

발간등록번호

11-1240000-001631-14

『인공지능산업실태조사』 2022년 정기통계품질진단 결과보고서

2022 Regular Assessment Report

한국통계진흥원

2022. 12.

본 보고서는 한국통계진흥원이 통계청으로부터 위탁을 받아 진단한 결과입니다. 보고서의 내용은 한국통계진흥원(연구진)이 진단한 내용이며, 통계작성기관의 확인을 거쳐 작성했습니다.

제 출 문

통계청장 귀하

본 보고서를 “『인공지능산업실태조사』 2022년 정기통계품질진단”
과제의 최종보고서로 제출합니다.

2022 년 12 월 15 일

연 구 원 : 한국통계진흥원 황규배 연구위원

표 본 연 구 원 : 한국통계진흥원 오유진

조사표·유사통계
연 구 원 : 한국통계진흥원 조준기

M D 연 구 원 : 한국통계진흥원 조준기

목 차

결과보고서 요약문	1
정기통계품질진단 흐름도	2
제 1 장 진단대상통계 개요	3
제 2 장 통계품질진단 결과	5
제 1 절 통계작성절차별 진단결과	5
1. 통계작성기획 진단결과	5
2. 통계설계 진단결과	8
3. 자료수집 진단결과	11
4. 통계처리 및 분석 진단결과	14
5. 통계공표, 관리 및 이용자서비스 진단결과	17
6. 통계기반 및 개선 진단결과	21
제 2 절 품질차원별 진단결과	23
1. 관련성	23
2. 정확성	24
3. 시의성/정시성	24
4. 비교성/일관성	25
5. 접근성/명확성	25
제 3 절 진단결과 종합표	26

제 3 장 개선과제별 개선방안	27
제 1 절 조사표 개선 검토	28
1. 현황 및 문제점	28
2. 세부 개선과제 내용	28
제 2 절 현장조사 파라미터 수집·관리	29
1. 현황 및 문제점	29
2. 세부 개선과제 내용	29
제 3 절 조사 참여율 제고 방안 수립·시행	30
1. 현황 및 문제점	30
2. 세부 개선과제 내용	30
제 4 절 마이크로데이터 서비스 제공	31
1. 현황 및 문제점	31
2. 세부 개선과제 내용	31
제 5 절 조사 관리용 문서 세분화	32
1. 현황 및 문제점	32
2. 세부 개선과제 내용	32
제 6 절 통계 생산·관리체계 단순화	33
1. 현황 및 문제점	33
2. 세부 개선과제 내용	33
제 7 절 개선과제 요약	34

제 4 장 발전전략 및 중장기 로드맵	35
제 1 절 인공지능산업 고유 분류체계 정립	35
제 2 절 표본조사로 전환	36
제 3 절 장기발전 제안사항 요약	37
붙임1) 자료수집 체계 점검 결과	39
붙임2) 표본설계 점검 결과	49
붙임3) 조사표 설계 및 유사통계 비교·분석 점검 결과	63
붙임4) 이용자 요구사항 반영실태 진단 결과	73
붙임5) 마이크로데이터 품질 점검 결과	81
붙임6) 공표자료 오류 점검 결과	87
부 록. 통계품질진단 개요	91
1. 통계품질진단의 개념	91
2. 통계품질진단 체계	92
3. 통계품질 수준 측정	97

표 목 차

<표 1> 인공지능산업실태조사(2020 기준) 개요	3
<표 2> 통계작성기획 진단결과	7
<표 3> 통계설계 진단결과	9
<표 4> 자료수집 진단결과	12
<표 5> 통계처리 및 분석 진단결과	15
<표 6> 통계공표, 관리 및 이용자서비스 진단결과	18
<표 7> 통계기반 및 개선 진단결과	22
<표 8> 진단결과 종합표	26
<표 9> 개선과제 요약	34
<표 10> 장기발전 제안사항 요약	37

그 립 목 차

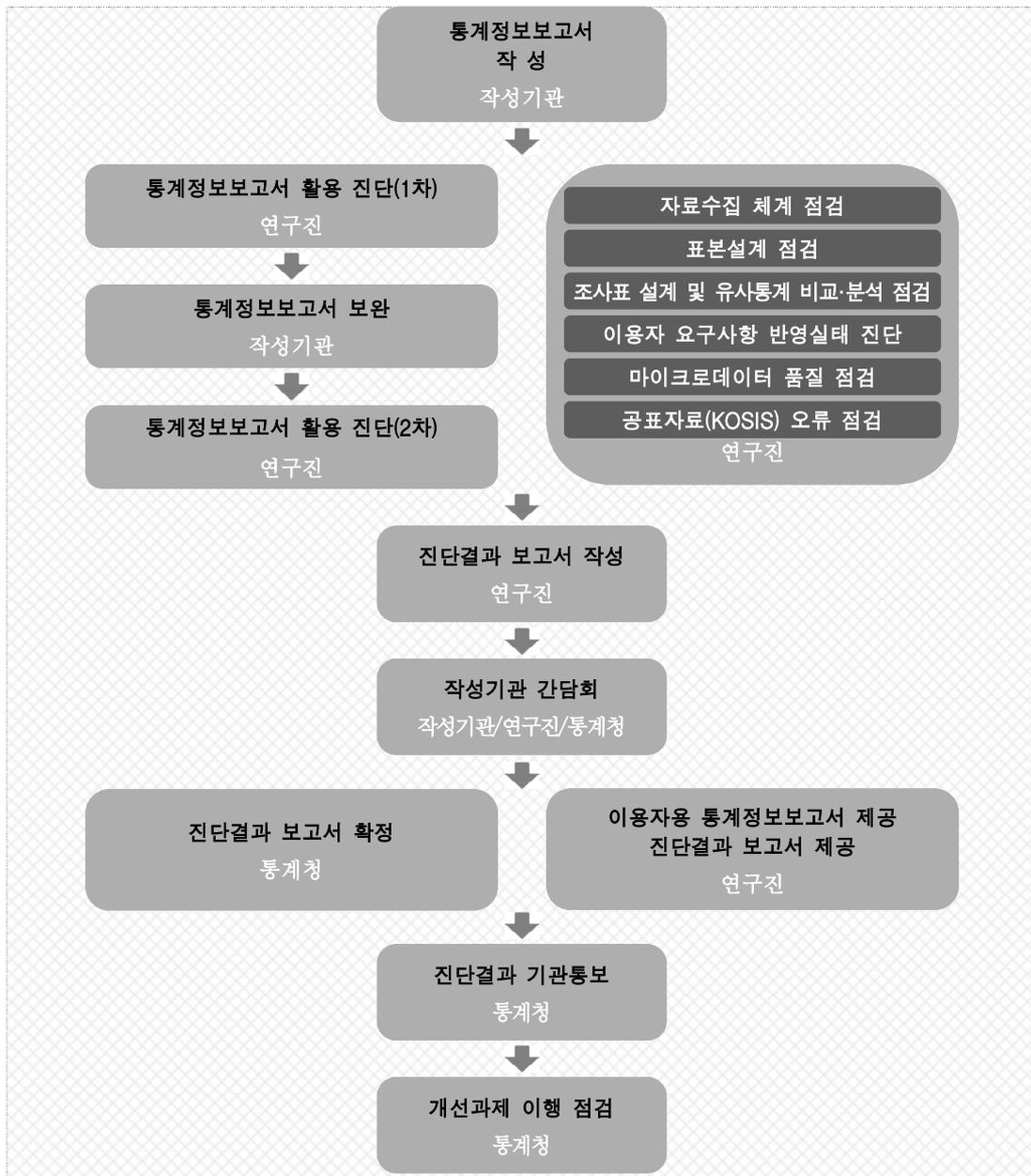
<그림 1> 통계품질진단 흐름도	2
<그림 2> 『인공지능산업실태조사』 품질차원별 진단점수	23

결과보고서 요약문

진단통계명	「인공지능산업실태조사」 (과학기술정보통신부)
주 제 어	인공지능산업(매출, 인력현황, 기술분야, 주력사업, 투자액)
진 단 기 간	2022.02. ~ 2022.12.
진 단 기 관	통계청, 한국통계진흥원
연 구 진	황규배, 오유진, 조준기
<p>이번 진단에서 활용한 통계는 2022.04.에 공표된 2021년(2020 기준) 인공지능 산업실태조사이다. 본 통계의 조사 기간은 2021년 9월~11월이다.</p> <p>본 진단은 인공지능산업실태조사의 전반적인 품질 상태를 살펴보고, 본 조사를 통해 제공되는 국가통계에 대한 신뢰성을 제고할 수 있는 방안을 제시하기 위해 수행되었다. 통계품질진단은 통계작성기관에서 작성한 통계정보보고서를 기반으로 한 통계작성절차별 작성실태 점검, 자료수집 체계 점검, 표본설계 점검, 조사표 설계 및 유사통계 비교·분석 점검, 이용자 요구사항 반영실태 진단, 마이크로데이터 품질 점검 및 공표자료 오류 점검을 근거로 종합적인 평가를 진행하였다.</p> <p>인공지능산업실태조사에 대한 통계작성절차별 진단결과, 통계작성기획 5.0점, 통계설계 5.0점, 자료수집 4.3점, 통계처리 및 분석 5.0점, 통계공표, 관리 및 이용자 서비스 4.2점, 통계기반 및 개선 5.0점으로 평가되었다. 자료수집, 통계공표, 관리 및 이용자 서비스에 대한 평가가 상대적으로 낮았다. 자료수집의 경우 조사 진행 관련 파라미터 수집·관리가 이루어지지 않았으며, 조사항목별 작성요령 및 유의사항이 일부 미흡하였으며, 주요 질의응답 오류사례 이에 대한 관리가 잘 이루어지지 않았다. 통계공표, 관리 및 이용자 서비스는 마이크로데이터 서비스가 제공되지 않고 있으며, 통계공표 시점이 사전에 공지되고 있지 않아, 감점 요인으로 작용하였다.</p> <p>품질차원별 진단결과는 관련성 4.9점, 정확성 4.7점, 시의성/정시성 4.0점, 비교성/일관성 5.0점, 접근성/명확성에서는 3.3점으로 진단되었다. 특히 접근성/명확성 차원에서의 진단결과가 낮았는데 이는 마이크로데이터가 제공되고 있지 않은 것이 주요 원인이었다.</p> <p>그리고 자료수집 체계 점검에서는 조사 참여율 제고, 조사관리용 문서 세분화, 표본조사로 전환 시점 검토를, 이용자 요구사항 반영실태 진단 결과 인공지능 산업 고유의 분류체계 정립, 마이크로데이터 서비스 제공을 주요 개선과제로 도출하였다.</p> <p>이를 토대로 단기과제로 조사표 개선 검토, 마이크로데이터 서비스 제공, 조사 관리용 문서 세분화, 통계 생산·관리체계 단순화, 조사 참여율 제고 방안 수립·실행을 도출하였으며, 현장조사 파라미터 수집·관리를 중기 개선과제로 선정하였다. 장기발전 제안사항으로는 ①인공지능산업 고유 분류체계 정립, ②표본조사로 전환을 제안한다.</p>	

정기통계품질진단 흐름도

정기통계품질진단은 하단의 진단절차에 따라 진행되며, 본 보고서는 진단 결과를 종합정리한 진단결과 보고서이다. 통계품질진단의 개념 및 체계, 수준 측정에 대한 자세한 설명은 보고서 마지막 부분의 부록을 통해 확인할 수 있다.



<그림 1> 통계품질진단 흐름도

제 1 장 진단대상통계 개요

<표 1> 인공지능산업실태조사(2020 기준) 개요

기본 정보	작성유형	• 조사통계
	통계종류	• 일반통계
	승인번호	• 127016
	승인일자	• 2019년 11월 4일
	법적근거	• 통계법 제 18조 4항
	조사목적	• 인공지능산업의 현황 파악과 인공지능 산업 생태계 조성 및 활성화를 위한 정책 수립의 기초 자료로 활용
일반 특 성	주요연혁	<ul style="list-style-type: none"> • 2017~2018년 : 시범조사 수행 및 인공지능산업 분류체계 작성 • 2019년 : 통계작성 승인 • 2020년 : 조사항목의 일반현황, 기술 및 사업 현황, 매출 현황, 인력 현황, 투자 및 특허 현황, 애로사항 및 건의사항 등은 유지하되 일부 문항 현행화, GPU 규모 및 수요, 컴퓨팅 인프라 비중 등 신규 문항 추가 등 조사항목 변경 • 2021년 : 코로나 19로 인한 비대면화 현황, 기술 및 사업 현황, 인력 현황, 투자 및 개발 현황 신규 문항 추가 등 조사항목 변경, 조사단위 변경(사업체→기업체)
	조사주기	• 1년
	조사대상 범위	• 인공지능산업 관련 사업을 영위하는 기업체
	조사대상 지역	• 전국
	조사항목	• 7개 부문 44항목 : 비 대면화 현황(6개 항목), 일반현황(1), 기술·사업 현황(14), 매출 현황(4), 인력 현황(9), 투자·개발 현황(8), 애로사항·건의 사항(2)
	자료수집방법	• 방문 유치조사(인터넷 웹 조사 병행)
	조사체계(위탁,용역포함)	• 조사대상 → 조사회사 → 조사관리기관(지능정보산업협회) → 수행기관(소프트웨어정책연구소) → 과학기술정보통신부(작성기관)
조사대상기간/ 조사 기준시점	<ul style="list-style-type: none"> • 조사대상 기간: 2020년 1월 1일 ~ 12월 31일 • 조사 기준시점: 2020년 12월 31일 	

	조사실시기간	<ul style="list-style-type: none"> 2021년 9월 ~ 11월
결과 공표	공표주기	<ul style="list-style-type: none"> 1년
	공표시기	<ul style="list-style-type: none"> 조사 기준년도 익익년 4월(전년 대비 2개월 당겨짐)
	공표범위	<ul style="list-style-type: none"> 전국
	공표방법	<ul style="list-style-type: none"> 간행물(인공지능산업실태조사 보고서), 국가통계포털(KOSIS), 소프트웨어정책연구소 홈페이지, 지능정보산업협회 홈페이지 등
조사 통계 특성	전수/표본구분	<ul style="list-style-type: none"> 전수조사
	모집단	<ul style="list-style-type: none"> 인공지능 기술개발 및 인공지능 적용 제품/서비스/플랫폼의 생산, 유통, 활용, 부가서비스(조사/분석, 컨설팅, 중개) 과정에서 가치를 창출하는 기업체
	표본추출틀	<ul style="list-style-type: none"> 인공지능산업 분류체계 및 ICT 통합모집단을 기반으로 모집단 정비조사
	추출단위	<ul style="list-style-type: none"> 기업체
	조사대상 규모	<ul style="list-style-type: none"> 1,365개 기업
통계 활용	마이크로데이터 보유	<ul style="list-style-type: none"> 보유
	마이크로데이터 제공	<ul style="list-style-type: none"> 미제공 - 제공 검토 중에 있음
	행정자료 활용 여부	<ul style="list-style-type: none"> 해당 없음
	KOSIS 제공 여부	<ul style="list-style-type: none"> 제공
	국제기구제출 여부	<ul style="list-style-type: none"> 미제출
	자료 이용시 주의사항	<ul style="list-style-type: none"> 2021년부터 본 조사의 조사단위를 사업체에서 기업체로 변경함에 따라 2021년 이전 통계와의 시계열 비교가 불가능한 점 유의 2021년 인공지능산업 실태조사의 기준시점은 2020년이나, 시의성 있는 통계 자료 제공을 위해 인력 현황과 GPU 보유 현황은 조사 당시 시점(2021년 6월)을 기준으로 조사한 결과이므로 자료 활용 시 주의 필요 본 보고서에 수록된 2020년 매출 규모는 확정치인 반면, 2021년 매출 규모는 잠정치로 2022년 조사 실시 후에 보완될 수 있음 2021년 인공지능산업 실태조사부터 매출 및 인력 규모 등 모든 결과는 가중값을 적용한 추정치를 사용 본 보고서의 통계표에 사용된 부호는 다음과 같음 : 조사값 중 『-』 은 해당 없음, 『0』 은 해당 영역의 값이 0임을 의미 통계표 및 도표내의 숫자는 반올림되었으므로 세부 항목의 합이 전체 합계와 일치하지 않을 수 있음

제 2 장 통계품질진단 결과

제 1 절 통계작성절차별 진단결과

1. 통계작성기획 진단결과

인공지능산업실태조사는 통계이용자가 해당 통계에 대해 잘 이해할 수 있도록 하고 있는 것으로 진단되었다.

통계작성을 위한 법적 근거 및 조사의 개발 배경, 통계의 변경 및 개편 사항을 상세하게 안내함으로써 통계이용자들이 본 통계에 대한 개발 배경과 변경 내역, 본 통계에 대한 사전 검토내용 등을 쉽게 이해할 수 있도록 하고 있다. 특히 본 조사를 최초 기획할 때 국내외 유사 조사에 대한 사전 검토를 충분히 하였고 2년간 사전 조사를 진행한 후 본 조사를 진행함으로써 조사의 사전 기획단계를 충실히 진행한 것으로 진단되었다.

또한, 조사 및 공표주기, 조사방법, 통계작성절차 등을 상세하게 제시함으로써 이용자들이 통계가 어떤 절차와 방법으로 작성되었는지 이해할 수 있게 하고 있다.

그리고 해당 통계는 이용자들을 위해 통계 작성목적에 이해하기 쉽게 안내하고 있으며 조사통계위원회를 운영하여 주요 이용자들의 의견을 수렴하고 반영하는 노력도 여러 차례 진행한 것으로 확인되었다. 통계의 작성목적과 해당 통계가 주로 활용되는 분야, 주요 이용자 유형별 용도들도 세부적으로 제시함으로써, 해당 통계가 목적에 부합하게 분야별로 활용되고 있음을 확인할 수 있었다.

통계작성문서화면에서는 통계 작성기본계획서는 관리되고 있었지만 업무 편람이 구비되어있지 않아 이에 대한 개선이 필요한 것으로 진단되었다.

□ 시사점

이용자 요구사항 반영실태 진단 결과, 본 통계는 인공지능산업에 대한 세부적이며 구체적인 자료를 제공하는 국내 유일한 조사로 통계이용자들에게 매우 신뢰성이 높고 중요한 통계로 인식되고 있었다.

또한, 본 조사를 최초 기획할 때 국내외 유사 조사에 대한 사전 검토를 충분히 진행한 것으로 확인되어 이를 반영하여 정성평가 상 가점을 부여하였다.

<표 2> 통계작성기획 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1. 통계명 ~ 7. 통계작성 문서화 (관련성)		5/5
통계명	제외	
통계작성기관/부서명	제외	
법적근거	1/1	
조사방법	1/1	
조사 및 공표주기	1/1	
조사일정 및 일정별 수행업무 제시	3/3	
통계작성 기본계획서 첨부 업무편람(직무편람) 첨부	1/1 0/1	
8. 통계연혁 (관련성)		5/5
작성통계의 최초개발 시기	2/2	
작성통계의 개발 배경	2/2	
통계의 개념 분류, 설계, 과정, 내용, 방법, 표본, 기준년, 가중치 등의 변경 또는 개편이력 관리	3/3	
9. 통계의 작성목적 (관련성)		5/5
통계작성 목적의 명확성	1/1	
주된 활용분야에 대한 명시 국내 또는 해외 관련 통계, 유사 사례 사전 검토	3/3 2/2	
10. 주요 이용자 및 용도 ~ 11. 이용자 의견수렴 (관련성)		5/5
주요 이용자 관리	1/1	
주요 이용자 유형별 용도 파악	2/2	
최근 이용자 또는 전문가 대상 의견수렴 실시 내용과 주요 결과 기록	2/2	
최근 이용자 또는 전문가 대상 의견수렴 요구사항 및 요구 반영 결과	3/3	
정성평가		0.2

※ 5점척도점수는 진단 지표에 대한 항목 점수

※ '해당없음'이 포함된 경우 5점척도점수의 구간기준이 변동될 수 있음

* 1.통계명~7.통계작성문서화: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)

* 8.통계연혁: 6점 이상(5), 5점(4), 3~4점(3), 2점(2), 1점 이하(1)

* 9.통계의작성목적: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)

* 10.주요이용자및용도~11.이용자의견수렴: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)

* 정성평가: -0.5점 ~ +0.5점

2. 통계설계 진단결과

인공지능산업실태조사는 조사항목 및 분류체계, 조사표 설계 및 조사표 변경, 조사모집단 구성 등 조사설계와 관련된 관리가 대체로 잘 이루어지고 있는 것으로 진단되었다. 그리고 조사항목과 조사항목체계, 통계에서 사용하는 분류체계에 대한 설명이 자세하게 제시되어 있어 통계이용자들이 조사항목과 통계의 세부내용을 잘 이해할 수 있도록 하고 있다. 한편 조사표 구성을 위한 내·외부 전문가 회의(조사통계위원회) 결과를 기반으로 조사표를 개선하고자 하는 노력을 진행하고 있음을 확인할 수 있었다.

그리고 목표 모집단과 조사 모집단을 명확하게 규정하고 있었고, 두 집단 간의 차이를 잘 기술하고 있는 것으로 진단되었다.

□ 시사점

이용자 요구사항 반영실태 진단 결과, 통계이용자들은 본 통계가 인공지능산업 분야를 전통적인 산업분류체계(KSIC)를 기초로 세분화하고 있으므로 인하여 인공지능산업 세부현황을 전체적으로 반영하는데 부분적으로는 한계가 있는 것으로 인식하고 있었다. 앞으로 인공지능산업은 지속 발전할 것이므로 이를 충분히 반영할 수 있는 분류체계의 정립이 필요할 것으로 사료된다. 이러한 점을 반영하여 정성평가에서 감점을 부여하였다.

