



NUOVI PERCORSI DI SVILUPPO
DELLA CAPACITÀ AMMINISTRATIVA
DELLA REGIONE SICILIANA



ASSESSORATO REGIONALE
DELLE AUTONOMIE LOCALI E DELLA FUNZIONE PUBBLICA
DIPARTIMENTO DELLE AUTONOMIE LOCALI

Linea 2.5 "Rafforzamento della capacità di
attuazione dei Fondi SIE da parte degli Enti Locali"

Ciclo "I Responsabili per la Transizione al Digitale e l'innovazione negli Enti Locali"

Webinar – 7 aprile 2021

I #dati nella Pubblica Amministrazione

- Il patrimonio informativo della PA
- Gli Open data
- Esempi di utilizzo dei dati aperti

Ciro Spataro c.spataro@comune.palermo.it



Unione Europea



Repubblica Italiana



Regione Siciliana

FSE FONDO SOCIALE EUROPEO
SICILIA 2020
PROGRAMMA OPERATIVO



Formez**PA**

CESARE BATTISTI

(Trento, 1875-1916) è stato un patriota, giornalista, geografo, politico



[..] Di più avrei potuto fare, specialmente nel campo statistico, se non ci fosse nel nostro paese, e nei privati e negli enti morali, una tal quale ritrosia a **confidare al dominio del pubblico dati**, fatti e notizie. [...]

“Il Trentino” - **1898**

slide a cura di [Maurizio Napolitano](#)





Conoscenza aperta

La conoscenza è aperta quando chiunque ha libertà di accesso, uso, modifica e condivisione ad essa – avendo al massimo come limite misure che ne preservino la provenienza e l'apertura.

[Definizione dell'Open Knowledge Foundation](#)





Il potere dei dati in formato aperto

I dati, che in sé sono freddi, possono **trasformarsi in informazioni**, in forme di conoscenza, creando valore.

I dati, se sono opportunamente trattati, processati, e gestiti informaticamente, sono in grado di **descrivere, diagnosticare e predire fenomeni**.

L'utilizzo dei dati supporta, nel senso che aiuta, i decisori (politici, amministratori) a **proporre soluzioni strategiche per superare problemi o migliorare la qualità della vita** nella società.



Il patrimonio informativo della PA

Tutte le PA posseggono dati per le funzioni “pubbliche” che svolgono quotidianamente attraverso l’uso di applicativi (programmi informatici).

Posseggono una mole impressionante di dati su vari ambiti:

- Ambiente
- Amministrazione
- Attività Economiche
- Economia
- Cultura
- Turismo
- Territorio
- Elezioni
- Istruzione
- Mobilità e Sicurezza
- Opere Pubbliche
- Reti
- Sanità e Sociale
- Sport
- Urbanistica
- Salute
- Ricerca
-

I dati delle PA sono principalmente disponibili:

- **nei database degli applicativi informatici**
- **in fogli excel / access dei dipendenti pubblici, quando non esistono applicativi informatici**

Questo patrimonio di dati, fatti salvi i dati contenente privacy degli individui, deve essere reso pubblicamente disponibile sul web (GOVERNO APERTO - TRASPARENZA - OPEN DATA)

A cosa servono i dati?

Perché è importante che siano resi pubblici?

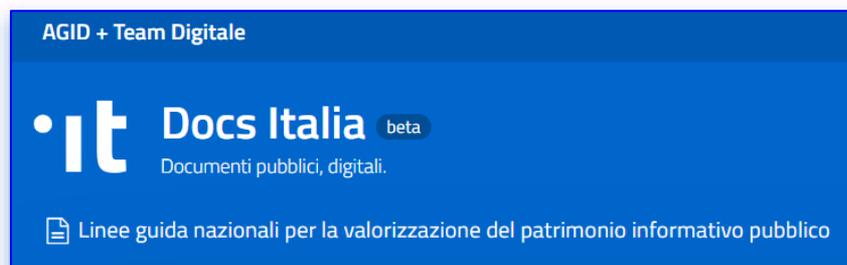


- per capire meglio l'andamento dei fenomeni nella nostra vita
- per fare monitoraggio sull'andamento di un progetto, di un attività, di un fenomeno, della ricerca scientifica/medica
- per consentire agli amministratori degli enti pubblici (assessori, sindaci, presidenti regione, dirigenti) di prendere decisioni efficaci basate sulla conoscenza precisa dei dati e non sulle sensazioni personali
- per consentire la necessaria trasparenza dell'attività amministrativa quotidiana degli enti pubblici (enti pagati con le tasse dei cittadini)
- per consentire la crescita di un mercato (economia) di soggetti privati (aziende, start up) che utilizzando i dati pubblici permette la generazione di servizi utili nella vita quotidiana dei cittadini (mobilità, ambiente, sociale,...)
- per alimentare un giornalismo (data journalism) sempre più basato sull'analisi dei dati pubblici (es. [Data Room di Milena Gabanelli](#) - [Data Ninja](#) - [Infodata_ilssole24ore](#)) così da fornire alla comunità analisi quanto più oggettive e meno soggettive, necessarie per il mantenimento della democrazia!



