



πυμίζ

Documentación API pública Nymiz

Versión del documento: v3.5.1154



Índice

1. Login	4
2. Anonimización de texto	5
3. Anonimización de documentos	6
4. Revertir seudonimización de texto plano	6
6. Anonimización de bases de datos	8
a. Fase de análisis (opcional)	8
b. Fase de anonimización/seudonimización	9
7. Revertir seudonimización de base de datos	10
8. Revertir datos de una misma columna seudonimizados.	10
9. Procesamiento por lotes	11
a. Iniciar procesamiento por lotes	12
b. Subir archivo al procesamiento por lotes	12
c. Subir carpeta al procesamiento por lotes	13
d. Obtener archivos del procesamiento por lotes	13
e. Comenzar procesamiento por lotes	14
f. Obtener el status del procesamiento por lotes	14
g. Descargar zip del procesamiento por lotes	14
10. Anonimización de JSON	14
11. Extracción de contenido de documentos en crudo (OCR a JSON)	15
12. Supervisión Manual Embebida	16
13. Objeto de configuración para los documentos y texto plano	17
14. Listado de infotipos soportados	19
15. Encabezados	20

1. Login

Todas las operaciones de la API pública están securizadas. Para poder hacer uso de las funcionalidades ofrecidas, es necesario obtener primeramente un token que nos autentique.

Endpoint:

(http POST): <https://api.nymiz.com/auth/signin>

Parámetros de entrada:

```
{
  "username": "[usuario de registro]",
  "password": "[contraseña de registro]"
}
```

Salida:

Si el usuario está registrado en el sistema y la contraseña es válida, se devolverá la siguiente información

```
{
  "token": "[token de seguridad]",
  "type": "Bearer",
  "id": "[id de usuario]",
  "idCustomer": [id de cliente],
  "idLicense": 0,
  "email": "[email de usuario]",
  "nameSurname": "[nombre y apellido de usuario]",
  "roles": [
    "[rol/es del usuario]"
  ],
  "idProduct": 0,
  "status": false,
  "verification": true
}
```

2. Anonimización de texto

Endpoint:

(http POST): <https://api.nymiz.com/text>

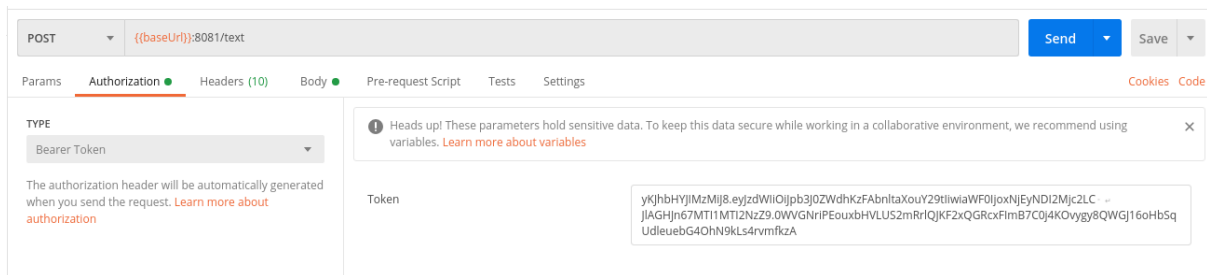
Parámetros de entrada:

- **text:** texto a anonimizar
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

Salida: texto anonimizado o seudonimizado

Ejemplo con aplicación de pruebas Postman:

Pestaña "Authorization" con el JWT obtenido en el Login



POST `{{baseUrl}}:8081/text` Send Save

Params **Authorization** Headers (10) Body Pre-request Script Tests Settings Cookies Code

