

WIKIPROJECT COVID-19

Alessandra Boccone

Tania Maio

Centro Bibliotecario di Ateneo
Università degli Studi di Salerno

GenOA Week

10 Novembre 2022

Biblioteche per l'open science



1.

Utenti,
bibliotecari e
Wikipedia

Durante la pandemia



Il bibliotecario durante il lockdown





*Lo ha fatto in un momento in cui l'utente si è trovato particolarmente disorientato, trovandosi di fatto per la prima volta in una **biblioteca senza pareti** nella quale ha potuto contare su un supporto mediato dallo schermo di un computer.*

*In questo panorama i **limiti del nostro sistema informativo** sono risultati eclatanti:*

una grande mole di informazioni provenienti dalle fonti più disparate, che il pubblico non ha gli strumenti per poter valutare e vagliare (Vedi i vaccini).

*Inoltre il **digital divide**, visto «nel paradosso dell'attuale situazione in cui tutti potrebbero avere tutta*

l'informazione liberamente disponibile, mentre invece ancora l'accesso all'informazione è un bene competitivo, cioè fa la differenza per quei pochi che riescono davvero ad avere l'informazione», si è confermato un fattore fortemente discriminante.

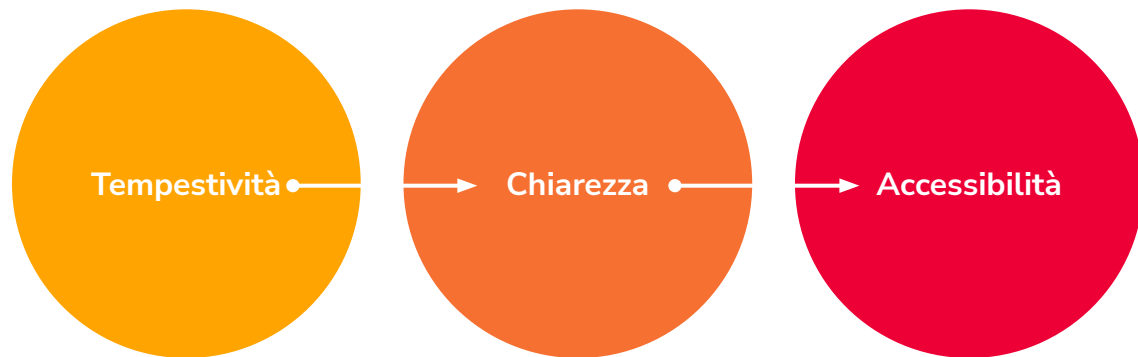
*Everything that
happens in the world
happens on Wikipedia*

Il disorientamento che ha colto l'utente accademico nel reperimento delle risorse online, necessarie per esigenze di studio e di ricerca, è lo stesso che ha riguardato l'utente comune nei giorni della pandemia, quando trovare **notizie aggiornate, attendibili e verificabili** riguardo il virus, la modalità di contagio, le complicazioni, l'estensione dell'epidemia si è rivelato difficile e demoralizzante.

Proprio per questo motivo milioni di utenti si sono rivolti alle pagine di **Wikipedia** per trovare le notizie di cui avevano bisogno

Wikipedia fonte di informazioni sulla salute

- Così come era accaduto per l'epidemia di Ebola in Africa, i lettori di tutto il mondo, con un importante incremento rispetto al passato, hanno fatto moltiplicare il numero di visualizzazioni delle pagine di Wikipedia che trattano di argomenti medici e scientifici. Tale predilezione era già stata analizzata nei mesi precedenti in numerosi studi.
- Tali ricerche hanno svelato il volto degli utenti che scelgono Wikipedia come fonte di informazione scientifica: si tratta di **medici, professionisti delle professioni sanitarie, studenti di medicina**, ma anche ovviamente **persone comuni** che scelgono l'enciclopedia online principalmente per la tempestività



Esplorando i
[dati](#) del
2020....

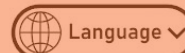
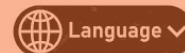
6,950 Articoli di Wikipedia

Scritti sul COVID-19



188 Lingue

in cui articoli sul COVID-19 esistono

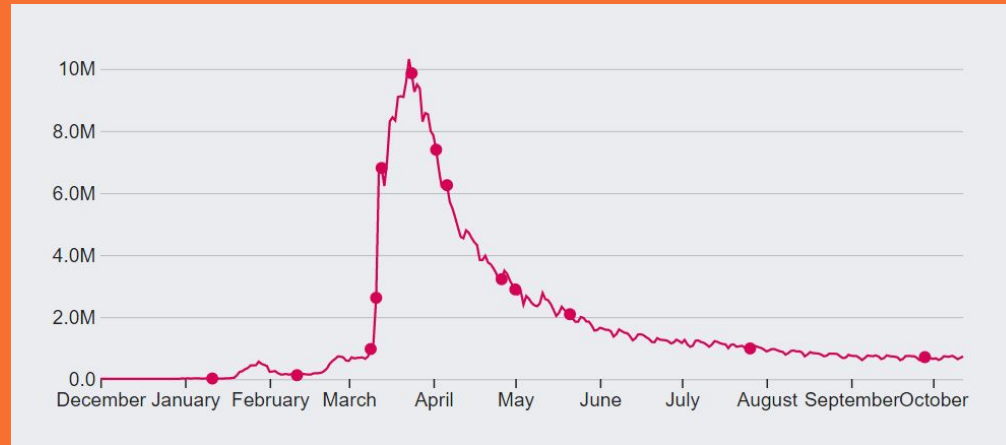


983,395 Edits

Fatti su articoli sul COVID-19 articles, una media di **110**
edits ogni ora



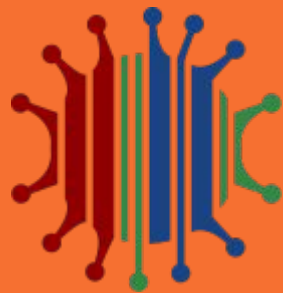
579,190,316



Visualizzazioni degli articoli sul COVID-19 nel mondo

Il Wikiproject Covid-19

15 marzo 2020



Una rete di volontari che condividono l'impegno di rendere Wikipedia una fonte di informazione accurata e attendibile sulla pandemia.



Censire

Tenere traccia degli articoli relativi al Covid



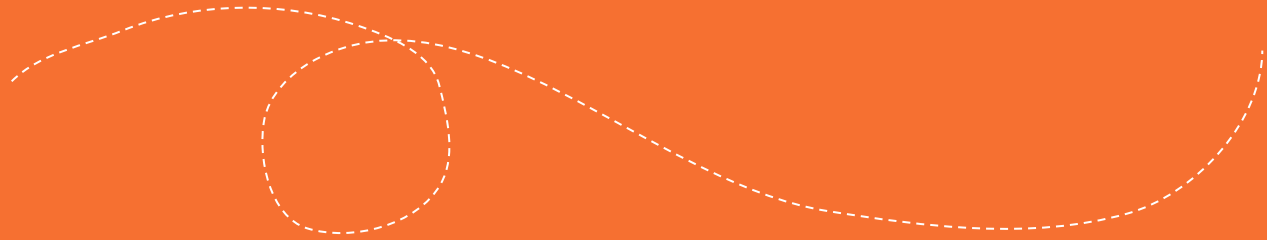
Sorvegliare

Curare una lista di articoli che solo i contributori esperti possono modificare



Verificare

Verificare le fonti esterne da cui trarre i dati



Verifiability, not truth

La verificabilità è una condizione necessaria per l'inserimento di informazioni in Wikipedia

Uso delle fonti nel WikiProject Covid - 19

A quali fonti rivolgersi per scrivere di argomenti scientifici in Wikipedia?