<표 3> 통계설계 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1-1. 조사 항목 ~ 1-2. 적용 분류체계 (비교성)		
주요 용어 및 항목별 명확한 정의의 적절성	2/2	5/5
주요 용어의 정의나 개념 등에 대한 국내 또는 국제기준 비교	2/2	
조사표 첨부	1/1	
조사항목의 체계	2/2	
통계에서 사용하는 분류체계 개요 및 내용의 적절성	1/2	
국내 또는 국제기준의 표준분류체계 사용 여부 또는 미사용 사유	2/2	
1-3. 조사표 구성 (정확성)		
내·외부 전문가 회의 개최	1/1	5/5
내·외부 전문가 회의 결과 반영 여부	3/3	
첨부된 조사표에 수록된 사항의 수	4/5	
1-4. 조사표 설계 및 변경 절차 ~ 1-5. 조사표 변경이력 (관련성)		
조사표 설계, 변경 절차나 방법의 적절성	3/3	5/5
조사표 변경 이력 관리	2/2	
조사표 변경 이유 기록·관리	1/1	
변경승인일자 기록·관리	2/2	
2-1. 목표모집단과 조사모집단 (정확성)		
목표모집단 정의	2/2	5/5
조사모집단 정의	2/2	
목표모집단 및 조사모집단 차이의 적절성	2/2	
2-2. 표본추출틀(표본조사) (정확성)		
표본추출틀로 사용되는 자료의 출처	해당없음	해당없음
표본추출틀로 선정한 이유	해당없음	
표본추출틀의 구축(갱신) 과정, 내용, 주기 등 제시	해당없음	
모집단 변동에 따른 표본추출틀 주기적 개편 시 개편의 주기, 필요성, 방법 및 절차, 결과 등 제시	해당없음	
3-1. 표본설계 방법 및 결과 ~ 3-2. 표본관리 (정확성)		
표본추출방법의 적절성	해당없음	해당없음
표본크기 결정의 타당성	해당없음	
표본추출 결과의 타당성	해당없음	
표본설계보고서 첨부	해당없음	

필수진단항목 (품질차원)		진단결과	
		진단점수/ 배점점수	5점척도점수
	표본설계보고서에 모수 및 분산 추정방법	해당없음	
	조사대상의 생멸, 전입, 전출 등 표본 내 변동이 발생한 경우, 수정·보완하는 방법	해당없음	
추가진단항목		추가점수 (진단점수/배점점수)	
1-1. 주요 항목의 조사목적		0/0.1	
1-1. 부정확한 응답 가능성이 있는 조사항목 검토		0/0.1	
1-3. 조사방법을 혼합하여 이용하는 경우 조사방법별로 조사표의 구성 내용, 특징 및 설계 시 고려한 다양한 요소 검토		0/0.1	
1-6. 응답자 유형별 응답 소요시간 등 검토		0/0.1	
2-1. 조사모집단의 과대포함, 과소포함 등 포함오차에 대한 분석 또는 검토		0/0.1	
2-2. 분류별, 지역별 기타 하위모집단별 추출단위 분포, 관련 통계량, 상관관계 등 기록 및 관리		0/0.1	
2-2. 표본틀에 한계가 있는 경우 그 내용과 보완 등의 검토 또는 조치 결과		0/0.1	
3-2. 동일 대상을 연속 조사하는 경우 평소 조사 대상자 관리방법		0/0.1	
정성평가		-1	

- * 1-1.조사항목~1-2.적용분류체계: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)
- * 1-3.조사표구성: 8점 이상(5), 6~7점(4), 4~5점(3), 2~3점(2), 1점 이하(1)
- * 1-4.조사표설계및변경절차~1-5.조사표변경이력: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 2-1.목표모집단과조사모집단: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 2-2.표본추출틀(표본조사): 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 3-1.표본설계방법및결과~3-2.표본관리: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)
- * 정성평가: -1점~+1점

3. 자료수집 진단결과

인공지능산업실태조사의 전체 조사과정은 몇몇 미흡한 점이 있지만 대체로 잘 관리되고 있는 것으로 진단되었다. 조사방법은 충분한 검토를 통해 적절하게 선택되었으며, 조사원 채용 및 교육, 조사 사전홍보, 조사지침서 준비, 현장조사 관리체계 구축 및 운영, 단위 무응답 대처 등 조사 진행 과정 전반이 대체로 잘 관리되고 있는 것으로 사료된다.

하지만 조사진행 관련 파라미터가 수집·관리가 이루어지지 않았으며, 조사항목별 작성요령 및 유의사항이 일부 미흡하며, 주요 질의응답 오류사례 이에 대한 관리가 잘 이루어지지 않아 이에 대한 개선이 필요한 것으로 진단되었다.

□ 시사점

자료수집 체계 점검 결과, 본 통계는 전수조사이지만 조사 참여율이 2021년 조사기준 50%(조사참여율 2019년 64%, 2020년 57%, 2021년 50%)로 낮은 수준이었다. 조사대상 중 가장 중요한 집단인 종업원 1,000인 이상 대기업의 조사 참여율이 54.2%로 대상 집단의 중요성을 감안할 때 조사 참여율이 낮은 편이며, 세부적으로는 1,000인 이상 종업원이 근무하는 인공지능서비스 기업에서 조사 참여율이 36.4%로 아주 낮아 편이었다.

조사대상 기업이 매년 늘어나는 추세를 감안 할 때 조사 참여율을 높이기 위한 추가적 노력이 없다면 조사 참여율은 앞으로 더 낮아질 가능성이 높을 것이다. 조사대상 기업의 조사 참여율을 높이기 위한 추가적인 노력이 필요할 것으로 사료된다. 또한 조사관리용 문서는 조사지침서에 조사 진행 관리에 필요한 전반적인 내용이 수록되어 있었지만 조사과정별로 보다 세분화하여 조사관리용 문서를 작성할 필요가 있을 것으로 사료된다. 세분화가 필요한 문서로는 조사 거부 기업별 상세 거부 원인 분석 문서, 에디팅 지침서, 리체크 및 검증 결과 내용 리스트, 조사원 평가 결과 문서 등이다.

조사표 설계 및 유사통계 비교·분석 점검 결과, 조사표 안내문 수록사항 보완 및 2개 조사표 문항에 대한 작성요령 구체화, 1개 문항에 대한 기준시점 명기 등이 필요한 것으로 진단되었다. 이러한 점을 반영하여 정성평가에서 감점을 부여하였다.

<표 4> 자료수집 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1. 조사방법 (정확성)		5/5
조사방법 선택에 대한 검토(조사비용, 조사인력, 조사기간, 조사체계 등)	2/2	
선택한 조사방법에 대한 조사과정의 적절성	3/3	
2-1. 조사원 채용 및 처우 ~ 2-3. 조사원 업무량 (정확성)		5/5
조사원 채용 방법 및 과정의 적절성	2/2	
조사원 자격요건, 지위, 급여수준, 지급방법, 부가혜택 등의 적절성	2/2	
조사원 교육훈련에 대한 일정	2/2	
조사원 교육훈련 내용의 적절성	2/2	
교육시간의 적정성 검토	1/1	
교육훈련 교재 첨부	1/1	
조사기간 중 교체된 조사원에 대한 교육 실시	해당없음	
조사원 대상 비밀보호 의무 교육 또는 서약서 작성	1/1	
업무량 배정 시 응답소요시간, 조사난이도, 평균 접촉시도 또는 방문횟수, 조사기간 등 고려사항	2/2	
3-1. 조사업무 흐름도 ~ 3-2. 조사준비 및 준비조사 (정확성)		5/5
조사실시에 대한 조사업무 흐름도 관리의 적절성	2/2	
조사 홍보 실시 내용과 방법	1/1	
응답자(조사대상) 사전 통지	1/1	
조사구 확인 또는 조사명부 보완	2/2	
3-3. 조사항목별 조사 방법 (정확성)		3/5
주요 조사항목별 작성요령 및 유의사항의 적절성	1/3	
조사표 기입에 필요한 조사지침서 첨부	1/1	
3-4. 현장조사 관리 (정확성)		5/5
현장조사 관리 체계	1/1	
현장조사 관리 방법	2/2	
현장조사 관리자 1인당 조사원수 등 관리	1/1	
현장조사 관리자 역할의 적절성	2/2	
현장조사 파라미터 기록·관리 여부	0/1	
조사기간 중 작성기관이 조사위탁기관이나 조사원을 대상으로 실시지도(지도점검) 실시	1/1	
3-5. 조사 질의응답 체계 (정확성)		2/5
현장조사 질의 및 응답체계 운영 방법의 적절성	1/3	

필수진단항목 (품질차원)		진단결과	
		진단점수/ 배점점수	5점척도점수
주요 질의 응답·오류사례 추적 및 관리		0/2	
	현장조사 사례집 첨부	1/1	
4-1. 응답자, 4-3. 무응답 대처 ~ 4-4. 표본대체 (정확성)			5/5
적격 응답자의 지위, 지정 이유의 타당성	2/2		
항목 무응답 대처 방법	2/2		
단위 무응답 대처 방법	2/2		
표본대체 허용 기준	해당없음		
표본대체 절차 및 방법	해당없음		
표본대체 기준, 절차 및 방법의 적절성	해당없음		
5. 사후조사 (정확성)			해당없음
조사 실시 후 사후조사(모니터링) 실시	해당없음		
사후조사(모니터링) 수행 결과 분석 및 사후 조치 방안	해당없음		
6. 행정자료 활용 목적 및 내용 ~ 7. 활용 행정자료 특성 및 입수체계(관련성)			해당없음
행정자료 활용에 대한 목적, 필요성, 활용 정도 파악	해당없음		
행정자료 이용 시 발생하는 이용제한 사항 및 사유 파악	해당없음		
활용하는 행정자료의 내용 및 항목 파악	해당없음		
활용하는 행정자료의 원래 수집 목적에 대한 파악 (관리/제공기관 기준)	해당없음		
활용하는 행정자료의 원래 수집과정 및 내용, 관리기관에 대한 파악(관리/제공기관 기준)	해당없음		
행정자료 입수 방법 및 경로의 기록·관리(통계작성기관 기준)	해당없음		
행정자료 입수주기 또는 갱신주기 및 정시성에 대한 기록· 관리(통계작성기관 기준)	해당없음		
행정자료 활용 법적근거(통계작성기관 기준)	해당없음		
추가진단항목		추가점수 (진단점수/배점점수)	
1. 조사의 효율성, 정확성 등의 제고를 위하여 조사 방법별 응답비율, 응답자 특성, 추정치에 미치는 영향 등 분석·검토		0/0.1	
2-1. 우수 조사원을 채용하기 위하여 적용한 방법이나 조치		0.1/0.1	
2-2. 조사원의 업무지식 숙지 정도에 대한 평가 및 평가 조치(재교육 실시 등)		0/0.1	
4-2. 기억응답과 관련된 검토 여부(조사대상 기간(또는 시점)과 조 사시기 사이의 간격, 응답에 필요한 기록물(영수증, 장부 등) 활용가능성 등)		0.1/0.1	
정성평가		-1	

- * 1.조사방법: 5점(5), 4점(4), 2~3점(3), 1점(2), 0점(1)
- * 2-1.조사원채용및처우~2-3.조사원업무량: 14점 이상(5), 11~13점(4), 5~10점(3), 2~4점(2), 1점 이하(1)
- * 3-1.조사업무흐름도~3-2.조사준비및준비조사: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 3-3.조사항목별조사방법: 4점(5), 3점(4), 2점(3), 1점(2), 0점(1)
- * 3-4.현장조사관리: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 3-5.조사질의응답체계: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 4-1.응답자, 4-3.무응답대처~4-4.표본대체: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)
- * 5.사후조사: 4점(5), 3점(4), 2점(3), 1점(2), 0점(1)
- * 6.행정자료활용목적및내용~7.활용행정자료특성및입수체계: 12점 이상(5), 9~11점(4), 5~8점(3), 2~4(2), 1점 이하(1)
- * 정성평가: -1점 ~+1점

4. 통계처리 및 분석 진단결과

인공지능산업실태조사 통계처리 및 분석 진단결과 주요 과정들이 잘 관리되고 있는 것으로 진단되었다.

자료입력을 위한 코드체계 및 입력작업은 잘 관리된 것으로 진단되었다. 자료 내검은 조사 완료된 조사표에 대한 조사원에 의한 1차 내검(육안 검증), 검증원에 의한 진위 검증 등을 거쳐 최종적으로 입력될 자료가 확정되었다. 자료입력은 온라인조사 입력 방식과 유사한 프로그램화된 방식으로 이루어져 입력과 동시에 입력자료에 대한 범위 내검 및 논리 내검이 이루어졌다. 그리고 데이터 분석 단계에서 전산팀에 의해 최종적으로 자료 내검이 이루어진 것으로 확인되었다.

또한, 해당 통계는 사후 가중치와 통계추정 및 분석 과정에 대한 정보를 제시하고 있으므로 통계의 모수추정 및 분석이 수행되었음을 확인할 수 있었다.

그러나 자료입력이 프로그램화된 방식으로 이루어져 자료입력 매뉴얼이 별도 작성·관리되어 있지 않아 이에 대한 보완이 필요한 것으로 사료된다.

□ 시사점

표본설계 점검 결과, 해당 통계는 전수조사이지만 무응답으로 인한 편향을 보정하기 위해 사후가중치를 제시하고 있으며 주요 가중치 분포, 가중치를 이용한 총계 추정식도 함께 제시하고 있다. 다만, 총계 추정식에 사용되는 가중치 W_{khi} 에 대한 산출과정이 제시되어 있지 않으므로 이를 보완할 필요가 있을 것으로 진단되었다. 이러한 점을 고려하여 정성평가에서 감점을 부여하였다.

<표 5> 통계처리 및 분석 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1. 행정자료의 매칭방법 (정확성)		해당없음
조사통계자료와 행정자료 간 매칭변수	해당없음	
조사통계자료와 행정자료 간 매칭방법	해당없음	
조사통계자료와 행정자료 간 매칭허용 한계 검토	해당없음	
조사통계자료와 행정자료 간 매칭비율 수치 파악	해당없음	
3. 자료코딩 ~ 4. 자료입력 (정확성)		5/5
자료 코드체계 및 코딩(부호화) 방법의 적절성	2/2	
조사결과 자료의 전산입력 방법의 적절성	2/2	
입력 시 오류 검출을 위해 적용한 방법의 적절성	2/2	
입력매뉴얼(지침서) 첨부	0/1	
자료 입력 교육 실시 여부와 교육 일정 및 방법	1/1	
5. 자료내검 (정확성)		5/5
조사현장 내검 내용 및 방법, 오류자료 처리방법의 적절성	2/2	
입력결과 내검 내용 및 방법, 오류자료 처리방법의 적절성	2/2	
전산내검 범위, 논리내검 적용대상 및 적용내용의 타당성	3/3	
내검매뉴얼(지침서) 첨부	1/1	
6-1. 주요 항목무응답 실태 ~ 6-3. 단위무응답 실태 (정확성)		5/5
주요 항목에 대하여 항목 무응답률 수치 제시	1/1	
주요 항목에 대하여 항목 무응답률 산출 산식	1/1	
주요 항목의 항목무응답을 대체하는 경우 대체방법의 적절성	3/3	
단위무응답률 수치 제시	2/2	
단위무응답률 산출 산식	1/1	
주요 하위그룹별 및 무응답 사유별 무응답률 검토	1/1	
7-1. 가중치 조정 ~ 7-2. 통계추정 산식 및 내용(정확성)		5/5
무응답 가중치 조정	1/1	
사후가중치 조정	1/1	
무응답 가중치 구체적인 조정과정 및 방법의 적절성	2/2	
사후가중치 구체적인 조정과정 및 방법의 적절성	1/2	
추정하고자 하는 주요 모수	1/1	
추정치를 계산하는 산식	2/2	
8. 표집오차 추정 방법 및 결과(표본조사) (정확성)		해당없음
주요 항목에 대한 분산, 표준오차 등의 추정 방법	해당없음	
주요 항목에 대한 상대표준오차, 신뢰구간 등의 적절성	해당없음	
주요 항목의 오차 특성과 이용 시 고려사항	해당없음	
9-1. 지수 유형 및 산출산식 ~ 9-2. 지수 가중치 및 갱신 (정확성)		해당없음
사용된 지수의 유형 및 지수의 장단점, 선정 이유의 타당성	해당없음	
사용된 지수의 산출 산식	해당없음	
지수작성 목적으로 조사대상 선정기준, 절차, 선정된 항목	해당없음	
지수작성 가중치 산출에 이용된 자료의 명칭 및 개요	해당없음	
가중치 산출 산식 및 과정, 갱신주기 및 이유	해당없음	
9-3. 지수개편 ~ 9-4. 디스플레이터(정확성)		해당없음

필수진단항목 (품질차원)		진단결과	
		진단점수/ 배점점수	5점척도점수
	지수개편의 주기	해당없음	
	지수개편의 목적 및 필요성, 방법, 절차, 내용의 적절성	해당없음	
	과거자료 접속방법	해당없음	
	디스플레이터의 개요, 특성, 적정성	해당없음	
	디스플레이터의 불변화 방법	해당없음	
10-1. 계절조정 의미 및 적용 방법 ~ 10-3. 계절조정 시계열	보정(비교성)	해당없음	
계절조정의 의미와 필요성, 방법 및 버전	해당없음		
계절조정 과정, 과정별 적용 방법, 내용, 산출물 등 관리	해당없음		
계절조정 시계열 보정의 주기, 이유, 보정의 내용, 방법	해당없음		
추가진단항목		추가점수 (진단점수/배점점수)	
2. 활용하는 행정자료를 점검 또는 보완하는 경우 내용, 방법, 결과 등의 기록·관리		0/0.1	
5. 자료 내용검토(에디팅) 시스템 구축		0.1/0.1	
5. 확인된 오류의 유형, 내용, 원인 등에 대한 분석		0/0.1	
5. 이상치를 처리하는 경우, 이상치의 기준, 식별 및 처리방법, 처리결과 등 기록·관리		0.1/0.1	
6-1. 항목특성별, 응답자 유형별 등 항목무응답 분포와 특징, 편향 발생 및 분산 증가 가능성 등 분석		0/0.1	
6-2. 항목 무응답 대체시 대체비율, 대체값의 추정치 기여도, 대체값의 자료 표기 방법 등 분석		0/0.1	
6-3. 단위무응답에 의한 편향 발생 및 분산 증가 가능성 검토		0/0.1	
6-4. 항목 또는 단위무응답 발생 시, 응답자와 무응답자의 성향으로 인해 발생할 수 있는 편향을 줄이기 위한 조치		0/0.1	
6-4. 측정 또는 처리오차에 대한 추정 또는 연구 사례 유무		0/0.1	
8. 마이크로데이터 이용자가 스스로 표집오차를 계산할 수 있도록 관련 방법을 제공하는 경우 이에 대한 사용방법		0/0.1	
정성평가		-0.1	

- * 1. 행정자료의매칭방법: 6점 이상(5), 5점(4), 3~4점(3), 2점(2), 1점 이하(1)
- * 3. 자료코딩~4. 자료입력: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 5. 자료내검: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 6-1. 주요항목무응답실태~6-3. 단위무응답실태: 8점 이상(5), 6~7점(4), 4~5점(3), 2~3점(2), 1점 이하(1)
- * 7-1. 가중치 조정~7-2. 통계추정산식및내용: 8점 이상(5), 6~7점(4), 4~5점(3), 2~3점(2), 1점 이하(1)
- * 8. 표집오차추정방법및결과: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 9-1. 지수유형및산출산식~9-2. 지수가중치및개산: 12점 이상(5), 9~11점(4), 5~8점(3), 2~4점(2), 1점 이하(1)
- * 9-3. 지수개편~9-4. 디스플레이터: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 10-1. 계절조정의미~10-3. 계절조정시계열보정: 8점 이상(5), 6~7점(4), 4~5점(3), 2~3점(2), 1점 이하(1)
- * 정성평가: -1점~+1점

5. 통계공표, 관리 및 이용자서비스 진단결과

인공지능산업실태조사는 통계공표 및 그와 관련된 생산 및 자료관리가 어느 정도 잘 이루어지고 있었지만, 이용자 서비스면에서는 다소 미흡한 것으로 진단되었다.

구체적으로, 공표되는 통계의 주요 분류수준, 통계 이용 시 주의사항, 주요 통계표 등을 잘 제시하고 있었다. 공표자료 오류 점검 결과 통계표의 표시 형식은 적절하며 통계수치도 정확한 것으로 진단되었다. 마이크로데이터 품질을 점검한 결과 KOSIS 집계표자료와 100% 일치하는 것으로 확인되었다.

그리고 통계조사 결과보고서를 위탁기관 및 수행기관 홈페이지에 게시하고 있으며, 통계청 국가통계포털에 통계 결과를 게시하고 있었다. 또한, 인공지능산업실태조사에 대해 이용자들이 알아야 할 주요 정보들이 KOSIS 국가통계포털에 거의 빠짐없이 잘 수록되어 있었다. 그러나 마이크로데이터가 이용자들에게 공개되고 있지 않았다.

자료의 수집, 처리 및 보관 과정의 비밀보호 및 보안관리는 지침에 의해 잘 이루어진 것으로 확인되었다.

정시성/시의성 면에서 해당 통계는 사전에 계획된 공표 일정에 맞추어 통계공표가 잘 이루어지고 있었지만, 통계공표가 사전에 공지되지 않았다. 그리고 조사 기준시점에서 발표 시점까지 12개월 이상이 소요되어 시의성은 낮은 편이었다.

