Infosys
Responsible AI Office

Infosys Responsible AI Toolkit - Privacy Module

API usage Instructions

Contents

About Privacy

The Privacy tenet of *Infosys Responsible AI toolkit* facilitates redact the PII and any sensitive data in an organization by detecting the privacy entities and giving options to analyze, anonymize and encrypt using models like Tesseract, Easy OCR and Computer Vision. We analyze privacy in images, videos, and text by reading through above document types, processing & applying the redacting techniques.

Dependencies

The PRIVADMIN_API environment variable is used when the adminconnection flag is set to True. It should contain the base URL of the Admin service responsible for account and portfolio data mapping. This allows the Privacy module to fetch account-level configurations such as:

- Account Master ID
- Threshold Score
- Associated recognition and encryption metadata

Without this configuration, the module may fall back to default or static data behavior, limiting its ability to dynamically determine privacy settings based on account-level controls.

Features & API End Points

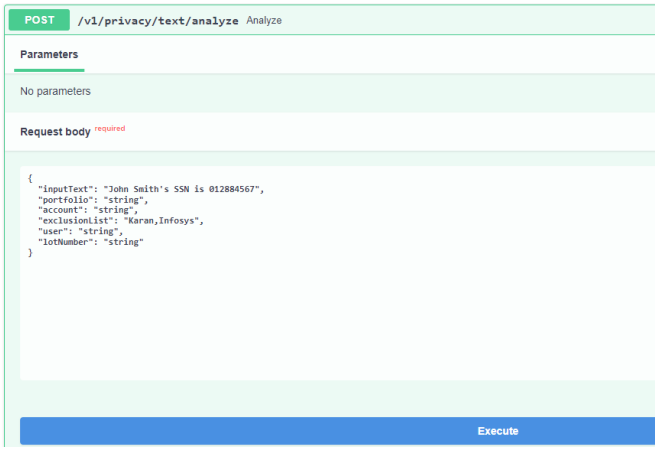
Privacy_main APIs

Text Analyze

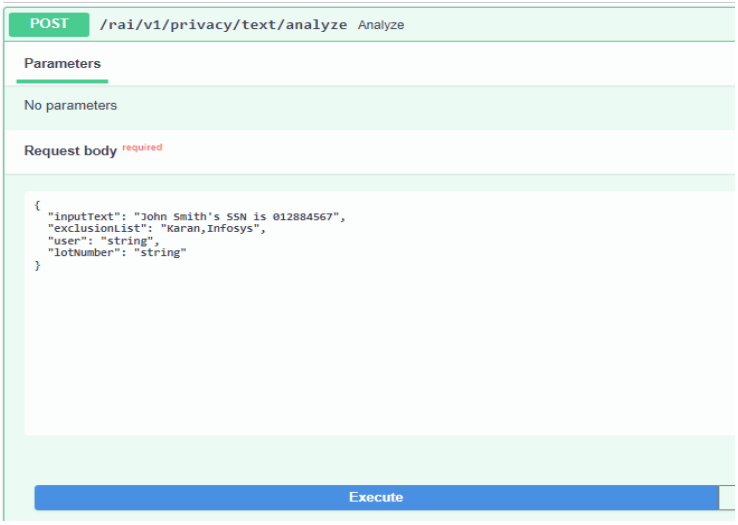
Endpoint – /rai/v1/privacy/text/analyze
Using this API, we can check if the input text contains any PII entities or not.

Input : Replace the input text with the prompt you want to check for PII entities. In exclusion list mention the PII entities which don't need to be blocked. The fields portfolio and account are optional (we can remove these from the Json if these are not needed), and we can create these from the admin portal. The fields user and lotNumber will be allocated at user login if they are using the application.

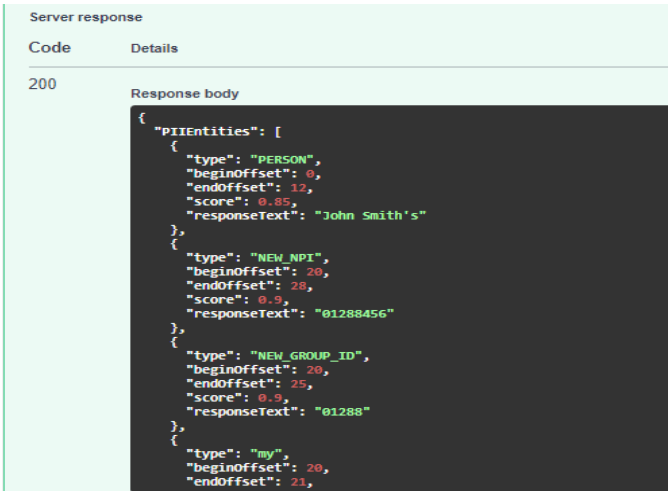
With all fields -



Without optional fields -



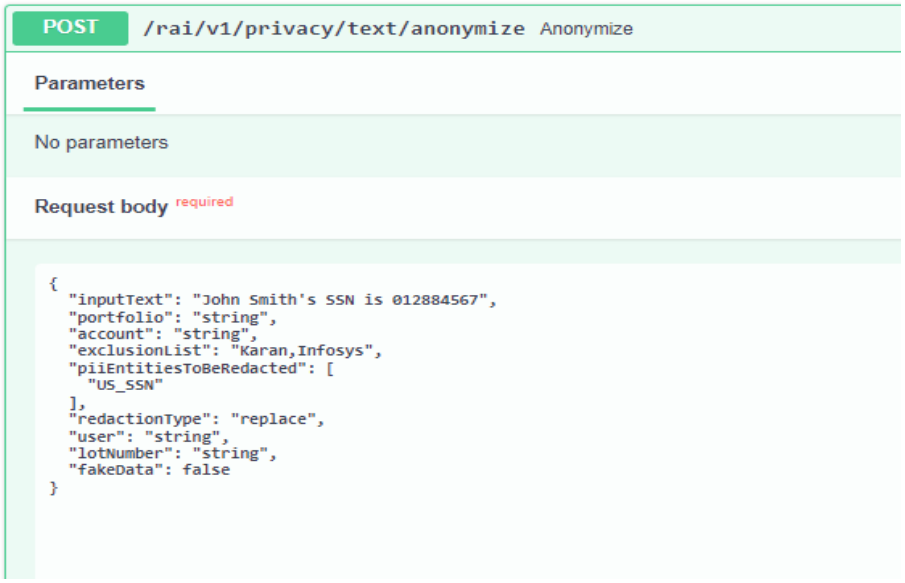
Response: Returns the list of PII Entities detected with position and its score.



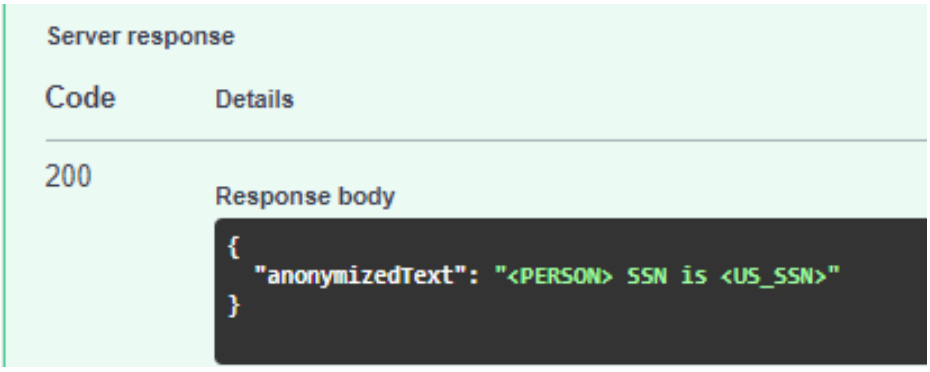
Text Anonymize

Endpoint – /rai/v1/privacy/text/anonymize
Using this API, we can anonymize all PII entities in the input text.

Input : Replace the input text with prompt to be checked, give the PII entities to be redacted in the list. Give fakeData as True if you want to mask the PII detected with fake data. The fields portfolio and account are optional (we can remove these from the Json if these are not needed), and we can create these from the admin portal. The fields user and lotNumber will be allocated at user login if they are using the application. We can give redaction type for the type of anonymization we want and can give PII entities in the list which we want to be redacted. In exclusion list mention the PII values which don't need to be anonymized and identified.



