



REGIONE
PUGLIA

Open Data

Manuale servizi REST esposti da Ckan e procedura Harvest

Versione 1.3

Indice

1. Introduzione	3
2. REST Service CKAN gestione Dataset.....	3
2.1 Ottenimento API KEY	4
2.2 Specifiche dei servizi REST	4
2.2.1 get_licenses	5
2.2.2 get_resource_licenses.....	5
2.2.3 get_frequencies	6
2.2.4 get_themes	7
2.2.5 package_create	7
2.2.6 package_update.....	11
2.2.7 resource_create	12
2.2.8 resource_update.....	14
3 Tipologie di Harvester CKAN.....	14
4 REST Service CKAN interrogazione catalogo.....	15
4.1 package_list.....	15
4.2 package_show	15
4.3 organization_list.....	16
4.4 organization_show	16

1. Introduzione

Il documento descrive le modalità di fruizione dei servizi REST, nativi e non, esposti dal sistema CKAN ed utili alla creazione di dataset attraverso l'invocazione da sistemi terzi. Il documento descrive inoltre i dettagli di utilizzo della procedura di harvesting realizzata, utile all'importazione in CKAN di dataset provenienti da un file CSV reso disponibile online.

Per completezza sono riportati anche una serie di servizi REST del sistema CKAN che consentono l'interrogazione del catalogo Open Data

Per completezza d'informazione o per quanto eventualmente non riportato nel presente documento, si può far riferimento alla **Deliberazione della Giunta Regionale n.1122 del 26/05/2015** con oggetto **"Open Data - Approvazione Linee Guida regionali e Piano Annuale Open Data 2015"** (pubblicata su B.U.R.P n.87 del 22/06/2015) ed ai suoi allegati, con particolare riferimento a:

- **"Principi generali per lo sviluppo degli Open Data della Regione Puglia" (Allegato A);**
- **"Linee Guida per la pubblicazione degli Open Data - La scelta della Regione Puglia"; (Allegato B);**
- **"Linee Guida per la pubblicazione degli Open Data - Tutorial" (Allegato C);**

e alla **Deliberazione della Giunta Regionale n.422 del 28/03/2022** con oggetto **"L.R. 24 luglio 2012, n. 20 "Norme sul software libero, accessibilità di dati e documenti e hardware documentato" – Deliberazione della Giunta Regionale del 31.10.2012, n. 2183 – Deliberazione della Giunta Regionale del 26.05.2015, n. 1122 – Atto di indirizzo per l'apertura del patrimonio informativo regionale e aggiornamento della relativa normativa regionale"** (pubblicata su B.U.R.P n. 54 del 16/05/2022)

Tali Deliberazioni, con relativi allegati, sono ricercabili e scaricabili dal portale <https://burp.regione.puglia.it/>.

2. REST Service CKAN gestione Dataset

CKAN espone nativamente le API che permettono l'interoperabilità con sistemi esterni attraverso l'invocazione di servizi REST, ed in particolare consente l'inserimento e modifica di dataset e risorse. Sono state inoltre estese le API native, introducendo ulteriori servizi web utili allo scopo, che permettono di ottenere la lista di valori disponibili per le licenze dei dataset, le licenze delle risorse, i temi ed i valori di frequenza di aggiornamento.

L'invocazione dei servizi per l'inserimento/modifica di dataset e risorse (servizi nativi) necessitano di una modalità di autenticazione tramite API KEY. Nessuna autenticazione è necessaria invece per i servizi di sola lettura (servizi estesi).

È possibile visionare la documentazione dei servizi REST nativi di CKAN direttamente sul sito ufficiale al link: <https://docs.ckan.org/en/2.9/api/index.html>

I servizi REST, nativi e non, sono esposti sul path: [https://dati.puglia.it/ckan/api/3/action/\\$nome_servizio](https://dati.puglia.it/ckan/api/3/action/$nome_servizio).