Indice citazionale

articoli scientifici con alto indice citazionale



Condivisi

largamente condivisi sui social media

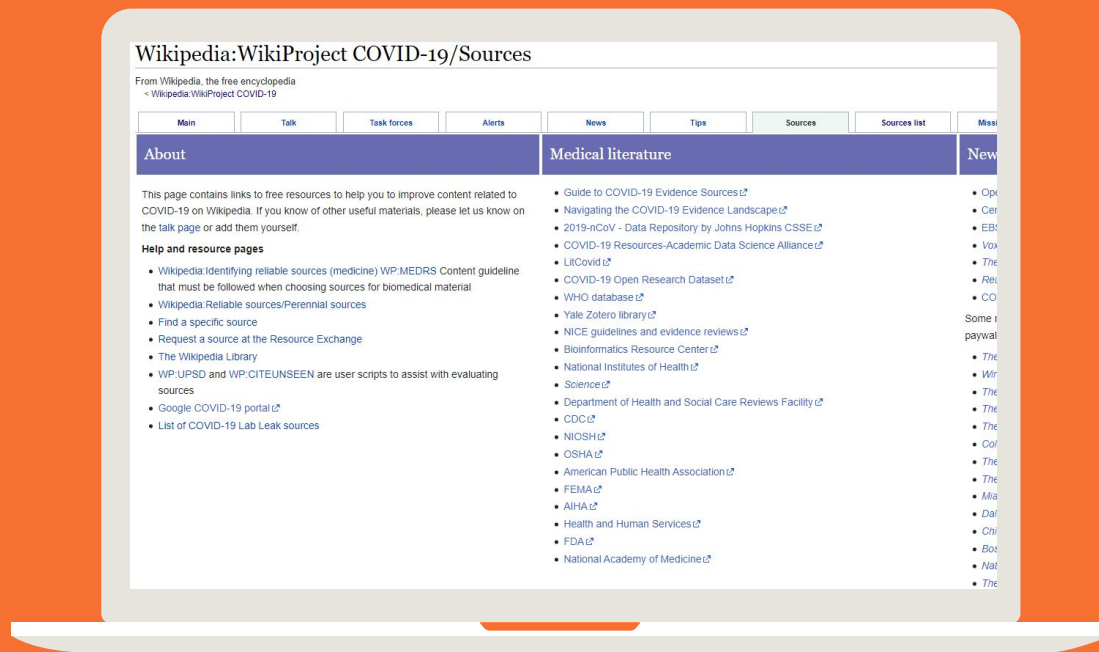


Peer- reviewed

sottoposti a revisione paritaria

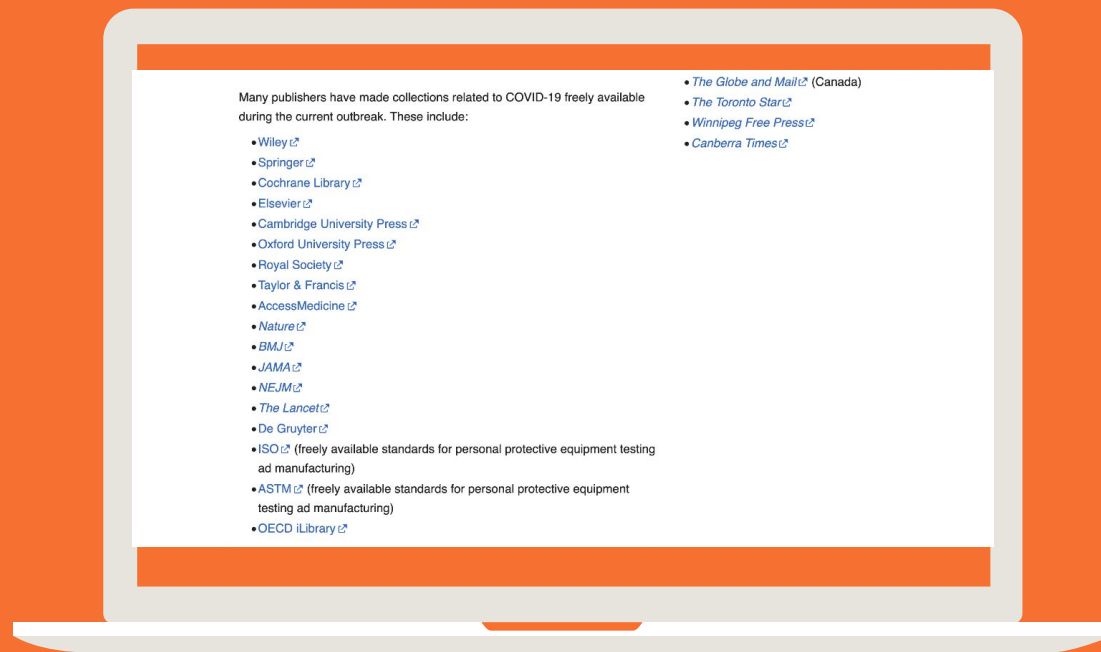
Il ruolo dei bibliotecari

Curano una selezione di fonti autorevoli a cui rivolgersi per ampliare e aggiornare le voci relative al Covid-19



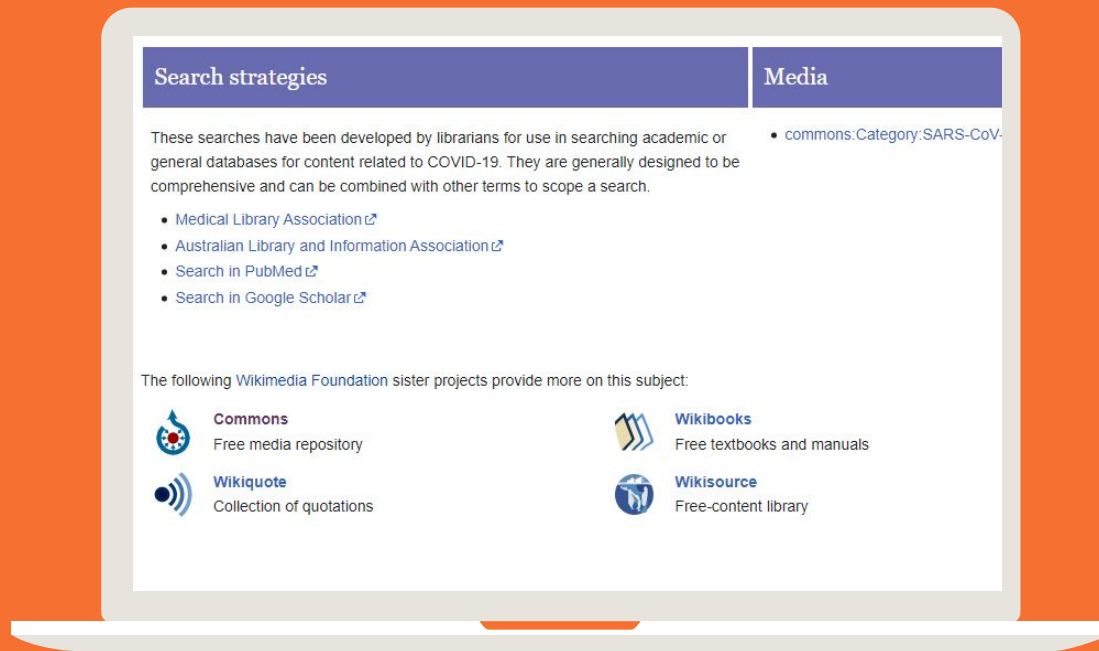
Il ruolo dei bibliotecari

Raccolgono e informano sulle iniziative di solidarietà digitale sul Covid - 19.



