

ONBOARDING DIGITAL

DOCUMENTACIÓN

Autor	Nicolás Aldecoa Rodrigo
Revisores	Adrián Pérez Víctor Vivas Cedeño
Fecha de Entrega	26/11/2021
Versión	1.0.7

ÍNDICE

Datos y Validaciones	3
DNI	3
Selfies	3
Operación	3
Detalles y Ejemplos	4
Ejemplo de Operación Satisfactoria	4
Anti-Spoofing (selfies)	5
Match de Contexto	6
Match Frente vs. Dorso	6
Módulo DocAlign	7
Descripción y recomendaciones	7
Parámetros Configurables	8
Panel de Control	8
RENAPER: DNI vigente	8
Modo Flexible	8
Tolerancia Match Facial	9
Tolerancia Anti-Spoofing (selfies)	9
Consultas a la API	10
Método	10
Endpoint 1	10
Endpoint 2	10
Endpoint 3 – Biometría Simple	12
Endpoint 3 – Biometría Múltiple	13
Notas	14
Aclaraciones	14
Tabla: Códigos de Prueba de Vida	14

DATOS Y VALIDACIONES

Para cada imagen de DNI recibida:

- **FRENTE:** indica qué entidad fue reconocida en la imagen recibida como frente ("front", "back", desconocida).
- **DORSO:** indica qué entidad fue reconocida en la imagen recibida como dorso ("front", "back", desconocida).
- **CALIDAD DE IMAGEN:** se devolverá una estimación automática de la calidad general de la imagen recibida (buena, aceptable o mala). Esta es obtenida a partir de mediciones de características propias de la fotografía como resolución, calidad del enfoque, iluminación, grado de compresión digital, etc.
- **DOCUMENTO SOSPECHOSO:** será verdadero si nuestros modelos de anti-spoofing detectan actividad sospechosa en la fotografía (recaptura, imagen alterada digitalmente, etc.). Si la imagen no presenta anomalías, el contenido de este campo será falso. La tolerancia de esta validación puede ajustarse desde el panel de configuración de usuario.

Para cada selfie recibida:

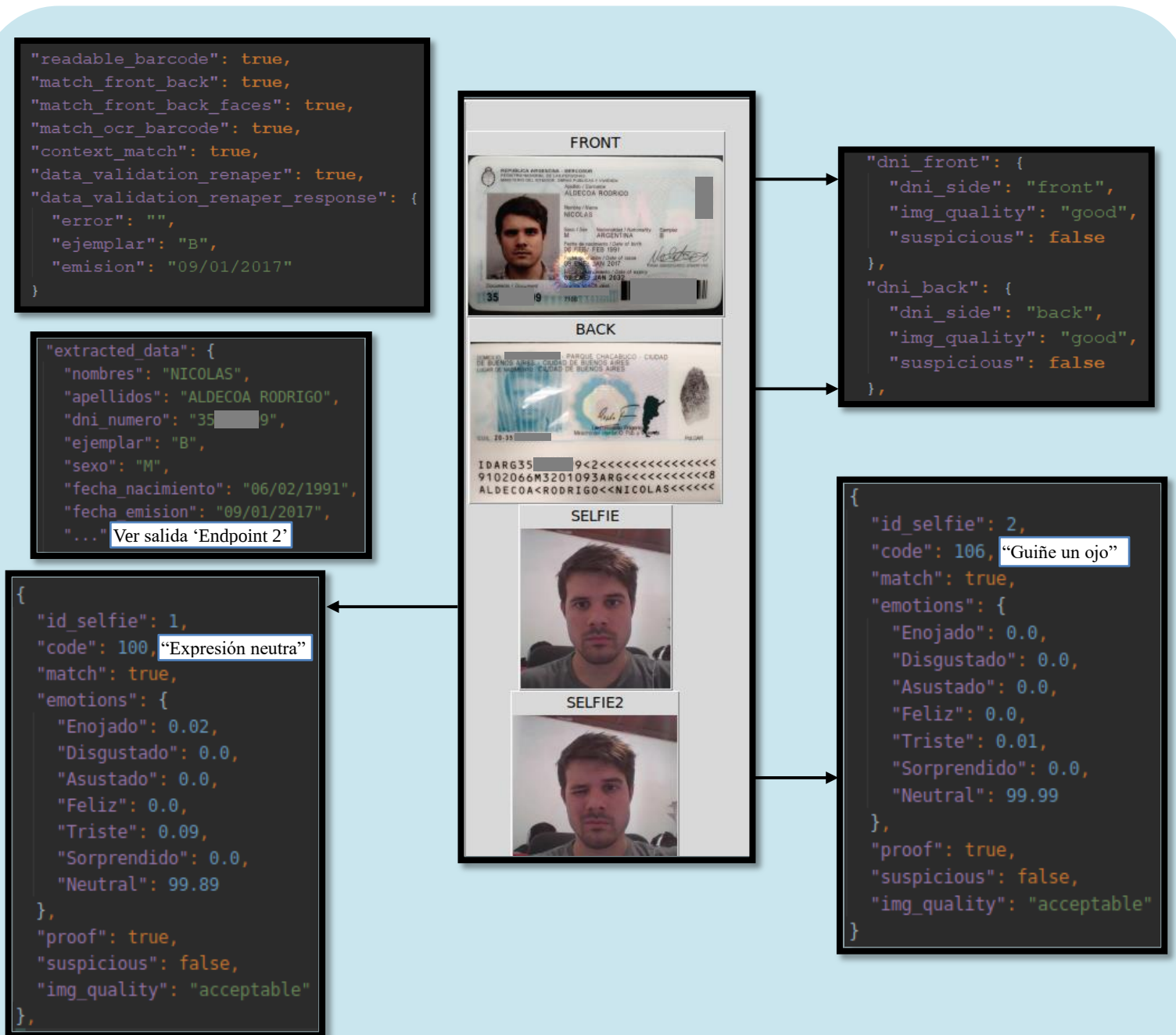
- **CALIDAD DE IMAGEN:** ídem imágenes de DNI.
- **MATCH:** indica el resultado de la verificación de identidad de la selfie contra la foto carnet del frente del DNI. La tolerancia de este control puede ajustarse desde el panel de configuración de usuario.
- **EMOCIONES:** devuelve el resultado de la estimación de emociones a partir de las expresiones faciales del individuo. A cada una de las 7 emociones contempladas se le asocia una probabilidad expresada como porcentaje.
- **PRUEBA DE VIDA:** indica si el gesto detectado en la selfie corresponde con el gesto solicitado al usuario. *Ver tabla de códigos de solicitud al final del documento.*
- **SELFIE SOSPECHOSA:** ídem imágenes de DNI. La tolerancia de esta validación puede ajustarse desde el panel de configuración de usuario.

Para la operación en general:

- **CÓDIGO DE BARRAS LEGIBLE:** indica si nuestro sistema fue capaz de leer el código de barras PDF-417 (buscará en el frente en el caso de un DNI de formato nuevo, o en el dorso si trata de un DNI "viejo").
- **MATCH CONTEXTUAL:** será verdadero si el sistema estima que todas las selfies recibidas pertenecen a un mismo contexto sin ser idénticas. Devolverá "false" cuando las fotos no presenten similitud de contexto o cuando sean exactamente iguales.
- **MATCH FRENTE-DORSO:** será verdadero si los datos del frente del DNI corresponden con los del dorso.
- **MATCH FACIAL FRENTE-DORSO:** será verdadero si el rostro del frente del DNI corresponde con el del dorso.
- **MATCH OCR-CÓDIGO DE BARRAS:** será verdadero si la información extraída mediante OCR (texto impreso) es consistente con la extraída del PDF-417. La comparación de datos posee un margen de tolerancia a errores puntuales que son esperables en el OCR.
- **VALIDACIÓN DE DATOS CONTRA RENAPER:** nuestro servicio también realiza una consulta a RENAPER para verificar la vigencia del DNI y la validez de los datos extraídos. Este procedimiento genera dos campos en nuestra respuesta: uno de validación, que será verdadero cuando la respuesta de RENAPER sea satisfactoria; y otro que contendrá dicha respuesta en su formato original.
- **DATOS EXTRAÍDOS:** datos extraídos del DNI, estructurados como {"campo": "valor"}. Si el código de barras es legible, la información del mismo será utilizada para corregir la del OCR: en el caso de que un dato no hubiera podido ser extraído por el OCR y se encuentre presente en el código de barras, será autocompletado. En cambio, si el dato se obtuvo por ambos medios, sólo se corregirá el OCR con el PDF-417 si el MATCH OCR-CÓDIGO DE BARRAS fuese verdadero.
Las fechas tendrán el formato estandarizado "DD/MM/AAAA" siempre que sea posible. Si no, el campo estará precedido por el texto "<best_try>" (por ejemplo: "<best_try>N / JUN 2020").
- **CONTROL OK:** este campo fue agregado para simplificar la operatoria de nuestros clientes. Tendrá valor verdadero únicamente cuando todas las validaciones configuradas por el usuario en su panel de control como "requeridas" sean satisfactorias.

