

# 2018 일반 연구 최종 점검

사회환경연구부 진대용

# 1. Open Data Map

# (1) Open Data Map

- ◆ Open Data Map : 환경연구 활용 빈도가 높은 온라인 데이터 hub
  - 카테고리 별 목록 / 링크 / 내용소개 / 키워드를 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 1967개 문서 중 1925개 문서 활용
    - 미분류 문서카테고리 재분류 (277개)
    - 복수 카테고리 문서 제외 (12개)
    - 다운로드가 안되거나 txt 문서로 전환이 안되는 문서 제외 (30개)
- ◆ 문서 카테고리 재분류
  - 총 22가지 카테고리 → 11개 카테고리

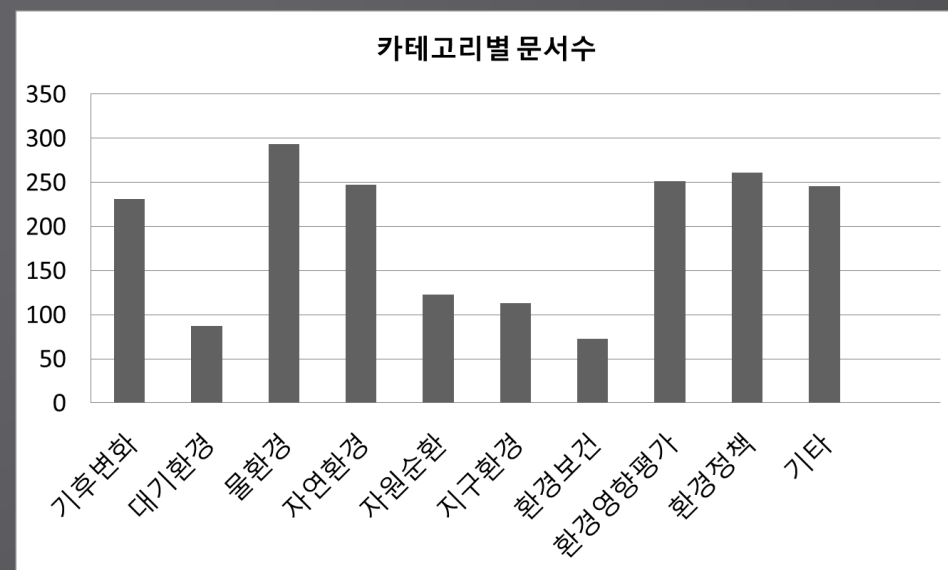
워크시트 | 질의 작성기

```
SELECT R.*, D.SERVICE_URI
FROM DLIDB3.IR_REPORT_VIEW R JOIN DLIDB3.IR_FILE_VIEW D
ON R.C_ID = D.C_ID;
```

질의 결과 x | 질의 결과 1 x

SOL | 50개의 행이 인출됨(0.108초)

CLASS_NAME	RECORD_STATE	CREATE_DATE	MODIFY_DATE	DC_IDENTIFIER	DC_TYPE	KEL_TITLEKOREA
1 6 수시연구	u	19960206	20120116	A 판1185 1995 WO-04	수시연구	우리나라 환경관련 예산정책의 개선방안
2 4 기본연구	u	19960109	20120113	A 판1185 1995 RE-13	기본연구	산업별 종업원수의 수요·수령·수출현황파악 및 재이용에
3 4 기본연구	u	19960109	20120113	A 판1185 1995 RE-15	기본연구	종합환경정보망 개발사업 (III)
4 2 기타보고서	u	19970502	20120116	A 판1185 1997-1	기타보고서	「지방의제21」 모델 개발연구
5 9 기술현황보고서	n	19950127	20111114	A 판1185 1993 AR-03	기술현황보고서	오염지표생물을 이용한 연안환경관리
6 9 기술현황보고서	n	19950127	20111114	A 판1185 1993 AR-01	기술현황보고서	직접여과법
7 9 기술현황보고서	n	19950127	20111114	A 판1185 1993 AR-02	기술현황보고서	다이크산에 관한 검토 및 분석기술
8 9 기술현황보고서	u	19950127	20170329	A 판1185 1993 AR-04	기술현황보고서	환경개선 부담금제도 개선방안.
9 4 기본연구	u	19950127	20170703	A 판1185 1993 RE-01	기본연구	환경기술 연구개발 관리체계 구축방안 (I)
10 4 기본연구	u	19950127	20170703	A 판1185 1993 RE-02	기본연구	환경기술 연구개발 관리체계 구축방안 (II)
11 4 기본연구	u	19950127	20170703	A 판1185 1993 RE-03	기본연구	환경기술 연구개발 관리체계 구축방안 (III)
12 4 기본연구	u	19950127	20120113	A 판1185 1993 RE-04	기본연구	한국형 선진환경산업의 육성책 개발을 위한 기초조사(I)
13 4 기본연구	u	19950127	20120113	A 판1185 1993 RE-05	기본연구	한국형 선진환경산업의 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

