

메가아시아 탐색을 위한 새로운 양적 지역연구방법론의 필요성과 가능성: 국내 변수중심 지역연구를 중심으로*

허정원** 서울대학교 아시아연구소 HK'연구교수

박선영 서울대학교 아시아연구소 HK'연구교수

장효진 서울대학교 지리학과 박사수로

심우진*** 서울대학교 아시아연구소 HK'연구교수

본 논문은 KCI에 등재된 지역학 학술지의 논문 중 변수중심의 양적방법론을 이용한 기존 연구를 정리하여 데이터와 방법론의 특징과 동향, 한계를 파악하고, 이를 바탕으로 새로운 양적방법론적 접근의 가능성을 타진하는 것을 목적으로 한다. 이를 위해 등재 학술지를 중심으로 지역 혹은 국가를 분석한 연구들의 키워드와 활용 데이터, 방법론 등을 추출하고 지역별·시계열별 분석과 비교 대상이 된 지역 간의 관계 동향을 분석하였다. 그 결과, 2000년대 이후 지역학 학술지의 발간이 크게 늘었으며 특히 아시아를 대상지역으로 하는 학술지가 크게 증가하였다. 하위 지역별로 보면 한국과 지리적으로 인접한 동아시아와 동남아시아를 대상으로 한 학술지의 발간이 증가하였다. 국가를 대상으로 한 연구가 지역을 대상으로 한 연구에 비해 더 많이 나타났고, 이를 지역별로 살펴보면 아시아와 유럽을 대상으로 한 연구가 많은 것을 확인하였다. 또한 연구에 활용한 데이터의 출처는 국제기구 혹은 정부 통계가 주를 이루고 있어 한정된 기구의 데이터를 의존하는 경향과 그에 따른 한계를 볼 수 있다. 그러나 스케일 의존도가 낮아 지역·국가·지방 비교에 모두 활용할 수 있고 경계에 구속받지 않는 공간자료의 활용은 드물게 나타나, 여러 형태의 데이터를 포괄적으로 고려하는 방법론의 개발과 적용을 통한 새로운 데이터 중심 연구방법의 가능성을 확인할 수 있었다.

주제어 메가아시아, 양적방법론, 비교지역연구, 공간자료, 매크로 데이터

I. 지역연구에서 양적 연구방법의 의미와 필요성

지역연구는 특정 국가나 지역의 맥락을 깊이 있게 이해함으로써 연구문제에 대한 답을 얻고자 하는 학제적인(interdisciplinary) 학문이다(이희연·최재현, 1998; 안승

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A6A3A02065553).

** 주저자

*** 교신저자

국 외, 2002). 지역연구는 주로 정치학, 인류학, 지리학 등 다양한 학문분야의 협업이 이루어지는 복합적인 학문으로써 각 학문의 방법론을 활용하여 지역의 사회현상에 대한 입지적 또는 공간적 접근 방법이 강조된다(이중희, 2001; 이창식, 2016). 답사 혹은 현지조사를 통하여 관심 국가 혹은 지역의 공간적·시간적 맥락성을 탐구하여 깊이 있는 이해를 구하는 단일사례연구가 전통적인 지역연구의 방법이다. 그런데 최근 지역연구 방법론 중 두 지역 이상을 비교함으로써 지역 간 일반성이나 지역 내 특수성을 파악하여 궁극적으로 지역성을 밝히는 것을 목적으로 하는 연구방법의 필요성이 대두되고 있다(이송호, 2009).

전통적으로 사회과학 연구방법론은 사례중심의 질적방법론적 접근과 변수중심의 양적방법론적 접근으로 구분되어 왔다(홍경준, 2002; 김인숙, 2007; 강근복, 2009). 변수중심의 양적방법론적 접근은 변수 지향적이며 이론 중심적인 방법론이다. 양적 연구의 기본적인 토대를 이루는 것은 검증할 가설을 설정하고 관측을 통해 모집단을 폭넓게 선정하며 이들의 상호연관성을 탐구하는 것이다. 이를 통해 다수의 자료를 기초로 한 변수들을 분석함으로써 지역 간의 상관관계를 규명하여 일반성을 도출하는 것을 목적으로 한다. 그러나 인간의 정신적·주관적 영역과 같이 계량화하기 어려운 분야는 연구가 어렵다는 점, 사회·문화 현상을 인간의 가치나 의도와 분리하여 연구하기 때문에 인간의 내면에 대한 심층적인 이해가 어렵다는 점 등의 인과성을 입증하지 못한다는 한계점이 있다(강근복, 2009; 박승배, 2009).

반면, 사례중심의 질적방법론적 접근은 중요한 사실을 설명하고, 일련의 일어나는 현상에 대한 원인을 탐구하는 방법론이다. 질적 연구는 어떠한 현상을 설명하기 위해 구체적인 데이터를 수집하고 이 자료를 중심으로 현상을 설명하는 연구방법론이다. 질적연구방법론은 양적연구방법론의 한계를 극복하기 위해 보완되었고, 1970년대 이후에 현상과 인간을 대상으로 하는 연구에 주로 적용되고 있다. 하지만 사례중심의 질적방법론은 복잡성을 중시하고 특수성을 조망하는데 탁월하지만, 개별 연구자의 주관이 개입되거나 미시적·개인적인 연구 등으로 일반화하기 어렵다는 문제로 인해 연구 효용성의 논란이 있다(홍서영, 서태열, 2014). 따라서 사례를 포괄하는 이론적 일반화를 도출하는 데는 한계점을 갖는다(조용기, 2001; 이희영, 2005; 최명민, 2007).

기존의 지역연구는 주로 질적방법론으로 접근해 왔다. 데이터를 이용한 변수 중심의 실증분석 방법인 양적방법론이 상대적으로 거시적인 특징을 갖기 때문에 지역에 대한 깊은 맥락적 이해를 추구하는 지역연구에 있어서는 적합하지 않다는 견해가 있었기 때문이다(김경수, 2006). 그러나 지역연구 영역에서 양적방법론에 대한 부적합하다는 의문은 양적방법론 자체에 대한 의문이라기보다는 지역성을 탐구하기에 적합한 양적으로 많고 촘촘한 지역데이터를 활용하기에 기획과 수집에 막대한 비용이 소요되고 연구자가 직접 수집하기 어렵고 가용데이터의 범위가 지극히 한정되어 있기 때문이다. 또한 지역연구자들은 깊이 있는 연구를 통하여 지역의 독특성과 맥락성에 대한 통찰을 얻는 것을 목적으로 하기에 대상 지역에 대한 비교연구의 필요성이 그동안 비교적 낮게 평가되어 왔다. 그러나 하나의 국가나 지역연구의 결과만 제시하기보다 비교연구방법을 통하여 단일 사례로부터 관찰한 연구 질문의 중요성과 논지의 타당성을 높이고 궁극적으로 사회과학 이론의 발전에도 기여할 수 있다는 점에서 지역연구의 방법론을 택하는 것은 의미가 있다(신재혁, 2019).

