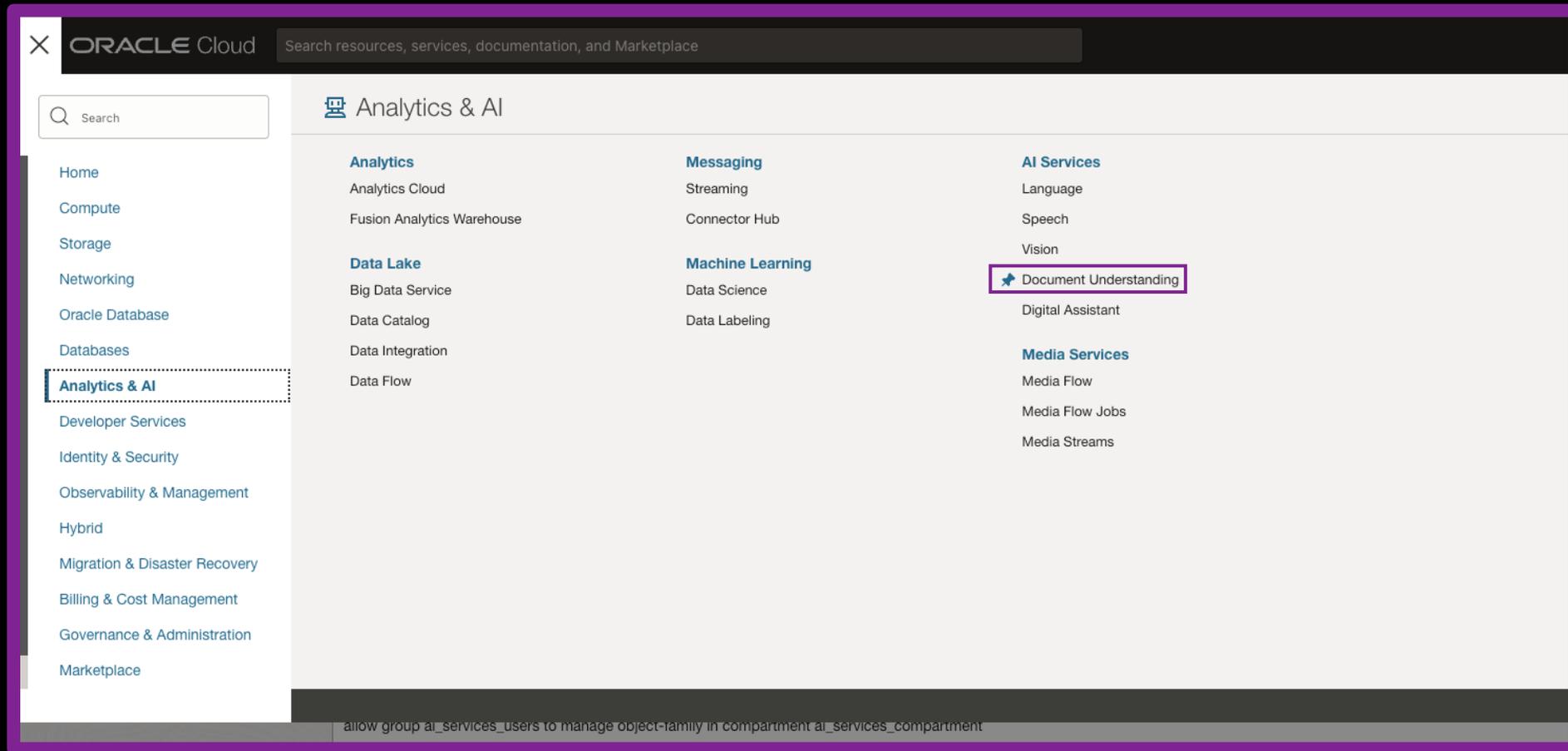
How can we use it?- Ways of using the document understanding service

Ways of using the document understanding service

- OCI UI/Cloud Console
- CLI
- SDKs for various languages
- REST API

How can we use it?- Ways of using the document understanding service

OCI UI / Cloud Console



How can we use it?- Ways of using the document understanding service

OCI UI / Cloud Console

The screenshot displays the OCI Document Understanding console interface. The main area shows a document classification task for a bank statement. The document source is set to 'Demo files', and the selected file is 'Statement'. The document content is displayed in a preview window, showing a bank statement from '1 THIS FIRST BANK' for a 'FIRST CHOICE ACCOUNT'. The statement includes the address 'MANDARIN BRANCH, 4444 THIS STREET, ANYTOWN, STATE 00000-0000', customer information for 'JOHN Q. CUSTOMER', and account details 'ACCOUNT 12345-678910, STATEMENT PERIOD 1-1-2000 TO 2-1-2000'. A table at the bottom shows a 'BEGINNING BALANCE' of 500.000.

