



# Draft URL specification for NZ Legislation website

Date	18 September 2025
Version	1.0
Author	Andy Neale / <a href="mailto:contact@pco.govt.nz">contact@pco.govt.nz</a>
Release notes	For draft release for public feedback

## Purpose

This specification proposes the structure, allowed values, and rules for public URLs on [legislation.govt.nz](https://legislation.govt.nz). It is designed to:

- Improve human readability and predictability of URLs
- Assign unique URLs to every resource to support long-term persistence (HTTP URIs)
- Allow lossless mapping to Akoma Ntoso (LegalDocML) URIs
- Enable for future extensions such as granular referencing across versions of legislation

## Definitions

### Resource

A resource is a specific version of a piece of legislation in a chosen language. You can also treat fragments (e.g., a part, section, or definition) as discrete resources when you link to them.

Example: <https://www.legislation.govt.nz/act/public/1986/121/1986-04-17/en/> identifies an English resource for that dated version.

### Representation

A representation is a legislation resource made available as a specific media type such as XML, JSON, or PDF. Also called a manifestation in the FRBR model and used in the Akoma Ntoso standard.

Example: `../en/file.xml`, `../en/file.json`, `../en/file.pdf`

### Canonical URI

A URI (Uniform Resource Identifier) is an identifier for a resource. In this spec, URIs are web addresses (URLs). The canonical URI is the single, stable URL we use for references, bookmarking, and long-term persistence. It aims to be unique across the full collection of NZ legislation.

Form: `{/TYPE}/{SUBTYPE}/{YEAR}/{NUMBER}/{DATE}/{LANGUAGE}/`

### Container

A container is a URI path that **MUST** end with a trailing `/`. It denotes the home location for a specific version of legislation in a given language; and **MAY** link to representations such as XML, JSON, or PDF files.

Example: `../en/` **MAY** contain `../en/file.xml`

## Resolver

A convenience URL that locates a resource but is not the canonical identifier. Resolvers (e.g., /latest/ or language-omitted forms) MUST return a 302/307 redirect to the canonical dated-language URL.

Example: /.../latest/ → 302 → /.../1986-04-17/en/

## URI Model

HTTPS://LEGISLATION.GOV.TZ / {TYPE} / {SUBTYPE} / {YEAR} / {NUMBER} / {VERSION} / {LANGUAGE}

SEGMENT VALUES	GUIDANCE
<b>TYPE</b>  act   bill   secondary-legislation   amendment-paper	  /.../act/
<b>SUBTYPE</b>  act: public   imperial   local   private   provincial  bill: government   local   members   private  secondary-legislation: publisher  amendment-paper: government   members	 The data source for <b>secondary-legislation: publisher</b> should be the acronym field of the Public Service Commission's organisation dataset <sup>1</sup>  A process will need to be established for agreeing new acronyms if the agency is not listed.  /.../secondary-legislation/acc/
<b>YEAR</b>  YYYY	 Date sources for YYYY:  – <b>act</b> : year enacted – <b>bill</b> : year introduced – <b>secondary-legislation</b> : year made – <b>amendment-paper</b> : year introduced  Where regnal years <sup>2</sup> such as 25-Eliz-2 are present the segment should use <b>year of assent</b> .  /.../bill/imperial/1267/
<b>NUMBER</b>  {base number} + {dedupe} + {bar number} + {split letter}	 Elements of NUMBER:  – <b>base number</b> : the instrument's main number for that year (e.g., 120). For PCO-published legislation the base number is generated by the PCOs internal drafting systems. For agency-published legislation (where there may be independent numbering systems), the PCO will need to generate a unique base number from the sequential date of first data collection (harvest). – <b>dedupe</b> : the base number may have a letter suffix (A-Z) to deduplicate historical number clashes (e.g., 96A, 96B).

