 <b>한국과학기술정보연구원</b> <small>Korea Institute of Science and Technology Information</small>	<h1>보도자료</h1>	<a href="http://www.kisti.re.kr">http://www.kisti.re.kr</a>
		<b>배포 즉시 보도 가능합니다.</b>
대전(본원): 대외협력실 이종성 042 - 869 - 0976 / 최영진 0947 <b>문의: 미래기술분석센터 윤진혁 선임연구원 (02-3299-6061 / 010-5295-4152)</b>		
배포번호 : 2018-50 배포일자 : 2018.12.19	매수 : 보도자료 3매	배포처 : 대외협력실

## 집단지성 형성과정의 비밀, 빅데이터로 풀어낸다

- 온라인, 오프라인 집단지성 모두 소수 저자의 독점적 영향력이 증가 -
- 규제 없는 지식생성 과정에서 소수의 여론 독점 경향성 우려 -

- 온라인 환경이 가짜 뉴스, 여론 조작 같은 공격에 왜 취약점을 보일까?  
인터넷상에는 누구나 의견을 자유로이 표출 가능한 것처럼 보이지만 소수의 힘으로 여론을 쉽게 독점하는 것이 가능하다는 것이 밝혀졌다.
- 한국과학기술정보연구원(원장 최희윤, 이하 KISTI)은 경남과학기술대학교(이하 경남과기대), 한국과학기술원(이하 KAIST) 연구진과 함께 인터넷 백과사전, 논문, 특허 등의 진화 양상을 빅데이터 분석으로 집단지성(용어설명1) 형성 과정의 규칙성을 밝혀냈다고 밝혔다. 또 이를 바탕으로 모든 집단지성에서 지속적으로 소수 기여의 영향력이 커지는 지식의 독점화가 점점 심해지고 있다는 진단을 내놓았다.
- 연구 결과는 18일 세계적인 과학저널 네이처(Nature)의 자매지 네이처 인간행동(Nature Human Behaviour) 온라인 판에 실렸다.
- 최근 인공지능과 빅데이터가 두각을 나타내며 인간 지성 형성 과정의 데이터 기반 분석 시도가 증가하고 있으나 지식 형성 과정의 장기성과 복잡성으로 인해 데이터가 방대하고 분석이 어렵다. 이로 인해 기존에는 소규모의 표본을 통해 전체적인 경향성을 추정하였으며, 전체 데이터의 분석을 시도하지 못했다.

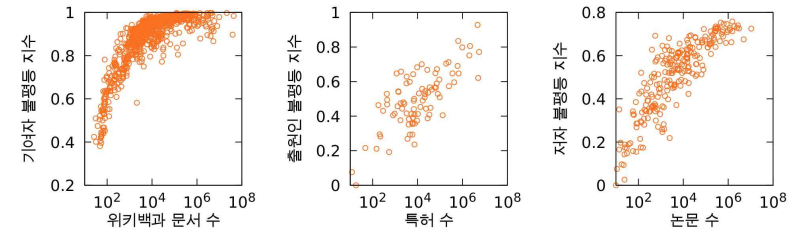
- KISTI 미래기술분석센터 윤진혁 선임연구원, 경남과기대 교양학부 이상훈 교수, KAIST 물리학과 정하웅 교수로 구성된 연구팀은 복잡성이 높은 데이터에서 규칙성을 찾기 위해 복잡계(complex systems) 방법론을 도입하여 대규모 집단지성 분석을 시도하였다.
- 연구팀은 먼저 273개 언어로 쓰여진 863개 위키미디어 프로젝트 각각의 성장을 측정하여 변화 양상을 분석하였다. 그 결과, 모든 데이터에서 시간이 지날수록 성장 속도가 둔화되고 있음을 발견하였으며, 신규 기여자의 유입이 점점 줄어드는 것에서 그 원인을 찾았다.
- 이러한 현상이 독점의 영향이라는 가설을 세우고, 기여자들 사이의 기여 불평등을 정량화해 불평등 지수(Gini index)의 변화를 추적하였다. 그 결과, 지식이 축적될수록 지식 생성의 불평등 지수가 높아졌으며, 소수의 독점적 영향력이 증가해 기여자의 행동을 대부분 지배하는 ‘독점화 현상’을 발견하였다. 이러한 독점 집단은 집단지성 생성 초기에 나타나 현재까지 지속적으로 독점적 영향력을 미치며, 신규 기여자가 이런 독점 계층에 진입하는 것이 거의 불가능하다는 것도 밝혀냈다.
- 연구진은 이러한 불평등의 창발 현상(용어설명3)을 재현하는 행위자모형 또한 개발하였는데, 모형을 통해 지식이 축적될수록 미래의 집단지성 지식 독점이 더 심해지는 현상이 뚜렷하게 예측되었다. 현 상태를 방지한다면 온라인상의 정보가 소수에 의해 독점되어 왜곡 될 수 있다는 것이다.
- 연구팀은 논문과 특허에서도 독점화 현상이 나타난다는 것을 확인하였다. 논문과 특허가 많이 발표된 국가들일수록 소수 연구자에 더 의존한다는 것이다. 신규 참여가 더 쉽다고 알려진 인터넷 백과사전이 상대적으로 폐쇄적인 것으로 알려진 논문과 특허에 비해서 더 독점화 경향이 강하게 나타나는 역설적 현상 또한 발견하였는데, 이는 진입장벽과 규제가 없는 경우에 더 쉽게 소수가 독점할 수 있음을 암시한다.