최근 국제기구, 개별국가, 지역공동체, 민간 기업 등의 여러 주체들이 국가와 지역에 대한 다양한 데이터를 수집하여 활용하고 있으며 ICT 기술의 발달로 이러한 데이터를 연구자들뿐만 아니라 일반인들도 공유하여 사용할 수 있게 되었다(안종욱 외, 2013; 구자용, 2015). 또 설계와 수집에 많은 비용이 소요되는 기존의 정량적 데이터베이스와는 달리 대량의 정형 혹은 비정형 데이터가 공간적 경계를 넘나들며 생산되는 빅데이터의 출현은 지역연구에서 새로운 방법론의 가능성을 의미한다(유성민, 2015; 조남경, 2019). 빅데이터는 데이터의 크기만을 의미하는 것이 아니라 대용량 데이터를 분석하여 가치 있는 정보를 추출하고 생성된 지식을 바탕으로 현상을 분류하여 미래를 예측하는 정보화 기술까지 포괄하는 개념이다(송태민, 2013). 빅데이터 분석은 이론에 근거한 가설을 입증하기 위하여 세심하게 준비된 표본, 즉 데이터를 통하여 모집단을 추정·검정하는 전통적인 양적방법론과 다르다. 빅데이터 분석은 주어진 데이터의 분석과 예측이 이루어지기에 양적방법론의 이론에 근거한 인과관계 추정모델링과는 다르다(조남경, 2019).

본 연구는 한국에서 수행된 변수중심의 양적방법론을 이용한 지역비교연구를 정리하여 기존 데이터와 방법론의 특징과 동향, 한계를 파악하고 이를 바탕

으로 새로운 변수중심의 양적방법론적 접근 가능성을 타진하는 것을 목적으로 한다.

II. 지역학 학술지에 나타난 양적 연구방법의 특징과 동향

1. 연구 자료와 방법

본 연구는 지역학 학술지 중 변수중심의 양적방법론을 활용한 논문들을 대상으로 한다. 연구자료는 한국학술지인용색인(Korea Citation Index, KCI)¹에서 주제분류를 지역학으로 구분해 놓은 학술지의 논문들을 대상으로 구축했다. 이들 중 KCI 등재지(우수등재 포함)를 기준으로 보면 총 46개의 학술지가 있지만 『한국도서연구』, 『대구경북연구』, *Seoul Journal of Korean Studies* 등 국내 지역연구만을 대상으로 하는 학술지 3개는 제외하였다.

최종적으로 43개 등재 학술지의 1만 3,999개 논문을 대상으로 양적 연구를 활용한 논문 798개를 추출하였다. 이후 각각의 논문에서 밝히고 있는 연구지역, 연구 방법, 데이터 출처, 논문 키워드, 출판연도, 초록 등을 DB로 구축하였다.

구축한 DB를 활용하여 학술지명에서 파악할 수 있는 관심 연구지역의 변화, 연구 대상지의 스케일, 연구지역의 시계열적 변화, 연구 방법과 출처의 변화, 연구의 대상지역으로 살펴본 각 지역의 관련성 변화 등의 분석을 실시하였다. 각각의 방법과 용어 정의는 다음과 같다.

첫째, 학술지의 창간연도와 주 연구지역을 통하여 한국 지역연구의 변화과정과 특징을 살펴보기 위해 43개 지역학 등재 학술지의 창간연도와 주 연구지역을 정리하여 활용하였다. 주 연구지역은 학술지명에 특정 지역이나 국가가 명시된 경우 해당 지역으로 구분하였고 명시하지 않은 경우에는 ‘일반’ 항목으로 구분하였다.

둘째, 연구 대상지의 스케일은 국가와 지역으로 구분했으며 아시아, 동유럽,

¹ <https://www.kci.go.kr/kciportal/main.kci>

표 1 아시아 5개 지역과 이에 속한 국가

아시아 지역	국가
동아시아	한국, 중국, 일본, 몽골, (대만), (홍콩), (마카오)
동남아시아	동티모르, 라오스, 말레이시아, 미얀마, 베트남, 브루나이, 싱가포르, 인도네시아, 캄보디아, 태국, 필리핀
남아시아	네팔, 몰디브, 방글라데시, 부탄, 스리랑카, 아프가니스탄, 인도, 파키스탄
서아시아	레바논, 바레인, 사우디아라비아, 시리아, 아랍에미레이트, 아르메니아, 아제르바이젠, 예멘, 오만, 요르단, 이라크, 이란, 이스라엘, 조지아, 카타르, 쿠웨이트, 키프로스, 터키, 팔레스타인
중앙아시아	우즈베키스탄, 카자흐스탄, 키르기스스탄, 타지키스탄, 투르크메니스탄

마그립 지역, 아세안 등 국가가 아닌 여러 스케일의 지역으로 접근한 연구는 지역으로 구분하였다.

셋째, 연구 대상 지역의 시계열적 변화를 파악하기 위해 국가 단위의 연구인 경우, UN에서 밝히고 있는 지역·국가 기준에 의거하여 지역별로 재분류하였다.² 단, UN 기준상 남아시아로 분류된 이란의 경우 아시아연구소 내부 의견을 반영하여 서아시아로 집계하였다. 이를 통해 아시아에 속한 국가들을 지역별로 구분하면 표 1과 같다.

넷째, 데이터의 출처는 국제, 국가, 설문, 자체로 크게 4가지로 구분하였다. World Bank, OECD, IMF 등의 국제기구 데이터를 이용한 경우는 국제, 각국 정부나 기관의 데이터를 이용한 경우는 국가로 분류하였다. 연구자가 서베이를 통해 인식조사 등을 수행한 경우는 설문으로, SNS, 유튜브, 문헌 등의 자료를 자체 구축하여 이용한 경우는 자체로 구분한다.

다섯째, 지역 간 관계성을 살펴보기 위해 두 국가 혹은 지역 이상을 사례로 하는 연구를 대상으로 해당 지역들을 추출하고 이들 사이의 연결을 수치로 집계했다. 이 과정에서 한 연구에서 여러 지역을 비교할 경우 동일 지역이 중복으

² UN은 통계 활용을 위한 지역 및 국가 표준코드를 통해 해당 국가가 속한 대륙과 지역을 구분한다. 이는 흔히 M49 표준(standard)으로 불리는데 Standard Country or Area Codes for Statistical Use of the UN 온라인 출판물 간행번호인 Series M, No. 49에서 유래했다(<https://unstats.un.org/unsd/methodology/m49/>).

로 집계되는 문제를 고려하여 자료를 구축하였고 이를 2004년부터 2021년까지 크게 세 시기로 구분하고 지역들의 관계 변화를 살펴보았다.

2. 분석 결과

43개 학술지에서 주 연구지역을 추출하여 학술지 출간 시기별 변화를 살펴본 결과는 그림 1과 같다.

