Redshift

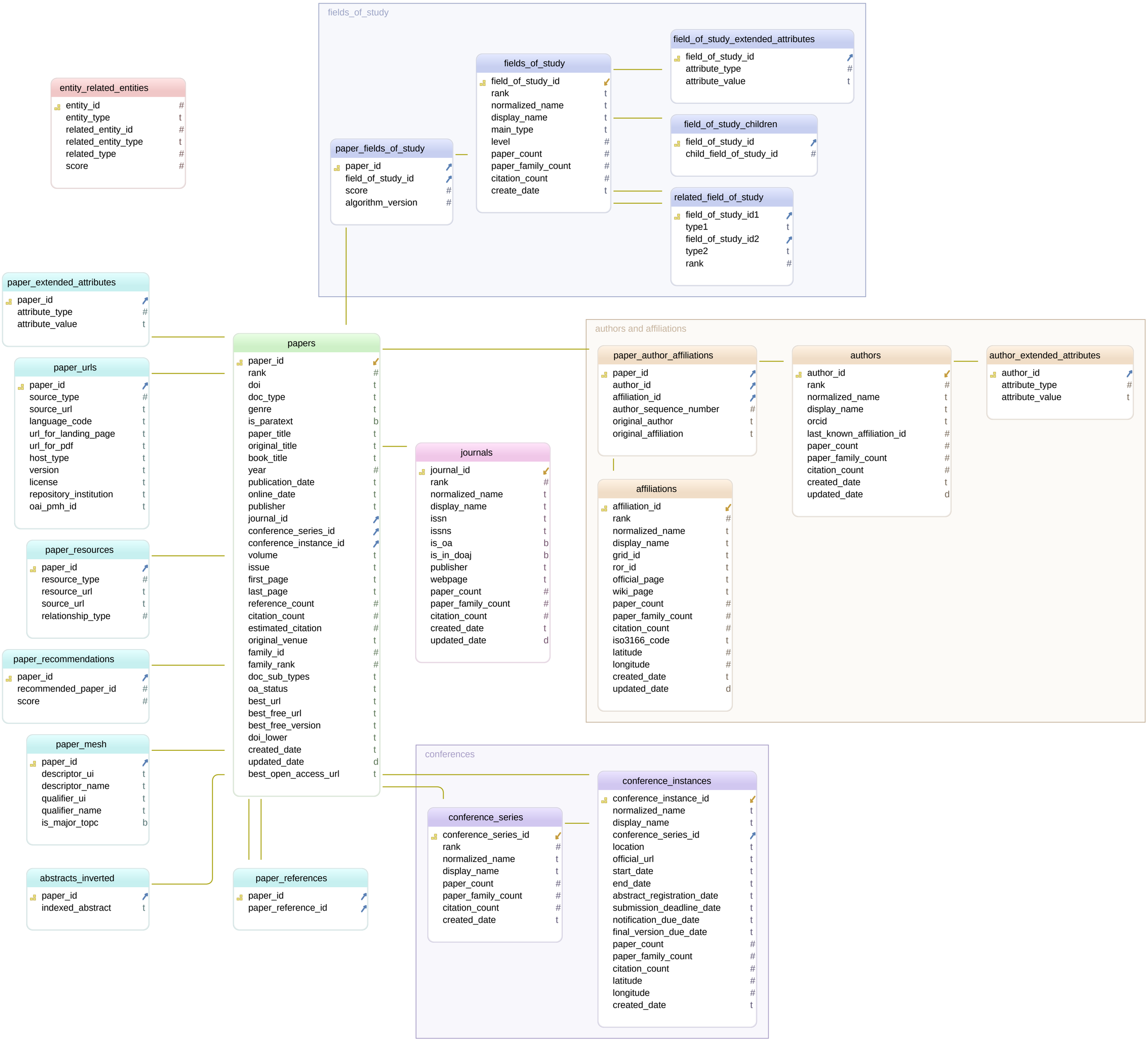
4-11-2021 null

Layouts

1. Export Tables 1

Tables

outs.abstracts_inverted [1] 2
outs.affiliations [1] 2
outs.author_extended_attributes [1] 2
outs.authors [1] 2
outs.conference_instances [1] 3
outs.conference_series [1] 3
outs.entity_related_entities [1] 4
outs.field_of_study_children [1] 4
outs.field_of_study_extended_attributes [1] 4
outs.fields_of_study [1] 4
outs.journals [1] 5
outs.paper_author_affiliations [1] 5
outs.paper_extended_attributes [1] 5
outs.paper_fields_of_study [1] 6
outs.paper_mesh [1] 6
outs.paper_recommendations [1] 6
outs.paper_references [1] 6
outs.paper_resources [1] 7
outs.paper_urls [1] 7
outs.papers [1] 7
outs.related_field_of_study [1] 8


