

2026-09

글로벌 로봇산업 정책·산업동향



KIRIA 로봇산업 정책동향 

로봇산업 정책동향

중국의 체화지능 훈련·실증 인프라
- 로봇 데이터·훈련장·운영체계 중심



한국로봇산업진흥원

KOREA INSTITUTE FOR ROBOT INDUSTRY ADVANCEMENT

로봇산업 정책동향

Contents

I. 체화지능 훈련 인프라의 형성과 분석 범위	1
II. 중국의 구축 현황과 운영모델	4
III. 로봇 훈련데이터의 생산과 활용	9
IV. 실제환경 검증과 표준화	14
V. 지속가능한 운영과 발전방향	19
VI. 중국 체화지능 훈련장의 주요 변화	23
주요참고자료	24

요약

- ✦ 중국의 체화지능(具身智能) 훈련장은 데이터를 모으는 시설을 넘어 모델 개선과 배치 전 검증을 연결하는 산업 기반으로 역할이 넓어지는 단계

 - ▶ 실제 물리적 상호작용 데이터는 인터넷 자료만으로 확보하기 어려워 별도의 데이터 생산체계가 필요하며, 수집한 데이터를 학습에 활용할 수 있는 형태로 전환하는 역량도 중요
 - ▶ 중국정보통신연구원(CAICT) 등이 발간한 '26년 8월 보고서는 훈련장 확산과 함께 훈련 시나리오·과제의 동질화, 데이터 처리 병목, 거래 불확실성, 지속가능한 사업모델 미형성을 주요 과제로 진단
- ✦ 중앙정부는 「AI+제조」, 2026년도 실경실훈(实景实训) 특별행동, 체화지능 중간시험(中试) 기지, 업종 고품질 데이터셋 정책을 통해 훈련·검증 인프라와 데이터 공급을 산업정책에 연결

* 실경실훈(实景实训): 실제 업무환경에서의 훈련·검증

 - ▶ '26년 6월 말 가동 시설은 70곳 초과, 건설·계획 시설은 46곳이며, 18개 성급지역에서 40건이 넘는 관련 지방정책 확인
 - ▶ 훈련장은 운영 목적에 따라 종합형과 산업 특화형으로 나뉘며, 전자는 여러 산업의 데이터와 범용 역량을, 후자는 특정 업무의 심층 훈련을 지향하고 정부 유도형·기업 자체 구축형과는 별도 분류축을 구성
- ✦ 분야별 적용여건을 보면 제조업은 작업경계와 경제성을 확인하기 쉬워 초기 적용에 유리하지만, 가정·의료·특수환경은 작업경계와 안전·환경 재현 조건이 복잡해 검증 부담이 커지는 구조

 - ▶ 공개된 훈련장 중 제조 시나리오를 포함한 비중은 86%이며, 한 시설이 여러 산업 시나리오를 포함할 수 있어 산업별 비중은 시장점유율과 구별
 - ▶ 생활서비스의 공간·물체 다양성과 의료·응급의 안전 요구는 시나리오 추가만으로 해결되지 않으며, 실제 작업을 대표할 환경 설계가 관건
- ✦ 실물 로봇 없이 수집한 사람의 행동기록과 실물 로봇의 실행기록은 비용·확장성·물리적 일치성에서 상호 보완적 역할을 수행

- ▶ 1인칭 영상·말단 조작·동작 포착 자료에는 로봇에 맞춘 변환이 필요하고, 원격조작 자료는 본체의 상태·동작을 반영하지만 장비·인력 부담이 큰 편
- ▶ 현장 배치 후 실패·개입·복구 기록은 모델의 약점을 드러내는 자원이지만, 현재 배치된 로봇 수만으로 대량 공급을 기대하기에는 한계
- ◆ 훈련장의 경쟁력은 원시 데이터를 정제·정렬·검수한 뒤 실제 로봇의 작업 개선으로 연결하는 전 과정의 일관성에서 형성
- ▶ 시간축·좌표계·작업단계가 다른 자료는 정렬 없이 합산하면 활용성이 낮아지므로, 생산부터 모델 검증까지 연결해 수집 목적을 명확히 할 필요
- ▶ 센서 오류로 사용할 수 없는 기록과 정상적으로 기록된 작업 실패는 구분해야 하며, 후자는 복구 전략과 예외상황 학습을 위한 보존 대상
- ◆ 실제환경 검증은 성공 동작을 보여주는 시연보다 실패 원인과 적용 경계를 확인하는 과정이며, 고객 현장의 위험을 사전에 줄이는 데 의의
- ▶ '26년도 실경실훈 특별행동은 사용자와 기술공급자의 공동 참여를 통해 작업 수요, 훈련, 응용검증과 상시배치를 연결하는 정책 방향을 제시
- ◆ 훈련장의 중장기 지속성은 데이터 판매보다 훈련·검증 서비스와 고객 수요를 결합한 사업모델 확보에 좌우
- ▶ 대규모 종합시설의 건설·운영 부담은 적지 않으며, 공개된 생산능력이나 판매가격만으로 가동률·판매율·수익성 추정은 부적절
- ▶ 3단계 발전 구상은 일률적인 성장 시간표가 아니며, 일부 시설은 전 과정 운영보다 수집 도구나 시험서비스에 특화될 전망
- ◆ 표준화는 훈련장 평가와 데이터의 호환성·품질을 연결하는 방향으로 구체화되나, 연구 제안 단계, 제정 진행 중인 표준, 의견수렴 중인 초안은 서로 구분해 해석할 필요
- ▶ 국가 휴머노이드 로봇 산업 표준체계 의견수렴안은 데이터·모델·시험평가·응용·안전을 포괄하며 '28년 핵심표준 100건 이상을 목표로 제시

I

체화지능 훈련 인프라의 형성과 분석 범위

1 물리적 행동 데이터의 공급 공백

- ✦ 체화지능(具身智能)은 환경을 인식하는 데 그치지 않고 물리적 행동과 그 결과를 학습해야 하므로, 기존 인공지능과 다른 데이터 생산 기반 필요
 - 언어·영상 자료에는 사물과 상황에 대한 지식이 축적돼 있지만, 로봇 관절의 움직임과 힘·촉각 변화가 작업 결과에 미친 영향은 기록이 불충분
 - 모델이 물체를 식별하더라도 이를 어느 방향에서 잡고 어느 정도의 힘으로 이동할지는 별도의 문제이므로, 인식 성능과 실제 작업능력을 구분할 필요
- ✦ 중국정보통신연구원(CAICT) 공동보고서는 고품질 실제 데이터의 부족을 체화지능 발전의 가장 큰 제약으로 보면서도, 데이터·모델·연산 규모 확장의 효과를 전제조건과 함께 설명
 - 체화지능의 발전경로는 데이터·모델·연산 규모를 늘리면 성능이 향상된다는 가정에 기초하며, 모든 작업에서 동일한 법칙이 검증됐다는 단정은 배제
 - 글로벌 실제 데이터 수십만~백만 시간 수준은 개략적 규모 서술로, 통일된 기준의 전수통계와는 성격이 다름
- ✦ 물리적 작업기록은 환경과 본체의 특성이 함께 포함되므로, 데이터 양을 늘리는 것과 여러 작업에 재사용할 수 있게 만드는 것은 별개의 과제
 - 같은 집기 작업이라도 카메라 시점, 관절 구성, 그리퍼 형태가 달라지면 행동 표현과 센서 기록이 달라져 다른 로봇에 이전하려면 추가 처리가 필요
 - 훈련 목적에 적합한 작업 분포와 품질을 확보하지 못하면 단순 반복수집이 늘어도 새로운 환경에 대한 적응력이 그만큼 향상된다고 보기에는 근거 부족
- ✦ 따라서 훈련장 분석의 출발점은 시설의 크기가 아니라 어떤 데이터 공백을 해소하고 그 성과를 어떻게 작업능력으로 확인하는지의 문제
 - 생산체계와 검증체계를 함께 살펴야 설비 투자, 데이터 처리비용, 모델 개선과 실제 활용 사이의 관계를 파악할 수 있으며, 이는 후속 장의 공통 분석축

