

# Add a New Dataset to data.gov.ie via the Website UI

---

*Author: Deirdre Lee, Derilinx*

*Date: 22/07/2015*

## **Table of Contents**

Introduction.....	2
Adding a Dataset.....	2
Tab 1: Name.....	2
Tab 2: Data Files.....	3
Tab 3: Description.....	4
Tab 4: Licence.....	4
Tab 5: Publisher.....	5
Tab 6: Theme & Keywords.....	5
Tab 7: Time & Date.....	6
Tab 8: Geographic Coverage.....	7

## Introduction

For the complete Guide for Publishers, see <https://data.gov.ie/guide-for-publishers>

Publication on an Open Data Portal opens the door to innovative data-reuse. Data.gov.ie does not host the actual datasets, but instead is a catalogue of metadata, with pointers to the data hosted elsewhere on the Web, for example, on the website of a public organisation.

Data can be manually published on data.gov.ie via the 'Add Dataset' online form. Alternatively, data can be programmatically published on data.gov.ie via the API or a data harvester.

## Adding a Dataset

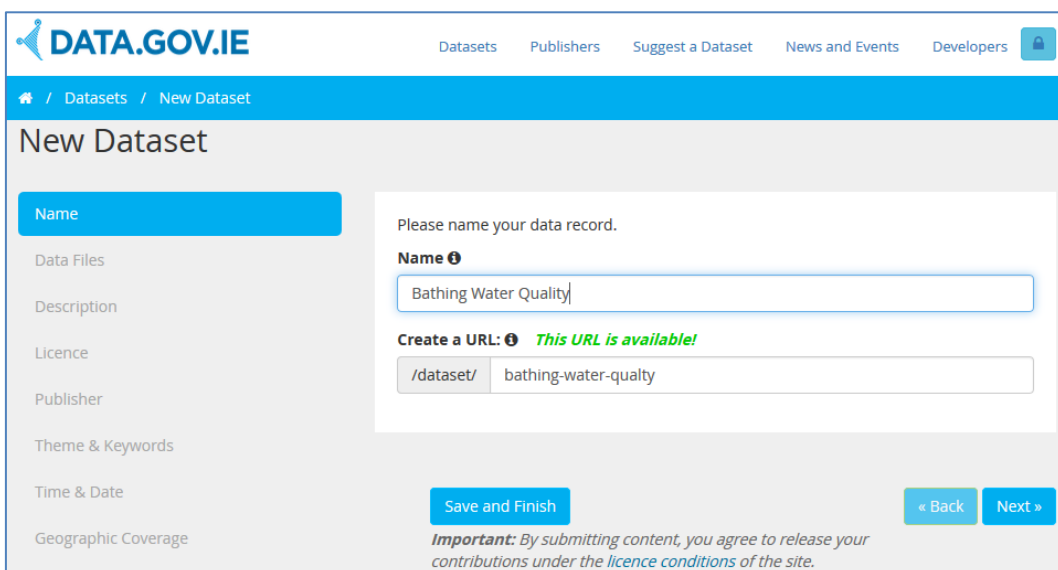
In order to add a dataset to data.gov.ie, the data publisher requires a user account. An account can be created for Public Sector Bodies by contacting [opendata@per.gov.ie](mailto:opendata@per.gov.ie) To log into your user account, go to <https://data.gov.ie/user/login>

Once logged-in, the data publisher can access the 'New Dataset' page at <https://data.gov.ie/dataset/new> . By stepping through the online form, the data publisher is prompted to enter all the necessary metadata. The user-friendly interface makes it easy for non-technical users to publish datasets one at a time. However, using the web interface is not an efficient way to publish multiple data sets, or to periodically update existing datasets. For this, programmatic publication via the API should be used.

Note, by submitting content on the data.gov.ie Open Data Portal, the data publisher agrees to release the data under the Open Data licence conditions outlined at <https://data.gov.ie/licence>. The following guide describes each step of adding a dataset.

### Tab 1: Name

The first tab is 'Name'. Enter the title of the dataset. The URL will be automatically generated, but it must be unique. If there is a conflict, a message will appear to say 'This URL is not available'. If this happens, change the title.