□ 시사점

이용자 요구사항 진단(FGI) 결과 마이크로데이터가 제공되고 있지 않으므로 이용자의 필요에 맞게 데이터를 분석하는데 제약이 있는 것으로 지적되어 이에 대한 개선이 필요할 것으로 사료된다. 이러한 점을 반영하여 정성평가에서 감점을 부여하였다.

<표 6> 통계공표, 관리 및 이용자서비스 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1-1. 공표통계 해석방법 (관련성)		4/5
주요 분류 수준별 세분화된 공표통계의 적절성	2/2	
통계 공표의 적정성(상대표준오차 등) 검토	1/3	
주요 통계표, 그래프	2/2	
공표되는 통계의 해석방법 및 이용 시 유의사항	2/2	
연도별(시계열) 통계결과 및 분석결과 관리	2/2	
1-2. 공표통계 정확성 (정확성)		5/5
공표된 통계표 형식, 단위표기, 주석 등의 적절성	3/3	
공표된 통계수치의 정확성	3/3	
2-1. 조사대상 기간/조사 기준시점과 공표 시기 (시의성)		4/5
조사대상 기간/조사 기준시점과 통계 공표 시점 제시	1/1	
조사과정별 소요되는 기간의 적절성	2/2	
조사기준 시점과 통계결과의 최초 공표일 간의 차이	3/5	
2-2. 공표일정 (정시성)		4/5
사전에 공개된 통계공표 일정과 공개방법	2/2	
통계공표 일정을 작성기관 홈페이지 등에 예고	0/2	
예고된 통계 공표일정 준수	5/5	
3-1. 통계 작성방법의 비교성 ~ 3-3. 국가간 비교성 (비교성)		5/5
통계의 개념 동일 여부	1/1	
분류체계 동일 여부	1/1	
조사 기준시점 동일 여부	1/1	
조사 실시시기 동일 여부	1/1	
변경된 경우, 변경 전·후 비교분석 결과	해당없음	
시계열 단절이 발생한 경우, 발생 원인과 변경된 자료 이용 시 고려사항 검토	2/2	
작성통계와 동일한 조사목적에 갖는 외국 통계 명칭과 개요	해당없음	
작성통계와 동일한 조사목적에 갖는 외국통계와 직접 비교 가능한지 여부, 가능하지 않은 사유 및 이용 시 고려사항 등에 대한 검토	1/1	
국제 기구에 제공하는 경우, 국제기구명, 제공항목 등 제시	해당없음	
3-4. 동일영역 통계와 일관성 ~3-6. 잠정치와 확정치의 일관성 (일관성)		해당없음
작성통계와 동일하거나 유사한 조사내용 혹은 항목을 포함한 조사의 명칭과 개요	해당없음	
두 통계간 차이 발생 시 차이가 나는 내용, 정도, 이유 등과 이용 시 고려사항에 대한 검토	해당없음	
동일한 내용을 조사하는 작성주기가 다른 통계의 명칭과 개요	해당없음	
두 통계간 차이 발생 시 차이가 나는 내용, 정도, 이유 등과 이용 시 고려사항에 대한 검토	해당없음	

필수진단항목 (품질차원)		진단결과	
		진단점수/ 배점점수	5점척도점수
	작성통계의 잠정치와 확정치의 차이	해당없음	
	두 수치가 차이가 나는 요인 및 이용 시 고려사항 검토	해당없음	
4-1. 통계의 이용자 서비스 (접근성)			4/5
	통계공표 방법의 다양화(브리핑 제공, 보도자료 제공, 보고서 간행물 제공, 홈페이지 제공)	2/3	
	국가통계포털(KOSIS) 수록	2/2	
4-3. 통계설명자료 제공 (명확성)			5/5
	통계 설명자료에 대한 소재 정보	2/2	
	국가통계포털(KOSIS) 통계설명자료에 정보 제공	-	
	통계설명자료 제공(통계개요)	3/3	
	통계설명자료 제공(조사관리)	3/3	
	통계설명자료 제공(표본설계/표본조사, 통계추정 추계 및 분석)	해당없음	
	통계설명자료 제공(지수편제)	해당없음	
	통계설명자료 제공(참고자료) 간행물 또는 작성기관 홈페이지 등에 통계설명자료 제공(KOSIS 설명자료 외)	2/3 3/3	
5-1. 마이크로데이터 생성·관리 (정확성)			5/5
	마이크로데이터 생성 방법	2/2	
	마이크로데이터 관리 방법	2/2	
5-2. 마이크로데이터 서비스 (접근성)			1/5
	마이크로데이터 제공	0/2	
	마이크로데이터 요구 및 제공 방법, 구입 소요시간, 구입비용, 자료제공 포맷, 자료제공 레이아웃, 미제공 항목에 대한 설명 및 제공과 관련된 인터넷 주소 제시	해당없음	
	마이크로데이터 미제공 사유	0/3	
	마이크로데이터 제공/미제공 관련 내부 규정(지침)	0/1	
5-3. 마이크로데이터 일치율 (정확성)			10/10
	마이크로데이터 점검용 자료 제출	5/5	
	마이크로데이터 일치율 점검 결과	5/5	
6-1. 자료 수집, 처리 및 보관 과정의 비밀보호 ~ 6-3 자료 보안 및 접근제한(관련성)			5/5
	자료수집과정에서 응답자 비밀보호 지침/조치	2/2	
	자료처리과정에서 응답자 비밀보호 지침/조치	2/2	
	자료보관과정에서 응답자 비밀보호 지침/조치	2/2	
	공표자료에서 응답자 비밀보호를 위한 조치/방법	2/2	
	마이크로데이터 제공 과정에서 응답자 비밀보호 조치/방법	2/2	
	자료 유실, 유출, 훼손 등 예방을 위한 자료보안 지침/조치	2/2	
추가진단항목		추가점수 (진단점수/배점점수)	

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1-1. 성인지와 관련하여 공표하는 관련 통계 항목 등	0.1/0.1	
2-1. 기간 단축 가능성 검토	0/0.1	
3-3. 주요 통계내용을 국가 간 비교하여 통계표, 그래프 등 제시	0/0.1	
3-6. 잠정치와 확정치 차이를 줄이기 위한 연구 또는 검토	0/0.1	
3-7. 통계 자료 공표 후 오류가 발견되어 수정한 경우, 내용, 사유, 조치과정, 결과 등 기록·관리	0/0.1	
4-1. 통계서비스 경로별 이용자 접속횟수나 마이크로데이터 제공실적 등에 대한 모니터링 및 분석 결과	0/0.1	
5-2 이용자 맞춤형 통계산출 서비스를 제공하는 경우, 요구방법, 소요시간 및 비용 등 명시	0/0.1	
정성평가	-1	

- * 1-1.공표통계및해석방법: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)
- * 1-2.공표통계정확성: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 2-1.조사대상기간/조사기준시점과공표시기: 7점 이상(5), 6점(4), 3~5점(3), 2점(2), 1점 이하(1)
- * 2-2.공표일정: 8점 이상(5), 6~7점(4), 4~5점(3), 2~3점(2), 1점 이하(1)
- * 3-1.통계작성방법의비교성~3-3.국가간비교성: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)
- * 3-4.동일영역통계와일관성~3-6.잠정치와확정치와의일관성 13점 이상(5), 10~12점(4), 5~9점(3), 2~4점(2), 1점 이하(1)
- * 4-1.통계의이용자서비스: 5점(5), 4점(4), 2~3점(3), 1점(2), 0점(1)
- * 4-3.통계설명자료제공: 18점 이상(5), 14~17점(4), 7~13점(3), 3~6점(2), 2점 이하(1)
- * 5-1.마이크로데이터생성·관리: 4점(5), 3점(4), 2점(3), 1점(1), 0점(1)
- * 5-2.마이크로데이터서비스: 5점 이상(5), 4점(4), 3점(3), 2점(2), 1점 이하(1)
- * 5-3.마이크로데이터일치율: 실제 측정점수 반영(0~10점)
- * 6-1.자료수집차별및보관비정의비밀보호 6-3.자료보안및접근제한 11점 이상(5), 8~10점(4), 5~7점(3), 2~4점(2), 1점 이하(1)
- * 정성평가: -1점 ~ +1점

6. 통계기반 및 개선 진단결과

인공지능산업실태조사는 통계위탁기관이 전문조사기관에 조사를 위탁하여 진행하고 있다. 조사 위탁은 통계조사 민간위탁지침을 준수하여 용역업체의 선정 및 진행관리가 이루어진 것으로 확인되었다.

통계위탁기관인 소프트웨어정책연구소는 조사 수탁기관으로부터 조사기획서, 모집단명부, 조사원 교육사항, 조사표, 조사결과 원 자료파일 및 설계 파일, 에디팅 보고서, 현장조사 평가보고서, 최종보고서 등의 자료제출을 포함하여 민간위탁지침에 따라 사업수행을 진행한 것으로 진단되었다.

또한, 자체적으로 통계품질을 높이기 위해 개선과제를 선정하여 이를 이행한 것으로 확인되었다.

□ 시사점

통계위탁기관 담당자 면담결과 통계의 생산 및 관리가 작성기관(과학기술정보통신부) → 통계위탁기관(소프트웨어정책연구소) → 수행기관(지능정보산업협회) → 조사수행기관으로 다단계로 이루어져 통계의 생산·관리의 책임 및 역할을 보다 명확하게 하기 위해서 생산·관리체계를 지금보다 단순화시킬 필요가 있을 것으로 사료된다. 이러한 점을 반영하여 정성평가에서 감점을 부여하였다.

<표 7> 통계기반 및 개선 진단결과

필수진단항목 (품질차원)	진단결과	
	진단점수/ 배점점수	5점척도점수
1. 기획 및 분석 인력 (정확성)		
통계업무 담당 부서명, 업무별 담당인력 구성 및 통계업무 담당년수, 업무 관련 전공 여부 등의 기술	2/2	5/5
외부 위탁 또는 용역사업으로 통계 생산하는 경우, 수탁 기관의 관련 업무 인력구성 및 통계담당년수 등의 적절성	1/1	
최근 1년간 전문성 제고를 위하여 통계 관련 교육과정을 이수한 내역(교육구분, 과정명, 교육기관, 참여인원수)	1/1	
3. 통계위탁 조사 (정확성)		
통계작성을 민간 위탁하여 작성하는 경우, 제안요청서, 제안서, 사업계획서 등 통계조사 민간위탁지침 반영	2/2	5/5
조사 완료 후 수탁기관으로부터 조사와 관련하여 제출받고 있는 자료 목록	-	
조사기획서(사업계획서)	1/1	
(표본조사) 표본설계서 및 예비표본 포함 명부	해당없음	
(전수조사) 모집단 명부 일체	1/1	
조사원 교육관련 사항(지침서, 사례집 등)	1/1	
조사표 원본(또는 폐기 등에 관한 계획)	1/1	
조사결과 원자료(마이크로데이터) 파일, 파일설계서	1/1	
에디팅(내용검토) 요령서	1/1	
현장조사 평가보고서	1/1	
자료처리 보고서	0/1	
최종보고서	1/1	
4. 통계 품질관리 및 개선 (관련성)		
통계품질제고 가능성에 대한 검토 결과나 개선 계획 또는 추진실적에 대한 기록·관리	2/2	5/5
최근 3년간 통계에 대한 학계, 언론, 국회 등 외부 지적 사례 내용, 관련 해명, 개선 등의 조치사항	해당없음	
과거 정기(수시)통계품질진단 결과에 따른 개선과제 관리 및 이행내역(중점관리과제, 기관관리과제 포함)	해당없음	
추가진단항목	추가점수 (진단점수/배점점수)	
2. 전체 및 주요항목, 활동별 사업예산 내역을 산출근거와 함께 제시 또는 예산 증액 필요성, 절감 가능성 등에 대한 분석·검토	0/0.1	
정성평가	-0.5	

* 1.기획및분석인력: 4점(5), 3점(4), 2점(3), 1점(2), 0점(1)

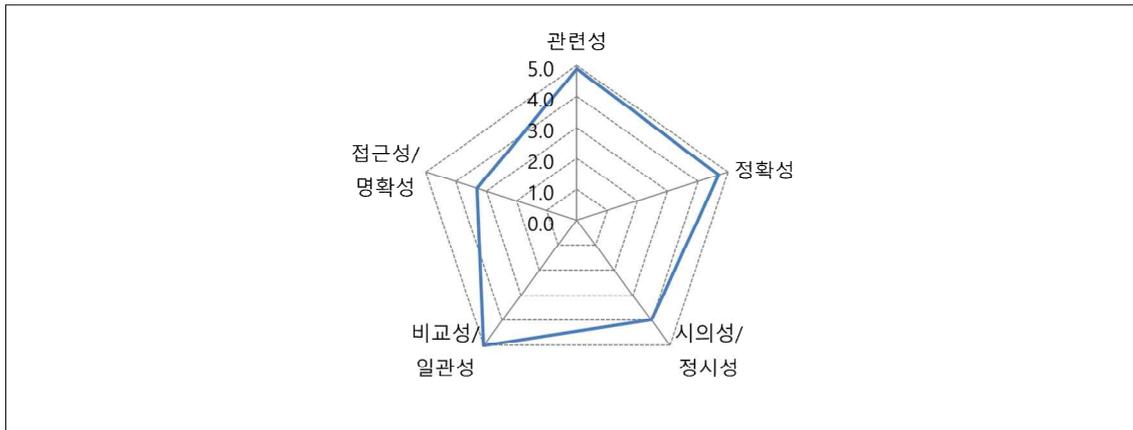
* 3.통계위탁조사: 10점 이상(5), 8~9점(4), 4~7점(3), 2~3점(2), 1점 이하(1)

* 4.통계품질관리및개선: 4점(5), 3점(4), 2점(3), 1점(2), 0점(1)

* 정성평가: -0.5점 ~ +0.5점

제 2 절 품질차원별 진단결과

통계작성절차별 진단을 토대로 인공지능산업실태조사의 품질차원별 점수를 도출한 결과, 관련성 척도 4.9점, 정확성 척도 4.7점, 시의성/정시성 척도 4.0점, 비교성/일관성 척도 5.0점, 접근성/명확성 척도 3.3점으로 진단되었다.



<그림 2> 「인공지능산업실태조사」 품질차원별 진단점수(방사형 그래프)

1. 관련성

통계이용자에게 얼마나 의미 있고 유용한 통계를 작성하여 제공하고 있는가를 진단하는 관련성 차원에서, 인공지능산업실태조사는 5.0점 중 4.9점으로 진단되었다. 이와 관련해 인공지능산업실태조사는 통계의 조사개요, 통계의 작성목적, 통계공표 및 해석방법, 조사표의 설계 및 변경 이력 등에 대한 관리가 잘 이루어지고 있었다. 또한 조사표 개선을 위해 이용자들의 의견을 수렴·반영하는 절차를 거치고 있어서 통계의 유용성을 높이는 노력을 하였음을 확인할 수 있었다. 또한, 공표자료에 대한 비밀보호 및 자료보관에 대한 보안 조치도 적절하게 잘 이루어지고 있는 것으로 확인되었다.

다만 업무편람이 구비되어있지 않아 이에 대한 보완이 필요한 것으로 진단되었다.

2. 정확성

인공지능산업실태조사의 정확성 차원의 품질은 5.0점 중 4.7점으로 진단되었다. 인공지능산업실태조사는 조사표 설계 및 조사 모집단에 대한 구성 등이 적절하게 이루어지고 있고 자료수집과정, 자료 내검, 자료처리 과정 등이 일정 수준에서 잘 관리되고 있는 것으로 진단되었다. 통계추정과 관련된 부분에서도 가중치를 적용하여 적절하게 이루어진 것으로 진단되었다. 공표자료 오류 점검 결과, 표시형식 및 통계수치의 정확성 면에서 문제가 발견되지 않았다. 통계의 위탁도 통계조사 민간위탁지침을 반영하여 이루어진 것으로 확인되었다.

그러나 현장조사 파라데이터 수집·관리가 이루어지지 않았으며, 조사관리를 더 강화하기 위해서 조사관리용 문서를 보다 세분화하여 관리할 필요가 있으며, 조사표도 조사표 안내문, 기준시점 표시 등에서 일부 개선의 여지가 있는 것으로 사료된다.

3. 시의성/정시성

인공지능산업실태조사의 시의성/정시성 차원의 품질은 5.0점 중 4.0점으로 진단되었다.

해당 통계는 사전에 계획된 공표 일정에 맞추어 통계공표가 잘 이루어졌지만, 통계공표 시점이 홈페이지 등에 공개되고 있지 않았으며, 조사 시점에서 발표 시점까지 12개월 이상이 소요되어 시의성은 낮은 편이었다.

4. 비교성/일관성

비교성/일관성 차원의 품질은 5.0점 중 5.0점으로 진단되었다. 인공지능산업실태조사는 주요 개념 및 용어를 통계정보보고서에 구체적으로 명시하고 있으며, 적용분류 체계도 기존의 산업분류방식(KSIC)을 기반으로 이루어지고 있는 것으로 확인되었다. 통계의 작성기준(개념, 분류체계 등)이 매년 동일하게 적용되고 있으며, 조사기준시점, 조사실시 시기도 매년 동일하게 적용되고 있는 것으로 진단되었다.

다만 2021년 공표통계부터 조사대상 단위가 사업체에서 기업체로 변경되어 이전 자료와 시계열분석에는 한계가 있는 것으로 확인되었다.

5. 접근성/명확성

인공지능산업실태조사를 얼마나 쉽게 접근할 수 있는지 진단한 접근성 및 통계가 어떻게 만들어졌는지에 대한 정보를 잘 제공하는지 진단한 명확성 차원의 품질은 5.0점 중 3.3점으로 진단되었다. 인공지능산업실태조사는 작성기관 홈페이지, 보고서, 통계청 국가통계포털 등을 활용해 이용자들이 해당 통계에 접근할 수 있도록 유도하고 있었지만, 보도자료를 통한 홍보는 이루어지지 않고 있었다.

KOSIS 국가통계포털 통계 설명자료에는 해당 통계에 대한 필요한 정보를 충실하게 잘 제공하고 있는 것으로 확인되었다.

마이크로데이터 서비스는 제공을 위해 준비를 하고 있지만 진단 현재일 기준 서비스가 제공되지 않고 있다.

제 3 절 진단결과 종합표

『 인공지능산업실태조사 』 통계정보보고서 등을 기반으로 6개 통계작성절차별 품질지표들을 진단하였고 이를 기반으로 5개 품질 차원별 진단결과도 함께 도출하였다. 최종 진단결과 종합점수는 다음과 같다.

<표 8> 진단결과 종합표

작성 절차 / 품질 차원	1. 통계작성기획	2 통계설계	3 자료수집	4 통계처리 및 분석	5 통계공표 관리 및 이용자서비스	6 통계기반 및 개선	평점 (5점척도)
관련성	5.0	5.0	-		4.5	5.0	4.9
정확성		5.0	4.3	5.0	5.0	5.0	4.7
시의성/ 정시성					4.0		4.0
비교성/ 일관성		5.0		-	5.0		5.0
접근성/ 명확성					3.3		3.3
평점 (5점척도)	5.0	5.0	4.3	5.0	4.2	5.0	4.6
가중치 적용	8.2	16.7	17.6	23.9	22.5	5.8	94.7
추가점수 (정성평가 포함)	0.2	-1.0	-0.8	0.1	-0.9	-0.5	-2.9
총계	8.4	15.7	16.8	24.0	21.6	5.3	91.8

* 평점은 세부진단항목에 대한 평균으로 작성절차별(또는 품질차원별) 평균과는 차이가 있음

제 3 장 개선과제별 개선방안

지금까지 국가통계의 품질 향상 및 신뢰도 제고와 통계이용자 친화적인 통계 생산을 위하여 『인공지능산업실태조사』에 대한 품질진단을 실시하였다. 품질 진단은 관련성, 정확성, 시의성/정시성, 비교성/일관성, 접근성/명확성의 5개 차원에 대해 통계정보보고서 활용 점검, 이용자 요구사항 반영실태(FGI) 진단, 자료수집 체계 점검, 표본설계 점검, 조사표 설계 및 유사통계 비교·분석 점검, 마이크로데이터 품질 점검, 공표자료 오류 점검이라는 7가지 절차를 통해 수행하였다. 제3장에서는 각 진단에서 도출한 개별 개선과제에 대해 개선방안을 제시하고자 한다.

제 1 절 조사표 개선 검토

1. 현황 및 문제점

조사표는 통계품질에 아주 직접적인 영향을 준다. 좋은 조사표는 질문 문항이 명확하고 이해하기 쉽게 기술되어야 하며, 문항 항목 분류체계, 문항 설명문 등이 응답자들이 질문 문항에 응답하는데 어려움을 주지 않게끔 적절하게 기술되어 있어야 한다. 해당 통계에 대한 조사표 설계 및 유사통계 비교·분석 점검 결과, 조사표 안내문에 법적근거 및 문의시 연락처가 없었으며 ‘GPU 보유 현황 및 수요’ 문항에 대한 작성 지침이 명확하지 않은 것으로 진단되었다. 해당 문항은 GPU 모델별 ‘현재 보유 규모’와 ‘향후 희망 규모’를 조사하고 있다. 14번 문항인 ‘향후 희망 규모’는 현재 보유 GPU를 포함하여 기입하는 것인지, 현재 보유 GPU를 제외하고 추가로 필요한 GPU를 기입하는 것인지 명확하지 않았다.

또한 ‘GPU 보유 현황’ (문항 14) 등 문항별 기준시점에 대해 보다 명확한 안내가 필요한 것으로 진단되었다.