먼저, 학술지 출간 시기별 개수를 살펴보면 1980년대 이전에는 5개, 1980~1989년에 10개, 1990~1999년에 10개, 2000~2009년에 21개, 2010년 이후에 7개가 발간되었다. 이와 같이 지역학 학술지는 2000~2009년 사이에 가장 많이 발간된 것을 알 수 있다. 특히 아시아를 대상으로 하는 학술지가 2000년대 이후 급격히 증가하는 것은 주목할 만하다. 1997년 아시아발 금융위기를 성공적으로 극복한 후, 한국은 사회주의권 붕괴와 함께 시작된 신자유주의적 글로벌 경제질서 개편의 물결 속에서 시장을 개방함과 동시에 해외로 진출하기 위하여 해외 지역연구에 대한 수요가 급증하였다. 주로 정치·외교와 경제·무역의 연구가 주를 이루었던 1990년대와 달리 2000년 이후 아시아 출신 이주 노동자와 결혼 이주자의 증가로 사회·문화 분야의 연구도 급증하였다(김형준·전제성, 2019; 정범모, 2019). 이 시기에는 학술논문뿐 아니라 연구 기관의 정책보고서도 크게 증가하였고 발간기관 또한 다양화되었다(김형준·전제성, 2019). 2000년 이후 아시아 출신의

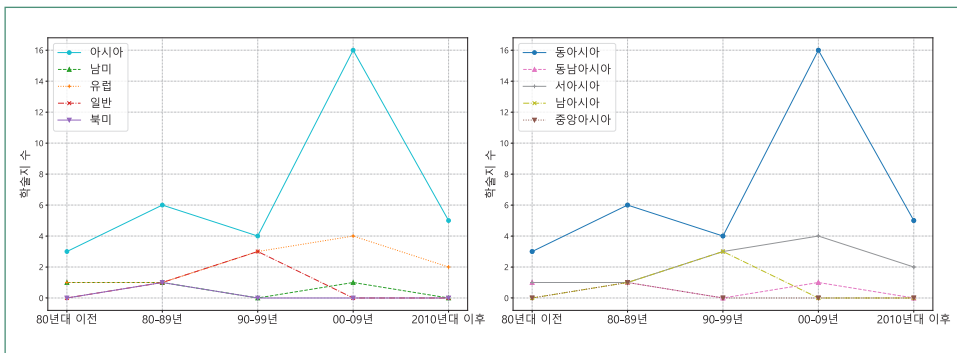


그림 1 지역학 학술지 출간 시기별 주 연구지역의 변화(대륙(왼), 아시아 지역(오))

유학생이 급증하여 아시아 출신 학자들의 석박사 학위논문이 늘어난 것도 중요한 요인이다(김형중, 2019; 김형준·전제성, 2019).

지역학 학술지의 주 연구지역을 대륙별로 나누어 보면 아시아가 34개로 가장 많고 유럽이 10개로 그 뒤를 잇는다. 북미의 경우는 1980~1989년 사이에 1개만 나타나는데 이는 미국이나 캐나다 등을 대상으로 하는 연구는 매우 많으나 북미 지역만을 주 연구지역으로 하는 학술지는 지역학으로 분류되어 있지 않기 때문으로 보인다.

다음으로 아시아를 다섯 지역으로 나누어 시기별로 살펴보면 동아시아는 2000~2009년에 9개, 동남아시아는 1980~1989년에 2개와 2000~2009년에 2개, 서아시아는 2000~2009년에 3개, 남아시아는 1990~1999년에 2개, 중앙아시아는 1980~1989년·2000~2009년에 각 1개, 북미는 1980~1989년에 1개, 남미는 1980년대 이전·1980~1989년·2000~2009년에 각 1개, 유럽은 2000~2009년에 4개로 나타난다.

지역연구 중 양적방법론을 이용한 798개 논문에서 연구지역을 국가 단위로 접근한 연구와 지역 단위로 접근한 연구의 시계열적 차이를 살펴보면 그림 2와 같다.



그림 2 연구지역 스케일(국가, 지역)의 시계열 변화

전체로 보면 지역을 대상으로 하는 논문은 148개, 국가를 대상으로 하는 논문은 636개이다. 이를 시계열별로 살펴보면 지역을 대상으로 하는 논문은 2003년부터 2020년까지 꾸준히 증가하고 있지만, 2012~2015년에 10개에서 14개로 높아졌다가 이후 7~8개로 낮아지고, 다시 2020년에 21개로 가장 높아지는 것을 확인할 수 있다. 반면, 국가를 대상으로 하는 논문은 지역을 대상으로 하는 논문과 비교할 수 없을 만큼 연도별로 많은 것을 확인할 수 있다. 이는 데이터의 물리적 단위와 관련한다. 대부분의 연구에서 국제기구나 정부가 국가 단위로 생산한 1차 자료를 데이터로 활용하기 때문에 지역연구의 양적인 접근에서 스케일의 한계가 분명히 드러나는 것이다. 실제로 분석 대상 논문 중 지역 대상 논문의 대다수는 EU, ASEAN, NAFTA 등 지역공동체가 존재하고 이를 통해 지역 단위의 데이터를 얻을 수 있는 지역으로 한정된다.

다음으로 798개의 양적 지역연구 논문에서 밝히고 있는 연구지역을 대륙스케일의 지역으로 나누어 시계열적 변화를 보면 그림 3과 같다. 한 논문이 복수의 연구지역을 대상으로 한 경우 중복으로 집계했다.

먼저 지역별 총합으로 보면 아시아는 917개, 아메리카는 116개, 유럽은 408

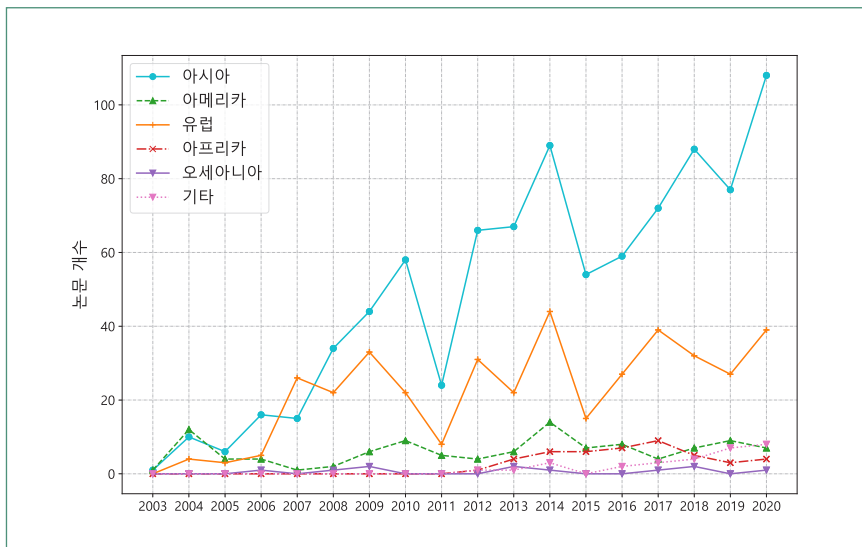


그림 3 양적방법론을 활용한 지역연구의 시계열별 연구지역(대륙)

개, 아프리카는 45개, 오세아니아는 12개, 기타지역(베콩강 지역 등)은 30개로 나타나 아시아 지역이 가장 많고 뒤를 이어 유럽, 아메리카 순으로 확인되었다. 이를 연도별로 살펴보면 아시아, 유럽 지역의 변화가 두드러지게 나타난 것을 볼 수 있다. 즉 아시아와 유럽은 지역연구의 대상지로 꾸준히 연구되어 왔고 그 수가 급격히 증가하는 추세인 반면, 아메리카와 아프리카, 오세아니아는 그렇지 않은 모습을 확인할 수 있다. 이는 물리적·심리적 거리가 상대적으로 가까운 아시아와 유럽이 다른 대륙에 비해 양적 지역연구의 대상으로 많이 활용되고 있는 것으로 보이지만, 해당 시기의 국제 정세, 경제 교류 상황 등 여러 요인들의 시기적 특성을 종합적으로 고려해야 정확한 이유를 파악할 수 있을 것이다.