The screenshot shows the 'New Dataset' form on the Data.gov.ie website. The page title is 'New Dataset' and the breadcrumb is 'Data.gov.ie / Datasets / New Dataset'. The form is divided into a left sidebar with tabs for 'Name', 'Data Files', 'Description', 'Licence', 'Publisher', 'Theme & Keywords', 'Time & Date', and 'Geographic Coverage'. The 'Name' tab is selected. The main content area contains the following elements:

- A heading: 'Please name your data record.'
- A label: 'Name' with a required field icon (Ⓜ).
- A text input field containing 'Bathing Water Quality'.
- A label: 'Create a URL:' with a required field icon (Ⓜ) and a green status message: 'This URL is available!'.
- A text input field containing '/dataset/ bathing-water-quality'.
- At the bottom, there are three buttons: 'Save and Finish', '« Back', and 'Next »'.
- A disclaimer: 'Important: By submitting content, you agree to release your contributions under the licence conditions of the site.'

## Tab 2: Data Files

The second tab is 'Data Files'. Enter the details of each of the data files available for the dataset. For example, a dataset may be available in multiple formats: csv, json, xml, xls, pdf, etc. According to the Open Data Technical Framework<sup>1</sup>, the dataset should be available in at least one open, machine-readable format. If your data comes in multiple formats, each should be added on a separate row.

Add the file title, URL and format of the data file. The URL points to the direct link where the data file can be downloaded from.

Your dataset may be creating a single file record (which points to a static dataset) or you may be creating a timeseries record (which points to a dynamic dataset, where multiple files are released over time, for example where a new data file is added to the dataset monthly).

**DATA.GOV.IE** Datasets Publishers Suggest a Dataset News and Events Developers

Home / Datasets / New Dataset

### New Dataset

Name

**Data Files**

Description

Licence

Publisher

Theme & Keywords

Time & Date

Geographic Coverage

#### Data Files

In this tab you enter the links to the files containing the actual data at the heart of this record.

Your record may be creating a single file record (a record which points to just one file) or you may be creating a timeseries record (where multiple files are released over time, for example where a new data file is added to the record monthly).

If your data comes in multiple formats, use a line for each.

**NB: Do not put links to web pages here.** These and other useful documents that are not actually the data should go in the 'Additional Resources' section.

This record is:

A single file record

A timeseries record

	File Title: ⓘ	URL: ⓘ	Check	Format: ⓘ
↑↓	Bathing Water Quality CSV	www.splash.ie/bathingwaterc	Check	CSV
	Bathing Water Quality PDF	www.splash.ie/bathingwaterc	Check	PDF

Check all URLs...

+ Add Row

<sup>1</sup> <https://data.gov.ie/technical-framework#recommended-formats-for-open-data>

## Tab 3: Description

The third tab is 'Description'. Enter a general description of the dataset. The landing page indicates the website where the dataset is available online. Note this differs from the URL, the direct download link, defined in the previous tab. Conforms to refers to what standards the data in the dataset is modelled against, for example, W3C standards, ISO standards, or INSPIRE. A list of recommended standards is available in the Open Data Technical Framework<sup>2</sup>. Language refers to the main language of the dataset.

**DATA.GOV.IE** Datasets Publishers Suggest a Dataset News and Events Developers

/ Datasets / New Dataset

### New Dataset

- Name
- Data Files
- Description**
- Licence
- Publisher
- Theme & Keywords
- Time & Date
- Geographic Coverage

#### Description

Please add a description of the data set. What does the data contain/show? Why was it produced?

Local authorities are responsible for bathing water quality in their areas in addition to monitoring bathing water quality and making information available to the public on water quality during the summer bathing season.

(It is often displayed with the package title. In particular, it should start with a short sentence that describes the data set succinctly, because the first few words alone may be used in some views of the data sets. Here is the place to state if there are any limitations or deficiencies to the data in order to enable users to evaluate the information; even incomplete data may be adequate for some users).

#### Landing Page

www.splash.ie

Website URL where the data is available

#### Conforms To

INSPIRE

Data standard(s) that the data conforms to, e.g. INSPIRE

#### Language

English

Valid options are: *English* (default), *Irish*, or *Gaeilge*

## Tab 4 – Licence and Rights

The fourth tab is 'Licence and Rights'. Enter the Licence associated with the dataset. The recommended licence to associate with all Open Data from the Irish public sector is the CC-BY 4.0 Licence.

