### 인덱스 관리

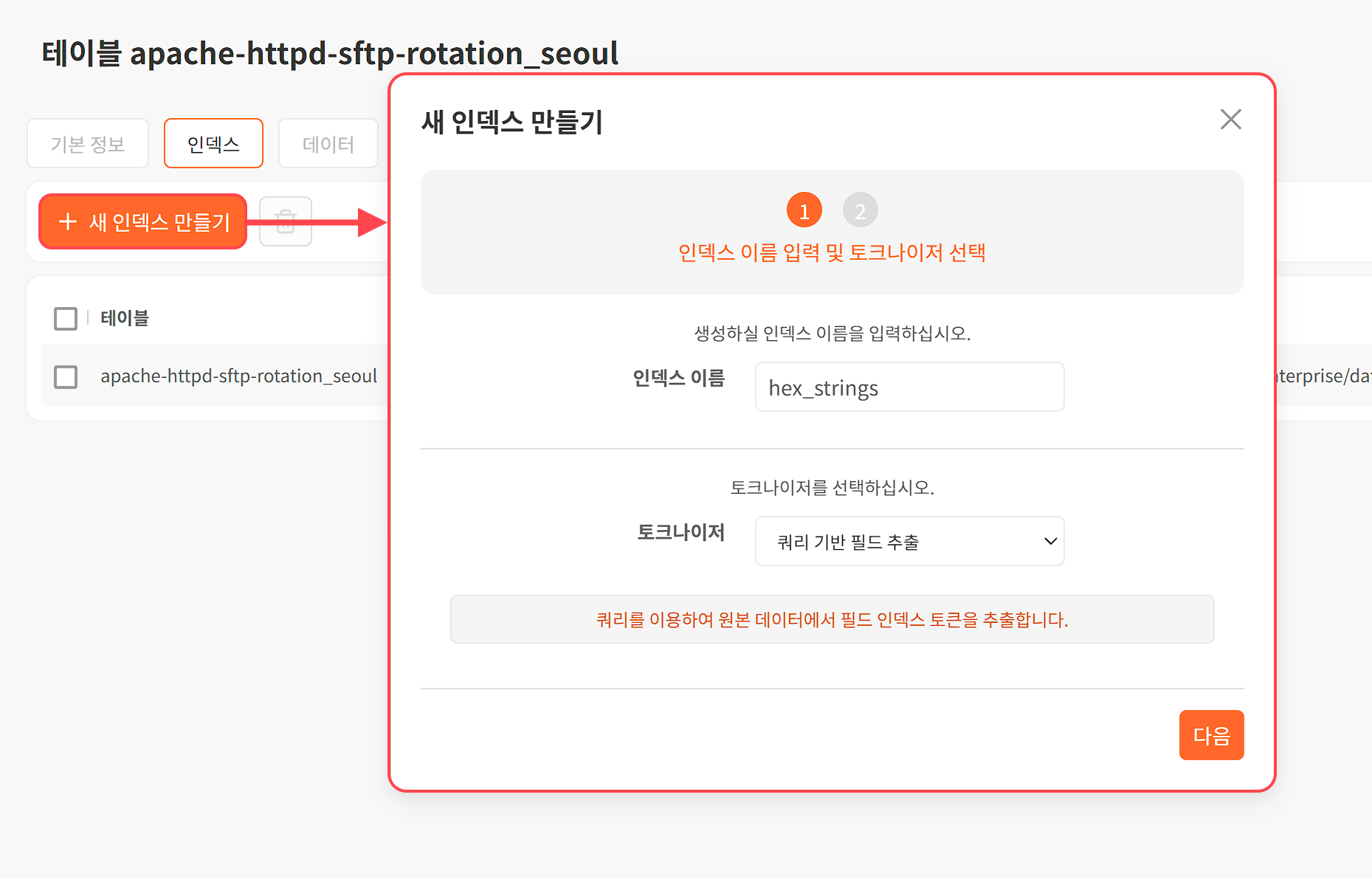
인덱스는 테이블에 저장된 데이터를 빠르게 검색하거나 필요한 데이터를 추출하는데 사용됩니다. 로그프레소 플랫폼이 제공하는 인덱스는 블룸 필터와 역인덱스의 조합으로 구성되고, 하나의 테이블에 여러 개의 인덱스를 적용할 수 있습니다.

