



구 분			
열람·서명자	빅데이터심의위원	성명 김준호	(서명 또는 인) 
	빅데이터심의위원	성명 김진수	(서명 또는 인) 
확 인 자	빅데이터담당관	성명 윤충식	
작 성 자	빅데이터담당관	성명 박창현	

## 2024년 1차 빅데이터심의위원회 회의록

2024. 4.

**디지털정책관**

(빅데이터담당관)

# 2024년 1차 「빅데이터심의위원회」 회의록

## 〈회의개요〉

- ◆ 일 시 : '24. 4. 18.(목) 15:00~17:00
- ◆ 장 소 : 서울특별시 본청 영상회의실(6층)
- ◆ 참석위원 : 12명
  - 외부위원(9명) : 김준호, 김진수, 안한희, 이경전, 이영섭, 정지영, 진혜진, 최은정, 최진혁
  - 내부위원(3명) : 김기현 안전총괄관, 박진영 디지털정책관, 송호재 노동공정상생정책관
- \* 배석 : 이상용 디지털수석, 윤충식 빅데이터담당관, 담당 팀장 등
- ◆ 진행순서
  - 참석자 소개, 위원장 인사말씀, 위원장 및 부위원장 선출
  - 빅데이터 정책 등 주요사업 안건 공유
    - 서울시 데이터기반 행정 종합계획 등 심의
  - 안건에 대한 추진과제 제안 및 토론(자문)
- ◆ 상정안건
  - 심의 1건
    - 【1】 서울시 데이터기반 행정 종합계획(안) 수립 보고
  - 보고(자문) 3건
    - 【1】 서울시 데이터 관련 정부평가 보고
    - 【2】 빅데이터 표준분석모델 후보과제 선정 자문
    - 【3】 수도권 생활이동(통근·통학인구) 데이터 개발·활용 보고

## □ 위원 발언내용

### <간사>

- 개회 및 진행안내
- 위원 및 내부참석자 소개
- 위원장 및 부위원장 선출

### <위원장>

- 인사말씀
- 회의록 서명의원 선출

### <간사>

- (심의안건) 서울시 데이터기반 행정 종합계획
- (보고안건) 서울시 데이터 관련 정부평가 보고
- (보고안건) 빅데이터 표준분석모델 후보과제 선정 자문
- (보고안건) 수도권 생활이동(통근·통학인구) 데이터 개발·활용 보고

### <위원장>

- 보고 받은 안건에 대해 논의하는 시간을 갖도록 하겠음.

### <000위원>

- 공공데이터를 활용하여 새로운 사업과 서비스를 유도하여 예비 창업자에게 기회를 제공하는 사업에 대해 설명 요청.

### <간사>

- 공공데이터를 활용한 창업경진대회가 열리며, 시민들은 앱이나 웹을 개발하여 실제 비즈니스를 시작할 수 있고, 선발된 참가자들은 상금과 지원 혜택을 받게 됨.
- 서울시 창조경제혁신센터와 한국장학재단은 창업 과정을 지원하고, 창업 단계에 따라 맞춤형 프로그램을 제공함.

### <000위원>

- 생성형 AI 기반 챗봇 서비스 관련하여 사업설명과 볼 수 있는 사이트 설명요청.

### <간사>

- 공공 데이터를 활용한 AI 챗봇 서비스가 공공 부문에서 개발 중이며, 자연어 처리를 통해 질문에 답변하고 요약하며, 그래프도 생성하여 정보를 제공할 예정임. 이 서비스는 9월에 빅데이터서비스 플랫폼을 통해 시민들에게 공개될 예정이며, 학교 숙제에 활용되는 등 다양한 활용이 기대됨.

### <000위원>

- 로컬 브랜드 사업의 매출액 데이터가 전체 상권에서 얻어지므로 특정 분야의 매출을 파악하기 어려워, 특정 분야의 매출액을 파악하는 방안 필요.
- 전통시장 활성화를 위해 시장 매출 데이터 빠르게 확보하여 카드 및 QR 코드 사용 현황 파악하고, 사업 성과 측정을 위한 보조 수단 개발 필요.

### <간사>

- 신용카드의 매출 매입 자료를 기반으로 상권 분석을 진행하고, 업종별 필요한 데이터를 추가로 수집하여 맞춤형 분석 지원을 계획 중임. 더불어 추가적인 지불 수단에 대한 연구도 진행할 예정임.