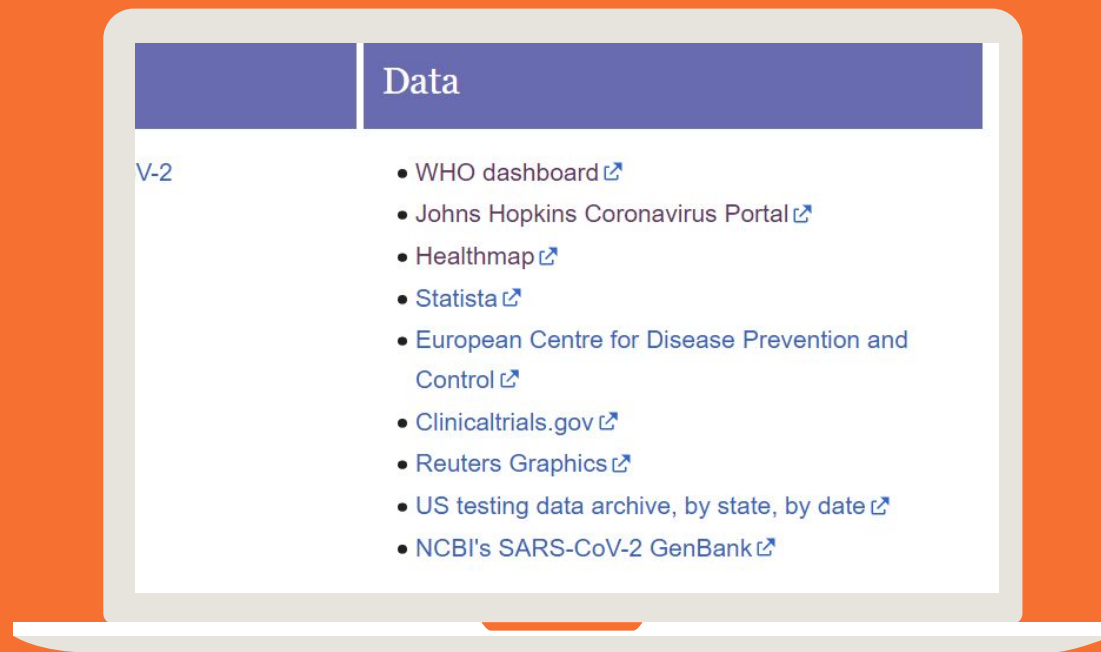
Il ruolo dei bibliotecari

Elaborano e pubblicano strategie di ricerca pronte per essere usate in banche dati scientifiche, come PubMed o indici generalisti come Google Scholar.