2. 세부 개선과제 내용(실행방법 포함)

조사표 관련 다음의 보완 검토 사항들을 작성기관이 검토하여 보완의 필요성이 있는 부분에 대해서는 개선할 것을 제안한다.

2021년 조사표 보완 검토사항

문항	내용
안내문	- 안내문 내 법적근거 및 문의시 연락처 명기
문항 14	- 현재 보유 GPU를 포함하여 응답하는 것인지, 현재 보유 GPU를 제외하고 추가적으로 필요한 GPU를 응답하는 것인지 구체적인 작성요령 명시
기준시점 안내	- 문13 등 문항별 기준시점을 명확히 제시하여 기준시점의 혼동 없이 정확한 응답을 할 수 있도록 설명자료 및 조사표를 보완

제 2 절 현장조사 파라미터 수집·관리

1. 현황 및 문제점

현장조사 파라미터는 조사진행과정에서 현장조사 수행 전반의 상황을 파악하기 위해 조사표 자료 외 부가적으로 수집되는 데이터이다. 통계조사를 표준화하고 품질을 높이기 위해서는 현장조사 파라미터에 근거하여 현장조사 과정이 구체적으로 기록으로 남아있어야 한다. 그리고 이러한 데이터는 다음 회차 조사의 조사 진행관리를 위해 분석, 평가되어야 한다. 이러한 데이터로는 조사원 방문 또는 접촉회수, 조사표 회수기간, 조사원 방문시간대, 조사표 응답 소요시간, 응답거절 또는 접촉실패 사유 등이 대표적인 것들이다.

자료수집 체계 점검 결과 인공지능산업실태조사는 현장조사 파라미터의 수집 및 관리가 미흡하였다. 이에 대한 개선이 필요할 것이다.

2. 세부 개선과제 내용(실행방법 포함)

진단 대상통계는 조사원 방문 유치조사 및 온라인 조사를 위주로 조사가 이루어짐으로 최종 조사표 회수까지 조사대상 기업에 대한 조사원의 지속적인 접촉과 독려가 필요한 조사이다. 그러므로 다음과 같은 현장조사 파라미터의 수집이 필요할 것으로 사료된다.

- ① 조사원(컨택원) 컨택(접촉) 회수
- ② 조사원 컨택 시간대
- ③ 조사표 작성에 참가한 응답자 수
- ④ 조사표 응답소요 시간
- ⑤ 응답거절 사유
- ⑥ 접촉실패 사유
- ⑦ 조사표 회수까지 소요기간

제 3 절 조사 참여율 제고 방안 수립·시행

1. 현황 및 문제점

인공지능산업실태조사는 전수조사로서 일정 수준의 조사 참여율이 확보되어야 제공 통계의 신뢰성이 확보될 수 있다.

하지만 인공지능산업실태조사는 조사 참여율이 2021년 조사기준 50%(2019년 64%, 2020년 57%, 2021년 50%)로 낮은 수준이었다. 조사대상 집단 중 가장 중요한 집단인 종업원 1,000인 이상 대기업의 조사 참여율이 54.2%로 대상 집단의 중요성을 감안할 때 조사 참여율이 낮은 편이며, 세부적으로는 1,000인 이상 종업원이 근무하는 인공지능서비스 기업에서 조사 참여율이 36.4%로 아주 낮은 편이었다. 인공지능산업은 성장 산업이므로 매년 조사대상 기업이 증가할 것이다. 이로 인해 조사 참여율을 높이기 위한 추가적 노력이 없다면 해당 통계의 조사 참여율은 앞으로 매년 낮아질 것으로 예상된다. 적정 수준의 조사 참여율을 확보하기 위해서는 조사대상 기업의 조사 참여를 독려하기 위한 추가적인 노력이 필요할 것으로 사료된다.

2. 세부 개선과제 내용(실행방법 포함)

인공지능산업실태조사의 조사 참여율을 높이기 위해 다음과 같은 방안을 수립하여 시행할 것을 제안한다.

- ① 조사 종료 후 조사 참여기업에 감사 메일 (Thanks Mail) 발송
- ② 조사대상 기업이 관심을 가질 만한 조사결과 내용을 5페이지 내외로 3~4개 유형으로 제작하여 조사 참여기업에 제공, 본 통계의 필요성·유용성 홍보
- ③ 조사진행에 어려움이 많았던 조사대상 사업체를 원인 분석하여 문서화, 차기 조사에 활용
- ④ 조사통계위원회 위원풀 확대 및 위원회를 통한 조사 미참여 기업 관리

제 4 절 마이크로데이터 서비스 제공

1. 현황 및 문제점

이용자 요구사항 진단(FGI) 결과, 본 통계이용자들은 마이크로데이터를 직접 활용하여, 보다 세부적으로 분석하기를 원하고 있었다. 현재 인공지능산업실태조사의 작성기관은 마이크로데이터를 이용자들에게 공개하기 위해 준비 중에 있지만, 진단일 현재까지 이용자들에게 마이크로데이터 서비스를 제공하고 있지 않다. 마이크로데이터를 이용자들에게 조속히 공개하여 본 통계에 대한 이용자 편의성을 높여야 할 것이다. 마이크로데이터 서비스가 제공된다면 이용자들은 정책반영, 연구논문 작성, 현황 파악 등 목적에 따라 다양하게 활용할 수 있게 되어, 통계이용자의 만족도를 높일 수 있다.

2. 세부 개선과제 내용(실행방법 포함)

국가승인통계의 경우 마이크로데이터를 제공하는 방식으로는 통계청의 마이크로데이터 제공 사이트인 MDIS(<http://mdis.kostat.go.kr>)를 활용하는 방식과 본 통계의 작성기관이나 위탁기관의 홈페이지를 활용하는 자체 서비스 방식이 있다. 해당 통계의 접근성을 높이기 위해서는 이들 두 가지 방식을 모두 활용하는 것이 바람직할 것으로 사료된다.

또한 마이크로데이터 제공에 따른 마이크로데이터 비밀보호, 자료제공 포맷, 제공 파일형태, 미제공 항목 여부, 비용부과 여부 등 마이크로데이터 제공 관련 내부 규정도 마련해야 할 것이다.

제 5 절 조사 관리용 문서 세분화

1. 현황 및 문제점

인공지능산업실태조사의 경우 조사관리용 문서로 조사지침서 및 내검지침서를 구비하고 있었지만 조사관리에 필요한 보다 세분화된 문서는 관리되고 있지 않았다. 조사지침서에 조사 진행관리에 필요한 전반적인 내용이 수록되어 있었지만 조사관리를 철저하게 하기 위해서는 조사 과정별로 보다 세분화하여 조사관리용 문서를 작성·관리할 필요가 있을 것으로 사료된다.

2. 세부 개선과제 내용(실행방법 포함)

해당 통계의 조사과정을 보다 철저하게 관리하기 위해서는 ①에디팅 지침서, ②리체크 및 검증결과 문서, ③입력 매뉴얼, 조사원 평가결과 기록문서, ④조사참여 거부 기업별 거부 원인 상세 분석 문서 등의 조사관리용 문서가 작성되어 활용되어야 할 것이다.

제 6 절 통계 생산·관리체계 단순화

1. 현황 및 문제점

통계위탁기관 담당자 면담결과 인공지능산업실태조사 통계의 생산 및 관리는 작성기관(과학기술정보통신부) → 통계위탁기관(소프트웨어정책연구소) → 수행기관(지능정보산업협회) → 조사수행기관으로 4개 단계 구조로 이루어지고 있었다. 통계 생산 및 관리 단계가 여러 단계를 거칠수록 통계 생산 및 관리와 관련된 역할 및 책임 소재 면에서 긍정적인 면보다 부정적인 면이 많을 것으로 사료된다. 통계의 생산 및 관리 구조를 지금보다 단순화할 필요가 있을 것으로 사료된다.

2. 세부 개선과제 내용(실행방법 포함)

인공지능산업실태조사통계의 전문성 및 작성기관의 인력 현황을 고려해 볼 때 해당 통계의 생산·관리 체계는 작성기관(과학기술정보통신부)→ 통계위탁기관(소프트웨어정책연구소) → 조사수행기관으로 단순화시켜 운영시키는 것이 바람직할 것으로 사료된다.

제 7 절 개선과제 요약

지금까지 제시한 개선과제를 요약한 내용은 <표 9>와 같다.

<표 9> 개선과제 요약

단계	개선과제	실행방법	기대효과	관련 품질차원	출처	비고 (예상문제점 등)
단기	조사표 개선 검토	- 일부 항목 수정 및 추가 검토 (조사표 안내문, 문항별 기준시점)	- 조사표 문항 응답 용이성 및 응답 정확성 확보	정확성	조사표설계 및 유사통계 비교·분석 점검, (3. 자료수집)	
	마이크로 데이터 서비스 제공	- MDIS제공 및 마이크로데이터 서비스 지침마련	- 마이크로데이터 이용 및 관련 연구 활성화	유용성	FG, (5. 통계공표 관리 및 이용자서비스)	
	조사 관리용 문서 세분화	- 입력매뉴얼, 에디팅 지침서 등 세분화된 지침서 마련	- 조사과정 관리 강화	정확성	자료수집 체계 점검, (3. 자료수집, 4. 통계처리 및 분석)	
	통계 생산·관리 체계 단순화	- 통계 생산·관리 체계를 작성기관 → 통계 위탁기관 → 조사수행기관으로 단순화	- 조사관리 효율성 제고	정확성	(6. 통계기반 및 개선)	
	조사 참여율 제고 방안 수립·시행	- 조사 종료후 감사 메일 (Thanks Mail), 조사결과 브리프 제공, 조사 진행 어려운 기업 원인 분석 및 문서화	- 조사 대상기업 조사 관심도 및 참여율 제고	정확성	자료수집 체계 점검, (3. 자료수집)	
중기	현장조사 파라 데이터 수집·관리	- 파라데이터 수집 및 관리방법 개선 (조사 응답관련)	- 파라데이터 수집 및 관리 방법 개선 (조사 응답관련)	정확성	자료수집 체계 점검, (3. 자료수집)	

※ 단기 : 1년 이내, 중기 : 2년, 장기 : 2년 이상

제 4 장 발전전략 및 중장기 로드맵

제 1 절 인공지능산업 고유 분류체계 정립

인공지능산업은 성장하는 첨단 산업 분야이므로 향후 어디까지 적용 범위를 확대하여 발전할지 예측하기가 어렵다.

인공지능산업실태조사는 현재까지는 인공지능사업을 기존의 산업분류체계(KSIC)에 기초하여 세분화하고 있다. 이로 인하여 해당 통계는 인공지능산업 세부현황을 전체적으로 반영하기에는 한계가 있는 것으로 통계 이용자들로부터 인식되고 있다. 해당 통계가 기존의 산업분류방식(KSIC)을 기반으로 인공지능산업을 세분화하고 있기 때문에 실제로 인공지능 사업을 영위하고 있지만, 분류기준의 한계로 조사 대상에서 제외되는 기업이 있을 수 있다는 것이다.

본 통계는 이러한 문제를 과학기술정보통신부 AI국 수탁 기업 명부를 활용하여 일부 보완하고 있지만 이러한 점을 완전히 해결하기에는 미흡한 것으로 사료된다. 현재 서비스업, 제조업 등에서도 인공지능사업을 영위하는 일부 기업이 있을 수 있는데 이러한 기업들이 모두 조사 모집단에 포함되고 있지는 못하고 있기 때문이다.

이러한 문제를 해결하기 위해서는 기존의 분류체계를 바탕으로 인공지능산업 전체를 포괄할 수 있는 고유의 분류체계를 검토하여 새롭게 정립할 필요가 있을 것으로 사료된다. 인공지능산업의 고유분류체계는 기업 및 학계 전문가 의견 수렴, 해외 분류사례 등을 참고하여 장기적인 면을 고려하여 정립되어야 할 것이다.

(출처: 차원별-정확성, 작성과정별-2. 통계설계)

제 2 절 표본조사로 전환

인공지능산업실태조사는 해당 산업이 아직 성장 초기 단계라 지난 3년간 전수조사를 진행하였다. 전수대상 기업의 최근 몇 년간 증가 추세를 고려하면 (2019년 424개, 2020년 933개, 2021년 1,365개) 본 통계는 해를 거듭할수록 조사 모집단이 지속적으로 증가할 것으로 예상된다. 조사 참여율이 높지 않은 상태에서는 조사 모집단이 어느 정도 이상 규모가 되면 전수조사를 진행하는 것이 조사의 효율성 및 신뢰성 면에서 부정적인 면이 더 많을 수 있다.

그러므로 조사 모집단 수가 일정 수준(예: 4,000개)이상 되면 표본조사로 조사 방법을 전환하여 조사 효율성 및 신뢰성을 확보하여야 할 것이다.

(출처: 차원별-정확성, 작성과정별-6. 통계기반 및 개선)

제 3 절 장기발전 제안사항 요약

〈표 10〉 장기발전 제안사항 요약

장기 발전을 위한 제안	실행방법	기대효과	예상되는 문제점	비고
인공지능산업 고유 분류체계 정립	- 기업 및 학계 전문가 의견 수렴, 해외 분류사례 참고	- 조사 모집단의 대표성 제고	- 분류체계 정립까지 기간이 많이 소요	
표본조사로 전환	- 조사모집단 일정 수준 (예:4,000개) 도달 시 표본조사로 전환	- 통계 신뢰도 제고	-	

붙임1

자료수집 체계 점검 결과 (조사통계용)

통 계 명	인공지능산업실태조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
면 접 일 시	2022년 5월 11일
연 구 원	황규배
연구보조원	-

제1부 **점검계획**

1. 점검 방법

1) 점검 목적 및 내용

- 점검 목적
 - 조사원, 조사 기획자, 조사 관리자를 대상으로 인터뷰를 진행하여 조사 진행 및 관리 상태를 점검하며, 문제점 및 개선점을 발견하고자 함

- 점검 내용
 - 자료수집방법 확인, 사전 조사준비, 조사홍보, 조사원 교육, 조사 진행관리, 자료 입력 및 검증, 조사표 및 조사대상자 리스트 관리 등 자료수집 체계와 관련된 중요 사항

2) 진행방식

- 사회자의 질문에 묻고 대답하는 인터뷰 방식으로 진행

2. 면담 대상자 및 면담 내용

일시	면담 대상자		장소	주요 점검사항
5.11	이OO 부장	조사기획	메가 리서치 회의실	조사기획 및 조사관리와 관련된 전반적인 사항
	이OO 과장	조사진행 관리 • 감독		조사진행(실사) 및 에디팅, 검증, 자료 입력관련 세부사항
	오OO	조사원		조사원교육, 조사진행시 애로사항, 조사지침서 등 필요 사항 인식 정도 등

제2부 **점검 결과 요약**

점검 자료목록	현황 및 문제점	개선 의견
<p>조사방법 (조사표 회수자료, 조사응답비율, 응답자 특성, 분석결과자료)</p>	<ul style="list-style-type: none"> - 본 조사는 방문 유치조사(55%) 및 온라인 웹조사(45%)로 웹 조사는 조사표 작성시 작성자료는 조사회사 서버에 데이터 형식으로 기록되었고, 방문 유치조사 조사표는 조사원에 회수되어 입력원에 의해 온라인 CAPI 방식으로 입력되었다. - 진단대상 통계는 전수조사로 2021년 조사의 경우 전수조사 대상 기업 1,365개 중 684개 기업체가 조사되어 조사 참여율이 50.1%로 낮은 편이다. - 분석은 설립연도, AI 주사업 여부, 주사업분야, 종사자 규모, 매출액 규모로 구분하여 통계치를 제공하고 있음 	<ul style="list-style-type: none"> - 조사 참여율 제고
<p>조사원 채용 및 처우 (조사원 채용과정 및 계획문서)</p>	<ul style="list-style-type: none"> - 조사원은 15명을 선발하였다. 과거 해당 통계 조사진행 경험자를 우선 선발(10명)하였고 나머지 5명은 사업체조사 경력이 있는 조사원을 선발함 - 완성기준 조사 사업체 진행 난이도에 따라 1건당 난이도에 따라 2만원에서 최대 5만원까지 조사원에게 지급됨(교통비별도 지급) - 조사회사 소속면접원 내에서 조사원을 채용하였으므로 조사원 채용관련 상세 계획 문건은 없음 	<ul style="list-style-type: none"> - 조사원 채용관련 상세 계획문건 관리

<p>조사원 교육훈련 (조사원 교육자료 교육세부일정 및 계획/결과, 보안교육 및 서약서, 조사원 평가 결과, 재교육일정 등)</p>	<ul style="list-style-type: none"> - 코로나 상황으로 인해 면접원 교육은 화상 교육으로 이루어짐 - 조사원교육은 작성기관 담당자 참관하에 2시간 진행됨 - 조사원교육자료, 세부일정 및 계획 자료확인함 - 교육시 보안교육을 한 후 서약서를 받음(서약서 확인함) - 조사진행 중 탈락 조사원이 없어서 신규 조사원을 선발하지 않았으므로 재교육이 필요가 없었음 - 조사원별 첫 회수 조사표를 대상으로 에디팅 방법을 다시 점검/교육함 - 조사 후 개별조사원의 별도 평가는 이루어지지 않음 	<ul style="list-style-type: none"> - 조사종료 후 조사원 평가 관리 필요
<p>조사원 업무량 (응답소요시간, 조사난이도, 조사기간 등 참고자료)</p>	<ul style="list-style-type: none"> - 조사표작성시 소요시간은 업체 규모에 따라 최소30분~ 최대 2시간이 소요되는 것으로 파악됨 - 조사 난이도는 중간 정도이지만 GPU보유 및 수요 문항에서 응답자가 모르는 경우가 많아 무응답이 상대적으로 많음 (무응답률 보유문항 52% 수요문항 58%) - 1개 사업체당 조사기간은 사업자 특성에 따라 편차가 매우 심함. 대기업 일수록 조사기간이 오래 소요됨 (여러 부서가 응답 해야하므로 조사표작성에 2~3일 소요) - 전체 실사진행 기간은 2개월 소요됨 	<ul style="list-style-type: none"> - GPU 보유 및 수요 문항 응답률 제고

<p>조사준비 및 준비조사 (홍보내역, 응답자 사전 통지, 조사명부보완)</p>	<ul style="list-style-type: none"> - 조사 시작 전 조사 대상 회사에 조사홍보로 협조공문과 함께 안내메일 발송 - 모집단 정비조사를 통해 조사모집단 명부 확정 - 조사진행과정에서 소수 부적격 업체를 최종 조사모집단에서 제외시킴(2021년 조사모집단수 1,365개)(명부보완 내역 확인함) 	<ul style="list-style-type: none"> - 해당사항없음
<p>조사항목별 조사방법 (조사 지침서 및 항목별 내검지침서)</p>	<ul style="list-style-type: none"> - 조사지침서 구비 - 항목별 내검지침서 확인 - 현장 내검 및 CAPI 방식으로 진행되어 대부분의 내검은 프로그램화되어 자동적으로 이루어 짐. 현장 육안 검증은 지침에 의해 이루어짐 	<ul style="list-style-type: none"> - 해당사항없음
<p>현장조사관리 (현장조사관리지침, 현장조사 파라데이터 자료)</p>	<ul style="list-style-type: none"> - 현장조사 관리지침 및 현장 - 조사 사례집에 해당되는 내용이 개략적으로 조사지침서에 수록되어 있음 - 조사진행관리와 관련된 파라데이터가 수집되지 않음 (재방문회수, 첫방문후 회수시까지의 기간, 방문시간대 등) 	<ul style="list-style-type: none"> - 조사진행관련 파라데이터 수집 및 관리 필요
<p>조사질의 응답체계 (현장조사 질의응답체계 운영, 주요질의 응답, 오류사례, 현장조사사례집)</p>	<ul style="list-style-type: none"> - 현장조사질의 응답체계가 운영되었고, 주요 질의에 대한 응답내용이 기록 관리됨. 오류사례는 기록관리 되고 있지 않음 - 현장조사사례집이 조사지침서에 준비되어 면접원교육에 전달되고 활용됨 	<ul style="list-style-type: none"> - 오류사례 기록관리 필요

<p>기억응답</p>	<ul style="list-style-type: none"> - 해당사항 없음 - 기억응답을 방지하기 위하여 사업체별 인력 및 재무현황 자료 첨부함 	<ul style="list-style-type: none"> - 해당사항없음
<p>무응답대처 (항목, 단위무응답 대처 지침, 사례)</p>	<ul style="list-style-type: none"> - 항목 무응답의 경우 조사표 회수 후 전화로 리체크하여 최소화시킴 - 대기업 응답률을 높이기 위해 조사통계위원 위원들 중 대기업 재직 위원의 도움을 받음 - 무응답 발생시 대처 지침을 확인함 - 전수조사로 단위무응답 발생시 대체가 없음 	<ul style="list-style-type: none"> - 해당사항없음
<p>표본대체</p>	<ul style="list-style-type: none"> - 전수조사로 해당사항 없음 	<ul style="list-style-type: none"> - 해당사항없음
<p>사후조사 (모니터링실시계획 자료, 모니터링대상명부, 선정내역, 검증항목 등)</p>	<ul style="list-style-type: none"> - 무응답 항목 리체크가 많아 문항리체크와 검증작업을 동시 진행함 	<ul style="list-style-type: none"> - 검증항목 및 검증 결과 관련 문서화 작업 필요

제3부 자료수집 체계 점검 결과

1. 자료수집 체계 점검 개요 및 설계

가. 점검 개요

조사원, 조사 기획자, 조사 관리자를 대상으로 자료수집과 관련된 인터뷰를 하여 조사 진행 및 관리 실태를 점검하며, 문제점 및 개선점을 발견하고자 함.

