PAVS

Physical AI Verification & Scoring

물리 채점 · 검증 · 성과 보고 공개 규격 스위트

OPEN SPECIFICATION SUITE

Part 0 — 백서 (informative)

발행본 v1.1 · 2026. 10. 2.

발행: ㈜이머시브캐스트 · 제안 단계의 공개 규격(공식 표준화 기구의 표준이 아님)

immersivecast.com/pavs

목 차

요약

0. 이 문서에 대하여

1. 배경 — 자율주행의 선례와 Physical AI 의 표준 공백

1.1 시뮬레이션 채점이 제도가 된 가장 성숙한 선례

1.2 로봇에는 그 등가물이 없다

1.3 평가 의무는 지금 부과되고 있다

2. 문제 — 성공률 한 숫자는 능력의 증명이 아니다

3. 답 — 채점의 신뢰 사슬 다섯 가지

4. 사슬 1·2 — 문항과 계측

4.1 문항: 시나리오는 자기완결적 도메인 언어로

4.2 계측 — 값의 존재와 측정의 구별

5. 사슬 3 — 판정: 미측정은 제 3 의 상태다

6. 사슬 4 — 검증: 채점기를 채점한다

7. 사슬 5 — 보고: 점수가 아니라 성능 경계

8. 조항의 유래 — 사례집(Part 2)의 위상

9. 적합성 — 시험 방법이 딸린 요구

9.1 모든 요구에 시험 방법을 둔다

9.2 적합성 클래스

9.3 참조 구현

9.4 규격을 만든 쪽이 먼저 시험받는다

10. 생태계 내 위치

11. 공개 규격화 로드맵

참고문헌

부록 A (informative) — 자율주행 선례와 이 규격의 대응

부록 B (informative) — 참조 표준·KS 대응표

B.1 국내 표준 접점에 대한 입장

부록 C (informative) — 인용 시 표현 경계

부록 D (informative) — 용어 해설 (평이 설명)

개정 고지 — v1.0 → v1.1 (2026-10-02)

문의

> **PAVS** (Physical AI Verification & Scoring) — 물리 채점·검증·성과 보고를 위한 공개 규격. **문서 지위**: 발행본 v1.1 · 2026-10-02 · 최초 공개 v1.0 2026-09-09 · 대응 규격 Part 1 v1.3(시험 단언 72 · ★핵심 8) · 공개 위치 immersivecast.com/pavs. 이 문서(Part 0)는 전체가 **informative**(설명적)이다. 규범 조항은 Part 1 에 있다. 이번 판의 변경 사항은 문서 끝의 개정 고지에 둔다. **이해관계 고지**: 발행 주체 ㈜이머시브캐스트는 인증기관·인정기관이 아니며, 이 규격으로 평가받는 구현(Part 4)의 저자와 같다. 중립성은 규격 본문에 특정 제품에 유리한 요구를 두지 않는 방식으로만 담보된다. **집필 원칙**: 본 백서는 규격이 지향하는 목표 상태를 서술한다. 수치는 논증에 필요한 것만 싣고, 사례는 조항이 생긴 근거로만 인용한다. **인용 규약**: 외부 사실은 출처와 시점을 병기한다. 출처 목록은 참고문헌에 둔다.

요약

■ 배경

- 자율주행은 시나리오 도메인 언어와 시뮬레이션 검증을 국제 규정·평가 프로토콜(UNECE R157·NATM·Euro NCAP)의 정식 방법론으로 확립
- 로봇·Physical AI 에는 그 등가물 부재 — 기존 로봇 표준은 물리 시험 전제, 휴머노이드는 관련 안전표준 제정 개시 단계(발행까지 수년 전망 — 업계 추정)
- 평가 «의무»는 제도가 선행 부과 중(EU 기계류 규정 2027-01 전면 적용 · 국내 AI 기본법 2026-01-22 시행 · 실외이동로봇 운행안전인증 2023-11 시행 · 피지컬 AI 국가 전략 2026-07 공개)

■ 문제 인식

- 성공률 단일 수치·데모 영상 중심의 성과 보고는 능력 증명으로 불충분 — 조건 변화에 취약(90%+→0% 붕괴 보고), 관행 시행수의 통계적 무력, 자율성 판별 불가
- 채점 도구·모델 자체의 신뢰성 미검증 시 높은 성적도 무의미 — 실측 사고 기록의 다수가 «오류 표시 없이, 타당해 보이는 형태로, 조용히 틀리는» 유형

■ 제안

- 채점의 신뢰 사슬 5 단(문항–계측–판정–증거–보고)을 요구사항으로 규정하는 공개 규격 PAVS 제안
- 전 요구사항을 시험 단언 72 건으로 정의하고 72 건 전부에 시험 방법 규정(주입 36 · 심사 33 · 시연 3) — 적합성은 이 절차로 수행한 시험의 기록으로 판정, 문서 검토만으로는 판정하지 않음
- 모든 조항에 유래 근거 대응 — 자체 운영 사고, 공개 문헌의 실패 기록, 또는 선례의 수용 조건(대응표 Part 1 부속서 B · 사고 서술 Part 2)
- 발행 기관이 자기 구현에 먼저 적용 — 2026-09-04 자기 시험 결과 72 건 중 충족 7 건 · ★핵심 8 건 중 3 건, 「PAVS 지향」 표기, 미충족 전건 공개

■ 구성 및 활용

- Part 0 백서(설계 사상) · Part 1 규격 본문(요구·적합성) · Part 2 사례집(해설) · Part 3 참조 구현·재현 부록 · Part 4 솔루션 프로파일(구현 예 — 미공개, 차기 공개)
- 활용처: 검증 체계 구축·조달 요구서 참조·평가 제도 설계의 방법론 층위 — 공개 스펙+참조 구현 → 컨소시엄 → 공식 표준화의 단계적 경로 상정

0. 이 문서에 대하여

PAVS 는 Physical AI — 물리 세계에서 지각하고 움직이고 조작하는 인공지능 시스템 — 의 능력과 안전을 **시뮬레이션에서 채점하고, 그 채점을 검증하고, 그 결과를 보고하는 방법**을 정의하는 공개 규격 스위트다. 본 백서에서는 이 규격이 필요한 배경과 세계 지형을 살펴보고, 규격의 설계 사상과 구성을 제시한다. 스위트는 다섯 부분으로 구성

된다 — 규격 본문은 제품·조직에 중립이고, 구현의 이름은 프로파일(Part 4)에서만 밝힌다. Part 4 는 아직 공개되지 않았으며 차기에 공개한다.

Part 0 백서	Part 1 규격 본문	Part 2 사례집	Part 3 참조 구현	Part 4 솔루션 프로파일
필요성 · 설계 사상 (informative · v1.1)	용어 · 요구사항 72 · 적합성 (normative · v1.3)	조항이 생긴 근거 사례 38 건 (해설 · 사례집 v1.1)	기전 · 시연 3 종 · 재현 패키지 (informative · v0.4, 이번 묶음과 함께 공개)	구현 하나의 형상 · 경계 · 제공 형태 (informative · 미공개, 차기)

규범 어휘: Part 1 은 shall(요구)/should(권고)/may(허용)/can(가능성)의 4 단 어휘를 사용하며, must 는 문서 외부의 법적 제약을 지칭할 때만 쓴다(ISO/IEC Directives Part 2 의 관례를 따른다). Part 1 본문은 비고·보기·informative 표기 절을 제외하고 모두 규범이다. 본 백서를 포함한 나머지 부는 전부 설명 문서다.

지위의 정직한 표기: PAVS 는 제안 단계의 **공개 규격(open specification)**이다. 공식 표준화 기구(ISO·IEC·KS)의 표준이 아니며, 본 문서에서 "표준 시나리오", "표준 문항" 등의 표현은 **이 규격이 정의하는 대상**을 가리킨다. 기존 공식 표준(ISO 9283, ISO 10218, ISO 18646, ISO 3691-4 등)은 **참조 표준**으로 인용하며, 이들에 대한 적합성 인증을 의미하지 않는다.

발행 주체와 이해관계: 발행 주체 ㈜이머시브캐스트는 인증기관·인정기관이 아니며, 이 규격으로 평가받는 구현(Part 4)의 저자와 같다. 규격을 쓴 곳과 그 규격으로 평가받는 곳이 같으므로, 중립성은 규격 본문에 특정 제품에 유리한 요구를 두지 않는 방식으로만 담보되고 발행 주체의 독립성으로는 담보되지 않는다(Part 1 12.1). 제 3 자 발행 주체로의 이관은 이 규격의 목표이며 현재 판에서는 이루어지지 않았다.

1. 배경 — 자율주행의 선례와 Physical AI 의 표준 공백

> **자율주행은 시뮬레이션 검증을 제도의 정식 방법론으로 만들었다. 로봇·Physical AI 에는 그 층위가 비어 있고, 평가 의무는 이미 부과되고 있다.**

1.1 시뮬레이션 채점이 제도가 된 가장 성숙한 선례

자율주행은 시뮬레이션 검증 결과를 제도가 인정한 가장 성숙한 선례다. 타 산업의 시뮬레이션 수용 — 항공 시뮬레이터 자격, 의료기기의 전산 모델링 증거 — 은 도구 자격·개별 증거의 층위에 머무는 반면, 자율주행은 시나리오 기반 검증을 양산 제품 형식승인의 방법론으로 명문화했다.

- **UNECE R157**(2021)은 레벨 3 자동차선유지시스템의 국제 형식승인에 "가상 시뮬레이션 시나리오 검증 → 폐쇄 트랙 → 실도로"의 단계형 구조를 도입했고, 시뮬레이션 도구의 범위·유효성(validity) 입증과 툴체인 검증(물리시험과의 상관)을 요구한다.
- **UNECE NATM**(2021-06 채택)은 시나리오 카탈로그와 5 개 검증 기둥(pillar — 가상시험 포함)으로 구성된 다기둥 방법론을 공식화했다.
- **Euro NCAP 2026 프로토콜**(보도자료 2025-11-26, 가상시험 프로토콜은 2025-03 선공개·2026-01 시행)은 능동안전(ADAS) 평가에 가상시험 데이터를 정식 편입한 첫 사이클이다(수동안전 일부 가상시험 데이터 제출은 2024 년부터 의무). 도입 동인은 시나리오 조합의 폭증(약 186% 증가 — Euro NCAP 기술진 인터뷰 기준)으로 전수 실물시험이 비현실적이 된 것이고, 수용 조건은 시뮬레이션-실차 상관성의 입증 — 정합 점수·지표 허용오차·클러스터별 충족률의 정량 기준(일부 잠정치) — 이다.