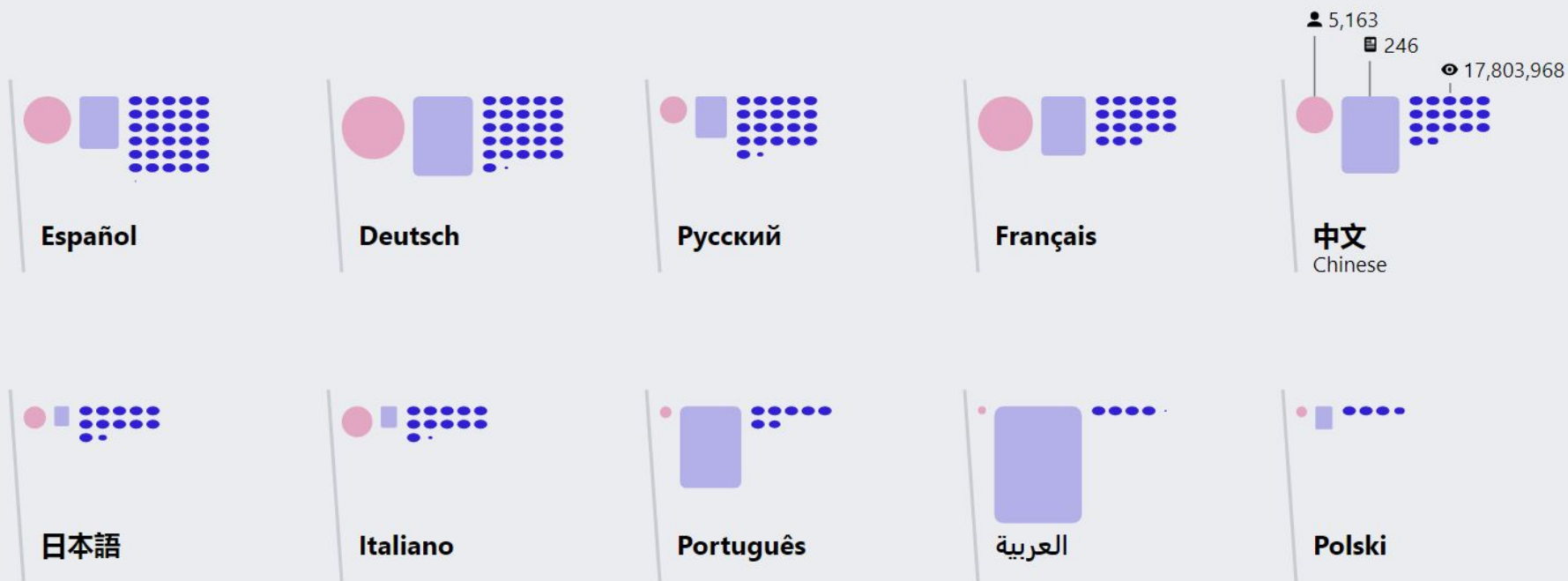


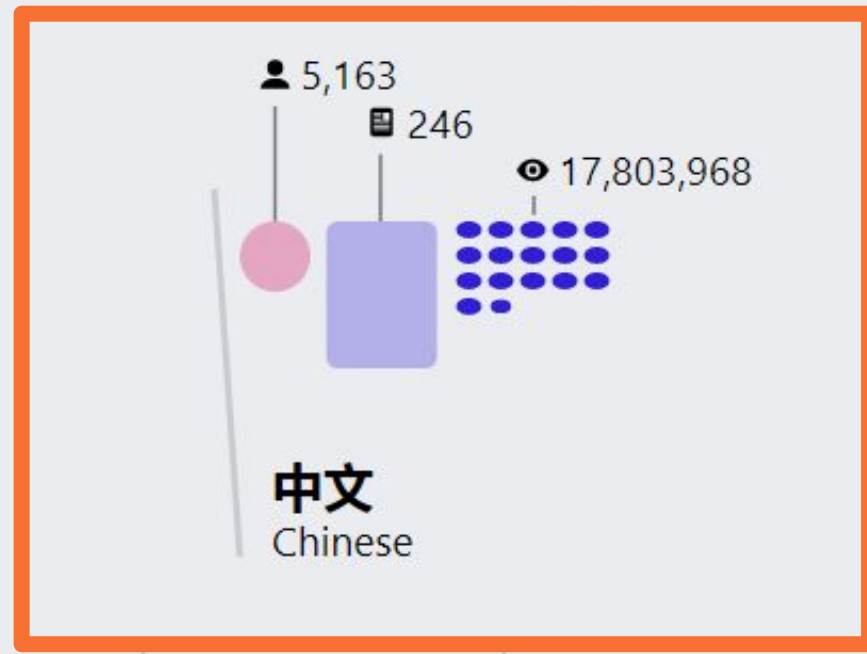
Il ruolo dei bibliotecari

Segnalano i portali da cui trarre, interrogare, visualizzare i dati sulla pandemia in maniera libera.



La Wikimedia Foundation ha elaborato i dati relativi alle 15 versioni linguistiche di Wikipedia con il maggior numero di page view, considerando il numero di contributori al progetto, il numero di voci pubblicate e le visualizzazioni ricevute, per capire quanto il bisogno informativo degli utenti nei diversi paesi sia stato soddisfatto dalla quantità e qualità di articoli pubblicati

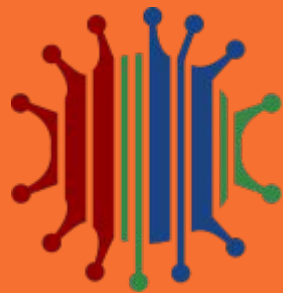




Nonostante la censura operata dal governo cinese su Wikipedia, la WP in lingua cinese mostra un alto numero di editori che desiderano condividere le info sul Covid-19, alto numero di pagine e un forte interesse da parte dei lettori.

Il Wikiproject Covid-19

15 marzo 2020



I pochi articoli presenti nella Wikipedia in Afrikaans hanno ottenuto più di 2 milioni di visualizzazioni, dato che permette di capire il grande interesse nutrito dagli utenti africani rispetto alla quantità minima di informazioni presenti nella loro lingua. Proprio durante le settimane del lockdown numerose iniziative hanno visto coinvolti i bibliotecari africani, tra queste la “African librarians week” (#AfLibWk), una campagna di sensibilizzazione in collaborazione con AfLIA27 e Wikimedia per incitare la comunità bibliotecaria locale a colmare le lacune nei contenuti riguardante il proprio continente in Wikipedia, aggiungendo informazioni e riferimenti mancanti negli articoli relativi. Alla campagna, unita alla più ampia iniziativa internazionale “#1lib1ref”, hanno partecipato più di 800 bibliotecari, molti dei quali hanno scelto di creare o migliorare le pagine relative all’epidemia Covid-19 nel proprio paese.

AFRICAN LIBRARIANS WEEK

"PROMOTING AFRICAN SCHOLARS TO THE WORLD"

24th - 30th May 2020

[#1Lib1Ref](#) [#AflibWk](#)

The African Librarians Week is a new outreach initiative by AfLIA, in collaboration with Wikimedia, to engage library communities to help fill key gaps in African content on Wikipedia by adding content and missing references to articles about Africa on the online encyclopedia.....

[Learn More](#)

AflibWk Dashboard Statistics

19

Voci create

3.94K

Voci modificate

27.8K

Modifiche totali

844

Editori

659K

Parole aggiunte

10K

References Added

40.3M

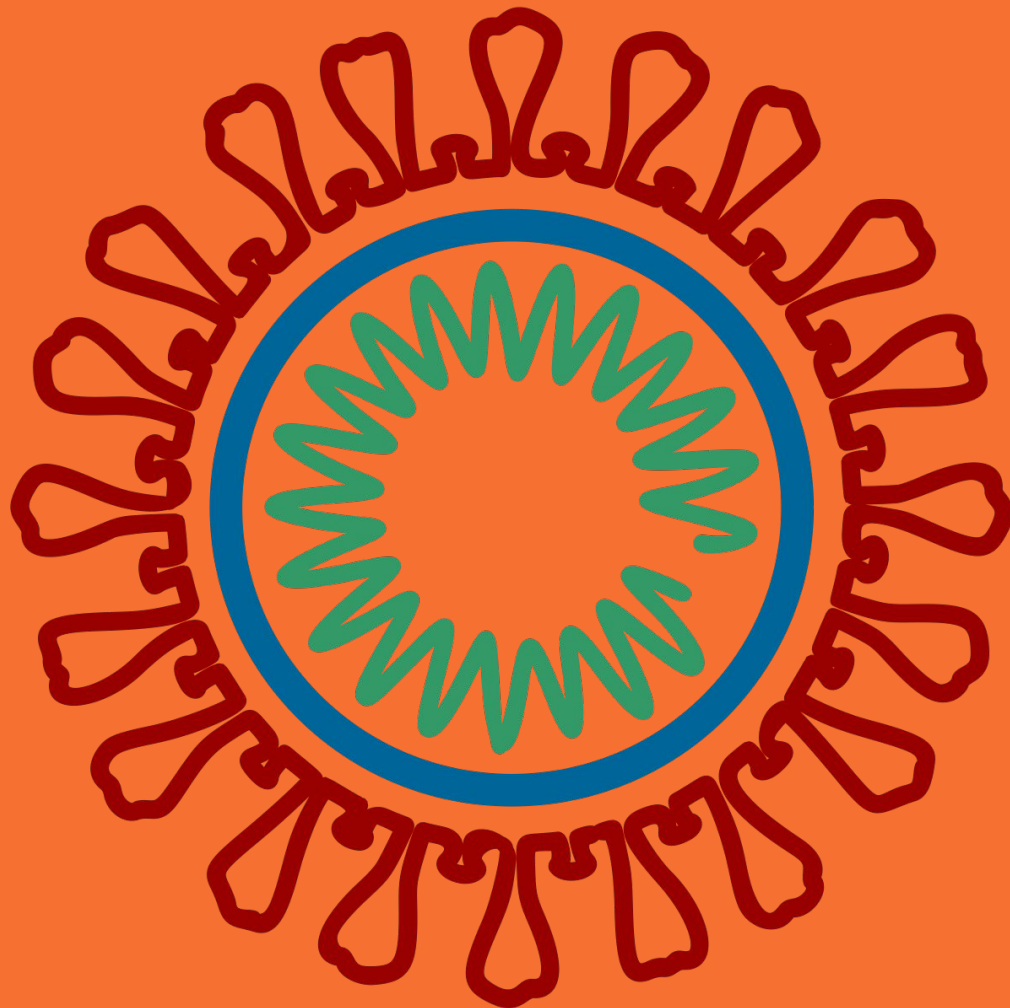
Visite alle voci

81

Caricamenti Commons

3.

