

Análisis del PKP Beacon: Publicaciones Académicas con Software PKP

Informe Consolidado con Validación Beacon v6 (2024-2025)

Patricio Pantaleo

2025-11-22

Table of contents

1	Análisis del PKP Beacon: Publicaciones Académicas con Software PKP	2
1.1	Informe Consolidado con Validación Beacon v6 (2024-2025)	2
1.2	Resumen	2
1.3	1. Introducción	3
1.3.1	1.1 Contexto	3
1.3.2	1.2 Antecedentes	3
1.3.3	1.3 Objetivos	3
1.3.4	1.4 Marco Conceptual	3
1.4	2. Materiales y Métodos	4
1.4.1	2.1 Fuente de Datos	4
1.4.2	2.2 Metodología de Análisis Actualizada	5
1.5	3. Resultados	7
1.5.1	3.1 Análisis Global v6	7
1.5.2	3.2 Análisis Específico de Chile - Validación v6	8
1.5.3	3.3 Validación Cruzada: Comparación v5 vs v6	10
1.5.4	3.4 Archivos Generados v6	11
1.5.5	3.5 Análisis de Calidad Editorial: Informes Dialnet	11
1.6	4. Discusión	14
1.6.1	4.1 Validación del Ecosistema Global	14
1.6.2	4.2 Caracterización del Ecosistema Chileno Actualizada	14
1.6.3	4.3 Implicaciones para Indexación en Dialnet	14
1.6.4	4.4 Contribución al Acceso Abierto Regional	15
1.6.5	4.5 Limitaciones y Consideraciones	15
1.7	5. Limitaciones del Estudio	15
1.7.1	5.1 Limitaciones Metodológicas	15
1.7.2	5.2 Limitaciones Técnicas	15
1.7.3	5.3 Limitaciones de Alcance	15
1.8	6. Conclusiones	16
1.8.1	5.1 Validación del Ecosistema Robusto	16
1.8.2	5.2 Liderazgo Institucional Confirmado	16

1.8.3	5.3 Crecimiento Sostenido Validado	16
1.8.4	5.4 Metodología Consolidada	16
1.8.5	5.5 Posicionamiento Global	16
1.9	7. Próximos Pasos	16
1.9.1	7.1 Evaluación Dialnet	16
1.9.2	6.2 Enriquecimiento Continuo	17
1.9.3	6.3 Análisis Comparativo Regional	17
1.10	Agradecimientos	17
1.11	Referencias	17
1.11.1	Estudios sobre JUOJS y PKP Beacon	17
1.11.2	Estudios sobre indexación y acceso abierto	17
1.11.3	Estudios sobre publicación científica chilena y latinoamericana	18
1.11.4	Marco teórico y metodológico	18
1.11.5	Recursos técnicos	18
1.11.6	Estándares y criterios	18
1.11.7	Datasets utilizados	19
1.12	Colaboraciones	19
1.13	Anexos	19
1.13.1	Anexo A: Diccionario de Datos del PKP Beacon	19
1.13.2	Anexo B: Scripts Desarrollados	20
1.13.3	Anexo C: Proceso de Evaluación en Dialnet	20
1.13.4	Anexo D: Reproducibilidad	21

1 Análisis del PKP Beacon: Publicaciones Académicas con Software PKP

1.1 Informe Consolidado con Validación Beacon v6 (2024-2025)

1.2 Resumen

Este estudio analiza el ecosistema global de revistas académicas que utilizan Open Journal Systems (OJS) mediante el dataset del PKP Beacon, con un enfoque específico en Chile. Se procesaron **86,282 instalaciones OJS globales** (Beacon v6), de las cuales **432 corresponden a Chile**. El análisis incluye validación cruzada entre versiones del beacon y enriquecimiento con datos de OpenAlex para evaluar visibilidad e impacto académico. La **validación v6** confirma **319 instalaciones chilenas activas** con criterio actualizado (>5 pub/2024), representando **14,162 publicaciones en 2024** con un **crecimiento del 16.5%** respecto a 2023.

1.3 1. Introducción

1.3.1 1.1 Contexto

El Public Knowledge Project (PKP) desarrolla software de código abierto para la publicación académica, siendo Open Journal Systems (OJS) su aplicación más utilizada globalmente. El PKP Beacon es un sistema de recopilación de datos que permite identificar, inspeccionar y catalogar instalaciones públicas de este software.

1.3.2 1.2 Antecedentes

Estudios previos:

Khanna et al. (2022): - Dataset 2020: 25,671 revistas - 36.5% de 70,214 beacons OJS operativos - 996,000 artículos publicados en 2020 - 5.4 millones de artículos históricos acumulados

Khanna et al. (2024): - Dataset actualizado: 47,625 revistas con ISSN validados - Aumento del 98% respecto a 2020 - 2,962,418 artículos publicados entre 2020-2023 - 10.6 millones de artículos históricos totales

Beacon v6 (2024-2025): - **86,282 instalaciones OJS globales** (nuevo récord) - **54,453 instalaciones activas** en 2024 (63.1%) - **2.36 millones de publicaciones** en 2024 - **Crecimiento global:** +516,768 publicaciones vs 2023

1.3.3 1.3 Objetivos

1. **Analizar el ecosistema global** de revistas que utilizan OJS con datos actualizados v6
2. **Caracterizar las instalaciones chilenas** y validar con criterios 2024
3. **Evaluar la visibilidad académica** mediante integración con OpenAlex
4. **Preparar datos para evaluación** en sistemas de indexación como Dialnet

1.3.4 1.4 Marco Conceptual

1.3.4.1 JUOJS (Journals Using OJS)

Las JUOJS emplean un PKP Beacon opcional que permite a PKP notificar sobre actualizaciones y transmite información de indexación incluyendo título, ISSN, número de artículos, títulos y resúmenes.