#### 인덱스 생성

인덱스를 생성하는 방법은 다음과 같습니다.

테이블 트리에서 인덱스를 생성할 테이블을 선택합니다.

**인덱스** 탭 패널에서 도구 모음에 있는 **새 인덱스 만들기**를 누릅니다.



**새 인덱스 만들기** 창에서 **인덱스 이름**을 입력하고, **토크나이저**를 선택한 후 **다음**을 선택합니다.

* **인덱스 이름**에 사용할 수 있는 문자는 영어, 숫자, 밑줄(\_)입니다.
* **토크나이저**는 다음 중에서 선택합니다(기본값: 쿼리 기반 필드 추출).
* **[쿼리 기반 필드 추출](index-management#outline3)**: 원본 데이터에서 쿼리를 이용해 필드 인덱스 토큰을 추출
* **[구분자 및 IP](index-management#outline4)**: 구분자를 이용해 풀텍스트 토큰을 추출하고, 추가로 IPv4 주소 형식의 토큰을 추출

토크나이저는 원본 로그를 풀텍스트 인덱싱할 때 사용될 풀텍스트 토큰을 만드는 역할을 수행합니다.

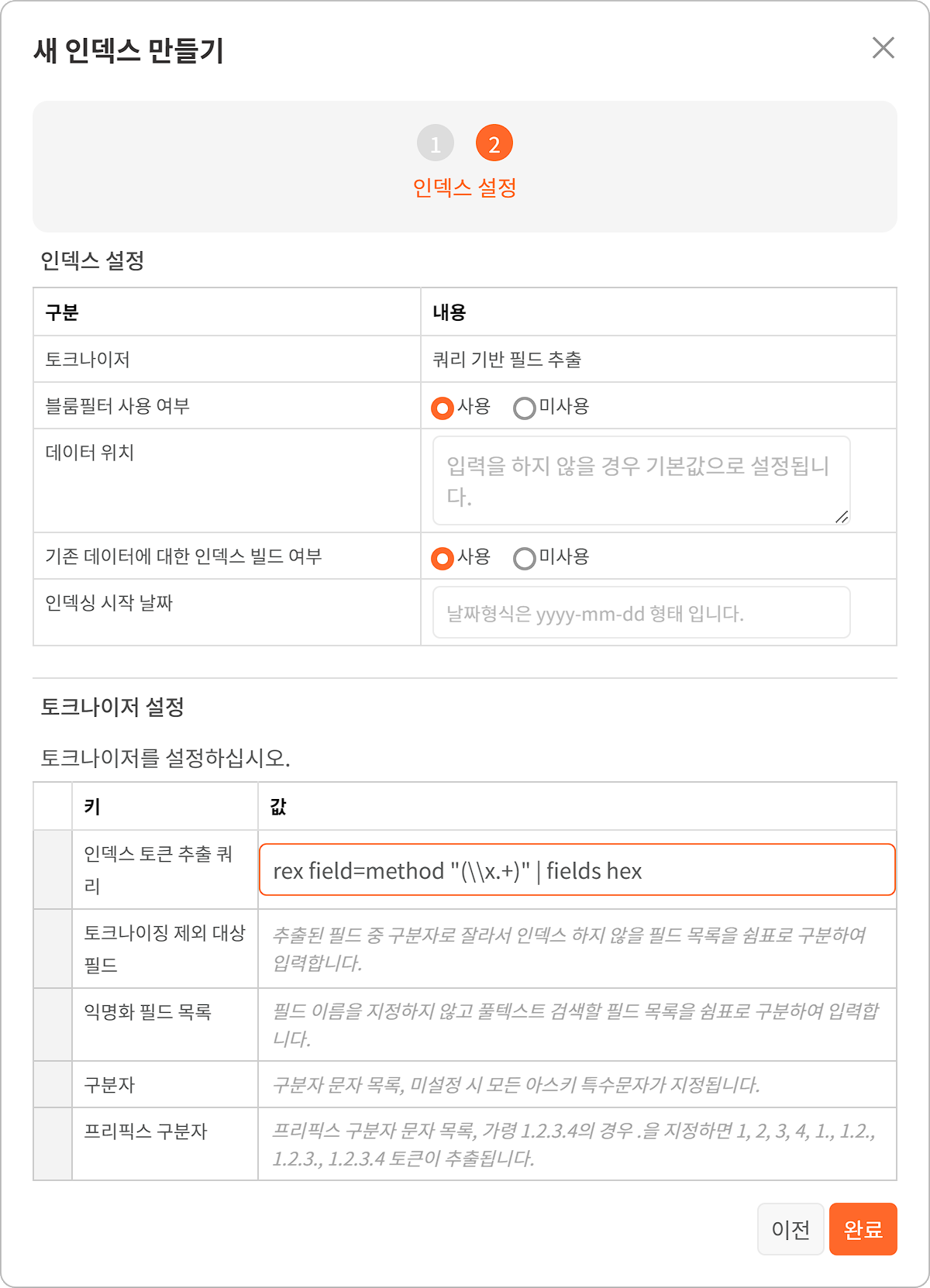
인덱스 구성에 필요한 내용과 토크나이서 속성을 입력하고 **완료**를 누릅니다.