fields_of_study
entity_related_entities
entity_id #
entity_type t
related_entity_id #
related_entity_type t
related_type #
score #
fields_of_study
field_of_study_id
rank t
normalized_name t
display_name t
main_type t
level #
paper_count #
paper_family_count #
citation_count #
create_date t
field_of_study_extended_attributes
field_of_study_id
attribute_type #
attribute_value t
field_of_study_children
field_of_study_id
child_field_of_study_id #
related_field_of_study
field_of_study_id1
type1 t
field_of_study_id2
type2 t
rank #
paper_fields_of_study
paper_id
field_of_study_id
score #
algorithm_version #
paper_extended_attributes
paper_id
attribute_type #
attribute_value t
paper_urls
paper_id
source_type #
source_url t
language_code t
url_for_landing_page t
url_for_pdf t
host_type t
version t
license t
repository_institution t
oai_pmh_id t
papers
paper_id
rank #
doi t
doc_type t
genre t
is_paratext b
paper_title t
original_title t
book_title t
year #
publication_date t
online_date t
publisher t
journal_id
conference_series_id
conference_instance_id
volume t
issue t
first_page t
last_page t
reference_count #
citation_count #
estimated_citation #
original_venue t
family_id #
family_rank #
doc_sub_types t
oa_status t
best_url t
best_free_url t
best_free_version t
doi_lower t
created_date t
updated_date d
best_open_access_url t
journals
journal_id
rank #
normalized_name t
display_name t
issn t
issns t
is_oa b
is_in_doaj b
publisher t
webpage t
paper_count #
paper_family_count #
citation_count #
created_date t
updated_date d
authors and affiliations
paper_author_affiliations
paper_id
author_id
affiliation_id
author_sequence_number #
original_author t
original_affiliation t
authors
author_id
rank #
normalized_name t
display_name t
orcid t
last_known_affiliation_id #
paper_count #
paper_family_count #
citation_count #
created_date t
updated_date d
author_extended_attributes
author_id
attribute_type #
attribute_value t
affiliations
affiliation_id
rank #
normalized_name t
display_name t
grid_id t
ror_id t
official_page t
wiki_page t
paper_count #
paper_family_count #
citation_count #
iso3166_code t
latitude #
longitude #
created_date t
updated_date d
paper_resources
paper_id
resource_type #
resource_url t
source_url t
relationship_type #
paper_recommendations
paper_id
recommended_paper_id #
score #
paper_mesh
paper_id
descriptor_ui t
descriptor_name t
qualifier_ui t
qualifier_name t
is_major_topc b
abstracts_inverted
paper_id
indexed_abstract t
paper_references
paper_id
paper_reference_id
conferences
conference_series
conference_series_id
rank #
normalized_name t
display_name t
paper_count #
paper_family_count #
citation_count #
created_date t
conference_instances
conference_instance_id
normalized_name t
display_name t
conference_series_id
location t
official_url t
start_date t
end_date t
abstract_registration_date t
submission_deadline_date t
notification_due_date t
final_version_due_date t
paper_count #
paper_family_count #
citation_count #
latitude #
longitude #
created_date t

Export Tables

Table abstracts_inverted

Inverted abstracts

	paper_id	bigint	FOREIGN KEY REFERENCES papers.papers_id
	indexed_abstract	varchar	Inverted index, see https://en.wikipedia.org/wiki/Inverted_index

Indexes

	Distribution	paper_id
	Sorting	paper_id

Foreign Keys

Vir	fk_abstracts_inverted_papers (paper_id) ref papers (paper_id)

Table affiliations

Base table for affiliations

	affiliation_id	bigint	PRIMARY KEY
	rank	integer	DEPRECATED; no new ranks are being added.
	normalized_name	varchar	
	display_name	varchar	
	grid_id	varchar	DEPRECATED; ror_id is the new standard identifier for organizations
	ror_id	varchar	NEW; ROR for this organization, see https://ror.org, https://ror.org/:ror_id
	official_page	varchar	
	wiki_page	varchar	
	paper_count	bigint	
	paper_family_count	bigint	DEPRECATED; same value as paper_count.
	citation_count	bigint	
	iso3166_code	varchar	Two-letter country codes, see https://en.wikipedia.org/wiki/ISO_3166-2
	latitude	real	
	longitude	real	
	created_date	varchar	
	updated_date	timestamp	NEW; set values updated from new ror data