### <000위원>

- 안전분야에서는 체계적인 다중 밀집지역 인파관리 시스템을 구축하고자 노력하고 있으며, 초기에는 데이터 기반으로 추진했으나 정확도와 시간적 제약으로 CCTV 기반의 AI 활용 시스템으로 전환하였음. 이번 고도화 프로젝트에서는 용역을 추진 중이며, 빅데이터 분석과제 명칭을 혼동을 방지하기 위해 '인파관리 상황 지원'으로 변경이 필요함.

### <000위원>

- 다양한 정책들 중 우선순위에 밀린 정책들을 검토하고, 데이터 수집과 활용에 대해 중요성을 강조하며 협업을 제안함. 특히, 서울시 외에도 민간 데이터나 해외 데이터를 확보하여 시민에게 유익한 정보를 제공하고, 벤치마킹을 통해 공무원들이 쉽게 분석하고 정책을 개선할 수 있도록 제안함. 이를 위해 서울시 내부 데이터뿐만 아니라 다양한 외부 데이터를 수집하고 활용하는 계획을 논의하고 있음.

### <간사>

- 공공데이터는 민간 경제에 쓸 수 있는 개방데이터, 데이터 기반 정책 의사결정 지원을 위한 데이터 두가지 목적이 있음. 수집된 데이터는 시민 개방과 분석 결과 활용을 위해 활용되며, 연간 데이터는 생활 인구, 통신사, 카드사, 은행 등에서 융합 데이터를 수집하여 개방함.
- 또한, AI 학습용 데이터의 중요성에 대해 강조하며, 서울시에서는 AI 종합 계획을 통해 다양한 분야에서 필요한 데이터를 수집하고 있음.

### <위원장>

- 민간 데이터 활용을 위해 안전한 가명 정보 결합이 중요하며, 서울시가 민간 데이터를 활용하는 경우, 자치구가 동일한 데이터를 활용하고자 할 때 별도의 계약이 필요한지 질문함.

### <간사>

- 별도 계약은 필요없으며, 같이 시/구가 같이 쓰고 있음.

### <000위원>



○서울시의 데이터 활용과 관련된 성과 및 노력은 어떻게 되는지, 특히, 외국인 관련 데이터 확보와 서비스 제공 노력에 대한 변화와 공유 방식이 중요한 관심사임. 이를 위해 서울시는 데이터 관련 성과를 투명하게 보고하고, 공무원과 시민들과의 공유를 강화할 필요가 있음.

#### <간사>

- 서울시는 데이터 개방 성과로 작년에는 7,800개에서 8,100개로 향상되어 KPI를 달성했으며, 현재 공공데이터포털에는 공공기관/지자체를 포함해 8만여 개의 데이터가 있는데, 서울시는 이중 약 10% 이상을 확보하여 상당한 물량을 보유하고 있음.
- 민간과의 융합데이터 생성 및 실적으로는 생활인구, 생활이동, 생활물류 등을 KT, CJ 등과 협력하여 추진 했음.

#### <000위원>

- 서울시는 숙박 데이터와 같은 민간 데이터 확보에 대한 어려움을 인지하고, 관광객이 아닌 거주 외국인에 대한 데이터 수집 노력이 필요함을 인식하고 있음. 이를 위해 서울시는 외국인 거주자 데이터를 포함한 다양한 데이터에 접근하기 위한 계획과 노력이 필요. 이는 외국인이 시민으로서의 역할을 수행하는 데 필요한 정보를 보다 잘 이해하고 고려하기 위함.

#### <간사>

- 서울시는 매년 약 10억에서 12억 원의 예산을 투입하여 민간 데이터를 구매하고 협업하여 채널을 확보하고 있음. 특히 관광 분야에서는 숙박관련 업체와 채널 확보를 위한 노력을 하며, 매년 2개씩의 발굴 KPI를 설정하여 추진하고 있음.

#### <000위원>

- 외국인들이 카드를 이용할 때 주로 신용카드를 등록하며, 로밍을 통해 이동하는 것으로 알려져 있음. 특히 동남아지역에서 학생 및 직장인들이 계속적으로 방문하는 추세이며, 이러한 외국인들의 생활 패턴에 대한 데이터 수집 및 분석이 국가 전반적으로 이루어지고 있다고 함.

