Inland Revenue

Build Pack: Document Service

Date: 10/11/2020
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 continuing, please consult
www.ird.govt.nz/digital-service-providers/services-catalogue
for business-level context, use cases and links to relevant policy.
The information available here explains how to integrate with Inland Revenue's 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. 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.

1.3.2   Authentication options

1.3.2.1   OAuth

When using OAuth the interaction with IR is transacted under the identity of a myIR user. OAuth requires the presence of a myIR user, as this person must be available to supply their user ID, password and consent at run-time in order to be authenticated. OAuth is especially suited to cloud-based applications where the transacting parties are application users rather than providers.

1.3.2.2   JWT

The alternative to OAuth is JWT, which does not require the presence of a myIR user. Authentication is based on the verification of a digital signature that (provably) belongs to a customer. In order to digitally sign their messages, the customer must acquire a digital certificate from a trusted certificate authority, or generate a self-signed certificate, and supply it to Inland Revenue during the on-boarding process.

JWT is therefore appropriate when the following conditions apply:

- The interaction with Inland Revenue is conducted under the identity of an organisation, as opposed to a person AND
- The organisation has the technical and operational capability to securely obtain and manage digital certificates AND
- The organisation's interactions with Inland Revenue can occur in the absence of specific people due to staffing issues such as out-of-hours non-availability, staff turnover and absence from work.
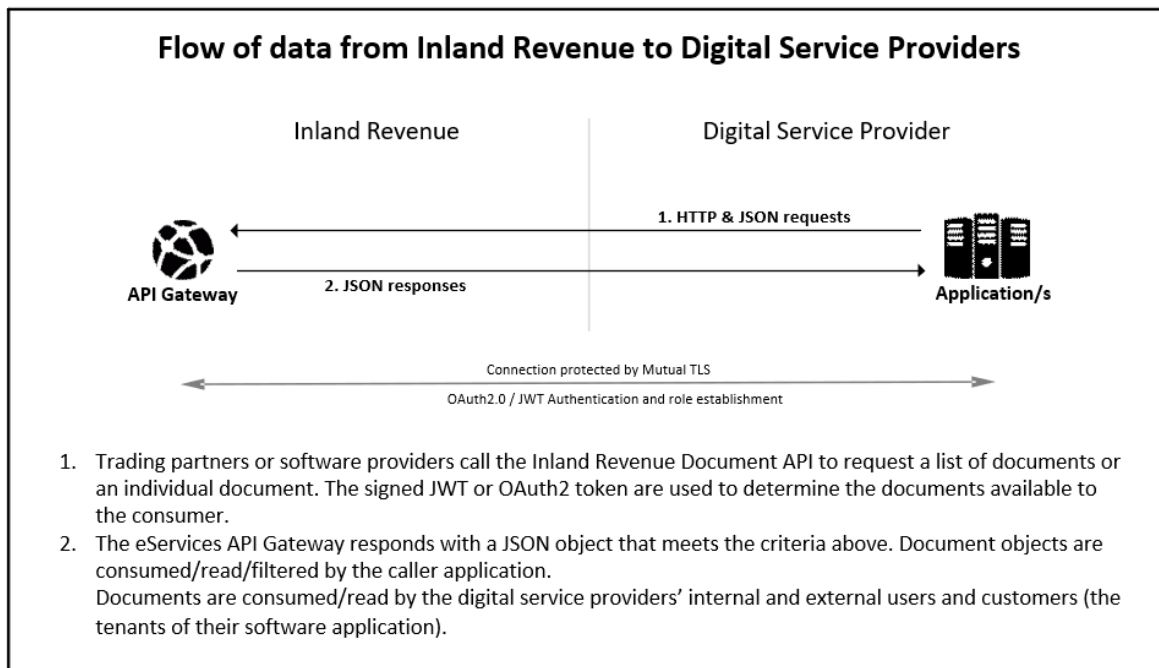
These factors tend to limit the use JWT to larger corporations and public sector organisations. It is not suitable for cloud-based applications as it requires all application users to have their own digital certificates—this is administratively burdensome and requires these users to lodge their private keys with their application provider, which is insecure.

