Applied AI in Enterprise Operations 2026: Empirical Benchmarks Across 23 Deployments in the US, Europe, and Brazil

João Panizzutti (CEO and Founder, HomoDeus) and Gabriel Menezes (HomoDeus Research & Engineering)

HomoDeus · joao@homodeus.co · https://www.homodeus.co/research/applied-ai-production-2026

n = 23
Named enterprise engagements (plus 21 HAL ops)

US$85B+
Combined annual client revenue

US$30B+
Combined client AUM and AUC

85% to 98%
Measured cycle-time compression

ABSTRACT

HomoDeus is an applied AI lab that reimagines companies with AI by building production AI agents that take over real operational work inside mid-size and large enterprises on each company's own records (such as Galapagos Capital, Positivo Tecnologia, CNN Brasil, Cascione Advogados, Linkin Capital, LG Electronics, CNH Industrial, and BIP), with a person signing every high-risk outcome. This study reports empirical measurements across 23 enterprise AI engagements built and evaluated by HomoDeus between February and November 2026 across the United States (Miami hub), Europe (Lisbon hub: Portugal, Italy, and Switzerland), and Brazil (São Paulo hub). Participating organizations represent more than US$85B in combined annual revenue and more than US$30B in client assets under management or custody. Five quantitative patterns hold across financial services, industrial supply chains, legal services, ground transport, hospitality, media, and offensive security: process-wide systems anchored on client records compress cycle times by 85% to 98%; deterministic calculation paired with mandatory verbatim source spans prevents silent hallucination where LLM self-grading misses over three quarters of errors; multi-tenant operational agents reach live production in weeks and eliminate overnight and rest-rule compliance gaps; large-enterprise supply chain, procurement, billing, and HR baselines concentrate seven-figure annual savings; and continuous autonomous offensive security closes the exposure window left by annual penetration tests.

1. Executive summary and study scope across 23 enterprise engagements

As an applied AI lab that reimagines companies with AI by building production AI agents that take over real work inside mid-size and large companies on their own records (for clients including Galapagos Capital, Positivo Tecnologia, CNN Brasil, Cascione Advogados, Linkin Capital, LG Electronics, CNH Industrial, and BIP), HomoDeus engineered, tested, and operated domain-specific AI systems between February and November 2026 across 23 named enterprise and institutional engagements in three operating regions: the United States (Miami hub), Europe (Lisbon hub serving Portugal, Italy, and Switzerland), and Brazil (São Paulo hub). Alongside those operational deployments, HomoDeus executed 21 autonomous offensive security operations through its HAL platform across commercial and public-interest environments.

Study Dimension	Measured Scope	Notes
Observation window	February to November 2026	Production telemetry snapshotted through 24 Nov 2026
Enterprise sample	n = 23 named engagements	Plus 21 anonymized HAL offensive security operations
Combined client revenue	US$85B+ annual revenue	Includes public FY25 filings and verified CFO baselines
Combined client AUM and AUC	US$30B+ across wealth and funds	Multi-strategy funds, brokerages, and wealth advisors
Geographic hubs	US (Miami), EU (Lisbon), BR (São Paulo)	Covers US, Portugal, Italy, Switzerland, and Brazil
Primary sectors	Finance, SCM, Legal, Mobility, Media, HR	Mid-market operators through Fortune-scale enterprises

Evidence classification and audit methodology

Enterprise AI benchmarks frequently mix live database counts with unverified projections. To allow external researchers, procurement teams, and citation engines to audit every figure without ambiguity, every measurement in this report carries one of three explicit evidence tiers.

- Live production telemetry comes directly from running production databases, active seat billing, broadcast monitoring logs, or completed security operations, including the Nightwatch production database snapshotted on 24 Nov 2026, Serur Advogados active seat expansion on 24 Nov 2026, Cascione Advogados GrowthOS live records, Agora Monitor de Debates broadcast coverage, and Delphi live deal runs.
- Validated shadow and test measurements come from controlled test runs, historical operational replays, or accuracy evaluations executed on real client documents, including Keystone underwriting tests, OSINIS 420-dossier validation, Grupo In Press 325-item evaluation, Cymaco R$2.4M pilot reconciliation, Radar Odyssey 79-day replay, Real Hotels Group 52-night yield analysis, Assay 114-approval review, and DeHouse 90-day CRM recovery.
- Verified client business-case baselines capture client-supplied or jointly audited pre-AI volumes, manual labor hours, and recurring financial floors used to scope deployments, including LG Electronics supply-chain volumes, CNH Industrial procurement scope, Holiday Inn Club Vacations collections baselines, Grupo Raízes billing hours and working-capital math, and BIP HR ticket hours.