* **블룸필터 사용 여부**: 블룸 필터 사용 여부(기본값: **사용**). 디스크 공간을 추가로 소모하지만 검색 성능을 높일 수 있습니다.
* **데이터 위치**: 인덱스 데이터를 저장할 루트 디렉터리 경로. 지정하지 않으면 로그프레소 환경 변수 araqne.data.dir으로 지정한 디렉터리 아래에 있는 araqne-logstorage/index에 저장됩니다.
* **기존 데이터에 대한 인덱스 빌드 여부**: 테이블에 저장된 기존 데이터에 대해 인덱스를 생성할 지 여부를 선택(기본값: **미사용**)
* **인덱싱 시작 날짜**: **기존 데이터에 대한 인덱스 빌드 여부**를 **사용**으로 설정하면 나타나는 옵션입니다. 인덱싱을 시작할 기간을 입력합니다. 입력하지 않으면 전체 데이터에 대해 인덱싱합니다. 미래의 날짜를 지정하면 해당 날짜 이전에 발생한 데이터에 대해서는 인덱싱을 수행하지 않습니다.
* **토크나이저 설정**: 토크나이저의 종류([쿼리 기반 필드 추출](https://docs.logpresso.comnull), [구분자 및 IP](https://docs.logpresso.comnull))에 따라 각각 다른 속성을 입력합니다.

인덱스 생성 완료 화면에서 생성된 결과를 확인합니다.

