Inland Revenue
Te Tari Taake

Inland Revenue

# Build Pack: Document Service

**Date:** 06/04/2020
**Version:** v1.0

## Contents

# 1 Overview

## 1.1 This solution

Inland Revenue has a suite of digital services available for consumption by our service providers that supports efficient, electronic business interactions with Inland Revenue. The Document Service described in this build pack document enables documents to be retrieved by external software platforms as well as be submitted by external software platforms.

Documents can comprise of mail generated by Inland Revenue for communicating with customers, as well as documents request by Inland Revenue from a customer, and documents a customer voluntarily provides to Inland Revenue. Documents from Inland Revenue are PDFs, while documents provided to Inland Revenue can be of many different file formats.

Before you continue, please be sure to consult
**http://www.ird.govt.nz/software-providers/**
for the products that use this service, business-level context and use cases, links to relevant policy, and information on how to integrate with Inland Revenue's products and services.

## 1.2 Intended audience

The solution outlined in this document is intended to be used by payroll providers, tax practitioners, KiwiSaver providers, banks and other financial institutions (referred to throughout the remainder of this document as 'Digital Service Providers').

The reader is assumed to have a suitable level of technical knowledge in order to comprehend the information provided. A range of technical terms and abbreviations are used throughout this document, and while most of these will be understood by the intended readers, a glossary is provided at the end.

## 1.3 Prerequisites

| Party | Requirement | Description |
|---|---|---|
| **Digital Service Provider** | Acquire a X.509 certificate from a competent authority for the Test and Production environments | This is required when using mutual TLS with cloud-based service providers or financial institutions.<br>Note that the same certificate cannot be used for the Test and Production environments. |

### 1.3.1 Mutual transport layer security and certificates

Mutual transport layer security (TLS) is implemented for this API. This requires the use of a publicly-issued X509 certificate from one of the trusted certificate authorities. Inland Revenue does not issue certificates to external vendors for web service security implementations.

Inland Revenue has the following minimum requirements for accepting public X509 keys:

- Minimum Key Length: 2048
- Signature Algorithm: SHA256[RSA]
- Self-signed certificates are not accepted
- Certificates issued by a private/internal certificate authority are not accepted.

In general, shorter-lived certificates offer a better security posture since the impact of key compromise is less severe but there is no minimum requirement for certificate expiry periods.

Below is a list for examples of certificate authority providers with no recommendations or rankings incorporated. It is recommended that a business researches which certificate authority meets their requirements:

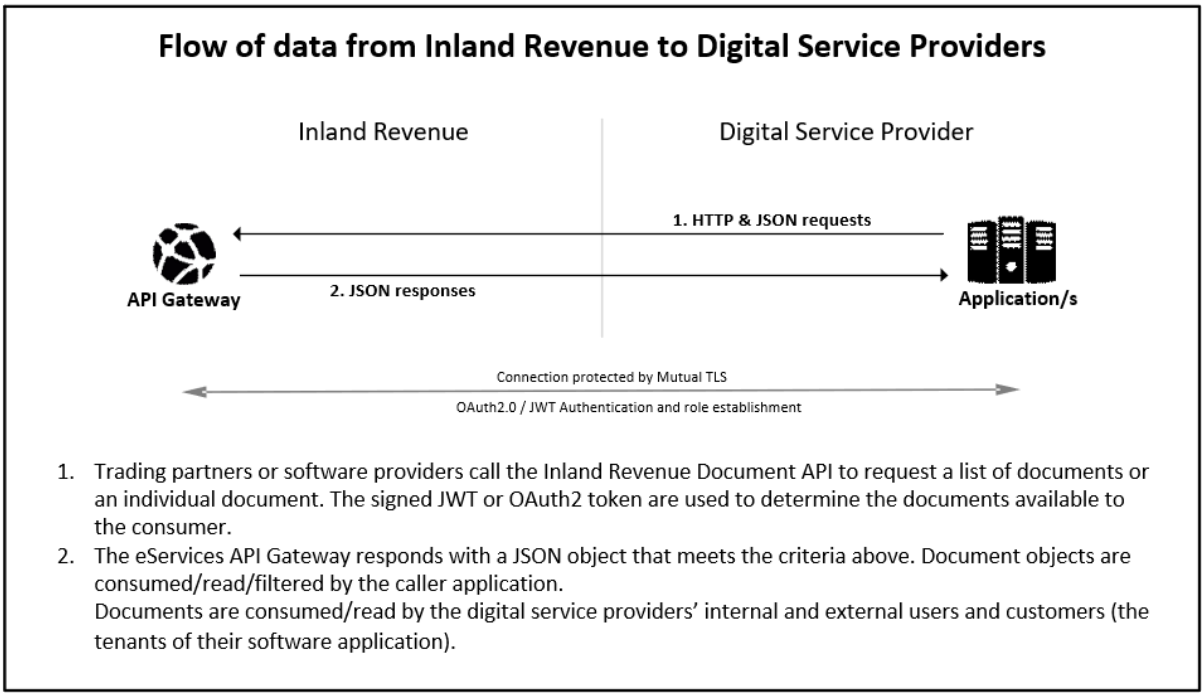
- Comodo
- GeoTrust
- DigiCert
- GlobalSign
- Symantec
- Thawte
- IdenTrust
- Entrust
- Network Solutions
- RapidSSL
- Entrust Datacard
- GoDaddy.

