

제4차 산업혁명과 소프트파워이슈리포트 2018-제18호

로봇저널리즘의 이해와 전망

윤인아 책임(twinkly24@nipa.kr), 홍보협력TF팀 기획평가단

2018. 5. 8.

목 차

- I. 로봇저널리즘 정의와 개념
- Ⅱ. 국·내외 로봇저널리즘의 사례
 - 1. 해외 활용 사례
 - 2. 국내 활용 사례
 - 3. 주요 이슈

Ⅲ. 시사점



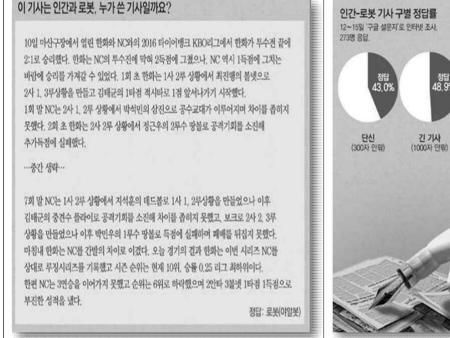
I . 로봇저널리즘 정의와 개념

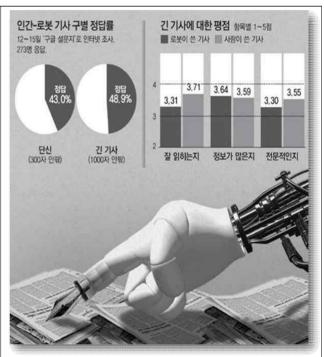
1 로봇저널리즘이란?

- □ 4차 산업혁명의 핵심기술인 인공지능(AI)은 저널리즘 영역에도 예외 없이 적용되어 '로봇 저널리즘'이라 불리는 AI 기사 생산 시대가 도래
 - **(로봇저널리즘)** 컴퓨터를 뜻하는 로봇(Robot)과 뉴스보도를 의미하는 저널리즘 (Journalism)의 합성어로
 - 뉴스 기사를 만드는 모든 과정에 컴퓨터 알고리즘이 관여, 미리 만들어진 알고 리즘을 통해 로봇(소프트웨어)이 자체적으로 정보들을 분류하고 정리한 뒤 의미를 해석해 스스로 기사를 작성하는 것
 - * 알고리즘 : 소프트웨어가 수행하는 일련의 작업의 방식과 규칙, 어떤 값이나 값의 집합을 입력받아 또 다른 값의 집합을 출력하는 잘 정의된 계산 절차 (Cormen, 2009)
 - (로봇 저널리즘의 시작) 뉴스를 생성하는 원가를 줄이고 인터넷 포털을 통해 페이지 조회 수를 늘리기 위한 방안으로 알고리즘이 데이터를 분석해 기사를 만들기 시작함 (Karlsson & Clerwall, 2013)
 - 캐나다의 경제정보업체인 톰슨 파이낸셜(Thomson Financial)의 알고리즘이 기업의수익과 관련된 기사를 자동으로 생성했으며 (Momus,2006), 이후 로봇 저널리즘이란 트렌드를 인식한 해외 언론사들은 이미 AI를 활발하게 사용하고 있음
 - 국내에서는 서울대학교 언론정보학과 hci+d lab에서 개발한 알고리즘이 프로 야구 경기 결과를 기사로 작성(2014), 국내 최초로 로봇저널리즘 연구가 이뤄 졌으며 이후 언론사에서도 로봇 저널리즘을 활용하기 시작함
- (로봇 저널리즘 활용 분야) 현재 로봇 저널리즘이 가장 많이 활용되는 분야는 데이터 값이 상대적으로 단순하고 분명한 금융, 증권, 스포츠, 기상 등이며 로봇이 작성할 때 빠른 속도와 정확성을 기대할 수 있음
 - (로봇 저널리즘2.0로의 진화) 4차 산업혁명이 도래하면서 지난 2~3년 간 로봇 저널리즘은 미디어 환경에 큰 변화를 가져왔으며, 언론사는 단순 기사 작성을 넘어서 생산성과 효율성을 개선하기 위한 연구를 지속적으로 수행