![Inland Revenue Te Tari Taake logo]
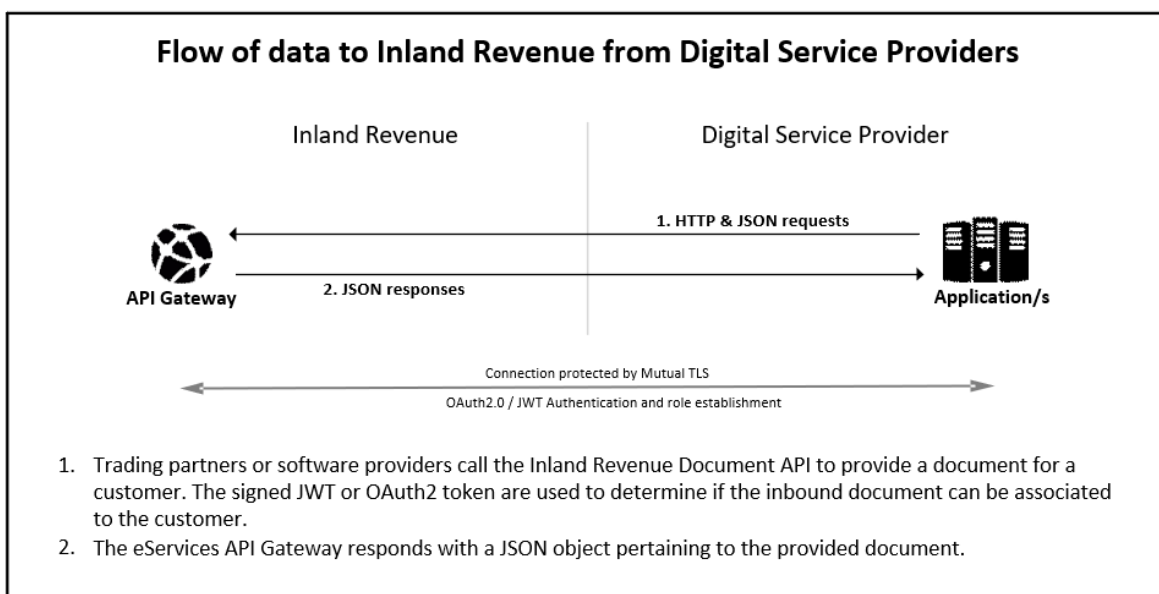
2   Solution design

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.

Flow of data from Inland Revenue to Digital Service Providers

1. Trading partners or software providers call the Inland Revenue Document API to request a list of documents or an individual document. The signed JWT or OAuth2 token are used to determine the documents available to the consumer.
2. The eServices API Gateway responds with a JSON object that meets the criteria above. Document objects are consumed/read/filtered by the caller application.
   Documents are consumed/read by the digital service providers' internal and external users and customers (the tenants of their software application).

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

Flow of data to Inland Revenue from Digital Service Providers

1. Trading partners or software providers call the Inland Revenue Document API to provide a document for a customer. The signed JWT or OAuth2 token are used to determine if the inbound document can be associated to the customer.
2. The eServices API Gateway responds with a JSON object pertaining to the provided document.

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 visible on myIR), or a list of documents linked to a return.

Note, certain types of letters generated by Inland Revenue must be posted and are not visible on myIR or through the Document Service.

2.3.1   Request payload

Field	Description
IRDNumber	IRD number of the customer for which the document list should be retrieved. If provided, only documents generated by Inland Revenue will be returned.
ExtID	Reference to a START object to which documents can be linked. If provided, only documents submitted through the Document Service will be returned.
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. This date must not be before 1753-01-01.
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. This date must not be before 1753-01-01.

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
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 retrieve a specific document identified by including the unique DocumentID as a query parameter.

Documents generated by Inland Revenue will always contain a MailType value and will not have a FileName. Instead, these documents will always have a file extension of .pdf.

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. This is only applicable to documents generated by Inland Revenue.
Description	A description of the document or the title of a letter
AccountType	Account type associated with the document, if applicable

Field	Description
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. This is only applicable to documents submitted through the Document Service.
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.

Multiple documents can be submitted for a given DocumentLocationID or ExtID/ExtIDType. When possible, a DocumentLocationID should be used, as this may enable certain automated actions to be performed and may speed up the processing of the document.

Note: The DocumentID returned by this operation must be stored, as documents submitted with a DocumentLocationID will not be displayed in the Document List operation.

Note: If a document is submitted for a return, the status of the return must not be 'Submitted'.

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

Field	Description
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 Secure Hash Algorithm (SHA-2) NOTE: ECDSA is preferred but RSA will be supported.	FIPS 180-3	SHA-256 (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 • 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) • OAuth2 authorisation using tokens generated by XIAMS
Firewalling in production	• No IP address restrictions • Access limited by certificate enrolment
Firewalling in non-production environments	• No IP address restrictions • 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 ES256, ES384, ES512
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 subject. The subject must be the data owner.	Valid myIR logon, or null
iat	Required	Issued at. The number of seconds since Unix epoch 1 Jan 1970, UTC.	Must not precede the signing certificate issue date. 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. 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

3.1 End points

Current environment information for this service—including the end points for each environment—is available within the relevant Software Development Kit (SDK).

To access the SDK, do one of the following:

- Go to https://github.com/InlandRevenue and select this service
- Go to https://developerportal.ird.govt.nz and click the link to the SDK within the Gateway Service documentation (please register first).

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.

To access the latest OpenAPI definition for this service, please do the following:

- Login to the developer portal at https://developerportal.ird.govt.nz (register first)
- Download and view the OpenAPI definition within the Gateway Service documentation.

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.

Acronym/term	Definition
SOAP	Simple Object Access Protocol—a set of standards for specifying web services. GWS uses SOAP version 1.2
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	24/11/20	2.4	• Removed extraneous "either"
	10/11/20	2.3.1	• Clarified dates cannot be before 1753-01-01
	29/09/20	1.3.2	• New section added – 'Authentication options'
	08/17/20	2.5	• Added a note clarifying what the status of a return may not be in to be used for submitting a document
	08/09/20	2.3, 2.4, 2.5	• Updated the descriptions of these operations and their request fields to clarify how the operations should be used
	06/08/20	2.7.2.1	• Typo corrected in value values field for 'alg'
		1.1	• Updated URL in boxed text
		3	• Entire section updated with new information on where to find end points and OpenAPI specs
	06/04/20		• V1.0 released