# 2 Solution design

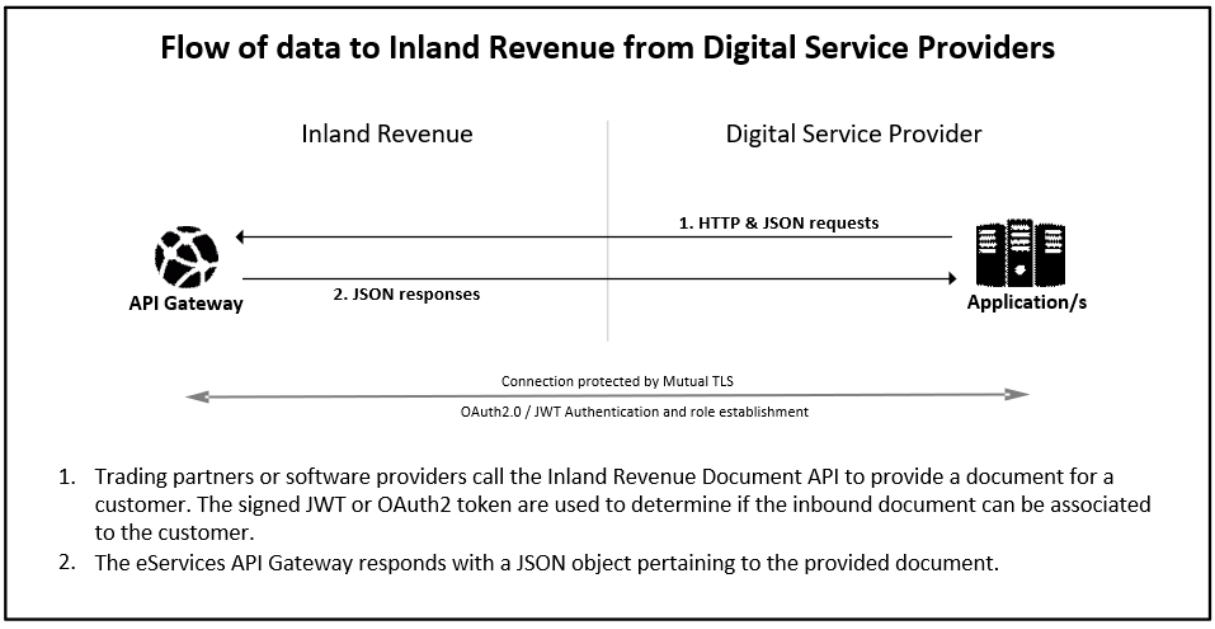
## 2.1 Architecture

Inland Revenue is offering a suite of web applications in order to facilitate interactions via software packages. This API will be used by approved organisations to interact with the Inland Revenue document service.

