

The Agent Lifecycle vs. the Literature

The Auto-Agent Intent & Execution Lifecycle (v3.96.0) vs Regulatory Standards and Published Research

Prepared: August 16, 2026

Scope: Cross-reference of AGENT_LIFECYCLE_SPEC.md (v3.96.0, HEAD) against (i) the regulatory canon for automated trading risk controls — SEC Rule 15c3-5, MiFID II RTS 6, FINRA Notice 15-09, CFTC Reg AT lineage, FIA best-practice guides, and the Knight Capital enforcement order; (ii) order-lifecycle engineering standards (FIX order-state model, idempotency practice); (iii) human-in-the-loop automation research; and (iv) exit-management and operational-risk literature. Roughly 120 primary sources across four research sweeps; consolidated bibliography in Section 9. Companion to the prior engine-wide and component reviews; overlapping ground is referenced, not repeated.

1. Executive Summary

This is the strongest of the three reviews for the codebase. The lifecycle's core mechanics independently reinvent, or deliberately implement, most of what regulation and standards require of institutional automated trading: the intent hash is a faithful idempotency-key/CIOrdID implementation of 15c3-5's mandated duplicate-order control [1][26]; the fill-race guard is exactly FIX Appendix D's "too late to cancel" semantics [32]; day-only orders mirror the reasoning that led NYSE to abolish native GTC [38]; the mode ladder (shadow → semi-auto → full-auto-paper) is a private-scale implementation of RTS 6's testing-and-phased-deployment regime [5][21]; the subtract-only ML gate, PANIC demotion, and halt-on-missing-evidence rules are textbook IEC-61508 safe-state asymmetries [58]; every-cycle reconciliation with block-on-break *exceeds* the FIA baseline and directly implements the Knight lesson [14][20]; and the typed production confirmation is the "hard stop" intervention class — the only approval-friction design with strong effect sizes in the clinical-decision-support literature [48][49].

Against that, the review found **3 divergences**, **8 gaps**, and **9 refinements**. The highest-priority items:

- 1. The kill switch blocks new orders but does not cancel working ones.** RTS 6 Article 12 defines kill functionality as the ability to *cancel immediately any or all unexecuted orders* [5]; FIA and FINRA guidance agree [8][14]. A working day order at trip time survives until fill or expiry. Day-orders-only bounds the tail, but cancel-on-trip is the standard.
- 2. Per-order price/notional collars are incomplete.** 15c3-5(c)(1)(ii) mandates *price and size* parameters per order [1][3]; the agent's structure floors act as an economic collar for defined-risk spreads, but there is no explicit max-notional-per-order or price-vs-mark deviation check — and the manual path's 100-contract cap admits a \$900k notional inside the rules. Options obvious-error thresholds are the published calibration for collars, and they cut against relying on exchanges to bust bad fills [30].
- 3. Partial fills are not first-class.** The poll maps `partially_filled` to `WORKING` but no `cum/leaves` quantity flows through the fill-race guard, the reaper, or the position-open chain — the FIX state model treats `CumQty/LeavesQty` as core state precisely because multi-leg packages part-fill [32]. A half-filled condor at day end currently resolves through the honesty of the reconciliation pass rather

than by design.

- 4. Reaping by absence is weaker than the standard.** Orders absent from a broker listing after their session are reaped as expired; the FIX-era standard is an authoritative per-order status query (35=H analog — "get order by id"), since a late fill misclassified as expiry is a position-tracking corruption the current design would only catch at reconciliation [32][34].
- 5. No LULD/halt awareness.** Nothing in the lifecycle refuses to trade while the underlying is in a limit state or halted — the one condition where quotes are least reliable and obvious-error treatment is least favorable [28][30]; the Flash Crash record is the canonical argument [22].
- 6. The daily loss kill needs an outer layer and a pre-breach tier.** The –2% daily kill matches prop-industry convention (static, realized+unrealized, hard lockout) [24][63], but industry practice layers it inside a multi-day/max-drawdown governor (~5-6%) with alerts *before* breach [16][63] — short-vol losing days cluster.

The most striking overall finding: several of the lifecycle's most distinctive choices — the human-exit kill-switch bypass, the never-reprice-human-orders-unattended rule, the assistant-stages-never-executes boundary, the below-limit standing authorization — turn out to have *direct* analogs in published standards (fail-operational exceptions in functional safety [58]; RTS 6's human-oversight requirements [5]; the IOSCO/FCA "accountable human for every order" principle [54][55]; delegation-of-authority matrices [56]). These were good instincts that the literature retroactively endorses, with tightening suggestions noted in §6.