1.3.4.2 Criterio de revista activa

Se emplea el estándar del Directory of Open Access Journals (DOAJ) de **cinco artículos al año** para distinguir revistas con actividad editorial sostenida de instalaciones inactivas o en prueba. Este criterio, está establecido en Chavarro, D., Alperin, J. P., & Willinsky, J. (2025) donde se dice:

“This study relies on data gathered from the PKP Beacon, a feature introduced into OJS in 2015 that allows PKP to notify OJS users about security updates and software upgrades. The beacon also notifies PKP of the web location of each journal’s metadata harvesting API, which we subsequently used to collect journal metadata, such as the journal name and ISSN, and article metadata, such as titles, abstracts, publication dates, and DOIs. As in our previous study, we considered a journal to be “active” in a given year if it published at least five documents that year, a threshold established by the Directory of Open Access Journal (DOAJ) in 2020. We began with the OJS journal list updated in 2024 (Khanna et al., 2024). This updated list included 50,920 journals with an ISSN, and that were deemed to be active in 2023 (a 98% increase in the number reported for 2020; Khanna et al., 2022).”

1.4 2. Materiales y Métodos

1.4.1 2.1 Fuente de Datos

1.4.1.1 2.1.1 Dataset Principal PKP Beacon v6

Dataset: Khanna, S., Raoni, J., Smecher, A., Alperin, J. P., Ball, J., & Willinsky, J. (2024). “Details of Publications Using Software by the Public Knowledge Project.” Harvard Dataverse.

Características v6: - **Archivo:** beacon_v6.csv (87,170 observaciones, 31 campos) - **Instalaciones OJS:** 86,282 (99.0%) - **Instalaciones OMP:** 810 (0.9%) - **Instalaciones OPS:** 78 (0.1%) - **Nuevas columnas:** - record_count_2024 - Publicaciones 2024 - record_count_2025 - Publicaciones 2025 - region - Región geográfica PKP - admin_email - Email de administrador - country_doaj - País según DOAJ

1.4.1.2 2.1.2 Dataset PKP Beacon v5 (Versión Inicial)

Dataset: Khanna, S., Raoni, J., Smecher, A., Alperin, J. P., Ball, J., & Willinsky, J. (2024). “Details of Publications Using Software by the Public Knowledge Project.” Harvard Dataverse.

Características v5: - **DOI:** <https://doi.org/10.7910/DVN/OCZNVY> - **Archivo:** beacon.csv (67,138 observaciones, 25 campos) - **Instalaciones OJS:** 66,527 (99.1%) - **Instalaciones OMP:** 562 (0.8%) - **Instalaciones OPS:** 49 (0.1%) - **Datos hasta:** 2023 (último año completo) - **Criterio actividad:** >5 publicaciones en 2023 - **Chile identificado:** 396 instalaciones totales - **Chile activas v5:** 309 instalaciones (78.0%)

1.4.1.3 2.1.3 Validación Cruzada v5-v6

- **Crecimiento total:** 67,138 → 87,170 (+29.8%)
- **Crecimiento OJS:** 66,527 → 86,282 (+29.7%)
- **Chile:** 396 → 432 instalaciones (+9.1%)
- **Metodología:** Análisis comparativo para validación de tendencias

1.4.2 2.2 Metodología de Análisis Actualizada

1.4.2.1 2.2.1 Flujo Metodológico v5 (Versión Inicial)

1. Procesamiento Base v5

```
# Separar beacon por tipo de aplicación
python3 scripts/1_split_beacon.py
# Genera: beacon_ojs.csv (66,527 instalaciones)

# Análisis global OJS
Rscript scripts/2_analisis_ojs_mundial.R
# Genera: visualizations/top15_paises_barras.png
```

2. Análisis Chile v5

```
# Generar dataset completo de Chile (todos los criterios de país)
Rscript scripts/3_analisis_chile.R
# Genera: visualizations/chile_todas_instalaciones.csv (396)

# Filtrar solo instalaciones activas (>5 pub/2023)
Rscript scripts/4_chile_juojs_filtrado.R
# Genera: visualizations/chile_juojs_activas.csv (309 - DATASET PRINCIPAL v5)
```

3. Evaluación Dialnet v5

```
# Generar URLs OAI desde dataset JUOJS v5
python3 scripts/5_generar_urls_dialnet.py
# Genera: visualizations/chile_oai_urls_limpio.csv (202 URLs únicas)

# Verificar completitud de informes descargados
python3 scripts/6_find_missing_reports.py

# Procesar informes HTML de Dialnet descargados
python3 scripts/8_procesar_informes_dialnet.py
# Genera: visualizations/dialnet_informes_procesados.csv

# Analizar tipos de errores por gravedad
python3 scripts/9_analizar_tipos_errores.py
# Genera: análisis de 40 tipos de errores en 95 informes
```

1.4.2.2 2.2.2 Flujo Metodológico v6 (Validación Actualizada)

0. Limpieza de Datos v6 (Preprocesamiento con Comandos)

Problemática identificada: El archivo beacon_v6.tab original presentaba saltos de línea incrustados dentro de campos de texto que rompían la estructura del archivo.