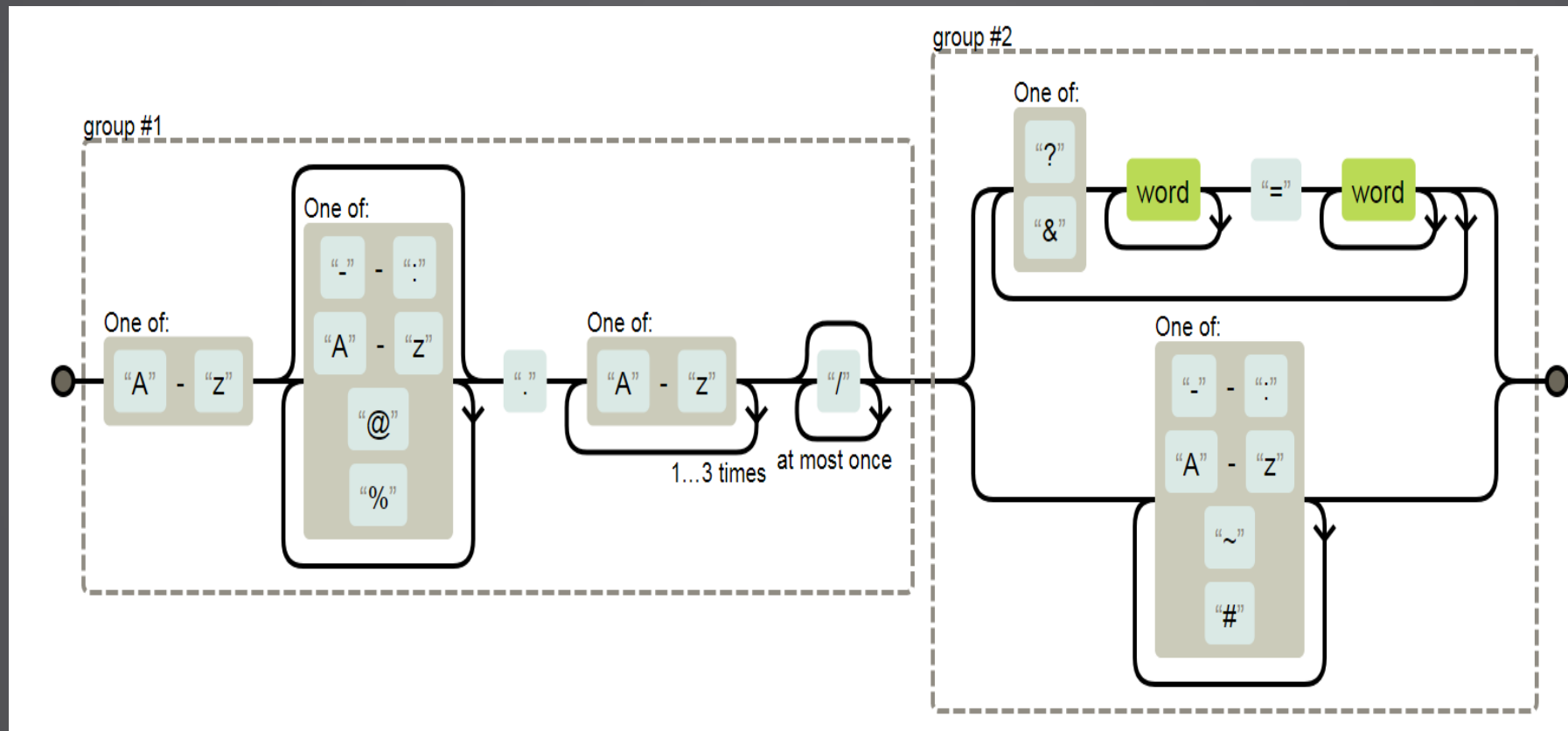
### 1단계: 정규표현식

- 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



# Open Data Map 카테고리 분석

각 카테고리 별 데이터가 사용된 문서의 갯수

순위	데이터	환경정책	환경보건	물환경	대기환경	자원순환	환경영향평가	기후변화	자연환경	지구환경	기타	합	분류	연구자 검증
1	<a href="https://www.me.go.kr">https://www.me.go.kr</a>	20	4	20	9	11	15	19	26	5	20	149	종합	종합
2	<a href="http://www.law.go.kr">http://www.law.go.kr</a>	2	3	9	0	7	11	11	18	4	14	81	종합	종합
3	<a href="https://www.epa.gov">https://www.epa.gov</a>	7	4	16	4	8	5	4	12	0	10	70	종합	종합
4	<a href="https://kosis.kr">https://kosis.kr</a>	1	5	10	1	5	1	18	2	1	20	64	종합	종합
5	<a href="http://water.nier.go.kr">http://water.nier.go.kr</a>	2	0	25	0	2	8	6	5	1	5	55	물환경	물환경
6	<a href="https://www.wamis.go.kr">https://www.wamis.go.kr</a>	4	1	20	0	0	2	12	11	1	2	53	물환경	물환경
7	<a href="http://www.kma.go.kr">http://www.kma.go.kr</a>	2	5	9	2	2	2	20	5	4	1	52	기후변화	기후변화, 대기환경
8	<a href="https://www.eiass.go.kr">https://www.eiass.go.kr</a>	0	1	1	0	3	21	6	6	2	9	50	환경영향평가	환경영향평가
9	<a href="http://www.kosis.kr">http://www.kosis.kr</a>	1	4	5	0	1	3	13	3	1	11	42	기후변화, 기타	종합
10	<a href="https://www.index.go.kr">https://www.index.go.kr</a>	0	0	5	0	3	2	14	7	2	7	40	기후변화	종합
11	<a href="https://www.airkorea.or.kr">https://www.airkorea.or.kr</a>	3	8	2	1	2	1	8	3	3	5	37	기후변화, 환경보건	기후변화, 대기환경
12	<a href="https://egis.me.go.kr">https://egis.me.go.kr</a>	0	0	2	0	1	6	8	14	2	4	37	자연환경, 기후변화	종합
13	<a href="http://www.env.go.jp">http://www.env.go.jp</a>	4	2	1	4	3	4	2	3	4	7	35	종합	종합
14	<a href="http://www.moleg.go.kr">http://www.moleg.go.kr</a>	4	0	4	1	3	6	3	10	0	3	34	종합	종합
15	<a href="http://kosis.nso.go.kr">http://kosis.nso.go.kr</a>	10	1	2	4	1	2	1	8	2	1	32	환경정책	종합
16	<a href="https://ecos.bok.or.kr">https://ecos.bok.or.kr</a>	3	1	1	0	4	0	8	0	1	14	32	기타	기타
17	<a href="http://www.nier.go.kr">http://www.nier.go.kr</a>	2	0	4	1	1	2	9	5	0	7	31	종합	종합
18	<a href="https://www.gims.go.kr">https://www.gims.go.kr</a>	1	1	10	0	0	3	4	7	1	0	27	물환경, 자연환경	물환경
19	<a href="http://www.nso.go.kr">http://www.nso.go.kr</a>	17	1	0	0	2	0	2	2	2	1	27	환경정책	종합
20	<a href="http://airemiss.nier.go.kr">http://airemiss.nier.go.kr</a>	2	3	0	2	0	1	8	0	0	8	24	기후변화, 기타	기후변화, 대기환경