□ 경남과기대 이상훈 교수는 “인류가 점점 압도적으로 많은 비율의 정보를 온라인/디지털 미디어에서 얻게 되었는데, 그것은 그 정보의 형성 및 취득 과정 또한 매 순간 모두 기록(digital footprint)으로 남는다는 의미이기도 하다” 는 점을 강조하였고, “이제는 지금까지 주로 수동적으로 참고해왔던 그러한 정보의 모든 기록을 분석하고 근본 원리를 파악하여, 장점은 부각시키고 바로잡을 것은 바로잡을 때가 되었다. 우리의 연구가 그러한 흐름의 좋은 예시가 되었으면 한다.” 는 희망을 밝혔다.

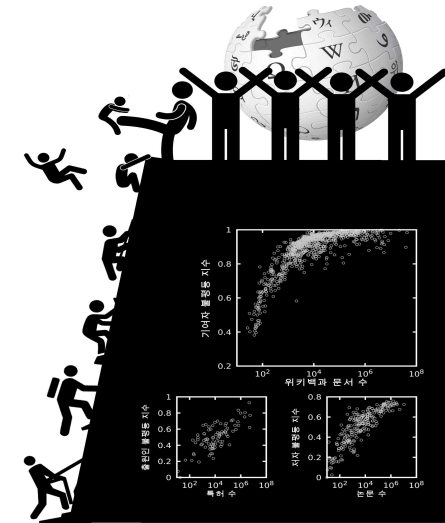
□ KISTI 윤진혁 박사는 “복잡계 과학을 통한 새로운 접근법과 빅데이터를 통해 인간 지성과 여론 형성 과정의 구조를 파악할 수 있었다” 며, “대부분의 정보가 인터넷에서 유통되는 현재, 풍성하고 협동적인 환경을 유지하고 소수의 독점화를 줄이려면 새로운 참여자들의 적극적 활동을 지원하고 독과점에 대해 제재할 필요가 있다” 고 제언했다.

□ 이번 연구는 과학기술정보통신부 산하 한국연구재단 전략공모사업, 신진연구자지원사업 및 중견연구자지원사업 지원으로 수행됐다.

별첨 : 연구결과 개요 및 이력사항



<위키백과와 논문, 특허 모두에서 축적된 지식의 양이 많아질수록 소수 저자의 영향력이 강해지는 현상을 관측하였다. 위키백과 불평등 지수는 0.9 이상으로 매우 높으며, 대부분 0.8 이하의 불평등지수를 보이는 논문이나 특허에 비해서 매우 높다.>



<위키미디어 2억 7천만여건, 스코푸스(SCOPUS) 등재 논문 4천만 건, 세계 출원 및 등록 특허 9천만 건의 집단지성을 분석한 결과 위키백과, 논문, 특허 모두에서 축적된 지식의 양이 많아질수록 소수 저자의 영향력이 강해지는 현상을 관측하였다. 위키백과 불평등 지수는 0.9이상으로 매우 높다. 오랜 기간 위키백과는 전세계 누구나 지식을 공유할 수 있는 이상향적 모습으로 그려져 왔다. 하지만 실제 연구결과 소수 기자가 거의 독점적으로 콘텐츠를 생산하고 있으며, 새로운 유저는 이런 독점적 유저의 카르텔에 진입할 수 없다.>

