

# COVID-19 Explorer Documentation

## COVID-19 iKnow Content Navigator - Documentation for end users

### Background information

This explorer for the [COVID-19 data set](#) (CORD-19, 2020) is built on the InterSystems' iKnow NLP technology. The iKnow technology analyses written text and identifies the "Concepts" and the "Relations" between those Concepts. Concepts and Relations are called "Entities". More information about the technology can be found in the wiki session of our GitHub repository: <https://github.com/interSystems/iknow/wiki>

### Using the iKnow Content Navigator

#### Start the exploration

The iKnow Content Navigator is not a traditional search tool. It enables you to explore what is in the data set even if you don't know what you are looking for. You can start diving into the data by selecting an entity from the Top Entities or - if you do know what you are looking for - you can start typing in the exploration box in the top left corner. The Navigator will start suggesting actual entities from the data set that match what you are typing so you can explore even faster. Then click "Explore!".

The screenshot shows the COVID-19 Explorer interface with the following components and annotations:

- Domain Explorer COVID-19 Body Text**: The main header of the application.
- Search Bar**: Located at the top left, with an annotation: "Considered entity – selected from a list or your own 'search term'".
- Explore! Button**: An annotation points to it: "Click here to show results".
- Filters**: Annotations point to filter icons: "Click to use the explored entity as a filter" and "Show and add filters".
- Reset Button**: An annotation points to it: "Reset (removes all filters)".
- Top Entities Table**: An annotation points to it: "List of the most frequent or dominant entities". It has tabs for "frequency" and "dominance". An annotation points to the first column: "First column: frequency, second column: spread".
- Similar Entities Table**: An annotation points to it: "Click to use the explored entity as a filter".
- Related Entities Table**: An annotation points to it: "Toggle between 'related' and 'proximity'". It has tabs for "related" and "proximity". An annotation points to it: "Entities that occur in the context of the explored entity".
- Articles Tab**: An annotation points to it: "Scroll through pages of results".
- Articles View**: An annotation points to it: "Click to open document view".
- Terms in italics are explained below.**: A note at the bottom of the interface.

**Frequency:** the number of occurrences of the entity in the data set

**Dominance:** a metric expressing entity relevance within the data set

**Spread:** the number of distinct documents in which the explored entity appears

**Filter:** a means to limit the number of articles that is taken into consideration. See below for details.

**Related** versus **proximity**: "related" shows all other concepts the entity being explored is directly related to (with a single relational element between). The frequency and spread numbers now express how often the concepts appear in a direct relationship. If you toggle to "proximity" in the title bar, the relationship can now also be an indirect one, with the concepts appearing at a longer distance within the same sentence.

**How to explore the results:** The full-width widget below will show the **articles** (documents), **paths** (sentence subsections), **CRCs** (concept-relation-concept triples) and **CCs** (concept-related concept) in which the entity appears, with the seed entity highlighted in green. In the Articles view, text fragments containing the seed entity are displayed, with negative sections marked in red.

**Document view:** Pop-up window with additional functionality. See below for details.

## Explore the context

When looking at sources or paths in the main bottom widget, you can click the eye icon to consult the complete source text and optionally highlight indexing results. Using the controls at the bottom, you can also look up the source's metadata and generate summaries of the text. In other words, after initial exploration points you to a couple of documents, you can quickly find the original context and read a summary of the entire document.

The screenshot shows a document viewer interface. At the top, the document title is "Articles Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study". Below the title is the document text, which includes the sentence: "Data using the same terms in Chinese, with no time restrictions. We found only previously published research discussing the radiological characteristics of severe acute respiratory syndrome coronavirus or Middle East respiratory syndrome coronavirus, and preliminary studies on the evolution of SARS-CoV-2 in Wuhan. We found no published work about the radiological characteristics of patients with pneumonia caused by SARS-CoV-2 infection. We analysed chest CT images from 81 patients confirmed to have COVID-19 pneumonia caused by SARS-CoV-2 infection, and described the imaging findings across different follow-up periods of the disease, including subclinical and clinical stages up to 3 weeks after symptom onset. This study is, to the best of our knowledge, the largest case series to date on the CT findings of COVID-19 pneumonia worldwide, and provides detailed information on the radiological, clinical, and laboratory features of these patients. Our findings provide insight in the evolution of the disease and its corresponding imaging changes, and suggest potential associations between imaging findings and patient outcomes, which requires further study. Implications of all the available evidence SARS-CoV-2 has substantial capacity to infect humans, and is capable of human-to-human transmission. Infection can lead to severe respiratory symptoms, including acute respiratory distress syndrome, and the mortality rate is around 2%. Combining imaging assessments with clinical and laboratory findings could help to identify SARS-CoV-2 infections early. CT imaging features recorded from our cohort are".

At the bottom of the document text, there are four buttons: "Original paper", "similar", "indexing results", and "100 %".


Annotations point to the following elements:


- Document title**: Points to the title of the document.
- Document text with highlighting of the explored entity, negation, measurements and time indications**: Points to the text content, specifically highlighting "pneumonia", "no time restrictions", "previously published", "no published work about the radiological characteristics of patients with pneumonia", "from 81 patients", "3 weeks after symptom onset", and "100 %".
- Link to URL of The Allen Institute for AI**: Points to the "Original paper" button.
- List similar documents**: Points to the "similar" button.
- Toggle between this view and the fully iKnow-indexed document**: Points to the "indexing results" button.
- Create and show a x% summary of the document**: Points to the "100 %" button.

## Focus by filtering

The navigator contains three types of filters.

1. Each entity, CRC or CC can be used to filter the results, i.e. to limit the exploration to the articles that contain the selected entity, CRC or CC. To

define an entity, CRC or CC as a filter, just click the  button next to the element in the navigator. The filter button in the upper right corner

will turn green: .

2. The navigator also contains a number of predefined terms - called "sets" - that can be used as filters.
3. (Moreover, the metadata PaperID, Source and Title can be added as filters. - not available yet)

All types of filters can be managed through the dialogue box that opens if you click the filter button. The entities, CRCs and CCs you selected will appear automatically in the upper part. Use the dropdown menus and click the "add" buttons to define additional filters. Activate them by clicking "Apply".

Filters

Entity: virus

add filter here

✕

AND

Set: covid-19

add filter here

✕

add filter here

Predefined set:

transmission

▼

add

Metadata filter:

PaperID

▼

=

▼

add

Clear

Apply

The filter icon will remain highlighted as long as any filter conditions are active and the number of sources that are being explored is updated to reflect the filter conditions.

Reference for the COVID-19 dataset:

COVID-19 Open Research Dataset (CORD-19). 2020. Version 2020-03-20. Retrieved from <https://pages.semanticscholar.org/coronavirus-research>. Accessed 2020-03-27. doi:10.5281/zenodo.3715505