

연구논문

## 정부 데이터 역량이 데이터기반행정에 미치는 영향에 관한 실증연구: 인식조사자료를 중심으로\*

윤 건\*\*

- I. 서론
- II. 이론적 논의
- III. 연구설계
- IV. 분석결과
- V. 결론

### 〈 요약 〉

2020년 6월 데이터기반행정법 제정 이후 정부에서 데이터기반행정이 본격화되고 있다. 데이터기반행정은 과학적 데이터 분석 결과를 정부의 행정과 정책과정에 적극적으로 활용하여 정부의 과학성을 높이고 이를 통해 정부조직의 성과를 제고하려는 디지털 전환 시대의 정부 혁신 전략이라고 할 수 있다. 본 연구에서는 정부의 데이터 기반행정이 성공하기 위해 데이터 역량이 필수적이라고 보고 실제 정부 데이터 역량이 데이터기반행정에 통계적으로 유의미한 영향을 미치는지 실증적으로 검증하고자 하였다. 선행연구 검토를 통해 정부 데이터 역량을 개념화하여 분석모형을 설정하고 한국행정연구원 ‘공공데이터 정책활용 실태조사’ 데이터를 활용하여 분석하였다. 분석 결과, 조직 역량 요소에서는 데이터 기획과 데이터 문화, 인적 역량 요소에서는 개인 데이터 스킬, 기반 역량 요소에서는 데이터 제도 등의 변수가 통계적으로 유의미하게 나타났다. 이러한 결과를 바탕으로 일정한 정책적 시사점을 제시하였다.

[주제어: 정부역량, 데이터 역량, 정부 데이터 역량, 데이터기반행정]

\* 이 논문은 2021년도 한신대학교 교내 연구비 지원에 의해 수행된 연구임. 한국행정연구원에서 생산된 자료를 활용하였으며, 한국행정연구원 연구자료관리규칙에 의거 사용허가를 받았음. 이 논문은 2021년도 한국행정학회 하계학술대회 발표논문을 발전시킨 것임.

\*\* 한신대학교 공공인재학부 교수(kyoon2010@hs.ac.kr)

## I. 서론

2020년 6월 「데이터기반행정 활성화에 관한 법률」(이하 데이터기반행정법)이 제정되면서 데이터기반행정이 본격화되고 있다. 데이터기반행정은 디지털 전환(digital transformation)과 AI·데이터시대에 적극적으로 대응하고 증거기반의 과학적 행정을 구현하기 위한 정부의 핵심 전략으로 이해된다. 실제 이는 미국의 ‘증거기반정책법(Foundations for Evidence-Based Policymaking Act)’이나 영국의 ADR UK, 일본의 데이터중심행정정책의 사례와 같이(김법연, 2020; 윤건·김철우, 2021) 이 시대 정부들이 공통적으로 지향하는 방향성이 되고 있다. 실시간으로 수집되는 대용량의 데이터 분석은 행정과 정책의 시간적 적실성과 효과성을 높이고, 개인화된 맞춤형 데이터의 적용은 국민 개개인의 행정 생활을 더욱 편리하고 풍성하게 할 것이다.

데이터기반행정이 활성화되기 위해서는 정부조직에 다양한 요건이 갖추어져야 한다. 개별 데이터기반행정 프로젝트에 대한 기획, 이를 총괄하는 리더의 데이터 스킬, 다른 조직과의 데이터 협업, 데이터기반행정이 일상화된 조직문화 등의 조직 요소, 조직구성원 개개인의 데이터 스킬, 그러한 스킬을 갖춘 데이터 인력의 규모, 그러한 스킬을 갖추 수 있도록 지원하는 데이터 교육 등의 인적 요소, 데이터기반행정이 작동할 수 있도록 지원하는 법이나 거버넌스와 같은 데이터 제도, 데이터기반행정에 필요한 데이터 예산, 데이터 표준화나 시스템 통합과 같은 데이터 기술 등의 기반 요소 등이 필요하다.

다만 이러한 요소들이 실제 데이터기반행정에서 얼마나 중요하고 필요한 것인지는 명확하지 않다. 데이터기반행정을 활성화하고자 하는 정부의 입장에서는 이러한 요소들이 실제 데이터기반행정에 영향을 미치는지, 어떠한 요소가 보다 중요하고 지금 당장 필요한 것인지에 대한 정보가 필요하다. 본 연구는 이러한 관점에서 이러한 요소들을 정부 데이터 역량으로 개념화하고 실제 이것이 데이터기반행정에 영향을 미치는지, 또한 어떠한 요소가 보다 많은 영향을 미치는지 실증분석하는 것을 목적으로 한다.

이를 위해 우선 정부 데이터 역량과 데이터기반행정에 대한 이론적 논의를 살펴보고, 연구 분석틀과 모형을 제시한다. 다음으로 한국행정연구원의 ‘공공데이터 정책 활용 실태조사’ 데이터를 활용하여 실증분석을 시도한다. 이 데이터는 국책연구기관인 한국행정연구원이 한국갤럽과 함께 데이터기반행정에 대한 공공데이터 실무담당자

와 국책연구기관 연구원 330명을 대상으로 조사한 것으로 데이터기반행정과 관련된 실증연구를 수행하기 적절한 신뢰성 높은 데이터라고 할 것이다. 마지막으로 분석결과를 토대로 정책적 함의를 제시한다.

## II. 이론적 논의

### 1. 데이터기반행정

우리나라 맥락에서 데이터기반행정 개념은 데이터기반행정법에서 출발한다.<sup>1)</sup> 데이터기반행정법 제2조에서는 “공공기관이 생성하거나 다른 공공기관 및 법인·단체 등으로부터 취득하여 관리하고 있는 데이터를 수집·저장·가공·분석·표현하는 등의 방법으로 정책 수립 및 의사결정에 활용함으로써 객관적이고 과학적으로 수행하는 행정”으로 데이터기반행정을 정의하고 있다. 이러한 정의에는 데이터 활용 주체로서의 공공기관, 데이터 소스의 다양성(직접 생성뿐만 아니라 다른 주체가 생성한 데이터 포함), 데이터 프로세스, 데이터 활용 유형, 데이터기반행정의 수행 양식 등, 다양한 요소가 포함되어 있다. 이러한 각각의 요소에 대해 몇 가지 논의해 보면 다음과 같다.

첫째, 데이터 활용 주체는 공공기관으로 기본적으로 조직의 관점에 입각해 있으며, 국가, 지방자치단체, 공운법상 공공기관, 지방공기업법상 지방공사/지방공단, 각급 학교, 특수법인 등, 공공영역의 대부분의 조직이 포함된다. 다만 이러한 공공기관 용어가 법적, 실무적 용어로서의 특성이 강하기 때문에 본 논문에서는 통칭하여 ‘정부’ 혹은 ‘정부조직’이라는 용어를 사용하고자 한다. 둘째, 이러한 데이터기반행정의 주체들이 활용하는 데이터 소스는 다양해야 한다. 정책융합의 환경에서 적실성 있는 정책을 마련하기 위해 개별 기관은 서로 다른 정책 영역의 기관이 보유한 데이터를 필요로 한다. 특정 정책목적으로 만들어진 데이터가 다른 정책목적에 활용되는 것이다. 또한 행정이 모든 데이터를 직접 생성하는 것은 가능하지도 바람직하지도 않기 때문에 민간이 생성한 데이터를 활용하는 것도 필요하다.

1) 데이터기반행정과 유사한 개념으로 ‘데이터기반정책’이 활용되는데, 두 개념은 ‘행정’과 ‘정책’이라는 용어 차이에서 기인하는 것으로 데이터가 활용되는 실제 상황에서 구분이 어렵기 때문에 본 연구에서는 두 개념을 동일한 의미로 사용한다.

셋째, 데이터가 행정에 활용되기 위해서는 기본적으로 수집 및 저장-가공 및 분석-표현 등, 일련의 단계가 필요하다. 다만 이것보다 선행되어야 하는 것은 기획 과정이다. 수집과 분석, 표현의 전과정에 대한 기획을 통해 프로세스를 관리하고 통제해야 한다. 또한 프로세스가 너무 세분화되어 있기 때문에 수집은 저장을 포함하는 것으로, 가공은 분석에 포함되는 것으로 하여 단계를 줄이는 것이 바람직하다. 넷째, 데이터는 정책의 전단계, 즉 정책형성-정책집행-정책평가의 각 단계에서 활용될 수 있다. 정책수립과 의사결정에 한정하는 것은 바람직하지 않다. 또한 데이터 활용에 있어 정책과정의 단계별 차이가 크지 않으므로 이를 적극적으로 개념 정의에 이용할 필요는 없을 것으로 보인다.

