

(양식) 데이터명세서

구분	조사항목	유형	답변내역	단답형, 서술형 등 작성 예시
1	대학명	단답형		(예: OO대학교)
2	소속기관/부서	단답형		(예: OOOO학과 연구실, OOOO 연구소)
3	담당자 성명/직책	단답형		(예: 홍길동/연구원)
4	담당자 연락처	단답형		(예: 02-0000-0000 / 010-0000-0000)
5	메일주소	단답형		(예: OOOO@nia.or.kr)
6	데이터 개방여부	선택형		
7	데이터 개방 플랫폼	단답형		(예: AI허브, 공공데이터포털, 모두의 말뭉치, KAMP 등)
8	중복성 자체검토	선택형		AI-Hub, 공공데이터포털, 원·윈도우 자체 검토 후 작성
9	데이터셋 명칭	단답형		(예: 자율주행 도로 환경 및 장애물 인지 데이터)
10	소속 도메인	선택형		
11	데이터 규모/용량	단답형		(예: 자연 객체 데이터 1,500,000건 / 500GB + 인공 구조물 데이터 2,000,000건 / 500GB), (예: □□□데이터 1,000,000건 / 500GB)
12	데이터 구축 시기	선택형		
13	샘플데이터 구성비	단답형		(예: 자연 5건, 시가지 5건, 건축물 5건)
14	라벨링 구축 형태	선택형		
15	개방을 위한 권리 확보 여부	선택형		
16	개인정보 비식별화 여부	선택형		
17	기술 표준 준수	선택형		
18	데이터 매칭 구조	선택형		
19	클래스 분포 상태	선택형		
20	상세 중빙 목록	재출형		데이터 구축 결과서, 품질관리 결과서, 데이터 설명서, 라벨링 가이드, 메타데이터 정의서 등 제출한 산출물 내역 작성
21	재사용 잠재력	선택형		
22	권장 학습 모델	선택형		
23	주요 예상 수요처	서술형		본 데이터 개방 시 활용 가능한 구체적인 기업이나 서비스 시나리오 기재
24	원천 수집 경로	선택형		
25	데이터 업데이트 주기	선택형		
26	국정과제 및 정부정책 연계성	서술형		(예: 국정과제 20 'AI 3대 강국 도약' 중 '핵심 분야인 법률 AI 고도화에 필수적인 자산인, 특히 하급심 판결문은 민간 수집이 불가능하여 공공 서비스 혁신(법률 복지) 기여도가 매우 높음
27	기술 트렌드 및 산업특화 가치	서술형		(예: 본 데이터는 단순 로그를 넘어 '교장 원안-조치' 간의 CoT(추론) 구조와 공정 전문 지식을 포함하여 재조 특화 LLM 고도화에 최적화되어 있어, 개방 시 인간의 지능형 예지보전 에이전트 서비스 상용화 기여 가능)