인터뷰를 통해 자료수집방법, 조사원 모집 및 교육/평가, 사전 조사준비 및 홍보, 조사 모집단 확정 및 조사 모집단 현황, 조사진행 관리, 자료 입력 및 검증 등 자료수집 체계와 관련된 중요 내용들을 점검하고 확인함.

나. 점검 설계

조사관리 기획 및 관리자 1명, 조사 진행 관리자 1명, 조사원 1명을 심층면접함.

2. 점검 결과

가. 현황 및 문제점

(1) 현황

2021년 인공지능산업실태조사는 기업체 방문 유치조사(일부 온라인조사 병행)로 전수조사이다. 현장조사 관리가 적절하게 이루어져 자료수집 과정은 전반적으로 잘 관리되고 있는 것으로 판단되지만 전수대상 기업체 중 실제 조사가 이루어진 비율은 50% 정도로 조사 대상 기업들의 조사 참여율이 낮은 편임.

조사표는 재무 및 R&D 인력에 의해 주로 작성됨. 조사표 문항 중 GPU 보유 및 수요 문항에 대한 응답률이 낮음.

나. 주요 개선의견

(1) 조사 참여율 제고

전수조사이지만 조사 참여율이 50%로 낮은 수준임. 조사참여율을 높이기 위해(특히 대기업) 다음의 방법들을 적용 검토할 필요가 있음.

- ① 조사 시작전 조사 대상기업에 대한 협조공문을 조사작성기관인 과학기술정보통신부에서 직접 발송(현재는 조사회사에서 발송)하여 조사대상기업에 대한 조사 협조 요청
- ② 매년 조사시 조사진행에 어려움이 많았던 조사대상 사업체를 원인 분석하여 문서화하여 차기 조사 시 활용
- ③ 조사 종료후 조사작성기관인 과학기술정보통신부 담당자가 조사참여 감사메일(Thanks Mail) 발송
- ④ 조사결과를 가공하여 조사대상업체가 관심을 가질 만한 내용을 5페이지 내외로 3~4개 유형으로 만들어 조사 참여자에게 제공하여 본 통계의 유용성 홍보

(2) 표본조사로 전환시점 검토

전수 대상기업의 최근 몇 년간 증가 추세를 고려하면(2019년 : 424개, 2020년 : 933개, 2021년 : 1,365개) 본 통계는 해를 거듭할수록 전수 대상기업이 지속적으로 증가할 것이고 이로 인하여 조사 참여율은 더 낮아질 가능성이 높을 것으로 판단된다.(조사참여율 2019년 64%, 2020년 57%, 2021년 50%) 현재의 전수조사 방식에서 표본조사 방식으로 전환할 시점을 검토할 필요가 있을 것임.

(3) 조사관리용 문서 세분화(고도화)

현재 본 통계의 조사관리용 문서는 조사지침서가 거의 유일함. 이 조사지침서에 조사진행 관리에 필요한 전반적인 내용이 수록되어 있지만 조사과정별로 보다 세분화하여 조사관리용 문서를 작성할 필요가 있음. 세분화 필요가 있는 문서는 조사 거부기업별 상세 거부원인 분석 문서, 에디팅 지침서, 리체크 및 검증 결과 내용 리스트, 조사원 평가결과 기록문서 등으로 사료됨.

(4) GPU 보유 및 수요 문항 응답률 제고 방안 수립

GPU 보유 및 수요 문항의 응답률이 보유문항 48%, 수요문항 42%로 매우 낮은 편이다. GPU 보유 및 수요 현황을 보다 현실에 맞게 수정하여 응답률을 높일 수 있게 문항 수정이 필요할 것이다. 응답 기업과 업계 전문가 의견을 수렴해 문항 수정을 하여야 할 것임.

다. 근거자료 확인 목록

[매뉴얼 III.자료수집] 진단항목	근거자료 목록	확인결과
1. 조사방법	· 조사 응답 비율, 응답자특성 · 분석결과 자료	분석표에서 분석결과 확인함. 잘 관리됨
2-1 조사원 채용 및 처우	· 채용 과정 및 계획 문서	조사원 채용 관련 현황 확인함
2-2 조사원 교육훈련	· 조사원 교육자료 · 교육 세부일정 및 계획/결과 · 보안 교육 및 서약서 · 조사원 평가 결과 · 재교육 일정 등	조사원 평가 기록관리를 하지 않음 나머지 사항은 잘 관리됨
2-3 조사원 업무량	· 응답소요시간, 조사난이도, 조사기간 등 참고자료	참고자료 확인함
3-2 조사준비 및 준비조사	· 홍보 내역 · 응답자 사전 통지서 · 조사구 또는 명부 보완내역	관련 근거자료 확인함(잘 관리됨)
3-3 조사항목별 조사방법	· 조사 지침서 · 항목별 내검지침(추가 확인)	조사지침서, 항목별 내검지침 내용 확인
3-4 현장조사 관리	· 현장조사 관리 지침 · 현장조사 파라미터 세부자료 (방문 또는 접촉시도 횟수, 방문요일 및 시간대, 조사 성공/실패 등) · 실사지도(지도점검) 결과자료	실사지도 점검 결과문서화 안됨 현장조사 파라미터 수집이 안됨
3-5 조사 질의응답 체계	· 현장조사 질의 응답 체계 운영방법 · 주요 질의 응답, 오류사례 · 현장조사 사례집	질의응답체계운영함. 현장조사사례 및 주요 질의사례는 조사지침서에서 확인함 오류사례 기록관리가 안됨
4-2 기억응답	· 기억응답에 활용된 참고자료	참고자료 확인함
4-3 무응답 대처	· 항목, 단위 무응답 대처 지침, 사례	조사지침서에서 확인함. 단위무응답은 해당사항 없음
4-4 표본대체	· 표본대체 기준 및 방법 · 표본대체 목록 현황 자료	해당사항 없음
5. 사후조사	· 모니터링 실시 계획자료 · 모니터링 대상 명부, 표본선정내역, 질문지, 검증항목 및 오차범위 등 · 모니터링 결과자료 및 사후 조치 사례	모니터링결과 자료 문서화 되지 않음

붙임2

표본설계 점검 결과

통 계 명	인공지능산업실태조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
점 검 일 시	2022년 8월 24일
연 구 원	오유진
연구보조원	심주용, 전해정

제1부 점검 개요

I. 점검 개요

- 표본설계 점검 시 검토한 자료
 - 인공지능산업실태조사 통계정보 보고서
(조사개요, 작성목적, 조사설계, 통계추정 및 분석)
 - 2021 인공지능산업실태조사 보고서

II. 조사 개요

조 사 명	인공지능산업실태조사	
작 성 기 관 명	과학기술정보통신부	
작 성 주 기	1년	
전수/표본조사	전 수 (●)	표 본 ()
표본설계주체	자체설계()	외부용역()
조 사 목 적	인공지능산업의 현황 파악과 AI 산업 생태계 조성 및 활성화를 위한 정책 수립의 기초자료로 활용	
조 사 대 상	인공지능산업 관련 사업을 영위하는 전국 기업체	
조 사 방 법	방문조사, 인터넷 웹 조사, 전화조사	

제2부 **점검 결과 요약**

- 점검 결과 주요 문제점 및 개선의견 정리

구 분	점검결과	개선의견
모집단	- 모집단 정의를 제시하고 있음	-
조사모집단	- 조사모집단 현황 및 분포를 제시하고 있음	-
추정	- 사후가중치 및 추정식을 제시하고 있음	- 추정식에 사용된 가중치 W_{khi} 산출과정을 제시할 필요가 있음
무응답처리	- 무응답 대체방법 및 현황을 제시하고 있음	-

제3부 표본설계 점검 결과

1. 표본설계 점검 개요

인공지능산업실태조사의 통계명, 승인번호, 작성기관, 조사목적, 조사대상, 조사방법은 다음과 같다.

- (1) 통계명 : 인공지능산업실태조사(작성주기 : 1년)
- (2) 승인번호 : 제127016호
- (3) 작성기관 : 과학기술정보통신부
- (4) 조사목적 : 인공지능산업의 현황 파악과 AI 산업 생태계 조성 및 활성화를 위한 정책 수립의 기초자료로 활용
- (5) 조사대상 : 인공지능산업 관련 사업을 영위하는 전국 기업체
- (6) 조사방법 : 방문조사, 인터넷 웹 조사, 전화조사
- (7) 표본설계연도 : 2021년

이번 표본설계 진단은 2021년도 기준 인공지능산업실태조사가 전수조사임을 감안하여, 모집단, 조사방법, 가중치 및 추정, 응답률 등의 항목에 대해서만 진단하였으며, 작성기관에서 작성한 정보보고서와 통계 간행물 등을 토대로 진단을 실시하였다.

2. 점검 결과

가. 모집단 및 표본추출틀

(1) 현황

인공지능산업실태조사에 대한 모집단 설명은 다음과 같다.

○ 목표모집단

- 인공지능 기술 개발 및 인공지능 적용 제품/서비스/플랫폼의 생산, 유통, 활용, 부가서비스(조사/분석, 컨설팅, 중개) 과정에서 가치를 창출하는 기업체
- 사업적 범위 : 인공지능 소프트웨어, 인공지능 서비스, 인공지능 하드웨어 관련 산업 활동을 포함
- 대상적 범위 : 정보통신업(58~63)에서 인공지능 소프트웨어·서비스를 생산·제공하는 기업체와 제조업(26)에서 인공지능 하드웨어를 생산하는 기업체

<대상적 범위(분류체계 코드별)>

대분류	중분류	KSIC Rev. 10
1. 인공지능 소프트웨어 개발 및 공급업	11. 인공지능 시스템	■(58221)시스템 소프트웨어 개발 및 공급업
		■(58211)유선 온라인 게임 소프트웨어 개발 및 공급업
	12. 인공지능 응용 소프트웨어 개발 및 공급업	■(58212)모바일 게임 소프트웨어 개발 및 공급업
		■(58219)기타 게임 소프트웨어 개발 및 공급업
		■(58222)응용소프트웨어 개발 및 공급업
2. 인공지능 구축·관리 및 관련 정보 서비스업		■(62010)컴퓨터 프로그래밍 서비스업
		■(62021)컴퓨터 시스템 통합 자문 및 구축 서비스업
		■(62022)컴퓨터시설 관리업
		■(62090)기타 정보 기술 및 컴퓨터 운영관련 서비스업
		■(63111)자료처리업
		■(63112)호스팅 및 관련 서비스업
		■(63120)포털 및 기타 인터넷 정보매개 서비스업
		■(63910)뉴스 제공업
		■(63991)데이터베이스 및 온라인 정보 제공업
	■(63999)그 외 기타 정보 서비스업	
3. 인공지능 연산 및 처리 부품/장치 제조업		■(26111)메모리용 전자집적회로 제조업
		■(26112)비메모리용 및 기타 전자집적회로 제조업
		■(26121)발광 다이오드 제조업
		■(26129)기타 반도체소자 제조업
기타		

○ 조사모집단

- ICT 통합모집단을 대상으로 모집단 정비조사를 통해 확인된 1,365개의 인공지능 기업체
 - 기본 명부 작성 : ICT 통합모집단에서 인공지능산업분류에 해당하는 기업체를 대상으로 기본 명부 작성
 - 모집단 정비 : 기본 명부 대상으로 데스크 리서치, 유선 판별조사 등을 수행하여 조사 모집단 확정
- ① 사전 조사 : ICT통합모집단을 기반으로 분류체계 기반 선별, 데스크리서치, 유선 판별 조사를 통해 인공지능 사업체 선별
- 분류체계 기반 선별(30만개 → 54,358개) : ICT통합모집단에서 인공지능산업 분류체계-표준산업분류 연계표에 해당하는 기업체 선별
 - 데스크리서치(54,358개 → 34,619개) : ICT통합 모집단 정보 및 홈페이지 확인을 통해 인공지능 사업을 영위하지 않는 기업체(19,739개)를 제외
 - 유선판별 조사(34,619개 → 1,156개) : 34,619개 사업체 중 전화번호 정보가 있는 사업체를 대상으로 유선판별 조사를 시행하여 1,156개의 인공지능 기업체 발견
- * ICT 통합모집단 및 추가 검색을 통해 전화번호 정보를 확보한 15,723개 기업체
- ② 추가 명부 활용(257개 기업체) : 과학기술정보통신부 AI국 수탁 기업 명부를 활용하여 ICT 통합모집단을 보완
- 조사과정 중 확인된 폐업 및 인공지능 비즈니스 사업을 영위하지 않는 기업체를 제외한 1,365개 기업체를 최종 조사모집단으로 함



구분	종사자 수 구간	모집단 수(개)	
전체	합계	1,365	
	1,000인 이상	24	
	100~1,000인 미만	112	
	10~100인 미만	505	
	10인 미만	724	
	1. 인공지능 소프트웨어	합계	916
		1,000인 이상	13
		100~1,000인 미만	63
		10~100인 미만	340
		10인 미만	500
	2. 인공지능 서비스	합계	432
		1,000인 이상	11
		100~1,000인 미만	46
		10~100인 미만	159
		10인 미만	216
	3. 인공지능 하드웨어	합계	17
1,000인 이상		0	
100~1,000인 미만		3	
10~100인 미만		6	
10인 미만		8	

(2) 점검결과

모집단에 대한 점검결과, 목표모집단은 ‘인공지능 기술 개발 및 인공지능 적용 제품/서비스/플랫폼의 생산, 유통, 활용, 부가서비스(조사/분석, 컨설팅, 중개) 과정에서 가치를 창출하는 기업체’로 정의하고 있고, 조사모집단은 ‘ICT 통합모집단을 대상으로 모집단 정비조사를 통해 확인된 1,365개의 인공지능 기업체’로 정의하고 있다. 사전 조사와 추가 명부를 활용한 모집단 정비 과정을 제시하고 있으며, 모집단 현황 및 분포를 제시하고 있다.

나. 추정

(1) 현황

인공지능산업실태조사에 대한 가중치 작성 및 추정식은 다음과 같다.

- 사후가중치 적용 필요성
 - 본 조사는 전수조사로 진행되었으나, 응답률이 50%에 그치는 등 조사된 표본의 분포가 모집단 분포와 차이가 있으므로 모집단을 대표할 수 있도록 사후증화 방법에 따라 가중치 적용
- 가중치
 - 가중치는 3개 업종(인공지능 소프트웨어, 서비스, 하드웨어)과 4개 종사자 규모(1,000인 이상, 100~1,000인 미만, 10~100인 미만, 10인 미만)를 증화변수로 지정
 - 모집단 정비조사를 통해서 확정된 모집단의 종사자 수 등에 관한 종합적인 정보를 알고있는 경우 설계 가중치와 무응답 조정 계수를 곱한 가중치를 적용하여 추정한 매출액이나 종사자 수의 비를 사용하여 모수 추정에서 편향을 보정하기 위해서 아래와 같은 사후증화보정계수를 가중치로 보완하였음

$$BF_{ij} = \frac{\text{모집단의 종사자수}}{\text{조사된 표본기업체의 가중 종사자수}}$$

○ 주요 가중치 분포

구분	종사자 수 구간	모집단 수(개)	응답수(개)	가중치
전체	합계	1,365	684	2.00
	1,000인 이상	24	13	1.85
	100~1,000인 미만	112	62	1.81
	10~100인 미만	505	281	1.80
	10인 미만	724	328	2.21
1. 인공지능 소프트웨어	합계	916	492	1.86
	1,000인 이상	13	9	1.44
	100~1,000인 미만	63	42	1.50
	10~100인 미만	340	190	1.79
	10인 미만	500	251	1.99

2. 인공지능 서비스	합계	432	181	2.39
	1,000인 이상	11	4	2.75
	100~1,000인 미만	46	17	2.71
	10~100인 미만	159	87	1.83
	10인 미만	216	73	2.96
3. 인공지능 하드웨어	합계	17	11	1.55
	1,000인 이상	0	0	-
	100~1,000인 미만	3	3	1.00
	10~100인 미만	6	4	1.50
	10인 미만	8	4	2.00

○ 총계 추정식

$$\text{추정식 } \hat{Y}_k = \hat{Y}_{kc} + \hat{Y}_{ks} = \hat{Y}_{kc} + \sum_{h=1}^L \sum_{i=1}^{n_{ksh}} W_{khi} y_{khi}$$

(2) 점검결과

추정방법에 대한 점검결과, 전수조사이지만 무응답으로 인한 편향을 보정하기 위해서 사후가중치를 제시하고 있으며 주요 가중치 분포, 가중치를 이용한 총계추정식도 함께 제시하고 있다.

다만, 총계추정식에 사용되는 가중치 W_{khi} 에 대한 산출과정이 제시되어 있지 않으므로 이에 대해 제시할 필요가 있다.

다. 무응답처리

(1) 현황

인공지능산업실태조사에 대한 무응답처리방법은 다음과 같다.

○ 무응답 대처

- 응답 유도를 통한 무응답 최소화
 - 항목 무응답 응답 유도 방안
 - 무응답 발생 시 리체크를 통해 전화로 재조사 진행함
 - 업체의 재무정보와 같이 민감한 사항은 차후 조사를 진행할 수 있도록 함
 - GPU 관련 항목을 제외한 항목의 무응답률은 1% 미만
 - 단위 무응답 응답유도 방안
 - 컨택원이 응답자에게 2주 혹은 한 달 정도의 시간을 두고 5회 이상 연락
 - 조사담당 슈퍼바이저와 연구진이 직접 전화 또는 방문하여 설득
 - 단위 무응답 처리 기준
 - 다음의 상황일 때 단위 무응답으로 처리함
 - ‘부채중’, ‘바쁨’, ‘내부 결재 중’, ‘상급자와 논의’ 등의 사유는 일정을 확인한 후 일정에 맞춰 연락. 일정을 정확히 확인해주지 않는 경우 2주 정도의 시차를 두고 5회 연락
 - ‘단순 거절’ 은 한 달 정도 후 3회 연락
 - ‘강력 거절’ 은 1회 컨택 후 종료
 - 최대한 정보 확인 후 확인되지 않은 기업체는 ‘정보 확인 안됨’ 으로 구분

○ 항목 무응답

- 산출산식

$$\text{항목 무응답률(\%)} = \frac{\sum \text{항목 무응답 기업체 수}}{\sum \text{해당년도 조사 성공 기업체 수}} \times 100$$

- 2021년 인공지능산업실태조사 조사항목에서 무응답률이 높은 항목은 ‘GPU 보유’ 문항이며 그 외 문항은 0.8% 수준임

* GPU 보유 규모 및 수요 문항 무응답률(21년 기준) : 평균 52.2%

** GPU 보유 규모 및 수요 외의 문항 무응답률(21년 기준) : 평균 0.8%

- 각 조사 항목별 무응답 현황은 다음과 같음

<매출 현황>

문항	변수	무응답수	무응답률	대체방법
2021 전체 매출액 및 인공지능 분야 매출액	전체 매출액	1	0.1	중시점 회귀대체 Hot-deck
	인공지능 매출 발생 여부	1	0.1	
	인공지능 매출	3	0.4	
	인공지능 시스템 소프트웨어	1	0.1	
	인공지능 응용 소프트웨어	3	0.4	
	인공지능 구축·관리 및 관련 정보 서비스	3	0.4	
	인공지능 연산 및 처리 부품/장치	1	0.1	
	공공	1	0.1	
	민간	3	0.4	
	개인	3	0.4	

<인력 현황>

문항	변수	무응답수	무응답률	대체방법
2021 전체 인력 및 인공지능 분야 인력	전체 종사자	2	0.3	중시점 회귀대체 Hot-deck
	인공지능 종사자	22	3.2	
	인공지능 여성 종사자	39	5.7	
	인공지능 외국인 종사자	2	0.3	
21년 직업별 인공지능 인력 현황 현재 상용근로자	인공지능 프로젝트 관리자	2	0.3	중시점 회귀대체 Hot-deck
	인공지능 컨설턴트	2	0.3	
	인공지능 개발자	3	0.4	
	인공지능 아키텍처 설계 및 분석가	3	0.4	
	인공지능 SW 개발자	2	0.3	
	인공지능 HW 개발자	2	0.3	
	기타	2	0.3	
	인공지능 시스템 운영·관리자	2	0.3	
	데이터 가공·처리 종사자	2	0.3	
	인공지능 데이터 분석가	3	0.4	
합계	2	0.3		
21년 직업별 인공지능 인력 현황 입시 및 일용 근로자	인공지능 프로젝트 관리자	2	0.3	중시점 회귀대체 Hot-deck
	인공지능 컨설턴트	2	0.3	
	인공지능 개발자	2	0.3	
	인공지능 아키텍처 설계 및 분석가	2	0.3	
	인공지능 SW 개발자	2	0.3	
	인공지능 HW 개발자	2	0.3	
	기타	2	0.3	
	인공지능 시스템 운영·관리자	2	0.3	
	데이터 가공·처리 종사자	2	0.3	
	인공지능 데이터 분석가	2	0.3	
합계	2	0.3		