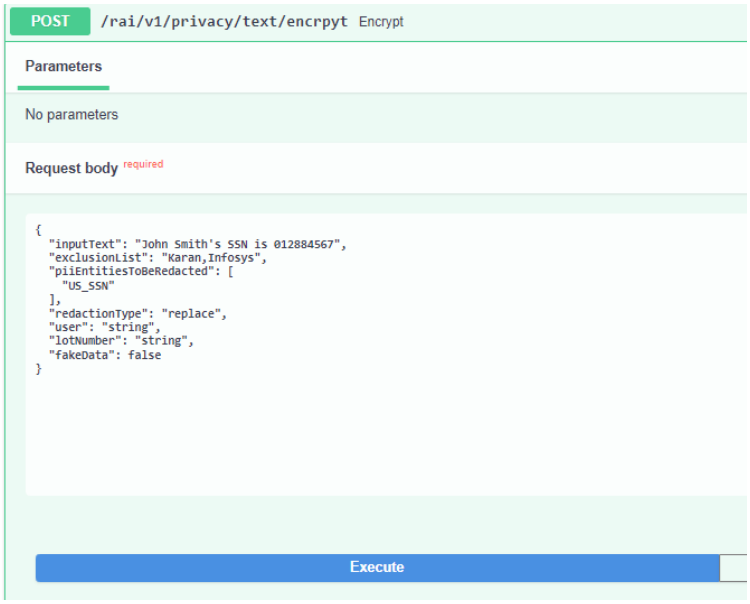
Response : Mask the PII entities detected in the inputText.



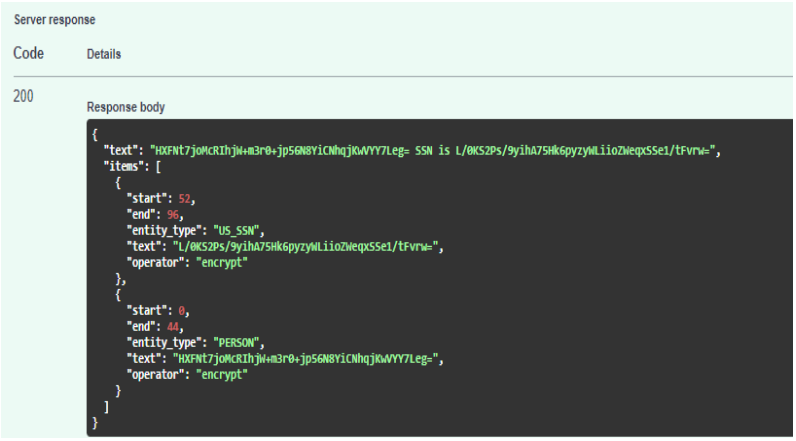
Text Encrpyt

Endpoint – /rai/v1/privacy/text/encrpyt
Using this API, we can encrypt all PII entities in the input text.

Input : Replace the input text with prompt to be checked. Give fakeData as True if you want to mask the PII detected with fake data. The fields portfolio and account are optional (we can remove these from the Json if these are not needed), and we can create these from the admin portal. The fields user and lotNumber will be allocated at user login if they are using the application. We can give redaction type for the type of anonymization we want and can give PII entities in the list which we want to be identified and encrypted.



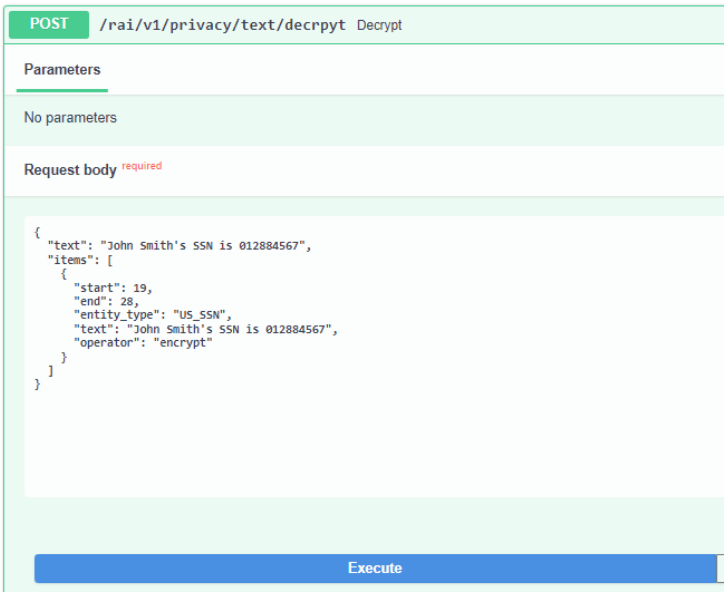
Response : Returns the encrypted text with masked PII entities.



Text Decrypt

Endpoint – /rai/v1/privacy/text/decrpyt
Using this API, we can encrypt PII entities of a specific type at given location in the text.

Input : Replace the input text with prompt to be checked, give the PII entities to be encrypted in the list.



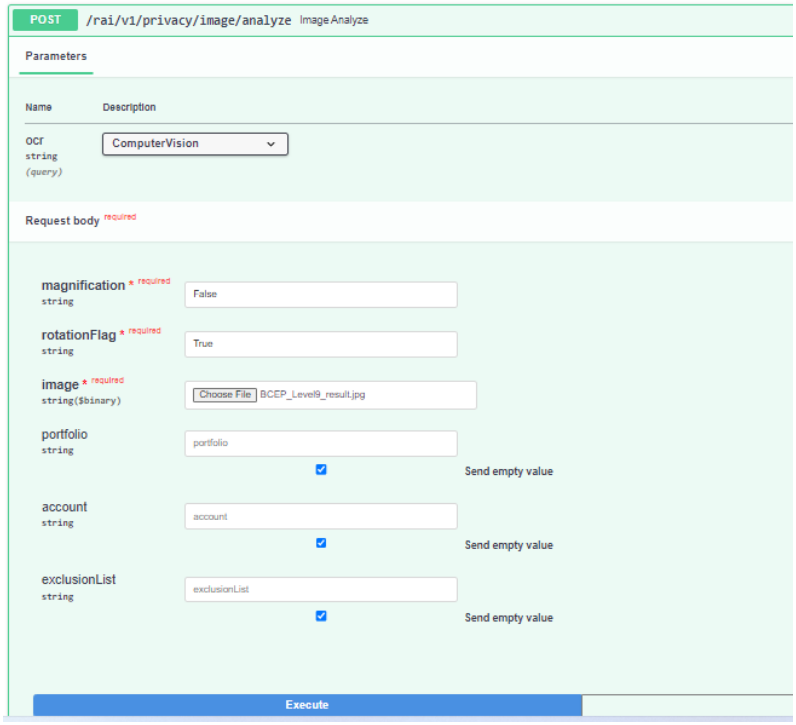
Response: Decrypts the encrypted text and gives the output.

```
{
  "decryptedText": "John Smith's SSN is <US_SSN>"
}
```

Image Analyze

Endpoint – /rai/v1/privacy/image/analyze
Using this API, we can analyze the uploaded image for any PII entities.

Input :
We can select EasyOcr, tesseract or ComputerVision to analyze the image. Upload the image file, if we want the image to be magnified give 'magnification' as 'true' otherwise give 'false'. If we want to fix the rotation of an unknown image file give 'rotationFlag' as 'true' otherwise 'false'. The fields 'portfolio', 'account' and 'exclusion list' are optional. Portfolio and account can be created from admin portal and in exclusion list we can mention the PII values which we don't want to be analyzed.



Response: Returns the list of PII entities detected from the image.