2 분산 수집·훈련장·현장 환류의 관계

- ✦ 실제 데이터의 공급경로는 분산 수집, 집중형 훈련장, 현장 배치 후 환류로 정리할 수 있으며 세 경로는 서로 대체하기보다 서로의 제약을 보완
- 분산 수집은 사람이 생활하거나 일하는 여러 장소를 활용하므로 공간과 물체의 다양성을 확보하기 쉽고, 별도 시설 구축 부담도 상대적으로 낮은 수준
 - 반면 참여자의 숙련도와 작업 해석이 달라 동일 과제의 실행방식이 흔들릴 수 있어, 다양한 환경을 확보하는 이점과 품질관리 부담이 동시에 발생

표 1. 실제 데이터 공급경로의 기능 분담

구분	주된 강점	관리해야 할 제약
분산 수집	다양한 생활·작업환경 접근	참여자 교육, 작업 해석과 기록 품질
집중형 훈련장	동일 조건의 반복과 실패 재현	환경 대표성, 설비·운영비 부담
배치 현장 환류	실제 운영의 실패·복구 정보	배치 규모, 고객 안전·정보보호

자료: CAICT 보고서를 기능 중심으로 재구성

- 집중형 시설은 통일된 작업절차와 품질검사로 수집의 일관성을 높이고, 고객 현장에서 반복하기 어려운 오류와 경계조건을 시험하는 데 유리
- 현장 환류는 실사용 분포를 제공하지만 현재의 신뢰성·도입 규모·고객 제약을 고려하면 초기 대량 데이터 공급을 전적으로 맡기기는 어렵다는 것이 보고서의 진단
- ✦ 훈련장은 분산 수집과 현장 환류 사이에서 수집과 검증을 조직하는 수단이며, 장기적으로도 모든 실제 작업환경을 한 건물에 복제하는 모델에 국한되지 않는 개념

3 훈련장의 기능과 자료 활용 범위

- ✦ 훈련장은 데이터 공급, 모델 개발, 시스템 검증과 산업 협력을 함께 지원하지만, 시설마다 실제로 제공하는 서비스 범위에는 차이
- 데이터 수집공장(数据采集工厂)·수집센터(采集中心)·실훈련장(实训场) 등의 명칭은 기능의 강조점과 발전단계를 반영하므로, 명칭만으로 전 과정 훈련·검증 역량을 보유했다고 판단하기에는 무리

- ▶ 일부 시설은 데이터를 생산·납품하는 데 집중하고 다른 일부는 모델 훈련, 본체 적용, 시나리오 검증까지 지원하며, 중간시험 기능은 별도로 확장되는 영역
- ✦ 본 보고서는 중국 로봇산업 정책 전체가 아니라 훈련장이 산업 인프라로 조직되는 방식과 그 성립조건에 분석 범위를 한정
 - ▶ 구축 현황에서는 운영주체·산업 시나리오·지역 협업을, 데이터에서는 수집방식·정제·품질·활용을, 정책에서는 검증과 표준화의 접점을, 운영에서는 비용·서비스·거래의 지속조건을 중심으로 검토
 - ▶ 기업의 모델 성능이나 개별 제품 경쟁력을 순위화하지 않으며, 사례는 시설 운영방식과 데이터 처리의 차이를 설명하는 범위에서 활용
- ✦ 본 보고서는 중국정보통신연구원 인공지능연구소·휴머노이드로봇(상하이)유한공사, 체화지능 테스트실험실이 공동 작성한 '26년 8월 훈련장 연구보고서를 중심으로 분석하고, 국가 휴머노이드 로봇 산업 표준체계 의견수렴안과 실경실훈 특별행동 통지를 보완자료로 활용
 - ▶ 8월 18일 공개된 CAICT 보고서는 4월부터 심층면담·설문 등을 거쳐 작성됐으며, 본 보고서는 PDF 전체 56쪽 가운데 본문 인쇄쪽수 1~48쪽을 중심으로 재구성
 - ▶ 시설 수와 유형 비중은 공개자료에 기반한 불완전 집계이고 사례별 시점이 다르므로, 이를 9월 현재의 전수통계나 검증된 경영성으로 확대 해석하는 것은 배제
 - ▶ CAICT 보고서는 산업 현황·전망을, 표준체계 의견수렴안과 실경실훈 특별행동 통지는 각각 제도적 접점과 정부 정책 요구를 확인하는 근거로 활용

II

중국의 구축 현황과 운영모델

1 확산 규모와 지역 입지

- ✦ 중국에서는 '24년 기업의 자발적인 대규모 실제환경 데이터 수집시설 구축이 먼저 나타났고, '25년 치린(麒麟) 훈련장 가동 이후 중앙정부 정책과 지방의 시나리오 개방·산업지원이 결합하면서 훈련장 확산 기반이 확대

 - ▶ 공업정보화부 등 8개 부처의 「AI+제조 특별행동 실시의견」은 체화지능 제품혁신과 함께 휴머노이드 로봇 중간시험(中试, 연구개발 성과의 산업화를 위한 공학적 검증) 기지·훈련장 건설, 대표 생산라인과 제조현장 적용을 명시
 - ▶ 공업정보화부·국무원 국유자산감독관리위원회의 '26년도 실경실훈(实景实训) 특별행동은 산업·서비스·특수 분야의 실제 작업단위를 훈련·시험·검증 공간으로 활용하고 응용검증을 상시배치로 연결
 - ▶ 국가발전개혁위원회는 체화지능 응용 중간시험 기지 건설을 가속하고 소프트웨어·하드웨어 생태계와 훈련 기반시설의 연계를 강화하는 방향을 제시
 - ▶ 국가데이터국의 업종 고품질 데이터셋 건설행동은 인공지능이 의사결정·실행, 체화지능, 물리적 상호작용으로 확장되는 흐름에 맞춰 데이터 공급경로와 유형을 넓히도록 요구
- ✦ 지방 차원에서는 '24년부터 '26년 6월까지 18개 성급지역에서 40건이 넘는 관련 정책이 발표됐고, 가동 시설은 70곳 초과, 건설·계획 시설은 46곳으로 집계
- ✦ 체화지능 훈련장은 장강삼각주·징진지·주강삼각주에 집중되며, 장강삼각주는 상하이로 중심으로 장쑤·저장·안후이의 산업 기반과 연결돼 가장 밀집한 집적지를 형성

 - ▶ 징진지에서는 베이징의 연구자원과 텐진·스자좡·바오딩 등의 훈련시설이 이어지고, 베이징 스징산은 단계적으로 훈련장을 확장하며 실제 데이터 생산기반을 확충
 - ▶ 주강삼각주는 방대한 제조업 기반과 응용 수요를 토대로 광저우·포산·선전·주하이·둥관·후이저우 등에 여러 훈련장을 구축하고 범용·산업 특화 시나리오를 각각 배치
- ✦ 훈련장 입지는 혁신자원이 밀집한 핵심도시뿐 아니라 실제 산업현장과 낮은 운영비용을 활용할 수 있는 다양한 지역 도시(多层次城市)로도 확산되며, 지역 산업의 작업환경 자체가 데이터 생산자원으로 부상

 - ▶ 지방정부는 산업기금·전용정책·시나리오 개방·산업단지 등을 활용해 훈련장을

유치하고 로봇 연구개발·시험, 데이터서비스, 장비제조 등 연관기업의 집적을 함께 도모

- ▶ 상하이 치린(麒麟) 훈련장이 있는 장강(张江) 모리사구(模力社区)에는 체화지능·인공지능 상하류 기업 200곳 이상이 모여, 훈련장이 데이터·검증시설을 넘어 산업자원 협업과 생태 집적의 거점으로 기능하는 사례를 형성