이 길의 기반에는 **시나리오를 도메인 언어로 쓰는 계보**가 있다. ASAM OpenSCENARIO/OpenDRIVE 는 시험 시나리오를 사람이 읽고 기계가 실행하는 언어로 만들었고, 규정 문안(R157 의 시험 시나리오)이 실행 가능한 코드가 되는 계보를 세웠다.

1.2 로봇에는 그 등가물이 없다

로봇·Physical AI 의 표준 지형은 다르다.

- **측정 항목의 어휘는 이미 있다.** ISO 9283 은 매니퓰레이터의 자세 정확도·반복정밀도를, ISO 18646 시리즈는 서비스 로봇의 이동·내비게이션·조작 등의 성능을(다리 이동·하지 착용형 등 신규 파트 제정 진행 중), ISO 3691-4 는 무인 산업차량의 안전 요구를, ISO 10218:2025 는 산업용 로봇 안전(협동 운용 통합)을 정밀하게 규정한다.
- **그러나 이들 표준은 물리 시험을 전제한다.** "이것을 재라"는 있으나 "시뮬레이션에서 이것을 어떻게 재현하고, 그 측정을 어떻게 믿고, 결과를 어떻게 보고하는가"를 규정한 표준은 확인 범위(2026-08, 공개 자료 기준)에서 발견되지 않았다 — 기능안전 분야의 소프트웨어 도구 신뢰 관례는 있으나, 로봇 성능·안전 문항의 시뮬레이션 채점·보고 층위는 다루지 않는다.
- **인접 규범은 존재하며, 이 규격은 그것을 계승한다.** ISO/IEC 9646 계열은 적합성 시험의 3 상 판정(pass/fail/inconclusive)과 구현 적합성 진술(ICS)을 확립했다. ISO/IEC TS 42119-2:2025 는 AI 시스템의 시험 프로세스를, ISO/IEC 17025 는 시험소의 역량·불확도·판정규칙을, ASME V&V 40 과 NASA-STD-7009 는 계산 모델의 신뢰도 평가를 규정한다. 이 규격은 이들을 계승하고 그 아래 층(시뮬레이션 물리량 계측의 출처와 판정의 성립)을 더한다(Part 1 부속서 D). 선행 표준이 규정했으나 이 규격이 아직 갖추지 않은 사항(불확도 구간의 판정규칙, 리스크 비례 신뢰도)도 같은 부속서에 밝힌다.
- **휴머노이드는 표준 자체가 만들어지는 중이다.** IEEE 휴머노이드 표준 스터디그룹(2025-09 보고서)은 기존 로봇 안전 표준이 동적 불안정 보행체를 다루지 못한다고 명시했고, 관련 ISO 안전표준 제정(동적 안정 모바일 로봇 — 휴머노이드·다족 포함, ISO 25785-1)은 개시 단계(2025-05 개시 발표)이며, 발행까지 수년의 기간이 업계에서 전망된다.

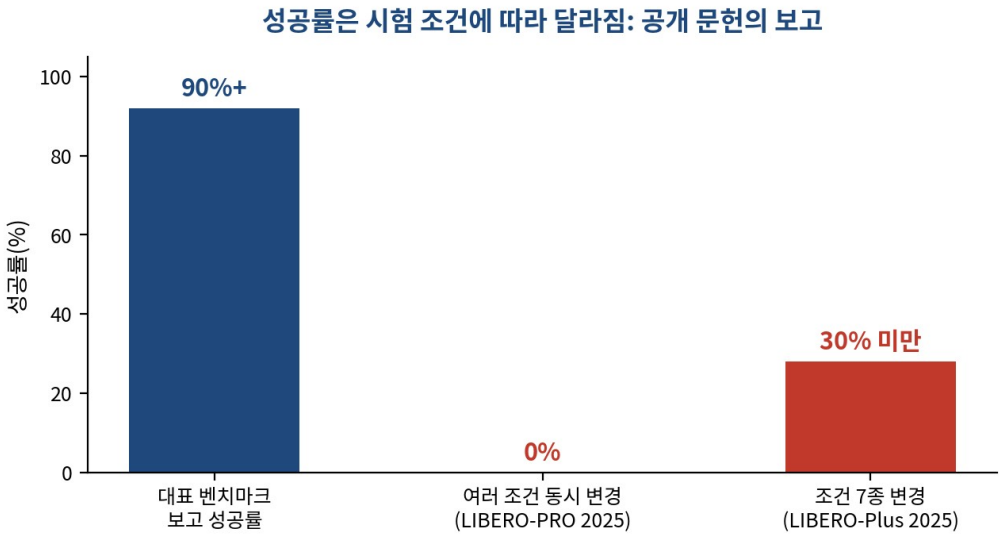
1.3 평가 의무는 지금 부과되고 있다

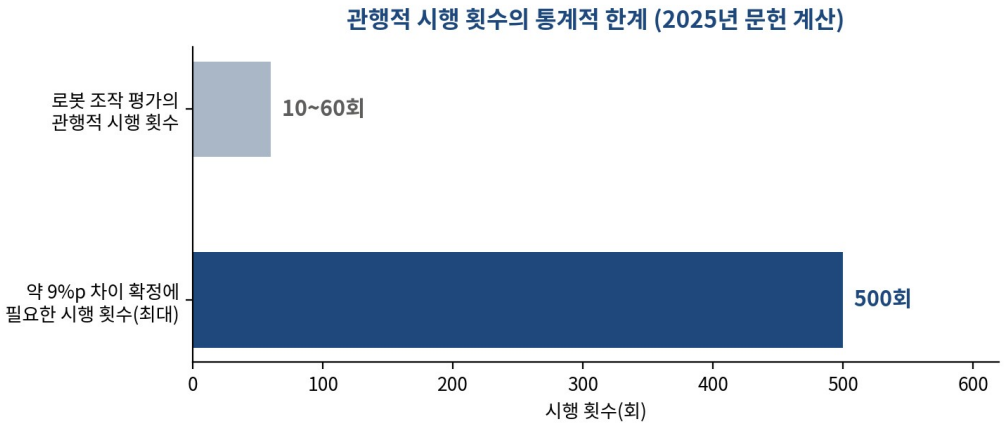
방법론의 공백과 무관하게, 평가의 **의무**는 제도가 이미 부과하는 중이다.

- EU 기계류 규정(2023/1230)은 2027-01-20 전면 적용되며, 학습에 따라 거동이 바뀌는(완전·부분 자기진화) 머신러닝 안전 기능 컴포넌트에 제 3 자 적합성 평가를 요구한다 — 시험 방법의 조화표준은 정비 중이다.
- 한국은 AI 기본법(「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」, 2026-01-22 시행)으로 같은 방향에 있고, 정부는 「피지컬 AI 핵심 경쟁력 확보 전략」(과학기술정보통신부, 2026-07-01 공개 — 데이터 확보·핵심기술·서비스 확산·생태계 조성의 4 대 축)으로 방향을 공식화했다. 실외이동로봇 운행안전인증(지능형로봇법, 2023-11 시행)은 "채점을 통과해야 공공 공간 진입권을 얻는" 구조를 이미 운영한다 — 다만 그 시험은 실기(實機) 기반이며, 인증 기준에 시뮬레이션 결과를 인정하는 조항은 확인되지 않았다(2026-09 조사).
- 국가 가상 실증 **인프라**는 준비되고 있다 — 국가로봇테스트필드는 실증지원 플랫폼에 가상환경(Simulation) 실증평가를 명시하고 첫 과제로 디지털 트윈화 프레임워크 기술개발(2024~2028)을 착수했다. 인프라(장소·장비)와 **채점 방법론**(문항·임계·증거 규격)은 다른 층위이며, 본 규격이 다루는 것은 후자다 — 인프라 투자와 경쟁하지 않고 정합한다.

> **비고** — 이 절의 논증 구조: "재야 할 것"(기존 표준)과 "재야 할 의무"(신규 제도)는 갖춰지는 중인데, "시뮬레이션으로 재고 믿고 보고하는 방법"이 비어 있다. PAVS 는 그 층위를 채우는 규격이다.

2. 문제 — 성공률 한 숫자는 능력의 증명이 아니다





> **성공률 한 숫자는 조건의 함수이고, 데모 영상은 자율성조차 증명하지 못한다 — 채점은 문항·계측·판정·증거·보고의 사슬이다.**

Physical AI 의 성과는 오늘날 대부분 두 형태로 보고된다: **데모 영상**과 **벤치마크 성공률**. 둘 다 채점이 아니다. 최근 연구가 이를 정량으로 보였다.

1. **성공률은 조건의 함수다.** 대표 벤치마크에서 90% 이상을 기록한 모델들이 결합 섭동 평가에서 성공률 0%로 붕괴했고, 물체 위치 이동 단일 축만으로도 일정 변위 이상에서 0%로 급락했다(LIBERO-PRO, 2025). 의미 없는 문자열 지시에도 정상 지시와 동일한 궤적으로 "성공"했다 — 저자들의 결론은 과제 이해가 아니라 "행동 순서의 암기"였다. 온건한 섭동 7 축(배치·시점·초기상태·조명 등)만으로 성공률이 95%에서 30% 미만으로 급락하는 결과도 보고되었다(LIBERO-Plus, 2025).

2. **관행적 시행수는 통계적으로 무력하다.** 로봇 조작 정책 비교의 관행은 10~60 회 시행이다. 약 9%p(59% 대 68%)의 성공률 격차를 확정하는 데 최대 500 회가 필요하다는 계산이 제시되었다(2025). 결과가 애매할 때 시행을 추가하는 관행은 통계 보증을 무효화한다.