쿼리 기반 필드 추출 토크나이저

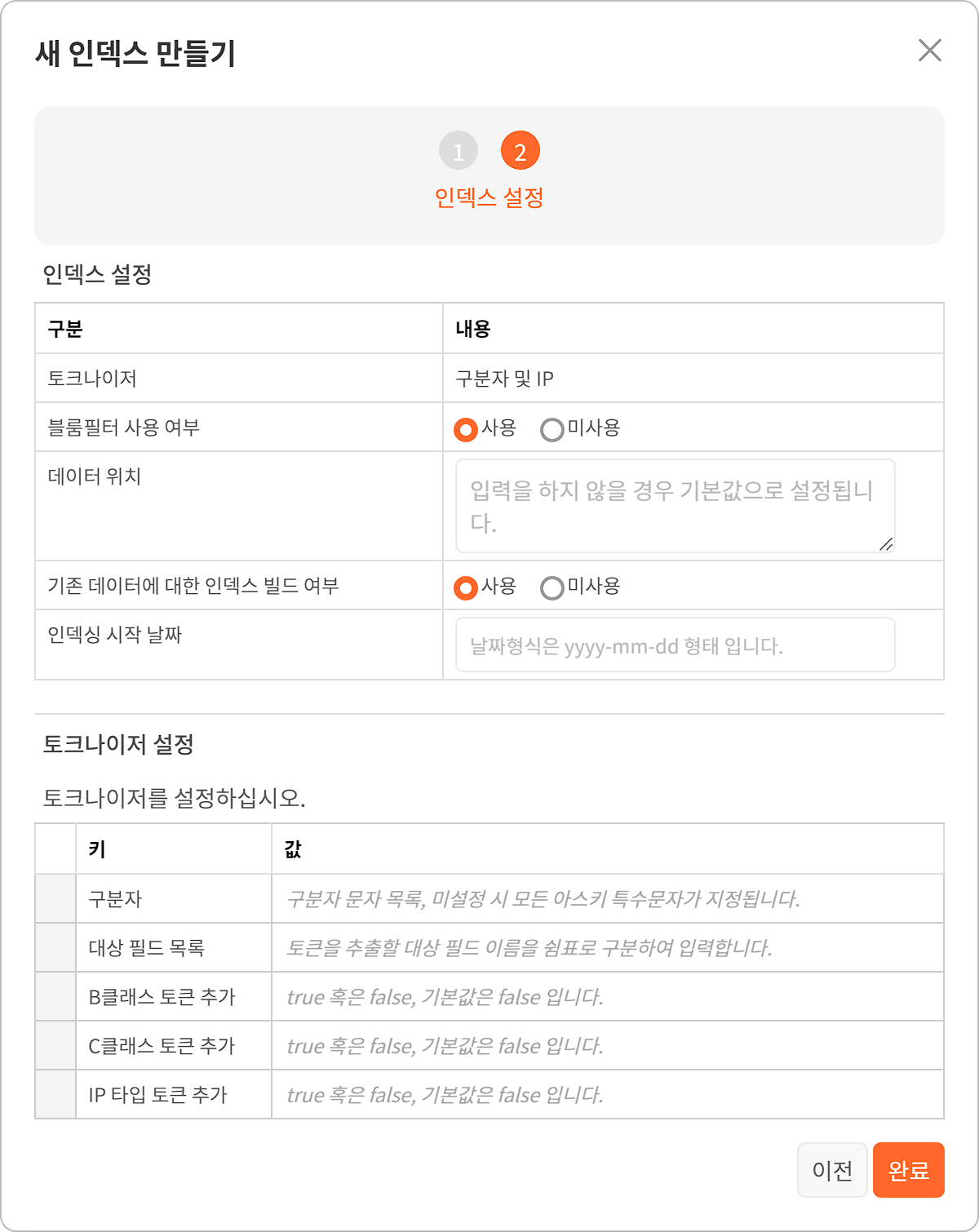
인덱스에 쿼리 기반 필드 추출 토크나이저를 설정할 때 지정해야할 속성은 다음과 같습니다.



* **인덱스 토큰 추출 쿼리**: 인덱스 토큰 추출에 사용할 쿼리를 입력. 이 쿼리는 테이블에 저장된 데이터를 입력으로 받아서 처리합니다.
* **토크나이징 제외 대상 필드**: 인덱스를 생성하지 않을 필드 목록을 쉼표(,)로 구분하여 입력
* **익명화 필드 목록**: 풀텍스트 검색을 실행할 때 필드 이름을 지정하지 않아도 검색할 필드 목록을 쉼표(,)로 구분하여 입력
* **구분자**: 인덱스 토큰을 추출할 때 구분자로 사용할 문자 목록. 지정하지 않으면 모든 ASCII 특수 문자를 구분자로 인식합니다.
* **프리픽스 구분자**: 프리픽스 구분자 문자 목록. 가령 마침표(.)를 프리픽스 구분자로 지정할 경우, 필드 값 "1.2.3.4"에서 생성되는 인덱스 토큰은 '1', '2', '3', '4', '1.', '1.2.', '1.2.3.', '1.2.3.4'입니다.

구분자 및 IP 토크나이저

인덱스에 구분자 및 IP 토크나이저를 설정할 때 지정해야할 속성은 다음과 같습니다.