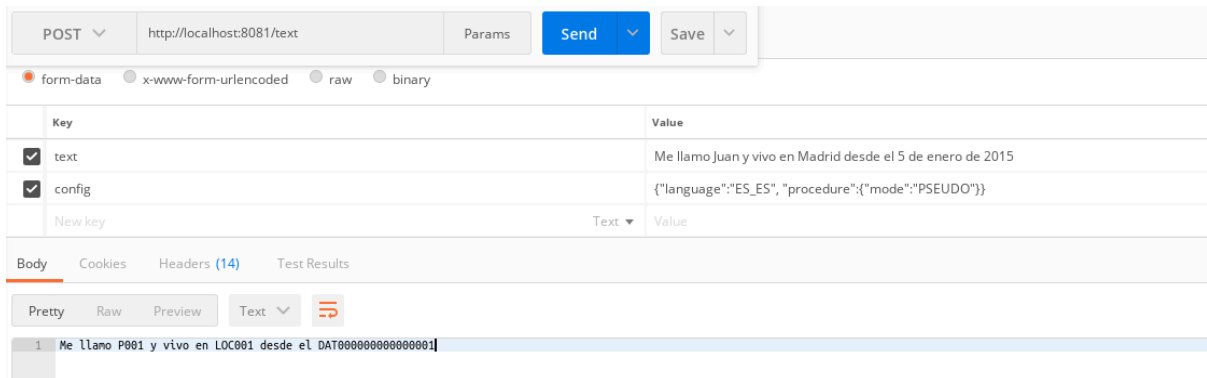
TYPE: Bearer Token

The authorization header will be automatically generated when you send the request. [Learn more about authorization](#)

Heads up! These parameters hold sensitive data. To keep this data secure while working in a collaborative environment, we recommend using variables. [Learn more about variables](#)

Token: `yKjhbHYjIMzMij8.ejzdwIiOjpb3j0ZWdhKzFAbnItaXouY29tiwiaWF0JoxNjEyNDI2Mjc2LC0tJlAGHjIn67MTI1MTI2NzZ9.0WVGnRiPEouxhVLU52mRrIQjKF2xQGRcxFimB7C0j4KOvygy8QWVGj16oHb5qUdleuebG40hN9kLs4rvmfkzA`

Parámetros de entrada, y texto anonimizado



POST `http://localhost:8081/text` Params Send Save

form-data x-www-form-urlencoded raw binary

Key	Value
<input checked="" type="checkbox"/> text	Me llamo Juan y vivo en Madrid desde el 5 de enero de 2015
<input checked="" type="checkbox"/> config	{"language":"ES_ES", "procedure":"{"mode":"PSEUDO"}}}
New key	Text Value

Body Cookies Headers (14) Test Results

Pretty Raw Preview Text

```
1 Me llamo P001 y vivo en LOC001 desde el DAT000000000000001
```

3. Anonimización de documentos

Se aceptan los siguientes formatos: txt, doc, docx, pdf, jpg, png, tiff, xls, xlsx

Endpoint:

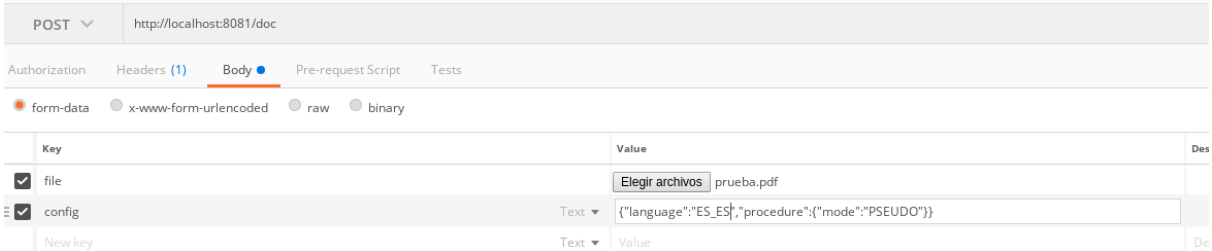
(http POST): <https://api.nymiz.com/doc>

Parámetros de entrada:

- **file:** documento a anonimizar/seudonimizar/seudo reversible
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

Salida: documento anonimizado oseudonimizado

Ejemplo con aplicación de pruebas Postman:



Key	Value	Des
<input checked="" type="checkbox"/> file	<input type="text" value="Elegir archivos"/> prueba.pdf	
<input checked="" type="checkbox"/> config	Text <input type="text" value='{"language":"ES_ES","procedure":{"mode":"PSEUDO"}}'/>	
<input type="checkbox"/> New key	Text <input type="text" value="Value"/>	De

4. Revertirseudonimización de texto plano

Endpoint:

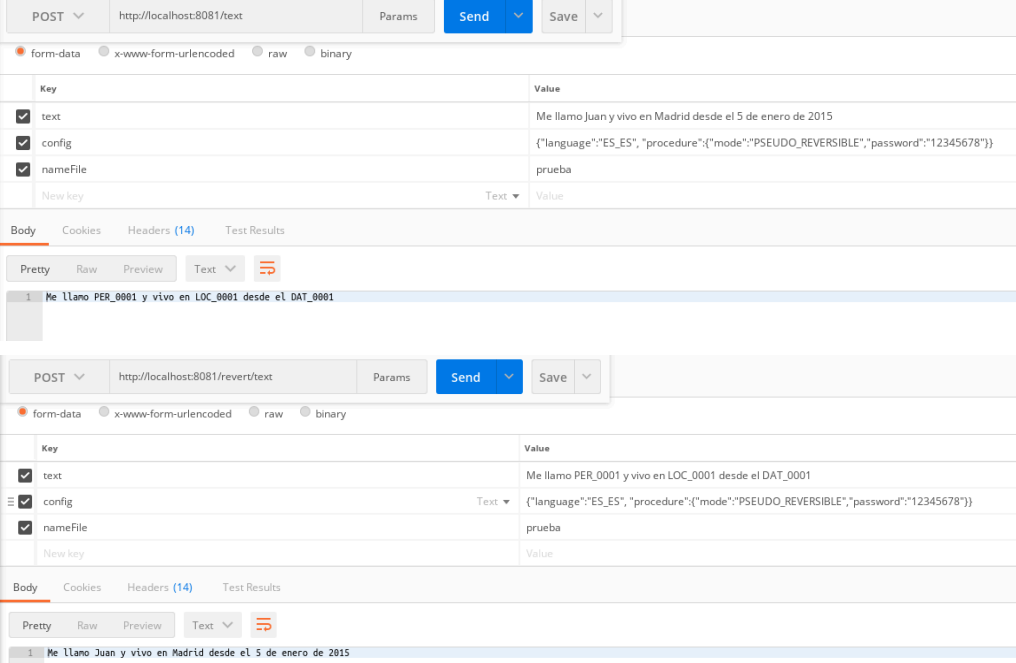
(http POST): <https://api.nymiz.com/revert/text>

Parámetros de entrada:

- **text:** texto a revertir
- **nameFile:** nombre del documento original
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

Salida: texto plano con laseudonimización revertida. Es decir, el texto original.

Ejemplo con aplicación de pruebas Postman



The first screenshot shows a POST request to `http://localhost:8081/text`. The body is configured as form-data with the following fields:

Key	Value
text	Me llamo Juan y vivo en Madrid desde el 5 de enero de 2015
config	{ "language": "ES_ES", "procedure": { "mode": "PSEUDO_REVERSIBLE", "password": "12345678" } }
nameFile	prueba

The response body is:

```
Me llamo PER_0001 y vivo en LOC_0001 desde el DAT_0001
```

The second screenshot shows a POST request to `http://localhost:8081/revert/text`. The body is configured as form-data with the following fields:

Key	Value
text	Me llamo PER_0001 y vivo en LOC_0001 desde el DAT_0001
config	{ "language": "ES_ES", "procedure": { "mode": "PSEUDO_REVERSIBLE", "password": "12345678" } }
nameFile	prueba

The response body is:

```
Me llamo Juan y vivo en Madrid desde el 5 de enero de 2015
```

5. Revertir seudonimización de documentos

Endpoint:

(http POST): <https://api.nymiz.com/revert/doc>

Parámetros de entrada:

- **file:** fichero seudonimizado
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

Salida: documento con la seudonimización revertida. Es decir, el documento original.

6. Anonimización de bases de datos

El proceso de anonimización / seudonimización de bases de datos consta de dos pasos.

a. Fase de análisis (opcional)

Endpoint:

(http POST): <https://api.nymiz.com/analyzer/analyze>

Parámetros de entrada:

- **dataBaseName:** nombre de la base de datos a analizar
- **jdbcUrl:** cadena de conexión con formato JDBC
- **(header):** [Encabezado](#)

Ejemplo:

```
{
  "dataBaseName": "db_prueba",
  "jdbcUrl": "jdbc:postgresql://0.0.0.0:5432/prueba",
  "Password": "4iipooi342423",
  "userId": "intento",
  "schema": ""
}
```

Salida:

```
{
  "name": "db_prueba",
  "tables": {
    "datos_simulados": {
      "name": "datos_",
      "columns": {
        "nombre": {
          "name": "nombre",
          "type": "varchar"
        },
        "identificador": {
          "name": "identificador",
          "type": "int4"
        }
      }
    }
  },
  "pk": true
}
```


b. Fase de anonimización/seudonimización

Endpoint:

(http POST): [https://api.nymiz.com/anonymizer/anonymize?type=\[MODO\]](https://api.nymiz.com/anonymizer/anonymize?type=[MODO])

El parámetro MODO puede tener los siguientes valores:

- **PSEUDO**:seudonimización mediante tokens
- **ANON**:anonimización
- **PSEUDO_REVERSIBLE**: operación “reversible”: elegir esta opción si se desea tener la posibilidad de recuperar la base de datos original a partir de la base de datos seudonimizada.

