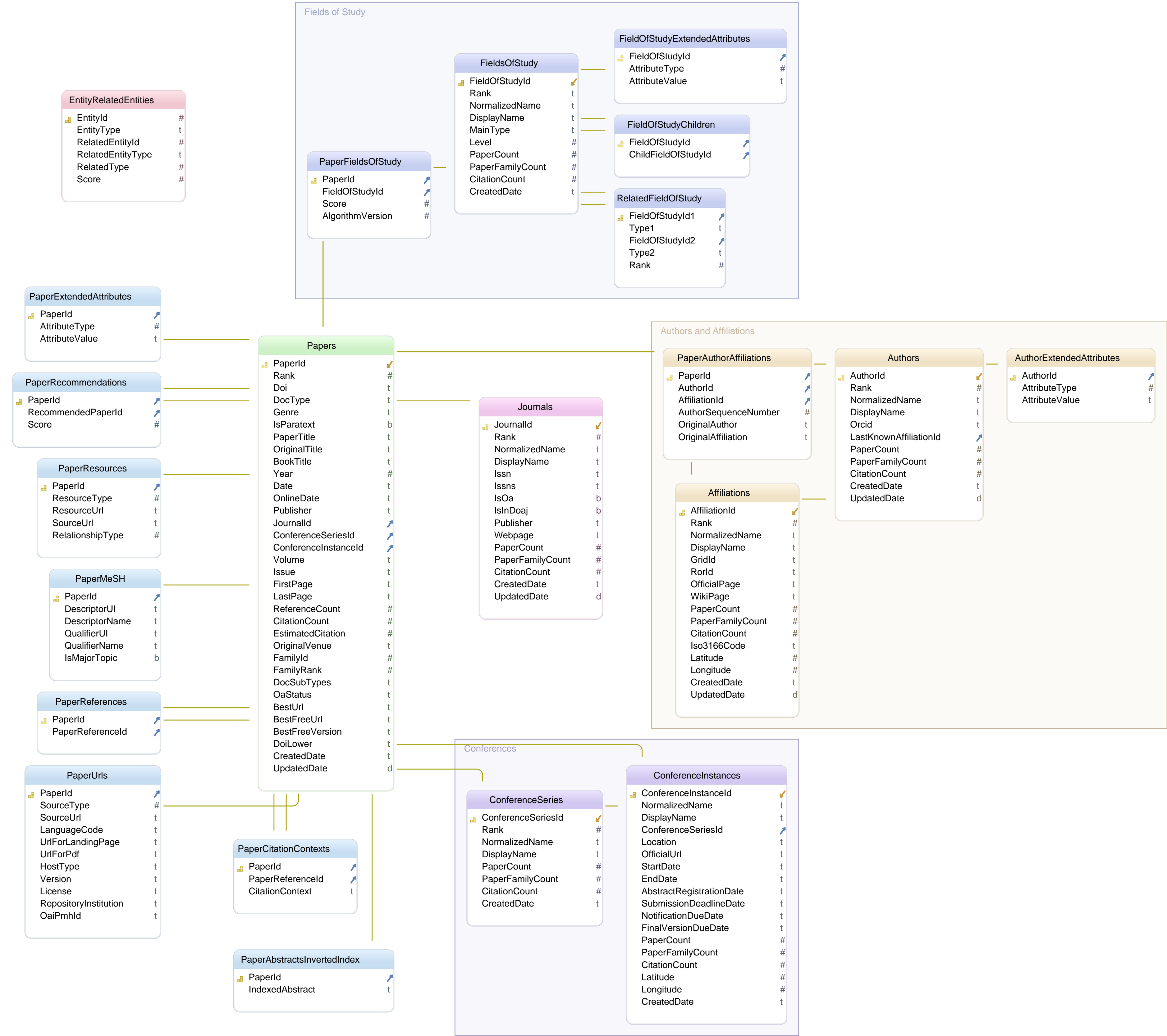


Layouts

| | |
|---------------------|---|
| 1. Main Layout..... | 1 |
|---------------------|---|