2 건설주체와 서비스 범위

- ✦ 훈련장은 정부가 자원을 제공하고 전문기관이 운영하는 형태와 기업이 자체 수요에 따라 구축하는 형태로 나뉘며, 소유와 운영의 역할을 따로 볼 필요
- ▶ 정부 유도형에는 공간·자금·정책 지원을 바탕으로 혁신센터, 연구기관, 로봇기업 등이 참여하며, 공유시설을 통한 기업 접근성 개선과 산업집적을 추구
- ▶ 기업 자체 구축형은 자사 모델·제품의 개발주기에 맞춘 데이터를 확보하는 데 유리하지만, 외부에 공개하는 범위는 각 기업의 사업전략에 따라 차이

표 2. 건설주체별 운영 동기와 확인사항

유형		주요 운영 동기	확인할 서비스 범위
정부 유도형		공통 기반 제공·지역 산업 지원	외부 이용 조건, 실제 이용자와 운영 책임
기업 자체 구축형	기술기업 자체 구축형	모델 개선·본체 적응·자체 데이터	자사 수요와 외부 제공의 구분
	수요기업 협력형	기존 업무의 자동화·현장 적합성	사용자의 작업목표와 검증 참여

자료: CAICT 보고서 내용을 토대로 재구성

- ✦ 수요기업 협력형은 선전 위수(玉树) 체화지능 로봇 즈창(智创)센터처럼 환경미화·청소 사업자와 로봇기업이 공동 구축해 기존 업무를 청소·정리 훈련과제로 구성하는 방식
- ✦ 중국 밖에서도 기업·연구기관이 자체 훈련환경을 마련하는 사례가 나타나며, 보고서는 중국이 이 모델을 먼저 규모화하고 있다고 평가
- ▶ LG전자가 서울 양재 R&D 단지를 가정·공장 환경을 재현한 로봇 실제환경 훈련센터로 활용하고, LG CLOiD 가정용 로봇의 물리세계 조작 데이터를 수집하는 사례가 소개
- ▶ 미국에서는 Figure AI가 BotQ 생산시설에서 공급한 로봇을 연구개발·데이터 수집·

현장 적용에 활용하고, Apptronik은 Robot Park에서 수집·훈련·배치를, Tutor Intelligence는 100대 Sonny 로봇의 Data Factory에서 수집·원격조작·평가·훈련을 수행

▶ 독일 TUM과 NEURA Robotics의 RoboGym도 산업·가정 환경에서 휴머노이드 데이터 수집과 훈련을 진행

✦ 건설주체를 비교할 때는 공공성이나 소유형태보다 누가 작업수요를 정의하고 생산한 데이터와 검증결과를 실제 개발·배치에 사용하는지를 함께 확인할 필요

3 종합형과 산업 특화형의 분화

✦ 종합형과 산업 특화형은 시설 규모가 아니라 여러 산업에 걸친 범용성 확보와 특정 작업의 깊이 있는 적응이라는 목적의 차이로 구분

▶ 세 개 이상 산업과 여러 세부 시나리오를 포함하는 종합형이 60%를 넘고, 단일 산업의 여러 과제에 집중하는 특화형은 20% 미만으로 집계

✦ 종합형은 다양한 작업·물체·환경을 연결해 범용 모델 학습과 시나리오 간 적응을 지원하는 것이 장점

▶ 종합형 훈련장 가운데 제조·가정·상업서비스를 함께 구축한 조합은 44%로 제시

▶ 지방정부 관점에서는 제조·에너지·항만·농업 등 여러 지역 산업에 공동서비스를 제공할 여지가 있으나, 폭넓은 분류와 과제의 깊이는 별개

✦ 산업 특화형은 사용자가 요구하는 핵심 절차와 성능조건을 구체화하고, 실제 업무에 대한 반복 검증을 심화하는 방식으로 차별화

▶ 칭다오의 돌봄·재활 분야 검증센터가 돌봄시설·의료기관·지역사회 양로·재가 양로·단지서비스의 다섯 범주 아래 11개 세부 시나리오를 구성한 사례가 소개

▶ 이 사례는 단일 분야 안에서도 서로 다른 환경과 사용자 요구가 존재함을 보여주며, 산업 특화가 곧 하나의 단순 동작만 수집한다는 뜻이 아니라는 점도 시사

✦ 두 유형의 역할은 상호 보완적이지만 동일한 평가표를 적용하면 범용성과 전문성 중 한쪽의 성과가 제대로 드러나지 않을 소지

▶ 종합평가 구상도 두 유형의 역할에 따라 평가 중점을 달리하는 방향을 제시

▶ 시설 유형은 데이터 고객, 검증 대상 업무와 기존 지역시설의 기능 분담에 따라 구분

4 산업 시나리오의 분포와 진입조건

- ▶ 현재 훈련장의 시나리오는 제조업에 가장 집중돼 있으며, 작업경계·환경 통제성·경제적 가치가 초기 적용의 선택에 영향

표 3. 산업별 훈련과제와 적용 제약

영역	대표 과제	주요 검증조건
제조	분류·운반·조립·검사	작업속도, 품질과 연속운전
가정	청소·의류 정리·생활 보조	공간·물체 다양성, 사생활 보호
상업서비스	진열·포장·배달·안내	반복 업무와 고객 상호작용
의료·돌봄	보조·돌봄·물품 취급	대인 안전, 신뢰성과 업무 경계
특수·응급	순찰·점검·위험 탐지	지형·환경 재현, 수집 위험

자료: CAICT 보고서 내용을 토대로 재구성

- ▶ 공개된 시설 중 제조 시나리오 포함 비중은 86%, 특수·응급 시나리오 포함 비중은 약 25%로 제시되며, 한 시설이 여러 산업을 중복 포함
- ▶ 제조업 내부에서는 물류·창고·자재운반이 공개된 제조 시나리오 가운데 약 40%를 차지
- ▶ 가정과 상업서비스는 비정형 환경 및 사람과의 접촉이 늘어나는 영역으로, 제조업 시험결과의 직접 적용에는 한계
 - ▶ 가정은 물체 종류와 공간배치 변화가 크고 작업의 연쇄성이 높으며, 상업서비스는 고객 간섭과 상호작용이 빈번해 다른 실패요인을 포함
- ▶ 의료·돌봄과 특수·응급은 활용 필요성이 크더라도 안전과 환경 재현비용 때문에 단계적인 검증이 요구되는 분야
 - ▶ 실제 환자·위험현장에 대한 실험의 제약을 고려해야 하며, 보조작업 시험을 고위험 핵심업무 수행능력까지 검증한 것으로 확대 해석은 불가

