

# Scopus Custom Data

## SCOPUS CUSTOM DATA DOCUMENTATION – csv/txt formats

---

Version 3  
January 2021

*Peter Berkvens*  
Senior Product Manager  
Scopus Content and Databases  
[p.berkvens@elsevier.com](mailto:p.berkvens@elsevier.com)

**CONTENTS**

I. Introduction .....	3
II. Scopus Custom Data XML flavours .....	3
III. Other formats .....	4
IV. csv format .....	4
V. txt format.....	7

## **I. Introduction**

### **Custom Data facts and figures**

Scopus is the largest Abstract and Indexing database worldwide. It is important to bear in mind that the database is continuously growing at ~8% CAGR (Compound Annual Growth Rate). Currently the database contains over 81 million records. Records with publication dates before 1970 do not have references.

Traditional Scopus Custom data deliveries are done in XML format. XML ensures that articles get delivered in a highly structured format and makes it possible to host the data in a database. If you are interested in XML please refer to the respective XML documentation available on request. The XML flavours that can be delivered are listed hereunder.

## **II. Scopus Custom Data XML flavours**

Scopus Custom Data is offered in four different flavours depending on the way customers want to use and expose the data, and/or on the requirements customers have for incorporation of the Scopus data in their database.

### **Full XML**

Scopus Custom Data's primary deliverable meant to be used by Bibliometricians and Data Analysts. Format cannot be used for exposure to other parties than the Bibliometric/Analytical teams handling the data. This XML 'flavour' contains all below mentioned elements and sub-elements and is the richest possible format.

### **Institutional Repository XML format**

Scopus Custom data used to add data to Institutional Repositories. Since this data is accessible for much larger groups of users, fields that fall under copyright restrictions are left out. The latter refers to the following elements:

- (1) All data covered by the <enhancement> element such as descriptor groups containing controlled vocabulary
- (2) All data covered by the <tail> element containing citing References
- (3) All data covered by the <abstract> element containing the non-Elsevier Abstracts
- (4) All data covered by <correspondence>, <additional-srcinfo>, <citation-language> and <abstract-language> elements

Simply said this means that this format will not contain References, Controlled Vocabulary and other than Elsevier Abstracts.

### **Light XML format**

For clients that want simple metadata sets. The following elements are left out:

- (1) All data covered by the <enhancement> element such as descriptor groups containing controlled vocabulary
- (2) All data covered by the <tail> element containing citing References
- (3) All data covered by the <abstract> element containing the non-Elsevier Abstracts

### **Light XML format plus Abstracts**

For clients that want simple metadata sets incl. Abstracts. The following elements are left out:

- (1) All data covered by the <enhancement> element such as descriptor groups containing controlled vocabulary
- (2) All data covered by the <tail> element containing citing References

### **III. Other formats**

In addition to the above-mentioned XML flavours, Scopus Custom Data is also available in txt and csv format. Both are identical to the data formats delivered when using the ‘all available information’ option listed under the export link in scopus.com results lists. These simple formats can be exported into for instance Excel and Access and do not require knowledge of database technology. Full export format in scopus.com is limited to 2000 items per export action. There are no limits in numbers of items delivered through Scopus Custom Data csv or txt format options.

### **IV. csv format**

The Scopus Custom data csv format offers the following metadata categories as column headers in csv/excel:

#### **Author names**

Names and initials of all authors of an article. No links between author and address.

*Format:* Columbus C., Cook, J.

#### **Affiliations**

Affiliations of all authors. No links between author and address.

*Format:* Department of Agriculture, University of Madrid, Madrid, Spain; Department of Chemistry, Free University, P.O. Box 12, Amsterdam, Netherlands

#### **Authors with affiliations**

Authors and affiliations linked together.

*Format:* Columbus C., Department of Agriculture, University of Madrid, Madrid, Spain; Cook, J., Department of Chemistry, Free University, P.O. Box 12, Amsterdam, Netherlands

#### **Title**

Title of article.

#### **Year**

Year of publication of article.

#### **Source title**

Journal/Proceedings/Book Series title in which article has been published.

#### **Volume**

Volume number of issue in which article has been published.

#### **Issue**

Issue number of issue of volume in which article has been published.

#### **Art. No.**

Article number of article used in journals that do not have traditional page numbering anymore. Example: PLoS One article number ‘e90514’.

**Page start**

First page of article.

**Page end**

Last page of article.

**Page count**

Number of pages of article if available from source.

**Cited by**

Cited-by count taken from Scopus the moment the data was extracted.

**Link**

Link to article in scopus.com.

**Abstract**

Complete Abstract of article when available in source.

*Format:* Certain foodstuffs exhibit matrix interference effects on the vitamin B1 analysis prescribed in the official methods of the European Union, AOAC International, and Japan. In this study, we demonstrated that one of the problematic polyphenols in green tea or cocoa was tannin. From these results, the method was found to be effective for vitamin B1 analysis regardless of the presence of interference matrices. © 2012 Elsevier Ltd. All rights reserved.

**Author Keywords**

Uncontrolled keywords assigned to the article by author(s).

**Index Keywords**

Controlled keywords originating from for instance EMBASE or Compendex thesaurus.

**Molecular Sequence Numbers**

GENBANK numbers.

*Format:* GENBANK: EU637012, JN834015, JN834016, JN834017, JN834018, JN834019, JN834020, JN834021, JN834022, JN834023, JN834024, JN834025, JN834026

**Chemicals/CAS**

CAS registry numbers.

*Format:* 25 hydroxyvitamin D, 64719-49-9; 25-hydroxyvitamin D, 64719-49-9; Vitamin D, 1406-16-2

**Tradenames**

Set of tradenames of a specific type that occur in a document. Only used if article is indexed for a particular database containing Tradenames.

