

UW-COSMOS



Shanan Peters, Miron Livny, Shivaram Venkataraman
Theodoros Rekatsinas

Ian Ross, Daven Quinn, Iain McConnell

ASKE-E UW-COSMOS Tasks

- **Task 1: Scale and Enhance xDD Infrastructure and APIs**
 - **Task 1.A. Objective: Expand Corpus**
 - **Task 1.B. Objective: Develop Container Template for Collaborator Code**
 - **Task 1.C. Objective: Augment xDD Corpus with Container Outputs**
 - **Task 1.D. Objective: Improve and Scale xDD API** (and add other features)
- **Task2: Scale and Enhance COSMOS Retrieval and API**
 - **Task 2.A. Improve Visual Segmentation** (preprints are challenging, need ML method)
 - **Task 2.B. Incorporate Body-Text Content into Table and Figure Retrieval**
 - **Task 2.C. Automatic Knowledge Base Construction** (Marius, table extraction)
 - **Task 2.D. Release Public* COSMOS API Over COVID-19 xDD Set**

xDD and UW-COSMOS data infrastructure overview

Text-based search and retrieval over full text
of all 14M full texts in xDD corpus

xdd.wisc.edu/api

- [api/snippets](#): string-based search over document full-text to retrieve document metadata and contextual snippets highlighting search matches.
- [api/articles](#): similar to snippets, but article metadata focused
- [api/journals](#): search for specific journals to assess coverage
- [api/](#): additional routes and documentation

continuous
updating

xDD
API

xDD
sets

Subsets of xDD corpus for specific
projects: xdd.wisc.edu/sets

periodic updating

sets/xdd-covid-19

154,777 full-text docs

xDD-defined COVID19 doc set (CORD19,
EMMAA, search-term defined): full content only

restrict API results to set:
&dataset=xdd-covid-19

Table, figure, equation, retrieval over sets
via COSMOS AI assistant (cosmos.wisc.edu)

xdd.wisc.edu/sets/xdd-covid-19/cosmos/api/

- [search](#): string-based search over document figures, tables, equations based on associated text.
- other routes available, most *require API key*

xdd.wisc.edu/sets/xdd-covid-19/word2vec/api/

- trained embeddings
- several vector operations

xdd.wisc.edu/sets/xdd-covid-19/doc2vec/api/

- document-level embeddings, doc discovery based on similarity

ASKE-
ID

UW
COSMOS
API

word2vec
API

doc2vec
API

- service for
generating/
retrieving
metadata and
links to artifacts



xDD automated
full-pub fetching
10³s per day

xDD corpus
14M full texts
+
PMC abstracts

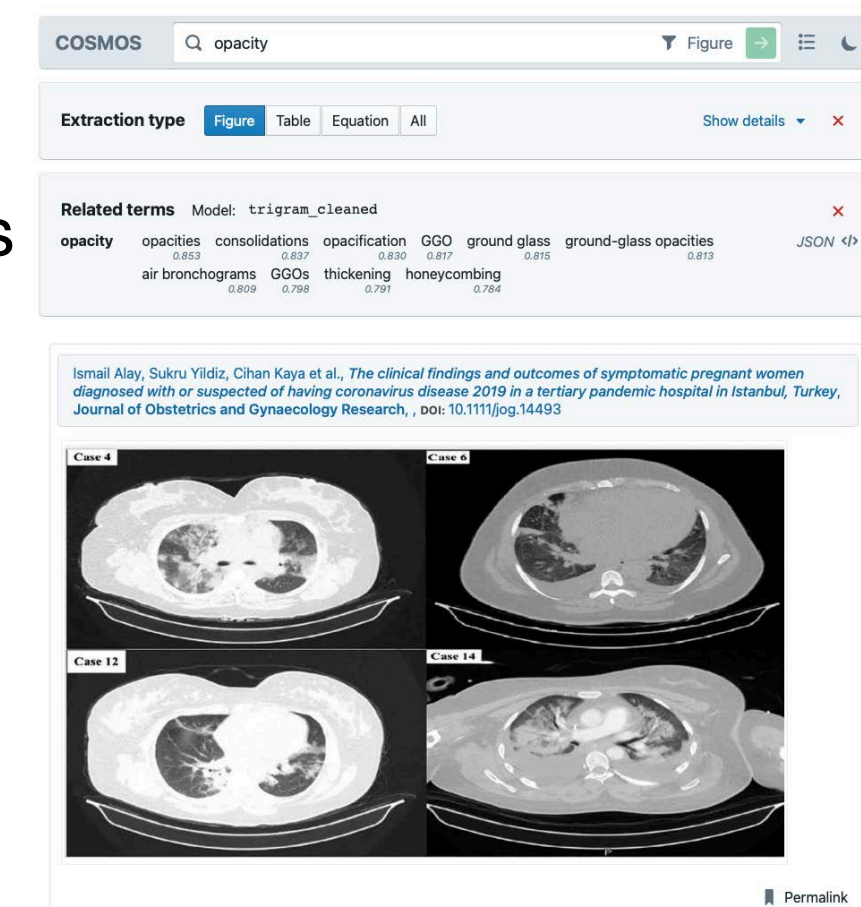


open access
full-pub

abstract-only
text



Major support by DARPA ASKE, DOE, USGS
2014-2018 by NSF-ICER 1343760



Task 1:

xDD Infrastructure Upgrades

xDD Infrastructure Upgrades - Overview

- Kubernetes
- Introduced sets
- Upgrades and updates
- Pubmed abstracts
- API updates
 - Corpus annotation from user containers (known entities)