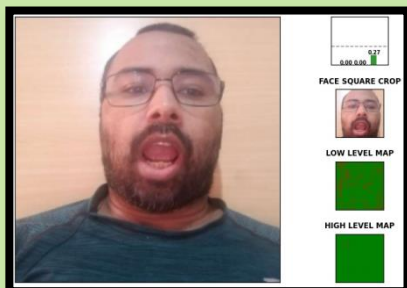
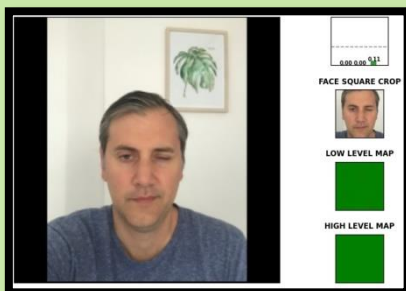
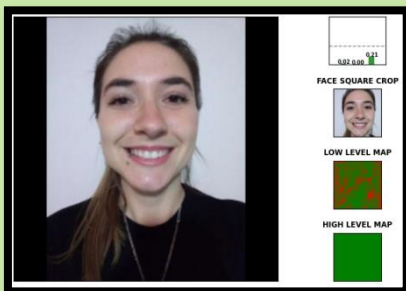
DETALLES Y EJEMPLOS

Ejemplo de Operación Satisfactoria

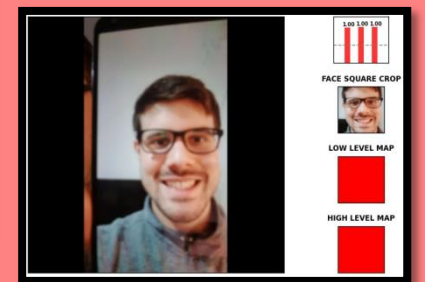
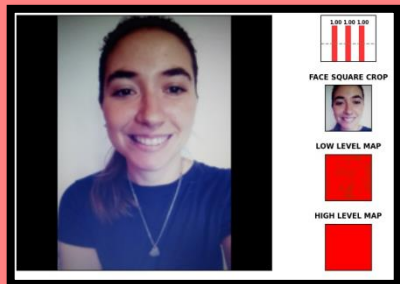
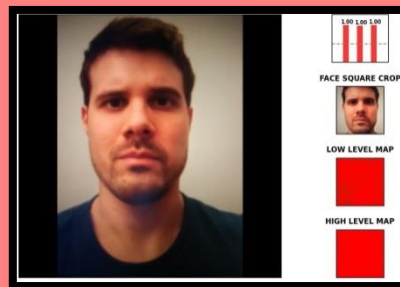


✓ Para combatir robos de identidad, todas las selfies y DNIs pasan por nuestros modelos de anti-spoofing que son continuamente entrenados para detectar intentos de fraude con fotos recapturadas de los ID de la víctima o de sus perfiles de redes sociales. A continuación se muestran algunos ejemplos de intentos de autenticación legítimos e ilegítimos efectuados por miembros del equipo de EntreConsultas y la respuesta del módulo de Anti-Spoofing de selfies.