* **구분자**: 인덱스 토큰을 추출할 때 구분자로 사용할 문자 목록. 지정하지 않으면 모든 ASCII 특수 문자를 구분자로 인식합니다.
* **대상 필드 목록**: 토큰을 추출할 대상 필드 목록을 쉼표(,)로 구분하여 입력
* **B클래스 토큰 추가**: B 클래스 단위 토큰의 생성 여부를 나타내는 불리언 값(기본값: false)
* **C클래스 토큰 추가**: C 클래스 단위 토큰의 생성 여부를 나타내는 불리언 값(기본값: false)
* **IP 타입 토큰 추가**: IP 주소 타입 토큰의 생성 여부를 나타내는 불리언 값(기본값: false)

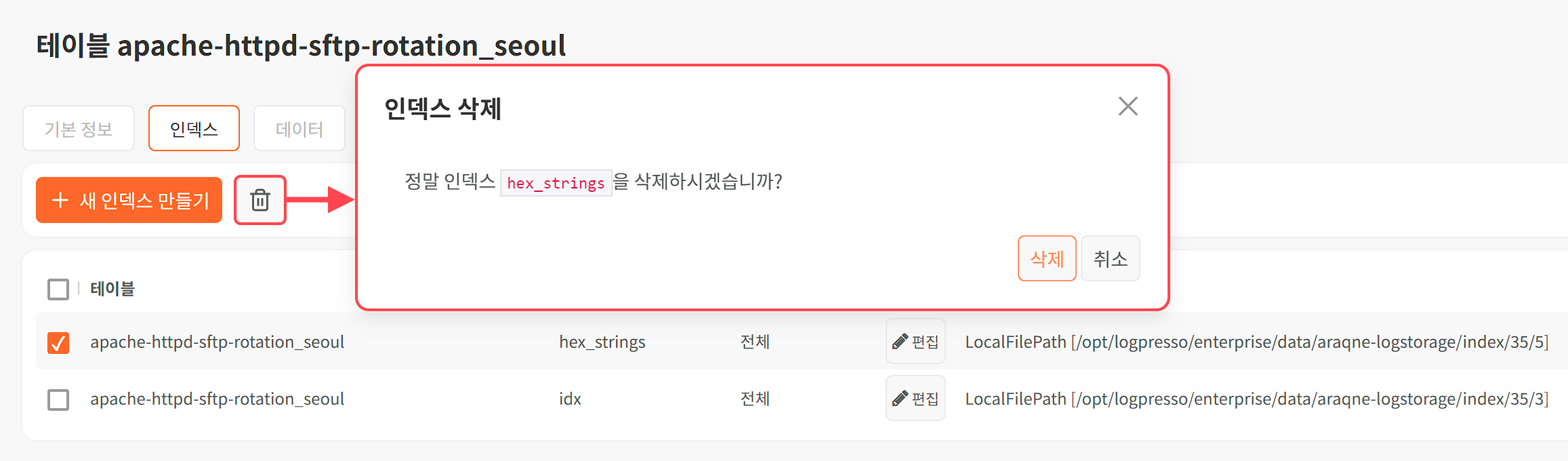
#### 인덱스 삭제

테이블에서 인덱스를 삭제하는 방법은 다음과 같습니다.

테이블 트리에서 인덱스를 삭제할 테이블을 선택합니다.