#### <위원장>

- 안건1 수정안에 대한 의결에 이의가 있으십니까?

#### <위원>

- (다같이) 없음

#### <위원장>

- 오늘 첫 번째 안건 서울시 데이터기반 행정 종합계획안은 수정한 심의안을 통과하는 것으로 그렇게 하겠음.

#### <위원장>

- 그 다음에 보고 안건 3건에 대해서 우리가 토론을 하도록 하겠음. 위원님들께서 질문이 있으

시면 질문해 주시기 바랍니다.

#### <000위원회>

- 서울시의 지능형 CCTV 입지분석 모델 후보 과제는 범죄 예방에 초점을 맞추었지만, 재난안전 예방과 관리에 대한 내용이 미흡하다는 의견이 제시함. 이 모델의 우선순위 결정 시 시민의 생명과 안전을 보호하는 측면이 반영되어야 한다는 주장함.

#### <간사>

- 서울시는 CCTV 입지 분석을 통해 범죄 예방에 중점을 두고 있으며, 최근에는 지능형 CCTV 의 확산을 추진하고 있음. 이러한 기술은 범죄 예방과 쓰레기 투척 등의 예측에 활용되고 있으나, 움직임 학습 과정에서 아직 오차가 있어 개발이 필요함.

#### <간사>

- 표준분석모델 4개의 후보 과제를 자료 검색과 지자체 선호도 조사를 통해 선정했음. 심의회를 통해 전문가 자문을 받아 모델 개발 시 애로사항을 반영하고, 5월에 2개의 모델을 최종 확정 하여, 11월까지 진행할 예정임.
- 전기차 충전소 입지 분석에서 잘 쓰는 지역과 적게 쓰는 지역의 모델을 하나로 통합하면 예측 능력이 떨어질 우려가 있어, 테스트 베드 2개를 선정하여 편차를 검증하고 최종적으로는 1개의 모델을 채택하는 방향으로 진행될 예정임.
- 4개의 모델 중 2개를 선정하여 서울시 표준 분석 모델을 만들 예정임.

#### <위원장>

- 행안부가 전국 지자체에 표준 분석 모델을 만들어 제공하는데, 예산 중복을 피하기 위해 관광이나 안전과 같은 분야에 대한 모델을 각 지자체가 개별적으로 만드는 것이 아닌, 표준 모델을 활용함. 이 모델은 원천이 되어 각 지자체가 필요에 맞게 변형하여 사용할 수 있음. 예를 들며, 서울시도 자치구 간 외국인 분석하면, 표준 모델을 활용하여 자치구에서 만든 분석을 나중에 다른 구에서도 활용할 수 있음.

#### <000위원회>

- 표준 분석 모델사업인 행안부의 '혜안'과 중첩되거나 유사성이 있는지 확인이 필요

#### <간사>

- 서울시는 전국 표준 분석 모델을 적용했으나 일부 데이터 불일치로 인해 제대로 작동하지 않았음. 따라서 서울시에 특화된 모델을 개발하고 행안부에 충분한 설명을 통해 승인을 받음. 이로써 서울시는 표준 모델 대신 서울에 맞는 모델을 사용할 수 있게 되었음.

#### <000위원회>

- 민원이 불법 주정차 단속구간에서 가장 중요한 부분이며, 교통량과 사고 기록은 중요한 데이터임. 그러나 재난 및 안전에 관한 데이터도 반드시 고려되어야 하며, 소방차 및 구급차의 진

입 경로 등 시민 안전을 우선적으로 고려해야 함.

#### <간사>

○ 저희는 화재 발생 시 출동 시간 내에 도달하는 골든타임 모델을 사용함. 이 모델은 불법 주정차와 도로 폭 등을 고려하여 작성되었으며, 이미 행안부에서 만든 모델을 기반으로 하고 있음. 만약 불법 주정차가 발생한다면 해당 모델을 참고하여 반영 예정.

#### <위원장>

○ 행안부는 수년간의 수요 조사를 통해 다양한 노하우를 축적해왔음. 이번 서울시의 분석 모델 개발에서는 이러한 전문 지식을 활용하여 서울시에 특화된 모델을 개발될 것으로 기대함.

