

XXI Workshop de Proyectos Digitales REBIUN

Universidad de Castilla-La Mancha, 28 y 29 de
mayo de 2026

**Análisis de problemas de
funcionamiento producidos
por BOTS e I.A.**

Datos preliminares.

Análisis de problemas de funcionamiento producidos por BOTS e I.A.

Catalina Guzmán (UCO)

Oriol Olivé (UdG)

Ana Poveda (UC3M)

José Antonio Vidal (UPCT)

REBIUN Línea 2. Grupo de Repositorios

URI: <https://hdl.handle.net/20.500.11967/6745>

Miembros de la acción

Catalina Guzmán (UCO)

Oriol Olivé (UdG)

Ana Poveda (UC3M)

José Antonio Vidal (UPCT)

Análisis de problemas de funcionamiento producidos por BOTS e I.A.

45

Respuestas de 71 instituciones
REBIUN

63%

Tasa de respuesta a la
encuesta

42

Usan DSpace (93%)

82%

Reportan impacto por bots

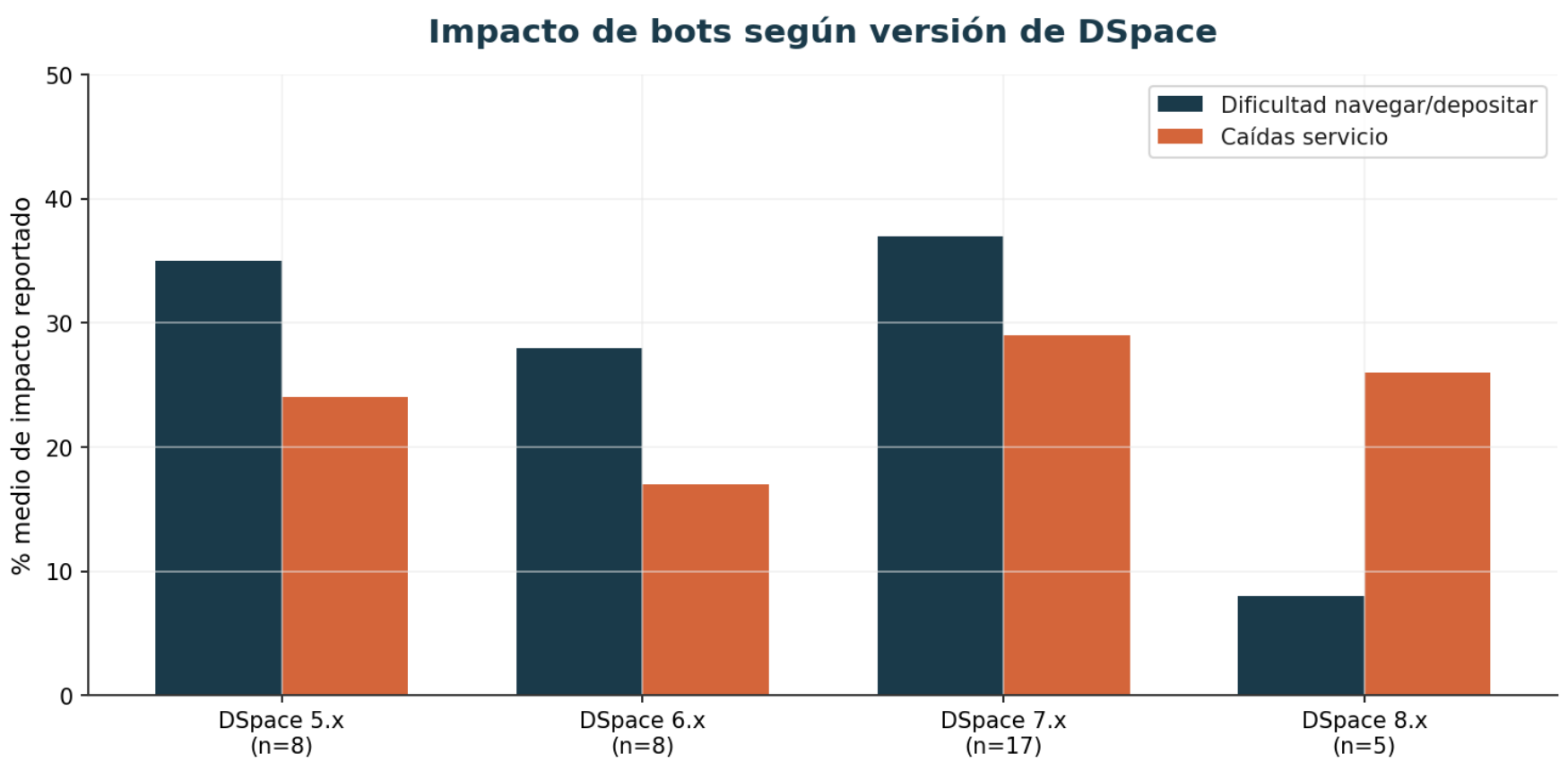
Según los datos:

La mayoría de los repositorios españoles (82%) sufre impacto real por el tráfico de bots de IA.