3. **영상은 자율성조차 증명하지 못한다.** 상용 휴머노이드의 공개 데모가 원격조종 논란에 반복적으로 휘말렸다(2024~2025 — 일부는 정황 기반 해석으로 해당 제조사의 공식 인정이 없고, 다른 일부는 제조사 공식 안내와 사

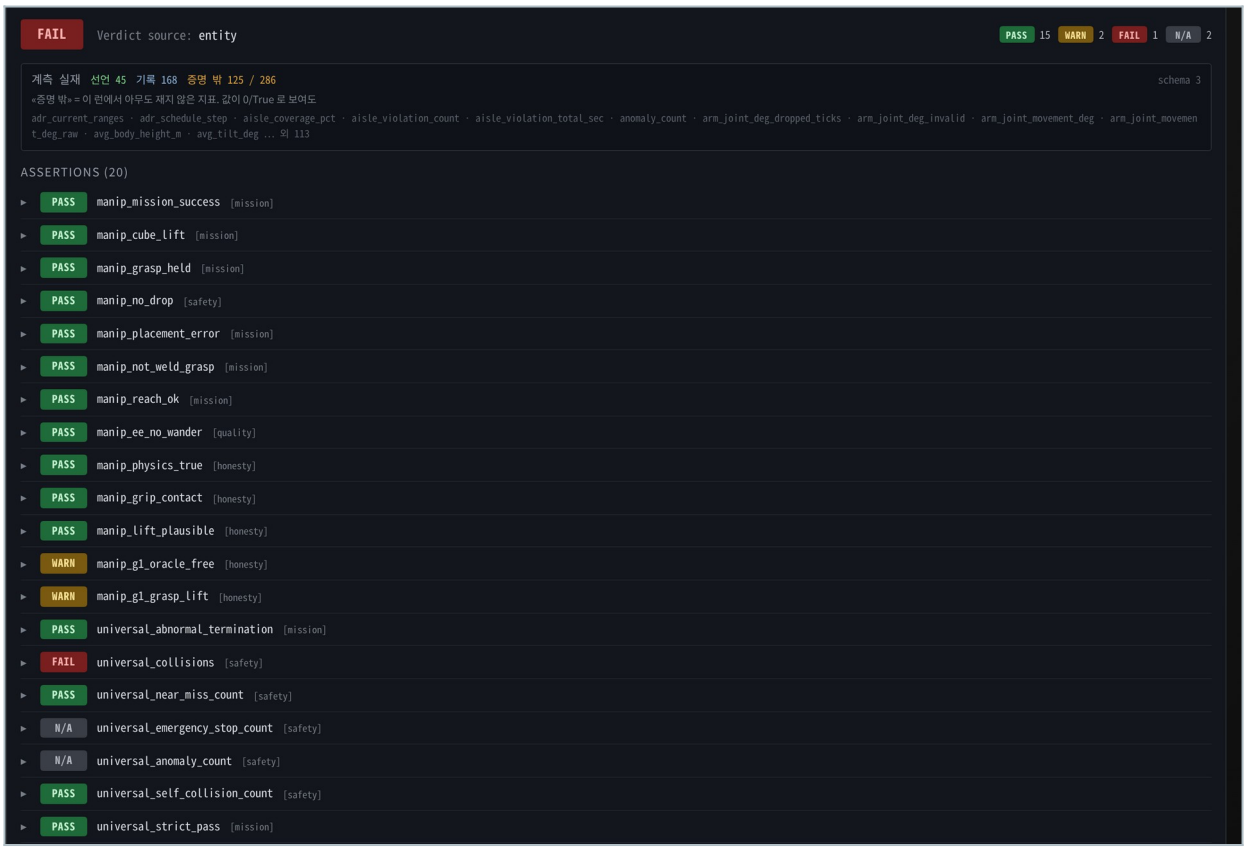
용 승인 절차에 원격 지원 동의가 명시되어 있다). 조작 주체의 로그 없이, 영상만으로는 "누가 했는가"도 판별되지 않는다.

4. **채점 도구 자체가 판정을 바꾼다.** 같은 천(변형체) 조작 과제에 대해 4 개 물리 엔진의 실측 대비 오차·안정성이 서로 다름을 실측한 교차 시뮬레이터 벤치마크가 동료심사를 거쳐 발표되었다(RA-L 2024). 공개 보행 학습 프레임워크(Humanoid-Gym)는 한 물리 엔진에서 학습한 정책을 다른 엔진에서 다시 검증하는 절차를 기본으로 둔다.

산업계의 대규모 평가 실무도 같은 방향을 가리킨다. 대규모 blind A/B, 초기조건 통제, 루브릭 기반 부분점수, 그리고 **평가 자체의 품질 관리**(재검율·루브릭 불일치율 계측)를 갖춘 평가 체계가 공개되며 성공률만으로는 정책 성능의 전모를 말하지 못한다고 명시했다(2025).

> **비고** — 학계의 수렴 방향은 분명하다: ① 단일 성공률 탈피(부분점수·실패 모드) ② 채점 도구 신뢰성의 계측 ③통계 보증 ④섭동·강건성의 문항화 ⑤평가의 제 3 자화. PAVS 는 이 방향을 **하나의 운영 가능한 규격**으로 통합하는 시도이며, 독자적 발견이 아니라 국제 연구가 요구하는 것을 실행 형식으로 옮긴 것이다.

3. 답 — 채점의 신뢰 사슬 다섯 가지



한 시험의 판정 화면 — 문항이 각각 합격·경고·불합격·해당없음으로 갈린다. 위쪽 상자는 이 시험에서 측정되지 않은 지표 목록을 그대로 공개한다. 값이 0 이나 «참»으로 보여도 측정된 값이 아니라는 표시다.

> **시험 성적의 신뢰는 다섯 요건이 순차로 충족될 때 성립한다. PAVS 의 구조가 그 다섯 요건이다.**

시험 성적의 신뢰는 다섯 요건이 순차로 충족될 때 성립한다. PAVS 의 구조는 이 다섯 요건을 그대로 반영한다.

#	사슬	요구 내용	Part 1 의 대응 절
1	**문항** (Item)	측정 대상의 정의 — 문항의 명확성·실행 가능성·변별력이 검증된다	시나리오·문항 요구
2	**계측** (Measurement)	값의 출처 — 채점에 사용되는 값이 실측임이 증명된다	계측 출처(provenance) 요구
3	**판정** (Verdict)	합격 기준 — 미측정은 합격도 불합격도 아닌 제 3 의 상태로 처리된다. 필수 문항에서 계측이 없으면 «계측 부재» 사유의 불합격으로 처리할 수 있고, 미측정 건수는 별도로 보존·표기된다	판정 규칙 요구
4	**검증·증거** (Evidence)	판정의 근거 — 채점기 자신이 검증되고 결과가 재현된다	검증·재현성 요구
5	**보고** (Reporting)	결과의 기술 방식 — 성적서가 자기 한계와 귀속을 명시한다	성과 보고 요구

각 사슬은 다음 절에서 목표 상태로 서술한다. 운영 요구(실패한 방법의 기록부·기록 규율)는 다섯 사슬을 지탱하는 횡단 지원층으로, Part 1 에서 별도 군(10 장)을 이룬다. 4~7 장은 Part 1 규범 조항의 요지를 옮긴 것이며, 이 백서는 요구사항을 만들지 않는다. 요구의 정본 문면은 Part 1 이다.

4. 사슬 1·2 — 문항과 계측

> **규격의 실체는 문서가 아니라 실행 가능한 시나리오 묶음(코퍼스)이고, 값의 출처를 묻지 않는 채점은 거짓 합격과 거짓 불합격 양쪽으로 틀린다.**

4.1 문항: 시나리오는 자기완결적 도메인 언어로

시험 시나리오는 사람이 읽고 기계가 실행하는 **도메인 언어**로, 자기완결적으로 기술되어야 한다. 제 3 자가 동일한 기술(記述)만으로 동일한 시험을 재현할 수 있어야 하며, 시나리오 로직을 외부 보조 코드에 분산하거나 하드코딩하여서는 안 된다. 이는 자율주행이 OpenSCENARIO 로 확립한 관행의 로봇판이다.

규격의 실체는 문서가 아니라 코퍼스다. 실행 가능한 시나리오, 대조군, 기대 판정을 한 묶음으로 갖춘 집합(코퍼스)이 규격의 실물이며, 규격의 산출물에는 기계가 읽는 시나리오 라이브러리가 포함되어야 한다(ASAM 이 도메인 모델 라이브러리를 공식 산출물로 동봉하는 관례).

문항 자체도 검증 대상이다. 문항의 가치는 채점을 수행했다는 사실이 아니라 합격과 불합격을 실제로 가려내는 능력(변별력)에 있다. 어떤 입력에도 합격이 나오는 문항(늘 참인 명제, 곧 항진명제와 같은 문항)은 결함을 잡지 못하고, 불합격이 한 번도 나온 적 없는 문항은 합격과 불합격의 경계가 맞는지 확인되지 않은 상태다. 규격은 시험 이력으로 문항을 넷으로 분류하고(변별함 · 상수합격(늘 합격) · 상수불합격(늘 불합격) · 자격없음), 한쪽 판정만 나와 온 문항의 경계는 **반대 방향 대조군**(원 시나리오에서 자극만 반대 방향으로 바꾼 시나리오)으로만 확인할 것을 요구한다. 임계값을 바꾸어 원하는 판정을 만드는 방식은 허용되지 않는다.

정격 능력은 지속 가능성으로 확인한다. 제조자가 공표한 정격 능력(하중·속도·연속 운전 등)마다, 그 능력이 선언된 조건에서 지속되는지를 확인하는 문항이 있어야 한다. 한 번의 성공은 정격 충족의 근거가 되지 못한다. 열·전류·마모처럼 시뮬레이션이 재현하지 않는 요인에서 오는 한계는 실물 시험으로만 확인할 수 있는 문항으로 표기한다(**PAVS-ITM-10**).

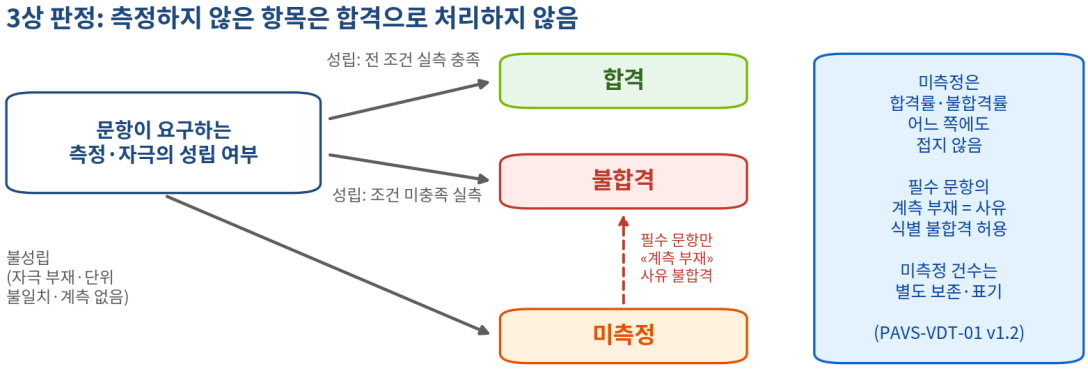
4.2 계측 — 값의 존재와 측정의 구별

채점에 쓰이는 모든 값은 **측정의 증명**(provenance)을 동반하여야 한다.

