Neuro-Symbolic Agents for Regulated Process Automation: Challenges and Research Agenda

Alexander Rombach[1,2] and **Chantale Lauer**[1,2] and **Nijat Mehdiyev**[1,2]
[1] German Research Center for Artificial Intelligence (DFKI), Campus D3 2, 66123 Saarbrücken, Germany
[2] Saarland University, Campus D3 2, 66123 Saarbrücken, Germany
{alexander_michael.rombach, chantale.lauer, nijat.mehdiyev}@dfki.de

Abstract

LLM-based agents are entering regulated industries where they automate judgment intensive quality management processes. We argue that symbolic structures already embedded in these domains, including regulations, typed process models, and compliance constraints, should be treated not merely as external monitoring mechanisms but as core architectural components that shape the agent's decision-making and behavior. We propose compliance-by-construction as a complementary paradigm to guardrail-based monitoring: a structural foundation that prevents control-flow violations, while guardrails remain essential for catching semantic errors. We identify a structured set of neuro-symbolic research challenges on foundational and capability level and show that addressing them jointly enables compliance-by-construction. We call on the neuro-symbolic community to engage with regulated process automation as a high impact research domain.

1 Introduction

Neuro-symbolic AI has matured into a productive research paradigm. Garcez and Lamb [2023] position it as the "third wave" of AI, integrating robust neural learning with symbolic reasoning and explainability. Surveys document progress across planning, knowledge representation, and natural-language understanding [Hitzler *et al.*, 2022; Raedt *et al.*, 2020; Wan *et al.*, 2024], while a systematic review of 167 publications from 2020–2024 identifies learning/inference (63%) and knowledge representation (44%) as dominant themes, with explainability (28%) and metacognition (5%) significantly underexplored [Colelough and Regli, 2025]. Yet a persistent concern is the scarcity of applied domains where neuro-symbolic integration is not only beneficial but rather *necessary*, particularly where the stakes of failure are high enough to justify the additional architectural complexity.

We argue that regulated process automation is such a domain. In industries like pharmaceutical manufacturing, biotechnology, and medical devices, quality management (QM) processes are governed by dense regulatory frameworks, such as EU-GMP, ISO 13485, and the EU AI Act (Regulation (EU) 2024/1689), and executed through formally modeled workflows in electronic quality management systems (eQMS). These processes combine a rich symbolic scaffold of regulations, typed process models, and compliance constraints with activities that require contextual, evidence-based reasoning over unstructured documents, which LLMs excel at.

The standard approach to deploying LLM agents in such environments is to add guardrail schema validators that check agent outputs, governance monitors that detect drift, and human-in-the-loop (HITL) gates at critical decision points [Derouiche *et al.*, 2025]. Recent work on symbolic guardrails demonstrates their value: ShieldAgent [Chen *et al.*, 2025] extracts safety policies into probabilistic logic rule circuits for trajectory verification, while QuadSentinel [Yang *et al.*, 2025] addresses multi-agent safety via sequent-based supervisory control. However, they address compliance in a post-hoc manner: they monitor and catch violations but cannot structurally prevent them. In high-risk regulated processes where a detected-but-occurred violation is itself a regulatory finding, an additional architectural layer is needed to reduce the range of failures that guardrails must catch.

Our thesis. The symbolic structures already present in regulated industries should not serve only as an external monitoring layer but as *primary architectural components* that co-constitute the agent's reasoning. We argue for compliance-by-construction as a complementary paradigm: an architectural foundation that structurally prevents control-flow violations (wrong sequencing, missing approvals, skipped mandatory steps), while guardrails remain essential for catching semantic errors that no structural guarantee can eliminate. Realizing this layered approach surfaces a structured set of neuro-symbolic research challenges that we identify in this paper.

Our discussion draws on the consortium research project that develops multi-agent AI for pharmaceutical quality management. While we draw domain context, the challenges we identify are relevant to any scenario where autonomous agents must operate under formal regulatory constraints.

2 Why Regulated Process Automation Structurally Demands Neuro-Symbolic Integration

2.1 Domain Context: Pharmaceutical Quality Management

Consider deviation management, a core QM process governed by EU-GMP Annex 15 and ISO 13485 §8.5.2. When a nonconformity is detected in pharmaceutical manufacturing (e.g., a raw-material contamination during an incoming goods inspection, an out-of-specification test result, an equipment malfunction during batch production), a structured workflow is initiated. The deviation is captured and classified by severity; immediate containment actions are defined to prevent further impact; root-cause analysis (RCA) is performed using structured methods such as Ishikawa diagrams, 5-Whys, or 8D reports; corrective and preventive actions (CAPA) are proposed, approved by designated authorities, implemented, and verified for effectiveness; and the case is formally closed with a complete, tamper-proof audit trail.

This process is conceptually modeled in Business Process Model and Notation (BPMN) and executed through eQMS platforms. It involves multiple roles (quality manager, subject matter expert, CAPA owner, approver), multiple artifact types (deviation records, investigation reports, CAPA plans, effectiveness reviews), and multiple decision points where regulatory constraints must be satisfied. Critically, separation-of-duties requirements mandate that the person who investigates a deviation cannot be the same person who approves the resulting CAPA.