Scorecard

Lifecycle element	Verdict	Headline
Intent hash idempotency + preview tokens	Aligned	Faithful 15c3-5 duplicate-control / CIOrdID / idempotency-key implementation
Transactional state machine	Aligned	The atomic check-and-transition all idempotency literature requires
Fill-race guard on cancel	Aligned (partial)	FIX "too late to cancel" semantics; needs quantity, not boolean
Partial-fill handling	Gap	Cum/leaves not first-class through guard, reaper, open chain
Day-order reaper	Deviation	Confirm terminal status per order, not absence from a listing
Kill switch	Deviation	Blocks new orders; RTS 6 kill = cancel working orders too
Price/notional collars	Gap	No per-order max notional or price-vs-mark check; manual path widest
LULD/halt awareness	Gap	No no-trade condition on halted/limit-state underlyings
Daily loss kill (–2%)	Aligned (partial)	Convention-consistent; add pre-breach alert + multi-day governor
Reject-counter kill (3/day)	Aligned	RTS 6 Art. 15 throttle-and-disable pattern
Reconciliation + block-on-break	Ahead	Exceeds FIA baseline; verified-fetch honesty is exemplary
Force-close discipline	Ahead	Independent-source verification, no fabricated P&L
60 s polling as sole channel	Deviation	Standard is streaming/drop-copy primary, polling as reconciliation
Reprice cadence (60 min clock)	Refinement	Event-driven on mid-drift; fill hazard decays in minutes
Paper fill ladder	Refinement	Right shape (GLFT concession); wrong units (% of mid vs spread), linear not accelerating; conflates pricing policy with fill simulation
Mode ladder + typed confirmation	Aligned	RTS 6 phased deployment; hard-stop friction evidence

Lifecycle element	Verdict	Headline
Below-limit standing auth	Aligned (partial)	DoA-matrix pattern; add aggregate daily cap + re-affirmation
Subtract-only ML gate, PANIC demotion, evidence halts	Aligned	IEC 61508 safe-state asymmetry, near verbatim
Human-exit kill-switch bypass	Aligned (conditional)	Defensible fail-operational exception; must be drilled and narrowly scoped
Exit rules (50%/2x/21/7 DTE, ex-div, put-EEP)	Aligned	Every guard has direct literature support; 2x stop is consistency, not tail defense
Mark quality feeding triggers	Refinement	Add crossed/one-sided quote rejection + trigger persistence
Notifications & dedup	Aligned (partial)	ISA-18.2-consistent; add positive heartbeat ("silence ≠ success")
Process controls (drills, self-assessment, deploy verification)	Gap (process)	RTS 6 Art. 9 / Knight lessons — documentation and drills, not code

2. Methodology

Reference point: AGENT_LIFECYCLE_SPEC.md (v3.96.0, from full source reads at HEAD with ~18 test modules' pinned invariants). Four research sweeps: regulatory/industry risk-control standards; order-lifecycle engineering (FIX, idempotency, microstructure of order revision, quoting theory); human-in-the-loop automation and safe autonomy; exit management and operational risk. A retail agent trading through a broker is a *customer* — 15c3-5/RTS 6 obligations legally sit with brokers and venues — so regulatory texts are used here as the canonical *design templates* they have become, not as compliance obligations. Verdicts: **Aligned** / **Partial** / **Deviation** / **Divergent** / **Gap**, plus **Ahead** where the design exceeds the documented baseline.

3. Risk-Control Stack vs. the Regulatory Canon

3.1 What maps cleanly

The lifecycle implements, without apparently having been designed from the texts: aggregate exposure caps hard-linked to order blocking (20% open risk *including active intents* — precisely the control whose absence sank Knight [20]); per-order size caps; duplicate-order prevention as a mandated control (15c3-5(c)(1)(ii) [1][3]) via the intent hash; a reject-count throttle-and-disable (RTS 6 Art. 15's "repeated executions" pattern [5]); reconciliation with block-on-break (RTS 6 Art. 17 endorses exactly this, pushed toward continuous [5]); limits living outside the agent's write path (15c3-5's "direct and exclusive control" [3]); fail-closed defaults on missing evidence (FIA: "no order where no size limit has been set" [14]); and market-hours-only operation with quiet windows (consistent with the volatility-mitigant logic of IOSCO Rec. 2 [18]).