연구 대상 지역에서 아시아 지역만을 분석한 결과를 보면 그림 4와 같다. 먼저 총합으로 보면 아시아 지역(917개)에 기타지역을 제외한 907개 중 동아시아 532개, 동남아시아 176개, 서아시아 57개, 남아시아 109개, 중앙아시아 33개로 확인되었다. 아시아 지역들 중에서도 한국 연구자들의 주 관심 대상은 동아시아와 동남아시아임을 확인할 수 있다.

이를 연도별로 살펴보면 아시아 지역들 모두 지역연구의 대상으로 꾸준히 증가하고 있음을 알 수 있다. 특히 동아시아는 중국 관련 연구들의 급증으로 인

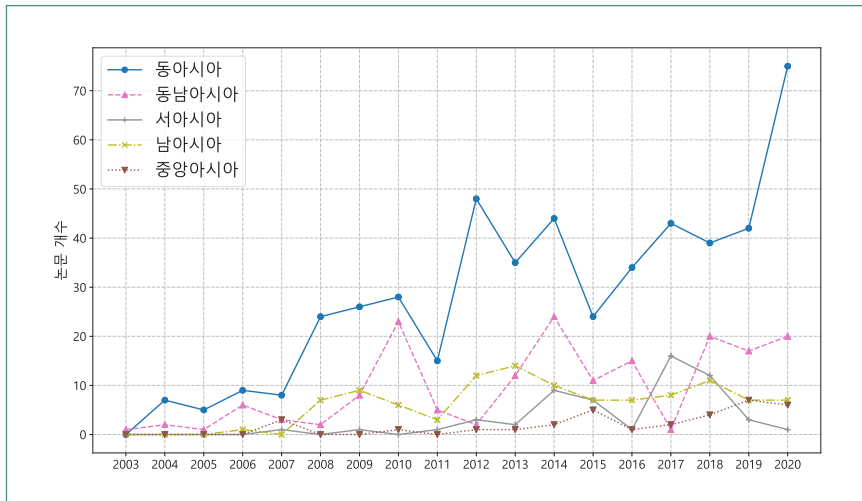


그림 4 양적방법론을 활용한 지역연구의 시계열별 연구대상 지역(아시아 지역)

해 가장 두드러진 상승 추세를 보인다. 또한 최근 동남아시아와 중앙아시아에 대한 관심이 지역연구로 이어지는 결과도 확인할 수 있으며 서아시아의 경우 2010년대 중반 이후 관련 연구가 급증한 모습도 주목할 만하다.

연구 대상인 798개 논문에서 사용한 데이터의 출처를 국제 데이터, 국가 데이터, 설문 데이터, 자체제작 데이터로 구분하여 살펴본 결과는 그림 5와 같다. 먼저 전체 데이터를 보면, 국제 데이터는 366개, 국가 데이터는 256개, 설문 데이터는 123개, 자체 데이터는 151개로 확인되었다. 양적 지역연구를 위한 데이터로 국제기구와 정부가 생산하는 1차 자료가 약 70%를 차지하고 있는 결과를 통해 양적 지역연구가 국가 단위로 이루어질 수밖에 없는 이유를 짐작할 수 있다. 다음으로 연도별로 보면 국제 데이터의 활용은 지속적으로 증가하는 추세를 보인다. 국가 데이터는 2020년에 26개로 가장 높게 나타나고, 2003·2005년에 각 2개, 2011년에 8개, 2015년에 12개로 보였다. 설문 데이터는 2005년부터 논문에서 활용되고 있는데, 2014년에 17개로 가장 높게 보였고 2010년에 2개, 2013년에 5개, 2015년에 8개, 2018년에 9개로 낮은 수치를 나타낸다. 자체 데이터는 2018년에 20개로 가장 많은 수치를 보였고, 2005년에 1개, 2011년에 6개, 2014년에 6개로 낮은 수치를 보였다. SNS, 유튜브 등을 이용하는 자체 데이터에서 주목할 부분은 2000년대 중반 이후 꾸준히 증가하고 있다는 것이다.

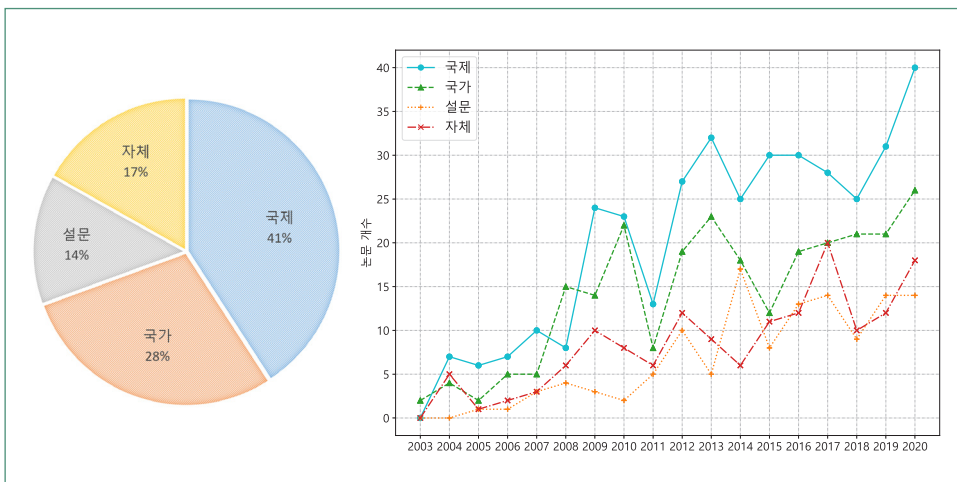


그림 5 양적 지역연구에 활용된 데이터의 출처별 비율(원)과 시계열 추이(오)

또한 798개 양적 지역연구 논문에서 밝히고 있는 연구 대상지를 지역 단위로 분류하여 지역 사이의 관계성 세 시기로 나누어 분석한 결과를 보면 그림 6과 같다.

먼저 2004년에서 2009년 사이에 지역 간 관계를 보면 동아시아와 유럽 사이는 총 9건으로 가장 많이 관련지어 연구를 진행한 것을 알 수 있다. 또한 동아시아와 동남아시아를 연결 지어 설명한 연구가 총 7건으로 그 뒤를 잇는다. 이 시기에는 한국과 중국이 속한 동아시아를 기준으로 타 지역과 연결 짓는 시도가 대다수이고 타 지역끼리 연관시켜 분석한 연구는 적었다.

