

2018 일반 연구 중간 점검

사회환경연구부 진대용

1. Open Data Map

Open Data Map

◆ Open Data Map : 환경연구 활용 빈도가 높은 온라인 데이터 hub

- 카테고리 별 목록 / 링크 / 내용소개를 제공
- KEI 보고서에서 자주 인용된 데이터 source 우선 파악

◆ KEI 도서관 DB로 부터 KEI에서 발간된 문서정보 추출

- 문서 경로, 문서 제목, 키워드, 카테고리 등

◆ 문서 수집 및 변환

- PDF → TXT
- 총 1979개 문서 중 1949개 문서 활용
 - 다운로드가 안되거나 txt 문서로 전환이 안되는 30개 문서 제외

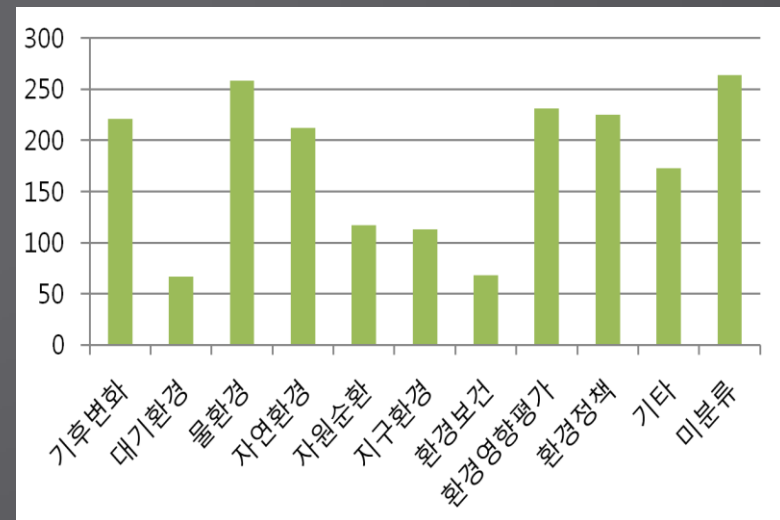
◆ 문서 카테고리 재분류

- 총 22가지 카테고리 → 11개 카테고리

```

SELECT R.*, D.SERVICE_VNR
FROM DL10B3_IR_REPORT_VIEW R JOIN DL10B3_IR_FILE_VIEW D
ON R.C_ID = D.C_ID;
    
```

CLASS_NAME	RECORD_STATE	CREATE_DATE	MODIFY_DATE	DC_IDENTIFIER	DC_TYPE	KEL_TITLEKOREA
1 6 수시연구	u	19960206	20120116	환1185 1995 WD-04	수시연구	우리나라 환경관련 예산정책의 개선방안
2 4 기본연구	u	19960109	20120113	환1185 1995 RE-13	기본연구	산업별 공업용수의 수요-수량-수질현황과 대책
3 4 기본연구	u	19960109	20120113	환1185 1995 RE-15	기본연구	종합환경보안 개발사업 (II)
4 2 기타보고서	u	19970502	20120116	환1185 1997-1	기타보고서	「지방의제2」 모델 개발연구
5 9 기술현황보고서	u	19950127	20111114	환1185 1993 AB-03	기술현황보고서	오염지표성질을 이용한 연안환경관리
6 9 기술현황보고서	u	19950127	20111114	환1185 1993 AB-01	기술현황보고서	직업교육법
7 9 기술현황보고서	u	19950127	20111114	환1185 1993 AB-02	기술현황보고서	다미육선에 관한 검토 및 분석기술
8 9 기술현황보고서	u	19950127	20170329	환1185 1993 AB-04	기술현황보고서	환경개선 부담금제도 개선방안
9 4 기본연구	u	19950127	20170703	환1185 1993 RE-01	기본연구	환경기술 연구개발 관리체계 구축방안 (I)
10 4 기본연구	u	19950127	20170703	환1185 1993 RE-02	기본연구	환경기술 연구개발 관리체계 구축방안 (II)
11 4 기본연구	u	19950127	20170703	환1185 1993 RE-03	기본연구	환경기술 연구개발 관리체계 구축방안 (III)
12 4 기본연구	u	19950127	20120113	환1185 1993 RE-04	기본연구	한국형 선진화공산업의 육성책 개발을 위한 기초조사 (I)
13 4 기본연구	u	19950127	20120113	환1185 1993 RE-05	기본연구	한국형 선진화공산업의 육성책 개발을 위한 기초조사 (II)



