

# 지식 공유의 파레토 비율 및 불평등 정도와 가상 지식 협업: 위키피디아 행위 데이터 분석\*

박현정

이화여자대학교 경영연구소  
(sparrowpark@ewha.ac.kr)

신경식

이화여자대학교 경영대학  
(ksshin@ewha.ac.kr)

전체 결과의 80%가 전체 원인의 20%에 의해 일어난다는 파레토 법칙(Pareto principle)은 상위 20%의 핵심 고객에 대한 우선적인 마케팅을 비롯하여 기업 경영의 많은 부분에서 적용되어 왔다. 파레토 법칙과는 대조적으로, 80%의 사소한 다수가 20%의 핵심적인 소수보다 우월한 가치를 창출한다는 롱테일 법칙(Long Tail theory)은 ICT(Information and Communication Technology)의 발전과 함께 새로운 경영 패러다임으로 주목 받아오고 있다. 본 연구의 목적은 경영 현장에서 양대 흐름을 형성해온 이러한 법칙들이 변화무쌍한 글로벌 가상화 환경에서 기업의 핵심적인 성공 요인이라고 할 수 있는 가상 지식 협업에는 어떻게 관련되는지를 규명하는 것이다. 이를 위해, 대표적인 가상 지식 협업 커뮤니티인 위키피디아에서 품질 최상위 등급인 피쳐드 아티클(Featured Article) 레벨로 승급된 2,978개의 아티클에 대한 협업 행위를 분석하였다. 즉, 각 아티클 그룹에서 편집 횟수 기준 상위 20%에 속하는 참여자들의 총 편집 횟수가 전체 편집 횟수에서 차지하는 비율인 파레토 비율(Pareto ratio)이 지식 협업 효율성과 어떤 관계를 가지고 있는지를 도출하였다. 그리고, 이러한 연구를 편집 참여를 통한 지식 공유에 대한 전체적인 불평등 정도를 나타내는 지니 계수(Gini coefficient)의 영향 및 그룹의 작업 특성을 반영하도록 확장하였다. 결과적으로, 지식 공유의 파레토 비율과 지니 계수가 증가하면 지식 협업 효율성도 높아지지만, 이러한 변수들이 일정 수준 이상으로 증가하면 오히려 지식 협업 효율성이 낮아지는 역 U자(inverted U-shaped) 관계가 있음을 확인하였다. 그리고, 이러한 관계는 인지적 노력을 상대적으로 더 많이 요구하는 학문적인 특성의 작업에서 더 민감하게 작용하는 것으로 보인다.

**주제어** : 가상 지식 협업, 지식 공유, 파레토 비율, 지니 계수, 온라인 커뮤니티

논문접수일 : 2014년 8월 6일    논문수정일 : 2014년 8월 24일    게재확정일 : 2014년 8월 27일  
투고유형 : 국문일반    교신저자 : 신경식

## 1. 서론

이탈리아의 경제학자이자 사회학자인 빌프레도 파레토(Vilfredo Pareto)는 1906년에 이탈리아 인구의 20%가 이탈리아 국토의 80%를 보유하고 있음을 발견하였다. 경영 컨설턴트인 조지프 주란(Joseph Moses Juran)이 사회현상에서 발견되

는 이러한 경험적인 법칙을 파레토 법칙(Pareto principle)이라 명명하였고, 이를 일반화시켜 20%의 주요문제를 해결하면 나머지 80%는 저절로 해결된다는 ‘주요한 소수와 사소한 다수(the vital few and the trivial many)’라는 개념을 주장하였다. 80대20 법칙이라고도 불리는 파레토 법칙은 20%의 충성스러운 고객들이 전체 매출의 80%를

\* 이 논문(또는 저서)은 2013년도 정부재원(교육부)으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013S1A3A2054667).

창출하고, 20%의 핵심제품이 기업 전체 매출의 80%를 차지하며, 20%의 기업구성원이 전체 업무의 80%를 수행하는 현상 등이 발견되면서 비즈니스의 황금 룰로 인식되어 왔다. 특히, 마케팅 분야에서는 매출의 대부분을 차지하는 상위 소수 소비자들을 공략하는 마케팅 전략이 활성화되기도 했다.

이와는 대조적으로, 2004년 미국의 인터넷 비즈니스 관련 잡지 와이어드(Wired)의 편집장 크리스 앤더슨(Chris Anderson)에 의해 명명된 롱테일 법칙은 80%의 ‘사소한 다수’가 20%의 ‘핵심적인 소수’보다 더 많은 가치를 창출한다는 이론으로, 역파레토 법칙이라고도 불린다 (Anderson, 2004; 2006). 인터넷의 발전에 따라 물리적인 거리 및 시간의 한계를 초월하여 대규모 고객 집단에 대한 접근이 용이해지고 전신비용이나 물류비용이 획기적으로 감소하면서 유통구조의 혁신이 이루어졌다. 소비자들은 적은 탐색 비용으로 훨씬 다양한 상품에 대한 정보를 획득할 수 있고 다른 소비자들과 쉽게 정보를 공유함으로써 과거에는 상상하지 못했던 여러 가지 이점을 누릴 수 있게 되었다. 이러한 변화에 따라 롱테일 마켓의 잠재력이 부각되면서 이베이, 야후, 구글, 아마존, 아이튠즈(iTunes) 스토어 등 인터넷 비즈니스 기업들이 롱테일을 비즈니스 전략으로 채택하여 성공을 거두고 있다. 대표적인 온라인 서점 아마존의 전체 수익 가운데 절반 이상은 오프라인 서점에서는 잘 팔리지 않아 진열대에 비치하지도 않는 비주류의 희귀한 책들에 의해 창출되고, 인터넷 포털 구글의 주요 수익원은 포춘 500대 기업으로 선정된 대기업들이 아니라 꽃 배달 업체나 제과점 등 소규모 광고주라는 사실은 이미 널리 알려진바 있다. 롱테일 현상은 TOD(Television on demand)나 페이퍼뷰

(pay per view) 서비스와 같이, 소외되었던 롱테일 일에 위치해있는 소수의 기호를 발굴함으로써 상품의 다양화를 촉진하고 다양한 공급자에게 기회를 제공함으로써 문화의 다양성에도 영향을 미치고 있다(Anderson, 2005). 또, 다수의 롱테일 유저들이 ICT(Information and Communication Technology)로 연결됨으로써 유저 주도 혁신(user driven innovation), 집단지성(collective intelligence), 크라우드소싱(crowdsourcing), 크라우드캐스팅(crowdcasting) 등 소셜 네트워크 메커니즘에 의한 생산 모델이 가능하게 되었다.

본 연구는 이와 같이 경영 현장에서 양대 흐름을 형성하고 있는 파레토 법칙과 롱테일 법칙이 가상 지식 협업의 성공에는 어떻게 관련되는지를 규명해보고자 하는 목적에서 출발하였다. 경영 환경의 글로벌화와 가상화에 따라 가상 지식 협업의 중요성은 빠르게 증가하고 있지만, 이에 대한 체계적인 연구는 아직 부족한 실정이다. 지식 공유는 장려할 수 있을 뿐 강요할 수 없는 특성을 가지고 있다 (Gibbert and Krause, 2002). 또, 개인 행동의 합리성(rationality) 가정에 의하면 지식을 공유하기 보다는 다른 사람들이 공헌한 지식을 사용하기만 할 가능성이 높는데, 모든 사회 구성원이 이렇게 무임승차하는 대안을 선택하면 결국 개인의 효용이 감소하게 되는 소셜 딜레마 (Dawes, 1980; Kollock, 1999) 문제가 발생한다. 이러한 이유로 개인의 지식 공유를 촉진하려는 목적에서 지식 공유를 유발하는 요인들에 대한 많은 연구들이 수행되었다 (Bock and Kim, 2002; Bock et al., 2005; Hsu et al., 2007; Kankanhalli et al., 2005; Lin, 2001; Wasko and Faraj, 2005; Yang and Lai, 2010).

그런데, 여러 사람들이 함께 참여하여 이루어지는 지식 협업은 지식 공유뿐만 아니라 지식의

축적(accumulation), 변환(transformation) 및 통합(integration) 등을 통한 지식의 창조 과정이라고 정의할 수 있다 (Faraj et al., 2011, Grant, 1996). 이러한 그룹 차원의 관점에서 보면, 지식 공유 양(quantity) 기준 상위 20%나 하위 80%가 수행하는 역할이나 참여자들간의 상대적인 지식 공유 분포도 절대적인 개인의 지식 공유 양 이상으로 지식 협업에서 중요한 의미를 가지고 있을 가능성이 높다. 따라서, 본 연구는 다음과 같은 질문들에 대한 답을 찾고자 한다.

- 지식 협업에서는 상위 20%나 하위 80%의 참여자들 중 어느 편이 더 중요한 역할을 담당하고 있을까? 또는 상위 20%에 속하는 참여자들도 더욱더 많은 지식을 공헌할수록 그룹의 전체적인 지식 협업 효율성은 계속 좋아지는 것일까?
- 그룹 구성원들의 상대적인 지식 공유 불평등 정도는 그룹 협업의 성공에 어떤 영향을 미칠까?
- 협업 작업의 인지적 또는 지식 특성에 따라 이러한 관계는 어떻게 달라질까?

이를 위해 대표적인 가상 지식 협업 커뮤니티인 영어 버전 위키피디아에서 아티클 최상위 등급인 피쳐드 아티클(Featured Article) 레벨에 도달한 2,978개의 아티클 협업 행위 및 아티클 등급 데이터를 분석하였다. 즉, 참여자들의 지식 공유 분포를 포함하여 아티클 협업 프로세스의 효율성에 영향을 미칠 수 있는 주요 변수들과 해당 아티클이 생성된 이후 피쳐드 아티클 레벨로 승급 될 때까지 소요된 시간과의 관계를 도출하였다.

본 논문의 내용 구성은 다음과 같다. 2장에서는 위 질문들과 관련된 기존연구들에 대해 살펴볼 것이다. 3장에서는 온라인 가상 환경에서 지

식 공유의 파레토 비율 및 불평등 정도가 지식 협업의 효율성에 어떻게 작용할지, 그리고 이러한 관계가 작업 특성에 따라 어떻게 달라질지에 대한 가설을 도출할 것이다. 4장에서는 가설에 대한 분석 결과를 정리하고 5장에서 이론 및 실제 적용을 위한 시사점과 연구의 한계점을 제시할 것이다.