- 로봇이 쓰는 기사는 이미 독자들에게 깊숙이 자리 잡아 사람이 썼는지 로봇이 썼는지 구분하지 못하는 단계에 도달한 상황
- 동아일보의 '인간과 로봇, 누가 쓴 기사일까요? '라는 온라인 실험에 따르면 총 273명 응답, 정답률 45.9%로 절반 이상의 사람들이 로봇의 기사와 사람의 기사를 구분하지 못함

< 인간과 로봇, 누가 쓴 기사일까요? >





출처: 놀랍다, 로봇기자가 이런 표현까지 쓰다니 기사 발췌 (동아일보 2016.4.18.)

- 자동 기사작성 알고리즘인 '워드스미스'(Wordsmith)를 개발한 오토메이티드 인사이츠(Automated Insights)는 자동으로 단순 기사를 작성하는 것을 넘어 음성인식 인공지능과 결합, 음성으로 인간과 대화할 수 있는 서비스도 개발 중
- 또한 가짜뉴스와 진실에 기반을 둔 뉴스를 구분해주는 인공지능 기술도 빠르게 발전하고 있음
 - * 2016년 트럼프의 대통령 당선에 이른바'가짜뉴스(fake news)'가 지대한 영향을 미쳤다고 분석 (오하이오 주림대학 연구결과(2018.4))

□ 로봇저널리즘의 기사작성 과정

< 로봇 저널리즘의 기사작성 과정 5단계>



 서울대학교 연구팀에 의하면 로봇에 의한 자동 뉴스 작성은 대부분 ① 데이터 수집 ②이벤트 추출 ③ 중요이벤트 선별 ④기사의 분위기 결정 ⑤ 뉴스 기사 생성의 단계로 이루어짐

① 데이터 수집 (Date Crawling)

- 알고리즘을 통해 원시 데이터 (raw date)를 수집하는 단계로 수집된 대부분의 데이터는 분석에 적합한 형태가 아닌 경우가 많으며, 수집된 데이터를 정제하고 분석을 위한 적절한 형태로 변화하는 과정

② 이벤트 추출 (Event Extraction)

- 수집된 데이터의 분석이 이루어지는 단계, 텍스트 마이닝(text mining), 의미망 분석(semantic analysis) 등의 텍스트 분석 기법을 활용, 수집된 데이터에 의미를 부여하는 과정

③ 중요 이벤트 선별 (key Event Detection)

- 추출한 이벤트들의 중요도를 선별하는 단계로 통계적 분석 기법, 기계학습 등의 알고리즘을 통해 중요 이벤트가 무엇인지 선별함
- 스포츠나 금융정보와 같은 경우 상대적으로 통계적 분석 기법을 적용하기 용이 한 분야이기 때문에 초기 로봇 저널리즘 연구에서 많이 탐구됨

④ 기사의 분위기 결정 (Mood Detection)

- 중요 이벤트의 선별이 이루어진 후에는 이를 기반으로 기사의 분위기를 결정, 하나의 이야기에는 여러 관점이 포함될 수 있는데 이를 종합해서 화자가 글을 통해 전달하고 싶은 주제를 전달하는 방식으로 기사의 분위기를 만듬

⑤ 뉴스 기사 생성 (News Article Generation)

- 지금까지의 단계에서 처리한 데이터를 가지고 실제 문장을 생성해 기사로 탄생하는 마지막 단계, 이전 단계에서 선별한 중요 이벤트를 설명할 수 있는 적절한 문장을 선택해 이야기를 만들어냄

Ⅱ. 로봇저널리즘의 사례

1 해외 활용 사례

□ 내러티브 사이언스 '퀼(Quill)'

