



FONDAZIONE  
BRUNO KESSLER

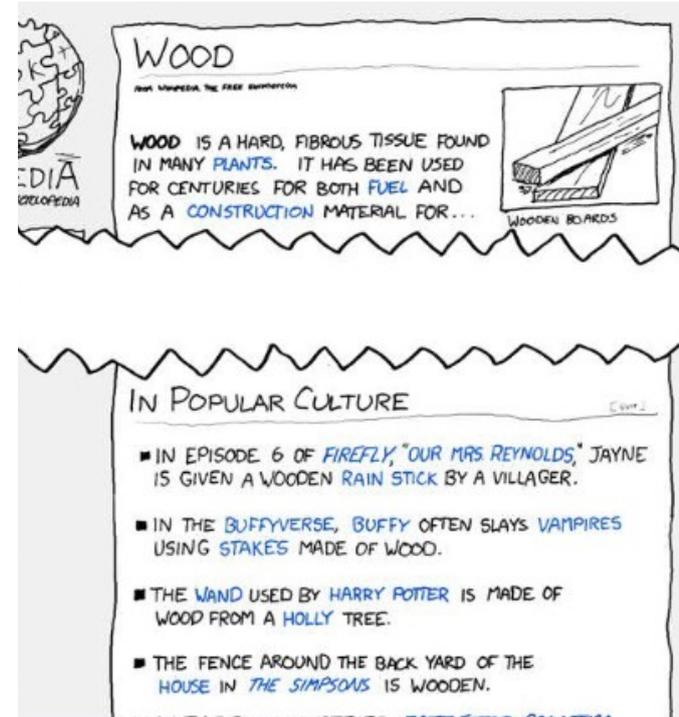
# Open Data e Progetti Collaborativi

Digital Commons Lab  
Fondazione Bruno Kessler

Cristian Consonni  
15 aprile 2014

# Storia

- **1999** Richard Stallman:  
*"Il world wide web ha le potenzialità per svilupparsi in un'enciclopedia universale che copra tutti i campi della conoscenza"*
- **2001** Jimmy Wales e Larry Sanger **creano Wikipedia**. In primavera ha inizio l'edizione in lingua italiana.



Credits: xkcd.com, cc-by-nc-2.5

# I Cinque Pilastri di Wikipedia



1. Wikipedia è un'enciclopedia



2. Wikipedia ha un punto di vista neutrale



3. Wikipedia è libera



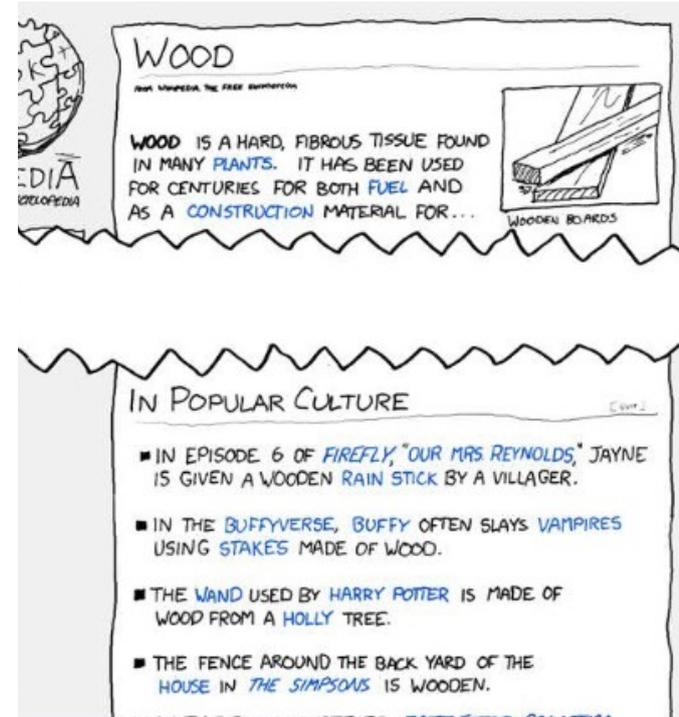
3. Wikipedia ha un codice di condotta



5. Wikipedia non ha regole fisse

# Storia

- **1999** Richard Stallman:  
*"Il world wide web ha le potenzialità per svilupparsi in un'enciclopedia universale che copra tutti i campi della conoscenza"*
- **2001** Jimmy Wales e Larry Sanger **creano Wikipedia**. In primavera ha inizio l'edizione in lingua italiana.