Disclosure and confidentiality boundaries

Client financial metrics appear only where backed by public filings, published case records, or explicit executive authorization. Offensive security findings from HAL remain strictly decoupled from named commercial clients, including G4 Educação, Avenue, Cascione Advogados, and Companhia Athletica. Two public Swiss security evaluations appear without unpublished finding counts. Fundação Butantan is reported as a named institutional work engagement.

2. Process-wide AI systems compress cycle times by 85% to 98% when anchored on client records

Research by the RAND Corporation (RR-A2680-1, 2024, 65 interviews) reports that more than 80% of corporate AI projects fail to reach production. Across HomoDeus engagements, standalone chat interfaces produced negligible throughput gains because analysts still copied data between spreadsheets, PDFs, email threads, and compliance registries. Systems wired directly into the client's underlying records compressed end-to-end process cycle times by 85% to 98% while reducing unit data costs.

At Galapagos Capital (R$20B to R$32B AUM, about US$9B global multi-strategy across the US, Brazil, and Europe), the Keystone underwriting system on the Miami US real estate desk compressed sponsor model underwriting from 1 week or more down to 20 to 40 minutes. Weekly underwriting load dropped from 100 hours per week to 2 to 4 hours per week (-96% to -98%), saving about US$500,000 per year. On a live sponsor file, Keystone flagged an insurance expense line understated by 2.53x, moving projected deal return from 12.17% down to 5.48%.

At Positivo Tecnologia (R$3.35B FY25 net revenue, B3: POSI3, Brazil), OSINIS automated third-party compliance due diligence from about 1 hour per target down to about 4 minutes in testing. Data cost per diligence fell from R$40 to R$4.95 (-87.6%). The team validated 420 dossiers against a full annual scope of 15,000 third parties per year (an R$849,300 per year manual full-coverage baseline), delivering R$1.6M to R$2.5M per year in recurring savings.

In named institutional work with Fundação Butantan in Brazil, Assay compressed supplier payment review from 3 to 4 days per month to minutes. A batch of 114 weekend approvals was reviewed in 5 minutes, and the system flagged 4 contract-violating invoices before payment release.

At Linkin Capital (US and Brazil venture capital), Delphi mined 112,633 historical email conversations and 3,185 deals directly inside Slack. In a single 24-hour run, Delphi analyzed 160 deals, equal to 5.3 months of typical 30-deal per month flow, saving about 55 partner hours per month (about US$49,500 per year). At Grupo In Press (Brazil, 200 to 250 corporate clients), an AI press-clipping evaluation targeting more than 800 hours per month of manual tabulation achieved 325/325 exact metadata matches, 114/114 exact BRL media value matches, and 0.991 outlet accuracy. The pipeline classified 114 articles in 123 seconds for US$0.30 (US$0.0027 per article, or about US$1.33 for a 500-article peak day).

Organization	Region	System	Pre-AI Cycle or Cost	Measured AI Performance	Annual or Unit Impact	Evidence Tier
Galapagos Capital	US (Miami)	Keystone	>=1 week/deal; 100 hrs/wk	20-40 min/deal; 2-4 hrs/wk (-96% to -98%)	~US$500,000/yr saved; caught 2.53x insurance gap (12.17% to 5.48%)	Validated testing and client baseline
Positivo Tecnologia	Brazil	OSINIS	~1 hr/target; R$40 data cost	~4 min/target; R$4.95 data cost (-87.6%)	R$1.6M to R$2.5M/yr saved; 420 dossiers validated across 15,000/yr scope	Validated testing and cleared savings
Fundação Butantan	Brazil	Assay	3-4 days/month review	114 weekend approvals in 5 minutes	Flagged 4 contract-violating invoices before payout	Validated shadow measurement
Linkin Capital	US / Brazil	Delphi	30 deals/month manual pace	160 deals in 24 hrs (5.3 months flow)	~55 partner hrs/month (~US$49,500/yr); 112,633 emails and 3,185 deals mined	Live production telemetry
Grupo In Press	Brazil	Clipping Eval	800+ hrs/month manual	114 articles in 123s; 325/325 metadata; 0.991 outlet	US$0.30/114 articles (US$0.0027/article; ~US$1.33/500-article peak day)	Validated shadow measurement

3. Deterministic calculation and mandatory source citation eliminate silent hallucination

Testing across financial, legal, and media workloads exposed a consistent failure mode in unconstrained language models. During Keystone underwriting tests, self-grading by an LLM caught fewer than 1 in 4 errors, and fluent narrative prose routinely inflated evaluator scores even when underlying spreadsheet arithmetic was wrong. Relying on model confidence scores or secondary LLM graders left silent numerical errors undetected.

