



Scientific Electronic Library Online

Certification Criteria for XML Article text Tagging
Services according to SciELO Publishing Schema
(SciELO PS)

January 2021

In the past few years, the SciELO Program has been prioritizing lines of action to strengthen professionalization and internationalization of publishing and publication services for indexed journals. To achieve this, it is updating and extending the common services of publishing, publication and interoperability that are operated by the SciELO Platform in partnership with international and national companies.

Among SciELO platform services are the text structuring in XML language for storage in databases, meta-data extraction, bibliometric calculations, and interoperability with other Web sources. In 2013, the methodology of the platform's XML text tagging and structuring was updated with the [SciELO Publishing Schema \(SciELO PS\)](#) based on standard [Journal Publishing Tag Library NISO JATS Version 1.1](#).

[SciELO PS](#) is fully compatible with the standards JATS: Journal Article Tag Suite, but has additional data elements that are essential for SciELO indexing functions:

- Tagging needs institutional affiliation from all authors, necessary to generate the scientific production indicators per institution and its main units (e.g: University, College, etc.);
- Tagging of data elements that compose the references of the documents cited in articles, indispensable for SciELO's bibliometric indexes, and
- Identification of funding sources for the research.

With the objective of seeking correct application of the article texts structuring schema, SciELO establishes minimum criteria for the companies that render text-tagging services for SciELO journals.

IMPORTANT: Each institution is responsible for monitor the documentations launch versions, as well as introduce the modifications on XML production as indicated.

1. How to apply for certification

For certification services for tagging text, it is necessary to contact the SciELO team by mail scielo@scielo.org and request of sample articles for XML production as [SciELO PS](#).

It will be forwarded to the applicant a set of PDF files for XML production containing different types of documents of SciELO Network journals.

The set of files, when finished, will be sent to the team SciELO by sendspace or wetransfer, communicating its sent by email: scielo@scielo.org.

2. Sample for submission

The sample files must be marked according to the latest version of [SciELO PS](#) observed [SciELO Brasil Specific Rules](#). Each sent must contain the material produced in .zip format (*do not use .rar*).

The folder structure of the files and the items sent from the marked sample must comply with the rules specified **specifically** in item **2. Pasta de Entrega** in the [Package Delivery Guide for Publication in SciELO](#). (*Ignore only “Sumário PDF/Table of Contents”*)

3. Evaluation Process:

After receiving the data package (XML, PDF, digital assets), the SciELO team will evaluate the compliance of files [SciELO PS](#), the result presented by the validators ([see item 6](#)) and verify the marking of articles elements in the XML document against the PDF.

It will be sent detailed report for the data package correction, if it has.

The data package may be re-submitted after submission of the evaluation report, one more time.

The package may be disapproved on its first shipment if a large amount of problems in the files is evaluated, equivalent to 50% or more items with incorrect markings and structure. The evaluation of the XPM report sent with the package will be of great importance for the evaluation of the sample.

In case of doubt about marking the article, generation or validation of XML documents, you may contact the SciELO team through [discussions lists](#) “scielo-xml” and “scielo-discuss”..

4. Criteria for approval and certification services as tagging text according [SciELO PS](#)

It will be approved the data package is:

- In accordance with [SciELO PS](#);
- Valid validation tools offered, without making errors ([see item 6](#)); and
- With its contents identified / marked appropriately (Figure 1).

