



AGID

Agenzia per l'Italia Digitale

FormezPA

FORMAZIONE AGID – FORMEZ SULLA TRANSIZIONE DIGITALE DELLA PA

**Progetto Informazione e formazione per la transizione digitale della PA
nell'ambito del progetto «Italia Login – la casa del cittadino»**

(A valere sul PON Governance e Capacità Istituzionale 2014-2020)



Dati aperti dalla teoria alla pratica: identificazione e analisi

Francesco Piero Paolicelli



dati.gov.it

i dati aperti della pubblica amministrazione

Cerca nel sito



Dati ▾

Fare Open Data ▾

Monitoraggio ▾

Sviluppatori

Scrivi alla redazione

Cerca nei metadati



[Ricerca avanzata](#)

naviga i dati per categoria tematica



Agricoltura, pesca,
silvicoltura e prodotti
alimentari



Economia e finanze



Istruzione, cultura e sport



Energia



Ambiente



Governo e settore
pubblico



Salute



Tematiche internazionali

Individuazione



- dati nativi, cioè dati generati dalle amministrazioni, perché prodotti dai vari uffici durante l'adempimento delle proprie funzioni istituzionali;
- dati mashup, cioè dati provenienti da diverse fonti e soggetti a operazioni di integrazione. *In entrambi i casi, l'apertura può essere sollecitata da soggetti esterni interessati attraverso specifiche richieste*
- sulla base della disposizione di cui all'art. 7 del **[D-LGS-33-2013]**, " *i documenti, le informazioni e i dati oggetto di pubblicazione obbligatoria ai sensi della normativa vigente, resi disponibili anche a seguito dell'accesso civico di cui all'articolo 5*" del medesimo decreto;
- ai sensi dell'art. 50-quater del **[CAD]**, tutti i dati acquisiti e generati nella fornitura dei servizi svolti in concessione da pubbliche amministrazioni. L'obbligo del concessionario di rendere disponibili all'amministrazione concedente detti dati come dati di tipo aperto DEVE essere inserito nei contratti e nei capitolati con i quali le pubbliche amministrazioni affidano lo svolgimento dei servizi in concessione.
- Per i nuovi dati che rientrano nell'ambito di applicazione del Decreto, il processo di produzione e pubblicazione deve essere guidato dal principio dell'**apertura fin dalla progettazione e per impostazione predefinita**, come, tra l'altro, disposto dall'art. 6 comma 4 del Decreto (open data by design).

Individuazione



Ricognizione - Per i dati nativi, all'interno dei singoli uffici o dei vari settori dell'amministrazione vanno quindi individuati quali dati, tra tutti quelli prodotti, si vogliono rendere aperti anche per abilitare nuove forme di riutilizzo dell'informazione in quanto utili all'esterno. Per i dati mashup, le amministrazioni possono raccogliere e integrare informazioni da diverse fonti interne ed esterne che concorrono alla formazione del dato. Per tale tipologia di dati, la parte più importante è la definizione delle modalità di accesso a partire dalle politiche dei singoli produttori dei dati e le relative modalità di rilascio e aggiornamento dei dati.

Analisi dei vincoli giuridici - Alla fase di ricognizione fa seguito l'analisi giuridica delle fonti del dato, fondamentale per garantire sostenibilità nel tempo del processo di produzione e pubblicazione dei dati considerando i possibili vincoli che possono impedire o limitare (anche temporalmente) l'apertura di dati e documenti, evidenziando limitazioni d'uso, finalità di competenza, determinazione dei diritti e dei termini di licenza.