The screenshot shows the 'New Dataset' form on the Data.gov.ie website. The 'Licence & Rights' tab is selected. The 'Licence' dropdown menu is set to 'Creative Commons Attribution (CC-BY 4.0)'. Below the dropdown, there is a note: 'Note. Circular 12/2016: Licence for Re-Use of Public Sector Information adopts CC-BY as the standard PSI licence, and notes that the open standard licence identified in this Circular supersedes PSI General Licence No:2005/08/01 (more info)'. The 'Rights' text area is empty. At the bottom of the form, there is an important notice: 'Important: By submitting content, you agree to release your contributions under the licence conditions of the site.' and three buttons: 'Back', 'Next', and 'Save and Finish'. The footer of the page includes contact information for the Open Data Unit, Department of Public Expenditure & Reform, and mentions that the site is powered by CKAN and powered by Derilinx.

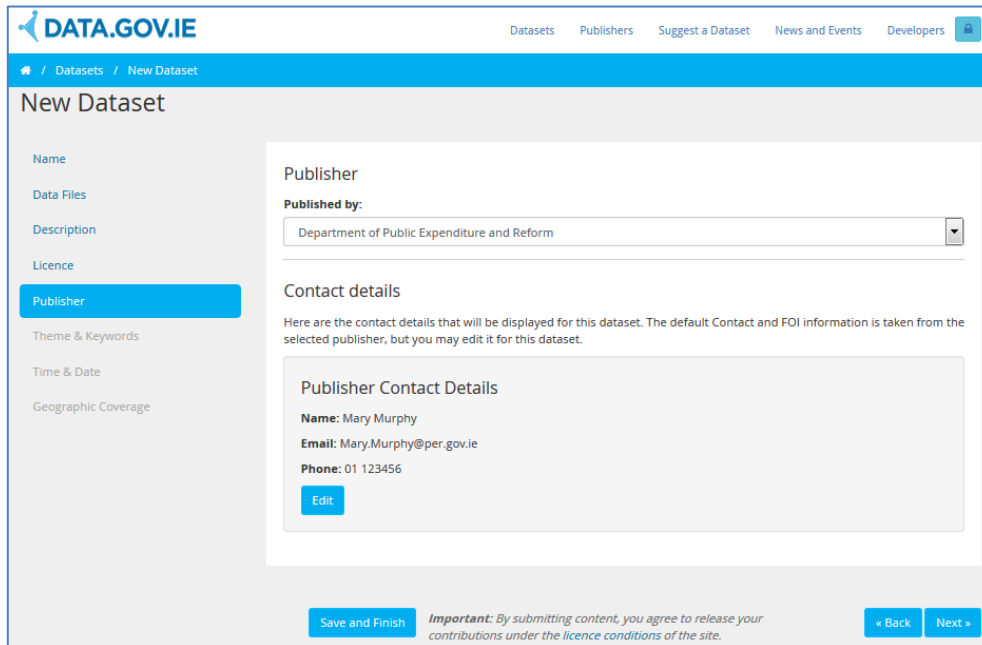
A statement about the intellectual property rights (IPR) held in or over a Resource, a legal document giving official permission to do something with a resource, or a statement about access rights.

<sup>2</sup> <https://data.gov.ie/technical-framework#recommended-standards-for-open-data>

