

8.

Оновлення/видалення контенту

- 8.1. Частота оновлення
- 8.2. Неповна індексація
- 8.3. Не працює автоматичне оновлення
- 8.4. Не працює оновлення змін до документу

8.1. Частота оновлення

Як часто ви оновлюєте вміст проіндексованих постачальників контенту? Ми оновлюємо всі проіндексовані постачальники контенту двічі на місяць. З більшими інтервалами весь вміст повністю збирається і переіндексується.

8.2. Неповна індексація

Чому джерело проіндексовано неповністю? Індексція запису/документа, який додається до інтерфейсу OAI постачальника контенту, в Харвестер може зайняти до 6 тижнів. Якщо дані не оновлюються протягом тривалого періоду часу, це, як правило, пов'язано з проблемами в інтерфейсі OAI постачальника контенту. Також може статися так, що записи постачальника контенту надаються неправильно або взагалі не надаються через інтерфейс OAI постачальника контенту. Документи можуть бути надані коректно/повністю через веб-інтерфейс постачальника контенту, але оскільки ми індексуємо лише метадані OAI, веб-інтерфейс не має для нас значення. Ви можете повідомити про помилки за допомогою нашої контактної форми. Якщо ви є менеджером постачальника контенту, будь ласка, використовуйте наш валідатор OAI OVAL для перевірки на відповідність та дотримуйтесь наших “золотих” правил для менеджерів репозитаріїв.

8.3. Не працює автоматичне оновлення

Не працює автоматичне оновлення доданих документів або журналів. Мій журнал вже індексується в Харвестер. Чи потрібно мені надсилати вам повідомлення про публікацію нової статті/випуску?

Ні, контент всіх індексованих постачальників контенту автоматично оновлюється на регулярній основі. Новий випуск/стаття, опублікований у джерелі, що вже індексується в Харвестер, індексується автоматично. Якщо він не був проіндексований через 6 тижнів, то, як правило, є проблема з інтерфейсом OAI постачальника контенту. У такому випадку, будь ласка, надішліть нам повідомлення з детальною інформацією (принаймні назву та URL постачальника контенту).

Чи потрібно повідомляти вас про додавання нового журналу до журнальної платформи, яка вже індексується в Харвестер?

Ні, контент усіх проіндексованих постачальників контенту автоматично оновлюється на регулярній основі. Якщо журнал додається до ресурсу, який вже індексується в Харвестер, він буде автоматично проіндексований. Якщо він не був проіндексований через 6 тижнів, то, як правило, є проблема з інтерфейсом OAI контент-провайдера. У такому випадку, будь ласка, надішліть нам повідомлення з детальною інформацією (принаймні, назву та URL-адресу вашого журналу і журнальної платформи, частиною якої він є).

Я опублікував документ у джерелі, яке індексується в Харвестер, але не можу знайти свій документ у Харвестер. Чому?

Потрібен певний час, щоб нові документи також були проіндексовані в Харвестер. Якщо ваш документ не проіндексовано через 6 тижнів, це, як правило, пов'язано з проблемою інтерфейсу OAI постачальника контенту. У такому випадку, будь ласка, надішліть нам повідомлення з детальною інформацією (принаймні назву та URL постачальника контенту, назву та URL документа).

8.4. Не працює оновлення змін до документу

Я вніс зміни до документів у джерелі. Чому ці зміни не оновлюються в Харвестер?

Будь-яка наступна зміна запису має бути позначена в інтерфейсі OAI постачальника контенту шляхом надання змінених записів під час інкрементного збору. Усі постачальники індексованих даних оновлюються в Харвестер на регулярній, інкрементній основі. Це означає, що ми використовуємо запит за датою (від), щоб перевірити, чи оновлювався ваш інтерфейс OAI з моменту останнього запуску індексування. Якщо змінені записи доставлені некоректно, оновити індекс Харвестер неможливо, і документ залишиться незмінним, а отже, некоректним в індексі, доки ми повністю не переіндексуємо постачальника контенту, що трапляється лише через довші проміжки часу.