환경연구 활용 빈도가 높은 온라인 데이터 Source 파악

- ◆ KEI 발간된 문서 DB의 문서들로부터 데이터 리스트 추출
 - 전처리
 - 정규표현식 활용 데이터 리스트 추출
 - 추가규칙 적용
 - 문서 카테고리별 카운트

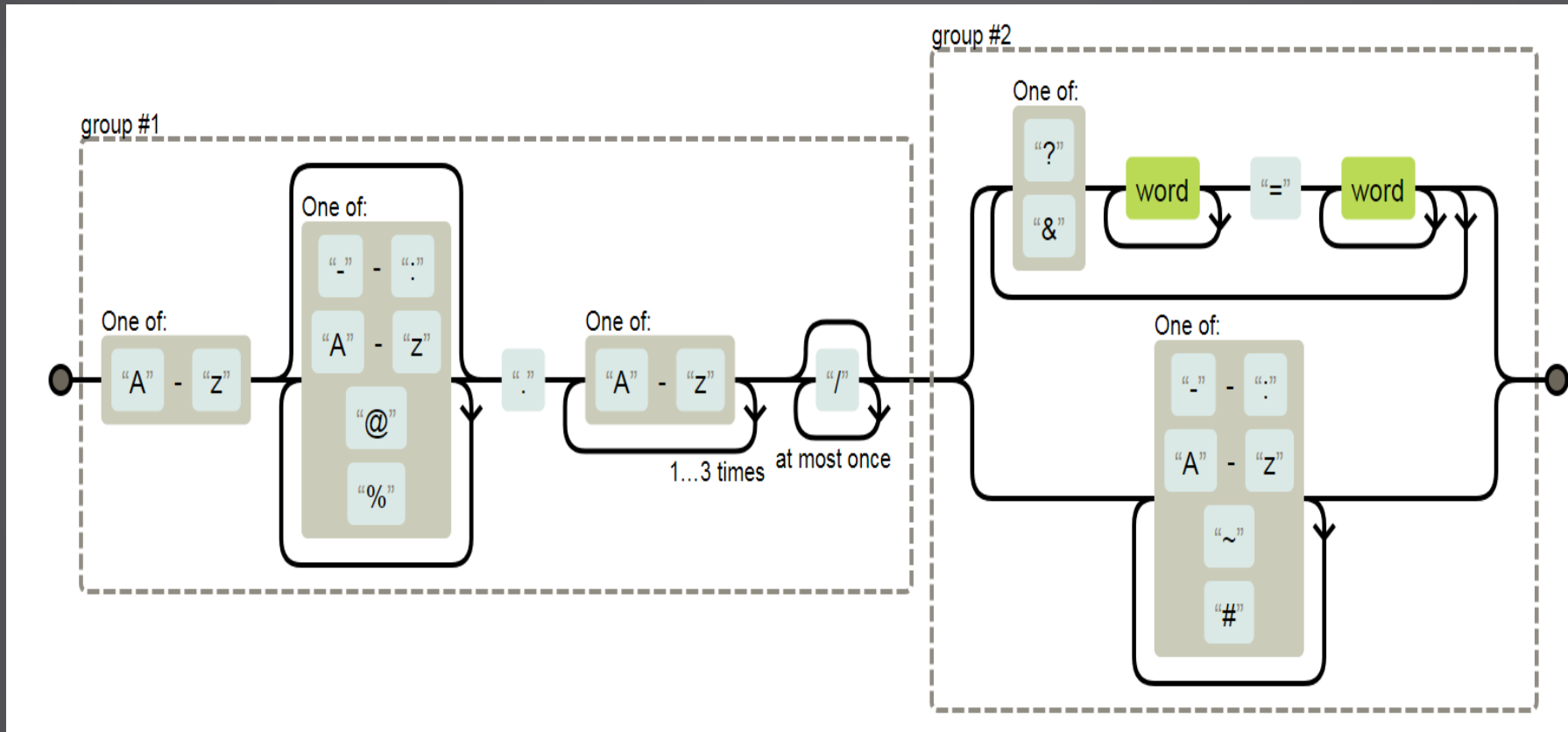
1단계: 정규표현식

- 정규표현식 핵심규칙
 - 영어로 시작
 - 콤마(.) 이후 1글자 - 3글자의 영어로 구성된 부분
예) .kr .com 등
 - 그 이후에 내용이 있을경우 반드시 '/'