# Open Data Map 웹 서비스 구축

---

- ◆ 카테고리별 데이터 소스 랭킹 알고리즘 적용
  - 데이터소스가 카테고리에 등장한 카운트 횟수 활용
  - 모든 카테고리에서 자주 나오는 데이터소스에 대해서는 패널티 부여
- ◆ R shiny를 활용한 웹서비스 구축
  - 오픈데이터 맵 서비스를 웹으로 활용
  - 카테고리별 순위화된 데이터 소스 출력
  - 데이터 소스에 대한 순위, 웹사이트 주소, 제목, 설명, 한국어 키워드, 영문 키워드 제공
  - 사용자 편의를 위한 검색기능, 출력 문서 개수 설정 등 추가

# Open Data Map 웹 서비스 구축

기후변화	대기환경	물환경	자연환경	자원순환	지구환경	환경보건	환경영향평가	환경정책	기타
Show 10	entries								
Search: <input type="text"/>									
Rank ▲	Website	타이틀	설명	KR_Keywords	EN_Keywords				
1	http://www.kma.go.kr	'하늘을 친구처럼, 국민을 하늘처럼' 기상청	대한민국의 날씨와 기후에 대해 조사를 하고 관찰을 하여 앞으로 어떤 날씨가 있을지 예측하는 기관	환경영향평가, 전략환경영향평가, GIS, 환경 자료 조사, 지리 정보,만경강, 용담댐, 외부유입수량, 하천유지유량, 증가방류,저탄소 사회, 소비행태, 구조분해분석(SDA), 다지역...	Urban ecology,Health,Socioeconomic valuation,Climatic changes,Urban poor,Water resources development...				
2	https://kosis.kr	국가통계포털	국가승인통계 전체를 데이터베이스로 한 곳에 구축하여 국민들이 한 번의 접속으로 쉽고 편리하게 통계자료를 이용할 수 있도록 지원하는 이용자 중심의 One-Stop 국가통계포털 서비스	지역 기후경쟁력, 기후변화 산업 영향, 기후변화 리스크, 기후변화 적응대책,초미세먼지, 건강영향, 대기환경규제지역, 배출량 전망, 관리 방향,물환경목표, 오염원변화, 수질변화, 수...	Regional climate competitiveness, Climate change impacts on industry, Climate change risk, Climate c...				
3	https://www.me.go.kr	환경부	자연환경, 생활환경의 보전, 환경오염방지, 수자원의 보전·이용 및 개발에 관한 사무를 관장하는 대한민국의 중앙행정기관	환경평가, 육상풍력발전, 수상태양광발전, 개발 잠재량, 신재생에너지,녹색생활, 교통, 건물, 여성, 녹색의식,사후환경영향조사, 사후모니터링, 전과정 환경영향평가, 사후환경관리, 협...	Environmental Assessment, Onshore Wind Power Generation, Floating Photovoltaic Power Generation, Ren...				
4	https://www.climate.go.kr	기후정보포털	기후정보포털은 기후변화 과학정보를 중점적으로 소개하며, 영향평가와 정책에 관련된 정보를 동시에 제공하고, 기후변화 연구에 필요한 자료들의 데이터 베이스를 구축하여 기후변화 탐지 모...	적응보고제도, 사회기반시설 기후변화 영향, 기후변화 리스크평가,가뭄, 기후변화, 적응대책, 피해 유형, 가뭄 리스크 관리 유형,기후변화, 취약생태계, 구상나무, 환경안보, 기후변화...	Drought, Climate Change, Adaptation Policy, Types of Drought Effects, Types of Drought Risk Manageme...				
5	https://www.index.go.kr	국가지표체계	정부기관에서 엄선한 740개 지표를 통하여 다양한 방면에서의 우리나라 현위치를 보여 줍니다. e-나라지표 시스템에서 제공하는 지표들은 국가 공식 승인 통계자료뿐만 아니라, 현황이나 ...	갯벌매립사업, 환경영향평가, 환경평가 모니터링, 사후환경관리,국토개발정책, 공간환경정책, 국토계획, 환경계획,대심도 지하공간, 도심지역, 지속 가능성, 지반환경영향, 위해성 소통...	Reclamation Project, Environmental Impact Assessment, Environmental Assessment Monitoring, Post-Envi...				
6	http://kostat.go.kr	통계청	국가통계 발전을 선도하며, 신뢰받는 통계생산으로, 각 경제주체에게 유용한 통계정보 제공	기후변화 적응, 민간부문, 산업계, 리스크, 적응 역량,석탄화력, LNG 복합화력, 탄소중립 프로그램, 배출허용총량,기후변화 리스크 평가, 리스크 목록, 기후변화 영향, 기후변화 ...	Climate Change Adaptation, Private Sector, Industrial Sector, Risk, Adaptative Capacity,thermal powe...				
7	http://www.safekorea.go.kr	국민재난안전포털	대국민 재난정보의 일관되고 통합적인 제공을 위해, 기존 국가재난정보센터, 재난심리 상담정보센터, 재난훈련시스템을 통합 제공	기후변화, 사회경제적 영향, 분류 체계, 주요 지표, 통계,기후변화, 사회경제적 영향, 분류 체계, 주요 지표, 통계,기후변화, 취약성, 평가지표, 영향분석,도시 회복력, 도시 기...	Climate Change, Socio-economic Impacts, Framework, Indicators, Dataset, Adaptation,Climate Change, S...				
8	http://www.kosis.kr	KOSIS 국가통계포털	국가통계포털(KOSIS, Korean Statistical Information Service)은 국내·국제 북한의 주요 통계를 한 곳에 모아 이용자가 원하는 통계를 한 번에 찾을...	저탄소사회 정책, 소비행태 분석, 온실가스 유발배출량, 배출증감 요인분해, 다지역투입산출분석, 연산가능 일반균형모형,개발사업, 수질오염, 인식조사, 지방자치단체, 총량관리,환경가치...	ecology,Urban poor,Development Project, Local Government, Pollution Load, Staff Survey, Water Qualit...				
9	http://www.wamis.go.kr	국가수자원관리종합정보	물관련 정보를 대국민 서비스하기 위해 구축 운영되고 있는 인터넷 기반의 포털 시스템	지중 환경, 지중 자연환경, 지중 생활환경, 지중 환경 구성요소, 지중 환경 가치,지하수 오염, 관리, 정화, 제도적 지원, 사후조치,물-효율적 경제사회, 물이용 인식, 물이용 갈...	Subsurface Environment, Subsurface Natural Environment, Subsurface Living Environment, Subsurface En...				
10	http://www.greengrowth.go.kr	녹색성장위원회	녹색성장위원회는 정부의 녹색성장 정책을 심의·조율하고, 사회 각계의 다양한 의견을 수렴하는 국무총리 소속 기구입니다.	녹색생활, 교통, 건물, 여성, 녹색의식,기후변화, 에너지, 온실가스, 감축, 국가 계획,환경협력, 중견국협의체, 중견국 외교, 한?ASEAN 환경협력, 거버넌스,기후변화 적응, ...	Climate Change, Greenhouse-gas, Energy, Emission Reduction, National Plans,Environmental Cooperation...				
Showing 1 to 10 of 1,000 entries									
Previous 1 2 3 4 5 ... 100 Next									