The diagram below illustrates the flow of data from Inland Revenue to Digital Service Providers.




The diagram below illustrates the flow of data to Inland Revenue from Digital Service Providers.



## 2.2 Supported message types

This service supports the following message types:

- **READ:** Retrieve a list of document identifiers from Inland Revenue. Requires either an IRD number or a return submission key.
- **READ/{documentID}:** Retrieve a single document from Inland Revenue. Requires a document identifier.
- **CREATE:** Submit a single document to Inland Revenue. Requires either a document location identifier or a return submission key.
- **UPDATE:** Inform Inland Revenue that a previously-submitted document is invalid. Requires a document identifier.

## 2.3 Document List

These are used by the consumer to either retrieve a list of documents associated with an IRD number (for example, letters generated by Inland Revenue and published to myIR), or a list of documents linked to a return.

### 2.3.1 Request payload

| Field | Description |
|---|---|
| **IRDNumber** | IRD number of the customer for which the document list should be retrieved |
| **ExtID** | Reference to a START object to which documents can be linked |
| **ExtIDType** | Enumerates the area in START where the documents referenced by ExtID is organised |
| **FromDateTime** | The earliest point in time up to which documents can be selected based on their date-time stamp—may be useful as a form of pagination |
| **ToDateTime** | The latest point in time up to which documents can be selected based on their date-time stamp—may be useful as a form of pagination |

A list of the valid values for **ExtIDType** is as follows:

| Type | Description |
|---|---|
| **RTNSUB** | The submission key of a return submitted through the Returns Service |

### 2.3.2 Response payload

| Field | Description |
|---|---|
| **DocumentID** | An identifier of a document that can be retrieved through the document service |
| **DateTimeCreated** | Date the document was created or submitted |
| **MailType** | Internal identifier of a letter type, if applicable |
| **Description** | A description of the document or the title of a letter |

| Field | Description |
|---|---|
| **AccountType** | Account type associated with the document, if applicable |
| **AccountID** | Identifier of the account associated with the document, if applicable |
| **FilingPeriod** | The end of the filing period to which the document corresponds, if applicable |
| **Redirected** | Indicates if the mail item was redirected, if applicable |
| **FileName** | The file name of the document, if applicable |
| **DocumentCategory** | The category of the document provided upon submission, if applicable |

## 2.4 Document Retrieve

This is used by the consumer to either retrieve a specific document identified by including the unique DocumentID as a query parameter.

### 2.4.1 Request parameters

| Field | Description |
|---|---|
| **DocumentID** | An identifier of a document that can be retrieved through the document service |

### 2.4.2 Response payload

| Field | Description |
|---|---|
| **DocumentID** | An identifier of a document that can be retrieved through the document service |
| **DateTimeCreated** | Date the document was created or submitted |
| **MailType** | Internal identifier of a letter type, if applicable |
| **Description** | A description of the document or the title of a letter |
| **AccountType** | Account type associated with the document, if applicable |
| **AccountID** | Identifier of the account associated with the document, if applicable |
| **FilingPeriod** | The end of the filing period to which the document corresponds, if applicable |
| **FileName** | The file name of the document, if applicable |
| **DocumentCategory** | The category of the document provided upon submission, if applicable |
| **Document** | A base64 encoded string containing the byte array of the document |

## 2.5 Document Create

Used by the consumer to submit a new document to Inland Revenue, either as a result of receiving a request from Inland Revenue via a notification or to link to a previously filed return.

The maximum file size of a document submitted through the create service is 9,000,000 bytes before Base64 encoding.