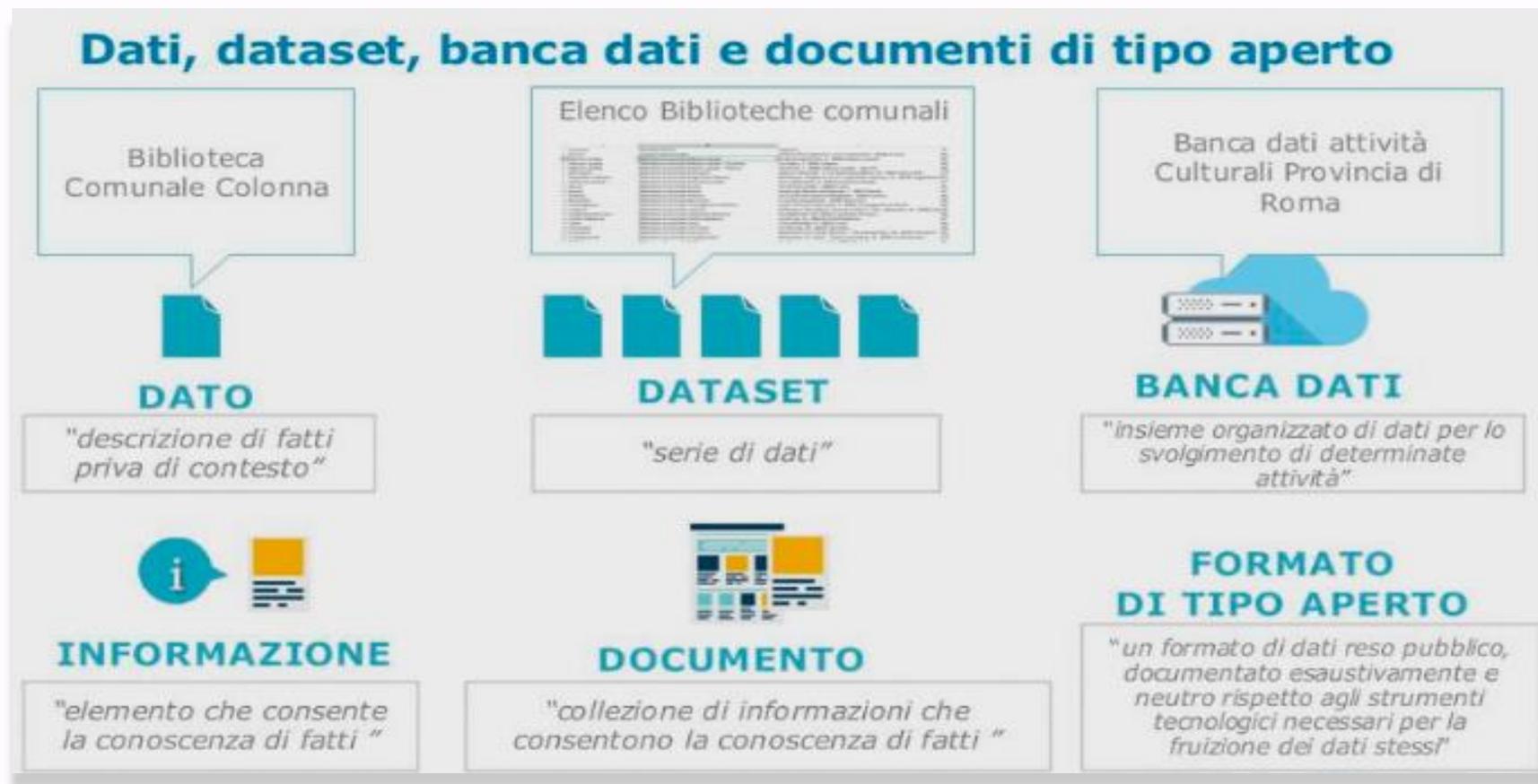
Linee guida per la valorizzazione del patrimonio informativo pubblico

Linee guida nazionali sulla valorizzazione del patrimonio informativo pubblico, a cura dell'AGID



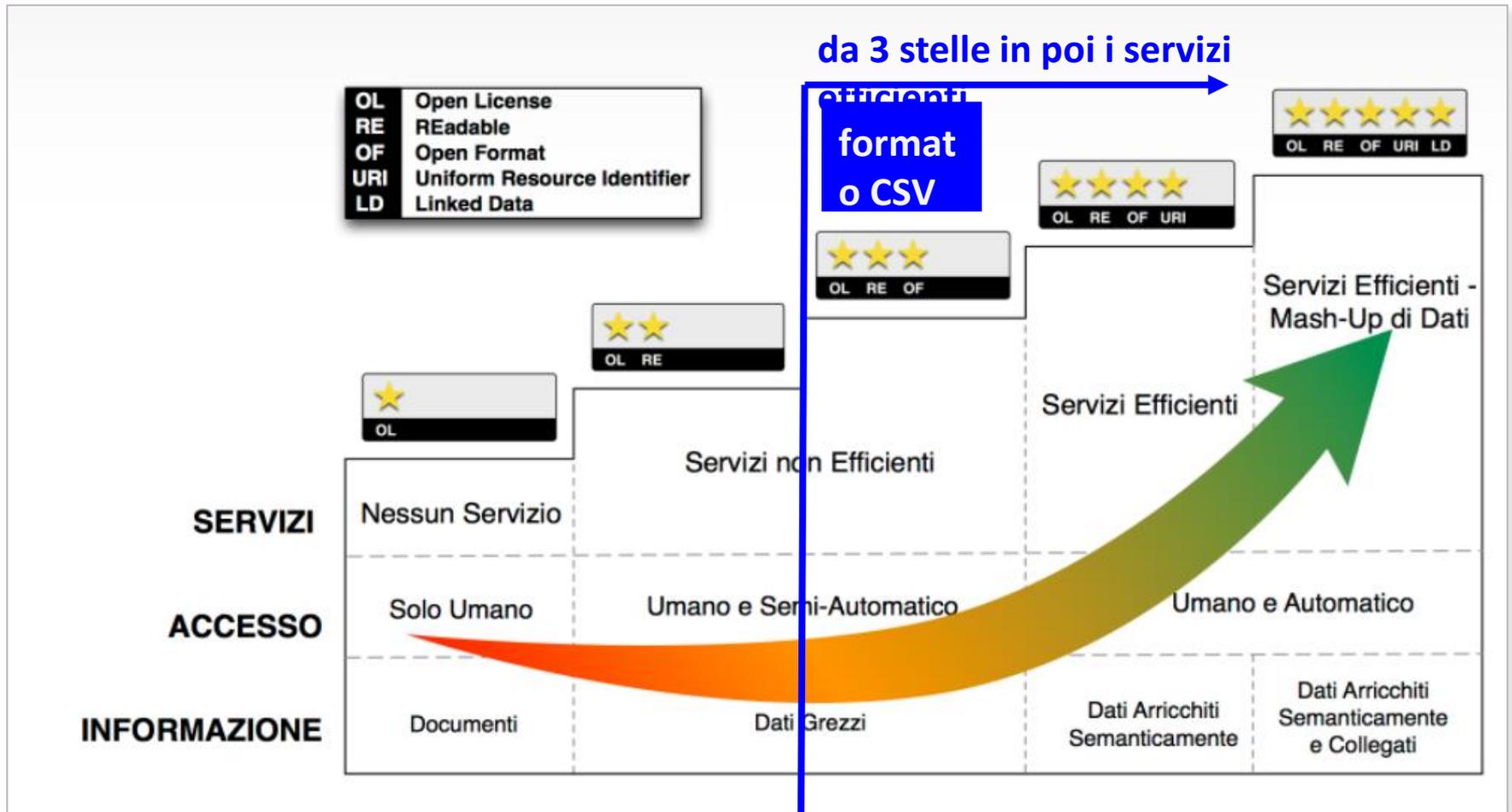
Sono le regole di riferimento per costruire, ottenere e pubblicare dataset in formato aperto sul portale istituzionale dell'ente. Non dobbiamo inventarci nulla, basta seguire le regole stabilite, che di seguito vengono illustrate

