

시스템 사양

서버환경 :

- HP/UX on Itanium : 11iv3 (11.31)
- IBM AIX R64 on POWER architecture 7.1
- IBM z/OS : V1R11 and higher
- Linux x64 (64-bit) : Novell SuSE 11 SP1; Red Hat Enterprise Linux 6.1; Oracle Linu x 6.
- Microsoft Windows on x 64 (64-bit) :
Desktop : Windows 7*x 64 SP1; Windows 8**x 64
Server : Windows Server 2008 x 64 SP2 Family; Windows Server 2008 R2 SP1 Family; Windows Server 2012 Family
- Solaris on SPARC : Version 10 Update
- Solaris on x 64 (x64-86) : Version 10 Update 9; Version 1

미드-티어

- HP/UX on Itanium
- IBM AIX on POWER
- Linux x64 (x86-64)
- Microsoft Windows x 64 (x86-64)
- Solaris (SPARC and x64)

브라우저 환경

- Internet Explorer 9 : Windows 7 (32-bit and x 64 32-bit Web browsers)
- Internet Explorer 10 : Windows 7 and Windows 8 (32-bit and x 64 32-bit Web browsers)
- Internet Explorer 11 : Windows 7 and Windows 8 (32-bit and x 64 32-bit Web browsers)
- Firefox 6 and up : Windows 7 and Windows 8 (32-bit and x64 32-bit Web browsers); Linux x64 : RHEL 6 and SLES 11 (32-bit Web browsers)
- Chrome 15 and up : Windows 7 and Windows 8 (32-bit and x64 32-bit Web browsers); Linux x64 : RHEL 6.1 and SLES 11 SP 1 (32-bit Web browsers)

* Windows 7 supported editions are : Professional, Ultimate and Enterprise.
** Windows 8 supported editions are : Standard, Pro and Enterprise.

클라이언트 환경 (SAS® Text Miner for Desktop)

- Microsoft Windows (32-bit) : Windows 7*x86-64; Windows 8**x86-64
- Microsoft Windows (64-bit) : Windows 7*x64 SP1; Windows 8**x64

*Windows 7 supported editions are: Professional, Ultimate and Enterprise.
** Windows 8 supported editions are: Standard, Pro and Enterprise.

For More Information

상세한 시스템 사양 및 텍스트 분석 관련 백서, 관련 매뉴얼 등 보다 자세한 내용은 온라인 포털에서 확인해보실 수 있습니다.
www.sas.com/korea