Tables

| | |
|--|---|
| outs.Affiliations [1]..... | 2 |
| outs.AuthorExtendedAttributes [1]..... | 2 |
| outs.Authors [1]..... | 2 |
| outs.ConferenceInstances [1]..... | 3 |
| outs.ConferenceSeries [1]..... | 3 |
| outs.EntityRelatedEntities [1]..... | 3 |
| outs.FieldOfStudyChildren [1]..... | 4 |
| outs.FieldOfStudyExtendedAttributes [1]..... | 4 |
| outs.FieldsOfStudy [1]..... | 4 |
| outs.Journals [1]..... | 4 |
| outs.PaperAbstractsInvertedIndex [1]..... | 5 |
| outs.PaperAuthorAffiliations [1]..... | 5 |
| outs.PaperCitationContexts [1]..... | 6 |
| outs.PaperExtendedAttributes [1]..... | 6 |
| outs.PaperFieldsOfStudy [1]..... | 6 |
| outs.PaperMeSH [1]..... | 6 |
| outs.PaperRecommendations [1]..... | 7 |
| outs.PaperReferences [1]..... | 7 |
| outs.PaperResources [1]..... | 7 |
| outs.PaperUrls [1]..... | 7 |
| outs.Papers [1]..... | 8 |
| outs.RelatedFieldOfStudy [1]..... | 9 |



Main Layout

Table Affiliations

| | | |
|---|-----------|---|
| Base table for affiliations/institutions (mag/Affiliations.txt) | | |
| AffiliationId | bigint | PRIMARY KEY |
| Rank | integer | FROZEN; no new ranks are being added. |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| GridId | varchar | FROZEN; ror_id is the new standard identifier for organizations |
| RorId | varchar | NEW; ROR for this organization, see https://ror.org , https://ror.org:ror_id |
| OfficialPage | varchar | |
| WikiPage | varchar | |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count. |
| CitationCount | bigint | |
| Iso3166Code | varchar | Two-letter country codes, see https://en.wikipedia.org/wiki/ISO_3166-2 |
| Latitude | real | |
| Longitude | real | |
| CreatedDate | varchar | |
| UpdatedDate | timestamp | NEW; set values updated from new ror data |

Indexes

| | |
|--------------|---------------|
| Distribution | AffiliationId |
| Sorting | AffiliationId |

Table AuthorExtendedAttributes

| | | |
|---|---------|---|
| Additional author name representations (mag/AuthorExtendedAttributes.txt) | | |
| AuthorId | bigint | FOREIGN KEY REFERENCES Authors.AuthorId |
| AttributeType | integer | Possible values: 1=Alternative name |
| AttributeValue | varchar | |

Indexes

| | |
|--------------|----------|
| Distribution | AuthorId |
| Sorting | AuthorId |

Foreign Keys

Vir Fk_AuthorExtendedAttributes_Authors (AuthorId) ref
Authors (AuthorId)

Table Authors

| | | |
|--|-----------|---|
| Base table for authors (mag/Authors.txt) | | |
| AuthorId | bigint | PRIMARY KEY |
| Rank | integer | FROZEN; no new ranks are being added |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| Orcid | varchar | NEW; ORCID identifier for this author |
| LastKnownAffiliationId | integer | |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count |
| CitationCount | bigint | |
| CreatedDate | varchar | |
| UpdatedDate | timestamp | NEW; set when changes are made going forward |

Indexes

| | |
|--------------|----------|
| Distribution | AuthorId |
| Sorting | AuthorId |

Foreign Keys

Table Authors

Vir Fk_Authors_Affiliations (LastKnownAffiliationId) ref Affiliations (AffiliationId)

