

표절 방지 시스템

iThenticate

1. iThenticate 소개

iThenticate은 iParadigms사에서 1996년 출시한 학생 보고서 및 논문 중복 검사 솔루션인 Turnitin의 기술력을 바탕으로 대상을 확장하여 학회, 기업, 연구기관 및 정부기관을 위한 표절방지 시스템이다. 10여 년 전, 교육기관의 표절 확산 문제에 고심하던 연구자, 교사, 수험자 및 컴퓨터 프로그래머 그룹이 협력하여 세계 최초로 고안된 기술을 소개하였다. 그 후, 콘텐츠 검증 및 표절 방지의 효과적인 방법을 찾고자 하는 교육자 사이에서 급속히 유명해지기 시작하였으며, iThenticate은 출시 즉시 지적 재산권 보호 솔루션 및 신뢰성 있는 표절 감시 프로그램을 찾는 출판사, 연구 기관, 정부 기관, 금융 기관, 산업기관 및 법률 회사의 관심을 받고 있다. 2008년 iThenticate의 향상된 버전을 출시하였으며, 이용자들을 위하여 꾸준히 서비스를 개선하며, 정보 업계의 요구를 반영하고 있다.

2. iThenticate 주요 기능

iThenticate 시스템은 3가지 주요기능을 포함하고 있다.

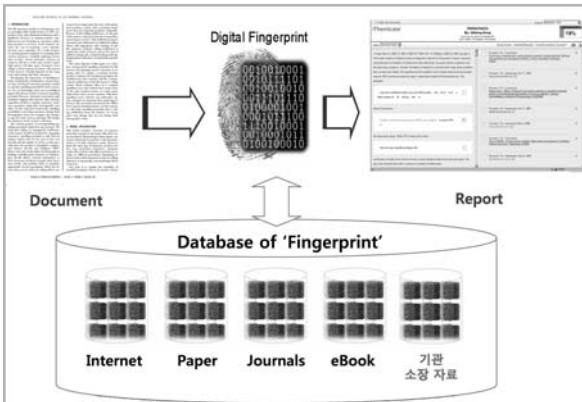
첫째, 문서 표절 및 중복 게재 방지의 기능이다. iThenticate가 보유하고 있는 전 세계 주요 출판사 및 학회 article, 주요 대학 석·박사 논문 및 인터넷 자료의 데이터베이스와 이용자가 확인하고자 하는 문서를 단어 대 단어(Word by word)로 비교하여 유사성 검사를 시행, 이를 통해 해당 문서와 유사한 자료를 5분 내외로 찾아내어 주어 문서의 표절여부를 확인할 수 있을 뿐 아니라 문서가 중복으로 게재되는 것을 미연에 방지할 수 있다.

둘째, 지적 재산권 보호의 기능이다. iThenticate은 적절한 인용 절차가 무시된 채 이용자의 아이디어나 저작물이 표절되는 것을 방지하여原作者의 지적 재산권을 보호하여 준다.

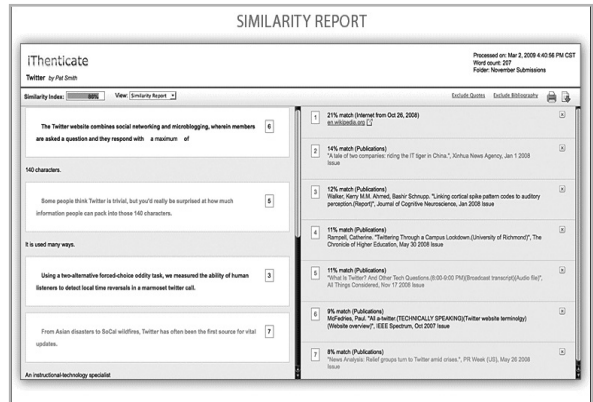
셋째, Document의 평가 및 분석의 기능이다. iThenticate 시스템을 활용해 빠르고 정확하게 문서를 비교·분석함으로써 문서 평가 및 편집 과정에서 Editor 및 Peer Reviewer의 업무 시간을 단축할 수 있다.

3. iThenticate 기반 기술 환경

iParadigms사는 이미 1994년부터 Dr. John Barrie를 주축으로 UC Berkeley의 연구자들과 함께 문서의 재사용 및 표절을 검사하는 시스템을 수년간 개발하여 온 선두기업으로써 그동안 축적한 기술력을 바탕으로 문서 내 중복된 자료를 탐지하는 시스템을 제공하고 있다. iParadigms사의 표절 탐지 핵심 기술인 “디지털 핑거프린트” 방식은 문서에 포함된 단어 각각에 지문과 같이 식별 가능한 디지털 코드를 부여하여 각 단어 단위를 기준으로 유사도를 분석한다. 이용자가 표절 검사를 위해 업로드 한 문서는 이미 iParadigms에 코드화 되어 축적되어 있는 저널, 인터넷 아카이브, eBook, 논문 등과 단어, 구, 문단 단위로 유사도를 분석하여 유사도 분석 과정을 유사성 보고서(Similarity Report)형태로 제공한다. 이 때 비교 대상이 되는 데이터베이스에는 iParadigms가 수년간 축적해 온 135억 페이지 이상의 웹 아카이브 콘텐츠 및 야후, 네이버 등의 뉴스 콘텐츠를 비롯하여 GALE OneFile, EBSCO Database, Emerald Journals등의 주요 출판사와 학생 보고서가 포함되어 있으며, CrossRef를 통해 국내외 과학기술 학회를 비롯하여 해외의 우수 출판사 AIP/APS, ACM, Wiley-Blackwell, Elsevier, IEEE, Nature, Springer등 114곳 이상의 출판사_학회의 자료가 포함되어 있다. 이러한 다양한 정보원을 통한 iParadigms사의 데이터베이스에는 90,000여 종 이상의 정기 간행물에서부터 2,500만 건 이상의 저널 논문이 포함되어 있으며, 이 중 대다수는 iParadigms에만 축적된 독점적인 자료들이다.



〈그림 1. 디지털 핑거프린트〉



〈그림 2. 유사성 검사 결과 화면〉

4. iThenticate 향후 기술 발전 추이

iThenticate은 원문의 내용을 자국어 또는 영어로 번역하여 출처 없이 사용하는 번역 관련 표절(Translated Plagiarism)을 탐지해내는 기술 역시 6개월 안에 시스템에 도입할 예정이며, 이러한 서비스를 바탕으로 지난 2004년 네이처에서 다뤄 국제적인 이슈가 되었던 “영국 캠브리지 대학의 한국인 과학자 표절” 유사한 사건을 미연에 방지할 수 있는 효과가 예상된다.