#### <000위원>

○ 서울시의 데이터 모델링과 표준화에는 연합학습 및 로컬라이제이션의 원칙이 적용될 수 있음. 각 구는 자체적인 데이터를 관리하고 구별 특성을 반영한 모델을 개발할 수 있으며, 동시에 표준화된 데이터와 모델을 공유함. 연합학습을 통해 구별 데이터의 특성을 고려하여 중앙 모델을 발전시키고, 로컬모델을 구에게 제공함으로써 혁신적인 데이터 관리가 가능해짐. 이를 통해 서울시의 데이터 활용과 표준화가 현대적이고 효과적으로 이루어질 수 있음.

#### <000위원>

○ 자치구에 모델 개발을 맡기는 것은 번거로울 뿐 아니라 비용도 많이 듦. 따라서 시가 모델 개발을 담당하고, 자치구는 데이터 관리에 집중하는 것이 좋겠음. 이를 통해 데이터 중심의 AI를 구현하고, 모델 개발에 필요한 비용과 시간을 절약할 수 있음.

#### <000위원>

○ 영상데이터에 대한 활용에 대해 궁금함.

#### <간사>

○ 도로 영상을 활용하여 포트홀을 감지하고, 이를 기반으로 포트홀이 재발생할 시기를 예측하며 실시간으로 모니터링함. 예측 결과를 기반으로 도로 보수 계획을 수립하고 적용되는 걸로 알고 있음.

#### <000위원>

○ 서울시의 버스와 택시에 장착된 블랙박스를 통해 수집된 영상 데이터는 교통 상황 및 사고 발생 등의 정보를 실시간으로 업데이트하는 데 활용될 수 있음. 사생활 보호와 데이터 보호 문제를 고려하여, 해당 기업들과의 협약을 통해 데이터를 확보하고 활용하는 방안을 모색할 필요가 있음.

#### <000위원>

○ 페더레이트 러닝(Federated learning)은 민간기업에게 AI를 제공하고, 데이터는 그들이 가지고



학습하여 모델을 개선하는 방식으로 데이터 보호와 활용을 조화시킬 수 있음. 서울시와 자치구, 민간 기업과의 협력을 통해 교통 정보를 예측하는 AI를 개발하고, 데이터 보호와 성능 향상을 고려하여 시스템을 구축하는 접근 방식이 필요함. 프라이버시와 데이터 보호를 고려하면서도 효율적인 교통 정보 시스템을 구축하기 위해 새로운 접근 방식을 고려해야 함.

#### <000위원>

○ 공공데이터를 활용한 창업경진대회에 대해 알려주세요.

#### <간사>

○ 올해부터 행안부가 주최하는 공공데이터 활용 창업 경진대회는 예선전 후 본선이 진행되며, 공무원 대상의 우수 사례 경진대회와 일반 시민 및 학생 대상의 빅데이터 공모전도 별도로 진행됨. 창업 경진대회는 창업 지원 프로그램과 연계하여 참가자들이 실제 창업을 위한 프로그램을 받을 수 있도록 지원함.

#### <000위원>

○ 데이터를 실시간으로 활용하는 서비스를 제공하기 위해서는 API 형태로 제공되어야 하며, 과거 데이터를 쌓는 것보다는 실시간 데이터 반영이 중요함. 공공 데이터 활용 서비스의 개발과 관리는 지속적으로 실시간 데이터를 제공하는 API를 통해 이루어져야 하며, 이를 통해 서비스의 효율성과 신속성을 높일 수 있음.

#### <000팀장>

○ 열린데이터광장의 공공데이터가 7,800개 개방되어 있는데, API형태는 5,000여건이 해당됨.

#### <위원장>

○ 수도권 생활 이동 데이터 개발 관련하여 수도권이면 서울시뿐만 아니라 인천/경기도인데, 협의가 잘 되고 있습니까?

#### <간사>

○ 작년에 협의회를 설립하여 KT와 협력하여 250m 격자 형태의 데이터를 생성했고, 이를 공무원은 격자 단위로, 시민들은 동 단위로 개방할 예정이며, 교수님들께는 학생들이 활용할 수 있는 과제를 제공할 것을 요청함.