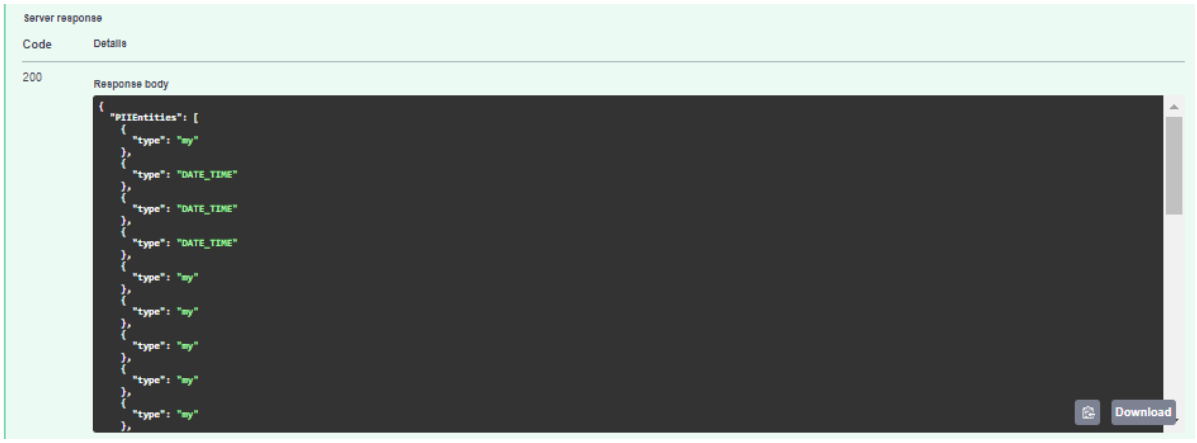
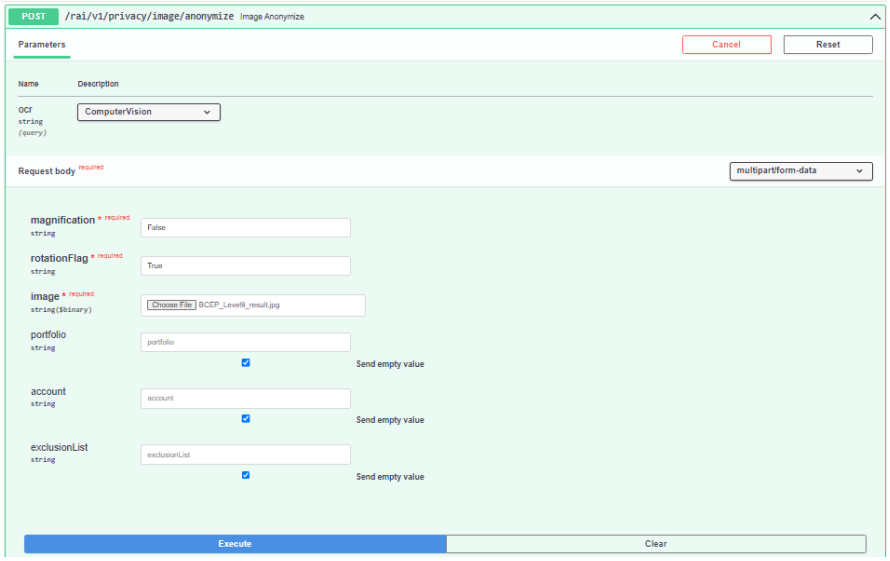


Image Anonymize

Endpoint – /rai/v1/privacy/image/anonymize

Using this API, we can anonymize the PII entities present in the uploaded image.

Input : We can select EasyOcr, tesseract or ComputerVision to analyze the image. Upload the image file, if we want the image to be magnified give 'magnification' as 'true' otherwise give 'false'. If we want to fix the rotation of an unknown image file give 'rotationFlag' as 'true' otherwise 'false'. The fields 'portfolio', 'account' and 'exclusion list' are optional. Portfolio and account can be created from admin portal and in exclusion list we can mention the PII values which we don't want to be anonymized.



Response: Returns the anonymized detected PII entities from an extracted text by drawing bounding boxes in the place of text for image in base-64 encoded format.

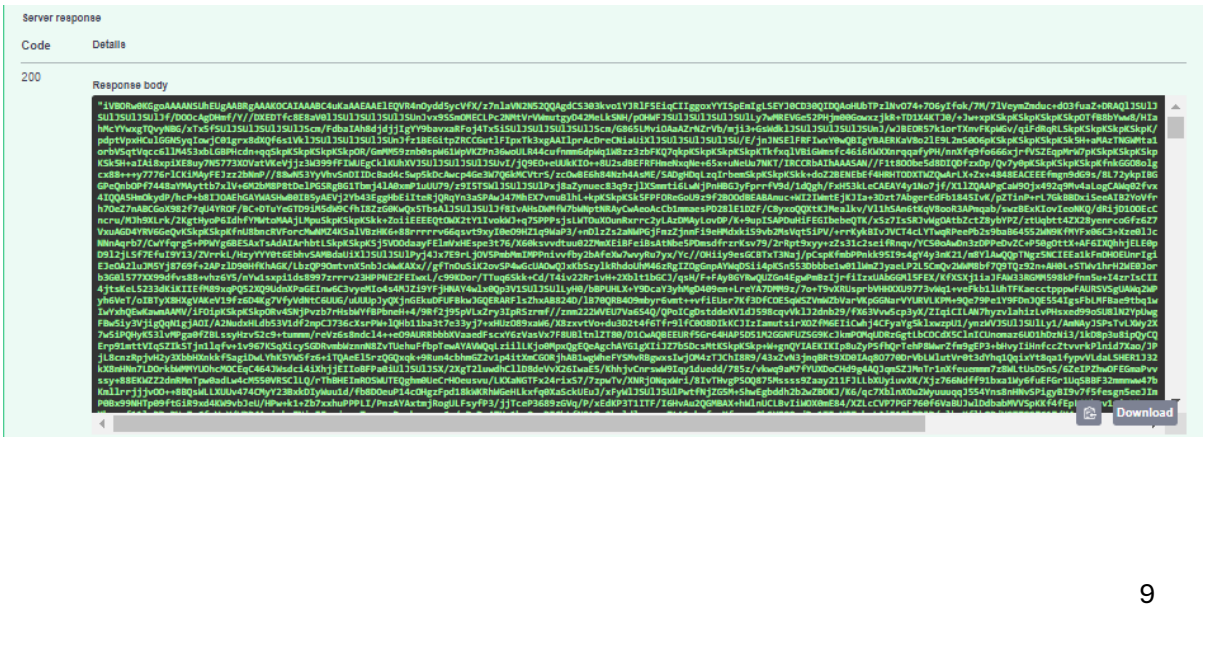
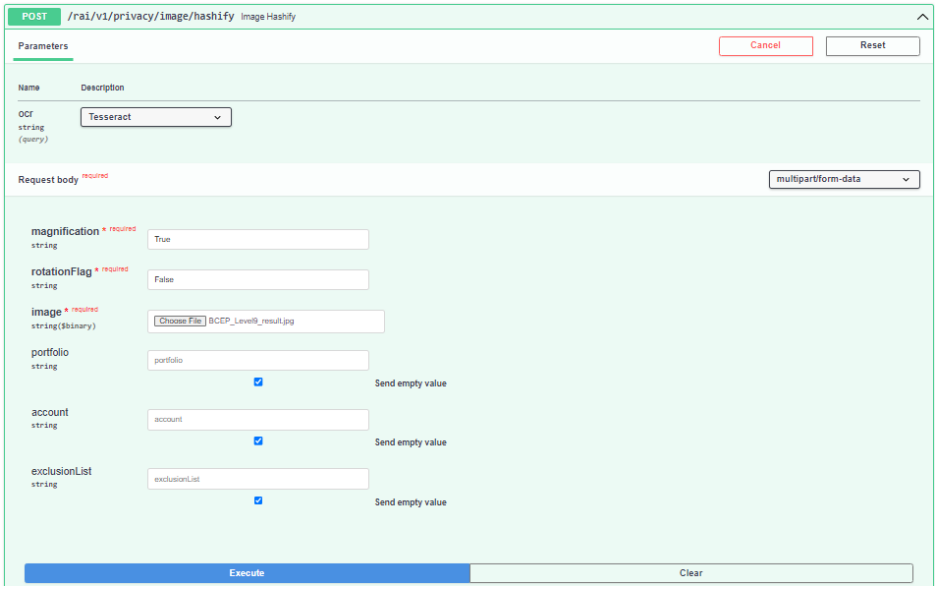