Document Understanding

Document classification

Categorize a document as a predefined type such as resume, invoice, or receipt

Document source
 Demo files Local files Object storage

Demo files: Statement Demo files: Receipt

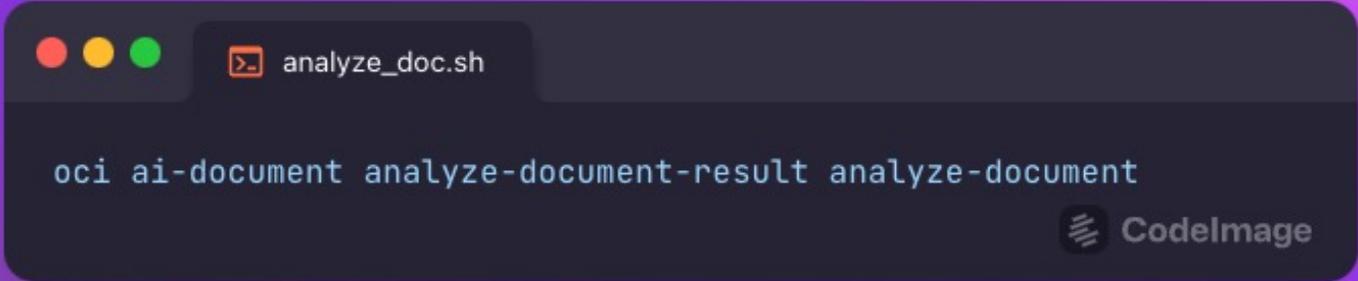
Results

Label	Confidence
BANK STATEMENT	100.00%
OTHERS	0.00%
INVOICE	0.00%
PAYSLIP	0.00%
RESUME	0.00%

► Request
► Response

How can we use it?- Ways of using the document understanding service

CLI – Command Line Interface



```
analyze_doc.sh  
oci ai-document analyze-document-result analyze-document
```

Codelmage

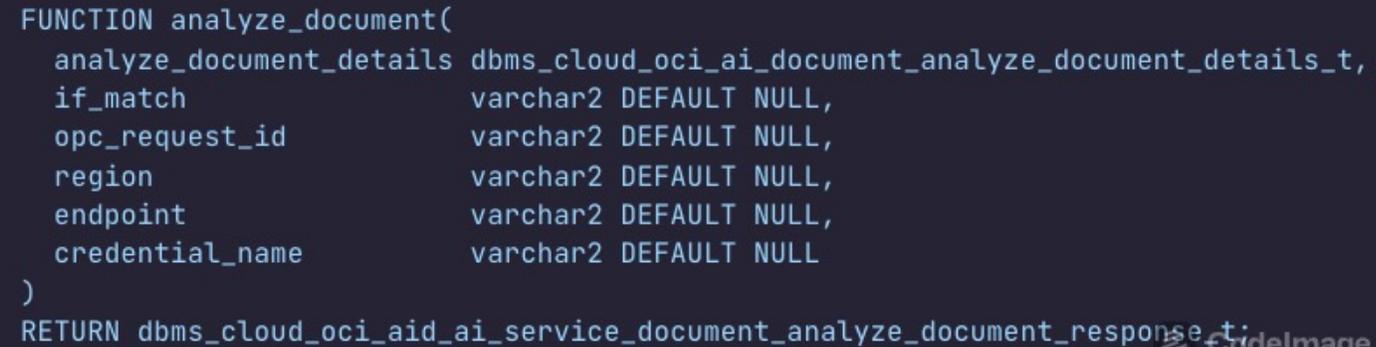
How can we use it?- Ways of using the document understanding service

Various SDKs (Software Development Kits)

- Java
- Python
- JavaScript and Typescript
- .NET
- Go
- Ruby
- PL/SQL

How can we use it?- Ways of using the document understanding service

Various SDKs



```
FUNCTION analyze_document(  
  analyze_document_details dbms_cloud_oci_ai_document_analyze_document_details_t,  
  if_match                 varchar2 DEFAULT NULL,  
  opc_request_id          varchar2 DEFAULT NULL,  
  region                  varchar2 DEFAULT NULL,  
  endpoint                varchar2 DEFAULT NULL,  
  credential_name         varchar2 DEFAULT NULL  
)  
RETURN dbms_cloud_oci_aid_ai_service_document_analyze_document_response_t;
```

How can we use it?- Ways of using the document understanding service

REST API

The screenshot shows the Oracle REST API Explorer interface for the `AnalyzeDocument` endpoint. On the left is a navigation tree with the following items: Data Labeling Management API, Data Safe API, Data Science API, Database Management API, Database Migration API, Database Service API, Database Tools, DevOps API, Digital Assistant Service Instance API, DNS API, Document Understanding API (expanded), Document Understanding API Home, AnalyzeDocumentResult (expanded), AnalyzeDocumentResult Reference, **AnalyzeDocument** (selected), Model, ProcessorJob, Project, WorkRequest, WorkRequestError, WorkRequestLogEntry, Data Types, EM Warehouse API, and Email Delivery API.