Indexes

	Distribution	affiliation_id
	Sorting	affiliation_id

Table author_extended_attributes

Additional author name representations

	author_id	bigint	FOREIGN KEY REFERENCES Authors.AuthorId
	attribute_type	integer	Possible values: 1=Alternative name
	attribute_value	varchar	

Indexes

	Distribution	author_id
	Sorting	author_id

Foreign Keys

Vir	fk_author_extended_attributes_authors (author_id) ref authors (author_id)

Table authors

Base table for affiliations

	author_id	bigint	PRIMARY KEY
	rank	integer	DEPRECATED; no new ranks are being added
	normalized_name	varchar	

Table authors

	display_name	varchar	
	orcid	varchar	NEW; ORCID identifier for this author
	last_known_affiliation_id	integer	
	paper_count	bigint	
	paper_family_count	bigint	DEPRECATED; same value as paper_count
	citation_count	bigint	
	created_date	varchar	
	updated_date	timestamp	NEW; set when changes are made going forward

Indexes

	Distribution	author_id
	Sorting	author_id

Table conference_instances

Base table for Conference Instances

	conference_instance_id	bigint	PRIMARY KEY
	normalized_name	varchar	
	display_name	varchar	
	conference_series_id	bigint	FOREIGN KEY REFERENCES ConferenceSeries.ConferenceSeriesId
	location	varchar	
	official_url	varchar	
	start_date	varchar	
	end_date	varchar	
	abstract_registration_date	varchar	
	submission_deadline_date	varchar	
	notification_due_date	varchar	
	final_version_due_date	varchar	
	paper_count	bigint	
	paper_family_count	bigint	DEPRECATED; same value as paper_count
	citation_count	bigint	
	latitude	real	
	longitude	real	
	created_date	varchar	

Indexes

	Distribution	conference_instance_id
	Sorting	conference_instance_id

Foreign Keys

Vir	fk_conference_instances_conference_series (conference_series_id) ref conference_series (conference_series_id)

Table conference_series

Base table for Conference Series

	conference_series_id	bigint	PRIMARY KEY
	rank	integer	DEPRECATED; no new ranks are being added
	normalized_name	varchar	
	display_name	varchar	
	paper_count	bigint	
	paper_family_count	bigint	DEPRECATED; same value as paper_count
	citation_count	bigint	
	created_date	varchar	

Indexes

	Distribution	conference_series_id
	Sorting	conference_series_id

Table entity_related_entities

DEPRECATED

entity_id	bigint	
entity_type	varchar	Possible values: af (Affiliation), j (Journal), c (Conference)
related_entity_id	bigint	
related_entity_type	varchar	Possible values: af (Affiliation), j (Journal), c (Conference)
related_type	integer	Possible values: 0 (same paper), 1 (common coauthors), 2 (co-cited), 3 (common field of study), 4 (same venue), 5 (A cites B), 6 (B cites A)
score	real	Confidence range between 0 and 1. Larger number representing higher confidence.

Indexes

Distribution	entity_id	
Sorting	entity_id	

Table field_of_study_children

Relationship between fields_of_study.

field_of_study_id	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId
child_field_of_study_id	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId

Indexes

Distribution	field_of_study_id	
Sorting	field_of_study_id	

Foreign Keys

Vir	fk_field_of_study_children_fields_of_study (field_of_study_id) ref fields_of_study (field_of_study_id)

Table field_of_study_extended_attributes

Other identifiers for fields_of_study.

field_of_study_id	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId
attribute_type	bigint	Possible values: 1 (AUI from UMLS https://www.nlm.nih.gov/research/umls/licensedcontent/umlsarchives04.html#2018AA), 2 (source url), 3 (CUI from UMLS https://www.nlm.nih.gov/research/umls/licensedcontent/umlsknowledgesources.html)
attribute_value	varchar	

Indexes

Distribution	field_of_study_id	
Sorting	field_of_study_id	

Foreign Keys

Vir	fk_field_of_study_extended_attributes_fields_of_study (field_of_study_id) ref fields_of_study (field_of_study_id)

Table fields_of_study

Base table for Fields of Study

field_of_study_id	bigint	PRIMARY KEY
rank	varchar	DEPRECATED; no new ranks are being added.
normalized_name	varchar	
display_name	varchar	
main_type	varchar	
level	integer	Possible values: 0-5
paper_count	bigint	
paper_family_count	bigint	DEPRECATED; same value as paper_count.
citation_count	bigint	