Parámetros de entrada:

- (header): **Encabezado** Será necesario añadir el token de autenticación a la llamada
- **Ejemplo de JSON de configuración:**

```
{
  "configuration":{
    "jdbcurlFrom":"[cadena de conexión",
    "useridFrom":"[usuario con acceso a la base de datos]",
    "passwordFrom":"[contraseña para acceder a la base de datos]",
    "Password_token":"[contraseña a indicar en caso de seudonimización reversible]",
    "tables":[
      {
        "name":"[nombre de la tabla (se pueden indicar varias tablas)]",
        "columns":[{"name":"[nombre de la columna]","type":"[tipo de la columna]"}]
      }
    ]
  },
}
```

Salida: se obtiene la base de datos anonimizada

7. Revertir seudonimización de base de datos

Endpoint:

(http POST): <https://api.nymiz.com/anonymizer/anonymize/revert>

Parámetros de entrada:

- (header): **Encabezado** Será necesario añadir el token de autenticación a la llamada
- **Ejemplo de JSON de configuración:**

```
{
  "configuration":{
    "jdbcurlFrom":"[cadena de conexión] ",
    "useridFrom":"[usuario con acceso a la base de datos]",
    "passwordFrom":"[contraseña para acceder a la base de
datos]",
    "Password_token":"[contraseña a indicar en caso de
seudonimización reversible]"
  }
}
```

Salida: se obtiene la base de datos original.

8. Revertir datos de una misma columna seudonimizados.

Permite obtener los datos revertidos de una columna seudonimizada previamente.

Endpoint:

(http POST): <https://api.nymiz.com/anonymizer/anonymize/revert>

Parámetros de entrada:

- (header): **Encabezado** Será necesario añadir el token de autenticación a la llamada
- **Ejemplo de JSON de configuración:**

```
{
  "configuration":{
    "jdbcurlFrom":"[urlConexion]:[puerto]/[baseDatos]",
    "schema":"[nombre esquema]",
    "useridFrom":"[usuario con acceso]",
    "passwordFrom":"[password de usuario]",
    "password_token":"[password de cifrado usado]",
    "table":"[tabla de busqueda]",
    "column":"[columna de busqueda]",
    "cypherText":"[palabra cifrada a buscar]"
  }
}
```

Salida: se obtienen los elementos de la columna indicada revertidos.

```
{
  "configuration":{

    "jdbcurlFrom":"[urlConexion]:[puerto]/[baseDatos]",
    "table":"[tabla de busqueda]",
    "column":"[columna de busqueda]",
    "cypherText":"[palabra cifrada a buscar]",
    "plainText":"[texto revertido]"
  }
}
```

9. Procesamiento por lotes

Se aceptan múltiples archivos de los siguientes formatos: txt, doc, docx, pdf, jpg, png, tiff, xls, xlsx

El procesamiento por lotes tiene varios pasos consecutivos:

- **Iniciar procesamiento por lotes**

Se genera un procesamiento por lotes. Éste será el objeto con el que trabajaremos durante todo el proceso. La respuesta nos generará un identificador único que usaremos en el resto de llamadas.

- **Subir archivo al procesamiento por lotes**

Añadir un archivo al procesamiento por lotes. Se indicará el identificador único del procesamiento por lotes creado. Será necesario realizar una llamada por cada archivo a añadir al procesamiento por lotes.

- **Subir carpeta al procesamiento por lotes**

Añadir un archivo en una carpeta al procesamiento por lotes. Se indicará el identificador único del procesamiento por lotes creado. Será necesario realizar una llamada por cada archivo a añadir al procesamiento por lotes.

- **Obtener archivos del procesamiento por lotes**

Obtenemos el listado de archivos añadidos al procesamiento por lotes. Se indicará el identificador único del procesamiento por lotes creado.

- **Comenzar procesamiento por lotes**

Una vez todos los archivos han sido subidos al procesamiento por lotes comenzará a procesarlos con esta llamada. Se indicará el identificador único del procesamiento por lotes creado.