3.2 The kill switch — DEVIATION

RTS 6 Article 12 defines kill functionality as the ability "as an emergency measure, to cancel immediately any or all of its unexecuted orders" [5]; FIA's 2015/2024 guides and FINRA 15-09 describe kill switches the same way [8][14][16]. The agent's kill switch blocks *new* submissions (and auto-exits queue behind it), but working day orders at trip time keep working until filled or expired. Day-orders-only caps the horizon at the close; the standard caps it at the trip. **Recommendation:** on

any kill-switch trip (reject counter or daily loss), cancel all working entry orders at their stamped brokers (working *exit* orders may reasonably survive — they reduce risk, the same asymmetry the design already applies to approvals).

3.3 Per-order price and notional checks — GAP

15c3-5(c)(1)(ii) requires per-order *price* and *size* parameters; RTS 6 Art. 15 mandates price collars, **maximum order values**, and volumes [1][5]. The agent's structure floors (credit/width ≥ 0.20 , debit/width ≤ 0.45) and max-loss caps function as an economic collar for defined-risk spreads — a genuine partial credit unusual at retail. Missing: (a) an explicit **max premium/notional per order** — most exposed on the manual path, where 100 contracts of a \$90 option pass every current check; (b) a **price-vs-mark deviation collar** calibrated *tighter* than the options obvious-error minimum-deviation table (\$0.25–\$2.00 by premium tier [30]) — the published lesson being that exchanges *adjust* rather than bust most bad options fills, so a bad fill usually stands; (c) an **intrinsic-value sanity check** (the retail analog of exchange price-reasonability). The v3.96 delayed-chain refusal (409 on yfinance-priced tickets) is the right family of control and shows the pattern to extend.

3.4 Halt/LULD awareness — GAP

Nothing refuses to act while the underlying is halted or in an LULD limit state — conditions under which options quotes are unreliable, obvious-error treatment is unfavorable, and the Flash Crash record shows automated flows do their damage [22][28][29]. One quote-level check (underlying halted / limit state / quote flagged) added to the entry filters and the reprice paths closes it.

3.5 Message and re-entry throttles — PARTIAL

The 60 s tick and 40-chain-fetch cap discipline message rates, but there is no explicit **daily order-count ceiling** (Reg AT/RTS 6 pattern [10][5]). Same-symbol **re-entry cooldown** after an exit is mostly covered by the per-symbol position cap and — accidentally — by the intent hash: identical legs after a FILLED intent are blocked until the 180-day purge. That accident cuts both ways: it throttles same-structure re-entry (good) and it can *silently* block a legitimately repeated identical trade (surprising). Making the re-entry policy explicit (a configurable cooldown, and day-scoping the entry hash) converts an accident into a control.

3.6 Dead-man behavior — PARTIAL

FIA's cancel-on-disconnect/heartbeat canon [14][16]: on loss of the controlling system, working orders should die. The agent's crash story is: leases expire (55 s/360 s), day orders expire at the close, and reconciliation catches drift — a coarse dead-man. The literature-consistent upgrade is explicit: a startup pass that cancels unknown working orders (crash-recovery sweep), and a documented out-of-band runbook (the broker app is the independent order-management channel — write it down).

3.7 Knight's process lessons — GAP (process, not code)

The Knight order's specific findings [20][21] map to process controls the repo doesn't evidence: **deployment verification** (startup assertion of code+config version — the "8th server" lesson); **kill-first incident response** (a written runbook: trip the switch before diagnosing; never roll back while order flow is enabled); **routed, actionable alerts** (97 ignored emails — the agent's inbox/email/push channels are good; a *positive daily heartbeat* is the missing piece, §6.3); and an **annual written self-assessment** of the control stack with evidence of each block firing (15c3-5(e)/RTS 6 Art. 9 [3][5]). For a solo operator these are documentation and drills, not engineering.

4. Order-Lifecycle Engineering vs. FIX-Era Standards

4.1 Idempotency — ALIGNED (with two refinements)

The intent hash (price excluded, so a reprice chains like FIX cancel/replace under one economic order) plus single-use preview tokens plus transactional from-state guards is a faithful composite of the CIOrdID/idempotency-key standard [26][32][40]. Refinements from the standards: (a)

query-before-retry on ambiguous submits — a place-order timeout currently records ORDER_REJECTED; if the broker actually accepted, the error is caught only later by reconciliation (defense-in-depth exists, but the standard resolution is an immediate order-status query before recording failure); (b) **hash scoping** — see §3.5's FILLED-blocks-re-proposal edge.