Proceso de limpieza realizado:

```
# Comandos utilizados para limpieza:
# 1. Remover saltos de línea problemáticos
sed ':a;N;$!ba;s/\n\([^0-9]\)/\1/g' beacon_v6.tab > beacon_v6_temp.tab

# 2. Conversión de TAB a CSV
sed 's/\t/,/g' beacon_v6_temp.tab > beacon_v6_clean.csv

# 3. Limpieza de archivos temporales
rm beacon_v6_temp.tab

# Resultado: beacon_v6_clean.csv (87,170 registros limpios)
```

Validación post-limpieza: - **Registros preservados:** 100% (87,170/87,170) - **Estructura validada:** 31 columnas consistentes - **Formato:** CSV estándar con separadores de coma - **Problema resuelto:** Saltos de línea incrustados eliminados

1. Procesamiento Base

```
# Separar beacon v6 por tipo de aplicación
python3 scripts_v6/1_split_beacon_v6.py
# Input: beacon_v6_clean.csv (datos limpiados manualmente)
# Genera: beacon_v6_ojs.csv (86,282 instalaciones)

# Análisis global v6
Rscript scripts_v6/2_analisis_mundial_v6.R
# Genera: visualizations_v6/top15_paises_barras_v6.png
```

2. Análisis Chile v6

```
# Filtrar instalaciones chilenas con criterios v6
Rscript scripts_v6/3_analisis_chile_v6.R
# Genera: visualizations_v6/chile_todas_instalaciones_v6.csv (432)

# Aplicar criterio JUOJS actualizado (>5 pub/2024)
Rscript scripts_v6/4_chile_juojs_filtrado_v6.R
# Genera: visualizations_v6/chile_juojs_activas_v6.csv (319 - DATASET PRINCIPAL)
```

3. Evaluación Dialnet v6

```
# Generar URLs OAI desde dataset JUOJS v6
python3 scripts_v6/5_generar_urls_dialnet_v6.py
# Genera: visualizations_v6/chile_oai_urls_v6.csv (225 URLs únicas)
```

4. Análisis y Visualizaciones

```
# Generar tablas y análisis institucional
Rscript scripts_v6/6_tablas_chile_v6.R
# Genera: tablas CSV y análisis institucional
```

5. Validación Cruzada v5-v6

```
# Comparación precisa entre versiones por dominio
python3 scripts/7_comparar_urls_v5_v6.py
# Compara: chile_oai_urls_limpio.csv vs chile_oai_urls_limpio_v6.csv
# Clasifica: cambios sustanciales vs menores (www, paths)
```

1.4.2.3 2.2.3 Criterio JUOJS: Evolución v5→v6

Versión v5: - **Criterio temporal:** >5 publicaciones en **2023** - **Filtrado geográfico:** country_consolidated, country_issn, country_tld, country_ip - **Chile resultante:** 309 instalaciones activas (78.0%) - **URLs generadas:** 202 únicas (107 duplicados eliminados)

Versión v6 (Actualizada): - **Criterio temporal:** >5 publicaciones en **2024** (vs 2023 en versión anterior) - **Filtrado mejorado:** Incluye country_doaj para mayor precisión geográfica - **Eliminación de duplicados:** Automática por dominio (94 duplicados eliminados)

1.5 3. Resultados

1.5.1 3.1 Análisis Global v6

1.5.1.1 3.1.1 Distribución General

- **Total instalaciones procesadas:** 87,170
- **Instalaciones OJS:** 86,282 (99.0%)
- **Instalaciones OMP:** 810 (0.9%)
- **Instalaciones OPS:** 78 (0.1%)
- **Países representados:** 154 países activos

1.5.1.2 3.1.2 Instalaciones OJS Activas Globalmente v6

- **Total instalaciones OJS activas globalmente:** 54,453 (63.1% del total)
- **Publicaciones 2024:** 2,365,938 (+28.4% vs beacon anterior)
- **Crecimiento 2023→2024:** +516,768 publicaciones (+28.4%)
- **Cambio en actividad:** +3,074 instalaciones activas vs 2023

1.5.1.3 3.1.3 Top 10 Países por Instalaciones Activas 2024

País	Instalaciones Activas	Pub. 2024	Crecimiento 2023→2024
Indonesia (ID)	23,045	713,408	+122,211
Brasil (BR)	4,050	228,789	+48,856
Desconocido	3,260	161,446	+87,204
Estados Unidos (US)	1,714	87,977	+26,968
India (IN)	1,630	95,726	+22,334
España (ES)	1,295	57,190	+13,336
Tailandia (TH)	1,172	43,923	+5,298
Ucrania (UA)	1,052	73,371	+13,299
Rusia (RU)	954	74,346	+6,280
Pakistán (PK)	893	43,070	+11,921

1.5.1.4 3.1.4 Distribución por Regiones PKP v6

Región PKP	Instalaciones	Pub. 2024	Países
East Asia & Pacific	25,767	873,834	22
Europe & Central Asia	9,112	543,986	47
Latin America & Caribbean	8,485	377,023	22
Other	3,325	164,692	22
South Asia	3,099	151,324	8
North America	2,110	103,790	2
Sub-Saharan Africa	1,525	57,713	33
Middle East & North Africa	1,030	64,297	19

1.5.2 3.2 Análisis Específico de Chile - Validación v6

1.5.2.1 3.2.1 Estadísticas Generales Actualizadas

- **Total instalaciones identificadas:** 432 (vs 396 en versión anterior)
- **Instalaciones JUOJS activas (>5 pub/2024):** 319 (73.8%)
- **Instalaciones filtradas:** 113 (5 pub/2024)
- **URLs únicas generadas:** 225 (eliminados 94 duplicados)

1.5.2.2 3.2.2 Productividad 2024 (Validación v6)

- **Total publicaciones 2024:** 14,162
- **Total publicaciones 2023:** 12,156
- **Crecimiento neto:** +2,006 publicaciones (+16.5%)
- **Promedio por instalación activa:** 44.4 pub/instalación
- **Total histórico acumulado:** 144,529 publicaciones

1.5.2.3 3.2.3 Top 10 Instalaciones Más Productivas 2024 (v6)