- 노스웨스턴대학교의 스탯몽키(Stats Monkey)*에서 출발한 회사(2010년)로 기사 작성 알고리즘인 '퀼(Quill)' 개발
 - * 스탯몽키 : 2009년 미국 노스웨스턴대 학생들의 연구 프로젝트, 야구통계를 이용한 경기 결과 자동 뉴스 작성
 - 기업의 실적, 금융, 스포츠 경기 결과 등 데이터를 입력하면 퀼 알고리즘에 따라 분석, 자동으로 뉴스를 작성하며 데이터 입력 후 작성까지 10초 이내 소요
- 주요 사례 : 포브스(Fobes)
 - 격주 마다 발행하는 미국의 경제 잡지로 2000년대 후반부터 경영난을 겪던 중 퀼을 활용(2012), 증권 시황, 스포츠 경기 결과 등의 기사 작성
 - 하루 평균 30여 건의 기사 작성으로 뉴스의 보도 범위가 확대되고 고품질의 뉴스 제공 및 수익 증가 효과 창출

□ 오토메이티드 인사이츠 '워드스미스 (Word smith)'

- 시스코 엔지니어 출신 로비 알렌(Robbie Allen)이 설립한 스탯시트에서 오토 메이티드 인사이츠로 사명 변경(2011)
 - 아마존 클라우드 기반의 플랫폼으로 스포츠를 넘어 모든 구조화된 데이터를 기반으로 수백만 개의 뉴스를 만들어 낼 수 있음
- 앞서 언급했던 것처럼 최근 자동으로 기사를 작성하는 것을 넘어 아마존의 음성인식 인공지능 '알렉사(Alexa)'와 결합, 음성으로 인간과 대화하며 데이 터의 의미를 분석할 수 있는 서비스 개발 중

ㅇ 주요 사례 : AP(Associated Press)

- 미국의 세계적인 통신사 AP는 매 분기마다 기업 발표 수치들을 뽑아내 300개 정도의 뉴스를 작성하던 중 워드스미스 플랫폼 도입 (2014.06)
- 워드 스미스 활용으로 분기당 평균 3,750개의 기업 실적 뉴스를 작성, 기존 뉴스 작성 대비 10배 이상 증가하는 효과를 누림



자료: (좌) Forbes홈페이지 캡처. (우) AP홈페이지 캡처

□ LA타임스 '퀘이크봇(Quakebot)'

- 미국 LA타임즈에서 활용, LA 주변에서 발생하는 지진정보를 실시간으로 수집 하고 자동으로 기사 작성
 - 미국지질조사국(USGS)에서 진도 3.0 이상의 지진 발생 정보가 송출되면 장소, 시간 ,지도, 해당 지역의 최근 지진 발생 내역을 포함해 기사 작성

□ 중국 텐센트 '드림라이터((Dreamwriter)' / '샤오난 (小南)'

- (드림라이터) 중국 최대 인터넷 기업인 텐센트가 개발, 중국 경제동향과 전문가 의 전망을 포함하고 있는 1000자 분량의 경제 기사를 1분 만에 작성해 자사 포털 사이트 '큐큐닷컴(qq.com)' 게재
- (샤오난) 중국 북경대 교수 '샤오 쥔(Wan Xiaojun)'이 개발, 일간지 남방도시보 에 기차역 승객 운송 현황에 관한 300자의 기사를 작성하여 게재
 - 편집부에서 제공한 예문, 열차 편명, 표 잔여수량 등의 수치를 기반으로 1초 만에 기사 작성

< '퀘이크봇(Quakebot), 샤오난(小南) 기사 예시 >



자료 : (좌) LA타임즈 홈페이지 캡처, (우) 1초 만에 300자 기사 작성한 기자로봇 발췌(아주경제 2017.1.19.)