Un esempio di path è il seguente: https://dati.puglia.it/ckan/api/3/action/get_licenses

2.1 Ottenimento API KEY

Per creare/modificare dei contenuti in CKAN, occorre invocare i servizi in maniera autenticata. L'autenticazione in CKAN si ottiene con il passaggio ad ogni invocazione di una chiave, denominata "API KEY" che identifica un utente del sistema. Ogni utente censito nel sistema ha dunque una propria "API KEY" che può essere utilizzata nell'invocazione dei servizi REST affinché tale operazione venga effettuata in maniera autenticata. L'API KEY da utilizzare nell'invocazione del servizio REST deve essere quella di utente che abbia in CKAN i permessi per effettuare l'operazione richiesta. Se non si conosce un utente che abbia i privilegi necessari o se si preferisca non utilizzare un utente corrispondente ad una persona censita, è possibile censire un nuovo utente corrispondente al sistema terzo chiamante, da cui poi recuperare l'API KEY.

Una volta scelto o censito l'utente col quale si intendono invocare i servizi REST, è possibile ottenerne la relativa API KEY accedendo alla webapp di CKAN con tale utente, e recarsi alla pagina "\$ckanbaseurl/user/\$username" dove \$username sarà quello dell'utente. In fondo alla pagina, sulla sinistra sarà riportato, sotto l'etichetta "Chiave API" il codice alfanumerico di nostro interesse.

Un esempio dell'endpoint del servizio è: <https://dati.puglia.it/ckan/user/mariorossi>

2.2 Specifiche dei servizi REST

Tutti i servizi esposti, restituiscono come risposta all'invocazione un JSON nel formato sotto indicato.

Proprietà	Descrizione
success	Contiene true o false in base all'esito della richiesta

result	Contiene il contenuto informativo richiesto
error	JSON che contiene il dettaglio degli errori rilevati

Nei paragrafi successivi sarà descritto il contenuto atteso nella proprietà "result" restituita da ogni specifico servizio.

2.2.1 get_licenses

Restituisce tutte le licenze presenti nel sistema, associabili ai dataset.

Endpoint: https://dati.puglia.it/ckan/api/3/action/get_licenses

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è a sua volta un JSON Array di due elementi, il primo elemento costituisce il codice identificativo della licenza, il secondo costituisce la descrizione leggibile.

Esempio di risultato:

```
[
  ["other-nc", "Altro (Non Commerciale)"],
  ["other-pd", "Altro (Public Domain)"],
]
```

2.2.2 get_resource_licenses

Restituisce tutte le licenze presenti nel sistema, associabili alle risorse.

Endpoint: https://dati.puglia.it/ckan/api/3/action/get_resource_licenses

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è un JSON nel formato sotto indicato.

Proprietà	Descrizione
value	Identificativo della licenza (es. https://w3id.org/italia/controlled-vocabulary/licences/A21_CCBY40)
text	Descrizione leggibile della licenza (es. "Creative Commons Attribuzione 4.0 Internazionale (CC BY 4.0)")
depth	Ogni licenza ha una struttura gerarchica, questa proprietà indica la profondità gerarchica della licenza. Si tratta di un valore numerico che parte da 1. (es. 2)

2.2.3 get_frequencies

Restituisce tutte i possibili valori di frequenza di aggiornamento presenti nel sistema.

Endpoint: https://dati.puglia.it/ckan/api/3/action/get_frequencies

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è un JSON nel formato sotto indicato.

Proprietà	Descrizione
value	Si tratta del codice identificativo (es. "BIHOURLY")
text	Descrizione leggibile (es. "ogni due ore")

2.2.4 get_themes

Restituisce tutti i temi presenti nel sistema, associabili ai dataset.

Endpoint: https://dati.puglia.it/ckan/api/3/action/get_themes

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è un JSON nel formato sotto indicato.

Proprietà	Descrizione
value	Codice identificativo del tema (es. "AGRI")
text	Descrizione leggibile del tema (es. "Agricoltura, pesca, silvicoltura e prodotti alimentari")

2.2.5 package_create

Permette di inserire un nuovo dataset in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.create.package_create

Endpoint: https://dati.puglia.it/ckan/api/3/action/package_create

Method: POST

Autenticazione: occorre aggiungere nell'header della richiesta, una chiave denominata "X-CKAN-API-Key" e come corrispettivo valore inserire l'API KEY ottenuta come descritto al paragrafo 2.2.