### 2.5.1 Request payload

| Field | Description |
|---|---|
| **DocumentLocationID** | An identifier that is used to properly route a document submitted through the document service |
| **ExtID** | Reference to a START object to which documents can be linked. |
| **ExtIDType** | Enumerates the area in START where the documents referenced by ExtID is organised |
| **DocumentCategory** | The category of the document provided upon submission, if applicable |
| **Description** | A description of the document |
| **FileName** | The file name of the document, including the file extension |
| **Document** | A base64 encoded string containing the byte array of the document |

File extension must be one of the following: gif, jpg, jpeg, pdf, png, csv, doc, docx, xls, xlsx, ppt, pps, pptx, ppsx, odt, ods, odp.

### 2.5.2 Response payload

| Field | Description |
|---|---|
| **DocumentID** | An identifier of a document that can be retrieved through the Document Service |

## 2.6 Document Update

This is used by the consumer to mark a previously supplied document as submitted in error. The document is identified by including the unique DocumentID as a query parameter.

### 2.6.1 Request parameters

| Field | Description |
|---|---|
| **DocumentID** | An identifier of a document that can be retrieved through the Document Service |

### 2.6.2 Request payload

| Field | Description |
|---|---|
| **Reason** | Text description of the reason for marking a previously supplied document as submitted in error |

### 2.6.3 Response payload

No response payload will be returned. The successful HTTP response (200/202) will be treated as acknowledgment of success.

## 2.7 Security

The API will use and require a unique identifier to be provided to establish the calling party identity and authentication required by the access model. This design will use JSON Web Tokens (JWT) and OAuth2.0 tokens and protocol to establish the calling party's identity. The OAuth2.0 method requires a user to logon, while the other is a machine-to-machine credential. Each HTTPS header contains the authorisation attribute JWT/OAuth:

1. A signed JSON Web Token (JWT) token. This will establish a registered digital services provider identity via the asymmetric public key held in the key store established during onboarding.
2. An OAuth2.0 token that is a customer- or intermediary-level XIAMS user account recognised by START.

The Notification Service uses an HTTPS transport layer, with HTTP1.1 transport protocol supported.

Regarding transport layer security (TLS), note that while TLS1.3 is now an industry standard, it is not yet widely adopted, as doing so requires upgrades to perimeter security devices and software. Inland Revenue will upgrade to TLS1.3 once it is adopted widely enough, and where practical, external software partners should also anticipate upgrading to this version. TLS1.0 and TLS1.1 are not supported by myIR or Gateway Services.

Asymmetric keys of approved strength must be used. Inland Revenue requires the following ciphers and key strengths to be used:

| | | | |
|---|---|---|---|
| **Encryption:** | Advanced Encryption Standard (AES) | FIPS 197 | 256-bit key |
| **Hashing:** | Elliptic Curve Digital Signature Algorithm (ECDSA) using P-256 or<br>Secure Hash Algorithm (SHA-2)<br>NOTE: ECDSA is preferred but RSA will be supported. | FIPS 180-3 | SHA-256<br>(or greater) |

Gateway Services will use this token in the HTTP header of a message in the same manner that an OAuth token has been used, namely:

```
"Authorization: {JWTAccessToken}"
```

*Refer to the Identity and Access Services build pack for more information.*

| | End point for machine-to-machine connections |
|---|---|
| **Purpose** | • End point for digital service providers to connect to |
| **Client application type** | • Cloud applications or in-house servers |
| **Constraints** | • Only for source locations with client-side TLS certificates<br>• On the cloud end point Inland Revenue has controls to shield service providers from issues caused by heavy usage from other providers |
| **Mutual TLS** | • Inland Revenue explicitly trusts the certificate the service provider associates with the TLS connection as client for Mutual TLS connections and uses it to identify the web service's sending party |
| **Minimum TLS version** | • 1.2 |
| **URL** | • Contains …/gateway/.. |
| **Port** | • 4046 |
| **Web service consumer identification** | • Machine-to-machine authentication using client-signed JSON web tokens (JWT)<br>• OAuth2 authorisation using tokens generated by XIAMS |
| **Firewalling in production** | • No IP address restrictions<br>• Access limited by certificate enrolment |
| **Firewalling in non-production environments** | • No IP address restrictions<br>• Access limited by certificate enrolment |

**Delegated permissions:** The services will allow one to access or provide documents for a customer to which the calling user (as represented by the JWT or OAuth2 token) has access. There may be additional accounts this identity does not have access to, but those will not be mentioned. If an account or data within it is targeted by the request parameters but the user does not have permission, an error will be returned.