- 값의 출처(실측/명령값/기본값/추정)가 선언되고, 실측이 아닌 값은 실측 문항의 채점에 쓰여서는 안 된다.
- 결측은 0 이 아니다. 표본이 없으면 값은 "없음"이어야 하며, 없는 값을 0 이나 기본값으로 채워서는 안 된다 — 급정지를 수행하지 않았음에도 정지 시간 0 초로 기록되어 만점이 산출되는 구조를 금지한다.
- 계측 채널은 실제로 신호에 반응하는지(**살아있음**, liveness)를 확인받아야 한다. 입력 상태가 변하는데 계측값이 그대로이면 그 값은 유효한 신호가 아니라 계측 고장의 표시다.
- 계측 채널이 정상인지는 채널 안에 설정된 값을 다시 읽어 보는 방식(자기 일관성 검산)으로 확인하지 않는다. **크기를 아는 기준 자극을 넣고, 그 자극이 계측 경로의 맨 끝 출력(최종 기록)까지 도달하는지** 보는 독립 검사로 확인한다.
- 물리량을 직접 재지 않고 계산 설정에 따라 값이 달라지는 대리 지표(예: 시뮬레이션 계산 간격에 반비례하는 접촉력 수치)는 비교할 수 있는 범위와 조건을 밝혀야 하며, 조건이 다른 결과끼리의 절대값 비교에 써서는 안 된다.

> **보기** — 값의 출처를 묻지 않으면 양방향으로 틀린다. 한 실측 사례에서, 최고속도 지표가 실은 주행 명령값에 계수를 곱한 것이었다: 실제 1.4 m/s 로 달린 로봇의 최고속도가 서류상 4.7 m/s 로 기록되어(2026-08 실측) 안전 조항의 거짓 불합격이 산출될 수 있는 상태였고, 존재하지 않는 과속이 위반 계수기를 실제로 켰다. 반대 방향으로는, 아무도 재지 않은 충돌 횟수 "0"이 안전 문항을 통과시켰다(거짓 합격). 값의 출처 선언은 이 양방향 오류를 같은 장치로 막는다.

5. 사슬 3 — 판정: 미측정은 제 3 의 상태다



> **재지 않은 것을 통과로 접지 않는다 — 미측정은 합격도 불합격도 아닌 제 3 의 상태다.**

판정 규칙의 핵심은 하나다: 재지 않은 것을 통과로 접지 않는다.

- 판정은 합격/불합격의 2 상이 아니라 **합격/불합격/미측정**의 3 상이어야 한다. 이 3 상 판정은 ISO/IEC 9646 의 pass/fail/inconclusive 를 계승한 것이다. 미측정은 합격으로 계상되지 않으며, 합격률·불합격률의 분모·분자 어디에도 접히지 않는다. 다만 필수 문항(그 계측 없이는 합격이 성립하지 않는 문항)에서 계측이 없으면 «계측 부재» 사유의 불합격으로 처리할 수 있다. 이 경우에도 미측정 건수는 계측 부재 불합격과 구별되는 별도 상태로 보존·표기한다.
- 합격은 필수 조건 전부의 **실측 충족이라는 긍정 증거**로만 산출되어야 한다. 판정 로직의 기본값이나 조건 미달 경로가 합격을 반환하여서는 안 된다.
- 성공 지표는 그것이 전제하는 **물리적 사건의 독립 증거와 교차검증**되어야 한다. 로봇의 실행 장치가 실제로 움직이지 않았는데(no-op) 성공이 계상되는 구조를 금지한다.
- 판정을 내리는 단계가 여럿이면, 운영자는 그 단계들의 엄격성 순서를 문서로 밝혀야 한다. 더 엄격한 단계가 낸 불합격·경고를 더 관대한 단계가 합격으로 덮어쓸 수 없다.
- 측정 실패는 대체값·모의값으로 메우지 않고 실패로 드러내어 보고한다(fail-loud). 물리 시뮬레이션으로 채점한다고 밝힌 시험에서, 물리 계산 경로의 결함을 피하려고 물리 계산을 거치지 않는 경로(예: 위치를 직접 지정해 움직이는 운동학 경로)로 낸 결과를 증거로 쓰는 것을 금지한다.
- 채점기는 피시험체의 성능 한계를 코드에 상수로 내장하여서는 안 된다. 모든 제한은 피시험체의 선언 사양(기계 판독 가능한 매니페스트)에서 유래하여야 한다 — 채점기가 피시험체를 왜곡하는 구조의 금지다.
- **판정 기준의 결함이 수정되면, 결함 기간의 기존 판정 전수를 새 기준으로 소급 감사하고 결과를 기록하여야 한다.**

> **보기** — 마지막 조항의 탄생 사고: 한 채점 체계에서 성공 판정의 기본 경로가 조건 미달에도 True 를 반환하고 있었다. 수정 후 소급 감사 결과, 결함 기간의 합격(총 401 건) 중 89%가 거짓 합격이었다(2026-05 소급 감사). 소급 감사 의무가 없었다면 이 수치는 존재하지 않았고, 거짓 합격은 실적으로 남았을 것이다.

6. 사슬 4 — 검증: 채점기를 채점한다

> **채점 체계의 신뢰는 주장이 아니라 채점기 자신이 통과한 시험에서 나온다.**

채점기 검증의 요구는 다음과 같다.

- **음성 대조군의 상설화**: 일부러 위반하는 시나리오가 정확히 그 문항으로 불합격해야 채점기가 합격이다. 기대 판정(expect)과의 자동 대조를 회귀 스위트에 상주시킨다.
- **적격성 시험**: 채점기 변경 시, 알려진 정답 세트(골든셋)의 재채점이 기존 판정과 일치함을 증명하여야 한다(판정 이동은 세어서 공개).
- **비교 시험의 실효값 확인**: A/B 시험의 독립변수는 설정 의도가 아니라 시스템이 실제 적용한 실효값으로 각 시험군에서 확인되어야 한다. 실효값 미확인 시험은 반증이 아니라 미검증이다.
- **피시험체 탓으로 돌리기 전 계측기 점검**: 시험 실패를 피시험체의 결함으로 결론 내리기 전에, 계측 계통이 정상인지를 먼저 반증 시험으로 확인하여야 한다. 계측 결함은 피시험체 결함처럼 보이기 쉽다.
- **모델 사양의 진위 확인(모델 진본성)**: 피시험체의 시뮬레이션 모델(제원·기하·물성)은 제조자가 확정한 사양과 대조 검증되어야 한다. 제조자가 같은 항목에 여러 값을 공표하면 기준으로 삼은 값과 그 이유를 밝히고, 모델 값이 기준과 다른 항목마다 차이의 근거를 기록한다. 근거가 기록되지 않은 차이는 미검증으로 취급하며, 대조에서 불일치가 확인된 모델의 결과는 적합 주장의 근거로 쓰지 못한다(**PAVS-VRF-05**). 같은 로봇·같은 과제에서 공개 배포 제원과 제조자 확정 제원이 정반대 판정(실기로는 불가능한 자세의 거짓 합격 ↔ 실행 거부)을 낳는 것이 실측으로 확인되어 있다(2026-08 실측). 제 3 자를 채점하는 체계에서 모델 사양의 진위는 채점의 전제다.
- **실패한 방법의 기록**: 실패한 검증·계측 방법과 그것이 틀렸다는 근거는 공식 기록부(실패 접근 등기부)에 남기고, 새 조사를 시작하기 전에 그 기록부를 조회하는 단계를 절차로 둔다. 잘못된 진단을 정정할 때도 원래 기록을 지우지 않고 정정 기록과 연결해 보존한다. 오진의 이력도 다음 검증의 근거 자료가 된다.
- **재현성**: 모든 판정은 명령 한 번으로 다시 계산해 재현할 수 있어야 하며, 시드·구성·채점 코드 판번이 결과에 고정 기록되어야 한다. 채점 코드가 바뀌면 그 코드로 산출된 기존 판정이 여전히 유효한지 재평가하여야 한다.
- **자기 검사의 상시 실행**: 음성 대조군·주입 검사 등 채점 체계의 자기 검사 장치는 정해진 주기나 채점 체계가 바뀌는 시점에 자동으로 실행되어야 한다. 실행되지 않은 검사의 결과를 «위반 없음»으로 보고하여서는 안 되며, 자동 실행이 보호 대상 판정 경로 전체를 덮지 못하면 빠진 범위를 밝힌다(**PAVS-VRF-16**).

> **보기** — 모델 사양의 결함은 오류 없이 지나간다. 공개 배포된 한 로봇 모델에서 팔 링크 전부의 중력이 꺼져 있었다. 시뮬레이션은 오류 없이 돌았고, 그 모델로 잰 정밀도는 팔 자체의 무게가 빠진 값이었다. 시험 하중을 받는 링크의 중력 상태를 확인하는 검사(이상이 있으면 실행을 실패로 드러내는 fail-loud 검사)가 이를 드러냈다. 조치는 모델 정규화, 재측정·재판정, 링크별 중력 상태의 기록과 사전 점검의 순서로 이루어졌다(2026-10, 사례집 D12).

> **보기** — 자기 검사 조항이 생긴 경위: 규격의 이전 판은 자기 검사 장치를 «갖출» 것만 요구했다. 실제로는 구현되어 있으나 자동 실행 지점에 연결되지 않은 검사기가 있었다. 2026-08-26 자기 점검이 이를 드러냈고, 그 공백을 메우려 **PAVS-VRF-16** 을 신설했으며, 발행 기관 자신의 미충족으로 계상했다.

> **비고** — 시뮬레이션 결과를 신뢰할 근거로 자율주행 제도는 "도구의 적합성 입증"을 요구했다(R157). PAVS 도 같은 구조를 따른다. 시뮬레이션과 실기의 일치(정합)는 주장이 아니라 **측정의 대상**이다(상관을 계량하는 학계 선례: SIMPLER 2024). 교정은 파라미터를 늘려 값을 맞추는 방식보다, 모델에 빠져 있는 물리 현상(누락 물리 — 예: 관절 마찰)을 찾아 넣는 방식을 우선한다. 교정값은 교정에 쓰지 않고 따로 남겨 둔 시험 조건(보류 시험, held-out)에서도 맞을 때에만 정

> 식 값으로 채택한다.