- **Obtener el status del procesamiento por lotes**

Podemos consultar el estado de los procesamientos por lotes de nuestro usuario para saber si están aún en proceso o han terminado y si ha procesado todos los archivos o alguno ha fallado. Se indicará el identificador único del procesamiento por lotes creado.

- **Descargar zip del procesamiento por lotes**

Una vez un procesamiento por lotes ha concluido podemos descargarlo en formato zip indicando su identificador único en la propia dirección de la llamada.

a. Iniciar procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/initprocess>

Parámetros de entrada:

- **data:** objeto de configuración del procesamiento por lotes
- **(header):** [Encabezado](#)

Ejemplo de data:

```
{
  'bucket_configuration_input': 's3-nymiz',
  'bucket_configuration_output': 's3-nymiz',
  'batch_description': 'TESTING'
}
```

Salida:

```
{
  'id': 111,
  'startTimestamp': None,
  'endTimestamp': None,
  'statusDetails': None,
  'inputBucket': 'batch-processing-temporary-files-dev',
  'outputBucket': 'batch-processing-temporary-files-dev',
  'bucketConfigurationInput': 's3-nymiz',
  'bucketConfigurationOutput': 's3-nymiz',
  'batchDescription': 'TESTING',
  'totalFilesize': 0,
  'totalFilecount': 0,
  'successFilecount': None,
  'errorFilecount': None,
  'idUser': 1,
  'idCustomer': 1,
  'batchProcessFileList': None
}
```

El valor ID de la respuesta es el identificador único del procesamiento por lotes que hemos creado.

b. Subir archivo al procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/upload>

Parámetros de entrada:



- **file:** documento a incluir en el procesamiento por lotes
- **data:** objeto de configuración del procesamiento por lotes
- **(header):** [Encabezado](#)

Ejemplo de data:

```
{
  'bucket_configuration': 's3-nymiz',
  'key': '111'
}
```

c. Subir carpeta al procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/upload>

Parámetros de entrada:

- **file:** documento a incluir en el procesamiento por lotes
- **data:** objeto de configuración del procesamiento por lotes
- **(header):** [Encabezado](#)

Ejemplo de data:

```
{
  'bucket_configuration': 's3-nymiz',
  'key': '111',
  'subfolder': 'subfoldername'
}
```

d. Obtener archivos del procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/files>

Parámetros de entrada:

- **data:** objeto de configuración del procesamiento por lotes
- **(header):** [Encabezado](#)

Ejemplo de data:

```
{
```

```
'bucket_configuration': 's3-nymiz',  
'key': '111',  
'filelimit': '0'  
}
```

e. Comenzar procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/startprocess>

Parámetros de entrada:

- **key:** ID del procesamiento por lotes
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

f. Obtener el status del procesamiento por lotes

Endpoint:

(http POST): <https://api.nymiz.com/batch/status>

Parámetros de entrada:

- **(header):** [Encabezado](#)

Salida: Misma respuesta que en el initprocess pero para todos los procesamientos por lotes

g. Descargar zip del procesamiento por lotes

Endpoint:

(http POST): [https://api.nymiz.com/batch/downloadZip/\[ID\]](https://api.nymiz.com/batch/downloadZip/[ID])

Parámetros de entrada:

- **(header):** [Encabezado](#)

Salida: Archivo Zip con todos los archivos procesados.

10. Anonimización de JSON

Se aceptan los siguientes formatos: .json

Endpoint:

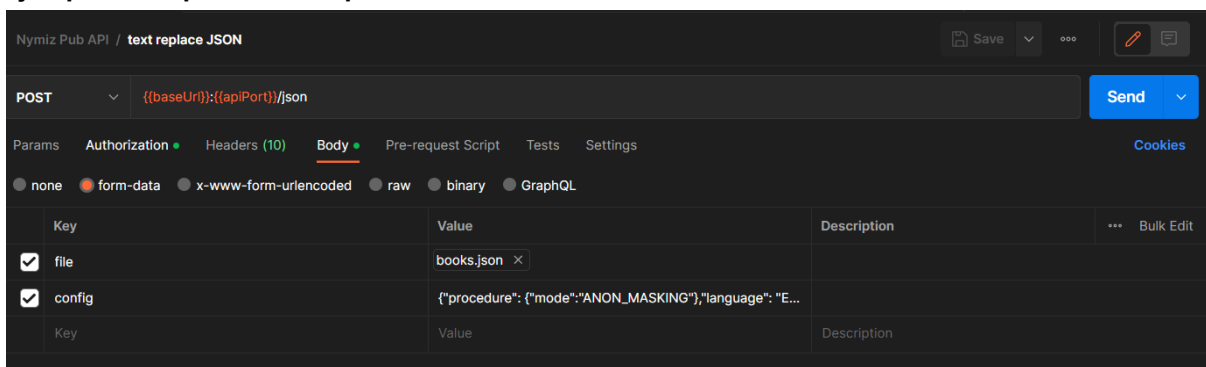
(http POST): <https://api.nymiz.com/json>