문항	변수	무응답수	무응답률	대체방법
21년 직업별 인공지능 인력 현황 합계	인공지능 프로젝트 관리자	3	0.4	중시점 회귀대체 Hot-deck
	인공지능 컨설턴트	3	0.4	
	인공지능 개발자	3	0.4	
	인공지능 아키텍처 설계 및 분석가	3	0.4	
	인공지능 SW 개발자	3	0.4	
	인공지능 HW 개발자	3	0.4	
	기타	3	0.4	
	인공지능 시스템 운영·관리자	3	0.4	
	데이터 가공·처리 종사자	3	0.4	
	인공지능 데이터 분석가	3	0.4	
	합계	2	0.3	
21년 직업별 인공지능 인력 현황 부족 인력	인공지능 프로젝트 관리자	2	0.3	중시점 회귀대체 Hot-deck
	인공지능 컨설턴트	2	0.3	
	인공지능 개발자	2	0.3	
	인공지능 아키텍처 설계 및 분석가	3	0.4	
	인공지능 SW 개발자	2	0.3	
	인공지능 HW 개발자	2	0.3	
	기타	2	0.3	
	인공지능 시스템 운영·관리자	2	0.3	
	데이터 가공·처리 종사자	2	0.3	
	인공지능 데이터 분석가	2	0.3	
	합계	2	0.3	
22년 채용 예정 인력	인공지능 프로젝트 관리자	1	0.1	중시점 회귀대체 Hot-deck
	인공지능 컨설턴트	1	0.1	
	인공지능 개발자	2	0.3	
	인공지능 아키텍처 설계 및 분석가	2	0.3	
	인공지능 SW 개발자	2	0.3	
	인공지능 HW 개발자	2	0.3	
	기타	1	0.1	
	인공지능 시스템 운영·관리자	2	0.3	
	데이터 가공·처리 종사자	1	0.1	
	인공지능 데이터 분석가	2	0.3	
	합계	2	0.3	
플랫폼 노동자	플랫폼 노동자	3	0.4	Hot-deck
인공지능 학력별 인력	합계	3	0.4	중시점 회귀대체 Hot-deck
	전문대학 졸업 이하	3	0.4	
	대학교 졸업	3	0.4	
	대학원 석사 졸업	3	0.4	
	대학원 박사 졸업	3	0.4	
인공지능 경력별 인력	합계	3	0.4	중시점 회귀대체 Hot-deck
	3년 미만	3	0.4	
	3년 이상 ~ 7년 미만	3	0.4	
	7년 이상 ~ 10년 미만	3	0.4	
	10년 이상	3	0.4	

<투자 및 개발 현황>				
문항	변수	무응답수	무응답률	대체방법
2021 자금 사용금액	기업 전체 연구개발 투자액	7	1.0	Hot-deck
	인공지능 부문 연구개발 투자액	17	2.5	
	인건비	32	4.7	
	기타 경상비	31	4.5	
	기계 장치	31	4.5	
	토지 건물	31	4.5	
	컴퓨터 소프트웨어	31	4.5	
	기타 지식재산물	31	4.5	
외부 투자 유치 실적	외부 투자 유치 실적 여부	1	0.1	무응답처리
2021 외부 투자 유치 건수 및 금액	외부 투자 유치 건수	6	0.9	Hot-deck
	외부 투자 유치액	12	1.8	
3개년도 외부 투자 유치 방법	IPO(기업 공개/상장/주식발행)	1	0.1	Hot-deck
	벤처캐피탈/엔젤투자	1	0.1	
	정부 정책 지원금	1	0.1	
	은행 등 일반 금융	1	0.1	
	기타	1	0.1	
	기타내용	1	0.1	
현재까지 전체 누적 건수	전체특허출원	20	2.9	중시점
	전체특허등록	20	2.9	회귀대체
	인공지능 특허출원	23	3.4	Hot-deck
	인공지능 특허등록	23	3.4	

○ 항목 무응답 대체

- 재조사 : 재조사 가능한 항목 무응답은 다시 조사하여 보완
- 1차 조사 후 항목 무응답률은 평균 2%이며, 재조사를 통해 항목 무응답률을 0.8%로 낮춤
- 재조사가 불가능한 경우 기업체 기본 현황/재무현황
 - 금융감독원의 전자공시자료와 신용평가정보회사의 정보를 활용하여 입력
 - 상장기업 중 5개사의 재무정보를 전자공시자료를 통해 매출액, 종사자 정보 등을 확보하여 보완하였으며, 비상장사 중 50여개사의 매출액 정보를 신용평가정보회사의 데이터를 확인하여 데이터를 보완함
- 재조사가 불가능한 경우 그 외 문항
 - 항목 유형 및 과거 정보 유무에 따라 방법론을 적용하여 무응답 대체

○ 단위 무응답

- 단위 무응답 원인

- 불능
 - 강력거절 : 담당자와 컨택 후, 기간을 두고 5회 이상 컨택을 시도하였음에도 거절한 경우
 - 조사 잘못된 경우 : 조사완료 후 조사가 잘못된 경우(리체크 불가상황, 인공지능 관련 사업 미진행)
 - 조사 중단 : 조사를 시작하였으나 조사 도중 중단된 경우
- 삭제
 - 휴업 및 폐업 : 업체가 폐업이나 휴업된 경우
 - 장기부채 : 통화가 전혀 되지 않거나 결번인 경우
 - 업체 변경 : 업체가 바뀐 경우
 - 해당없음 : 인공지능 사업(프로젝트)을 진행하지 않은 경우
- 단위 무응답 현황

구분	모집단	응답기관	미응답기관 (거절 등)	단위무응답률(%)
합계	1,365	684	681	49.9

단위 무응답 구분		단위 무응답 (계)	부채중 등	단순거절	강력 거절	정보 확인 안됨
계		681 (100.0)	148 (21.7)	208 (30.5)	105 (15.4)	220 (32.3)
종사자 구간별	1.1000인 이상	11	1	7	3	0
	2.100~1000인 미만	50	14	23	7	6
	3.10~100인 미만	224	57	93	30	44
	4.10인 미만	297	62	71	63	101
	확인안됨	99	14	14	2	69

- 단위 무응답 산출산식

$$\text{단위 무응답률(\%)} = \frac{\sum \text{단위 무응답 기업체 수}}{\sum \text{해당년도 조사 대상 기업체 수}} \times 100$$

(2) 점검결과

무응답처리에 대한 점검결과, 항목무응답과 단위무응답에 대한 대처방안을 제시하고 있다. 그리고 주요항목에 대한 무응답 현황과 그에 따른 대체방법을 제시하고 있으며, 또한 유형별로 단위무응답 현황을 제시하고 있다.

붙임3

조사표 설계 및 유사통계 비교분석 점검

통 계 명	인공지능산업실태 조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
연 구 원	조준기
연구보조원	최다빈, 박혜원

제1부 점검 개요

I. 점검 개요

- 조사표 설계 및 유사통계 비교·분석 점검 시 검토한 자료
 - 통계정보보고서
 - 조사표
 - 조사지침서
 - 통계자료(KOSIS, 보고서)
 - FGI 이용자 의견

II. 조사 개요

조 사 명	인공지능산업실태조사	
작 성 기 관 명	과학기술정보통신부	
작 성 주 기	1년	
작성기준년도	2020년	
전수/표본조사	전 수 (●)	표 본 ()
조 사 목 적	○ 인공지능산업 생태계 조성 및 활성화 정책 수립을 위한 기초자료 확보	
조 사 대 상	○ 인공지능 기술 개발 및 인공지능 적용 제품·서비스·플랫폼의 생산, 유통, 활용, 부가서비스(조사/분석, 컨설팅, 중개) 과정에서 가치를 창출하는 기업체	
조 사 방 법	○ 방문조사, 인터넷 웹 조사, 전화조사 등 병행	
주요조사항목	○ 인공지능 기업체의 일반 현황, 기술 및 사업 현황, 매출 현황, 인력 현황, 투자 및 개발 현황, 애로 사항 및 건의 사항 등 세부 항목을 조사	

제2부 점검 결과 요약

구 분	점검결과	개선 의견	비 고
주요 용어 및 항목별 정의	- 각 용어 및 항목에 대한 정의가 적절함	-	
조사표 구성	- 조사표 수록사항 10개 중 8개 확인 - 조사표에 법적근거, 문의사항 연락처 명시 필요	- 조사표 수록사항 보완	
조사표 설계 및 변경 절차	- 조사표 설계 및 변경 절차가 적절함	-	
조사항목의 적정성	- 조사항목 구성 및 질문 방식이 적절함	-	
응답항목 및 지시문의 적정성	- 응답항목 및 문항이동을 나타내는 지시문이 적절함 - 문항이동을 나타내는 지시문이 적절함	-	
기준시점의 적정성	- 조사항목별 기준시점이 대체로 적절함 - 설명자료와 조사표 간 조사 기준시점이 대응하지 않은 문항 존재	- 조사항목별 기준시점 보완 검토	
조사표 변경 이력 관리	- 조사표 변경 이력 관리가 적절함	-	
조사항목별 작성요령 및 유의사항	- 조사항목별 작성요령 및 유의사항이 대체로 적절함 - GPU 규모 관련 문항의 작성요령 예시 추가 필요	- 조사표 보완 검토	
동일영역 통계와 일관성	- 동일영역 통계의 명칭 및 개요를 제시함 - 동일영역 통계에 대해 구체적으로 검토함	-	
유사통계항목 간 수치의 정확성	- 인공지능 관련 공급기업의 실태를 파악하는 통계는 본 통계가 유일함	-	해당 사항 없음

제3부 조사표 설계 및 유사통계 비교·분석 점검 결과

1. 점검 개요

「조사표 설계 및 유사통계 비교·분석 점검」은 응답자의 응답 부담 경감, 이해도 제고를 위해 조사표 설계 과정에서 발생할 수 있는 측정오차를 점검하는 과정으로, 자료수집의 정확성 진단을 강화하기 위하여 점검하는 과정이다.

조사표는 조사목적에 부합하는 정보를 응답자로부터 얻기 위하여 고안된 질문들을 모아놓은 표이다. 조사표는 자료수집과정에서 아주 핵심적인 역할을 한다. 자료가 조사표의 질문에 근거하여 수집되기 때문에 조사표는 자료 품질에 직접적인 영향을 준다.

유사통계는 서로 다른 통계더라도 동일한 공표항목이 존재하는 통계를 말한다. 예를 들어 동일한 영역에서 조사통계 간 유사한 통계 항목이 존재할 수 있으며, 보고·가공통계에서 공표하고 있는 항목이 조사통계에서도 조사 후 공표되는 항목이 있을 수 있다. 통계마다 목적, 대상 범위, 표본설계가 다르므로 완벽하게 동일한 결과를 제공하지는 않는다. 그러나 유사한 내용을 공표하고 있다면 어느 정도 일관성이 있어야 이용자가 신뢰할 수 있다.

진단에서는 통계정보보고서를 기반한 절차적 점검과 조사표 항목 점검 및 유사통계 비교·분석 등을 실시하였다.

가) 조사표 설계 적정성 진단

통계정보보고서 및 기타 설명자료 등을 기반으로 주요 용어 및 항목별 정의, 조사표 구성, 조사표 설계 및 변경 절차, 조사표 변경 이력을 점검한다. 그리고 조사표 점검 및 FGI 의견을 토대로 조사항목 구성 및 질문 방식의 적정성, 응답항목 및 지시문의 적정성, 기준시점의 적정성, 조사항목별 작성요령 및 유의사항 등을 점검한다.

나) 유사통계 비교·분석 점검

점검대상이 공표하고 있는 통계 중 동일하거나 유사한 통계가 있는지 파악한다. 점검통계와 유사한 항목이 있는 통계 간의 작성기관, 작성목적, 작성대상 및 범위, 작성단위, 작성주기, 기준시점, 공표시기, 표본조사 여부, 작성규모를 비교하고 유사항목의 결과값 및 추이가 유사한지 파악한다.

2. 점검 결과

가) 조사표 설계 적정성 진단

(1) 주요 용어 및 항목별 정의

인공지능산업실태조사는 인공지능 산업 및 직업, 인공지능 기술분야, 사업 품목 등 조사표에 활용되는 모든 용어의 정의와 개념을 보고서 및 조사표를 통해 구체적으로 제시하고 있으므로 적절하다고 판단된다.

(2) 조사표 구성

조사표 수록사항인 조사명, 조사목적, 법적근거, 국가승인통계로고, 작성 승인번호, 응답자 협조사항, 조사협조 감사인사, 조사기관, 응답자 비밀보호 정책, 문의사항 연락처 10가지 항목의 수록 여부를 확인한 결과, 법적근거와 문의사항 연락처 2가지가 명시되어 있지 않은 것으로 나타났다.

<그림 1> 조사표 개요



2021년 인공지능산업 실태조사

승인번호
제 127016 호

안녕하십니까?

귀사의 무궁한 발전을 기원합니다.
과학기술정보통신부와 지능정보산업협회에서는 인공지능산업 실태를 파악하고, 생태계 조성 전략 수립 및 지원 방안을 마련하기 위하여 「2021년 인공지능산업 실태조사」를 실시하고 있습니다.

귀사가 응답해 주신 내용은 향후 인공지능산업 관련 국가 정책 자료로 활용되오니 바쁘시더라도 잠시 시간을 내어 조사원의 질문에 답변해 주시면 감사하겠습니다.

2021. 9.

통계 작성 기관 : 과학기술정보통신부 조사 기관 : 지능정보산업협회
수행 기관 : 소프트웨어정책연구소 조사 위탁 기관 : ㈜메가리서치

◆ 응답 및 작성 방법

- 이 조사표에 기재된 내용은 통계법 제33조에 따라 비밀이 보장되며 통계작성 이외의 목적으로는 절대로 사용하지 않습니다.
- 응답은 귀사의 인력, 기술, 자금 등 경영의 전반적인 상황을 잘 파악하고 계신 대표이사나 임원 및 관리자(부장급)께서 직접 작성하셔야 정확하게 기입할 수 있는 내용으로 구성되어 있습니다.
- 본 조사의 대상 기간은 2020년 1월 1일 ~ 2020년 12월 31일입니다.
특별한 응답 지시문이 없을 경우 모든 응답은 2020년 12월 31일을 기준으로 작성해 주십시오.
단, 인력과 관련된 문항은 각 문항 설명에 따라, 현재 등의 시점을 기준으로 작성해 주십시오.
- 질문 내용 중 특별한 안내문이 없는 한 모든 문항에 대해 주십시오.
- 종업원 수, 금액 또는 % 부분에서 해당 사항이 없는 경우 '0' 또는 '-'를 표시해 주십시오.
- 각 문항마다 설명과 작성 안내문을 참고하여 작성해 주시기 바랍니다.

기업명			
부서/직위		연락처	
응답자 성명		이메일 주소	@

(3) 조사표 설계 및 변경 절차

인공지능산업실태조사는 기존 조사항목에 대한 자료 이용 분야별 적합성을 검토하고 조사통계위원회를 개최해 신규 조사항목 발굴 및 조사표 개선을 위한 전문가 의견수렴을 하는 것으로 나타났다. 또한, 기업 대표 및 종사자를 대상으로 사전조사를 실시하여 조사항목별 위험요인을 사전에 파악하고 있으며, 최종적으로 통계청 변경승인을 통해 조사표를 확정하고 있으므로 조사표 설계 및 변경 절차는 적절하다고 판단된다.

(4) 조사항목의 적정성¹⁾

인공지능산업실태조사는 인공지능산업 생태계 조성 및 활성화 정책 수립을 위한 기초자료를 확보하기 위해 매년 조사항목을 추가·보완하고 있다. 본 통계는 내부 검토 및 전문가 자문을 통해 조사항목의 타당성 및 신규 문항 추가에 대해 검토하고 있어 조사항목 구성 및 질문 방식은 적절한 것으로 판단된다.

다만, 이용자 FGI에서 직업분류별 인공지능 인력 현황뿐만 아니라 직업분류별 학력 수준 및 전공 분야 관련 통계에 대한 수요가 있는 것으로 나타났으며, 인공지능 학습용 데이터 확보 방식뿐만 아니라 어떤 분야의 학습용 데이터를 주로 확보하는지에 대한 조사항목이 필요하다는 의견이 제시되었다.

(5) 응답항목 및 지시문의 적정성²⁾

인공지능산업실태조사 조사표의 응답항목은 독립적이고 상호배타적이며, 문항이동을 나타내는 지시문이 명확히 제시되어 있으므로 응답항목 및 지시문은 적절한 것으로 판단된다.

1) '(4) 조사항목의 적정성'에 작성된 의견은 한국통계진흥원 통계품질센터 연구진의 의견으로 통계청 견해가 아님

2) '(5) 응답항목 및 지시문의 적정성'에 작성된 의견은 한국통계진흥원 통계품질센터 연구진의 의견으로 통계청 견해가 아님

(6) 기준시점의 적정성

2021년 본 조사의 대상 기간은 2020년 1월 1일부터 2020년 12월 31일까지이며, 특별한 응답 지시문이 없을 경우 모든 응답은 2020년 12월 31일을 기준으로 작성하도록 <그림 1> 조사표에 안내하고 있다. 다만, ‘GPU 보유 현황 및 희망 규모’ 문항은 기준시점에 대한 지시문이 없음에도 조사기준시점이 조사 시점(2021년 6월 30일)인 것으로 확인되었다. 또한, ‘플랫폼 노동자 수’ 문항은 ‘인력 현황’ 문항과 연계되는 항목임에도 2020년을 기준시점으로 조사가 이루어져 문항 간 기준시점의 일관성이 떨어지는 것으로 나타났다. 이렇게 조사표상 확인되는 문항별 기준시점이 설명자료와 대응하지 않는 문항이 존재하므로 작성기관은 아래 <표 1> 내용을 검토하여 조사표 및 설명자료를 보완할 필요가 있다.

<표 1> 기준시점 검토내용

문항	항목	내용
Q13 Q14	GPU보유 현황 및 희망 규모	조사시점(2021년 6월 30일)을 기준으로 응답하도록 조사표에 명시 필요
Q20-1	데이터 가공·처리 종사자 중 플랫폼 노동자 수	인력 현황 관련 문항임에도 2020년 한 해 를 기준으로 조사하고 있음

<그림 2> 인공지능산업실태조사 일러두기

- 본 조사는 2021년 하반기에 실시되었으며, 매출액 조사기준 시점은 2020년 12월(확정치), 2021년 12월(실적 추정치)입니다.
- 회계 기준(다음 해 4월 말) 전 조사 시작에 따른 잠정치 결과

- 본 조사의 기준시점은 2020년 12월이나, 시의성 있는 통계 자료 제공을 위해 인력 현황과 GPU 보유 현황 등은 조사 당시 시점(2021년 6월)을 기준으로 조사한 결과이므로 자료 활용 시 주의가 필요합니다.

(7) 조사표 변경 이력 관리

인공지능산업실태조사는 보고서를 통해 통계작성 변경이력에 대한 기록·관리가 잘 이뤄지고 있다. 본 통계는 조사표의 변경내용 및 사유를 이전 조사표와 비교하여 기록·관리하는 것으로 확인되었다.

(8) 조사항목별 작성요령 및 유의사항

인공지능산업실태조사는 ‘GPU 보유 현황 및 수요’ 문항에 대해 구체적인 작성요령 및 지침이 필요하다고 판단된다.

해당 문항은 GPU 모델별 ‘현재 보유 규모’와 ‘향후 희망 규모’를 조사하고 있다. 여기서 14번 문항인 ‘향후 희망 규모’는 현재 보유 GPU를 포함하여 기입하는 것인지, 현재 보유 GPU를 제외하고 추가로 필요한 GPU를 기입하는 것인지 명확하지 않다. 예를 들어 현재 보유 GPU가 10장, 향후 추가로 필요한 GPU가 15장인 경우, ‘향후 희망 규모’ 문항에 현재 보유한 GPU를 포함하여 25장으로 응답해야 하는지, 추가로 필요한 15장만을 응답해야 하는지 명확하지 않다. 조사가 유치조사 또는 온라인 조사로 이뤄지는 만큼, 조사원이 부재하므로 해당 문항에 대한 구체적인 작성요령을 조사표에 명시할 필요가 있다.

<그림 3> GPU 보유 현황 및 수요 관련 문항

Q13-14. 귀사의 인공지능 계산을 위한 GPU 보유* 현황과 수요**를 아래 표에 작성하여 주십시오.

* 현 시점에서 가장 널리 활용되는 학습용(training) 연산처리장치는 GPU이기 때문에 GPU의 규모로 현황과 수요를 파악하고자 함

** GPU 규모에 대한 수요 부분은 향후 3년 이내 추진하고자 하는 인공지능 제품 및 서비스 개발에 필요한 규모 작성

※ 성능별 보기 참조

GPU 모델명	성능 요약	계산 능력	Q13. 현재보유 규모	Q14. 향후 희망 규모
① GTX 10 시리즈 (GTX 1070, 1080 등)	이론성능 : 5.7 ~ 10.6 TFLOPS 메모리 : 8GB ~ 11GB	이미지 학습 : 초당 100여장 기계번역1) : 1배 성능 향상	_____장	_____장
② RTX 20 시리즈 (RTX 2070, 2080 등)	이론성능 : 6.5 ~ 11.7 TFLOPS 메모리 : 8GB ~ 11GB	이미지 학습 : 초당 150여장 기계번역 : 1.5배 성능 향상	_____장	_____장
③ TITAN X (10시리즈 고성능)	이론성능 : 10.1 TFLOPS 메모리 : 12GB	이미지 학습 : 초당 200여장 기계번역 : 2배 성능 향상	_____장	_____장
④ TITAN RTX (20시리즈 고성능)	이론성능 : 16.3 TFLOPS 메모리 : 24GB	이미지 학습 : 초당 300여장 기계번역 : 3배 성능 향상	_____장	_____장
⑤ Tesla P100 (P4, P40 등 포함)	이론성능 : 10.6 TFLOPS 메모리 : 16GB ~ 24GB	이미지 학습 : 초당 200여장 기계번역 : 2배 성능 향상	_____장	_____장
⑥ Tesla V100	이론성능 : 14 TFLOPS 메모리 : 16GB ~ 32GB	이미지 학습 : 초당 300여장 기계번역 : 3배 성능 향상	_____장	_____장
⑦ Tesla T4	이론성능 : 8 TFLOPS 메모리 : 16GB	이미지 학습 : 초당 180여장 기계번역 : 1.8배 성능 향상	_____장	_____장
⑧ Tesla A100	이론성능 : 19.5 TFLOPS 메모리 : 40GB	이미지 학습 : 초당 500여장 기계번역 : 5배 성능 향상	_____장	_____장
⑨ 위 모델에 없는 경우	모델명 _____		_____장	_____장

1) 기계번역 성능 기준 : K80 활용 시 1배 (LSTM + RNN 활용)

나) 유사통계 비교·분석 점검

(1) 동일영역 통계와 일관성

본 통계는 인공지능 분야 관련 통계라는 점에서 인공지능융합실태조사(미승인통계)가 동일영역 통계로 확인되었으며, 작성기관은 동일영역 통계 간 현황 및 차이점을 구체적으로 파악하고 있는 것으로 나타났다.