## 용어 설명

### 1. 집단지성

- 하나의 개체가 아닌 여러 개체들이 협력과 경쟁을 통하여 지식을 축적하는 과정과 결과를 의미한다.
- 개인 능력의 합보다 더 큰 능력을 발휘하는 창발 현상(emergence)이 나타나며, 대표적인 예로 위키백과, 오픈소스 프로젝트 등을 들 수 있다

### 2. 위키미디어 프로젝트

- 위키미디어 프로젝트는 집단지성을 통해 온라인 상에 누구나 무료로 자유롭게 사용할 수 있는 지식 공유의 장을 만드는 프로젝트이다.
- 캘리포니아 샌프란시스코의 비영리재단인 위키미디어 재단 (Wikimedia Foundation, Inc.) 에서 운영 중이며, 위키백과, 위키낱말사전, 위키책, 위키뉴스 등 900개 이상의 프로젝트를 운영하고 있다.

### 3. 창발 현상

- 개별 구성 요소에서는 발생하지 않는 현상이 구성 요소가 구조화되었을 때 자발적으로 발현되는 현상을 뜻한다.
- 주식시장의 흐름 형성, 대도시의 계층 구조, 인터넷의 멱함수(power-law) 법칙과 허브의 존재 등 사회를 구성하는 많은 요소들에서 나타나며, 외부의 영향 없이도 자기조직화(self-organization)를 통하여 시스템 스스로 변화하는 특징이 관측된다.

## 연구결과 개요

Early onset of structural inequality in the formation of collaborative knowledge in all Wikimedia projects

(Nature Human Behaviour, 온라인 게재일: 한국시간 2018. 12. 18.  
doi:<https://doi.org/10.1038/s41562-018-0488-z>)

### □ 연구의 중요성 및 차별성

위키백과는 지난 20여년간 인터넷 상 최대의 지식 창구로서, 사용자의 자발적인 기여를 통해 집단지성으로 발전하는 이상향적인 모습으로 알려져 있다. 하지만 최근 다양한 사례와 연구결과들은 집단지성이 지닌 취약성으로 인해, 이러한 위키백과가 소수 집단의 독점이 가능하며, 실제 영문 위키백과가 극소수에 영향력에 좌우되고 있다는 현상을 보고하였다 (참고자료 1, 2). 하지만 이러한 연구는 대부분 영어 위키백과 한가지에 국한되어 있었으며, 실제 독점현상이 일반적인 집단지성의 형성과정에도 모두 적용되는지에 대한 근거를 제시하지 못하였다.

본 연구에서는 위키백과를 포함하는 위키재단의 863개 프로젝트, 전세계 200여국의 논문 및 특허에 대한 빅데이터 분석을 통하여 세계 최초로 이러한 현상의 완전한 이해에 성공하였다. 본 연구진의 발견은 (1) 위키백과 및 모든 온라인 상 집단지성은 이러한 영향력 독점화에 취약하며, (2) 이러한 독점현상에 의하여 위키백과 등 집단지성 성장이 점점 비효율적이 되어가며, (3) 이러한 집단지성의 독점화는 기존의 대표적인 오프라인 지식체계인 논문과 특허에서도 나타나나, 그 정도는 더 쉽게 수정 가능한 온라인 미디어가 더 크게 나타나며, (4) 신규 참여자들이 이미 존재하는 독점자가 만든 사회 시스템에 융화되는 과정을 통해 이러한 독점 현상이 나타날 수 있다는 것을 증명하였다. 이 발견은 개인의 소셜미디어 등으로 누구나 여론 형성에 참여할 수 있는 현대사회가 오히려 과거보다 여론의 조작과 독점이 쉬우며 자유로운 온라인 환경이 가짜뉴스나 여론조작 등에 더욱 취약할 수 있다는 것을 나타낸다. 결론적으로 본 연구는 빅데이터를 이용하여 인간 지성 형성에 대한 이해를 높였으며 또한 최근 문제가 되고 있는 가짜뉴스나 여론조작 등 온라인 상의 사회적 문제 해결을 위한 기반 지식과 근거를 제시하였다.