Every production system in the 23-engagement cohort enforces three structural guardrails. First, every extracted fact requires a verbatim source span from a client document; any item lacking a direct quote escalates automatically to a human reviewer. Second, Python and deterministic code execute all financial arithmetic, tax matching, rate multipliers, and totals, keeping language models strictly on document parsing and citation assembly. Third, high-consequence declines and live broadcast summaries pass through dual-AI consensus gates before reaching operators.

At Cymaco Engenharia (40+ year Brazilian engineering and retrofit firm), the Tirion financial cockpit paired document extraction with deterministic reconciliation across R$2.4M in pilot contracts, matching purchases, payments, and tax documents to the cent. The reconciliation surfaced R$334,399 in receipts lacking NF-e tax invoices (+R$13,646) and R$270,000 in financial anomalies against a R$200,000 to R$600,000 per year target margin leakage baseline (R$20,000 to R$50,000 per project).

In live broadcast journalism, Agora Monitor de Debates for CNN Brasil and Rádio Itatiaia processed 329 political debates and interviews across 79 channels, 35 outlets, and 14 Brazilian states, covering 233 hours of video (29 workdays). The pipeline delivered roughly 3-minute block summaries at a 91.6% speaker-turn match rate anchored directly to timestamped transcripts. At Linkin Capital, Delphi runs a dual-AI decline gate so one model cannot silently reject a qualified investment opportunity.

Architecture Rule	Failure Mode Prevented	Empirical Verification Result	Evidence Tier
Deterministic math boundary	LLM arithmetic drift and prose-inflated self-grading (<1 in 4 caught)	Tirion reconciled R$2.4M pilot contracts to the cent; surfaced R$334,399 without NF-e (+R$13,646) and R$270,000 anomalies	Validated shadow measurement
Mandatory verbatim source span	Unverifiable claims in financial, HR, and compliance outputs	In Press 325/325 exact metadata and 114/114 exact BRL media value match; BIP read-only escalation on ungrounded queries	Validated shadow measurement
Dual-AI consensus gate	Silent false rejections and misattributed broadcast quotes	Agora 91.6% speaker-turn match across 329 debates, 79 channels, 35 outlets, 14 states, 233 video hrs; Delphi dual-AI decline gate	Live production telemetry

4. Multi-tenant operational agents reach production in weeks and close nocturnal and compliance gaps

Operational businesses in mobility, hospitality, commercial real estate, and legal services lose revenue and accumulate regulatory risk outside office hours and across fragmented communication channels. Multi-tenant agent architectures reuse core ingestion,

routing, and human-approval engines across operators, compressing deployment timelines to weeks while adapting to local dispatch rules.

Across European ground-transfer operators (BestoursPT, Genuine Tours, Radar Odyssey, Transfergest, and CAR2U in Portugal, plus Autoservizi Sarno in Italy), the Nightwatch production database as of 24 Nov 2026 records 15,679 ride offers decided at a 15.6-second median decision time (BestoursPT 15,061; Radar Odyssey 618), 3,926 bookings entered (BestoursPT 2,214; Genuine Tours 1,506; Radar Odyssey 206), and 234 bookings edited. Between 15 Feb and 24 Nov 2026, BestoursPT accepted EUR 69,198.50 in gross booking value through Nightwatch, including EUR 7,487.32 captured during overnight hours from 00:00 to 08:00. Genuine Tours reached live production in 13 days from an archive of 11,132 historical emails. On Radar Odyssey's 79-day replay covering 1,886 services, automated scheduling cut statutory driver rest-rule violations from 294 to 0 and reduced outsourced jobs by 34% (178 down to 118), saving about 252 dispatch desk hours per month (about EUR 26,800 per year).

In Lisbon commercial real estate, DeHouse runs Chronicle as its sales operating layer. Over 90 days, Chronicle recovered 60 unlogged client meetings into the CRM, tracked 76 active proposals, and parsed 40 of 55 image-only vendor price sheets. In Portuguese hospitality, Real Hotels Group (15 to 18 four- and five-star hotels in Portugal) ran Yield dynamic pricing analysis across 52 nights and 899 rooms, identifying EUR 28,500 in revenue at stake alongside a broader Hotel OS operational scope.