5 중심시설과 분산 훈련장의 협업

- ✦ 지역 간 협업은 훈련장 수 확대를 넘어 데이터 규격과 관리능력을 공유하는 조직적 협업체계로 전환
 - 상하이 치린 훈련장이 여러 지역의 분산 훈련장을 연결하는 '1+N' 구상을 통해 데이터 집계와 규격 통일·역량 공유를 추진하는 사례가 제시
 - 광둥성은 성급 종합훈련장을 능력 중추로 두고 여러 도시·산업별 분산 훈련장을 연결하는 '1+1+N' 체계를 제시해, 성정부 주도의 협업망을 추진
- ✦ 기업 중심 협업에서는 중심 데이터 수집시설이 파트너 훈련장에 수집방법을 제공하고 결과를 회수해 공통 개발체계와 연결하는 방식도 확인
- ✦ 현장마다 물체와 업무가 달라지는 것은 유용한 다양성이 될 수 있지만, 센서 규격과 작업 의미가 불일치하면 자료를 합치는 데 추가 비용이 발생
 - 중앙에서 모든 시나리오를 동일하게 만드는 것도 지역 산업의 특징을 잃게 할 수 있어, 공통 처리규칙과 현장 고유조건의 구분이 핵심 과제
- ✦ 협업망의 실효성은 연결된 거점 수뿐 아니라 공통 규격 아래 생산된 데이터가 실제 학습·검증에 함께 사용되고 지역별 현장조건이 서비스에 반영되는지에서 확인 가능
 - 작업정의, 성공·실패 기준, 수집절차와 품질관리가 일치해야 다른 지역에서 받은 자료를 재처리하는 부담을 줄이고 비교 가능한 결과를 축적
 - 데이터 생산방법 자체를 기업의 핵심 노하우로 보는 경우가 많아 기관 간 정보공유에 제약
- ✦ 분산형 실제환경(分布式实景) 지점은 기존의 지역별 시설망을 넘어 생산라인·가정·매장 등 실제 업무공간과 훈련능력을 연결하는 구상
 - 이 구상에서는 중심시설이 공통 도구와 검증역량을 제공하고 각 지점은 실제 업무의 변화를 반영하는 역할을 맡을 수 있으나, 구체적 분담은 운영모델에 좌우

III

로봇 훈련데이터의 생산과 활용

1 수집기술의 선택과 비용·가치

- ✦ 실제 데이터 수집기술은 로봇 본체 없이 사람의 행동을 기록하는 무본체(无本体) 수집 방식과 실물 로봇의 실행을 기록하는 방식으로 구분
 - ▶ 무본체 수집은 사람의 환경 적응력과 장비의 이동성을 활용하지만, 로봇이 실행할 동작으로 변환하는 처리가 필요
 - ▶ 실물 로봇 수집은 관절·말단·센서 정보를 직접 기록해 본체 일치성이 높지만, 장비·인력·공간의 투입 부담이 주요 제약

표 4. 실제 데이터 수집방식의 비교

수집방식	주된 활용	처리·운영 부담
1인칭 영상	상황·행동 이해와 공간 예측	행동 추정, 로봇 동작으로 변환
말단 조작	집기 등 말단의 정밀 상호작용	도구·본체의 형태와 동작 정렬
동작 포착	전신·손의 복잡한 움직임	골격 해석과 동작 재지정
원격조작	본체와 일치하는 실제 실행	로봇·조작인력·공간 비용
자율작업 환류	실패·개입·복구 학습	실제 배치 규모와 기록체계

자료: CAICT 보고서 내용을 토대로 재구성

- ✦ 훈련장이 데이터를 상품으로 판매하는 경우, 단순 비맞춤형 무본체 수집 데이터를 시간당 약 100위안, 실물 로봇 데이터를 품질·비용 차이에 따라 시간당 수백 위안에서 1천 위안 이상으로 제시
- ✦ 생산원가 측면에서 원격조작은 규모화 운영 후 시간당 유효 데이터 비용이 약 275위안(IT쥬프(IT桔子) 자료 인용)으로 제시되지만, 소규모 수집에서는 시간당 수천 위안까지 올라 장비 감가·인력·시나리오 비용의 영향이 큰 구조

- ✦ 따라서 표면적인 시간당 가격만 비교하기보다 행동 변환·정제·검수와 본체 적합성까지 포함한 최종 학습자료의 비용을 함께 산정할 필요
- 시나리오 이해에는 폭넓은 무본체 자료, 실제 제어 적응에는 실물 로봇 데이터, 예외상황 개선에는 현장 환류를 조합하는 방식이 비용과 활용가치의 균형을 좌우

2 사람의 행동을 로봇에 연결하는 정렬

- ✦ 사람의 행동 데이터를 로봇에 활용하려면 시각적으로 유사한 동작을 모으는 데 그치지 않고 서로 다른 행동공간을 맞추는 과정이 필요
 - 사람의 손·팔과 로봇의 관절·말단은 구조와 움직임 범위가 다르므로, 사람에게 자연스러운 동작과 로봇의 실행 가능성은 별개
 - 특히 영상만으로는 본체 상태와 실행 가능한 동작 라벨이 충분하지 않아, 이동경로 복원·손 자세 추정·동작 변환 등 여러 후처리 단계를 수반
- ✦ CAICT 보고서는 행동을 공통 표현으로 바꾸는 방식과 수집장비를 실제 본체에 가깝게 구성하는 방식을 구분해 설명
 - 위치·자세·그리퍼 개폐처럼 비교적 공통적인 표현을 활용하면 다른 본체로 이전할 여지가 생기지만, 변환 이후에도 실제 실행의 제약 확인이 필수
 - 카메라와 그리퍼 등 수집장비를 로봇의 말단 구성에 맞추면 변환 오차를 줄일 수 있으나, 그 조건이 달라지는 다른 본체에도 동일하게 적용되는지는 별도 문제
- ✦ 동작 포착에서도 정밀한 측정과 로봇에 적합한 행동 생성은 구별되며, 장비 방식에 따라 공간 구성과 오차의 성격에 차이
 - 광학 방식은 여러 카메라를 배치해 움직임을 복원하므로 측정공간 구축의 제약이 있고, 관성 방식은 이동성이 좋지만 위치 추정의 누적오차가 고려 대상
 - 전신 또는 손의 움직임을 정확히 측정했다라도 골격 해석과 동작 재지정 과정에서 정보 손실이나 본체의 운동 제한과 충돌할 가능성이 존재
- ✦ 데이터의 범용성은 특정 하드웨어와 무관하다는 선언보다 어느 표현과 조건에서 이전할 수 있는지를 설명할 때 실질적 의미가 성립
 - 사전학습에 쓰는 행동 경험과 특정 작업에 적응시키는 후속 학습자료의 역할을 분리하면, 대량의 사람 데이터와 소량의 로봇 실행자료를 결합하는 이유가 명확

3 작업정의와 반복수집의 공정화

- ✦ 데이터 생산은 장비 가동 전에 어떤 작업능력을 개선할 것인지 정하는 단계에서 시작하며, 사용 목적에 따라 수집목표에도 차이
 - 연구용 자료는 시나리오 다양성과 전체 길이가 중요한 경우가 있지만, 현장 적용용 자료는 해당 작업의 목표 성능에 기여하는지가 핵심 판단기준
 - 필요한 시나리오와 과제·규모를 정한 뒤 정밀도·작업 길이·모달리티와 본체 조건에 맞춰 수집방식을 선택해야 과도한 장비 투입과 재수집을 줄이는 데 유리
 - 위험환경이나 드문 사건은 합성·시뮬레이션으로 보완할 수 있으나, 실제 물리조건과의 일치성을 확인하지 않은 합성자료가 실제 기록을 완전히 대체하기에는 한계
- ✦ 반복수집은 사람의 행동을 단순히 여러 번 촬영하는 것이 아니라, 로봇 학습에 필요한 작업절차를 같은 기준으로 반복 실행하는 과정
 - 의류 정리 데이터를 예로 들어, 사람의 평소 동작을 그대로 따라 하면 로봇 학습에 필요한 일관된 행동자료가 되지 않을 수 있다고 지적
 - 따라서 작업자는 물체를 다루는 순서와 동작 흐름을 이해하고, 시작조건·작업단계·종료조건을 같은 기준으로 해석해 수집할 필요
- ✦ 동일 조건의 반복과 환경 변화를 통한 다양성 확보는 함께 설계해야 하며, 변화를 모두 오류로도, 모두 유용한 사례로도 보지 않는 선별이 전제
 - 물체·위치·조명 등 과제의 본질과 관련된 변화는 일반화 학습에 도움이 될 수 있지만, 센서 설정 차이나 임의의 조작습관은 불필요한 잡음이 될 소지
 - 어떤 차이를 의도적으로 포함했는지 기록해야 이후 성능 차이의 원인을 해석할 수 있으며, 이는 시나리오 설계와 품질검사를 연결해야 하는 이유
- ✦ 데이터 생산라인은 작업지시에서 수집·정제·검수·저장을 거쳐 훈련·검증으로 이어지는 연속 공정
 - 수집만 빠르게 진행되고 후속 단계의 처리속도가 부족하면 원시 데이터가 축적돼도 이용 가능한 데이터 공급은 늘지 않으므로, 병목을 공정 전체에서 확인할 필요