다섯째, 데이터기반행정의 수행 양식의 기본적 특성은 과학성으로서, 데이터기반행정의 수행에 과학적 방법론이 적용되어야 하는 것이다. 다만 사회과학의 방법론은 자연과학의 방법론과 정확하게 일치하기 어렵다. 사회과학에서는 엄밀하게 설계된 주관적 방법론을 포함한다. 예를 들어 인터뷰 실시와 인터뷰 내용에 대한 객관화된 분석 역시 과학적 방법론으로 인정되어야 하는 것이다. 여섯째, 과학적으로 수행하는 행정을 통해 이루고자 하는 바가 개념 정의에 포함될 필요가 있다. 데이터기반행정법 제1조에 따르면 데이터기반행정은 데이터기반의 과학적 행정을 통해 공공기관의 책임성, 대응성, 신뢰성 제고를 직접적 목적으로 하는데 여기에는 정책의 효과성 및 적실성 제고가 포함될 필요가 있다.

이상의 논의를 바탕으로 본 연구에서는 데이터기반행정을 '정부조직이 책임성, 대응성, 신뢰성, 효과성, 적실성 등의 가치를 제고하기 위해 다양한 소스의 데이터를 기획/수집/분석/표현 등의 과정을 통해 과학적인 방법으로 활용하여 행정과 정책을 수행하는 것'으로 정의하고자 한다.

데이터기반행정법 등장을 전후로 하여 다양한 연구들이 수행되고 있다. 이하에서는 데이터기반행정 선행연구 분석을 실시한다. 이를 통해 데이터기반행정을 바라보는 관점과 시각, 데이터기반행정의 핵심요소와 체계화 방식 등을 확인하고, 본 연구의 의의를 제시한다. 데이터기반행정 선행연구들은 크게 데이터기반행정에 대한 법제도적 연구와 조사데이터에 기반한 실증 연구로 대별될 수 있다.

우선 데이터기반행정 연구에서는 데이터기반행정의 제도적 기반으로서 법적 연구가 많이 나타나는데, 이는 법시행 초기라는 시점의 영향으로 이해된다. 김법연(2020)은 데이터기반행정과 관련된 법제 동향을 연구하였다. 그는 데이터기반행정 개념, 등

장배경, 미국/일본/영국 등의 해외 입법 및 정책 동향, 데이터기반행정법의 주요 내용 등을 제시한다. 데이터기반행정을 정부혁신의 관점에서 지지하고 법시행의 과제로서 일선 공무원의 혁신적 사고, 민간과의 협업, 프라이버시와 국가감시 문제 등을 제시한다. 정성호·전주열(2020)은 데이터기반행정과 데이터기반행정법의 의의와 전제, 한계 등을 증거기반정책의 관점, 재정과 재정법의 관점에서 논의한다. 데이터기반행정법의 한계로 법률적용 문제, 재정감독 책임 주체 문제, 거버넌스 문제 등을, 데이터기반행정의 전제조건으로 데이터와 데이터를 활용하는 행정의 신뢰성을 제시하고 재정 관점에서 적용 대상의 명확화, 재정정보의 투명성, 재정효과 데이터 산출, 재정감독 책임 주체 명확화, 기관 간 시스템 연계 등을 제안한다. 윤건·김철우(2021) 역시 증거기반정책의 관점에서 데이터기반행정법을 바라보면서 이와 관련된 미국과 영국의 법·제도적 사례를 비교분석하여 우리나라 데이터기반행정법의 발전 방향을 제시한다. 미국의 증거기반정책법 사례에서는 전략, 증거, 평가, 통계 등의 연계성, 영국의 2017년 디지털경제법에 기반한 ADR UK 사례에서는 행정데이터 분석의 정책활용, 정부/대학/통계청/연구회(ESRC) 등 핵심 주체의 역할과 관계 등이 제시된다. 선지원(2021)은 데이터기반행정의 법적 문제를 탐구하였다. 데이터기반행정은 적극행정, 행정의 효율성과 정확성, 행정서비스의 품질 제고, 행정의 투명성과 통제 가능성 제고를 돕는다. 데이터기반행정은 증거기반행정에서 출발해 인공지능을 통한 자동화된 적극행정으로 확장되었다. 데이터기반행정에서는 편향성, 차별, 프라이버시, 자기결정권 등의 역기능 문제가 발생할 수 있으므로 알고리즘 통제나 적절한 거버넌스 등, 정교한 제도 설계와 균형 잡힌 시각이 필요하다.

다음으로 데이터기반행정과 관련한 조사데이터에 기반한 실증연구들을 살펴보면 다음과 같다. 서재호(2020)는 지방자치단체 공무원에 대한 인식조사를 중심으로 데이터기반행정 영향요인을 실증적으로 분석하였다. 이 연구에서는 데이터기반행정 영향요인을 문화적 요인, 기술적 요인, 인적 요인, 리더십 요인, 제도적 요인 등, 다섯 가지 요인으로 구분하였는데, 분석결과 문화적 요인, 기술적 요인, 제도 요인이 통계적으로 유의미한 것으로 나타나, 데이터 관리 조직원의 태도형성, 조직 전체 차원의 데이터 플랫폼 구축과 IT기술 도입, 데이터 협업장치 등을 제안하였다. 윤건·김윤희(2020)는 데이터기반행정의 맥락에서 공공데이터의 융합의 영향요인과 매개효과를 실증분석하였다. 이 연구에서는 공공데이터 융합 영향요인을 제도, 역량, 기술 등의 세 가지 차원으로 구분하여 제시하였는데, 분석결과 제도요인(개인정보보호), 역량요인(데이터협

업), 기술요인(데이터표준화, 데이터품질)이 공공데이터 융합에, 그리고 역량요인이 데이터기반행정에 직접적, 간접적(공공데이터 융합을 통해)으로 통계적으로 유의미한 영향을 미치는 것으로 나타났다. 윤건·김철우(2020)는 데이터기반행정 구현을 위해 필요한 요소를 데이터기반행정 주체별(국가공무원, 지방공무원, 국책연구기관 연구원)로 분석하였다. 이 연구에서는 데이터기반행정의 핵심 요소를 데이터 프로세스, 데이터기반행정 기반요소, 데이터 정책 활용 등으로 구분하고, 주체간 차이 분석, 데이터 프로세스 경험과 데이터기반행정 기반 인식의 데이터 정책 활용에의 영향 분석 등을 실시하였다. 분석결과 데이터 경험의 상대적 차이(낮은 지방공무원 데이터 경험), 낮은 데이터 인력/협력 충분성 인식, 데이터기반행정 기반요소와 데이터 정책 활용 간 통계적으로 유의미한 관계 등이 나타났다.

본 연구는 데이터기반행정에 관한 추가적 실증연구로서의 의의를 갖는다. 이러한 데이터기반행정의 실증연구는 데이터기반행정의 이론적 발전에 중요하다. 실증연구 과정에서 이론적 모델을 정립하는 과정이 수반되기 때문이다. 또한 데이터기반행정에 대한 실증연구는 데이터기반행정 시행 초기인 현 시점에서 정책적으로 큰 도움을 줄 수 있다. 데이터기반행정을 활성화하기 위한 제도적 기초로서 데이터기반행정법이 시행되었으나 우리나라의 데이터기반행정의 현실은 녹록치 않다. 선행연구들의 지적대로 데이터기반행정법의 지속적인 개선을 통한 제도적 보완도 중요하지만, 실제 데이터기반행정 활성화 정책을 실행하기 위해서는 어디서부터 어떻게 접근해야 하는지도 중요하다. 데이터기반행정과 관련된 실증연구들은 이러한 정책의 우선순위 설정에 도움을 줄 것이다. 본 연구는 기존 실증연구들의 한계를 극복하려는 의의도 갖는다. 즉 기존 데이터기반행정 실증연구에서는 개별 세부 요소들의 영향을 직접 파악하는 것이 미흡하였다. 서재호(2020)의 연구에서는 문화, 기술, 인적, 제도, 기술 등의 요소로 구분하였으나 그 하위 요소별 영향을 파악하는데 한계가 있었고, 윤건·김윤희(2020)의 연구에서도 역량 요인과 데이터기반행정의 실증적 관계를 보여주었으나, 역량 요인의 세부 하위 요소별 관계는 분석하지 못했다. 윤건·김철우(2020)의 연구에서도 기반 요소와 데이터기반행정의 관계의 통계적 유의미성을 보여주었으나, 역시 기반 요소의 세부 하부 요소별로는 보여주지 못했다. 본 논문에서는 데이터기반행정의 세부 요소별로 데이터기반행정에 어떻게 영향을 미치는지 실증적으로 보여줌으로써 데이터기반행정의 이론과 정책적 발전에 기여하고자 한다.