Credits: xkcd.com, cc-by-nc-2.5

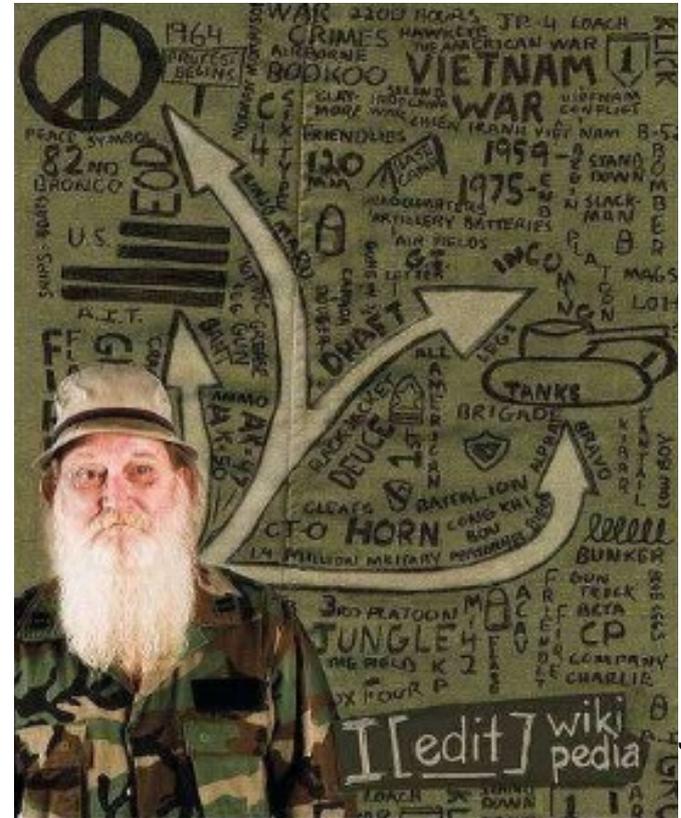
# Chi partecipa? Tutti!

Alcune regole fondamentali e di base: **licenza cc-by-sa-3.0** e nessuna violazione di copyright

Oltre ad esse, **tutte le regole** e le convenzioni sono definite **autonomamente** dalla comunità stessa.

Le **decisioni** vengono prese tra i **partecipanti attivi** al momento cercando la soluzione dal più ampio **consenso** possibile.

La **validazione dei contenuti** è fatta dagli stessi partecipanti.



Credits: Quartermane on flickr.com,  
cc-by-3.0

# I limiti

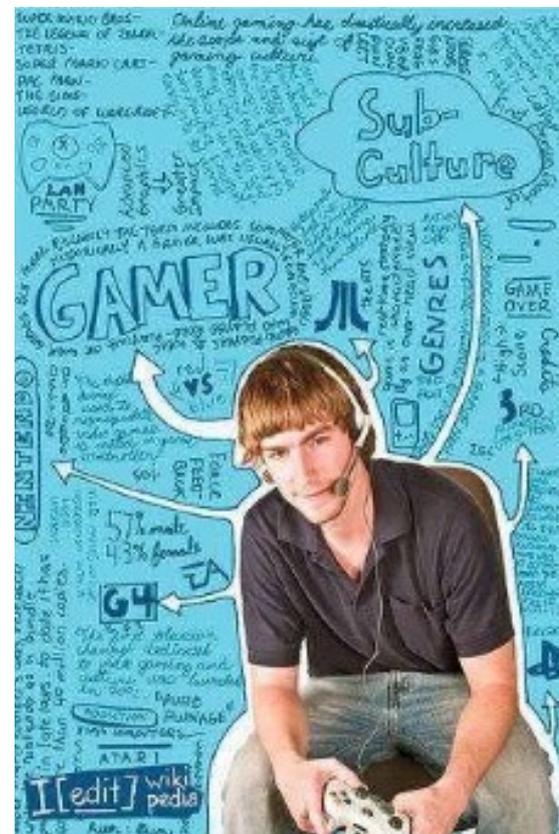
Perché permettere a **chiunque** di **modificare** le voci?

Problemi:

- ♦ **Vandalismo**, c'è chi partecipa in maniera non costruttiva;
- ♦ **Incompletezza**, non c'è una pianificazione editoriale;
- ♦ **Affidabilità**, non c'è un controllo e una validazione sistematica;

La risposta è la **comunità**

La crescita della comunità permette di **monitorare ed eliminare** i vandalismi, **ampliare** l'estensione degli argomenti trattati, sottoporre le voci ad un *controllo continuo*.



Credits: Quartermame on flickr.com, cc-by-3.0

A top-down view of several hands of different skin tones cupped together, holding a large pile of white, uncooked rice grains. The hands are arranged in a circle, with the rice in the center. The background is dark, making the white rice and the skin tones stand out.

Domanda classica:  
se tutti possono fare modifiche,  
cosa garantisce la qualità?

Risposta classica:  
il controllo fra pari  
("appena l'errore è visto,  
può essere corretto")

# Amministratori

Wikipedia:Amministratori

## Chi sono:

- sono utenti (registrati) come gli altri
- Hanno la possibilità di compiere azioni tecniche aggiuntive:
  - ✓ (protezione delle pagine)
  - ✓ (blocco utenti/IP)
- sono eletti e riconfermati annualmente



# Cosa sono i dati?

## Dato [dà-to]

- s.m.
  - 1 - Ciò che viene proposto come oggetto di osservazione e studio SIN elemento: raccogliere tutti i d. a disposizione || d. di fatto, ciò che è dimostrato dai fatti e accertato dall'esperienza
  - 2 - inform. Qualunque informazione rappresentata in modo da poter essere trattata da un calcolatore: base di dati**
- sec. XIII

[http://dizionari.corriere.it/dizionario\\_italiano/D/dato.shtml](http://dizionari.corriere.it/dizionario_italiano/D/dato.shtml)