Kubernetes

Moved many services from bare-metal to k8s

- Benefits
 - Stability and consistency
 - Flexibility - can scale services, easily modify setups and environments, continually deploy updates/upgrades
 - Some services run on xDD-specific hardware, other on shared hardware within the cluster
 - Decoupling - Only allow network traffic between things that **need** to communicate, services run on multiple hosts
 - Centralized config - resource definitions are stored in git and anybody with access can read it and deploy changes
- Downsides
 - Took time to learn and transition
 - Fast-moving ecosystem - can become dependent on a third-party resource that becomes abandonware

Sets

Collections of documents within xDD to enable knowledge discovery over vertical application stacks

- Defined by:
 - Text content query
 - Metadata query (journal name, publisher, publication year, etc)
 - DOI list
 - Combinations of the above
- Enables:
 - Filtered searches in standard xDD endpoints (e.g. restricting text snippets searches to relevant documents)
 - Transformations and extractions via software tools (e.g. COSMOS processing, word2vec training, user-provided applications)
 - Available as static files or webapp endpoint

Sets - Example

xdd-covid-19

- Definition: Union of CORD-19 (via DOI matching) and internally-derived keyword-based document sets
- Enabled:
 - https://xdd.wisc.edu/api/articles?full_results=true&dataset=xdd-covid-19 - metadata of all matching documents
 - https://xdd.wisc.edu/api/snippets?term=ACE2&full_results=true&dataset=xdd-covid-19 - contextual usage of "ACE2" within matching documents
 - <https://xdd.wisc.edu/sets/xdd-covid-19/> - Basic stats and available products
 - <https://xdd.wisc.edu/sets/xdd-covid-19/cosmos/> - API to COSMOS results
 - <https://xdd.wisc.edu/sets/xdd-covid-19/word2vec/> - API to word2vec embeddings
 - <https://xdd.wisc.edu/sets/xdd-covid-19/doc2vec/api/> - API to doc2vec embeddings

Upgrades and updates

- New dedicated hardware
- Upgraded all Centos6 machines to RHEL7
- Software version updates
 - mongodb, postgresql, python

Pubmed Abstracts

- Now a first-class corpus within xDD
- Allows searching of article metadata and contextual snippets from PM abstracts with the same interface as the primary fulltext corpus, simultaneously (default) or separately:
 - [https://xdd.wisc.edu/api/articles?
term=remdesivir&corpus=fulltext&max=10](https://xdd.wisc.edu/api/articles?term=remdesivir&corpus=fulltext&max=10)
 - [https://xdd.wisc.edu/api/articles?
term=remdesivir&corpus=pubmed abstracts&max=10](https://xdd.wisc.edu/api/articles?term=remdesivir&corpus=pubmed%20abstracts&max=10)
- Also enables abstracts as input into xDD transformation/extraction pipelines (NLP, word/document embeddings)

Additional xDD API updates

Many updates, most requested from ASKE collaborators, examples:

- ASKE-ID as parameter to retrieve ingested ASKE-ID docs
- Inclusive parameter to require multiple search times to all be present
- Journal name filter on /snippets route
- known_terms - Recall xDD-known dictionary terms for each document returned
- Require additional terms at the document level for /snippets results
- Other bug fixes and minor updates

API updates - known_entities

- Method to incorporate knowledge (third-party or xDD-derived) into document-level returns within xDD API
 - external links when applicable
- Entities, objects, relationships, ...
- Examples:
 - EMMAA statements, linking back into HMS EMMAA model
 - Drug mentions, from a list provided with MITRE with their internal IDs
 - "How is *phosphatidylethanolamine* talked about within the corpus and what drugs, with IDs, are mentioned in these documents?"
 - [https://xdd.wisc.edu/api/snippets?
term=phosphatidylethanolamine&known_entities=drugs](https://xdd.wisc.edu/api/snippets?term=phosphatidylethanolamine&known_entities=drugs)

Docker container template

<https://github.com/UW-xDD/xdd-docker-recipe>

- GitHub-hosted blueprint for producing a working xDD application
- ASKE-E implementation limited to HMS reader deployed over xdd-covid-19 document set
 - ~1M relevant documents processed for grounded INDRA statements, returned to HMS EMMAA model
 - HMS has started work on Table/Figure contributions for evidence using COSMOS API

Task 2: Scale and Enhance COSMOS Retrieval and API

COSMOSView training dataUsagePermalinkNext image >

Page Header

4P. LI ET AL.

