



문서번호	기획예산과-3042
결재일자	2016.3.11.
공개여부	대시민공개
방침번호	

주무관	구정기획전문관	기획예산과장	기획재정국장	
윤수	안지훈	이윤영	03/11 박기웅	
협조	전산정보과장 이창호 정보화기획팀장 박윤영			

기고문, 인사말씀 등의 효율적 통계 분석을 위한
텍스트 분석기 프로그램(워드 클라우드) **개발 계획**

2016. 3.

사 전 검 토 사 항

::: 해당사항에 표시하시기 바랍니다.

항 목	검 토 여 부
사 업 구 분	신규사업 <input checked="" type="checkbox"/> 공약(약속)사업 <input type="checkbox"/> 계속사업 <input type="checkbox"/> 인센티브/공모사업 <input type="checkbox"/>
소 통 분 야 교 려 사 항	<ul style="list-style-type: none"> ● 구 민 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/> ● 전 문 가 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/> ● 이 해 당 사 자 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/>
기 타 교 려 사 항	일 자 리 <input type="checkbox"/> 환경영향 <input type="checkbox"/> 안 전 <input type="checkbox"/> 유지비용 <input type="checkbox"/> 바른 공공언어 <input checked="" type="checkbox"/> 성 인 지 <input type="checkbox"/> 취약계층 <input type="checkbox"/> 장 애 인 <input type="checkbox"/> 디 자 인 <input type="checkbox"/> 갈등발생 요인 <input type="checkbox"/>
타자원 활용	<ul style="list-style-type: none"> ● 중 앙 부 처 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/> ● 서 울 시 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/> ● 기 업 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/> ● 민 간 단 체 : 유 <input type="checkbox"/> () 무 <input checked="" type="checkbox"/>
언 론 홍 보 계 획	기획보도 <input type="checkbox"/> 보도자료 <input type="checkbox"/> SDTV <input type="checkbox"/> 성동뉴스레터 <input type="checkbox"/> 성동구소식지 <input type="checkbox"/> 기 고 문 <input type="checkbox"/> 전자행정서비스 <input type="checkbox"/> S N S <input type="checkbox"/> 기타(리플릿 등) <input type="checkbox"/> 없 음 <input checked="" type="checkbox"/>
<ul style="list-style-type: none"> ● 홍 보 제 목 : ● 중점 홍보사항 - 	

기고문, 인사말씀 등의 효율적 통계 분석을 위한

텍스트 분석기 프로그램(워드 클라우드) 개발 계획

텍스트 분석기 프로그램을 자체 개발하여 기고문, 인사말씀 등의 키워드를 통계분석하고 시각화하여 구정 홍보 방향의 일관성 유지 및 비전 수립 등에 활용하고자 함

I 추진 방향

- 문장에서 단어를 추출하여 키워드를 분석(사용빈도수 측정)하고 시각화하여 자료를 축적
- 전산정보과 프로그램 개발 전문가를 활용하여 자체개발(비에산)
- 기고문, 연설문 작성 시 참고자료 활용

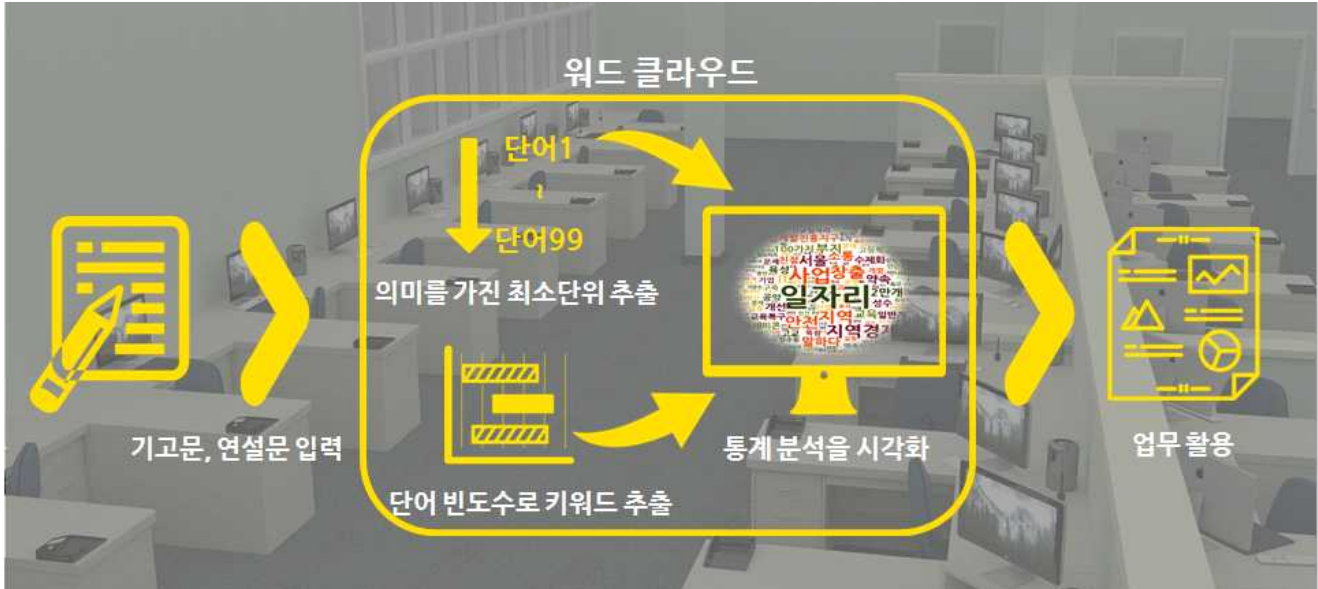
II 추진 개요

- 추진기간: 2016년 3월 내
- 이용대상: 기고문, 연설문 등 모니터링 업무
- 사업비: 비예산(자체개발: 전산정보과 프로그램 개발 전문가 활용)
- 사업내용: 텍스트 통계 및 시각화 프로그램 개발
- 기대효과
 - 키워드 분석을 통해 기고문, 연설문 작성 시 핵심 가치 반영
 - 키워드를 시각화하여 홍보자료의 활용성 제고(인포그래픽¹⁾)
 - 보도 자료를 분석하고 모니터링하여 데이터 관리의 효율성 제고

1) 인포그래픽: 인포메이션(information)과 그래픽(graphic)의 합성어로 많은 정보를 한눈에 파악할 수 있도록 지도, 다이어그램, 이미지, 차트 등을 활용하여 만든 자료를 의미

III

시스템 주요기능



① 텍스트 입력

- 분석 대상 자료(기고문, 연설문 등)를 텍스트(문자열)형태로 입력

② 텍스트 분석

- 단어 추출: 의미를 가진 최소 단위(형태소)로 분리
- 키워드 추출: 형태소별 사용 빈도 수를 산출하여 키워드 추출

③ 데이터 시각화

- 사용된 빈도 수에 따라 변환률을 이용하여 인포그래픽이나 히스토그램²⁾으로 표현

IV 추진 일정

- 프로그램 기획 및 제작: 2016년 3월

V 행정 사항

- 워드 클라우드 프로그램 자체 개발: 전산정보과
- 프로그램 운영 및 활성화: 기획예산과(구정기획단)
 - 운영결과 분석 후 시스템 고도화 추진. 끝.

2) 히스토그램: 도수 분포를 그림으로 나타낸 것