4 정제·주석과 실패 데이터의 품질관리

- ✦ 체화지능 데이터 정제는 영상의 선명도만 확인하는 작업이 아니라 시간·좌표·동작·힘 정보를 연결해 실제 사건의 순서를 복원하는 과정

- ▶ 카메라 영상과 관절 움직임의 시간이 어긋나면 행동과 결과를 잘못 연결할 수 있어, 서로 다른 센서의 기록을 같은 시간축에서 해석할 수 있도록 맞추는 작업이 선행
- ▶ 주석(라벨링)은 작업을 최소 스킬 단위로 나누거나 시간정보·언어 설명 또는 술어 논리를 기준으로 구분하는 등 여러 방식으로 수행
- ✦ 데이터가 읽히고 처리된다는 의미의 유효성과 학습 목적·품질·준법 조건을 충족한다는 의미의 적합성은 구별할 필요
 - ▶ 영상·음성·환경정보에는 사람과 장소의 정보가 포함될 수 있고 산업현장에는 핵심 공정이 노출될 수 있어, 처리기술과 접근 권한의 병행 관리가 중요
 - ▶ 동일 과제의 자료라도 작업단계 구분과 성공 판정이 다르면 통합 활용이 어려우므로, 주석의 정확성뿐 아니라 그 의미가 일치하는지도 검토 대상
- ✦ 실패 기록은 일률적으로 폐기할 저품질 자료가 아니며, 기록 자체의 결함과 로봇이 실제로 작업에 실패한 사건과는 구별
 - ▶ 센서 단절이나 핵심 정보의 누락으로 원인을 알 수 없는 기록은 활용이 제한되지만, 실패 직전·직후 상태와 개입·복구가 남아 있으면 개선 학습에 기여 가능
 - ▶ 실패를 보존하더라도 원인과 회복조치를 구분해야 하며, 성공사례와 무분별하게 합치면 데이터 양은 늘어도 학습 목적이 불명확해질 우려
- ✦ 품질관리는 납품 직전의 선별로 끝나지 않고 문제를 수집과 주석 단계로 되돌리는 반복체계일 때 생산효율 개선으로 연결
 - ▶ 베이징 사례에서는 동작 완결성·주석 정확성·시나리오 일반화성을 다층적으로 검토하고, 검사결과를 앞 공정의 작업절차 개선에 활용
 - ▶ 자동 주석과 지능형 검수는 반복작업 부담을 줄이는 수단으로 제시되지만, 자동화 수준과 상관없이 최종 자료가 목표 작업에 적합한지는 별도 확인 사항

5 저장에서 모델·현장 검증까지

- ✦ 저장체계는 대량 자료를 보관하는 용량뿐 아니라 학습에 필요한 조건의 자료를 빠르게 찾고 연결하는 기능까지 필요
 - ▶ CAICT 보고서는 여러 대의 로봇이 영상·라이다·관성·힘·촉각 기록을 동시에 생산하면 하루 데이터 증가량이 테라바이트 규모에 이를 수 있다고 설명
 - ▶ 대용량의 원시 데이터와 반복적으로 사용하는 학습자료의 접근 빈도가 다르므로, 저장 기반과 고속 접근 계층을 구분하는 구조가 활용 사례로 제시

- ✦ 자료의 조직방식은 프로젝트·유형·생애주기별 관리, 원시 데이터와 벡터의 연결, 시계열 또는 다중모달 표현 등 여러 형태로 구성

 - 이들 방식의 선택은 검색과 재사용 목적에 따라 달라질 수 있으며, 파일을 한곳에 모으는 것만으로는 다른 본체·시나리오와의 데이터 의미 차이 해소에 한계
 - 수집·정제·검수 결과와 실제 학습자료의 연결은 모델 개선·실패 원인을 데이터 생산 단계까지 추적하는 기반
- ✦ 훈련장 데이터의 최종 가치는 납품량보다 모델과 본체가 목표 작업을 더 잘 수행하는지에서 확인되므로, 내부 검증이 핵심 연결고리

 - 생산자료와 공개자료를 학습 목적에 맞게 조합하고 모델 성능의 변화를 확인한 뒤, 본체에 적용해 실제 작업조건에서 목표 지표를 시험하는 절차가 필요
 - 검증에서 나온 분포 편향과 시나리오 누락은 수집목표·동작·주석 규칙의 조정 근거
 - 모델 적용을 외부 고객에게 전적으로 맡기는 단절형 서비스 체계는 데이터 효용을 확인하기 어렵고, 향후 현장 환류가 늘면 경쟁력이 약해질 수 있다는 것이 보고서의 경고
- ✦ 공정 관리에서는 수집 가능한 규모, 품질검사를 통과한 규모와 실제 훈련에 사용한 규모를 같은 지표처럼 취급하지 않는 것이 중요

 - '26년 천만 시간급 연간 데이터 생산 전망은 가동시설의 잠재 생산능력에 관한 수치로, 이미 확보·검수된 고품질 데이터의 실제 보유량과는 구별

IV

실제환경 검증과 표준화

1 반복시험과 배치 전 적합성 확인

- ✦ 훈련장의 검증 기능은 성공적인 시연을 확보하는 것보다 어떤 조건에서 실패하고 어떻게 개선할 수 있는지 확인하는 데 중점

 - ▶ 물체 종류·배치·조명·사람의 간섭·지시방식과 작업 난도(難度)를 조절하면 특정 오류를 반복해서 관찰할 수 있고, 결과를 추가 수집과 재훈련으로 연결 가능
 - ▶ 난도별 시나리오 구성은 단일 물체·고정 위치에서 다중 물체·무작위 위치·동적 가림으로 조건을 넓히는 방식
 - ▶ 고객 현장에서는 생산일정과 안전·사생활·운영책임 때문에 반복실패를 허용하기 어려워, 초기에는 위험을 통제할 수 있는 공간의 역할이 중요
- ✦ 배치 전 검증은 본체의 개별 성능을 넘어 환경·작업·운영조건이 결합된 시스템의 능력을 확인하는 과정

 - ▶ 작업 성공률, 실행효율, 안정성, 강건성, 안전성과 시나리오 적합성을 단일 동작의 성공 여부를 넘어 확인할 평가 항목으로 제시
 - ▶ 특정 물체를 고정 위치에서 집는 시험과 주변 사람이 움직이는 업무공간에서 연속작업을 수행하는 시험은 난도가 달라 결과의 동일 수준 비교는 무리
- ✦ 훈련공간의 통제성은 원인 분석에 유리하지만, 고객의 업무 대표성은 별도 점검 과제

 - ▶ 시연에 유리한 작업만 선택하면 현장에서 빈번한 예외와 긴 작업연쇄를 충분히 반영하지 못할 수 있어, 업무 파트너의 지식이 시나리오 구성에 반영되는 것이 관건
- ✦ 모델 평가와 데이터 생산을 연결하려면 실패 발견 이후의 조치가 명확해야 하며, 같은 시험의 반복만으로는 개선 효과 보장에 한계

 - ▶ 실패가 데이터 부족, 행동 표현의 불일치, 본체 제약 또는 환경조건 중 어디에서 비롯됐는지 구분해야 추가 수집과 장비·모델 개선의 방향을 결정 가능
 - ▶ 이런 분석은 훈련장을 최종 인증기관으로 간주한다는 뜻이 아니며, 시설이 제공하는 시험서비스와 해당 산업의 공식적인 승인 절차는 별개

