

Building Digital Workers as Managed Assets

From operating problem to role, record, evidence, and lifecycle

Fabio A. Thiers, MD, PhD

ConverGen AI, New York, NY, USA

ABSTRACT

A Digital Worker should be built and managed as an operational asset: a bounded role with a named owner and supervisor, current organizational context, explicit identity and permissions, local validation evidence, defined authority, and a reversible lifecycle. This article presents a repeatable build sequence for healthcare administration and adjacent clinical operations, centered on the Digital Worker Record as the persistent management object. The sequence begins with an operating problem, translates it into a role, validates performance in shadow mode, moves through supervised production, and adjusts authority as evidence accumulates. The purpose is not to slow deployment, but to make Digital Labor repeatable, observable, supervisable, and governable across an organizational portfolio.

Keywords: healthcare AI; Digital Workers; Digital Labor; Digital Operations; managed assets; local validation; shadow mode; lifecycle management; evidence-gated authority

Key points

- **Digital Workers should be built as operational assets, not deployed as disconnected AI features.** The unit of management is the bounded role operating under organizational authority.
- **The Digital Worker Record is the persistent management object.** It connects role scope, context, access, supervision, validation, performance, incidents, authority, and lifecycle in one maintained record.
- **Authority should move with local evidence.** Shadow validation, supervised production, monitoring, revalidation, redesign, and retirement are normal lifecycle stages rather than governance afterthoughts.

Start with the operating problem, not the model

Digital Worker creation should begin with work the organization can describe and measure. A backlog item such as “use AI for prior authorization” is too broad. A more applicable operational problem is something like: “authorization packets for a defined service line are delayed because payer requirements vary, missing evidence is discovered late, and staff spends substantial time reconciling policy with chart content”.

This framing matters because healthcare AI translation often fails when technical development is separated from workflow, implementation context, evaluation, and stakeholder use.¹⁻³ Before envisioning the sourcing of an AI agent, the organization should define the queue, volume, delays, error types, handoffs, systems, current-state baseline, material consequences, and human burden. It can then determine which material acts are suitable for a bounded Digital Worker role.

The distinction is operational rather than technical. Many AI systems can execute individual tasks, so organizations still tend to deploy and manage them primarily as models, software components, or isolated automations. As AI begins performing more elaborate work on behalf of the organization, a different management unit becomes necessary. Accountable work requires a persistent organizational role through which supervision, authority, evidence, operational history, and lifecycle can be managed consistently over time. Building a Digital Worker therefore begins by defining that managed role rather than simply configuring an AI capability.

Not every backlog item should become a Digital Worker. Some problems require process redesign, policy simplification, staffing, payer negotiation, or elimination of unnecessary work. Digital Labor should not be used to make a poorly designed process run faster. Analyses of administrative AI in prior authorization and billing similarly warn that automation can increase transaction volume or shift costs when workflows and incentives remain unchanged.⁴ The first build decision is therefore operational: is the work worth doing, and which portion should become a digital role?

Translate the backlog into a role

A backlog item becomes a Digital Worker only when the organization can state the job. The role description should identify the workflow, expected outputs, permitted inputs, excluded cases, named supervisor, escalation conditions, success measures, and consequences of error. It should distinguish routine execution from exception work and define the material actions for which approval remains human.

For example, a documentation-preparation Digital Worker might retrieve active payer policy, identify required documentation, compare those requirements with the chart, draft a packet summary, and flag missing evidence. It might be prohibited from submitting the packet without approval, altering clinical documentation, contacting the patient, or resolving ambiguous medical-necessity questions without escalation.

Role translation is a management act: it converts a technology idea into accountable work.

Healthcare work-system research makes the same point from the human side: performance and safety emerge from the interaction of people, tasks, tools, organization, and environment rather than from a technology in isolation.^{5,6} Role translation also reveals whether the proposed role is too broad. If one role requires unrelated authorities, source hierarchies, supervisors, or risk profiles, it is probably several Digital Workers rather than one. In healthcare, scope should define capability—not the reverse.