Payload: JSON nel formato sotto indicato

Result: JSON contenente i dati relativi al dataset appena creato. Assente in caso di errore nell'operazione. Contiene la proprietà "id" (stringa alfanumerica) da utilizzare, se necessario, come package_id nei servizi che ne necessitano (es. "resource_create").

Error: JSON contenente il dettaglio degli errori rilevati, nel formato sotto descritto. Assente in caso di successo dell'operazione.

Payload	
Proprietà	Descrizione
name (obbligatorio)	Nome del dataset Deve essere una stringa di lunghezza compresa tra 2 e 100 caratteri e può contenere esclusivamente caratteri alfanumerici in minuscolo, - e _
frequency (obbligatorio)	Identificativo della frequenza (recuperabile tramite il servizio "get_frequencies")
modified (obbligatorio)	Data di ultima modifica del dataset
identifier (obbligatorio)	Identificativo del dataset su CKAN
notes (obbligatorio)	Descrizione del dataset
theme (obbligatorio)	Identificativo del tema da associare al dataset (recuperabile tramite il servizio "get_themes")
owner_org (obbligatorio)	Identificativo dell'organizzazione cui si intende associare il dataset (recuperabile tramite il servizio nativo "organization_list")

title	Il titolo del dataset (di default viene posto uguale a quando indicato come "name")
private	True/false indica se si intende rendere o meno il dataset privato
author	Il nome dell'autore del dataset
author_email	Indirizzo e-mail dell'autore del dataset
maintainer	Il nome del manutentore del dataset
maintainer_email	Indirizzo e-mail del manutentore del dataset
license_id	Identificativo della licenza (recuperabile tramite il servizio "get_licenses" o in alternativa utilizzando il servizio nativo "license_list")
url	URL di provenienza del dataset
version	Stringa di massimo 100 caratteri
state	Può assumere il valore "active" o "deleted". Soltanto i dataset in stato "active" saranno visibili nei risultati di ricerca e nella lista dei dataset. Questo parametro viene ignorato qualora l'API KEY inserita per l'autenticazione appartenga ad un utente che non sia autorizzato a modificare questo parametro.
type	Tipologia di dataset
resources	Lista di risorse, nel formato indicato per il servizio "resource_create"
tags	Lista di oggetti nel formato indicato sulla documentazione ufficiale per il servizio (tag_create)
extras	Lista di oggetti JSON aventi obbligatoriamente il formato key/value. key/value.

	Es. [{"key": "Informazione Extra di esempio", "value": "Questa è l'informazione che voglio aggiungere"}, {"key": "Altra informazione extra", "value": "Contenuto della seconda informazione extra"}]
relationships_as_object	Lista di oggetti nel formato indicato nella documentazione ufficiale per il servizio "package_relationship_create"
relationships_as_subject	Lista di oggetti nel formato indicato nella documentazione ufficiale per il servizio "package_relationship_create"
groups	Lista dei gruppi cui deve appartenere il dataset. Occorre indicare per ogni elemento della lista almeno uno tra i parametri "id" e "name". I gruppi esistenti sul sistema sono recuperabili tramite il servizio nativo "group_list".

Error (presente se success è false)	
Proprietà	Descrizione
__type	Stringa che indica il tipo di errore (es. "Validation Error")

<p>\$nome_proprieta (nome della proprietà per cui si sono verificati problemi)</p> <p>Es. "owner_org": ["An organization must be provided"]</p> <p>N.B.Se l'errore è relativo ad una proprietà annidata rispetto al JSON atteso nel payload, viene riportata l'intera gerarchia (JSON) del dato oggetto di errore.</p> <p>Es. Problema rilevato nel campo extras -> key</p> <pre>"extras": [{ "key": ["Esiste un campo di schema con lo stesso nome"] }]</pre>	<p>Array di stringhe contenente la lista degli errori rilevati per questa proprietà</p>
<p>\$nome_proprieta2 (nome di un'altra eventuale proprietà per cui si sono verificati problemi)</p>	<p>Array di stringhe contenente la lista degli errori rilevati per questa proprietà</p>
<p>.....</p>	