<sup>1</sup> <https://catalogue.data.govt.nz/dataset/mog-events/resource/bfc80102-20f8-4a37-815a-2c07c667101c>

<sup>2</sup> [https://en.wikipedia.org/wiki/Regnal\\_years\\_of\\_English\\_and\\_British\\_monarchs](https://en.wikipedia.org/wiki/Regnal_years_of_English_and_British_monarchs)

	<ul style="list-style-type: none"> <li>– <b>bar number</b>: for Bills only, this is -1, -2, -3, -4, -5 to show the parliamentary stage.</li> <li>– <b>split letter</b>: for split Bills only, a letter is added after the bar number if multiple Bills are created (e.g., 3A, 3B).</li> </ul> <p>./.../bill/members/1957/96A-3B/</p>
<b>VERSION</b>  YYYY-MM-DD + {dedupe}	<p>Legislation published by the PCO currently includes a sequential version number however this is fragile and doesn't extend to agency-published legislation. The proposal to use date as a version stamp, that will span all NZ legislation, is based on the Akoma Ntoso standard.</p> <p>Date sources for VERSION:</p> <ul style="list-style-type: none"> <li>– <b>act</b>: as-enacted date first version; consolidation date (Reprint as at date) for any subsequent versions.</li> <li>– <b>bill</b>: as-reported date from the appropriate committee.</li> <li>– <b>secondary-legislation</b>: as-made date first version; consolidation date for any subsequent versions.</li> <li>– <b>amendment-paper</b>: released date.</li> </ul> <p>Resolving clashes:</p> <ul style="list-style-type: none"> <li>– <b>dedupe</b>: the date may have a letter suffix (A-Z) to deduplicate date clashes (e.g., 2021-04-17B).</li> <li>– Use of a deduping letter will be very rare, however is specified in case there are ever multiple versions of the same legislation published in a single day.</li> </ul> <p>./.../act/public/1986/121/1986-04-17A/</p>
<b>LANGUAGE</b>  en   mi	<p>ISO 639-1</p> <ul style="list-style-type: none"> <li>– <b>en</b> = English</li> <li>– <b>mi</b> = te reo Māori</li> </ul> <p>We propose adopting the more modern convention of 2-letter language codes instead of Akoma Ntoso's 3-letter language codes. This is because search engines and many user agent negotiations assume 2-letter codes by default, and they are shorter, slightly cleaner URLs. If the language segment is omitted, then fallback resolution rules apply.</p> <p>./.../act/public/1986/121/1986-04-17/mi/</p>

### URI Rules

CONCEPT	RULE
Leading zeros	No leading zeros e.g. <b>123</b> not 0123 (except in dates)

Singular	Specify singular not plural form of TYPE e.g <b>act</b> not acts
Case	Specify lower case e.g. <b>act</b> not Act except for uppercase dedupe letter and split Bills in number and version such as /bill/members/1957/96A-3B/
Trailing slash	Always include trailing slash i.e. <b>always include /</b> unless there is a filename in the URL, as trailing slashes resolve more predictably and indicate that this is a container for different formats such as XML, JSON, or PDF. <ul style="list-style-type: none"> <li>– redirect non-slash form</li> <li>– when the path ends with a concrete file name (e.g., .pdf, .json, .xml), <b>no trailing slash</b> and no redirect</li> </ul>

## Endpoints

PATTERN	GUIDANCE
<p><code>./file.{FORMAT}</code></p> <p>xml   json   pdf</p>	<p>The /file.FORMAT endpoint can be added to the end of a URI to request download of a specific media type. If language is omitted, then fallback to resolution rules. On download the standard filename will be rewritten to the form:</p> <p>{LEGISLATION-TITLE}-{VERSION}-{LANGUAGE}.{FORMAT}</p> <p>For example:</p> <ul style="list-style-type: none"> <li>– <code>./act/public/1986/120/1989-06-23/en/file.pdf</code></li> <li>– will be rewritten as:</li> <li>– ‘Local Government Amendment Act 1986 1989-06-23 English.pdf’</li> </ul>
<code>/latest/</code>	<p>The /latest/ endpoint will resolve to the latest version of the resource. If language is omitted, then fallback to resolution rules. If multiple versions use a dedupe letter then resolve to the version with the highest letter for that DATE, even if the letter is omitted. TBC whether to include prospective consolidations.</p> <p>For example:</p> <ul style="list-style-type: none"> <li>– <code>./act/public/1986/121/latest/</code></li> <li>– will resolve to:</li> <li>– <code>./act/public/1986/121/1986-04-17C/en/</code></li> </ul>

## Resolution rules

It is desirable for the different URL elements to be addressable.