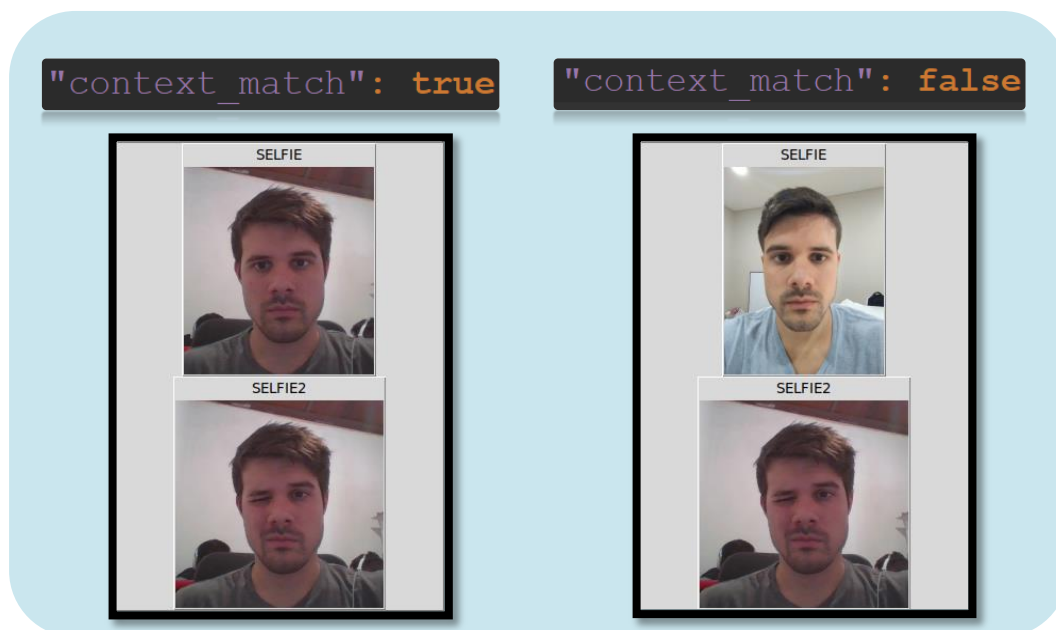
"suspicious": false



"suspicious": true



- ✓ En los casos en que se envíen archivos de imagen a la plataforma en lugar de sólo permitir la toma de fotos en el momento con la cámara del celular (no recomendado), los ladrones de identidad pueden utilizar fotos descargadas de redes sociales para intentar pasar las pruebas de vida. Para prevenir esto, nuestro servicio verifica que el contexto de todas las selfies sea consistente.



- ✓ Además de la validación mediante cruce de datos legibles entre el frente y dorso del documento, nuestro sistema realiza una comparación del rostro de la foto carnet del frente contra la imagen monocolor modulada en guilliches que figura en el dorso. Esto impide que se pueda burlar el sistema de reconocimiento facial reemplazando la foto del frente por la de un impostor.



MÓDULO: DocAlign

Nuestro servicio posee un módulo de alineamiento que fue desarrollado con el objetivo de posibilitar la obtención de datos a partir de imágenes que contienen documentos mal encuadrados, rotados y con perspectiva oblicua. DocAlign consiste en un proceso de dos pasos: uno de reconocimiento/localización, y otro de alineamiento/estandarización.



Aunque DocAlign resulta indispensable cuando el usuario de nuestro servicio no tiene manera de controlar la calidad del encuadrado de las imágenes que toman sus clientes, hemos comprobado en múltiples ensayos que los mejores resultados se obtienen cuando es posible llevar a cabo los siguientes procedimientos:

- Desactivar en el panel de control el módulo DocAlign, ya que se enviarán las imágenes ya encuadradas.
- Desplegar en el front-end una máscara de relación de aspecto (ancho sobre alto) igual a 1,6 para que el cliente encuadre el documento apaisado con la ayuda de la aplicación, idealmente obligando a que la foto sea tomada con el smartphone paralelo al documento.
- Enviar a nuestro servicio la imagen recortada (sólo los píxeles correspondientes al interior de la máscara, o sea el documento, con una resolución óptima de 1152x720 px).

Esto es porque tanto la localización como la transformación de la imagen tienen sus respectivas chances de error que, aunque juntas acumulan una probabilidad menor al 5%, pueden fallar con fotografías de mala calidad, credenciales que se encuentren en mal estado (lo cual suele ser común con los DNI formato "viejo") o cuando los bordes del documento quedan fuera de la imagen.

En caso de que las imágenes enviadas a la API sean de buena calidad y se haya hecho el trabajo de alinear el documento en el momento de tomar la fotografía, es recomendable desactivar DocAlign para reducir el tiempo de consulta y evitar las posibilidades de errores intrínsecos al método.

Por estas razones recomendamos a nuestros usuarios que si ya están realizando un trabajo de encuadrado y control de calidad en el front-end para la toma de imagen, evalúen el servicio con DocAlign activado y desactivado para determinar qué configuración se adecua mejor a su proceso.

PARÁMETROS CONFIGURABLES

- ✓ Desde su panel de control, nuestros usuarios pueden configurar el servicio acorde a sus preferencias, monitorear su flujo de operaciones, acceder a estadísticas de validaciones y realizar pruebas contra la API usando una interfaz gráfica (ver <https://obd.entreconsultas.com:3000>).