Table fields_of_study

create_date	varchar	
Indexes		
Distribution	field_of_study_id	
Sorting	field_of_study_id	

Table journals

Base table for Journals

journal_id	bigint	PRIMARY KEY
rank	integer	DEPRECATED; no new ranks are being added
normalized_name	varchar	
display_name	varchar	
issn	varchar	the ISSN-L for the journal (see https://en.wikipedia.org/wiki/International_Standard_Serial_Number#Linking_ISSN)
issns	varchar	NEW; JSON list of all ISSNs for this journal (example: '["1469-5073","0016-6723"]')
is_oa	boolean	NEW; TRUE when the journal is 100% OA
is_in_doaj	boolean	NEW; TRUE when the journal is in DOAJ (see https://doaj.org/)
publisher	varchar	
webpage	varchar	
paper_count	bigint	
paper_family_count	bigint	DEPRECATED; same value as paper_count
citation_count	bigint	
created_date	varchar	
updated_date	timestamp	NEW; set when changes are made going forward
Indexes		
Distribution	journal_id	
Sorting	journal_id	

Table paper_author_affiliations

Links between papers

paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
author_id	bigint	FOREIGN KEY REFERENCES Authors.AuthorId
affiliation_id	bigint	FOREIGN KEY REFERENCES Affiliations.AffiliationId
author_sequence_number	integer	1-based author sequence number. 1: the 1st author listed on paper, 2: the 2nd author listed on paper, etc.
original_author	varchar	
original_affiliation	varchar	
Indexes		
Distribution	paper_id	
Sorting	paper_id	
Foreign Keys		
Vir	fk_paper_author_affiliations_papers (paper_id) ref papers (paper_id)	
Vir	fk_paper_author_affiliations_affiliations (affiliation_id) ref affiliations (affiliation_id)	
Vir	fk_paper_author_affiliations_authors (author_id) ref authors (author_id)	

Table paper_extended_attributes

Extra paper identifiers

paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
attribute_type	integer	Possible values: 1=PatentId, 2=PubMedId, 3=PmcId, 4=Alternative Title
attribute_value	varchar	

Table paper_extended_attributes

Indexes		
	Distribution	paper_id
	Sorting	paper_id
Foreign Keys		
Vir	fk_paper_extended_attributes_papers (paper_id) ref papers (paper_id)	

Table paper_fields_of_study

Linking table from papers to fields

	paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	field_of_study_id	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId
	score	real	Confidence range between 0 and 1. Bigger number representing higher confidence.
	algorithm_version	integer	NEW; version of algorithm to assign fields. Possible values: 1=old MAG (no longer added), 2=OpenAlex
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_fields_of_study_papers (paper_id) ref papers (paper_id)		
Vir	fk_paper_fields_of_study_fields_of_study (field_of_study_id) ref fields_of_study (field_of_study_id)		

Table paper_mesh

MeSH headings assigned to the paper by PubMed

	paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	descriptor_ui	varchar	see https://en.wikipedia.org/wiki/Medical_Subject_Headings
	descriptor_name	varchar	see https://en.wikipedia.org/wiki/Medical_Subject_Headings
	qualifier_ui	varchar	see https://en.wikipedia.org/wiki/Medical_Subject_Headings
	qualifier_name	varchar	see https://en.wikipedia.org/wiki/Medical_Subject_Headings
	is_major_topc	boolean	see https://en.wikipedia.org/wiki/Medical_Subject_Headings
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_mesh_papers (paper_id) ref papers (paper_id)		

Table paper_recommendations

Paper recommendations with score

	paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	recommended_paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	score	real	Confidence range between 0 and 1. Bigger number representing higher confidence.
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_recommendations_papers (paper_id) ref papers (paper_id)		

Table paper_references

Paper references

Table paper_references

	paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	paper_reference_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_references_papers (paper_id) ref papers (paper_id)		
Vir	fk_paper_references_papers_0 (paper_reference_id) ref papers (paper_id)		

Table paper_resources

DEPRECATED

	paper_id	bigint	FOREIGN KEY REFERENCES papers.papers_id
	resource_type	integer	Bit flags: 1=Project, 2=Data, 4=Code
	resource_url	varchar	Url of resource
	source_url	varchar	List of urls associated with the project, used to derive resource_url
	relationship_type	integer	Bit flags: 1=Own, 2=Cite
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_resources_papers (paper_id) ref papers (paper_id)		