URL	RESPONSE
<code>/act/public/1986/120/1986-04-17/mi/</code>	Te reo Māori text of that version, else return 404
<code>/act/public/1986/120/1986-04-17/en/</code>	English text of that version, else return 404

/act/public/1986/120/1986-04-17/	Redirect to explicit language variant in Accept-Language header defaulting to /en/ if unavailable.
/act/public/1986/120/	Versions tab of public Act 120 in 1986
/act/public/1986/	Search results for all public Acts from 1986
/act/public/	Search results for all public Acts
/act/	Search results for all Acts
/act/public/1986/120/latest/file.xml	XML file of latest version for explicit language variant in Accept-Language header defaulting to /en/ if unavailable.
/act/public/1986/120/1986-04-17/file.xml	XML file for explicit language variant in Accept-Language header defaulting to /en/ if unavailable.
/act/public/1986/120/1986-04-17/mi/file.xml	XML file of Te reo Māori text of that version. Return a 404 error if there is no /mi/ variant

## Further details to confirm

While feedback is primarily being sought on the approach to URIs and endpoints, there are several other areas of future development where comment is welcome:

### 1. Akoma Ntoso

The PCO is currently considering whether to adopt the Akoma Ntoso standard for its XML drafting process. That standard has a form of URI construction that is very different to the current model in use at the PCO, but this specification proposes a model that is much closer. While there is no requirement for adopting the Akoma Ntoso URI model for website URLs, we think it is useful to get a close as possible. For comparison, this is what an Akoma Ntoso based URL might look like compared to the same Act using this document's proposed URLs:

- **Akoma Ntoso:** <https://www.legislation.govt.nz/akn/nz/act/1986-23-10/8/eng@1986-04-17/>
- **This proposal:** <https://www.legislation.govt.nz/act/public/1986/8/1986-04-17/en/>

### 2. Granular Hierarchical Identifiers

It is intended that the proposed URIs will eventually be extended to support Granular Hierarchical Identifiers (GHIs) to make fragment levels fully addressable, for example:

- [www.legislation.govt.nz/act/public/1986/121/1986-04-17/en/part1/para5/sect1/def4/b/i/](https://www.legislation.govt.nz/act/public/1986/121/1986-04-17/en/part1/para5/sect1/def4/b/i/)

The naming convention for GHIs is still to be confirmed. While this can be implemented using today's technology on a resource-by-resource basis, it is not yet clear how to ensure these GHIs persist over time through different versions when un-numbered elements (such as `./.../def4/b/i/`) can be impacted by changing structures above them. This has important implications for how both people and machines may use this data.

### 3. Permanent IDs

We'll keep using our canonical URIs for legislation, but in the future, we also want to add a short permanent ID for each item as a handle that never changes. Even if numbering, publishers, or paths evolve each handle will **permanently redirect** to the canonical URI of the resource it denotes, so references can remain valid for decades. This will require several different systems to adopt the same official numbering scheme so NZ legislation numbers can span the publication of both primary and secondary legislation as well as being retrospectively assigned. For example:

- <https://www.legislation.govt.nz/id/498379204/>

- would resolve to
- <https://www.legislation.govt.nz/act/public/1986/8/1986-04-17/en/>

4. **Whether to map to IRIs?**

An Internationalised Resource Identifier (IRI) can support macrons in the URIs, while this spec does not currently support that.

5. **Additional endpoints**

For example, a `/current/` endpoint could resolve to the current in force version of a resource. It would return in force legislation as-at today and would return a 404 error if repealed or expired.

An `/as-at/{YYYY-MM-DD}` endpoint could then resolve to the version in force on a specified date, falling back to the last version before that date if needed.

6. Header, redirect, status code, and caching details are yet to be specified