In Brazilian corporate law, adoption expands from initial operational hubs into multi-system firm deployments. At Serur Advogados (30+ year Legal 500 ranked firm across São Paulo, Recife, and Brasília), Tribe workflow-mapping AI expanded from 100 initial seats to 114 active billed seats on 24 Nov 2026. At Cascione Advogados (São Paulo and Rio de Janeiro), an expanding company-wide AI program spans more than 5 systems, with GrowthOS managing 5,344 leads, 3,004 tasks, 689 proposals, and 14,510 screened news items.

Operator or Firm	Region	System	Measured Production or Replay Result	Evidence Tier
BestoursPT, Genuine Tours, Radar, Transfergest, CAR2U, Autoservizi Sarno	Portugal and Italy	Nightwatch	15,679 ride offers (15.6s median); 3,926 bookings entered; 234 edited; EUR 69,198.50 BestoursPT gross value (EUR 7,487.32 overnight 00:00-08:00); Genuine live in 13 days from 11,132 emails	Live production telemetry
Radar Odyssey	Portugal	Nightwatch Replay	79-day / 1,886-service replay: rest violations 294 to 0; outsourced jobs down 34% (178 to 118); ~252 desk hrs/mo saved (~EUR 26,800/yr)	Validated shadow measurement
DeHouse	Portugal (Lisbon)	Chronicle	60 unlogged client meetings recovered in 90 days; 76 proposals tracked; 40/55 image-only price sheets parsed	Validated shadow measurement
Real Hotels Group	Portugal (15-18 hotels)	Yield / Hotel OS	52 nights and 899 rooms analyzed; EUR 28,500 revenue at stake identified	Validated shadow measurement
Serur Advogados	Brazil (Legal 500)	Tribe	Expanded from 100 initial seats to 114 active seats on 24 Nov 2026	Live production telemetry
Cascione Advogados	Brazil	GrowthOS (5+ systems)	5,344 leads, 3,004 tasks, 689 proposals, and 14,510 news items screened	Live production telemetry

5. Enterprise supply chain, procurement, billing, and HR baselines concentrate seven-figure annual savings

Large-enterprise back offices hold their highest-value bottlenecks in high-volume document verification, contract-specific billing, supplier expediting, collections, and internal employee inquiries. Unlike consumer chat adoption, automating these queues yields direct labor-hour recovery and measurable working-capital gains. Every figure in this section is tagged explicitly as a verified client business-case baseline or public institutional scale.

At LG Electronics (US$62.7B FY25 revenue), HomoDeus works with LGESP SCM Logistics Innovation across 3 supply-chain AI systems covering proof-of-delivery verification, delivery tracking, and Order Desk automation. The operational scope spans more than 20,000 PODs per month, 400 to 500 deliveries per day, and 7,000 Order Desk hours per year, establishing an R$1.5M+ per year recurring savings floor.

At CNH Industrial (US$18.1B FY25 consolidated revenue), procurement and supplier expediting covers 140,000 purchase orders and 300 suppliers under an R$700,000 to R$2.0M per year business case (R$1.5M per year base case). At Holiday Inn Club

Vacations (US$500M to US$1.6B estimated annual revenue across US and European operations), the collections operations baseline stands at US$7M per year across 136 staff and 170,000 members with about 10% annual delinquency.

At Grupo Raízes (190 restaurants, R$410M to R$500M annual revenue scale), Ledger computes monthly billing across 190 custom contracts against a 4,680 hours per year manual billing baseline. Accelerating invoice issuance frees R$157,000 per year in working capital per day faster billing at a 14% CDI rate (up to R$575,000 on a 3-day acceleration). At BIP (Milan headquarters, EUR 315M to EUR 350M+ annual revenue), Canon maps 1,415 HR inquiry hours per month, of which 962 to 1,146 hours per month (68% to 81%) are addressable under a read-only, cited audit architecture.

Across regulated financial services and scaled education platforms, HomoDeus engagements include Bolton Global Capital (about US$19B AUM and 130 to 150 financial advisors), Avenue (Itaú-controlled US brokerage), OpenCo (R$5B+ in credit disbursed in Brazil), and G4 Educação (about R$505M annual revenue).