SAS® Text Analytics

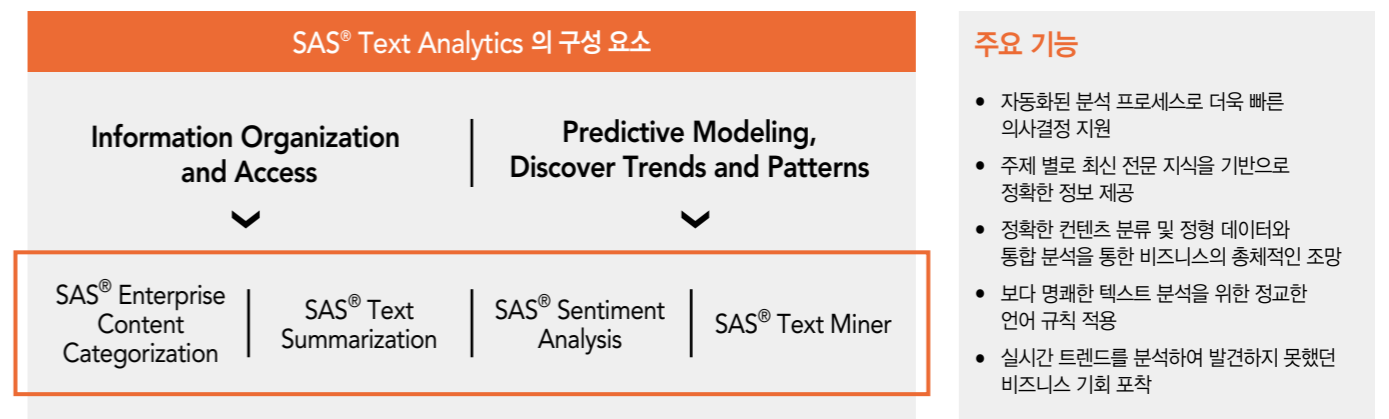
높은 완성도로 검증된 첨단 비정형 데이터 분석 솔루션

> fact sheet

SAS® Text Analytics 소개

매일 쏟아지는 비정형 텍스트 데이터를 어떻게 분석하고 활용할 것인가에 대해 많은 조직들이 고민하고 있습니다. 전체 비즈니스 데이터의 90%를 차지하고 있는 비정형 데이터로 비즈니스 가치를 극대화하기 위해서는 검증된 분석 능력이 필수적입니다.

SAS® Text Analytics는 정교한 언어 규칙 (Linguistic Rules) 과 고급 통계 분석 기술로 사람이 인지하는 것과 비슷한 수준으로 텍스트를 이해하고, 문맥 상 논리적인 모순과 모호성을 개선하여 명쾌하게 분석하는 솔루션입니다. 또한 기존의 정형 데이터와 결합하여 분석함으로써, 비즈니스에 대한 총체적인 시각을 확보할 수 있습니다. 이로써 여태껏 숨겨져 있던 텍스트 데이터의 패턴, 감성, 심리 및 비즈니스 연관성을 정확하게 이해할 수 있습니다.



SAS® Text Analytics 특징

- **다양한 언어로 된 비즈니스 프로세스의 효과적인 분석**
Teragram 기술을 기반으로 한 자연어 처리 (NLP) 기술과 고급 언어 기법으로 대량의 콘텐츠를 자동으로 분류하고 분석하여 중요한 정보를 파악해줍니다. 이렇게 도출된 정보는 체계적으로 분류되어 저장됨으로써 실시간 비즈니스 프로세스에 활용할 수 있습니다.
- **텍스트 데이터에서 발견하는 새로운 지식과 정보**
텍스트 데이터를 체계적인 포맷으로 변환하여 데이터 분류를 용이하게 해줄 뿐 아니라, 데이터 간의 연관성을 명확하게 규명하고 카테고리 별로 분류함으로써 새로운 인사이트를 쉽게 도출할 수 있도록 도와줍니다.
- **소셜 미디어의 실시간 자동 분석을 통한 양질의 마케팅 정보 수집**
자동 분석 프로세스로 소셜 미디어에서 오가는 수많은 정보를 실시간으로 찾아내어 분석하고 그 이면에 숨겨져 있는 정서나 심리를 해석할 수 있도록 도와줍니다. 마케팅, 브랜드 및 물류 관리 부서에서는 시간 소모가 많은 검색이나 수작업 없이 가치 있는 정보를 얻을 수 있습니다.
- **문서 리포지토리 (repository) 간의 연관성 파악**
따로 수집된 서로 다른 데이터를 공통 업무 용어에 따라 분류하여 텍스트 리포지토리 및 기타 비즈니스 자산에 적용할 수 있습니다. 이로써 조직은 미리 정의된 용어로 새로운 텍스트 데이터를 분류하여 자동으로 관련된 주제를 신속 정확하게 검색할 수 있습니다.
- **텍스트 기반 정보와 정형화 데이터의 통합**
정형화 데이터와 비정형화 데이터를 동시에 분석하는 통합 예측 모델링 환경에서 보다 완성도 높은 인사이트를 확보할 수 있습니다. 두 가지 유형의 데이터를 모두 활용하여 분석함으로써, 비즈니스를 전체적으로 조망하여 더욱 정확하고 현명한 의사결정을 돕습니다.

SAS® Enterprise Content Categorization

텍스트 데이터의 자동 분류로 콘텐츠에 대한 신뢰도 향상

SAS® Enterprise Content Categorization은 자연어 처리 (NLP : Natural Language Processing) 및 고급 언어 기법으로 텍스트 문서의 주요 토픽이나 문구를 파악합니다. 많은 양의 여러 언어로 된 콘텐츠를 파악된 주제에 따라 자동으로 분류하여 리포지토리에 저장할 수 있습니다.