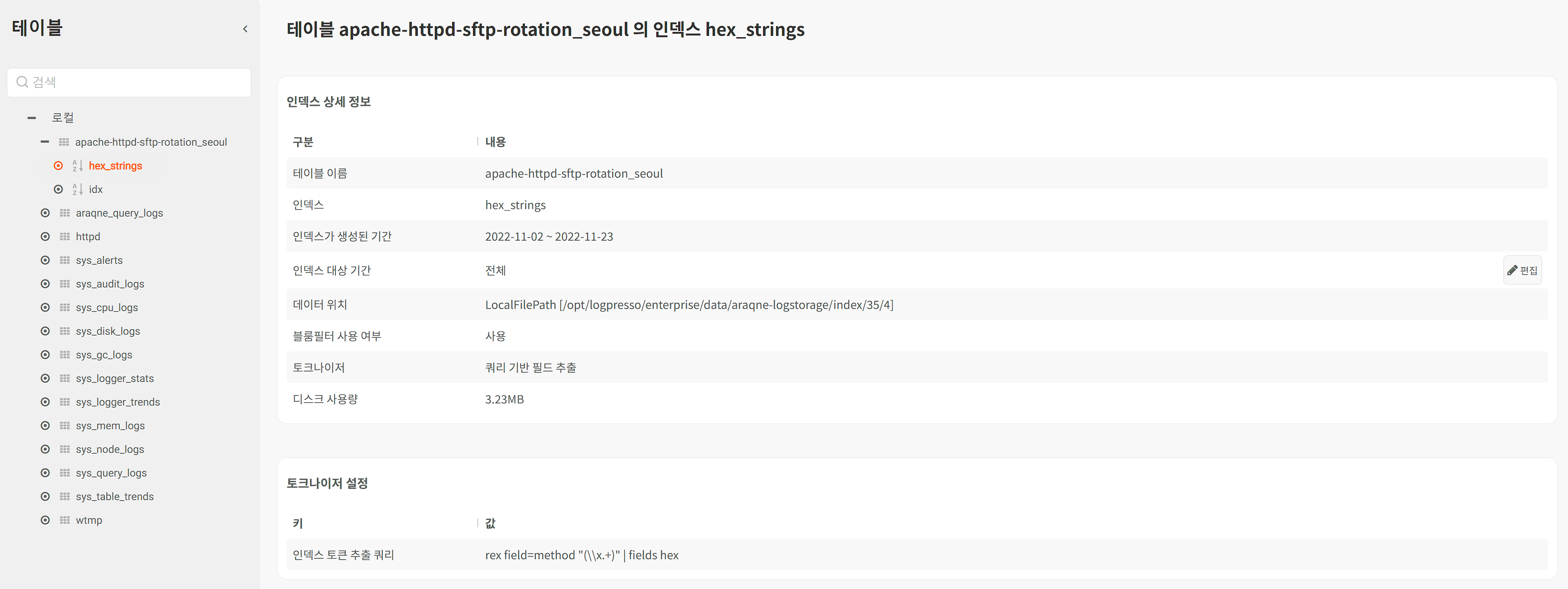
**인덱스** 탭 패널에서 도구 모음에 있는 **휴지통**을 누릅니다.