Wikidata
Wikiproject
Covid-19:
biblioteche e
bibliotecari





WIKIDATA



È una base di conoscenza:

secondaria

Raccoglie dati strutturati da tutti i progetti Wikimedia e fonti esterne
in formato *proprietà:valore*

libera

Il contenuto di Wikidata è disponibile sotto licenza libera, esportabile usando formati standard e può essere
interconnesso ad altri insiemi di dati aperti
Fornisce supporto, oltre che ai progetti Wikimedia, a molti altri siti e servizi
(vedi [Google Knowledge Graph](#))

multilingue

Inserimento e fruizione sono possibili nelle oltre 300 lingue
usate nei progetti Wikimedia (+ machine readable)

collaborativa

Inserimento e gestione dei dati sono supervisionati dalla comunità
che decide sulle linee guida generali del progetto
e su quelle specifiche sui contenuti.



collaborativa

collegata

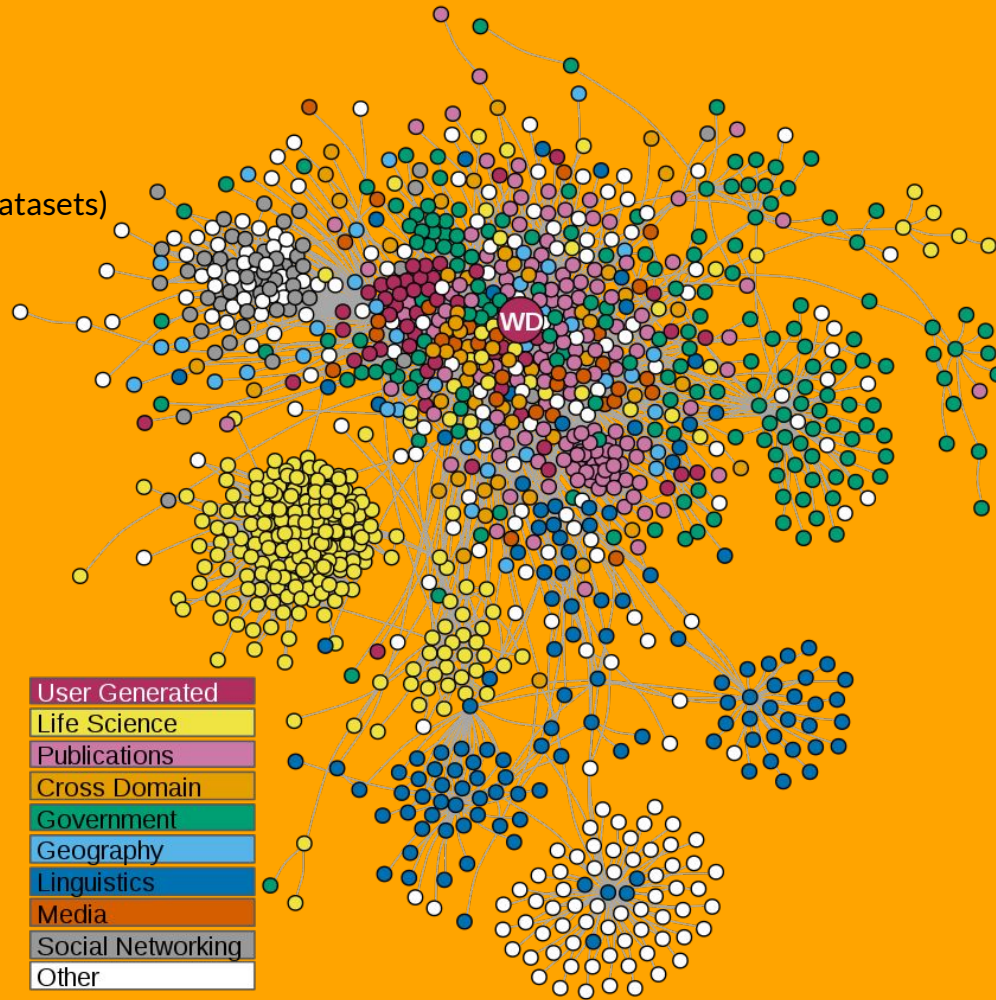
+24.000 utenti attivi

+104.000.000 item

strutturata

LOD Cloud 2020

(fonte <https://lod-cloud.net/datasets>)



label

Marie Curie (Q7186)

item identifier

description

Polish-French physicist and chemist
Maria Skłodowska-Curie | Maria Salomea Skłodowska | Madame Curie
▶ In more languages

aliases

Statements

property

award received

Nobel Prize in Physics

value

point in time	1903
together with	Henry Becquerel Pierre Curie
prize money	35,339 Swedish krona

qualifiers

rank

▼ 2 references

reference URL	http://www.nobelprize.org/nobel_prizes/physics/laureates/1903/index.html
retrieved	3 August 2015
publisher	Nobel Foundation
language of work or name	English
title	The Nobel Prize in Physics 1903 (English)
reference URL	https://www.nobelprize.org/nobel_prizes/about/amounts/

opened references

statement group

Willard Gibbs Award

point in time	1921
---------------	------

▶ 1 reference

collapsed reference

+ add reference

+ add value



Wikidata Wikiproject Covid 19

«a place to collect Wikidata resources related to Covid-19,
SARS-COV-2» (16 marzo 2020)

Obiettivi



Creazione di data model specifici per la standardizzazione dei vari tipi di item relativi all'epidemia



Creazione e l'ampliamento di elementi, dichiarazioni, categorie e template rilevanti per la rappresentazione dell'epidemia e del virus



Strutturazione di query SPARQL che permettono statistiche precise sul coronavirus e sull'andamento della pandemia