Individuazione



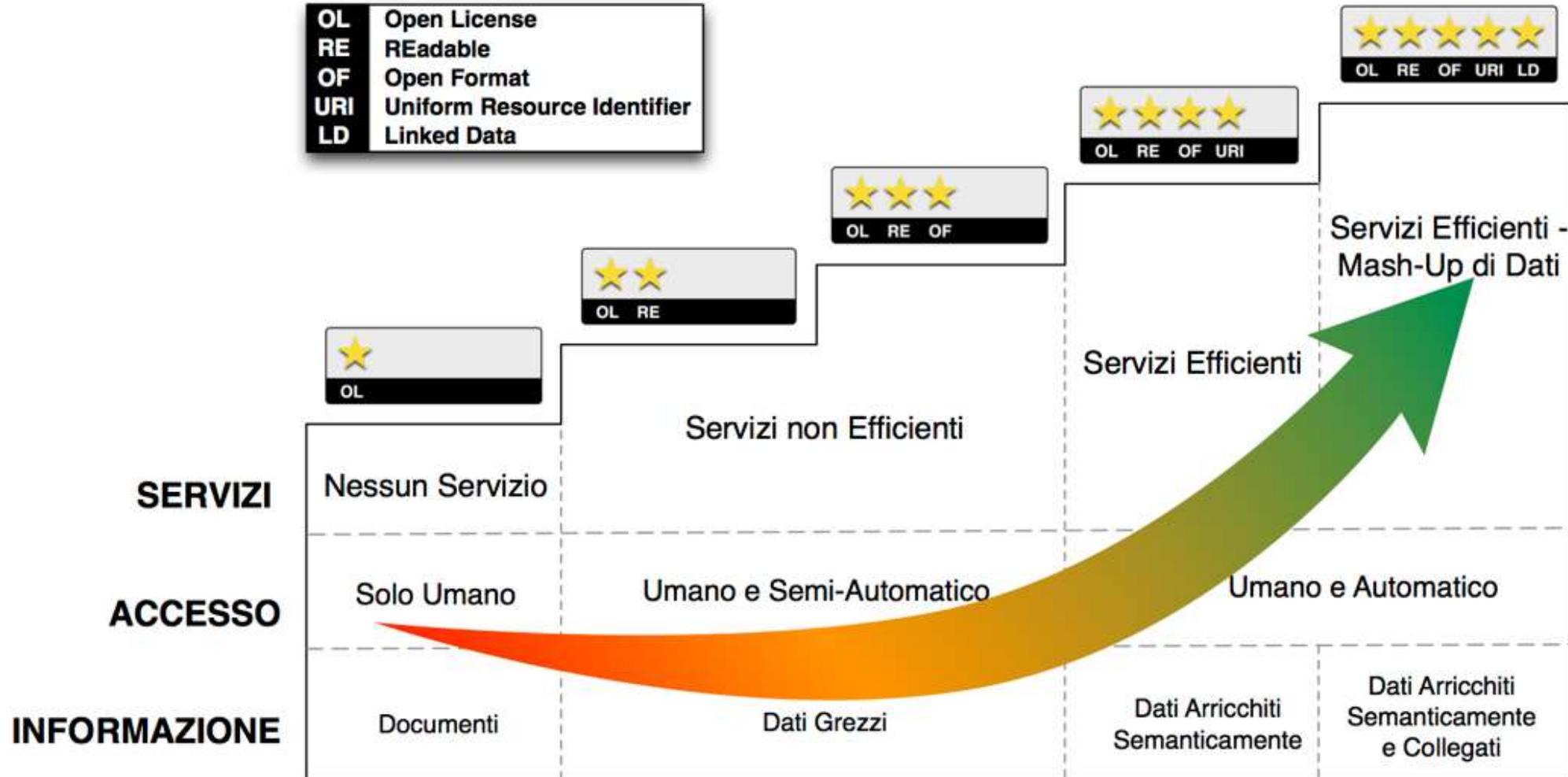
Definizione priorità e percorso di apertura - La ricognizione dei dati consente di avere contezza del patrimonio informativo, mentre l'analisi dei vincoli consente di fare una prima selezione dei dati che possono essere resi disponibili per il riutilizzo. Per i dati non soggetti a vincoli, è necessario individuare criteri e modalità di apertura anche in base ad una eventuale **scala di priorità** basata, per esempio, su un approccio di tipo "demand-driven" che tenga conto dell'impatto economico e sociale nonché del livello di interesse e delle necessità degli utilizzatori.