## 2. 관련 연구

### 2.1 가상 지식 협업(Virtual Knowledge Collaboration)

지식 공유는 자신의 지식을 타인과 기꺼이 공유하려는 정도에 의해 결정된다 (Gibbert and Krause, 2002). 지식 공유를 지식을 공헌하는 측면과 사용하는 측면으로 나누어 생각해볼 수 있지만, 지식의 사용은 충분히 자발적으로 이루어지기 때문에 대부분의 지식 공유 관련 연구들은 지식 공헌을 다루고 있다. 지식 협업은 앞에서 제시한 바와 같이, 지식 공유를 포함하는 지식의 창조 행위라고 할 수 있다. 가상 지식 협업은 다양한 커뮤니케이션 틀들에 의해 지원되는 상호작용을 통해 가상 환경에서 이루어지는 지식 협업을 지칭한다고 할 수 있다. 지식 협업 프로세스의 성공 요인(success factor)은 조직의 지식 창조 프로세스를 설명하는 이론적인 프레임워크/framework)로 잘 알려진 조직 지식 창조 이론(organizational knowledge creation theory)을 통해 살펴볼 수 있다. 이 이론에서는 새로운 지식의 생성을 촉진하는 5가지 요인을 의도(intention), 자율성(autonomy), 환경 변화(fluctuation) 또는 창조적 혼돈(creative chaos), 정보의 중복(redundancy

of information), 필수적 다양성(requisite variety or minimax internal diversity)으로 제시하고 있다 (Nonaka, 1994; Nonaka et al., 1996). 이 중 의도, 자율성, 환경 변화 또는 창조적 혼돈은 지식 창조의 필요성을 제공하고 개인의 헌신(commitment)을 유도하는 중요한 역할을 한다. 정보의 중복은 조직 구성원들이 자기 분야가 아닌 다른 분야에 대해서도 알고 이해하는 부분이 있어야 지식 창조 프로세스가 촉진된다는 것을 의미한다. 또, 필수적 다양성은 환경 변화에 적응하기 위한 새로운 지식을 창조하기 위해서는 조직 내부의 다양성이 높을수록 효과적인 반면에, 조직의 통합(integration)을 위해서는 이질성으로 작용할 수 있는 다양성이 낮을수록 바람직하기 때문에 결국 적정 수준의 다양성이 좋다는 의미를 내포하고 있다 (Nonaka, 1994; Nonaka et al., 1996).

## 2.2 지식 공유 파레토 비율(Pareto ratio)과 지니 계수(Gini coefficient)

지식 협업에서 지식을 공유한 양(quantity)을 기준으로 볼 때 상위 20%의 참여자들이 공헌한 지식의 양이 전체 지식에서 차지하는 비율, 즉 파레토 비율을 계산할 수 있다. 그리고, 소득의 불평등 정도를 나타내는 지니 계수 (Gini, 1997)를 응용하면 참여자들간에 지식 공유의 불평등 정도를 계산할 수 있다. 이러한 두 변수를 계산할 때, 공헌된 지식의 절대적인 양은 측정하기가 쉽지 않으므로, 지식 공유에 대한 다른 연구들에서와 같이 (Kankanhalli et al., 2005; Wasko and Faraj, 2005) 지식을 공헌한 횟수를 기준으로 계산하기로 한다. 지식 협업 프로세스는 참여자를 노드로, 참여자들간의 지식 공헌 선후 관계를 링크로 하는 네트워크로 표현될 수 있다. 예를 들

어, 위키피디아의 한 기사를 에디터 A가 수정한 후, 에디터 B가 다시 수정하면, 노드 B에서 노드 A로 향하는 링크가 생성된다. 이렇게 연속적인 모든 에디팅 행위를 노드와 링크로 나타내면 하나의 네트워크 그래프가 완성된다. 결국, 기사 편집 프로세스는 노드 집합  $N$ 과 링크 집합  $L$ 로 이루어진 그래프  $G = (N, L)$ 이며, 노드  $n_i$  해당하는 에디터  $i$ 의 편집 횟수를  $ed(n_i)$ 라 할 때, 파레토 비율(Pareto ratio)  $pr$ 은  $pr = \frac{\sum_{n_i \in s_{20}} ed(n_i)}{\sum_i ed(n_i)}$  (where  $S_{20} = \{n_i | ed(n_i) > P_{80}\}$ ,

$P_{80}$  is the 80th percentile for  $ed(n_i)$ ,  $i=1, 2, \dots, |N|$ ) 으로 정의된다 (Park et al., 2013). 위키피디아 피쳐드 기사(Featured Article) 협업에서는 파레토 비율 평균이 약 76%로 개인의 지식 공유만을 기준으로 하면 파레토 법칙이 대략 성립한다고 할 수 있다. 그런데, 최소 36%에서 최대 98%까지 기사 그룹마다 차이가 있고, 그룹의 지식 협업 차원에서는 이러한 파레토 비율이 어떠한 영향을 미칠지에 대한 연구는 아직까지 거의 없었다 (Appendix <Table A1> 참조).

다음으로, 지니 계수를 계산하기 위해서는  $ed(n_i)$ 를 감소하지 않는 순서(non-decreasing order)로 정렬할 필요가 있다. 정렬 후의 노드를  $m_i$ 라 하면, 지니 계수(Gini coefficient)  $gc$ 는  $gc = \frac{2\sum_i i \times y_i}{|N|} - \frac{|N|+1}{|N|}$  (where  $y_i = \frac{ed(m_i)}{\sum_i ed(n_i)}$  and  $ed(m_i) \leq ed(m_{i+1})$ , for  $i=1, 2, \dots, |M|$ )로 정의된다 (Park et al., 2013).

## 2.3 작업 그룹의 다양성(diversity)과 그룹 성과(group performance)

작업 그룹의 다양성은 사회적(social) 다양성,

가치(value)의 다양성, 그리고 정보(information)의 다양성으로 분류할 수 있다 (Jehn et al., 1999). 사회적 다양성은 인종, 성(sex), 나이 등 소셜 멤버십의 차이를 의미하고, 가치의 다양성은 선호하는 목적(goal), 목표(target), 가치 등의 차이를, 그리고, 정보의 다양성은 교육, 경험, 훈련 등의 차이에서 비롯되는 정보와 전문지식(expertise)의 차이를 의미한다.

그 동안 이러한 다양성과 그룹 성과간의 관계에 대한 많은 연구들이 있었는데, 이러한 연구들도 크게 세가지 카테고리로 정리해볼 수 있다. 첫째, 다양성의 긍정적인 측면을 제시한 연구들이 있다. 다양성이 높은 그룹은 지식과 통찰력(insights)이 풍부하여 동질적인 그룹보다 그룹 성과가 좋았다 (Amabile, 1994; Cox and Blake, 1991; Gruenfeld et al., 1996; Hoffman, 1978; Hoffman and Maier, 1961; Jackson, 1992; Nemeth, 1986; Oldham and Cumminings, 1998; Smith et al., 2005). 그리고, 전문지식의 다양성이 높은 그룹이 복잡하고 어려운 문제를 해결함에 있어 더욱 우수한 성과를 보였다 (Bantel and Jackson, 1989; Jackson et al., 2003). 둘째, 다양성의 부정적인 측면을 제시한 연구들은 다양성이 높은 그룹들이 흔히 겪게 되는 구성원간 마찰(conflicts)이나 의사소통의 어려움에 주목함으로써, 동질적인 그룹이 보다 효율적이라고 주장했다 (Ancona and Caldwell, 1992; O'Reilly et al., 1989; Steiner, 1972). 셋째, 긍정적인 면과 부정적인 면을 모두 제시한 연구들이 있다. 이 중에는 앞의 두 부류 연구들이 사실은 서로 다른 부분을 다루고 있는 것 일뿐, 서로 모순되는 것이 아님을 지적한 흥미로운 연구도 있었다 (Reagans and Zuckerman, 2001). 즉, 긍정적인 결과를 보인 연구들은 다양한 구성원들이 그룹 외부의 관계로

부터 획득하게 되는 정보 원천의 다양성에 주목한 반면에, 부정적인 결과를 보고한 연구들은 그룹 내부의 상호작용에 집중하여 구성원간 마찰이나 의사소통 문제를 다루었다는 것이다. 이 외에, 전문지식(expertise)의 다양성과 그룹 성과 및 학습 행위간에 (Van der Vegt and Bunderson, 2005), 교육의 다양성과 그룹의 정보 이용간에 (Dahlin et al., 2005), 그리고 근무기간(tenure)의 다양성과 그룹 혁신(innovation) 간에 (Chi et al., 2009) 역 U자(inverted U-shaped) 관계가 있음을 보인 연구들이 있다. 예를 들어, Chi et al. (2009)의 연구에서는, 정보 처리 이론(information processing theory)이나 의사결정 능력(decision-making capacity) (Milliken and Martins, 1996) 관점과 같은 맥락에서 역 U자 관계가 나타나는 이유를 설명하고 있다. 즉, 그룹 구성원들의 근무기간 다양성이 적정 수준 이하일 때에는 이질적인 아이디어와 관점이 의사결정 능력의 향상으로 연결되므로 그룹 혁신에 도움이 되지만, 다양성의 정도가 너무 높아지면 구성원간의 불일치(disagreement)가 증가하여 팀 프로세스에 대한 좌절과 불만족이 상승함으로써 그룹 혁신에 해가 된다는 것이다.

#### 2.4 작업 특성(Work Characteristics)과 그룹 성과(Group Performance)

현재까지 작업(work) 특성에 관한 많은 연구들이 있었는데 주요한 작업 특성은 대략 다음 세 가지로 정리해볼 수 있다. 첫째, 작업 자체가 어떻게 수행되는지를 파악하여, 특정 업무(job)와 관련된 과업(task)의 범위(range) 및 성질(nature)을 설명하는 ‘업무 특성’이다 (Hackman and Oldham, 1976). 이러한 업무 특성 모델(Job

Characteristics Model)은 수많은 연구들에서 인용되어왔으며, 과업 자율성(task autonomy), 과업 다양성(task variety), 과업 정체성(task identity), 과업 의미(task significance), 피드백(feedback) 등의 다섯 가지 과업 특성을 제시했다. 업무 특성 모델의 기본적인 아이디어는 조직 구성원의 동기부여(motivation), 만족도, 성과에 긍정적인 영향을 줄 수 있도록 업무 특성을 디자인하는 것이었다 (Morgeson and Humphrey, 2006; Morgeson and Humphrey, 2008). 둘째, 과업에서 요구되는 지식(knowledge), 기술(skill), 그리고 능력(ability)의 종류를 반영하는 ‘인지적(cognitive) 또는 지식(knowledge) 특성’이다 (Morgeson and Humphrey, 2006; Morgeson and Humphrey, 2008). 이 특성은 정보기술의 발전으로 복잡한 지식 작업이 증가하면서 더욱 중요해지고 있는데, 업무 복잡성(job complexity), 정보 처리(information processing), 문제 해결(problem solving), 기술 다양성(skill variety), 전문화(specialization) 등의 특성이 여기에 속한다. 셋째, 사회적 상호작용과 사회적 환경을 작업 디자인을 결정하는 중요한 요인으로 고려하여, 이와 관련된 업무의 구조적 특성에 주목하는 ‘사회적(social) 특성’이다 (Humphrey et al., 2007). 사회적 특성은 인지적 또는 지식 특성과 비슷하게, 최근에 혼자 수행하기 어려운 복잡한 과업의 증가로 팀(team) 운영이 활성화되면서 더욱 중요해지고 있다. 상호의존성(interdependence), 사회적 지원(social support), 조직 외부 상호작용(interaction outside the organization), 그리고 타인으로부터의 피드백(feedback from others) 등을 들 수 있다.