Dati, dataset, banche dati, che differenza c'è?



slide di una presentazione del FormezPA di alcuni anni fa

I livelli del modello per i dati aperti



da 1 a 5 stelle i formati open data, più alto è il numero, maggiore è l'interoperabilità tra i dataset (scala di Tim Berners Lee)



- **PDF** = 1 stella (formato non leggibile da macchine e formato)
- **EXCEL** = 2 stelle (formato tabellare proprietario di Microsoft)
- **CSV** = 3 stelle (formato tabellare non proprietario)
- **RDF** = 4 stelle (formato per la codifica, lo scambio e il riutilizzo di metadati strutturati e consente l'interoperabilità semantica tra applicazioni che condividono le informazioni sul Web)
- **LOD** = 5 stelle (sono dati collegati - sul web - che permettono l'interoperabilità con altri dataset con cui condividono i metadati)

`<rdf:RDF`

```
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
xmlns:au="http://description.org/schema/">
```

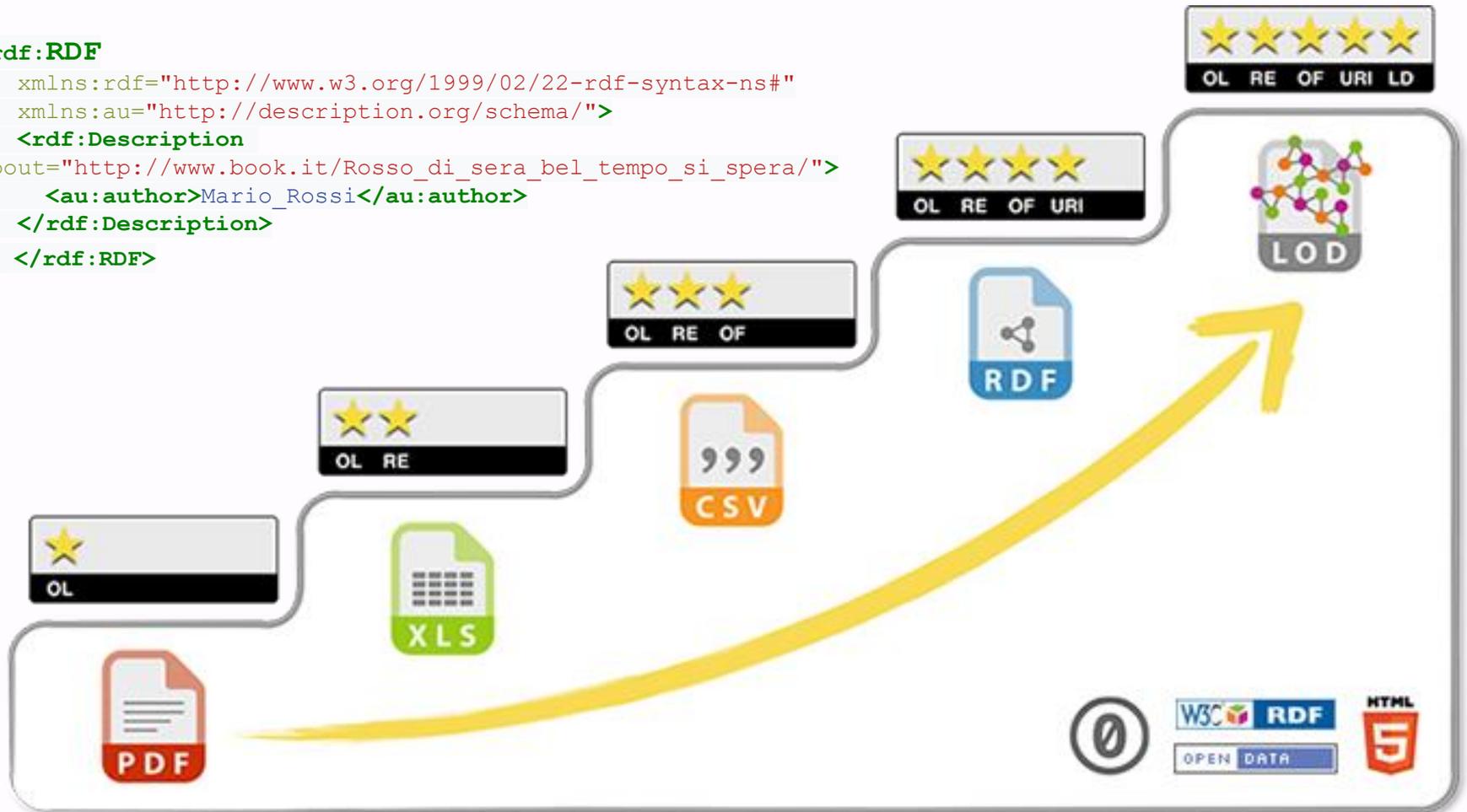
`<rdf:Description`

```
about="http://www.book.it/Rosso_di_sera_bel_tempo_si_spera/">
```

```
<au:author>Mario_Rossi</au:author>
```