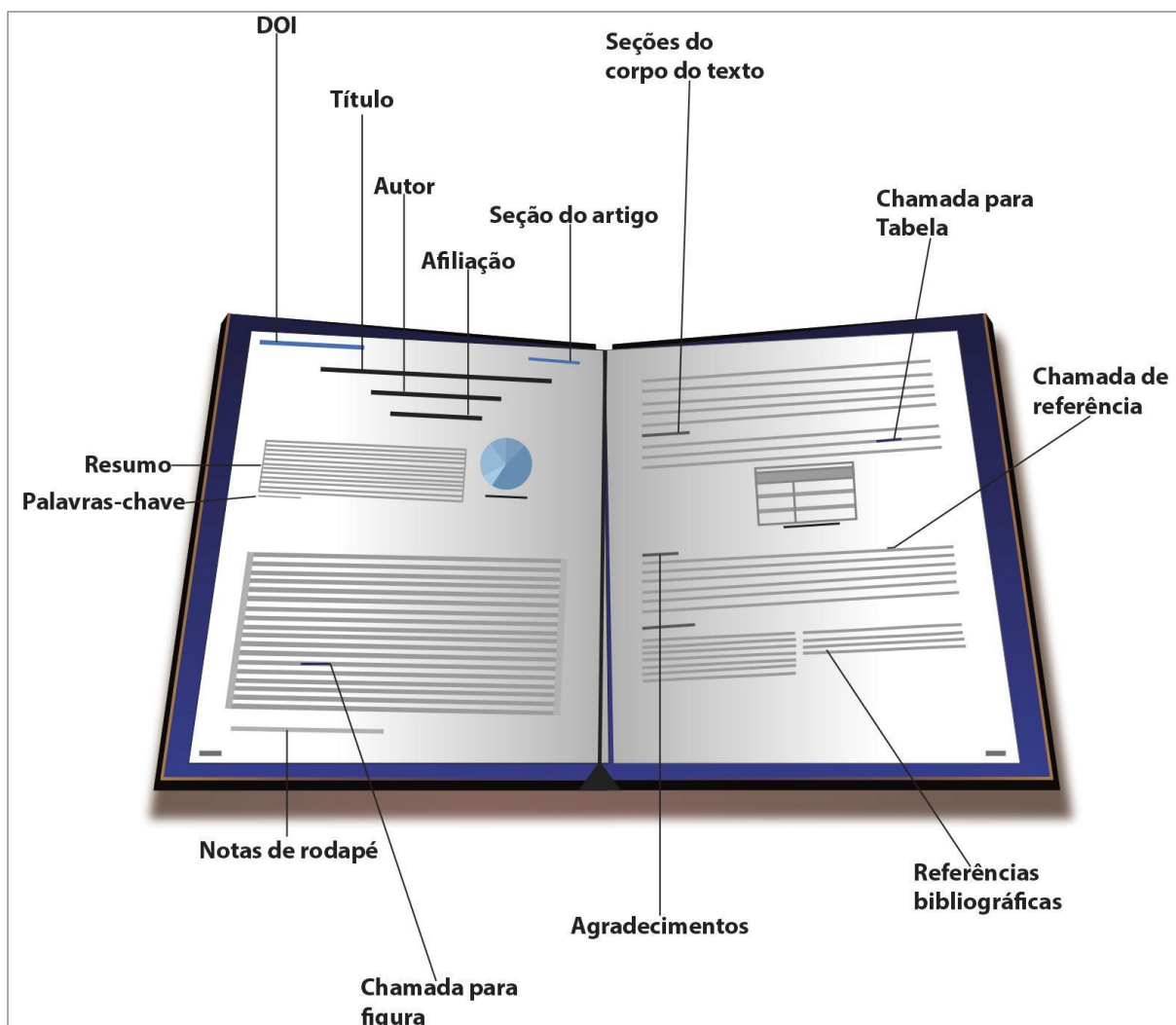


Figure 1: Content to be marked and evaluated for certification

In case of failure, a new application for submission may be made after **six months** after the last communication by SciELO.

5. Continuing Evaluation

Data packages sent to SciELO, after certification, will be continuously evaluated in considering:

- XML support structure according to the [SciELO PS](#) (Most current version)
- Frequent errors;
- Delivery time material, including corrections.

In case of unsatisfactory performance in one or more items, certification will be discredited and will only be possible to request a reevaluation **six months** after the communication of the loss of certification.

6. Tools SciELO

They were developed by the SciELO team the following support tools for producing XML packages:

- [PC-Programs](#) - Set of tools SciELO Methodology to manage journals, numbers and data items, to publish on the site;
- [Package Maker](#)- Validation Tool for SciELO data packages. It presents content dialing errors and conformity of the structure of XML;
- [Packtools](#) - Python library that adds features and utilities for a data package handling SciELO PS and XMLs SciELO PS;
 - [StyleChecker](#) - Web-based tool that presents detailed report on the conformity of an XML data in relation to the SciELO PS specification;

7. Documentation and Support Material

In the list below are indicated document references and tools that assist in understanding and in generating XML files.

- *SciELO Publishing Schema* (SciELO PS):
<https://scielo.readthedocs.io/projects/scielo-publishing-schema/pt_BR/latest/>
- Welcome to SciELO PC Programs' documentation:
<http://docs.scielo.org/projects/scielo-pc-programs/en/latest/pt_how_to_use_the_programs.html>
- JATS 1.1
<<http://jats.nlm.nih.gov/publishing/tag-library/1.1/>>
- SciELO Style Checker
<<http://manager.scielo.org/tools/validators/stylechecker/>>
- Criteria, policies and procedures for the admission and permanence of scientific journals in the SciELO Brazil Collection
<<https://www.scielo.br/avaliacao/20200500%20Criteria%20SciELO%20Brazil%20en.pdf>>
- Release models: To check the specifications documents as erratum, continued publication ahead of print etc. See the link:
<<https://www.scielo.org/pt/sobre-o-scielo/metodologias-e-tecnologias/>>

How to cite this document

SciELO. *Certification Criteria for XML Article text Tagging Services according to SciELO Publishing Schema (SciELO PS)* [online]. SciELO, 2021 [cited DD Month YYYY]. Available from:
