

# XML JATS in OJS: contributions towards integrating markup process and galley generation in diverse contexts

**Soler, Santiago** PREBI-SEDICI UNLP and CESGI, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina, Argentina

**Villarreal, Gonzalo L.** PREBI-SEDICI UNLP and CESGI, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina

**Messineo, Tomás** PREBI-SEDICI UNLP, Argentina

**Ruiz, Adela** Universidad Nacional de La Plata, Argentina

**García, Dolores** Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina

**Correa, Lucía** Universidad Nacional de La Plata, Argentina



EDUCACIÓN  
PÚBLICA  
Y GRATUITA



UNIVERSIDAD  
NACIONAL  
DE LA PLATA

## Context

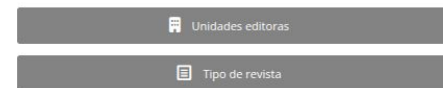
UNLP is a public university located in La Plata city, Argentina

More than **100 journals** edited by our institution

Distributed model: one central journals portal + journals portals in some faculties

A central figure of “Scientific Journals Coordination”, for all portals and journals => DOI, Markup, support, software, Q&A, institutional directives, indexation...

Collaboration with the institutional repository **SEDICI** and vice versa



➤ CAPACITACIÓN PARA LA IMPLEMENTACIÓN DEL SISTEMA DE LA UNLP PARA EL MERCADO DE REVISTAS CIENTÍFICAS

<https://portalderevistas.unlp.edu.ar>



## The Importance of JATS XML

JATS XML enables the implementation of standardized metadata schemas for content publication:

- ✓ Digital preservation
- ✓ Machine readable
- ✓ Generation of different formats (PDF, ePUB, HTML)

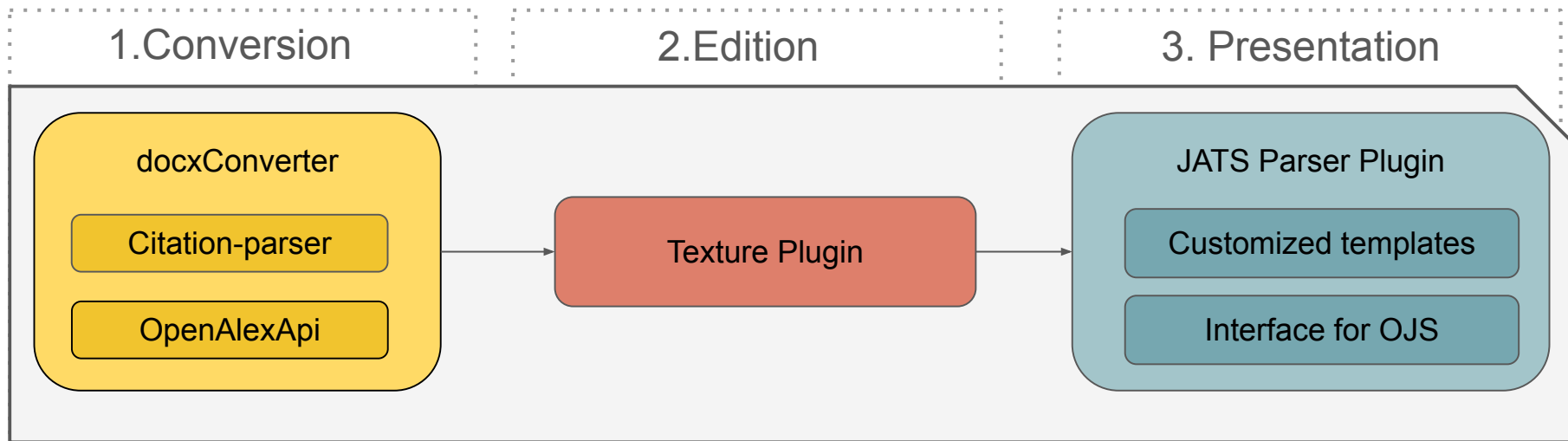
Today we (mostly) use Marcalyc (Redalyc, AmeliCA) and Scielo Makup

We would like **to gain independence** from any organization, and **to achieve sovereignty** over our editorial processes.