Create the Digital Worker Record

The Digital Worker Record is the persistent build and operating artifact that prevents a role from becoming an informal bundle of prompts, tools, credentials, and assumptions. It should exist before production and remain current throughout the Digital Worker’s operating life.

At minimum, the record should include the role name, owner, supervisor, workflow, job family, scope, allowed and prohibited actions, authorized systems, identity and access profile, context requirements, source hierarchy, authority band, validation

record, monitoring plan, performance measures, evidence requirements, escalation rules, incident history, change log, dependencies, and retirement criteria.

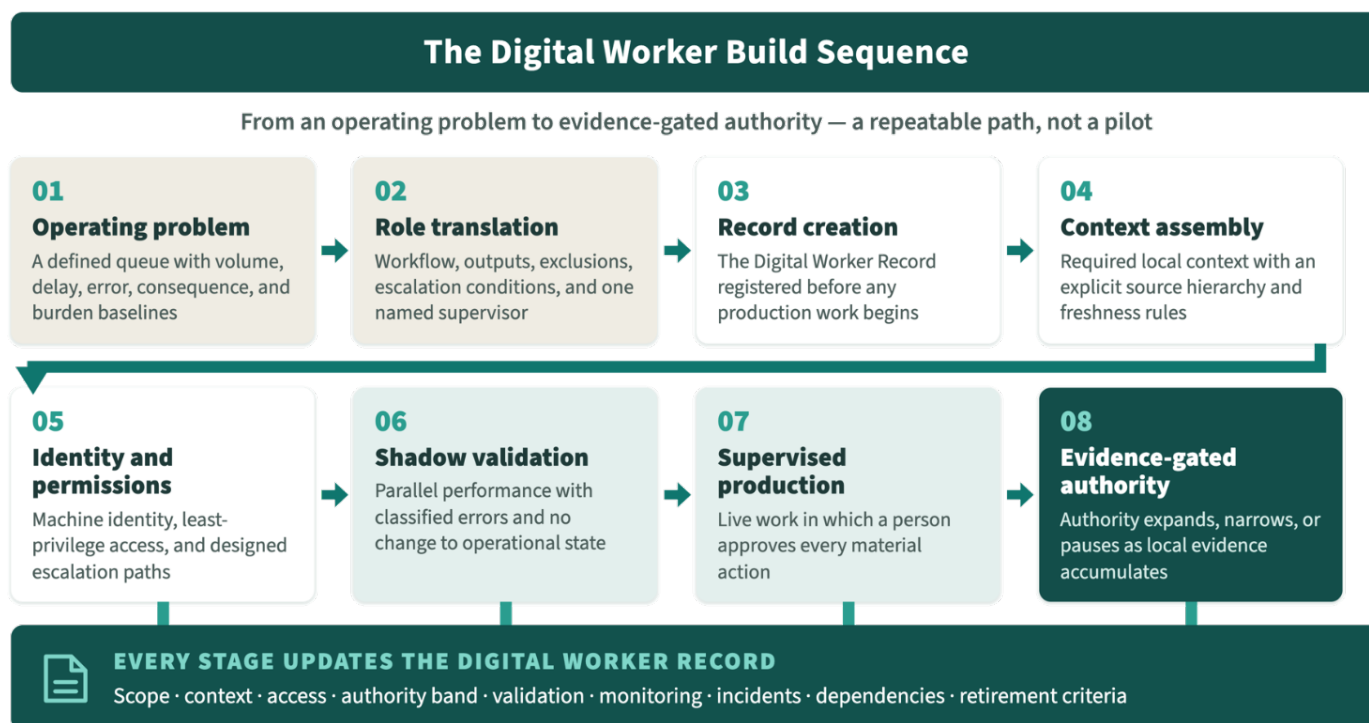


FIGURE 1 - Digital Worker build sequence. A repeatable build sequence moves from operating problem to role design, record creation, context assembly, identity and permissions, shadow validation, supervised production, and evidence-gated authority. Each stage contributes to the maintained Digital Worker Record.

