




Fighting COVID-19 with Artificial Intelligence

Slava Tykhonov,
DANS-KNAW R&D, Royal Netherlands
Academy of Arts and Sciences



Spanish National Research Council (CSIC) webinar / 20.04.2021

About me: DANS-KNAW projects (2016-2021)

- CLARIAH+ (ongoing)
- EOSC Synergy (ongoing)
- SSHOC Dataverse (ongoing)
- CESSDA Dataverse Europe 2018
- Time Machine Europe Supervisor at DANS-KNAW
- PARTHENOS Horizon 2020
- CESSDA PID (PersistentIdentifiers) Horizon 2020
- CLARIAH
- RDA (Research Data Alliance) PITTS Horizon 2020
- CESSDA SaW H2020-EU.1.4.1.1 Horizon 2020



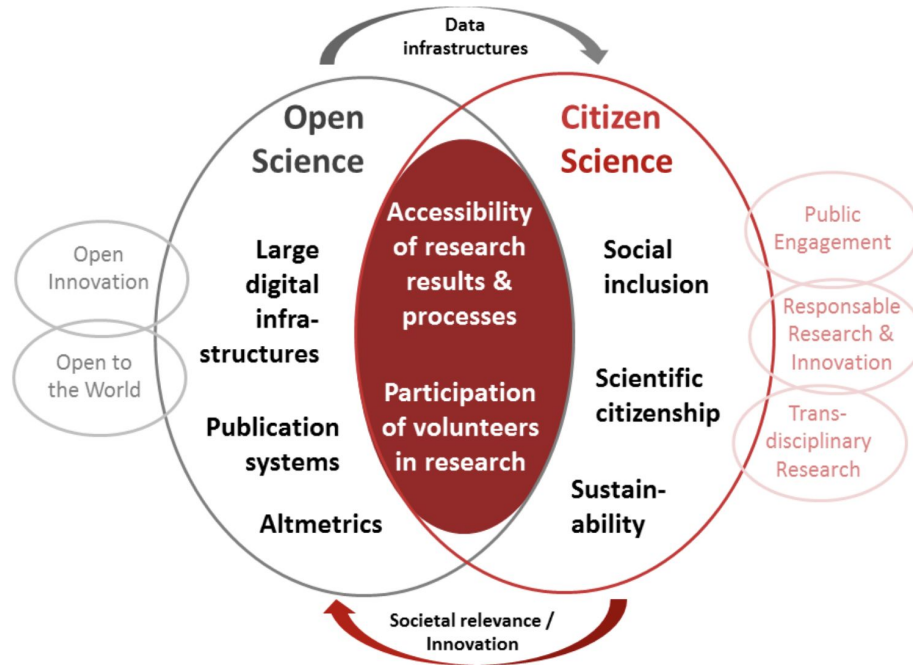
Source: [LinkedIn](#)

7 weeks in 2020 lockdown in Spain



[Resistere \(I will resist\)](#)

Moving towards Open Science



Source: Citizen Science and Open Science Core Concepts and Areas of Synergy (Vohland and Göbel, 2017)

About CoronaWhy

www.coronawhy.org



[Home](#)

[Daily Progress](#)

[Calendar](#)

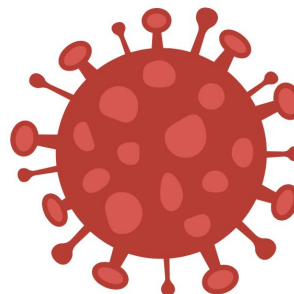
[JOIN THE FIGHT!](#)

FIGHTING CORONAVIRUS WITH ARTIFICIAL INTELLIGENCE

We are a globally distributed, volunteer-powered research organisation, trying to assist the medical community's ability to answer key questions related to COVID-19

[JOIN THE FIGHT](#)

[LEARN MORE](#)



Who we are?

[Artur Kiulian](#) started CoronaWhy because he realized that we are all in this together. Now CoronaWhy is an international group of 900+ volunteers whose mission is to improve global coordination and analysis of all available data pertinent to the COVID-19 outbreak and ensure all findings reach those who need them.

It's impossible to list everyone out but we will eventually.

1300+ people
registered in the
organization, more
than 300 actively
contributing!

COVID-19 Open Research Dataset Challenge (CORD-19)

It's all started from [this](#) (March, 2020):

"In response to the COVID-19 pandemic and with the view to boost research, the Allen Institute for AI together with CZI, MSR, Georgetown, NIH & The White House is collecting and making available for free the COVID-19 Open Research Dataset (CORD-19). This resource is updated weekly and contains over 287,000 scholarly articles, including 82,870 with full text, about COVID-19 and other viruses of the coronavirus family." ([Kaggle](#))

Collaborators

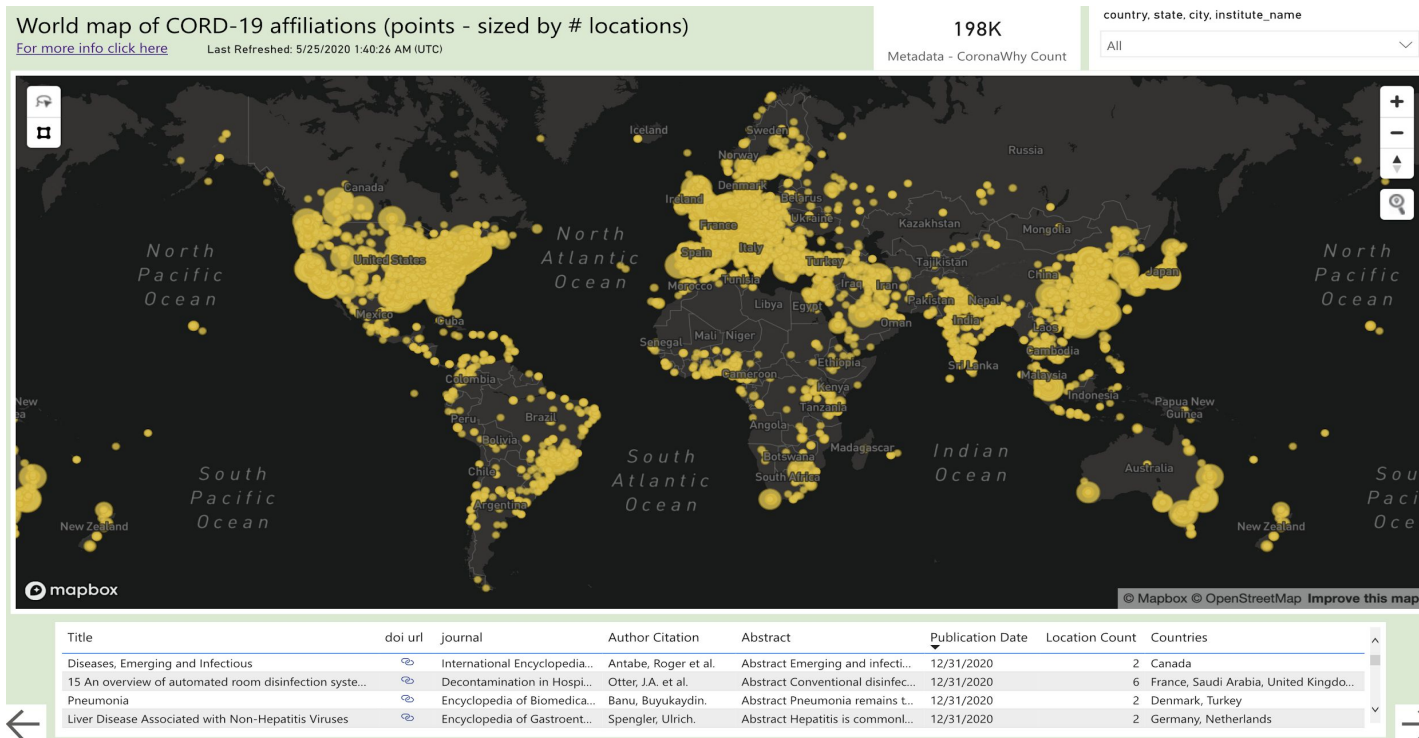
CORD-19 was made possible by the [Semantic Scholar team](#) at the [Allen Institute for AI](#) in collaboration with the following institutions:



CoronaWhy Community Tasks (March-April 2020)

1. [Task-Risk](#) helps to identify risk factors that can increase the chance of being infected, or affects the severity or the survival outcome of the infection
2. [Task-Ties](#) to explore transmission, incubation and environment stability
3. [Match Clinical Trials](#) allows exploration of the results from the [COVID-19 International Clinical Trials](#) dataset
4. [COVID-19 Literature Visualization](#) helps to explore the data behind the AI-powered literature review
5. [Named Entity Recognition](#) across the entire corpus of COVID-19 papers with full text

CORD-19 affiliations recognized with Deep Learning



Source: [CORD-19 map visualization](#) and [institution affiliation data](#)