### 2.7.1 OAuth

HTTP headers intended for OAuth access services will be have the JWT prefixed with "Bearer ".

| HTTP header | Example value |
|---|---|
| **Authorization** | Bearer {JWTAccessToken} |

*Refer to the Identity and Access Services build pack for more information on authorisation flows.*

### 2.7.2 M2M JWT

Authorisation intended for M2M (machine-to-machine) communication will not use "Bearer " flag on the HTTP header and only contain the JWT. The JWT will contain a field "startLogon" which can resolve to a myIR logon. The M2M JWT will be identified by a value of "M2M" in the Key ID ("kid"). The M2M JWT will be signed with a self-signed certificate, for which the public key was provided during onboarding.

| HTTP header | Example value |
|---|---|
| **Authorization** | {JWTAccessToken} |

Example data structure used for M2M authorisation:

```
Base64Url encoded {
    "alg": <algorithm value>,
    "typ": "JWT",
    "kid": "M2M"
}
.
Base64Url encoded {
    "sub": <token subject>,
    "iss": <issuer value>,
    "startLogon": <myIR_user>,
    "iat": <epoch issued value>,
    "exp": <epoch expired value>
}
.
JWS Signature (
    base64UrlEncode(header) + "." + base64UrlEncode(payload)
)
```

#### *2.7.2.1 Header*

| Field | Requirement | Description | Valid values |
|---|---|---|---|
| **alg** | Required | Signature or encryption algorithm | RS256, RS384, RS512<br>ES256, ES384, RS512 |
| **typ** | Required | Type of token | JWT |
| **kid** | Required | Key ID | M2M |

#### *2.7.2.2 Payload*

| Field | Requirement | Description | Valid values |
|---|---|---|---|
| **sub** | Required | Subject (to whom the token refers) | SHA-1 Thumbprint/fingerprint of signing certificate |
| **iss** | Required | Issuer who created this token | eg CompanyNameA |
| **startLogon** | Required | The myIR logon of a representative of the token | Valid myIR logon, or null |

| Field | Requirement | Description | Valid values |
|---|---|---|---|
| | | subject. The subject must be the data owner. | |
| **iat** | Required | Issued at. The number of seconds since Unix epoch 1 Jan 1970, UTC. | Must not precede the signing certificate issue date.<br>Example: 1560144847 |
| **exp** | Required | Expiration time. The number of seconds since Unix epoch 1 Jan 1970, UTC. | Must not exceed eight hours from the **iat** (issued at) time value.<br>Example: 1574323940 |

#### *2.7.2.3 startLogon*

A myIR logon can be provided in order to use the myIR delegation model for identifying customers for whom documents can be accessed. If the myIR logon is provided, then documents will only be shown and accepted for customers the logon can access. If a myIR logon is not used, the field should be included with a value of null, and the subject will determine the documents shown and allowed to access.

#### *2.7.2.4 sub*

A subject must be provided, which is the thumbprint of the signing certificate, and can be used to determine which documents can be accessed.

The subject will always be used to validate the signature of the JWT but will only be used for determining which documents to access when value for **startLogon** is not provided. The subject can be used for access in two distinct situations, when the subject is a KiwiSaver scheme provider, or when the subject is a tax preparer:

- If the subject is a KiwiSaver scheme provider, documents can be provided and accessed for the current members of the scheme.
- If the subject is a tax preparer, documents can be provided and access for customers currently linked to the tax preparer.

# 3 End points and OpenAPI specifications

**IMPORTANT**
For the authoritative definitions, please refer to the OpenAPI specifications at https://www.ird.govt.nz/software-providers/

## 3.1 End points

Onboarding instructions are available at https://www.ird.govt.nz/software-providers/.

