

소수자를 향한 언론의 관점, 28년

뉴스빅데이터 활용 프로젝트

권오성 한겨레 데이터베이스 기자

이재호 한겨레21 기자

고정관념 (STEREOTYPE)

- 월터 리프먼
- 우리 마음 속의 그림

기획 소개

- 2018년은 ‘소수자의 해’였다고 해도 과언이 아님
- ‘미투 운동’ 성 불평등에 대한 근본적 문제제기
- 제주도 예멘 난민, 난민에 대한 공동체의 생각 급변 계기
- 전국 각지에서 첫 퀴어 축제
- 소수자에 대한 우리 생각은 어떻게 바뀌어 왔을까?
- 빅카인즈의 방대한 뉴스 텍스트를 분석해 단초를 발견
- 시각화 작업을 통해 이해하기 쉬운 뉴스 기획으로 구현

분석 대상

- 빅카인즈 DB

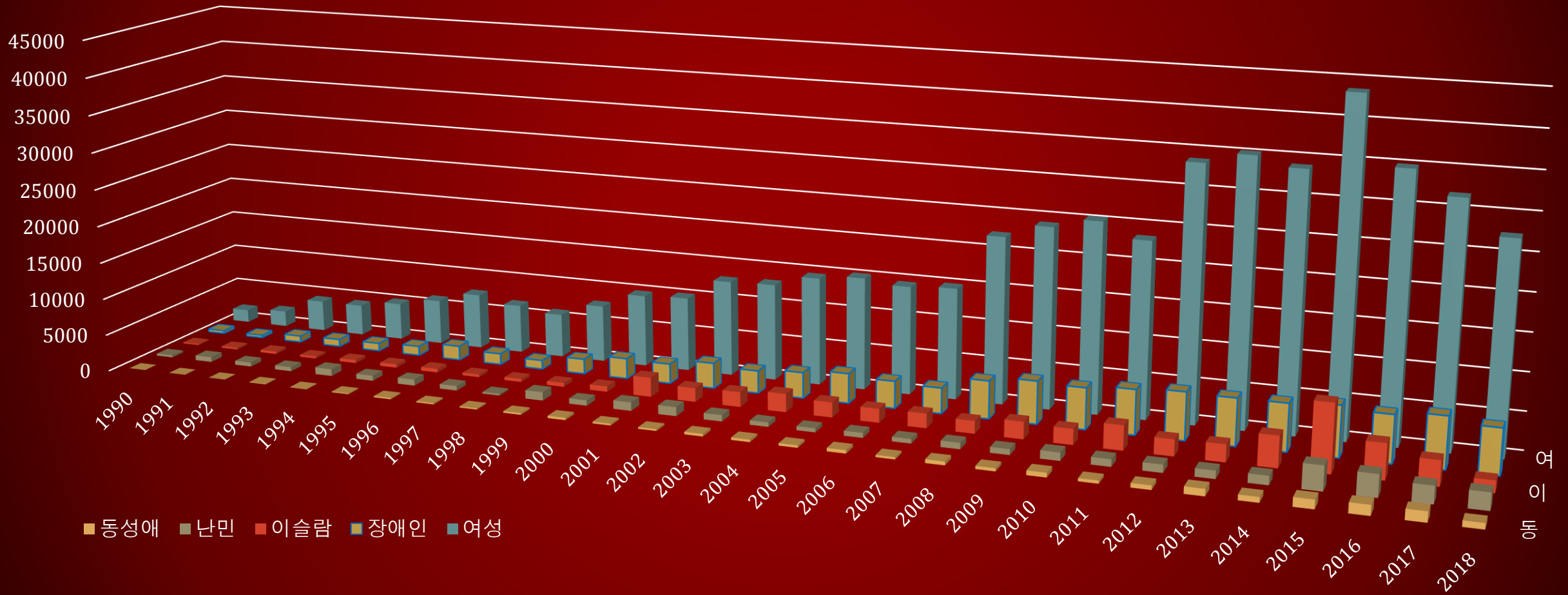
분석 대상

- 우리 사회의 대표적인 소수자 키워드 다섯
- 여성, 동성애, 장애인, 난민, 이슬람
- 시계열상 일관된 조건을 위해 1990~2018년(10월) 대상 기간 동안 누락 없이 전체 기사가 데이터베이스화 된
- 중앙 종합일간지 대상
- 경향신문, 국민일보, 서울신문, 세계일보, 한겨레, 한국일보