<sup>3</sup> <https://data.gov.ie/technical-framework#open-data-licence>

## Tab 5: Publisher

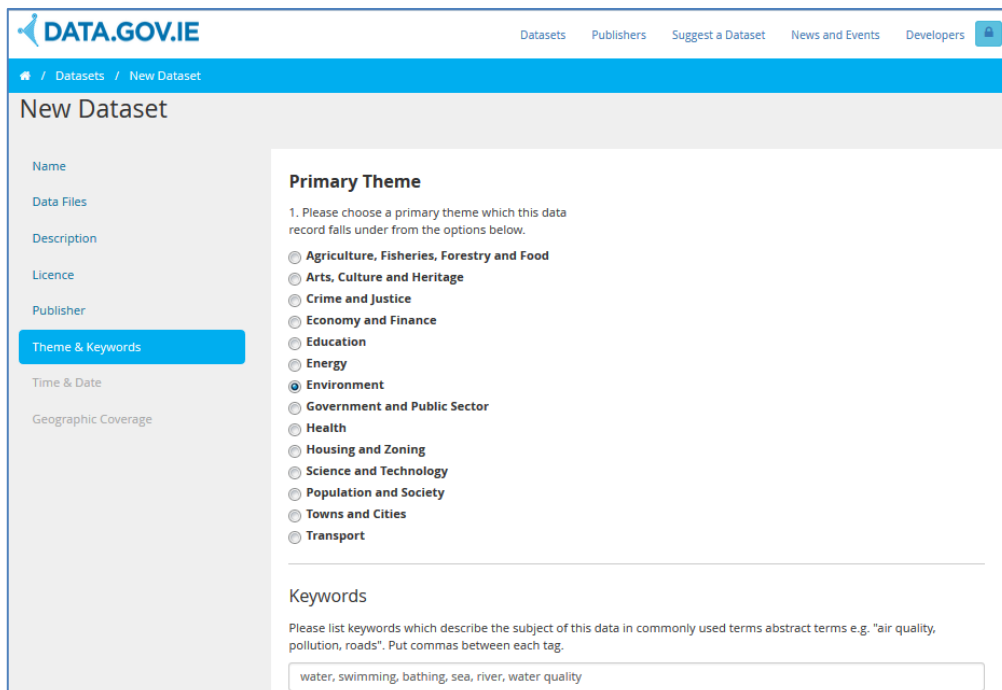
The fourth tab is 'Publisher. From the drop-down list, select the organisation who is publishing the data. If your organisation is not in the drop-down list, please contact [opendata@per.gov.ie](mailto:opendata@per.gov.ie). Also enter contact details that will be displayed for this dataset. The default contact details are taken from the selected publisher, but this can also be added for each dataset.



The screenshot shows the 'New Dataset' form on the DATA.GOV.IE website. The 'Publisher' tab is selected in the left-hand navigation menu. The main content area is divided into two sections: 'Publisher' and 'Contact details'. In the 'Publisher' section, there is a 'Published by:' dropdown menu with 'Department of Public Expenditure and Reform' selected. The 'Contact details' section contains a text box with the following information: 'Name: Mary Murphy', 'Email: Mary.Murphy@per.gov.ie', and 'Phone: 01 123456'. Below this information is an 'Edit' button. At the bottom of the form, there is a 'Save and Finish' button, a warning message: 'Important: By submitting content, you agree to release your contributions under the licence conditions of the site.', and 'Back' and 'Next' buttons.

## Tab 6: Theme & Keywords

The fourth tab is 'Theme & Keywords. Enter the primary theme for the dataset. Only one theme can be selected per dataset. List the keywords that describe the subject of this data in commonly used, abstract terms e.g. "air quality, pollution, roads". Commas should be used to separate keywords.



The screenshot shows the 'New Dataset' form on the DATA.GOV.IE website. The 'Theme & Keywords' tab is selected in the left-hand navigation menu. The main content area is divided into two sections: 'Primary Theme' and 'Keywords'. In the 'Primary Theme' section, there is a heading 'Primary Theme' followed by the instruction: '1. Please choose a primary theme which this data record falls under from the options below.' Below this is a list of radio button options: 'Agriculture, Fisheries, Forestry and Food', 'Arts, Culture and Heritage', 'Crime and Justice', 'Economy and Finance', 'Education', 'Energy', 'Environment' (which is selected), 'Government and Public Sector', 'Health', 'Housing and Zoning', 'Science and Technology', 'Population and Society', 'Towns and Cities', and 'Transport'. In the 'Keywords' section, there is a heading 'Keywords' followed by the instruction: 'Please list keywords which describe the subject of this data in commonly used terms abstract terms e.g. "air quality, pollution, roads". Put commas between each tag.' Below this is a text input field containing the text: 'water, swimming, bathing, sea, river, water quality'.

## Tab 7: Time & Date

The fourth tab is 'Time & Date'. Data released is the date the dataset was issued, this might be prior to when the dataset is being added to data.gov.ie. Data updated indicates the last modification date of the dataset. The temporal coverage shows what time period the dataset covers, for example, a particular month or year. If more context is required, this can be added under temporal notes.