*Format:* cellcept, Hoffmann La Roche, Netherlands; neoral, Novartis, Netherlands; sandimmune, Novartis

**Manufacturers**

Set of manufacturers. Only used if article is indexed for a particular database containing manufacturers.

*Format:* Hoffmann La Roche, Netherlands; Novartis, Netherlands

**Funding Details**

Funding body/sponsor, acronym and Grant number.

*Format:* AFRI 2011-67009-20049, NSF, National Science Foundation; Environmental Protection Agency; 9816232, NSF, National Science Foundation; 0070183, NSF, National Science Foundation

**References**

All references from an article.

*Format:* Bernhard, J.M., Characteristic assemblages and morphologies of benthic foraminifera from anoxic, organic-rich deposits: Jurassic through Holocene (1986) Journal of Foraminiferal Research, 16, pp. 207-215; Gibson, T.G., Buzas, M.A., Species diversity: Patterns in modern and miocene foraminifera of the eastern margin of north america (1973) Geological Society of America Bulletin, 84, pp. 217-238; Glover, A.G., Smith, C.R., Paterson, G.L.J., Wilson, G.D.F., Hawkins, L., Sheader, M., Polychaete species diversity in the central Pacific abyss: Local and regional patterns, and relationships with productivity (2002) Marine Ecology Progress Series, 240, pp. 157-170

**Correspondence Address**

Addresses of corresponding authors.

*Format:* Jones, L.S.; Department of Industrial and Physical Pharmacy, College of Pharmacy, Purdue University, West Lafayette, IN, United States; email: xjones@pppp.edu

**Editors**

Names of Editors of for instance a Proceedings or Special Issue.

**Sponsors**

Names of Sponsors of for instance a Proceedings or Special Issue.

**Publisher**

Publisher. Fully indexed in Scopus after 2014.

**Conference name**

Name of Conference.

**Conference date**

Date Conference was held.

**Conference location**

Location Conference was held.

**Conference code**

Code identifying a Conference.

**ISSN**

ISSN of serial publication article was published in.

**ISBN**

ISBN of Book item originates from.

**CODEN**

Six-letter code assigned by ASCII identifying a serial publication.

**DOI**

Digital Object Identifier. Unique CrossRef assigned number/letter combination of an item.

*Format:* 10.1287/ijoc.2013.0549

**PubMed ID**

Unique PubMed assigned ID.

*Format:* 22980817

**Language of Original Document**

Language the original document was published in.

*Format:* English

**Abbreviated Source Title**

Abbreviated source title. Not standardized but most of the time following ASCII.

*Format:* Anal. Chim. Acta

**Document Type**

Document type (e.g., Article, Review, Book Chapter, Conference Paper)

*Format:* Article

**Source**

Default is 'Scopus'. Database the data has been extracted from.

**V. txt format**

The Scopus Custom data txt format is a flat file containing all above metadata presented in a flat untagged format.

The following page shows an example txt format output file of one particular article.

EXPORT DATE:04 Mar 2014

An, H.a , Pospelov, M.a b , Pradler, J.c

New stellar constraints on dark photons

(2013) Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics, 725 (4-5), pp. 190-195. Cited 3 times.

[http://www.scopus.com/inward/record.url?eid=2-s2.0-](http://www.scopus.com/inward/record.url?eid=2-s2.0-84883054151&partnerID=40&md5=a5d96ffce18a5a624272ad39d7b72b74)[84883054151&partnerID=40&md5=a5d96ffce18a5a624272ad39d7b72b74](http://www.scopus.com/inward/record.url?eid=2-s2.0-84883054151&partnerID=40&md5=a5d96ffce18a5a624272ad39d7b72b74)

AFFILIATIONS: Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada;

Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 1A1, Canada;

Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21210, United States

ABSTRACT: We consider the stellar production of vector states  $V$  within the minimal model of "dark photons". We show that when the Stückelberg mass of the dark vector becomes smaller than plasma frequency, the emission rate is dominated by the production of the longitudinal modes of  $V$ , and scales as  $\kappa^2 m_V^2$ , where  $\kappa$  and  $m_V$  are the mixing angle with the photon and the mass of the dark state. This is in contrast with widespread assertions in the literature that the emission rate decouples as the forth power of the mass. We derive ensuing constraints on the  $(\kappa, m_V)$  parameter space by calculating the cooling rates for the Sun and horizontal branch stars. We find that stellar bounds for  $m_V < 10\text{eV}$  are significantly strengthened, to the extent that all current "light-shining-through-wall" experiments find themselves within deeply excluded regions. © 2013 Elsevier B.V.

FUNDING DETAILS: NSERC, Natural Sciences and Engineering Research Council of Canada

REFERENCES: Holdom, B., (1986) Phys. Lett. B, 166, p. 196;

Okun, L.B., (1982) Sov. Phys. JETP. Zh. Eksp. Teor. Fiz., 83, p. 892;

Boehm, C., Fayet, P., (2004) Nucl. Phys. B, 683, p. 219., arxiv:hep-ph/0305261;

Redondo, J., Ringwald, A., (2011) Contemp. Phys., 52, p. 211., arxiv:1011.3741

CORRESPONDENCE ADDRESS: An, H.; Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada; email: han@perimeterinstitute.ca

ISSN: 03702693

CODEN: PYLBA

DOI: 10.1016/j.physletb.2013.07.008

LANGUAGE OF ORIGINAL DOCUMENT: English

ABBREVIATED SOURCE TITLE: Phys Lett Sect B Nucl Elem Part High-Energy Phys

DOCUMENT TYPE: Article

SOURCE: Scopus