The main content area displays the following details for the `AnalyzeDocument` endpoint:

- Endpoint:** `POST /20221109/actions/analyzeDocument` (with a tag `AISERVICEDOCUMENT`)
- Description:** Perform different types of document analysis.
- Request:** A text input field containing `AnalyzeDocument`.
- Parameters Table:**

Name	Where	Description
if-match	header	<ul style="list-style-type: none">Required: noType: string <p>For optimistic concurrency control. In the PUT or DELETE call for a resource, set the <code>if-match</code> parameter to the value of the etag from a previous GET or POST response for that resource. The resource will be updated or deleted only if the etag you provide matches the resource's current etag value.</p>
opc-request-id	header	<ul style="list-style-type: none">Required: noType: string <p>The client request ID for tracing.</p>
- Body:** (Section header, content not visible)

The Document Understanding REST API

A closer look

Synchronous

- Make **POST** Request to the **analyzeDocument** Endpoint with the document attached
- Directly receive response with the result of the analysis

Asynchronous

- Make **POST** Request to the **createProcessorJob** Endpoint with the document attached or an object storage location of the document
- Use GET Requests to **GetProcessorJob** to query the status of the Processorjob
- Retrieve Results via the **Object Storage Service API**

Generate API Keys

The screenshot displays the Oracle Cloud dashboard interface. At the top, there is a navigation bar with the Oracle Cloud logo, a search bar, and the region 'Germany Central (Frankfurt)'. Below the navigation bar, the dashboard is divided into several sections:

- Service links:** A section with three columns: 'PINNED' (Document Understanding AI Services, APEX Application Development Developer Services, Instances Compute, Virtual cloud networks Networking), 'RECENTLY VISITED' (Identity Identity & Security, Policies Identity, Buckets Object Storage & Archive Storage, Domains Identity, Compartments Identity, Billing Billing & Cost Management), and 'RECOMMENDED' (Overview Identity, Dedicated Virtual Machine Hosts Compute, Autonomous Transaction Processing Autonomous Database, Logging Logging, Kubernetes Clusters (OKE) Containers & Artifacts, Region Management Account Management).
- Quickstarts:** A section with a 'View my deployments' button and six quickstart cards: 'Predict the result of the next race' (25-30 mins), 'Deploy a WordPress website' (6-8 mins), 'Deploy a low-code app on Autonomous Database using APEX' (3-5 mins), 'Deploy RStudio in a container' (10-12 mins), 'Deploy a baseline landing zone' (7-9 mins), and 'Visualize and analyze Strava data on Autonomous Database' (2-4 mins).
- Usage and Profile:** A right-hand sidebar showing 'Your tenancy faalneu', 'All services operational', 'Usage' (Analyze costs, Subscription 30281945, Pay As You Go, €0.00, 20 of 30 days elapsed), 'Cost savings opportunities' (Estimated savings: 0), 'Experience previews', 'Marketplace', and 'OCI mobile app'. A 'Profile' dropdown menu is also visible, showing 'Default/fa.al.neu@gmail.com', 'Identity domain: Default', 'My profile', 'Tenancy: faalneu', 'Service user Console', 'Console settings', and 'Sign out'.

Generate API Keys

The screenshot shows the Oracle Cloud Identity console for user Fabian Neureiter. The user profile is active and includes fields for OCID, Username, Email, and various contact information. The API keys section is currently empty, with an 'Add API key' button highlighted by a red box. The breadcrumb trail indicates the path: Identity > My profile > API keys.

Oracle Cloud Search resources, services, documentation, and Marketplace Germany Central (Frankfurt) [dropdown] [refresh] [notifications] [help] [globe] [profile]

Identity > My profile > API keys [profile icon]

Fabian Neureiter

[Edit my profile](#) [Change password](#) [Link support account](#) [Security](#) [More actions](#) [dropdown]

User information | Tags | Capabilities

OCID: ...xq4d6q [Show](#) [Copy](#) **Email:** fa.al.neu@gmail.com

Username: fa.al.neu@gmail.com **Instant messaging address:** -

Prefix: - **Home phone number:** -

First name: Fabian **Mobile phone number:** -

Middle name: - **Federated:** No

Last name: Neureiter **My Oracle Support Account:** -

Suffix: - **Authenticated by:** IDCSApp

> [User preferences](#)

> [Work information](#)

> [Other information](#)

API keys

[Add API key](#) [Delete](#)

<input type="checkbox"/>	Fingerprint	Created
No items found.		