Collaboration with other organizations

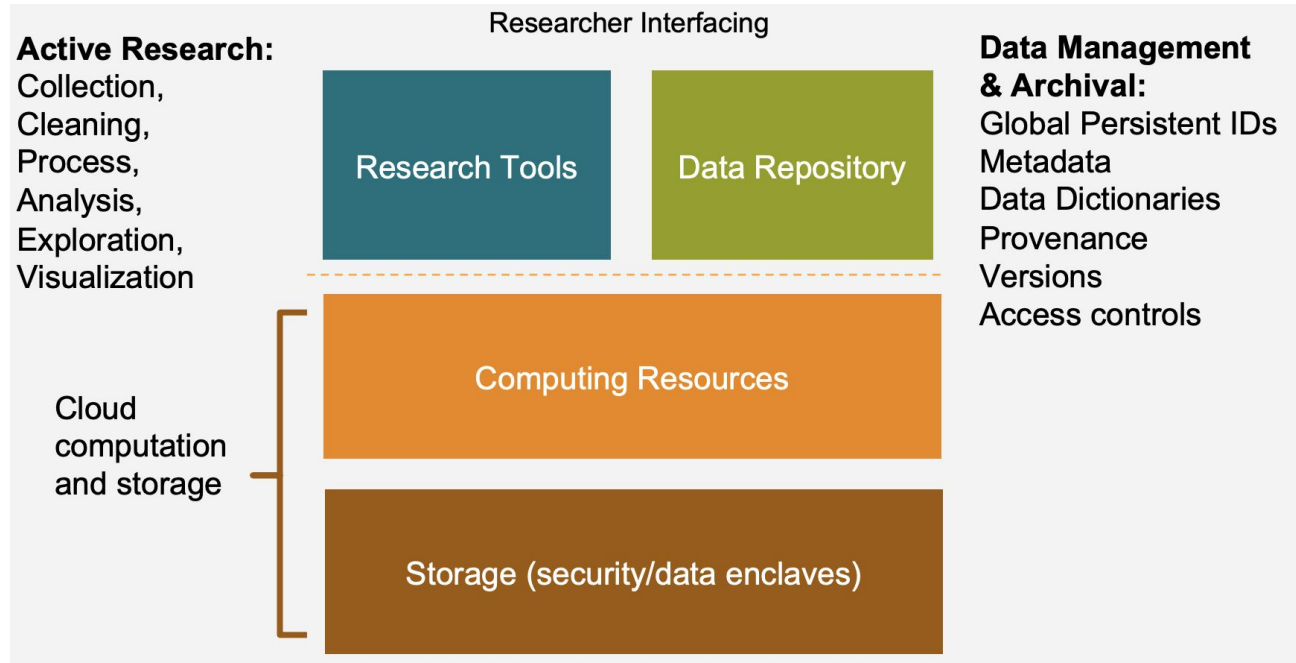
- Harvard Medical School, INDRA integration
- Helix Group, Stanford University
- NASA JPL, COVID-19 knowledge graph and GeoParser
- Kaggle, coronamed application
- Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS, knowledge graph
- dcyphr, a platform for creating and engaging with distillations of academic articles
- CAMARADES (Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies)

We've got almost endless data streams...

Collaborations problems

- too much datasets shared without taking into account FAIR principles
- limited use of common domain specific ontologies (MeSH, Wikidata, ...)
- lack of the usage of controlled vocabularies
- collaboration in research between academic and industrial contexts is very low
- interoperability layer between different tools is limited or absent
- there is no common data infrastructure suitable for all use cases

Collaborations and best practice data sharing



Merce Crosas, "[Harvard Data Commons](#)"

WHAT IS FAIR ?

Findable:

- FM-F1A** **FM-F1B**
F1 (meta)data are assigned a globally **unique** and **persistent** identifier;
- FM-F2**
F2 data are described with **rich metadata**;
- FM-F3**
F3 metadata clearly and explicitly include the **identifier of the data** it describes;
- FM-F4**
F4 (meta)data are registered or **indexed** in a searchable resource;

Interoperable:

- FM-I1**
I1 (meta)data use a formal, accessible, shared, and broadly applicable **language for knowledge representation**.
- FM-I2**
I2 (meta)data use **vocabularies that follow FAIR principles**;
- FM-I3**
I3 (meta)data include **qualified references** to other (meta)data;

Accessible:

- A1 (meta)data are retrievable by their identifier using a standardized communications protocol;
- FM-A1.1**
A1.1 the protocol is **open, free, and universally implementable**;
- FM-A1.2**
A1.2 the protocol allows for an **authentication and authorization** procedure, where necessary;
- FM-A2**
A2 metadata are accessible, **even when the data are no longer available**;

Reusable:

- R1 meta(data) are richly described with a plurality of accurate and relevant attributes;
- FM-R1.1**
R1.1 (meta)data are released with a clear and **accessible data usage license**;
- FM-R1.2**
R1.2 (meta)data are associated with **detailed provenance**;
- FM-R1.3**
R1.3 (meta)data meet domain-relevant **community standards**;

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)

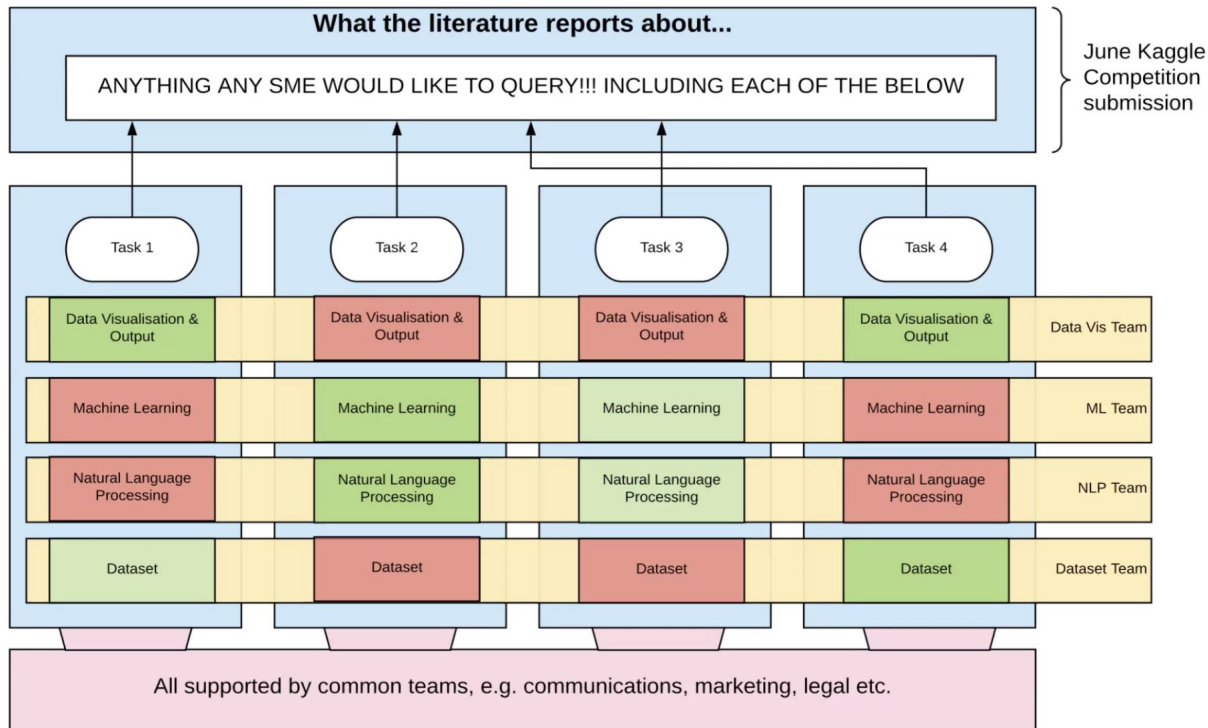
<http://fairmetrics.org>

<https://github.com/FAIRMetrics/Metrics/blob/master/ALL.pdf>



[DANS-KNAW](#) is one of worldwide leaders in FAIR Data ([FAIRsFAIR](#))

We need a horizontal platform to serve vertical teams



Source: CoronaWhy organization

What is Dataverse and How It Can Help?

- Open source data repository developed by IQSS of Harvard University
- Great product with very long history (from 2006) created by experienced and Agile development team
- Clear vision and understanding of research communities requirements, public roadmap
- Well developed architecture with rich APIs allows to build application layers around Dataverse
- Strong community behind of Dataverse is helping to improve the basic functionality and develop it further.
- DANS-KNAW is leading SSHOC WP5.2 task to deliver production ready Dataverse repository for the European Open Science Cloud (EOSC) communities CESSDA, CLARIN and DARIAH.

Federated Dataverse data repositories worldwide

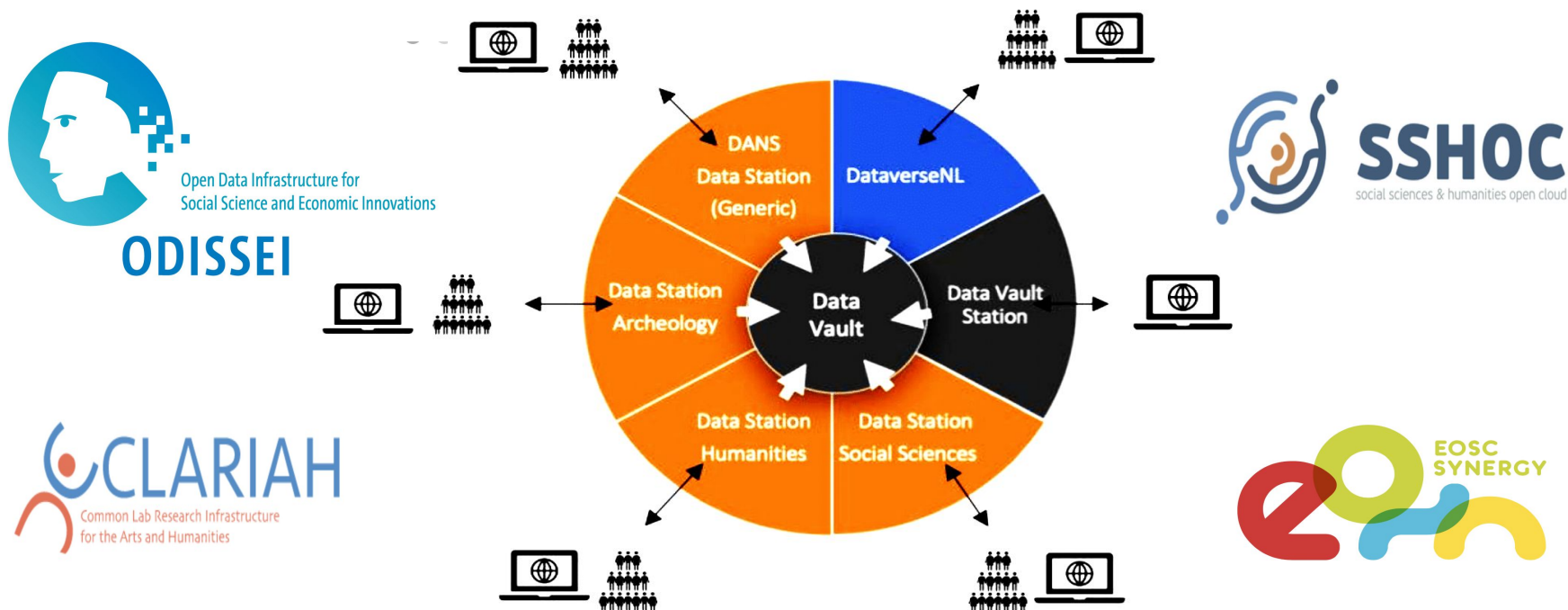


Developed at Harvard's Institute for Quantitative Social Science (IQSS) with contributions from the Dataverse community (<https://dataverse.org>)

- **Open-source**
- **63** installations
- **6** continents
- **7K** Dataverse collections
- **135K** datasets
- **800K** files
- **28M** file downloads
- **Metadata** shared across repositories

Source: Merce Crosas, [Harvard Data Commons](#)

Data Stations - Future Data Services



Dataverse is API based data platform and a key framework for Open Innovation!