관련 보도 및 선행연구

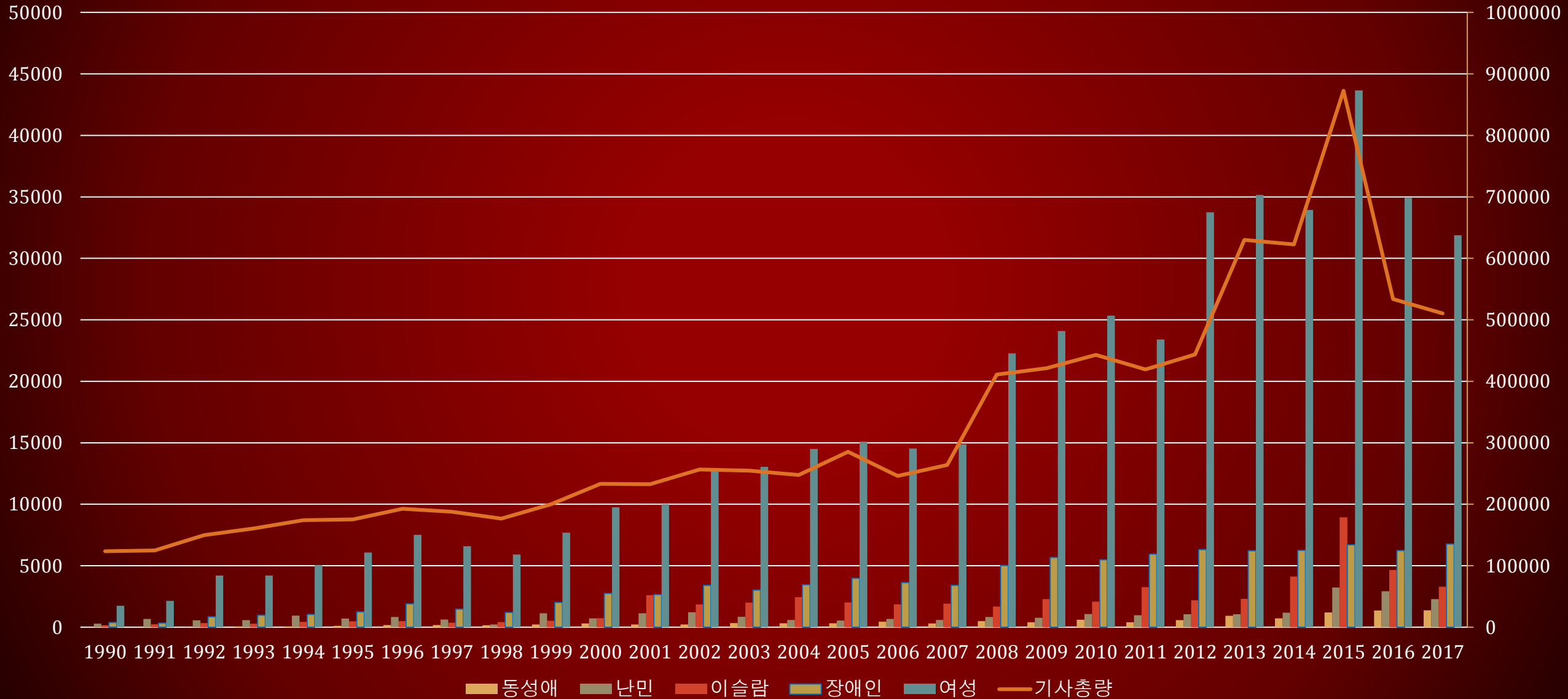
- ‘여성 1호’ 보도, ‘여성혐오’ 보도, 여성 정치인 보도에 관한 연구
- 성소수자 언론보도 프레임 분석(2007년 동아일보, 한겨레, 경향신문)
- 종합일간지 지적장애 보도 분석
- 언론사별 장애 보도 경향(개인 노력에 초점 VS. 정부지원 시스템 접근)
- 최근에 떠오른 난민문제는 선행연구나 언론보도 분석 부족
- 한국인의 아랍, 이슬람 이미지 및 관련 언론보도 인식연구(2016년)