2 실경실훈 정책과 운영의 연결

- ✦ 실경실훈은 실제 업무환경의 훈련과 검증을 통해 제품의 현장 적용을 촉진하려는 접근이며, 별도 시설 신설을 전제로 하지 않는 방식
 - ▶ 훈련장을 통제된 수집공간에서 실제환경의 데이터·모델·시스템 검증을 지원하는 시설로 확장해 설명하고, 경우에 따라 중간시험 기능까지 포함
 - ▶ 실제 업무의 대표성과 반복검증의 가능성을 함께 확보해야 하므로, 시나리오 제공자의 요구와 훈련장 운영자의 기술 역량을 연결하는 것이 중요
 - ▶ '26년도 실경실훈 특별행동은 사용자와 기술공급자가 실제 작업 수요를 중심으로 협력하고 검증을 배치로 연결하는 정부 정책 근거
 - ▶ 공업정보화부·국무원 국유자산감독관리위원회의 통지는 6월 3일 작성·6월 8일 공개됐으며(문서번호: 工信厅联科函[2026]256号), 응용검증에서 실제 작업 성공률·효율 향상률·안전신뢰성·경제적 타당성을 핵심 지표로 제시
 - ▶ 정책은 '26년 말까지 대표 시나리오에서 응용검증과 상시배치를 추진하고 고가치 응용 시나리오 100개 이상을 도출해 만 대급 규모의 배치역량을 형성하는 것을 목표로 설정
- ✦ 정책의 응용 지향성을 훈련장에 연결하려면 시설 내부의 데이터 생산과 외부 업무 개선 사이에 확인 가능한 연결이 전제
 - ▶ 사용자 단위(用户单位)와 완성로봇 기업 또는 응용서비스사를 주체로 모델·부품 등 공급망 기업과 연구기관이 혁신응용연합체(创新应用联合体)를 구성하며, 사용자 단위는 실제 훈련공간을 개방하고 배치목표와 시나리오 요건을 구체화
 - ▶ 검증된 해결책은 사용자 단위·동종 시나리오·연관 영역으로 상시배치를 확대하도록 설계돼, 데이터 생산·작업기술 개발·응용검증이 단계별로 끊기지 않게 정책으로 연결
- ✦ 특별행동의 연합체 구조는 실제 사용자가 제시하는 작업목표와 공급자의 기술역량을 결합해 검증 가능한 현장 서비스로 전환하는 데 초점
 - ▶ 특별행동의 목표치와 지원체계는 개별 훈련장의 달성 실적으로 대입하기보다, 훈련·검증 인프라가 수요기업의 상시배치와 연결되는 정책 방향을 보여주는 근거로 해석

3 국가 표준체계 초안의 훈련장 접점

- ✦ 국가 휴머노이드 로봇 산업 표준체계 건설지침은 '26년 8월 24일 작성·25일 공개된 의견수렴안(공시기간 8.25~9.23.)으로, 훈련장과 데이터·시험평가·응용·안전의 제도적 접점을 보완하는 자료

- ▶ 이 의견수렴안은 '26년 2월 표준화기술위원회가 발표한 「휴머노이드 로봇·체화지능 표준체계(2026판)」와 별개의 정부 의견수렴 문서
- ▶ 초안은 '28년까지 핵심표준 최소 100건을 마련하고, 표준 보급·실행에 참여하는 기업이 200개를 넘도록 한다는 목표를 제시하며, 이는 확정 실적이 아닌 정책 목표

표 5. 표준체계의 여섯 영역과 훈련장 관련 내용

영역	초안의 주요 내용	훈련장과의 접점
기초 공통	분류·중간시험·시험평가·신뢰성	평가조건과 결과 비교
뇌형·지능컴퓨팅 (类脑与智算)	데이터·기초모델·세계모델·개발	수집부터 모델 적용까지 연결
신체·부품	감지·구동·통신·에너지 등	센서·본체 제약의 확인
완성로봇·시스템	인터페이스·운영체제·설계· 시뮬레이션 플랫폼	여러 장비의 연계·검증
응용	제조·민생서비스·특수 분야	업무별 과제와 성능조건
안전·윤리	물리·데이터·알고리즘·행동 등	전 과정 위험과 책임 경계

자료: 국가 휴머노이드 로봇 산업 표준체계 의견수렴안을 토대로 재구성, 훈련장 접점은 본 보고서의 해석이며, 확정된 시설별 의무와 무관

- ✦ 데이터와 시험평가를 하나의 표준체계 안에 두는 구성은 시설의 생산물과 실제 작업능력 사이의 일관성을 확보하려는 방향과 부합
- ▶ 다만 휴머노이드 중심의 표준체계를 체화지능의 모든 본체와 시설에 동일하게 적용할 수 있는지는 표준별 범위와 후속 내용의 확인이 필요

4 데이터·훈련장 평가의 표준화

- ✦ CAICT 보고서는 훈련장 표준화 수요를 시나리오 구축, 데이터 생산, 모델 훈련서비스, 중간시험 서비스, 훈련장 건설·운영, 안전·준법의 6개 영역으로 구분
- ▶ 시나리오 표준화는 유형·과제 정의·환경요소·복잡도 등급을 맞춰 서로 다른 시설의 수집조건과 검증결과를 비교할 수 있게 하는 데 초점

- ▶ 데이터 표준화는 수집·품질평가·저장관리·권리확인·유통의 공통규칙을 통해 여러 기관 자료의 일관성·호환성과 재사용성을 높이는 역할
 - ▶ 모델 훈련서비스 표준화는 훈련환경·과제·모델·하드웨어 접속요건을 정리해 이용자가 다른 시설의 자원을 반복적으로 사용할 때 적응비용을 낮추는 방향
 - ▶ 중간시험 표준은 시스템통합·공학검증·소량생산 일관성 확인을, 건설·운영 표준은 분류·등급·관리·평가를 다뤄 데이터 생산만으로 포괄하기 어려운 기능을 별도 규범화
 - ▶ 안전·준법 영역은 물리·시스템·데이터·운영 안전을 전 생애주기에서 다뤄 훈련·검증 서비스의 기술요건과 책임경계를 함께 정리하는 방향
- ✦ ITU-T SG21은 '26년 2월 19일 체화지능 포커스그룹(FG-EAI)을 설립했으며, 산하 WG2는 다중모달 데이터 프레임워크 담당
- ▶ 이와 함께 SG21에서 데이터 형식·이용규칙 통일과 실제환경 원격조작·다중모달 데이터 수집 표준화가 추진되고 있다고 소개
- ✦ 중국에서는 MIIT/TC1 체화지능 그룹을 중심으로 데이터셋 품질과 훈련장 종합평가의 표준화가 병행
- ▶ 첫 체화지능 데이터셋 품질 산업표준 YD/T 6771-2026은 '26년 7월 15일 발표돼 11월 1일 시행 예정
 - ▶ 2026-0074T-YD 훈련장 종합평가 기술규범은 보고서 작성 당시 제정 진행 중
 - ▶ 두 건 모두 AIIA(중국인공지능산업발전연맹, 中国人工智能产业发展联盟)에서도 발표돼 선행 시범 적용을 추진
- ✦ 훈련장 종합평가 구상은 시나리오 구축, 설비·기반시설, 데이터 전 과정, 운영관리, 위험통제를 함께 살피고 종합형·산업 특화형에 서로 다른 중점항목을 적용
- ▶ 종합형에서는 산업·시나리오·본체 간 일반화 능력과 고품질 데이터 생산을, 특화형에서는 핵심작업의 수집·훈련·적용 연계와 배치능력을 중점 평가하며, 두 유형 모두 1~5성 등급평가로 시설 역량과 위상을 명확히 해 구축을 촉진하려는 취지
 - ▶ 표준화는 시설 수와 면적을 비교하는 기준보다 데이터 품질, 서비스 기능, 작업 검증을 같은 체계 안에서 연결해 훈련장의 실제 능력을 점검하는 수단으로 확대되는 흐름