# OJS

## Steps for an JATS XML markup workflow





## Integration of strategies for markup

To have a **complete** markup tool, and **integrated** into OJS

Adapted to the needs of the UNLP but extensible to different requirements from other institutions

EDUCACIÓN  
PÚBLICA  
Y GRATUITA

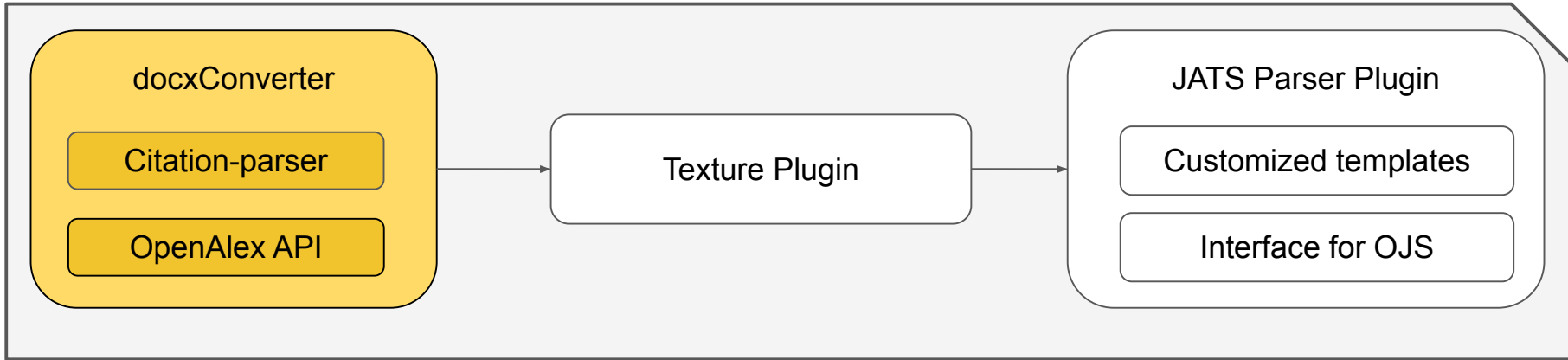


UNIVERSIDAD  
NACIONAL  
DE LA PLATA

# OJS

## docxConverter

DOCX to JATS XML converter



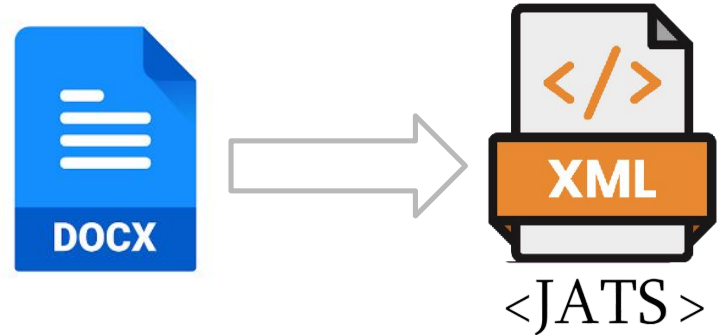


# docxConverter

## DOCX to JATS XML converter

Convert files in DOCX format to JATS XML format

The biggest limitation of the tool is that it doesn't convert references, which are one of the most time-consuming part of the markup process



1. Forked repository <https://github.com/sedici/docxConverter>





## OpenAlex integration for references

References may be incomplete, due to requirements of citation standards

Becerril-García, A., Aguado López, E. y Macedo García, A. (2023). Marcalyc: software para la marcación XML JATS para las revistas científicas de acceso abierto diamante. *Palabra Clave*, 12(2), e179. <https://doi.org/10.24215/18539912e179>

Detecting the DOI in a citation enable us to:

- ✓ Retrieve all metadata and improve the citation
- ✓ Detect errors (non existing DOI)
- ✓ Warn about discrepancies (incorrect metadata, missing authors, misspellings in journal/article title or author names)