Coming back to FAIR principles

FM [AID*]	Question	Dataverse Q'aire	Dataverse Optimized
Identifier type	1	DOI	DOI
F1A	2		
F1B	Not tested in Q'aire		
F2A	4A		
F2A	4B		
F3	5B		
F4	6A		
F4	6B		
A1.1	7A		
A1.2	8A		
A1.2	8B	N/A	N/A
A2	9		
I1	10		
I2	11		
I3	12		
R1.1	13		
R1.2	14A		

DATAVERSE FAIR SUMMARY

- Strong support for Findable, Accessible, and Reusable principles
- Weak for Interoperable principles
- In agreement* with FAIR test results (*F3 was fixed after test)
- There is no FAIR “compliance”
- Instead, it’s a process and can always be improved

Source:

[Mercè Crosas,](#)

[“FAIR principles and beyond: implementation in Dataverse”](#)

Benefits of the Common Data Infrastructure

- It's distributed and sustainable, suitable for the future
 - maintenance costs will drop massively, as more organizations will join, less expensive it will be to support
 - maintenance costs could be reallocated to the training and further development of the new (common) features
 - reuse of the same infrastructure components will enforce the quality and the speed of the knowledge exchange
 - building a multidisciplinary teams reusing the same infra can bring us new insights and unexpected views
 - Common Data Infrastructure plays a role of the **universal gravitation** layer for Data Science projects
- (and so on...)


Dataverse as data integration point

The screenshot shows the Dataverse COVID-19 Data Hub interface. At the top, the Dataverse logo is on the left, and navigation links (Search, User Guide, Support, Sign Up, Log In) are on the right. Below the header, the title "COVID-19 Data Hub" is displayed. A metrics bar shows "3,429 Downloads" and links for "Contact" and "Share". A disclaimer states: "Information and Data hub produced by all CoronaWhy research groups. Please join us if you want to help in the fight against COVID-19. Disclaimer: at the moment all materials published on this site are available for public for the demonstration purposes, without DOI Persistent Identifiers." Below this is a carousel of featured datasets: "CoronaWhy Task Ties", "Pandemics in History", "INDRA", and "COVID-19 Knowledge Graphs". A search bar is present with a "Find" button and a link to "Advanced Search". On the left sidebar, filters for "Dataverses (24)", "Datasets (508)", and "Files (22,755)" are shown, along with "Dataverse Category" (Research Group, Laboratory, Research Project, Organization or Institution) and "Publication Year" (2020). The main content area displays "1 to 10 of 532 Results". The first result is "NASA JPL Knowledge Graph extracted from CORD-19 collection" by McGibbney, Lewis John, dated Aug 24, 2020, with a DOI link. The second result is "NASA Data Hub (NASA)" dated Aug 24, 2020, describing collaborative efforts. The third result is "INDRA statements" by Gyori, Benjamin M., dated Aug 24, 2020, with a DOI link.


- Started in April, 2020
- Used by CoronaWhy vertical teams for the data exchange and share
- Intended to help researchers to make their data FAIR
- One of the biggest COVID-19 data archives in the world with 700k files (public and private)
- New teams are getting own data containers and can reuse data collected and produced by others

Visit <https://datasets.coronawhy.org>

Dataset from CoronaWhy vertical teams

 Dataverse


Search ▾ User Guide Support Sign Up Log In

 CoronaWhy Task Risk (CoronaWhy)

COVID-19 Data Hub > CoronaWhy Task Risk >

COVID-19 risk factors

Version 1.0



Mayya Lihovodov; Pranjalya Tiwari; Ansun Sujoe; Guillermo Blanco; Iason Konstantinidis; Kriti Mahajan; Robbie Edwards; Vijay Datta; Michael Wang; Lukasz Gagala; Brandon Eychaner; Mohammad Tanweer; Anrew Wood; Kevin Lee; Samtha Reddy; Mark Koranda; Ruslan Oliynyk, MD; Mike Honey; Randall Brown, MD; Artur Kiulian, 2020, "COVID-19 risk factors", <https://doi.org/10.5072/FK2/3OZLV6>, COVID-19 Data Hub, V1

Cite Dataset ▾ Learn about [Data Citation Standards](#).

Access Dataset ▾

Contact Owner Share

Dataset Metrics ⓘ

184 Downloads ⓘ

Description ⓘ

A major topic of interest among researchers is the study of the various risk factors related to COVID-19. A risk factor is anything that increases the chance of being infected, or affects the severity or the survival outcome of the infection. Many of the papers in the dataset are studies on the severity and outcome of the infection, without, however, any systematic documentation that would be easily searchable.

The focus of this study is to extract and present in a meaningful and easily accessible way scientific papers that are related to risk factors associated with viral diseases through a procedure that can be automated as much as possible.

At the current stage, a semi-automated approach is implemented using manual review of retrieved papers. It is important to note that through the proposed procedure a small subset of papers is manually reviewed, the ones that are identified as most probable to be relevant to a specific risk factor. This brings the volume of papers for review down to less than 100-200 instead of multiple thousands, rendering the review task feasible in much shorter timeframes.

Also, at the current stage the paper extraction is limited to the following factors:

Environmental: Pollution, Population Density, Humidity, Temperature
Comorbidity: Heart diseases
Demographics: Senior age
Lifestyle: Smoking

The above risk factors were identified as being the most important by the medical community. An extensive list of risk factors is provided under section 4 below and is subject of a future extension of this study.

Source: [CoronaWhy Dataverse](#)

COVID-19 data files verification

The screenshot displays the Dataverse web application interface. At the top, the navigation bar includes the Dataverse logo, links for 'Add Data', 'Search', 'User Guide', 'Support', 'Sign Up', and 'Log In'. The main content area shows the dataset details for 'coronavirus' under the subject 'Medicine, Health and Life Sciences'. Below this, there are tabs for 'Files', 'Metadata', 'Terms', and 'Versions', with 'Files' being the active tab. A 'Change View' section offers 'Table' and 'Tree' views. A search bar is present with the placeholder text 'Search this dataset...'. Below the search bar, there are filters for 'File Type: All', 'Access: All', and 'File Tag: All'. The main list shows '1 to 10 of 17 Files'. Three files are visible, each with a document icon, a title, a description, and a 'Download' button. The first file is 'age-distribution-died-and-survivors.xlsx', the second is 'age-distribution-status.xlsx', and the third is 'age-distribution-status.xlsx'. Each file entry includes a detailed description, a data snapshot URL, and a list of tags such as 'patients_deceased', 'dutchcovid19data', 'age_group', 'patients_recovered', 'patients_in_hospital', 'recovered_patients', 'patients_in_icu', and 'deceased_patients'.

Description coronavirus

Subject Medicine, Health and Life Sciences

Files Metadata Terms Versions

Change View Table Tree

Search this dataset... Find

Filter by
File Type: All Access: All File Tag: All Sort

1 to 10 of 17 Files Download

[age-distribution-died-and-survivors.xlsx](#)
data/hospitalized/age-distribution-died-and-survivors.xlsx/
MS Excel Spreadsheet - 5.0 KB - May 30, 2020 - 0 Downloads
MD5: 7343b296eaf59c98f0426e1a03234221
Data snapshot from <https://raw.githubusercontent.com/Sikerdebaard/dutchcovid19data/master/data/hospitalized/age-distribution-died-and-survivors.xlsx>
[patients_deceased](#) [dutchcovid19data](#) [age_group](#) [patients_recovered](#) Download

[age-distribution-status.xlsx](#)
data/hospitalized/age-distribution-status.xlsx/
MS Excel Spreadsheet - 5.1 KB - May 30, 2020 - 0 Downloads
MD5: ea9604d071e3597748ec185b19d956cc
Data snapshot from <https://raw.githubusercontent.com/Sikerdebaard/dutchcovid19data/master/data/hospitalized/age-distribution-status.xlsx>
[patients_deceased](#) [dutchcovid19data](#) [patients_in_hospital](#) [age_group](#) [patients_recovered](#) Download

[age-distribution-status.xlsx](#)
data/age-distribution-status.xlsx/
MS Excel Spreadsheet - 5.1 KB - May 30, 2020 - 0 Downloads
MD5: 9c418de86b10f0aef872200885b5a4a3
Data snapshot from <https://raw.githubusercontent.com/Sikerdebaard/dutchcovid19data/master/data/age-distribution-status.xlsx>
[dutchcovid19data](#) [patients_in_hospital](#) [recovered_patients](#) [patients_in_icu](#) [deceased_patients](#) [age_group](#) Download


We do a verification of every file by importing its contents to dataframe.

All column names (variables) extracted from tabular data available as labels in files metadata

We've enabled Dataverse data previewers to browse through the content of files without download!