2 국내 활용 사례

□ 프로야구로봇

- 서울대학교 언론정보학과 이준환 교수 연구팀 개발, 프로야구 정보를 실시간 으로 작성하여 페이스북을 통해 정보 제공(2015)
- 국내 스포츠기자들이 사용하는 승리로 장식했다, 안방에서 승리를 내주었다와 같은 표현을 활용, 실제 기자가 쓴 것과 같은 언어를 선택해 인간 기자가 쓴 기사라는 판단을 이끌어내기도 함
- 단순히 경기승패를 알려주는 기사에서 벗어나 어떤 선수가 결정적인 역할을
 했는지 어떻게 승패가 결정되었는지 주요 경기내용을 추출하여 작성

□ 파이낸셜뉴스 'IamFNBOT'

- 서울대학교 이준환·서봉원 교수 연구팀과의 협업으로 파이낸셜 뉴스를 통해 코스피 뉴스 기사 작성
 - 코스피 시황과 주가가 오른 업종, 내린 업종을 정확히 분석하며 국내 최초로 언론사를 통해 공식적으로 게재

< 프로야구로봇과 'IamFNBOT'의 기사 예시>



□ 테크홀릭 '테크봇'

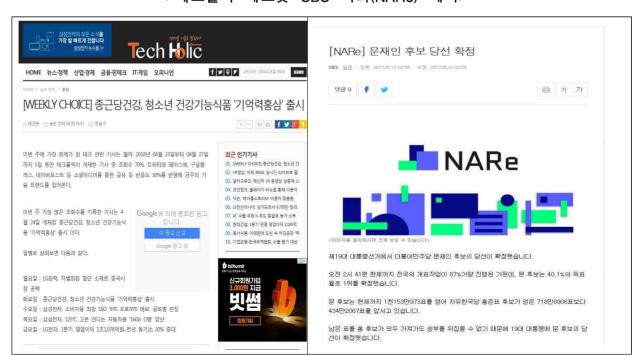
○ IT 관련 전문 미디어인 테크홀릭의 '테크봇' 알고리즘을 활용, 인기 조회 수와 소셜 반응 등 다양한 지표를 기준으로 매주 핫 이슈 기사를 자동으로 작성

□ 매일경제 '엠로보 (M-Robo)'

- 복잡한 투자 정보를 이해하기 쉽게 확인시켜주고 투자 효율을 높여주는 매일경제 미디어 그룹의 로봇저널리즘 프로그램
- 국내 상장 기업에 관련된 핵심뉴스를 텍스트 기사, 그림, 도표형식으로 표현해 기사 작성

□ SBS '나리 (NARe)'

- 2017년 대선 SBS 로봇기자, 인공지능 로봇기자의 약자로 SBS와 서울대 이준환· 서봇워 교수 공동 개발
 - 중앙선거관리위원회가 제 19대 대선 결과를 실시간으로 제공, 투개표 데이터를 자동적으로 계산하고 분석해 기사 작성

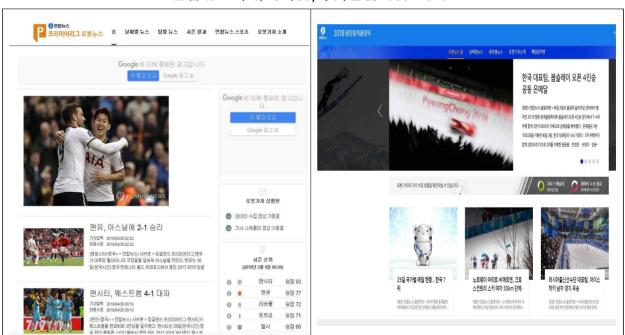


< 테크홀릭 '테크봇' SBS '나리(NARe)' 예시>

출처 : (좌) 테크홀릭 홈페이지 (우) sbs news 홈페이지

□ 연합뉴스 '싸커봇(Soccerbot)', '평창올림픽뉴스봇'