2단계: 추가 규칙

- 추가 규칙
 1. 프로토콜이 없는 경우 -> http:// 추가
 2. https://, http:// -> https://로 통일
 3. .kr, .com, .gov 등 국가기관 코드 없는 경우 제외

STEP 1 : 정규표현식



파악된 온라인 데이터 Source 범주화 : 인용된 문헌의 범주

각 카테고리 별 데이터가 사용된 문서의 갯수

순위	데이터	환경정책	환경보건	물환경	대기환경	자원순환	환경영향평가	기후변화	자연환경	지구환경	미분류	기타	합	분류	연구자 검증
1	https://www.me.go.kr	17	3	20	8	11	15	19	23	5	10	18	149	종합	종합
2	http://www.law.go.kr	1	3	8	0	5	10	11	17	4	8	14	81	종합	종합
3	https://www.epa.gov	5	4	15	4	8	4	4	11	0	6	9	70	종합	종합
4	https://kosis.kr	0	5	10	1	5	1	18	2	1	2	20	65	종합	종합
5	http://water.nier.go.kr	1	0	25	0	2	8	6	4	1	4	4	55	물환경	물환경
6	http://www.kma.go.kr	2	5	9	2	2	2	21	4	4	1	1	53	기후변화	종합
7	https://www.wamis.go.kr	4	1	20	0	0	2	11	11	1	1	2	53	물환경	물환경
8	https://www.eiass.go.kr	0	1	1	0	3	20	6	6	2	10	2	51	환경영향평가	환경영향평가
9	http://www.kosis.kr	1	4	4	0	1	3	13	3	1	1	11	42	기후변화, 기타	종합
10	https://www.index.go.kr	0	0	5	0	3	2	14	7	2	0	7	40	기후변화	종합
11	https://www.airkorea.or.kr	2	8	2	1	2	1	9	3	3	3	4	38	기후변화, 환경보건	기후변화, 대기환경
12	https://egis.me.go.kr	0	0	2	0	1	6	8	13	2	1	4	37	자연환경, 기후변화	종합
13	http://www.env.go.jp	3	2	1	4	3	4	2	2	4	6	4	35	종합	종합
14	http://www.moleg.go.kr	3	0	4	1	3	6	3	8	0	3	3	34	종합	종합
15	http://kosis.nso.go.kr	10	0	2	3	1	2	1	6	2	5	0	32	환경정책	종합
16	https://ecos.bok.or.kr	3	1	1	0	4	0	7	0	1	1	14	32	기타	기타
17	http://www.nier.go.kr	2	0	3	1	1	2	9	5	0	1	7	31	종합	종합
18	http://www.nso.go.kr	17	0	0	0	2	0	2	2	2	2	0	27	환경정책	종합
19	https://www.gims.go.kr	1	1	10	0	0	3	4	7	1	0	0	27	물환경, 자연환경	물환경
20	http://airemiss.nier.go.kr	2	3	0	0	0	1	8	0	0	2	8	24	기후변화, 기타	기후변화, 대기환경

향후 계획

- ◆ 오픈 데이터맵 데이터 리스트 보완
 - 카테고리 미분류 문서 → 카테고리별 분류
 - 카테고리 별 단순 빈도수 정렬 → 전체 빈도수 및 비율 등을 고려한 랭킹 알고리즘 적용
 - 원내 설문조사를 활용한 데이터 리스트 보완
- ◆ 오픈 데이터맵 원내 서비스화
 - 오픈 데이터맵 웹 서비스화
 - 오픈 데이터맵 서비스 원내 테스트 및 공개

2. 연구동향 파악 서비스

LDA 토픽 모델링 및 연구동향 분석 서비스

- ◆ 2017년 '텍스트마이닝을 이용한 KEI 연구동향 분석' 연구결과 재생
 - 전 연구의 소스코드와 동일 데이터를 활용하여 동일 결과를 플랫폼 위에서 출력
 - 인코딩 문제, 각종 라이브러리 세팅 등 작업 환경 차이로 인한 문제 해결
 - 기존 연구 일반화 및 서비스 구축을 위해 수정이 필요한 부분을 분석
- ◆ 동일 포맷 다른 데이터 입력 시 토픽 모델링 결과를 출력할 수 있도록 기존 소스코드 일반화
 - 형태소 분석기 추가, 파라미터 지정 부분 분리, 시각화 관련 부분 수정 등
 - 코드 간소화 : 파일과 인자를 입력 받으면 한번의 실행으로 모든 결과가 도출되게 하는 작업
- ◆ 웹서비스 구축: R shiny를 활용한 웹 기반 GUI 인터페이스 구축
 - 원내 서비스를 위한 테스트 단계