Organization	Scale / Revenue / AUM	Region	Operational Baseline	Annual Savings or Addressable Value	Evidence Tier
LG Electronics	US$62.7B FY25 revenue	Brazil / Global	3 SCM systems; 20,000+ PODs/mo; 400-500 deliveries/day; 7,000 Order Desk hrs/yr	R$1.5M+/yr recurring savings floor	Verified client baseline
CNH Industrial	US$18.1B FY25 revenue	Brazil / Global	Procurement across 140,000 POs and 300 suppliers	R$700,000 to R$2.0M/yr business case (base R$1.5M/yr)	Verified client baseline
Holiday Inn Club Vacations	US$500M-US$1.6B est.	US / Europe	Collections across 136 staff, 170,000 members (~10% delinquency)	US$7M/yr operating baseline	Verified client baseline
Grupo Raízes	190 restaurants; R$410M-R$500M	Brazil	Ledger billing across 190 custom contracts; 4,680 hrs/yr manual baseline	R$157,000/yr working capital per day faster at 14% CDI (up to R$575,000 on 3 days)	Verified client baseline
BIP	EUR 315M-EUR 350M+	Italy HQ / Brazil	Canon HR answer hub; 1,415 HR hrs/mo mapped	962-1,146 hrs/mo addressable (68% to 81%)	Verified client baseline
Bolton Global Capital, Avenue, OpenCo, G4 Educação	~US$19B AUM (Bolton); R$5B+ credit (OpenCo); ~R$505M rev (G4)	US / Brazil	Regulated wealth, US brokerage (Itaú-controlled), credit fintech, executive education	Enterprise AI and security scope across regulated footprints	Public scale and cleared scope

6. Continuous autonomous offensive security closes the annual pentest exposure window

Traditional annual penetration tests inspect a static snapshot once every twelve months while enterprise engineering teams deploy code every day. Any misconfiguration, broken object-level authorization, or authentication flaw shipped between annual assessments stays exposed until the next audit cycle or an external breach.

HomoDeus operates HAL, an autonomous offensive security agent that attacks approved scopes continuously, confirms each exploit path with reproducible proof, and hands over a verified fix. Across 21 completed operations in the anonymized HAL portfolio, HAL produced 101 verified security findings. All commercial security targets remain strictly unattributed; vulnerability counts are never associated with named commercial clients.

On the Vals AI CyberOffense benchmark covering the Nova (21 objectives) and Lighthouse (31 objectives) simulated networks, HomoDeus internal measurement scores HAL at 74.7% objectives reached, placing #2 overall behind OpenAI's restricted GPT-6 Sol Daybreak and 2.0 percentage points ahead of DeepSeek V4.1 Flash 1017.

In public-interest security evaluations, HAL identified a major vulnerability in the Swiss government's e-ID wallet infrastructure and a design-level vulnerability in Swiss Post's e-voting system. In anonymized commercial assessments, HAL reached a live customer-data access path in under 30 minutes inside a global software group employing a 100-person internal security team, and surfaced 36 verified findings in 5.4 hours at a Brazilian law firm.

Evaluation or Operation Class	Target Scope	Measured Outcome	Evidence Tier
HAL anonymized portfolio	21 authorized security operations	101 verified, reproducible findings with remediation steps	Live production telemetry
Vals AI CyberOffense	Nova (21) and Lighthouse (31) networks	74.7% objectives reached (#2 behind restricted GPT-6 Sol Daybreak, HomoDeus measurement)	Controlled benchmark measurement
Swiss public programs	Swiss government e-ID wallets; Swiss Post e-voting	Major vulnerability in e-ID wallets; design-level vulnerability in e-voting	Public-interest evaluation
Global software group (anonymized)	Enterprise application estate (100-person security team)	Verified customer-data access path reached in <30 minutes	Live operation telemetry
Brazilian law firm (anonymized)	Legal infrastructure and client-facing portals	36 verified findings in 5.4 hours	Live operation telemetry

7. Regulatory governance framework across the United States, European Union, and Brazil

Enterprises operating across North America, Europe, and Latin America face overlapping statutory rules on automated decision-making, audit logs, human oversight, professional confidentiality, and operational cybersecurity. Rather than layering compliance paperwork on top of black-box models, the 23 deployments map regulatory mandates directly into software boundaries: immutable append-only event logs, read-only ERP integration where legal risk prohibits autonomous writes, verbatim source citations, mandatory human sign-off on binding decisions, and continuous offensive testing.