Table ConferenceInstances

FROZEN; Base table for Conference Instances (mag/ConferenceInstances.txt)

| | | |
|--------------------------|---------|--|
| ConferenceInstanceId | bigint | PRIMARY KEY |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| ConferenceSeriesId | bigint | FOREIGN KEY REFERENCES ConferenceSeries.ConferenceSeriesId |
| Location | varchar | |
| OfficialUrl | varchar | |
| StartDate | varchar | |
| EndDate | varchar | |
| AbstractRegistrationDate | varchar | |
| SubmissionDeadlineDate | varchar | |
| NotificationDueDate | varchar | |
| FinalVersionDueDate | varchar | |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count" |
| CitationCount | bigint | |
| Latitude | real | |
| Longitude | real | |
| CreatedDate | varchar | |

Indexes

| | |
|--------------|----------------------|
| Distribution | ConferenceInstanceId |
| Sorting | ConferenceInstanceId |

Foreign Keys

Vir Fk_ConferenceInstances_ConferenceSeries (ConferenceSeriesId) ref ConferenceSeries (ConferenceSeriesId)

Table ConferenceSeries

FROZEN; Base table for Conference Series (mag/ConferenceSeries.txt)

| | | |
|--------------------|---------|---|
| ConferenceSeriesId | bigint | PRIMARY KEY |
| Rank | integer | FROZEN; no new ranks are being added |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count" |
| CitationCount | bigint | |
| CreatedDate | varchar | |

Indexes

| | |
|--------------|--------------------|
| Distribution | ConferenceSeriesId |
| Sorting | ConferenceSeriesId |

Table EntityRelatedEntities

Relationship between papers, authors, fields of study. (advanced/EntityRelatedEntities.txt)

| | | |
|-------------------|---------|--|
| EntityId | bigint | |
| EntityType | varchar | Possible values: af (Affiliation), j (Journal), c (Conference) |
| RelatedEntityId | bigint | |
| RelatedEntityType | varchar | Possible values: af (Affiliation), j (Journal), c (Conference) |
| RelatedType | 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) |

Table EntityRelatedEntities

| | | |
|--------------|----------|---|
| Score | real | Confidence range between 0 and 1. Larger number representing higher confidence. |
| Indexes | | |
| Distribution | EntityId | |
| Sorting | EntityId | |

Table FieldOfStudyChildren

Relationship between Fields of Study (advanced/FieldOfStudyChildren.txt)

| | | |
|---------------------|--------|---|
| FieldOfStudyId | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |
| ChildFieldOfStudyId | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |

Indexes

| | |
|--------------|----------------|
| Distribution | FieldOfStudyId |
| Sorting | FieldOfStudyId |

Foreign Keys

Vir Fk_FieldOfStudyChildren_FieldsOfStudy (FieldOfStudyId) ref FieldsOfStudy (FieldOfStudyId)

Vir Fk_FieldOfStudyChildren_FieldsOfStudy_0 (ChildFieldOfStudyId) ref FieldsOfStudy (FieldOfStudyId)

Table FieldOfStudyExtendedAttributes

Other identifiers for Fields of Study (advanced/FieldOfStudyExtendedAttributes.txt)

| | | |
|----------------|---------|---|
| FieldOfStudyId | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |
| AttributeType | bigint | Possible values: 1 (AUI from UMLS https://www.nlm.nih.gov/research/umls/licensedcontent/umls_archives04.html#2018AA), 2 (source url), 3 (CUI from UMLS https://www.nlm.nih.gov/research/umls/licensedcontent/umls_knowledgesources.html) |
| AttributeValue | varchar | |

Indexes

| | |
|--------------|----------------|
| Distribution | FieldOfStudyId |
| Sorting | FieldOfStudyId |

Foreign Keys

Vir Fk_FieldOfStudyExtendedAttributes_FieldsOfStudy (FieldOfStudyId) ref FieldsOfStudy (FieldOfStudyId)