다음으로 2010년에서 2015년 사이를 보면 동아시아와 동남아시아가 14건, 동아시아와 남아시아가 14건, 동아시아와 유럽이 10건, 동아시아와 북미가 8건, 동아시아와 서아시아가 4건, 동아시아와 남미가 4건으로 나타났다. 이는 동아시아와 타 지역 사이의 관계를 양적으로 파악하려는 시도가 2004~2009년에 비해 증가했다는 것을 알 수 있는 동시에 동아시아를 중심으로 그 연결성이 타 지역으로 확대되어 가는 것을 확인할 수 있는 부분이다. 특히 동아시아를 기준으로 유럽과 동남아시아를 주로 비교했던 시기에서 아시아의 다른 지역들과 관련시키려는 시도가 증가한 것은 주목할 만하다. 또한 유럽과 북미, 북미와 남미 등 동아시아를 제외한 타 지역 사이의 관계성도 증가했다.

마지막으로 2016년에서 2021년 사이에 지역 간 관계를 보면 동아시아와 유럽이 18건, 동아시아와 북미가 13건, 동아시아와 남아시아, 러시아가 각 11건, 동아시아와 동남아시아가 10건, 동아시아와 서아시아가 8건, 동아시아와 중앙아시아, 아프리카가 각 4건으로 아시아와 타 지역의 관계가 2010~2015년보다 더 다양해지고 깊어진 것을 확인할 수 있다. 또한 아시아의 지역들 간 연결성이 더 강해졌다는 것도 확인할 수 있다. 특히 기존에는 연결 지어 연구하지 않았던 러시아와 중앙아시아의 관계성도 증가한 것으로 나타났다.

각 시기별 특징을 정리해 보면 첫 번째 시기(2004~2009)에는 동아시아와 유럽, 동아시아와 동남아시아 지역의 사이의 관계성이 두드러지게 나타났고 두 번째 시기(2010~2015)에는 동아시아를 기준으로 타 지역으로의 관계가 다양해지고 깊어졌으며 마지막 세 번째 시기(2016~2021)에는 아시아 지역들 사이를 연결 짓는 시도가 증가한 것이 가장 큰 특징이라 할 수 있다.

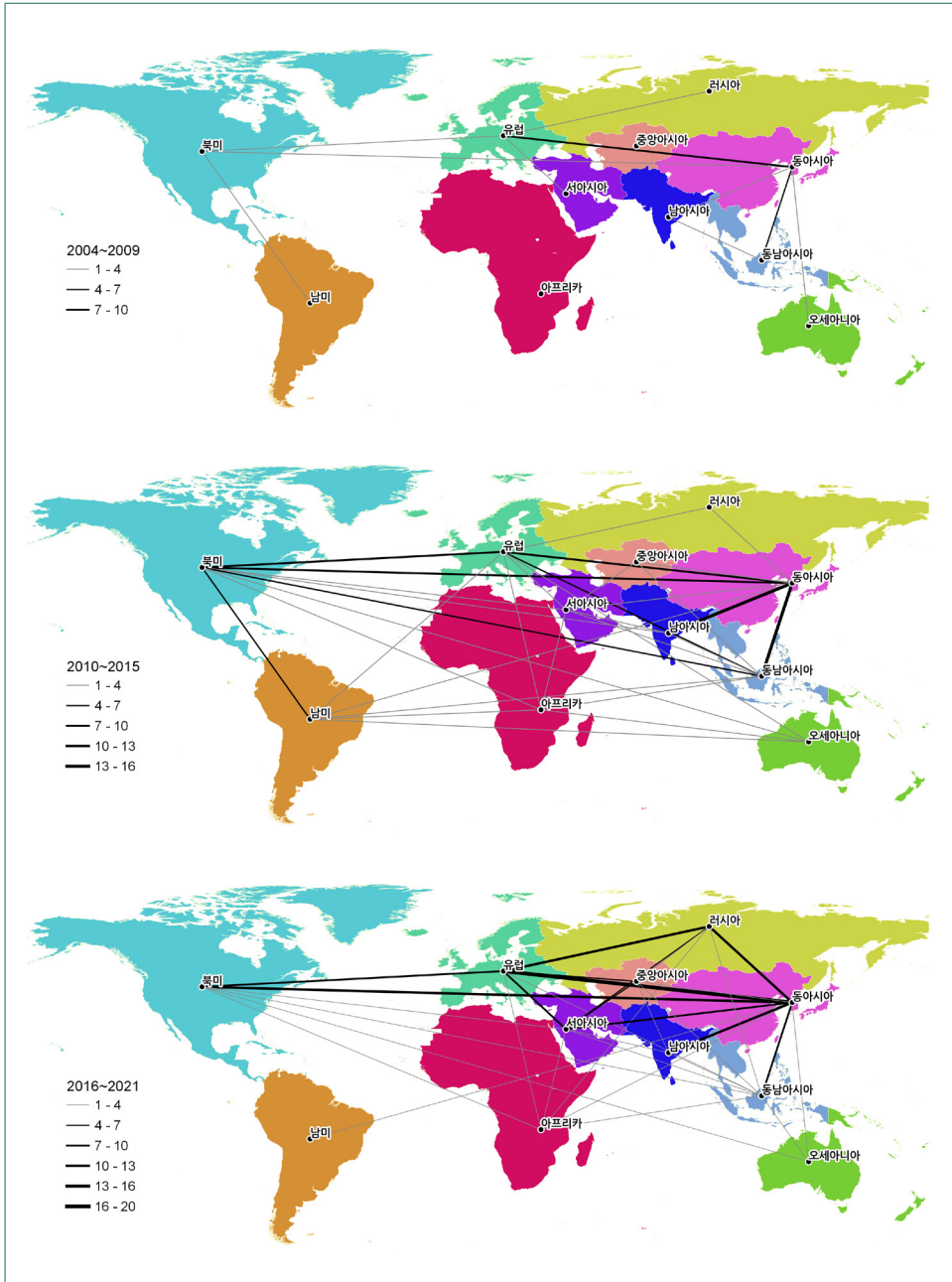


그림 6 양적 지역연구의 대상 지역(국가)로 살펴본 지역 간 관련성(2004~2009(상), 2010~2015(중), 2016~2021(하))

III. 메가아시아 탐색을 위한 새로운 지역연구방법 가능성 제안

이 장에서는 앞선 양적 지역연구방법의 특징과 동향을 분석한 결과를 기반으로 한 주요 논의사항을 정리함으로써 메가아시아 탐색을 위한 새로운 양적방법론의 가능성을 제안하고자 한다. 주요 논의사항은 크게 세 가지로 구분할 수 있다.

첫째 양적 지역연구에 있어서 국가별 데이터(매크로 데이터)의 높은 의존도, 둘째 지리적 근접성과 같은 공간적 맥락을 고려할 수 있는 공간자료 활용의 부재, 셋째 기술의 발전과 함께 대폭 증가한 데이터의 다양성과 지역연구로의 활용가능성 등이다.