7. 사슬 5 — 보고: 점수가 아니라 성능 경계

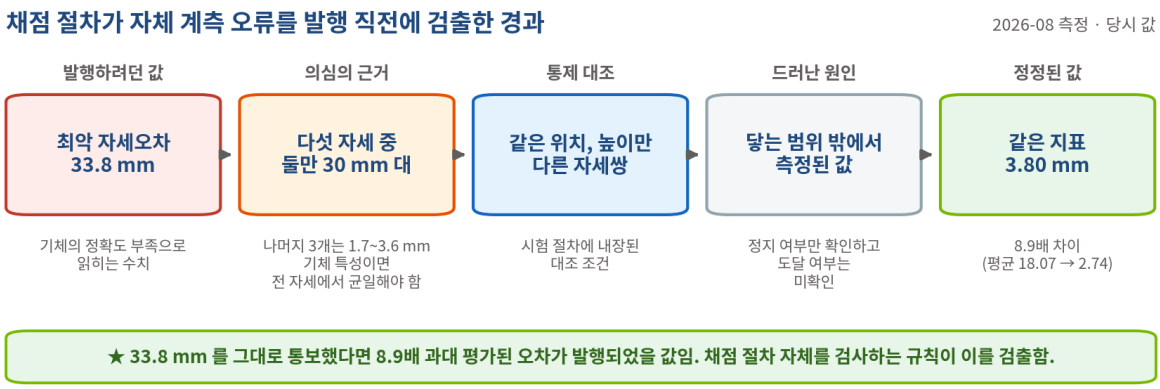
> **성과 보고의 표준은 «몇 점»이 아니라 «이 점수의 신뢰 범위»를 문서가 스스로 기술하게 하는 것이다.**

성과 보고의 요구는 다음과 같다.

- **귀속 공개**: 성적서는 각 문항의 판정 근거가 권위 있는 참조 표준에서 왔는지, 발행자의 자체 정의인지 문항 단위로 표기하여야 한다. 발행 체계에 등재되지 않은 표준 이름, 곧 존재하지 않는 표준 이름으로 성적서를 발행하는 것을 금지한다.
- **커버리지 공개**: 측정한 항목과 측정하지 못한 항목을 비율이 아니라 **목록**으로 명시한다.
- **자기선언의 한계**: 시나리오가 스스로 밝힌 성공(자기선언)은 판정을 올리지 못한다. 자기선언만으로 나온 판정은 성과 문서·적합 표기의 근거가 될 수 없다(순환 논증의 차단). 자기선언이 판정을 내리는 방향(시나리오가 스스로 불합격을 밝힌 경우)은 허용되며, 자기선언이 총평에 기여한 경우 그 사실을 문항 단위로 표기한다.
- **경계 설명**: 합격의 "뜻"(런타임 보장인지, 능력 경계인지, 자극이 없어 시험되지 않은 것인지)을 문항 단위로 실어야 한다. 그 설명과 모순되는 실측이 나오면 자동으로 지목되는 장치를 두는 것이 권고된다(**PAVS-REP-14**).
- **통계 규율**: 성공률류 수치는 시행수·신뢰구간과 함께만 보고한다. 개선 주장은 널대조군 결과를 함께 보고해야 성립한다. 널대조군은 주장된 개선 효과만 뺀 채 같은 처리 경로를 그대로 거친 비교 조건으로, 처리 경로 자체가 만드는 차이와 실제 개선 효과를 구별하게 한다.
- **한계의 명시**: 알려진 단순화·근사·휴면 결함은 유효범위와 왜곡 방향(낙관/비관)을 명시해 공개한다.
- **실기 대조 단계의 표기(★핵심)**: 시뮬레이션 성적을 보고하는 문서는 그 시뮬레이션 모델이 실기와 어디까지 대조되었는지를 표기하여야 한다 — ①자기 일관성만 확인됨 ②다른 계산 방식·모델과의 교차 대조까지 확인됨 ③같은 조건의 실기 계측과 대조되어 차이가 선언된 범위 안에 있음. 상관 증거가 없으면 «없음»으로 적고, 표기 자체를 생략하지 않는다(**PAVS-REP-15**).
- **완료 선언의 규율**: 보고서가 «완료»를 주장할 때에는 약속한 산출물마다 달성·유예·누락 여부를 정해진 형식으로 함께 밝혀야 하며, 부분 달성을 완료로 표기하여서는 안 된다.
- **무결성**: 발행된 성적서는 덧붙이기만 가능한 기록부(append-only 원장)로 관리하여, 기록의 변조·삭제·순서 변경이 기계적으로 탐지되도록 하여야 한다.
- **이해상충의 선언**: 채점 발행자가 피시험체의 조정(adaptation)·훈련에 관여한 경우, 그 사실과 독립성 범위를 성적서에 반드시 선언한다. "중립"은 조정에 관여하지 않은 건에 한해서만 표기할 수 있다.

> **보기** — 같은 시험 실행에 대해 두 경로가 정반대의 성적서를 낸 실측 사례가 있다: 정직성 장치가 붙은 발급기는 불합격을, 다른 화면 경로는 "합격 — 운용 인증됨"을 반환했다. 후자는 빈 입력으로도 동일한 성적서를 생성했다 — 부재를 0으로 치환하고, 기본값을 날조하고, 요청자가 표준 이름을 골랐기 때문이다. 귀속 공개·존재하지 않는 표준 이름의 차단·커버리지 공개 조항은 이 사건에서 나왔다.

8. 조항의 유래 — 사례집(Part 2)의 위상



사례 한 건(2026-08) — 채점기가 발행 직전에 자기 계측의 결함을 잡아낸 과정. 이 규격의 조항은 이런 사건에서 나왔다.

> **이 규격의 조항은 실제로 일어난 오류에서 나왔다. 모든 조항에 유래 근거가 대응한다 — 자체 운영 사고, 공개 문헌의 실패 기록, 또는 선례의 수용 조건이다.**

항공 안전 규정이 사고 조사에서 나왔듯, PAVS 의 조항도 기록된 사건에서 일반화되었다. 모든 조항에는 유래 근거가 대응한다 — 자체 운영 사고, 공개 문헌의 실패 기록, 또는 선례의 수용 조건(대응표는 Part 1 부속서 B, 사고 서술은 Part 2). 그중 자체 운영 사고는 그 조항이 없을 때 **오류 표시 없이 잘못된 결과가 산출된 실제 사건**이다.

Part 2 사례집은 그 기록이며, 각 사례는 다음 4 단 구조로 서술된다.

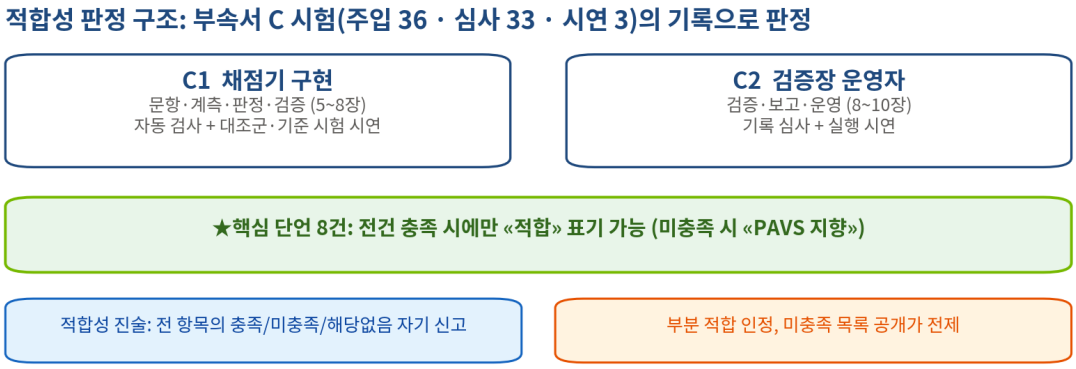
> **문제 상황 → 표준 부재 시의 오류 유형(거짓 합격/거짓 불합격) → 그것을 검출하는 장치 → 재현 방법**

사례는 성과의 나열이 아니다. 다수의 사례가 "채점 체계 자신의 결함을 채점 체계의 검증 장치가 잡아낸" 기록이다.

자기 자신에게 동일한 기준을 적용하는 구조가 이 규격이 요구하는 목표 상태다. 대표 사례군은 다음과 같다.

- 값의 출처(명령값을 측정으로 오인 → 거짓 불합격 / 부재를 0 으로 → 거짓 합격)
- 단위·의미의 혼동(속도 값이 힘 문항을 통과시킴 — 참조 표준의 조항 번호가 같아도 판본에 따라 다른 물리량을 가리킬 수 있음)
- 문항 이름과 실제 계측 대상의 불일치(이격 거리 문항이 실제로는 이동거리를 재고 있었음)
- 판정 기본값이 만든 거짓 합격과 소급 감사(결함 기간 합격의 89%가 거짓 합격 — 2026-05 소급 감사)
- 채점기의 음성 대조군(채점기가 불합격을 낼 수 있는지 확인하는 시험)
- 모델 사양의 진위(공개 배포 제원과 제조자 확정 제원의 정반대 판정 · 팔 링크 중력이 꺼진 공개 모델)
- 계측의 살아있음(고장 난 계측값을 믿고 제어한 결과가 계측 없이 제어한 결과보다 나빴던 사례)
- 안전 장치의 작동 검증(장치가 호출되는지가 아니라 호출되었을 때 실제로 동작하는지 — 주입 시험과 위약 대조)
- 자기 검사의 미실행(검사기는 구현되어 있었으나 자동 실행 지점에 연결되지 않았음)
- 제 3 자 안전의 문항화(시험 기준선에 제 3 자 조건이 없으면 같은 로봇이 같은 경로에서 합격도 불합격도 된다)

9. 적합성 — 시험 방법이 딸린 요구



> **모든 요구에 단언 식별자를 부여하고, 72 건 전부에 시험 방법을 규정한다 — 주입 36 · 심사 33 · 시연 3. 적합성은 이 절차로 수행한 시험의 기록으로 판정하며 문서 검토만으로 판정하지 않는다.**

9.1 모든 요구에 시험 방법을 둔다

ISO 의 원칙 — "요구는 객관적으로 검증 가능하여야 하며, 검증 가능한 요구만 표준에 포함한다" — 을 PAVS 는 다음 단계까지 확장한다.

Part 1 의 모든 요구에는 시험 단언(test assertion) 식별자가 부여되고, 72 건 전부에 시험 방법이 규정된다 — 주입 36 · 심사 33 · 시연 3. 적합성은 이 절차로 수행한 시험의 기록으로 판정하며 문서 검토만으로 판정하지 않는다. "규격의 실체는 코퍼스"라는 설계 사상을 적합성 판정에 적용한 형태다.

