

텍스트 마이닝으로 분석한 한은소식

Text Mining

한은소식은 1978년 10월 15일 제1호를 시작으로 여러 코너를 통해 경제, 문화, 취미활동 등 다양한 주제의 관심사를 다뤘었다. 본 분석에서는 텍스트 자료 확보가 용이한 2004년 이후로 한정하여 한은칼럼, 행우광장, 총재 신년사, 생각하는 경제 등 네 코너를 살펴보았다. 분석방법으로는 텍스트 마이닝에 널리 이용되는 단어기반 빈도 분석(Bag of words)을 이용하였다.

우선 형식이 비슷한 한은칼럼과 행우광장 원고를 살펴보자. 자유주제 수필 형식인 한은칼럼과 행우광장이야말로 한은맨의 평소 관심과 가치관이 그대로 드러나는 코너이다. 지금 당장 한 편의 수필을 부탁받은 당신 무엇을 말할 것인가. 2004년 이후 한은칼럼 147편과 행우광장 145편을 살펴보면 가장 많이 등장하는 단어는 '사람' 그리고 '생각'이다(차트1). 다만 한은칼럼에서는 경제, 사회, 조직 등의 단어가 그리고 행우광장에서는 마음, 여행, 시작 등의 단어가 자주 등장하여, 한은칼럼이 조금 더 무거운 주제를 다루고 있음을 짐작할 수 있다 (차트2).

더 자세한 분석을 위해 두 코너에 나타난 한은맨의 관심사를 Latent Semantic Indexing 모형을 이용하여 네다섯 가지의 주제로 나눠보았다. 한은칼럼의 주제는 크게 '삶'(42%), '경제/금융'(11%), '기업/사회/소득'(27%), '디지털/미래'(5%), '조직/당행/화폐'(15%) 등 다섯 가지로 분류되었으며 그 비중은 연도별로 다소 변동은 있으나 꾸준히 유지되었다 (차트3).¹⁾ 반면 행우광장의 주제는 크게 '가족'(10%), '여행'(25%), '문화/스포츠'(37%), '삶/행복/가치'(28%) 등 네 가지로 분류되었으며, 최근 들어 여행 관련 원고가 크게 늘어났음을 알 수 있다 (차트4). 이 중 여행 및 문화/스포츠의 관련 키워드를 뽑아보면 여행지로는 '제주'가 그리고 문화/스포츠 활동으로는 '영화'가 가장 많이 언급되었다 (차트5). 또 재밌는 점으로, 두 코너 모두에서 '자신'과 '마음'이 삶 부문 최빈도 단어에 포함되었는데, 두 단어는 다른 주제에는 자주 등장하지 않는다. 삶에 대한 글쓰기가 결국 '자신'의 '마음'을 살피는 자아성찰의 과정이었기 때문으로 해석해 본다. 조금 더 흥미로운 분석을 위해 시계열 분석에 용이한 총재 신년사를 생각하는 경제와 함께 살펴보았다. 한 해의 경제 전반 및 한국은행 가정사를 아우르는 총재 신년사와 외부인사들이 주로 기고하는 생각하는 경제에는 어떤 차

한은소식이 이번 달 500호를 맞이한다. 40여 년에 걸쳐 한국은행 안팎의 다양한 관심사를 담아 온 한은소식. 한은소식 500호를 기념하여 그동안의 기고 원고를 최신 텍스트 마이닝 기법을 이용해 분석해 보고자 한다.



차트1)

이가 있을까. 우선 총재 신년사의 주요 키워드를 살펴보자. 빈도수 기준 키워드 1위는 단연 '경제'(457회)가 차지했고, 그 뒤를 '정책'(201회), '한국은행'(156회), '중앙은행'(143회), '상황'(120회), '금융'(119회), '변화'(119회), '노력'(117회), '조직'(102회), '통화정책'(100회) 등이 뒤따랐다 (차트6). 이들 중 흥미로운 키워드 몇 개를 골라 연도별 추이를 살펴보았다. 연도별로 보면 2008년 금융위기 이후 '위기' 및 '글로벌/세계'에 대한 언급이 한동안 높게 유지되었으며 해당 기간 동안 '물가', '금융', '성장/고용' 등에 대한 언급은 상대적으로 줄어든 것으로 나타났다 (차트7). 이러한 인식이 생각하는 경제에서도 비슷하게 나타났을까? 생각하는 경제는 주로 기자와 교수의 원고로 채워지는 만큼 그 주제가 다양했다. 2004년 이후 137편의 원고를 살펴보면, 가장 많이 등장하는 10개 단어로 '경제', '정부', '사람', '미국', '기업', '문제', '사회', '한국', '시장', '정책' 등이 포함되었으며 '중국', '세계', '소득', '산업' 등도 20위 안에 들었다 (차트8). 총재 신년사에서 살펴본 것과 같은 키워드로 연도별 추이를 살펴보면, '위기'에 대한 언급은 2009년 두세 편의 원고에서 집중적으로 사용된 이후 그 사용이 미미하였다 (차트9). 다양한 주제가 다뤄진 만큼, 연 10편 내외의 원고로 시계열을 분석하기에는

1) 한은칼럼 및 행우광장 2004~2007년 자료의 경우 텍스트 형식으로 남아있는 원고가 많지 않아서 해당 수치의 연도별 대표성이 떨어지는 점을 유의해주시 바랍니다.