인공지능산업실태조사는 AI 품목을 공급하는 기업의 실태를 파악하는 데 초점을 두고 있는 반면, 인공지능융합실태조사는 타 산업의 AI에 기반한 디지털전환 수요를 지표화하는 데 초점을 두고 있어 조사목적 및 조사대상 자체의 차이가 있는 것으로 나타났다.

<표 2> 동일영역 통계 현황

구분	인공지능산업실태조사	인공지능융합실태조사
작성기관	과학기술정보통신부	과학기술정보통신부
통계종류	조사통계	미승인통계
작성목적	인공지능산업 생태계 조성 및 활성화 정책 수립을 위한 기초자료 확보	국내 기업의 AI 활용 실태를 비교적 정확하게 파악하고 융합의 저해 요인을 분석함으로써, 실효성 있는 AI 융합 정책 수립을 위한 기초 통계자료 확보
작성대상 및 범위	인공지능 기술개발 및 인공지능 적용 제품·서비스·플랫폼의 생산, 유통, 활용, 부가서비스(조사/분석, 컨설팅, 중개) 과정에서 가치를 창출하는 기업체	AI 공급기업을 제외한 제조, 금융, 교육, 의료, 에너지 등 AI 도입·활용해 부가가치를 창출하는 AI 수요기업
작성단위	기업체	기업체
작성주기	1년	-
표본/전수	전수조사	표본조사
작성규모	1,365개	2,000개

(2) 유사통계항목 간 수치의 정확성

인공지능 관련 품목 및 서비스를 공급하는 기업의 실태를 파악하는 통계는 인공지능산업실태조사 유일하므로 유사통계항목 간 수치의 정확성 점검은 해당사항이 없다.

3. 주요 개선의견

(1) 조사표 수록사항 보완

본 통계 조사표에 수록되어 있어야 하는 항목 중 범적근거와 문의사항 연락처 2가지가 명시되어 있지 않으므로 이를 보완할 필요가 있다. 조사와 관련하여 답변이 가능한 담당자의 연락처와 범적근거인 ‘통계법 제18조에 의해 승인된 조사통계’를 조사표 개요에 추가하여 응답자에게 안내할 필요가 있다.

(2) 조사표 보완 검토

조사항목별 작성요령 및 유의사항 점검 결과에 따라 다음과 같이 조사표를 보완 및 검토할 필요가 있다.

<표 3> 2021년 조사표 보완 검토사항

문항번호	내용
13, 14	- 현재 보유 GPU를 포함하여 응답하는 것인지, 현재 보유 GPU를 제외하고 추가적으로 필요한 GPU를 응답하는 것인지 구체적인 작성요령 명시
20-1	- ‘지난 1년간’으로 조사기준시점 변경 검토

(3) 조사항목별 기준시점 보완 검토

설명자료 상 2021년 본 조사의 대상 기간은 2020년 1월 1일부터 2020년 12월 31일까지이며, GPU 보유 현황과 인력 현황 문항은 조사 시점(2021년 6월 30일)을 기준으로 하고 있으나, 이러한 설명자료 내용과 조사표에서 확인되는 기준시점이 서로 대응하지 않은 문항이 확인되었다. 작성기관은 문항별 기준시점을 명확히 제시하여 이용자는 정확한 정보를 받고, 응답자는 기준시점의 혼동 없이 정확한 응답을 할 수 있도록 설명자료 및 조사표를 보완할 필요가 있다.

붙임4

이용자 요구사항 반영실태 진단 결과

통 계 명	인공지능산업실태 조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
면 접 일 시	2022년 5월 18일
연 구 원	황규배
연구보조원	-

제1부 회의 준비 및 진행

I. 회의 준비과정

1. 참석자 선정	
<ul style="list-style-type: none"> ● 참석자 선정방법 <ul style="list-style-type: none"> - 작성 위탁기관인 소프트웨어정책연구소 담당 연구원의 지원을 받아 작성된 통계이용자 리스트를 활용하여 실제 인공지능산업실태조사 데이터를 활용하여 연구 및 업무 또는 정책수립을 지원하는 교수 및 소프트웨어기업의 임원, 정책 연구소 연구원을 참석자로 선정 	<ul style="list-style-type: none"> ● 참석자 현황 <ul style="list-style-type: none"> - 정책고객 관련 정책기관 연구위원 1명 - 교수 2명 - 관련 비즈니스 기업 임원 2명
<ul style="list-style-type: none"> ● 실시 장소 토즈모임센터 광화문점 	
<ul style="list-style-type: none"> ● 소요 시간 2시간 	

2. 회의 참석자 명부			
연번	소 속	성 명	직 위
1	한국지능정보사회진흥원 빅데이터플랫폼팀	남OO	책임
2	한경대학교 법경영학부	허O	교수
3	중앙대학교 다빈치교양대학	이OO	교수
4	데이터앤밸류	이OO	대표이사
5	와이즈넷	김OO	상무

II. 회의 진행

회의 진행	
<ul style="list-style-type: none"> - 사전에 준비해둔 질문지를 참고하여 진행하였고 부가적으로 회의를 진행함에 따라 추가적으로 질문지에 없는 질문들을 필요에 의해 질문하여 내용을 좀 더 보강하는 방식으로 회의를 진행하였음 	<ul style="list-style-type: none"> ● 사회자 : 황규배 ● 기록자 : 황규배 ● 관찰자 : 양윤서 ● 녹음 · 녹화 여부 : 녹음

제2부 회 의 록

작성절차별	이용자 요구사항	개선 의견
1. 통계작성기획	- 조사모집단리스트에 대한 대표성이 애매하다.	- 인공지능 관련 기업 조사모집단 리스트 확정을 위해 보다 다양한 소스(자료원)를 활용 - 조사모집단 확정과정에 대한 상세한 정보제공
	- 생산하는 서비스나 제품에 활용되는 인공지능기술의 매출기여도를 측정하여 해당 기업의 인공지능기술 매출을 보다 객관적으로 파악할 필요가 있다.	- 인공지능산업 사업규모를 객관적으로 측정할 수 있는 산식 정립
	- 인공지능사업 분야를 전통적인 산업분류체계를 기초로 세분화하고 있기 때문에 산업 세부현황을 생생하게 반영하는데 한계가 있다.	- 인공지능산업 세부분류체계 점검 및 개선안 검토
	- 지역별 인공지능산업 육성 및 지원을 위해서는 지역별 자료가 필요하다.	- 집계표 구분 변수로 지역변수 도입
	- 해외자료와 비교가능한 콘텐츠를 제공해 주었으면 좋겠다.	- 인공지능산업관련 해외자료 수집 및 가공
3. 자료수집	- 인공지능사업 애로사항에 대한 정량적 자료가 미흡하다.	- 인공지능사업 애로사항에 대한 조사표 보기 문항 세분화하여 폐쇄형 질문 문항으로 구성
	- 인공지능 산업에 대한 해당 회사의 인식도 관련 자료가 있었으면 좋겠다.	- 조사표에 인공지능 사업에 대한 종사자들의 인식도 문항 추가
	- 인공지능 산업 종사자들의 어떤 교육에 대해 필요성을 느끼는지 교육 수요 자료가 필요하다.	- 조사표에 교육관련 요구에 대한 문항 추가 검토
	- GPU 보유 및 수요 자료의 경우 클라우드 임대로 변하고 있기 때문에 이러한 추세를 반영할 필요가 있다.	- GPU 보유 현황 및 수요 문항 개선방안 마련

	<ul style="list-style-type: none"> - 공표보고서의 내용이 부족하다. 	<ul style="list-style-type: none"> - 공표보고서 분석 내용 보강
<p>5. 통계공표 관리 및 이용자서비스</p>	<ul style="list-style-type: none"> - 보다 세부분석 및 산업에 대한 통찰력 발견을 위해서는 마이크로데이터가 이용자들이 접근할 수 있어야 할 것이다. 	<ul style="list-style-type: none"> - 마이크로데이터 제공 검토
	<ul style="list-style-type: none"> - 마이크로데이터의 제공이 현실적으로 어렵다면 웹기반 이용자 맞춤형 집계표데이터 서비스가 되었으면 좋겠다. 	<ul style="list-style-type: none"> - 이용자 맞춤형 웹기반 집계표서비스 제공 검토
	<ul style="list-style-type: none"> - 조사기준시점과 통계공표시점이 1년5개월의 시차로 긴 편이다. 통계공표시점을 조금이라도 당겨주었으면 좋겠다. 	<ul style="list-style-type: none"> - 조사기준시점에서 통계공표 기간까지 기간 단축 검토
<p>6. 통계기반 및 개선</p>	<ul style="list-style-type: none"> - 마이크로데이터 제공, 이용자 맞춤형 집계표 서비스 제공, 통계 생산기간 단축 등을 위해서는 인력보강 및 예산 확보가 있어야 할 것이다. 	<ul style="list-style-type: none"> - 데이터 관리 전문인력 및 예산확보

제3부 이용자 요구사항 반영실태 진단 결과

1. 이용자 요구사항 반영실태 진단 개요 및 설계

가. 진단 개요

통계이용자는 이용하는 통계로부터 기대하는 자료를 충분히 얻기를 원한다. 품질이 우수한 통계일수록 이용자가 원하는 자료를 많이 제공하게 된다.

통계이용자가 통계자료에 대해 얼마나 만족하는지를 조사하면 통계품질을 높이는데 활용될 수 있다. 이러한 목적으로 본 진단에서는 통계이용자의 통계에 대한 만족 수준과 요구사항을 파악한다.

이를 위해 본 통계를 연구 및 정책수립 지원에 활용하는 5인을 초청하여 구조화된 인터뷰 질문지를 활용하여 그룹 토의 형식으로 2022년 4월 27일 진행하였으며, 토의 내용을 참석자의 동의 후 녹음하였다.

나. 진단 설계

- 진단방법: FGI(Focus Group Interview)

통계이용자 5인을 그룹으로 구성하여 특정 주제에 대해서 계획적이면서 자유롭게 토론하는 형태로 진행

참석자 구성: 해당 통계의 데이터를 직접 이용해 연구 활동 및 정책 제안을 하는 이용자로 본 통계에 대한 이용도가 높은 이용자를 선정하기 위해 작성기관이 제공한 이용자 리스트 10명과 한국통계진흥원이 자체 보유한 이용자 리스트 4명을 확보하여 진단자가 유선으로 사전 컨택하여 5명을 최종 선정. 교수 3명, 정책연구소 연구원 2명으로 구성

2. 진단 결과

가. 현황 및 문제점

(1) 현황

인공지능산업실태조사는 디지털 전환 가속화에 따라 미래 혁신성장을 지속적으로 담보하기 위한 정책지원 방안 마련을 위한 기초 통계를 제공하기 위해 2017년부터 2년간 시범조사를 거쳐 2019년 통계의 중요성을 인증받아 국가승인통계로 지정되었다.

본 통계는 인공지능산업 관련 기업 현황 일반, 매출 및 인력현황을 통계자료로 제공하고 있어 인공지능산업관련해 세부적으로 구체적인 통계를 제공하는 국내 유일한 조사로 통계이용자들에게 매우 중요한 통계로 인식되고 있었으며, 통계의 신뢰도면에서도 이용자들로부터 긍정적인 평가를 받고 있었다.

(2) 문제점

본 통계의 이용자들은 본 통계는 ①인공지능산업 분야를 전통적인 산업분류 체계(KSIC)를 기초로 세분화하고 있기 때문에 인공지능산업 세부현황을 생생하게 반영하는데 한계가 있는 것으로 인식하고 있었다. ②마이크로데이터가 제공되고 있지 않으므로 연구자의 필요에 맞게 데이터를 분석하는데 제약이 있다. ③ICT통합모집단을 기초로 인공지능산업분류에 해당하는 기업체를 대상으로 기본 명부를 작성, 이를 대상으로 확인 작업을 거쳐 인공지능사업을 영위하는 기업들을 선별하고 여기에 과학기술정보통신부 AI국 수탁 기업 명부를 활용하여 해당업체를 추가하여 조사모집단을 최종적으로 확정하는 데 이러한 방식으로 수집된 조사모집단이 인공지능산업 전체를 포괄하는 데는 한계가 있는 것으로 인식하고 있었다.

나. 주요 개선의견

(1) 인공지능산업 고유 분류체계 정립

인공지능산업은 첨단 신사업이므로 향후 어디까지 적용 범위를 확대하여갈지 현재로서는 예측하기가 어렵다. 이러한 이유로 기존의 산업분류방식을 기반으로 인공지능 산업을 세분화한다면 실제 인공지능 사업을 영위하고 있지만 분류 기준의 한계로 빠지는 기업이 존재할 수 있다. 현재도 서비스업, 제조업 등에서도 인공지능사업을 할 수 있는데 이러한 업체가 모두 조사모집단에 포함되고 있지는 못하고 있다. 이러한 문제를 해결하기 위해서는 기존의 분류체계를 바탕으로는 인공지능산업 전체를 포괄할 수 있는 분류기준을 새롭게 검토하여 정립할 필요가 있을 것으로 사료된다.

(2) 대표성 있는 조사모집단 리스트 구축

인공지능산업 전체를 포괄하는 조사모집단 리스트를 구축하여야 할 것이다. 이를 위해 인공지능산업 분류체계가 산업 전체를 반영할 수 있게 재정립되어야 하며 이를 근거로 조사모집단 리스트가 수집되어야 할 것으로 판단된다.

(3) 마이크로데이터 서비스 제공

마이크로데이터를 이용자에게 제공하여 연구자들이 인공지능산업에 대해 보다 심층적으로 분석할 수 있게 하여야 할 것이다. 마이크로데이터 제공을 위해서는 자체 플랫폼에서 제공할 것인가 아니면 통계청의 마이크로데이터 플랫폼을 활용할 것인가를 내부적으로 판단하여 결정하면 될 것으로 사료된다.

붙임5

마이크로데이터 품질 점검 결과

통 계 명	인공지능산업실태 조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
연 구 원	조준기
연구보조원	박연진, 최다빈

제1부 **점검 개요**

I. 점검 개요

- 마이크로데이터 품질 점검 시 검토한 자료
 - 통계정보보고서(통계작성 기획, 통계설계, 통계처리 및 분석)
 - 통계보고서
 - 조사표, 항목 및 코드집, 가중치
 - 통계승인사항

- 마이크로데이터 품질 점검 내용
 - 관리 주체, 마이크로데이터 메타자료 현황 점검
 - 공표자료와 마이크로데이터 집계치의 일치율 점검

II. 마이크로데이터 개요

조 사 명	인공지능산업실태조사	
작 성 기 관 명	과학기술정보통신부	
작 성 주 기	1년	
작성기준년도	2019년	
전수/표본조사	전 수 (●)	표 본 ()
조 사 대 상	○ ICT통합모집단을 대상으로 모집단 정비조사를 통해 확인된 1,365개의 인공지능 기업체	
주 요 조 사 항 목	○ 일반현황: 기업명, 대표자 성별, 설립연도, 인공지능 시작연도 등 ○ 기술 및 사업 현황: 보유 기술 분야, 주력 기술 분야 등 ○ 매출 현황: 최근 3개년 전체 매출액, 인공지능 분야 매출 발생 여부 등 ○ 인력 현황: 전체 종사자수, 인공지능 부문 종사자수, 인공지능 여성 종사자수 등 ○ 투자 및 특허 현황: 최근 3개년 기업 전체 연구개발 투자액 등 ○ 애로사항 및 건의사항: 부문별 애로사항 정도	

제2부 점검 결과 요약

구 분	점검결과	개선 의견	비 고
마이크로데이터 생성·관리 현황	- 작성기관, 위탁기관, 용역기관에서 생성·관리하고 있는 것으로 확인됨	-	
마이크로데이터 서비스 현황	- 마이크로데이터를 제공하고 있지 않음	- MDIS(통계청)를 통해 마이크로데이터 제공 검토	- MDIS 서비스 예정
마이크로데이터 관련 메타자료 보유 현황	- 가중치를 포함한 마이크로데이터, 조사표, 코드집 및 파일설계서, 공표용 보고서 등의 메타자료를 보유하고 있는 것으로 확인됨	-	
일치율	- 보고서 통계표와 마이크로데이터 간의 일치 여부 점검 결과 65개의 통계표 중 65개(100%)의 통계표가 일치함	-	

제3부 마이크로데이터 품질 점검 결과

1. 점검 개요

마이크로데이터 품질 점검은 통계작성기관이 보유 및 관리하고 있는 마이크로데이터 및 관련 메타자료를 제공받아 기초점검 및 실질점검(일치율 점검)을 실시하였다.

기초점검은 관리기관 적합성과 메타자료 적정성(누락자료, 파일형태, 주요항목의 이상여부)을 점검하며, 실질점검은 현재 공표된 보고서와의 수치비교를 통하여 마이크로데이터 정합성을 점검하는 것이다.

점검결과는 관리기관 적합성, 메타자료 적정성에 대하여 점검 의견으로 정리하였고, 마이크로데이터 오류에 대한 원인을 분석하였다. 그리고 마이크로데이터 품질 점검 과정에서 도출된 문제점 및 개선요구사항 등을 종합하여 정리 및 분석하였다.

인공지능산업실태조사는 1년 주기로 조사가 이루어지며 조사 결과는 조사기준년도 익익년도 3월에 공표하고 있다. 최신 공표 자료는 2020년도 조사자료(2022년 4월 공표)이나 마이크로데이터 품질 점검 일정에 의해 2019년도 조사자료를 점검하였다.

2. 점검 결과

(1) 마이크로데이터 생성·관리 현황

인공지능산업실태조사 마이크로데이터는 작성기관인 과학기술정보통신부와 위탁기관인 소프트웨어정책연구소 그리고 용역기관에서 1년주기로 생성하고, 관리하는 것으로 확인하였다.

인공지능산업실태조사는 원자료에서 조사요류, 입력오류 및 논리오류 등에 대한 검토를 실시하며 최종 마이크로데이터 조건을 고려하여 생성하며, 마이크로데이터는 통계청 마이크로데이터과에서 점검 완료 후 MDIS를 통해 표준화된 형태로 보존하여 관리하는 것으로 나타났다.

(2) 마이크로데이터 서비스 현황

인공지능산업실태조사 마이크로데이터는 현재 서비스되지 않으나, 통계청 MDIS를 통해 2019~2020년도 기준 마이크로데이터를 제공 예정인 것으로 나타났다.

향후 인공지능산업실태조사는 엑셀 형태의 마이크로데이터, 항목 및 코드집, 한글 형태의 조사표, 공표용 보고서, 조사지침서, 내검규칙을 이메일로 제공할 예정인 것으로 확인되었다.

(3) 마이크로데이터 관련 메타자료 보유 현황

인공지능산업실태조사의 경우 조사표, 공표용 보고서 등의 메타자료를 보유하고 있는 것으로 확인하였다. 공표용 보고서는 소프트웨어정책연구소 홈페이지를 통해 접근·활용할 수 있다.

(4) 일치율

인공지능산업실태조사는 보고서와 마이크로데이터를 통해 재현한 통계표 간 일치율 점검 결과, 65개의 통계표 중 65개(100%)의 통계표가 모두 일치하는 것으로 나타났다.

<일치율 점검 결과>

계	점검 집계표 수(개)		일치율(%)
	일치 수	불일치 수	
65	65*	0	100

*소수점 차이 포함(0.2 이하)

3. 주요 개선의견

(1) MDIS(통계청)를 통해 마이크로데이터 제공 검토

통계의 접근성을 높이기 위해 자체 제공뿐만 아니라 통계청 MDIS를 통한 마이크로데이터 서비스 활용을 검토할 필요가 있다. 마이크로데이터 서비스 시 이용자들은 정책반영, 연구목적, 현황파악 등 목적에 따라 다양하게 활용할 수 있어 이용자 만족도를 높일 수 있다.

