

PubMed to VIVO XSL Information

Overview

The definition of attributes from Pubmed and their transformation to elements in the VIVO ontology's schema. PubMed uses the Medline schema for storing citations. Medline is the Medical Literature Analysis and Retrieval System Online (Medlars Online).

Development Documents

Medline/PubMed Resource Guide - <http://www.nlm.nih.gov/bsd/pmresources.html#data>

Medline/PubMed XML Data Elements - http://www.nlm.nih.gov/bsd/licensee/data_elements_doc.html

XSLT Tutorial - http://www.w3schools.com/xsl/xsl_languages.asp

Xpath Tutorial - <http://www.w3schools.com/xpath/default.asp>

XSL-FO Tutorial - <http://www.w3schools.com/xslfo/default.asp>

Medline to VIVO

See also the [attached xsl file](#) (last updated 2013-12-09) for mapping Scopus and MEDLINE bibliographic records to VIVO. Note that there are some minor differences between the mappings below and the attached file.

Element /Attribute	Name	Description	VIVO /Scoring	Equivalence in RDF	Appears In
E	Medline Citation	Top Level element in a MedlineCitationSet and contains one publication	V	<div><rfd:Description></div>	RDF Element
A	Owner	Attribute of Medline Citation, the owner of the article is the party responsible for creating and validating the citation. There are only 8 possible owners	NLM, NASA, PIP, KIE, HSR, HMD, SIS, NOTNLM	N/A	(VD: not necessary unless needed for scoring)
A	Status	The status is the stage the citation is currently in there are seven values; Completed, In-Process, PubMed-not-Medline, In-Data-Review, Publisher, MEDLINE, OLDMEDLINE	S	<div><score:PubMedStatus></div>	Main Citation (VD: not needed unless for scoring. We only want to pull in "completed")
E	PMID	The Pubmed unique identifier	V	<div><bibo:pmid></div>	Main Citation
E	DateCreated	The date the processing of the citation began. It has three sub elements Year, Month, and Day	S	<div><score:DateCreated></div>	Main Citation
E	DateCompleted	Date processing of the record ends	S	<div><score:DateCompleted></div>	Main Citation

E	DateRevised	Identifies the date a change has been made. Only occurs in records that are MEDLINE, OLDMEDLINE, PubMed-not-MEDLINE	S	<score:DateRevised>	Main Citation
E	Article	The wrapper around the non medline information about an article, like Authors and Title	N/A		
A	PubModel	This describes how the publication was distributed and is apart of Article. There exist four values; Print, Print-Electronic, Electronic, Electronic-Print	UnK		(VD: Change to N/A)
E	Journal	The envelope around the Journal information about an Article	V	<rdf:Description rdf:about="http://viovweb.org/PubMed/journal/j{ISSN}>	(VD: This is only an envelope. Not needed?)
E	ISSN	The ISSN, International Standards Serial Number is an 8 charcater value that identifies the journal	V	<bibo:ISSN>	Journal Entry
A	ISSNType	Journals that have multiple distributions have multiple ISSNs, there are two types; Print and Electronic	UnK		(VD: change to N/A)
E	JournalIssue	Contains information about the specific issue the article resides in			Journal Entry (VD: N/A)
A	CitationMedium	The medium that was used to process the article. Two values; Internet or Print	UnK		(VD: N/A)
E	Volume	The volume number that the article resides	V	<bibo:volume>	Main Citation
E	Issue	Identifies the issue that the journal resides in	V	<bibo:issue>	Main Citation
E	PubDate	The publication date for the article has sub elements, year, month, and day and medlinedate	UnK		(VD: data should be parsed into core:year, core:yearMonth, and core:date)
E	MedlineDate	Dates that can not be parsed will be found in this sub element of PubDate	N/A		
E	Title	The title of the Journal that the article was published in	V	<core:Title>	Journal Entry
E	ISOAbbreviation	Used for linking from GenBank to PubMed			(VD: N/A)
E	ArticleTitle	The entire title of the Journal Article	V	<core:Title>	Main Citation
E	Pagination	The pages that the article occurred on	V	<bibo:pages><bibo:pageStart><bibo:pageEnd>	Main Citation
E	ElocationID	Procides an electronic location for items which lack standard page numbers	UnK		(VD: N/A)
E	Abstract	Wrapper around information about the Abstract	N/A		