4.2 Partial fills — GAP

FIX makes CumQty/LeavesQty first-class because multi-leg packages part-fill [32]. The poll maps partially_filled → WORKING (correct holding state) but: the fill-race guard checks a boolean; the reaper infers from absence; and the open chain assumes full quantity. Between two polls an order can go WORKING → PARTIAL → CANCELED (Deribit's documented trap: a cancel *and* a fill notification for the same order [33]). **Recommendation:** carry filled quantity on every poll read; open positions at actual filled contracts; treat cancel-with-partial as a fill event for the filled portion.

4.3 Reaper and poll-gap tolerance — DEVIATION

Confirm terminal status by per-order query rather than listing absence (§1 item 4). More broadly, the standard architecture is streaming order events with polling as the reconciliation layer [34][36]; at retail cadence, poll-only is workable *if* the poller tolerates collapsed transitions (it does: WORKING→FILLED direct is legal) and *if* the cancel path polls immediately around the cancel (it does — the fill-race guard) — the remaining exposure is the quantity issue above and the reaper inference.

4.4 Reprice cadence and the ladder — REFINEMENT

Fill hazard decays fast: execution-time distributions are hypersensitive to distance-from-mid, and most of the fill probability at a given price resolves early in an order's life (Lo-MacKinlay-Zhang [37]); revision in real books is event-driven on quote movement (Fong-Liu [39]; Hasbrouck-Saar's fleeting-order world [35]). The 60-minute clock with a 3-reprice lifetime cap is a defensible *monitoring-cost* policy, but the literature-consistent upgrade is **event-driven repricing**: trigger the stale pass when the mid has drifted beyond tolerance, with the clock as fallback. For the ladder, quoting/liquidation theory (Avellaneda-Stoikov; Guéant-Lehalle-Fernandez-Tapia's limit-order liquidation — exactly the "work then concede toward the touch" problem [41][42]) endorses the shape and corrects the parameters: concessions denominated in **spread units** (5% of mid is several spreads on a wide option, sub-tick on a tight one), an **accelerating** schedule near the deadline rather than linear steps, and an explicit **never-beyond-natural** cap.

4.5 Paper-fill realism — REFINEMENT

The ladder conflates the *pricing policy under test* with *fill simulation*. Standards separate them: simulate fills with a distance-based hazard (cross-rule minimum: the paper order fills only when the opposite side crosses the limit), and let the ladder be strategy [37][44][45]. This matters most on the **simulated plane**, which fills at mid instantly — the documented optimism bias (IB's own paper-trading disclaimer [45]). The engine's per-plane slippage telemetry and sim-vs-sandbox gap metric are precisely the calibration instruments the literature asks for; wire them back into a sub-100% mid-fill probability for the simulated plane.

5. Human Oversight vs. the HITL Literature

The mode ladder maps onto the canonical levels-of-automation scale (shadow $\approx$ LOA 2–4, semi-auto $\approx$ LOA 5 "computer executes if human approves," full-auto-paper = LOA 7 quarantined from consequences) and follows Parasuraman-Sheridan-Wickens' core prescription: automate sensing/analysis aggressively, keep real-money action implementation at LOA 5 [46][47]. Typed production confirmation is supported by the strongest evidence in the approval literature — passive alerts are overridden at 90%+ rates while hard stops change behavior [48][49] — and the tiered friction (one-tap reprice accepts, \$250 carve-out) matches the "proportional hard stops" caveat. The assistant's stage-but-never-execute boundary and the accountable-human principle are exactly where 2023–26 regulator thinking on AI agents landed (IOSCO CR/01/2025; FCA/SM&CR: no "the model did it" defense [54][55]).

Literature-backed additions:

- 1. Kill-switch drills** — RTS 6/FCA expect kill functionality *tested*; the Central Bank of Ireland's thematic review found untested kill switches the modal failure [21][22b]. The human-exit bypass path especially: an exception path that overrides the kill switch is itself a safety function and must meet the same integrity bar (IEC 61508 framing [58]). Schedule drills for both.
- 2. Aggregate cap + anti-splitting + re-affirmation on the standing authorization** — delegation-of-authority practice caps *cumulative* per-period exposure, forbids structuring under the threshold, and expires standing consents for periodic re-signature [56]. The 20% open-risk cap partially covers aggregation; a daily count/total for below-limit auto-executions and a periodic re-affirmation of the typed enable complete the pattern.
- 3. Positive heartbeat** — ISA-18.2/aviation alerting require the alarm system's own health to be monitored: deduped silence and broken silence are indistinguishable to the human [51][52]. A daily "agent alive, N evaluated, M queued, 0 alerts" digest closes the loop the notification dedup opens.
- 4. Approval-behavior monitoring** — automation-complacency research says consent behavior degrades under consistent reliability [49][50]; tracking the human's approval rate vs. agent confidence (rubber-stamp drift) is the operational mitigation, and the attestation record (who approved, what they saw, exact order) is the accountability artifact regulators actually want [55].