- (外커봇) 연합뉴스가 자체 개발한 알고리즘을 통해 영국 프리미어리그(EPL) 축구 전(全)경기의 데이터를 수집・분석한 뒤 경기 종료 후 1~2초 만에 기사 게재
 - 2017-2018 정규시즌에 치러지는 380건의 전체 경기를 취재해 작성 중임
- (올림픽뉴스봇) 프리미어리그에 이어 '올림픽봇'이 지난 평창 동계올림픽 취재 현장 에서 맹활약을 펼침
 - 올림픽 경기종료 후 기사 작성을 시작해 1~2초만에 웹사이트에 게재
 - 선수들의 결승전 진출 여부, 경기결과, 쇼트 트랙 금메달 결정 속보 등 기자가 작성한 기사보다 최소 1~2분, 길게는 1시간 이상 빨리 기사화함



< 연합뉴스 (좌)싸커봇. (우)올림픽봇 예시>

출처: 연합뉴스 홈페이지

□ 전자신문 'ET봇'

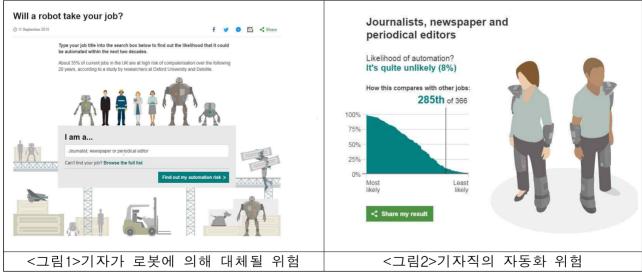
○ 전자신문의 증권 분석 로봇기자인 ET봇은 실시간으로 기업 매출이나 이익 증감률을 자동 계산하고 기업 공시를 실시간으로 처리, 투자의견까지 덧붙여 기사를 작성, 단순 기사 작성에서 더 나아가 증시 전반의 정보를 추출해 비교하거나 가공해 증시 분석 콘텐츠를 선보임

로봇저널리즘의 주요 이슈

□ 로봇이 기자를 대체 할 수 있는가?

- ㅇ 로봇저널리즘의 도입으로 기자와 미디어에 끼치는 영향이 점차 커지고 있는 상황에서 로봇이 기자를 대체할 수 있을지에 대한 논의가 지속되고 있으며 로봇 기자는 인간 기자 대체재가 아닌 조력자라는 반응 우세
- 영국 BBC 웹사이트에 'will a robot take your job? '의 제목으로 기사 게재 (2015.9.11)
 - 옥스퍼드대학교와 회계법인 딜로이트(Deloitte)의 공동 연구에 따르면 영국에 있는 일자리의 35%가 20년 안에 컴퓨터 자동화로 위협에 빠지게 될 것이 라고 발표함
 - * 결과도출 기준 : 일자리의 특성상 사회적 분별력, 협상력, 설득력, 타인에 대한 배려심, 독창성, 순수예술, 조작능력 등 핵심 스킬 요구 정도를 기준으로 결과 도출
 - 특정 직업을 입력하면 그 직업의 자동화 위험을 퍼센트로 환산해 순위를 보여준 결과 텔레마케터, 은행 창구 직원 등 낮은 수준의 상호작용이 요구 되는 직업은 자동화 위험률이 높아질 것으로 예상함
 - 그러나 기자직의 자동화 위험은 8%로 분석 대상이었던 366개 직업 중 285위 를 차지, 자동화 위험이 상대적으로 매우 낮은 직업에 속한 것으로 나옴