Table FieldsOfStudy

Base table for Fields of Study (advanced/FieldsOfStudy.txt)

| | | |
|------------------|---------|---|
| FieldOfStudyId | bigint | PRIMARY KEY |
| Rank | varchar | FROZEN; no new ranks are being added. |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| MainType | varchar | |
| Level | integer | Possible values: 0-5 |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count." |
| CitationCount | bigint | |
| CreatedDate | varchar | |

Indexes

| | |
|--------------|----------------|
| Distribution | FieldOfStudyId |
| Sorting | FieldOfStudyId |

Table Journals

Base table for Journals (mag/Journals.txt)

| | | |
|------------------|-----------|--|
| JournalId | bigint | PRIMARY KEY |
| Rank | integer | FROZEN; no new ranks are being added |
| NormalizedName | varchar | UPDATED; slightly different normalization algorithm |
| DisplayName | varchar | |
| Issn | varchar | UPDATED; 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: <code>['1469-5073','0016-6723']</code>) |
| IsOa | boolean | NEW; TRUE when the journal is 100% OA |
| IsInDoaj | boolean | NEW; TRUE when the journal is in DOAJ (see https://doaj.org/) |
| Publisher | varchar | |
| Webpage | varchar | |
| PaperCount | bigint | |
| PaperFamilyCount | bigint | FROZEN; same value as "paper_count" |
| CitationCount | bigint | |
| CreatedDate | varchar | |
| UpdatedDate | timestamp | NEW; set when changes are made going forward |

Indexes

| | |
|--------------|-----------|
| Distribution | JournalId |
| Sorting | JournalId |

Table PaperAbstractsInvertedIndex

Inverted abstracts (nlp/PaperAbstractsInvertedIndex.txt).

| | | |
|-----------------|---------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES papers.papers_id |
| IndexedAbstract | varchar | Inverted index, see https://en.wikipedia.org/wiki/Inverted_index |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperAbstractsInvertedIndex_Papers (PaperId)
ref Papers (PaperId)

Table PaperAuthorAffiliations

Links between papers, authors, and affiliations/institutions. NOTE: It is possible to have multiple rows with same (PaperId, AuthorId, AffiliationId) when an author is associated with multiple affiliations. (mag/PaperAuthorAffiliations.txt)

| | | |
|----------------------|---------|--|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| AuthorId | bigint | FOREIGN KEY REFERENCES Authors.AuthorId |
| AffiliationId | bigint | FOREIGN KEY REFERENCES Affiliations.AffiliationId |
| AuthorSequenceNumber | integer | 1-based author sequence number. 1: the 1st author listed on paper, 2: the 2nd author listed on paper, etc. |
| OriginalAuthor | varchar | |
| OriginalAffiliation | varchar | |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperAuthorAffiliations_Papers (PaperId) ref
Papers (PaperId)

Vir Fk_PaperAuthorAffiliations_Authors (AuthorId) ref
Authors (AuthorId)

Vir Fk_PaperAuthorAffiliations_Affiliations (AffiliationId)
ref Affiliations (AffiliationId)

Table PaperCitationContexts

FROZEN; citation contexts (nlp/PaperCitationContexts.txt)

| | | |
|------------------|---------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES papers.papers_id |
| PaperReferenceld | bigint | FOREIGN KEY REFERENCES papers.papers_id |
| CitationContext | varchar | |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperCitationContexts_Papers (PaperId) ref Papers (PaperId)

Vir Fk_PaperCitationContexts_Papers_0 (PaperReferenceld) ref Papers (PaperId)

Table PaperExtendedAttributes

Extra paper identifiers (mag/PaperExtendedAttributes.txt)

| | | |
|----------------|---------|--|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| AttributeType | integer | Possible values: 1=PatentId, 2=PubMedId, 3=Pmclid, 4=Alternative Title |
| AttributeValue | varchar | |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperExtendedAttributes_Papers (PaperId) ref Papers (PaperId)