`</rdf:Description>`

`</rdf:RDF>`



I dati per considerarsi “aperti”, in base a standard internazionali, devono essere



- **Completi.** I dati devono comprendere tutte le componenti (metadati) che consentano di esportarli, utilizzarli on line e off line, integrarli e aggregarli con altre risorse e diffonderli in rete.
- **Primari.** Le risorse digitali devono essere strutturate in modo tale che i dati siano presentati in maniera sufficientemente granulare, così che possano essere utilizzate dagli utenti per integrarle e aggregarle con altri dati e contenuti in formato digitale
- **Tempestivi.** Gli utenti devono essere messi in condizione di accedere e utilizzare i dati presenti in rete in modo rapido e immediato, massimizzando il valore e l'utilità derivanti da accesso e uso di queste risorse;
- **Accessibili.** I dati devono essere resi disponibili al maggior numero possibile di utenti senza barriere all'utilizzo, quindi preferibilmente attraverso il solo protocollo Hypertext Transfer Protocol (HTTP) e senza il ricorso a piattaforme proprietarie. Devono essere inoltre resi disponibili senza alcuna sottoscrizione di contratto, pagamento, registrazione o richiesta.
- **Leggibili da computer.** Per garantire agli utenti la piena libertà di accesso e soprattutto di utilizzo e integrazione dei contenuti digitali, è necessario che i dati siano processabili in automatico dal computer.
- **In formati non proprietari.** I dati devono essere codificati in formati aperti e pubblici, sui quali non vi siano entità (aziende o organizzazioni) che ne abbiano il controllo esclusivo. Sono preferibili i formati con le codifiche più semplici e maggiormente supportati.
- **Liberi da licenze che ne limitino l'uso.** I dati aperti devono essere caratterizzati da licenze che non ne limitino l'uso, la diffusione o la redistribuzione. Esempi di licenze utilizzate diffusamente a livello mondiale per il riuso dei dati sono le [Creative Commons](#).
- **Riutilizzabili.** Affinché i dati siano effettivamente aperti, gli utenti devono essere messi in condizione di riutilizzarli e integrarli, fino a creare nuove risorse, applicazioni e servizi di pubblica utilità.
- **Ricercabili.** I dati devono essere facilmente identificabili in rete, grazie a cataloghi e archivi facilmente indicizzabili dai motori di ricerca.

formati dei dataset
open data

<u>tipo di dataset, acronimo, descrizione</u>	tipo di dato	estensione del file	stelle del formato secondo la classificazione di Tim Berners Lee
Comma Separated Value (CSV) Formato testuale per l'interscambio di tabelle, le cui righe corrispondono a record e i cui valori delle singole colonne sono separati da una virgola (o punto e virgola)	Dato tabellare	.csv	3
Geographic Markup Language (GML) Formato XML utile allo scambio di dati territoriali di tipo vettoriale	Dato geografico vettoriale	.gml	3
GeoJSON (http://geojson.org) un formato di testo aperto, per la codifica di oggetti geografici e dei correlati attributi non spaziali, scritto in JSON (JavaScript Object Notation).	Dato geografico vettoriale	.geojson, .topojson, .json	3
Keyhole Markup Language (KML) Formato basato su XML creato per gestire dati territoriali in tre dimensioni.	Dato geografico vettoriale	.kml	3
Open Document Format per dati tabellari (ODS) Formato per l'archiviazione e lo scambio di fogli di calcolo	Dato tabellare	.ods	3
Resource Description Framework (RDF) Basato su XML, e' lo strumento base proposto da World Wide Web Consortium (W3C) per la codifica, lo scambio e il riutilizzo di metadati strutturati e consente l'interoperabilità tra applicazioni che si scambiano informazioni sul Web	Dato strutturato	.rdf	4
ESRI Shapefile (SHP) Lo Shapefile ESRI è un popolare formato vettoriale per sistemi informativi geografici. Il dato geografico viene distribuito normalmente attraverso tre o quattro files (se indicato il sistema di riferimento delle coordinate). Il formato è stato rilasciato da ESRI come formato (quasi) aperto	Dato geografico vettoriale	.shp, .shx, .dbf, .prj	3
Tab Separated Value (TSV) Formato testuale per l'interscambio di tabelle, le cui righe corrispondono a record e i cui valori delle singole colonne sono separati da un carattere di tabulazione	Dato tabellare	.tsv	4
Extensible Markup Language (XML) E' un formato di markup, ovvero basato su un meccanismo che consente di definire e controllare il significato degli elementi contenuti in un documento o in un testo attraverso delle etichette (markup)	Dato strutturato	.xml	4
Un documento in formato Adobe PDF	File non strutturati	.pdf	1
Immagine in formato grezzo	File non strutturati	.gif, .jpg, .png	1
Dati in formato tabellare e proprietario, ad esempio la tabella di Microsoft Excel	Dato tabellare	.xls	2
Open data Text. Documento di testo in formato non proprietario. Gestibile con programmi open source quali Open Office e Libre Office	File non strutturati	.odt	3