La versión de Dspace influye: **DSpace 5.x presenta un 100% de impacto frente al 40% de DSpace 8.x.**

La comunidad está dispuesta a colaborar pero necesita herramientas compartidas para gestionar de forma coordinada.

DSpace 5.x reporta 100% de impacto por bots; DSpace 8.x lo reduce al 40%

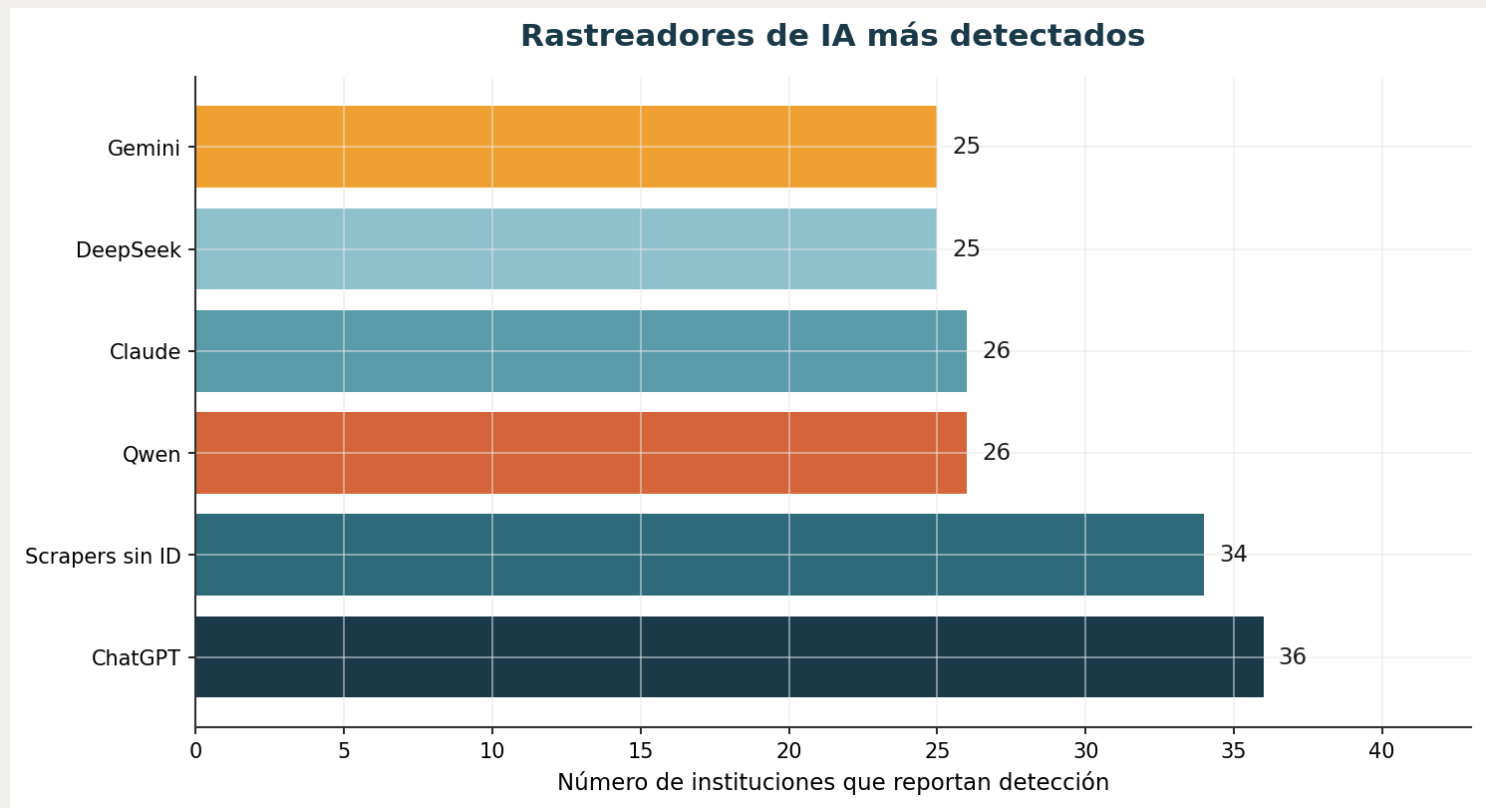


Impacto por versión de DSpace

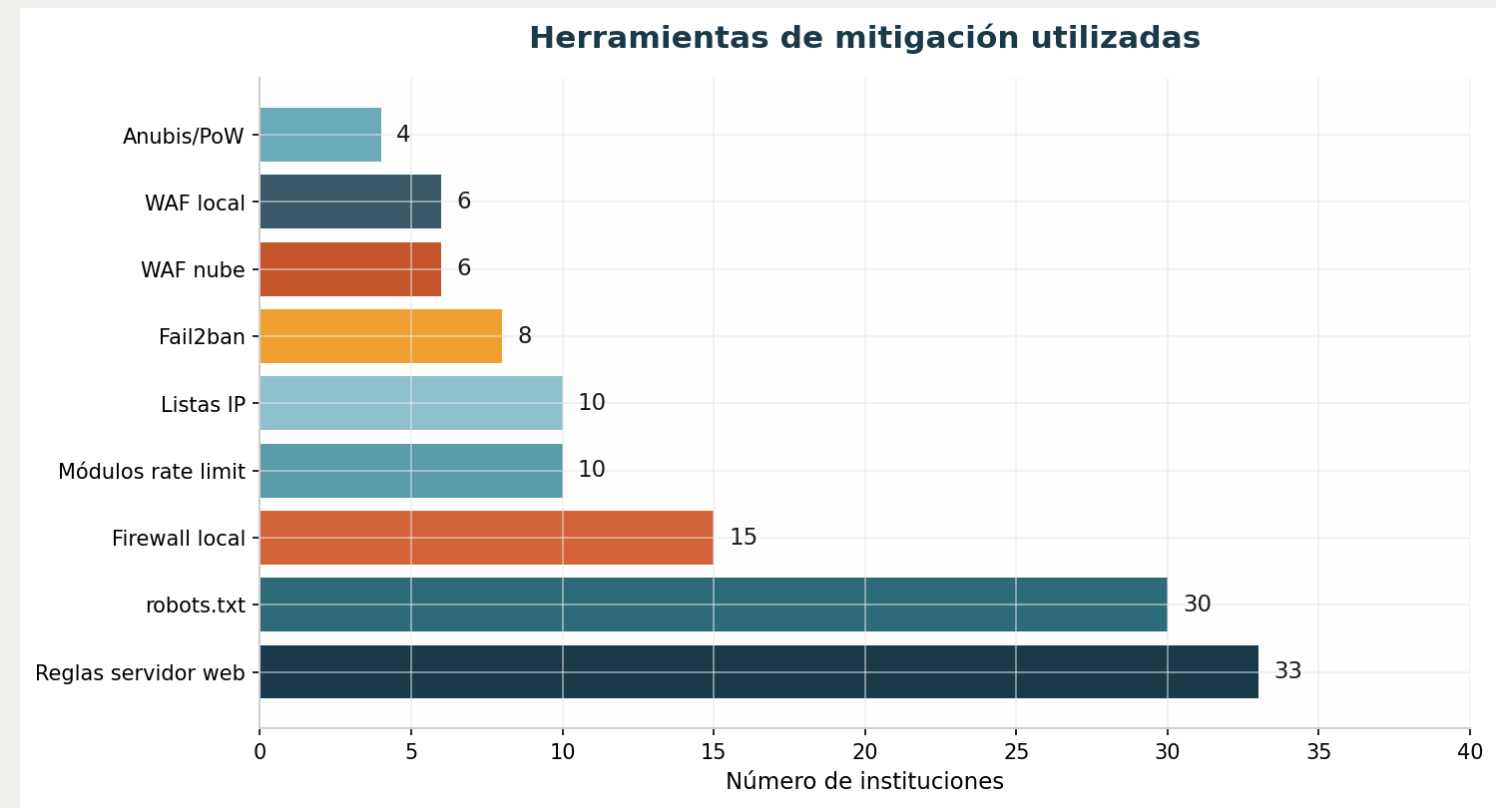
Versión	n	Dificultad	Caídas	Sin impacto
DSpace 5.x	8	100%	100%	0%
DSpace 6.x	8	75%	50%	25%
DSpace 7.x	17	82%	76%	18%
DSpace 8.x	5	40%	40%	60%