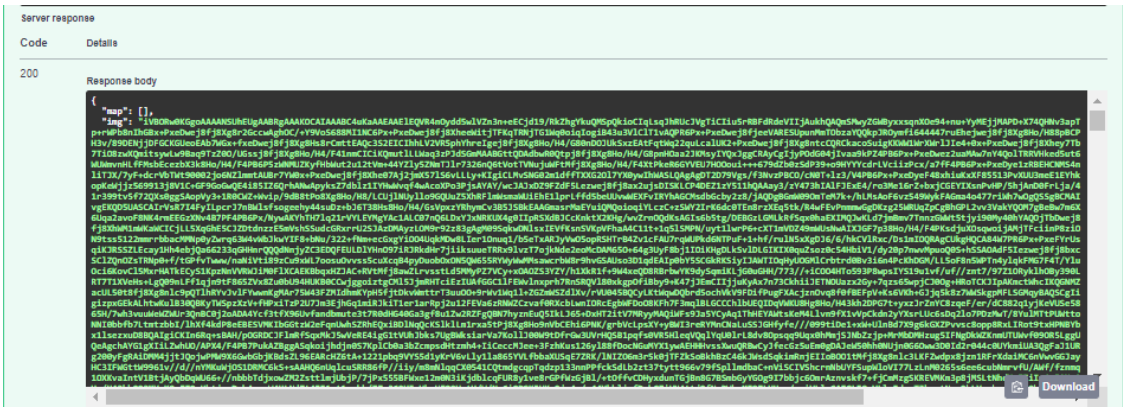
Image Hashify

Endpoint – /rai/v1/privacy/image
Using this API, we can hash out the PII entities detected in the image we uploaded.

Input : We can select EasyOcr, tesseract or ComputerVision to analyze the image. Upload the image file, if we want the image to be magnified give 'magnification' as 'true' otherwise give 'false'. If we want to fix the rotation of an unknown image file give 'rotationFlag' as 'true' otherwise 'false'. The fields 'portfolio', 'account' and 'exclusion list' are optional. Portfolio and account can be created from admin portal and in exclusion list we can mention the PII values which we don't want to be hashed in the image.



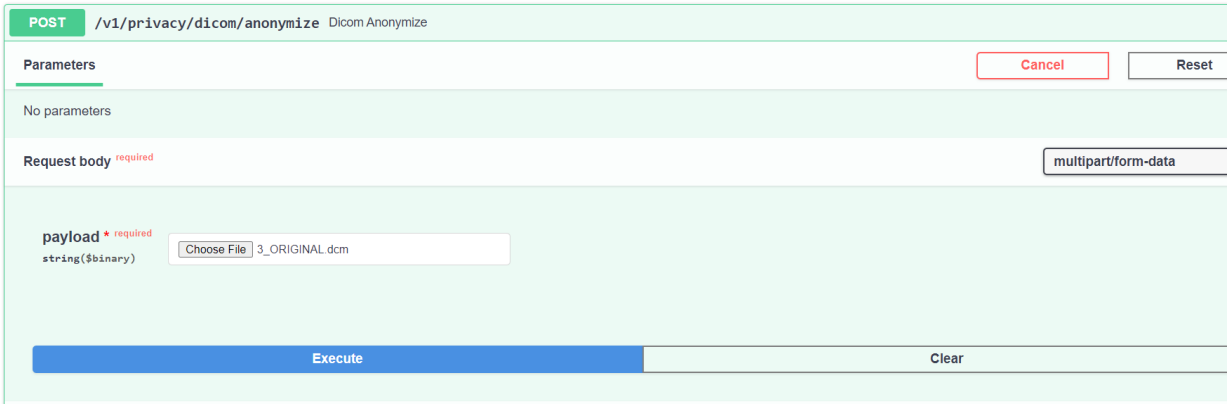
Response: Returns the image in base-64 encoded format with hashed PII entities.



Dicom Anonymize

Endpoint – /rai/v1/privacy/dicom/anonymize
Using this API, we can identify and anonymize the PII entities present in dicom images.

Input : Upload the .dcm image file

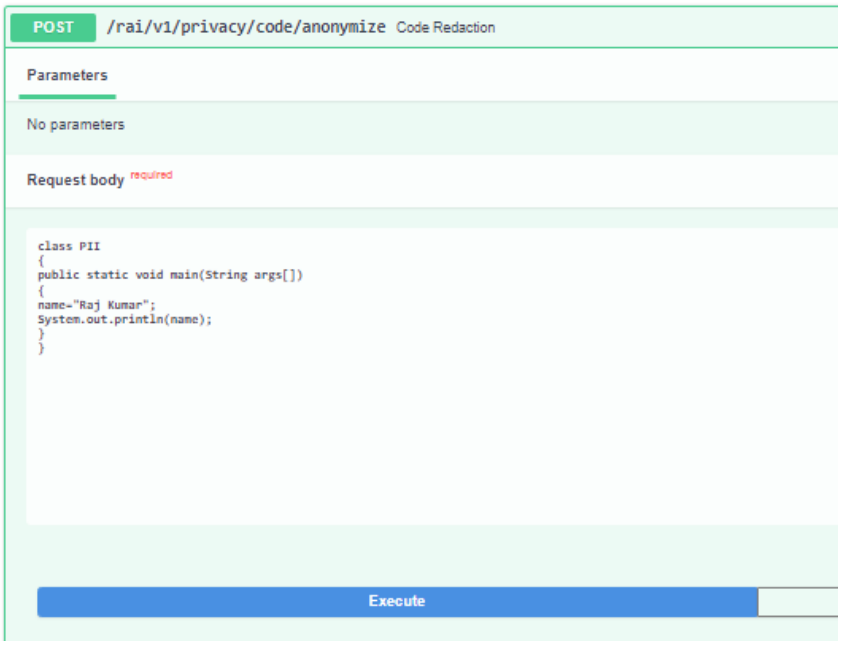


Response: Returns the dicom image in base-64 encoded format with masked PII entities.

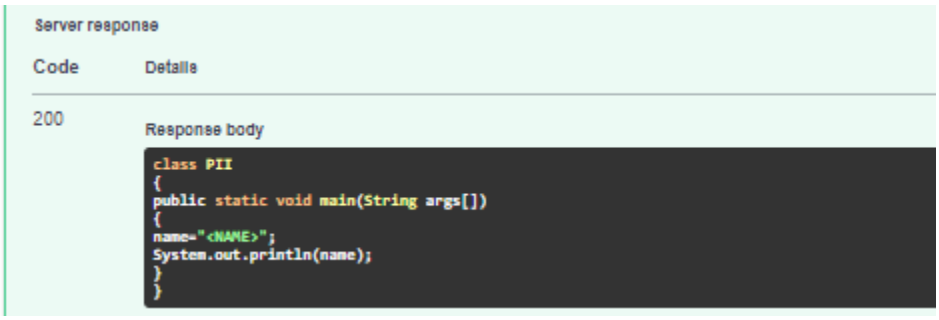
Code Anonymize

Endpoint – /rai/v1/privacy/code/anonymize
Using this API, we can identify and anonymize the PII entities present in the code that we entered as text.

Input : Enter the code that we want to check for PII entities.



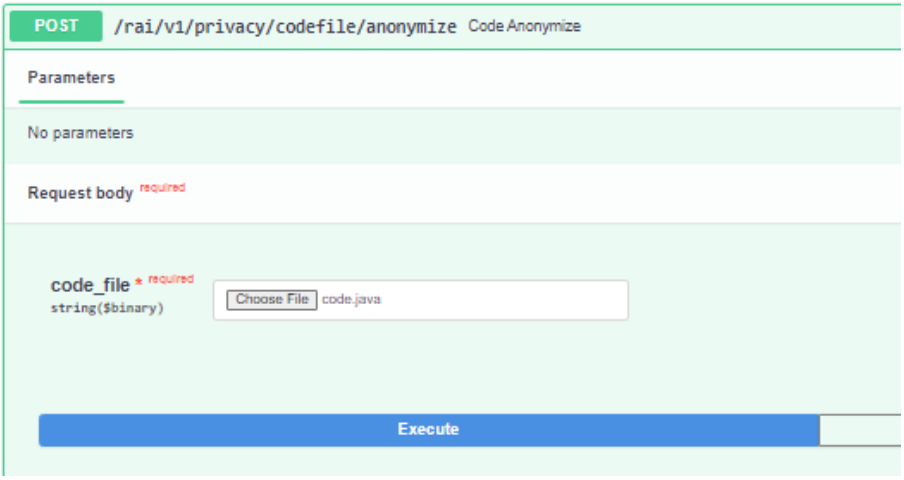
Response: Detects the PII entities in the input code and mask it.