We're starting internal challenges to build ML models for the metadata classification

Dataverse content in Jupyter notebooks

 Dataverse content, data and metadata import to Pandas Dataframes ☆

File Edit View Insert Runtime Tools Help [All changes saved](#)

+ Code + Text

RAM Disk

Editing

<>

Now read the content of first file to dataframe

```
FILEID=pdfiles.loc['id'][0]
fileURL = "%s/api/access/datafile/%s" % (BASE_URL, FILEID)
df = pd.read_csv(fileURL)
df.head()
```

	Unnamed: 0	Risk Factor	Title	Keyword/Ngram	No of keyword occurrence in Paper	paper_id	URL	Sentences	Authors	Correlation	Design Methodology	S
0	0	age	Coronavirus-like particles in nonhuman primate...	['older age group']	1	3cf9a172522a7db0df9e436029707bb6e3e0ff8c	[https://www.ncbi.nlm.nih.gov/pmc/articles/PM...	['It might be assumed that coronaviruses are n...	['Smith, G. C.; Lester, T. L.; Heberling, R. L...	It might be assumed that coronaviruses are not...	NaN	
1	1	age	Estimates of the severity of coronavirus disea...	['60 years and over', 'older age group']	7	ac2a1ba62fdf52eb276bf42b22ed3d927b5330b1	[https://doi.org/10.1016/s1473-3099(20)30243-...	['Reported cases in Wuhan were more frequent l...	['Verity, Robert; Okell, Lucy C; Dorigatti, Il...	Reported cases in Wuhan were more frequent in ...	In cases reported outside of mainland China , ...	V nur obser in I
2	2	age	Infections in travellers returning to Turkey f...	['65 years old']	3	f33e3be8c6ec1d348cf8983037dcf8adb25e7f94	[https://www.ncbi.nlm.nih.gov/pmc/articles/PM...	['Seventy four (40 %) of them were ≥ 65 years ...	['Erdem, H.; Ak, O.; Elaldi, N.; Demirdal, T.;	A total of 185 Turkish patients were recruited...	NaN	
3	3	age	The use of corticosteroid as treatment in SARS...	['patients older than']	1	52f5440ec7a22706f95be3c6f5e0ed2e940e4945	[https://doi.org/10.1016/j.jinf.2004.09.008'...	['A total of 80 patients older than 18 years o...	['Auyeung, Tung Wai; Lee, Jenny S.W.; Lai, Win...	NaN	NaN	
4	4	age	Burden, seasonal pattern and symptomatology of...	['older age group']	5	0957f96f8188f65cc145464dc7882abd259e0f5f	[https://doi.org/10.1016/j.cmi.2015.05.027', ...	['On comparison of the two age groups , viral...	['Wei, L.; Chan, K.-H.; Ip, D.K.M.; Fang, V.J....	On comparison of the two age groups , viral a...	NaN	

Source: [Dataverse examples on Google Colabs](#)

COVID-19 Data Crowdsourcing

Dataset in CoronaWhy Dataverse #579

 Open k-goncharova opened this issue 2 days ago · 4 comments

New issue



k-goncharova commented 2 days ago



Hello, Your dataset was added to CoronaWhy (<https://www.coronawhy.org/>) Data Lake on Dataverse as a piece of common COVID-19 dataframe <https://datasets.coronawhy.org/dataset.xhtml?persistentId=doi:10.5072/FK2/A20BEO>

Would you be willing to help with maintenance of your dataset in Dataverse, e.g. adding the relevant metadata and keeping the dataset up-to-date? That will help to make the dataset findable and accessible for medical science community.

Assignees

No one assigned

Labels

None yet

Projects

None yet

Milestone

No milestone

Linked pull requests

Successfully merging a pull request may close this issue.

None yet

Notifications

Customize



You're not receiving notifications from this thread.

2 participants



swsoyee commented 2 days ago

Owner



Hi, @k-goncharova

Sure, I will keep updating my dataset, and what should I do in Dataverse?



k-goncharova commented 2 days ago

Author



Great, thank you. Please register in CoronaWhy Dataverse <https://datasets.coronawhy.org/> and I'll give you the Dataset creator permissions to your dataset. Then please add relevant metadata - description and keywords to your dataset.



swsoyee commented 2 days ago

Owner



Okay, my user's name is swsoyee .

I'm a little busy these days, so the progress maybe slow, please forgive me.

CoronaWhy data management team review of all harvested datasets and trying to identify the important data.

We're approaching github owners by creating issues in their repos and inviting them to help us.

More than 20% of data owners were joining CoronaWhy community or interested to curate their datasets.

Bottom-up data collection works!

How about sustainability and interoperability of this data infrastructure?

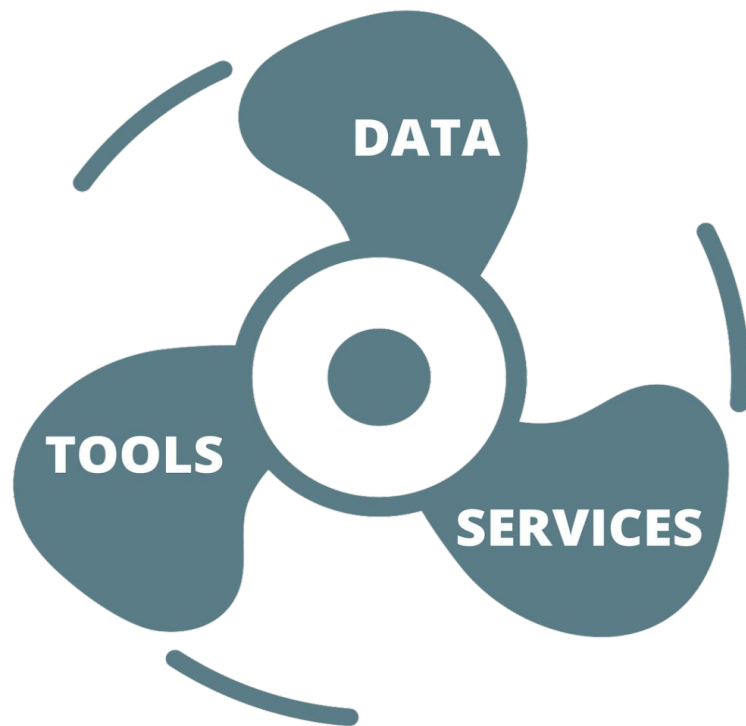
Interoperability in the European Open Science Cloud (EOSC)

- **Technical interoperability** defined as the “ability of different information technology systems and software applications to communicate and exchange data”. It should allow “to accept data from each other and perform a given task in an appropriate and satisfactory manner *without the need for extra operator intervention*”.
- **Semantic interoperability** is “the ability of computer systems to transmit data with unambiguous, shared meaning. Semantic interoperability is a requirement to enable *machine computable logic, inferencing, knowledge discovery, and data*”.
- **Organisational interoperability** refers to the “way in which organisations align their business processes, responsibilities and expectations to achieve commonly agreed and mutually beneficial goals. Focus on the *requirements of the user community* by making services available, easily identifiable, accessible and user-focused”.
- **Legal interoperability** covers “the broader environment of *laws, policies, procedures* and cooperation agreements”

FAIR data and services

a solid and sustainable infrastructure for the 'core' of the propeller image:

- mapping tables, protocols and other community emerging standards should not only **find a 'home'** (such as for instance FAIRsharing), but should also be **collectively endorsed** and used in practice by much more **coherent communities**.
- to support the process of coordination within and across implementation, training and certification networks to **minimise reinvention of redundant infrastructure components**, including such things as thesauri and domain specific or generic ontologies protocols



Our goals to increase interoperability on the global scale

Provide a custom FAIR metadata schema for European research communities:

- CESSDA metadata (Consortium of European Social Science Data Archives)
- Component MetaData Infrastructure (CMDI) metadata from CLARIN linguistics community

Connect metadata to ontologies and external controlled vocabularies:

- link metadata fields to common ontologies (Dublin Core, DCAT)
- define semantic relationships between (new) metadata fields (SKOS)
- select available external controlled vocabularies for the specific fields
- provide multilingual access to controlled vocabularies

All contributions should go to Dataverse source code and available worldwide!

The importance of standards and ontologies

Generic controlled vocabularies to link metadata in the bibliographic collections are well known: ORCID, GRID, GeoNames, Getty.

Medical knowledge graphs powered by:

- Biological Expression Language (BEL)
- Medical Subject Headings (MeSH®) by U.S. National Library of Medicine (NIH)
- Wikidata (Open ontology) - Wikipedia

Integration based on metadata standards:

- MARC21, Dublin Core (DC), Data Documentation Initiative (DDI)

The most of prominent ontologies already available as a Web Services with API endpoints.

SKOSMOS framework to discover ontologies

Skosmos

Vocabularies About Feedback Help | in English

Global Research Identifier Database GRID

Content language English

Search

Alphabetical

Hierarchy

Groups

A Á Â Ã Ä Å Æ Ç D E É F G H

I Í J K L Ł M N O Ö Ø P Q

R Š S Š Š Ţ T U Ú Û Ü V W X Y Z

Ž H * 0-9

Dassault Aviation (France)

Dassault Systèmes (Canada)

Dassault Systèmes (France)

Dassault Systèmes (Germany)

Dassault Systèmes (Japan)

Dassault Systèmes (United Kingdom)

Dassault Systèmes (United States)

Dat (Norway)

Data & Society Research Institute

Data Access Technologies → Model Driven Solutions (United States)

Data Archiving and Networked Services

Data Assurance and Communication Security

Data Assurance and Communication Security Research Center → Data Assurance and Communication Security

Data Fusion International (Ireland)

Data Fusion Research Center

Data Harbor (United States)

Data Management (Italy)

Data Management Services (United States)

Data Numerica Institute (United States)

Data Observation Network for Earth → DataONE

Data One Global (United States)

Data Power Decisions (United States)

Data Respons (Norway)

Data Sciences International (United States)

Data Security Council of India

Data Storage Institute

Data Voice Exchange (United States)

Data-Mate (Finland)

DATA4 (France)

Data61

Data:Lab Munich (Germany)

Databank (Italy)


Datchassi (Sweden)

DataCite

Datacon (Czechia)

Datacorp (United States)

PREFERRED TERM



ENTRY TERMS

Netherlands Institute for Permanent Access to Digital Research Resources

DATE

2005-01-01

IDENTIFIER

grid.500519.8

HOMEPAGE


<https://dans.knaw.nl/en>

EXACTLY MATCHING

<http://www.wikidata.org/entity/Q13570995>

CONCEPTS

URI



<http://www.grid.ac/institutes/grid.500519.8> 


Download this concept:


[RDF/XML](#) [TURTLE](#) [JSON-LD](#)













- SKOSMOS is developed in Europe by the National Library of Finland (NLF)
- active global user community
- search and browsing interface for SKOS concept
- multilingual vocabularies support
- used for different use cases (publish vocabularies, build discovery systems, vocabulary visualization)

The same metadata field linked to many ontologies

 **Dataverse** Search ▾ User Guide Support  **Dataverse Admin** ▾

Subject 

Keyword 

Vocabulary 	Term 	
<input type="text" value="unesco"/>	<input type="text" value="Family"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL 		
<input type="text" value="http://skos.um.es/unescothes/C01489"/>		
Vocabulary 	Term 	
<input type="text" value="thesaurus"/>	<input type="text" value="family labour"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL 		
<input type="text" value="http://vocabulaires.irstea.fr/thesaurus/T"/>		
Vocabulary 	Term 	
<input type="text" value="agrovoc"/>	<input type="text" value="families"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL 		
<input type="text" value="http://aims.fao.org/aos/agrovoc/c_2785"/>		
Vocabulary 	Term 	
<input type="text" value="iptc"/>	<input type="text" value="fam"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL 	<ul style="list-style-type: none">• family• family planning• famine	

Language switch in Dataverse will change the language of suggested terms!