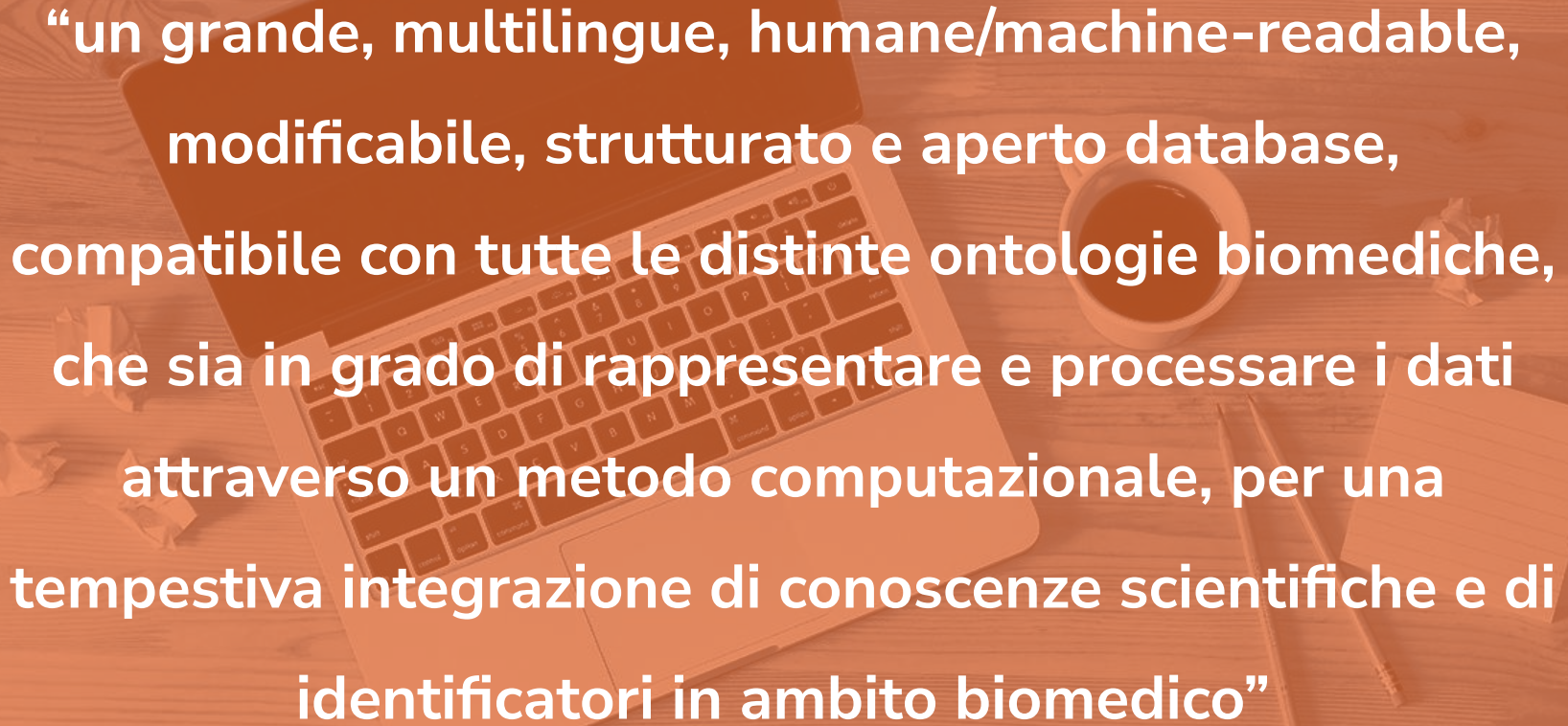
Controllo della qualità delle informazioni su Wikidata della letteratura scientifica ed elaborazione di strategie di accesso alle informazioni



Monitoraggio e aggiornamento delle pagine relative ai focolai a livello internazionale e/o divisi per nazione



Focus list in cui vengono raggruppati tutti gli item relativi al Covid-19

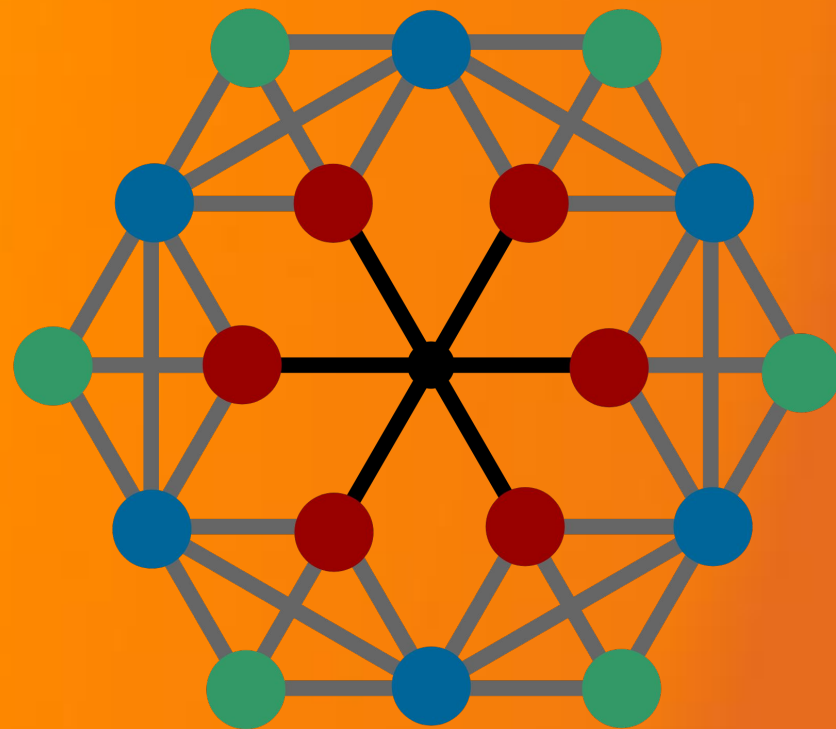
A wooden desk with a laptop, a cup of coffee, and some crumpled paper. The text is overlaid on the image in white.

“un grande, multilingue, humane/machine-readable, modificabile, strutturato e aperto database, compatibile con tutte le distinte ontologie biomediche, che sia in grado di rappresentare e processare i dati attraverso un metodo computazionale, per una tempestiva integrazione di conoscenze scientifiche e di identificatori in ambito biomedico”



Wikidata è una
possibile soluzione:

- sincronizzazione istantanea con tutte le risorse RDF
- disponibilità al riuso e all'analisi dei dati attraverso il suo SPARQL endpoint



WIKIDATA

A protocol for adding knowledge to Wikidata, a case report.

Andra Waagmeester (1) ORCID: [0000-0001-9773-4008](https://orcid.org/0000-0001-9773-4008)

Egon L. Willighagen (2) ORCID: [0000-0001-7542-0286](https://orcid.org/0000-0001-7542-0286)

Andrew I Su (3) ORCID: [0000-0002-9859-4104](https://orcid.org/0000-0002-9859-4104)

Martina Kutmon (2, 4) ORCID: [0000-0002-7699-8191](https://orcid.org/0000-0002-7699-8191)

Jose Emilio Labra Gayo (5), ORCID: [0000-0001-8907-5348](https://orcid.org/0000-0001-8907-5348)

Daniel Fernández-Álvarez (5), ORCID: [0000-0002-8666-7660](https://orcid.org/0000-0002-8666-7660)

Peter J. Schaap (6) ORCID: [0000-0002-4346-6084](https://orcid.org/0000-0002-4346-6084)

Lisa M. Verhagen (7), ORCID: [0000-0002-4130-580X](https://orcid.org/0000-0002-4130-580X)

Jasper J. Koehorst (6) ORCID: [0000-0001-8172-8981](https://orcid.org/0000-0001-8172-8981)

Affiliations:

1. Micelio, Antwerpen, Belgium.
2. Department of Bioinformatics - BiGCaT, NUTRIM, Maastricht University, The Netherlands
3. Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, United States
4. Maastricht Centre for Systems Biology - MaCSBio, Maastricht University, The Netherlands
5. WESO Research group, University of Oviedo, Spain
6. Department of Agrotechnology and Food Sciences, Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Wageningen, the Netherlands
7. Intravacc, PO Box 450, 3720 AL, Bilthoven, The Netherlands