Parámetros de entrada:

- **file:** documento a anonimizar
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
 - Nota: Es importante incluir al menos un jsonpath para su anonimización
 - Sólo se puede anonimizar en modo ANON_MASKING
- **(header):** [Encabezado](#)

Salida: documento anonimizado

Ejemplo con aplicación de pruebas Postman:



11. Extracción de contenido de documentos en crudo (OCR a JSON)

Endpoint:

(http POST): <https://api.nymiz.com/ocr-json>

Tipos de documentos soportados:

- doc
- docx
- pdf
- tif / tiff
- jpg / jpeg
- png

Parámetros de entrada:

- **file:** documento para la extracción de información.
- **(header):** [Encabezado](#)

Salida: Salida de OCR en formato JSON del contenido del documento.

12. Supervisión Manual Embebida

Endpoint:

(http POST): <https://api.nymiz.com//url-viewer>

Tipos de documentos soportados:

- pdf

Parámetros de entrada:

- **file:** documento para la extracción de información.
- **config:** configuración de la anonimización. Ver [Objeto de configuración para los documentos y texto plano](#)
- **(header):** [Encabezado](#)

Salida: URL para abrir el documento (adjunto en el parámetro **file**) en Supervisión Manual. Esta url se puede abrir en un navegador o en un iframe.

(http POST): <https://api.nymiz.com//url-files>

Parámetros de entrada:

- **uuid:** Id de la sesión de la Supervisión Manual. Este parámetro se debe pasar como Query Param y no en el body.
- **(header):** [Encabezado](#)

Salida: Objeto con los siguientes parámetros:

- **linearizedFileUrl:** Url del documento a procesar
- **nymiztaskUrl:** Url del json con los datos del documento.

13. Objeto de configuración para los documentos y texto plano

Permite establecer los parámetros por medio de los cuales se llevará a cabo el proceso de protección.

Ejemplo:

```
{
  "language": "es-ES",
  "procedure": {
    "mode": "ANON"
  },
  "forceNLPCellByCell": false,
  "skipHeaderRow": true,
  "excluded_entities": [
    "DAT"
  ],
  "blacklist": [
    "Julio"
  ],
  "whitelist": [
    "Nymiz"
  ],
  "jsonpath": [
    "$.store.book[0].title",
    "$.store.book[*].author"
  ],
  "options": {
    "ocrPremium": false,
    "containsVerticalText": false
  }
}
```

En este caso se ha elegido especificar que el contexto del lenguaje es español de España, el procedimiento es seudonimización y la reversión se protege con la contraseña "123456". También, que se deberán excluir de la seudonimización los infotipos de nombres de persona y las fechas. Por último, se establece que se deberá anonimizar la palabra "julio" sin importar cómo sea detectada pero que en cambio, la palabra "Nymiz" se debe dejar en abierto, sin seudonimizar.

Referencia

- **language:** opcional, permite optimizar la extracción de texto durante el proceso de OCR sobre imágenes y PDFs y se usará en el futuro para casos de modelos de NLP especializados. Utilizar las combinaciones habituales de ISO 639-1 y 3166-1 (ej.: en_us, en_gb, es_es, pt_br, pt_pt, etc.)
- **procedure**
 - **mode:**
 - ANON_MASKING → anonimización por masking/blacklining (opción por defecto)
 - ANON_TOKEN → tokenización/clasificación ("PER", "DAT", "PHO", etc.)
 - ANON_SYNT → Sustitución por datos sintéticos de misma naturaleza y similar extensión