0 selected No API key to display < Page 1 >

Generate API Keys

The screenshot displays the Oracle Cloud Identity console interface. On the left, the user profile for Fabian Neureiter is visible, including a green circular profile picture with the letter 'U' and the status 'ACTIVE'. Below the profile, there are sections for 'User information' (OCID, Username, Prefix, First name, Middle name, Last name, Suffix) and 'API keys' (with 'Add API key' and 'Delete' buttons). The main area is titled 'Add API key' and contains a note: 'An API key is an RSA key pair in PEM format used for signing API requests. You can generate the key pair here and download the private key. If you already have a key pair, you can choose to upload or paste your public key file instead. [Learn more](#)'. There are three radio buttons: 'Generate API key pair' (selected), 'Choose public key file', and 'Paste a public key'. Below this, a 'Public key' section contains an information icon and a message: 'Download the private key. It will not be shown again. After you download it, [change the file permissions](#) so only you can view it'. Two download buttons are present: 'Download private key' and 'Download public key'. At the bottom of the form, there are 'Add' and 'Cancel' buttons.

Generate API Keys

Configuration file preview Help

Note: This configuration file snippet includes the basic authentication information you'll need to use the SDK, CLI, or other OCI developer tool. Paste the contents of the text box into your `~/.oci/config` file and update the `key_file` parameter with the file path to your private key. [Learn more](#)

Select API key fingerprint

ef: [REDACTED]

Configuration file preview *Read-only*

```
[DEFAULT]
user=ocid1.user.oc1.[REDACTED]
fingerprint=ef:[REDACTED]
tenancy=ocid1.tenancy.[REDACTED]
region=eu-frankfurt-1
key_file=<path to your private keyfile> # TODO
```

[Copy](#)

Paste the contents of the text box into your `~/.oci/config` file.

Close

LIVE Demo

Using the API within APEX

(Current) limits of the service

(Current) Limits of the service

Biggest and most significant drawback

- Currently, no languages other than English are supported (!)
- Sadly makes OCI Document Understanding not usable for our Use Case



https://www.reddit.com/r/memes/comments/z6bsm1/germany_sad/

(Current) Limits of the service

Workaround by Dirk Strack

The screenshot shows the GitHub repository page for 'Document-AI-Scanner' by user 'dstrack'. The repository is public and has 3 watchers, 1 fork, and 3 stars. The main branch is 'main' with 1 branch and 0 tags. The repository contains several files and folders, including a README, LICENSE, and various source files. The README is currently selected and displays the following content:

Document AI Scanner

Overview

The Document AI Scanner is an Oracle APEX 23.2 application that leverages the Oracle Cloud - Document Understanding AI Service in a modern APEX user interface. The aim of this application is to process several PDF or image files in one job and provide the results of the analysis as a ZIP file for download or export to an accounting database.

The repository also includes a table of recent commits:

Commit	Message	Time
f3a425e	Maintenance for better compatibility:	last month
		27 Commits
	Update issue templates	3 months ago
	improved documentation	2 months ago
	improved dokumentation	2 months ago
	Maintenance for better compatibility:	last month
	remove default parameters from Document_Scan_AI_Con...	2 months ago
	Initial commit	5 months ago
	Update README.md	3 months ago

Alternatives

Cloud Providers and custom services

- Azure Document Intelligence
 - <https://learn.microsoft.com/de-de/azure/ai-services/document-intelligence/overview?view=doc-intel-3.1.0>
- Amazon Textract
 - <https://aws.amazon.com/de/textract/>
- Custom, locally hosted and trained Models

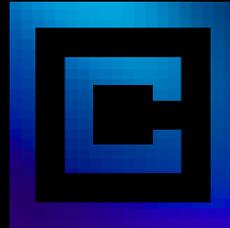
Useful Resources

Useful Resources

- How To and sample app from Oracle:
 - <https://blogs.oracle.com/apex/post/automate-invoice-handling-by-integrating-your-oracle-apex-app-with-oci-document-understanding>
- Service Overview:
 - <https://docs.oracle.com/en-us/iaas/Content/document-understanding/using/home.htm>
- REST API Reference:
 - <https://docs.oracle.com/en-us/iaas/api/#/en/document-understanding/20221109/>
- Document AI Scanner by Dirk Strack
 - <https://github.com/dstrack/Document-AI-Scanner>

Say Hy_

+49 (0) 531 23767-0
info@hyand.com



© 2024 – The developed thoughts and ideas are the intellectual property of Hyand and are subject of copyright law. Reproduction, transfer to third parties or use – even of parts – is only permitted with the express of Hyand.