*«Most people don't just wrangle data for fun. They have a story to tell or a problem to solve.»*

<http://schoolofdata.org>

**Data**



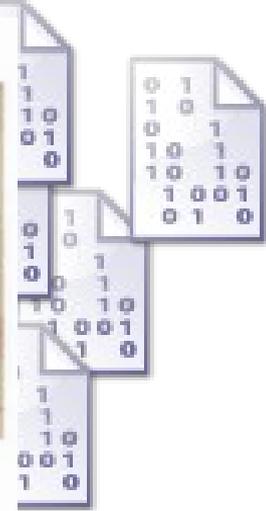
**Information**



**Presentation**



**Knowledge**



EpicGraphic.com

# Open Data (I): Perché i dati?



«Software is what the 21st century is made of.  
What steel was to the economy of the 20th century.  
What steel was to the power of the 20th century  
What steel was to the politics of the 20th century, software is now.  
It's the crucial building block, the component out of which  
everything else is made.  
And when I speak of everything, else I mean, of course,  
freedom.»

Why Political Liberty Depends on Software Freedom More  
Than Ever  
Eben Moglen @ 2011 FOSDEM conference in Brussels on  
Feb 5, 2011  
<http://www.softwarefreedom.org/events/2011/fosdem/moglen-fosdem-keynote.html>

# Open Data (II): Free Software

## ***Le quattro libertà fondamentali:***

- Libertà di eseguire il programma, per qualsiasi scopo (libertà 0).
- Libertà di studiare come funziona il programma e di modificarlo in modo da adattarlo alle proprie necessità (libertà 1). L'accesso al codice sorgente ne è un prerequisito.
- Libertà di ridistribuire copie in modo da aiutare il prossimo (libertà 2).
- Libertà di migliorare il programma e distribuirne pubblicamente i miglioramenti da voi apportati (e le vostre versioni modificate in genere), in modo tale che tutta la comunità ne tragga beneficio (libertà 3). L'accesso al codice sorgente ne è un prerequisito.

Tratto da “*Definizione di Software Libero*”  
<http://www.gnu.org/philosophy/free-sw.it.html>



# Open Data (III): caratteristiche

- Availability of source data;
- Use of a free format;
- No technical restrictions (DRM);
- No other restrictions or limitations;

Tratto da *“Defining Free Cultural Works”*

<http://freedomdefined.org>

- Attribution = Citare la fonte
- Share Alike = Condividi allo stesso modo (se opera derivata)
- Non commercial non è Open Data



«A piece of content or data is open if anyone is free to use, reuse, and redistribute it – subject only, at most, to the requirement to attribute and share-alike.»

*“The Open Definition”*

<http://opendefinition.org>

# La forma dei dati

- Human-readable

## Internet

Da Wikipedia, l'enciclopedia libera.

*Wikidata: Internet (Q75), rete mondiale di reti di computer ad accesso pubblico*

Alias: Nessuno

 *Disambiguazione – Se stai cercando la tipologia di rete di computer, vedi **Internet (Informatica)**.*

**Internet**<sup>[1]</sup> (contrazione della locuzione inglese *interconnected networks*, ovvero "reti interconnesse") è una rete mondiale di reti di computer ad accesso pubblico, attualmente rappresentante il principale mezzo di comunicazione di massa, che offre all'utente una vasta serie di contenuti potenzialmente informativi e servizi.



Dati leggibili facilmente da un umano, per esempio, una pagina di Wikipedia.

- Machine-readable

«Formats that are machine readable are ones which are able to have their data extracted by computer programs easily. [...] Common machine-readable file formats are CSV files.»

<http://schoolofdata.org/handbook/appendix/glossary/#term-machine-readable>

# Acquisizione di dati

La data acquisition consiste nell'ottenere dei dati in formato **machine-readable**

Scaricare dataset da portali open-data (facile)  
Scraping di pagine web (medio)  
Scraping di PDF (difficile)

# Template

Un *template* ('templIt, "modello" in italiano, ma letteralmente "sagoma", "calco") è una pagina speciale creata per essere inclusa in altre pagine tramite il meccanismo dell'inclusione.

- Avvisi
- Infobox (tabelle informative)
- navigazione

`{{S|storia}}` → questa voce sull'argomento storia è un abbozzo (stub)

# Wikidata: cos'è?



# Un nuovo modo di condividere dati strutturati

Un database

- Secondario
- Libero
- Collaborativo
- Multilingua



# WIKIDATA

## Le licenze di Wikidata



(dati)



(pagine di servizio)

# Lo scopo di Wikidata

- Centralizzare i collegamenti fra progetti
- Centralizzare i dati fondamentali
- Creare delle query automatiche
- Rendere i dati riutilizzabili (anche da terzi)

# Quali dati?

- Collegamenti ad altri elementi
- Codici alfanumerici
- Immagini da Wikimedia Commons
- Coordinate temporali (gg/mm/aaaa, hh:mm)
- Coordinate geografiche
- Indirizzi URL
- Quantità (con unità di misura)

# Glossario brevissimo

- Elemento (Qxxx)
- Proprietà (Pxxx)



Rovereto

< ETICHETTA

[edit]

Italian comune

< DESCRIZIONE

[edit]

Also known as:

< ALIAS

[add]

< "liberamente modificabile" (cit.)