Table Caption

Table Effects of vitamin B6 on plasma AChE activity and PLP concentrations in rats fed with isocarbophos.

	AChE activity (nmol/min/ml)	PLP concentration (nmol/L)	sBP (mmHg)	dBp (mmHg)
Saline	2.11 ± 0.37	20.4 ± 5.3	128 ± 21	76 ± 25
Vitamin B6	2.16 ± 0.36	28.9 ± 2.4*	124 ± 20	72 ± 13
Eliprodil	2.09 ± 0.29	20.1 ± 4.7	130 ± 17	86 ± 9
Isocarbophos	1.98 ± 0.30	22.9 ± 3.5	137 ± 29	84 ± 10
Isocarbophos + vitamin B6	2.07 ± 0.31	27.6 ± 3.7#	126 ± 23	77 ± 13
Isocarbophos + vitamin B6 + Eliprodil	2.15 ± 0.27	26.1 ± 5.1#	138 ± 15	79 ± 17

Table Note

After isocarbophos treatment, blood pressure was determined by the tail cuff method. Blood was collected to measure the levels of plasma AChE activity and PLP concentrations in rats. Data were expressed as means ± SD. N was 6 in each group. One-way ANOVA was used for statistical analyses. **p* < 0.05 VS Saline. #*p* < 0.05 VS Isocarbophos.

Table Caption

Table Effects of vitamin B6 on the blood flow of posterior cerebral artery (PCA) in rats fed with isocarbophos.

Groups	Right PCA		Left PCA	
	V _s (cm/s)	V _d (cm/s)	V _s (cm/s)	V _d (cm/s)
Saline	26.32 ± 4.21	12.35 ± 2.41	26.85 ± 4.66	13.54 ± 2.55
Vitamin B6	24.63 ± 5.39	11.26 ± 3.41	27.52 ± 5.98	12.54 ± 3.63
Eliprodil	21.58 ± 4.26	9.66 ± 1.58	20.51 ± 3.64*	10.49 ± 2.53
Isocarbophos	15.74 ± 3.18*	8.57 ± 1.36*	17.41 ± 3.15*	8.26 ± 1.64*
Isocarbophos + vitamin B6	26.79 ± 3.54 [#]	13.66 ± 2.08 [#]	26.12 ± 4.66 [#]	13.52 ± 2.36 [#]
Isocarbophos + vitamin B6 + Eliprodil	15.25 ± 3.64 [§]	7.65 ± 1.54 [§]	15.88 ± 3.32 [§]	7.24 ± 1.62 [§]

Table Note

After isocarbophos treatment, blood flow of PCA was determined by TCD before sacrifice. Data were expressed as means ± SD. One-way ANOVA was used for statistical analyses. N was 6 in each group. **p* < 0.05 VS Saline. [#]*p* < 0.05 VS Isocarbophos. [§]*p* < 0.05 VS Isocarbophos + vitamin B6.

Body Text

Supplementation of vitamin B6 significantly recovered the blood flow of PCAs in isocarbophos-treated rats. Furthermore, the addition of eliprodil, which is a selective NR2B antagonist (27), abolished the effects of vitamin B6 on the blood flow of PCAs. There were no obvious differences of vascular thickness among groups as determined by HE staining (Figure 1B). Taking these data together, it can be suggested that vitamin B6 via NR2B activation preserved the cerebral blood flow in isocarbophos-treated rats.

Section Header

Vitamin B6 improves isocarbophos-impaired learning and memory, which is suppressed by eliprodil in rats

Body Text

The typical phenotype of dementia is the impairment of the spatial memory (28,29). Thus, the learning and memory behaviors were evaluated by the MWM test. As shown in Table 3, after 12-week isocarbophos administration, swimming distance in quadrant, ratio of swimming distance in quadrant to total swimming distance, swimming time in quadrant, swimming time in quadrant to total swimming time, and number of escape were increased in isocarbophos-treated rats, compared with saline-treated rats. Conversely, swimming time and distance out of the quadrant were increased in isocarbophos-treated rats. However, the supplementation of vitamin B6 significantly normalized these alterations in isocarbophos-treated rats, which were ablated by eliprodil.

Body Text

vitamin B6 on the protein levels of NR2B and PSD95 in hippocampal PSDs in isocarbophos-treated rats. As depicted in Figure 2A–2C, 12-week treatment of isocarbophos dramatically decreased the protein levels of NR2B and PSD95 but not NR2A, compared with saline-treated rats. Although this phenotype induced by isocarbophos was reversed by vitamin B6, co-treatment of eliprodil reduced the levels of both NR2B and PSD95. Interestingly, eliprodil alone lightly reduced both levels of NR2B and PSD95 in rat hippocampus.