Revista/Contexto	Dominio	Pub. 2024	Crecimiento
Revista Chilena de Derecho	revistachilenaderecho.uchile.cl	1,708	+1,671
Boletín de la Universidad de Chile	revistasdex.uchile.cl	1,689	+1,529
Boletín Jurídico del Observatorio	revistadelaconstruccion.uchile.cl	958	-1,723
REVISTA CUHSO	cuhso.uct.cl	510	+470
MGC / Revista de Gestión Cultural	revistasdex.uchile.cl	430	+218
Ingeniare. Revista Chilena de Ingeniería	revistalimite.uta.cl	321	+321
Revista de Trabajo Social	revistadelaconstruccion.uchile.cl	298	+277
Revista de Historia Social y de las Mentalidades	www.revistas.usach.cl	221	+190
Apuntes de Teatro	revistaapuntes.uc.cl	214	-804
Progress Annals: Journal of Progressive Research	academiaone.org	136	+83

1.5.2.4 3.2.4 Análisis de Crecimiento 2023-2024 (v6)

- **Instalaciones que crecieron:** 163 (51.1%)
- **Instalaciones que decrecieron:** 135 (42.3%)
- **Instalaciones estables:** 21 (6.6%)
- **Crecimiento promedio:** +6.3 pub/instalación

1.5.2.5 3.2.5 Distribución Institucional Actualizada (v6)

Institución	Instalaciones	%	Pub. 2024	Crecimiento
Universidad de Chile	67	21.0%	3,338	+1,646
Pontificia Universidad Católica	27	8.5%	4,129	-438
Universidad de Valparaíso	15	4.7%	227	-23
Universidad de Concepción	12	3.8%	353	-36
Universidad Austral de Chile	7	2.2%	213	-59
Universidad de Santiago	6	1.9%	299	+177
Universidad de La Frontera	4	1.3%	-	-
Otras instituciones	181	56.7%	5,603	+739

1.5.2.6 3.2.6 Distribución Geográfica PKP (v6)

- **Latin America & Caribbean:** 317 instalaciones (99.4%)
- **Europe & Central Asia:** 1 instalación (0.3%)
- **North America:** 1 instalación (0.3%)

1.5.2.7 3.2.7 URLs OAI Generadas para Dialnet (v6)

- **Total URLs únicas:** 225 (vs 202 en versión anterior)
- **URLs Universidad de Chile:** 58 (25.8%)
- **Publicaciones 2024 representadas:** 11,540 (81.5% del total)
- **Crecimiento neto representado:** +1,843 publicaciones

1.5.3 3.3 Validación Cruzada: Comparación v5 vs v6

1.5.3.1 3.3.1 Análisis Comparativo por Dominios (Script 7)

Metodología de comparación: Se ejecutó análisis preciso mediante `scripts/7_comparar_urls_v5_v6.py` que compara URLs por dominio normalizado, clasificando cambios como sustanciales o menores (www, paths, protocolos).

Resultados de estabilidad: - **URLs v5:** 208 → **URLs v6:** 225 (+17 URLs netas) - **Dominios efectivamente iguales:** 185/208 (88.9% estabilidad) - **Dominios con cambios menores:** 7 (variaciones www, paths) - **Dominios con cambios sustanciales:** 0 - **Crecimiento neto real:** +17 dominios (40 nuevos - 23 removidos)

Clasificación de cambios menores (7 dominios): - **ajvs.cl:** www.ajvs.cl → ajvs.cl - **biotaxa.org:** /rce/oai → /index.php/index/oai - **joralres.com:** www.joralres.com → joralres.com - **reidu.cl:** reidu.cl → www.reidu.cl - **revistapleyade.cl:** revista-pleyade.cl → www.revistapleyade.cl - **sisomosamericanos.cl:** sisomosamericanos.cl → www.sisomosamericanos.cl - **teps.cl:** teps.cl → www.teps.cl

1.5.3.2 3.3.2 Decisión Metodológica: Continuidad Operativa v5

Justificación: Dada la alta estabilidad (88.9%) entre versiones y que los datos v5 ya están procesados con informes Dialnet descargados, **se continúa el procesamiento sobre resultados v5** para análisis de calidad y evaluación. Paralelamente, se descargan informes Dialnet para URLs v6 nuevas.

Implicaciones operativas: - **Dataset principal:** chile_oai_urls_limpio.csv (v5) - 208 URLs - **Procesamiento paralelo:** chile_oai_urls_limpio_v6.csv - 225 URLs - **Validación cruzada:** Confirmada estabilidad metodológica

1.5.3.3 3.3.3 Cambios en el Dataset

- **Instalaciones totales Chile:** 396 → 432 (+9.1%)
- **Instalaciones activas:** 309 → 319 (+3.2%)
- **Criterio temporal:** 2023 → 2024 (actualización metodológica)
- **URLs únicas:** 202 → 225 (+11.4%)

1.5.3.4 3.3.4 Consistencia de Resultados

- **Universidad de Chile mantiene liderazgo:** 68 → 67 instalaciones
- **Distribución institucional estable:** Misma institución líder
- **Crecimiento sostenido:** Confirmado en ambas versiones

1.5.4 3.4 Archivos Generados v6

1.5.4.1 3.4.1 Datasets Principales

- `visualizations_v6/chile_todas_instalaciones_v6.csv` - 432 instalaciones completas
- `visualizations_v6/chile_juojs_activas_v6.csv` - 319 instalaciones activas (DATASET PRINCIPAL v6)
- `visualizations_v6/chile_oai_urls_v6.csv` - 225 URLs para evaluación Dialnet

1.5.4.2 3.4.2 Análisis Mundial v6

- `visualizations_v6/tabla_paises_mundial_v6.csv` - 154 países activos
- `visualizations_v6/regiones_mundial_v6.csv` - 8 regiones PKP
- `visualizations_v6/top15_paises_barras_v6.png` - Visualización países
- `visualizations_v6/regiones_circular_v6.png` - Visualización regiones

1.5.4.3 3.4.3 Tablas Chile v6

- `visualizations_v6/tabla_chile_todas_top30_v6.csv`
- `visualizations_v6/tabla_chile_juojs_top30_v6.csv`
- `visualizations_v6/tabla_chile_crecimiento_top15_v6.csv`
- `visualizations_v6/tabla_instituciones_v6.csv`

1.5.5 3.5 Análisis de Calidad Editorial: Informes Dialnet

1.5.5.1 3.5.1 Procesamiento de Informes (Script 8)

Metodología: Se procesaron **95 informes HTML** de Dialnet descargados manualmente, extrayendo información estructurada sobre errores de calidad y completitud de metadatos mediante `scripts/8_procesar_informes_dialnet.py`.