### In other languages

Rovereto [edit]

Deutsch

italienische Gemeinde [edit]

Rovereto [edit]

español

municipio italiano [edit]

Rovereto [edit]

français

commune italienne [edit]

#### navigation

- [Main page](#)
- [Community portal](#)
- [Project chat](#)
- [Create a new item](#)
- [Item by title](#)
- [Recent changes](#)
- [Random item](#)
- [Help](#)
- [Donate](#)

#### search

Search

Go

Search

## Pagine di Wikipedia collegate a questo elemento (38 siti)

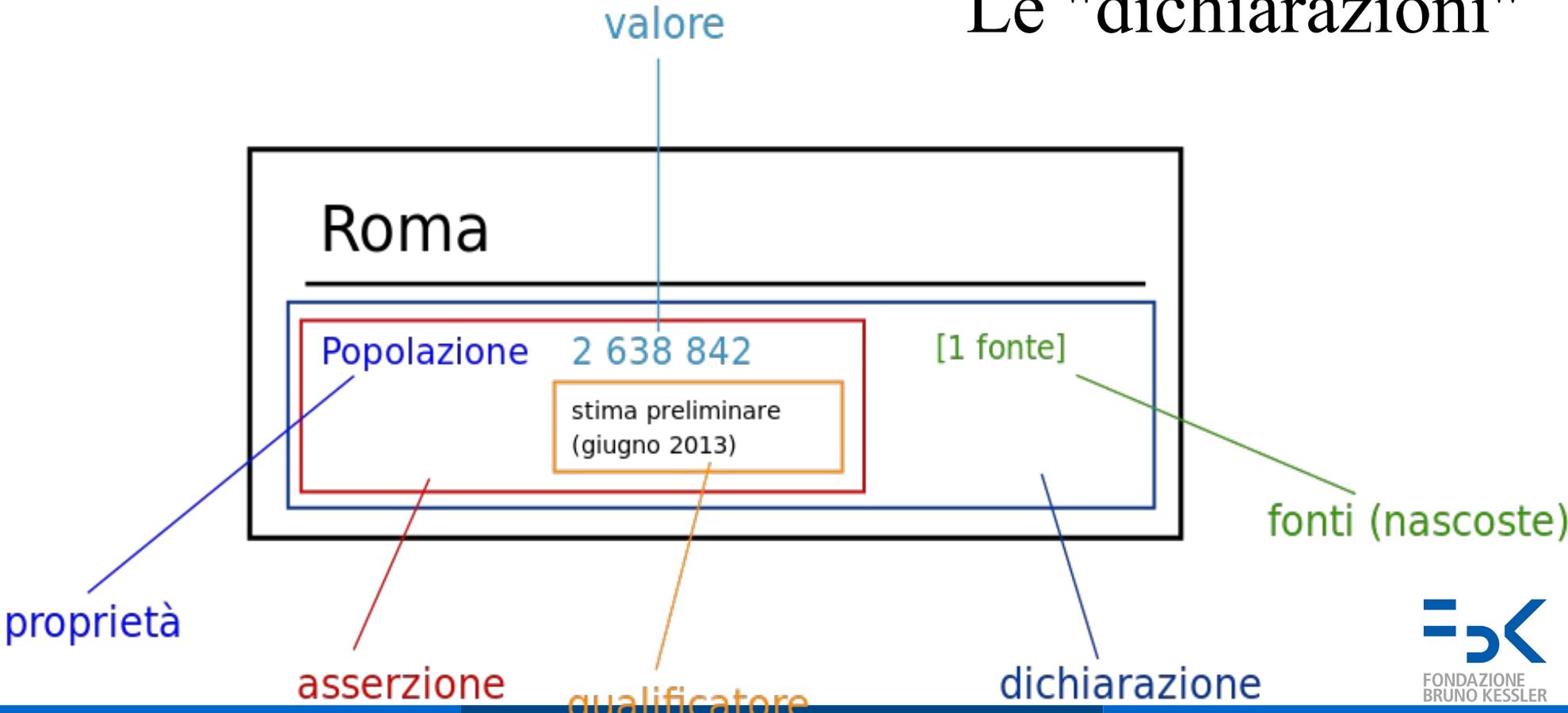
Lingua	↕	Codice	▲	Voce collegata	↕
Boarisch		barwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
català		cawiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
česky		cswiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
Deutsch		dewiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
emiliàn e rumagnòl		emlwiki		<a href="#">Ruveraid</a>	[ <a href="#">modifica</a> ]
English		enwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
Esperanto		eowiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
español		eswiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
eesti		etwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
français		frwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
Bahasa Indonesia		idwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
italiano		itwiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
日本語		jawiki		<a href="#">ロヴェレート</a>	[ <a href="#">modifica</a> ]
қазақша		kkwiki		<a href="#">Роверето</a>	[ <a href="#">modifica</a> ]
Latina		lawiki		<a href="#">Roboretum (Tridentinum-Alta Athesia / Tirolis Meridionalis)</a>	[ <a href="#">modifica</a> ]
lumbaart		lmowiki		<a href="#">Roveredo</a>	[ <a href="#">modifica</a> ]
lietuvių		ltwiki		<a href="#">Roveretas</a>	[ <a href="#">modifica</a> ]
Bahasa Melayu		mswiki		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]

vèneto	vecwiki	<a href="#">Roveredo</a>	[ <a href="#">modifica</a> ]
Tiếng Việt	viwiki	<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
Volapük	vowiki	<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
Winaray	warwiki	<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
中文	zhwiki	<a href="#">罗韦雷托</a>	[ <a href="#">modifica</a> ]
			[ <a href="#">aggiungi</a> ]