6. Exit Management & Operational Risk

6.1 The exit stack — ALIGNED (strongly)

Every guard has direct support: the **21-DTE/7-DTE** pair (ATM gamma scales $1/\sqrt{T}$ — $\sim 1.5\times$ at 21 DTE and $\sim 4.7\times$ at 2 DTE vs 45; assignment/pin/contrary-exercise risk concentrates in the final days, with the documented weekend-gap catastrophe class [64][66][71]); the **ex-div guard** (Pool-Stoll-Whaley: a professional arbitrage population exercises against short ITM calls at ex-div — \$491M harvested from those who don't manage it [67]); the **put early-exercise advisory** (Barraclough-Whaley's \$1.9B of forfeited interest; the $K(1 - e^{-rT})$ boundary is the textbook condition, and Jensen-Pedersen argues for triggering *early*) [68][69]. The 7-DTE hard guard is stricter than common retail practice and correctly so for unattended operation.

6.2 The 2 $\times$ stop, correctly framed — ALIGNED (partial)

Stops on short premium are consistency devices, not tail defenses: systematic backtests find credit-multiple stops drag EV slightly while capping ordinary losses, and in the February 2018 spike **every stop variant was gapped through** — marks blew past triggers and wide spreads prevented

execution near them [60][61][62]. Two implications the design should absorb: (a) sizing and the daily kill carry the tail, not the stop — already the design's structure; (b) IV spikes produce *transient* mark excursions through 2× that revert — a **persistence filter** (N consecutive marks beyond trigger) is supported before firing the urgent exit, especially given §6.3.

6.3 Mark quality feeding the triggers — REFINEMENT

Mids from wide/crossed/one-sided options quotes are the noisiest input in the lifecycle, and they feed both the 2× stop and the −2% kill. The all-or-nothing per-leg policy is the *correct* call (mixed-timestamp legs create phantom P&L), and industry marks are deliberately not mids (OCC's binomial marks [70]). Additions: reject crossed/locked and one-sided quotes as mark inputs; sanity-band the mid against a model price; require trigger persistence; and **fail toward alert** when marks are unavailable for many consecutive cycles — quote degradation is exactly the crisis signature (Feb-2018), and prolonged mark silence should escalate rather than idle.

6.4 Daily loss architecture — ALIGNED (partial)

The −2% static daily kill on realized+unrealized matches prop-industry convention (range 1–2.2%, unrealized counted, hard lockout, human re-enable) [63][24]. Missing layers per the same convention and FIA guidance: a **pre-breach alert tier** (warn at −1%) [16] and a **multi-day drawdown governor** (~5–6% cumulative that de-risks beyond one session) — short-vol losing days cluster, and a daily-reset-only limit is the documented gap [63]. Marking-glitch trips are the acceptable failure direction (halts are cheap) — the same asymmetry FIA endorses.

6.5 Get-me-out execution — ALIGNED

Width-capped marketable limits for urgent exits are the textbook-correct design: market orders in options are uniformly warned against, and multi-leg packages generally lack trade-through protection on complex books — a multi-leg market order has no price backstop at all [72][73]. The escalation ladder (mid → +5%·step → width cap) is the recommended shape; the instant-width path only fires when no mark exists, which is the right emergency posture.

6.6 Reconciliation & force close — AHEAD

Every-cycle reconciliation with per-symbol blocking exceeds the FIA "frequent reconciliation" baseline [14][16]; the verified-fetch honesty (None ≠ empty book), the single unscoped error blocking everything, and the force-close preconditions (live independent verification of zero on every leg; **no fabricated P&L**) implement the independent-source-of-truth principle better than the guidance asks. This is the part of the lifecycle to hold constant while fixing the rest.

