



Data Farm

Data Farm srl

C.so Martinetti 4/6

16149 Genova

tel +39010 0985000

email info@data-farm.it

P.I./C.F. 02841630995

R.E.A. GE-514829

LiveDataLinker – introduction

Data Farm developed LiveDataLinker, an integrated platform that allows users to access dynamic contents provided through Open Data Platforms and to post-process such contents to deliver live analyses.

Dynamic contents are those characterized by change, activity, and progress. Very often the data exposed by Public Administrations belong to the category of Dynamic Data, as they are periodically updated to reflect real world changes. For example, traffic, weather, goods and people movements, economic values and indicators, and other real-world data strongly depend on time and consequently lend themselves to be represented as time series.

Unfortunately, the most widespread tool to process data, namely the spreadsheet, is prevalently proposed and used as a personal productivity support tool rather than as a collaboration/social tool. Data Farm believes that a switch of mentality is necessary. Such a switch aims at enabling the huge quantity of people that can use spreadsheets to extend their activity from the category of data consumers to the category of data providers and to share their analyses over a public network space.

To obtain such a goal, Data Farm integrates several technologies and platforms as described below.

In summary Data Farm aims at introducing the concept of Dynamic Open Data by enriching current platforms, namely CKAN and Idra, by adding registration to content updates and connection with T-Space for automatic Open Data inspection through MS Excel and Analysis socialization over the T-Space.

LiveDataLinker – Installation guide

The LiveDataLinker service is distributed as a Java web application (war file). To use it, it must be deployed in a Java servlet container (e.g., [tomcat](#)).

The service integrates with the T-Space platform, and in order to operate it is necessary to register a user on any instance of SpreadSheetSpace.

The service integrates with the T-Space platform, and in order to operate it is necessary to register a user on any instance of SpreadSheetSpace (e.g., the [public instance](#) or an on-premises instance).

Finally, it will be necessary to create a file with configuration parameters called "sss_opendata.config," containing the url of the SpreadSheetSpace server used and the AuthToken of the user created.

```
sss_server= https://www.spreadsheetspace.net:443
```

```
sss_token=ae38ac4f098440d19b220eb12d39c120
```

LiveDataLinker – Rest API reference

In addition to the native integration available on CKAN, it is possible to integrate the LiveDataLinker service with any platform used by invoking the publishOpenData API provided by the system.



Data Farm

Data Farm srl
C.so Martinetti 4/6
16149 Genova
tel +39010 0985000
email info@data-farm.it
P.I./C.F. 02841630995
R.E.A. GE-514829

The WOODS middleware service exposes a REST API that is used by the Open Data portals to forward users subscription requests to specific datasets. The REST API accepts HTTP POST requests containing a set of parameters in their body using a JSON format (Figure 1).