E	Abstract Text	Actual Text of the abstract, not always available	V	<bibo:abstract>	Main Citation
E	Copyright Information	Appears for articles where the publisher or author claim copyright on the citation	N/A		
E	Affiliation	A string containing the affiliation information for the 1st Author	S	<score: Affiliation >	Main Citation
E	AuthorList	Wrapper around the list of authors for a citation	N/A		
E	Author	The Author of the Article in order as they appear in the publication sub elements include lastname, forename, suffix and initials	V	<rdf: Description><rdf:Type foaf: Person />	Individual Author
E	LastName	The Surname of the author	V	<foaf:lastName>	Individual Author
E	ForeName	The remainder of the name excluding the suffix, this is often only the initials of an individual spaced out	S	<score:foreName>	Individual Author
E	Suffix	A persons suffix NOT honorific	S	<score:suffix>	Individual Author
E	Initials	Often identical to the ForeName without spaces, this contains 2 initials for a person	S	<score:initials>	Individual Author
E	Language	Three letter description of the language the article was published in	UnK		(VD: N/A)
E	Databank	Contains information pertaining to the registration of several types of data. this data includes: molecular sequence data, clinical trial numbers, gene expression/molecular abundance data, and PubChem. This field is being ignored for now.	N/A		
?	DOI	This is the DOI, as listed in PubMed.	?		(VD: use bibo:doi)
E	GrantList	Wrapper for the list of grants related to the publication.	N/A		
E	Grant	Wrapper for the information about a single grant. Expect Inclusion of this data property in milestone 4.	UnK		Individual Grant (VD: N/A)
E	GrantID	The contract number of the grant	UnK		Individual Grant (VD: Is it possible to link a publication entity to a grant entity using the grant ID property "core: sponsorAwardId")
E	Acronym	The 2 letter acronym of the grant	UnK		Individual Grant (VD: N/A)
E	Agency	The agency that supplied the grant	UnK		Individual Grant (VD: N/A)
E	Country	Country of the grant agency	UnK		Individual Grant (VD: N/A)
E	PublicationTypeList	Wrapper containing a list of article types	N/A		
E	PublicationType	V	see below for map of types to vivo classes	Main Citation	
E	VernacularTitle	The title of the publication in its originating language.	UnK		(VD: N/A)

E	ArticleDate	the date the publisher made an electronic version of the article	UnK		(VD: N/A)
E	Country	the country the article was published in	UnK		(VD: N/A)
E	MedlineTA	The Medline Title Abbreviation for the Journal. The site http://www.nlm.nih.gov/pubs/factsheets/constructitle.html contains all of the rules used by NLM to abbreviate a title.	UnK		(VD: N/A)
E	NlmUniqueID	the accession number for the journal the article was written in.	UnK		(VD: N/A)
E	ISSNLinking	the ISSN designated by the ISSN Network to enable co-location or linking among the different media versions of a continuing resource	UnK		(VD: N/A)
E	ChemicalList	The wrapper for a list of chemicals	UnK		(VD: N/A)
E	Chemical	the chemical being specified in the NLM record	UnK		(VD: use core:freetextKeyword)
E	RegistryNumber	a unique 5-9 digit number in a hyphenated format assigned by the Chemical Abstract Services.	UnK		(VD: N/A)
E	Name of Substance	the name of the substance that the registry number identifies	UnK		(VD: use core:freetextKeyword)
E	CitationSubset	The subset that Medline records from certain sources or on certain topics were assigned.	UnK		(VD: N/A)
E	CommentsCorrections	This section details various citation notes, its unknown how this would be used by scoring and even more unknown if its needed in the main citation on VIVO.	UnK		(VD: N/A)
E	GeneSymbolList	The wrapper for gene's listed in a publication	UnK		(VD: N/A)
E	GeneSymbol	The abbreviated form of the gene name as reported in the literature	UnK		(VD: N/A)
E	MeshHeadingList	Wrapper for the list of Mesh Terms associated with an article	N/A		
E	MeshHeading	The individual Mesh Term for a publication, it has a descriptor and qualifier.	S	<div> <rdf:Description><rdf:Type score:MeshTerm> </div>	Mesh Term
E	DescriptorName	The main term, if its a central concept of the paper its marked as MajorTopic.	S	<div> <score:Descriptor> </div>	Mesh Term
E	QualifierName	A sub heading of the Descriptor Name.	S	<div> <score:Qualifier </div>	Mesh Term
E	NumberOfReferences	Number of Bibliographic references in the article	N/A		(VD: N/A)
E	PersonalNameSubjectList	The names of people that appear in the article (such as biographical notes or obituaries).	N/A		(VD: N/A)
E	OtherID	ID's of other collaborating partners with NLM. This information is not used at the present time but may be added later.	UnK		(VD: N/A)
E	OtherAbstract	Alternative abstracts created by a collaborating partner or other entity	UnK		(VD: N/A)
E	KeywordList	A list of Keywords assigned by a collaborating data producer, or in the case of oldmedline before there was mesh by NLM itself.	UnK		(VD: use core:freetextKeyword)
E	SpaceFlightMission	If a citation was created by NASA it may come with information on the space flight mission that the research was conducted on.	UnK		(VD: N/A)