5 안전 요구와 적용 경계

- ✦ 체화지능 훈련은 데이터 처리와 실제 물리적 움직임이 함께 발생하므로, 정보보호와 작업공간의 안전을 분리해서 다루기 어려운 영역

- ▶ 사람의 영상·음성·주변 환경과 산업 공정정보가 수집될 수 있고, 로봇의 동작은 사람·설비·물체에 직접 영향을 주므로 단계별로 위험의 성격에 차이
- ▶ CAICT 보고서는 물리·시스템·데이터·운영 안전을 전 생애주기의 관리과제로 설명하며, 이는 준법 검토가 데이터 판매 단계에 한정될 수 없음을 시사
- ✦ 국가 표준체계 초안은 모델의 내부 안전, 외부로 나타나는 행동 안전과 산업별 적용 안전을 별도 항목으로 구분
 - ▶ 알고리즘 안전은 개발·학습·배치·갱신 과정의 통제와 이상행동 식별 등을, 행동 안전은 지각·계획·운동·조작과 비상대응·검증방법을 포괄
 - ▶ 적용 안전은 산업별 위험, 작업경계, 사람과 환경의 보호, 배치 검수와 운영감시 등을 포함하므로, 동일 로봇이라도 사용 환경별로 확인조건이 상이
- ✦ 데이터가 표준 형식에 맞는다는 사실과 고객 현장에서 안전하게 사용할 수 있다는 결론 사이에는 여러 검증단계가 존재
 - ▶ 기록 형식과 품질이 적합해도 특정 작업의 예외상황이 충분히 포함되지 않을 수 있고, 모델이 이를 학습했더라도 본체 제약으로 실행 결과가 달라질 소지
 - ▶ 따라서 데이터 품질평가, 모델·시스템 시험, 산업 현장의 적용 판단은 서로 연결하되 하나의 성적서가 모든 단계를 대신하는 구조로 해석하지 않는 것이 타당
- ✦ 안전 요구는 표준체계 초안의 방향과 산업별 규정·검증조건을 구분해 해석할 필요

V

지속가능한 운영과 발전방향

1 설비 투자와 운영비의 구조

- ✦ 종합형 훈련장은 실제환경·본체·수집장비·연산·저장·운영체계를 함께 갖추는 자산집약적 사업으로, 초기 시설비만으로는 비용 평가가 불충분

 - CAICT 보고서는 수천 제곱미터 이상 시설의 건설비가 수천만~1억 위안 이상에 이를 수 있다고 설명하며, 제시된 시설 규모·장비 구성·생산량·비용 추정만으로 개별 시설의 판매율·가동률을 판단하거나 공통 건설단가·실제 수입을 산정하기에는 한계
 - 환경의 종류가 늘고 재현도가 높아질수록 본체·장비뿐 아니라 공간 개조, 소품, 통신과 연산·저장 수요가 함께 늘어나는 상호연동 구조
- ✦ 본체 구매는 주요 단일 지출항목이지만, 실제 사업비는 작업 종류와 배치규모에 따라 달라져 로봇 대수만으로는 산정 근거 부족

 - 100대 배치 예시는 대당 수십만 위안을 전제로 본체 구매비가 수천만 위안에 이를 수 있음을 설명한 가정으로, 특정 시설의 실집행액과 구분
 - 정밀 동작 포착, 환경 재현과 시험장비는 요구조건에 따라 비용차가 크며, 기존 공간을 활용하는 방식과 완전히 복제하는 방식의 부담에도 차이
- ✦ 가동 후에는 인력·설비 유지보수·공간 운영·연산 사용료가 지속적으로 발생하므로, 생산능력과 실제 매출의 간격이 중요

 - 연간 운영비가 수백만 위안에 이를 수 있다고 설명하고, 자체 연산설비 구축과 외부 자원 사용 모두 지속적인 비용을 수반하는 선택으로 제시
 - 연간 데이터 생산량 10만 시간, 건설투자 5천만 위안, 생산한 데이터를 모두 판매하는 가장 이상적인 조건을 가정해도 데이터 판매만으로 투자회수에 3년 이상이 걸리는 사례를 제시
- ✦ CAICT 보고서는 서비스 수요 확인을 시설계획의 선행 요소로 제시

2 데이터 판매에서 훈련서비스로

- ✦ 데이터 판매는 비교적 빠르게 매출을 창출할 수 있지만, 다양한 고객의 학습목적을 충족하고 재구매로 이어지게 하는 조건이 필요
- 기성 데이터와 맞춤형 데이터는 수집·주석 복잡도와 고객 요구가 달라 가격차가 크며, 데이터 길이가 같아도 본체 적합성과 처리수준이 다를 여지
- 현재 주요 구매자는 중국의 선도 체화지능 모델·본체 기업이고 그다음이 대학·연구기관이므로, 대규모 시설 건설과 폭넓은 구매시장 형성은 별개

표 6. 훈련장이 모색하는 서비스 다변화

서비스	고객에게 제공하는 가치	운영상 확인조건
데이터 공급	목적별 학습자료 확보	품질·권리·본체 적합성
훈련 대행	장비·자료·연산을 묶은 개발 지원	훈련범위·자원 이용·결과 확인
공동개발	고객 과제에 맞는 해결책	역할·성과·권리의 합의
적합성 검증	시장 진입 전 작업·환경 시험	시험조건과 적용한계 명시
중간시험·인재훈련	공학적 검증과 작업역량 축적	전문인력·고객 수요·서비스 품질

자료: CAICT 보고서 내용을 토대로 재구성

- ✦ 훈련서비스는 시설·수집도구·본체·연산을 갖추기 어려운 개발자에게 데이터 생산부터 검증까지 연결된 자원을 제공
- 요금은 훈련시간·연산 사용량·데이터 이용량 등과 연결될 수 있지만, 실제 수익성은 이용률과 서비스 수행비용에 좌우되므로 방식의 도입만으로는 수익성 보장에 한계
- ✦ 서비스 다변화에서는 수익항목의 수보다 축적된 시설·데이터·업무지식이 고객 문제의 해결에 반복적으로 쓰이는지가 중요

3 데이터 거래와 신뢰 형성

- ★ 데이터 거래에는 공급량뿐 아니라 구매 전 품질·용도·권리조건을 판단하기 어려운 정보 비대칭도 제약 요인
 - ▶ 범용 거래플랫폼은 체화지능 고객이 요구하는 표본 확인·품질시험·권한관리에 한계가 있다는 것이 보고서의 진단이며, 본체 적합성 비교도 같은 한계에 해당
 - ▶ 수집방식·시나리오·주석 수준이 다른 상품을 시간이나 파일 크기만으로 비교하면 실제 사용가치를 판단하기 어려워, 구매자와 판매자의 협의비용이 증가
- ★ 데이터를 만드는 과정에는 공간 제공자·본체 기업·작업자·운영자 등이 참여하므로 활용과 유통의 권리조건을 명확히 하는 것이 중요
 - ▶ 생산절차와 동작 기준 자체를 핵심 노하우로 보는 기업은 자료 공유가 기술적 우위를 약화시킬 수 있다고 우려하며, 규격 통일만으로 이 문제 해소에는 한계
 - ▶ 거래 과정에서 사용범위와 접근권한이 분명하지 않으면 후속 분쟁과 추가 활용의 제한이 발생할 수 있어, 품질평가와 별도의 신뢰 조건이 필요
 - ▶ '26년 초 장쑤성 데이터거래소에서는 전국 첫 체화지능 데이터셋 거래가 성사됐고, 상하이의 대형모델 기업이 구매하고 주장즈닝커지(箴境智能科技)가 데이터를 공급한 사례가 소개
 - ▶ '26년 4월 징둥은 징둥원(京东云) 체화지능 데이터 거래플랫폼을 공개해 데이터 유형·식별코드·수집방식·시나리오·과제등급과 구매자의 사용권한을 상품정보로 제시
 - ▶ FastUMI Pro 데이터마트와 바이두 지능형 클라우드(百度智能云)의 데이터마트처럼 여러 시나리오·과제의 상품을 한곳에 모아 분류·가격정보를 제공하는 방식도 병행 탐색
- ★ 거래경로는 중국 내 데이터거래소·전문 플랫폼에서 국경 간 유통지원까지 확장되지만, 개인정보·중요데이터 포함 여부와 권리조건에 따라 적용 가능한 준법경로에 차이
 - ▶ '26년 5월 텐진 파시니(帕西尼) 계열사가 자유무역시험구 데이터 국외 이전 네거티브 리스트 관리체계 아래 개인정보·중요데이터를 포함하지 않는 고가치 데이터 상품을 해외시장에 공급한 사례가 소개
 - ▶ 이 사례들은 거래 인프라가 형성되기 시작했음을 보여주지만, 상품 등록이나 국외 판매 자체가 품질·권리·수요의 지속성을 보장하지 않으므로 서비스 이용과 재구매까지가 확인 대상