※ [참고] 점검 집계표 일치 여부

<점검 집계표별 일치 여부>

구분	통계표명	일치여부
	대표자 성별	일치
	설립연도	일치
	인공지능 사업 시작연도	일치
	사업체 형태	일치
	상장 여부	일치
	기업부설 연구소 유무	일치
	AI 주사업 여부	일치
	보유 중인 인공지능 기술 분야 (1순위)	일치
	보유 중인 인공지능 기술 분야 (1+2순위)	일치
	인공지능 부문 주력 사업 분야 (1순위)	일치
	인공지능 부문 주력 사업 분야 (1+2순위)	일치
	인공지능 사업화 준비 단계	일치
	대표 인공지능 제품·서비스의 응용 산업 분야(중복응답)	일치
	AI 모델 개발에 있어 사용하는 도구 형태 비중	일치
	인공지능 계산을 위한 GPU 보유 현황	일치
	인공지능 계산을 위한 GPU 수요 규모	일치
	인공지능 R&D를 위한 향후 컴퓨팅 인프라 형태 비중	일치
	인공지능 학습용 데이터 주 확보 방식 (1순위)	일치
	인공지능 학습용 데이터 주 확보 방식 (1+2순위)	일치
보고서 (65개)	향후 3년간('20~'22년) 가장 유망할 것 같은 인공지능 응용 산업 분야	일치
	3개년도(2018년~2020년) 전체 매출액 현황	일치
	3개년도(2018년~2020년) 인공지능 매출 발생 여부	일치
	3개년도(2018년~2020년) 인공지능 매출액	일치
	3개년도(2018년~2020년) 인공지능 사업 분야별 매출액-인공지능 시스템 소프트웨어	일치
	3개년도(2018년~2020년) 인공지능 사업 분야별 매출액-인공지능 응용 소프트웨어	일치
	3개년도(2018년~2020년) 인공지능 사업 고객별 매출액-공공(B2G)	일치
	3개년도(2018년~2020년) 인공지능 사업 고객별 매출액-민간(B2B)	일치
	3개년도(2018년~2020년) 인공지능 사업 고객별 매출액-개인(B2C)	일치
	최근 3개년 전체 종사자 - 2018년	일치
	최근 3개년 전체 종사자 - 2019년	일치
	최근 3개년 전체 종사자 - 2020년(P)	일치
	3개년도(2018년~2020년) 인공지능 분야 종사자	일치
	인공지능 특허 현황 - 출원기준	일치
	인공지능 특허 현황 - 등록기준	일치
	전체 특허 현황 - 등록기준	일치
	전체 특허 현황 - 출원기준	일치
	3개년도(2018년~2020년) 인공지능 부문 연구개발 투자액	일치
	3개년도(2018년~2020년) 외부투자 유치 실적 유무	일치
	인공지능 사업 운영상 느끼는 애로사항 - 전체 항목 비교	일치
	인공지능 사업 운영상 느끼는 애로사항 - ① 데이터 확보 및 품질 문제	일치

*점검한 총 65개 중 유형별 일부 집계표만 나열함

붙임6

공표자료 오류 점검 결과

통 계 명	인공지능산업실태조사
승 인 번 호	127016
작 성 기 관	과학기술정보통신부
연 구 원	황규배
연구보조원	-

제1부 점검 결과 요약

1. KOSIS 통계표 점검

- 기준자료명: 2021년 인공지능산업실태조사 집계표
- 점검자료명: 인공지능산업실태조사 KOSIS 데이터
- 작성기준년도: 2019년

통계표명	점검결과	개선의견	반영여부
보유중인 인공지능기술 중 주력분야	통계표 불명확 (항목을 분류로 구성)	통계표 재구성 (순위별>순위별,기술분야별)	반영
인공지능부문 주력사업분야	통계표 불명확 (분류와 항목이 항목으로 구성)	통계표재구성 (순위별>순위별,사업분야별)	반영
인공지능 매출발생여부	원본 표현 오류	‘평균’, ‘합계’를 ‘있음’, ‘없음’으로 수정	반영
인공지능 분야 종사자	수치 불일치	2019년 수치, 2020년 보고서 기준 확정치로 수정	반영
학력별 인공지능 인력현황	통계표 불명확 원본 오류(전체 수치 생략)	통계표 수정 원본오류(전체 추정치부재) KOSIS에 보정	반영
인공지능 관련 인력 채용 시 주요 고려사항	통계표 불명확 원본 오류(전체 수치 생략)	통계표재구성 원본오류(전체 추정치부재) KOSIS에 보정	반영

제2부 공표자료 오류 점검 결과

1. 점검 개요

「통계정보보고서」의 공표 관련 내용을 검토하고, 국가통계포털(KOSIS) 공표자료 유무와 국제기구에 자료를 제공하는지 파악한다. 진단대상 통계의 기준자료(점검 시점을 기준으로 가장 최근에 발간된 보도자료, 통계보고서 등의 통계간행물 또는 통계표 입력 시 사용한 원본보고서)를 지정하고, KOSIS 통계표와 국제기구 자료를 대상으로 아래의 사항들을 점검한다.

(1) 통계표 형식 및 내용 점검

기준자료와 KOSIS 통계표의 형식 및 내용, 용어, 단위, 주석, 출처, 항목명 등을 점검한다.

(2) 통계표 수치자료 점검

기준자료와 KOSIS 통계표에 수록된 내용을 비교하여 수치를 점검한다. 단순오류나 오타뿐만 아니라 과거 시계열, 다른 통계표 등과 비교하여 논리적 타당성을 점검한다.

(3) 국제기구 자료 제공 일치 여부 점검

OECD, ITU, ILO, UN 등 국제기구에 통계자료를 제공하는 경우 국제기구 요구자료 및 제공현황을 파악하고, 국제기구에 제출한 자료와 국제기구의 간행물이나 DB 등에 서비스되는 자료의 일치 여부를 비교하고 그 원인을 파악한다.

2. 점검 결과

(1) 통계표 형식 및 내용 점검

점검 전 KOSIS에 제공된 통계표 중 전체수치(전체 추정 값)가 제시되지 않은 통계표가 있었는데 점검 결과 이러한 문제가 적절하게 수정되어 있었다. 일부 통계표의 경우는, 통계표 구성이 불명확한 것이 있었는데 점검결과 통계표를 재구성하여 이러한 문제가 개선된 것으로 확인되었다.

(2) 통계표 수치자료 점검

진단대상이 되는 기준연도 2019년 통계수치는 점검결과 정확한 것으로 확인되었다.

(3) 국제기구 자료 제공 일치 여부 점검

인공지능산업실태조사는 국제기구에 통계자료를 제공하지 않는 것으로 확인하여, 점검에서 제외하였다.

부 록. 통계품질진단 개요

1. 통계품질진단의 개념

현대적 의미의 통계품질은 ‘통계가 이용자에게 얼마나 이용하기 적합하게 작성 및 제공되고 있는가를 나타내는 특성’으로서 통계품질관리는 ‘통계이용자들에게 통계를 사용하는데 적합하도록 생산하는 방법뿐만 아니라 이용자에게 만족을 주면서 가장 경제적인 방법으로 통계를 작성·보급·관리하기 위한 모든 수단을 통합하는 체계’를 말한다.

따라서, 통계품질진단이란 생산된 통계가 이용자에게 얼마나 유용하게 사용되고 있는지를 살펴보는 과정으로서 국가 정책 결정의 기초 자료로 이용되는 국가승인통계에 대한 품질수준을 진단하여 국가통계의 품질 향상 및 신뢰도 제고를 목적으로 한다.

통계청에서는 통계품질의 수준을 관련성, 정확성, 시의성/정시성, 비교성/일관성, 접근성/명확성이라는 5가지 차원으로 정의하고 있으며, 통계품질진단은 5가지 차원의 품질수준이 어느 정도인지를 측정하고 각 차원의 품질수준을 높이기 위해 통계를 어떻게 개선해야 하는지 그 방향을 제시하고자 하는 것이다.

또한, 통계청이 제시한 통계품질진단의 과정은 첫째, 통계정보보고서를 활용한 품질진단, 둘째, 자료수집 체계 점검, 셋째, 조사표 설계 및 유사통계 비교·분석 점검, 넷째, 표본설계 점검, 다섯째, 이용자 요구사항 반영실태 진단, 여섯째, 마이크로데이터 품질 점검, 일곱째, 공표자료 오류 점검으로 이루어지며, 이러한 과정을 통해 통계생산과정에 대한 품질관리에 기초한 보다 정확하고 신뢰성이 높은 우수한 통계를 생산함과 동시에 이렇게 생산된 통계가 향후 이용자의 요구를 충족시킬 수 있도록 하는데 통계품질진단의 필요성과 궁극적인 목적이 있다.

2. 통계품질진단 체계

가. 통계정보보고서 작성

통계의 중요성이 강조되고 이용이 활성화되면서 통계자료와 함께 해당 통계의 작성 방법 등의 정보 요구도 높아졌다. 그 동안의 품질진단에서는 통계 작성 절차에 따른 양적·질적 정보를 「통계정보보고서」로 작성하여 통계 이용자에게 제공하였다. 또한, 통계생산자가 통계생산의 기반자료로 활용하여 절차적 품질 수준을 향상하도록 하였다.

이에 새롭게 생산된 통계도 이용자용 가이드이자 생산자용 편람으로 사용하기 위한 「통계정보보고서」를 작성하여야 하며, 지속적으로 생산하는 통계는 기존에 작성된 「통계정보보고서」를 보완하여 활용하여야 한다.

나. 통계정보보고서 활용 진단

이용자의 정확한 이해와 활용, 통계제반과정 및 산출물에 대한 정보 등 각 과정에 대한 품질정보 제공을 위한 통계정보보고서는 총 6장으로 구성되어 있다. 진단에서는 「통계정보보고서」에 수록되어 있는 6개의 작성절차별로 품질지표를 구성하여 통계의 품질수준을 측정하며, 기본적인 통계작성절차를 준수하는지 여부도 점검한다.

(1) 제1장 통계작성기획

통계 이용자의 입장에서 통계의 특성과 필요성 등 핵심적인 내용이 통계 개요에 수록되어 있는지 점검하고, 통계작성절차 전반에 대하여 진단한 결과를 작성한다. 또한 통계에 대한 작성목적이 명확한지, 통계의 주된 활용 분야가

무엇인지 등을 진단하고, 통계를 이용하는 이용자에 대한 관리 및 의견수렴 등에 대한 점검 결과를 기반으로 진단결과를 작성한다.

(2) 제2장 통계설계

통계는 작성목적에 맞게 조사내용 및 조사표를 설계하여야 하며, 응답자에게 조사목적에 부합하는 정보를 얻기 위해 노력하여야 한다. 이를 위해 응답자가 쉽게 응답할 수 있도록 용어나 분류 기준 등을 국내 또는 국제기준을 적용하는지 점검하고, 조사표의 기본 구성요소에 대한 수록 여부 등을 진단한다. 또한, 통계는 시대가 변함에 따라 진화하고 발전하여야 한다. 이에 따라, 조사표의 변경이력 등이 관리되고 있는지 진단한다.

또한, 조사를 위해서는 모집단과 표본추출틀에 대한 정의가 명확하게 설정되어야 하며, 특히 표본 조사의 경우 표본설계 및 모집단과 표본추출틀의 주기적인 갱신 등을 검토하고 진단결과를 작성한다.

(3) 제3장 자료수집

통계를 작성하기 위해서는 조사표를 이용하여 응답자로부터 응답을 받아내는 것이 가장 중요한 작업이다. 시대가 변함에 따라 자료를 수집하는 방식도 변화하고 있으며, 응답률 등을 고려하여 다양한 방식으로 조사를 실시하고 있다.

특히, 면접조사의 경우, 조사원의 채용 및 교육 등은 조사의 성공 여부를 좌우할 정도로 중요하다. 조사를 위한 업무, 조사준비, 홍보, 명부보완 등을 체계적으로 관리하고 있는지를 진단하고, 현장에서 발생할 수 있는 문제에 대한 관리방안 등이 마련되어 있는지도 진단한다. 그리고 무응답이 발생한 경우, 적절한 대체 방법이 강구되어 있는지를 점검하고, 사후조사 실시 여부 및 결과 조치방안을 확인한다. 위의 사항을 종합적으로 검토하여 진단결과를 작성한다.

또한 조사환경이 열악해짐에 따라 행정자료를 활용하여 다양한 방식으로 조사 자료의 보완 및 점검을 실시하고 있다. 이에 통계에 활용하는 행정자료의 활용 목적 및 내용, 특성 등을 파악하여 본 통계작성에 활용하는지에 대해 검토하고 진단 결과를 작성한다.

(4) 제4장 통계처리 및 분석

수집된 자료를 시스템적으로 검토하고 작성하기 위해, 코딩 및 코드체계 등이 정립되어 있는지와 입력된 자료를 기반으로 자료를 내검하는 방식과 무응답의 유형에 따른 실태 등을 점검한다. 수집된 자료 중 행정자료를 활용하는 경우, 행정자료의 매칭방법 등을 대해 검토하고 진단 결과를 작성한다. 즉, 통계로 작성되기 위해 사용되는 자료의 처리과정 전반에 대하여 점검한 후 진단 결과를 작성한다.

수집된 자료에 대한 기본적인 정제작업이 완료되면, 이것을 기반으로 통계를 추정하고 분석하게 된다. 통계추정을 위해선 표본설계 당시와 동일하게 조사되지 못한 부분을 가중치 조정 등을 통해 추정을 실시하고, 주요 항목들에 대한 변동계수 등이 기획의도와 동일하게 도출되고 있는지 등을 검토한다.

특히 지수를 작성하는 통계의 경우, 지수 유형 및 산식 등을 점검하고 개편 여부 등을 점검한다. 또한, 계절조정이 필요한 통계의 경우, 계절조정과정 및 내용에 대하여 점검한다. 이 모든 과정에 대하여 점검하고 진단결과를 작성한다.

(5) 제5장 통계공표, 관리 및 이용자서비스

통계가 작성되면 그 통계결과를 공표하여 이용자가 유용하게 활용할 수 있도록 해야 하고, 이용에 혼란을 줄 수 있는 사항은 사전에 공지하여 이용에 어려움이 없도록 조치하여야 한다. 따라서 공표일정, 통계설명자료 제공현황, 마이크로데이터 제공현황, 비밀보호 및 보안사항 등을 점검하고 진단결과를

작성한다. 또한 통계작성방법 유지, 시계열 단절 여부 등과 동일영역 통계와의 일관성 등도 점검하고 진단결과를 작성한다.

(6) 제6장 통계기반 및 개선

통계를 작성하는 환경에 대한 진단 또한 통계의 품질에 직접적인 영향을 미친다. 통계를 기획하고 분석하는 인력 현황과 위탁에 의해 작성되는 경우, 통계청에서 제시한 통계조사 민간위탁 지침의 준수여부와 통계품질향상을 위한 노력 등을 점검하고 진단결과를 작성한다.

다. 자료수집 체계 점검

자료수집 체계 점검은 조사기획자, 조사관리자, 조사원 등 자료수집 과정에 직접적으로 관여하는 사람들을 대상으로 자료가 정확히 수집되었는지, 절차적 오류는 없는지 등을 점검한다. 특히, 자료수집 과정에서 나타날 수 있는 자료수집 오류의 가능성을 체계적으로 점검하고, 발생한 또는 발생 가능한 문제점을 찾아 개선방안을 도출하여 자료수집 과정에서의 품질을 개선하려는 과정이다.

라. 표본설계 점검

표본설계 점검에서는 진단통계의 모집단, 표본추출틀, 표본추출방법, 목표오차, 표본규모, 가중치, 추정식, 주요 항목별 공표 범위 등 표본설계와 관련한 일련의 과정을 정밀 검토하여, 모집단을 잘 대표하는 통계자료가 생산되고 있는지 점검한다.

마. 조사표 설계 및 유사통계 비교·분석 점검

조사표 설계 점검에서는 주요 용어 및 항목별 정의, 조사표 구성, 조사표 설계 및 변경 절차, 설문응답 지시문, 응답보기의 포괄성·상호배타성을 만족하는지 점검한다. 그 다음 각 항목별 기준시점에 일관성, 조사표 변경 이력, 조사항목별 작성요령 및 유의사항을 점검한다.

유사통계 비교·분석 점검은 공표하고 있는 통계 중 동일하거나 유사한 통계가 있는지 검토한다. 점검통계와 유사한 항목이 있는 통계간의 작성기관, 작성목적, 작성대상 및 범위, 작성단위, 작성주기, 기준시점, 공표시기, 표본조사 여부, 작성규모를 비교하고 유사항목의 결과값 및 추이가 유사한지 점검한다.

바. 이용자 요구사항 반영실태 진단

통계 이용자는 이용하는 통계로부터 기대하는 정보를 충분히 얻기 원하므로, 품질이 우수한 통계는 이용자가 원하는 정보를 많이 제공할 수 있어야 한다. 따라서 통계 이용자가 해당 통계자료에 대해 얼마나 만족하는지를 살펴보는 것이 필요하다. 이를 위해 진단 대상통계와 관련하여 정책수립 및 평가, 학술연구 등에 직접 활용한 경험이 있는 전문 또는 일반이용자로 구성된 이용자 요구사항 반영실태 진단(FGI)을 실시하여 통계이용자의 통계에 대한 만족 수준과 요구사항 반영수준이 충분히 반영되는지를 진단한다.

사. 마이크로데이터 품질 점검

이용자의 유용한 마이크로데이터 활용을 위하여 충분한 메타데이터(파일설계서, 코드북 등) 및 정확한 마이크로데이터 제공이 필요하다. 이를 위해 마이크로데이터 품질 점검에서는 데이터의 정확성 진단을 목적으로 마이크로데이터 관리체계 및 메타자료 점검, KOSIS 공표항목 기준 집계표 일치율을 점검한다.

아. 공표자료 오류 점검

작성절차에서는 오류가 없는 통계일지라도 공표되는 과정에서 오류가 발생한다면 통계품질을 떠나 잘못된 통계를 사용하게 된다. 공표자료 오류 점검에서는 통계서비스의 질을 향상시키기 위해 KOSIS에 제공되는 통계표에 대한 수치, 단위표기, 주석 등을 점검하고, 국제기구 제공 통계의 경우에는 기관에서 제공한 수치와 국제기구에서 보고서 및 DB를 통해 발표한 수치를 상호비교하여 불일치한 수치 유무를 점검한다.

3. 통계품질 수준 측정

(1) 관련성

관련성이란 이용자 관점에 초점을 둔 측면으로 통계의 포괄범위와 개념, 내용 등이 이용자 요구에 부합되는 정도를 의미한다. 즉, 통계이용자에게 얼마나 의미 있고 유용한 통계를 작성하여 제공하고 있는가와 관련된 개념이다. 여기서는 통계의 작성목적에 명확히 설정하고 이를 달성하기 위하여 이용자 파악, 전문가 자문회의, 이용자 만족도 조사 등 이용자 요구를 지속적으로 파악하여 통계에 반영하고 있는지와 관련한 사항을 중심으로 점검한다.

(2) 정확성

정확성이란 측정하고자 하는 모집단의 특성을 추정함에 있어 이 추정된 값이 미지의 참값에 얼마나 근접하는가의 정도를 의미한다. 정확성과 관련한 품질진단에서는 표본설계, 표본오차, 비표본오차, 자료수집방법, 면접소요시간 등을 중심으로 발생 가능한 표본오차 및 비표본오차의 크기와 발생원인 등을 탐색하고 오차를 최소화하기 위한 방안을 마련하고 있는지를 점검한다.

(3) 시의성 및 정시성

시의성은 작성기준시점과 결과공표시점간의 차이를 나타내는 통계의 현실 반영도와 관련된 개념으로서 작성기준시점과 결과발표시점이 근접할수록 시의성이 높은 통계이다.

정시성은 공표한 날짜와 사전에 계획된 공표 날짜 사이의 시간 지체 정도를 나타내며, 예고된 공표시기를 정확히 준수하는가에 대한 개념이다. 여기서는 통계작성주기, 작성기준시점과 공표일까지의 소요기간, 공표예정일과 실제공표일의 차이, 공표지연 사유 등을 중심으로 점검한다.

(4) 비교성 및 일관성

비교성은 시간 흐름과 영역에 따라 비교되는 정도를 의미한다. 즉, 시간이나 공간이 달라도 통계자료가 공통된 기준(통계개념, 측정도구, 측정과정 및 기초자료)으로 집계되어 서로 비교 가능한지를 진단하는 차원이다. 따라서 비교성에서는 지리적 및 비지리적 영역 또는 시간적 통계를 비교할 때 통계작성에 적용된 개념, 정의와 측정방법의 차이가 주는 영향 등을 중심으로 점검한다.

일관성이란 동일한 경제·사회현상에 대해 서로 다른 기초자료나 작성방법, 작성주기(공표주기)에 의해 작성된 통계자료들이 서로 얼마나 유사성을 지니는가에 대한 정도를 의미한다. 따라서 서로 다른 기초자료나 작성방법에 의해 작성되었더라도 동일한 현상을 반영하는 통계자료들은 서로 유사한 결과를 보여야 한다. 일관성에서는 잠정자료와 확정자료, 연간자료와 분기(월) 자료를 비교한 내적일관성 여부와 다른 통계자료와 유사한 결과를 보이는지 비교한 결과 등을 중심으로 점검한다.

- * 비교성과 일관성은 유사한 개념이다. 일관성은 통계 간 결과가 유사한지 보는 것이고, 비교성은 통계에서 사용한 개념, 분류, 기준 등이 유사하여 비교가능한지를 보는 것이다.

(5) 접근성 및 명확성

접근성은 이용자가 통계자료에 대해 손쉽게 접근할 수 있는 정도를 말하며, 명확성은 통계가 어떻게 만들어졌는지에 대한 정보제공 수준을 말한다. 통계자료의 데이터베이스화, 간행물 및 보도자료 홈페이지 게시, SNS를 통한 속보 전송 등 다양한 방법으로 자료를 제공하고 이용자의 검색이 용이하도록 하는 것은 통계의 접근성을 높이는 활동이다. 여기서는 이용자들이 통계자료를 쉽게 이용할 수 있도록 이용자 친화적인 절차로 통계정보를 제공하고 있는지, 이용자를 위한 적절한 정보와 지원을 하고 있는지 등을 중심으로 점검한다.

2022년 정기통계품질진단 진단결과보고서

발 행 일 2022년 12월
발 행 인 통계청장 한훈
발 행 처 통계청 통계정책국 품질관리과
 대전시 서구 청사로 189
인 쇄 처 위드나래



안 내

1. 연구보고서의 내용을 발표 또는 인용할 때에는 반드시 올바른 인용 및 출처표시 방법을 준수해야 합니다.
2. 연구보고서의 지식재산권은 통계청에 있습니다.