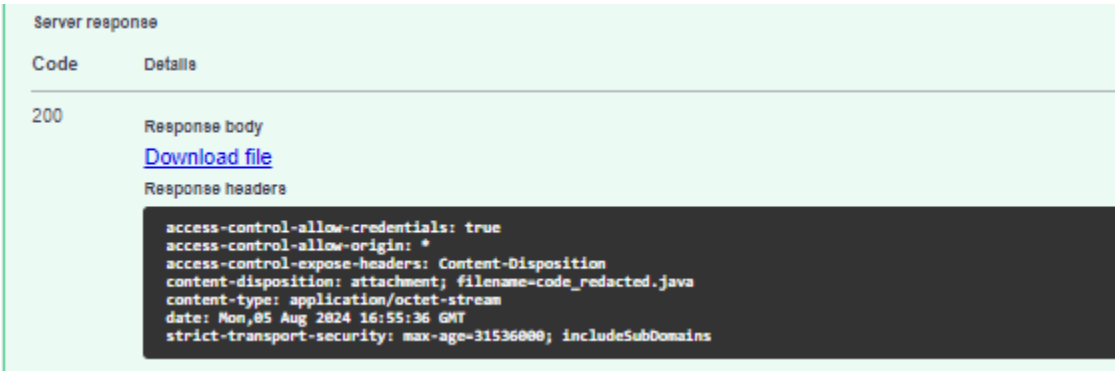
Code File Anonymize

Endpoint – /rai/v1/privacy/codefile/anonymize
Using this API, we can identify and anonymize the PII entities present in the code file that we uploaded as input

Input : Upload the code file in which you want to find and anonymize the PII entities.



Response: Returns the "Download File" link with PII entities masked in the code file.



Downloaded file –

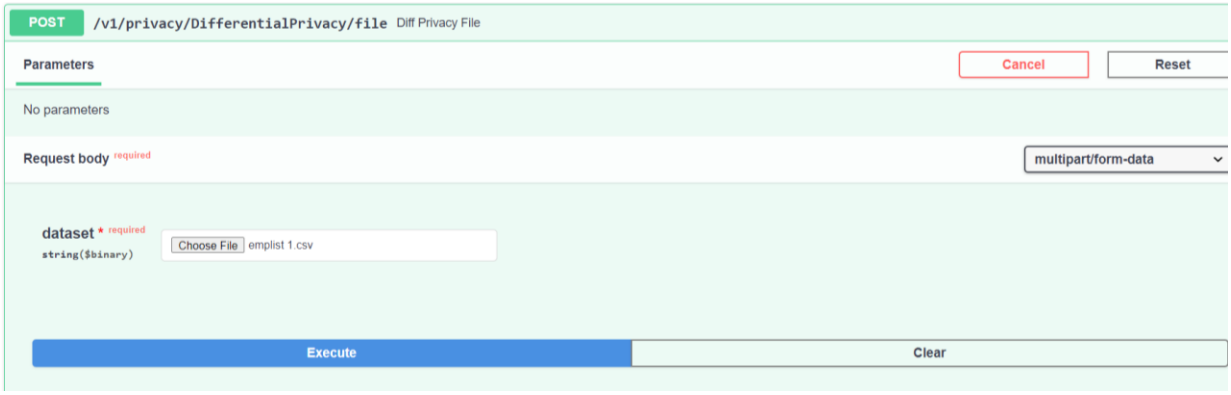
```
class PII
{
public static void main(String args[])
{
name="<NAME>";
System.out.println(name);
}
}
```

Differential Privacy

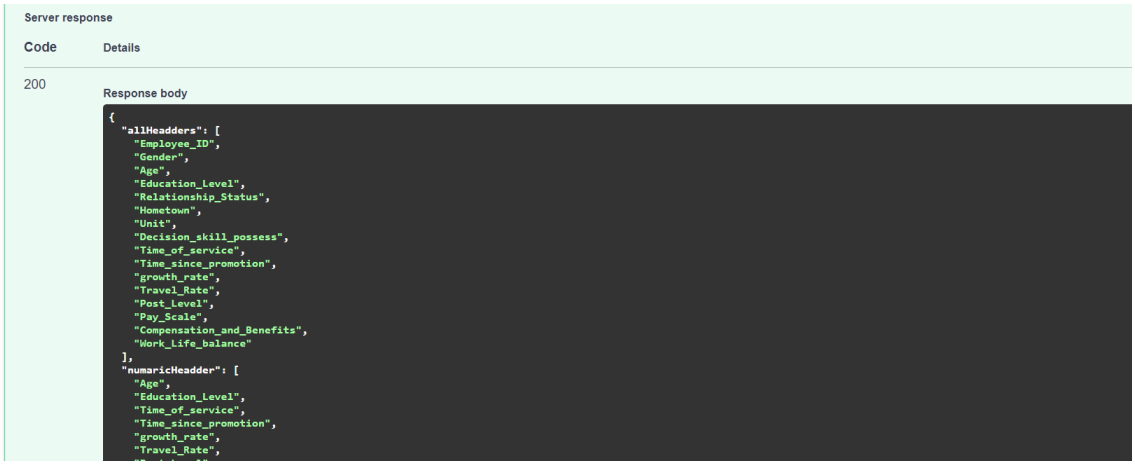
Endpoints – /rai/v1/privacy/DifferentialPrivacy/file
/rai/v1/privacy/DifferentialPrivacy/anonymize
Using the first API, we can upload the file we want to check for differential privacy and using the second API we can add suppression, noise etc. to the file values.

Input : Upload the file you want to check for differential privacy (using /rai/v1/privacy/DifferentialPrivacy/file) . Example – Here uploading a .csv file-

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Employee	Gender	Age	Education	Relationsh	Hometowi	Unit	Decision_s	Time_of_s	Time_sinc	growth_ra	Travel_Rat	Post_Leve	Pay_Scale	Compensa	Work_Life_balance	
2	EID_22713	F	32	5	Single	Springfielc	R&D	Conceptua	7	4	30	1	5	4	type2	1	
3	EID_9658	M	65	2	Single	Lebanon	IT	Directive	41	2	72	1	1	1	type2	1	
4	EID_22203	M	52	3	Married	Springfielc	Sales	Directive	21	3	25	0	1	8	type3	1	
5	EID_7652	M	50	5	Single	Washingtc	Marketing	Analytical	11	4	28	1	1	2	type0	4	
6	EID_6516	F	44	3	Married	Franklin	R&D	Conceptua	12	4	47	1	3	2	type2	4	
7	EID_20283	F	22	4	Married	Franklin	IT	Behaviora	3	1	53	0	3	6	type2	1	
8	EID_21014	M	42	3	Married	Washingtc	Purchasing	Analytical	6	4	35	1	3	4	type2	1	
9	EID_7693	F	41	2	Married	Springfielc	Sales	Conceptua	4	4	35	1	4	8	type2	1	
0	EID_13232	M	31	1	Single	Springfielc	IT	Analytical	7	3	73	2	3	8	type2	3	