과업 다양성, 과업 정체성, 과업 의미가 클수록 성공적인 그룹 성과로 이어지며 (Li et al., 2009; Stewart, 2006), 과업 자율성은 책임감을 높

여줌으로써 효과적인 그룹 성과를 유도하고 (Hackman and Oldham, 1976; Spreitzer et al., 1999), 과업 관련 피드백은 가상(virtual) 팀에서 동기부여, 만족도, 성과의 향상에 도움을 준다고 한다 (Geister et al., 2006). 그리고, 작업 특성에 따라 작업 그룹의 다양성 효과가 달라진다는 연구들도 있다. 예를 들어, 업무 복잡성이 높을 때에는 정보 다양성이 그룹 성과에 미치는 긍정적인 효과가 더욱 크다고 한다 (Jehn et al., 1999; Pelled et al., 1999). 또, 과업의 상호의존성(interdependence)이 높을 때, 그룹의 정보 다양성이나 가치 다양성이 그룹 성과에 미치는 영향력이 더 크다고 한다 (Jehn et al., 1999; Williams and O'Reilly, 1998).

### 3. 연구 가설

#### 3.1 지식 공유 파레토 비율 및 불평등 정도와 협업 효율성

H1-1: 온라인 커뮤니티에서 지식 협업 그룹의 지식 공유 파레토 비율과 협업 효율성 간에는 역 U자(inverted U-shaped) 관계가 존재한다.

H1-2: 온라인 커뮤니티에서 지식 협업 그룹의 지식 공유 불평등 정도와 협업 효율성 간에는 역 U자(inverted U-shaped) 관계가 존재한다.

온라인 커뮤니티에서 상대적으로 더 많은 지식을 공헌하는 사람들은 지식 공헌 행위가 자신의 평판(reputation)이나 지위(status) (Constant et al., 1996; Lakhani and von Hippel, 2003; Lakhani

and Wolf, 2005; Wasko and Faraj, 2005)를 높여 줄 수 있는 수단이라고 인식하기 때문에, 또는 상호(reciprocal) 관계를 기대하거나 자아가치개념(sense of self-worth) (Brock et al., 2005), 지식 자기효능감(knowledge self-efficacy)과 다른 사람을 돕는 즐거움 (Hsu et al., 2007; Kankanhalli et al., 2005; Kollock, 1999; Wasko and Faraj, 2000)을 느낄 수 있는 등의 이유로 더 강하게 동기부여 된다고 할 수 있다. 또, 개인의 가치(value)와 조직의 가치가 일치할 때, 개인의 조직에 대한 몰입도가 증가하고 (O'Reilly, 1989), 이것은 더 많은 지식 공헌으로 이어질 수 있다 (Anthony et al., 2009; Raymond, 2001). 특히, 전문지식이나 기술을 더 많이 소유한 사람일수록 자신의 지식을 공헌하려는 경향이 더 강하다 (Wasko and Faraj, 2005). 이것은 지식을 많이 공헌하는 사람들일수록 당면 문제나 아티클 주제에 대한 전문 지식이나 기술을 많이 소유하고 있을 가능성이 높음을 시사해주고 있다.

일반적으로 강하게 동기부여 된 개인들의 지식 공유 비율이 증가할수록 지식 협업의 결과는 더욱 좋아질 것으로 예상할 수 있다. 그런데, 예를 들어, 상위 20%의 지식 공유 비율이 높아지면 나머지 80%에 해당하는 참여자들로부터 기대할 수 있는 정보의 다양성이 제대로 반영되지 않을 가능성이 높아지기 때문에 이러한 비용을 따져보지 않을 수 없다. 새로운 지식을 창조하기 위해서는 선행 지식의 투입이 필요하고 (McFadyen and Cannella, 2004), 진행 과정에서 발생 가능한 문제 해결을 위해 다양한 아이디어와 관점이 요구되므로 (Highsmith, 1978; MacKay et al., 1992; Simon, 1960), 정보의 다양성은 지식 협업의 성공에 핵심적인 역할을 할 것이기 때문이다 (Nonaka, 1994; Nonaka et al., 1996).

물론, 이러한 정보의 다양성 때문에 참여자들 간에 의견의 불일치가 발생하고 이를 위한 조정 비용이 증가한다면 협업 효율성은 감소할 수 밖에 없다. 따라서, 정보 다양성의 긍정적인 효과는 협업 참여자들이 서로 다른 타인과 상호작용할 때 수반되는 어려움을 성공적으로 관리하는 정도에 따라 달라질 것이다 (Tsui and O'Reilly, 1989). 사람들은 타인과의 상호작용에서 유사성(similarity)을 선호한다는 유사성 매력 이론(similarity-attraction theory) (Byrne, 1971)이 있다. 비슷한 맥락에서 사회화 이론(socialization theory) (Van Maanen and Schein, 1979)은 집단 구성원의 유사성을 강조한다. 또, 선택 이론(selection theory) (Chatman, 1991)은 가치(value)와 인구통계학적인 유사성을 작업 환경의 효과성을 위한 기본적인 조건으로 가정하고 있다. 이러한 이론들에 의하면 다양성보다는 유사성이 협업을 위해 더욱 필요한 요인으로 간주될 수 있다. 그런데, 가상 환경에서는 사회적 실재감(presence) (Short et al., 1976)과 사회적 맥락(context)의 단서 (Kiesler et al., 1984; Sproull and Kiesler, 1986)가 부족하여 기본적으로 친밀한 인간관계를 형성하기 어렵기 때문에 (Kiesler et al., 1984; Short et al., 1976; Sproull and Kiesler, 1986), 사회적 다양성이나 유사성의 효과가 직접적으로 발현되지 않을 가능성이 크다. 또, 위키피디아와 같은 커뮤니티는 신뢰할만하고 (reliable) 검증 가능하며(verifiable) 동시에 가치 중립적인(neutral) 아티클 생성을 목표로 하며, 이를 위한 편집 가이드라인도 마련되어 있으므로, 가치의 다양성이나 유사성의 영향 정도가 아티클 그룹마다 크게 달라질 것으로 생각되지 않는다.

가상 지식 협업의 대표적인 예라 할 수 있는

오픈 소스 모델에 관한 한 연구에서는 비교적 적은 수의 공식적으로 조직화된 전문가 그룹에 의한 전통적인 소프트웨어 개발 모델을 성당(cathedral) 모델(이후 ‘전문가 주도형’으로 칭함)로, 이후 ICT의 발전으로 가능하게 된 피어 프러덕션(peer production) 모델, 즉 오픈 소스 모델을 바자(bazaar) 모델(이후 ‘대중 주도형’으로 칭함)에 비유했다 (Raymond, 2001). 대중 주도형은 주어진 문제에 대한 많은 참여자들의 다양한 관점을 통해 빠른 속도로 버그를 발견하고 수정한다. 그런데, 대중 주도형은 기존 프로그램을 보완하는 것과 같은 비교적 쉽거나 명확히 정의된 병렬적인 작업에 효율적이고, 새로운 프로그램을 처음부터 개발하는 것과 같은 난이도가 높은 작업에는 적합하지 않다 (Raymond, 2001). 리눅스가 성공할 수 있었던 것은 많은 사람들이 참여했을 뿐만 아니라 참여자들의 작업을 조정 및 통합하여 다음 버전 작업을 위해 다시 배포하는 전문적인 프로그래머들이 있었기 때문이라고 한다 (Carr, 2007). 요컨대, 많은 사람들의 다양성에 기반한 대중 주도형과 전문적인 프로그래머들에 의한 전문가 주도형의 결합이 성공을 이뤄낸 주요 요인이었다는 것이다.

위키피디아의 경우에는 전문적인 프로그래머와 같이 공식적으로 구성된 전문가 조직은 거의 존재하지 않는다고 할 수 있다. 그런데, 지식 협업의 효율성 측면에서 전문가와 대중의 조화가 중요하다는 것은 리눅스와 기본적으로 다르지 않을 것으로 생각된다. 결과적으로, 위키피디아에서는 대략 상위 20%에 속하는 참여자들이 자발적으로 이러한 성당(전문가)의 역할을 수행하고, 나머지 하위 80%의 참여자들이 바자(대중) 기능을 담당할 때 협업의 성과가 높아질 것으로 예측할 수 있다. 물론, 역할 비율을 확정적으로

단언할 수는 없고, 자연적으로 창출되는 역할의 조화가 온라인 커뮤니티의 유동성(fluidity)으로 인한 긴장(tension)에 대응해 얼마나 생산적으로 이루어지느냐가 중요한 요인이 될 것이다 (Faraj et al., 2011).

온라인 커뮤니티에서 긴장감을 유발하는 중요한 자원(resource)중의 하나는 강한 동기 부여 또는 열정(passion)이다. 열정은 온라인 커뮤니티에서 지식을 공헌하는 참여자들에게서 발견되며, 참여자들이 커뮤니티의 지식 베이스를 발전시키는데 노력을 집중하도록 유도한다 (Faraj et al., 2011). 강하게 동기 부여된 열정적인 참여자들은 모범적인 행위로 다른 참여자들을 고무시켜 더 많은 참여를 유도할 수 있고 (Bono and Illies, 2006), 더욱 능동적으로 서로 다른 아이디어들로부터 새로운 조합을 찾아내므로 (Chen et al., 2009), 지식 협업에 긍정적인 영향을 미친다. 그런데, 지나친 열정은 협업을 저해하는 갈등을 유발하기 쉽고 (O'Connor and Adair, 2003), 상대적으로 덜 동기 부여된 참여자들에게 소외감을 주게 되어 결국 참여를 저해하는 부정적인 결과를 가져올 수 있다 (Faraj et al., 2011).