**인덱스 삭제** 창에서 삭제할 인덱스를 확인하고 **삭제**를 누릅니다.



#### 인덱스 상세 정보

**인텍스** 탭 패널에서 인덱스 항목을 누르면 아래 그림과 같이 인덱스 상세 정보 및 토크나이저 설정을 보여줍니다.

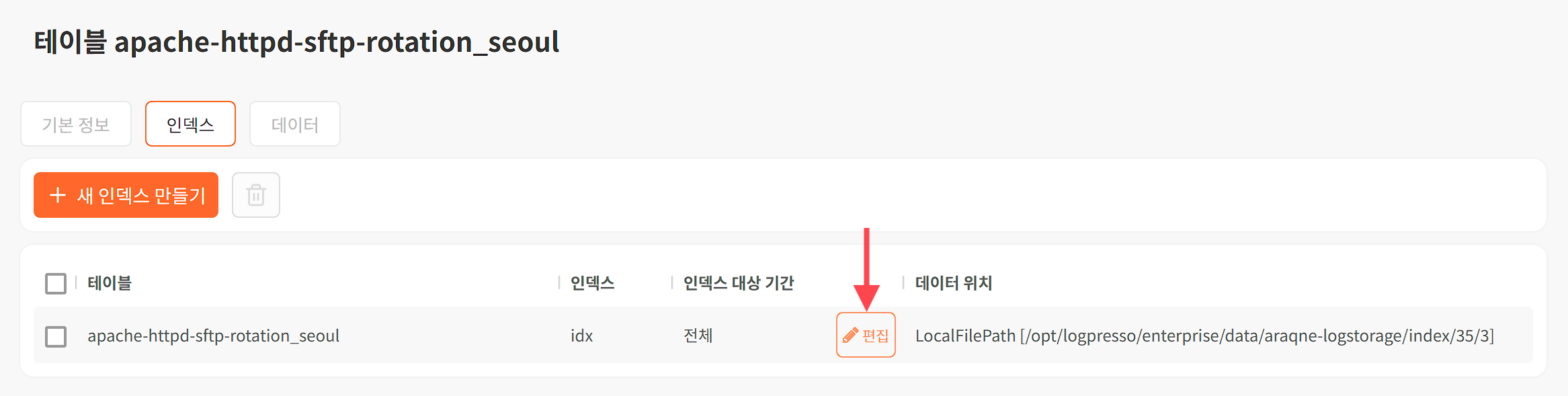


**토크나이저 설정**에 보이는 키인 **hex\_strings**는 [아파치 웹 서버](https://logpresso.store/ko/apps/apache-httpd) 앱이 사용하는 테이블에 사용자가 임의로 추가한 토크나이저입니다.

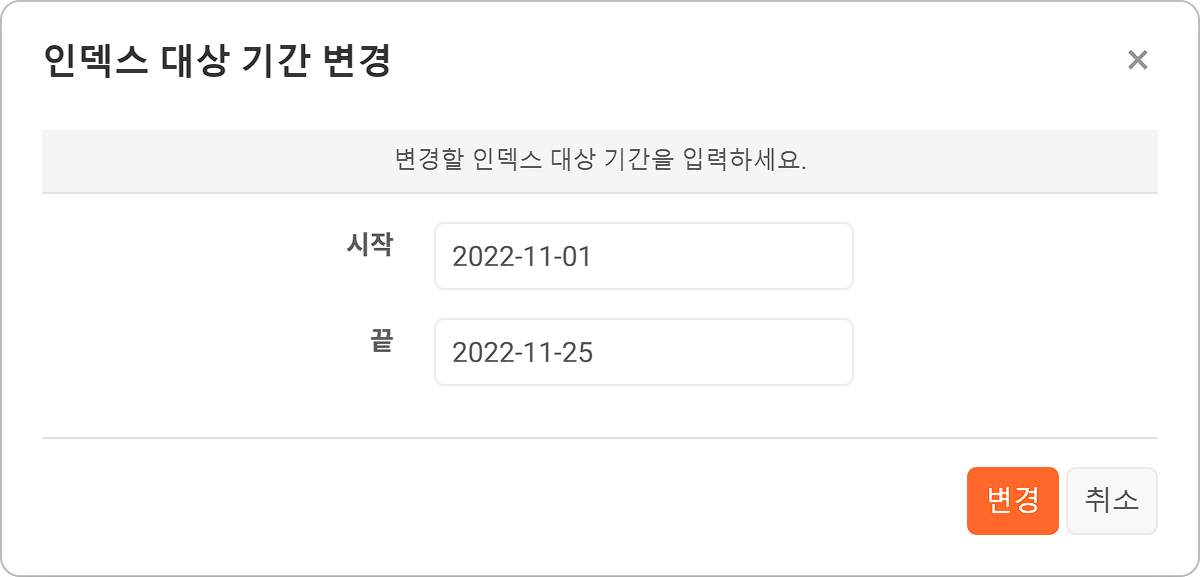
#### 인덱스 대상 기간 변경

현재 입력이 수행중인 테이블에 대해 데이터를 보존하면서 특정 인덱스가 지정한 기간에 대해서만 인덱싱을 수행하도록 지정할 수 있습니다.

인덱스 탭 패널에서 인덱스 대상 기간을 변경할 인덱스를 확인하고, 기간을 변경할 인덱스 행에서 **편집**을 누릅니다.



시작과 끝 항목에 기간을 입력하거나 월력을 사용하여 기간을 선택한 후 **변경**을 누릅니다.



기간 변경이 정상적으로 적용되었는지 확인합니다.