여성, 동성애, 장애인, 난민, 이슬람

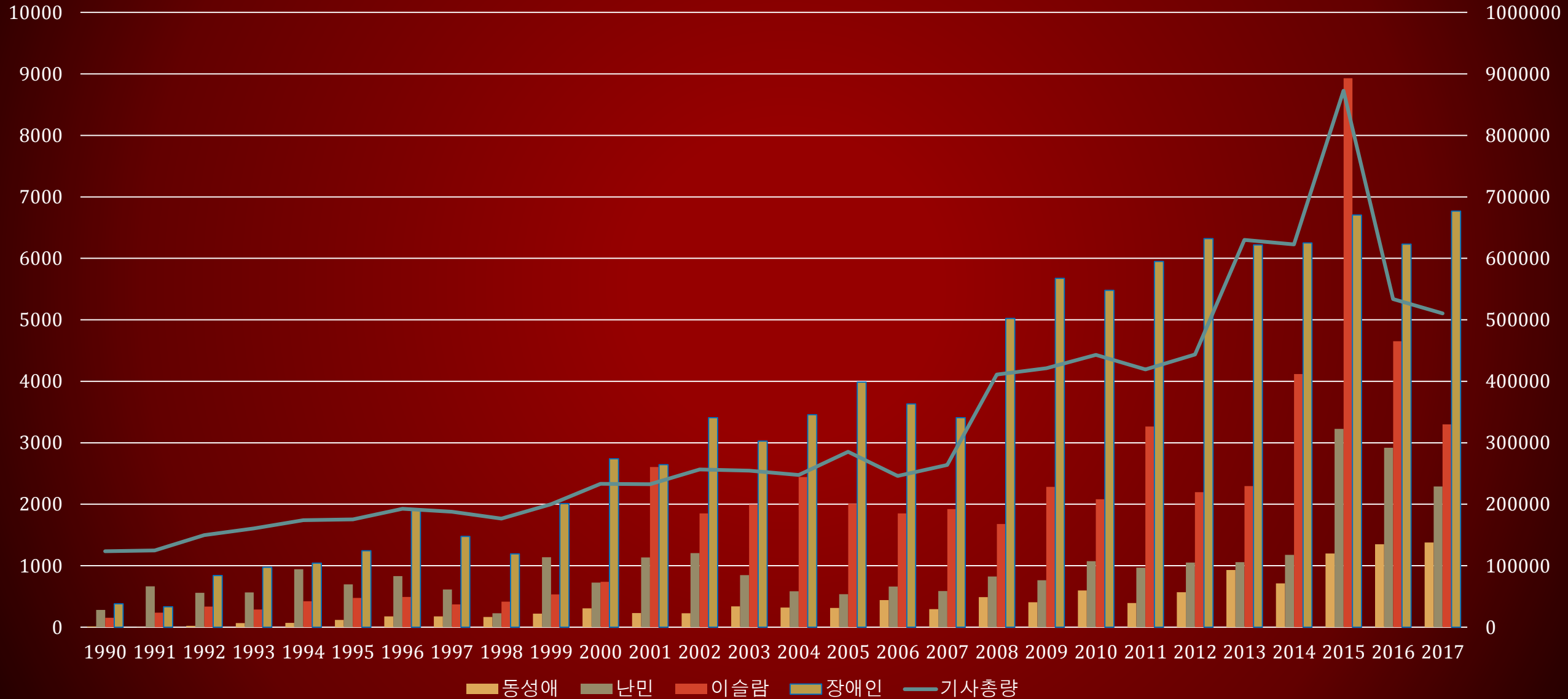


총량: 동성애 1.2만, 난민 3만, 이슬람 5만, 장애인 10만, 여성 49만

전체 기사수 추이 대비 비교



전체 기사수 추이 대비 비교 (여성 제외)