정리하자면, 지식 공유에 대한 동기 부여 정도가 높은 참여자들의 지식 공헌 비율이 높아질수록 지식 협업의 효율성은 높아지지만, 일정 수준을 넘어서면 정보 다양성의 이점을 포기해야 하거나 상대적으로 동기 부여 정도가 낮은 구성원들의 참여를 제한하게 되어 다시 감소할 것으로 예상된다. 이러한 현상은 상위 20%의 파레토 비율이나 불평등을 나타내는 지니 계수에 의해 비슷하게 설명될 수 있을 것이다. 실제로, 파레토 비율과 지니 계수의 상관계수는 Appendix의 <Table A2>에서와 같이 0.9로 상당히 높은 수준이다. 그런데, 파레토 비율과 지니 계수는 이들



이 의미하는 상징적인 의미가 서로 다를 수 있고, 각각의 효과를 따로 확인해보는 것도 필요하다고 판단하여 위와 같이 독립적인 가설을 설정하였다.

### 3.2 작업 특성을 고려한 지식 공유 파레토 비율 및 불평등 정도와 협업 효율성

H2-1: 온라인 커뮤니티에서 지식 협업 그룹의 지식 공유 파레토 비율이 협업 효율성에 미치는 영향의 한계 효과(marginal effect)는 학문적인 작업에서 더욱 강하게 나타난다.

H2-2: 온라인 커뮤니티에서 지식 협업 그룹의 지식 공유 불평등 정도가 협업 효율성에 미치는 영향의 한계 효과(marginal effect)는 학문적인 작업에서 더욱 강하게 나타난다.

21세기에 접어들면서 조직에서 수행하는 작업은 더욱 복잡해지고 인지적인 노력을 더 많이 요구하며, 유연한 작업 환경과 협업이 중요해지고, 작업 그룹의 구성은 더욱 다양해지는 등 과거와는 근본적으로 다르게 변화하고 있다 (Oldham and Hackman, 2010). 이에 따라, 지식 작업자(knowledge worker)들의 중요성이 점점 커지고, 조직의 전체 구성원 중에서 지식 작업자가 차지하는 비율도 계속 증가하고 있다 (Davenport, 2006; Levenson, 2012). 업무 복잡성이 높을 때에는 정보 다양성이 그룹 성과에 미치는 긍정적인 효과가 더욱 크기 때문에 (Bantel and Jackson, 1989; Jackson et al., 2003; Jehn et al., 1999; Pelled et al., 1999), 작업의 인지적 또는 지식 특성에 따라 지식 공헌의 파레토 비율과 불평등 정

도가 협업 효율성에 미치는 영향력의 크기도 달라질 수 있다. 위키피디아에는 전문적인 저널에 게재된 논문들을 참조하여 작성되는 아티클들과 그렇지 않은 아티클들이 있다. 전문적인 내용을 다루고 있는 아티클의 편집 작업은 업무의 복잡성(job complexity)이나 정보 처리(information processing) 및 문제 해결(problem solving)의 정도, 요구되는 기술의 다양성(skill variety)과 전문화(specialization) 정도가 그렇지 않은 경우에 비해 훨씬 높을 것으로 생각된다. 그러므로, 위의 가설을 상정하였다.

## 4. 연구 방법 및 연구 결과

### 4.1 데이터 수집 및 분석 방법

영어 위키피디아에서 2012년 6월 현재 피쳐드 아티클(Featured Article) 레벨에 도달한 2,978 개의 아티클에 대해 아티클 생성시점부터 피쳐드 아티클 도달 시점까지의 아티클 편집 히스토리를 크롤링(crawling)하였다. 크롤링은 자바 프로그램으로 위키피디아에서 제공하는 HTTP API를 통해 수행하였고, 수집된 데이터의 저장은 SQLite 프로페셔널을 사용하였다. 그리고, SQL과 자바 프로그램을 이용하여 데이터를 분석하였다.

아티클 편집 히스토리로부터 각각의 에디터가 아티클 편집에 몇 번 참여했는지를 추출할 수 있다. 이 자료를 이용하면, 아티클 마다 상위 랭킹 20%에 속하는 에디터들의 총 편집 횟수 비율인 파레토 비율과 지식 공유의 불평등 정도를 나타내는 지니 계수를 구할 수 있다. 위키피디아에 등록된 에디터들은 유저 네임과 ID가 편집 히스

〈Table 1〉 Control Variables

Control Vars.	Definition
ZNumOfEditors	the number of editors involved in an article
ZRegUserRatio	the ratio of the number of registered editors involved in an article to the number of editors involved in the article (%)
ZDiscUserRatio	the ratio of the number of editors who have participated in an article talk to the number of editors involved in the article (%)
ZInterArrivalTime	the average time length between successive edits of an article (days)
ZRevisitingPeriod	the average time length over which editors of an article visit the article for its editing (days)
ZNumOfSections	the number of sections a featured article has
ZNumOfReferences	the number of references a featured article refers to

토리에 기록되어 고유하게 식별 가능하지만, 익명으로 참여하는 에디터들은 사용 IP 주소만 기록되므로, 서로 다른 IP 주소는 서로 다른 에디터로 간주하였다. 또, 오타 수정 같은 단순 작업을 위한 자동화 프로그램인 ‘bots’나 ‘scripts’에 의한 편집은 분석에서 제외하였다. 결국, 2,978개의 아티클 협업 프로세스에 참여한 736,806명의 유저 행위를 분석하였다. 이 수치는 각 아티클의 협업 프로세스를 독립적으로 분석했으므로 아티클간 중복을 허용하여 계산한 결과이고, 고유한 유저 수는 428,357명이다.

지식 협업 효율성은 아티클 편집 히스토리로 부터 각 아티클의 생성 시점과 피쳐드 아티클 (Featured Article) 레벨로 승급된 시점을 추출하여 피쳐드 아티클 승급에 소요된 일수(number of days)로 표현된다. 특히, 분석 방법으로 준-모수 (semi-parametric) 비례 해저드(hazard) 모델인 콕스(Cox) 회귀분석을 사용하였으므로, 피쳐드 아티클 승급에 소요된 일수가 바로 종속변수가 된다. 콕스 회귀분석은 과거에 해당 이벤트가 발생하지 않았다는 조건하에 임의의 시각에 해당 이벤트가 발생할 가능성을 모델링하는 방법으로,

아티클 승급을 하나의 이벤트로 간주하여 본 연구에 적용될 수 있다. 독립변수들을 아티클 그룹의 특성을 나타내는 관련 변수들로 구성하면, 어떤 아티클 그룹 특성이 피쳐드 아티클 승급에 어떤 영향을 미치는지를 분석할 수 있다.

모델 빌딩 프로세스는 두 블록으로 진행되었다. 즉, 통제(control) 독립변수들만을 포함하는 첫 번째 블록 모델에 대한 결과를 얻고, 다시 통제 독립변수와 관심(focal) 독립변수를 모두 포함하는 두 번째 블록 모델을 분석하였다. 통제 독립변수들에 대한 정의는 <Table 1>과 같다. 얼마나 많은 에디터들이(ZNumOfEditors) 얼마나 자주(ZInterArrivalTime) 그리고 얼마나 오랫동안(ZRevisitingPeriod) 한 아티클을 방문하며 편집에 참여하는지, 그리고 해당 아티클의 주제는 얼마나 완성하기에 어려운지(ZNumOfSections, ZNumOfReferences), 익명이 아닌 등록 에디터들은 얼마나 되는지(ZRegUserRatio), 협업 프로세스 과정에서 얼마나 많은 에디터들이 토론에 참여했는지(ZDiscUserRatio)에 따라 협업 프로세스의 효율성이 달라질 것으로 판단하였다.

모든 독립변수들은 아티클 생성 시점부터 피

쳐드 레벨로 승급되기 바로 직전까지의 기간에 대해 계산되었고, 영향력의 상대적인 비교를 위해 정규화되었다. 아티클 완성의 난이도를 반영하는 변수로 아티클의 길이도 고려했었지만 예비 분석 결과 유의하지 않아 제외하였다. 그리고, 분석 과정에서 재방문기간의 제곱 값과 협업 효율성간에 유의한 관계가 발견되어 이 변수 (ZRevisitingPeriod2)도 회귀분석에 포함시켰다.

다음으로, 인지적(cognitive) 또는 지식(knowledge) 특성의 정도에 따라 2,978개의 아티클을 두 개의 그룹으로 분류하였는데, 이에 대한 기준은 각 아티클의 레퍼런스 리스트를 크롤링하여, 이 중 SCI, SSCI, A&HCI, 또는 SCIE 논문이 포함되어 있는지 여부로 결정하였다 (2013년 7월 20일 기준). SCI, SSCI, A&HCI, 또는 SCIE 레퍼런스가 포함된 학문적인(academic) 특성의 아티클을 생성하기 위해서는 그렇지 않은 경우에 비해, 상대적으로 더 높은 업무 복잡성(job complexity)이 예상되고 더 많은 정보 처리(information processing), 문제 해결(problem solving), 기술 다양성(skill variety), 전문화(specialization) 등이 요구될 것으로 판단하였다. 전체 2,978개 아티클 중 596개가 학문적인 아티클로, 전체 736,806명의 유저들 중 167,328명의 유저가 이러한 아티클의 편집에 참여한 것으로 나타났다.

## 4.2 가설 검증 결과

### 4.2.1 지식 공유 파레토 비율과 협업 효율성

<Table 2>는 피쳐드 아티클(Featured Article) 전체와 이들을 학문적인 특성의 정도에 따라 분류한 두 그룹의 아티클들에 대한 세 가지 콕스 회귀분석 결과를 보여주고 있다. 전체적으로 거

의 모든 통제변수들이 양측 검정(two-tailed test)의 유의수준 0.001에서 유의(significant)하다. 그리고, 관심 변수인 파레토 비율에 대한 이차함수 관계도 유의수준 0.01과 0.1에서 유의하다. <Table 2>에서 이전 블록의 -2LogLikelihood와 현재 블록의 -2LogLikelihood 차이는 현재 블록의 카이스퀘어 증가치(improvement chi-square change)와 같다. 이 값은 자유도가 현재 블록에 새로 투입된 변수의 수인 카이스퀘어 분포(chi-square distribution)를 따르는 것으로 알려져 있다. 이러한 카이스퀘어 증가치(improvement chi-square change)가 역시 모두 유의하므로, 관심 변수를 추가했을 때의 증가된 설명 효과가 유의함을 알 수 있다. 또, 모든 회귀계수가 영(zero)이라는 귀무가설(null hypothesis)은 전반적인 카이스퀘어 통계치(overall chi-square statistics)에 의해 기각되어 모든 모델이 유의수준 0.001에서 유의하다. SPSS에서는 콕스 회귀분석에 대한 R<sup>2</sup> 값을 제공하지 않아 오픈 소스인 R의 ‘survival’ 패키지를 이용해 의사(pseudo) R<sup>2</sup> 값을 구했다. 콕스 회귀분석에 대한 비례성 가정(proportionality assumption)은 각각의 독립변수를 카테고리 변수로 변환하여 카플란-마이어 커브(Kaplan-Meier curve) 방법 (Park, 2007)에 의해 검증함으로써 만족됨을 확인하였다.