The Digital Worker record draws from documentation traditions such as model cards, dataset datasheets, supplier declarations, internal audits, and AI lifecycle frameworks, all of which emphasize transparency and accountability.⁷⁻¹³ The Digital Worker Record adds the local work lens, including: what role is this system performing for this institution, under whose authority, using which context and systems, and with what evidence? It is the management object through which a Digital Worker can be supervised individually and governed as part of a portfolio.

Without this persistent record, AI capabilities remain difficult to supervise consistently as they change over time. Models, prompts, tools, and configurations may evolve independently, but the organizational role, its authority, accumulated evidence, and operational history must remain continuous if the Digital Worker is to be managed as a durable organizational asset.

Assemble context and source hierarchy

A Digital Worker cannot perform local healthcare work from general knowledge alone. It needs the current organizational context required by the role. That context may include patient and case facts, active payer policy, plan benefit rules, local SOPs, contract terms, documentation templates, workflow state, external portal status, and escalation norms.

The build process should establish an explicit source hierarchy. Which source prevails when payer policy conflicts with an internal summary? How fresh must a policy be before submission? Which EHR field is authoritative for the relevant patient fact? When does a portal value supersede a cached record? Which conflicts require the Digital Worker to stop and request review?

Source hierarchy is a reliability control. Missing, stale, conflicting, or low-authority context should be represented as an operational state that triggers defined behavior—not as an invitation for confident inference.

Design permissions and escalation

Digital Workers should operate through explicit machine identity and least-privilege access. The access profile should match the role rather than the technical convenience of deployment. A Digital Worker that prepares a draft may need read access and permission to update a queue, but not authority to submit, message, or close a case. Credentialing and onboarding should establish who supervises the role, what it may access, what it may do, and how its activity will be reviewed before productive work begins.

Permissions should be paired with designed escalation paths. Escalation is not a failure of automation; it is the expected response to uncertainty, exception, missing context, conflicting evidence, elevated consequence, or prohibited action. A useful escalation carries the case summary, source evidence, completed steps, uncertainty category, proposed next action, and a specific question for the supervisor.

Security and operational safety converge at this boundary. Agentic systems that can use tools and credentials increase the importance of identity, access control, monitoring, containment, and credential revocation.^{14,15} In healthcare, those controls must also preserve privacy and confidentiality obligations for protected health information.¹⁴

Validate in shadow mode

Shadow mode is the safest first test of a Digital Worker in local conditions. The role is performed in parallel without changing operational state or communicating externally. Outputs are compared with current human work, known outcomes, reviewer decisions, and role-specific measures.

Shadow validation should classify errors and exceptions rather than ask only whether the output was correct. Did the Digital Worker retrieve the wrong source, use stale context, miss a local exception, draft an unsupported statement, over-escalate, under-escalate, lose workflow state, or fail to recognize that the case was outside scope? These failure classes identify whether remediation belongs in the model, context, role boundary, interface, workflow, or supervision design.

Healthcare AI evaluation guidance increasingly emphasizes intended use, setting, human-AI interaction, error analysis, deployment context, and lifecycle monitoring.^{16–19} Shadow mode makes those principles operational by testing the Digital Worker against the local role before authority is granted. Aggregate performance can conceal a defective proxy or uneven performance across patient groups, so validation should also examine what the role is optimizing and for whom.²⁰ Performance claims do not simply transfer between settings; recurring local validation in the deploying organization's own case mix is the more relevant standard.²¹

Move to supervised production

Supervised production allows a Digital Worker to contribute to live work while a person approves material action. This stage should be designed rather than improvised. The supervisor should see the source evidence, proposed action, uncertainty, policy version, context completeness, and expected consequence. Approval, rejection, override, correction, and escalation should become structured operating data.

The organization should measure both Digital Worker performance and reviewer burden. Correct drafts that require exhausting verification may not improve the workflow. Excessive escalation may create a new queue. Under-escalation may make the role appear efficient while increasing escaped material error. Supervision is therefore part of the operating cost and reliability profile of the role, not an unmeasured externality.