- 설정된 언어 규칙에 따른 문서의 자동 분류로 실시간 텍스트 분석 수행
- 계층 분류 체계를 정의하여 관련 주제 별로 그룹화
- 브라우저가 지원되는 간편하고 직관적인 방식
- 정보에 대한 접근, 탐색, 공유의 효율성 강화로 간접비 절감
- 간편한 인터페이스 제공 및 CSV, XLS, TXT 등 다양한 형식의 파일로 결과 공유

SAS Enterprise Content Categorization add-on 제품 - 추가 기능을 통한 더욱 강력하고 효과적인 콘텐츠 분류

- SAS Document Duplication Detection : 중복되거나 비슷한 콘텐츠 내용을 파악하여 가장 최신 버전의 콘텐츠 확인 및 효과적인 문서 관리
- SAS Text Summarization : 방대한 콘텐츠에서 중요한 여러 문장을 추출 하여 효율적인 정보 열람 가능

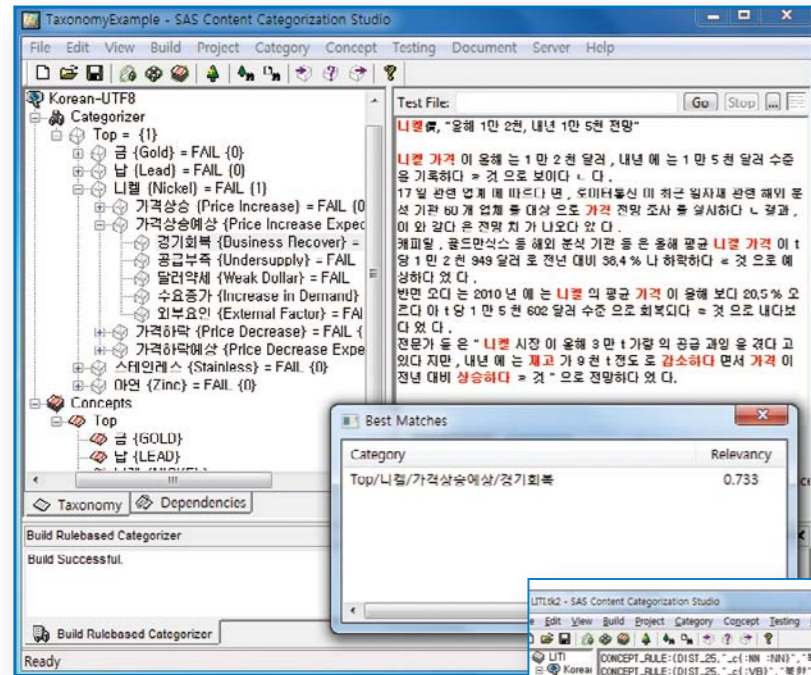
- SAS Search and Indexing : 언어 기법이나 규칙을 기반으로 하는 검색 및 인덱싱
- SAS Crawler : 오픈된 사이트의 웹 크롤러
- SAS Content Categorization Information Workbench : 대화형 정보 워크벤치
- SAS Content Alerts : 설정된 주제의 새로운 콘텐츠에 대한 실시간 알림
- SAS Industry Taxonomy Rules : 산업 별 분류 규칙 제공