Table PaperFieldsOfStudy

Linking table from papers to fields, with score (advanced/PaperFieldsOfStudy.txt)

| | | |
|------------------|---------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| FieldOfStudyId | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |
| Score | real | Confidence range between 0 and 1. Bigger number representing higher confidence. |
| AlgorithmVersion | integer | NEW; version of algorithm to assign fields. Possible values: 1=old MAG (FROZEN), 2=OpenAlex |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperFieldsOfStudy_Papers (PaperId) ref Papers (PaperId)

Vir Fk_PaperFieldsOfStudy_FieldsOfStudy (FieldOfStudyId) ref FieldsOfStudy (FieldOfStudyId)

Table PaperMeSH

MeSH headings assigned to the paper by PubMed (advanced/PaperMeSH.txt)

| | | |
|----------------|---------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| DescriptorUI | varchar | see https://en.wikipedia.org/wiki/Medical_Subject_Headings |
| DescriptorName | varchar | see https://en.wikipedia.org/wiki/Medical_Subject_Headings |
| QualifierUI | varchar | see https://en.wikipedia.org/wiki/Medical_Subject_Headings |
| QualifierName | varchar | see https://en.wikipedia.org/wiki/Medical_Subject_Headings |
| IsMajorTopic | boolean | see https://en.wikipedia.org/wiki/Medical_Subject_Headings |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Table PaperMeSH

Vir Fk_PaperMeSH_Papers (PaperId) ref Papers (PaperId)

Table PaperRecommendations

Paper recommendations with score (advanced/PaperRecommendations.txt)

| | | |
|--------------------|--------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| RecommendedPaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| Score | real | Confidence range between 0 and 1. Bigger number representing higher confidence. |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperRecommendations_Papers (PaperId) ref Papers (PaperId)

Vir Fk_PaperRecommendations_Papers_0 (RecommendedPaperId) ref Papers (PaperId)

Table PaperReferences

Paper references and, in reverse, citations (mag/PaperReferences.txt)

| | | |
|------------------|--------|---------------------------------------|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| PaperReferenceId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperReferences_Papers (PaperId) ref Papers (PaperId)

Vir Fk_PaperReferences_Papers_0 (PaperReferenceId) ref Papers (PaperId)

Table PaperResources

FROZEN; no longer updated. Data and code urls associated with papers (mag/PaperResources.txt)

| | | |
|------------------|---------|---|
| PaperId | bigint | FOREIGN KEY REFERENCES papers.papers_id |
| ResourceType | integer | Bit flags: 1=Project, 2=Data, 4=Code |
| ResourceUrl | varchar | Url of resource |
| SourceUrl | varchar | List of urls associated with the project, used to derive resource_url |
| RelationshipType | integer | Bit flags: 1=Own, 2=Cite |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperResources_Papers (PaperId) ref Papers (PaperId)

Table PaperUrls

Urls for the paper (mag/PaperUrls.txt)

| | | |
|-------------------|---------|--|
| PaperId | bigint | FOREIGN KEY REFERENCES Papers.PaperId |
| SourceType | 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 |
| SourceUrl | varchar | |
| LanguageCode | varchar | |
| UrlForLandingPage | varchar | NEW; URL for the landing page, when article is free to read |

Table PaperUrls

| | | |
|-----------------------|---------|---|
| UrlForPdf | varchar | NEW; URL for the PDF, when article is free to read |
| HostType | 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) |
| RepositoryInstitution | varchar | NEW; name of repository host of URL |
| OaiPmhId | varchar | NEW; OAH-PMH id of the repository record |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

Vir Fk_PaperUrls_Papers (PaperId) ref Papers (PaperId)