ChatGPT domina el 80% de detecciones; la mitigación sigue siendo reactiva

Rastreadores de IA más detectados



Herramientas de mitigación utilizadas

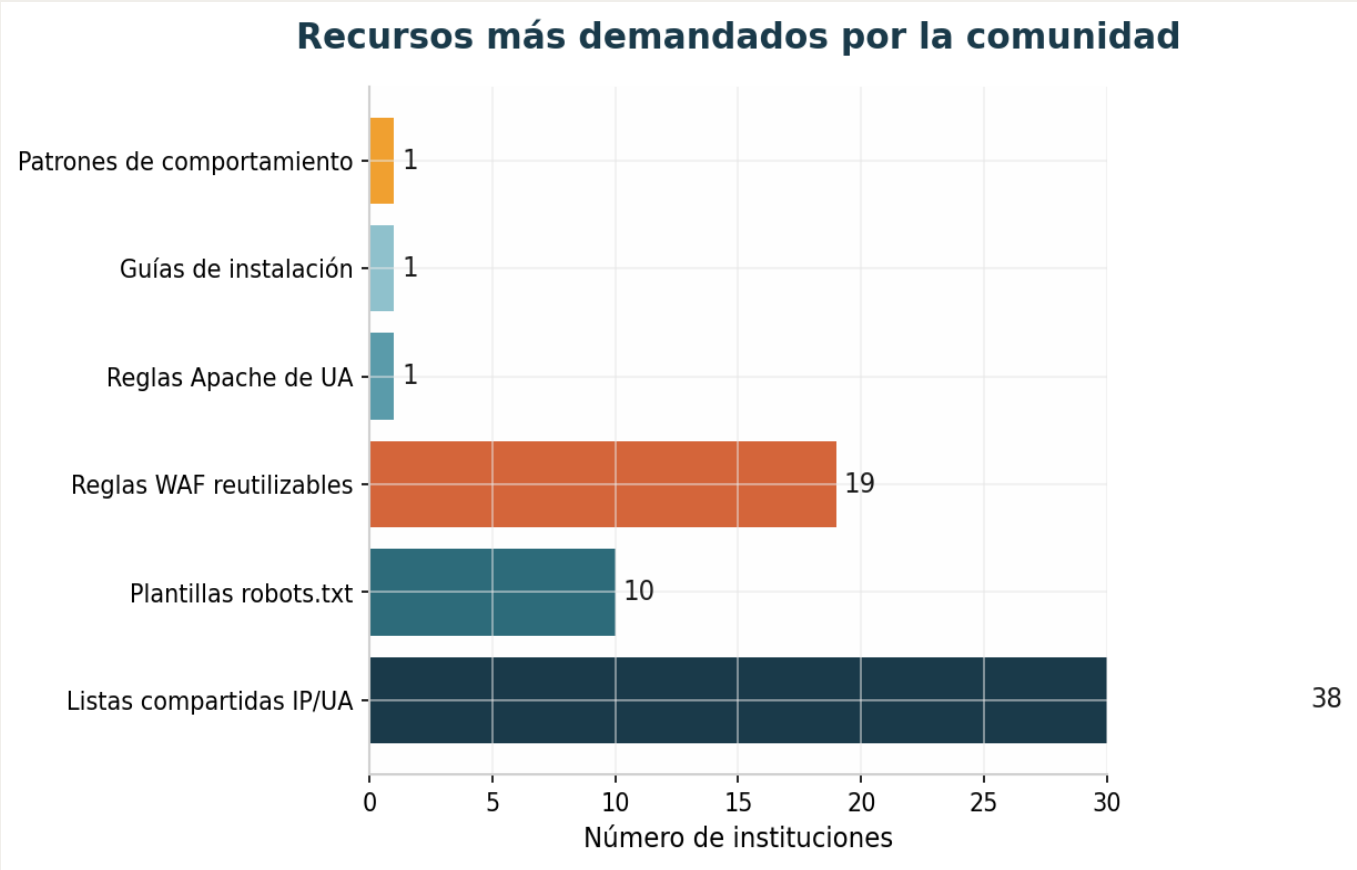


76% reportan scrapers sin identificar — el desafío más difícil de mitigar

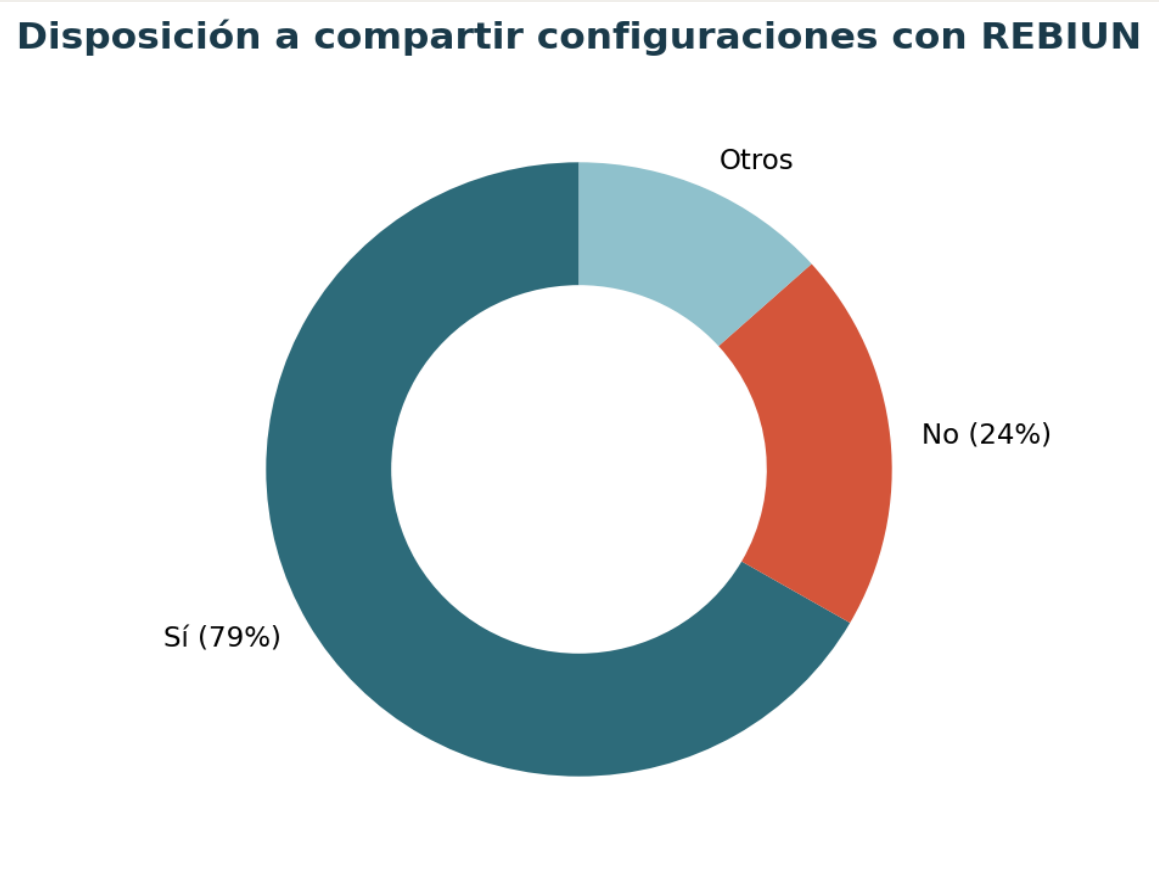
49% bloquean accidentalmente usuarios humanos — complejo balance seguridad/acceso

El 79% compartiría configuraciones; priorizan listas de IPs y reglas WAF

Recursos más demandados



Disposición a compartir con REBIUN



27 instituciones demandan listas compartidas de IPs y User-Agents problemáticos

17 instituciones solicitan reglas WAF reutilizables — segunda demanda prioritaria

Hallazgos y recomendaciones clave:

- 1 DSpace 8.x** — Reduce el impacto de bots del 100% al 40% gracias a "Manage AI bots" nativo y mejor rate limiting.
- 2 ChatGPT y scrapers sin ID son la amenaza principal** — El 80% detecta ChatGPT y el 76% reporta scrapers no identificables, lo que exige soluciones basadas en comportamiento, no solo en firma.
- 3 La comunidad demanda colaboración** — El 79% compartiría configuraciones. Prioridades: listas compartidas de IPs/UA (27), reglas WAF reutilizables (17), plantillas robots.txt (10).
- 4 Crear red de colaboración REBIUN** — El 49% bloquea usuarios legítimos por error. Una red para compartir listas, reglas WAF y patrones de comportamiento reduciría este riesgo y fortalecería la defensa colectiva.