#### <000위원>

○ 공공데이터포털에서 제공되는 데이터는 CSV 형식과 기계 형식으로도 제공되며, API를 통해 사용할 수 있어서 학생들을 포함한 다양한 사용자들이 다양한 형태로 데이터를 활용하고 있음.  
○ 공공데이터 운영실태 평가 결과에서 개방데이터 제공 주기 준수율이 상대적으로 낮게 나타났는데 그 이유는 무엇인가요?

#### <간사>



○서울시의 450개 시스템에 대한 권한이 실국에 있어, 업데이트 요청이 필요한데, 실국 담당자들과의 연락이 어려워 시스템 관리가 어려움을 겪고 있음.

#### <000위원>

○빅데이터 공모전에서 상장과 예산을 제공하는데, 개발된 솔루션이 단순히 개발 후 종료되는 것이 아니라 실제 서비스와 연계되거나 활용되어야 함을 강조하며, 지속적인 활용과 연계를 통해 더 큰 가치를 창출하고자 함.

#### <간사>

○행안부에서 4년 동안 운영한 공모전의 결과로 약 70%의 활용이 된 것으로 나타남, 학생들의 의견이 표준 분석 모델로 채택되거나 실국에서 사업으로 진행되는 등 다양한 형태로 활용되고 있음. 공모전 출품작은 서울시의 입지 분석 모델 등 정책 적용에 높은 영향을 미침.

#### <간사>

○공모전 출품작의 지적재산권을 공동으로 공유하고, 성과 공유 대회와 시상식 등을 통해 행안부는 인센티브를 제공하며, 스마트 라이프 워크와 같은 국제 행사에서 발표 및 시민들에게 알릴 기회를 마련하여 활동을 지속하고 있음.

#### <000위원>

○다양한 기관에서 진행되는 공모전에서 심사자들이 겹치는 제출물이나 중복을 식별하는 것이 어려운데, 초기부터 공공기관 간 협업을 강화하여 정보 및 제출물을 공유하고 중복을 최소화하는 방안이 필요하다는 의견이 제시됨. 서울시와 참가자 간의 정보 공유를 확대하여 공공기관 간의 협업을 강화하면 공모전의 효율성을 높일 수 있을 것으로 예상됨.

#### <간사>

○시상 전 후보에 대한 공문 발송하여, 타 기관에서 유사한 제출물로 상을 탄 적이 있는지 검증하고, 이를 통해 중복된 제출물을 탈락시키는 행정 절차를 시행하여 공정성을 유지하고 있음.

#### <000수석>

○서울시의 데이터가 특정분야 보유량이 부족한 상황에서 과제를 발굴하는 것이 어려운 점을 언급하며, 중앙정부와의 협력이 필요하다고 지적함. 특정분야 데이터 양의 한계로 인해 과제 발굴이 어려워져 있으며, 중앙정부와의 협력을 통해 이를 극복할 필요성을 강조함.

#### <위원장>

○로드맵 수립 시에는 짧은 기간 내에 달성 가능한 결과와 중장기적 목표를 명확히 구분하여 고려해야 함. 이를 통해 담당자들이 성과를 보여주고 예산 등의 압박을 최소화할 수 있을 것으로 예상됨.

#### <000수석>

- AI 전문가와 교수들이 데이터를 활용하여 복지 분야의 고독사 등의 문제를 해결할 수 있는 AI 기술 개발에 대한 중요성을 강조함. 이를 통해 빅데이터와 AI 기술이 융합되어 사회 문제 해결에 기여 할 수 있을 것으로 기대됨.

#### <000위원>

- 서울시는 AI 행정의 원년을 선포하고 AI 기술을 종합적으로 활용하기 위한 계획 발표하였으며, 각 실국별로 흩어져 있던 사업을 중앙에서 관리하고 품질 및 신뢰도를 높이기 위한 심의 절차와 가이드를 도입했다. 빅데이터와 AI가 한 몸처럼 생각되며, 이를 위해 많은 비용과 인력을 투입하여 공개와 활용을 촉진하고 있으며, 향후 더 많은 성과를 이루고 행정에 적극적으로 활용하기 위해 노력할 것이라고 밝혔음.

#### <위원장>

- 이상으로 이제 오늘 2024년 제1차 서울시 빅데이터 심의위원회를 마치도록 하겠음. 참석해 주신 위원님들께 다시 한 번 감사드립니다. 이것으로 마치도록 하겠음.