Section Header

Vitamin B6 increases CaMK-II expression in rat hippocampus

Body Text

CaMK-II has been demonstrated to potentiate NMDA receptor function by increasing the fraction of receptors in the synaptosomal membranes (32). Thus, we investigated whether isocarbophos altered the CaMK-II protein expression in the hippocampus PSDs. As shown in Figure 3A and 3B, compared with the saline group, both CaMK-II protein level and phosphorylated (p-CaMK-II) decreased in the hippocampus PSDs of the isocarbophos group. Similarly, the supplementation of vitamin B6 increased the CaMK-II protein level in isocarbophos-treated rats, which was ablated by eliprodil. Together, it reveals that vitamin B6 might activate NR2B/PSD95/CaMK-II signaling in hippocampal PSDs.

Downloaded by [University of Wisconsin - Madison] at 16:03 04 January 2018

COSMOS Updates - Pipeline - Overview

Major updates in the last year

- Major refactor (late 2020)
- New model training (February 2021)
- Table context enrichment (August 2021)
- Additional fixes and improvements
 - Entity linking module added via SciSpacy
 - Text normalization

Major refactor

Late 2020

- Vastly improved modularization (preprocess, ingest, detect, postprocess, retrieval)
- Command-line interface
- Canonical docker images, docker-compose orchestration files and suggested environment definitions
 - GPU or CPU-based
- DASK for task distribution across workers
- Default to unicode extracted directly from PDF (OCR still available as fallback)
- Retrieval options: Elasticsearch query, BERT reranking

New model training

Februrary 2021

- New weight files for object detection
 - Major new annotation push
- Minor updates to object segmentation
- Updates to rule-based post-processing
- Also included improved documentation and fixes re: training new models

Table context enrichment

Initially added in early 2021; complete in August 2021

- Links tables to mentions within text, supplying additional textual information to aid recall
- Example:
 - With enrichment, a query of "germination rate" would include Table 1 due to the inclusion of "rate" in the associated (orange) body text

Table 1. Germination of *Tilletia tritici*, *Tilletia horrida*, and *Tilletia indica* teliospores after exposing teliospores in sori or loose teliospores (dry or on agar) to volatiles of *Muscodor albus* for 2 or 5 days.

Treatment	<i>Tilletia tritici</i>			<i>Tilletia horrida</i>		<i>Tilletia indica</i>	
	exp I	exp II	exp III	exp I	exp II	exp I	exp II
Nonfumigated control	87.5	86.5	78.0	37.5	25.0	+	+
Teliospores in sori fumigated 5 days	76.0	84.0	84.0	38.5	24.0	+	+
Dry loose teliospores fumigated 5 days	2.5	10.0	24.5	43.0	18.0	+	+
Teliospores on agar fumigated for 2 days*	0.7	3.5	21.5	16.3	19.0	+	+
Teliospores on agar fumigated for 5 days*	0	0	0	0	0	–	–

Note: Mean percent germination of 100–500 spores on each of 2 plates. Zero percentages based on observation of >5000 spores. *Tilletia indica* was rated as positive (+) or negative (–). Counts were done up to 3–5 weeks after plating teliospores.
*Fumigation began immediately after plating spores.

Germination of *T. tritici* and *T. horrida* teliospores was assessed periodically after fumigation by microscopic examination of 100–500 spores in each of the 2 plates, and then averaging the percent germinated. *Tilletia indica* plates were rated as either positive or negative for germination because of the limited number of opportunities to assess them in Arizona and because of the growth of mycelia in plates, making actual counts of germination difficult. Assessment of teliospore germination continued for 3–5 weeks after plating teliospores. In plates of all 3 species, where no germination occurred, >5000 spores were examined.

Teliospores in nonfumigated control plates germinated normally ~57 days after plating. Teliospores of *T. tritici* commonly germinate at high rates, whereas those of *T. indica* and *T. horrida* normally germinate at lower rates as is reflected in the data (Table 1). Teliospores of all 3 species that were plated on agar were killed completely when fumigated for 5 days as evidenced by the failure of teliospore

Additional new COSMOS features

- Entity linking (late 2020)
 - SciSpacy
 - Recognizes Unified Medical Language System (UMLS) entities
 - Enables recall of objects via UMLS IDs
- Text normalization (August 2021)
 - Cleans up ligatures, mojibake

COSMOS Updates - API

Major updates in the last year, many requested by ASKE collaborators

- Document-level recall and filters
 - Bring back tables, figures, equations for a DOI, xDD ID, or ASKE-ID
 - Filter documents based on occurrence of additional term(s)
- dataset_id filtration
 - Can unify backend into one Elasticsearch but filter to a dataset at the API instance level
- Optional API key requirement
 - Grants object recall to specific users

COSMOS Updates - API

Major updates in the last year, many requested by ASKE collaborators

- Store and filter by object size
 - Can restrict full-page artifacts
- Image type
 - Pipeline by default outputs fill-resolution PNGs. Added parameter to swap to compressed (JPG) or thumbnails
- Parameters to choose search field
 - Object, associated caption, associated body text, or combinations
- Added hooks for object-level filtration (e.g. placeholders for reproduced copyrighted figures)
- Bug fixes

THE COSMOS VISUALIZER

The **COSMOS visualizer** is our primary public interface for interacting with visual extractions subset from documents using the COSMOS entity-extraction pipeline.

xDD sets

dolomites
geothermal
iodp
mars
nam-k-carbon
pbdb
xdd-covid-19



COSMOS visualizer

The **xdd-covid-19** knowledge base consists of 4,646,618 entities extracted from 154,777 scientific publications.