Showing 1 to 10 of 1,000 entries

Previous

1

2

3

4

5

...

100

Next



# Open Data Map 웹 서비스 구축

기후변화

대기환경

물환경

자연환경

자원순환

지구환경

환경보건

환경영향평가

환경정책

기타

Show 10 entries

Search:

Rank	Website	타 이 틀	설명	KR_Keywords	EN_Keywords
1	https://www.me.go.kr	환경 부	자연환경, 생활환경의 보전, 환경오염방지, 수자원의 보전·이용 및 개발에 관한 사무를 관장하는 대한민국의 중앙행정기관	환경평가, 육상풍력발전, 수상태양광발전, 개발 잠재량, 신재생에너지, 녹색생활, 교통, 건물, 여성, 녹색의식, 사후환경영향조사, 사후모니터링, 전과정 환경영향평가, 사후환경관리, 협...	Environmental Assessment, Onshore Wind Power Generation, Floating Photovoltaic Power Generation, Ren...
2	http://www.epa.gov/airmarkets/trading/so2market/transtable.html				Air quality management- Korea (South),Emissions trading
3	http://www.yeosu.go.kr/site/Home				Emissions trading,Emissions trading
4	http://www.gwangyang.go.kr/site/Home				Emissions trading,Emissions trading
5	http://www.nhlbi.nih.gov/resources/docs/02_chdbk.pdf				
6	http://www.deh.gov.au/atmosphere/airtoxics/publications/sok				
7	http://www.epa.gov/cgi-bin/broker				Emissions trading
8	http://www.suncheon.go.kr/home/information/statistics_info/2005				Emissions trading
9	http://gwangyang-jeonnam.stat.go.kr/cgi-bin				Emissions trading
10	http://yeosu.go.kr/gallery/stat				Emissions trading

Showing 1 to 10 of 1,000 entries

Previous

1

2

3

4

5

...

100

Next

<http://192.168.1.51:8787/p/6414/>

# 향후 계획

---

## ◆ 사이트 접속 여부 업데이트 필요 (추가/대체 필요)

– 패턴은 일치하지만 웹사이트가 아닌 경우

- <http://go.kr>

– 웹사이트가 없어진/대체된 경우

- <http://kosis.nso.go.kr>

- <http://www.nso.go.kr>

- <http://www.moct.go.kr>

– 기타 경우

- <https://www.wamis.go.kr> (접속 X) vs <http://www.wamis.go.kr> (접속 O)

- <http://www.law.go.kr> vs <http://www.law.go.kr/main.html>

# 향후 계획

---

- ◆ 주기적(분기별/연별 등) 업데이트 가능하도록 수정(자동화) 필요
- ◆ 웹사이트 제목/설명 등 추가 작업 필요
  - 현재는 전체에서 50위까지
- ◆ 랭킹 알고리즘/카테고리 세부 조정
- ◆ 원내 서비스를 위한 서비스 베포 관련 작업
- ◆ 서비스 테스트