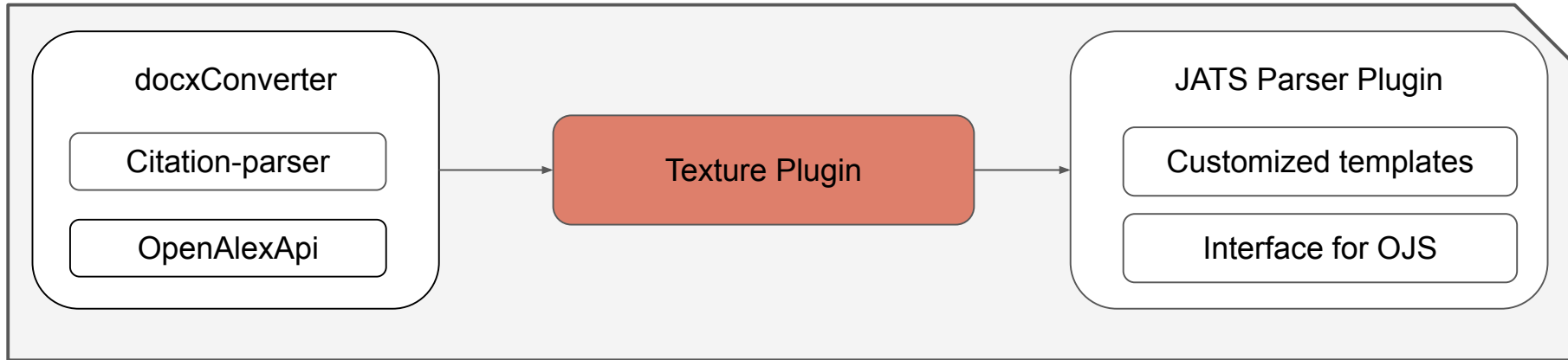


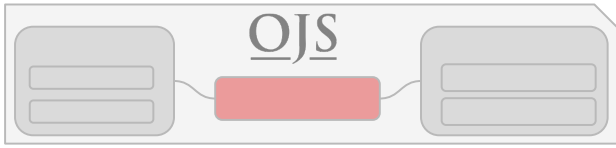
OpenAlex

# OJS

## Texture Plugin

JATS XML Edition



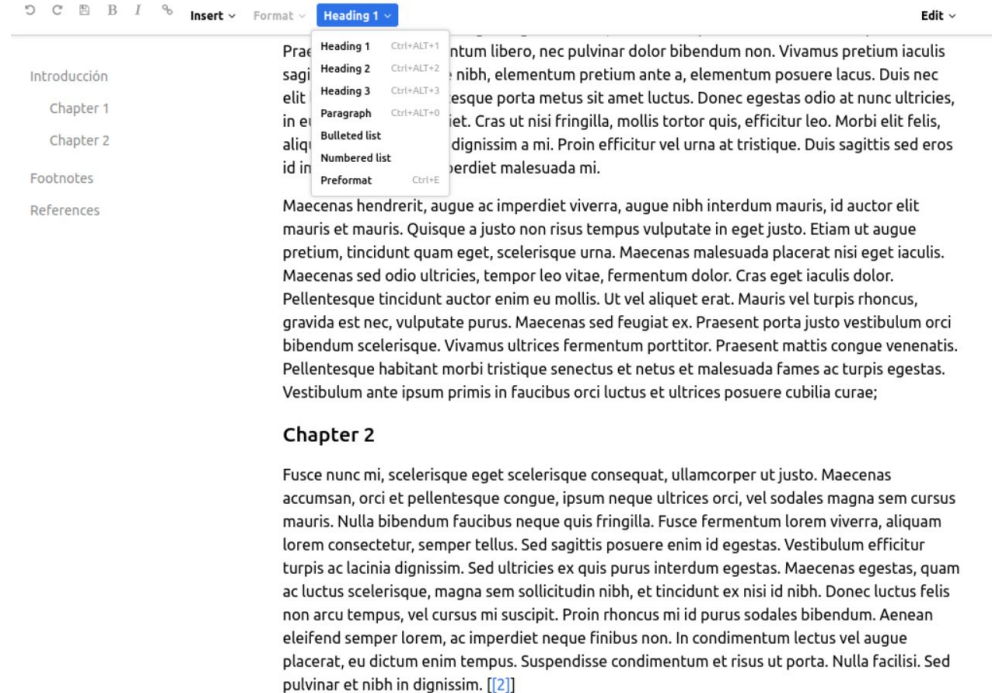


# Texture Plugin

## JATS XML editor

docxConverter output adapted to ensure compatibility with Texture

Migration of Texture plugin for OJS 3.4 (no longer supported by PKP)