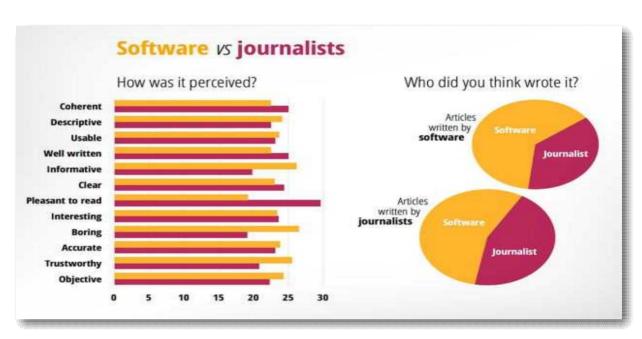
< 영국 BBC 'Will a robot take your job?' 기사 예시 > Will a robot take your job? Journalists, newspaper and f y O 🖸 < Share periodical editors



출처: http://www.bbc.com/news/technology-34066941

□ 로봇이 인간 수준의 글쓰기를 해낼 것인가?

- 기자 기사와 로봇 기사를 구분하기 위한 실험 결과 사람들은 기자 기사와 로봇 기사를 잘 구분해 내지 못하거나 품질 평가에서도 큰 차이가 나지 않는 것으로 나타남
 - 클러월(Clerwall,2014)의 연구에 따르면 소프트웨어가 작성한 글에 대해 총 12개 평가항목 중 절반이 넘는 7개 항목에서 기자의 글보다 우수한 것으로 평가



<로봇이 쓴 기사 vs 인간이 쓴 기사>

출처: Thomas Baekdal, https://onlinejournalismblog.com

- 서술적인(descriptive), 사용가능한(usable), 유익한(informative), 정확한(accurate), 지루한(boring), 신뢰할만한(trustworthy), 객관적인(objective) 항목에서 로봇 기사가 더 나은 평가를 받음
- 기자가 우위를 보인 항목은 총 5개로 일관적인(coherent), 잘 쓰인(well written), 보기 쉬운(clear), 읽기 좋은(pleasant to read), 흥미로운(interesting) 항목임
- 로봇 기사가 후한 점수를 받은 항목을 살펴보면 빠르고 정확한 연산 능력을 가진 컴퓨터 알고리즘의 특징이자 장점이 드러난 반면 사건의 배경이나 의미를 분석하는 등, 기자 본연의 임무에 가까운 항목에 대해서는 기자가 높은 평가를 받음

Ⅲ. 시사점

- □ 로봇저널리즘이의 기술 발달은 4차 산업혁명 시대의 핵심 화두로 등장하고 있는 인공지능(AI) 기술과 더불어 지속적으로 진화할 것으로 전망
 - 로봇 저널리즘은 향후 음성인식(speech recognition)과 이미지 및 패턴 인식 (face recognition), 자동 번역(automated translation) 기술 등과 결합
 - 인공지능(AI) 기술은 입력된 데이터를 논리적 연산(reasoning) 과정을 거쳐 기사를 생성하는 데 그치는 것이 아니라 센서를 통한 지각(perception) 능력 까지 갖춰 부분적으로는 취재의 영역까지 담당할 수 있게 될 것
 - 특히, 자동 번역 기능과 패턴을 인식할 수 있기 때문에 국제 뉴스 작성도 가능하며 현재와 같은 텍스트 중심의 단순한 자동 기사 작성에 머물지 않고 관련이미지와 동영상이 덧붙여진 기사 생성을 할 수 있을 것으로 보임
- □ 로봇이 기자를 대체할 수 있을지에 대한 우려 속에서도 로봇저널리즘은 기자의 업무를 보완해주는 방향으로 나아 갈 것
 - 스포츠 경기, 주가 동향, 재난 정보와 같이 숫자를 다루거나 통계를 기반으로 정확한 정보를 빠르게 전달해야 하는 경우 로봇 저널리즘이 큰 역할을 수행
 - 팩트 전달 등과 같이 정형화된 기사는 로봇저널리즘으로 대체, 기자는 보다 생산 적이고 심층적인 분석이 필요한 기사 작성 업무에 집중할 수 있도록 도와줌 으로써 업무의 효율성을 높일 수 있음
- 소셜 데이터와 개인의 활동 데이터 분석을 통해 개개인에게 더 의미 있는 뉴스의 발굴이 가능해져 맞춤형 정보 제공도 수월해 질 것
 - 금융 분야의 경우 기자가 개인에게 맞는 투자 정보를 일일이 제공하는 것은 어렵지만 맞춤형 정보 제공이 가능해짐으로 인해 기자뿐 아니라 개개인, 또한 그 분야의 전문가들에게도 도움이 될 것으로 전망
 - 최근 로봇 저널리스트뿐만 아니라 로봇 에디터(데스크)까지 도입한 미국 통신사 AP는 로봇 기사 프로그램 도입을 통해 기자의 생산성이 20% 정도향상됐다고 밝힘