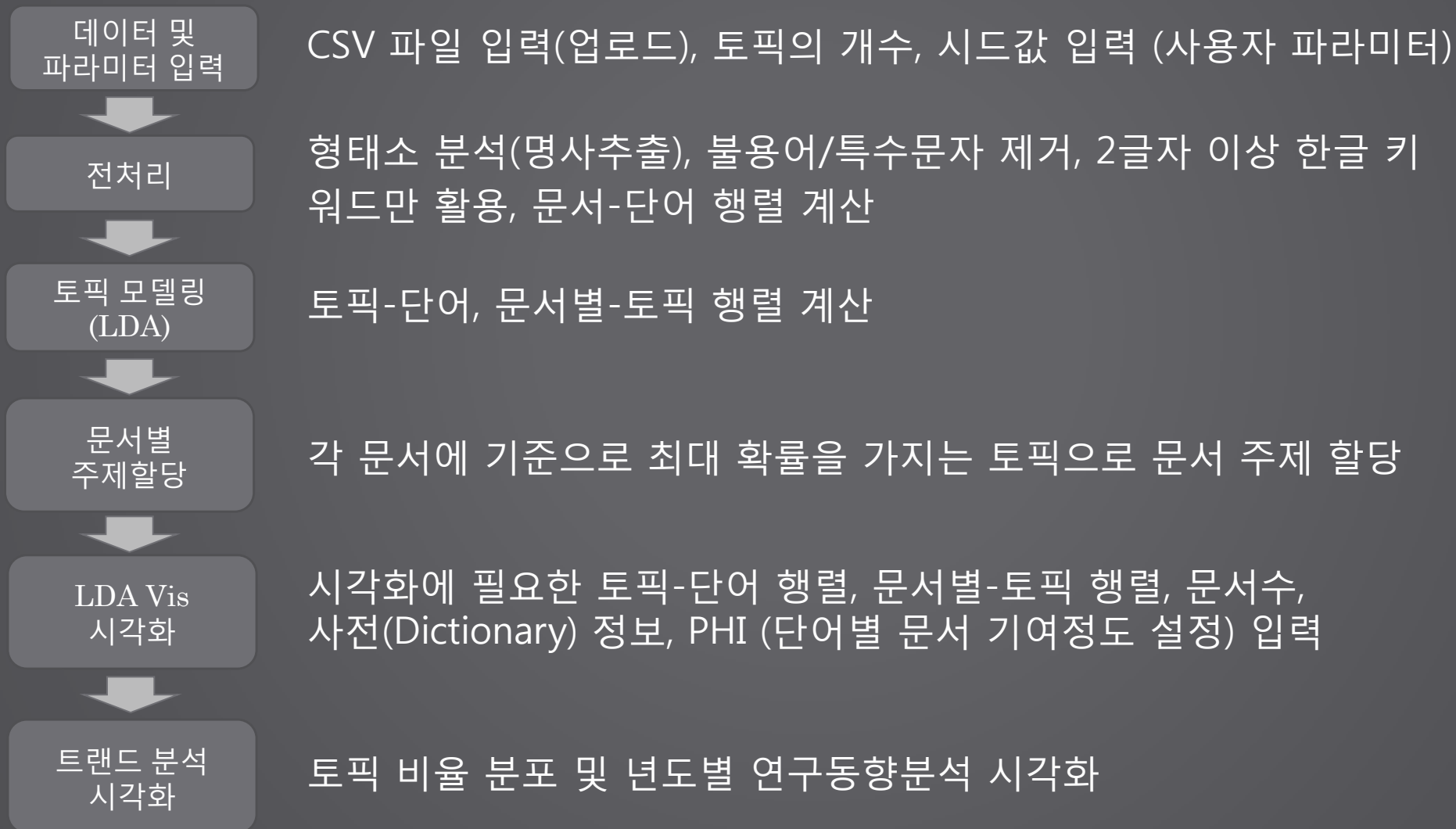
## 2. 연구동향 분석 서비스

# (1) LDA 토픽 모델링 및 연구동향 분석 서비스

---

- ◆ LDA 토픽 모델링 및 연구동향 분석 서비스
  - 텍스트(문서)들의 집합에 포함된 주제를 분석하고, 이를 토대로 연도별 동향을 파악하는 서비스
- ◆ 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
  - 코드 간소화 : 파일과 인자를 입력 받으면 한번의 실행으로 모든 결과가 도출되게 하는 작업
  - 형태소 분석기 추가, 파라미터 지정 부분 분리, 시각화 관련 부분 수정 등
- ◆ 웹서비스 구축: R shiny를 활용한 웹 기반 GUI 인터페이스 구축
  - 원내 서비스를 위한 테스트 단계
  - 파일 업로드 기능
  - 연도별 연구동향 분석 기능
  - 문서별 토픽 분석결과 및 문서들의 토픽 분포 추가

# (1) LDA 토픽 모델링 및 연구동향 분석 알고리즘 요약



## (1) LDA 토픽 모델링 및 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 윤	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 충남 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 차	2016	10	31
a13	제1장 서론 1. 연구의 배경 및 필요성 2. 연구의 범위 제2장 국내 지역사회 환경보	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 「정화관리	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

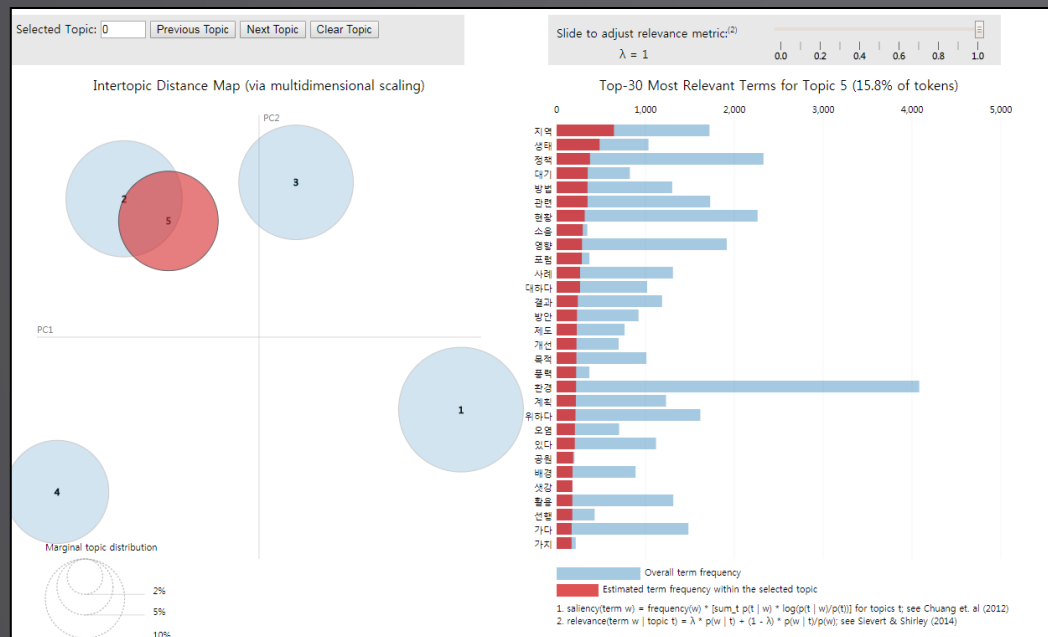
# LDA 토픽 모델링 웹서비스

[illegible]