2.2.6 package_update

Permette di modificare un dataset già esistente in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.update.package_update

Endpoint: https://dati.puglia.it/ckan/api/3/action/package_update

Method: POST

Autenticazione: occorre aggiungere nell'header della richiesta, una chiave denominata "X-CKAN-API-Key" e

come corrispettivo valore inserire l'API KEY ottenuta come descritto al paragrafo 2.2.

Le caratteristiche in termini di payload e risultato sono le medesime di quanto già visto per il servizio "package_create". Occorre indicare nel payload il campo "name" oppure "id" del dataset che si intende modificare.

È possibile ottenere la lista dei dataset presenti a sistema invocando il servizio nativo "package_list" ed possibile visualizzarne i dettagli invocando il servizio nativo "package_show".

2.2.7 resource_create

Permette di inserire una nuova risorsa in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.create.resource_create

Endpoint: https://dati.puglia.it/ckan/api/3/action/resource_create

Method: POST

Autenticazione: occorre aggiungere nell'header della richiesta, una chiave denominata "X-CKAN-API-Key" e come corrispettivo valore inserire l'API KEY ottenuta come descritto al paragrafo 2.2.

Payload: JSON nel formato sottoindicato

Result: JSON contenente i dati relativi alla risorsa appena creata. Assente in caso di errore nell'operazione.

Error: JSON contenente il dettaglio degli errori rilevati, nel formato sotto descritto. Assente in caso di successo dell'operazione.

Payload	
Proprietà	Descrizione
package_id (obbligatorio)	Id del dataset cui associare la risorsa (es. "d7c36c39-a637-42e3-b370-09c40eafe460")

name (obbligatorio)	Nome della risorsa
revision_id	Stringa
description	Descrizione della risorsa
format	Stringa che identifica il formato (es. "csv")
hash	Stringa rappresentante l'hash della risorsa
resource_type	Tipologia di risorsa
mimetype	Stringa rappresentante il mime type della risorsa (es. "text/csv")
mimetype_inner	Stringa rappresentante il mime type di dettaglio della risorsa (es. "text/csv")
cache_url	Stringa indicante un URL
size	Valore intero
created	Stringa che rappresenta la data di creazione, in formato ISO (es. 2019-12-31)
last_modified	Stringa che rappresenta la data di ultima modifica, in formato ISO (es. 2019-12-31)
cache_last_updated	Stringa che rappresenta la data di ultimo aggiornamento della cache, in formato ISO (es. 2019-12-31)

upload	Da utilizzare per effettuare l'upload del file della risorsa. In tal caso occorre anche indicare nell'header della richiesta che si tratta di un "multipart-form-data".
--------	---

2.2.8 resource_update

Permette di modificare una risorsa già esistente in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.update.resource_update

Endpoint: https://dati.puglia.it/ckan/api/3/action/resource_update

Method: POST

Autenticazione: occorre aggiungere nell'header della richiesta, una chiave denominata "X-CKAN-API-Key" e come corrispettivo valore inserire l'API KEY ottenuta come descritto al paragrafo 2.2.

Le caratteristiche in termini di payload e risultato sono le medesime di quanto già visto per il servizio "package_create". Occorre indicare nel payload il campo "id" relativo alla risorsa che si intende modificare. È possibile ottenere la lista delle risorse relative ad un dataset invocando il servizio nativo "package_show" che mostra tutti i dettagli di un dataset e delle sue risorse, oppure ricercare una risorsa utilizzando il servizio nativo "resource_search".

3 Tipologie di Harvester CKAN

Le tipologie di Harvester CKAN utilizzate per la raccolta di dati da sistemi terzi verso il portale Open Data regionale sono varie e la scelta dipende dal sistema dal quale effettuare la raccolta dei dati.