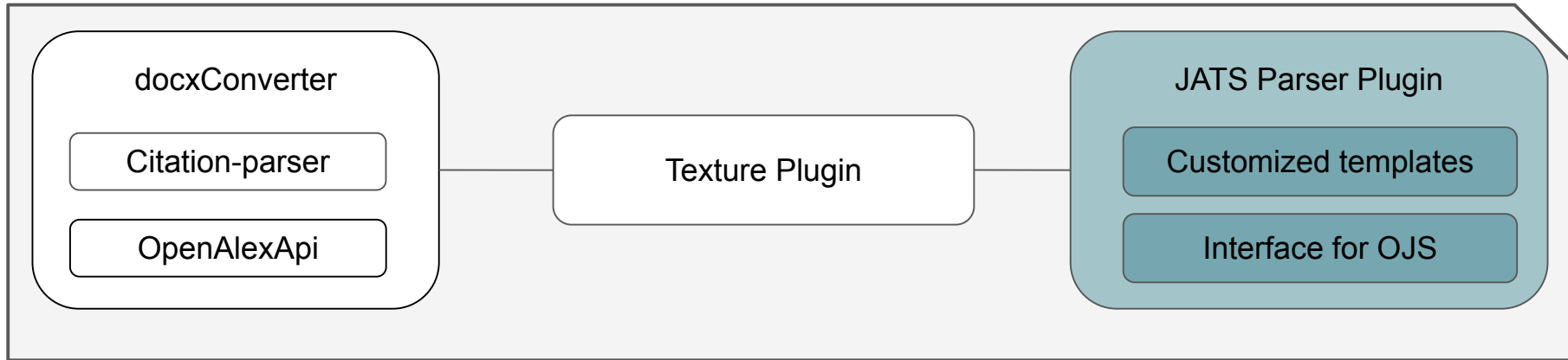


1. UNLP' improvements, merged into <https://github.com/pkp/texture>, are already available

# OJS

## JATS Parser Plugin

Presentation and generation of PDF





# JATS Parser Plugin

## PDF Generator

- ✓ OJS plugin for PDF generation
- ✗ No customizable templates
- ✗ No support for different templates
- ✗ No support for different citation styles

## Introduction

Vascular and cerebrovascular diseases remain an actual problem of modern society. This is explained by their significant prevalence, as well as high mortality rates. According to the Ministry of Health, 100-120 thousand strokes are registered annually in Ukraine, of which one third is among the working age population. In 2015, the incidence of acute impairment of cerebral circulation was 274.1 per 100 thousand population (the average rate of morbidity in Europe - 200 cases per 100 thousand population) (Mishchenko, 2017), (Zinchenko & Mishchenko, 2016). The American Association of Cardiologists reported that 795,000 acute impairments of cerebral circulation are recorded in the United States in 2017 year, that is, a new stroke occurs almost every 40 seconds ("Heart Disease and Stroke Statistics—2017 Update: A Report From the American Heart Association", 2017).

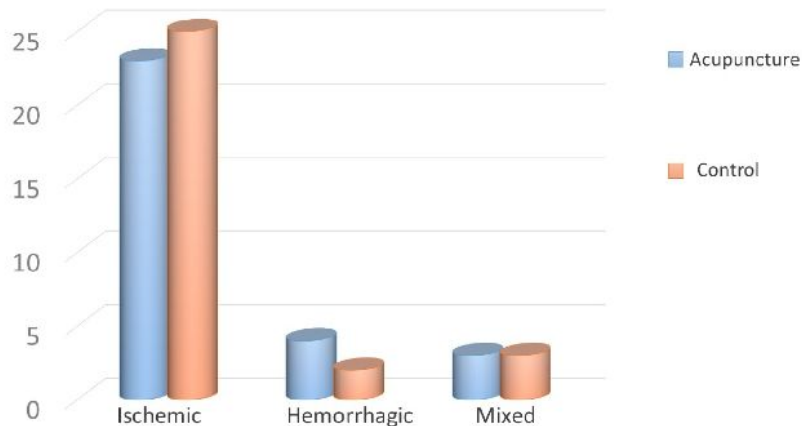


Figure 0.5. This is figure inside paragraph

Тут вказаний текст кирилицею для перевірки коректної конвертації шрифту.

According to Harrison R.A., more than 70% of patients with a stroke feel pain every day.



# Integration of customized templates

## PDF generation

✓ Added support for multiple templates for PDF output.