## 2. 정부 데이터 역량과 데이터기반행정

데이터기반행정의 선행연구들에서 제시되는 다양한 영향요인들은 데이터와 관련하여 정부조직이 갖추어야 할 역량으로 개념화될 수 있다. 본 연구에서는 데이터와 관련된 정부조직의 특성으로 이해할 수 있는 요소들을 정부 데이터 역량으로 개념화하고자 한다. 정부 데이터 역량 개념화를 위해 역량, 정부역량, 정부조직역량, 데이터 역량 등의 개념을 살펴보면 다음과 같다.

우리나라의 맥락에서 역량 연구는 개인, 조직, 정부, 국가 차원에서 다양하게 이루어지고 있으나 실제 그것이 지칭하는 영어적 표현이 일치하는 것은 아니다. 예를 들어 역량 연구에서 가장 대표적으로 활용되는 ‘competence’는 탁월한 성과를 가져오는 개인의 능력을 의미한다. 이러한 정의는 McClelland(1973)가 제시한 역량 개념에 기반한다. 그는 개인에 대한 평가를 ‘지능(intelligence)’이 아닌 실제 할 수 있는 ‘스킬(skill)’ 중심으로 할 것을 제안하였다. 이러한 역량 개념의 특징은 개인의 관점에서 제시된다는 것, 탁월한 성과와 연결된다는 것이다. 이후 조직연구에서는 조직 차원으로 확장된다. 대표적인 것이 Prahalad & Hamel(1990)의 핵심역량(core competence) 개념이다. 한 회사가 가지는 핵심역량은 다양한 시장 접근성, 최종 생산물로부터 얻는 고객의 인지된 혜택, 경쟁자의 모방 어려움 등의 특성을 갖는다. 이러한 정의는 회사를 하나의 주체로 이해하고 이것의 탁월한 성과를 가져오는 특성으로서 역량을 이해하고 있는 것이다.

정부역량 논의에서는 ‘capacity’가 활용된다. 미국 시라큐스대학교 캠벨공공행정연구소의 정부성과프로젝트(Government Performance Project, GPP)에서는 관리역량(management capacity) 개념을 활용한다(Donahue et al., 2000). 관리역량은 재정관리, 인적자원관리, 자산관리, 정보기술관리 등의 세부 관리적 요소를 정부성으로 연결시켜 최종적으로 정책결과를 만들어내는 역할을 한다. 윤건·심우현·박정원·김윤희(2018)의 연구에서는 정부역량을 “행정목적 달성에 필요한 역량”이라고 정의하고, 행정목적에 따라 요구되는 정부역량이 달라지는 것으로 보았다. 이 경우의 역량 개념은 정부조직 차원에서 정의되지만 역시 성과와 직결된다는 점에서 ‘competence’와 유사하다. 지방정부의 역량을 개념화한 이승중·윤두섭(2006)의 연구에서는 초기 역량 개념이 개인의 내적 특성이라는 한정된 의미로 활용되지만 이후 조직 역량을 포함하는 의미로 확장된다고 하면서, 지방정부의 역량을 “지방정부의 목적 달성을 위해 필요한

지방정부조직 및 공직자의 물적, 인적역량”으로 정의하여, 정부조직 차원과 정부조직의 개별 구성원으로서의 공무원 차원을 모두 포괄하여 정의한 바 있다. 결국 이러한 관점에 입각하면 영어 표현의 차이는 더 이상 큰 의미가 없음을 알 수 있다. 다만 본 연구에서는 정부역량 연구의 전통을 따라 ‘capacity’ 용어를 사용한다.

다음으로 정부 데이터 역량 개념화를 위해 데이터 역량에 대한 선행연구들을 검토해 본다. 선행연구들은 주로 데이터 역량을 직접적으로 정의하기보다는 세부 역량 중심으로 나열하는 경향을 나타낸다. 데이터 역량 관련 선행연구들은 크게 개인 관점과 조직 관점으로 구분되고, 조직 관점의 경우 회사조직과 정부조직으로 구분될 수 있다. 우선 개인의 관점에서 데이터 역량을 접근하는 연구는 다음과 같다. 서웅·안성진(2019)은 컴퓨팅 사고력의 요소로 데이터 수집, 데이터 분석, 데이터 표현의 세부 역량으로 수집방법 선택, 키워드 추출, 데이터 타입 이해, 탐색 기술 활용, 데이터 분류, 관계와 패턴 발견, 데이터 의미 탐색, 다양한 표현 방법 이해, 다양한 표현 방법 사용 등을 제시한다. 이태원·성행남·김은정(2020)은 NCS 기반 빅데이터분석 직무의 능력단위로 “탐색적 데이터 분석, 텍스트 데이터 분석, 분석용 데이터 구축, 머신러닝 기반 데이터 분석, 분석용 데이터 구축, 머신러닝 기반 데이터 분석, 통계기반 데이터 분석, 빅데이터 분석 결과 시각화” 등을 제시한다. 김효중·김희웅(2021)은 ‘LinkedIn’의 ‘AI’, ‘데이터사이언스’ 영문 키워드를 포함하는 채용공고 데이터에 대한 LDA 토픽모델링 분석을 통해 통계, 프로젝트경험, 머신러닝, 비즈니스문제해결, 관리, 프로그래밍 등, 6개의 토픽을 제시하였다.

다음으로 조직의 관점에서 데이터 역량을 접근하는 연구들을 살펴보면 다음과 같다. 먼저 회사조직에 대한 연구로서 Davenport et al.(2001)은 한 회사가 분석 역량(analytic capabilities)을 갖추기 위해 필요한 다섯 가지 역량으로 기술스킬, 통계적 모델링과 분석스킬, 데이터에 대한 지식, 비즈니스에 대한 지식, 커뮤니케이션 스킬 등을 제시한다. 김승현·박재홍·김이환·박주석(2015)의 연구에서는 기업의 데이터 분석 역량을 데이터 관리, 데이터 이해, 데이터 기반의 활동의 세 가지 요소로 구분하고 있는데, 이는 기본적으로 Kiron et al.(2011)의 분석 역량(analytics competencies)에 기반한다. Kiron et al.(2011)은 조직이 정밀한 데이터 분석 결과를 얻기 위해서는 정보 관리, 분석 스킬과 도구, 데이터 지향 문화의 세 가지 역량이 필요하다고 하였다. 그들은 세부 요소로 정보 관리의 경우 확고한 정보 기반, 표준화된 데이터 관리 관행, 접근 가능하고 활용 가능한 통찰 등을, 분석 스킬과 도구의 경우 핵심 요소로 개발된 스킬, 강

력한 도구와 솔루션 장착, 실행 지향 통찰 개발 등을, 데이터 지향 문화의 경우 사실 기반 리더십, 전략 자산으로 활용되는 분석, 통찰에 의해 지도되는 전략과 운영 등을 제시한다. 윤수영·오재인(2018)은 빅데이터 활용 및 지속적 사용의도에 영향을 미치는 요인으로 IT 자원, 인적 자원, 조직 자원의 세 가지 차원을 제시한다. IT 자원의 경우 보안성, 데이터관리, IT인프라 등을, 인적 자원의 경우 IT 스킬, 과업기술 적합성, 지식 공유 등을, 조직 자원의 경우 교육 및 훈련, 데이터 의사결정 문화 등을 제시한다.

정부 데이터 역량 연구로서 Chatfield et al.(2015)는 정부에서 빅데이터를 활용한 조직 전환의 역량적 측면의 도전 과제를 사회정치적 역량, 전략적 변화 역량, 분석 역량, 기술 역량의 네 가지 차원으로 제시한다. 사회정치적 역량의 경우 프라이버시와 보안을, 전략적 변화 역량의 경우 증거기반의사결정, 전략적 변화 리더십, 데이터 전략, 데이터 공유, 데이터 거버넌스 등을, 분석 역량의 경우 지식, 스킬, 비용(예산), 데이터 사이언티스트 등을, 기술 역량의 경우 데이터 질, 데이터 추출, 데이터 신뢰성, 데이터 확장성 등을 제시한다. 윤건·김윤희(2020)는 공공데이터 융합 영향요인으로 제도 차원, 역량 차원, 기술 차원을 제시하고, 제도 차원은 근거규정, 개인정보보호, 기관 이기주의 등을, 역량 차원은 데이터 리더십, 데이터 전문인력, 데이터 예산, 데이터 협업 등을, 기술 차원은 데이터 표준화, 데이터 검색, 데이터 품질, 데이터 충분성 등을 제시하였다. 윤건·김철우(2020)는 데이터기반행정 기반요소로 인적 요소와 물적 요소를 제시하고 있으며, 인적 기반으로는 인력, 교육, 문화, 협력, 리더역량, 구성원역량 등을, 물적 기반으로는 법령, 기획, 예산, 기술 등을 제시하고 있다.