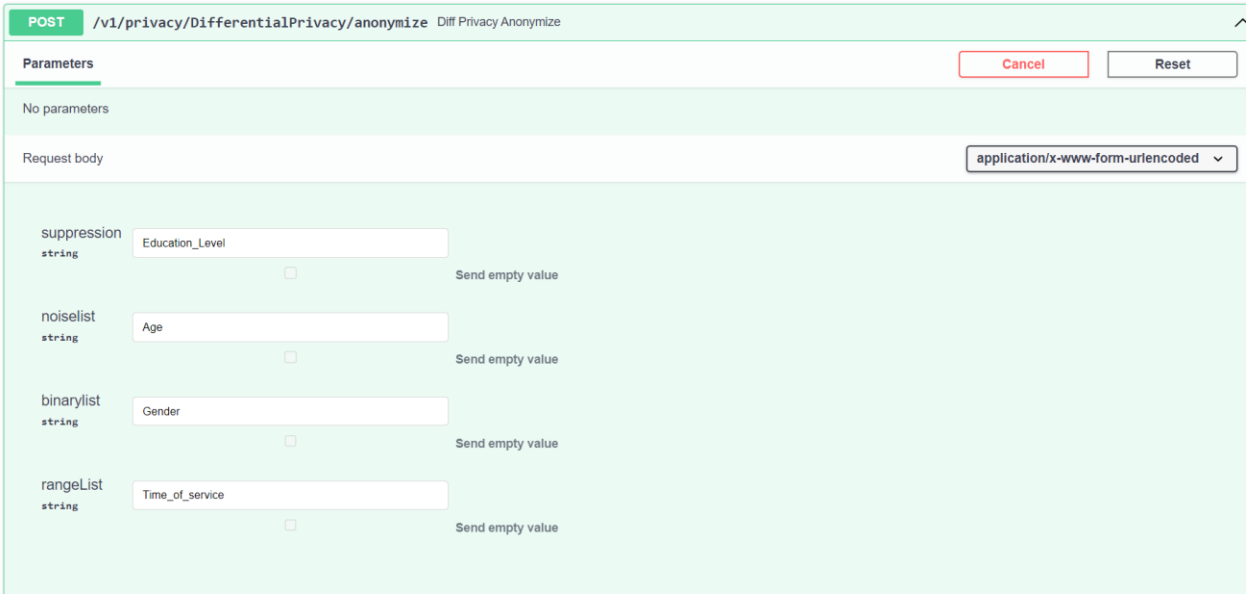


Response: List of column names

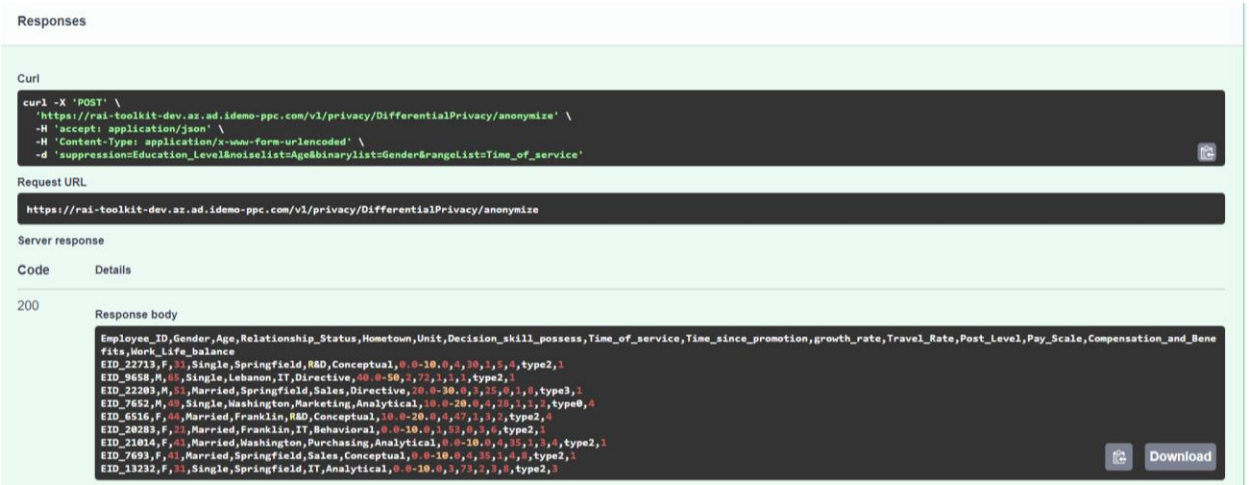


Set the column names –

1. In suppression give the column names which you want to remove or suppress
2. In noise list give the column names whose values you want to change by adding some noise or unwanted
3. In binary list give the column names which contain binary values(only two types of values like M or F, T or F) and whose values you want to anonymize by swapping the binary values at all rows. Example: If a column consists of M and F values, then replace M with F and F with M at all places.
4. In range list add the columns whose values you want to anonymize by converting them to a range.
   Using the second API, it will do these changes on the file uploaded using the first API –



Final Response : Gives the processed csv

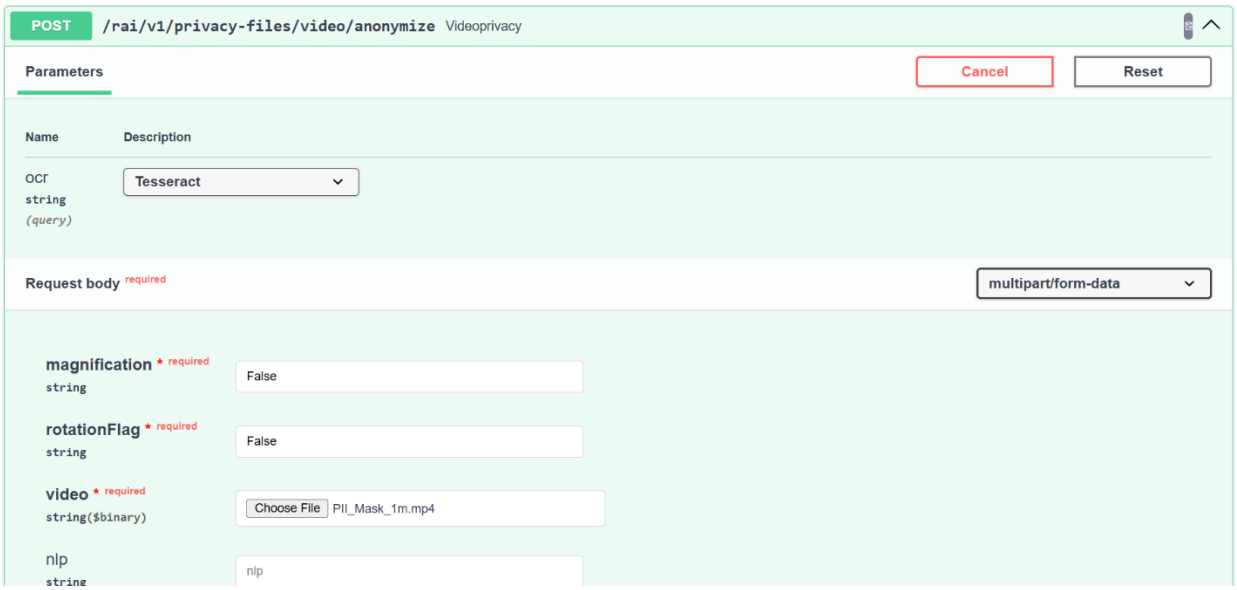


Privacyfiles_main APIs

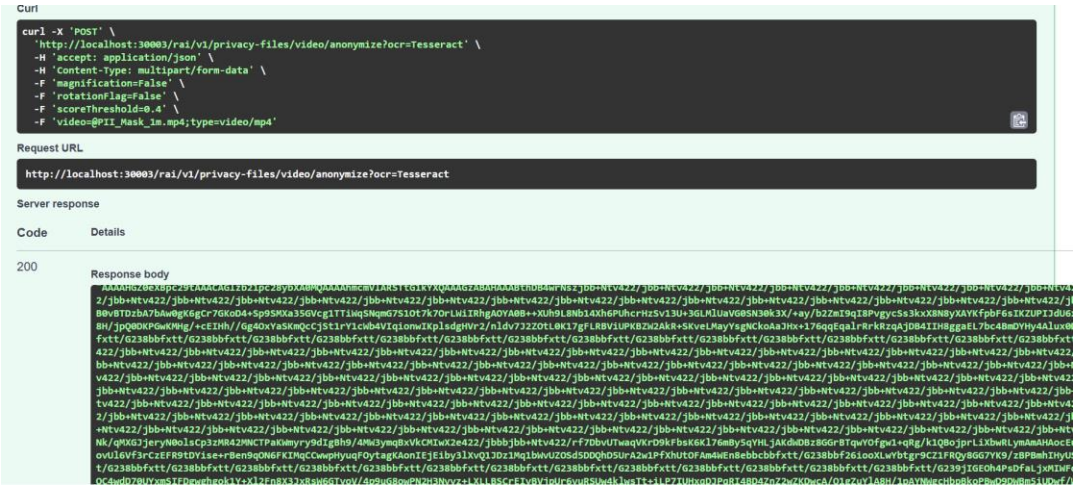
Video Anonymize

ENDPOINT: /rai/v1/privacy-files/video/anonymize
Using this API, Detect and anonymize PII entities in video frames by drawing the bounding boxes