Il PDF non è un formato dati leggibile da software



PRECISAZIONE

Molte PA pubblicano online (Amministrazione Trasparente) tabelle contenenti dati su formati PDF. Quando si tratta di dati in tabelle derivanti ad es. da formati excel, il formato idoneo alla pubblicazione per il riuso è il formato CSV.

Il d.lgs. 33/2013 (Amministrazione Trasparente) non impone il PDF come formato di pubblicazione. Pubblicare dati in formato aperto CSV permette a chiunque di riusare i dati con facilità e anche i software possono trattare il formato CSV senza problemi, al fine di effettuare analisi, e generare servizi utili alla collettività (app, mappe, infografiche). Può essere pubblicato il formato CSV quando si tratta di dati in formato tabellare.

Il PDF non è un formato strutturato



Regole per il formato CSV, molto usato nelle PA



- non unire le celle! Ogni cella appartiene ad un solo rigo e ad una sola colonna
- non mischiare nello stesso foglio dati e metadati, come ad es. la data di rilascio del dataset, o il nome dell'autore
- la prima riga di un foglio tabellare contiene l'intestazione. L'intestazione della colonna deve indicare chiaramente che unità di misura è utilizzata in quella colonna, per consentire la comprensione
- le rimanenti righe devono contenere solo dati, un dato per riga. Non includere dati aggregati come totali o medie
- i numeri nelle celle dovrebbero essere solo numeri. Non includere virgole, o asterischi dopo i numeri, o altro. Se hai bisogno di aggiungere annotazioni a qualche riga, usa colonne separate
- usa standard noti, es. identifica i paesi con lo standard [iso 3166](#) piuttosto che con i nomi
- non usare colori o spunti colorati per codificare informazioni.
- lascia le celle bianche se un valore non è disponibile
- se vuoi creare una presentazione dei dati molto gradevole da vedere, fallo creando un altro foglio di lavoro, all'interno dello stesso file, creando la relazione con le celle del primo foglio di lavoro

Il CSV è un formato strutturato

	A	B	C	D
1	Transaction Id	AP Invoice	AP Amount	AP Doc Date
2	1007	3001005	50	6-Oct-16
3	1006	3001005	60	6-Oct-16
4	1013	3001008	75	7-Oct-16
5	1014	3001008	166	7-Oct-16
6	1010	3001007	450	17-Oct-16
7	1011	3001007	430	17-Oct-16
8	1012	3001007	1038	17-Oct-16
9	1015	3001008	1300	7-Oct-16
10				
11				

L'importanza dei metadati in un dataset



metadati di questa presentazione del 7.04.2021

titolo_evento:		I dati nella PA
organizzatore_evento:	Formez	
tipo_evento:		webinar
sede_evento:		online
data_evento:		07_04_2021
ora_evento:	09.30	
codice_id_evento:		xxxx
nome_autore_presentazione:		Ciro Spataro
url_evento:	http://eventipa.formez.it/node/294207	
url_presentazione:	http://bit.ly/webinar-idatinellapa-7aprile2021	
licenza_presentazione:	CC_BY_SA	
ente_appartenenza_presentazione:	comune_Palermo	

I metadati sono informazioni importanti che permettono la facile identificazione e ricercabilità (nel tempo) del dataset e permettono di rendere il dataset quanto più interoperabile con altri dataset adeguatamente metadati.

[Si usa il profilo DCAT AP IT dei metadati dei dataset della PA.](#)

I metadati minimi dei dataset open data del portale di Palermo



- Titolo dataset
- Descrizione dataset
- Referente dataset (titolare - es. dirigente)
- Autore (es. funzionario ...)
- Formato originario della risorsa (formato open data: csv, xml, ...)
- Licenza del dataset
- URI permanente dove si trova il dataset (di solito è l'applicativo dove si pubblica che genera automaticamente l'URI)
- Copertura geografica dataset (territorio comunale, regionale, nazionale)
- Copertura temporale dataset (data inizio)
- Copertura temporale dataset (data fine)
- Tag parola chiave del dataset (ambiente - mobilità - istruzione - ...)
- Data di pubblicazione dataset (di solito è l'applicativo dove si pubblica che genera automaticamente la data)
- Frequenza aggiornamento dataset (mensile, bimestrale, semestrale, annuale)
- Note ritenute necessarie di accompagnamento alla pubblicazione del dataset

Licenze per il rilascio dei dati in formato aperto



Ogni dataset open data che l'ente pubblico deve pubblicare online, deve essere accompagnato da una licenza che specifica cosa si può fare con quel dataset



Le licenze Creative Commons sono tra le più diffuse



pubblico dominio, riutilizzabile da chiunque senza condizioni, anche a fini commerciali



condizione di citare autore (BY) e di citare autore e condividere con la stessa licenza (BY SA) anche a fini commerciali



tipologie di licenze "non aperte" in quanto non consentono un riutilizzo a fini commerciali. La PA non può imporre licenze non aperte nei dataset open data che pubblica online