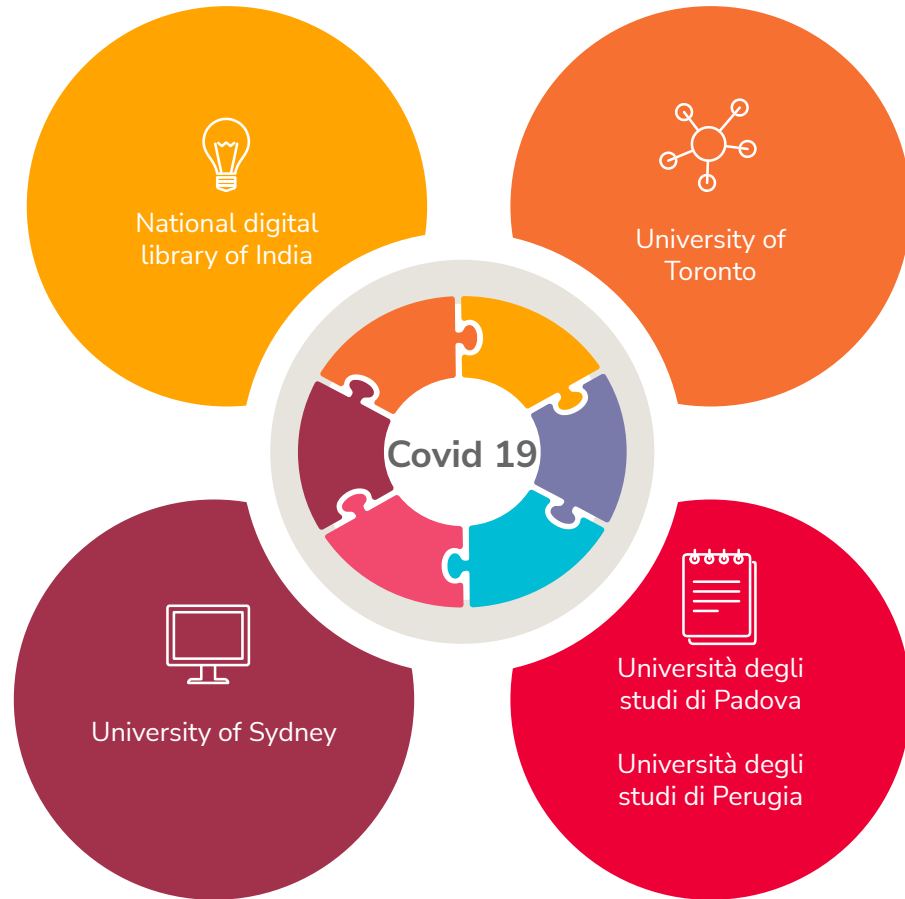


“invita le comunità di bibliotecari e archivisti a preservare i dataset e i vocabolari controllati in set di elementi LD e ad applicare i loro know-how, riferiti alla gestione dati e alla conservazione a lungo termine, ai dataset pubblicati in LD così che le terne di LD diventino immutabili, gli identificatori univoci URIs rimangano permanenti e nomi, titoli, intestazioni per soggetto permangano stabili nel tempo. Con il progredire delle tecnologie del Web semantico, le biblioteche e le istituzioni archivistiche avranno una buona opportunità per arricchire i loro attuali ruoli di gestori dell’universo bibliografico con la qualifica di autorità garanti della conservazione a lungo termine dei dataset in LD relativi al patrimonio culturale”

W3C Library Linked Data (LLDXG)

Library Linked Data Incubator Group Charter (<https://www.w3.org/2005/Incubator/lld/charter>)

- Repository istituzionali
- Metadati
- Letteratura accademica
- Liste di risorse open e full text
- Link a enti, centri di ricerca, aziende e gruppi



Wikidata Wikiproject Covid-19

Linked Data for Production working hour

- Fonti (#1lib1ref)
- Ampliamento/Creazione item
- Per esempio: autori degli articoli scientifici su SARS-Cov2, ospedali specializzati,...



Linked Data for Production

Wikidata Wikiproject Covid-19

Centro Bibliotecario di Ateneo
dell'Università degli Studi di Salerno:

- Soggettazione articoli scientifici
- Fonti (#1lib1ref)
- Ampliamento/creazione item autori
- Periodici scientifici
- Focus list



Wikidata tool

Mix'n'match



QuickStatements



VIAF gadget



Alllabels



Author
disambiguator

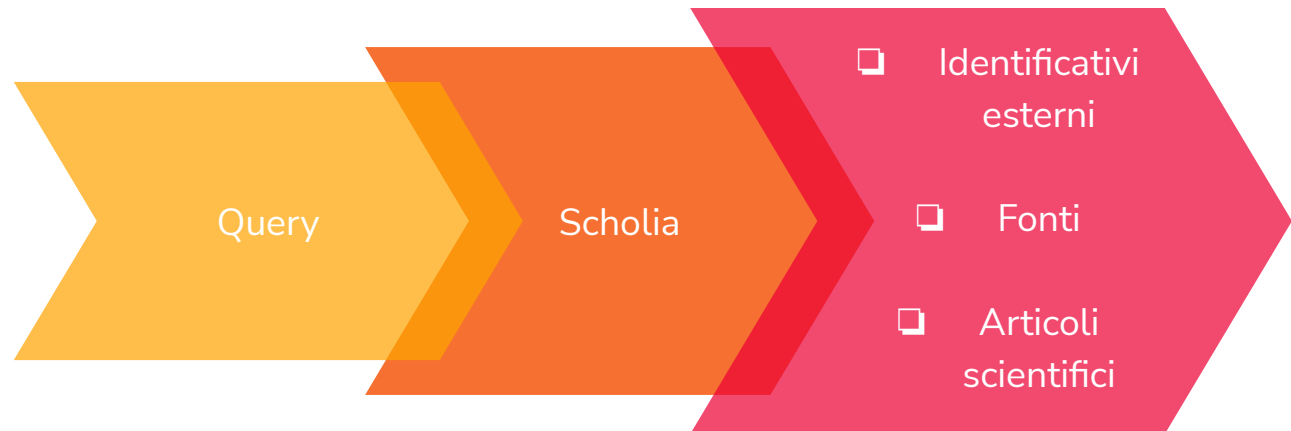


Useful Wikidata
Quick preset



SourceMD

Quali autori?



topic / Q82069695

Improve data

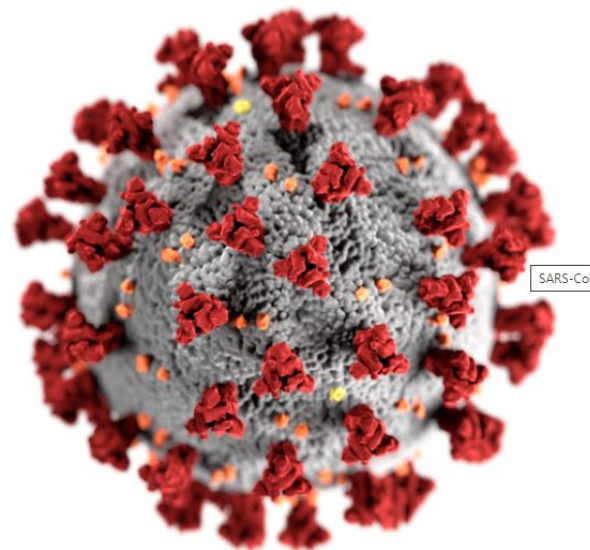
SARS-CoV-2 (Q82069695)