## Pagine di Wikivoyage collegate a questo elemento (2 siti)

Lingua	↕	Codice	^	Voce collegata	↕
Deutsch		dewikivoyage		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
italiano		itwikivoyage		<a href="#">Rovereto</a>	[ <a href="#">modifica</a> ]
					[ <a href="#">aggiungi</a> ]

# Le "dichiarazioni"



## Statements

Commons category	<a href="#">Rovereto</a> 	<a href="#">[edit]</a>
	▼ 0 sources	<a href="#">[add source]</a>
		<a href="#">[add]</a>

type of administrative division	<a href="#">comune of Italy</a> <i>q747074</i>	<a href="#">[edit]</a>
	▶ 1 source	
		<a href="#">[add]</a>

is in the administrative unit	<a href="#">Trentino</a> <i>q16289</i>	<a href="#">[edit]</a>
	▶ 1 source	
		<a href="#">[add]</a>

main type (GND)	<a href="#">geographical feature</a> <i>q618123</i>	<a href="#">[edit]</a>
	▶ 1 source	
		<a href="#">[add]</a>

## Dichiarazioni

categoria su Commons	<a href="#">Rovereto</a> 	<a href="#">[modifica]</a>
	▼ 0 fonti	<a href="#">[aggiungi fonte]</a>
		<a href="#">[aggiungi]</a>

tipo di divisione amministrativa	<a href="#">comune italiano</a> <i>q747074</i>	<a href="#">[modifica]</a>
	▶ 1 fonte	<a href="#">[aggiungi]</a>

unità amministrativa in cui è situato	<a href="#">provincia autonoma di Trento</a> <i>q16289</i>	<a href="#">[modifica]</a>
	▶ 1 fonte	<a href="#">[aggiungi]</a>

entità (GND)	<a href="#">oggetto geografico</a> <i>q618123</i>	<a href="#">[modifica]</a>
	▶ 1 fonte	<a href="#">[aggiungi]</a>

ويكيبيديا كومنز

[عدل]  Rovereto

source • ▾

[add source]

[أضف]

نوع التقسيم الإداري

q747074 كومونا

[عدل]

sources ١ ◀

[أضف]

مكانة في التقسيم الإداري

q16289 مقاطعة ترينتو

[عدل]

sources ١ ◀

[أضف]

نوع العنصر GND

q618123 معلم جغرافي

[عدل]

sources ١ ◀

[أضف]



## Утверждения

категория на Викискладе	<a href="#">Rovereto</a> 	<a href="#">[редактировать]</a>
	<span>▼</span> 0 источников	<a href="#">[добавить источник]</a>
<hr/>		<a href="#">[добавить]</a>

тип административной единицы	<a href="#">коммуна Италии</a> <i>q747074</i>	<a href="#">[редактировать]</a>
	<span>►</span> 1 источник	<a href="#">[добавить]</a>
<hr/>		<a href="#">[добавить]</a>

находится в административной единице	<a href="#">Тренто</a> <i>q16289</i>	<a href="#">[редактировать]</a>
	<span>►</span> 1 источник	<a href="#">[добавить]</a>
<hr/>		<a href="#">[добавить]</a>

GND-тип	<a href="#">географический объект</a> <i>q618123</i>	<a href="#">[редактировать]</a>
	<span>►</span> 1 источник	<a href="#">[добавить]</a>
<hr/>		<a href="#">[добавить]</a>

## 서술

공용 분류	Rovereto <a href="#">↗</a>	[편집]
	▼ 0 출처	[출처 추가]
<hr/>		[추가]

행정 구역 유형	코무네 <i>q747074</i>	[편집]
	▶ 1 출처	
<hr/>		[추가]