Esempi di grandi banche dati oggi disponibili sul web



PUBBLICO

- Portale europeo dei dati, più di 773.000 dataset <https://www.europeandataportal.eu/data/#/catalogues?locale=it>
- Agid, oltre 27.000 dataset <https://www.dati.gov.it/search/type/dataset>
- Rete Nazionale Dati Territoriali http://www.rndt.gov.it/RNDT/RNDT_v2/GuiAccess/ShowFormRicercaAccessibileDati
- Piattaforma Digitale Nazionale Dati - PDND (ex DAF) <https://pdnd.italia.it>
- Banca Dati Amministrazioni Pubbliche - i dati della finanza pubblica accessibili a tutti <https://openbdap.mef.gov.it/>
- UNICEF open data catalogo <https://data.unicef.org/> (anche <https://open.unicef.org/>)
- <https://www.kaggle.com/datasets>

PRIVATO

- Sciamlab, oltre 39.000 dataset <https://www.sciamlab.com/opendatahub/dataset>
- Hub di dataset resi disponibili da chiunque nel mondo <https://data.world/>
- Idra (Engeneering), oltre 35.000 dataset <https://idra.eng.it/#/viewCatalogues>
- Amazon cloud repository - molti open data di immagini satellitari <https://registry.opendata.aws/>
- Banca Mondiale <https://datacatalog.worldbank.org/>
- World Health Organization <http://apps.who.int/gho/data/node.home>



linee guida open data come “partecipazione” e “interazione”

- **2013** = 7 cittadini spontaneamente le scrivono e inviano via email al Comune che le approva in Giunta;
- **Team open data** (dipendenti, dirigenti);
- **2017** = linee guida vengono [aggiornate con la partecipazione online](#);
- **Riunioni pubbliche del Team open data**: partecipazione di soggetti competenti e interessati sui dati, che danno input e formazione al personale comunale sul processo di produzione/pubblicazione dei dati. [Ogni riunione termina con un report reso pubblico](#)



Archivio Dataset

Cerca Dataset Open Data



Seleziona un'Area Tematica dalla quale estrarre i Dataset oppure utilizza il motore di ricerca contestuale



Ambiente



Amministrazione



Cultura e Turismo



Dati sul territorio



Istruzione



Mobilità e Sicurezza



Opere Pubbliche



Reti



Sanità e Sociale



Sport



Urbanistica



Attività Economiche



Ciclabilità



Bilancio



Elezioni



AREE TEMATICHE

Archivio (1208)

- Ambiente (186)
- Amministrazione (42)
- Attività Economiche (6)
- Bilancio (8)
- Ciclabilità
- Cultura e Turismo (40)
- Dati sul territorio (129)
- Elezioni (574)
- Istruzione (41)
- Mobilità e Sicurezza (129)
- Opere Pubbliche (1)
- Reti
- Sanità e Sociale (17)
- Sport
- Urbanistica (16)

Come aumentare concretamente l'usabilità dei dati della PA



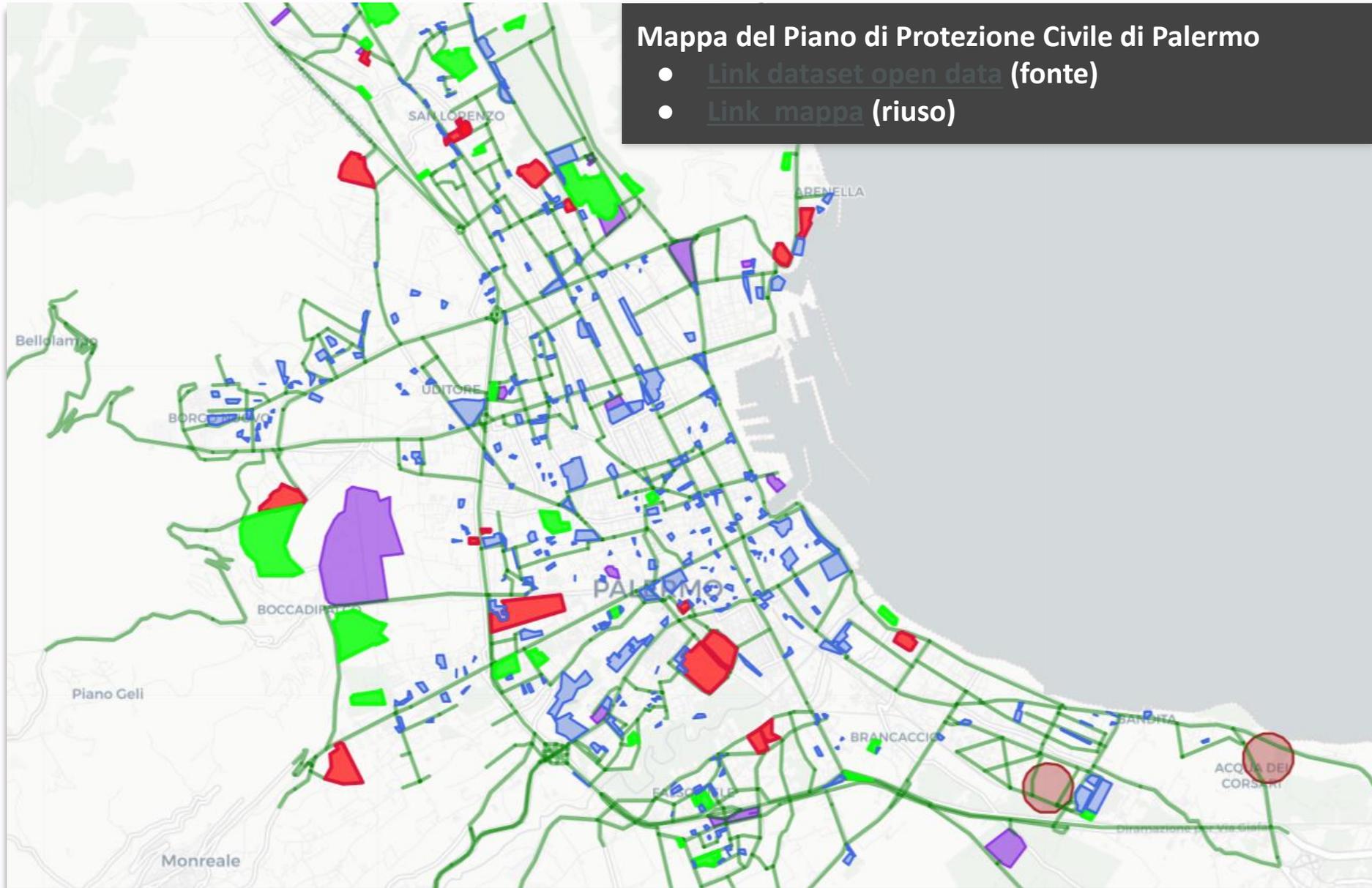
1. adottare “standard” nella fornitura degli applicativi che gestiscono i servizi pubblici delle PA: ad esempio chiedere al fornitore dell'applicativo di adottare i cosiddetti vocabolari controllati già definiti da AgID. Aumenta l'interoperabilità delle banche dati delle PA e stimola il riuso dei dati a fini commerciali;
2. adottare applicativi informatici attingendo dal catalogo del riuso software dell'AGID;
3. nei capitolati dei bandi pubblici per la fornitura di applicativi gestionali, richiedere (come requisito) API secondo quanto stabilito nelle Linee Guida sull'interoperabilità tecnica delle PA.