Use case: COVID-19 expert questions

"In response to the COVID-19 pandemic, the Epidemic Question Answering (EPIC-QA) track challenges teams to develop systems capable of automatically answering ad-hoc questions about the disease COVID-19, its causal virus SARS-CoV-2, related corona viruses, and the recommended response to the pandemic. While COVID-19 has been an impetus for a large body of emergent scientific research and inquiry, the response to COVID-19 raises questions for consumers."

```
▼ {
  question_id: "EQ001",
  question: "what is the origin of COVID-19",
  query: "coronavirus origin",
  background: "seeking range of information about the SARS-CoV-2 virus's origin, including its evolution, animal source, and first transmission into humans"
},
▼ {
  question_id: "EQ002",
  question: "how does the coronavirus respond to changes in the weather",
  query: "coronavirus response to weather changes",
  background: "seeking range of information about the SARS-CoV-2 virus viability in different weather/climate conditions as well as information related to transmission of the virus in different climate conditions"
},
▼ {
  question_id: "EQ003",
  question: "will SARS-CoV2 infected people develop immunity? Is cross protection possible?",
  query: "coronavirus immunity",
  background: "seeking studies of immunity developed due to infection with SARS-CoV2 or cross protection gained due to infection with other coronavirus types"
},
▼ {
  question_id: "EQ004",
  question: "what causes death from Covid-19?",
  query: "how do people die from the coronavirus",
  background: "Studies looking at mechanisms of death from Covid-19."
},
▼ {
  question_id: "EQ005",
  question: "what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?",
  query: "animal models of COVID-19",
  background: "Papers that describe the results of testing drugs that bind to spike proteins of the virus or any other drugs in any animal models. Papers about SARS-CoV-2 infection in cell culture assays are also relevant."
},
```

Source: [Epidemic Questions Answering](#)

COVID-19 questions in SKOSMOS framework

Skosmos

Vocabularies About Feedback Help | in English

COVID-19 QA

Content language English Search

Alphabetical

Hierarchy

Groups

A D H I U W

animal models of COVID-19 → are there any drugs that work for SARS-CoV or SARS-CoV-2 in animals?
are heart complications likely in patients with COVID-19?
are patients taking Angiotensin-converting enzyme inhibitors (ACE) inhibitors at increased risk for COVID-19?
are there any clinical trials available for the coronavirus
are there any drugs that work for SARS-CoV or SARS-CoV-2 in animals?
are there any existing drugs that could repurposed to treat COVID-19?
are there blood tests that detect antibodies to coronavirus?

> are there any clinical trials available for the coronavirus

PREFERRED TERM

are there any clinical trials available for the coronavirus 

BROADER CONCEPT

<https://www.wikidata.org/wiki/Q2252352>

RELATED CONCEPTS

<https://www.wikidata.org/wiki/Q84263196>

ENTRY TERMS

coronavirus clinical trials

SCOPE NOTE

seeking information on what COVID-19 clinical trials are available and who can enroll

URI

<https://skosmos.coronawhy.org/COVID19#CQ017> 

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

COVID-19 questions in Dataverse metadata

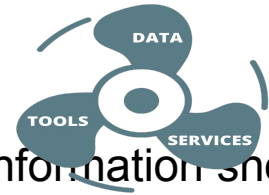
The screenshot shows the Harvard Dataverse interface for a dataset. The top navigation bar includes the Harvard Dataverse logo, a search bar, and links for Add Data, Search, About, User Guide, Support, and a user profile for Vyacheslav Tykhonov. The main content area displays a dataset titled 'percentage of cases reported), Accumulated deceased (Data obtained from the analysis of a daily variable percentage of cases reported)'. Below the title, there is a 'Date' field with a placeholder 'YYYY-MM-DD'. The 'Subject' field is set to 'Medicine, Health and Life Sciences' and 'Social Sciences'. The 'Keyword' field is empty. The 'Term' field is set to 'how has COVID-19 affected Spain'. The 'Vocabulary' field is set to 'COVID-QA'. The 'Vocabulary URL' field is set to 'https://coronawhy.org/COVIDQA#EM00I'. The 'Term' field is set to 'Spain'. The 'Vocabulary' field is set to 'Icsh'. The 'Vocabulary URL' field is set to 'http://id.loc.gov/authorities/names/n790'. The 'Term' field is set to 'Hospitalised patients with symptoms'. The 'Vocabulary' field is empty. The 'Vocabulary URL' field is set to 'Enter full URL, starting with http://'. The 'Term' field is set to 'Total amount of positive cases'. The 'Vocabulary' field is empty. The 'Vocabulary URL' field is set to 'Enter full URL, starting with http://'.

- COVID-19 ontologies can be hosted by SKOSMOS framework
- Researchers can enrich metadata by adding standardized questions provided by SKOSMOS ontologies
- rich metadata exported back to Linked Open Data Cloud to increase a chance to be found
- enriched metadata can be used for further ML models training

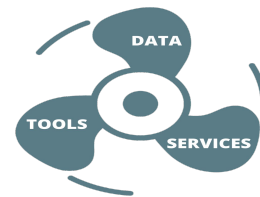
Source: [COVID-19 European data hub in Harvard Dataverse](#)

Vision of the Operating System for Open Science

- **Common Research and Data Infrastructure** should be distributed and robust enough to be scaled up and reused for other tasks like cancer research.
- Services should be build from Open Source components.
- Data processed and published in FAIR way, the provenance information should be a part of the Data Lake.
- Data evaluation and credibility is the top priority, we need tools for the expert community for the verification of our datasets
- The transparency of data and services should guarantee the reproducibility of all experiments and bring new insights in the multidisciplinary research



BeFAIR Open Science Framework



☰ README.md ✎

BeFAIR

BeFAIR (Be Findable, Accessible, Interoperable, Reusable) Open Science Framework.

BeFAIR is a Common Distributed Research Infrastructure where users can add and run any tools and components by themselves using Debian's way of managing services. All selected services should be available on a selected subdomain name and could be easily integrated together with [Dataverse](#), BeFAIR data repository.

BeFAIR was designed as out-of-the-box Distributed Networked Infrastructure that any research community can install with one command just as normal Operating System. The roadmap includes releases containing Open Data available for the different sciences, however [COVID-19 Data Hub](#) is our current priority.

Acknowledgements

BeFAIR infrastructure is standing the Shoulder of Giants. Please find below the acknowledgements for resources and contributions from the finished on ongoing projects.

Region	Project	Funding information	Component
European Union	CESSDA SaW	H2020-INFRADEV-1-2015-1, grant agreement #674939	Dataverse as a service
European Union	SSHOC	H2020-INFRAEOSC-04-2018, grant agreement #823782	Cloud Dataverse
European Union	EOSC Synergy	INFRAEOSC-05-2018-2019, grant agreement No 857647	SQAaaS service
European Union	EOSC-hub	H2020-EINFRA-12-2017, grant agreement #777536	DataTags as a service
United States	INDRA	Defense Advanced Research Projects Agency under award W911NF-14-1-0397	INDRA service
European Union	FAIRsFAIR	H2020-INFRAEOSC-2018-2020 Grant agreement 831558	F-UJI and FAIR Data Points

Basic Infrastructure of COVID-19 Museum

CoronaWhy/befair Unwatch 4 ☆

< Code Issues 21 Pull requests Actions Projects Wiki Security Insights Settings

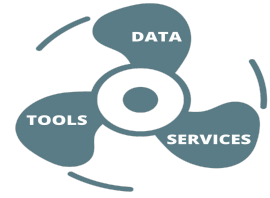
main befair / distros / covid19museum / Go to file Add file ...

ramok Refactor *.mk and distros Makefile 4195e63 22 days ago History

..		
.env	Mass rename distributive -> distro	last month
Makefile	Refactor *.mk and distros Makefile	22 days ago
airflow.env	Mass rename distributive -> distro	last month
airflow.yaml	Mass rename distributive -> distro	last month
dataverse-5.3.yaml	Mass rename distributive -> distro	last month
postgresql.yaml	Mass rename distributive -> distro	last month
solt.yaml	Mass rename distributive -> distro	last month
traefik-https.yaml	Mass rename distributive -> distro	last month

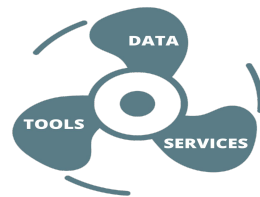
Visit <https://github.com/CoronaWhy/befair>

BeFAIR infra in the nutshell



- flexible and “fluid” infrastructure that allows to connect any dockerized services in easy way
- all operations (start/stop/...) done with **make** command in “Debian way”:
make up; make down
- contains a various distributives (distros) with interconnected services exposed and managed by traefik proxy
- users can enable/disable both services and distros there in a convinient way with menuconfig command
- users can use BeFAIR both locally, on a server on deployed on Cloud
- maintenance of Docker containers is the responsibility of their developers, BeFAIR is an integration part, services could be on/off

Supporting Semantic Web for Data

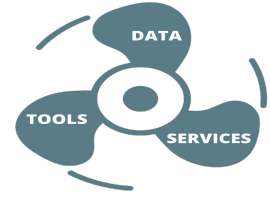


Historically most of datasets preserved in data silos (archives), not interlinked and lacking of standardization. There are **cultural**, **structural** and **technological** challenges.