본 연구에서는 정부역량 개념과 데이터 역량 개념을 종합하여 정부 데이터 역량을 '정부가 데이터기반행정의 목적을 달성하기 위해 필요한 데이터 관련 조직적, 인적, 기반적 차원의 역량'으로 정의한다. 이러한 정의의 특징은 다음과 같다. 첫째, 이승중·윤두섭(2006)의 통합적 관점을 따라 개인 차원의 역량과 조직 차원의 역량을 포괄적으로 다룬다. 둘째, 정부 데이터 역량의 목적(혹은 성과)을 데이터기반행정으로 명확하게 제시하였다. 셋째, 정부 데이터 역량을 조직 역량, 인적 역량, 기반 역량으로 구분하였다.

이와 관련하여 정부 데이터 역량의 유형 구분을 좀 더 세밀하게 살펴보면 다음과 같다. 조직 역량에는 데이터 기획, 리더 데이터 스킬, 데이터 협업, 데이터 문화 등이, 인적 역량에는 개인 데이터 스킬, 데이터 인력규모, 데이터 교육 등이, 기반 역량은 데이터 제도, 데이터 예산, 데이터 기술 등이 포함된다. 조직 역량과 인적 역량의 구분은

부서 수준 이상의 조직 차원에서 논의되는 경우 조직 역량으로 개별 구성원의 전문성 차원에서 논의되는 경우 인적 역량으로 하였다. 기반 역량은 조직 역량과 인적 역량 이외에도 선행연구들에서 제시된 제도, 기술, 예산 등, 조직 역량과 인적 역량의 기반이 되는 역량을 포함하는 것이다. 이러한 유형 분류는 이승중·윤두섭(2006)과 윤건·김철우(2021)의 인적 요소와 물적 요소의 구분을 보다 세분화하여 체계화한 것이다. 즉 인적 기반을 조직 역량과 인적 역량으로 구분하고, 물적 기반의 기획 역량은 조직 역량으로, 이외의 물적 기반 요소들을 기반 역량으로 미세조정하였다.

이러한 정부 데이터 역량은 데이터기반행정에 직접적 영향을 미칠 것으로 기대된다. 이는 앞서 제시된 실증연구 결과에서 어느 정도 예측할 수 있다. 김승현 외(2015)의 연구에서는 구조방정식모형 분석을 통해 데이터 분석 역량의 각 요소들이 데이터 분석의 유용성 혹은 사용자 만족을 통해 의사결정(업무처리 효율성 향상 및 업무 개선)에 통계적으로 유의미한 영향을 미치는 것으로 나타났다. 윤수영·오재인(2018)의 연구에서는 보안성, 과업기술 적합성, 데이터 의사결정 문화 등이 데이터 활용성과 지식역량과 인지된 기대성과를 매개로 빅데이터 사용의도에 통계적으로 유의미한 영향을 미치는 것으로 나타난다. 서재호(2020)의 연구에서는 문화적 요인, 기술적 요인, 제도 요인이 데이터기반행정에 통계적으로 유의미한 영향을 미치는 것으로 나타난다. 윤건·김윤희(2020)의 연구에서는 역량 요인이 데이터기반행정에 직접적인 영향을 미치는 것으로 나타났다. 윤건·김철우(2020)의 연구에서는 데이터기반행정 기반요소가 데이터의 정책 활용도에 통계적으로 유의미한 영향을 미치는 것으로 나타났다.

### III. 연구설계

#### 1. 연구 분석틀과 모형

본 연구는 정부 데이터 역량이 데이터기반행정에 미치는 영향을 실증적으로 분석하는 것을 목적으로 한다. 이를 위해 정부 데이터 역량을 조직 역량, 인적 역량, 기반 역량의 세 가지 차원의 하위 역량으로 구분하고 각각을 다시 세부 역량으로 구분하였다. 조직 역량의 경우 데이터 기획, 리더 데이터 스킬, 데이터 협업, 데이터 문화 등으로, 인적 역량의 경우 개인 데이터 스킬, 데이터 인력규모, 데이터 교육 등으로, 기반

역량의 경우 데이터 제도, 데이터 예산, 데이터 기술 등으로 구분하였다.

데이터 기획은 데이터기반행정을 통해 제공될 행정서비스나 정책문제의 해결을 위해 필요한 데이터에 대한 수집·분석·활용 계획을 수립하는 과정을 의미한다. 데이터 기획에서는 필요한 서비스나 해결하고자 하는 문제와 수단을 명확히 하고 분석 모델을 설정하며 분석에 필요한 데이터, 해당 데이터의 존재 유무, 데이터가 존재하지 않는 경우 대안적 데이터나, 혹은 신규 데이터 생성 방안 등을 제시해야 한다. 리더 데이터 스킬은 부서장급의 리더가 가지는 데이터 관련 기술 혹은 전문성 정도를 의미한다. 정부조직에서 리더는 조직의 업무와 조직 내 구성원에 대한 강력한 권한을 바탕으로 해당 부서의 성과를 만들어내는 책임을 진다. 리더의 데이터 스킬은 부서의 데이터 기획을 원활하게 하고 데이터 중심의 문화를 만들어가는 핵심 요소가 된다. 데이터 협업은 부서 간 데이터 협업에서 기관 간 데이터 협업까지 포괄한다. 데이터 협업은 필요한 데이터가 부서 내 혹은 기관 내에 존재하지 않는 경우, 개인이나 부서의 데이터 역량이 낮은 경우 중요하게 부각된다. 전자의 경우 분석할 데이터가 없는 경우이고, 후자의 경우 분석할 데이터가 있어도 데이터 분석 역량이 낮은 경우이다. 데이터 문화는 부서 혹은 기관 차원에서 데이터 중심의 업무 양식이 일상화된 상태를 의미한다. 데이터 문화는 형성되기까지 상당한 어려움이 있지만 일단 형성된 이후에는 데이터기반행정의 지속가능성을 높일 수 있다.

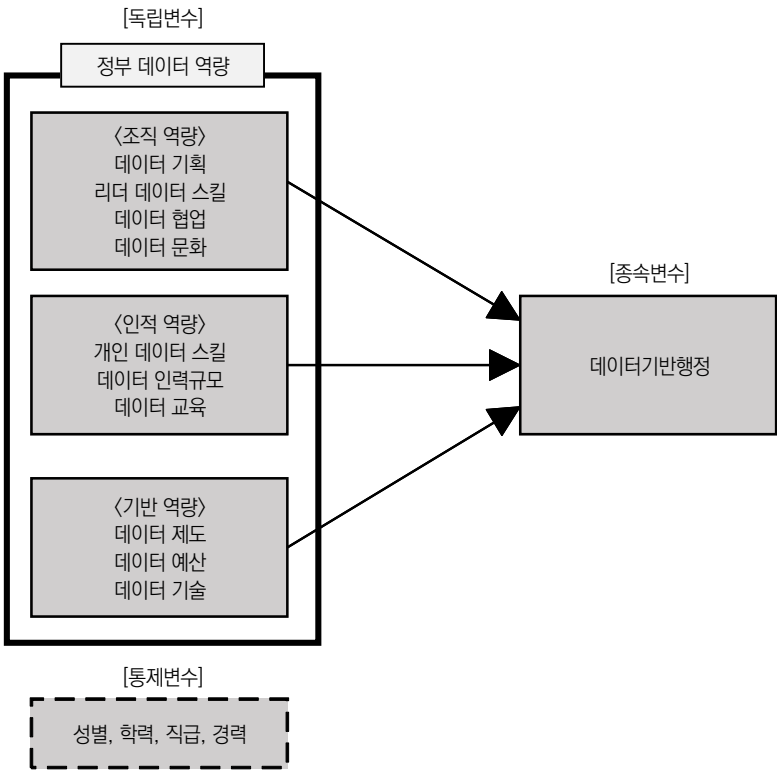
개인 데이터 스킬은 조직 내 개별 구성원의 데이터 관련 기술과 전문성을 갖춘 정도를 의미한다. 개인 관점의 데이터 역량 관련 선행연구 분석에서 제시된 다양한 분석 스킬을 개별 조직 구성원들이 갖추는 경우 기관 및 부서의 데이터 역량이 급속하게 증대될 것이다. 데이터 인력규모는 데이터 전문성을 갖춘 인력의 규모를 의미한다. 데이터 전문성을 갖춘 인력이 많아지는 경우 그렇지 않은 경우에 비해 데이터기반행정이 활발해질 것으로 기대할 수 있을 것이다. 데이터 교육은 데이터 전문성을 갖추도록 하기 위한 것이다. 데이터 교육의 체계와 프로그램이 적절하게 갖추어지면 개별 구성원들의 데이터 역량이 상승할 것으로 기대할 수 있다.