Each journal can now customized articles presentation

✓ Combines metadata (from OJS) with article content (from JATS XML)

**MEMBRETE**  
 DATOS DE LA REVISTA,  
 (VOL., NÚM. AÑO)  
 DOI Y URL DEL SITIO  
 (AMBOS ENLAZADOS)

**TITULACIÓN**  
 EN DIFERENTES IDIOMAS

**RESUMEN**  
**Y PALABRAS CLAVE**  
 EN DIFERENTES IDIOMAS



HILOS DOCUMENTALES  
 Vol. 4, núm. 8, e063, artículos, 2024  
<https://doi.org/10.24215/20184486e063>  
 ISSN: 2018-4486 | <https://revistas.unlp.edu.ar/HilosDocumentales>  
 Universidad Nacional de La Plata (Argentina)



**LOGOS**  
 DE LA REVISTA  
 Y DE LA INSTITUCIÓN

**Almanques y calendarios en el siglo XIX**  
 Impresos menores, lectores populares y tiempos seculares  
 en América del Sur (1825-1860)

Almanacs and Calendars in the 19th Century:  
 Minor Prints, Popular Readers and Secular Times in South America (1825-1860)

**Autor/a**  
 correo@untref.edu.ar  
 Universidad Nacional de Tres de Febrero, Argentina  
 Comisión de Investigaciones Científicas (CIC)

**Autor/a**  
 correo@unlp.edu.ar  
 Universidad Nacional de La Plata, Argentina  
 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

**DATOS DE AUTORÍA**  
 (NOMBRE, ORCID,  
 CORREO, AFILIACIÓN)

**Resumen:** Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem sequi sequi sequi sequi mmm.

**Palabras clave:** quisquam, magnam, beatae, omnis

**Abstract:** Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem sequi sequi sequi mmm.

**Keywords:** quisquam, magnam, beatae, omnis

Recibido 16/06/2024 - Aceptado 24/08/2024 - Publicado 06/09/2024

**LICENCIAS DE USO**



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional

**FECHAS**

RECIBIDO / ACEPTADO / PUBLICADO



**CARÁTULA**  
**DEL PDF**  
**GENERADO**  
**CON SUMARC**



# OJS extension to support citation output

## PDF Generation

✓ JATSParser enables the generation of a list of references in different citation styles (APA, IEEE, etc.)

✗ However, it doesn't support configuration of citation output **inside the body of the article**.

Parentetical citation: "The Sun is a star (Sandocam, 2025)."

Narrative citation: "According to Sandocam (2025), the Sun is a star"

Contexto	Referencias	Estilo de Cita
se sustenta en las discusiones fenomenológicas que resaltan la importancia de investigar desde la perspectiva propia de los sujetos y de su experiencia sobre los fenómenos sociales que los atraviesan (Taylor y Bogdan, 1987)	Taylor, S. J., y Bogdan, R. (1987). <i>Introducción a los métodos cualitativos de investigación. La búsqueda de sentido</i> . Paidós.	(Taylor y Bogdan, 1987)
Para el análisis de las relatorías, se utilizó la teoría fundamentada que se basa en dos estrategias: la comparación constante y el muestreo teórico (Strauss y Corbin, 1998)	Strauss, A., y Corbin, J. (1998). <i>Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada</i> . Universidad de Antioquia.	(Strauss y Corbin, 1998)
la virtualidad y la presencialidad; y otras. Una vez realizado el proceso de identificación, se pasó a la codificación selectiva de aquellas categorías periféricas que se relacionan con las centrales (Strauss y Corbin, 1998)	Strauss, A., y Corbin, J. (1998). <i>Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada</i> . Universidad de Antioquia.	(Strauss y Corbin, 1998)
El muestreo teórico, en tanto, consistió en construir un marco interpretativo y analítico para la información obtenida y sistematizada durante la comparación constante (Strauss y Corbin, 1998)	Strauss, A., y Corbin, J. (1998). <i>Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada</i> . Universidad de Antioquia.	(Strauss y Corbin, 1998)
obtenida y sistematizada durante la comparación constante [2]. Para las experiencias referentes a los afectos que produjo el regreso a la presencialidad, se recuperan las propuestas de Varela y Álvarez-Uría (1991)	Varela, J., y Álvarez-Uría, F. (1991). <i>Arqueología de la escuela</i> . La piqueta.	(1991)
sistematizada durante la comparación constante [2]. Para las experiencias referentes a los afectos que produjeron el regreso a	Rull S. (1993) <i>Fourault u la educación. Disciplinas u saber</i>	



