

Local government data schemas

Schema guidance

What is this document for?

This document explains how to use the [schema validator tool](#) (sometimes called the “CSV checker”) to check your data against schemas which define how to publish information in tabular or spreadsheet formats.

What is a schema?

A schema defines the format and structure of data.

Real world data often has complicated structures expressed in formats such as the extensible mark-up language (XML) requiring specialist tools and expertise to interpret. Because such expertise is not widely available in Local Government, many schemas are limited to defining data in tabular format.

Tabular format data is expressed as rows and columns using the Comma Separated Values (CSV) open format. It is read using spreadsheet software such as Microsoft Excel, Open Office or Google Drive spreadsheets.

Why have schemas in local government?

The Local Government Association is helping councils adopt common schemas across councils in order to improve consistency between councils datasets. Improved consistency means that datasets can be aggregated and compared more easily across areas.

Schemas are normally optional. They just represent suggestions as to how data can be structured and functional or legal requirements met. Inevitably local decisions will result in variances between the formats most suitable in specific circumstances.

How do schemas help with the Transparency Code?

The Department for Communities and Local Government [Transparency Code](#) mandates publication of certain datasets by English councils. The Code stipulates what data must be published and what is recommended. It does not specify precise formats.

For each dataset required by the code that is suitable for publishing in tabular format, the Local Government Association (LGA) provides guidance on the fields that might be used. [See here for the LGA's guidance.](#)

Online schemas define precise fields (representing columns in tabular datasets) that can be used to meet each requirement. Online tools let you validate datasets that confirm to schemas. Because different councils might choose different fields to comply with each requirement of the Code, the Validator tool lets you choose which optional fields you want before validating a dataset.

Local government data schemas

Schema guidance

Hence the schemas and supporting tools help you confirm your council is meeting the requirements of the Code. They also help you to make the data more comparable with data from other councils and more useful to data consumers.

Where are the schemas?

The effective service delivery (ESD) [Open Data pages](#) show schemas for the following datasets:

Transparency Code CSV schemas

- <http://schemas.opendata.esd.org.uk/Contracts>
- <http://schemas.opendata.esd.org.uk/LandAndBuilding>
- <http://schemas.opendata.esd.org.uk/OrganisationStructure>
- <http://schemas.opendata.esd.org.uk/Parking>
- <http://schemas.opendata.esd.org.uk/SeniorEmployees>
- <http://schemas.opendata.esd.org.uk/SeniorSalaryCount>
- <http://schemas.opendata.esd.org.uk/Spend>

Local open data incentive CSV scheme schemas

- <http://schemas.opendata.esd.org.uk/PlanningApplications>
- <http://schemas.opendata.esd.org.uk/PremisesLicences>
- <http://schemas.opendata.esd.org.uk/PublicToilets>

Inventory XML schema

- <http://schemas.opendata.esd.org.uk/Inventory>

Each schema page links to related documentation including guidance specific to the schema and/or the Validator page (see below).

Schema pages also specify the local government [Functions list](#) and [Services list](#) entries under which the datasets they describe should be categorised in inventories.

The scheme is designed to encourage councils to publish open data in a consistent format on popular themes.

What does the Validator do?

The Validator lets you check your dataset against a CSV schema.

Local government data schemas

Schema guidance

The tool shows which fields are required within a schema and which are optional. It lets you choose which optional fields you want to include. Then it lets you do any of the following:

- Generate a page documenting the schema to match your selection
- Generate a blank CSV spreadsheet (ie header row only) for all fields in your selection
- Generate a schema in the format required to validate your dataset with the Open Data Institute's (ODI's) [CSV Lint](#) tool
- Validate your dataset directly by either uploading it or providing its URL

The Validator shows schema versions, requirements, fields, mandates, cardinality and (in the "Document" page) field types. Each of these is explained in the following sections.

Schema versions

Schemas for the transparency code are numbered from 1.00 when they move from draft to official. Thereafter the major version number will be updated each time a new field is added or a backwards incompatibility introduced. Other changes will result in the minor version number being incremented (eg from 1.00 to 1.01).

Schema requirements – mandatory and optional

The "Requirement" column of the Validator page indicates the purpose of the data field(s). For the Transparency Code, requirements listed are those given in Annex A to the [Transparency Code Document](#).

Mandatory requirements are denoted by ✓. Requirements recommended (but not mandated) by the Code or by the LGA are denoted by a checkbox, which you should check if you intend to meet the requirement.

If a requirement has a cardinality that starts with 1 (rather than 0), each row of data must have a value in at least one of the requirement's fields.

Fields - mandatory and optional

The "Field" column of the Validator page indicates fields that can be used to satisfy each requirement.

Mandatory fields (for a mandated or chosen requirement) are denoted by ✓. Other fields are denoted by a checkbox, which you should check if you intend to include the field.

The field cardinality specifies the minimum and maximum number of values you may have in each row of your dataset.

Local government data schemas

Schema guidance

Cardinality

Cardinality gives the number of values that may be present in a field (or requirement) for one row. Cardinality is one of the following.

- 0..1 – zero or one values.
- 0..∞ – any number of values.
- 1 – one and only one value.
- 1..∞- at least one value.

If the field allows zero values, a null value is accepted. Null is represented by an empty field and NOT by the string “null” or “N/A”.

Where a single row contains more than one value for a field, values should be separated by the pipe (|) character.

Types of field

Field types, as shown in the “Document” page of the Validator are described below.

Boolean

True, Yes, False or No. These are not case sensitive.

Date

A date in the format dd/mm/yyyy. If time is included, this will be ignored.

Encoded list

Encoded lists are fixed sets of labels from which you select one or more values to populate a field.

Capitalisation (that is, use of upper and lower case letters) does not have to match, but any other variances from the encoded list specified will result in your data not being valid.

Number

A number, which may or may not have decimal places, except where stated otherwise. Note that zero is considered as a value that is distinguished from a null (or absent) value.

String

A single line of alphabetic and/or numeric characters.

Local government data schemas

Schema guidance

Text

One or more lines of text. Multiple lines can be delimited by carriage return and line feed characters. If these characters are used, the entire field value must be enclosed in quotes.

URI

Uniform resource identifier. URIs are always followed by a string field giving the label of the applicable URI.

URIs are machine readable codes that precisely identify things, like local authorities, wards, or types of service. URIs commonly used in UK local government can be looked up from this page: <http://uris.opendata.esd.org.uk>.

Getting more help

Our [help page](#) gives answers to common questions. The Knowledge Hub's [Local Open Data Community](#) includes a forum where the scheme is discussed. You need to register separately to participate in discussion. If your email address does not belong to a public sector organisation, email knowledgehub@capacitygrid.com to request access.

If the above don't address your query, email us at support@esd.org.uk.

Version	Date	Description
0.01	30 Nov 2014	Initial draft release for internal review
1.00	16 Dec 2014	First formal release following LGA review

Ref 07403