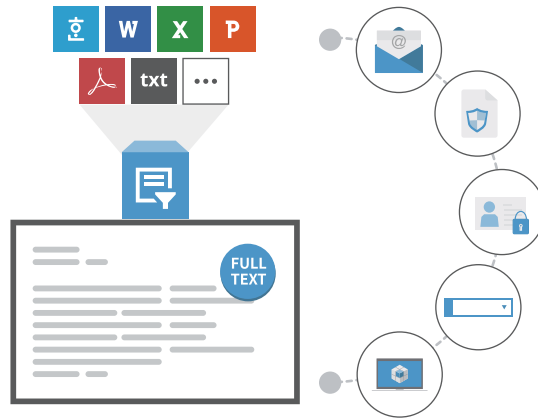


# 대한민국 No.1 텍스트 추출 솔루션

# 사이냅 문서필터

대형 포털 사이트, 빅데이터, 보안, 검색엔진, 메일 솔루션 등 3000여개 이상의 고객을 보유하고 시장점유율 95%를 자랑하는 국내 No.1 전문 텍스트 추출 솔루션



## 적용 분야



### 빅데이터/AI

빅데이터, AI학습데이터 구축 시 다양한 포맷의 대용량 문서파일을 식별, 텍스트 추출, 분석에 활용



### 개인정보 / 보안

개인정보보호, DLP, DRM 등 보안 솔루션에서 개인정보/기밀정보 검출, 감시, 유출방지를 위해 적용



### 검색

문서파일 식별 후, 색인DB 생성 또는 텍스트 미리보기 기능으로 활용



### E-mail

이메일 본문, 첨부파일 및 아카이빙 내용 검색, 스팸 차단 등을 위해 활용

## 레퍼런스

AhnLab	FAS00	DOUZONE	SOOSAN	SECU I	MarkAny®
SOMANSA	XCURENET	지란지교소프트	MOCOMSYS	(주)세토피아	IT NOMADS
VAV	KONAN TECHNOLOGY	WISEnut	diquest	I-BRICKS	인사이드정보
NAVER	kakao	SAMSUNG 삼성SDS	42MARU	이노비즈	Naravision
KISA 한국인터넷진흥원	NSR	NAON	DAOU 다투기술	Penta SECURITY cloud · iot · blockchain	Security (주)영세분서규리덕

## 특장점



400+ 포맷 식별  
50+ 포맷 추출



다양한 OS 플랫폼 지원



쉽고 편리한 연동

## 지원 포맷



MS Word



MS PowerPoint



MS Excel



한글



한소



한셀



MS Outlook



eml



mht



pst



PDF



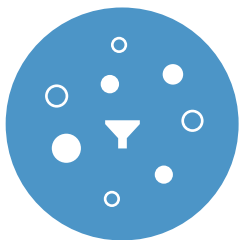
압축파일

- MS 오피스 기타포맷: potx, ppsx, ppsm, pptm, potm, pot, thmx, ppam, ppa, xlsx, xltx, xltm, xlam
- iWork: pages, key, numbers
- 핸디소프트 포맷: hwd, hwx, hwn
- ODF: odt, ott, ods, stc, odp, sti, sxw, sxc, sxi

- Text&Markup: txt, htm, html, rtf, xml
- Email: mht, eml, pst, msg
- Vector: dwg, dgn, mdi, xps, mdi
- 압축파일: zip, Tar, 7zip, RAR, Alzip, Bzip, Gzip, EGG
- Other: chm, wpd, mdb, dcm, mp3, ndoc

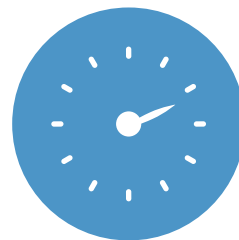
\* 그 외 다수 포맷 식별/추출 가능

## 주요 기능



### 다양한 기능

압축파일, 문자집합코드 자동탐지  
다국어 인코딩, OLE 객체 필터링,  
문서속성 정보 추출,  
문서 내 이미지 추출



### 빠른 속도와 안정성

대용량 파일 필터링 지원 (64bit)  
대형 포털에서 검증된 안정성



### 다양한 사용자환경

Windows, Linux, HP-UX,  
Oracle Solaris,  
AIX 32bit/64bit  
※MacOS M1/M2(옵션)



### 다양한 인터페이스

C/C++, Java, Python,  
Microsoft ifilter, C#