Solutions:

- Integrating Linked Data and Semantic Web technologies, forcing research communities to share data and add more interoperability following FAIR principles
- Create a standardized (meta)data layer for Large Scale projects like Time Machine and CoronaWhy
- Working on the automatic metadata linkage to ontologies and external controlled vocabularies in order to get it linked in the Linked Open Data Cloud
- Using the Knowledge Graph for the Machine Learning

Why Artificial Intelligence?



Human resources are very expensive and deficit, it's difficult to find appropriate expertise in-house.

Solution:

- Building AI/ML pipelines for the automatic metadata enrichment and linkage prediction
- applying NLP for NER, data mining, topic classification etc
- building multidisciplinary knowledge graphs should facilitate the development of new projects with economic and social scientists, they will take ownership of their own data if they see value (Clio Infra)

Using modern NLP frameworks (spacy, for example)

['GDP Official Exchange Rate ORG and GDP Purchasing Power Parity Comparison: East African Community ORG (EAC) A Comparative Study',
'<p>The strength of the book is that it is comprehensive, well written and accessible to students who don't have an in depth- understanding of
formal economics.'(Matthew Cole PRODUCT , Birmingham GPE University)<p>The main objectives of this Review
paper is to answer three CARDINAL questions in easy and soft language for an in depth understanding. which country is more advanced in East African
NORP community (EAC) using comparison of GDP Official exchange rate (GDP OER) and GDP Purchasing power party (GDP PPP), Is Rwanda GPE
contributing more to global Economy than other countries member of EAC, and the last one is the average person income or wealthier in Rwanda GPE
compared to other countries member of East African NORP community, all these issues are being addressed in this review paper
<p>The methodology used in this paper was a Descriptive ORG study, where EAC members where described based on historical
data published from world bank, IMF, NISR ORG , and then compared where comparison was based on GDP official exchange rate, PPP, expenditures
partners, Trade per capita, exports, imports and worldwide governance indicators.<p>The result showed that
Kenya GPE is the most advanced country in EAC members, with highest Nominal ORG GDP in US GPE dollar, times more than one point five
CARDINAL of Tanzania GPE , two CARDINAL point five times of Uganda GPE , seven CARDINAL times of Rwanda GPE and twenty
CARDINAL times of Burundi GPE . Even though country like Kenya GPE is one CARDINAL step ahead of Tanzania GPE in terms of GDP per
capita, Tanzania GPE had spent much in gross fixed capital formation compared to Kenya GPE and all other EAC members. And in terms of
governance Rwanda GPE is a head of its counterparts EAC members. Kenyans NORP are wealthier in EAC and Kenya GPE contributes more in
global economy in terms of exports and imports. <p>Key words: GDP OER, GDP PPP, EAC, Global economy, Governance
PERSON and Advanced.</p></jats:p>']

[' Chapter 14 LAW – Green GDP Accounting in China GPE : Controversies and Progress', '']

Recognised entities can form a metadata layer and stored in Dataverse

How to control Artificial Intelligence



Problem:

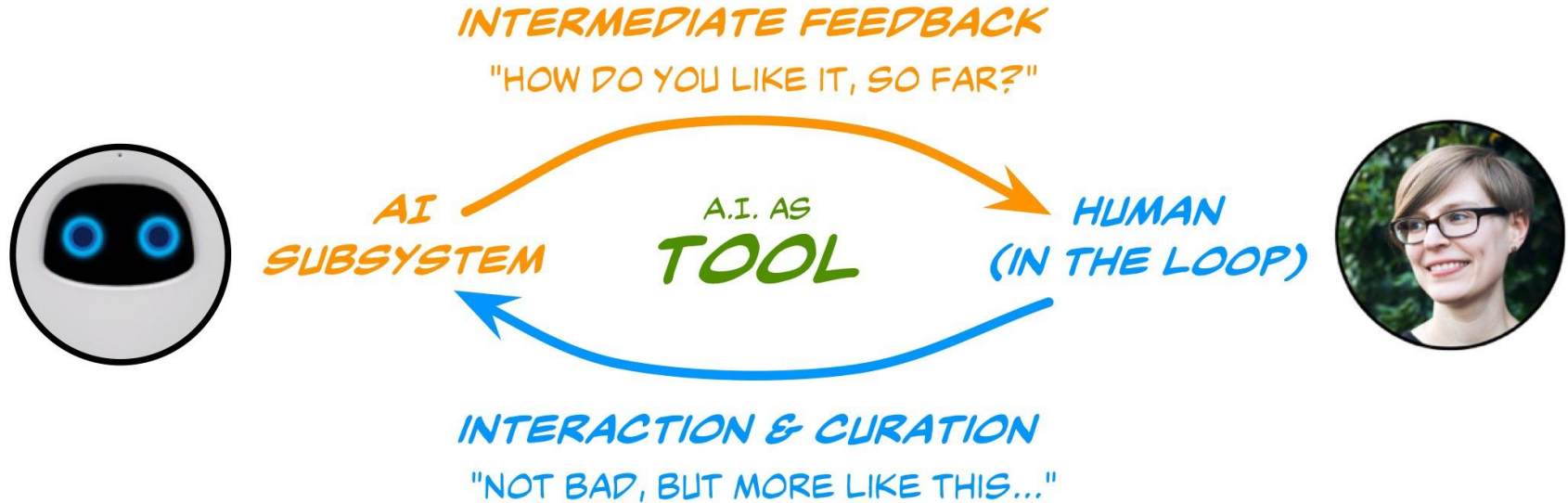
It's naive to fully trust Machine Learning and AI, we need to support a "human in the loop" processes to take a control over automatic workflows. Ethics is also important, fake detection problem.

Solution:

A lot of "human in the loop" tools already developed in research projects, we need to support the best for the different use cases, add the appropriate maturity, for example, with CI/CD and introduce them to research communities.


Human in the loop

“how do we build a smarter system?” to “how do we incorporate useful, meaningful human interaction into the system?”



General blueprint for a human-in-the-loop interactive AI system. Credits: [Stanford University HAI](#)

Semantic chatbot - ontology lookup service

 semanticbot

ABOUT

SUPPORT

about your dataset.
You can answer with the buttons in the bottom.
You can also check the progression of the process with the Toggle mapping button

Please choose a dataset from [Opendatasoft's data network](#)

coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata

I found some classes in your dataset, please check them to continue.

Field Category contains Category

Field Date contains Date

Field Location contains GeneLocation

Field Zone contains Place

I found some properties in your dataset, please check them to continue.

☐ Field **Category** is **category** of **something**

☐ Field **Date** is **date** of **something**

☐ Field **Location** is **Gene Location** of **gene**

Validate

Mapping

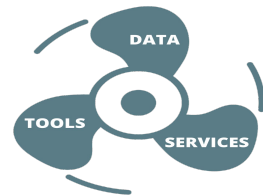
Schema

Get mapping

```
mappings:
  field-category:
    predicateobjects:
      - [a, 'http://www.w3.org/2006/vcard/ns#Category']
      - [a, 'http://www.w3.org/2006/vcard/ns#Explanatory']
      - ['http://www.w3.org/2000/01/rdf-schema#label', $(category)]
    source: dataset-source
    subject: https://data.opendatasoft.com/ld/resources/coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata/Category/$(category)
  field-date:
    predicateobjects:
      - [a, 'http://www.w3.org/2006/vcard/ns#Date']
      - [a, 'http://www.w3.org/2006/vcard/ns#RelatedType']
      - ['http://www.w3.org/2000/01/rdf-schema#label', $(date)]
    source: dataset-source
    subject: https://data.opendatasoft.com/ld/resources/coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata/Date/$(date)/
  field-location:
    predicateobjects:
      - [a, 'http://dbpedia.org/ontology/GeneLocation']
      - [a, 'http://www.w3.org/2002/07/owl#Thing']
      - ['http://www.w3.org/2000/01/rdf-schema#label', $(location)]
    source: dataset-source
    subject: https://data.opendatasoft.com/ld/resources/coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata/Location%20geo.
  field-zone:
    predicateobjects:
      - [a, 'http://schema.org/Place']
      - [a, 'http://purl.org/goodrelations/v1#Location']
      - [a, 'http://rdfs.co/juso/SpatialThing']
      - [a, 'http://schema.org/Thing']
      - ['http://www.w3.org/2000/01/rdf-schema#label', $(zone)]
    source: dataset-source
    subject: https://data.opendatasoft.com/ld/resources/coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata/Place/$(zone)/
  sources:
    dataset-source: [coronavirus-covid-19-pandemic-worldwide-data@bruxellesdata.json~jsonpath, '$.[*].fields']
```

Source: [Semantic Bot](#)

Hypothes.is annotations as a peer review service



Home Hypothesis Test Daily Progress Calendar

Ensuring global access to COVID-19 vaccines

AUTHORS: GAVIN YAMEY, MARCO SCHÄFERHOFF, RICHARD HATCHETT, MUHAMMAD PATE, FENG ZHAO, KACI KENNEDY

The current response to the coronavirus disease 2019 (COVID-19) pandemic involves aggressive implementation of suppression as case identification, quarantine and isolation, contact tracing, and social distancing. However, models developed by the Imperial College COVID-19 Response Team suggest that "transmission will quickly rebound if interventions are relaxed".¹ WHO warns of major outbreaks of COVID-19 worldwide.² The development of COVID-19 vaccines that can be used globally is therefore a priority pandemic.

This vaccine effort should be guided by three imperatives: speed, manufacture and deployment at scale, and global access. The World Bank and the Coalition for Epidemic Preparedness Innovations (CEPI), which funds development of epidemic vaccines, are conducting a global consultation on these goals.³ This consultation led to the launch of a COVID-19 Vaccine Development Taskforce that will look at how to finance and manufacture vaccines for global access.