7. Consolidated Findings

Deviations/Divergences:

#	Finding	Standard/Evidence	Action
D1	Kill switch doesn't cancel working entry orders	RTS 6 Art. 12; FIA; FINRA 15-09 [5][8][14]	Cancel working entries on trip (exits may survive)
D2	Reaper infers expiry from listing absence	FIX 35=H recovery pattern [32]	Per-order terminal-status query before reaping

#	Finding	Standard/Evidence	Action
D3	Poll-only order events at 60 s	Drop-copy/streaming-primary standard [34][36]	Acceptable at retail <i>with</i> §4.2/§4.3 fixes; document the trade

Gaps:

#	Finding	Evidence	Action
G1	Partial-fill quantity not first-class (guard, reaper, open chain)	FIX CumQty/LeavesQty [32][33]	Carry filled qty everywhere; open at actual fills
G2	No max-notional-per-order or price-vs-mark collar (manual path widest)	15c3-5(c)(1)(ii); RTS 6 Art. 15; obvious-error tables [1][5][30]	Add notional ceiling + deviation collar + intrinsic sanity
G3	No halted/LULD no-trade condition	Flash Crash; LULD/obvious-error mechanics [22][28][30]	Quote-state check in entry filters and reprice paths
G4	No multi-day drawdown governor or pre-breach alert tier	Prop/FIA layered-limit practice [16][63]	Warn at -1% ; cumulative $\sim -5\%$ multi-session de-risk
G5	No positive heartbeat (deduped silence $\neq$ health)	ISA-18.2; aviation alerting [51][52]	Daily "alive + summary" digest
G6	No explicit daily order-count ceiling or re-entry cooldown (hash accident covers partially)	Reg AT; RTS 6 Art. 15 [5][10]	Explicit ceilings; day-scope the entry hash
G7	Ambiguous submit failures recorded as rejection without a broker query	Idempotency practice [26][40]	Query order status before recording failure
G8	Process controls undocumented: kill drills, deployment verification, incident runbook, annual self-assessment	RTS 6 Arts. 9/12; Knight order [3][5][20]	Write and drill them

Refinements: event-driven repricing on mid-drift with the clock as fallback (R1, [37][39]); ladder concessions in spread units, accelerating, capped at natural (R2, [41][42]); separate paper-fill simulation from concession policy, sub-100% mid-fill probability on the simulated plane calibrated from the existing slippage telemetry (R3, [37][45]); trigger-persistence filter on the $2\times$ stop (R4, [60][62]); crossed/one-sided quote rejection + model sanity band on marks, escalate on prolonged mark silence (R5, [23][70]); aggregate daily cap + anti-splitting + periodic re-affirmation on the below-limit authorization (R6, [56]); approval-behavior (rubber-stamp) monitoring with attestation records (R7, [49][55]); startup crash-recovery sweep canceling unknown working orders + out-of-band runbook (R8, [14][16]); kill-switch and exit-bypass drills on a schedule (R9, [21][22b][58]).

Where the lifecycle is ahead of published practice: every-cycle verified reconciliation with per-symbol blocking and no-fabricated-P&L force close (§6.6); the intent-hash idempotency composite; the safe-state asymmetry set (subtract-only ML gate, PANIC demotion, evidence halts, no unattended price changes on human orders); the honest simulated plane (refuse rather than fabricate a fill); per-plane slippage telemetry separating sim optimism from live fill quality; day-scoped intents; and the

three-layer typed production confirmation with a bounded standing-consent carve-out that matches delegation-of-authority practice.

8. A Note on Regulatory Posture

None of the cited regulations bind a retail customer trading through a broker — 15c3-5 obligations sit with the broker, RTS 6 with EU investment firms. They are used here because they are the distilled, incident-tested consensus on what automated trading systems need, and because the agent's brokers' own controls (the second layer the Reg-AT-era "layered controls" doctrine assumes) should never be the *only* price/size backstop. The design pattern the whole canon converges on — layered, automated, fail-closed pre-trade controls beneath an accountable human — is the pattern the lifecycle already follows; this review's findings are about completing the layer, not changing the philosophy.

9. References

Numbering continuous; grouped by research area. URLs verified August 2026. Practitioner sources labeled (P).