Riusi: mappe per comprendere efficacemente i fenomeni



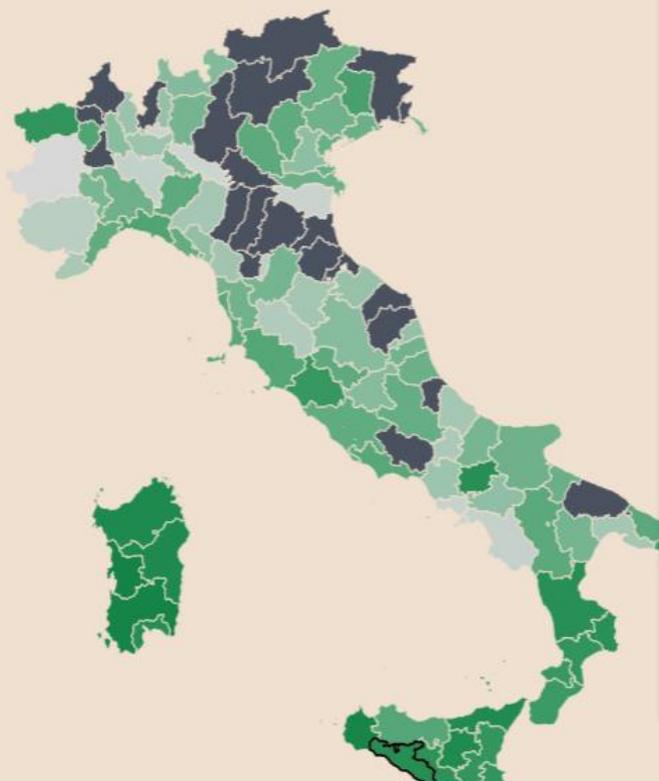
Mappa del Piano di Protezione Civile di Palermo

- [Link dataset open data \(fonte\)](#)
- [Link mappa \(riuso\)](#)