Table Papers

Main data for papers (mag/Papers.txt)

| | | |
|----------------------|---------|--|
| PaperId | bigint | PRIMARY KEY |
| Rank | integer | FROZEN; no new ranks are being added |
| Doi | varchar | Doi values are upper-cased per DOI standard at https://www.doi.org/doi_handbook/2_Numbering.html#2.4 |
| DocType | varchar | Possible values: Book, BookChapter, Conference, Dataset, Journal, Patent, Repository, Thesis, NULL : unknown. Patent is FROZEN; no new Patents are being added. |
| Genre | varchar | NEW; Crossref ontology for work type such as "journal-article, posted-content, dataset, or book-chapter" |
| IsParatext | boolean | NEW; indicates front-matter. See https://support.unpaywall.org/support/solutions/articles/44001894783 |
| PaperTitle | varchar | UPDATED; slightly different normalization algorithm |
| OriginalTitle | varchar | |
| BookTitle | varchar | |
| Year | integer | |
| Date | varchar | |
| OnlineDate | varchar | |
| Publisher | varchar | |
| JournalId | bigint | FOREIGN KEY references journals.journal_id |
| ConferenceSeriesId | bigint | FROZEN; no longer updated, no new Conference Series are being added. FOREIGN KEY references conference_series.conference_series_id. |
| ConferenceInstanceId | bigint | FROZEN; no longer updated, no new Conference Instances are being added. FOREIGN KEY references conference_instances.conference_instance_id. |
| Volume | varchar | |
| Issue | varchar | |
| FirstPage | varchar | |
| LastPage | varchar | |
| ReferenceCount | bigint | |
| CitationCount | bigint | |
| EstimatedCitation | bigint | UPDATED; new algorithm |
| OriginalVenue | varchar | |
| FamilyId | bigint | FROZEN; no longer updated. |
| FamilyRank | bigint | FROZEN; no longer updated. |
| DocSubTypes | varchar | Possible values: Retracted Publication, Retraction Notice |
| OaStatus | varchar | NEW; Possible values: closed, green, gold, hybrid, bronze (see https://en.wikipedia.org/wiki/Open_access#Colour_naming_system) |
| BestUrl | varchar | NEW; An url for the paper (see paper_urls table for more) |

Table Papers

| | | |
|-----------------|-----------|--|
| BestFreeUrl | varchar | NEW; Url of best legal free-to-read copy when it exists (see https://support.unpaywall.org/support/solutions/articles/44001943223) |
| BestFreeVersion | varchar | NEW; Possible values: submittedVersion, acceptedVersion, publishedVersion |
| DoiLower | varchar | NEW; lowercase doi for convenience linking to Unpaywall |
| CreatedDate | varchar | |
| UpdatedDate | timestamp | NEW; set when changes are made going forward |

Indexes

| | |
|--------------|---------|
| Distribution | PaperId |
| Sorting | PaperId |

Foreign Keys

| | |
|-----|---|
| Vir | Fk_Papers_Journals (JournalId) ref Journals (JournalId) |
| Vir | Fk_Papers_ConferenceSeries (ConferenceSeriesId) ref ConferenceSeries (ConferenceSeriesId) |
| Vir | Fk_Papers_ConferenceInstances (ConferenceInstanceId) ref ConferenceInstances (ConferenceInstanceId) |

Table RelatedFieldOfStudy

Relationships between fields of study (advanced/RelatedFieldOfStudy.txt)

| | | |
|-----------------|---------|--|
| FieldOfStudyId1 | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |
| Type1 | varchar | Possible values: general, disease, disease_cause, medical_treatment, symptom |
| FieldOfStudyId2 | bigint | FOREIGN KEY REFERENCES FieldsOfStudy.FieldOfStudyId |
| Type2 | varchar | Possible values: general, disease, disease_cause, medical_treatment, symptom |
| Rank | real | FROZEN; no new ranks are being added. |

Indexes

| | |
|--------------|-----------------|
| Distribution | FieldOfStudyId1 |
| Sorting | FieldOfStudyId1 |

Foreign Keys

| | |
|-----|---|
| Vir | Fk_RelatedFieldOfStudy_FieldsOfStudy (FieldOfStudyId1) ref FieldsOfStudy (FieldOfStudyId) |
| Vir | Fk_RelatedFieldOfStudy_FieldsOfStudy_0 (FieldOfStudyId2) ref FieldsOfStudy (FieldOfStudyId) |