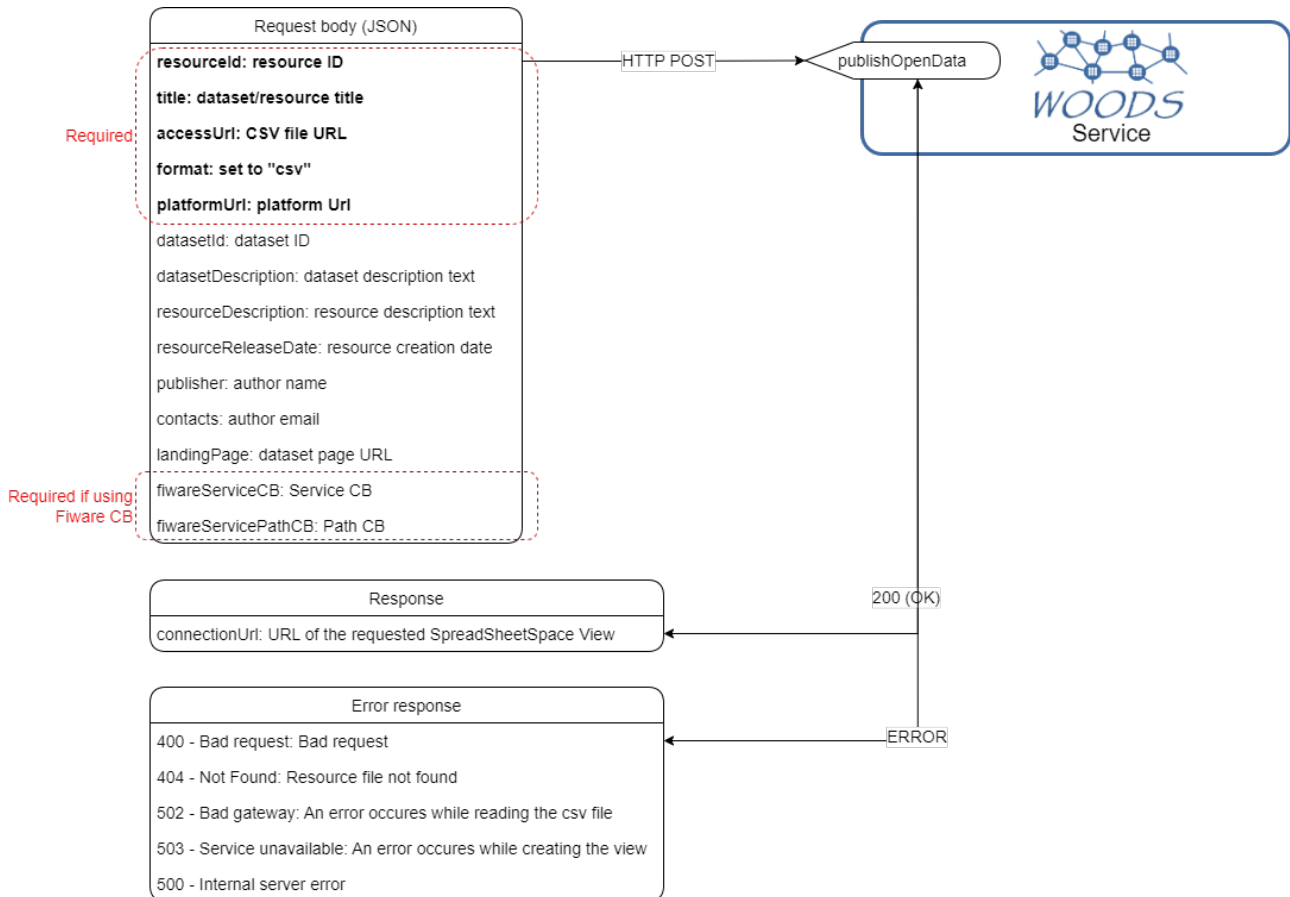


Figure 1 - `publishOpenData` API schema

As reported in the schema above, the required parameters are the following:

- `resourceId`: an unique identifier of the required resource, this is the ID used by the Open Data platform to identify a specific resource.
- `title`: the title of the requested dataset.
- `accessURL`: the URL from which it is possible to access, and therefore download, the resource's CSV file.
- `format`: the format of the file containing the required data, in this case it must be set to "csv", if other data types will be supported in the future, this parameter will be used to determine how to interpret the data.



Data Farm

Data Farm srl

C.so Martinetti 4/6
16149 Genova

tel +39010 0985000

email info@data-farm.it

P.I./C.F. 02841630995

R.E.A. GE-514829

- **platformUrl**: additional parameter referring to the platform that is invoking the service. This data is used, coupled with its “resourceId” and “datasetId”, to uniquely identify a specific resource. The WOODS service can be used by several different platforms at the same time (Hydra, CKAN, etc.) and for this reason, the only resource ID, which at the level of a single platform is certainly unique, is not enough.

The service may accept also the following optional parameters:

- **datasetId**: an identifier of the request dataset
- **datasetDescription**: the dataset’s description, if any.
- **resourceDescription**: the resource’s description, if any.
- **resourceReleaseData**: the creation date of the resource.
- **publisher**: the name of the dataset’s author.
- **contacts**: the email address of the dataset’s author.
- **landingPage**: the dataset web page URL on the source Open Data platform.
- **fiwareServiceCB**: Orion implements a simple multitenant/multiservice model-based and logical database separation; it is mandatory to identify the service/tenant.
- **fiwareServicePathCB**: indicates the path to reach the entity; the entity scope to be able to identify it if several are present.

These optional parameters are helpful to compose a single descriptive text of the dataset itself, this text will be set as a description for the SpreadsheetSpace View that will be created and will be associated to the dataset itself. This View will automatically be published and made available for other SpreadsheetSpace users in the T-Space platform and its details (title, description, etc.) might be retrieved also by browsing the T-Space web interface. In particular then, parameters like the publisher or the landingPage might be significative to correctly report the author’s information in the View’s page on the T-Space platform.

If the subscription process to the requested dataset succeeds, the service will respond with an HTTP response having the status code set to 200 (HTTP OK) and containing the requested connectionURL in its body using a JSON format.

If something instead goes wrong during the subscription process, the service will respond with an HTTP response having the status code set to a specific error code and containing a descriptive error message in its body, still using a JSON format.

CKAN integration

For details on the installation and configuration procedure of CKAN integration, please refer to the document "CKAN extension documentation.docx"