Input: Mandatory Fields: Image, Ocr method, Magnification, Rotation flag.
Optional Fields: portfolio, account, piiEntitiesToBeRedacted, exclusionList



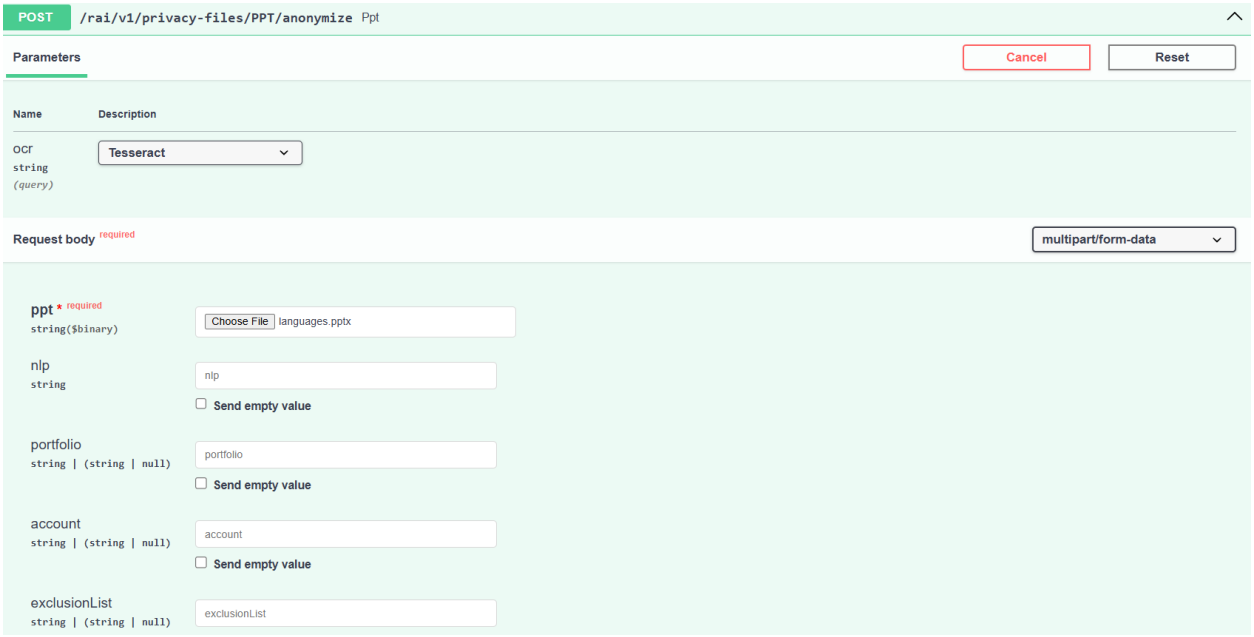
Response: Returns the video in base 64-encoded format by anonymizing the PII entities from the video.



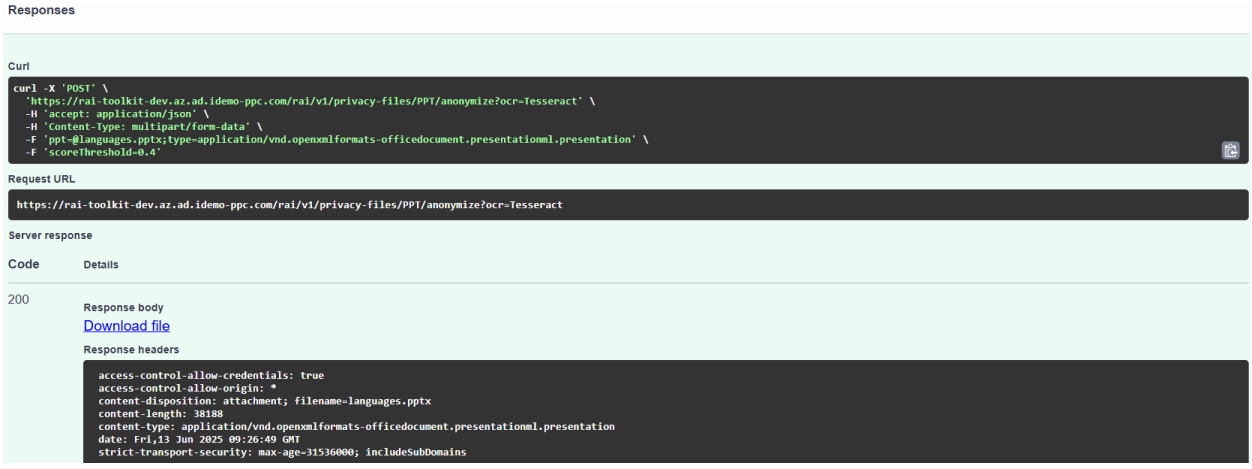
PPT Anonymize

ENDPOINT: /rai/v1/privacy-files/PPT/anonymize
Using this API, Detect and anonymize PII entities in ppt slides in text, image or tabular format.

Input: Mandatory Fields: Image, Ocr method. Optional Fields: portfolio, account, exclusionList



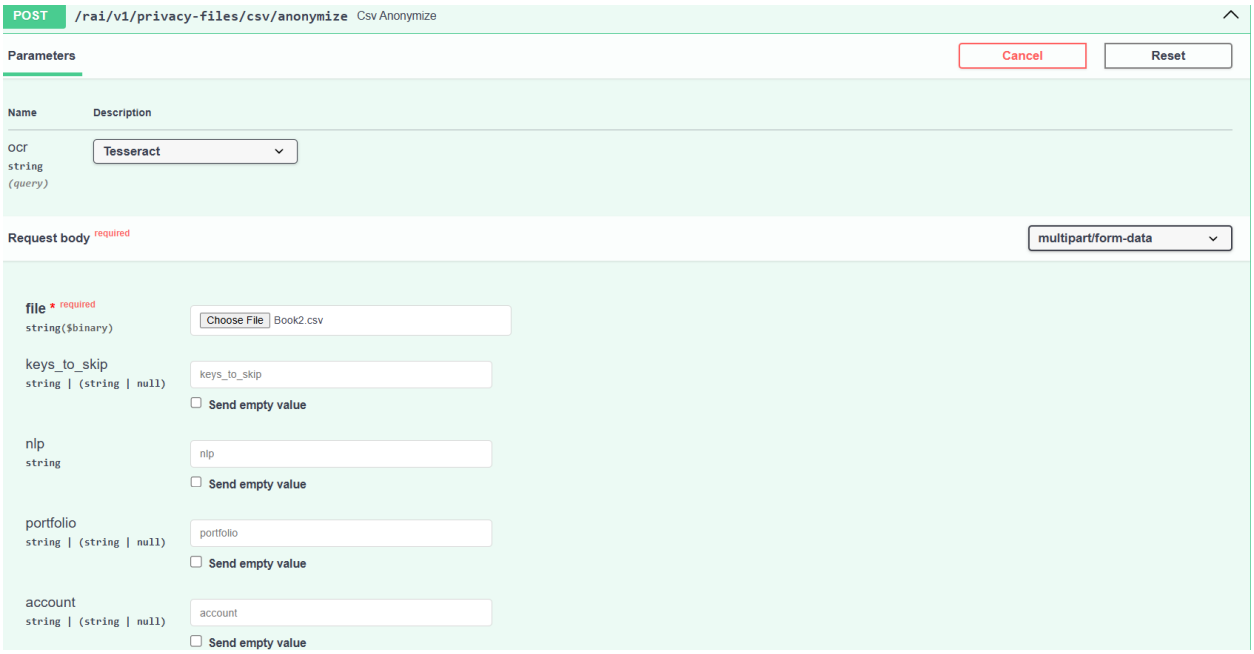
Response: Returns the "Download File" link with PII entities masked in the PPT.



