

# Infosys

---

## Responsible AI Office

### Infosys Responsible AI Toolkit - Privacy Module

#### API usage Instructions

#### Contents

About Privacy .....	2
Dependencies .....	2
Features & API End Points.....	3
Privacy_main APIs .....	3
Text Analyze .....	3
Text Anonymize .....	5
Text Encrypt .....	6
Text Decrypt.....	7
Image Analyze.....	8
Image Anonymize.....	9
Image Hashify.....	10
Dicom Anonymize .....	11
Code Anonymize.....	12
Code File Anonymize.....	13
Differential Privacy.....	14
Privacyfiles_main APIs.....	16
Video Anonymize .....	16
PPT Anonymize .....	17
CSV Anonymize.....	18
getRecognizer .....	19
loadRecognizer .....	20
Endpoints usage flow .....	21
Privacy-main .....	21
PRIVACYFILES-MAIN.....	22

## 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 -

POST	/rai/v1/privacy/text/analyze	Analyze
Parameters		
No parameters		
Request body <small>required</small>		
<pre>{   "inputText": "John Smith's SSN is 012884567",   "portfolio": "string",   "account": "string",   "exclusionList": "Karan,Infosys",   "user": "string",   "lotNumber": "string" }</pre>		
Execute		

Without optional fields -

POST	/rai/v1/privacy/text/analyze	Analyze
Parameters		
No parameters		
Request body <small>required</small>		
<pre>{   "inputText": "John Smith's SSN is 012884567",   "exclusionList": "Karan,Infosys",   "user": "string",   "lotNumber": "string" }</pre>		
Execute		

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

Server response	
Code	Details
200	<div>Response body</div> <pre>{   "PIIEntities": [     {       "type": "PERSON",       "beginOffset": 0,       "endOffset": 12,       "score": 0.85,       "responseText": "John Smith's"     },     {       "type": "NEW_NPI",       "beginOffset": 20,       "endOffset": 28,       "score": 0.9,       "responseText": "01288456"     },     {       "type": "NEW_GROUP_ID",       "beginOffset": 20,       "endOffset": 25,       "score": 0.9,       "responseText": "01288"     },     {       "type": "my",       "beginOffset": 20,       "endOffset": 21, </pre>

## 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.

**POST**
/rai/v1/privacy/text/anonymize
Anonymize

Parameters

No parameters

Request body required

```

{
  "inputText": "John Smith's SSN is 012884567",
  "portfolio": "string",
  "account": "string",
  "exclusionList": "Karan,Infosys",
  "piiEntitiesToBeRedacted": [
    "US_SSN"
  ],
  "redactionType": "replace",
  "user": "string",
  "lotNumber": "string",
  "fakeData": false
}

```

**Response :** Mask the PII entities detected in the inputText.

Server response

Code	Details
200	<div>Response body</div> <pre> {   "anonymizedText": "&lt;PERSON&gt; SSN is &lt;US_SSN&gt;" } </pre>

## Text Encrypt

### Endpoint – /rai/v1/privacy/text/encrypt

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.

**POST** /rai/v1/privacy/text/encrypt Encrypt

**Parameters**

No parameters

**Request body** required

```
{
  "inputText": "John Smith's SSN is 012884567",
  "exclusionList": "Karan,Infosys",
  "piiEntitiesToBeRedacted": [
    "US_SSN"
  ],
  "redactionType": "replace",
  "user": "string",
  "lotNumber": "string",
  "fakeData": false
}
```

**Execute**

**Response :** Returns the encrypted text with masked PII entities.

**Server response**

Code	Details
200	<p><b>Response body</b></p> <pre>{   "text": "HXFNT7j0McRtHjW+H3R0+Jp56N8Y1CNhQjKWYY7Leg= SSN is L/0KS2Ps/9yihA7SHk6pyzyWlIoZNeqSse1/tFvrm=",   "items": [     {       "start": 52,       "end": 96,       "entity_type": "US_SSN",       "text": "L/0KS2Ps/9yihA7SHk6pyzyWlIoZNeqSse1/tFvrm=",       "operator": "encrypt"     },     {       "start": 0,       "end": 44,       "entity_type": "PERSON",       "text": "HXFNT7j0McRtHjW+H3R0+Jp56N8Y1CNhQjKWYY7Leg=",       "operator": "encrypt"     }   ] }</pre>

## Text Decrypt

**Endpoint** – /rai/v1/privacy/text/decrypt

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.