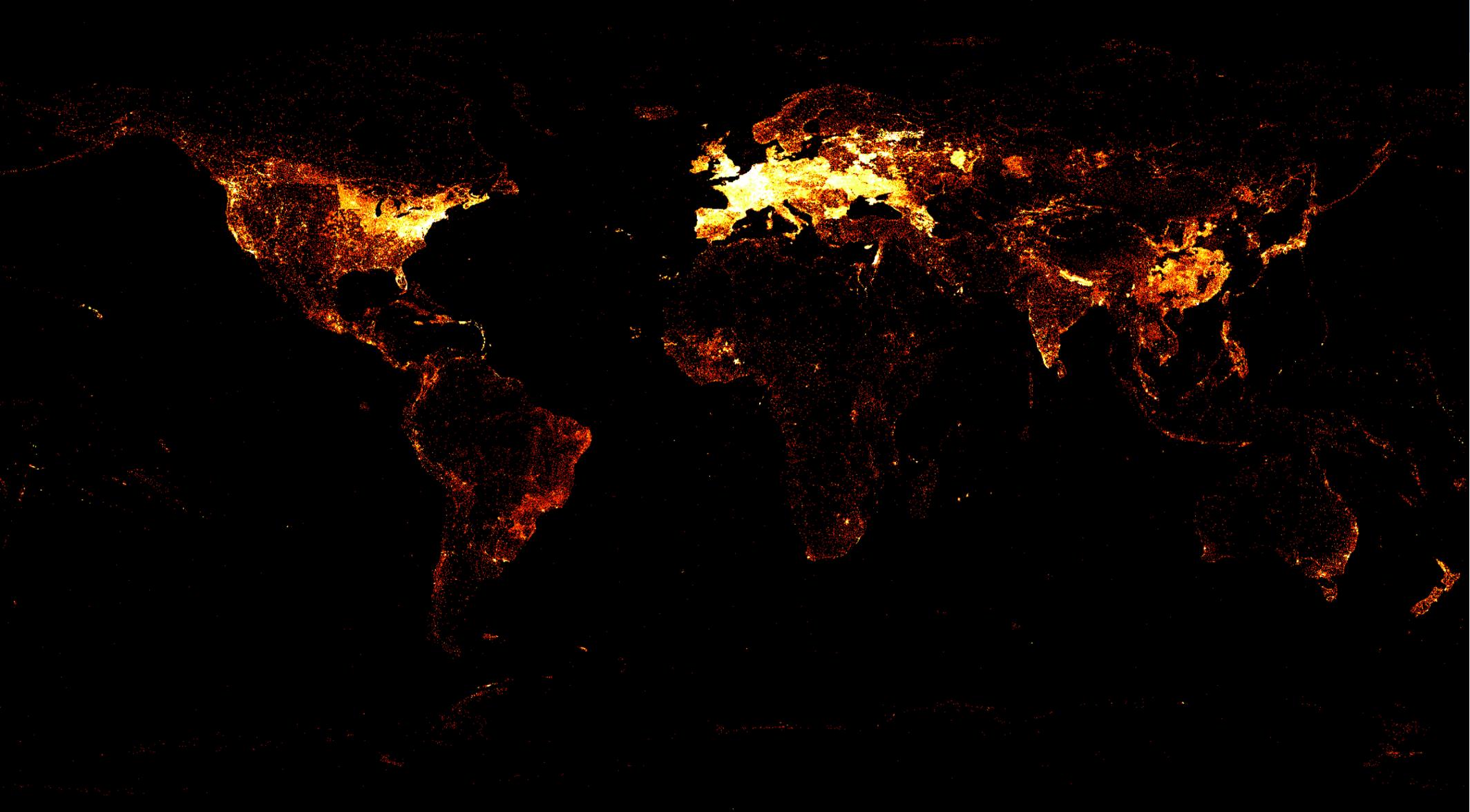
소재 행정 구역	트렌토 현 <i>q16289</i>	[편집]
	▶ 1 출처	
<hr/>		[추가]

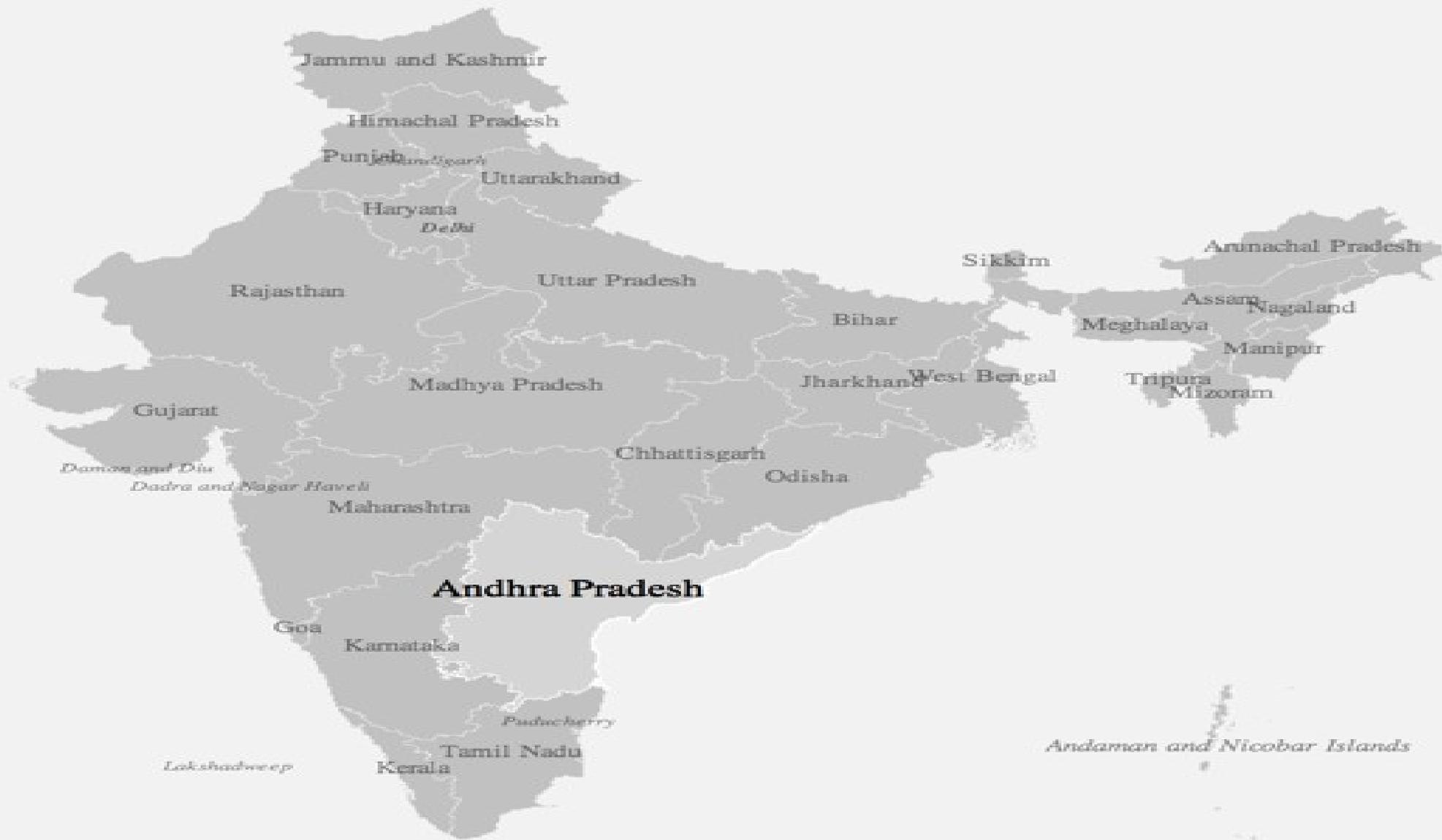
항목 유형	장소 <i>q618123</i>	[편집]
	▶ 1 출처	
<hr/>		[추가]