테스트 데이터 : 도서관 DB의 연구제목  
(총 1700개)

# (1) LDA 토픽 모델링 및 연구동향 분석 서비스 GUI

## LDA 토픽 모델링 시각화



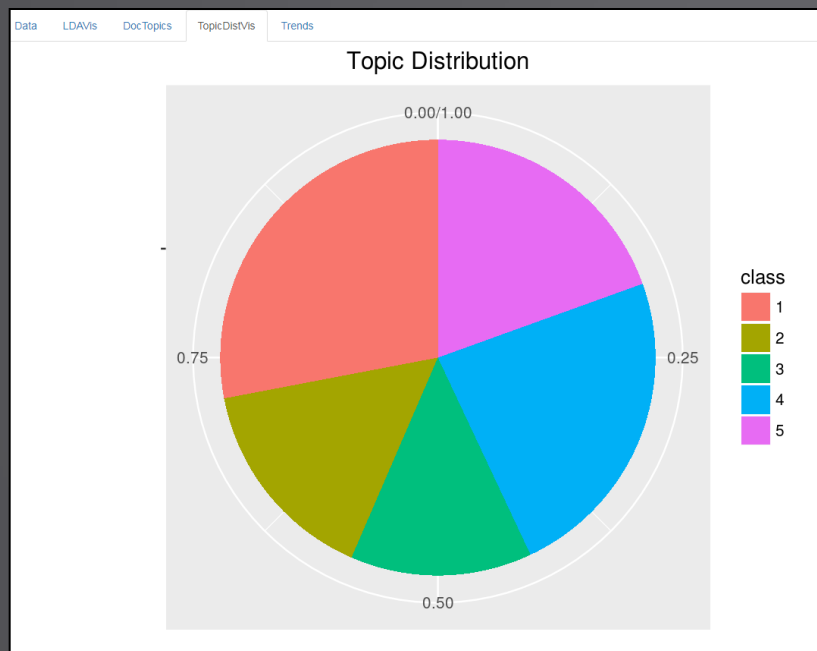
## 문서 토픽 정보

Data	LDAVis	DocTopics	TopicDistVis	Trends	
name					doc_topic
환경분야 공적개발원조(ODA) 사업평가 지침 마련을 위한 연구					3
제주 탄소제로성 추진전략 연구					5
저탄소 기후변화 적응 사회를 위한 사회·경제 변화 시나리오 개발					3
화학사고의 경제적 손실 추정을 위한 방법론 진단 및 선정 방안 연구 : 인적·생태적 피해액 추정을 중심으로					1
나노폐기물의 안전처리를 위한 관리전략 수립 연구					5
가뭄 단계에 따른 적응형 가뭄관리정책 연구 : 지역 차원의 비구조적 가뭄대처를 중심으로					2
국내 농산축 GIS기반 통합관리시스템 개발					3
기후변화에 따른 건강영향 평가·적응 기술 및 정책지원 시스템 개발(I-IV)					3
Post-2020 신기후체계 협상 적응의제 대응방안 연구					3
사물인터넷(IoT)을 활용한 스마트 물환경관리 방안 및 정책기반 마련 연구					3
중국의 '일대일로(一帶一路)' 대응 유라시아 지역 환경전략 연구					1
도시의 기후 회복력 확보를 위한 공간단위별 평가체계 및 모형 개발(II)					4
지역기반 환경보건정책 지원 방안 연구(II)					5
신기후체계의 기후변화 적응 및 손실과 피해에 관한 대응방안					4
대기환경비용을 고려한 친환경차 구매보조금 실효성 제고 연구 : 차종별 적정 보조금 수준 분석을 중심으로					4
실도로에서 경유차의 대기오염물질 초과 배출에 따른 사회적 비용 연구					5
토양정화 관련 부지의 최적 관리방안 연구					3
제3차 국가생물다양성전략 이행상황 점검					4
국가 지속가능성 평가 등에 관한 연구					1