각 시험 방법(규범 부속서 C)은 전제 조건, 절차, 합격 기준, 기록 항목, 그리고 **그 시험 자체가 잘못 통과할 수 있는 경로(함정)**를 담는다. 자동 검사로 지정된 단언은 원칙적으로 **주입 시험**(결함을 일부러 넣고 검출되는지 보는 시험)으로 시험한다. 검사기가 설치되어 있다는 사실은 그 검사기가 실제 결함에 반응한다는 증거가 아니기 때문이다. 기간에 걸친 기록의 성질을 보는 단언은 심사 시험으로, 결함 조건을 만들 수 없거나 만드는 것이 안전하지 않은 단언은 시연 시험으로 확인한다.

> **시험 방법이 없는 요구는 재현되지 않는다 — 같은 시스템을 두 시험자가 각각 시험해 다른 결과가 나오면 그 규격은 판정의 근거가 될 수 없다.**

9.2 적합성 클래스

적합성은 역할별로 정의된다(W3C 적합성 관례). 성적을 읽는 소비자를 위한 해석 지침은 규범 적합성 클래스가 아니라 informative 지침(본 백서와 사례집)으로 제공된다.

클래스	대상	적합성의 의미
C1 채점기 구현 (Scoring Engine)	시뮬레이션 채점 시스템	Part 1 5~8 장(문항·계측·판정·검증)의 요구 충족
C2 검증장 운영자 (Range Operator)	시험을 수행·발행하는 조직	Part 1 8~10 장(검증·보고·운영)의 요구 충족

적합성 주장 양식은 Part 1 11.2, 구현 적합성 진술(ICS)의 분모는 부속서 A(72 건)다. 도메인 프로파일(매니퓰레이션·이동·안전 등 도메인별 문항 선택집)은 별도 공개 문서로 선언한다(**PAVS-CNF-05**).

9.3 참조 구현

PAVS 에는 참조 구현이 있다 — 시나리오 도메인 언어 실행, 계측 출처 배선, 3 상 판정 처리부, 음성 대조군 스위트, 덧붙이기만 가능한 성적 기록부를 포함하는 운영 체계다. Part 3(참조 구현·재현 부록, v0.4 — 이번 묶음과 함께 공개)에 그 기전과 시연 3 종을 수록한다. 시연 A 는 CPU 전용으로 명령 1 회에 실행되며 GPU 가 필요하지 않고, 산출을 기준값과 자동 대조한다. 시연 A 의 재현 패키지는 요청 시 제공한다(실행 명령·내부 경로는 공동 시범 시험 상대에게 제공). 세 시연의 공개 수준이 서로 다르다는 사실은 Part 3 §5 에 표로 밝힌다.

참조 구현의 목표는 전 조항 실행 가능성의 증명이며, 현재 충족 현황(2026-09-04 시험 · 72 건 중 충족 7 건)은 적합성 진술로 전건 공개한다. 참조 구현은 채택자가 자기 환경에서 재검산할 수 있는 기준점의 역할도 한다.

규격 본문은 특정 제품·조직에 중립이므로 Part 3 은 구현을 익명으로 다룬다. 그 구현이 제품으로서 갖는 형상·경계·제공 형태는 **Part 4 솔루션 프로파일**에서 이름을 밝혀 서술한다(Part 4 는 미공개 — 차기 공개). 다른 구현도 같은 형식으로 자신의 프로파일을 작성할 수 있다. 미충족 단언을 명시한 적합성 주장의 실제 형태는 적합성 진술(9.4)이 보인다.

9.4 규격을 만든 쪽이 먼저 시험받는다

발행 기관은 자기 참조 구현에 이 규격을 먼저 적용했다. 시험은 2026-09-04 에 부속서 C 의 절차로 수행했다(대상 규격 Part 1 v1.2 · REP-15 는 2026-09-06 재시험 · v1.3 은 단언 요구가 같은 편집 정정판).

- 충족: 72 건 중 충족 7 건 · ★핵심 8 건 중 3 건(MEA-02 · VDT-01 · REP-15)
- ★핵심 미충족: MEA-01 · VDT-02
- ★핵심 판정 불가: VRF-01 · CNF-01 · CNF-02(시험 기록으로 충족을 확인하지 못한 상태)
- 결론: ★핵심 단언이 모두 충족되지 않았으므로 적합을 주장하지 않고 「PAVS 지향」으로만 표기
- 공개: 미충족 65 건 전건을 적합성 진술(2026-09-09 공개 · v1.1 2026-10-02 문면 정정)에 열거

이 결과는 발행 기관의 자기 시험이며 제 3 자 심사가 아니다.

자기 적용은 규격 자체도 바꾸었다. 기록은 다음 3 건이다.

- **PAVS-VRF-16 신설(2026-08-26)** — 자기 점검이 자동 실행 지점에 연결되지 않은 검사기를 드러냄
- **PAVS-REP-15 ★핵심 승격(2026-09-01)** — 발행 기관 자신이 실기 상관 증거 없이 성적 문서를 내고 있었음
- **PAVS-VDT-01·REP-04 문면 정정(2026-09-04)** — 자기 시험이 규격 문면과 규격 취지의 어긋남을 드러냄. 구현을 문면에 맞추면 판정이 관대해지는 방향이라 문면을 고쳤고, 요구 내용 변경으로 고지

10. 생태계 내 위치

> **PAVS 는 대체가 아니라 층위다. 기존 표준이 측정 대상을 규정한다면, PAVS 는 시뮬레이션에서의 측정·신뢰·보고 방식을 규정한다.**

PAVS 는 대체가 아니라 **층위**다.

- **기존 표준과**: ISO 9283·ISO 18646·ISO 3691-4·ISO 10218 이 "무엇을 잴 것인가"의 정본이다. PAVS 는 그 측정 항목을 "시뮬레이션에서 어떻게 재현·측정·채점·보고하는가"의 층위에서 **확장·보완**한다. 각 문항은

참조 표준 번호를 근거 표준으로 연결하며(anchoring), 한국 적용 시 KS 부합화 표준(KS B ISO 10218, KS B ISO 3691-4 등)을 병기한다 — 국제판과 KS 의 판본 시차가 있으므로 이중 표기가 정확하다.

- **물리 엔진·시뮬레이터와**: 특정 엔진의 규격이 아니다. 엔진 간 판정 차이가 실측되는 현실이 교차 엔진 대조와 모델 사양의 진위 확인 조항의 근거다. 시뮬레이터는 로봇을 개발하는 단계의 도구이고, 이 규격은 로봇을 출하·도입하는 단계에서 요구되는 증명을 다룬다 — 둘은 경쟁하지 않고 앞뒤 단계로 이어진다.
- **월드모델·합성 데이터와**: 생성형 증강은 채점의 대체가 아니라 보조 레이어이며, 그 채택 여부도 실측 비교(널대조군 동반)로 정하여야 한다는 것이 본 규격의 통계 규율이다.
- **국가 인프라와**: 가상 실증 인프라(테스트필드)와 채점 방법론은 상보적이다. 인프라 위에서의 채점 대상과 채점 방식이 본 규격이 다루는 영역이다.
- **법정 인증과**: PAVS 는 인증 제도가 아니다. 법정 인증(예: 실외이동로봇 운행안전인증)의 사전 검증·회귀 검증 층위에서 쓰이는 것을 목표 상태로 하며, 인정(accreditation)의 어휘를 참칭하지 않는다.

11. 공개 규격화 로드맵

> **표준화 경로는 공개 스펙 + 참조 구현 → 컨소시엄 → 공식 표준화의 3 단계다. 이 스위트가 그 1 단계에 해당한다.**

표준이 되는 길은 세 단계가 실증되어 있다.

1. 공개 스펙 + 참조 구현	2. 컨소시엄	3. 공식 표준화
스펙 공개 · 참조 구현으로 «전 조항 실행 가능» 증명(목표) · 채택 사례 확보 (사실상 표준)	시나리오 언어 인접 컨소시엄(ASAM 계보) 또는 독립 컨소시엄 — 경량 산출물 경로(PAS·TS·IWA)로 이행	ISO 경량 산출물로 번호 확보 후 IS 승격 또는 ISO/TC 299 신규 작업 제안 — ISO 18646 과 확장·보완 관계

산업용 3D 장면 기술 표준(OpenUSD 계열)이 같은 순서 — 공개 구현의 사실상 표준화 → 컨소시엄 결성 → 공식 표준화 추진 — 를 밟고 있다.

이 경로의 1 단계가 본 스위트다. **읽을 수 있는 규격(Part 0·1), 증거가 딸린 사례집(Part 2), 실행 가능한 참조(Part 3)**를 공개한다. 이름을 밝힌 구현 프로파일(Part 4)은 아직 공개하지 않았으며 차기에 공개한다.

참고문헌

본문이 인용한 외부 출처다. 확인 시점은 2026-08-25 1 차 출처 대조(2026-08-27 일부 재대조) 기준이며, 국내 인증 기준의 시뮬레이션 인정 조항 여부는 2026-09 조사 기준이다.