$i$  번째 케이스에 대한 시각  $t$ 에서의 해저드(hazard) 레이트(rate)  $h_i(t)$ 를  $h_i(t) = h_0(t) e^{b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip}}$  (단,  $h_0(t)$ 는 시각  $t$ 에서의 베이스라인 해저드(hazard),  $p$ 는 독립변수의 수)라고 할 때, 로그 해저드비(log hazard ratio)는  $h_i(t) = h_0(t) e^{b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip}}$ 로 표현된다.

(Table 2) Effect of Pareto Ratio on Collaboration Efficiency

Vars.	Total		Academic		Non-academic	
	Reduced	Full	Reduced	Full	Reduced	Full
	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))
ZNumOfEditors	-0.199*** (0.819)	-0.269*** (0.764)	-0.344*** (0.709)	-0.370*** (0.690)	-0.175*** (0.840)	-0.246*** (0.782)
ZRegUserRatio	0.836*** (2.308)	1.025*** (2.786)	1.013*** (2.753)	1.142*** (3.132)	0.791*** (2.206)	1.002*** (2.725)
ZDiscUserRatio	0.463*** (1.589)	0.440*** (1.553)	0.477*** (1.611)	0.490*** (1.632)	0.477*** (1.612)	0.449*** (1.567)
ZInterArrivalTime	-0.821*** (0.440)	-0.950*** (0.387)	-0.954*** (0.385)	-1.023*** (0.359)	-0.808*** (0.446)	-0.963*** (0.382)
ZRevisitingPeriod	-1.181*** (0.307)	-1.175*** (0.309)	-1.310*** (0.270)	-1.299*** (0.273)	-1.259*** (0.284)	-1.271*** (0.281)
ZRevisitingPeriod2	0.605*** (1.831)	0.607*** (1.834)	0.636*** (1.889)	0.623*** (1.865)	0.791*** (2.207)	0.816*** (2.262)
ZNumOfSections	-0.108*** (0.897)	-0.101*** (0.904)	-0.038 (0.963)	-0.024 (0.976)	-0.107*** (0.898)	-0.102*** (0.903)
ZNumOfReferences	-0.390*** (0.677)	-0.349*** (0.706)	-0.335*** (0.715)	-0.325*** (0.723)	-0.410*** (0.663)	-0.358*** (0.699)
ZParetoRatio		0.518* (1.679)		1.007* (2.738)		0.584* (1.794)
ZParetoRatio2		-0.730** (0.482)		-1.121* (0.326)		-0.822** (0.439)
Observations	2,978	2,978	595	595	2,383	2,383
-2LogLikelihood	38,498.180	38,433.476	5,721.101	5,714.541	29,793.825	29,729.060
Overall Chi-square	3,607.042***	3,684.837***	847.017***	852.242***	2,773.660***	2,848.960***
Improve. Chi-square	3,200.794***	64.705***	700.174***	6.560*	2,513.241***	64.765***
R Square (%)	65.9	66.6	69.2	69.5	65.2	66.1

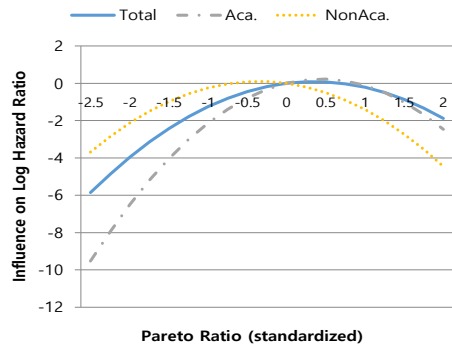
\*p<0.1, \*\*p<0.01, \*\*\*p<0.001 (two-tailed test)

-2LogLikelihood for Null Model (Total) : 41,698.974

-2LogLikelihood for Null Model (Academic) : 6,421.274

-2LogLikelihood for Null Model (Non-academic) : 32,307.066

<Table 2>의 세 가지 모델에서 통제변수들이 일정하다고 할 때, 로그 해저드비에 미치는 파레토 비율의 종합적인 효과는 <Figure 1>과 같다. 최적 파레토 비율은 학문적인 특성이 강한 아티클의 경우가 상대적으로 더 높고, 파레토 비율 1 단위 증가에 대한 로그 해저드비의 한계(marginal) 효과도 학문적인 특성이 강한 아티클이 더 큰 것으로 보인다. 이것은 <Table 2>의 파레토 비율 이차항(ZParetoRatio2)의 계수에서도 확인할 수 있다. 결과적으로, 가설 H1-1과 H2-1이 성립한다



(Figure 1) Relationship between Pareto Ratio and Collaboration Efficiency

고 할 수 있다.

#### 4.2.2 지식 공유의 불평등 정도와 협업 효율성

<Table 3>는 피쳐드 아티클(Featured Article) 전체와 이들을 학문적인 특성의 정도에 따라 분류한 두 그룹의 아티클들에 대한 세 가지 콕스 회귀분석 결과를 보여주고 있다. 앞서와 같이, 거의 모든 통제변수들이 유의수준 0.001에서 유의(significant)하다. 그리고, 관심 변수인 지니 계

수에 대한 이차함수 관계도 세 모델 모두 유의수준 0.001에서 유의하다. 카이스퀘어 증가치(improvement chi-square change) 역시 유의수준 0.001에서 유의하므로, 지니 계수를 추가했을 때 증가된 설명 효과가 유의함을 알 수 있다. 또, 전반적인 카이스퀘어 통계치(overall chi-square statistics)를 보면, 모든 모델이 유의수준 0.001에서 유의하다. R에 의해 구한 의사(pseudo) R<sup>2</sup> 값도 전반적으로 양호하다. 앞서와 마찬가지로, 콕스 회귀분석의 비례성 가정(proportionality

<Table 3> Effect of Gini Coefficient on Collaboration Efficiency

Vars.	Total		Academic		Non-academic	
	Reduced	Full	Reduced	Full	Reduced	Full
	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))	B (Exp(B))
ZNumOfEditors	-0.199*** (0.819)	-0.276*** (0.759)	-0.344*** (0.709)	-0.349*** (0.705)	-0.175*** (0.840)	-0.256*** (0.774)
ZRegUserRatio	0.836*** (2.308)	1.045*** (2.843)	1.013*** (2.753)	1.181*** (3.256)	0.791*** (2.206)	1.018*** (2.769)
ZDiscUserRatio	0.463*** (1.589)	0.432*** (1.541)	0.477*** (1.611)	0.477*** (1.610)	0.477*** (1.612)	0.442*** (1.556)
ZInterArrivalTime	-0.821*** (0.440)	-0.977*** (0.377)	-0.954*** (0.385)	-1.056*** (0.348)	-0.808*** (0.446)	-0.983*** (0.374)
ZRevisitingPeriod	-1.181*** (0.307)	-1.178*** (0.308)	-1.310*** (0.270)	-1.320*** (0.267)	-1.259*** (0.284)	-1.264*** (0.282)
ZRevisitingPeriod2	0.605*** (1.831)	0.610*** (1.841)	0.636*** (1.889)	0.636*** (1.889)	0.791*** (2.207)	0.810*** (2.247)
ZNumOfSections	-0.108*** (0.897)	-0.104*** (0.901)	-0.038 (0.963)	-0.027 (0.974)	-0.107*** (0.898)	-0.103*** (0.902)
ZNumOfReferences	-0.390*** (0.677)	-0.340*** (0.712)	-0.335*** (0.715)	-0.320*** (0.726)	-0.410*** (0.663)	-0.350*** (0.705)
ZGini		0.663*** (1.942)		1.702** (5.483)		0.586** (1.797)
ZGini2		-0.888*** (0.412)		-1.794*** (0.166)		-0.839*** (0.432)
Observations	2,978	2,978	595	595	2,383	2,383
-2LogLikelihood	38,498.180	38,416.653	5,721.101	5,705.494	29,793.825	5,666.894
Overall Chi-square	3,607.042***	3,691.090***	847.017***	853.766***	2,773.660***	698.367***
Improve. Chi-square	3,200.794***	81.527***	700.174***	15.607***	2,513.241***	64.112***
R Square (%)	65.9	66.8	69.2	70.0	65.2	66.3

\*p<0.1, \*\*p<0.01, \*\*\*p<0.001 (two-tailed test)

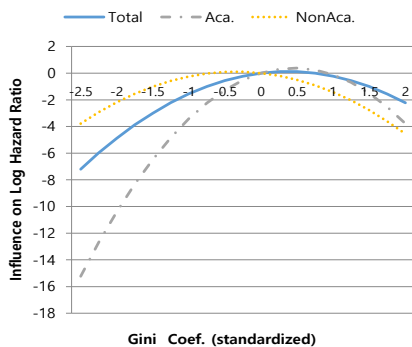
-2LogLikelihood for Null Model (Total) : 41,698.974

-2LogLikelihood for Null Model (Academic) : 6,421.274

-2LogLikelihood for Null Model (Non-academic) : 32,307.066

assumption)은 카플란-마이어 커브(Kaplan-Meier curve) 방법으로 검증하였다.

<Table 3>의 세 가지 모델에서 통제변수들이 일정하다고 할 때, 로그 해저드비에 미치는 지니 계수의 종합적인 효과는 <Figure 2>와 같다. 최적 지니 계수는 학문적인 특성이 강한 아티클의 경우가 상대적으로 더 높고, 지니 계수 1 단위 증가에 대한 로그 해저드비의 한계 효과도 학문적인 특성이 강한 아티클이 더 큰 것으로 보인다. 두 그룹의 차이는 파레토 비율에서보다 더 뚜렷하다. 결과적으로, 가설 H1-2와 H2-2가 성립한다고 할 수 있다.