Human-centered deployment research shows that real clinical environments can alter both system performance and staff experience.²² The same applies to administrative Digital Workers. Supervised production is where the organization learns how the role behaves under live workload, changing context, interruptions, and human review.

Coach, measure, and adjust

Digital Workers should improve through structured feedback. Overrides, corrections, near misses, tool failures, and incident reviews should not remain anecdotal. They should update prompts, rules, context assets, source hierarchy, permissions, escalation thresholds, interfaces, evaluation cases, or workflow design. This is continual quality improvement applied to Digital Labor: monitoring and updating are assigned responsibilities with owners and cadence rather than ad hoc reactions to complaints.²³

Useful measures include first-pass yield, cycle time, rework, escaped material error, near misses, override rate, escalation precision, false-escalation rate, reviewer burden, source-staleness rate, tool-call failure, correction incorporation, and

role-specific outcomes. A prior-authorization Digital Worker should not be evaluated by the same scorecard as a referral or denial-response role. The measurement system should follow the role and the consequences of its work. This is where Digital Labor begins to compound and scale: the organization learns not only how the Digital Worker performs, but also where the underlying workflow, context, and policies are poorly designed.

Manage the Digital Worker as an asset

A Digital Worker is a managed asset only when the organization can observe and govern its value, risk, dependencies, evidence, and lifecycle. It should therefore exist within a digital workforce portfolio. The portfolio view identifies which roles exist, which systems and workflows they touch, which teams they burden or relieve, which authority they hold, and which shared dependencies could create correlated failure.

Machine-learning operations and software-engineering research describe maintenance burdens created by data dependencies, entanglement, feedback loops, monitoring gaps, and changing environments.^{24–26} Digital Workers add operational dependencies: role scope, tool access, payer policy, local SOPs, supervisor availability, external portals, and evidence requirements. The managed asset is therefore not merely technical; it is the complete role operating within an AI-enabled work system.

Portfolio management also helps leaders prioritize shared infrastructure. When many Digital Workers depend on the same payer-policy source, identity pattern, source hierarchy, or supervisor queue, that component becomes organizational infrastructure. Its reliability, ownership, capacity, and change history should be managed accordingly.

Retire or redesign when the role no longer fits

A Digital Worker should be retired or redesigned when its role no longer fits the operating environment. Triggers may include policy change, workflow redesign, system migration, repeated near misses, unacceptable reviewer burden, tool instability, new regulatory requirements, poor outcomes, or loss of business value. Many are forms of dataset or environment shift, which deployed clinical AI has shown to be a routine operating condition rather than an edge case.²⁷ Regulatory guidance similarly assumes that AI-enabled systems will change over their operating life and expects modifications to be planned, bounded, and evidenced.^{28,29}

Retirement should be orderly. The organization should preserve evidence, close access, archive the record, transfer remaining work, document lessons, and decide whether the role should be redesigned. A Digital Worker that cannot be retired safely was not fully managed as an asset. The same discipline applies to incidents: detection, containment, credential revocation where necessary, rollback, root-cause review, remediation, and a fresh authority decision should be designed before they are needed. An organization that can reconstruct which Digital Worker acted, under which version, with which sources and approvals, has an operational event to manage; one that cannot has a control and accountability gap.