데이터 제도는 데이터 관련 법령 정비를 의미한다. 데이터 관련 법령에는 추진체계와 거버넌스, 기타 관련 제도적 규율이 포함된다. 특히 개인정보와 같은 민감데이터의 경우 제도적 규율이 적절하게 이루어질 필요가 있다. 데이터 예산은 데이터 과정에 필요한 금전적 요소라고 할 수 있다. 데이터 예산은 데이터기반행정에 필요한 인력과 시스템 확보를 용이하게 함으로써 데이터기반행정을 활성화시킬 것이다. 데이터 기술

은 데이터 장비나 설비와 같은 하드웨어적 요소뿐만 아니라 데이터 표준화와 같은 소프트웨어적 요소를 포괄한다.

본 연구에서는 이와 같은 각각의 세부 역량이 독립변수로서 종속변수인 데이터기반행정에 긍정적 영향을 미칠 것으로 보았다. 본 연구의 분석틀과 연구모형을 제시하면 다음 그림과 같다.

〈그림 1〉 연구 분석틀



〈그림 2〉 연구모형

$$DBPA = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \epsilon$$

〈종속변수〉 DBPA: 데이터기반행정

〈독립변수〉 [조직 역량]  $X_1$ : 기획,  $X_2$ : 리더 스킬,  $X_3$ : 협업,  $X_4$ : 문화

[인적 역량]  $X_5$ : 개인 스킬,  $X_6$ : 인력규모,  $X_7$ : 교육

[기반 역량]  $X_8$ : 제도,  $X_9$ : 예산,  $X_{10}$ : 기술

이러한 연구 분석틀과 모형, 앞선 선행연구 논의를 바탕으로 본 연구의 가설을 제시하면 다음과 같다.

- 〈가설1-1〉 데이터 기획이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설1-2〉 리더 데이터 스킬이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설1-3〉 데이터 협업이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설1-4〉 데이터 문화 역량이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설2-1〉 개인 데이터 스킬이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설2-2〉 데이터 인력규모가 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설2-3〉 데이터 교육이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설3-1〉 데이터 제도가 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설3-2〉 데이터 예산이 데이터기반행정에 정(+)의 영향을 미칠 것이다.
- 〈가설3-3〉 데이터 기술이 데이터기반행정에 정(+)의 영향을 미칠 것이다.

## 2. 분석 자료와 변수의 측정

본 연구에서는 정부 데이터 역량이 데이터기반행정에 미치는 영향을 실증적으로 분석함에 있어 한국행정연구원의 ‘공공데이터 정책 활용 실태조사’ 데이터를 활용하였다. 이 조사의 대상은 데이터기반행정의 실제적인 담당자로서 관련성이 가장 높은 공공데이터 담당 공무원과 국책연구기관 연구원이다. 표본크기는 330명으로 국가공무원 103명, 지방공무원이 124명, 연구원이 103명 표집되었다. 이 조사는 데이터기반행정 실태 및 인식과 공공데이터 융합 실태 및 인식과 관련된 다양한 설문문항을 포함하고 있다. 조사방법은 웹서베이 방식으로 이루어졌고, 2019년 7월 1일부터 15일까지 15일 동안 수행되었다. 조사 실시는 한국갤럽조사연구소에서 수행하였다.

〈표 1〉 분석자료 개요

|      |                                           |
|------|-------------------------------------------|
| 조사명  | 공공데이터 정책 활용 실태조사                          |
| 모집단  | 공공데이터 담당 공무원 및 국책연구기관 연구원                 |
| 표본크기 | 330명 (국가공무원 103명, 지방공무원 124명, 연구원 103명 등) |
| 조사내용 | 데이터기반행정 실태 및 인식, 공공데이터 융합 실태 및 인식         |
| 조사방법 | 웹서베이 방식                                   |
| 조사기간 | 2019년 7월 1일~15일(15일)                      |
| 조사기관 | 주관 기관: 한국행정연구원, 실시 기관: 한국갤럽조사연구소          |

출처: 윤건(2019: 159-161) 발췌 및 수정

본 연구에서 활용한 측정문항은 다음 표와 같다. 종속변수인 데이터기반행정은 '정부 공무원의 업무 수행시 데이터 활용에 대한 중요성 인식'으로 측정하였고, 각각의 독립변수들은 '데이터기반행정 활성화 관점에서 각 항목들의 중요성 인식'으로 측정하였다. 각 변수는 단일 문항으로 측정하였으며, 종속변수와 독립변수 모두 중요성 인식을 통해 측정함으로써 설문항목의 일관성을 확보하였다.

〈표 2〉 변수의 측정

| 구분       |          | 설문문항                                                      | 척도                                                                    |
|----------|----------|-----------------------------------------------------------|-----------------------------------------------------------------------|
| 종속<br>변수 | 데이터기반행정  | 귀하께서는 정부 공무원이 담당 업무를 수행하면서 데이터를 활용하는 것이 얼마나 중요하다고 생각하십니까? |                                                                       |
| 독립<br>변수 | -        |                                                           | 1: 전혀 중요하지 않다<br>2: 별로 중요하지 않다<br>3: 보통이다<br>4: 약간 중요하다<br>5: 매우 중요하다 |
|          | 조직<br>역량 | 데이터 기획                                                    |                                                                       |
|          |          | 리더 데이터 스킬                                                 |                                                                       |
|          |          | 데이터 협업                                                    |                                                                       |
|          |          | 데이터 문화                                                    |                                                                       |
|          | 인적<br>역량 | 개인 데이터 스킬                                                 |                                                                       |
|          |          | 데이터 인력규모                                                  |                                                                       |
|          |          | 데이터 교육                                                    |                                                                       |
|          | 기반<br>역량 | 데이터 제도                                                    |                                                                       |
|          |          | 데이터 예산                                                    |                                                                       |
|          |          | 데이터 기술                                                    |                                                                       |

본 조사데이터 응답자들의 인구통계학적 특성을 제시하면 다음 〈표 3〉와 같다. 성별로는 남성이 221명, 여성이 109명으로 남성이 여성에 비해 2배 정도 많게 나타난다. 학력별로는 대학이나 대학교 졸업 이하가 193명, 대학원 졸업 이상이 137명으로 대학원 졸업 이상이 일반적인 공무원 조사에 비해 높게 나타난다. 이는 국책연구기관 연구원이 포함된 결과로 볼 수 있다. 직급별로는 1~5급이 32명, 6~7급이 142명, 8~9급이 53명으로 6~7급에 해당하는 인원수가 62.6%로 다른 급수에 비해 압도적으로 높게 나타난다. 연구위원급과 연구원급은 각각 51명과 52명으로 거의 차이가 없다. 경력별로는 3년 미만인 67명, 3~10년 미만인 86명, 10년 이상 177명으로 경력 10년 이상 응답자가 53.6%로 다른 유형에 비해 두 배 이상 높게 나타난다.

〈표 3〉 응답자 인구통계학적 특성

| 구분 |           | 빈도(명) | 비율(%)  |
|----|-----------|-------|--------|
| 기관 | 중앙행정기관    | 103   | 31.2   |
|    | 지방자치단체    | 124   | 37.6   |
|    | 국책연구기관    | 103   | 31.2   |
| 성별 | 남성        | 221   | 67.0   |
|    | 여성        | 109   | 33.0   |
| 학력 | 대학(교)졸 이하 | 193   | 58.5   |
|    | 대학원졸 이상   | 137   | 41.5   |
| 직급 | 1~5급      | 32    | 14.1   |
|    | 6~7급      | 142   | 62.6   |
|    | 8~9급      | 53    | 23.4   |
|    | 연구위원급     | 51    | 49.5   |
|    | 연구원급      | 52    | 50.5   |
| 경력 | 3년 미만     | 67    | 20.3   |
|    | 3~10년 미만  | 86    | 26.1   |
|    | 10년 이상    | 177   | 53.6   |
| 합계 |           | 330   | 100.00 |

## IV. 분석 결과

### 1. 기술분석 결과

우선 변수별 기술분석 결과를 비교분석해 보면 다음과 같다. 첫째, 대부분의 변수들이 4점 이상으로 나타나 데이터기반행정과 정부 데이터 역량의 높은 중요성 인식을 확인할 수 있다. 둘째, 여러 변수 중에서 종속변수인 데이터기반행정 중요성 인식이 4.39로 가장 높게 나타난다. 셋째, 그 다음으로 중요성 인식이 높은 것은 데이터 인력규모(4.28), 데이터 협업(4.20)으로 나타났다. 넷째, 중요성 인식이 4점을 넘지 못한 변수는 데이터 기술과 데이터 제도로 그 중에서 데이터 기술이 3.86점으로 더 낮게 나타났다. 이러한 분석결과에 대해 몇 가지 해석을 제시하면 다음과 같다. 우선 데이터기반행정의 중요성 인식이 높게 나타난 것은 무척 고무적이다. 그리고 데이터 기술이나 데이터 제도와 같은 기반 요소에 비해 데이터 인력규모나 데이터 협업과 같은 조직 역량,

인적 역량 요소가 높게 나타난다. 우리나라 전자정부 수준이나 데이터기반행정법 제정을 고려했을 때 데이터 기술이나 데이터 제도의 문제는 큰 문제가 아니라고 인식하는 것으로 보인다. 이보다는 데이터 전문인력 확보가 어려운 문제, 개인정보보호 등으로 데이터의 공유나 상호 교환이 어려운 문제를 더 심각하게 인식하는 것으로 보인다.