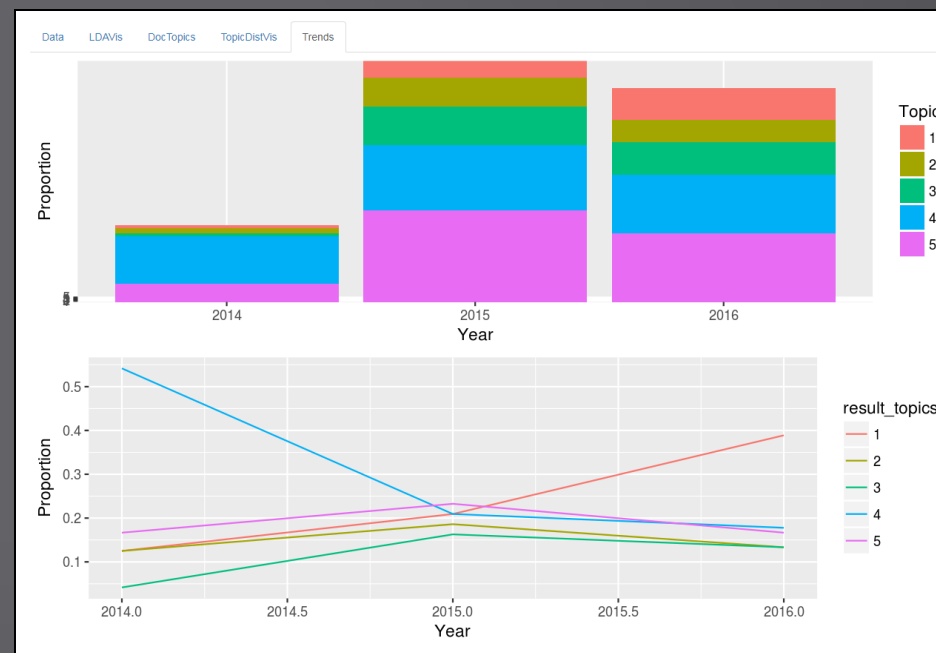


# (1) LDA 토픽 모델링 및 연구동향 분석 서비스 Output

문서 주제분포 시각화



연구동향분석 시각화



## (2) 키워드 네트워크 분석 서비스

### ◆ 키워드 네트워크 분석 서비스

- 환경 텍스트의 키워드를 파악하고 키워드 사이의 연관 빈도가 높은 단어들의 네트워크를 도출

### ◆ 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

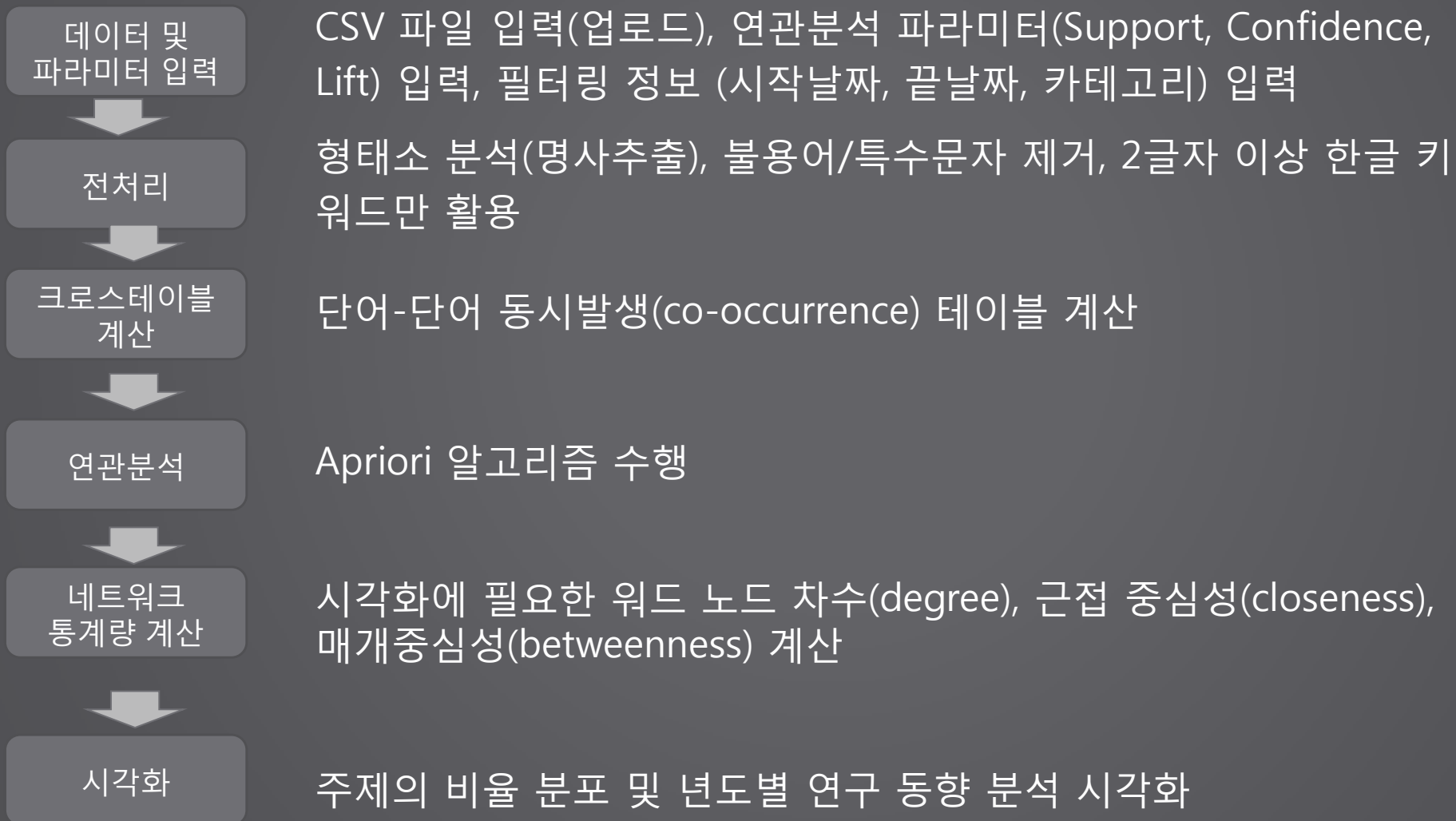
### ◆ 동일 포맷 다른 데이터 입력 시 연관분석 결과를 출력할 수 있도록 기존 소스코드 일반화

- 코드 간소화 : 파일과 인자를 입력 받으면 한번의 실행으로 모든 결과가 도출되게 하는 작업
- 연관분석 핵심코드 정리 및 함수화 및 각종 인자 파라미터화 (연관분석 및 시각화 부분 : 차수, 중심성, 근접중심성 등 고려 및 활용)
- 규칙 수정 : 일부 오류 수정, 빈도수 규칙 추가, LIFT 기준으로 규칙 정렬

### ◆ 웹서비스 구축: R shiny를 활용한 웹 기반 GUI 인터페이스 구축

- 원내 서비스를 위한 테스트 단계
- 파일 업로드 기능 추가
- 지지도(support), 신뢰도(confidence), 사용할 규칙 수 입력 기능
- 날짜 및 카테고리 필터링 기능 추가

## (2) 키워드 네트워크 분석 알고리즘 요약



# 키워드 네트워크 분석 서비스 GUI

입력 파일 포맷 (csv)