The screenshot displays a configuration panel with the following sections:

- Configuración:**
 - Modo Flexible: ☒
 - Usar DocAlign: ☒
 - Expresiones Faciales: ☒
- Control de validaciones:**
 - Activar Control OK: ☒
 - Frente DNI: ☒
 - Match Facial: ☒
 - Selfie Sospechosa: ☒
 - Barcode legible: ☒
 - Match OCR-Barcode: ☒
 - RENAPER - Último ejemplar vigente DNI: ☒
 - Dorso DNI: ☒
 - Prueba de vida: ☒
 - Contexto Selfies: ☒
 - Match frente-dorso: ☒
 - Match de rostros Frente-Dorso: ☒
 - DNI Sospechoso: ☒
- Parámetros:**
 - Tolerancia Match Facial: 0.70
 - Tolerancia AntiSpoofing Selfie: 0.50
 - Tolerancia AntiSpoofing DNI: 0.50

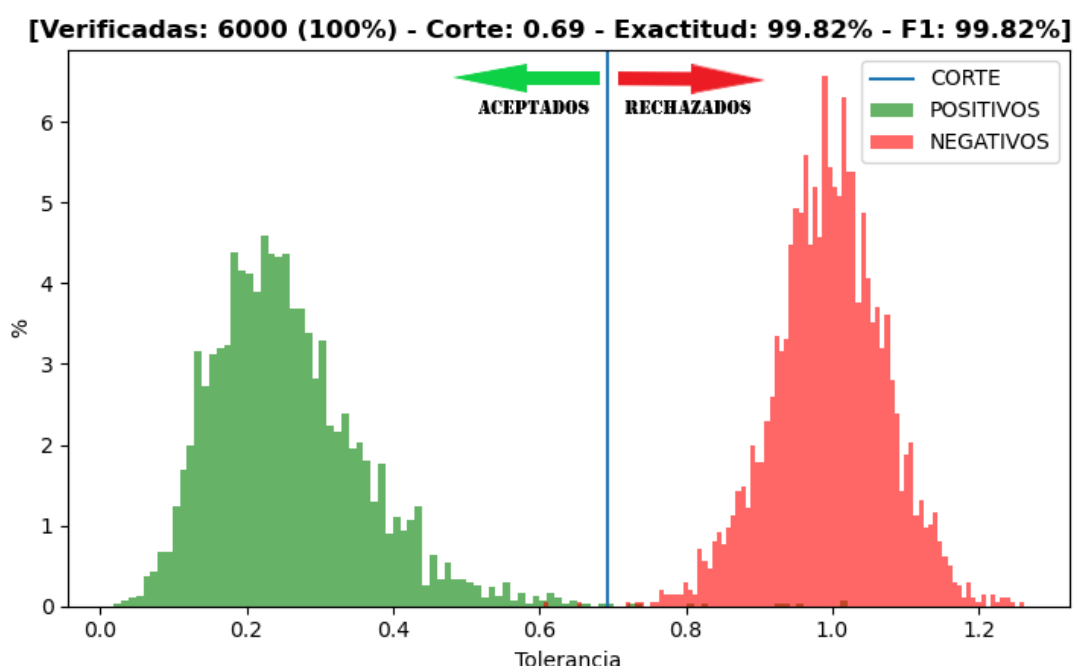
An "Aplicar cambios" button is located at the bottom right of the configuration area.

- ✓ La consulta de **vigencia de DNI** y validez de los datos extraídos que se realiza al Registro Nacional de las Personas (RENAPER) consiste en el envío de tres datos: “*número de DNI*”, “*sexo*” y “*fecha de nacimiento*”, a los que RENAPER responde con “*ejemplar*” y “*fecha de emisión*” del último DNI vigente en caso de que sean datos válidos, o con un mensaje de error de verificación en caso contrario. Nuestro servicio cruza los datos de la respuesta con los extraídos del DNI recibido y, en caso de que el “*ejemplar*” y la “*fecha de emisión*” sean los mismos, considerará el resultado de “*data_validation_renaper*” como verdadero. Además, nuestra respuesta contendrá en el campo “*data_validation_renaper_response*” la información recibida de RENAPER en su formato original.