Regulatory & industry risk-control standards

- [1] SEC, 17 CFR 240.15c3-5 — Risk Management Controls for Brokers or Dealers with Market Access; adopting release No. 34-63241 (2010), 75 FR 69792.
- [2] FINRA, Annual Regulatory Oversight Report — Market Access Rule findings (2026).
- [3] SEC Division of Trading and Markets, Responses to FAQs Concerning Rule 15c3-5.
- [5] Commission Delegated Regulation (EU) 2017/589 (MiFID II RTS 6): Arts. 5–9 (testing, deployment, self-assessment), 12 (kill functionality), 13 (surveillance), 15 (pre-trade controls), 16 (real-time monitoring), 17 (post-trade reconciliation).
- [8] FINRA, Regulatory Notice 15-09, "Supervision and Control Practices for Algorithmic Trading Strategies" (2015).
- [10] CFTC, Regulation Automated Trading NPRM, 80 FR 78824 (2015); withdrawn 2020, replaced by Electronic Trading Risk Principles, 17 CFR 38.251(e)–(g) (2020).
- [14] FIA, "Guide to the Development and Operation of Automated Trading Systems" (2015).
- [16] FIA, "Best Practices for Automated Trading Risk Controls and System Safeguards" (2024).
- [17] FIA PTG/EPTA, "Software Development and Change Management Recommendations" (2012).
- [18] IOSCO, "Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency," FR09/11 (2011); FR31/2015 on volatility mechanisms.
- [20] SEC, In re Knight Capital Americas LLC, Admin. Proc. 3-15570, Release 34-70694 (2013).
- [21] FCA, "Algorithmic Trading Compliance in Wholesale Markets" (2018); 2024 follow-up review.
- [22] CFTC & SEC Staffs, "Findings Regarding the Market Events of May 6, 2010" (2010).
- [22b] Central Bank of Ireland, Thematic Review of Algorithmic Trading (kill-switch testing deficiencies).
- [24] Clark, C. (Chicago Fed), "How Do Proprietary Trading Firms Control the Risks of High Speed Trading?" PDP 2012-1; Clark & Ranjan, PDP 2012-3.
- [26] FIX Trading Community, FIX 4.4 Appendix D — Order State Change Matrices; CIOrdID/PossDupFlag dedup semantics.
- [28] FINRA, Limit Up-Limit Down Plan; Rules 11890–11893 (clearly erroneous transactions).
- [29] Cboe, LULD FAQ; NYSE/Nasdaq clearly-erroneous guidance.
- [30] Nasdaq Options Obvious Error Policy; Cboe Rule 6.5 lineage (minimum-deviation tables \$0.25–\$2.00 by premium tier).

Order-lifecycle engineering

- [32] FIX specification: ExecutionReport (35=8) state machine, Order Status Request (35=H), Order Cancel/Replace (35=G), Multileg Cancel/Replace (35=AC), CxlRejReason "too late to cancel."
- [33] Deribit, "Order Management Best Practices" — cancel+fill same-order trap; replace-vs-cancel race semantics. (P)
- [34] CME Group, Drop Copy FAQ; FIA Drop Copy Recommendations (2013).

- [35] Hasbrouck, J. & Saar, G. (2009). "Technology and Liquidity Provision." *Journal of Financial Markets* 12(2) (fleeting orders: 27.7% canceled within 2 s).
- [36] Broker API architecture guidance: FIX vs REST/WebSocket for order events. (P)
- [37] Lo, A.W., MacKinlay, A.C. & Zhang, J. (2002). "Econometric Models of Limit-Order Executions." *JFE* 65(1), 31–71.
- [38] NYSE, Client Notice — Removal of Stop and GTC Orders (2015, effective Feb 2016).
- [39] Fong, K. & Liu, W.-M. (2010). "Limit Order Revisions." *Journal of Banking & Finance* 34(8), 1873–1885.
- [40] Stripe, "Designing Robust and Predictable APIs with Idempotency"; IETF Idempotency-Key header draft.
- [41] Avellaneda, M. & Stoikov, S. (2008). "High-Frequency Trading in a Limit Order Book." *Quantitative Finance* 8(3), 217–224.
- [42] Guéant, O., Lehalle, C.-A. & Fernandez-Tapia, J. (2012/2013). "Optimal Portfolio Liquidation with Limit Orders." *SIAM J. Financial Math.* 3(1); "Dealing with the Inventory Risk." *Math. Financial Economics* 7(4).
- [43] Almgren, R. & Chriss, N. (2000). "Optimal Execution of Portfolio Transactions." *Journal of Risk* 3, 5–39.
- [44] Huang, W., Lehalle, C.-A. & Rosenbaum, M. (2015). "Simulating and Analyzing Order Book Data: The Queue-Reactive Model." *JASA* 110(509), 216–229.
- [45] Interactive Brokers, Paper Trading limitations (simulated fills optimistic vs live); TradingView/NautilusTrader fill-rule documentation. (P)