<Figure 2> Relationship between Gini Coefficient and Collaboration Efficiency

### 4.3 연구 결과 요약

먼저, <Table 2>와 <Table 3>의 통제 변수에 대한 결과를 살펴볼 때 몇 가지 공통되는 사항이 있다. 첫째, 더 많은 에디터들이 참여할수록 피쳐드 아티클(Featured Article)로 승급되는데 더 오랜 시간이 소요된다 (ZNumOfEditors). 이것은 에디터들이 아티클 생성부터 승급 시점까지 책임감을 가지고 계속 참여하는 것이 아니고, 그룹 사이즈가 커질수록 이에 의해 야기되는 복잡성

과 혼돈의 정도가 심해지기 때문인 것으로 추측된다. 기존의 연구에서도 이와 같은 결과를 보고한 바 있다 (Nemoto et al., 2011). 둘째, 아티클 그룹 전체 에디터 수에 대한 등록 에디터 수의 비율(ZRegUserRatio)이 높아질수록 승급에 소요되는 시간이 짧아진다. 위키피디아에서는 커뮤니티에 등록을 하고 활동하거나 익명으로 참여할 수도 있다. 등록을 하는 경우에는 에디터의 유저 네임과 함께 모든 활동 내역이 기록되므로 평판이나 지위를 위해 동기 부여 되거나 (Constant et al., 1996; Lakhani and von Hippel, 2003; Lakhani and Wolf, 2005; Wasko and Faraj, 2005), 커뮤니티가 추구하는 가치에 부합하는 콘텐츠를 제공할 가능성이 높다 (Anthony et al., 2009; O'Reilly, 1989). 즉, 전반적으로 익명으로 참여할 때보다 공헌된 지식의 양이나 질이 더 우수하기 때문인 것으로 추측할 수 있다 (Anthony et al., 2009). 셋째, 아티클 그룹 전체 에디터 수에 대한 토론에 참여한 에디터 수의 비율(ZDiscUserRatio)이 높아질수록 승급에 소요되는 시간이 짧아진다. 위키피디아에서는 아티클 내용에 관련된 이슈들에 대한 자신의 아이디어나 의견을 아티클마다 마련된 아티클 토크(talk)이라는 토론 페이지에 올릴 수 있다. 토론에 참여하는 에디터들이 많을수록 아티클에 대한 공통된 이해 기반이 넓어지고 (Nonaka, 1994; Nonaka et al., 1996) 다른 사람들의 견해에 대해 생각해보므로써 자신의 지식을 발전시킬 수 있으므로 그룹의 인지적 자본(cognitive capital) (Nahapiet and Ghoshal, 1998; Wasko and Faraj, 2005)이 증가하여 지식 협업에 긍정적인 효과를 가져올 수 있다. 넷째, 참여자들의 평균적인 도착 사이 시간(ZInterArrivalTime)이 짧을수록 승급에 소요되는 시간이 짧아진다. 이것은 연속되는 편집 행위가

바로 전 편집행위에 대한 피드백으로 작용하여 좀더 즉각적인 반응이 더 강한 몰입을 유발하는 피드백 효과에 의한 것으로 생각해볼 수 있다 (Csikszentmihalyi et al., 2005). 다섯째, 에디터들이 특정 아티클을 재방문하는 기간(ZRevisitingPeriod)은 협업 효율성과 U자 형태의 이차 곡선 관계를 보인다. 변곡점 이하 구간에서는 재방문기간이 길어질수록 협업 효율성이 낮아지는 반면에 변곡점 이후 구간에서는 재방문기간이 길어질수록 협업 효율성이 증가한다. 여섯째, 아티클의 섹션 수(ZNumOfSections)나 아티클 편집에 참조된 레퍼런스 수(ZNumOfReferences)가 많을수록 승급에 소요되는 시간이 길어진다. 이것은 구조적-참조적인 복잡도가 높은 아티클일수록 완성하기가 어려울 것이라는 예상과 일치하는 결과라고 할 수 있다.

다음으로, 파레토 비율과 지니 계수는 가상 공간에서 이루어지는 지식 협업 효율성과 역 U자 형태의 이차 함수 관계를 보이며, 학문적인 특성이 강한 작업인 경우에 더 민감한 반응을 보인다.

## 5. 결론 및 미래 연구 방향

### 5.1 이론적인 시사점

본 연구를 통해 생각해 볼 수 있는 이론적인 시사점은 다음과 같이 세 가지로 정리해볼 수 있겠다. 첫째, 경영 현장에서 자주 접하게 되는 파레토나 롱테일 현상을 가상 지식 협업에 적용해 볼 때, 어느 한 쪽이 더 중요하다기보다는 상위 20%와 하위 80%의 보완적인 조화가 이루어질 때 협업 효율성이 최대가 되는 것으로 보인다.

이 결과는 전문가의 전문성과 대중의 다양성이 함께 어우러질 때 가상 지식 협업이 효과적임을 제시한 연구 (Carr, 2007)와 일맥상통하는 면이 있다. 다만, 이전 연구 (Carr, 2007)에서는 조직의 구조적인 측면에서 위키피디아에는 리눅스 개발 작업에서와 같은 전문가 기능이 없다고 했지만, 위키피디아 피쳐드 아티클 그룹의 협업 프로세스를 살펴보면, 상위 20%에 해당하는 참여자들이 자발적이고 유기적인 전문가 역할을 수행하고 있는 것으로 생각된다. 이러한 위키피디아의 전문가 기능은 온라인 커뮤니티의 유동성(fluidity)에 의해 자연스럽게 형성되고 변화하는 (Faraj et al., 2011) 것으로 리눅스의 공식적인 업무 할당에 의한 정적인 전문가 조직과는 차이가 있다. 그리고, 이전 연구에서는 주요 관리자들에 대한 인터뷰 내용을 중심으로 논의를 전개하였는데, 본 연구에서는 객관적인 행위 데이터 분석 결과를 활용하고 있다. 둘째, 지식 공유의 파레토 비율과 지식 공유 불평등 정도를 나타내는 지니 계수는 둘 다 지식 협업의 효율성에 대해 역 U자(inverted U-shaped) 형태의 영향을 미치며, 이것은 작업 그룹의 다양성 및 모티베이션 정도의 상대적인 차이에서 발생하는 트레이드오프(trade-off) 등에 의해 설명될 수 있다. 셋째, 인지적인 노력을 상대적으로 많이 요구하는 작업의 경우에는 그렇지 않은 경우보다 참여자들간 지식 공유의 불평등 정도가 협업 효율성에 더욱 민감한 영향을 미친다. 이것은 업무 복잡성이 높은 경우에 정보 다양성이 그룹 성과에 미치는 긍정적인 효과가 더욱 크다는 기존 연구 결과와 상통하는 면이 있다. 협업 효율성에 대한 최적 파레토 비율과 지니 계수 값은 인지적인 노력을 상대적으로 더 많이 요구하는 학문적인 작업의 경우에 더 높게 나타난다.

## 5.2 실용적인 시사점

본 연구는 온라인 커뮤니티 환경의 지식 협업을 관리하는 경영자들에게 대략 다음 세 가지의 시사점을 제공한다. 첫째, 전문가 기능을 담당할 수 있는 핵심 멤버들이 적극적으로 오랜 기간 동안 참여할 수 있도록 평판, 자기효능감, 다른 사람을 돕는 즐거움 등을 위한 인센티브 메커니즘을 구현하는 것이 중요하다. 그런데, 가상 지식 협업에서는 작업의 발전 단계에 따라 새로운 참여자들이 전문가 기능을 담당할 수 있도록 기존의 핵심 멤버들이 뒤로 물러나거나 커뮤니티를 떠나는 것도 필요하다. 온라인 커뮤니티와 같은 가상 지식 협업 플랫폼에서는 기존 멤버들이 공헌한 지식이 자동적으로 저장되어 새로운 참여자들이 이러한 지식을 활용할 수 있으므로 턴오버(turnover)에 의한 손실이 완화될 수 있다. 둘째, 새로운 참여자들이나 동기부여 정도가 낮은 참여자들에게도 개방적인 커뮤니티 문화가 중요하며, 인센티브 메커니즘은 이들의 아이디어가 충분히 반영될 수 있도록 지원하는 기제로도 활용될 수 있어야 한다. 셋째, 작업 특성에 따라 파레토 비율과 지니 계수가 협업 프로세스의 효율성에 미치는 영향의 양상이 달라질 수 있음을 인식하고, 유사한 특성의 기존 협업 프로세스에 대한 최적 값을 고려하여 해당 커뮤니티 환경에서 이루어지는 새로운 협업 프로세스를 관리할 필요가 있다.

## 5.3 연구의 한계점 및 미래 연구 방향

본 연구는 다음과 같은 한계점 및 지속적인 연구 주제를 남긴다. 첫째, 본 연구는 영어 위키피

디아에서 피쳐드 아티클(Featured Article) 레벨에 도달한 아티클들의 협업 프로세스를 대상으로 수행되었는데, 오랜 기간 동안 피쳐드 아티클 레벨에 도달하지 못한 아티클들의 협업 프로세스 특성도 비교 연구할 필요가 있을 것으로 보인다. 효율적이지 못한 협업 프로세스와의 비교를 통해 주요 변수 값들의 차이를 분석함으로써 가상 지식 협업의 성공 요인을 좀더 정확하게 규명할 수 있을 것이다. 둘째, 본 연구는 아티클을 학문적 특성에 의한 인지적 노력의 요구 정도에 따라 두 그룹으로 분류했는데, 협업 프로세스 과정에서 논란의 여지가 많은 정도나 최신성이 중요한 정도를 반영하는 주제별 분류 등으로 좀더 세분화해서 접근해 볼 수 있겠다. 주제에 따라 서로 다른 지식 및 생각의 표현과 통합이 필수적인 협업 프로세스의 효율성이 달라질 수 있을 것이다. 셋째, 영어 위키피디아 외에 다른 언어 버전의 위키피디아나 오픈 소스 커뮤니티 플랫폼에서도 같은 결과를 얻을 수 있는지 확인해 보는 연구가 필요하다. 플랫폼마다 커뮤니티 문화나 운영 컨텍스트가 다르기 때문에 관련 변수 값의 범위나 변수들의 역학적 관계가 달라질 수 있기 때문이다.

불확실성으로 가득한 환경적인 어려움 속에서 시간과 공간의 한계를 초월하여 유용한 지식을 창출할 수 있게 해주는 가상 지식 협업은 여러 분야에서 새로운 가능성을 열어가고 있다. 작업 특성에 따른 헤드(head)와 테일(tail)의 유기적인 조화의 중요성을 입증한 본 연구가 가상 지식 협업을 통해 지속 가능한 경쟁 우위와 혁신을 달성하고자 하는 수많은 기업들에게 실질적인 도움을 줄 수 있기를 바란다.