Alcuni criteri da considerare per assegnare la priorità nell'apertura potrebbero essere:

- le tipologie di dati individuate dal Decreto a cui è riconosciuto un enorme potenziale economico, quindi dati dinamici, serie di dati di elevato valore e dati della ricerca;
- le richieste pervenute dai riutilizzatori sulla base dell'art. 5 del Decreto e della procedura di cui al par. [Richieste di riutilizzo](#);
- i dati che le norme prevedono di rendere disponibili obbligatoriamente come dati di tipo aperto (v. par. [Individuazione e selezione](#)).

Tale attività potrebbe essere sistematizzata attraverso la definizione di un **percorso di apertura dei dati** da inserire nel **Piano Triennale ICT** che ciascuna amministrazione, secondo la roadmap definita dalle Linee d'Azione nel Piano triennale nazionale e le modalità operative fornite da AgID, è chiamata a definire anche utilizzando il format PT reso disponibile da AgID stessa. La predisposizione di tale Piano rientra tra i compiti che la **Circolare n. 3/2018** del Ministro per la pubblica amministrazione raccomanda di assegnare al RTD.

Modello 5 stars



Focus formato CSV

CSV (Comma Separated Values)

È un formato di file testuale utilizzato per rappresentare informazioni con struttura tabellare. Le righe delle tabelle corrispondono a righe nel file di testo CSV e i valori delle celle sono divisi da un carattere separatore, che, come indica il nome stesso, dovrebbe essere la virgola. Il CSV non è uno standard vero e proprio ma la sua modalità d'uso è descritta nell'RFC 4180. Nel rilascio di dati secondo il formato CSV, per agevolare i riutilizzatori, si raccomanda di dichiarare almeno 1) il separatore di campo utilizzato (e.g, virgola, punto e virgola); 2) se è stato usato un carattere per delimitare i campi di testo.

Analizziamo alcune indicazioni presenti nella [Linee Guida Europee Qualità dei dati](#) (Agosto 2021)

Focus formato CSV

•Utilizza un file per tabella

Ogni file CSV deve contenere solo una tabella. Se la tabella da pubblicare è composta da più fogli, è necessario creare un file CSV per ogni foglio.

•Assicurati che tutte le righe abbiano lo stesso numero di colonne

Ogni riga deve avere lo stesso numero di colonne e, quindi, di caratteri separatori. Se in una riga manca un valore, questo di solito viene interpretato come “null”. Ciò può comportare un trattamento errato dei dati. Se il CSV contiene righe con un numero diverso di colonne, bisognerebbe controllare se c'è un problema con valori di ‘escape’ non corretti (ad es. un valore che corrisponde al carattere separatore che in quel caso non va interpretato come tale).

•Indica le unità in una modalità facilmente elaborabile

L'unità di misura di un valore dovrebbe essere indicata nell'intestazione della colonna. Se l'unità cambia da un valore all'altro, allora bisognerebbe considerare una colonna dedicata con un'opportuna intestazione e non inserire l'unità insieme al valore stesso. Per le unità dovrebbero essere utilizzati i codici (URI) derivati da un vocabolario controllato.

•Evita gli spazi bianchi e informazioni aggiuntive nel file

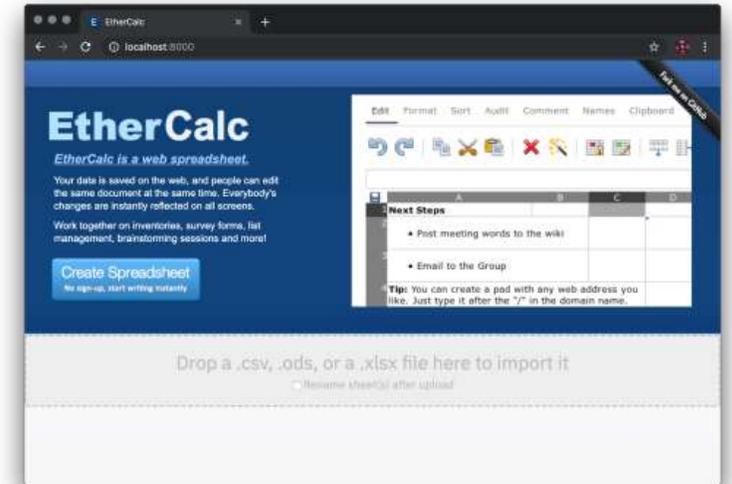
È importante assicurarsi che il file contenga solo i dati che appartengono alla tabella effettiva, come le intestazioni di colonna e i valori delle voci presenti nella tabella stessa. Nel file CSV, quindi, non devono essere presenti titolo della tabella, righe vuote o eventuali informazioni aggiuntive che aiutino l'utente a capire meglio i dati (queste ultime, che sono utilissime, vanno inserite nei metadati). Il file, inoltre, deve contenere una sola riga di intestazione.