먼저 앞장에서 살펴본 바와 같이 양적 연구방법을 이용한 지역연구의 대부분은 국가별 데이터에 의존한다. 양적 데이터의 수집은 데이터 구조의 설계, 일반화에 적합한 표본 추출, 실질적인 수집 과정까지 막대한 시간과 비용이 투입되며 국가와 지역의 장단기적인 발전을 계획하고 평가하는 데 필수적인 작업이다. 그러나 데이터 수집과정 자체는 특정 개인이나 기업에 즉각적인 이윤이 발생하지 않는 공공재로서 주로 국가나 국제기구가 주체가 되어 진행되기 때문에 대부분의 매크로 데이터는 국가 단위로 수집·관리된다. 해당 지역에 대하여 책무성을 가지는 정부, 즉 영토와 국민에 대한 이해가 높은 주체에 의해 수집되기에 비교적 데이터 품질이 높다. 또 그 경계가 비교적 명확하여 주제 영역별로 여러 국가의 다양한 데이터를 결합하고 비교분석하기에 용이한 장점도 있다. 그러나 국가별 데이터는 그 경계가 국가로 제한되어 국가 간의 경계지역에서 나타나는 특수한 현상을 포착할 수 없을 뿐만 아니라 국가들 사이의 지리적 접근성과 위치성도 고려할 수 없게 된다. 지리적 근접성과 위치성으로 대표되는 공간적인 맥락(시각)은 해당 지역의 지역성, 일반성, 특수성에 영향을 미치는 요소이므로 지역연구에서 굉장히 중요하다. 이는 거리와 위치에 따라 영향을 주고받는 정도가 달라지기 때문이다. 또한 ‘모든 것은 다른 모든 것과 관련되어 있으나, 가까운 것은 먼 것보다 더 관계가 깊다.’는 지리학 제1법칙에도 잘 나타나 있다 (Tobler, 1970).

이와 같은 한계를 지닌 국가별 데이터를 위주로 지역을 연구하다 보니 지역

연구이지만 지역 단위의 시각을 갖는 지역연구가 아닌 국가 단위의 지역연구가 진행될 수밖에 없다.

그러므로 양적 연구 시 국가별 접근이 아닌 지역별 접근을 통해 진정한 의미의 지역연구가 가능하게 하기 위해서는 국가별 데이터의 공간자료화 혹은 공간적 맥락이 반영된 데이터의 생산을 통해 연구 자료의 지역화가 선행되어야 할 것이다.

공간자료는 스케일 의존도가 낮아 지역-국가-지방 경계에 구속받지 않으므로 여러 스케일의 지역연구에 활용도가 높다. 그러므로 공간 자료화된 매크로 데이터는 연구자가 원하는 스케일의 데이터를 자유롭게 추출하여 해당 지역에 의미를 부여하거나 타 지역과의 비교대상으로 활용할 수 있다. 또한 공간화된 매크로 데이터를 통해 공간을 분석하면 전통적인 매크로 데이터로는 발견하기 어려운 관계와 현상을 파악하는 데 도움이 될 수 있다. 즉, 각종 수치데이터들을 공간정보화하는 것은 단순히 데이터의 시각화 이상의 정보를 생산해 낼 수 있으며 기 수집된 데이터의 효용을 극대화할 수 있을 것이다. 궁극적으로 국가 수준을 넘어선 지역으로의 메가아시아 탐색 연구에 유용할 것으로 기대한다.

이미 서로 다른 형태의 데이터를 통합시키고자 하는 시도는 2010년대 중반부터 꾸준히 진행되어 왔다. 새로운 시각을 열어 줄 것이라는 기대와 함께 빅데이터 접근방법이 각광받고 다양한 분야에서 여러 데이터에 대한 관심이 폭증하면서 서로 다른 형태의 데이터와, 목적에 부합하지 않는 데이터에 대한 고민 등이 생겨났다. 이를 해결하기 위해 데이터 통합(Data Blending)에 대한 개념이 대두되었고 2013년 말 2014년 초부터 지금까지 꾸준히 그 방법론이 개발되어 왔다. 데이터 통합(Data Blending)이란 여러 수치자료, 문서, 기록, 설문, 관찰연구 등과 같은 여러 유형의 1차 또는 2차 자료들로 데이터를 결합하는 과정을 의미한다(Singh et al., 2020). 데이터 통합에 대한 관심은 그림 7의 최근 10년간 'Data Blending'의 구글 검색량 변화 추이로도 알 수 있다.

정형데이터와 비정형데이터의 결합, 관측 데이터와 추정 데이터의 결합 등 여러 형태의 통합 시도되어 왔고 그중에는 공간정보와 매크로 데이터의 통합도 포함되어 있다.

대표적으로 유럽연합 통계국(Eurostat)은 2012년부터 공간정보와 각 국가별 통

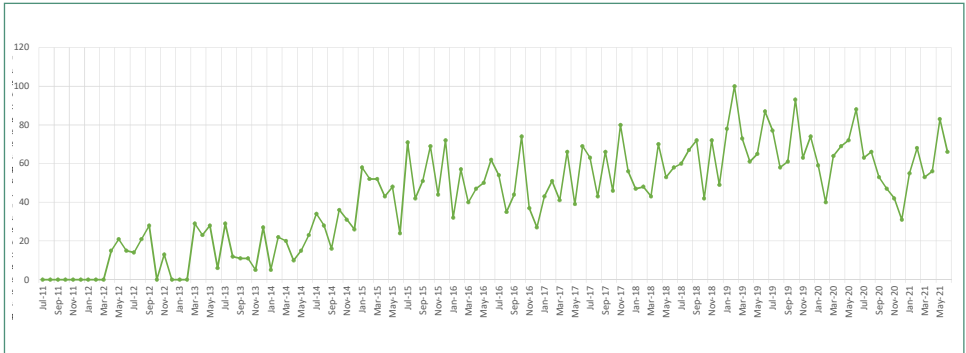
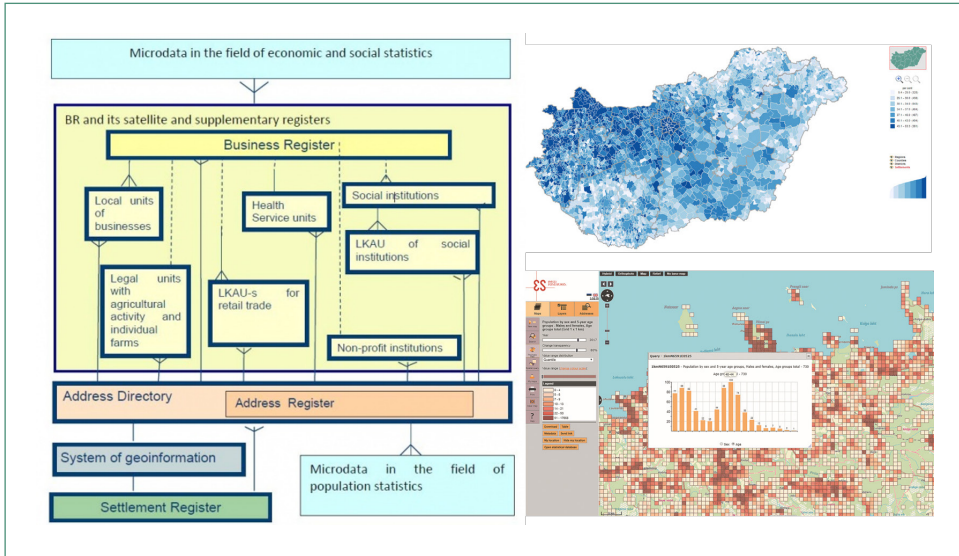


그림 7 최근 10년 동안 'Data Blending' 구글 검색량 변화

계수치들을 통합하려는 노력을 지속하고 있다. 유럽연합 통계국은 2012~2015 'Merging statistics and geospatial information' 이니셔티브를 통해 공간자료와 매크로 데이터의 결합을 통해 특정 스케일의 현상을 파악하기 위한 특정 스케일의 데이터 구축을 시도했고, 그 가능성을 높게 평가하여 2020 주요 비전 중 하나로 '국가 및 지역 정책 수립을 지원하기 위한 공간적 맥락이 반영된 통계 개발'을 명시한 바 있다(그림 8). 세부적인 EU 정책의 개발, 구현, 모니터링 및 평가에는 공간적 맥락이 포함된 세부 자료가 필요하기 때문이라 그 이유를 밝히고 있다.