Di seguito l'elenco disponibile:

- **CKAN:** Raccolta dati da una istanza remota di CKAN;
- **Generic DCAT RDF Harvester:** Raccolta dati per dataset DCAT contenuti in formato RDF
- **DCAT JSON Harvester:** Raccolta dati per dataset DCAT da un formato JSON
- **CKAN Harvester for DCATAPIT:** Versione speciale di CKAN Harvester, che mappa i gruppi in base ai temi

4 REST Service CKAN interrogazione catalogo

CKAN espone nativamente le API che permettono l'interoperabilità con sistemi esterni attraverso l'invocazione di servizi REST, ed in particolare consente l'interrogazione del catalogo

È possibile visionare la documentazione dei servizi REST nativi di CKAN direttamente sul sito ufficiale al link: <https://docs.ckan.org/en/2.9/api/index.html>

I servizi REST, nativi e non, sono esposti sul path: [https://dati.puglia.it/ckan/api/3/action/\\$nome_servizio](https://dati.puglia.it/ckan/api/3/action/$nome_servizio).

Un esempio di path è il seguente: https://dati.puglia.it/ckan/api/3/action/get_licenses

Tra le API di CKAN sono disponibili:

4.1 package_list

Restituisce i dataset pubblici presenti nel sistema.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.get.package_list

Endpoint: https://dati.puglia.it/ckan/api/3/action/package_list

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è un una stringa indicante il nome di un dataset.

4.2 package_show

Permette di visualizzare i dettagli di un dataset già esistente in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.get.package_show

Endpoint: https://dati.puglia.it/ckan/api/3/action/package_show?name_or_id={dataset_name_or_id}

Method: GET

Autenticazione: per visualizzare il dettaglio di un dataset privato occorre aggiungere nell'header della richiesta, una chiave denominata "X-CKAN-API-Key" e come corrispettivo valore inserire l'API KEY ottenuta come descritto al paragrafo 2.2. L'autenticazione deve appartenere ad un utente che ha i permessi di

visualizzazione del dataset privato. Se il dataset è pubblico è possibile invocare il servizio senza alcuna autenticazione.

Parametri: name_or_id ossia l'identificativo del dataset di interesse

Result: JSON con le caratteristiche del dataset

Le caratteristiche del JSON sono le medesime di quanto già visto per il servizio "package_create".

4.3 organization_list

Restituisce le organizzazioni presenti nel sistema.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.get.organization_list

Endpoint: https://dati.puglia.it/ckan/api/3/action/organization_list

Method: GET

Autenticazione: non necessaria

Parametri: non necessari

Result: JSON Array

Ogni elemento dell'array è un una stringa indicante il nome di un'organizzazione.

4.4 organization_show

Permette di visualizzare i dettagli di un'organizzazione censita in CKAN.

Documentazione ufficiale:

https://docs.ckan.org/en/2.9/api/index.html#ckan.logic.action.get.organization_show

Endpoint: https://dati.puglia.it/ckan/api/3/action/organization_show?id={organization_id}

Method: GET

Autenticazione: Non necessaria

Parametri: identificativo dell'organizzazione di interesse

Result: JSON con le caratteristiche dell'organizzazione



**REGIONE
PUGLIA**

Struttura Speciale "Comunicazione Istituzionale"

Lungomare N. Sauro, 33 - 70121 Bari

Tel: + 39 080 5402813

Mail: comunicazione@regione.puglia.it

Pec: comunicazione.istituzionale@pec.rupar.puglia.it

Ufficio per la Transizione al Digitale

Lungomare N. Sauro, 33 - 70121 Bari

Tel: 080 540 3727

Mail: resp.transizionedigitale@regione.puglia.it

Pec: resp.transizionedigitale@pec.rupar.puglia.it

www.regione.puglia.it



UNIONE EUROPEA



P  **PUGLIA**
FESR·FSE
2014/2020
Il futuro alla portata di tutti


InnovaPuglia™