## Alcune statistiche

- Nata il 30 ottobre 2012
- 15.000.000 di elementi creati
- 20.000.000 di dichiarazioni create
- 240.000.000 di righe di codice rimosse
- 76.500.000 di modifiche effettuate









ג'אמו וקשמיר

הימאצ'אל פראדש

ג'אמיינג'אנג'אם  
אוטראקהאנד

הריאנה  
דלמי

ראג'סטאן

אוטר פראדש

ביהר

סיקים

ארונצ'אל פראדש

מגהלאיה  
מאניפור

טריפורא  
מיזוראם

מערב בנגל

מאדהיה פראדש

צ'האטיסגאר

אוריסה

גוג'אראט

ראסטאן דיין

מאהאראשטרה

מאהאראשטרה

**אנדרה פראדש**

גואה

קארנאטקה

טאמיל נאדו

טאמיל נאדו

קרלה

קשארחהוופ

אי' אנדמן וניקובר



Джамму и Кашмир

Химачал-Прадеш

Пенджаб

Уттаракханд

Харьяна  
*Дели*

Раджастанхан

Уттар-Прадеш

Бихар

Сикким

Аруначал-Прадеш

Ассам

Нагаленд

Мегхалая

Манипур

Мадхья-Прадеш

Джаркханд

Западная Бенгалия

Трипура

Мизорам

Гуджарат

Чхаттисгарх

Орисса

*Пондичерри*

*Пандра и Никобар-Хавели*

Махараштра

**Андхра-Прадеш**

Гов

Карнатака

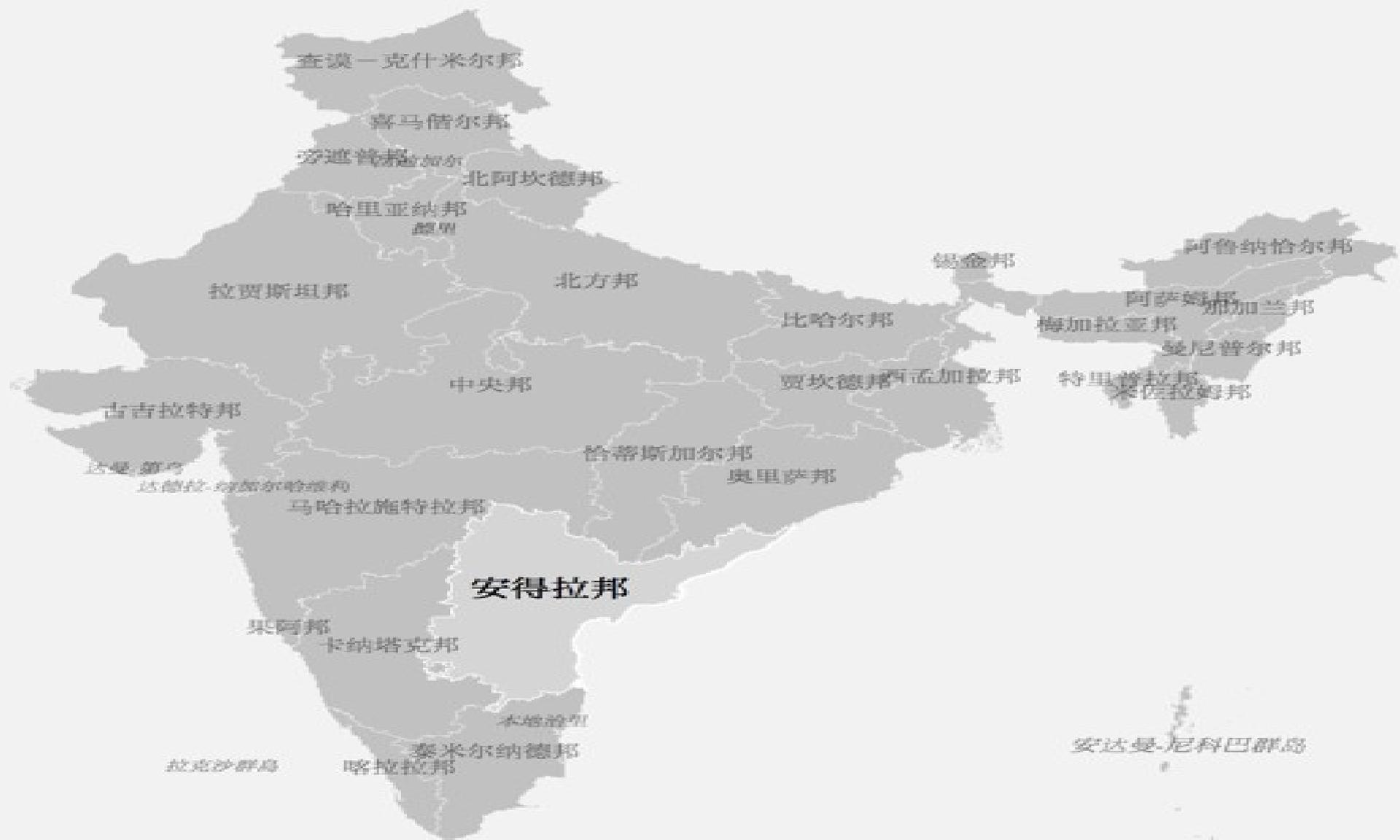
*Пудучерри*

Тамилнад

Керала

*Лакшадвип*

*Андаманские и Никобарские острова*



# Grazie dell'attenzione!

# Domande?

# Download

**Open Data e Pro**

Digital Cc  
Fondazione



**laborativi**

Cristian Consonni

15 aprile 2014

<http://www.slideshare.net/CristianCantoro/open-data-progetti>

# Contatti



- ◆ Cristian Consonni
- ◆ [consonni@fbk.eu](mailto:consonni@fbk.eu)
- ◆ [@CristianCantoro](#)
- ◆ CristianCantoro sui progetti Wikimedia



# Credits & licenza

- ◆ Questa presentazione è rilasciata con licenza CC-BY-SA
- ◆ Tutte le immagini utilizzate sono rilasciate con una licenza libera e sono tratte da Wikimedia Commons (se non diversamente indicato).
- ◆ Le immagini usate nelle slide 3 – 8 derivano da una presentazione tenuta da Lorenzo Losa a OSMit 2013:  