SAS® Ontology Management

단어의 의미에 따른 콘텐츠의 체계적 관리

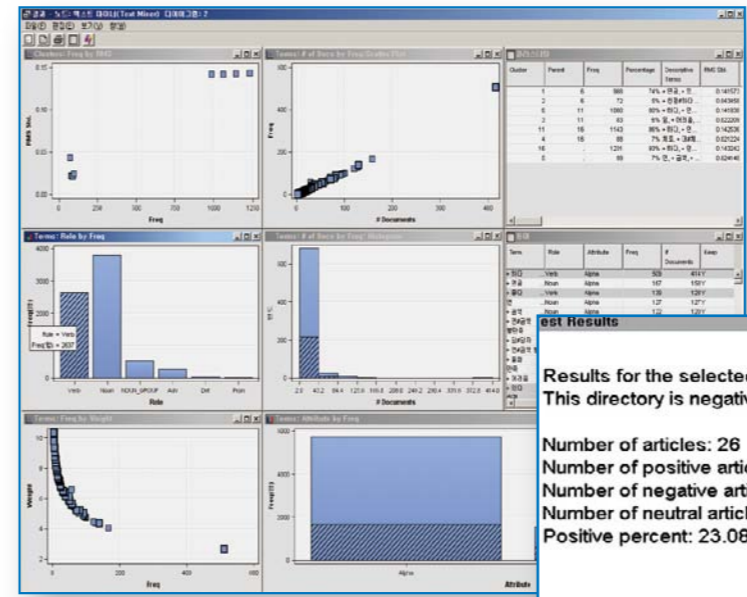
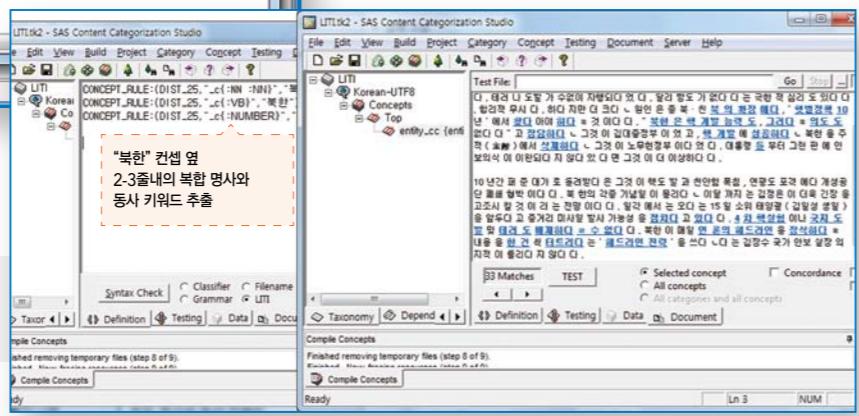
SAS® Ontology Management는 조직 내 시스템이나 프로젝트, 리포지토리 등 다양한 콘텐츠를 유사한 키워드에 따라 정리하고 처리합니다. 하나의 체계로 관리하는 온톨로지 (Ontology) 생성으로 보다 정교하고 쉽게 콘텐츠를 검토하여 콘텐츠 가치를 극대화 합니다.

- 텍스트 전반에 걸쳐 일관성 있는 용어 메타데이터 체계를 생성 · 유지
- 체계적인 공통의 개념 파악으로 신속하고 정확한 정보 검색
- 서로 단절 되어 있던 텍스트 리포지토리 간의 연관성 파악
- 콘텐츠의 품질 관리 및 변화 추적



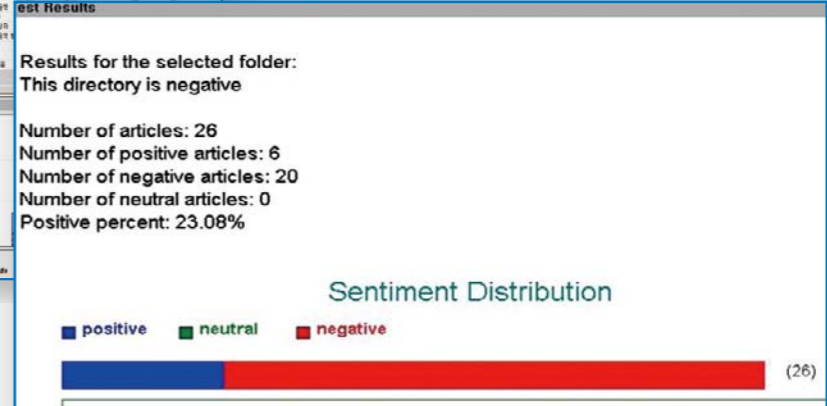
Enterprise Content Categorization으로 텍스트 의미 기반의 정확한 분류