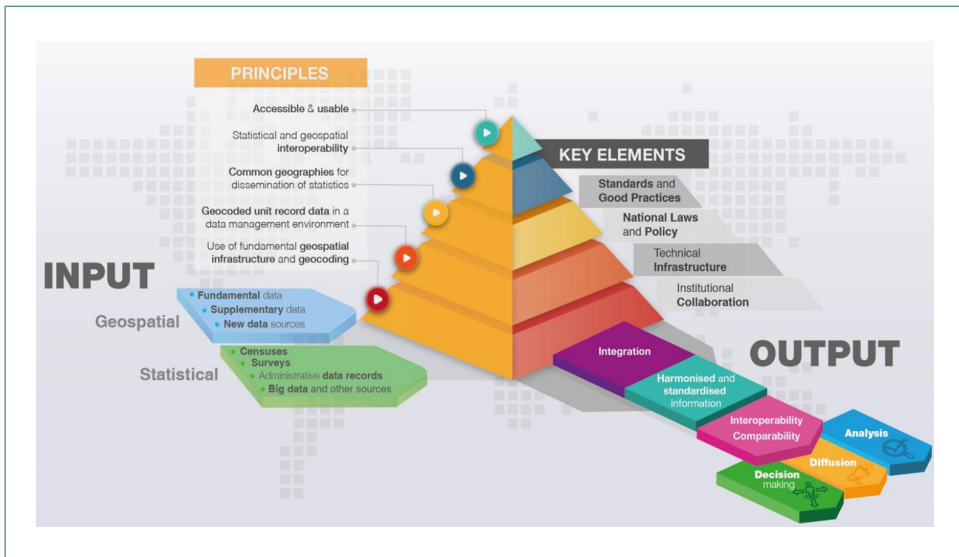
UN에서도 유럽연합 통계국과 비슷한 접근을 통해 공간자료와 매크로 데이터의 통합을 시도하고 있다. 이를 통해 국가별 통계 데이터를 지리 공간적으로 활성화함으로써 분석 및 의사 결정 프로세스를 위한 새롭고 더 나은 통합 데이터를 구축하고 국가 단위가 아닌 더 작은 영역에 대한 정보 증가 등을 기대한다(그림 9).

즉, 인구, 사회, 문화, 경제 등의 인문환경적 요소들을 나타내는 매크로 데이터를 공간자료화함으로써 여러 스케일의 지역정보를 만들어 낼 수 있고 더 나아가 이를 이미 공간자료로 존재하는 자연환경 데이터와 결합함으로써 사회, 경제 및 환경 문제에 대한 더 나은 이해가 가능할 것이다. 그리고 이러한 데이터는 국가 경계 기반이 아니므로 원하는 지역에 대한 정확한 정보를 추출해 낼 수 있어 의미 있는 지역단위 연구에 보다 효과적으로 활용할 수 있다. 즉, 매크로 데



출처: Eurostat, 2019

그림 8 'Merging Statistics and Geospatial Information' 프레임워크(왼쪽)와 결과(오른쪽)



출처: UN-GGIM, 2019

그림 9 'The Global Statistical Geospatial' 프로젝트 프레임워크

이터와 공간자료를 결합하는 작업은 이들을 분리해서 볼 때 가능한 것보다 훨씬 더 많은 가능성을 가질 수 있다.

또한, 최근 크게 주목받고 있는 빅데이터를 통한 아시아의 역사적·사회적 거시궤적을 보여 주는 자료들이 다양하게 발굴되고 있다. 이러한 현상은 KCI에 등재된 지역학 학술지 논문 중 양적 자료를 분석한 지역 연구에서도 서베이 데이터나 국가 데이터 외 연구자가 직접 만든 데이터의 사용이 증가하는 것으로 확인할 수 있다. 연구자가 직접 문헌자료 등을 코딩하여 분석하는 방법은 기존에도 많이 활용되어 왔으나 최근에는 그림이나 이미지 등을 연구자가 코딩하여 분석하기도 하고 빅데이터를 직접 분석하거나 이미지와 동영상 자료를 인공지능을 통해 코딩한 데이터를 사용하는 연구도 등장하고 있다. 이종혁(2021)은 중국 공산당 간부연구에서 기존의 중국공산당 간부의 기록을 모두 데이터화하여 머신러닝을 통해 학습시켜 당간부가 최고위 지위로 승진하는 것을 예측하는 모델을 개발하였다. 또한 허정원과 장주영(2020)은 휴대폰의 송출데이터를 통해 추계된 서울시 생활인구 빅데이터를 분석하여 코로나 기간 중 서울시 이주민 밀집 지역의 이동성 변화를 이주민과 원주민으로 나누어 분석하였다. 이러한 연구들은 공간과 맥락성을 반영한 충실한 양적 데이터 부족으로 인한 양적 연구의 한계를 극복하기 위한 노력의 예이다. 이처럼 아시아 내 소셜 미디어, 인터넷 뉴스 게시판과 커뮤니티 등에 데이터를 실시간으로 수집하여 가상 공간에 존재하는 아시아 지역들의 모습을 탐색할 수 있다.

정보통신기술과 항공우주기술이 발전함에 따라 사물인터넷 및 전자 상거래 데이터, 스마트폰 송출자료, 위성사진 등 여러 지역과 국가에 대한 데이터의 종류가 크게 증가하였다. 이에 따라 여러 형태의 데이터를 결합하여 해당 지역이 경험하는 특수성이나 타 지역과의 관계성, 그 안에서 발견되는 일반성을 통하여 지역을 다각적으로 이해하는 데 도움이 되는 새로운 방법을 고민하고 논의해 나가는 것은 매우 중요하다. 특히 국가 단위의 연구를 넘어선 지역의 역동성과 관계성을 탐색을 위하여 공간 기반의 데이터베이스 구축은 필수적이다. 또한 폭발적으로 증가한 새로운 데이터를 포함한 여러 소스에서 혼합된 데이터를 결합하여 새롭고 더 복잡한 방식으로 지역을 포괄적으로 이해할 수 있다. 특히, 소셜 미디어나 인터넷 게시판 등 가상공간에서 수집된 데이터는 특정 목적에 의해

과되지 않은 해당 지역의 특징과 인식을 그대로 보여 주기 때문에 활용 가치가 높고 상대적으로 적은 비용으로 다양한 지역을 대상으로 빠르게 수집 가능하여 다양한 주제의 지역연구에 활용할 수 있을 것이다.