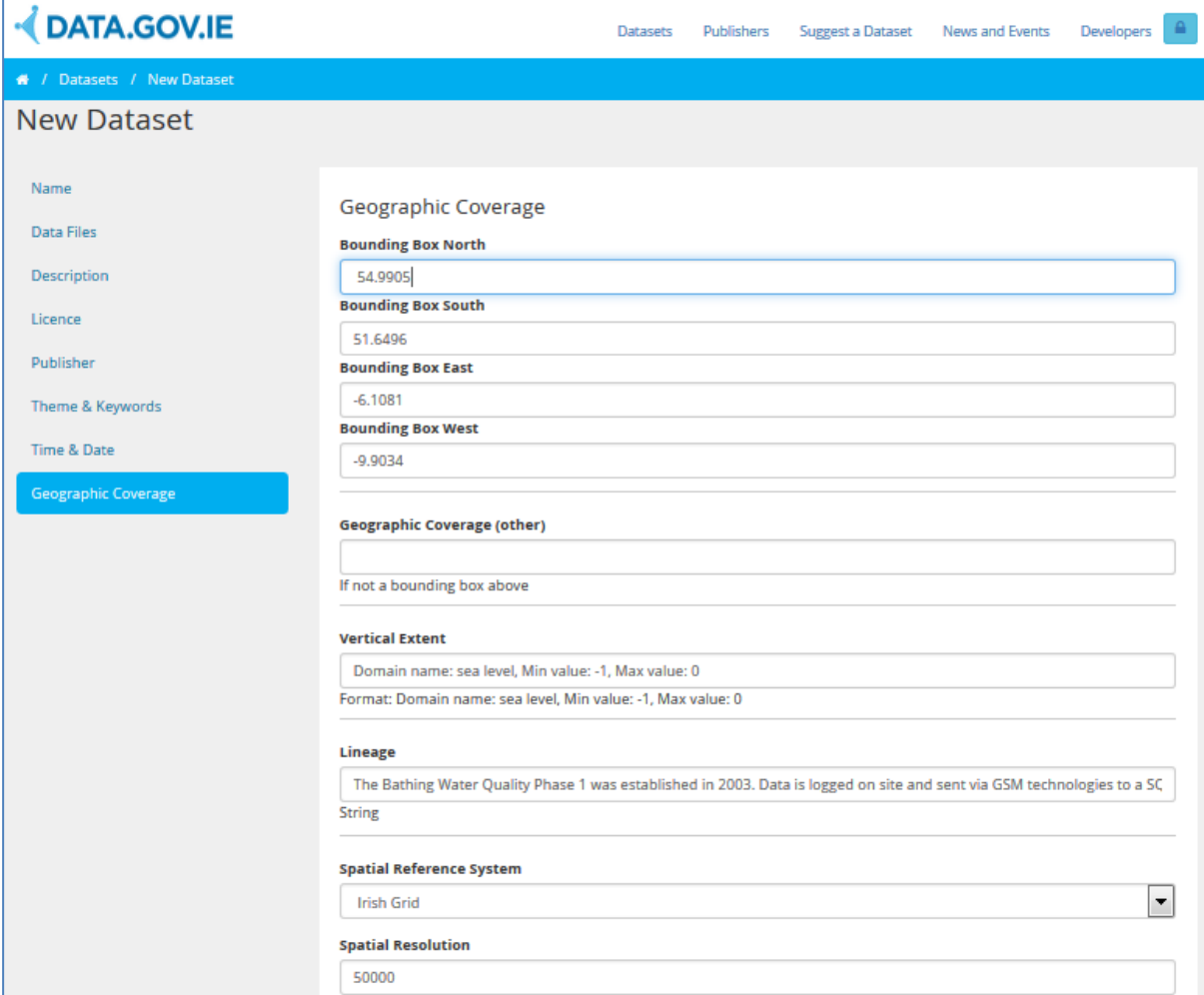
The screenshot shows the 'New Dataset' form on the data.gov.ie website. The page has a blue header with the logo and navigation links: Datasets, Publishers, Suggest a Dataset, News and Events, and Developers. Below the header is a breadcrumb trail: / Datasets / New Dataset. The main content area is titled 'New Dataset' and contains a sidebar on the left with menu items: Name, Data Files, Description, Licence, Publisher, Theme & Keywords, Time & Date (highlighted in blue), and Geographic Coverage. The main form area is titled 'Time & Date' and contains the following fields and instructions:

- Date released:** A text input field containing '01/02/2015'. Below it, the format is specified as 'Format: DD/MM/YYYY' and a note states: 'This is probably not the date that it is uploaded to data.gov.ie. Be careful not to confuse a new 'version' of some data with a new dataset covering another time period or geographic area.'
- Date updated:** A text input field containing '01/06/2015'. Below it, the format is specified as 'Format: DD/MM/YYYY' and a note states: 'This is not necessarily the date when it was updated on data.gov.ie. As with 'Date released', this is for updates to a particular dataset, such as corrections or refinements, not for that of a new time period.'
- Temporal coverage:** Two text input fields separated by a hyphen. The first field contains '01/01/2014' and the second contains '31/12/2014'. Below this, an example is given: 'e.g. 21/03/2007 - 03/10/2009 or 07:45 31/03/2006'. A note follows: 'If available, please indicate the time as well as the date. Where data covers only a single day, the second box can be left blank.'
- Temporal Notes:** A text input field containing 'Annual'.

At the bottom of the form, there are three buttons: 'Save and Finish', '« Back', and 'Next »'. A disclaimer is located between the 'Save and Finish' and '« Back' buttons: 'Important: By submitting content, you agree to release your contributions under the licence conditions of the site.'

## Tab 8: Geographic Coverage

The fourth tab is 'Geographic Coverage'. If there is geographic metadata associated with the dataset, enter it in this tab. Geographic coverage can refer to a bounding box or a particular geographic area, e.g. national, regional, etc. Vertical extent includes domain name, min value, and max value. Lineage provides more information about the provenance of the data, or context of how it was collected. Spatial reference system must refer to one of the recommended spatial reference systems from the Irish geospatial community<sup>4</sup>, as listed in the drop-down menu.



The screenshot shows the 'New Dataset' form on the DATA.GOV.IE website. The 'Geographic Coverage' tab is selected in the left-hand navigation menu. The form fields are as follows:

- Geographic Coverage**
  - Bounding Box North:** 54.9905
  - Bounding Box South:** 51.6496
  - Bounding Box East:** -6.1081
  - Bounding Box West:** -9.9034
  - Geographic Coverage (other):** (Empty text box)
  - If not a bounding box above
- Vertical Extent**
  - Domain name:** sea level, Min value: -1, Max value: 0
  - Format:** Domain name: sea level, Min value: -1, Max value: 0
- Lineage**
  - String:** The Bathing Water Quality Phase 1 was established in 2003. Data is logged on site and sent via GSM technologies to a SC
- Spatial Reference System**
  - Dropdown:** Irish Grid
- Spatial Resolution**
  - Text:** 50000

<sup>4</sup> <https://data.gov.ie/technical-framework#recommended-metadata-schema-for-open-data>



# Tab 9 API and Linked Data

Enter details of API and Linked data if applicable

The screenshot shows a web browser window with the URL <https://data.gov.ie/dataset/new#tab-api>. The page title is "New Dataset" and the user is logged in as "marian.beakey". The navigation menu includes Home, Datasets, Publishers, Suggest Data, Showcases, Contact, About, More, and Tools. The breadcrumb trail is "Home / Datasets / New Dataset".

The main content area is titled "New Dataset" and contains a sidebar with the following menu items: Name, Data Files, Description, Licence & Rights, Publisher, Theme & Keywords, Time & Date, Geographic Coverage, and "API & Linked Data" (which is highlighted in blue).

The "API" section contains the text: "Your dataset does not seem to have any associated API resources. Add one, if this exists." Below this is a horizontal line.

The "Linked Data" section contains the text: "Your dataset does not seem to have any associated Linked Data resources. Add one, if this exists." Below this is a horizontal line.

An "Important" notice states: "By submitting content, you agree to release your contributions under the licence conditions of the site." To the right of this notice are three buttons: "Back", "Next", and "Save and Finish".

The footer contains contact information for the Open Data Unit, Department of Public Expenditure & Reform, including social media handles (@GovDataIE, @IRLDeptPER), email (opendata@per.gov.ie), and phone number (+353 1 6767571). It also mentions "Powered by ckan" and "Site by Derilinx".

The Windows taskbar at the bottom shows the time as 15:26 on 17/01/2018.