차트2) 한은칼럼/행우광장 코너에 가장 많이 등장한 단어



차트5) 여행 및 문화/스포츠 부문 관련 키워드



차트3) 한은칼럼 연도별 주제 변화

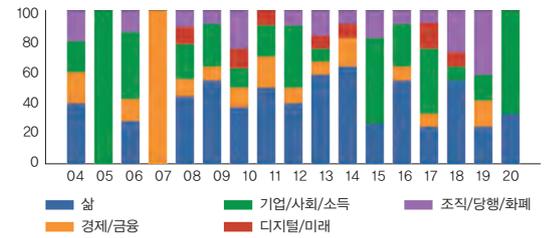


차트4) 행우광장 연도별 주제 변화

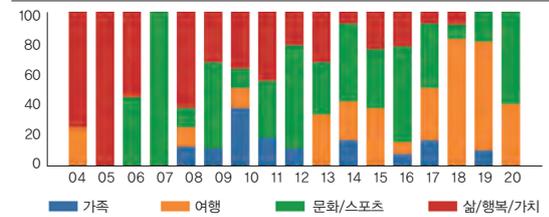


차트7) 총재 신년사 주요 키워드 비중 연도별 추이

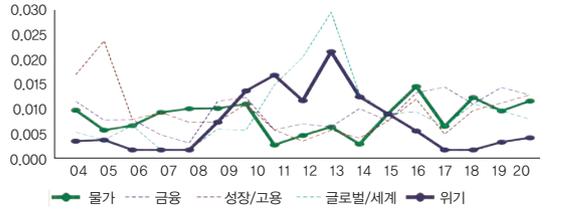


차트9) 생각하는 경제 주요 키워드 비중 연도별 추이

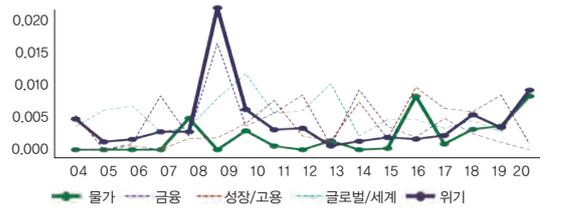


차트10) 생각하는 경제에서 언급된 주요 국가의 인용횟수

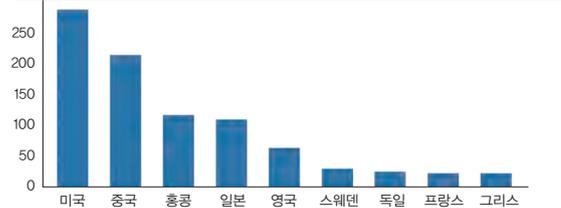


차트6)



차트8)



부족한 면이 있지만, 2009년 경제위기 및 2016년 저물가에 대한 관심이 총재 신년사에서 드러난 것과 비슷하게 생각하는 경제에서도 나타난 것으로 보인다. 한 가지 더 재밌는 특징으로, 생각하는 경제 코너에서는 다른 나라 인용 횟수가 여타 코너에 비해 월등히 많았는데, 언급된 국가들의 인용 패턴을 확인해 보았다. 총 7개 국가가 언급됐으며, 홍콩을 제외하면 예상과 비슷하게 미국, 중국, 일본, 영국 등이 높은 순위를 차지하였다 (차트10). 이들 국가의 연관 단어를 함께 살펴보면 미국은 '세계', '가격', '사회', '경제' 등과 함께 등장했고 중국은 '일본', '미국', '수출', '경제'와, 홍콩은 '달러', '중국', '아파트', '집값'과 그리고 일본은 '미국', '경제', '거품', '일본은행' 등과 함께 등장하여 이들 국가에 대한 집필자의 관심을 조금이나마 엿볼 수 있었다. 이상으로 한은소식 네 코너에 나타난 다양한 패턴을 텍스트 마이닝 기법을 이용해 살펴보았다. 한은소식은 한은맨의 관심과 가치관을 공유하는 장으로서 또 종종 무료함을 달래주는 매체로서 역할을 해왔다. 특히 한은칼럼 및 행우광장 코너에서는 삶에 대한 따뜻한 인식이, 그리고 생각하는 경제 코너에서는 외부인사들의 다양한 견해가 잘 전달된 것으로 보인다. 앞으로도 한은소식이 한국은행 안팎의 다양한 견해를 담아 계속 소통의 창구가 되어주길 기대해 본다.