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a strain of coronavirus that causes COVID-19, the respiratory illness responsible for the ongoing COVID-19 pandemic. The virus previously had a provisional name, 2019 novel coronavirus (2019-nCoV), and has also been called human coronavirus 2019. First identified in the city of Wuhan, Hubei, China, the World Health Organization declared the outbreak a Public Health Emergency of International Concern on 30 January 2020, and a pandemic on 11 March 2020. SARS-CoV-2 is a positive-sense single-stranded RNA virus that is contagious in humans. [\(Read more on English Wikipedia\)](#)

 Reasonator  Squid

Table of Contents

- [The topic in context](#)
- [Recently published works on the topic](#)
- [Publications per year](#)
- [Earliest published works on the topic](#)
- [Authors](#)
 - [Authors publishing about the topic](#)
 - [Author score](#)
 - [Co-author graph](#)
 - [Awards received by authors who published on the topic](#)
- [Topics](#)
 - [Co-occurring topics](#)
 - [Co-occurring topics graph](#)
 - [Co-occurring topics map](#)
- [Uses](#)
- [Venues and series publishing works about the topic](#)
- [Citations](#)
 - [Works cited from works on the topic](#)
 - [Authors cited from works on the topic](#)
 - [Map of organizations associated with works about the topic](#)
 - [Citation graph of works about the topic, aggregated by country](#)



Codogno (Q42298)...

comune italiano modifica

Enter a schema to check against e.g. E10

Visita la pagina principale [alt-shift-z]

► **Recom:** Proprietà assenti più rilevanti

► Proprietà assenti più rilevanti: ID

▼ In altre lingue

Configura

Lingua	Etichetta
italiano	Codogno

Tutte le lingue inserite

santo patrono

Biagio di Sebaste ...

modifica

► 1 riferimento

+ aggiungi valore

descritto nella fonte

Meyers Konversations-Lexikon (1888–1889) ...

modifica

dichiarazione è argomento di Q112772611 ...

▼ 0 riferimenti

+ aggiungi riferimento

+ aggiungi valore

sulla focus list di un progetto
Wikimedia

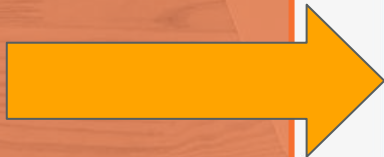
WikiProject COVID-19 ...

modifica

▼ 0 riferimenti

+ aggiungi riferimento

+ aggiungi valore



primary

Articles with a COVID-19 topic in the title but not as main subject [\[edit \]](#)

- In this section, queries are collected that search for publications with the strings "COVID-19", "COVID-19 pandemic", "SARS-CoV-2" in their title but without a corresponding main subject (P921) statement about COVID-19 pandemic (Q81068910) [\[en \]](#), COVID-19 (Q84263196) [\[en \]](#) or SARS-CoV-2 (Q82069695) [\[en \]](#), respectively.
- These queries provide their results in *QuickStatements* format, so as to facilitate running *QuickStatements* batches for topic tagging for these terms.

Notes of caution

- To see whether anyone else is currently working on this, check [recent QuickStatements batches](#) [\[en \]](#) or the relevant [RecentChangesLinked](#) pages ([COVID-19 pandemic](#) [\[en \]](#) / [COVID-19](#) [\[en \]](#) / [SARS-COV-2](#) [\[en \]](#)).
- Before sending the query output to *QuickStatements*, please check whether everything is OK, then uncomment the `?item ?title` line and rerun the query to get the results in *QuickStatements* format.
- SPARQL query results may be affected by [server lag](#) [\[en \]](#).
- When adapting these queries for any other topics, please pay attention not just to the search string but also to
 - the `ReplaceWithYourTargetID` and
 - `-haswbstatement:P921=` parts (which should both use the same identifier, e.g. [Q82069695](#) for SARS-CoV-2) and of course to
 - strings and terms that can have different meanings in different circumstances.

COVID-19 pandemic [\[edit \]](#)

- Properties used: [title \(P1476\)](#) [\[en \]](#) [\[en \]](#) [\[en \]](#)

```
SELECT
DISTINCT
?item ?title
(
  REPLACE(STR(?item), ".*Q", "Q") AS ?qid
  ("P921" AS ?property)
  ("Q81068910" AS ?ReplaceWithYourTargetID)
  ("S887" AS ?heuristic)
  ("Q69652283" AS ?deduced)
)

WHERE {
  hint:Query hint:optimizer "None".

{
  SERVICE wikibase:mwapi {
    bd:serviceParam wikibase:api "Search";
                    wikibase:endpoint "www.wikidata.org";
                    mwapi:srsearch "\"covid-19\" pandemic haswbstatement:P31=Q13442814 -haswbstatement:P921=Q81068910".
    ?page_title wikibase:apiOutput mwapi:title.
  }
}
BIND(IRI(CONCAT(STR(wd:), ?page_title)) AS ?item)

?item wdt:P1476 ?title.
FILTER CONTAINS(LCASE(?title), "covid-19").
FILTER CONTAINS(LCASE(?title), "pandemic").
}
```

Impatto del progetto

Modifiche effettuate nell'ambito del Wikidata Wikiproject Covid-19 nel periodo compreso tra il 18 giugno e il 31 luglio 2020 da

User:Taniamaio



User:Alessandra Boccone

Proprietà	Numero di <i>edit</i>
Argomento principale (P921)	6.446
Autore (P50)	788
Identificativi esterni (vari)	396
Qualifier addition (P12)	1.157
Reference addition (P11)	6.424
Sulla focus list di un progetto Wikimedia (P5008)	15.682
Item/Property creation (P18)	24
Item/Property change (P19)	953
Claim deletion (P15)	864
Totale	32.734



“[...] lo scambio di informazioni tra istituzioni affini richiede una prima fase di mappatura dispendiosa. In secondo luogo l’eredità di soluzioni tecnologiche – più o meno aperte e più o meno manutenibili – difficilmente permette di dialogare con fonti di dati esterne senza ulteriori costi e/o cambiamenti radicali. Non ultimo, il dialogo tra fonti richiede la revisione e l’aggiornamento delle basi di conoscenza, previa una sofferta fase di data cleansing. Il costo per aggiornare le descrizioni catalografiche è infatti dispendioso, in termini di risorse umane, expertise e tempo. Ciò comporta una drastica selezione delle informazioni da condividere con enti esterni. La conseguenza immediata è un netto ridimensionamento delle domande di ricerca e delle aspettative a cui l’integrazione potrà dare risposta”

Marilena Daquino,

Archivi fotografici per la storia dell’arte e semantic web: problemi, risorse e linee di ricerca,

«JLIS.it», 10 (2019), n. 2, p. 37-47: p. 39-40, <https://www.jlis.it/article/view/12533/11362>.

Criticità



Correttezza
e
verificabilità
dei dati



È
(anche)
compito
del
bibliotecario!



Grazie!



Domande?



Scriveteci!



Alessandra Boccone
aboccone@unisa.it



Tania Maio
tmaio@unisa.it