????????? ??????????

Я хотів би видалити документ з індексу. Чи можливо це зробити?

Харвестер - це пошукова система. Ми індексуємо лише ті документи, які надають у відкритий доступ постачальники контенту. Ми не можемо видаляти записи безпосередньо з нашого індексу, оскільки вони будуть повторно проіндексовані при наступному індексуванні джерела. Якщо ви вважаєте, що документ має бути вилучений з Харвестер, ви повинні зв'язатися з постачальником контенту і попросити його видалити документ зі свого джерела. Зверніть увагу, що документи також поширюються через агрегатори, а це означає, що на документ можуть бути посилання в різних джерелах, а отже, він буде проіндексований в Харвестер кілька разів. Для того, щоб вилучити документ з індексу Харвестер, його потрібно вилучити у всіх постачальників контенту. Зверніть також увагу, що іноді існують законодавчі вимоги щодо вилучення документів з індексу Харвестер.

?? ?????? ?????????? ??????????

Я видалив документ з джерела. Чому він не видаляється з Харвестер?

Якщо документ видалено з джерела, запис має бути позначений як видалений в інтерфейсі OAI і доставлений під час інкрементального збору. За жодних обставин запис не можна повністю видаляти з інтерфейсу OAI. Якщо документ не буде доставлено і позначено як видалений під час інкрементального збору, його неможливо буде видалити з індексу Харвестер, і документ залишиться помилковим в індексі. У цьому випадку його можна видалити лише шляхом повторного збору та повної переіндексації джерела. Це можна зробити лише через довші проміжки часу.

??????? ?????????? ??????????

Джерело/документ, очевидно, було вилучено з Харвестер. Чому це сталося?

Ми регулярно перевіряємо сервери всіх індексованих постачальників контенту. Якщо сервер / інтерфейс OAI не працює взагалі або працює дуже ненадійно протягом тривалого періоду часу, цей сервер тимчасово або назавжди видаляється з індексу. Це також робиться, якщо постачальник контенту більше не пропонує контент у відкритому доступі. Також може статися так, що окремі документи більше не доступні через інтерфейс OAI від постачальника контенту. Тоді ці документи також видаляються з Харвестер при оновленні контенту. У цьому випадку перевірте, чи інтерфейс OAI постачальника контенту працює правильно і чи документ все ще міститься в інтерфейсі OAI постачальника контенту. Якщо це так, ви можете надіслати нам повідомлення через контактну форму і попросити перевірити.

Джерело/видавець здається "профанацією"/"фальшивою наукою" або пропонує рекламу / спам. Чи можна видалити це джерело?

Ми усвідомлюємо, що хижацька публікація є величезною проблемою, і ми намагаємося перевірити, чи є, наприклад, журнал хижацьким, перш ніж (не) індексувати його, але це дуже складно, оскільки ми можемо виконати лише коротку перевірку – часто лише технічну перевірку постачальника контенту в цілому – і ми не можемо перевірити контент, наприклад, журналу, детально. Харвестер індексує лише ті джерела, які надають робочий інтерфейс OAI, і це часто є технічним бар'єром для таких журналів. Однак деякі статті з цих журналів індексуються через загальних ("відкритих для всіх") постачальників контенту або агрегаторів. Тільки постачальник контенту / агрегатор може видалити цей контент зі свого джерела.

На жаль, може статися так, що постачальник контенту (наприклад, журнальна платформа) припиняє надання своїх послуг і продає домен. Тоді може статися так, що новий оператор пропонує контент, який не має нічого спільного з тим, що ми індексували спочатку.

Як тільки ми помічаємо подібні проблеми, постачальник контенту негайно видаляється з нашого індексу. Ви можете допомогти нам, повідомивши про таких постачальників контенту через нашу контактну форму.