Focus formato CSV

Strumenti più comuni per la costruzione di file CSV, in assenza di sistemi interoperabili automatici (webservice)



Google Sheets



CSV Esempi

- [Popolazione scolastica Comune di Lecce 2021-2022](#)
- [PNRR PaDigitale2026 Comuni Dipartimento TD](#)
- [Catasto rifiuti 2021 per Regioni ISPRA](#)



Andiamo a vedere LIVE..

Metadatazione e inserimento su catalogo opendata locale

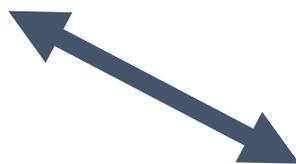


Google Sheets



ckan

Configuratore RDF per
pochi dataset



Andiamo a fare alcuni esempi ..

Linee Guida per i Cataloghi Dati - Requisiti

1. Catalogo

dcatapit:Catalog

- [1.2. Elementi obbligatori](#)
 - [1.2.1. Titolo](#) `dct:title`
 - [1.2.2. Descrizione](#) `dct:description`
 - [1.2.3. Editore](#) `dct:publisher`
 - [1.2.4. Data ultima modifica](#) `dct:modified`
 - [1.2.5. Dataset](#) `dcat:dataset`
- [1.3. Elementi raccomandati](#)
 - [1.3.1. Homepage](#) `foaf:homepage`
 - [1.3.2. Lingua](#) `dct:language`
 - [1.3.3. Data di rilascio](#) `dct:issued`
 - [1.3.4. Temi](#) `dcat:themeTaxonomy`

2. Dataset

dcatapit:Dataset

2. Elementi obbligatori

- [2.2.1. Identificativo del dataset](#) `dct:identifier`
- [2.2.2. Titolo del dataset](#) `dct:title`
- [2.2.3. Descrizione del dataset](#) `dct:description`
- [2.2.4. Ultima modifica del dataset](#) `dct:modified`
- [2.2.5. Temi del dataset](#) `dcat:theme`
- [2.2.6. Titolare del dataset](#) `dct:rightsHolder`
- [2.2.7. Frequenza aggiornamento dataset](#) `dct:accrualPeriodicity`
- [2.2.8. Distribuzione del dataset](#) `dcat:distribution`

3. Elementi raccomandati

- [2.3.1. Sottotema del dataset](#) `dct:subject`
- [2.3.2. Punto di contatto del dataset](#) `dcat:contactPoint`
- [2.3.3. Editore del dataset](#) `dct:publisher`

3. Distribuzione

dcatapit:Distribution

3.2. Elementi obbligatori

- [3.2.1. Formato Distribuzione](#) `dct:format`
- [3.2.2. URL di accesso](#) `dcat:accessURL`
- [3.2.3. Licenza](#) `dct:license`

3.3. Elementi raccomandati

- [3.3.1. Descrizione distribuzione](#) `dct:description`

Linee Guida per i Cataloghi Dati

4. Soggetto o Organizzazione

SOGGETTO M

Soggetto identificativo M

Soggetto nome M

6. Punto di contatto

dcatapit:Organization

PUNTO DI CONTATTO (ORGANIZZAZIONE) R

Punto di contatto (organizzazione) nome M

Punto di contatto (organizzazione) email M

5. Licenza

dcatapit:LicenseDocument

LICENZA M

Licenza tipo R



Raccomandato ma con massima importanza

NOVITA': Configuratore su Google Sheet



Agenzia per la
Coesione Territoriale



Istruzioni compilazione

Compilare attentamente tutti i campi:

- Catalogo
- Organizzazione
- Dataset
- Risorsa

* campi obbligatori ma si raccomanda la compilazione di ogni campo per la qualità dei metadati. Gli altri campi sono (R)=raccomandati oppure (O)=opzionali

Le date vanno inserite con il carattere - come separatore. Esempio 20-12-2020 , il configuratore le cambierà nella notazione inglese 2020-12-20 automaticamente

Istruzioni configurazione:

1. Cliccare File Crea una copia e accettare di migrare anche gli script
2. Nel foglio di lavoro così creato, ora appare nel menù "AGID". Cliccare e autorizzare il proprio account. (Avanzate, Sheet2RDF e Consenti)
3. Cliccare Estensioni - AppScript - (VECCHIO EDITOR) - Menu Pubblica - Nuova versione e fare Aggiorna. Chiudere lo script
4. Ora dal Googl Sheet, menù "AGID" è possibile scaricare il file RDF

ps: controllare se il file .rdf generato è conforme alle specifiche DCATAP_IT tramite il validatore --> <https://dati.gov.it/sviluppatori/validatore>

Agid testerà l'importazione automatica e se avrà risolto positivo, sarà importato su dati.gov.it il catalogo manuale creato di dati aperti

TITOLO*	DESCRIZIONE*	EDITORE*	CODICEIPA/PIVA*	DATARILASCIO*	WEBSITECATALOGO*	DATAMODIFICA*
PCM - Dipartimento dei Dati emergenza Ucraina 2	PCM - Dipartimento della Protezione Civile - Redazione		PCM	2022-05-20	https://github.com/pcm-dpc/UKR-2022	2022-07-17

Nome*	Telefono (diventa il punto di contatto telefono)	Website (diventa il punto di contatto web)	Email* (diventa il punto di contatto email)
PCM - Dipartimento della	800840840	emergenze.protezionecivile.gov.it/it/umanitarie	protezionecivile@pec.governo.it

NOVITA': Configuratore su Google Sheet

TITOLO*	dct:identifier	TEMA*	DESCRIZIONE*	dct:publisher	dct:identifier	ORGANIZZAZIONE	DATARILASCIO(O)	DATAMODIFICA*
Border entry	PCM border-entry	Giustizia, sistema giuridico e s	Fornisce i dati sugli ingressi in Ita In seguito alla guerra scoppiata in pcm---dipartime	PCM	PCM	PCM - Dipartimento dell	2022-05-20	2022-07-17
Subsistence allowance	PCM subsistence-allowance	Giustizia, sistema giuridico e s	Possono ottenere il contributo di sostentamento le persone proven Può essere richiesto entro il 30 settembre 2022 tramite la piattafor	PCM	PCM	PCM - Dipartimento dell	2022-05-20	2022-07-17
Expressions of interest diffused hospit	PCM expressions-of-interest-di	Giustizia, sistema giuridico e s	Si è chiuso alle ore 18 del 22 aprile 2022 il termine di presentazione Obiettivo: offrire misure di accoglienza diffusa e di accompagname Con decreto del Capo del Dipartimento del 22 aprile 2022 è stata n Delle 48 manifestazioni di interesse caricate sulla piattaforma Avvis All'esito dei lavori della Commissione, sul totale delle 47 manifesta:	PCM	PCM	PCM - Dipartimento dell	2022-05-20	2022-07-17
Humanitarian aid	PCM humanitarian-aid	Giustizia, sistema giuridico e s	L'impegno del sistema italiano di protezione civile nell'assistenza ui Dall'inizio dell'emergenza, il Dipartimento della Protezione Civile st Sulla base delle richieste di assistenza internazionale, il Dipartimen	PCM	PCM	PCM - Dipartimento dell	2022-05-20	2022-07-17
Temporary Protection	PCM temporary-protection	Giustizia, sistema giuridico e s	All'arrivo in Italia, le persone in fuga dalla guerra in Ucraina posson	PCM	PCM	PCM - Dipartimento dell	2022-05-20	2022-07-17