E	InvestigatorList	For citations created by NASA originally, and other data partners later, the investigator is the PI on the grant that the research was produced. This is stored in the same format as the Author list	UnK		(VD: N/A)
E	General Note	Descriptive information provided by an NLM partner.	UnK		(VD: N/A)
E	DeleteCitation	Most of the data is removed, but the stub is marked as deleted. We shouldn't be ingesting these publications.	N/A		

Publication Type Translation

Publication Characteristics (Publication Types) - Scope Notes

Abbreviations – DO NOT INCLUDE IN VIVO

Abstracts - DO NOT INCLUDE IN VIVO

Academic Dissertations - DO NOT INCLUDE IN VIVO

Account Books - DO NOT INCLUDE IN VIVO

Addresses – map to core:Conference Paper

Advertisements - DO NOT INCLUDE IN VIVO

Almanacs - DO NOT INCLUDE IN VIVO

Anecdotes - DO NOT INCLUDE IN VIVO

Animation - DO NOT INCLUDE IN VIVO

Annual Reports - DO NOT INCLUDE IN VIVO

Aphorisms and Proverbs - DO NOT INCLUDE IN VIVO

Architectural Drawings - DO NOT INCLUDE IN VIVO

Atlases – map to bibo:Reference Source

Bibliography – map to bibo:Academic Article

Biobibliography - map to bibo:Academic Article

Biography – map to bibo:Book

Book Illustrations - DO NOT INCLUDE IN VIVO

Book Reviews – map to core:Review

Bookplates - DO NOT INCLUDE IN VIVO

Broadsides - DO NOT INCLUDE IN VIVO

Caricatures - DO NOT INCLUDE IN VIVO

Cartoons - DO NOT INCLUDE IN VIVO

Case Reports – map to core:Case Study

Catalogs - DO NOT INCLUDE IN VIVO

Charts – map to bibo:Image

Chronology - DO NOT INCLUDE IN VIVO

Classical Article – map to bibo:Academic Article

Clinical Conference – map to bibo:Academic Article

Clinical Trial – map to bibo:Academic Article

Clinical Trial, Phase I - map to bibo:Academic Article

Clinical Trial, Phase II - map to bibo:Academic Article

Clinical Trial, Phase III - map to bibo:Academic Article

Clinical Trial, Phase IV - map to bibo:Academic Article

Collected Correspondence – map to bibo:Article and bibo:Book

Collected Works - map to bibo:Article and bibo:Book

Collections - DO NOT INCLUDE IN VIVO

Comment – map to bibo:Academic Article

Comparative Study – map to bibo:Academic Article

Congresses – map to bibo:Proceedings

Consensus Development Conference – map to bibo:Proceedings

Consensus Development Conference, NIH – map to bibo:Proceedings

Controlled Clinical Trial – map to bibo:Academic Article

Cookbooks – map to bibo:Book

Corrected and Republished Article - map to bibo:Academic Article

Database – map to core:Database

Dataset – map to vivo:Dataset

Diaries – map to bibo:Book

Dictionary – map to bibo:Reference Source

Directory - map to bibo:Reference Source

Documentaries and Factual Films – map to bibo:Film

Drawings – map to bibo:Image

Duplicate Publication – map to bibo:Article and bibo:Book

Editorial – map to core:Editorial Article

Encyclopedias - map to bibo:Reference Source

English Abstract - DO NOT INCLUDE IN VIVO

Ephemera - DO NOT INCLUDE IN VIVO

Essays – map to bibo:Article and bibo:Book

Eulogies - DO NOT INCLUDE IN VIVO

Evaluation Studies – map to bibo:Academic Article

Examination Questions - DO NOT INCLUDE IN VIVO

Exhibitions - DO NOT INCLUDE IN VIVO

Festschrift – map to bibo:Book

Fictional Works – map to bibo:Article and bibo:Book

Forms - DO NOT INCLUDE IN VIVO

Formularies – map to bibo:Book

Funeral Sermons - DO NOT INCLUDE IN VIVO

Government Publications – new publication class needed

Guidebooks – map to bibo:Reference Source

Guideline – map to bibo:Standard

