

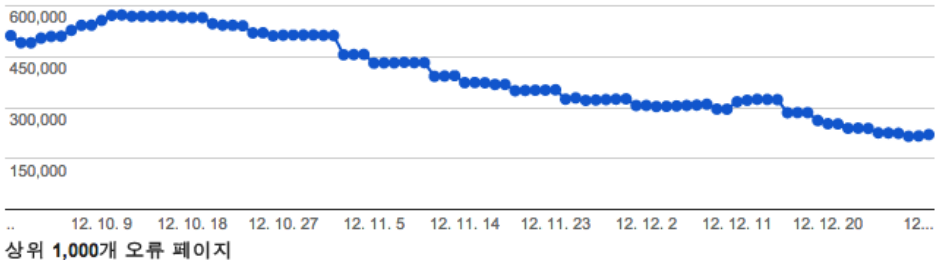
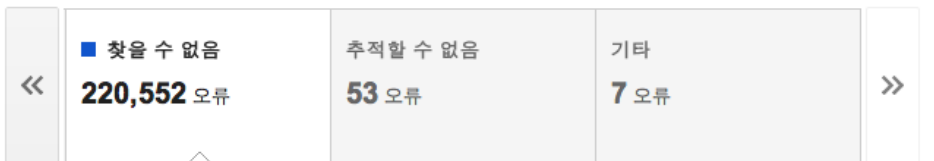
구글봇이 깨진 링크를 한 달 만에 60만 건 가져가

구글봇의 크롤 정보가 갱신되는 시간인 오후 한 시가 되면 어김없이 **구글 웹마스터 도구**에 들어갑니다. 최근 블로그를 리뉴얼 하면서 주소형식이 변경되는 일이 있었습니다. 그래서 이전의 주소로 접근했을 때 올바른 주소로 301 리디렉션하도록 했어요. 그런데 만드는 과정에서의 실수로 잘못된 주소가 대량 방출되었다는 사실을 뒤늦게 알았습니다. 한 달 만에, 무려 60만 건에 달하는 크롤 오류가 감지되었더군요.

URL 오류

웹

모바일



웹마스터 도구에서 보고한 오류 정보를 토대로 수정 작업을 완료하고

Googlebot에게 이제 다시 크롤해도 좋다고 알려줄 수 있습니다. 그런데 이는 오류 종류(500, 404, 401, 410 등)별로 하루 최대 1,000개이므로 60만 건을 모두 처리하려면 600일이 소요됩니다.(물론, 놔두면 구글봇이 알아서 처리한다고 합니다.) 그런데, 이게 은근 중독성인 게, 마치 롤플레이팅 게임에서 몸을 때려잡는 듯한 감칠맛까지 나더라고요.

구글이 개밥 주듯이 조금씩 내주는 정보에 만족할 수 없어서 사이트 전체의 깨진 링크를 검사해 주는 도구인 [Integrity](#)(맥용)를 이용해서 내부적으로 잘못된 주소가 생성되는지 확인했습니다. 공짜치고는 꽤 쓸만한 도구였어요. 전체 5만 여건의 내부링크가 검색되었는데 그중에서 잘못된 링크를 모두 찾아 고치고 수개월째 경과를 지켜보고 있습니다.

구글봇이 가져간 크롤오류가 바로잡히는 데에는 생각보다 오랜 시간이 걸렸습니다. 약 3개월에 걸쳐 조금씩 줄어들어 지금은 22만여 건으로 많이 줄어든 상태예요. 기존에 서비스되고 있는 라우트를 건드릴 때에는 정말 조심스럽게 작업에 임해야겠습니다.