국제 규정·평가 프로토콜

- UNECE, UN Regulation No. 157 — 자동차선유지시스템(ALKS) 형식승인 규정(2021)
- UNECE WP.29, New Assessment/Test Method for Automated Driving(NATM)(2021-06 채택)
- Euro NCAP, 2026 프로토콜 변경 보도자료(2025-11-26) · 가상시험 프로토콜(2025-03 공개)
- ASAM, OpenSCENARIO · OpenDRIVE
- Regulation (EU) 2023/1230 on machinery, EUR-Lex 32023R1230(제 54 조 · 전문 55)

표준·표준 문서

- ISO 9283:1998 · ISO 10218-1/-2:2025 · ISO 3691-4:2023 · ISO 18646 시리즈 · ISO 13482 · ISO 25785-1(제정 개시 단계)

- KS B ISO 9283 · KS B ISO 10218 · KS B ISO 3691-4:2021 · KS B ISO 13482
- ISO/IEC 9646 계열 · ISO/IEC TS 42119-2:2025 · ISO/IEC 17025 · ASME V&V 40 · NASA-STD-7009
- ISO/IEC Directives Part 2 · W3C QA Framework
- IEEE 휴머노이드 표준 스터디그룹 보고서(2025-09)

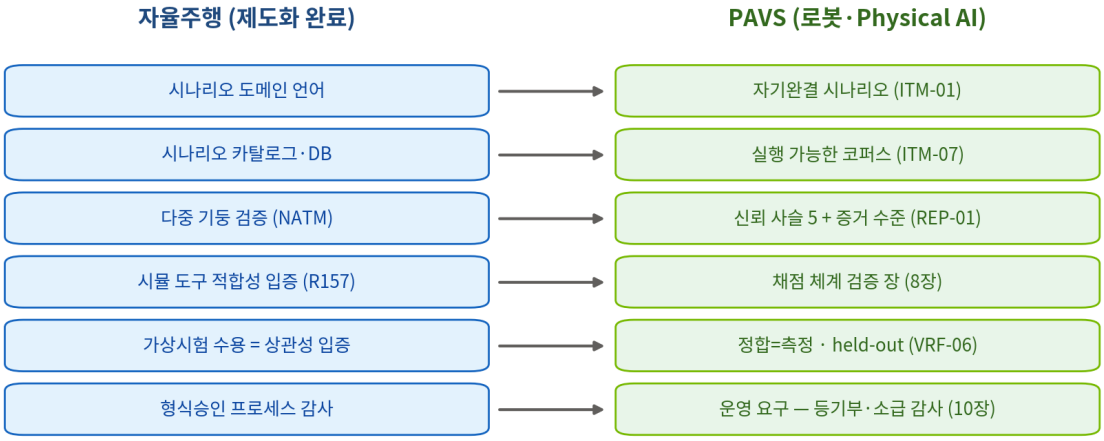
국내 법령·정책

- 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(2026-01-22 시행)
- 「지능형 로봇 개발 및 보급 촉진법」 제 40 조의 2 실외이동로봇 운행안전인증(2023-11-17 시행) · 한국로봇산업진흥원(KIRIA) 공시 심사항목
- 과학기술정보통신부, 「피지컬 AI 핵심 경쟁력 확보 전략」(2026-07-01)
- 국가로봇테스트필드, 실증지원 플랫폼 · 디지털 트윈화 프레임워크 기술개발(2024~2028)

학술 문헌

- LIBERO-PRO, arXiv 2510.03827(2025)
- LIBERO-Plus, arXiv 2510.13626(2025)
- Snyder 외, 로봇 조작 정책 비교의 시행수와 통계 검정, arXiv 2503.10966(2025)
- 교차 시뮬레이터 변형체(천) 조작 벤치마크, IEEE Robotics and Automation Letters(2024), arXiv 2310.09543
- Humanoid-Gym, arXiv 2404.05695
- Toyota Research Institute, 대규모 행동 모델(LBM) 평가 연구, arXiv 2507.05331(2025)
- SIMPLER, arXiv 2405.05941(CoRL 2024)

부록 A (informative) — 자율주행 선례와 이 규격의 대응



구조의 대응이며 제도적 지위의 동일성은 아님

자율주행 검증 생태계의 각 요소와 이 규격의 대응 관계다. 대응은 **구조의 유사성**이지 제도적 지위의 동일성이 아니다 — PAVS 는 제안 단계의 공개 규격이다.

자율주행 생태계의 요소	해결한 문제	PAVS 의 대응
시나리오 도메인 언어 (ASAM OpenSCENARIO/OpenDRIVE)	시험 시나리오를 사람이 읽고 기계가 실행하는 언어로 — 규정 문안이 실행 코드가 되는 계보	자기완결적 도메인 언어 시나리오 요구 (**PAVS-ITM-01**)
시나리오 카탈로그·DB (국가 선행연구 → 규정 편입 계보)	시험의 실체를 문서가 아니라 시나리오 집합으로	실행 가능한 코퍼스 — 시나리오+대조군+기대 판정(**PAVS-ITM-07**)

다중 기둥 검증 (UNECE NATM: 시뮬레이션·트랙·실도로·감사·운행 중 모니터링)	서로 다른 층위의 증거를 결합해 판정	신뢰 사슬 5(문항–계측–판정–증거–보고) + 증거 수준 L0~L3(**PAVS-REP-01**)
시뮬레이션 도구의 범위·유효성 입증과 툴체인 검증 (UNECE R157)	도구 신뢰 근거의 제도화	채점 체계 검증 요구(Part 1 8 장) — 음성 대조군·골든셋·모델 사양의 진위 확인·계측의 살아있음 확인
가상시험 수용 조건 = 실측 상관성 입증 (Euro NCAP 2026)	시뮬 증거의 자격을 실측 대조로	정합 주장에 측정 방법·표본·불확도 동반 · 교정에 쓰지 않은 별도 조건(보류 시험)을 통과한 값만 정식 채택(**PAVS-VRF-06**) · 실기 대조 단계 표기(**PAVS-REP-15**)
형식승인의 프로세스 감사(audit)	결과만이 아니라 절차의 증거	운영 요구(Part 1 10 장) — 실패한 방법의 기록부·소급 감사·기록 규율

부록 B (informative) — 참조 표준·KS 대응표

각 도메인 프로파일의 문항이 근거 표준으로 연결하는(anchoring) 참조 표준과 그 국내 부합화 현황이다. 확인 시점 2026-08, 공개 자료 기준. **이 규격은 표의 어느 표준에 대한 적합성 인증도 의미하지 아니한다.** 국제판과 KS 의 판본 시차가 있으므로 인용은 국제판 직접 참조 + KS 병기의 이중 표기를 원칙으로 한다.

도메인(프로파일 후보)	국제 참조 표준	KS 부합화(공개 검색 기준)	PAVS 가 다루는 층위
매니퓰레이터 성능	ISO 9283:1998	KS B ISO 9283 존재	자세 정확도·반복정밀도 등 측정 항목의 시뮬레이션 재현·채점·보고
산업용 로봇 안전	ISO 10218-1/-2:2025 (협동 요구 통합)	KS 는 2011 년판 기준 — 2025 년판 부합화 확인되지 않음(unknown)	안전 조항 문항의 근거 표준 연결 · 판본 시차의 이중 표기
AMR·무인 산업차량	ISO 3691-4:2023	KS B ISO 3691-4:2021(부합화 기준 =ISO/DIS 3691-4.2, IDT — 2023 년판 부합화 KS 미확인)	구역 속도·정지 등 안전 문항의 근거 표준 연결
서비스 로봇 성능	ISO 18646-1~4 (신규 파트 제정 진행 중)	부합화 현황 unknown	이동·내비게이션·조작 문항의 근거 표준 연결 — 확장·보완 관계
개인 지원 로봇 안전	ISO 13482	KS B ISO 13482 존재	서비스 환경 안전 문항의 근거 표준 연결
국내 제도 접점	—	지능형로봇법 실외이동로봇 운행안전인증(제 40 조의 2, 2023-11-17 시행 — 실기 기반 8 개 심사항목, KIRIA 공시 · 인증 기준에 시뮬레이션 결과 인정 조항은 확인되지 않음, 2026-09 조사)	법정 인증의 사전 검증·회귀 검증 층위(인증 참칭 없음)
산업안전 인증 제도	—	**미확인** — 산업안전보건법 안	확인 시 안전 문항의 근거 표

		전인증(고용노동부·안전보건공단) 체계에서 산업용 로봇·협동로봇의 대상 여부와 해당 항목명을 1 차 출처로 확인하지 못하였다	준 연결 대상
국가통합인증(KC)	—	**미확인** — KC 표시 대상 법령군에서 로봇 제품의 해당 범위(제품군·적용 법률)를 1 차 출처로 확인하지 못하였다	확인 시 제품 안전 문항의 근거 표준 연결 대상

> **비고 — 미확인의 표기.** 위 두 행은 조사하지 않아 비어 있는 것이 아니라 **1 차 출처 대조를 완료하지 못하여 미확인으로 표기한 것**이다. 이 규격의 계측 원칙(미측정을 측정으로 접지 않는다)을 규격 자신의 서술에도 적용한다. 확인되는 대로 개정판에 반영한다.

B.1 국내 표준 접점에 대한 입장

이 절은 국내 표준·인증 제도와의 관계에 대한 이 규격의 입장이다.

첫째, 층위가 다르므로 경쟁하지 아니한다. 국내외 표준은 **측정 대상과 합격 기준**을 정하고, 이 규격은 **그 측정 대상의 시뮬레이션 재현·측정·판정·보고 방식**을 정한다. 이 규격은 **합격 임계를 규정하지 아니하며**(1 장 범위), 임계의 출처는 성적서에 그대로 노출된다(**PAVS-REP-02**). 국내 표준이 임계를 공급하고 이 규격이 그 임계의 측정과 보고를 규율하는 분업이 성립한다.

둘째, 국내 매핑은 프로파일 층위의 작업이다. 이 규격은 도메인 프로파일 기제를 이미 규정하고 있다 — 프로파일은 문항 집합을 선별하되 요구를 완화할 수 없고(**PAVS-CNF-03**), 명칭·문항 목록·선정 근거·제외 문항을 공개 문서로 선언하여야 한다(**PAVS-CNF-05**). 국내 표준을 근거 표준으로 연결하는 문항 집합은 **규격 본문의 개정이 아니라 프로파일의 신설**로 수용된다. 규격 본문은 국가 중립을 유지한다.

셋째, 현황을 있는 그대로 기술한다. 위 표 기준으로 KS 부합화는 일부 확인·일부 미확인이고, 산업안전 인증 제도와 KC 의 적용 범위는 미확인이다. 국제판과 KS 판본 사이에 시차가 있으므로 인용 시 **국제 표준 번호와 KS 번호를 병기**한다.

넷째, 국내 매핑은 단독으로 수행할 사안이 아니다. 어느 국내 표준의 어느 항목을 문항으로 채택할지는 소관 기관·이해관계자의 합의 사항이며, 이 규격의 저자가 단독으로 정할 경우 그 선별 자체가 이해관계의 산물이 된다. 따라서 국내 프로파일의 문항 선별은 **소관 기관과의 공동 수행을 제안**하며, 이 규격은 그 공동 작업이 올라설 방법론 층위를 제공한다.

> **비고** — 이 절은 이 규격이 국내 표준에 대한 적합성·인증을 의미하지 아니한다. 표준 번호의 인용은 근거 표준으로의 연결이며, 인정(accreditation)의 어휘를 사용하지 아니한다.

부록 C (informative) — 인용 시 표현 경계

이 문서를 번역·발표·영업·디자인 목적으로 인용하는 중개자를 위한 표현 경계다. 원자료가 경계를 지켜도 인용문이 경계를 넘으면 동일한 사고가 된다. 스위트의 문안이 개정되면 이 표도 함께 개정한다.