Table paper_urls

MeSH headings assigned to the paper by PubMed

	paper_id	bigint	FOREIGN KEY REFERENCES Papers.PaperId
	source_type	integer	Possible values: 1=Html, 2=Text, 3=Pdf, 4=Doc, 5=Ppt, 6=Xls, 8=Rtf, 12=Xml, 13=Rss, 20=Swf, 27=Ics, 31=Pub, 33=Ods, 34=Odp, 35=Odt, 36=Zip, 40=Mp3, 0/999/NULL=unknown
	source_url	varchar	
	language_code	varchar	
	url_for_landing_page	varchar	NEW; URL for the landing page, when article is free to read
	url_for_pdf	varchar	NEW; URL for the PDF, when article is free to read
	host_type	varchar	NEW; host type of the free-to-read URL, Possible values: publisher, repository
	version	varchar	NEW; version of the free-to-read URL Possible values: submittedVersion, acceptedVersion, publishedVersion
	license	varchar	NEW; license of the free-to-read URL (example: cc0, cc-by, publisher-specific)
	repository_institution	varchar	NEW; name of repository host of URL
	oai_pmh_id	varchar	NEW; OAH-PMH id of the repository record
Indexes			
	Distribution	paper_id	
	Sorting	paper_id	
Foreign Keys			
Vir	fk_paper_urls_papers (paper_id) ref papers (paper_id)		

Table papers

Main data for papers

	paper_id	bigint	PRIMARY KEY
	rank	integer	DEPRECATED; no new ranks are being added

Table papers

	doi	varchar	Doi values are upper-cased per DOI standard at https://www.doi.org/doi_handbook/2_Numbering.html#2.4
	doc_type	varchar	Possible values: Book, BookChapter, Conference, Dataset, Journal, Patent, Repository, Thesis, NULL : unknown. Patent is DEPRECATED; no new Patents are being added.
	genre	varchar	NEW
	is_paratext	boolean	NEW
	paper_title	varchar	
	original_title	varchar	
	book_title	varchar	
	year	integer	
	publication_date	varchar	
	online_date	varchar	
	publisher	varchar	
	journal_id	bigint	FOREIGN KEY references journals.journal_id
	conference_series_id	bigint	DEPRECATED; no longer updated, no new Conference Series are being added. FOREIGN KEY references conference_series.conference_series_id.
	conference_instance_id	bigint	DEPRECATED; no longer updated, no new Conference Instances are being added. FOREIGN KEY references conference_instances.conference_instance_id.
	volume	varchar	
	issue	varchar	
	first_page	varchar	
	last_page	varchar	
	reference_count	bigint	
	citation_count	bigint	
	estimated_citation	bigint	DEPRECATED; is set equal to citation_count
	original_venue	varchar	
	family_id	bigint	DEPRECATED; no longer updated.
	family_rank	bigint	DEPRECATED; no longer updated.
	doc_sub_types	varchar	Possible values: "Retracted Publication", "Retraction Notice".
	oa_status	varchar	NEW; Possible values: closed, green, gold, hybrid, bronze (see https://en.wikipedia.org/wiki/Open_access#Colour_naming_system)
	best_url	varchar	NEW; An url for the paper (see paper_urls table for more)
	best_free_url	varchar	NEW; An url of legal free-to-read copy when it exists
	best_free_version	varchar	NEW; Possible values: submittedVersion, acceptedVersion, publishedVersion
	doi_lower	varchar	NEW; lowercase doi for convenience linking to Unpaywall
	created_date	varchar	
	updated_date	timestamp	NEW; set when changes are made going forward
	best_open_access_url	varchar	NEW; The best url for reading this paper for free

Indexes

	Distribution	paper_id
	Sorting	paper_id

Foreign Keys

Vir	fk_papers_conference_series (conference_series_id) ref conference_series (conference_series_id)
Vir	fk_papers_conference_instances (conference_instance_id) ref conference_instances (conference_instance_id)
Vir	fk_papers_journals (journal_id) ref journals (journal_id)

Table related_field_of_study

Relationships between fields of study

Table related_field_of_study

	field_of_study_id1	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId
	type1	varchar	Possible values: general, disease, disease_cause, medical_treatment, symptom
	field_of_study_id2	bigint	FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId
	type2	varchar	Possible values: general, disease, disease_cause, medical_treatment, symptom
	rank	real	DEPRECATED; no new ranks are being added.

Indexes

	Distribution	field_of_study_id1
	Sorting	field_of_study_id1

Foreign Keys

Vir	fk_related_field_of_study_fields_of_study (field_of_study_id1) ref fields_of_study (field_of_study_id)
Vir	fk_related_field_of_study_fields_of_study_0 (field_of_study_id2) ref fields_of_study (field_of_study_id)