## 참고문헌(References)

- Amabile, T. M., *The Atmosphere of Pure Work: Creativity in Research and Development*, In *The Social Psychology of Science*, W. R. Shadish and S. Fuler (eds.), Guilford Press, New York, 1994, 316~328.
- Ancona, D. G. and D. F. Caldwell, "Bridging the Boundary: External Activity and Performance in Organizational Teams," *Administrative Science Quarterly*, Vol.37, No.4(1992), 634~665.
- Anderson, C., *The Long Tail TV: Conclusion*, The Long Tail Blog, 2005. Available at [http://www.longtail.com/the\\_long\\_tail/2005/01/long\\_tail\\_tv\\_pa\\_2.html](http://www.longtail.com/the_long_tail/2005/01/long_tail_tv_pa_2.html) (Accessed 6 July, 2014).
- Anderson, C., *The Long Tail*, Wired, 2004. Available at <http://archive.wired.com/wired/archive/12.10/tail.html> (Accessed 6 July, 2014).
- Anderson, C., *The Long Tail: Why the Future of Business is Selling Less of More*, NY: Hyperion, New York, 2006.
- Anthony, D., S. W. Smith, and T. Williamson, "Reputation and Reliability in Collective Goods: The Case of the Online Encyclopedia Wikipedia," *Rationality and Society*, Vol.21, No.3(2009), 283~306.
- Bantel, K. A. and S. E. Jackson, "Top Management and Innovations in Banking: Does the Composition of the Top Team Make a Difference?" *Strategic Management Journal*, Vol.10, No.S1(1989), 107~124.
- Bock, G.-W. and Y.-G. Kim, "Breaking the Myths of Rewards: An Exploratory Study of Attitudes about Knowledge Sharing," *Information Resource Management Journal*, Vol.15, No.2 (2002), 14~21.
- Bock, G.-W., R. W. Zmud, Y.-G. Kim, and J.-N. Lee, "Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate," *MIS Quarterly*, Vol.29, No.1(2005), 87~111.
- Bono, J. E. and R. Ilies, "Charisma, Positive Emotions and Mood Contagion," *The Leadership Quarterly*, Vol.17, No.4(2006), 317~334.
- Byrne, D., *The Attraction Paradigm*, Academic Press, New York, 1971.
- Carr, N. G., "The Ignorance of Crowds," *strategy+business*, Vol.47(2007), 1~5.
- Chatman, J. A., "Matching People and Organizations: Selection and Socialization in Public Accounting Firms," *Administrative Science Quarterly*, Vo.36, No.3(1991), 459~484.
- Chen, X.-P., X. Yao, and S. Kotha, "Entrepreneur Passion and Preparedness in Business Plan Presentations: A Persuasion Analysis of Venture Capitalists' Funding Decisions," *Academy of Management Journal*, Vol.52, No.1(2009), 199~214.
- Chi, N. W., Y. M. Huang, and S. C. Lin, "A Double-Edged Sword? Exploring the Curvilinear Relationship between Organizational Tenure Diversity and Team Innovation: The Moderating Role of Team-Oriented HR Practices," *Group and Organization Management*, Vol.34, No.6 (2009), 698~726.
- Constant, D., L. Sproull, and S. Kiesler, "The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice,"

- Organization Science*, Vol.7, No.2(1996), 119~135.
- Cox, T. H. and S. Blake, "Managing Cultural Diversity: Implications for Organizational Competitiveness," *Academy of Management Executive*, Vol.5, No.3(1991), 45~56.
- Csikszentmihalyi, M., S. Abuhamdeh, and J. Nakamura, *Flow*, In *Handbook of Competence and Motivation*, A. Elliot et al. (eds.), The Guilford Press, New York, 2005, 598~698.
- Dahlin, K. B., L. R. Weingart, and P. J. Hinds, "Team Diversity and Information Use," *Academy of Management Journal*, Vol.48, No.6(2005), 1107~1123.
- Davenport, T. H., *Thinking for a Living: How to Get Better Performance and Results from Knowledge Workers*, Harvard Business School Press, Boston, 2006.
- Dawes, R. M., "Social Dilemmas," *Annual Review of Psychology*, Vol.31(1980), 169~193.
- Faraj, S., S. L. Jarvenpaa, and A. Majchrzak, "Knowledge Collaboration in Online Communities," *Organization Science*, Vol.22, No.5(2011), 1224~1239.
- Li, F., Y. Li, and E. Wang, "Task Characteristics and Team Performance: the Mediating Effect of Team Member Satisfaction," *Social Behavior and Personality: An International Journal*, Vol.37, No.10(2009), 1373~1382.
- Geister, S., U. Konradt, and G. Hertel, "Effects of Process Feedback on Motivation, Satisfaction, and Performance in Virtual Teams," *Small Group Research*, Vol.37, No.5(2006), 459~489.
- Gibbert, M. and H. Krause, *Practice Exchange in a Best Practice Marketplace*, In *Knowledge Management Case Book: Siemens Best Practices*, T. H. Davenport and G. J. B. Probst (eds.), Publicis Corporate Publishing, Erlangen, Germany, 2002, 89~1015.
- Gini, C., "Concentration and Dependency Ratios (in Italian)," English Translation in *Rivista di Politica Economica*, Vol.87(1997), 769~789.
- Grant, R. M., "Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration," *Organization Science*, Vol.7, No.4(1996), 375~386.
- Gruenfeld, D. H., E. A. Mannix, K. Y. Williams, and M. A. Neale, "Group Composition and Decision Making: How Member Familiarity and Information Distribution Affect Process and Performance," *Organizational Behavior and Human Decision Processes*, Vol.67, No.1(1996), 1~15.
- Hackman, J. R. and G. R. Oldham, "Motivation through the Design of Work: Test of a Theory," *Organizational Behavior and Human Performance*, Vol.16(1976), 250~279.
- Highsmith, J. A., "Solving Design Problems More Effectively," *MIS Quarterly*, Vol.2, No.4(1978), 23~30.
- Hoffman, L. R., *The Group Problem-Solving Process*, In *Group Processes*, L. Berkowitz (ed.), Academic Press, New York, 1978, 101~114.
- Hoffman, L. R. and N. R. F. Maier, "Quality and Acceptance of Problem Solutions by Members of Homogeneous and Heterogeneous Groups," *Journal of Abnormal and Social Psychology*, Vol.62, No.2(1961), 401~407.
- Hsu, M.-H., T. L. Ju, C.-H. Yen, and C.-M. Chang, "Knowledge Sharing Behavior in Virtual Communities: The Relationship between Trust, Self-Efficacy, and Outcome Expectations,"

- International Journal of Human-Computer Studies*, Vol.65, No.2(2007), 153~169.
- Humphrey S. E., J. D. Nahrgang, and F. P. Morgeson, "Integrating Motivational, Social, and Contextual Work Design Features: A Meta-analytic Summary and Theoretical Extension of the Work Design Literature," *Journal of Applied Psychology*, Vol.92, No.5 (2007), 1332~1356.
- Jackson, S., *Team Composition in Organizations*, In *Group Process and Productivity*, S. Worchel, W. Wood, and J. Simpson (eds.), Sage, London, 1992, 1~12.
- Jackson, S. E., A. Joshi, and N. L. Erhardt, "Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications," *Journal of Management*, Vol.29, No.6(2003), 801~803.
- Jehn, K. A., G. B. Northcraft, and M. A. Neale, "Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups," *Administrative Science Quarterly*, Vol.44, No.4(1999), 741~763.
- Kankanhalli, A., B. Tan, and K. Wei, "Contribution Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly*, Vol.29, No.1(2005), 113~143.
- Kiesler, S., J. Siegel, and T. W. McGuire, "Social Psychological Aspects of Computer-Mediated Communication," *American Psychologist*, Vol. 39, No.10(1984), 1123~1134.
- Kollock, P., *The Economies of Online Cooperation: Gifts and Public Goods in Cyberspace*, In *Communities in Cyberspace*, M. Smith and P. Kollock (eds.), Routledge, New York, 1999, 220~239.
- Lakhani, K. and E. von Hippel, "How Open Source Software Works: 'Free' User-to-User Assistance," *Research Policy*, Vol.32, No.6 (2003), 923~943.
- Lakhani, K. and R. Wolf, *Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects*, In *Perspectives in Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani (eds.), MIT, Cambridge, 2005.
- Levenson, A., "Talent Management: Challenges of Building Cross-functional Capability in High-performance Work Systems Environments," *Asia Pacific Journal of Human Resources*, Vol.50, No.2(2012), 187~204.
- Li, F., Y. Li, and E. Wang, "Task Characteristics and Team Performance: The Mediating Effect of Team Member Satisfaction," *Social Behavior and Personality*, Vol.37, No.10(2009), 1373~1382.
- Lin, N., *Social Capital*, Cambridge University Press, Cambridge, UK, 2001.
- Mackay, J. M., S. H. Barr, and M. G. Kletke, "An Empirical Investigation of the Effects of Decision Aids on Problem-Solving Processes," *Decision Sciences*, Vol.23, No.3(1992), 648~672.
- McFadyen, M. A. and A. A. Cannella, "Social Capital and Knowledge Creation: Diminishing Returns of the Number and Strength of Exchange," *The Academy of Management Journal*, Vol.47, No.5(2004), 735~746.
- Milliken, F. and L. Martins, "Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups," *Academy of Management Review*, Vol.21, No.2(1996), 402~433.

- Morgeson, F. P. and S. E. Humphrey, "The Work Design Questionnaire (WDQ): Developing and Validating Comprehensive Measure for Assessing Job Design and the Nature of Work," *Journal of Applied Psychology*, Vol. 91, No.6(2006), 1321~1339.
- Morgeson, F. P. and S. E. Humphrey, *Job and Team Design: Toward a More Integrative Conceptualization of Work Design*, In *Research in Personnel and Human Resource Management*, J. Martocchio (ed.), Emerald Group Publishing Limited, London, 2008, 39~92.
- Nahapiet, J. and S. Ghoshal, "Social capital, intellectual capital and the organizational advantage," *Academy of Management Review*, Vol.23, No.2(1998), 242~266.
- Nemeth, C., "Differential Contributions of Majority and Minority Influence," *Psychological Review*, Vol.93, No.1(1986), 23~32.
- Nemoto, K., P. A. Gloor, and R. Laubacher, "Social Capital Increases Efficiency of Collaboration among Wikipedia Editors," *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, (2011), 231~240.
- Nonaka, I., "A Dynamic Theory of Organizational Knowledge Creation," *Organization Science*, Vol.5, No.1(1994), 14~37.
- Nonaka, I., K. Umemoto, and D. Senoo, "From Information Processing to Knowledge Creation: A Paradigm Shift in Business Management," *Technology in Society*, Vol.18, No.2(1996), 203~218.
- Oldham, G. and A. Cummings, "Creativity in the Organizational Context," *Productivity*, Vol.39, No.2(1998), 187~194.
- Oldham, G. R. and J. R. Hackman, "Not What It Was and Not What It Will Be: The Future of Job Design Research," *Journal of Organizational Behavior*, Vol.31, No.2/3(2010), 463~479.
- O'Connor, K. M. and W. L. Adair, *Integrative Interests? Building a Bridge Between Negotiation Research and the Dynamic Organization*, In *Leading and Managing People in the Dynamic Organization*, R. S. Petersen and E. A. Mannix (eds.), Lawrence Erlbaum Associates, Mahwah, NJ, 2003, 163~184.
- O'Reilly, C., "Corporations, Culture, and Commitment: Motivation and Social Control in Organizations," *California Management Review*, Vol.31, No.4 (1989), 9~25.
- O'Reilly III, D. F. Caldwell, and W. P. Barnett, "Work Group Demography, Social Integration, and Turnover," *Administrative Science Quarterly*, Vol.34, No.1(1989), 21~37.
- Park, S. J., J. W. Kim, H. J. Lee, H. Park, D. Han, and P. Gloor, "Exploration of the Online Culture for Virtual Multinational Teams through Multilingual Network Analysis of Wikipedia," Paper Under Review.
- Pelled, L. H., K. M. Eisenhardt, and K. R. Xin, "Exploring the Black Box: An Analysis of Work Group Diversity, Conflict, and Performance," *Administrative Science Quarterly*, Vol.44, No.1 (1999), 1~28.
- Raymond, E. S., *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, O'Reilly, CA, 2001.
- Reagans, R. and E. W. Zuckerman, "Network, Diversity, and Productivity: The Social Capital of Corporate R&D Teams," *Organization Science*, Vol.12, No.4(2001), 502~517.
- Short, J., E. Williams, and B. Christie, *The Social Psychology of Telecommunications*, John Wiley & Sons, Ltd., London, 1976.