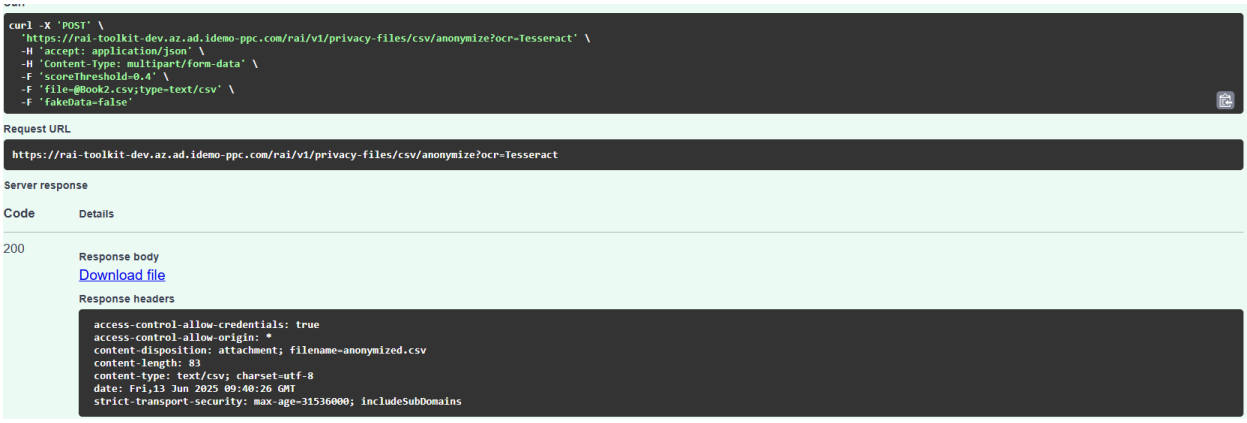
CSV Anonymize

ENDPOINT: /rai/v1/privacy-files/CSV/anonymize
Using this API, Detect and anonymize PII entities in CSV.

Input: csv file as a payload



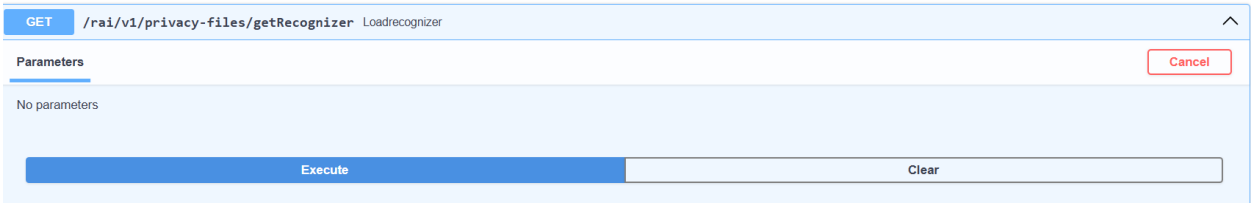
Response: Returns a “Download file” link with masked PII entities.



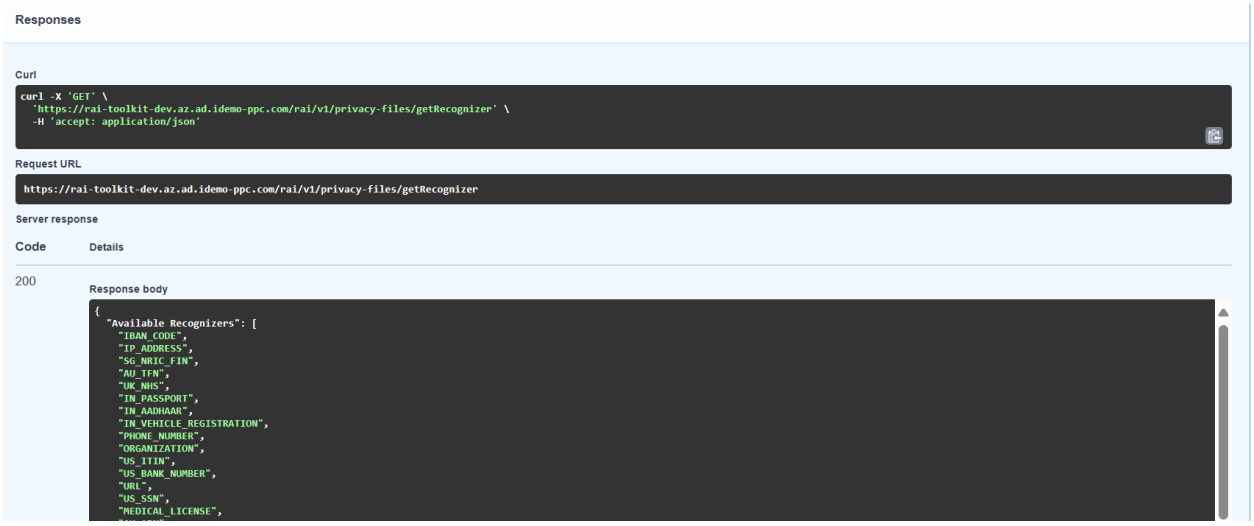
getRecognizer

ENDPOINT: /rai/v1/privacy-files/getRecognizer
Using this API, you can get a list of all supported PII entities

Input: No input required

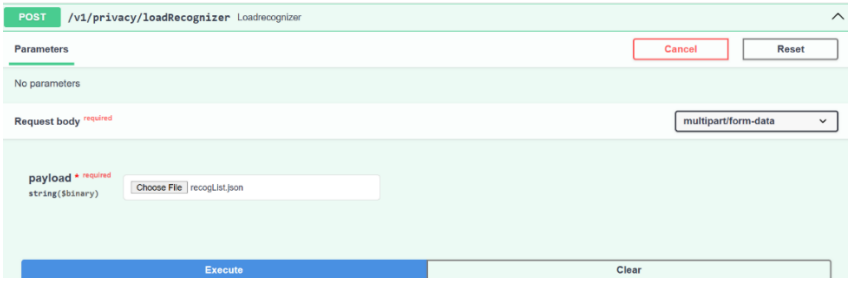


Response: Returns the list of all recognizers.



loadRecognizer

ENDPOINT: /rai/v1/privacy-files/loadRecognizer
Using this API, Dynamically registers **custom entity recognizers** (data or pattern-based) from a user-provided JSON payload for use in a text analysis pipeline.



Response: Returns the custom entity recognizers.

Endpoints usage flow

Privacy-main

Endpoint	Description
/rai/v1/privacy/text/analyze	Detects PII entities in a text and gives JSON report as an output.
/rai/v1/privacy/text/anonymize	Anonymizing detected PII entities in a text and gives anonymized text as an output.
/rai/v1/privacy/text/encrypt	Encrypts the PII data in a text using a key.
/rai/v1/privacy/text/decrypt	Decrypt the encrypted PII in the text using the encryption key.
/rai/v1/privacy/image/analyze	Detects PII entities from Image in an extracted text and gives list of PII Entities as an output.
/rai/v1/privacy/image/anonymize	Anonymizing detected PII entities from an extracted text by drawing bounding boxes in the place of text
/rai/v1/privacy/image/hashify	Hashes the PII text present in an Image which is mapped to account. Gives back the Json which contains the hash value with a key that is mentioned on the bounding boxes in the response image.
/rai/v1/privacy/dicom/anonymize	Anonymize PII entities in medical x-ray images.
/rai/v1/privacy/code/anonymize	Anonymize PII entities in code text
/rai/v1/privacy/codefile/anonymize	Anonymize PII entities in code file.
/rai/v1/privacy/DifferentialPrivacy/file	Load csv file and gives the column name which can be used in further api.
/rai/v1/privacy/DifferentialPrivacy/anonymize	Using the column name, user can apply differential privacy

PRIVACYFILES-MAIN

Endpoint	Description
/rai/v1/privacy-files/video/anonymize	Detect and anonymize PII entities in video frames by drawing the bounding boxes
/rai/v1/privacy-files/PPT/anonymize	Detect and anonymize PII entities in ppt slides in text, image or tabular format.
/rai/v1/privacy-files/DOCX/anonymize	Detect and anonymize PII entities in docx in text, image format.
/rai/v1/privacy-files/CSV/anonymize	Detect and anonymize PII entities in CSV.
/rai/v1/privacy-files/loadRecognizer	Dynamically registers **custom entity recognizers**
/rai/v1/privacy-files/getRecognizer	It detects all recognizers