〈표 4〉 기술분석 결과

| 구분    |         |           | count | mean | std  | min | 50% | max |
|-------|---------|-----------|-------|------|------|-----|-----|-----|
| 종속 변수 | 데이터기반행정 |           | 330   | 4.39 | 0.89 | 1   | 5   | 5   |
| 독립 변수 | 조직 역량   | 데이터 기획    | 330   | 4.20 | 0.83 | 1   | 4   | 5   |
|       |         | 리더 데이터 스킬 | 330   | 4.01 | 0.85 | 1   | 4   | 5   |
|       |         | 데이터 협업    | 330   | 4.20 | 0.82 | 1   | 4   | 5   |
|       |         | 데이터 문화    | 330   | 4.17 | 0.85 | 1   | 4   | 5   |
|       | 인적 역량   | 개인 데이터 스킬 | 330   | 4.02 | 0.85 | 1   | 4   | 5   |
|       |         | 데이터 인력규모  | 330   | 4.28 | 0.80 | 2   | 4   | 5   |
|       |         | 데이터 교육    | 330   | 4.02 | 0.81 | 1   | 4   | 5   |
|       | 기반 역량   | 데이터 제도    | 330   | 3.97 | 0.86 | 1   | 4   | 5   |
|       |         | 데이터 예산    | 330   | 4.16 | 0.85 | 1   | 4   | 5   |
|       |         | 데이터 기술    | 330   | 3.86 | 0.85 | 1   | 4   | 5   |

## 2. 상관분석 결과

다음은 데이터기반행정과 정부 데이터 역량 각 변수 간 상관분석 결과를 살펴본다. 첫째, 모든 변수 사이의 상관관계가 유의수준 0.001 수준에서 통계적으로 유의미하게 나타났다. 둘째, 데이터기반행정과의 상관계수를 비교해보면 데이터 기획, 데이터 문화, 개인 데이터 스킬 등이 0.5 이상으로 상관관계가 상당히 높게 나타난다. 셋째, 조직 역량의 모든 요소들이 0.4 이상으로 나타나 데이터기반행정과의 관련성이 높게 나타난다. 넷째, 상관성이 가장 낮은 것은 데이터 전문인력 확보 문제로 상관계수가 0.364로 나타나고, 데이터 교육 0.380, 데이터 기술 0.389 등으로 데이터기반행정과 상관성이 상대적으로 낮게 나타난다. 다만 0.3 이상도 높은 상관성을 나타내는 것이므로 해석에 유의할 필요가 있다.

〈표 5〉 상관분석 결과

| 구분         |          | 1        | 2-1      | 2-2      | 2-3      | 2-4      | 3-1      | 3-2      | 3-3      | 4-1      | 4-2      | 4-3 |
|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----|
| 1. 데이터기반행정 |          | 1        | .        | .        | .        | .        | .        | .        | .        | .        | .        | .   |
| 2. 조직 역량   | 2-1.기획   | 0.581*** | 1        | .        | .        | .        | .        | .        | .        | .        | .        | .   |
|            | 2-2.리더   | 0.420*** | 0.531*** | 1        | .        | .        | .        | .        | .        | .        | .        | .   |
|            | 2-3.협력   | 0.469*** | 0.586*** | 0.633*** | 1        | .        | .        | .        | .        | .        | .        | .   |
|            | 2-4.문화   | 0.503*** | 0.600*** | 0.625*** | 0.584*** | 1        | .        | .        | .        | .        | .        | .   |
| 3. 인적 역량   | 3-1.개인역량 | 0.500*** | 0.584*** | 0.759*** | 0.616*** | 0.681*** | 1        | .        | .        | .        | .        | .   |
|            | 3-2.전문인력 | 0.364*** | 0.486*** | 0.515*** | 0.560*** | 0.512*** | 0.569*** | 1        | .        | .        | .        | .   |
|            | 3-3.교육   | 0.380*** | 0.517*** | 0.645*** | 0.628*** | 0.568*** | 0.679*** | 0.644*** | 1        | .        | .        | .   |
| 4. 기반 역량   | 4-1.제도   | 0.467*** | 0.567*** | 0.528*** | 0.625*** | 0.536*** | 0.623*** | 0.548*** | 0.641*** | 1        | .        | .   |
|            | 4-2.예산   | 0.437*** | 0.607*** | 0.401*** | 0.513*** | 0.493*** | 0.540*** | 0.589*** | 0.457*** | 0.514*** | 1        | .   |
|            | 4-3.기술   | 0.389*** | 0.549*** | 0.577*** | 0.616*** | 0.518*** | 0.653*** | 0.511*** | 0.656*** | 0.692*** | 0.582*** | 1   |

\*\*\*:  $p < .001$ 

### 3. 위계적 회귀분석 결과

마지막으로 위계적 회귀분석을 실시하였다. 모델 1은 조직 역량, 인적 역량, 기반 역량의 제요소를 포함하지만 성별, 학력, 직급, 경력 등의 통제변수를 포함하지 않는 반면, 모델 2의 경우 통제변수를 포함한다는 점에 두 모델의 차이가 있다. 각 모델의 F 값이 통계적으로 유의미하게 나타나고 설명력을 나타내는 R2 값도 각각 0.407, 0.414로 낮지 않으며, 모델 1에서 모델 2로 변화할 때 값이 증가하는 것을 볼 때 전체적으로 위계적 회귀분석모형은 적절한 것으로 판단된다. 또한 모델 2를 중심으로 VIF 확인 결과 값이 가장 높은 변수는 개인 데이터 스킬 변수로 4.4에 불과하여 다중공선성 문제는 나타나지 않았다.

위계적 회귀분석의 주요 결과는 크게 두 가지로 요약될 수 있다. 하나는 데이터기반행정과 통계적으로 유의미한 관계를 나타내는 변수에 대한 것이고 다른 하나는 통계적으로 유의미한 변수들의 표준화 계수값 비교와 관련된 것이다. 전자의 경우 조직 역량 중 데이터 기획과 데이터 문화, 인적 역량 중 개인 데이터 스킬, 기반 역량 중 데이터 제도 변수가 데이터기반행정과 통계적으로 유의미한 것으로 나타난다. 다만 데이터 제도 변수의 계수값은 모델 2에서는 통계적으로 유의미하지 않았다. 이러한 결과는 기존 선행연구들과 어느 정도 일치한다. 서재호(2020) 연구에서는 본 연구와 같이 문화적 요인은 데이터기반행정에 통계적으로 유의미한 영향을 미친 반면, 리더십

요인이나 인적 요인(인력규모)은 통계적으로 유의미하지 않았다. 윤수영·오재인(2018)의 연구에서도 문화 요인은 빅데이터 사용의도에 영향을 미치는 것으로 나타난 바 있다. 다만 이들 연구에서 유의미하게 나타난 기술적 요인은 본 연구에서는 유의미하지 않았다. 이를 통해 문화적 요인의 강건성을 확인할 수 있다. 서재호(2020)의 연구에서는 문화적 요인의 영향력이 기술적 요인에 비해 높게 나타났다.

다음은 후자의 표준화 계수값 비교의 경우를 살펴보자. 모델 2를 기준으로 데이터 기획, 데이터 문화, 개인 데이터 스킬의 표준화 계수값을 비교해보면 개인 데이터 스킬의 값이 0.284로 가장 높고, 데이터 기획의 값이 0.244, 데이터 문화의 값이 0.173의 순서로 높게 나타난다. 통계적으로 유의미하지는 않았지만 데이터 제도의 표준화값이 그 다음으로 0.132의 값을 나타냈다. 이러한 결과는 데이터기반행정 활성화를 위해 정부 데이터 역량을 높이하고자 하는 경우 개인 데이터 스킬, 데이터 기획, 데이터 문화 형성, 데이터 제도 등의 우선순위를 높게 할 필요가 있음을 시사한다. 특히 기존 연구에서 실증적 검토가 미흡했던 개인 데이터 스킬과 데이터 기획의 영향력이 높게 나타난 것은 눈여겨볼 만하다. 기존 논의에서는 당장의 데이터기반행정 수행 관점에서 접근하여 개인 데이터 스킬보다는 데이터 전문인력 확보에, 데이터 기획보다는 데이터 분석에 보다 관심이 있었던 것이 사실이다. 흥미로운 것은 이 두 가지가 상호 연결된 특성이 있다는 것이다. 즉 데이터 전문인력의 경우 데이터 분석 스킬은 우수하지만 정책 정책과 행정 영역에 대한 지식이나 경험 등의 전문성 부족으로 인해 실제적인 데이터 기획에는 한계를 보일 수 있는 것이다.

〈표 6〉 위계적 회귀분석 결과

| 변수 구분<br>(종속변수:<br>데이터기반행정) |    | 모델 1   |              |      |       |        | 모델 2   |              |      |       |        |
|-----------------------------|----|--------|--------------|------|-------|--------|--------|--------------|------|-------|--------|
|                             |    | Coef.  | Std.<br>Err. | t    | P>t   | Beta   | Coef.  | Std.<br>Err. | t    | P>t   | Beta   |
| 상수                          |    | 1.182  | 0.248        | 4.8  | 0.000 | .      | 0.656  | 0.548        | 1.2  | 0.233 | .      |
| 조직<br>역량                    | 기획 | 0.362  | 0.068        | 5.3  | 0.000 | 0.337  | 0.276  | 0.094        | 2.9  | 0.004 | 0.244  |
|                             | 리더 | -0.016 | 0.076        | -0.2 | 0.830 | -0.016 | -0.046 | 0.114        | -0.4 | 0.688 | -0.040 |
|                             | 협력 | 0.108  | 0.072        | 1.5  | 0.135 | 0.100  | 0.122  | 0.106        | 1.2  | 0.250 | 0.106  |
|                             | 문화 | 0.146  | 0.068        | 2.1  | 0.033 | 0.139  | 0.191  | 0.097        | 2.0  | 0.051 | 0.173  |
| 인적<br>역량                    | 개인 | 0.178  | 0.084        | 2.1  | 0.035 | 0.170  | 0.317  | 0.124        | 2.6  | 0.011 | 0.284  |
|                             | 규모 | -0.036 | 0.071        | -0.5 | 0.609 | -0.033 | -0.077 | 0.095        | -0.8 | 0.419 | -0.068 |
|                             | 교육 | -0.083 | 0.079        | -1.1 | 0.294 | -0.076 | -0.048 | 0.120        | -0.4 | 0.691 | -0.040 |

|            |    |              |       |      |       |        |              |       |      |       |        |
|------------|----|--------------|-------|------|-------|--------|--------------|-------|------|-------|--------|
| 기반<br>역량   | 제도 | 0.152        | 0.070 | 2.2  | 0.031 | 0.147  | 0.149        | 0.098 | 1.5  | 0.127 | 0.132  |
|            | 예산 | 0.070        | 0.067 | 1.1  | 0.291 | 0.068  | 0.014        | 0.089 | 0.2  | 0.877 | 0.012  |
|            | 기술 | -0.112       | 0.074 | -1.5 | 0.133 | -0.106 | -0.129       | 0.117 | -1.1 | 0.271 | -0.108 |
| 통제         | 성별 |              |       |      |       |        | 0.071        | 0.109 | 0.7  | 0.517 | 0.035  |
|            | 학력 |              |       |      |       |        | 0.171        | 0.147 | 1.2  | 0.247 | 0.064  |
|            | 직급 |              |       |      |       |        | -0.011       | 0.101 | -0.1 | 0.918 | -0.007 |
|            | 경력 |              |       |      |       |        | 0.086        | 0.072 | 1.2  | 0.236 | 0.075  |
| F(p-value) |    | 21.89(0.000) |       |      |       |        | 10.69(0.000) |       |      |       |        |
| R2         |    | 0.407        |       |      |       |        | 0.414        |       |      |       |        |
| Adj. R2    |    | 0.388        |       |      |       |        | 0.375        |       |      |       |        |

## V. 결론

본 연구에서는 데이터기반행정법 시행 이후 정부조직에서 데이터기반행정이 본격화되고 있는 가운데 데이터기반행정에 대한 이론적, 실증적 연구의 발전과 데이터기반행정의 정책적 우선순위 설정에 기여하기 위해 정부 데이터 역량이 데이터기반행정에 미치는 영향을 실증적으로 분석하였다. 이를 위해 데이터기반행정 연구의 동향 분석과 본 연구의 의의 제시, 정부 데이터 역량 개념화 및 유형화, 정부 데이터 역량과 데이터기반행정 간 관계에 대한 선행연구 분석 등을 통해 연구 분석틀과 모형을 설정하고 한국행정연구원의 ‘공공데이터 정책 활용 실태조사’ 데이터를 활용하여 기술분석, 상관분석, 위계적 회귀분석 등, 실증분석을 수행하였다. 기술분석 결과 대부분의 변수들이 5점 만점에 4점 이상으로 높게 나타났고, 그 중에서도 데이터기반행정 중요성 인식, 데이터 인력규모, 데이터 협업 등의 중요성 인식이 높게 나타났다. 상관분석에서는 종속변수인 데이터기반행정과 모든 독립변수들의 상관성이 통계적으로 유의미하게 나타나고 있어 각 독립변수에 대한 관리가 어느 정도 의미가 있음을 확인할 수 있었다. 특히 데이터 기획, 데이터 문화, 개인 데이터 스킬 등이 0.5 이상으로 상당히 높은 상관성을 보여주었다. 이러한 결과는 위계적 회귀분석 결과에서도 동일하게 나타났다. 통제변수까지 포함한 모델 2에서 통계적으로 유의미한 변수가 바로 이 세 변수로 나타난 것이다. 표준화 계수를 분석한 결과 개인 데이터 스킬의 표준화 계수값이 가장 높게 나타났고, 데이터 기획, 데이터 문화의 순서를 나타내었다. 통제변수를 포함하지 않은 모델 1에서는 데이터 제도가 통계적으로 유의미한 관계를 나타내었다.

이러한 분석결과를 토대로 데이터기반행정 활성화를 위한 정책적 시사점을 제시하면 다음과 같다. 첫째, 정부조직의 개별 구성원들의 데이터 역량을 높이는 것이 중요하다. 데이터 시대가 급속하게 도래하면서 정부조직 구성원들이 이에 적응하는 것이 힘겨운 것이 현실이다. 모든 공무원의 데이터 역량을 높이는 것이 가능하지도 바람직하지 않다는 의견도 있다. 윤건·김철우(2020)는 데이터 분석 역량에 강점을 가지는 국책연구기관과 데이터 관리 역량에 강점을 가지는 행정기관과 적절한 역할 분담을 제안한다. 또한 인사혁신처에서 새롭게 신설한 데이터 직류 역시 이러한 관점에서 이해될 수 있다. 다만 이러한 논의가 개별 구성원들을 방치하는 것으로 이어져서는 안 된다. 데이터기반행정은 데이터 전문가들만이 수행하는 것이 아니라 정부의 전 구성원이 힘을 모아야 성공할 수 있는 것이다. 결국 데이터 전문가의 영입은 끝이 아니라 시작을 의미하며 이것이 개별 조직 구성원의 데이터 역량을 높이는 계기가 되어야 할 것이다.

둘째, 데이터 기획의 중요성이다. 데이터기반행정이 보다 체계적으로 운용되기 위해서는 데이터 기획이 무엇보다 중요하다. 정책문제를 해결하고 보다 나은 행정서비스를 제공하기 위해 어떠한 데이터를 필요로 하는지, 그러한 데이터가 존재하는지, 존재한다면 어디에 존재하는지, 그것을 실제 확보할 수 있는지, 데이터가 존재하지 않는다면, 그리고 데이터를 찾는 데 드는 시간적, 물적 비용이 많이 소요된다면, 데이터를 직접 생성할 수 있는지 등을 확정해야 한다. 그런데 흥미로운 것은 데이터 기획이 데이터 스킬만으로는 한계가 있다는 것이다. 효과적인 데이터 기획을 위해서는 개별 정책문제나 행정서비스 자체에 대한 지식과 경험을 통해 깊이 있는 내용 전문성을 함께 가지고 있어야 한다. 개별 분야의 정책이나 행정 내용에 대한 전문성이 적합한 데이터를 찾고 활용하는 데이터 기획의 과정에서 적절한 가이드의 역할을 할 것이다.