분석 방법

- 빈도분석: R
- 토픽 모델링: PYTHON

작업 화면

빈도분석: 키워드 별 특징

- 여성: 정치에서 범죄, 그리고 미투로
- 동성애: 혐오에서 인권으로
- 장애: 단조로워 더 소외된
- 난민: 국제 → 북한 → 국내
- 이슬람: 세계 뉴스에서 우리 뉴스로

여성: 연도별 상위 단어 ([링크](#))

여성: 분석

- 초기 '여성' 뉴스는 정치가 초점(후보, 대표, 민주, 자당, 신한국당)
- 1994년 '세계': 김영삼 정부 '세계화'
- 2000년 들어 다양한 주제가 부각되기 시작: 교수(2002), 영화(2005), 결혼(2008), 환자(2009), 시작(2011), 인기(2015)
- 2011년부터 가장 두드러진 단어: '경찰'
- '여성'은 주로 '범죄'에 대한 보도와 함께 등장
- 트럼프(2016), 대통령(2017)
- 2018 미투의 증거: 피해자(7), 성폭력(10)

동성애: 연도별 상위 단어

동성애: 분석

- ‘에이즈’, ‘감염’에서 ‘인권’, ‘혐오’에 도달하기까지
- 90년대 기사 예: “동성애는 공공연한 변태적 성행위”
- 동성애가 보다 가까이 느껴지기까지 가장 큰 공신은 ‘영화’
- 우리 언론에서 미국을 얼마나 중요하게 다루는지 한 단면
- 반대의 움직임도: ‘에이즈’ 2015 재등장, ‘교회’, ‘목사’
- 분석 대상 6종 가운데 2종이 기독교계 신문

장애: 연도별 상위 단어

장애: 분석

- 가장 단조로운 키워드
- ‘지원’, ‘사회’, ‘사람’의 지속적인 강세
- 다른 키워드와 달리 특정 이슈가 부각되기 힘들게 ‘장애인 지원’에 대한 기사가 꾸준히 다수 생산되는 구조로 추정
- ‘여성’과 유사하게 초반 정치 연관성은 나타남: 후보(1992)

난민: 연도별 상위 단어

난민: 분석

- ‘국제’ 뉴스→북한, 탈북자→국내 난민으로 보도내용 이동
- 북한의 경제상황 악화 후 탈북자 수 늘자 보도량도 증가
- 2001년 국내 ‘1호 난민’ 인정으로 국내 난민에도 관심
- 지속적인 탈북자 발생으로 탈북자 지원, 고용 관련 보도
- 2004년 난민과 관련, 인권을 언급하는 보도
- 시리아 난민 대량 발생으로 2015년 보도 정점
- 2018년 예멘 난민 유입으로 보도량 증가

이슬람: 연도별 상위 단어

이슬람: 분석

- 난민과 유사하게 ‘국제’보도에서 ‘사회’보도로 확장
- 이슬람과 관련된 보도 1990년 걸프전부터 시작
- 종교로서 이슬람 보다 국가분쟁·종교갈등과 관련돼 언급
- 1994년, 2001년 무역센터 테러 이후 부정적 보도 경향 공고화
- 한국의 중동지역 파병 전후 이슬람 보도량 증가
- IS(이슬람 국가) 이후 테러리즘과 관련한 이슬람 보도
- 최근에는 이슬람 여성, 난민 등과 관련한 보도 주목

토픽 모델링

- 진행중

결론 -빈도 분석

- 우리 사회의 소수자에 대한 관심은 지속적으로 높아짐(기사량)
- 인권 의식 높아지는 경향이 나타남(특히 동성애)
- 초반에 관찰(난민, 이슬람), 이용(여성, 장애인에서 나타나는 정치 언어), 심지어 혐오(동성애)하는 관점에서
- 우리의 문제로 보거나(난민, 이슬람), 그들의 목소리를 전하는 관점이 점차 늘어나는 경향 보임
- 미국 주류의 화제가 국내 언론에 강하게 투영(이슬람, 동성애)

결론 -토픽 모델링

결론: 언론에서 텍스트 마이닝의 의미

- 한국언론진흥재단과 사이버커뮤니케이션학회 주최 뉴스 빅데이터 분석 프로젝트 사례

Q&A

감사합니다