- □ 알고리즘의 기술적 한계로 인한 정보의 오류, 알고리즘의 객관성 등은 해결해야 할 과제
 - 알고리즘 역시 인간이 짜놓은 규칙이기 때문에 온전히 공정할 수 없고 쏟아 지는 데이터의 타당성에 대한 판단을 내리는 일은 무리이며
 - 미디어의 디지털화와 로봇 저널리즘이 비용 절감, 정확성 향상, 콘텐츠 다양화 등 산업계에 혜택을 주는 것은 사실이지만 알고리즘을 통한 기사 작성이 가지는 정확성과 신뢰성에 대한 문제를 어떻게 해결할 것인지에 대한 고민 필요
 - 로봇이 기자를 '대체한다, 못 한다'의 관점이 아닌 로봇저널리즘이 언론에 주는 시사점과 기자의 미래에 미칠 영향에 대한 검토가 필요한 시점

<참고 문헌>

- 1) 로봇저널리즘 가능성과 한계(한국언론진흥재단, 2015)
- 2) 로봇저널리즘: 알고리즘을 통한 스포츠 기사 자동 생성에 관한 연구 (서울대학교 김동환, 이준환)
- 3) 집중점검, 로봇저널리즘 현황과 전망 (이준환, 김동환 ,김영주, 정재민, 오세욱)
- 4) 데이터분석·기사 작성은 기본, 가짜 기사도 판별 (매일경제, 손재권)
- 5) 4차 산업혁명과 미디어정책 (한국언론진흥재단, 2017 김선호, 박아람)
- 6) 인공지능 기사 5인방을 소개합니다 (아시아경제, 2017.3.28.)
- 7) Robot reporter gets its first news article published (The Economic Times, 2017.5.29.)
- 8) 로봇저널리즘, 기사 정확성, 오류 감소에 도움 (연합뉴스, 2017.9.27.)
- 9) The New York Times Built a robot to help make article tagging easier (NimanLab, 2015.7.30.)
- 10) 중국에도 로봇 신문 기사 등장 (전자신문, 2017.1.20.)
- 11) People Think Computer Journalists Are More Trustworthy Than Human Ones (motherboard, 2014.3)
- 12) 로봇 저널리즘 국내 실험 사례(다독다독,2015.12.7.)
- 13) 저널리즘에도 4차 산업혁명이 시작됐다(일간투데이, 2017.3.31.)
- 14) 1초만에 300자 기사 작성한 기자로봇(아주경제, 2017.1.19.)
- 15) 인공지능 시대, 기자와 언론의 미래는 (매일경제, 손재권)
- 16) Welcome to your automated news future (blog.wan-ifra.org, 2015.3.10.)
- 17) 이제 로봇이 기사도 쓴다 (SK hynix, 2017.09.13.)

- ※ 본 이슈리포트의 내용은 무단 전재할 수 없으며, 인용할 경우, 반드시 원문출처를 명시하여야 합니다.
- ※ 본 자료의 내용은 필자의 개인의견으로 정보통신산업진흥원(NIPA)의 공식견해가 아님을 밝힙니다.