Cobertura del análisis: - **Total informes procesados:** 95 revistas chilenas - **Archivo generado:** `visualizations/dialnet_informes_procesados.csv` - **Datos extraídos:** Errores por gravedad, tipos de errores, porcentajes de metadatos

1.5.5.2 3.5.2 Resultados de Calidad Editorial

Estado general de errores: - **Revistas sin errores:** 0 (0.0%) - **Revistas con errores:** 95 (100.0%) - **Total errores detectados:** 59,602 - **Alta gravedad:** 35,811 (60.1%) - **Media gravedad:** 23,791 (39.9%) - **Promedio errores por revista:** 627.4

Ranking	Dominio	Total Errores	Alta Gravedad	Media Gravedad
---------	---------	---------------	---------------	----------------

1.5.5.3 3.5.3 Revistas con Mayor Número de Errores

Ranking	Dominio	Total Errores	Alta Gravedad	Media Gravedad
1	www.revistasdex.uchile.cl	5,583	2,626	2,957
2	revistas.uach.cl	3,881	1,607	2,274
3	revistas.uv.cl	2,865	1,988	877
4	rhv.uv.cl	2,865	1,988	877
5	www.iace.uv.cl	2,865	1,988	877

1.5.5.4 3.5.4 Completitud de Metadatos

Promedios de completitud por tipo de metadato:

Tipo de Metadato	Completitud Promedio
Títulos	100.0%
Enlaces	100.0%
Autores	99.3%
Resúmenes	87.0%
Afiliaciones	75.6%
Palabras clave	65.6%
Referencias	33.7%

1.5.5.5 3.5.5 Análisis de Tipos de Errores por Gravedad (Script 9)

Top 5 errores de alta gravedad más frecuentes:

Ranking	Tipo de Error	Ocurrencias
1	Artículo sin número de página inicial	12,201
2	Artículo sin ejemplar identificado en la importación	6,964
3	Artículo con título y subtítulo en mayúsculas	6,810
4	Artículo con resumen repetido para la misma revista	2,056
5	Artículo con autorías duplicadas	1,968

Top 5 errores de media gravedad más frecuentes:

Ranking	Tipo de Error	Ocurrencias
1	Autor con apellido 2 de longitud menor de 4 caracteres	6,468
2	Ejemplar marcado como no paginado y contiene artículos con página final	5,257

Ranking	Tipo de Error	Ocurrencias
3	Artículo con título repetido para la misma revista	2,721
4	Artículo con referencias cuyo literal es demasiado corto	2,080
5	Artículo con título repetido	1,843

Distribución de revistas por rangos de errores:

Rango de Errores	Nº Revistas	Porcentaje
Muy Alto (>2000)	9	9.5%
Alto (1000-2000)	10	10.5%
Medio (500-999)	8	8.4%
Bajo (100-499)	43	45.3%
Muy Bajo (<100)	25	26.3%

Revistas con mayor proporción de errores graves:

Ranking	Dominio	% Alta Gravedad	Errores Alta/Total
1	revistaderechovaldivia.cl	97.6%	1,160/1,188
2	intushistoria.uai.cl	97.6%	775/794
3	boletin.scha.cl	94.6%	2,390/2,526
4	nuestramerica.cl	92.9%	351/378
5	revistalimite.uta.cl	91.9%	846/921

Diversidad de tipos de errores:

Categoría	Tipos Únicos
Alta gravedad	23
Media gravedad	17
Total	40

1.5.5.6 3.5.6 Implicaciones para Indexación

Fortalezas identificadas: - **Metadatos básicos completos:** Títulos y enlaces al 100% - **Información de autoría:** 99.3% de completitud - **Contenido académico:** 87% de resúmenes disponibles

Áreas de mejora críticas: - **Referencias bibliográficas:** Solo 33.7% de completitud - **Palabras clave:** 65.6% requiere mejora para indexación - **Errores de alta gravedad:** 35,811 errores requieren atención prioritaria - **Problemas de paginación:** 12,201 artículos sin página inicial - **Estandarización de títulos:** 6,810 casos con mayúsculas incorrectas

Recomendaciones operativas: 1. **Priorización por calidad:** Enfocar esfuerzos en 25 revistas con <100 errores 2. **Mejora de metadatos:** Implementar estándares para referencias y palabras clave 3. **Capacitación editorial:** Programas de mejora de calidad para editores 4. **Corrección de paginación:** Atender 12,201 artículos sin numeración 5. **Estandarización de formato:** Corregir mayúsculas en títulos y subtítulos

1.6 4. Discusión

1.6.1 4.1 Validación del Ecosistema Global

La **validación v6** confirma el crecimiento acelerado del ecosistema OJS global, con **86,282 instalaciones** representando un aumento del **28.4%** respecto al beacon anterior. Este crecimiento es particularmente notable en:

- **Asia-Pacífico:** Indonesia lidera con 23,045 instalaciones activas
- **América Latina:** 8,485 instalaciones confirman la región como hub de acceso abierto
- **Diversificación geográfica:** 154 países con instalaciones activas