Human-in-the-loop automation

- [46] Sheridan, T.B. & Verplank, W.L. (1978). *Human and Computer Control of Undersea Teleoperators*. MIT.
- [47] Parasuraman, R., Sheridan, T.B. & Wickens, C.D. (2000). "A Model for Types and Levels of Human Interaction with Automation." *IEEE Trans. SMC-A* 30(3), 286–297; Endsley, M.R. & Kaber, D.B. (1999). *Ergonomics* 42(3).
- [48] Strom, B.L. et al. (2010). Hard-stop alert RCT. *Arch Intern Med* 170(17); Powers, E.M. et al. (2018). Hard-stop systematic review. *JAMIA* 25(11).
- [49] Parasuraman, R. & Riley, V. (1997). "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39(2); Parasuraman, R. & Manzey, D.H. (2010). "Complacency and Bias in Human Use of Automation." *Human Factors* 52(3).
- [50] Bainbridge, L. (1983). "Ironies of Automation." *Automatica* 19(6).
- [51] ANSI/ISA-18.2-2016, Management of Alarm Systems for the Process Industries; EEMUA 191.
- [52] 14 CFR §25.1322 + FAA AC 25.1322-1 (graded alerting, prioritization, inhibits); Cvach, M. (2012). Alarm fatigue review. *Biomed Instrum Technol* 46(4).
- [53] Lee, J.D. & See, K.A. (2004). "Trust in Automation: Designing for Appropriate Reliance." *Human Factors* 46(1); Sorkin, R.D. et al. (1988). "Likelihood Alarm Displays." *Human Factors* 30(4).
- [54] IOSCO (2025). *Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges*, CR/01/2025.
- [55] FCA, AI Update / "AI and the FCA" (SM&CR accountability); UK Treasury Committee report on AI in financial services.
- [56] Delegation-of-authority matrix practice: threshold tiers, aggregate caps, anti-splitting, periodic re-affirmation (treasury/AP control literature). (P)
- [57] Shadow/canary/champion-challenger deployment for ML and trading models (MLOps practice; forward-gated model replacement literature). (P)
- [58] IEC 61508/61511 functional safety: safe-state concept, safe-vs-dangerous failure asymmetry, fail-safe vs fail-operational design.

Exit management & operational risk

- [60] Kaminski, K. & Lo, A.W. (2014). "When Do Stop-Loss Rules Stop Losses?" *Journal of Financial Markets*; Han, Y., Zhou, G. & Zhu, Y. "Taming Momentum Crashes."
- [61] spintwig short-put/strangle stop-loss-vs-profit-taking backtests; 7 Circles syntheses. (P)
- [62] projectfinance, 11-year short-strangle management study (touch frequencies; Feb-2018 gap-through of all stop variants). (P)
- [63] Prop-firm daily-loss/drawdown design conventions (1–2.2% daily static, 4–6% max drawdown, unrealized counted, session lockout). (P)
- [64] Black-Scholes ATM $\gamma \propto 1/(S\sigma\sqrt{T})$ — gamma escalation magnitudes by DTE (analytic).
- [66] FINRA, "Trading Options: Understanding Assignment"; OCC exercise statistics.
- [67] Pool, V.K., Stoll, H.R. & Whaley, R.E. (2008). "Failure to Exercise Call Options." *Journal of Financial Markets* 11(1) (\$491M abandoned; dividend-spread arbitrage population).

- [68] Barraclough, K. & Whaley, R.E. (2012). "Early Exercise of Put Options on Stocks." *Journal of Finance* 67(4) (\$1.9B forfeited interest; interest-condition boundary).
- [69] Jensen, M. & Pedersen, L.H. "Early Option Exercise: Never Say Never" (frictions advance the boundary).
- [70] OCC/OIC, "Mark-to-Market: Where Every Option Finds Its True Value" (binomial marks, explicitly not mids); Stoikov, S. (2018). "The Micro-Price." *Quantitative Finance* 18(12).
- [71] Nasdaq/InvestorPlace, automatic exercise and after-hours risk (contrary-exercise window to ~5:30 pm ET; weekend-gap assignment case); Schwab assignment-risk and expiration-checklist guidance. (P)
- [72] projectfinance/SteadyOptions, market orders in options (documented hazards). (P)
- [73] CAT NMS multi-leg/complex order documentation; complex-order-book trade-through exclusion. (P)

Prepared from AGENT_LIFECYCLE_SPEC.md v3.96.0 (HEAD, same-day) and ~120 primary sources across four research sweeps. Regulatory texts are used as design templates, not compliance obligations, for a retail system. Practitioner-grade evidence labeled (P). Research commentary, not investment or legal advice.