Enter a query to search. Or try one of these examples:

incidence rate ACE2 coronavirus lungs opacity
peak infections death rate

The visualizer is deployed across seven document sets within the xDD infrastructure, each with 10^3 - 10^5 publications (of $\sim 10^7$ total).

COSMOS

death rate

Figure

Extraction type

FigureTableEquationAll

Hide details

Thresholds

Score

Post confidence

ElasticSearch

Match terms

Related terms

Model: trigram_cleaned

death rate

mortality rate

death rates

fatality rate

birth rate

mortality rates

number deaths

recovery rate

case-fatality rate

CFR

fatality

Configurability of extraction parameters and related terms

Code repository: <https://github.com/UW-COSMOS/cosmos-visualizer>
Live instance: https://xdd.wisc.edu/set_visualizer/

COSMOS

death rate

All types

, *Effect of age, comorbidity and remission status on outcome of COVID-19 in patients with hematological malignancies, Blood Cells, Molecules, and Diseases, 2020, DOI: 10.1016/j.bcmed.2020.102525*

Treatment received	
Antivirals	65 (50)
HCQs	8 (6.2)
Favipiravir	14 (10.8)
Remdesivir	32 (25)
HCQs + Favipiravir	2 (1.5)
HCQs + Remdesivir	7 (5.4)
Favipiravir+Remdesivir	2 (1.5)
Convalescent plasma	17 (13.1)
Corticosteroids	80 (61.5)
Anticoagulation	39 (30)

Permalink

, *A novel compartmental model to capture the nonlinear trend of COVID-19, Computers in Biology and Medicine, 134, 2021, DOI: 10.1016/j.compbio.2021.104421*