연구동향 분석 서비스 GUI

입력 파일 포맷 (csv)

id	Content	year	month	day
a1	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 목적 및 방법 제2장 사업평가 운	2016	6	30
a2	제1장 서론 1. 배경 및 목적 2. 연구추진방향 제2장 개념 및 선행사례 1. 개념 2. 선행	2016	6	24
a3	1. 연구개발과제의 개요 1-1. 연구개발 목적 1-2. 연구개발의 필요성 1-3. 연구개발	2016	5	31
a4	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 목적 3. 연구 내용 및 범위 제2장	2016	4	30
a5	제1장 서론 1. 연구의 목적 2. 연구의 배경 및 필요성 3. 연구 목표 및 범위 제2장	2016	4	30
a6	제1장 서론 1. 연구 배경 및 목적 2. 연구 내용 및 방법 제2장 2015년 중남 서북부	2016	3	31
a7	1. 연구개발과제의 개요 1-1. 연구개발 목적 1-2. 연구개발의 필요성 1-3. 연구개발	2016	3	31
a8	연구개발과제의 개요 1-1. 연구개발 목적 1-2. 연구개발의 필요성 1-3. 연구개발	2016	2	28
a9	제1장 서론 1. 연구의 배경 및 필요성 2. 연구방향 3. 연구수행 방법 제2장 제네바	2015	12	31
a10	제1장 서론 1. 연구의 배경 및 필요성 가. 사물인터넷; 혁명적 핵심기술로서 부상	2016	10	31
a11	제1장 서론 1. 연구의 필요성 및 목적 가. 연구의 배경 및 필요성 나. 연구의 목적	2016	10	31
a12	제1장 서론 1. 연구의 배경 및 목적 2. 연구의 대상 3. 연구의 구조 제2장 이론적	2016	10	31
a13	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 범위 제2장 국내 지역사회 환경	2016	10	31
a14	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 목적 제2장 파리협정의 핵심사항	2016	9	30
a15	제1장 서론 1. 연구의 필요성 및 목적 2. 연구의 범위 및 방법 3. 선행연구와의 차	2016	9	30
a16	제1장 서론 1. 연구의 필요성 및 목적 2. 연구 범위 3. 연구 내용 및 수행체계 제2	2016	9	22
a17	제1장 서론 1. 연구의 배경 및 목적 2. 연구의 접근 및 방법 제2장 국내 "정화콘란	2016	8	31
a18	제 1 장 서론 1. 과업 배경 및 필요성 2. 과업 내용 가. NBSAP 실행 관련 생물다양	2016	8	30
a19	지속가능발전 개념은 '87년 환경과 개발에 관한 세계위원회(WCED)가 발표한 보	2016	7	30
a20	제1장 서론 1. 배경 2. 필요성 및 목적 3. 연구 범위 4. 연구 내용 5. 연구 방법 및	2016	11	22
a21	제1장 서론 1. 연구의 필요성 및 목적 2. 시스템과 네트워크 언어 제2장 미래 환경	2016	11	6