연관 Keyword 추출



텍스트 마이닝을 통한 인사이트 도출

텍스트 상에서 드러나는 핵심 정서 파악



SAS® Sentiment Analysis

웹 상에서 오가는 핵심 정서의 정확한 파악

SAS® Sentiment Analysis는 웹사이트, 내부 콜 센터 데이터, 이메일 등에서 오고 가는 다양한 의견이나 리뷰를 실시간으로 분석하여 핵심 정서를 알아냅니다. 이로써 조직은 더욱 효과적인 커뮤니케이션을 지원하고, 보다 앞서서 최신 트렌드나 이슈를 파악할 수 있습니다.

- 고급 통계 모델과 언어 규칙 적용으로 디지털 콘텐츠의 실시간 감성 분석
- 지속적인 변화 모니터링 및 구체적인 피드백에 대한 빌드-인 리포팅
- 신규 타겟군 확보를 위한 피드백 소스 파악
- 고객 경험치 극대화로 경쟁력 있는 제품 전략 개발 및 유지
- 다양한 언어에 대한 지원 : 30여 개 이상의 언어

SAS® 를 선택해야 하는 이유

- 검증된 30개 언어 텍스트 분석 엔진
비즈니스에 있어 국가간의 장벽이 없어지고 경쟁이 심화되면서 다국어 콘텐츠의 관리 및 공유가 중요한 과제가 되었습니다. SAS Text Analytics는 조직이 콘텐츠 자산을 활용을 극대화하고, 다양한 지역에서 생성된 여러 콘텐츠의 리포지토리 또한 활용할 수 있도록 해줍니다.
- 정형 및 비정형 데이터의 통합 마이닝
방대한 양의 문서를 쉽게 요약하여 주요 내용을 파악할 뿐 아니라 기존의 정형 데이터와의 통합을 통해 비즈니스 가치를 이끌어낼 수 있도록 돕습니다. 모든 데이터의 확보에서부터 공유에 이르는 통합 데이터 라이프 사이클 전반에 걸쳐 효과적인 관리가 가능해집니다. 도와줍니다.
- 자동으로 주요 트렌드 및 패턴 발견
수십 년 역사의 자연어 처리 기술, 고급 언어 규칙, 통계 기법을 적용하여 자동으로 텍스트 데이터를 유용한 인사이트로 변환하고, 중요한 연관성이나 관계를 규명해줍니다.
- 실시간 분석이 가능한 엔진 성능
매일 새롭게 생겨나는 소셜 미디어 콘텐츠 및 기업내 상담 및 VOC 텍스트에 대하여 형태소 분석 및 다양한 자동 분류 및 문맥 추출 알고리즘을 바탕으로 복잡한 처리 과정 없이 빠르게 분석 정보를 얻을 수 있습니다. SAS® Text Analytics는 국내 여러 사례에서 하루 수만, 수십만의 문서를 분석 처리하는 경험을 가지고 있습니다.

SAS® Text Miner

비정형 데이터로부터 인사이트를 도출하는 텍스트 마이닝 솔루션

SAS® Text Miner는 텍스트 데이터 분석 모델링 툴로 구성된 고급 분석 솔루션으로, 각종 텍스트 문서로부터 유용한 정보를 검색하여 상황을 예측할 수 있게 해줍니다. 텍스트 문서 분석으로 지금까지 발견할 수 없었던 새로운 아이디어나 패턴 파악으로 보다 현명한 의사 결정을 지원합니다.

- 텍스트 프로파일링으로 다양한 문서의 새로운 트렌드 발견
- 언어 표현 논리연산 (Boolean Rule)의 자동 생성으로 간편하게 콘텐츠 분류
- 고성능 텍스트 마이닝 기법으로 대용량 문서의 초고속 분석
- 도출된 인사이트를 예측 모델에 직접 적용 및 결과에 대한 시각화
- 정형화 데이터 소스와의 통합으로 총체적인 시각 확보
- 대화형 인터페이스로 보다 상세한 내용에 대한 드릴 다운 (Drill down)