POST

/rai/v1/privacy/text/decrypt Decrypt

Parameters

No parameters

Request body required

```
{
  "text": "John Smith's SSN is 012884567",
  "items": [
    {
      "start": 19,
      "end": 28,
      "entity_type": "US_SSN",
      "text": "John Smith's SSN is 012884567",
      "operator": "encrypt"
    }
  ]
}
```

Execute

**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.

**POST** /rai/v1/privacy/image/analyze Image Analyze

Parameters

Name	Description
OCR string (query)	ComputerVision

Request body *required*

**magnification** *required*  
string  
False

**rotationFlag** *required*  
string  
True

**image** *required*  
string(Binary)  
Choose File BCEP\_Level9\_result.jpg

**portfolio**  
string  
portfolio ☒ Send empty value

**account**  
string  
account ☒ Send empty value

**exclusionList**  
string  
exclusionList ☒ Send empty value

Execute

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

Server response

Code Details

200

Response body

```
{
  "PIIEntities": [
    {
      "type": "my"
    },
    {
      "type": "DATE_TIME"
    },
    {
      "type": "DATE_TIME"
    },
    {
      "type": "DATE_TIME"
    },
    {
      "type": "my"
    },
    {
      "type": "my"
    },
    {
      "type": "my"
    },
    {
      "type": "my"
    },
    {
      "type": "my"
    },
    {
      "type": "my"
    }
  ]
}
```

Download



## 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.



## 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.

POST

/rai/v1/privacy/code/anonymize

Code Redaction

Parameters

No parameters

Request body required

```

class PII
{
public static void main(String args[])
{
name="Raj Kuman";
System.out.println(name);
}
}

```

Execute

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

Server response	
Code	Details
200	<div>Response body</div> <pre> class PII { public static void main(String args[]) { name="&lt;NAME&gt;"; System.out.println(name); } } </pre>

## 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.

POST /rai/v1/privacy/codefile/anonymize Code Anonymize

Parameters

No parameters

Request body required

**code\_file** \* required  
string(\$binary)  code.java

Execute

**Response:** Returns the “Download File” link with PII entities masked in the code file.

Server response

Code	Details
200	<p>Response body</p> <p><a href="#">Download file</a></p> <p>Response headers</p> <pre>access-control-allow-credentials: true access-control-allow-origin: * access-control-expose-headers: Content-Disposition content-disposition: attachment; filename=code_redacted.java content-type: application/octet-stream date: Mon, 05 Aug 2024 16:55:36 GMT strict-transport-security: max-age=31536000; includeSubDomains</pre>

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	Relationship	Hometown	Unit	Decision_skill_possess	Time_of_service	Time_since_promotion	growth_rate	Travel_Rate	Post_Level	Pay_Scale	Compensation_and_Benefits	Work_Life_balance	
2	EID_2271	F	32	5	Single	Springfield	R&D	Conceptual	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_2220	M	52	3	Married	Springfield	Sales	Directive	21	3	25	0	1	8	type3	1	
5	EID_7652	M	50	5	Single	Washington	Marketing	Analytical	11	4	28	1	1	2	type0	4	
6	EID_6516	F	44	3	Married	Franklin	R&D	Conceptual	12	4	47	1	3	2	type2	4	
7	EID_2028	F	22	4	Married	Franklin	IT	Behavioral	3	1	53	0	3	6	type2	1	
8	EID_2101	M	42	3	Married	Washington	Purchasing	Analytical	6	4	35	1	3	4	type2	1	
9	EID_7693	F	41	2	Married	Springfield	Sales	Conceptual	4	4	35	1	4	8	type2	1	
0	EID_1323	M	31	1	Single	Springfield	IT	Analytical	7	3	73	2	3	8	type2	3	

**POST** /v1/privacy/DifferentialPrivacy/file Diff Privacy File

**Parameters**

Cancel

Reset

No parameters

**Request body** required

multipart/form-data

**dataset** required string(\$binary)

Choose File

emplist 1.csv

Execute

Clear

**Response:** List of column names

Server response	
Code	Details
200	Response body <pre>{   "allHeaders": [     "Employee_ID",     "Gender",     "Age",     "Education_Level",     "Relationship_Status",     "Hometown",     "Unit",     "Decision_skill_possess",     "Time_of_service",     "Time_since_promotion",     "growth_rate",     "Travel_Rate",     "Post_Level",     "Pay_Scale",     "Compensation_and_Benefits",     "Work_Life_balance"   ],   "numericHeader": [     "Age",     "Education_Level",     "Time_of_service",     "Time_since_promotion",     "growth_rate",     "Travel_Rate"   ] }</pre>

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 –

POST
/v1/privacy/DifferentialPrivacy/anonymize
Diff Privacy Anonymize

Parameters
Cancel
Reset

No parameters

Request body
application/x-www-form-urlencoded

suppression
string
Education\_Level
☐
Send empty value

noiselist
string
Age
☐
Send empty value

binarylist
string
Gender
☐
Send empty value

rangeList
string
Time\_of\_service
☐
Send empty value

**Final Response :** Gives the processed csv

Responses

Curl

```
curl -X 'POST' \
'https://rai-toolkit-dev.az.ad.idemo-ppc.com/v1/privacy/DifferentialPrivacy/anonymize' \
-H 'accept: application/json' \
-H 'Content-Type: application/x-www-form-urlencoded' \
-d 'suppression=Education_Level&noiselist=Age&binarylist=Gender&rangeList=Time_of_service'
```

Request URL
https://rai-toolkit-dev.az.ad.idemo-ppc.com/v1/privacy/DifferentialPrivacy/anonymize