## □ 기타

- 본 연구 그룹 (KISTI, 경남과기대, KAIST 공동연구진)은 지난 수년간 집단지성의 성장과정을 빅데이터로 분석하는 분야의 국제적인 선구자임.
  - 2016년 발표한 “Intellectual interchanges in the history of the massive online open-editing encyclopedia, Wikipedia” (Phys. Rev. E. 2016)의 경우 Washington Post, Gizmodo등 10여개국 20여개 이상의 내·외신에 의하여 소개되었으며, 미국물리학회 Physics Focus에 주목할만한 논문으로 소개.
  - 이외에 “Build Up of a Subject Classification System from Collective Intelligence” (Sae Mulli, 2018), “Anatomy of Scientific Evolution” (PLOS ONE, 2015) 등의 연구 성과를 지속적으로 발표.
- Nature Human Behaviour (네이처 인간행동): Springer Nature에서 발매하는 Nature 자매지로, 심리학, 사회과학, 자연과학등의 융합을 통하여 인간 행동과 그 심리학적, 생물학적, 사회학적 기반 등 모든 측면에 대한 우수한 연구 결과를 출판하는 학술지. 인간의 인지, 상호작용, 능력의 발전 및 퇴보, 집단과 개인의 진화과정, 문화의 발전과정 등의 융합 과학적 방법 및 연구결과에 대해 주로 게재함 (참고자료 3, 4).
- Springer Nature의 의뢰로 논문 소개글을 Nature의 온라인 미디어 *Behavioural and Social Sciences at Nature Research*에 기고 (참고자료 5).

## □ 참고자료

- Washington Post, 2016-04-28, Wikipedia’s social structures resemble a bureaucratic corporation, studies say.
- APS Physics, 2016-01-22, Focus: Wikipedia Articles Separate into Four Categories.
- Nature Human Behaviour, <https://www.nature.com/nathumbehav/>.
- Nature Human Behaviour 소개자료, <https://media.springernature.com/full/springer-cms/rest/v1/content/12229100/data/v5>.
- How I met my collaborators, <https://go.nature.com/2rHnvC4>.

## 연구진 이력사항 (윤진혁 박사)

### 1. 인적사항

- 소속 : 한국과학기술정보연구원 (KISTI)  
미래기술분석센터

### 2. 학력

- 2010 ~ 2016 : 한국과학기술원 물리학과 (박사)
- 2005 ~ 2010 : 한국과학기술원 물리학과 (학사)

### 3. 경력사항

- 2016 ~ 현재 : KISTI 미래기술분석센터 선임연구원
- 2016 ~ 2016 : Naver Search 연구원

### 4. 전문분야정보

- 빅데이터, 복잡계, 네트워크과학, 통계물리학, 과학계량학



## 연구진 이력사항 (이상훈 교수)

### 1. 인적사항

- 소속 : 경남과학기술대학교 (GNTECH) 교양학부



### 2. 학력

- 2004 ~ 2010 : 한국과학기술원 물리학과 (박사)
- 2000 ~ 2004 : 한국과학기술원 물리학과 (학사)

### 3. 경력사항

- 2018 ~ 현재 : 경남과학기술대학교 교양학부 조교수
- 2018 ~ 현재 : 고등과학원 (KIAS) Associate Member
- 2015 ~ 2018 : 고등과학원 (KIAS) Research Fellow
- 2014 ~ 2015 : 성균관대학교 에너지과학과 BK21+ Research Professor
- 2012 ~ 2014 : University of Oxford Postdoctoral Research Associate
- 2010 ~ 2012 : Umeå University Postdoctoral Researcher
- 2010 ~ 2010 : 한국과학기술원 Postdoctoral Researcher

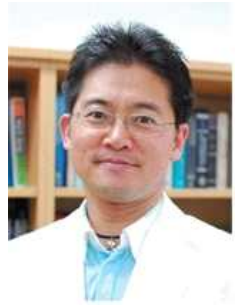
### 4. 전문분야정보

- 통계물리학, 네트워크 과학, 복잡계 및 빅데이터

## 연구진 이력사항 (정하웅 교수)

### 1. 인적사항

- 소속 : 한국과학기술원 (KAIST) 물리학과



### 2. 학력

- 1993 ~ 1998 : 서울대학교 물리천문학과 (박사)
- 1991 ~ 1993 : 서울대학교 물리천문학과 (석사)
- 1987 ~ 1991 : 서울대학교 물리천문학과 (학사)

### 3. 경력사항

- 2001 ~ 현재 : 한국과학기술원 물리학과 교수
- 1998 ~ 2001 : University of Notre Dame Research Assistant Professor
- 1998 ~ 1998 : 서울대학교 이론물리학연구소 (CTP) Postdoctoral Researcher
- 1997 ~ 1997 : KFA HLRZ Visiting Researcher

### 4. 전문분야정보

- 통계물리학, 복잡계, 네트워크 및 빅데이터