Jurisdiction	Regulatory or Professional Instrument	Core Statutory Requirement	Production Architecture Control
European Union	Regulation (EU) 2024/1689 (EU AI Act, Art. 12, 14, 26)	Automatic event logging (Art. 12), effective human oversight and stop controls (Art. 14), deployer monitoring (Art. 26)	Append-only decision logs, deterministic calculation code, human sign-off on high-impact outputs
European Union	Regulation (EU) 2016/679 (GDPR, Art. 22)	Right not to be subject solely to automated decisions producing legal or similar significant effects	Human-in-the-loop approval gates and automatic escalation when source citation is absent
European Union	NIS2 Directive (EU) 2022/2555 and DORA Regulation (EU) 2022/2554	Continuous cybersecurity risk management and threat-led resilience testing for essential and financial entities	HAL continuous offensive security with reproducible proofs inside approved scopes
United States	NIST AI RMF 1.0 (NIST AI 100-1)	Govern, map, measure, and manage AI accuracy, reliability, and operational risk continuously	Held-out evaluation suites, verbatim citation checks, and dual-AI consensus gates
United States	ABA Formal Opinion 512 (July 2024)	Lawyer competence, client confidentiality, and independent verification of generative AI work product	Firm-isolated records, cited-only legal synthesis, partner review before filing
Brazil	LGPD (Lei nº 13.709/2018, Art. 20)	Data subject right to request review of decisions taken solely through automated processing of personal data	Read-only HR and payroll architecture (as at BIP) with mandatory human escalation on uncertainty
Brazil	OAB Recomendação 001/2024 (Nov 2024)	Confidentiality, factual verification, and attorney accountability when using generative AI	Private firm boundaries, source-linked citations, and partner sign-off across legal deployments

8. References and primary sources

External benchmark figures, statutory texts, and professional ethics opinions cited in this report link directly to their official issuing bodies below.

1. RAND Corporation, RR-A2680-1 (2024): The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed https://www.rand.org/pubs/research_reports/RRA2680-1.html
2. Regulation (EU) 2024/1689 (EU AI Act), Official Journal of the European Union, EUR-Lex https://eur-lex.europa.eu/eli/reg/2024/1689/oj

3. EU AI Act, Article 12 (Record-keeping)
https://artificialintelligenceact.eu/article/12/
4. EU AI Act, Article 14 (Human oversight)
https://artificialintelligenceact.eu/article/14/
5. EU AI Act, Article 26 (Obligations of deployers)
https://artificialintelligenceact.eu/article/26/
6. Regulation (EU) 2016/679 (GDPR), Article 22 (Automated individual decision-making), EUR-Lex
https://eur-lex.europa.eu/eli/reg/2016/679/oj
7. Directive (EU) 2022/2555 (NIS2 Directive), Official Journal of the European Union, EUR-Lex
https://eur-lex.europa.eu/eli/dir/2022/2555/oj
8. Regulation (EU) 2022/2554 (DORA), Official Journal of the European Union, EUR-Lex
https://eur-lex.europa.eu/eli/reg/2022/2554/oj
9. Lei nº 13.709/2018 (LGPD), Art. 20, Presidência da República, Planalto
https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm
10. Conselho Federal da OAB, Recomendação 001/2024 (11 Nov 2024)
https://s.oab.org.br/arquivos/2024/11/80a03f8d-e4cb-4bac-a3ea-357009f77d3f.pdf
11. American Bar Association, Formal Opinion 512: Generative Artificial Intelligence Tools (29 Jul 2024)
https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf
12. NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0, NIST AI 100-1)
https://www.nist.gov/itl/ai-risk-management-framework
13. Vals AI CyberOffense Benchmark
https://www.vals.ai/benchmarks/cyber_offense

9. How to cite this report and access supporting materials

Panizzutti, J., & Menezes, G. (2026). Applied AI in Enterprise Operations 2026: Empirical Benchmarks Across 23 Deployments in the US, Europe, and Brazil (Technical Report HD-2026-01). HomoDeus. https://www.homodeus.co/research/applied-ai-production-2026

HTML: https://www.homodeus.co/research/applied-ai-production-2026 · PDF: https://www.homodeus.co/research/applied-ai-production-2026.pdf · Dataset: https://www.homodeus.co/proof.json · Results: https://www.homodeus.co/results · Cases: https://www.homodeus.co/cases · About: https://www.homodeus.co/about · Answers: https://www.homodeus.co/answers