Server response

Code
Details

200
Response body

```
Employee_ID,Gender,Age,Relationship_Status,Hometown,Unit,Decision_skill_possess,Time_of_service,Time_since_promotion,growth_rate,Travel_Rate,Post_Level,Pay_Scale,Compensation_and_Bene
fits,Work_Life_Balance
EID_22713,F,31,Single,Springfield,R&D,Conceptual,0.0-10.0,0.30,1.5,0,type2,1
EID_9658,M,40,Single,Lebanon,IT,Directive,40.0-50.0,0.72,1.1,0,type2,1
EID_22203,M,51,Married,Springfield,Sales,Directive,20.0-30.0,0.3,0.25,0.1,0,type3,1
EID_7652,M,40,Single,Washington,Marketing,Analytical,10.0-20.0,0.28,1.1,0,type0,0
EID_6515,F,40,Married,Franklin,R&D,Conceptual,20.0-20.0,0.4,0.7,1.1,0,type2,0
EID_20283,F,31,Married,Franklin,IT,Behavioral,0.0-10.0,0.1,0.3,0.1,0,type2,1
EID_21014,F,41,Married,Washington,Purchasing,Analytical,0.0-10.0,0.3,0.3,0.1,0,type2,1
EID_7693,F,41,Married,Springfield,Sales,Conceptual,0.0-10.0,0.3,0.3,0.1,0,type2,1
EID_13232,F,31,Single,Springfield,IT,Analytical,0.0-10.0,0.3,0.3,0.1,0,type2,1
```

Download







## 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

POST

/rai/v1/privacy-files/PPT/anonymize Ppt

Cancel

Reset

Parameters

Name	Description
ocr string (query)	Tesseract

Request body required

multipart/form-data

ppt \* required

string(\$binary)

Choose File | languages.pptx

nlp

string

nlp

☐ Send empty value

portfolio

string | (string | null)

portfolio

☐ Send empty value

account

string | (string | null)

account

☐ Send empty value

exclusionList

string | (string | null)

exclusionList

**Response:** Returns the “Download File” link with PII entities masked in the PPT.

Responses

Curl

```
curl -X 'POST' \
  'https://rai-toolkit-dev.az.ad.idemo-ppc.com/rai/v1/privacy-files/PPT/anonymize?ocr=Tesseract' \
  -H 'accept: application/json' \
  -H 'Content-Type: multipart/form-data' \
  -F 'ppt=@languages.pptx;type=application/vnd.openxmlformats-officedocument.presentationml.presentation' \
  -F 'scoreThreshold=0.4'
```

Request URL

```
https://rai-toolkit-dev.az.ad.idemo-ppc.com/rai/v1/privacy-files/PPT/anonymize?ocr=Tesseract
```

Server response

Code	Details
200	<div>Response body</div> <div><a href="#">Download file</a></div> <div>Response headers</div> <pre>access-control-allow-credentials: true access-control-allow-origin: * content-disposition: attachment; filename=Languages.pptx content-length: 38188 content-type: application/vnd.openxmlformats-officedocument.presentationml.presentation date: Fri, 13 Jun 2025 09:26:49 GMT strict-transport-security: max-age=31536000; includeSubDomains</pre>

## 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

POST

/rai/v1/privacy-files/csv/anonymize

Csv Anonymize

Cancel

Reset

Parameters

Name	Description
ocr string (query)	<div>Tesseract</div>

Request body required

multipart/form-data

file required

string(\$binary)

Choose File

Book2.csv

keys\_to\_skip

string | (string | null)

keys\_to\_skip

☐ Send empty value

nlp

string

nlp

☐ Send empty value

portfolio

string | (string | null)

portfolio

☐ Send empty value

account

string | (string | null)

account

☐ Send empty value

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

```
curl -X 'POST' \
  'https://rai-toolkit-dev.az.ad.idemo-ppc.com/rai/v1/privacy-files/csv/anonymize?ocr=Tesseract' \
  -H 'accept: application/json' \
  -H 'Content-Type: multipart/form-data' \
  -F 'scoreThreshold=0.4' \
  -F 'file=@Book2.csv;type-text/csv' \
  -F 'fakeData=false'
```

Request URL

https://rai-toolkit-dev.az.ad.idemo-ppc.com/rai/v1/privacy-files/csv/anonymize?ocr=Tesseract

Server response

Code	Details
200	<div>Response body</div> <div> <a href="#">Download file</a> </div> <div>Response headers</div> <div> <pre>access-control-allow-credentials: true access-control-allow-origin: * content-disposition: attachment; filename-anonymized.csv content-length: 83 content-type: text/csv; charset=utf-8 date: Fri, 13 Jun 2025 09:40:26 GMT strict-transport-security: max-age=31536000; includeSubDomains</pre> </div>

18



## 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.

POST /v1/privacy/loadRecognizer Loadrecognizer

Parameters Cancel Reset

No parameters

Request body required multipart/form-data

payload required  
string(\$binary) Choose File recogList.json

Execute Clear

**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 <b>custom entity recognizers</b>
/rai/v1/privacy-files/getRecognizer	It detects all recognizers