## 3.2 OpenAPI specifications

An OpenAPI file allows you to describe your entire API, endpoints, operations on each endpoint, and operation parameters. The included .yaml file can be used along with an OpenAPI editor such as editor.swagger.io to view technical specifications for this operation and generate example client code.

# 4 Glossary

| Acronym/term | Definition |
|---|---|
| **API** | Application Programming Interface—set of functions and procedures that allow applications to access the data or features of another application, operating system or other service. |
| **Authentication** | The process that verifies the identity of the party attempting to access Inland Revenue |
| **Authorisation** | The process of determining whether a party is entitled to perform the function or access a resource |
| **End points** | A term used to describe a web service that has been implemented |
| **FIPS** | Federal Information Processing Standard—a suite of IT standards from the US Federal Government |
| **Gateway** | Inland Revenue's web services gateway |
| **HTTP, HTTPS** | Hyper Text Transmission Protocol (Secure)—the protocol by which web browsers and servers interact with each other. When implemented over TLS1.2 HTTP becomes HTTPS. |
| **IAMS** | Identity and Access Management—a logical component that performs authentication and authorisation. Physically it is a set of discrete hardware and software products, plug-ins and protocols. Usually implemented as separate External IAMS (XIAMS) and Internal IAMS. |
| **IAS** | Identity and Access Service |
| **IP** | Internet Protocol—the principal communication protocol in the Internet protocol suite for relaying datagrams across networks |
| **IRD** | Inland Revenue Department (ie IRD Numbers) |
| **JWT** | JSON Web Token—a compact, URL-safe means of representing claims to be transferred between two parties |
| **M2M** | Machine-to-machine communication |
| **OAuth** | An HTTPS based protocol for authorising access to a resource, currently at version 2 |
| **OpenAPI specifications** | Formerly known as Swagger specifications—a specification for machine-readable interface files for describing, producing, consuming and visualising RESTful web services. |
| **Payloads** | This refers to the data contained within the messages that are exchanged when a web service is invoked. Messages consist of a header and a payload. |
| **Schemas** | An XML schema defines the syntax of an XML document, in particular of a payload. The schema specifies what a valid payload must or can contain, as well as validating the payload. |
| **SHA** | Secure Hashing Algorithm. There is a family of them that provide different strengths. SHA-2 is currently favoured over SHA-1, which has been compromised. |
| **SOAP** | Simple Object Access Protocol—a set of standards for specifying web services. GWS uses SOAP version 1.2 |

| Acronym/term | Definition |
|---|---|
| **SSL** | Secure Sockets Layer certificates—used to establish an encrypted connection between a browser or user's computer and a service or website |
| **START** | Simplified Taxation and Revenue Technology—Inland Revenue's new core tax processing application. It is an implementation of the GenTax product from FAST Enterprises. |
| **TLS1.2** | Transport Layer Security version 1.2—the protocol that is observed between adjacent servers for encrypting the data that they exchange. Prior versions of TLS and all versions of SSL have been compromised and are superseded by TLS1.2. |
| **URL** | Universal Resource Locator—also known as a web address |
| **X.509 certificate** | An international standard for encoding and describing a digital certificate. In isolation a public key is just a very large number, the X.509 certificate to which it is bound identifies whose key it is, who issued it, when it expires etc. When a counterparty's X.509 digital certificate is received, the recipient takes their public key out of it and store the key in their own key store. The recipient can then use this key to encrypt and sign the messages that they exchange with this counterparty. |
| **XIAMS** | External IAMS—an instance of IAMS that authenticates and authorises access by external parties, for example customers, trading partners etc, as opposed to internal parties such as staff |
| **YAML** | "YAML Ain't Markup Language"—a human-readable data-serialisation language commonly used for configuration files and in applications where data is stored or transmitted. |

# 5 Change log

This table lists all material changes that have been made to this build pack document since the release of v1.0 (most recent changes listed first). It does not encompass non-material changes, such as to formatting etc.

| Version | Date of change | Document section | Description |
|---|---|---|---|
| **1.0** | 06/04/20 | | • V1.0 released |