- PSEUDO_TOKENIZATION → seudonimización (proceso reversible)
- PSEUDO_SYNT → seudonimización por sustitución por datos sintéticos
- **password**: solamente en caso de seudonimización (proceso reversible) se debe proporcionar la password (salt) para el cifrado
- **excluded_entities**: listado de los infotipo/s que se deben excluir del proceso (ver [Listado de infotipos soportados](#))
- **blacklist**: excepciones negativas, palabras que serán protegidas (anonimizadas o seudonimizadas) sin importar si no son detectadas como datos personales
- **whitelist**: excepciones positivas, palabras que NO serán protegidas (anonimizadas o seudonimizadas) aunque sean detectadas como dato personal
- **jsonpath**: paths que determinan los elementos que quieren ser anonimizados dentro de la estructura del json. El path puede designar:
 - Un elemento concreto dentro de la estructura principal del archivo
 - Todos los elementos de un tipo dentro de la estructura principal del archivo
 - Una subestructura completa o parcial dentro de la estructura principal
- **Options**: Objeto con los flags para activar distintas herramientas para el procesado de documentos. Los posibles flags son:
 - **forceNLPCellByCell**: Para excel. Campo booleano (true o false). Activa o desactiva la opción de analizar celda por celda.
 - False: Nymiz analiza una cantidad limitada de campos en cada columna y determina el infotipo basándose en las muestras analizadas. Se anonimizan todas las celdas de la columna y todo su contenido.
 - True: Nymiz analiza cada celda en busca de información que anonimizar. Solo anonimiza aquellas palabras que considere que deben ser anonimizadas.
 - **skipHeaderRow**: Para excel. Campo booleano (true o false). Evita anonimizar la primera línea que por lo general se reserva para los nombres de columna.
 - **ocrPremium**: Para pdf: Campo booleano (false o true). Si el valor es true el OCR extrae de forma explícita el contenido que haya en tablas y layouts. Si el valor es false, el contenido que haya en tablas y layouts lo extrae como párrafos.
 - **containsVerticalText**: Para pdf: Campo booleano (false o true). Si el valor es true el OCR busca en las páginas si existe texto vertical en los bordes.
 - **partialMasking**: Campo booleano. Activa o desactiva el modo de anonimización por enmascaramiento parcial siguiendo las orientaciones AEPD.
 - **consistenceExceptions**: Campo booleano. Activa o desactiva la consistencia en las excepciones.
 - **qr**: Para PDF. Campo booleano. Activa o desactiva la detección y anonimización de códigos QR. En caso de anonimizar siempre se hará con tachón y esta acción no es reversible.
 - **barcode**: Para PDF. Campo booleano. Activa o desactiva la detección y anonimización de códigos de barras. En caso de anonimizar siempre se hará con tachón y esta acción no es reversible.
 - **holographicSignatur:e** Para PDF. Campo booleano. Activa o desactiva la detección y anonimización de firmas manuscritas. En caso de anonimizar siempre se hará con tachón y esta acción no es reversible.



- **removeSigns:** Para PDF. Campo booleano. Activa o desactiva la detección y anonimización de firmas digitales. En caso de estar activo se eliminará la firma del documento y esta acción no es reversible.
- **manualSupervision:** Para PDF. Campo booleano. Activa o desactiva la revisión manual del documento una vez finalizado el proceso.
- **profileDetection:** Campo booleano. Activa o desactiva la detección de perfiles jurídicos (letrados y abogados).



14. Listado de infotipos soportados

ADD	Dirección / domicilio
CSV	Código seguro de verificación
CUR	Moneda, cantidades monetarias
DAT	Fecha
IBA	IBAN, número de cuenta bancaria
IPD	Dirección IP
LOC	Localización geográfica (ciudades, provincias, etc.)
NFJ	NIF persona jurídica
NFP	NIF persona física
NIE	Número de Identidad de Extranjero
NOR	Norma, Decreto o Circular
NSS	Número de Seguridad Social / SSN
ORG	Nombre de empresa u organización
PAS	Pasaporte
PER	Nombre de persona
PHO	Número de teléfono
PLA	Matrícula
WEB	Sitio web y correo electrónico

15. Encabezados

A excepción del método de login el resto de métodos utilizan un token de autorización.

- **Authorization:** token obtenido en el Login, con el tipo "Bearer Token"
- **Content-Type:** Si enviamos un **documento** junto al objeto "**Config**" es necesario establecer el **content-type** en la llamada como **multipart/form-data**. En ocasiones, también es necesario establecer **el Content-Type del archivo** que se está añadiendo en ese Multipart si la librería o software usado no lo hace por sí mismo. Si no se añade es posible recibir un error de **UnsupportedMediaType**.