The lifecycle also supports disciplined expansion. Strong performance in one service line does not justify automatic expansion to another. Context, risk, volume, supervisors, systems, and source hierarchy may differ. Expansion is therefore a new evidence decision, and recurring local validation applies to each setting on its own terms.²¹

The Digital Worker Lifecycle

Build once, then govern continuously — authority moves in both directions as local evidence changes

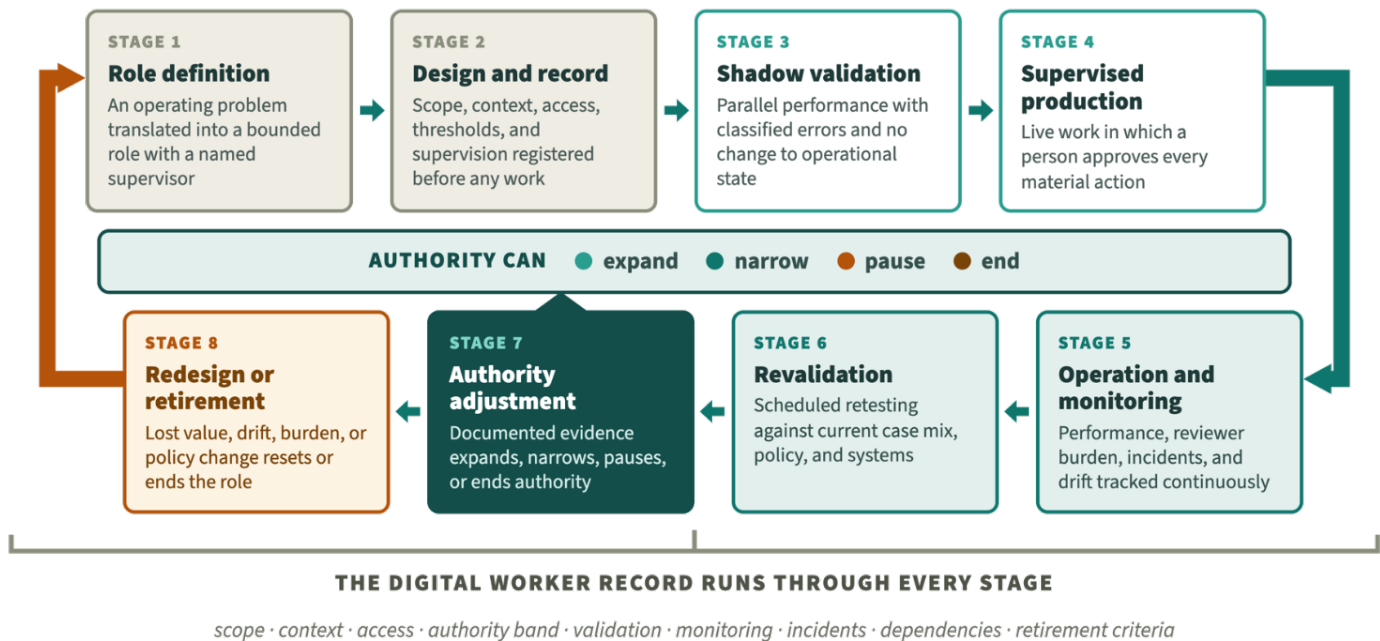


FIGURE 2 - Digital Worker lifecycle. A Digital Worker moves through proposed role, design and record, shadow validation, supervised production, operation and monitoring, revalidation, authority adjustment, and retirement or redesign. Authority can expand, narrow, pause, or end as local evidence changes.

The executive test

Before a proposed Digital Worker enters production, executives do not need to judge the underlying model in isolation. They need evidence that the role has been built as a managed asset. The

questions below test whether the operating problem, role boundary, record, validation, authority, and lifecycle are observable and controlled. The absence of an evidence-based answer is itself a finding.

Executive Test Before Production

Six questions that test whether the role was built as a managed asset — the absence of an answer is itself a finding

- 1 **OPERATING PROBLEM** What operating problem is this Digital Worker meant to solve?
EVIDENCE A defined queue with volume, delay, error, consequence, and burden baselines — not a technology aspiration
- 2 **ROLE BOUNDARY** Is the role narrow enough to define scope, context, access, exclusions, and supervision?
EVIDENCE A written role description with outputs, exclusions, escalation conditions, material actions, and one named supervisor
- 3 **THE RECORD** What must the Digital Worker Record contain before production?
EVIDENCE A registered record covering scope, context, access, authority, validation, monitoring, incidents, and retirement criteria
- 4 **VALIDATION** What evidence will shadow mode produce?
EVIDENCE Classified errors and exceptions measured against current work and consequences — not a single aggregate accuracy number
- 5 **AUTHORITY** Which material actions require human approval or remain prohibited?
EVIDENCE An enforced approval boundary, with overrides, corrections, and escalations captured as structured data
- 6 **LIFECYCLE** Which thresholds trigger expansion, narrowing, revalidation, redesign, or retirement?
EVIDENCE Role-specific thresholds reviewed on a schedule, with authority able to move in both directions