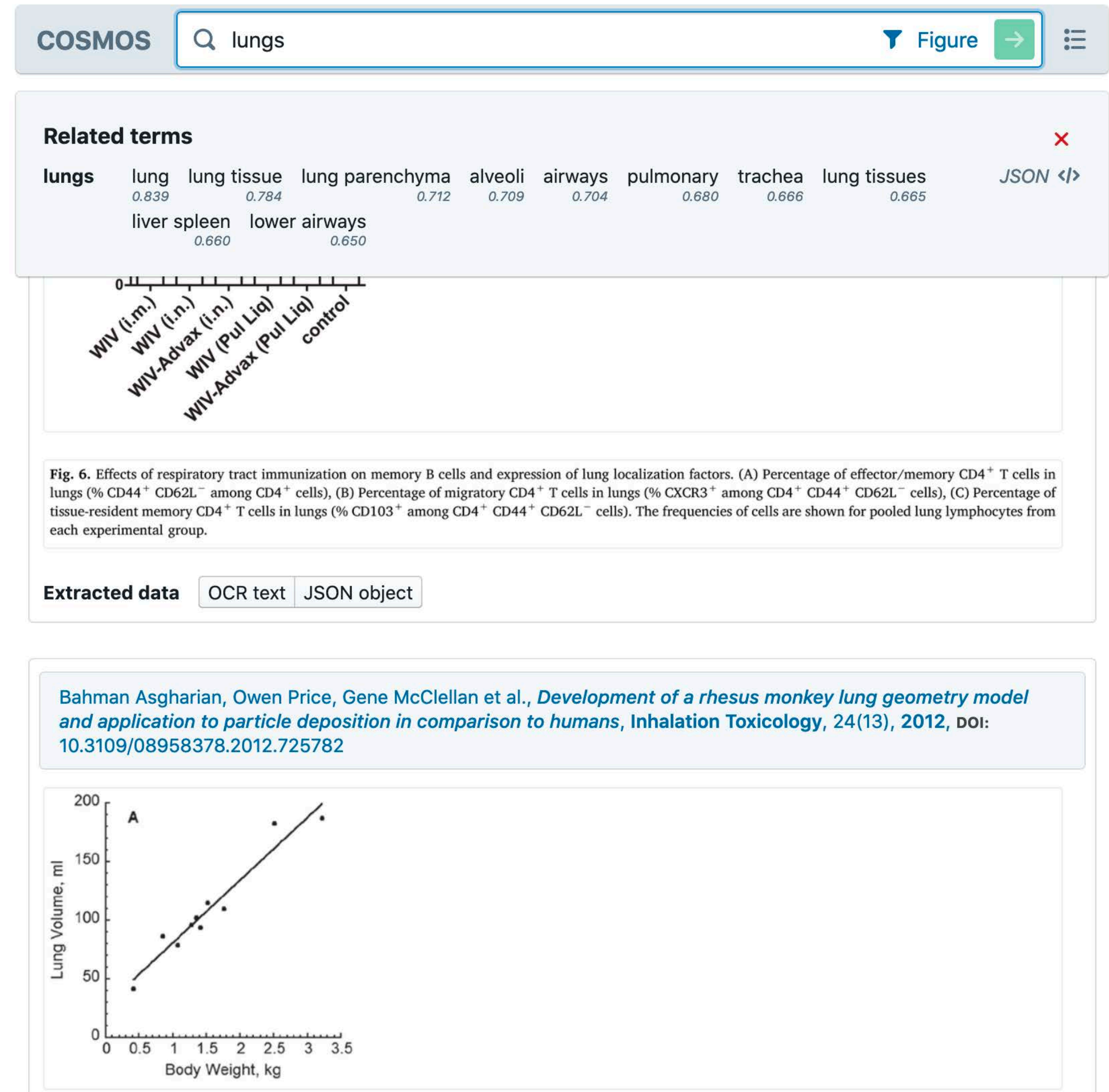
Permalink

Figures and tables extracted from the **xdd-covid-19** document set

Results visualizer

Improve public interface to extractions

- More performant integrated API over visual extractions
 - backed by ElasticSearch
 - Keyword search over figure/table captions, body text
- Bugfixes and improvements to visualizer user interface
- Incorporate user feedback on extraction quality



Model training

Maintain and build out annotation interface

- Maintain document-tagging functionality for model training and evaluation
- Support tagging more complex data types (e.g. linked information hierarchies)

solved, associated with the surface of black carbon, or partitioned into NOC. Thus, the distribution of PAHs can be expressed as

where K_p is now the PAH partition coefficient based on two particulate phases, NOC and BC; f_{BC} is the fraction BC in the soil or sediment; and K_{BC} is the BC-normalized partition coefficient. Recently, Bucheli and Gustaffson [21] and Accardi-Dey and Gschwend [17] revised this equation to include a nonlinear term:

in which n is the Freundlich exponent. This term recognizes the nonlinear adsorption interaction between the HOC and BC rather than the purely linear absorptive mechanism in Equation

Equation definition

Equation

$$K_p = f_{OC}K_{OC} + f_{BC}K_{BC}$$

Equation label

(4)

Symbol

Symbol definition

Symbol

Symbol definition

Symbol

Symbol definition

Equation

$$K_p = f_{OC}K_{OC} + f_{BC}K_{BC}C_D^{n-1}$$

Equation label

(5)

Symbol

Symbol definition

Other

Downloaded by [University of Wisconsin - Madison] at 16:03 04 January 2018

Page Header

4 P. LI ET AL.

Table Caption

Table Effects of vitamin B6 on plasma AChE activity and PLP concentrations in rats fed with isocarbophos.

	AChE activity (nmol/min/ml)	PLP concentration (nmol/L)	sBP (mmHg)	dBp (mmHg)
Saline	2.11 ± 0.37	20.4 ± 5.3	128 ± 21	76 ± 25
Vitamin B6	2.16 ± 0.36	28.9 ± 2.4*	124 ± 20	72 ± 13
Eliprodil	2.09 ± 0.29	20.1 ± 4.7	130 ± 17	86 ± 9
Isocarbophos	1.98 ± 0.30	22.9 ± 3.5	137 ± 29	84 ± 10
Isocarbophos + vitamin B6	2.07 ± 0.31	27.6 ± 3.7#	126 ± 23	77 ± 13
Isocarbophos + vitamin B6 + Eliprodil	2.15 ± 0.27	26.1 ± 5.1#	138 ± 15	79 ± 17

Table Note

After isocarbophos treatment, blood pressure was determined by the tail cuff method. Blood was collected to measure the levels of plasma AChE activity and PLP concentrations in rats. Data were expressed as means ± SD. N was 6 in each group. One-way ANOVA was used for statistical analyses. * $p < 0.05$ VS Saline. # $p < 0.05$ VS Isocarbophos.

Table Caption

Table Effects of vitamin B6 on the blood flow of posterior cerebral artery (PCA) in rats fed with isocarbophos.

Groups	Right PCA		Left PCA	
	V_s (cm/s)	V_d (cm/s)	V_s (cm/s)	V_d (cm/s)
Saline	26.32 ± 4.21	12.35 ± 2.41	26.85 ± 4.66	13.54 ± 2.55
Vitamin B6	24.63 ± 5.39	11.26 ± 3.41	27.52 ± 5.98	12.54 ± 3.63
Eliprodil	21.58 ± 4.26	9.66 ± 1.58	20.51 ± 3.64*	10.49 ± 2.53
Isocarbophos	15.74 ± 3.18*	8.57 ± 1.36*	17.41 ± 3.15*	8.26 ± 1.64*
Isocarbophos + vitamin B6	26.79 ± 3.54#	13.66 ± 2.08#	26.12 ± 4.66#	13.52 ± 2.36#
Isocarbophos + vitamin B6 + Eliprodil	15.25 ± 3.64 ^s	7.65 ± 1.54 ^s	15.88 ± 3.32 ^s	7.24 ± 1.62 ^s

Table Note

After isocarbophos treatment, blood flow of PCA was determined by TCD before sacrifice. Data were expressed as means ± SD. One-way ANOVA was used for statistical analyses. N was 6 in each group. * $p < 0.05$ VS Saline. # $p < 0.05$ VS Isocarbophos. ^s $p < 0.05$ VS Isocarbophos + vitamin B6.