- ✓ La activación del “**Modo Flexible**” indica a nuestro sistema que todos los procesos de extracción de datos, biometría y validaciones deben ser realizadas siempre que sea posible, sin importar el resultado negativo de validaciones previas. Ejemplos de validaciones que de estar el modo flexible desactivado abortarían el resto de los procesos sobre la imagen en cuestión son: tamaño del rostro detectado demasiado chico, rostro localizado demasiado cerca de los bordes de la imagen (posiblemente cortado), ángulo de toma de selfie inaceptable (parte del rostro no es visible), o que la fotografía esperada (frente de DNI, dorso de DNI o selfie) no sea reconocida como tal. Desactivar el “Modo Flexible” hará que ante cualquiera de estos casos se omita el resto de los procesos sobre la imagen, reduciendo así el tiempo de respuesta en transacciones con baja probabilidad de ser completadas exitosamente.

✓ Para ofrecer un mayor grado de personalización de nuestro servicio, en el **panel de control (configuración)** se irán disponibilizando gradualmente las opciones de modificar los parámetros de ciertos algoritmos de decisión involucrados en las validaciones.

- **Tolerancia Match Facial:** corresponde a la máximo grado de diferencia entre rostros que es tolerado en la verificación de identidad mediante biometría facial. A mayores valores de tolerancia, la validación se vuelve menos estricta, reduciendo la probabilidad de falsos negativos a expensas de aumentar el riesgo de falsos positivos, y viceversa. El valor por defecto para este parámetro es **0,70** y fue elegido tras realizar un exhaustivo análisis sobre 9250 validaciones de identidad con imágenes de tipo y calidad variables. Éste puede ser ajustado entre **0,50 (más estricto)** y **0,80 (más laxo)**. El comportamiento de la medición de distancias entre rostros con nuestro algoritmo se observa en el siguiente gráfico de histogramas*.



* Los resultados que se presentan en este gráfico corresponden a una evaluación realizada sobre 6000 validaciones con fotografías de buena calidad. Con imágenes de calidad variable entre aceptable y mala, la exactitud promedio decae a un 97.9%, manteniendo 0,70 como el valor óptimo de tolerancia (resultado obtenido a partir de 3250 validaciones de foto carnet de DNIs fotografiados contra selfies tomadas por 1625 usuarios distintos bajo condiciones variadas).

- **Tolerancia Anti-Spoofing:** este valor corresponde a la confianza mínima que debe haberse obtenido de nuestros modelos de anti-spoofing para que una imagen sea marcada como sospechosa. A mayores valores de tolerancia, la validación se vuelve menos estricta, reduciendo la probabilidad de falsos negativos a expensas de aumentar el riesgo de falsos positivos, y viceversa. El valor por defecto de este parámetro para selfies es **0,50** con un rango de ajuste permitido entre **0,40 (más estricto)** y **0,90 (más laxo)**. Para DNIs el valor por defecto de la tolerancia es también **0,50**, y puede ajustarse entre **0,10 (más estricto)** y **0,90 (más laxo)**.

CONSULTAS A LA API

URL: <https://obd.entreconsultas.com/api/validar/>

Header: {"Authorization": "Token <string_hexadecimal>"}

El contenido de la consulta debe tener método **POST**, contenido en formato **JSON** y respetar cualquiera de las estructuras de Entrada indicadas a continuación.

Aclaración: cualquiera de los campos del JSON de respuesta puede tener valor nulo ('null'). Esto indica que la validación correspondiente no pudo ser realizada o que no se pudo extraer la información correspondiente a dicho campo.

EndPoint 1

Entrada Frente DNI:

```
{
  "id_op": int,
  "dni_front": string (base 64)
}
```

Salida Frente DNI:

```
{
  "id_op": int,
  "dni_front": string ("front"/ "back"),
  "img_quality": string ("good"/ "acceptable"/ "bad"),
  "suspicious": boolean,
  "readable_barcode": boolean,1
  "ban_list_hit": boolean,2
  "control_ok": boolean,
  "description": string
}
```

EndPoint 2

Entrada Ambos Lados DNI:

```
{
  "id_op": int,
  "dni_front": string (base 64),
  "dni_back": string (base 64)
}
```

¹ "readable_barcode" tendrá valor null cuando se trate del frente de un DNI "viejo" que no contiene código PDF-417. Por esta razón, el valor de este campo no es tenido en cuenta para la validación "control_ok" en este Endpoint.