- Simon, H. A., *The New Science of Management Decision*, Harper & Brothers Publishers, New York, 1960.
- Smith, K. G., C. J. Collins, and K. D. Clark, "Existing Knowledge, Knowledge Creation Capability, and the Rate of New Product Introduction in High-Technology Firms," *Academy of Management Journal*, Vol.48, No.2(2005), 346~357.
- Spreitzer, G. M., S. G. Cohen, and J. G. Ledford, "Developing Effective Self-managing Work Teams in Service Organizations," *Group and Organization Management*, Vol.24, No.3(1999), 340~366.
- Sproull, L. and S. Kiesler, "Reducing Social Context Cues: Electronic Mail in Organizational Communication," *Management Science*, Vol.32, No.11(1986), 1492~1512.
- Steiner, I. D., *Group Process and Productivity*, Academic Press, San Diego, 1972.
- Stewart, G. L., "A Meta-analytic Review of Relationships Between Team Design Features and Team Performance," *Journal of Management*, Vol.32, No.1(2006), 29~55.
- Tsui, A. S. and C. A. O'Reilly, "Beyond Simple Demographic Effects: The Importance of Relational Demography in Superior-Subordinate Dyads," *Academy of Management Journal*, Vol.32, No.2(1989), 402~423.
- Van der Vegt, G. S. and J. S. Bunderson, "Learning and Performance in Multidisciplinary Teams: The Importance of Collective Team Identification," *Academy of Management Journal*, Vol.48, No.3(2005), 532~547.
- Van Maanen, J. and E. Schein, *Toward a Theory of Organizational Socialization*, In *Research in Organizational Behavior*, L. L. Cummings and B. M. Staw (eds.), CT: JAI Press, Greenwich, 1979, 209~264.
- Wasko, M. and S. Faraj, "It Is What One Does: Why People Participate and Help Others in Electronic Communities of Practice," *Journal of Strategic Information Systems*, Vol.9, No.2/3(2000), 155~173.
- Wasko, M. and S. Faraj, "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly*, Vol.29, No.1(2005), 35~57.
- Williams, K. Y. and C. A. O'Reilly, *Demography and Diversity in Organizations*, In *Research in Organizational Behavior*, B. M. Staw and R. M. Sutton (eds.), JAI Press, 1998, 77~140.
- Yang, H.-L. and C.-Y. Lai, "Motivations of Wikipedia Content Contributors," *Computers in Human Behavior*, Vol.26, No.6(2010), 1377~1383.

# Appendix

〈Table A1〉 Descriptive Statistics

Variable	Minimum	Maximum	Mean	Std. Dev.
1. NumOfEditors	2	5782	247.42	500.334
2. RegUserRatio (%)	25.05	100.00	70.41	18.81
3. DiscUserRatio (%)	0.00	80.00	11.31	8.71
4. InterArrivalTime (days)	0.02	58.67	3.86	4.10
5. RevisitingPeriod (days)	0.33	238.92	35.05	26.87
6. NumOfSections	0	43	14.02	6.11
7. NumOfReferences	0	456	67.45	50.82
Gini Coefficient	0.19	0.95	0.69	0.13
Pareto Ratio (%)	36.17	98.17	75.93	11.23

〈Table A2〉 Variable Correlations

Variable	1	2	3	4	5	6	7
1. NumOfEditors	1						
2. RegUserRatio	-.553	1					
3. DiscUserRatio	-.317	.616	1				
4. InterArrivalTime	-.271	.292	.036	1			
5. RevisitingPeriod	-.131	.228	.110	.337	1		
6. NumOfSections	.343	-.376	-.264	-.297	-.028	1	
7. NumOfReferences	.401	-.304	-.236	-.244	.035	.590	1
Gini Coefficient	-.358	.484	.304	-.171	.193	.032	.090
Pareto Ratio	-.354	.473	.282	-.163	.186	.024	.073

Abstract

## **Pareto Ratio and Inequality Level of Knowledge Sharing in Virtual Knowledge Collaboration: Analysis of Behaviors on Wikipedia\***

Hyun-jung Park\*\* · Kyung-shik Shin\*\*\*

The Pareto principle, also known as the 80-20 rule, states that roughly 80% of the effects come from 20% of the causes for many events including natural phenomena. It has been recognized as a golden rule in business with a wide application of such discovery like 20 percent of customers resulting in 80 percent of total sales. On the other hand, the Long Tail theory, pointing out that “the trivial many” produces more value than “the vital few,” has gained popularity in recent times with a tremendous reduction of distribution and inventory costs through the development of ICT(Information and Communication Technology). This study started with a view to illuminating how these two primary business paradigms – Pareto principle and Long Tail theory – relates to the success of virtual knowledge collaboration. The importance of virtual knowledge collaboration is soaring in this era of globalization and virtualization transcending geographical and temporal constraints. Many previous studies on knowledge sharing have focused on the factors to affect knowledge sharing, seeking to boost individual knowledge sharing and resolve the social dilemma caused from the fact that rational individuals are likely to rather consume than contribute knowledge. Knowledge collaboration can be defined as the creation of knowledge by not only sharing knowledge, but also by transforming and integrating such knowledge. In this perspective of knowledge collaboration, the relative distribution of knowledge sharing among participants can count as much as the absolute amounts of individual knowledge sharing. In particular, whether the more contribution of the upper 20 percent of participants in knowledge sharing will enhance the efficiency of overall knowledge collaboration is an issue of interest. This study deals with the effect of this sort of knowledge sharing distribution on the efficiency of knowledge collaboration and is extended to reflect the work characteristics. All analyses were conducted

---

\* This work was supported by the National Research Foundation of Korea Grant funded by the Korean government (NRF-2013S1A3A2054667).

\*\* Management Research Center, Ewha Womans University

\*\*\* Corresponding author: Kyung-shik Shin

School of Business, Ewha Womans University

52 Ewhayeodae-gil, Seodaemun-gu, Seoul, 120-750, Korea

Tel: +82-2-3277-2799, Fax: +82-2-3277-2766, E-mail: ksshin@ewha.ac.kr

based on actual data instead of self-reported questionnaire surveys. More specifically, we analyzed the collaborative behaviors of editors of 2,978 English Wikipedia featured articles, which are the best quality grade of articles in English Wikipedia. We adopted Pareto ratio, the ratio of the number of knowledge contribution of the upper 20 percent of participants to the total number of knowledge contribution made by the total participants of an article group, to examine the effect of Pareto principle. In addition, Gini coefficient, which represents the inequality of income among a group of people, was applied to reveal the effect of inequality of knowledge contribution. Hypotheses were set up based on the assumption that the higher ratio of knowledge contribution by more highly motivated participants will lead to the higher collaboration efficiency, but if the ratio gets too high, the collaboration efficiency will be exacerbated because overall informational diversity is threatened and knowledge contribution of less motivated participants is intimidated. Cox regression models were formulated for each of the focal variables—Pareto ratio and Gini coefficient—with seven control variables such as the number of editors involved in an article, the average time length between successive edits of an article, the number of sections a featured article has, etc. The dependent variable of the Cox models is the time spent from article initiation to promotion to the featured article level, indicating the efficiency of knowledge collaboration. To examine whether the effects of the focal variables vary depending on the characteristics of a group task, we classified 2,978 featured articles into two categories: Academic and Non-academic. Academic articles refer to at least one paper published at an SCI, SSCI, A&HCI, or SCIE journal. We assumed that academic articles are more complex, entail more information processing and problem solving, and thus require more skill variety and expertise. The analysis results indicate the followings; First, Pareto ratio and inequality of knowledge sharing relates in a curvilinear fashion to the collaboration efficiency in an online community, promoting it to an optimal point and undermining it thereafter. Second, the curvilinear effect of Pareto ratio and inequality of knowledge sharing on the collaboration efficiency is more sensitive with a more academic task in an online community.

**Key Words** : Virtual Knowledge Collaboration, Knowledge Sharing, Pareto Ratio, Gini Coefficient, Online Community

Received: August 6, 2014    Revised: August 24, 2014    Accepted: August 27, 2014



## 저자 소개



### 박현정

현재 이화여자대학교 경영연구소 연구교수로 재직 중이다. 한국과학기술원(KAIST) 경영과학과에서 공학사와 공학석사 학위를 취득하였으며, 서울대학교 경영학과에서 시맨틱 웹 자원의 랭킹에 관한 연구로 경영학 박사학위를 취득하였다. 주요 연구분야는 가상 지식 협업(Virtual Knowledge Collaboration), 소셜 네트워크 분석(Social Network Analysis), 비즈니스 인텔리전스(Business Intelligence), 시맨틱 랭킹(Semantic Ranking), 빅 데이터 분석 및 비즈니스 응용(Bid Data Analytics and Business Application) 등이다.



### 신경식

현재 이화여자대학교 경영대학 경영학부 교수로 재직 중이다. 신경식 교수는 연세대학교 경영학과를 졸업하고 미국 George Washington University에서 MBA, 한국과학기술원(KAIST)에서 인공지능, 지식기반 시스템 등 지능형 기법을 경영분야에 적용하는 연구로 경영공학 Ph.D.를 취득하였다. 주요 연구분야로는, 데이터 마이닝(Data Mining)과 비즈니스 인텔리전스(Business Intelligence), 빅 데이터(Big Data) 분석 / 비즈니스 애널리틱스(Business Analytics), 인공지능(Artificial Intelligence) 응용과 지식공학(Knowledge Engineering) 등이다.