ID_DATASET*	URLFILE*	NOME(O)
PCM border-entry	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/border-entry/dpc-ukr-2022--border-entry.csv	Border entry CSV
PCM border-entry	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/border-entry/dpc-ukr-2022--border-entry.json	Border entry JSON
PCM subsistence-allowance	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/subsistence-allowance/dpc-ukr-2022--subsistence-allowance.csv	subsistence-allowance CSV
PCM subsistence-allowance	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/subsistence-allowance/dpc-ukr-2022--subsistence-allowance.json	subsistence-allowance JSON
PCM expressions-of-interest-diffused-hospitality	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/expression-interest-diffused-hospitality/dpc-ukr-2022--expression-interest-d	expression-interest-diffused-f
PCM expressions-of-interest-diffused-hospitality	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/expression-interest-diffused-hospitality/dpc-ukr-2022--expression-interest-d	expression-interest-diffused-f
PCM humanitarian-aid	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/humanitarian-aid/dpc-ukr-2022--humanitarian-aid.csv	humanitarian-aid CSV
PCM humanitarian-aid	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/humanitarian-aid/dpc-ukr-2022--humanitarian-aid.json	humanitarian-aid JSON
PCM temporary-protection	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/temporary-protection/dpc-ukr-2022--temporary-protection.csv	temporary-protection CSV
PCM temporary-protection	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/temporary-protection/dpc-ukr-2022--temporary-protection.json	temporary-protection JSON
PCM humanitarian-aid	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/humanitarian-aid/dpc-ukr-2022--humanitarian-aid-geo.geojson	humanitarian-aid GEOJSON
PCM humanitarian-aid	▼ https://raw.githubusercontent.com/pcm-dpc/UKR-2022/master/data/humanitarian-aid/dpc-ukr-2022--humanitarian-aid-geo.json	humanitarian-aid TOPOJSON

NOVITA': Configuratore su Google Sheet



CKAN – Back end inserimento

LIVE

VALIDATORE RDF



AGID

Agenzia per
l'Italia Digitale

Validatore DCAT-AP_IT: Risultati query

Spiacenti! Abbiamo trovato i seguenti errori o warning (51)

Se presenti, le anomalie sono di seguito riportate in tabella:

Mostro da 1 a 10 di 51 risultati

Precedente 1 2 3 4 5 6 Successivo

Class Name	Rule ID	Rule Severity	Rule Description	Message	Subject	Predicate	Object
<input type="text" value="Search Clas"/>	<input type="text" value="Search"/>	<input type="text" value="Search Rule S"/>	<input type="text" value="Search Rule Des"/>	<input type="text" value="Search Message"/>	<input type="text" value="Search Subject"/>	<input type="text" value="Search Predicate"/>	<input type="text" value="Search Object"/>
Catalog	6	▲ error	dct:description is a required property for Catalog.	The dcatapit:Catalog http://dati.anticorruzione.it/opendata does not have a dct:description.	http://dati.anticorruzione.it/opendata	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://dati.gov.it/onto/dcatapit#Catalog
Organization	212	▲ error	vcard:hasURL should be a resource.	The vcard:hasURL is not a resource.	http://dati.anticorruzione.it/opendata/organizzazione/bec21674-7ced-465e-985c-279153ef8139	http://www.w3.org/2006/vcard/ns#hasURL	
Dataset	21	● warning	dct:subject is a recommended property for Dataset.	The dcatapit:Dataset http://dati.anticorruzione.it/opendata/dataset/28629794-d17d-403c-8c54-d69f7bed6552 does not have the recommended property dct:subject.	http://dati.anticorruzione.it/opendata/dataset/28629794-d17d-403c-8c54-d69f7bed6552	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://dati.gov.it/onto/dcatapit#Dataset

Validatore di sintassi: <https://www.w3.org/RDF/Validator/>

Validatore semantico → <https://dati.gov.it/sviluppatori/validatore>



AGID

Agenzia per
l'Italia Digitale

Formez PA

dati.gov.it (scrivi alla redazione)
dati@agid.gov.it

Francesco Piero Paolicelli