A Digital Worker is ready for production when every row has evidence, not intention

Close: repeatability is the practical advantage

The practical advantage of Digital Operations is repeatability. A repeatable build process does not make every Digital Worker complex; it prevents the organization from rediscovering role definition, access control, context assembly, validation, supervision, evidence, incident response, and retirement for every use case.

This discipline gives executives a more useful question than whether a pilot appeared impressive. They can ask whether the proposed Digital Worker has a bounded role, a maintained record, a named supervisor, current context, enforced access boundaries, local validation evidence, measurable reviewer burden, monitoring, incident response, and lifecycle criteria.

An institution that can repeatedly move an operating problem through role design, record creation, context assembly, credentialing, shadow validation, and supervised production is no longer accumulating disconnected pilots. It has created a production capability for managed Digital Labor. The scalable advantage is not any single model or Digital Worker; it is the institutional ability to build, supervise, improve, and retire bounded roles as evidence changes.

Glossary of key concepts

- **Managed asset.** A Digital Worker treated as an operational asset with defined value, risk, dependencies, evidence, authority, and lifecycle.
- **Role translation.** The conversion of an operating problem into a bounded Digital Worker role with defined outputs, exclusions, authority, and supervision.
- **Shadow mode.** A validation stage in which a Digital Worker performs the role in parallel without changing operational state or communicating externally.
- **Supervised production.** Live operation in which a human approves material action and approvals, overrides, corrections, and escalations become structured evidence.
- **Digital Worker Record.** The maintained institutional record connecting a Digital Worker's role, context, access, validation, performance, incidents, authority, dependencies, and lifecycle.

Live Article information

This PDF is a dated snapshot of a continuously maintained Live Article. Material updates undergo editorial review, and substantive public revisions are released when warranted.

The current maintained article and evidence base are available through the ConverGen Digital Operations Advisor: www.convergen.ai/advisor/

Peer-review status: This public edition has not undergone external peer review.

Competing interests: The author is the founder and CEO of ConverGen AI, which develops a Digital Operations System for healthcare organizations.

AI-assisted tools: AI-assisted tools supported research synthesis, drafting, editing, and figure production. The author reviewed and takes responsibility for all content.

Author ORCID: 0009-0008-1814-1983.