투고일: 2021년 6월 15일 | 심사일: 2021년 7월 13일 | 게재확정일: 2021년 7월 30일

참고문헌

- 강근복. 2009. “정책분석 및 평가에서의 양적질적 연구의 혼합.” 『정책분석평가학회보』 19권 4호, 43-67.
- 구자용. 2015. “공간정보 빅 데이터의 지도화와 공간적 분포 특성에 관한 연구: 서울시 지역의 트윗 데이터를 사례로.” 『국토지리학회지』 49권 3호, 349-360.
- 김경수. 2006. “지역학의 정체성과 방향성.” 『세계지역연구논총』 24권 1호, 5-25.
- 김인숙. 2007. “한국 사회복지 질적 연구: 동향과 의미.” 『한국사회복지학』 59권 1호, 275-300.
- 김형중. 2019. “한국의 말레이시아 연구.” 안청시·전제성 편. 『한국의 동남아시아 연구』, 205-238. 서울대학교출판문화원.
- 김형중·전제성. 2019. “한국의 인도네시아 연구.” 안청시·전제성 편. 『한국의 동남아시아 연구』, 23-64. 서울대학교출판문화원.
- 박승배. 2009. “질적연구와 양적연구의 혼합 논리로서 프래그머티즘과 교육학 연구방법에 대한 듀이의 입장 고찰.” 『교육과정연구』 27권 2호, 63-81.
- 송태민. 2013. “보건복지 빅 데이터 효율적 활용방안.” 『한국컴퓨터정보학회지』 21권 1호, 45-53.
- 신재혁. 2019. “싱가포르를 왜 민주화되지 않는가?: 비교연구 방법을 활용한 지역연구.” 『동남아시아연구』 29권 2호, 161-191.
- 안승국·김유경·이은정. 2002. “비교·지역연구에 있어서 최선의 연구전략: 변수중심연구 전략인가 사례중심연구전략인가.” 『국제·지역연구』 11권 4호, 57-74.
- 안종욱·이미숙·신동빈. 2013. “공간빅데이터 개념 및 체계 구축방안 연구.” 『한국공간정보학회지』 21권 5호, 43-51.
- 유성민. 2015. “IoT와 빅데이터 기술 연계 플랫폼 조사.” 『한국정보기술학회지』 13권 2

- 호, 19-25.
- 유평준. 2011. “지역정보화 연구의 경향 분석.” 『한국지역정보화학회지』 14권 4호, 33-64.
- 이성우·지우석·정진규. 2004. “서울시 정보화수준의 자치구별 격차 및 개선방안에 관한 연구.” 『서울도시연구』 5권 1호, 1-26.
- 이승호. 2009. “주요국 행정의 지역·비교연구.” 『한국행정학보』 43권 4호, 389-393.
- 이중희. 2001. “지역연구방법의 범위와 특성.” 『한국사회학회 사회학대회 논문집』, 107-121.
- 이중혁. 2021. “Race for Second Fiddle: Midterm Appraisals on Central Leaders of China(머신러닝을 이용한 중국 공산당 간부의 승진 예측 모델).” 서울대학교 아시아연구소, 정치외교학부 10-10 프로젝트 공동주최 HK+ 데이터스토리텔링 클러스터 워크숍.
- 이창식. 2016. “지역학 연구와 구술자료 활용.” 『지방사와 지방문화』 19권 2호, 7-38.
- 이한우. 2019. “한국의 베트남 연구.” 안청시·전제성 편. 『한국의 동남아시아 연구』, 63-126. 서울대학교출판문화원.
- 이희연·최재현. 1998. “지리학에서의 지역연구 방법론의 학문적 동향과 발전방향 모색.” 『대한지리학회지』 33권 4호, 557-574.
- 이희영. 2005. “사회학 방법론으로서의 생애사 재구성: 행위이론의 관점에서 본 이론적 의의와 방법론적 원칙.” 『한국사회학』 39권 3호, 120-148.
- 정법모. 2019. “한국의 필리핀 연구.” 안청시·전제성 편. 『한국의 동남아시아 연구』, 177-204. 서울대학교출판문화원.
- 조남경. 2019. “질적, 양적 연구를 넘어? 사회복지 빅데이터 연구방법의 모색.” 『한국사회복지학』 71권 1호, 7-25.
- 조용기. 2001. “질적 연구의 성격.” 『교육인류학연구』 4권 1호, 157-168.
- 최명민. 2007. “질적·양적 연구방법론의 혼합에 의한 의료사회복지사의 소진탄력성 및 소진위험성 척도개발 연구.” 『한국사회복지학』 59권 4호, 245-272.
- 허정원·장주영. 2020. “코로나19 확산시기 서울시 외국인 밀집지역의 지역특성과 생활인구 변화.” 『공간과 사회』 73, 99-137.
- 홍경준. 2002. “한국 사회복지학계 연구방법론의 동향.” 『비판사회정책』 13권 2002호, 13-32.
- 홍서영·서태열. 2014. “지리교육 연구의 대안적 연구 방법으로서 QCA(Qualitative Content Analysis)의 적용.” 『한국지리환경교육학회지』 22권 3호, 103-120.
- Eurostat. 2019. “Merging Statistics and Geospatial Information“ Experiences and

- Observations from National Statistical Authorities.” 2012-2015 Projects, Eurostat.
- Singh, L., M. Traugott, L. Bode, C. Budak, P. E. Davis-Kean, R. Guha, and S. Soroka. 2020, *Data Blending*. Massive Data Institute, Georgetown University.
- Tobler, W. 1970. “A Computer Movie Simulating Urban Growth in the Detroit Region.” *Economic Geography* 46, 234-240.
- UN-GGIM. 2019. *The Global Statistical Geospatial Framework*. United Nations New York.

Abstract

Necessity and Possibilities of a Data-driven Regional Research Methodology for Exploring Mega Asia: Implications from KCI Indexed Regional Study Articles with Quantitative Research Methods

Jungwon Huh Seoul National University Asia Center
Seonyoung Park Seoul National University Asia Center
Hyojin Jang Seoul National University
Woojin Shim Seoul National University Asia Center

The purpose of this paper is to explore a new data-driven research method for a comparative area study. The paper examines the data and analysis methods utilized by previous area studies with a quantitative method and proposes innovative research methods to explore the discourse of Mega Asia and comparative area studies. The authors searched research articles that compared countries and regions from KCI indexed journals and collected bibliographic data such as titles, keywords, abstracts, areas of interest, and main topics. Since 2000, the publication of regional study journals has significantly increased, and most newly published journals focus on Asia. Mainly, newly published Asia study journals specialized on East Asia and Southeast Asia regions, reflecting the growing interest in South Korea due to geographical proximity. More studies investigated countries than regions, and the most popular regions of interest among study subjects are Asia and Europe. Most of the studies relied on macro data from international organizations or country-level statistics produced by the governments. Few studies utilized spatial data with no scale dependency. Spatial data allow them to extract the information regardless of interstate boundaries. The findings reaffirm the need for an alternative

blended database with geographical coordinates and a data-driven approach with a data science perspective to analyze regional dynamics.

Keywords | Mega-Asia, data-driven research, comparative regional study, spatial data, data blending