CEPI estimates that developing up to three vaccines in the next 12–18 months will require an investment of at least US\$2 billion.⁴ This includes phase 1 clinical trials of eight vaccine candidates, progression of up to six candidates through phase 2 and 3 trials, regulatory and quality requirements for at least three vaccines, and enhancing global manufacturing capacity for three vaccines. CEPI does not include the costs of manufacture or delivery. Progress has been rapid. A phase 1 trial of a vaccine candidate, supported by the National Institutes of Health and CEPI, began on March 16, 2020,⁵ and 2 days later a clinical trial began in China.⁶ Clinical trials for three candidates will start soon.

Use of existing financing systems to support this work offers the benefits of speed and lower transaction costs than for new approaches. CEPI is supported by a World Bank financial intermediary fund that brings together public, philanthropic, and private capital to respond to global priorities.⁷ Through this fund, CEPI can act as a global mechanism for funding vaccine development until licensed or used under emergency use provisions. Mobilising \$2 billion in funding will require funding from all sources. Given the health, social, and economic consequences of COVID-19, there is a strong case for all governments to invest in vaccines.

Public

4tykhonov (edited May 7) May 7

Public

However, models developed by the Imperial College COVID-19 Response Team suggest that "transmission will quickly rebound if interventions are relaxed"

Transmission to normal life.

COVID-19 response

Imperial College COVID-19 Response Team

cchen1111 May 7

Public

the World Bank

Host of a global consultation on vaccine development

cchen1111 May 7

Public

12–18 months

time

cchen1111 (edited May 7) May 7

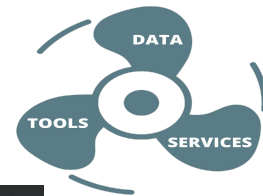
Public

US\$2 billion

cost of vaccine development

1. AI pipeline does domain specific entities extraction and ranking of relevant COVID-19 papers.
2. Automatic entities and statements will be added, important fragments should be highlighted.
3. Human annotators should verify results and validate all statements.

Doccano annotation with Machine Learning



doccano Edit Data Projects Logout

Search document

About 1 results (page 1 of 1)

Sort by

✓ An atypical RNA pseudoknot stimulator and an upstream attenu...

1/1

AMINO_ACID a CELL c CELLULAR_COMPONENT C-c CELL_LINE S-c

CELL_TYPE C-S-c CHEBI h CHEMICAL C-h CL l DISEASE d DNA C-d

GENE_OR_GENE_PRODUCT g GGP C-g GO S-g ORGANISM o PROTEIN p

RNA r SIMPLE_CHEMICAL s SO C-s TAXON t UMLS u

An atypical RNA pseudoknot stimulator and an upstream attenuation signal for

-1 ribosomal -1 ribosomal frameshifting of SARS SARS coronavirus. The -1

-1 ribosomal frameshifting requires the existence of an in cis RNA slippery sequence

cis RNA RNA slippery sequence and is promoted by a downstream stimulator RNA

An atypical RNA pseudoknot RNA pseudoknot pseudoknot with an extra stem formed by

complementary sequences complementary sequences within loop 2 of an

H-type pseudoknot is characterized in the severe acute respiratory syndrome coronavirus

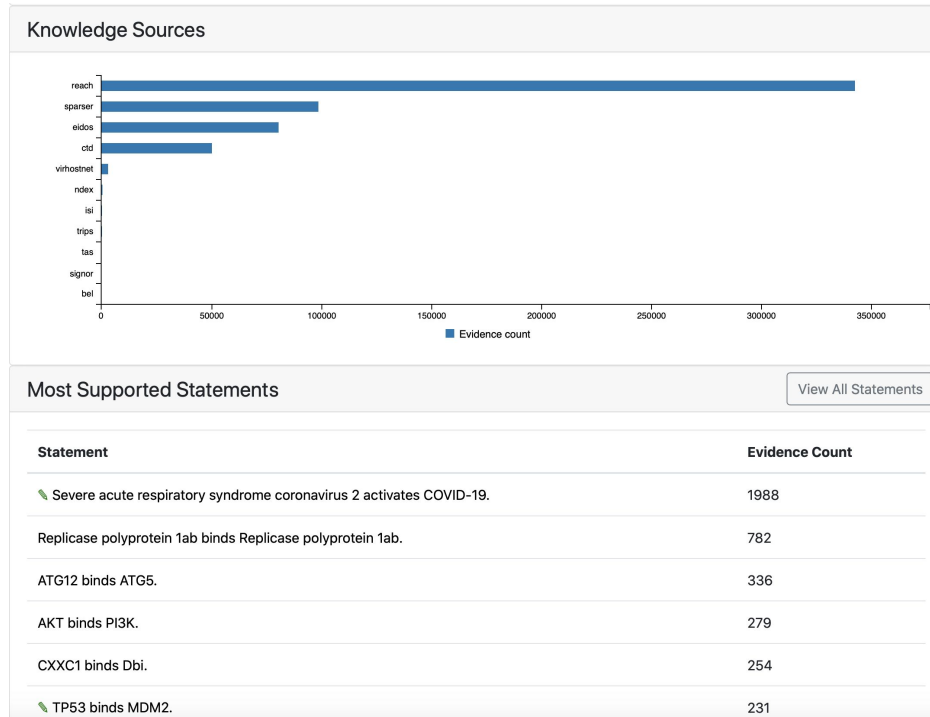
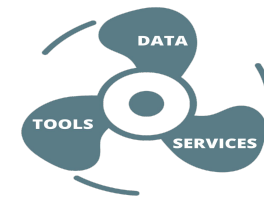
coronavirus (SARS CoV) genome. This pseudoknot can serve as an efficient

stimulator for -1 -1 frameshifting in vitro. Mutational analysis of the extra stem

suggests frameshift efficiency can be modulated via manipulation of the

Source: Doccano Labs

Statements extraction with INDRA

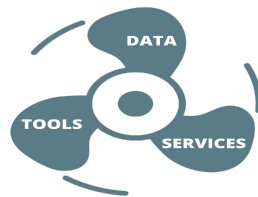


“INDRA (Integrated Network and Dynamical Reasoning Assembler) is an automated model assembly system, originally developed for molecular systems biology and currently being generalized to other domains.”

Developed as a part of Harvard Program in Therapeutic Science and the Laboratory of Systems Pharmacology at Harvard Medical School.

<http://indra.bio>

Knowledge Graph curation in INDRA



Effector	activates	transcription, DNA-templated.	0/50	JSON
Phenobarbital	increases the amount of	CYP3A4.	0/50	JSON
Oseltamivir	inhibits	Influenza, Human.	0/50	JSON
Arg-Val	activates	Asthma.	49/49	JSON
reach	It remains unknown if RV induces the development of wheeze and asthma or if asthma tics are more susceptible to RV infection.		18234348	
reach	Figure 3: A mouse model of RV -induced asthma exacerbation.		24278777	
	<div><div>✓ Correct</div><div>Entity Boundaries</div><div>Grounding</div><div>No Relation</div><div>Wrong Relation</div><div>Activity vs. Amount</div><div>Polarity</div><div>Negative Result</div><div>Hypothesis</div><div>Agent Conditions</div><div>Modification Site</div><div>Other...</div></div>	<div>Optional description (240 chars)</div> <div>Submit</div>		
reach	e of immune responses as well as differential regulation of different innate and adaptive has been implicated in the increased susceptibility of asthma tics to RV and in RV -		18234348	
reach	th acute RV -induced asthma Wark et al., found that increased serum IP-10 levels but or IL-8 was specifically associated with infection and correlated with the degree of		18234348	
reach	(b)A mouse model of RV -induced asthma exacerbation.		24278777	
reach	As RV is one the most common triggers of asthma exacerbations, it needs to be determined if blocking IL-4 and IL-13 could be useful in preventing experimental RV -induced exacerbation of asthma [24,50].		27088397	
reach	Some investigators have reported that RV -C caused more serious illness, especially wheezing and exacerbation of asthma , in some populations5, 9, 10, 11, 12, 13 compared with illnesses caused by RV -A and B.		31389049	

Building domain specific knowledge graphs

- We're collecting all possible COVID-19 data and archiving in our Dataverse
- Looking for various related controlled vocabularies and ontologies
- Building and reusing conversion pipelines to get all data values linked in RDF format

```
<http://www.protezionecivile.gov.it/dataset/covid19/national-trend/observations/20200224> a qb:Observation ;  
  rdfs:label "Andamento nazionale di giorno 24/02/2020"^^xsd:string ;  
  qb:dataset <http://www.protezionecivile.gov.it/dataset/covid19/national-trend> ;  
  sdmx-dimension:refArea <https://w3id.org/italia/controlled-vocabulary/territorial-classifications/countries/italy/ITA> ;  
  sdmx-dimension:refTime "2020-02-24"^^xsd:date ;  
  dpc:deads "7"^^xsd:int ;  
  dpc:healed "1"^^xsd:int ;  
  dpc:homeIsolation "94"^^xsd:int ;  
  dpc:hospitalizedWithSymptoms "101"^^xsd:int ;  
  dpc:intensiveCare "26"^^xsd:int ;  
  dpc:newPositive "221"^^xsd:int ;  
  dpc:swabs "4324"^^xsd:int ;  
  dpc:totalCases "229"^^xsd:int ;  
  dpc:totalHospitalized "127"^^xsd:int ;  
  dpc:totalPositive "221"^^xsd:int ;  
  dpc:totalPositiveVariation "0"^^xsd:int .
```

The ultimate goal is to automate the process of the Knowledge extraction by using the latest developments in Artificial Intelligence and Deep Learning. We need to organize all available knowledge in a common way.

CLARIAH conclusions



*"By developing these decentralised, yet controlled Knowledge Graph development practices we have contributed to increasing interoperability in the humanities and enabling new research opportunities to a wide range of scholars. However, **we observe that users without Semantic Web knowledge find these technologies hard to use, and place high value in end-user tools that enable engagement.** Therefore, for the future we emphasise the importance of tools to specifically target the goals of concrete communities – in our case, the analytical and quantitative answering of humanities research questions for humanities scholars. In this sense, usability is not just important in a tool context; in our view, we need to empower users in deciding under what models these tools operate." ([CLARIAH: Enabling Interoperability Between Humanities Disciplines with Ontologies](#))*

Chicken-egg problem: users are building tools without data models and ontologies but in reality they need to build a knowledge graph with common ontologies first!