말하려는 것	이렇게 옮긴다 (○)	이렇게는 안 된다 (×)
규격의 지위	"제안 단계의 공개 규격"	"국제 표준" · "KS/ISO 인증" (표준화

		기구 승인 함의)
"표준" 어휘	"본 규격이 정의하는 표준 문항·표준 시나리오"	"표준 인증 시나리오" (외부 기관 인증 함의)
참조 표준	"ISO ○○의 측정 항목을 근거 표준으로 연결한다(참조)"	"ISO ○○ 적합·인증을 획득했다"
참조 구현	"참조 구현의 목표는 전 조항 실행 가능성의 증명이며, 현재 충족 현황(72 건 중 충족 7 건)은 적합성 진술로 전건 공개한다"	"참조 구현이 전 단언을 실행한다" (현재형 단정) · 완성 제품·판매 중 서비스로의 단정
재현 패키지	"시연 A 의 재현 패키지는 요청 시 제공한다"	"누구나 내려받아 재현할 수 있다" (미게시 상태의 공개 단정)
성적 보고	"성적 보고 양식과 무결성 장치를 규정한다"	"성적서를 발급해 드립니다" (현재형 서비스 단정)
중립성	"이해상충은 성적서에 필수 선언된다" · "발행 주체는 인증·인정기관이 아니다"	"중립 제 3 자 기관" (현재형 지위 주장)
사례 수치	사건의 물리량 그대로 + 측정 시점 병기	반올림 격상 · 가능형의 단정 전환("산출될 수 있었다"→"판정을 받았다")
실패 사례	"체계가 자기 결함을 잡아 고쳤다"(발견→교정→재검증 3 단)	결함만 떼어 극화 · 교정 사실 생략
외부 연구	"저자들이 보고했다/주장했다"	외부 주장의 사실 단정
적합성	"핵심 단언 충족을 전제로 한 부분 적합 — 미충족 목록 공개"	"PAVS 통과·인증" (미충족 비공개 함의)
발행 기관의 현황	"자기 시험 결과 «PAVS 지향» — 미충족 단언 전건 공개"	"PAVS 적합 구현" · 제 3 자 심사 결과로의 표현

부록 D (informative) — 용어 해설 (평이 설명)

규범 정의는 Part 1 3 장이 정본이며, 이 해설은 이해를 돕기 위한 평이한 설명이다.

- **피시험체(SUT)** — 시험 대상이 되는 로봇·AI.
- **문항 / 시험 단언** — 채점의 개별 항목 / 그 항목을 구성하는, 참·거짓을 판정할 수 있는 한 문장. 이 규격의 조항 하나하나가 단언 번호(PAVS-XXX-NN)를 갖고, 72 건 전부에 시험 방법(주입·심사·시연)이 규정되어 있다.
- **계측 출처(provenance)** — 해당 수치가 실측값·명령값·기본값 중 무엇인지를 표시하는 항목. 출처가 선언되지 않은 값은 채점에 사용할 수 없다.
- **결측** — 측정이 수행되지 않은 상태. 0 이 아니라 값의 부재를 의미한다.
- **계측 부재** — 필수 문항(그 계측 없이는 합격이 성립하지 않는 문항)에서 계측이 이루어지지 않은 상태. 이 사유로 불합격을 낼 수 있으며, 그 경우 불합격 사유가 기록에서 식별되어야 한다.
- **3 상 판정** — 합격/불합격/미측정. 재지 않은 문항을 합격으로 처리하지 않는다. 필수 문항의 계측 부재는 사유를 밝힌 불합격으로 처리할 수 있으며, 이때에도 미측정 건수는 따로 보존·표기한다. ISO/IEC 9646 의 pass/fail/inconclusive 를 계승한 개념이다.
- **fail-loud** — 측정 실패 시 대체값으로 보정하지 않고 실패로 명시 보고하는 원칙.
- **음성 대조군** — 의도적으로 위반하도록 구성한 시험 대상. 채점기가 이를 통과시키면 채점기가 불합격이다.

- **널대조군** — 주장된 개선 효과만 뺀 채 같은 처리 경로를 그대로 거친 비교 조건. 처리 경로 자체가 만드는 차이와 실제 개선 효과를 구별한다.
- **위약 대조** — 울리면 안 되는 조건에서 경보가 울리지 않는지 확인하는 시험.
- **골든셋** — 판정이 확정된 기준 시험 집합. 채점기 변경 후 재채점하여 판정의 불변 여부를 확인한다.
- **실효값** — 시스템이 실제로 적용한 값. 설정했다고 적용된 것이 아니다.
- **경계 설명** — 해당 합격이 보장하는 범위를 문항 단위로 기술한 주석.
- **증거 수준(L0~L3)** — 코드가 있다(L0) → 단위 시험(L1) → 실운용 경로 종단 실행(L2) → 산출물 직접 확인(L3). 합격 발행은 L2 이상에서만.
- **모델 진본성(모델 사양의 진위 확인)** — 시뮬레이션 모델의 제원·기하·물성이 제조자 확정 사양과 맞는지, 다르다면 그 차이에 기록된 근거가 있는지 확인하는 것.
- **보류 시험(held-out)** — 교정에 쓰지 않고 따로 남겨 둔 시험 조건. 교정값은 이 조건에서도 맞을 때에만 정식으로 채택된다.
- **소급 감사** — 판정 기준의 결함을 고친 뒤, 결함이 있던 기간의 기존 판정 전부를 새 기준으로 다시 판정해 바뀐 판정의 방향과 규모를 기록하는 절차.
- **실패 접근 등기부** — 이미 실패한 접근 방법의 공식 기록부. 동일한 실패의 반복을 방지하는 장치이며, 오진의 정정도 원 기록을 삭제하지 않고 연결하여 보존한다.
- **매니페스트** — 피시험체의 식별 정보와 사양(관절·구동 한계)을 담은 선언 파일. 채점기는 이 범위를 벗어난 지령을 근사 처리하지 않고 거부한다.
- **프로파일** — 도메인별 문항 선택집(매니퓰레이션·이동·안전 등). 문항을 고를 수는 있어도 요구를 완화할 수는 없다.
- **성적서** — 검증장 운영자가 시험 결과를 담아 발행하는 문서. 법정 시험성적서를 뜻하지 않는다.
- **적합성 진술** — 구현이 규격의 각 단언을 충족하는지를 정해진 양식으로 밝힌 문서. 부속서 A 전 단언의 체크리스트(ICS)를 첨부하고 미충족 단언을 빠짐없이 공개해야 유효하다(Part 1 11.2·11.3).
- **PAVS 지향** — ★핵심 단언 중 하나라도 미충족인 구현이 «PAVS 적합» 대신 쓰는 구별 표현(Part 1 11.2).
- **★핵심 단언** — «PAVS 적합» 주장의 전제가 되는 최소 집합(**8 건** — 2026-09-01 REP-15 승격으로 8 건). 목록 정본은 Part 1 11.2.1 이다. 하나라도 미충족이면 「PAVS 지향」만 쓸 수 있다.

★핵심 단언	뜻
PAVS-MEA-01	채점에 쓰인 값마다 출처(실측·명령값·기본값·추정)를 밝히고, 실측이 필요한 문항은 실측값으로만 채점
PAVS-MEA-02	측정이 없을 때 0 이나 기본값으로 채우지 않음
PAVS-VDT-01	합격·불합격·미측정의 3 상 판정 · 미측정을 합격으로 접지 않음
PAVS-VDT-02	합격은 필수 조건 전부가 실측으로 충족되었다는 증거로만 산출
PAVS-VRF-01	일부러 위반하는 시나리오(음성 대조군)를 상설 유지하고 기대 판정과 자동 대조
PAVS-REP-15	시뮬레이션 모델이 실기와 대조된 단계와 상관 증거 상태를 표기 · 증거가 없으면 «없음»
PAVS-CNF-01	정해진 양식으로 적합성을 주장하고 미충족 단언을 빠짐없이 공개
PAVS-CNF-02	부속서 A 전 단언의 충족 여부 체크리스트(ICS) 첨부 · 해당 없음의 자격 제한

개정 고지 — v1.0 → v1.1 (2026-10-02)

설명 문서의 정정·보강이며, 규격의 요구 내용은 이 문서로 바뀌지 않는다(요구의 정본은 Part 1).

변경	내용	이유
규격 개정분 반영	대응 규격을 Part 1 v1.3 으로 갱신 · ITM-10(정격 능력의 지속 가능성) · VRF-05 강화(차이의 근거 기록, 불일치 모델 결과의 사용 금지) · VDT-01(계측 부재 불합격, 미측정 건수 보존) · REP-04(자기선언의 한계) · REP-15(★핵심, 실기 대조 단계 표기) · VRF-16 서술	최초 공개본이 이전 판의 문면을 설명하고 있었음
과장 정정	「문서 심사가 아닌 검사 실행」 → 시험 방법 유형(주입 36 · 심사 33 · 시연 3)과 시험 기록 기준 · 「조항–사건 대응 공백 0」 → 유래 근거 3 종 · 참조 구현 「전 단언 실행 가능」 → 목표와 현황의 구분	실제 상태보다 강한 서술
제도 인용 정정	보도 기반 국내 제도 인용 1 건 삭제 · EU 기계류 규정의 대상 정밀화 · 실외이동로봇 운행안전인증의 시뮬레이션 인정 조항 미확인 병기	1 차 출처로 확인된 항목만 유지
상태 정정	사례 수(사례집 v1.1 기준) · Part 3 v0.4 함께 공개 · Part 4 미공개 · 시연 A 재현 패키지 요청 시 제공 · 개발 장비 실행 시간 삭제 · 적합성 클래스의 장 범위	공개 문서 사이의 상충 해소
측정 시점 병기	최고속도 사례(2026-08) · 소급 감사 89%(2026-05) · 모델 사양 대조 실측(2026-08)	시점 없는 수치의 오독 방지
보강	이해관계 고지 · 인접 규범 계보 · VRF-16 유래 · 모델 사양 결함 보기 · 9.4 자기 시험 · 참고문헌 · 문의 · 용어 7 건	독자가 확인하여야 할 정보의 누락
가독성	은어·비유 정리(「계류」 → 「근거 표준으로 연결」 등) · 내부 절차 문구 삭제	일반 독자 대상 설명 문서

문의

이 규격과 백서에 대한 문의·이의·정오 요청은 contact@immersivecast.com 으로 받는다. 이의 처리 절차의 세부(접수 기록, 처리 결과와 사유의 공개 방식)는 차기 개정 고지에서 밝힌다.

PAVS Part 0 백서 · 부록 A~D 포함 · 외부 인용은 2026-08-25 1 차 출처 대조(08-27 일부 재대조) 기준 · 발행본 v1.1 · 2026-10-02