LDA 토픽 모델링 웹서비스

LDA Analysis

Choose CSV File

Browse... kei_original.csv Upload complete

of topics

5

SEED

2007

Display

Head

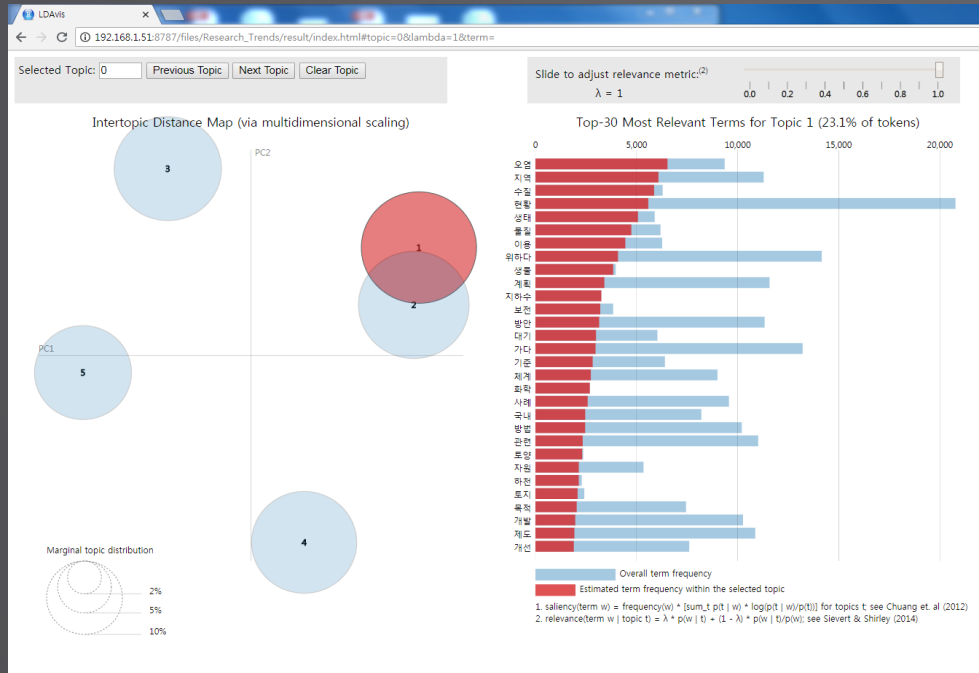
All

Do analysis

id	Content	year	month	day
a1	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 목적 및 방법 제2장 사업평가 운영현황 분석 1. KEITI의 환경개선 마스터플랜 수립 지원사업 평가 현황 가. KEITI의 ODA 사업 현황 나. KEITI의 사업평가 2. 국내 ODA 사후평가 운영 현황 가. 현황 분석 및 목표 설정 나. 사전평가의 연계성 다. 평가기준 및 범위 라. 정보공개 3. 기타 평가 운영현황 및 개선과제 가. 평가대상의 통합과 평가범위 확대 필요 나. 평가 대상사업 선정기준 미비 다. 자제평가 역할 강화 필요 4. 소결 21 제3장 국내외 기관 사업평가 지침 및 매뉴얼 분석 1. 국내 기관별 사업평가 지침 비교 2. 국내외 기관 사업평가 매뉴얼 가. OECD DAC 나. JICA 다. EDCF 라. KOICA 제4장 사업평가지침(안) 및 사업평가해설서 1. 지침 및 해설서 작성 방향 2. 환경분야 국제개발협력 사업평가지침(안) 3. 사업평가 담당자를 위한 해설 참고문헌 Abstract	2016	6	30
a2	제1장 서론 1. 배경 및 목적 2. 연구추진방향 제2장 개념 및 선행사례 1. 개념 2. 선행사례 가. 국내 탄소감축 정책 사례 나. 해외 탄소감축 정책 사례 다. 기후변화 적응 계획 사례 제3장 추진 여건 분석 1. 온실가스 배출 현황(2013) 2. 제주도의 사회경제적 환경 가. 인구증가 나. 전력사용량 및 설비 다. 자동차 운행 대수 라. 관광객 수 마. 폐기물 바. 인프라 구축(스마트그리드) 3. 온실가스 배출전망 4. 온실가스 감축에 유리한 환경여건 가. 개요 나. 신재생에너지발전 보급현황 및 계획 다. 태양광 자원 라. 풍력 자원 5. 기후변화 취약성 평가 가. 건강 분야 나. 산림 분야 다. 물관리 분야 라. 농업 분야 마. 해양/수산 분야 바. 재해 분야 제4장 비전 및 추진전략 1. 비전 및 지표 가. 탄소제로성 추진 비전 및 목표 나. 탄소제로성 추진 핵심사업 다. 지표 2. 추진전략 가. 재생에너지로 움직이는 청정에너지 자립성 나. 세계 전기자 산업의 매기 조성 다. 자연친화형 탄소제로의 글로벌 명품 관광 브랜드로 발전 라. 주민이 하나 되어 전 과정에서 저탄소 생활 실천 마. 기후변화 적응으로 안전한 제주 제5장 정책과제 도출 참고문헌 부록 목록 I. 추진전략별 주요 사업목록 부록 II. 온실가스 부문 정책 방향과 추진과제 부록 III. 제주 탄소제로성 조성사업의 경제적 파급효과 Abstract	2016	6	24
a3	1. 연구개발과제의 개요 1-1. 연구개발 목적 1-2. 연구개발의 필요성 1-3. 연구개발 범위 2. 국내외 기술개발 현황 2-1. 사회경제적 시나리오 개발 현황 2-2. 사	2016	5	31

연구동향 분석 서비스 Output

LDA 토픽 모델링 결과



연구동향분석 시각화