Handbooks – map to bibo:Reference Source

Herbals - DO NOT INCLUDE IN VIVO

Historical Article – map to bibo:Article

Humor - DO NOT INCLUDE IN VIVO

In Vitro - DO NOT INCLUDE IN VIVO

Incunabula – map to bibo:Book

Indexes – map to bibo:Reference Source

Instruction – map to bibo:Audio Visual Document

Interactive Tutorial – need new pub class for multimedia

Interview – map to bibo:Book and bibo:Article

Introductory Journal Article – map to bibo:Article

Journal Article – map to bibo:Academic Article

Juvenile Literature - DO NOT INCLUDE IN VIVO

Laboratory Manuals - DO NOT INCLUDE IN VIVO

Lecture Notes - DO NOT INCLUDE IN VIVO

Lectures – need new pub class for speeches??

Legal Cases – map to bibo:Legal Case Document

Legislation – map to bibo:Legislation

Letter - DO NOT INCLUDE IN VIVO

Manuscripts - DO NOT INCLUDE IN VIVO

Maps – map to bibo:Map

Meeting Abstracts – map to bibo:Conference Proceedings

Meta-Analysis – map to bibo:Academic Article

Monograph - DO NOT INCLUDE IN VIVO

Multicenter Study – map to bibo:Academic Article

News – map to bibo:Article

Newspaper Article – map to bibo:Article

Nurses' Instruction - DO NOT INCLUDE IN VIVO

Observational Study – map to bibo:Academic Article

Outlines - DO NOT INCLUDE IN VIVO

Overall - DO NOT INCLUDE IN VIVO

Patents – map to bibo:Patent

Patient Education Handout – map to bibo:Document

Periodical Index – map to bibo:Reference Source

Periodicals – map to bibo:Periodical

Personal Narratives - DO NOT INCLUDE IN VIVO

Pharmacopoeias – map to bibo:Reference Source

Phrases - DO NOT INCLUDE IN VIVO

Pictorial Works – map to bibo:Article and bibo:Book

Poetry - DO NOT INCLUDE IN VIVO

Popular Works – DO NOT INCLUDE IN VIVO

Portraits – map to ~~bibo:Image~~ bibo:Image vivo:Article

Postcards - DO NOT INCLUDE IN VIVO

Posters - DO NOT INCLUDE IN VIVO

Practice Guideline – map to bibo:Standard

Pragmatic Clinical Trial – map to bibo:Academic Article

Price Lists - DO NOT INCLUDE IN VIVO

Problems and Exercises - DO NOT INCLUDE IN VIVO

Programmed Instruction - DO NOT INCLUDE IN VIVO

Programs - DO NOT INCLUDE IN VIVO

Prospectuses - DO NOT INCLUDE IN VIVO

Publication Components - DO NOT INCLUDE IN VIVO

Publication Formats - DO NOT INCLUDE IN VIVO

Published Erratum - DO NOT INCLUDE IN VIVO

Randomized Controlled Trial – map to bibo:Academic Article

Research Support, American Recovery and Reinvestment Act - DO NOT INCLUDE IN VIVO

Research Support, N.I.H., Extramural - DO NOT INCLUDE IN VIVO

Research Support, N.I.H., Intramural - DO NOT INCLUDE IN VIVO

Research Support, Non-U.S. Gov't - DO NOT INCLUDE IN VIVO

Research Support, U.S. Gov't, Non-P.H.S. - DO NOT INCLUDE IN VIVO

Research Support, U.S. Gov't, P.H.S. - DO NOT INCLUDE IN VIVO

Resource Guides - DO NOT INCLUDE IN VIVO

Retracted Publication - DO NOT INCLUDE IN VIVO

Retraction of Publication - DO NOT INCLUDE IN VIVO

Review – map to core:Review and bibo:Book

Scientific Integrity Review - DO NOT INCLUDE IN VIVO

Sermons - DO NOT INCLUDE IN VIVO

Statistics – map to bibo:Book and bibo:Article

Study Characteristics – map to bibo:Article

Support of Research - DO NOT INCLUDE IN VIVO

Tables - DO NOT INCLUDE IN VIVO

Technical Report – map to bibo:Report

Terminology - DO NOT INCLUDE IN VIVO

Textbooks – map to bibo:Book

Twin Study – map to bibo:Academic Article

Unedited Footage - DO NOT INCLUDE IN VIVO

Union Lists - DO NOT INCLUDE IN VIVO

Unpublished Works - DO NOT INCLUDE IN VIVO

Validation Studies - map to bibo:Academic Article

Webcasts - need new pub class for multimedia