# OJS GUI for citation to **References** edition

Referencias

Tablas y Figuras

Contexto	Referencias	Estilo de Cita
se sustenta en las discusiones fenomenológicas que resaltan la importancia de investigar desde la perspectiva propia de los sujetos y de su experiencia sobre los fenómenos sociales que los atraviesan ( <a href="#">Taylor y Bogdan, 1987</a> )	Taylor, S. J., y Bogdan, R. (1987). <i>Introducción a los métodos cualitativos de investigación. La búsqueda de sentido</i> . Paidós.	(Taylor y Bogdan, 1987) ▾
Para el análisis de las relatorías, se utilizó la teoría fundamentada que se basa en dos estrategias: la comparación constante y el muestreo teórico ( <a href="#">Strauss y Corbin, 1998</a> )	Strauss, A., y Corbin, J. (1998). <i>Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada</i> . Universidad de Antioquia.	(Strauss y Corbin, 1998) ▾



# OJS interface for **Tables** and **Figures** citation edition

Referencias **Tablas y Figuras**

Contexto	Títulos	Estilo de Cita
Para la recuperación de las experiencias se invitó a estos grupos (60 estudiantes en total, 17 de la UAM-X y 43 de la BUAP) <b>Figura 1</b>	Tabla 1. Caracterización de los/as estudiantes entrevistados/as	Figure 1 Figure 1 Tabla 1. Caracterización de los/as estudiantes entrevistados/as Otro
virtualidad y la presencialidad; y otras. Una vez realizado el proceso de identificación, se pasó a la codificación selectiva de aquellas categorías periféricas que se relacionan con las centrales [2] <b>Tabla 1</b>	Cuadro 1. Resultado del proceso de codificación	Tabla 1

**Guardar**



## Compatibility with OJS 3.3 LTS and OJS 3.4

Even though we are using OJS 3.4 in our main Journals Portal, we ported our code to OJS 3.3 too

-> Long Term Support

-> Availability for much more OJS installations

Improvements shared with OJS worldwide community in:

- GitHub SEDICI repository
- GitHub PKP base repository



PKP PUBLIC  
KNOWLEDGE  
PROJECT



## SUMARC

Sistema UNLP para el Mercado de Revistas Científicas  
(*UNLP System for Scientific Journals Markup*)

This tool set was launched at UNLP as SUMARC project

We started with one single journal (our Guinea pig)

After several improvements and updates, SUMARC is currently being used by 10 UNLP' scientific journals

We plan to gradually incorporate more journals

# SUMARC

SISTEMA UNLP  
PARA EL MERCADO  
DE REVISTAS CIENTÍFICAS



# Current/Future projects

## 1. Support for more citation standards

## 2. A better PDF output system

So far, PDF generation is achieved via PHP code.

Even though we added the ability to have many PDF templates, the current system is hard to customize

-> requires PHP programming skills

-> hard to design templates from scratch

**Proposal: use TPL (Smarty) templates to generate HTML and CSS output => Print HTML/CSS as PDF**

The output file complies with the PDF/A standard (easier integration with institutional repository SEDICI; no need to convert files)

**MEMBRETE**  
DATOS DE LA REVISTA,  
(VOL., NÚM., AÑO)  
DOI Y URL DEL SITIO  
(AMBOS ENLAZADOS)

**TITULACIÓN  
EN DIFERENTES IDIOMAS**

**RESUMEN  
Y PALABRAS CLAVE  
EN DIFERENTES IDIOMAS**

**LICENCIAS DE USO**

**DATOS DE LA REVISTA**  
(VOL., NÚM., AÑO)  
DOI Y URL DEL SITIO  
(AMBOS ENLAZADOS)

**ALMANAQUES Y CALENDARIOS EN EL SIGLO XIX**  
Impresos menores, lectores populares y tiempos seculares  
en América del Sur (1825-1860)

**DATOS DE AUTORÍA**  
(NOMBRE, ORCID,  
CORREO, AFILIACIÓN)

**FECHAS**  
RECIBIDO / ACEPTADO / PUBLICADO

**LOGOS  
DE LA REVISTA  
Y DE LA INSTITUCIÓN**

**CARÁTULA  
DEL PDF  
GENERADO  
CON SUMARC**



## Help needed

We are looking for collaboration from institutions and individuals to:

- improve documentation and support for the global community (installation, usage)
- add support for more citation standards besides APA7 and IEEE
- complete the new templating system
- generate a large gallery of PDF templates
- provide training for editorial teams and IT staff



designed by  freepik

# Thank you!



**SUMARC** download and installation instructions

<https://portalderevistas.unlp.edu.ar>



**sumarc@revistas.unlp.edu.ar**