1.6.2 4.2 Caracterización del Ecosistema Chileno Actualizada

1.6.2.1 4.2.1 Fortalezas Confirmadas v6

- **Alta actividad editorial:** 73.8% de instalaciones activas (vs 78.0% anterior)
- **Crecimiento sostenido:** +16.5% en publicaciones 2024
- **Diversidad institucional:** 181 instalaciones en “otras instituciones” (56.7%)
- **Productividad destacada:** Promedio 44.4 artículos por instalación

1.6.2.2 4.2.2 Tendencias Identificadas

- **Concentración en líderes:** Universidad de Chile mantiene 21% del ecosistema
- **Crecimiento diferenciado:** Algunas instituciones decrecen mientras otras crecen
- **Nuevas instalaciones:** +36 instalaciones identificadas en v6

1.6.3 4.3 Implicaciones para Indexación en Dialnet

1.6.3.1 4.3.1 Optimización del Proceso

- **225 URLs únicas** (vs 202 anteriores) para evaluación
- **Eliminación automática** de 94 duplicados
- **Cobertura mejorada:** 81.5% de publicaciones 2024 representadas

1.6.3.2 4.3.2 Estrategia de Priorización

El análisis v6 permite priorizar instalaciones por: 1. **Productividad 2024:** Revista Chilena de Derecho (1,708 pub) 2. **Crecimiento:** Instalaciones con tendencia positiva 3. **Estabilidad institucional:** Universidad de Chile (58 URLs)

1.6.4 4.4 Contribución al Acceso Abierto Regional

Los **144,529 artículos históricos** acumulados (vs 133,931 anteriores) posicionan a Chile como contribuyente significativo al ecosistema latinoamericano, con un **crecimiento del 7.9%** en el acervo histórico.

1.6.5 4.5 Limitaciones y Consideraciones

1. **Criterio temporal:** Cambio de 2023 a 2024 puede afectar comparabilidad directa
 2. **Cobertura beacon:** Dependencia de instalaciones que mantienen beacon activo
 3. **Variabilidad institucional:** Algunas instituciones muestran decrecimiento
-

1.7 5. Limitaciones del Estudio

1.7.1 5.1 Limitaciones Metodológicas

- **Dependencia del PKP Beacon:** Solo incluye instalaciones con beacon activo
- **Criterio temporal:** Cambio de 2023 a 2024 afecta comparabilidad histórica
- **Cobertura geográfica:** Posible subrepresentación de instalaciones sin identificación clara de país

1.7.2 5.2 Limitaciones Técnicas

- **Endpoints OAI:** Variabilidad en configuración y disponibilidad
- **Duplicados:** Posibles instalaciones múltiples de la misma revista
- **Datos faltantes:** Instalaciones sin ISSN o metadatos incompletos

1.7.3 5.3 Limitaciones de Alcance

- **Evaluación Dialnet:** Proceso manual limita escalabilidad
 - **Visibilidad académica:** Análisis OpenAlex pendiente de integración completa
 - **Impacto editorial:** No incluye métricas de citación o factor de impacto
-

1.8 6. Conclusiones

1.8.1 5.1 Validación del Ecosistema Robusto

La **validación v6** confirma que Chile mantiene un ecosistema editorial académico **robusto y en crecimiento** con: - **432 instalaciones OJS** identificadas (+9.1% vs versión anterior) - **319 instalaciones activas** con criterio 2024 (73.8%) - **14,162 publicaciones en 2024** (+16.5% vs 2023)

1.8.2 5.2 Liderazgo Institucional Confirmado

La **Universidad de Chile** mantiene su posición dominante con: - **67 instalaciones** (21% del ecosistema) - **3,338 publicaciones en 2024** - **Crecimiento de +1,646 publicaciones**

1.8.3 5.3 Crecimiento Sostenido Validado

El **51.1% de instalaciones** muestra crecimiento positivo, confirmando la vitalidad del sistema editorial chileno en el contexto global de acceso abierto.

1.8.4 5.4 Metodología Consolidada

La **validación cruzada v5-v6** demuestra: - **Consistencia metodológica:** Resultados coherentes entre versiones - **Actualización exitosa:** Criterios 2024 reflejan estado actual - **Escalabilidad:** Procedimientos replicables para análisis futuros

1.8.5 5.5 Posicionamiento Global

Chile se posiciona como **actor relevante** en el ecosistema latinoamericano de acceso abierto, contribuyendo significativamente a la democratización del conocimiento científico regional.

1.9 7. Próximos Pasos

1.9.1 7.1 Evaluación Dialnet

- **Procesamiento completado:** 95 informes HTML analizados (Scripts 8 y 9)
- **Resultados disponibles:**
 - `dialnet_informes_procesados.csv` con métricas de calidad
 - Análisis de tipos de errores por gravedad completado
- **Próximos pasos:** Implementación de recomendaciones de mejora editorial
- **Expansión:** Procesamiento de URLs v6 nuevas (40 URLs adicionales)
- **Priorización:** Enfocar en 25 revistas con <100 errores para indexación rápida

1.9.2 6.2 Enriquecimiento Continuo

- **Integración OpenAlex** con datos 2024
- **Análisis longitudinal** de tendencias
- **Monitoreo** de nuevas instalaciones

1.9.3 6.3 Análisis Comparativo Regional

- **Benchmarking** con otros países latinoamericanos
 - **Identificación** de mejores prácticas
 - **Colaboración** regional en acceso abierto
-

1.10 Agradecimientos

Al Public Knowledge Project (PKP) por mantener el sistema de beacon que hace posible este análisis. A la Universidad de La Rioja y el equipo de Dialnet por facilitar el acceso a su sistema de evaluación de revistas. A las instituciones chilenas que mantienen instalaciones OJS activas, contribuyendo al ecosistema de acceso abierto regional.

1.11 Referencias

1.11.1 Estudios sobre JUOJS y PKP Beacon

Khanna, S., Ball, J., Alperin, J. P., & Willinsky, J. (2022). Recalibrating the scope of scholarly publishing: A modest step in a vast decolonization process. *Quantitative Science Studies*, 3(4), 912–930. https://doi.org/10.1162/qss_a_00228

Khanna, S., Raoni, J., Smecher, A., Alperin, J. P., Ball, J., & Willinsky, J. (2024). *Details of publications using software by the Public Knowledge Project (V4)* [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/OCZNVY>

1.11.2 Estudios sobre indexación y acceso abierto

Beigel, F. (2025). The transformative relation between publishers and editors: Research quality and academic autonomy at stake. *Quantitative Science Studies*, 6, 154–170. https://doi.org/10.1162/qss_a_00343

Chavarro, D., Alperin, J. P., & Willinsky, J. (2025). On the open road to universal indexing: OpenAlex and Open Journal Systems. *Quantitative Science Studies*, 6, 1039–1058. <https://doi.org/10.1162/QSS.a.17>

Simard, M.-A., Basson, I., Hare, M., Larivière, V., & Mongeon, P. (2025). Examining the geographic and linguistic coverage of gold and diamond open access journals in OpenAlex, Scopus, and Web of Science. *Quantitative Science Studies*, 6, 732–752. <https://doi.org/10.1162/qss.a.1>

1.11.3 Estudios sobre publicación científica chilena y latinoamericana

Collazo-Reyes, F. (2014). Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. *Scientometrics*, 98, 197–209. <https://doi.org/10.1007/s11192-013-1036-2>

Koch, T., & Vanderstraeten, R. (2018). Internationalizing a national scientific community? Changes in publication and citation practices in Chile, 1976–2015. *Current Sociology*, 67(5), 723–741. <https://doi.org/10.1177/0011392118807514>

Koch, T., Vanderstraeten, R. y Ayala, R. (2021). La internacionalización de la ciencia: revistas y comunidades chilenas en el mundo científico. *Estudios Sociales de la Ciencia*, 51(1), 121–138. <https://doi.org/10.1177/0306312720949709>

Krauskopf, M. y Pessot, R. (1980). Estudio preliminar sobre publicaciones y productividad científica en Chile. *Archivos de Biología y Medicina Experimentales*, 73, 195–208.

Sabaj, O., Valderrama, J. O., González-Vergara, C., & Pina-Stranger, A. (2015). Relationship between the duration of peer-review, publication decision, and agreement among reviewers in three Chilean journals. *HAL Archives Ouvertes*, halshs-01242627.

1.11.4 Marco teórico y metodológico

Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology*, 61(8), 1622–1634. <https://doi.org/10.1002/asi.21347>

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.

Paasi, A. (2015). Academic Capitalism and the Geopolitics of Knowledge. In J. Agnew, V. Mamadouh, A. J. Secor, & J. Sharp (Eds.), *The Wiley Blackwell Companion to Political Geography* (pp. 507–523). Wiley Blackwell. <https://doi.org/10.1002/9781118725771.ch37>

1.11.5 Recursos técnicos

- **PKP Official Site:** <https://pkp.sfu.ca/>
- **Dataset URL:** <https://pkp.sfu.ca/software/ojs/usage-data/>
- **Harvard Dataverse:** <https://doi.org/doi:10.7910/DVN/OCZNVY>
- **MARC Country Codes:** <https://www.loc.gov/marc/countries/>
- **OAI-PMH Protocol:** <https://www.openarchives.org/pmh/>
- **OpenAlex API:** <https://docs.openalex.org/>
- **Dialnet API:** <https://dialnet.unirioja.es/ws/dialnetcris-sandbox/v2/swagger-ui/index.html>

1.11.6 Estándares y criterios

- Directory of Open Access Journals (2020). Criterio de revista activa: 5 artículos por año
- ISO 3166-1 alpha-2: Códigos de países
- ISSN International Centre: Estándares de numeración de publicaciones seriadas

1.11.7 Datasets utilizados

Beacon v6: Khanna, S., Raoni, J., Smecher, A., Alperin, J. P., Ball, J., & Willinsky, J. (2024). *Details of publications using software by the Public Knowledge Project (V6)* [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/OCZNVY>

Beacon anterior: Khanna, S., Raoni, J., Smecher, A., Alperin, J. P., Ball, J., & Willinsky, J. (2024). *Details of publications using software by the Public Knowledge Project (V4)* [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/OCZNVY>

1.12 Colaboraciones

Geraldine Trujillo - Procesamiento y descarga manual de informes Dialnet - Evaluación sistemática de URLs en validador Nexus - Organización y carga de archivos HTML

Claude (Anthropic) - Procesamiento y optimización de scripts de análisis - Apoyo en redacción y estructuración del informe - Asistencia técnica vía Claude Web y Amazon Q Developer

1.13 Anexos

1.13.1 Anexo A: Diccionario de Datos del PKP Beacon

1.13.1.1 Campos descriptivos

Campo	Ejemplo	Descripción
application	OJS	Tipo de aplicación (OJS, OMP o OPS)
version	3.2.1.4	Versión del software
country_consolidated	CL	País resuelto/final asignado
set_spec	psykhe	Identificador de revista específica
context_name	Psykhe	Nombre completo de la revista
issn	0718-2228	ISSN de la revista

1.13.1.2 Campos de conteo y actividad

Campo	Ejemplo	Descripción
total_record_count	872	Total artículos publicados históricos
record_count_2023	41	Artículos publicados en 2023
last_completed_update	2024-10-11	Última sincronización
unresponsive_endpoint	0	Estado de respuesta (0=activo, 1=inactivo)

1.13.2 Anexo B: Scripts Desarrollados

1.13.2.1 Scripts de análisis general

- `scripts/1_split_beacon.py`: Separación por aplicación
- `scripts/2_analisis_ojs_mundial.R`: Análisis global
- `scripts/visualize_network_enhanced.py`: Análisis de redes
- `scripts/visualize_interactive.py`: Dashboards interactivos

1.13.2.2 Scripts específicos de Chile

- `scripts/3_analisis_chile.R`: Análisis chileno completo
- `scripts/4_chile_juojs_filtrado.R`: Filtrado JUOJS (dataset principal)
- `scripts/5_generar_urls_dialnet.py`: Generación URLs Dialnet
- `scripts/6_find_missing_reports.py`: Verificación completitud informes Dialnet
- `scripts/7_comparar_urls_v5_v6.py`: Comparación cruzada v5-v6
- `scripts/8_procesar_informes_dialnet.py`: Procesamiento informes HTML Dialnet
- `scripts/9_analizar_tipos_errores.py`: Análisis de tipos de errores por gravedad

1.13.2.3 Script de enriquecimiento

- `scripts/openalex.py`: Integración con OpenAlex API

1.13.3 Anexo C: Proceso de Evaluación en Dialnet

1.13.3.1 Estado actual

Proceso manual de envío de 246 URLs únicas de instalaciones chilenas a través del portal web de Dialnet para solicitud de informes de calidad.

1.13.3.2 Organización de informes descargados

Nomenclatura de archivos: Los informes de Dialnet se descargan y renombran siguiendo el patrón del dominio extraído de la URL OAI: - Se mantiene el formato exacto del dominio (incluyendo `www` si está presente) - Se omite `www` si no aparece en la URL original - Formato: `[dominio].pdf` o `[dominio].html`

Ejemplos de nomenclatura:

- <code>https://revistas.udec.cl/index.php/index/oai</code>	→	
<code>revistas.udec.cl.pdf</code>	- <code>https://www.estudiospublicos.cl/index.php/index/oai</code>	→
<code>www.estudiospublicos.cl.pdf</code>	- <code>http://revistas.uach.cl/index.php/index/oai</code>	→
<code>revistas.uach.cl.pdf</code>		

Estructura de carpetas:

```
dialnet/  
  revistas.udec.cl.pdf  
  www.estudiospublicos.cl.pdf  
  revistas.uach.cl.pdf  
  ...
```

1.13.3.3 Registro de errores del sistema Nexus

Metodología de registro: Cuando el sistema Nexus de Dialnet arroja errores durante la consulta, estos se registran en el archivo `visualizations/chile_oai_urls_limpio.csv` agregando una coma seguida del mensaje de error.

Formato de registro:

```
oai_url,error_message  
https://revistas.udec.cl/index.php/index/oai,Error al consultar la configuración de la revista  
https://www.biotaxa.org/index.php/index/oai,La url no corresponde a una revista recolectable p
```

Tipos de errores identificados: 1. **Error al leer el XML:** Problemas de formato o accesibilidad del endpoint OAI 2. **URL no recolectable:** Instalaciones que no cumplen estándares OAI-PMH 3. **Timeout de conexión:** Instalaciones temporalmente inaccesibles 4. **Configuración incorrecta:** Problemas en la configuración del servidor OJS

1.13.3.4 Limitaciones identificadas

- No existe API para automatizar solicitudes de informes
- Requiere evaluación manual por parte del equipo de Dialnet
- Proceso individual por cada URL de instalación
- Algunos endpoints OAI presentan problemas técnicos

1.13.3.5 Alternativas evaluadas

- Contacto directo para procesamiento en lote
- Priorización por índice de visibilidad
- Automatización de formularios (descartada por términos de servicio)
- Pre-validación técnica de endpoints OAI antes del envío

1.13.4 Anexo D: Reproducibilidad

1.13.4.1 Requisitos del sistema

- R 4.0+ con librerías: dplyr, ggplot2, readr, stringr
- Python 3.8+ con librerías: pandas, requests, urllib
- Sistema operativo: Linux/macOS (scripts optimizados para bash)

1.13.4.2 Repositorio de código

- **Ubicación:** /scripts/ y /scripts_v6/
- **Orden de ejecución:** Numeración secuencial (1-7)
- **Datos de entrada:** beacon.csv (v5) y beacon_v6_clean.csv (v6)
- **Datos de salida:** /visualizations/ y /visualizations_v6/

1.13.4.3 Estructura de archivos generados

```
OJSinTheWorld/Chile/
  beacon.csv                # Dataset v5 original
  beacon_v6_clean.csv       # Dataset v6 limpiado
  beacon_ojs.csv            # OJS filtrado v5
  beacon_v6_ojs.csv         # OJS filtrado v6
  scripts/
    1_split_beacon.py
    2_analisis_ojs_mundial.R
    3_analisis_chile.R
    4_chile_juojs_filtrado.R
    5_generar_urls_dialnet.py
    6_find_missing_reports.py
    7_comparar_urls_v5_v6.py
    8_procesar_informes_dialnet.py
    9_analizar_tipos_errores.py
  scripts_v6/
    1_split_beacon_v6.py
    2_analisis_mundial_v6.R
    3_analisis_chile_v6.R
    4_chile_juojs_filtrado_v6.R
    5_generar_urls_dialnet_v6.py
    6_tablas_chile_v6.R
  visualizations/           # Resultados v5
    chile_todas_instalaciones.csv
    chile_juojs_activas.csv
    chile_oai_urls_limpio.csv
    dialnet_informes_procesados.csv
  visualizations_v6/        # Resultados v6
    chile_todas_instalaciones_v6.csv
    chile_juojs_activas_v6.csv
    chile_oai_urls_limpio_v6.csv
  dialnet/                  # Informes descargados
    revistas.udec.cl.pdf
    www.estudiospublicos.cl.pdf
    ...
```