[http://conf.openstreetmap.it/wp-content/uploads/2013/10/osmit\\_20131004\\_free.pdf](http://conf.openstreetmap.it/wp-content/uploads/2013/10/osmit_20131004_free.pdf)
- ◆ I credits delle immagini usate in quella presentazione sono riportati nella slide seguente.

# Crediti

In questa presentazione sono state utilizzate le seguenti immagini:

[https://commons.wikimedia.org/wiki/File:Gum\\_met\\_potlood.jpg](https://commons.wikimedia.org/wiki/File:Gum_met_potlood.jpg)

Autore: User:Michiel1972 su nl.wikipedia.org

GNU Free Documentation Licence 1.2 o successiva

Creative Commons Attribution-ShareAlike 3.0 Unported

[https://commons.wikimedia.org/wiki/File:Granules\\_many\\_hands.jpg](https://commons.wikimedia.org/wiki/File:Granules_many_hands.jpg)

Autore: User:Lobkovs su Wikimedia Commons

Creative Commons Attribution-Share Alike 3.0 Unported

[https://commons.wikimedia.org/wiki/File:Jigsaw\\_puzzle\\_01\\_by\\_Scouten.jpg](https://commons.wikimedia.org/wiki/File:Jigsaw_puzzle_01_by_Scouten.jpg)

Autore: User:Scouten su Wikimedia Commons

Creative Commons Attribution-Share Alike 3.0 Unported

Gli screenshot seguono la licenza del software MediaWiki, la GNU General Public License.

I testi presenti sui progetti Wikimedia seguono la licenza Creative Commons Attribution-ShareAlike 3.0 Unported.

I loghi di Wikipedia e di Wikimedia sono marchi registrati della Wikimedia Foundation, che ne detiene il copyright.

Le vignette di Singloids seguono la licenza Creative Commons Attribution-NonCommercial-NoDerivs 2.5 Italia

Al netto di ciò, la struttura ed i contenuti di questa presentazione sono rilasciati sotto la licenza Creative Commons Attribution 3.0 Unported.