Riusi: infografiche per comprendere la realtà,

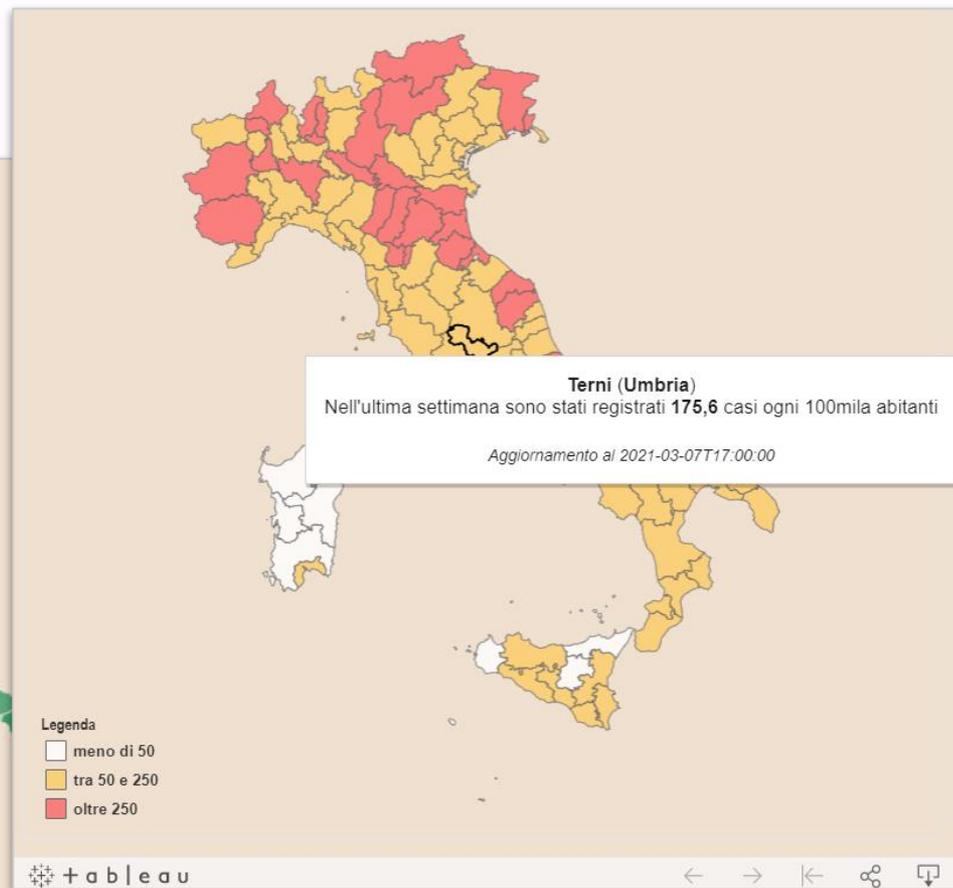
www.infodata.ilsole24ore.com



Agrigento (Sicilia)

Tra il 22 ed il 28 febbraio in questa provincia si sono registrati **43,7** casi di positività al coronavirus ogni 100 abitanti.

Questo significa che le scuole **rimangono aperte**.





#opendatasicilia
Palermo Hub

Home Categoria ▾

126 risorse

Titolo	Descrizione	Immagine	Categoria	Tag ▾
Uffici della Motorizzazione e Officine Autorizzate alla revisione	Uffici della Motorizzazione e Officine Autorizzate alla revisione - Marzo 2017		Mobilità	Mobilità, Motorizzazione, Officine autorizzate, and Uffici Motorizzazione
Sicilia, spiagge accessibili	Strutture balneari che tra percorsi dedicati e servizi necessari per le persone con disabilità motoria (anche temporanea) costituiscono in Sicilia una costellazione di spiagge accessibili in cui il mare è accessibile: da Palermo ad Agrigento, da Trapani a Catania, Siracusa e Pantelleria, il mare di Sicilia è un'esperienza godibile da tutti. (cit)		Cultura	Sicily accessible beaches, Sicilia spiagge accessibili, Sedia Job, Disabili, Vacanze, Mare, Spiagge, Dog club, Dog beach, and Spiaggia di Pluto
Italia - Qualità delle acque di balneazione nel 2016	Oltre l'85% dei siti di balneazione monitorati in tutta Europa nel 2016 soddisfa i requisiti più rigorosi per fregiarsi della qualifica di "eccellente", il che significa che sono per lo più esenti da inquinanti pericolosi per la salute umana e l'ambiente, secondo la relazione annuale sulla qualità delle acque di balneazione pubblicata oggi		Ambiente	Ambiente, Mare, and Spiagge

Portale curato e aggiornato dai membri della comunità civica OpendataSicilia



Riuso civico dei dati: aumentare la cultura sull'importanza dei dati per comprendere meglio la realtà che ci circonda



OpendataSicilia è una “community civica” che usa i dati e diffonde la **cultura** dell'importanza dei dati. E' utile per la PA e per la società. Contribuisce ad aumentare le **competenze digitali** di cittadini/dipendenti e sviluppa **senso civico!**

The graphic features a dark background with a glowing map of Sicily. The map is overlaid with a network of yellow nodes and white connecting lines, representing data connectivity. In the bottom left corner, there is a white hexagonal logo with a cloud icon, binary code (10110010, 01101111, 1, 101), and the text 'opendata sicilia'. Below the logo, there are three yellow horizontal bars containing the text: '#ODS2021', '📅 17 Aprile 2021', and 'IL VALORE DEI DATI'. At the very bottom, in white text, it says 'Partecipa al raduno di Open Data Sicilia'.

ods2021.opendatasicilia.it

il **17 aprile 2021** un raduno online in streaming della comunità OpendataSicilia, con interventi da varie parti d'Italia sul tema degli open data



Dati pubblici come motore dell'economia digitale

Un ingegnere di Apple_Italia telefona e dice: “il [dataset dello stradario comunale di Palermo](#) con licenza **CC BY SA** non lo possiamo riusare a fini commerciali. E’ possibile abbinare la licenza **CC BY?**”

Soluzione



PRESCRIZIONI PER L'UTILIZZO DEL DATASET

Licenza d'uso

CC BY 4.0 IT	▼
-- Seleziona un valore --	
CC BY 3.0 IT	
CC BY 4.0 IT	
CC BY-SA 2.0 IT	
CC BY-SA 4.0	

ANNOTAZIONI

Se una PA vuole incentivare l'economia con i dati pubblici, deve essere flessibile nell'adozione di licenze “aperte” e riusabili a fini commerciali, come in questi casi.



- [Linee guida nazionali sulla valorizzazione del patrimonio informativo pubblico, a cura dell'AGID](#)
- [Profilo metadattazione di riferimento per i dataset della Pubblica Amministrazione - DCAT-AP IT](#)
- [Portale nazionale dei dati aperti](#)
- [Portale open data del comune di Palermo](#)
- [Portale divulgazione e formazione su open data, costantemente aggiornato](#)
- [Le licenze aperte Creative Commons per i dataset open data](#)
- [Comunità civica per la cultura e riuso dei dati, OpendataSicilia](#)

Questa presentazione è rilasciata con licenza aperta **CC BY** ed è disponibile pubblicamente al link **bit.ly/webinar-idatinellapa-7aprile2021**