DATE	TITLE	CATEGORY
19960109	상·하수처리의 효율적 운영방안 연구(I)	물환경
19970501	폐자원 재활용의 경제성 분석	폐기물관리
19980716	우리나라 환경법체계 정비에 관한 연구 I	정책일반
20011101	배출부과금 부과제도 개선방안	정책일반
19960109	환경친화적 기업경영체제 발전방향에 관한 연구	정책일반
19960109	폐기물 처리시설 적정관리를 위한 폐기물 유통체계 개선방안	폐기물관리
19961002	지방자치제 하에서의 자치단체간 오염규제보조금제도 연구	물환경
19970501	연안역에서의 오염물질 유입 저감을 위한 총량규제 방안에 관한 연구	물환경
19970501	산성비 원인물질 저감방안에 관한 연구 III	대기환경
19970501	환경친화적 상품의 개발 및 생산촉진 방안 연구	정책일반
19980525	민간기업의 환경기초시설 운영제고 방안 연구	정책일반
19980525	대기보전정책과 에너지정책의 조화방안 연구	대기환경
20020301	상수원 관리지역 주민지원사업 추진결과에 대한 평가	물환경
20020708	개도국의 온실가스 감축의무 참여방식에 관한 연구	대기환경
19960109	환경보전의 국제화와 정책대응(II)	정책일반
19970501	지역대기질 보전대책 수립 및 시행에 관한 연구	대기환경
20120808	물 재이용 기본계획 수립을 위한 연구	물환경

키워드 네트워크 분석 웹서비스

## Keyword Association Analysis

Choose CSV File

Browse... title\_data.csv

Upload complete

Category

물환경

StartDate

20100101

EndDate

20171231

Do filtering

Data Result Plot

DATE	TITLE	CATEGORY
20120808	물 재이용 기본계획 수립을 위한 연구	물환경
20120808	물 재이용 기본계획 수립을 위한 연구(부록)	물환경
20130203	물문화 선진화의 정책방향 설계(I)	물환경
20130203	물이용 인식 선진화 정책 연구(I)	물환경
20121018	토양·지하수 정책 지원시스템 구축을 위한 중·장기 계획수립 연구	물환경
20111231	물관리 취약성과 물안보 전략 III	물환경

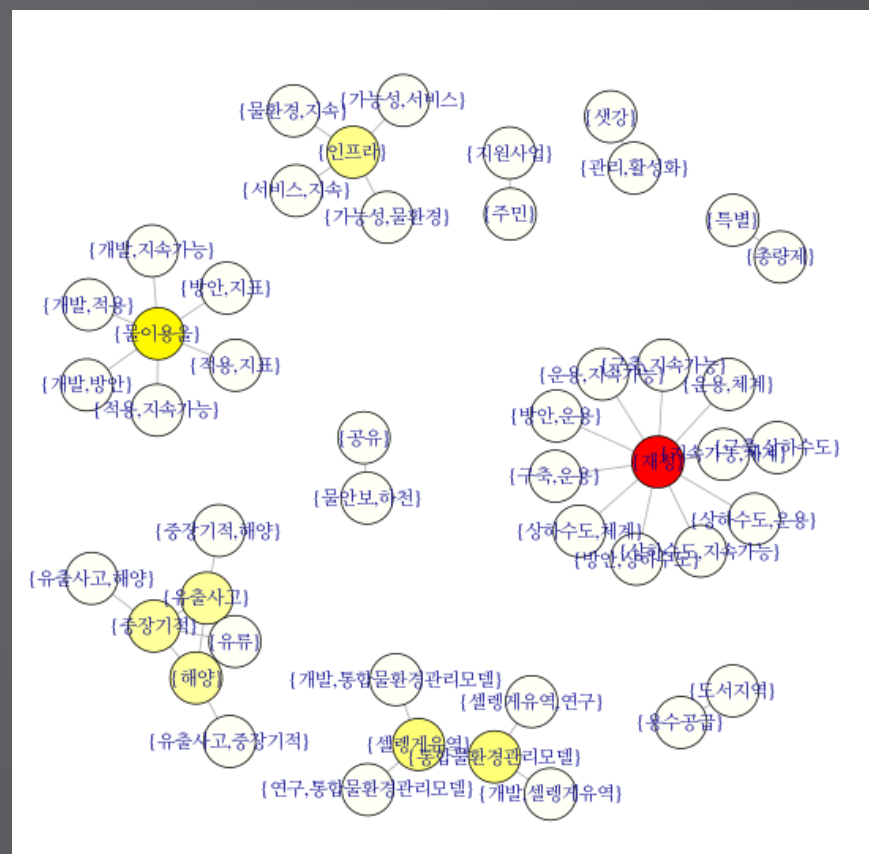
테스트 데이터 : 도서관 DB의 연구제목  
(총 1700개)

# 키워드 네트워크 분석 서비스

## 키워드 연관 분석 결과

Data	Result	Plot				
lhs	V2	rhs	support	confidence	lift	count
{도서지역}	=>	{용수공급}	0.01438849	1.00000000	69.5000000	2
{용수공급}	=>	{도서지역}	0.01438849	1.00000000	69.5000000	2
{셀렝게구역}	=>	{통합물환경관리모델}	0.01438849	1.00000000	69.5000000	2
{통합물환경관리모델}	=>	{셀렝게구역}	0.01438849	1.00000000	69.5000000	2
{주민}	=>	{지원사업}	0.01438849	1.00000000	69.5000000	2
{지원사업}	=>	{주민}	0.01438849	1.00000000	69.5000000	2
{해양}	=>	{증장기적}	0.01438849	1.00000000	69.5000000	2
{증장기적}	=>	{해양}	0.01438849	1.00000000	69.5000000	2
{해양}	=>	{유출사고}	0.01438849	1.00000000	69.5000000	2
{유출사고}	=>	{해양}	0.01438849	1.00000000	69.5000000	2
{해양}	=>	{유류}	0.01438849	1.00000000	69.5000000	2

## 키워드 연관 분석 시각화



# 향후 계획

---

- ◆ 원내 서비스를 위한 서비스 베포 관련 작업
- ◆ 서비스 테스트
- ◆ 시각화 결과 정교화