Body Text

Supplementation of vitamin B6 significantly recovered the blood flow of PCAs in isocarbophos-treated rats. Furthermore, the addition of eliprodil, which is a selective NR2B antagonist (27), abolished the effects of vitamin B6 on the blood flow of PCAs. There were no obvious differences of vascular thickness among groups as determined by HE staining (Figure 1B). Taking these data together, it can be suggested that vitamin B6 via NR2B activation preserved the cerebral blood flow in isocarbophos-treated rats.

Body Text

vitamin B6 on the protein levels of NR2B and PSD95 in hippocampal PSDs in isocarbophos-treated rats. As depicted in Figure 2A–2C, 12-week treatment of isocarbophos dramatically decreased the protein levels of NR2B and PSD95 but not NR2A, compared with saline-treated rats. Although this phenotype induced by isocarbophos was reversed by vitamin B6, co-treatment of eliprodil reduced the levels of both NR2B and PSD95. Interestingly, eliprodil alone lightly reduced both levels of NR2B and PSD95 in rat hippocampus.

Section Header

Vitamin B6 improves isocarbophos-impaired learning and memory, which is suppressed by eliprodil in rats

Body Text

The typical phenotype of dementia is the impairment of the spatial memory (28,29). Thus, the learning and memory behaviors were evaluated by the MWM test. As shown in Table 3, after 12-week isocarbophos administration, swimming distance in quadrant, ratio of swimming distance in quadrant to total swimming distance, swimming time in quadrant, swimming time in quadrant to total swimming time, and number of escape were increased in isocarbophos-treated rats, compared with saline-treated rats. Conversely, swimming time and distance out of the quadrant were increased in isocarbophos-treated rats. However, the supplementation of vitamin B6 significantly normalized these alterations in isocarbophos-treated rats, which were ablated by eliprodil.

Section Header

Vitamin B6 increases CaMK-II expression in rat hippocampus

Body Text

CaMK-II has been demonstrated to potentiate NMDA receptor function by increasing the fraction of receptors in the synaptosomal membranes (32). Thus, we investigated whether isocarbophos altered the CaMK-II protein expression in the hippocampus PSDs. As shown in Figure 3A and 3B, compared with the saline group, both CaMK-II protein level and phosphorylated (p-CaMK-II) decreased in the hippocampus PSDs of the isocarbophos group. Similarly, the supplementation of vitamin B6 increased the CaMK-II protein level in isocarbophos-treated rats, which was ablated by eliprodil. Together, it reveals that vitamin B6 might activate NR2B/PSD95/CaMK-II signaling in hippocampal PSDs.

Minimalist interface focused on locating and extracting data from tables

Search terms/phrases: (separate multiple entries with commas):
death rate

Require terms to co-occur: ☒ Case Sensitive: ☐ Type: Table

Get Results

Related Terms from Trained Embedding Model:

death rate: death rates (0.83), mortality rate (0.81), fatality rate (0.79), birth rate (0.76), number deaths (0.75), hospitalization rate (0.75), mortality rates (0.74), fatality (0.73), cfr (0.72), lethality rate (0.72), recovery rate (0.71), numbers deaths (0.71), case-fatality rate (0.71), probability death (0.7), fatality rates (0.7), case-fatality ratio (0.69), fatality ratio (0.68), incidence rate (0.68), death toll (0.68), risk dying (0.68), (see [similarity matrix](#) or raw [JSON](#))

Search Results for **death+rate**: (Page 0: 162 results)

Get Next Results

The Data set for Patient Information Based Algorithm to Predict Mortality Cause by COVID-19

Published: 2020. Data in Brief

Table 5
Death rate estimation in South Korea.

5-A: Death rate derived from the formula of trendlines						
Date	3/11/2020	3/12/2020	3/13/2020	3/14/2020	3/15/2020	3/16/2020
Death rate 1	3.95%	26.92%	30.57%	21.26%	5.35%	
current ratio	0.79%	0.82%	0.85%	0.88%	0.91%	0.94%
Death rate 2	23.35%	20.90%	18.45%	16.00%	13.55%	11.10%
Date	3/17/2020	3/18/2020	3/19/2020	3/20/2020	3/21/2020	
Death rate 1						
current ratio	0.97%	1.00%	1.03%	1.06%	1.09%	
Death rate 2	8.65%	6.20%	3.75%	1.30%		

EXTRACT DATA

- all mentions of death+rate in **JSON**

Data extraction: human-in-the-loop

Minimalist interface focused on locating and extracting data from tables

Extract data from [UW-COSMOS](#)