² "ban_list_hit" es un campo reservado para una validación que actualmente se encuentra en desarrollo. Su contenido será falso ('false') hasta próximo aviso.

Salida Ambos Lados DNI:

```
{
  "id_op": int,
  "dni_front": {
    "dni_side": string ("front" / "back"),
    "img_quality": string ("good" / "acceptable" / "bad"),
    "suspicious": boolean,
  },
  "dni_back": {
    "dni_side": string ("front" / "back"),
    "img_quality": string ("good" / "acceptable" / "bad"),
    "suspicious": boolean,
  },
  "readable_barcode": boolean,
  "match_front_back": boolean,
  "match_front_back_faces": boolean,
  "match_ocr_barcode": boolean,
  "data_validation_renaper": boolean,
  "data_validation_renaper_response": {1
    "error": string,
    "ejemplar": string,
    "emision": string
  },
  "extracted_data": {
    "nombres": string,
    "apellidos": string,
    "dni_numero": string,
    "ejemplar": string,
    "sexo": string,
    "fecha_nacimiento": string,2
    "fecha_emision": string,2
    "fecha_vencimiento": string,2
    "nacionalidad": string,
    "numero_tramite": string,
    "dni_numero2": string,
    "domicilio": string,
    "cuil": string,
    "lugar_nacimiento": string,
    "fecha_ingreso": string,
    "categoria_ingreso": string,
    "expediente": string,
    "disposicion": string,
    "fecha_radicacion": string,
    "codigo_dorso": string
  },
  "ban_list_hit": boolean,
  "control_ok": string,
  "description": string
}
```

¹ Este campo contiene la respuesta original que nuestro servicio recibió de RENAPER.

² Los campos indicados serán convertidos al formato del código de barras del DNI ("DD/MM/AAAA"). En caso de no ser posible a partir de los datos obtenidos, la cadena de caracteres comenzará con "<best_try>" y continuará con el texto que se haya podido leer mediante OCR.

EndPoint 3

Entrada Biometría Simple:

```
{
  "id_op": int,
  "dni_front": string (base 64),
  "selfies": [
    {
      "image": string (base 64),
      "id_selfie": int,
      "code": int
    }
  ]
}
```

Salida Biometría Simple:

```
{
  "id_op": int,
  "dni_front": {
    "dni_side": string ("front" / "back"),
    "img_quality": string ("good" / "acceptable" / "bad"),
    "suspicious": boolean,
  },
  "selfies": [
    {
      "id_selfie": int,
      "code": int,
      "match": boolean,
      "emotions": {
        "Enojado": float,
        "Disgustado": float,
        "Asustado": float,
        "Feliz": float,
        "Triste": float,
        "Sorprendido": float,
        "Neutral": float
      },
      "proof": boolean,
      "suspicious": boolean,
      "img_quality": string ("good" / "acceptable" / "bad"),
    }
  ],
  "context_match": null,1
  "ban_list_hit": boolean,
  "control_ok": boolean,
  "description": string
}
```

¹ "context_match" siempre tendrá valor nulo ('null') cuando la consulta sea realizada con una única selfie. En este caso, el valor de este campo no es tenido en cuenta para la validación "control_ok".

Entrada Biometría Múltiple (difiere en que la lista “selfies” posee 2 elementos):

```
{
  "id_op": int,
  "dni_front": string (base 64),
  "selfies": [
    {
      "image": string (base 64),
      "id_selfie": int,
      "code": int
    },
    {
      "image": string (base 64),
      "id_selfie": int,
      "code": int
    }
  ]
}
```

Salida Biometría Múltiple (la lista “selfies” posee 2 elementos y es posible realizar el match de contexto):

```
{
  "id_op": int,
  "dni_front": {
    "dni_side": string ("front" / "back"),
    "img_quality": string ("good" / "acceptable" / "bad"),
    "suspicious": boolean,
  },
  "selfies": [
    {
      "id_selfie": int,
      "code": int,
      "match": boolean,
      "emotions": {
        "Enojado": float,
        "Disgustado": float,
        "Asustado": float,
        "Feliz": float,
        "Triste": float,
        "Sorprendido": float,
        "Neutral": float
      },
      "proof": boolean,
      "suspicious": boolean,
      "img_quality": string ("good" / "acceptable" / "bad"),
    },
    {
      "id_selfie": int,
      "code": int,
      "match": boolean,
      "emotions": {
        "Enojado": float,
        "Disgustado": float,
        "Asustado": float,
        "Feliz": float,
        "Triste": float,
        "Sorprendido": float,
        "Neutral": float
      },
      "proof": boolean,
      "suspicious": boolean,
      "img_quality": string ("good" / "acceptable" / "bad"),
    }
  ],
  "context_match": boolean,
  "ban_list_hit": boolean,
  "control_ok": boolean,
  "description": string
}
```