셋째, 데이터 문화 형성 전략에 힘을 쏟아야 한다. 개인의 데이터 역량이나 데이터 기획조차도 중장기적으로 데이터 문화 형성에 영향을 받는다. 데이터 문화가 형성된다는 것은 조직의 모든 구성원들이 데이터에 기반하여 사고하고 판단하고 결정하는 것이 익숙해지는 것이다. 다만 분명한 것은 데이터 문화가 하루아침에 형성될 것으로 기대해서는 안 된다는 것이다. 데이터 문화 형성을 위해서는 중장기적 문화 형성 전략에 따라 체계적이고 일관되게 추진해야 한다. 그렇지 않으면 데이터기반행정의 어려움으로 인한 회의감이 나타날 때 데이터기반행정의 효율성이나 효과성, 신뢰성 등에 대한 기대가 급속도로 무너질 수도 있다.

마지막으로, 데이터 제도에 대한 지속적인 관심이 필요하다. 데이터기반행정법의 형성은 데이터기반행정의 이상을 구현하기 위한 첫 번째 단계에 불과하다. 실제 많은 선행연구들에서 데이터기반행정법의 미비점들을 지적하고 있다. 따라서 데이터기반행정법을 지속적으로 업데이트하고 업그레이드하여 이러한 미비점들을 지속적으로 개선해가야 할 것이다. 이러한 데이터 제도와 관련하여 데이터기반행정의 가장 큰 걸림돌은 개인정보 활용의 문제라고 할 수 있다. 실제 데이터기반행정의 보안성을 높게 유지하는 것은 활용의 비효율성을 낳는다. 규제 샌드박스나 규제 프리존과 같은 규제 혁신의 다양한 방법론을 적용하면 조금이나마 개선되는 점이 있을 수 있다. 그러나 이것이 인권의 문제와 결부되어 있기 때문에 근본적인 해결은 쉽지 않다는 점을 고려할 필요가 있다. 특히 데이터기반행정을 구현하고자 하는 정부조직의 구성원들은 데이터 시대가 될수록, 데이터기반행정을 통해 국민의 삶의 질을 높일 수 있는 환경이 갖추어져 갈수록 데이터의 소중함과 함께 데이터의 잘못된 사용에 대한 경각심도 함께 높아갈 필요가 있다.

본 연구는 정부 데이터 역량과 데이터기반행정 간 관계를 분석함으로써 데이터기반행정에 대한 실증연구를 증진시켰다는 의의가 있지만 이론적으로나 실증분석에서 몇 가지 분명한 한계를 가지고 있으므로 향후 연구에서는 이러한 점들을 보완하여 연구를 수행할 필요가 있다. 첫째, 설문조사 데이터의 한계성이다. 향후 실증분석에서는 다양한 객관적 자료에 기반하여 보다 객관화된 관계 분석을 시도하는 것이 바람직하다. 둘째, 개별 변수를 중요성 인식으로 측정 한계가 있다. 이는 2차 자료를 사용하는 연구로서의 한계라고 할 수 있다. 향후 조사설계에서는 측정문항 개발을 통해 개별 변수를 측정하는 보다 적절한 방안을 마련할 필요가 있다. 셋째, 설문조사 인식 데이터를 활용하면서 개별 변수에 하나의 단일 문항을 사용하면서, 복수의 설문문항을 사용하여 요인분석과 신뢰도분석 등을 통해 보다 엄밀하게 측정하지 못한 한계가 있다. 이 역시 2차 자료의 한계라고 할 수 있으므로 향후 연구에서는 설문문항 설계에 보다 세밀한 관심을 가질 필요가 있다.

## 참고문헌

김법연. (2020). 데이터기반행정 관련 국내 법제 동향. 『경제규제와 법』, 13(2): 188-197.

- 김승현·박재홍·김이환·박주석. (2015). 국내기업을 위한 빅데이터 적용방안 도출 및 기대효과 분석- 데이터 분석 및 활용과 업무사용자의 의사 결정 향상에 관한 연구 -. 「정보기술 아키텍처 연구」, 12(1): 159-170.
- 김효중·김희웅. (2021). AI·데이터 사이언스 분야 직무 역량 강화를 위한 커리큘럼 연구. 「정보화 정책」, 28(2): 34-56.
- 서웅·안성진. (2019). 초등학생의 데이터 수집, 분석, 표현 수업을 위한 세부역량 개발 및 적용. 「정보교육학회논문지」, 23(2): 131-139.
- 서재호. (2020). 지방자치단체 공무원의 데이터기반 행정 영향 요인에 대한 탐색적 연구: 자치구 공무원의 인식조사를 토대로. 「지방정부연구」, 23(4): 445-464.
- 선지원. (2021). 데이터 기반 행정의 법적 문제. 「행정법연구」, 65: 107-135.
- 윤건·김윤희. (2020). 공공데이터 융합의 영향요인과 매개효과에 관한 탐색적 연구: 데이터기반 행정의 맥락을 중심으로. 「한국지역정보학회지」, 23(3): 1-22.
- 윤건·김철우. (2021). 증거기반정책과 데이터기반행정법 미국과 영국의 법·제도 비교 및 함의. 「가천법학」, 14(3): 83-112.
- 윤건· 김철우. (2020). 지자체 차원의 데이터기반행정 구현을 위한 요인 분석: 핵심 구성요소와 주요 참여자를 중심으로. 「한국지방자치학회보」, 32(4): 71-101.
- 윤건·심우현·박정원·김윤희. (2018). 정부역량 측정을 위한 지표체계 구성에 관한 연구. 「한국 사회와 행정연구」, 29(3): 143-158.
- 윤수영·오재인. (2018). IT 자원, 인적 자원, 조직적 자원에 따른 빅데이터 사용의도에 미치는 요인에 관한 연구. 「한국경영과학회지」, 43(4): 67-94.
- 이승중·윤두섭. (2005). 지방정부의 역량에 관한 개념화 연구. 「한국지방자치학회보」, 17(3): 5-24.
- 이태원·성해남·김은정. (2020). 키워드 네트워크 분석을 이용한 MIS 교과정보와 NCS 기반 빅데이터 분석 직무역량에 대한 연구. 「정보시스템연구」, 29(4): 101-121.
- 정성호·전주열. (2020). 데이터 기반 행정법과 증거기반 정책결정. 「한국비교정부학보」, 24(3): 173-191.
- Chatfield, A., Reddick, C., & Al-Zubaidi, W. (2015). Capability challenges in transforming government through open and big data- Tales of two cities.Thirty Sixth International Conference on Information Systems.
- Davenport, T. H., Harris, J. G., De Long, D. W., and Jacobson, A. L. (2001). Data to Knowledge to Results- Building an Analytic Capability. California Management Review, 43(2): 117-138.
- Donahue, A. K., Selden, S. C., and Ingraham, P. W. (2000). Measuring government management capacity: A comparative analysis of city human resources management systems. J-PART, 10(2): 381-411.

- Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., and Haydock, M. (2011). Analytics: The widening divide. MIT Sloan Management Review Research Report, Fall, 1-20.
- McClelland, D. C. (1973). Testing for competence rather than for "Intelligence". American psychologist, 28(1): 1-14.
- Prahalad, C. K., & Hamel, G. (1990). The core competence of the corporation. Harvard Business Review, May/June: 1-15.

## Abstract and Key Words

### **An Empirical Study on the Effects of Governmental Data Capacity on Data-Based Public Administration**

Yoon, Kun  
Hanshin University

Since the enactment of the Data-based Public Administration Act in June 2020, data-based public administration (DBPA) has been in full swing in government organizations. DBPA can be said to be one of the government strategies in the era of digital transformation that actively utilizes the results of scientific data analysis in the government's administration and policy process to enhance the government's scientific prowess and thereby enhance the performance of government organizations. In this study, considering that data competence or data capacity/capability is essential for the success of data-based administration in government organizations, the purpose of this study is to empirically verify whether the data-based capabilities of government organizations have a statistically significant effect on DBPA. An analysis model was established through the analysis of previous studies, and descriptive analysis, correlation analysis, and hierarchical regression analysis were performed using the data of the Korea Institute of Public Administration's 'Public Data Policy Utilization Survey'. As a result of the analysis, variables such as data planning and data culture in the organizational capacity, personal data competency in the human capacity, and data institution in the foundational capacity were found to be statistically significant. Based on these analysis results, policy implications are presented.

[ Key Words: government capacity, data capacity, government data capacity,  
data-based public administration ]