4 역량 축적과 전문화의 경로

- ✦ CAICT 보고서는 훈련장 발전을 건설 열기, 역량 축적, 생태계 성숙의 세 단계로 설명하되, 단계 간 경계가 명확하지 않고 역량 축적기는 장기간 조정될 수 있다고 부연
- 초기에는 넓은 시나리오와 장비가 빠르게 늘지만 시나리오·과제의 동질화와 데이터 공정 미비, 검증·거래 연결 부족으로 생산물의 품질과 활용성이 달라질 여지
- 역량 축적기에는 종합형·특화형 분화와 협업망, 공정화·검증 환류·운영모델 탐색이 진행되며, 보고서는 중국이 이 전환기에 있다고 진단

표 7. 발전단계별 특징과 해석상 유의점

단계	보고서가 제시한 특징	해석상 유의점
건설 열기	시설·본체·시나리오 확대	규모와 품질·활용은 별개
역량 축적	공정 정비·기능 분화·검증 연계	기술·수요 변화로 장기간 조정 가능
생태계 성숙	표준·거래·전문서비스의 정착	현재 완료된 상태가 아닌 전망

자료: CAICT 보고서 내용을 토대로 재구성

- ✦ 모든 시설이 전 과정을 갖춘 종합서비스로 성장하는 것이 유일한 경로는 아니며, 운영여건에 따라 기능을 줄이고 전문성을 높이는 선택도 가능
- 비용을 감당하기 어렵거나 데이터 상품이 충분하지 않은 시설이 수집도구·모델·본체 시험 등 특정 영역으로 분화할 수 있다고 전망
- 전문화는 실패가 아니라 지속적인 수요가 있는 기능을 깊게 수행하는 경로가 될 수 있으나, 실제 성립 여부는 고객과 제공가치에 좌우
- 분산형 실제환경, 공정 자동화, 준법 거래, 통합서비스와 전주기 표준화의 다섯 방향은 서로 연결돼 있으나, 모든 시설에 동일한 순서와 비중을 요구하는 처방이 아니라 참고 기준

VI

중국 체화지능 훈련장의 주요 변화

1 시설 확대에서 검증 가능한 산업 기반으로

- ✦ 중국의 훈련장 확산은 로봇 개발에 필요한 물리적 경험을 개별기업의 내부 활동에서 공동 활용 가능한 산업서비스로 조직하려는 움직임

 - ▶ 데이터의 생산을 설비·인력·업무절차와 연결하고, 모델과 본체의 실패를 다시 수집과 훈련에 반영한다는 점에서 단순한 데이터 저장소와 구별
 - ▶ 공개된 현황은 건설 확대와 함께 운영모델·데이터 공정·거래방식·표준화가 동시에 실험되는 전환기를 보여주며, 보고서 역시 이를 건설 열기에서 역량 축적으로 넘어가는 단계로 진단
- ✦ 분석에서 확인되는 첫 번째 전환은 평가 초점이 양적 생산능력에서 학습 활용성과 실제 업무 대표성으로 이동하는 흐름

 - ▶ 보고서는 생산 확대 전망과 별개로 정렬·주석·검수 등의 처리 병목을 지적하며, 이는 수집 규모와 실제 학습에 사용하는 규모의 차이를 시사
 - ▶ 다양한 산업명을 갖추는 것보다 각 업무의 물체·순서·간섭·실패조건을 깊이 포함하는지가 중요하며, 이는 시나리오 동질화 문제와 직접 연결
- ✦ 두 번째 전환은 데이터 납품을 넘어 모델 개선과 현장 적합성 확인을 서비스에 포함하려는 방향이며, 수요기업과의 관계 비중이 확대

 - ▶ 장비와 연산자원의 공유만으로 해결되지 않는 업무 정의·운영제약·검증조건은 실제 사용자가 제공하는 지식과 참여를 통해 구체화 가능
 - ▶ 종합형·특화형과 중심시설·분산형 실제환경 지점은 동일한 시설모델을 복제하기보다 공통 데이터·검증 기능과 산업별 현장 적응을 분담하는 방향으로 조직될 가능성이 큼
- ✦ 세 번째 전환은 표준화와 운영수익을 기술개발 이후의 부수과제가 아니라 인프라 지속성의 핵심 조건으로 다루는 방향

 - ▶ 자료를 비교·결합하고 거래하려면 품질·권리에 대한 신뢰가 필요하고, 훈련·검증 서비스가 지속되려면 고객이 지불할 가치와 운영비의 관계 성립이 전제
- ✦ 향후 확인 대상은 데이터 품질과 검증능력이 훈련서비스·현장배치·거래 수요로 이어지는지 여부

주요 참고자료

- [1] 中国信息通信研究院人工智能研究所·人形机器人（上海）有限公司·具身智能测试实验室,
《具身智能训练场研究报告（2026年）》，2026.8., 공개 8.18., PDF 56쪽(본문 인쇄쪽수 1~48쪽).
https://www.caict.ac.cn/kxyj/qwfb/ztbg/202608/t20260818_735676.htm
- [2] 중국 공업정보화부 과기사(科技司), 「《国家人形机器人产业标准体系建设指南（2026版）》(征求意见稿)
의견수렴 공시」, 2026.8.24. 작성, 8.25. 공개(공시기간 8.25~9.23.), 의견수렴안 15쪽. 본 보고서의 분석
대상은 초안의 내용과 목표이며, 해당 문서는 의견수렴 단계로 최종 확정 전.
https://wap.miit.gov.cn/gzcy/yjzj/art/2026/art_cb9cca921b2946c595d897b14bf520a6.html
- [3]工业和信息化部办公厅·国务院国资委办公厅,
「关于联合开展2026年度人形机器人与具身智能实景实训专项行动的通知」, 工信厅联科函[2026]256号, 작성
2026.6.3., 공개 6.8.
https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2026/art_f291ccd3da4c47ce95741de63cc088e6.html
- [4] ITU, 「ITU Focus Group on Embodied AI for Multimedia Technologies」, FG-EAI 설립일 및 WG2
담당 분야 안내.
<https://www.itu.int/en/ITU-T/focusgroups/eai/Pages/default.aspx>
- [5] 全国标准信息公共服务平台, 「中华人民共和国行业标准备案月报 2026年第8号（总第316号）」, 2026.9.1.,
부표 1의 YD/T 6771-2026 등록사항.
<https://hbba.sacinfo.org.cn/rnDetail/87b39be6d031947ddb7ea82bb1b3a045f3b8fce04e0824396d490fd0f93d73a>