Linked Data integration challenges



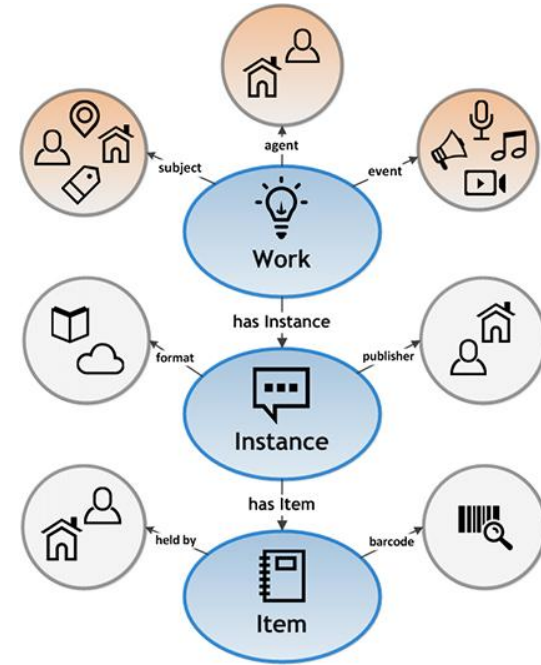
Word Wide Web Consortium

- datasets are very heterogeneous and multilingual
- data usually lacks sufficient data quality control
- data providers using different modeling schemas and styles
- linked data cleansing and versioning is very difficult to track and maintain properly, web resources aren't persistent
- even modern data repositories providing only metadata records describing data without giving access to individual data items stored in files
- difficult to assign and manually keep up-to-date entity relationships in knowledge graph

CoronaWhy has too much information streams that seems to be impossible to integrate and give back to COVID-19 researchers. So, do we have a solution?


BIBFRAME 2.0 concepts

- **Work.** The highest level of abstraction, a Work, in the BIBFRAME context, reflects the conceptual essence of the cataloged resource: authors, languages, and what it is about (subjects).
- **Instance.** A Work may have one or more individual, material embodiments, for example, a particular published form. These are Instances of the Work. An Instance reflects information such as its publisher, place and date of publication, and format.
- **Item.** An item is an actual copy (physical or electronic) of an Instance. It reflects information such as its location (physical or virtual), shelf mark, and barcode.
- **Agents:** Agents are people, organizations, jurisdictions, etc., associated with a Work or Instance through roles such as author, editor, artist, photographer, composer, illustrator, etc.
- **Subjects:** A Work might be “about” one or more concepts. Such a concept is said to be a “subject” of the Work. Concepts that may be subjects include topics, places, temporal expressions, events, works, instances, items, agents, etc.
- **Events:** Occurrences, the recording of which may be the content of a Work.



Source: [the Library of Congress](#)

CoronaWhy COVID-19 Portal

 CORONAWHY

[Home](#) [Daily Progress](#) [Solutions](#) [Data](#) [Team](#) [JOIN THE FIGHT!](#)

[Login](#) [Language](#)

All Fields Find [Advanced](#)

[Reset Filters](#) Language: English

Search Results

Suggested Topics within your search.


[coronavirus](#) 203,451 [Medicine](#) 172,412 [COVID-19](#) 133,230 [C1880229](#) 4,507 [DICOM Study](#) 4,507

[C0086418](#) 4,317 [Homo sapiens](#) 4,317 [more ...](#)

Showing 1 - 20 results of 203,461 for search "", query time: 0.43s


Sort Relevance

1


NO IMAGE AVAILABLE


[OnabotulinumtoxinA infiltration and nerve blocks in patients with headache and neuralgia: safety recommendations to prevent SARS-CoV-2 infection](#) [★ Save to List](#)
by Santos-Lasaosa, S., Porta-Etessam, J.
Published 2020
[Get full text](#)
[Book](#)

2


NO IMAGE AVAILABLE


[Financial markets under the global pandemic of COVID-19](#) [★ Save to List](#)
by Zhang, Dayong, Hu, Min, Ji, Qiang
Published 2020
[Get full text](#)
[Book](#)

3


NO IMAGE AVAILABLE

[Hepatic manifestations and impact of COVID-19 on the cirrhotic patient](#) [★ Save to List](#)
by Velarde-Ruiz Velasco, J. A., García-Jiménez, E. S., Remes-Troche, J. M.
Published 2020
[Get full text](#)
[Book](#)

4


NO IMAGE AVAILABLE

[SARS-CoV-2 pandemic on the activity and professionals of a General Surgery and Digestive Surgery Service in a tertiary hospital](#) [★ Save to List](#)
by Álvarez Gallego, M., Gortázar de las Casas, S., Pascual Migueláñez, I., Rubio-

Narrow Search

Institution ▲
[MyInstitution](#) 203,461


Library ▲
[Library A](#) 203,461

Format ▲
[Serial](#) 133,229
[Book](#) 70,222
[Journal](#) 10
[Microfilm](#) 3

Call Number ▲
[A - General Works](#) 10

Author ▲
[CoronaWhy Labs Labs](#) 102,036

Landing page of publications from CORON-19




[Home](#) [Daily Progress](#) [Solutions](#) [Data](#) [Team](#) [JOIN THE FIGHT!](#)

All Fields [Find](#) [Advanced](#)

[Reset Filters](#) Language: [English](#) Suggested Topics: [C1880229](#)

[Search](#) / [The Transformation of Enterovirus](#) / Holdings

[Cite this](#) [Text this](#) [Email this](#) [Print](#) [Export Record](#) [Save to List](#)



The Transformation of Enterovirus Replication Structures: a Three-Dimensional Study of Single- and Double-Membrane Compartments


All positive-strand RNA viruses induce membrane structures in their host cells which are thought to serve as suitable microenvironments for viral RNA synthesis. The structures induced by enteroviruses, which are members of the family Picornaviridae, have so far been described as either single- or do...


[Full description](#)


Main Authors:	Limpens, Ronald W. A. L., van der Schaar, Hilde M., Kumar, Darshan, Koster, Abraham J., Snijder, Eric J., van Kuppeveld, Frank J. M., Bárcena, Montserrat
Corporate Authors:	Leiden University Medical Center, Radboud University Nijmegen Medical Centre
Format:	Serial
Language:	English
Published:	mBio 2011
Subjects:	Act Relationship Type - transformation > C1554215 Three-dimensional > C0450363 Induce (action) > C0205263 Membrane > C0596901 Cells > C0007634 Cell Microenvironment > C1707328 RNA, Viral > C0035736 Anabolism > C0220781

Similar Items

 [SARS-Coronavirus Replication Is Supported by a Reticulovesicular Network of Modified Endoplasmic Reticulum.](#)
by: Knoop, Kévin, et al.
Published: (2008)

 [Three-Dimensional Architecture and Biogenesis of Membrane Structures Associated with Hepatitis C Virus Replication](#)
by: Romero-Brey, Inés, et al.
Published: (2012)

 [Packaging of Genomic RNA in Positive-Sense Single-Stranded RNA Viruses: A Complex Story.](#)
by: Comas-García, Mauricio
Published: (2019)

 [Role of Cellular Lipids in Positive-Sense RNA Virus Replication Complex Assembly and Function.](#)
by: Stapleford, Kenneth A., et al.
Published: (2010)

Source: [CoronaWhy Portal - VuFind page](#)

Raw structured data extracted from the page

id	http://vufind.labs.coronawhy.org/vufind/Record/o32e9o01(0)						
rdfs:usesVocabulary	http://schema.org/						
id	http://schema.org/CreativeWork(1)						
rdfs:type	rdfs:Class						
id	http://schema.org/Language(1)						
rdfs:type	rdfs:Class						
id	http://schema.org/Organization(1)						
rdfs:type	rdfs:Class						
id	http://schema.org/Product(1)						
rdfs:type	rdfs:Class						
id	http://vufind.labs.coronawhy.org/vufind/Record/o32e9o01#export-options(0)						
role	http://www.w3.org/1999/xhtml/vocab#menu						
id	http://vufind.labs.coronawhy.org/vufind/Record/o32e9o01#modal(0)						
role	http://www.w3.org/1999/xhtml/vocab#dialog						
id	http://vufind.labs.coronawhy.org/vufind/Record/o32e9o01#record(0)						
rdfs:type	schema:CreativeWork schema:Product schema:Thing rdfs:Resource						
rdfs:label	The Transformation of Enterovirus Replication Structures: a Three-Dimensional Study of Single- and Double-Membrane Compartments						
schema:author	<ul style="list-style-type: none">• Limpens, Ronald W. A. L.• Kumar, Darshan• Koster, Abraham J.• van der Schaar, Hilde M.• van Kuppeveld, Frank J. M.• Snijder, Eric J.• Bárcena, Montserrat						
schema:availableLanguage	<table><tr><td>rdfs:type</td><td>schema:Intangible schema:Thing rdfs:Resource schema:Language</td></tr><tr><td>rdfs:label</td><td>English</td></tr><tr><td>schema:name</td><td>English</td></tr></table>	rdfs:type	schema:Intangible schema:Thing rdfs:Resource schema:Language	rdfs:label	English	schema:name	English
rdfs:type	schema:Intangible schema:Thing rdfs:Resource schema:Language						
rdfs:label	English						
schema:name	English						
schema:creator	<ul style="list-style-type: none">• Radboud University Nijmegen Medical Centre• Leiden University Medical Center						
schema:datePublished	2011						
schema:keywords	<ul style="list-style-type: none">• Induce (action) > C0205263• Coxsackievirus Infections > C0010246• Ultrastructure > C0041623						

Source: [Structured Data Liner](#)

Thank you! Questions?

Slava Tykhonov

DANS-KNAW

vyacheslav.tykhonov@dans.knaw.nl