Notas:

- El campo "control_ok" tendrá valor 'true' cuando todas las validaciones que hayan sido activadas por el cliente en el panel de control resulten satisfactorias ('true' para todos los campos con excepción de los denominados "suspicious" que en casos legítimos se esperan con valor 'false').
- El campo "description" está reservado para proveer información adicional en algunos casos.

Aclaraciones:

1. Que el contenido de un campo sea nulo ('null') significa que el dato correspondiente no fue encontrado o, en el caso de una validación, que ésta no pudo ser realizada (no confundir con 'false' que es un resultado de un procedimiento que efectivamente fue realizado).

- Ejemplo 1: "readable_barcode": 'null'. No significa que el código de barras no fue encontrado, si no que no pudo ser buscado (por ejemplo, por falta de la imagen del lado correspondiente del DNI).
- Ejemplo 2: "proof": 'null' para una selfie indica que no se realizó la detección de gestos (por ejemplo, por no haberse encontrado un rostro en la imagen). En cambio, "proof: false" indica que se realizó la detección de gestos y no hubo coincidencia con los tolerados por el código solicitado.
- Ejemplo 3: "ejemplar": 'null' en "extracted_data" indica que no pudo encontrarse ese dato en el DNI.

2. Al enviar una selfie, se verificará que el gesto de la persona corresponda con el que representa el código enviado (ver tabla de gestos a continuación). Si no es necesario verificar ningún gesto, debe enviarse el código '200'. No confundir esto con la expresión neutra que sí es un tipo de gesto para el sistema y corresponde al código '100'. En caso de recibir '200' se asume que la prueba de vida es superada por defecto: no se realizará la detección de gestos y el campo "proof" tendrá valor verdadero en la respuesta siempre y cuando la selfie sea válida.

Códigos de Prueba de Vida

Detector de Gestos (Backend)						Manejo desde la API		
Hash Binario	Código Entero	Ojo Izquierdo Cerrado	Ojo Derecho Cerrado	Boca Abierta	Sonrisa	Código de Solicitud*	Descripción de la Solicitud	Códigos Enteros Aceptados
1100	1	X	X	X	X	100	"Mantenga una expresión neutra."	1
1101	2	X	X	X	✓	101	"Guiñe el ojo izquierdo."	9, 10, 11, 12
1110	3	X	X	✓	X	102	"Guiñe el ojo derecho."	5, 6, 7, 8
1111	4	X	X	✓	✓	103	"Abra la boca."	3, 4, 7, 8, 11, 12, 15, 16
1000	5	X	✓	X	X	104	"Guiñe el ojo izquierdo y abra la boca."	11, 12
1001	6	X	✓	X	✓	105	"Guiñe el ojo derecho y abra la boca."	7, 8
1010	7	X	✓	✓	X	106	"Guiñe un ojo."	5, 6, 7, 8, 9, 10, 11, 12
1011	8	X	✓	✓	✓	107	"Guiñe un ojo y abra la boca."	7, 8, 11, 12
0100	9	✓	X	X	X	108	"Sonría."	2, 4, 6, 8, 10, 12, 14, 16
0101	10	✓	X	X	✓	109	"Guiñe un ojo y sonría."	6, 8, 10, 12
0110	11	✓	X	✓	X	110	"Guiñe el ojo derecho y sonría."	6, 8
0111	12	✓	X	✓	✓	111	"Guiñe el ojo izquierdo y sonría."	10, 12
0000	13	✓	✓	X	X	112	"Guiñe un ojo." (tolera ambos cerrados)	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16
0001	14	✓	✓	X	✓	113	"Cierre los ojos."	13, 14, 15, 16
0010	15	✓	✓	✓	X	...	...	...
0011	16	✓	✓	✓	✓	200	NO EVALUAR GESTOS	-

* el código de solicitud debe ser generado en el momento en que el usuario realiza la biometría y enviado a la API en el campo "code" del JSON, asociado a la selfie codificada en base64.