References

1. Wiens, J. et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat. Med.* 25, 1337–1340 (2019). <https://doi.org/10.1038/s41591-019-0548-6>.
2. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. & King, D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* 17, 195 (2019). <https://doi.org/10.1186/s12916-019-1426-2>.
3. Sendak, M. P. et al. A path for translation of machine learning products into healthcare delivery. *EMJ Innov.* 4, 16–20 (2020). <https://doi.org/10.33590/emjinnov/19-00172>.
4. Peterson Health Technology Institute. Administrative AI: Current Use and Potential Impact (2026). <https://phti.org/administrative-ai-current-use-and-potential-impact/>. Accessed 23 Jul 2026.
5. Carayon, P. et al. Work system design for patient safety: the SEIPS model. *Qual. Saf. Health Care* 15 Suppl. 1, i50–i58 (2006). <https://doi.org/10.1136/qshc.2005.015842>.
6. Holden, R. J. & Carayon, P. SEIPS 101 and seven simple SEIPS tools. *BMJ Qual. Saf.* 30, 901–910 (2021). <https://doi.org/10.1136/bmjqs-2020-012538>.
7. Mitchell, M. et al. Model cards for model reporting. *Proc. Conf. Fairness Accountability Transparency* 220–229 (2019). <https://doi.org/10.1145/3287560.3287596>.
8. Gebru, T. et al. Datasheets for datasets. *Commun. ACM* 64, 86–92 (2021). <https://doi.org/10.1145/3458723>.
9. Raji, I. D. et al. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *Proc. 2020 Conf. Fairness Accountability Transparency* 33–44 (2020). <https://doi.org/10.1145/3351095.3372873>.
10. Arnold, M. et al. FactSheets: increasing trust in AI services through supplier's declarations of conformity. *IBM J. Res. Dev.* 63, 6:1–6:13 (2019). <https://doi.org/10.1147/JRD.2019.2942288>.
11. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1 (2023). <https://doi.org/10.6028/NIST.AI.100-1>.
12. International Organization for Standardization. ISO/IEC 42001:2023 Artificial intelligence - Management system (2023). <https://www.iso.org/standard/81230.html>. Accessed 23 Jul 2026.
13. International Organization for Standardization. ISO/IEC 23894:2023 Artificial intelligence - Guidance on risk management (2023). <https://www.iso.org/standard/77304.html>. Accessed 23 Jul 2026.
14. U.S. Department of Health and Human Services. Summary of the HIPAA Security Rule (2024). <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>. Accessed 23 Jul 2026.
15. OWASP Agentic Security Initiative. Agentic AI - Threats and Mitigations. OWASP (2025). <https://genai.owasp.org/resource/agentic-ai-threats-and-mitigations/>. Accessed 23 Jul 2026.
16. Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat. Med.* 26, 1364–1374 (2020). <https://doi.org/10.1038/s41591-020-1034-x>.
17. Rivera, S. C. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat. Med.* 26, 1351–1363 (2020). <https://doi.org/10.1038/s41591-020-1037-7>.
18. Vasey, B. et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat. Med.* 28, 924–933 (2022). <https://doi.org/10.1038/s41591-022-01772-9>.
19. Lekadir, K. et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ* 388, e081554 (2025). <https://doi.org/10.1136/bmj-2024-081554>.
20. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 447–453 (2019). <https://doi.org/10.1126/science.aax2342>.
21. Youssef, A., Pencina, M., Thakur, A., Zhu, T., Clifton, D. & Shah, N. H. External validation of AI models in health should be replaced with recurring local validation. *Nat. Med.* 29, 2686–2687 (2023). <https://doi.org/10.1038/s41591-023-02540-z>.
22. Beede, E. et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. *Proc. 2020 CHI Conf. Hum. Factors Comput. Syst.* 1–12 (2020). <https://doi.org/10.1145/3313831.3376718>.
23. Feng, J. et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. *NPJ Digit. Med.* 5, 66 (2022). <https://doi.org/10.1038/s41746-022-00611-y>.
24. Sculley, D. et al. Hidden technical debt in machine learning systems. *Adv. Neural Inf. Process. Syst.* 28, 2503–2511 (2015).

25. Amershi, S. et al. Software engineering for machine learning: a case study. Proc. IEEE/ACM 41st Int. Conf. Softw. Eng. Softw. Eng. Pract. 291–300 (2019). <https://doi.org/10.1109/ICSE-SEIP.2019.00042>.
26. Kreuzberger, D., Kuehl, N. & Hirschl, S. Machine learning operations (MLOps): overview, definition, and architecture. IEEE Access 11, 31866–31879 (2023). <https://doi.org/10.1109/ACCESS.2023.3262138>.
27. Finlayson, S. G. et al. The clinician and dataset shift in artificial intelligence. N. Engl. J. Med. 385, 283–286 (2021). <https://doi.org/10.1056/NEJMc2104626>.
28. U.S. Food and Drug Administration. Clinical Decision Support Software: Guidance for Industry and Food and Drug Administration Staff (2026). <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software>. Accessed 23 Jul 2026.
29. U.S. Food and Drug Administration. Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions (2025). <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence>. Accessed 23 Jul 2026.