See [all extracted data](#)

The Data set for Patient Information Based Algorithm to Predict Mortality Cause by COVID-19						
Published: 2020. Data in Brief						
Table 5						
Death rate estimation in South Korea.						
5-A: Death rate derived from the formula of trendlines						
Date	3/11/2020	3/12/2020	3/13/2020	3/14/2020	3/15/2020	3/16/2020
Death rate 1	3.95%	26.92%	30.57%	21.26%	5.35%	
current ratio	0.79%	0.82%	0.85%	0.88%	0.91%	0.94%
Death rate 2	23.35%	20.90%	18.45%	16.00%	13.55%	11.10%
Date	3/17/2020	3/18/2020	3/19/2020	3/20/2020	3/21/2020	
Death rate 1						
current ratio	0.97%	1.00%	1.03%	1.06%	1.09%	
Death rate 2	8.65%	6.20%	3.75%	1.30%		

Variable	Value	VMax	VMin	Local	Date	Notes
<div>Submit Data</div>						
Extracted Data for Object						
Variable	Value	VMax	VMin	Local	Date	Notes

Data extraction: human-in-the-loop

Minimalist interface focused on locating and extracting data from tables

All Extracted Data (CSV download)										
Variable	Value	VMax	Vmin	Local	Date	Notes	object link	object_id	gddid	Year
R0	3.04			Hefei, China	January 2020	mean value during initial stage of response	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
incubation period	9			Shenzhen, China	January 2020	IQR median days	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
R0	1.94			Hefei, China	January 2020	mean value during the last week	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
R0	3.03			Shenzhen, China	January 2020	mean value during initial stage of response	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
incubation period	4.0			Hefei, China	January 2020	IQR median days	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
R0	1.59			Shenzhen, China	January 2020	mean value during the last week	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
infection period	3	6	1	Hefei, China	January 2020	average, days median IQR	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020
infection period	3	6	1	Shenzhen, China	January 2020	average, days median IQR	see object	296630d0aea844e8b96388c7cb2b08737f5e75d7	5e834fbc998e17af8269f85f	2020

~700 data points



A System for Large Scale Graph Embeddings

Anze Xie*, Anders Carlsson*, Jason Mohoney*, Roger Waleffe*
Shanan Peters#, Theodoros Rekatsinas*, Shivaram Venkataraman*

*Computer Sciences, #Geoscience



ASKE-E UW-COSMOS Tasks

- **Task 1: Scale and Enhance xDD Infrastructure and APIs**
 - **Task 1.A. Objective: Expand Corpus**
 - **Task 1.B. Objective: Develop Container Template for Collaborator Code**
 - **Task 1.C. Objective: Augment xDD Corpus with Container Outputs**
 - **Task 1.D. Objective: Improve and Scale xDD API** (and add other features)
- **Task2: Scale and Enhance COSMOS Retrieval and API**
 - **Task 2.A. Improve Visual Segmentation** (preprints are challenging, need ML model)
 - **Task 2.B. Incorporate Body-Text Content into Table and Figure Retrieval**
 - **Task 2.C. Automatic Knowledge Base Construction** (Marius, table extraction)
 - **Task 2.D. Release Public* COSMOS API Over COVID-19 xDD Set**

xDD and UW-COSMOS data infrastructure overview

Text-based search and retrieval over full text
of all 14M full texts in xDD corpus

xdd.wisc.edu/api

- [api/snippets](#): string-based search over document full-text to retrieve document metadata and contextual snippets highlighting search matches.
- [api/articles](#): similar to snippets, but article metadata focused
- [api/journals](#): search for specific journals to assess coverage
- [api/](#): additional routes and documentation

continuous
updating

xDD
API

xDD
sets

Subsets of xDD corpus for specific
projects: xdd.wisc.edu/sets

periodic updating

sets/xdd-covid-19

154,777 full-text docs

xDD-defined COVID19 doc set (CORD19,
EMMAA, search-term defined): full content only

restrict API results to set:
&dataset=xdd-covid-19

Table, figure, equation, retrieval over sets
via COSMOS AI assistant (cosmos.wisc.edu)

xdd.wisc.edu/sets/xdd-covid-19/cosmos/api/

- [search](#): string-based search over document figures, tables, equations based on associated text.
- other routes available, most *require API key*

xdd.wisc.edu/sets/xdd-covid-19/word2vec/api/

- trained embeddings
- several vector operations

xdd.wisc.edu/sets/xdd-covid-19/doc2vec/api/

- document-level embeddings, doc discovery based on similarity

ASKE-
ID

UW
COSMOS
API

word2vec
API

doc2vec
API

- service for
generating/
retrieving
metadata and
links to artifacts



xDD automated
full-pub fetching
10³s per day

xDD corpus
14M full texts
+
PMC abstracts



open access
full-pub
abstract-only
text



Major support by DARPA ASKE, DOE, USGS
2014-2018 by NSF-ICER 1343760