The economic stakes are substantial. Industry estimates place the cost of poor quality in the pharmaceutical sector at 25–40% of operating expenses [BSI Group, 2021; Bruttin and Dean, 2004], significantly exceeding comparable quality-driven industries such as semiconductors (4–8%). The FDA completed 2,953 CGMP drug inspections worldwide in FY2023, with warning letter frequency rising 43% per 100 inspections between 2019–2023 [Auria Consulting Group, 2022]. A single compliance failure can trigger warning letters, consent decrees, or product recalls with direct patient-safety consequences.

2.2 Three Co-Present Properties

What makes this domain structurally distinctive, and therefore distinctively suited to neuro-symbolic research, is the co-presence of three properties that rarely coincide:

Pre-existing symbolic structure. Regulations, standard operating procedures (SOPs), BPMN models, controlled vocabularies, and typed document schemas constitute a symbolic knowledge base that is already maintained, versioned, and audited within the organization. Unlike many AI domains where symbolic structure must be discovered or learned from data, in this context, it already exists as an organizational artifact. The challenge is to make it computationally active within an agent architecture.

Demanding neural tasks. Each process activity requires reasoning no rule engine can perform: reading a supplier's corrective-action response letter and judging its adequacy against the identified root cause, synthesizing a root-cause hypothesis from a combination of laboratory reports, three prior deviation records involving the same raw material, and an audit finding from the previous quarter, or assessing whether a proposed CAPA addresses the systemic cause rather than the immediate symptom. These require natural-language understanding, multi-document reasoning, and domain judgment, which are precisely the areas where LLMs can provide support.

Hard verification requirements with legal consequences. Compliance properties are binary requirements with regulatory force. The EU AI Act, with high-risk obligations taking binding effect on August 2, 2026, adds further layers: Article 9 mandates risk management systems throughout the AI lifecycle, Article 14 requires human oversight mechanisms that are *technically embedded* in the system (not merely procedural), and Article 17 demands documented quality management procedures [European Union, 2024]. For AI systems used as safety components in medical devices, third-party conformity assessment is required under Annex II.

By contrast, most LLM agent deployment domains (e.g., web navigation, open-ended task completion, conversational assistance) offer demanding neural tasks but lack preexisting symbolic structure and hard verification requirements. This makes regulated process automation a uniquely demanding testbed for neuro-symbolic integration where the research challenge is not discovering symbolic structure but *integrating* it with neural capabilities under formal guarantees.

2.3 Why Classical BPM Is Insufficient

An objection arises naturally: classical business process management (BPM) solved process compliance decades ago through workflow engines and conformance checking [van der Aalst, 2016]. What is new?

The answer is that classical BPM assumes *human workers* or *deterministic scripts* at activity nodes. The process engine governs control flow; the executor, whether human or scripted, is either trusted (human judgment, verified post-hoc) or trivially verifiable (script output matches specification). The introduction of LLM agents as activity executors reopens the compliance problem in a new way, because the executor is now probabilistic, opaque, and capable of generating outputs that are syntactically valid but semantically wrong. An LLM might produce a well-formatted severity classification that is factually incorrect, or a grammatically perfect CAPA proposal that fails to address the root cause. The process engine can verify that the activity *occurred* and that its output matches a type schema, but it cannot verify semantic adequacy, nor can a post hoc guardrail do so reliably beyond probabilistic estimation.

This creates challenges that neither BPM alone nor LLM-agent research alone can address. They sit precisely at the neuro-symbolic boundary: the symbolic process structure must constrain the neural executor, the neural executor must provide calibrated signals such that the symbolic layer can reason about its trustworthiness. The interface between them must enforce contracts rich enough to catch meaningful errors but tractable enough to be verified automatically.

3 Research Challenges

We organize the challenges into two tiers (Figure 1). The **foundational** tier (§3.1–3.2) defines the architectural base: how regulatory knowledge becomes executable and how agents are bound to process structure. The **capability** tier (§3.3–3.5) builds on this base, addressing calibrated trust, evolving knowledge, and auditable explanations. This hierarchy reflects a dependency structure: the capability-tier challenges presuppose that the foundational tier has been at least partially addressed.

3.1 Foundational: Regulatory Knowledge Operationalization

The problem. Regulatory frameworks express requirements in natural language with domain-specific terminology, implicit preconditions, and context-dependent interpretation. EU-GMP Chapter 1 requires that deviations be investigated to determine the root cause. ISO 13485 §8.5.2 mandates correction to eliminate detected nonconformities. For these to be enforceable at runtime in an agentic system, they must be translated into formal, executable constraints (e.g., temporal properties, state invariants, transition guards) that a process engine can enforce and a model checker can verify.

Why it is neuro-symbolic. An LLM can read, interpret, and contextualize regulatory text with reasonable accuracy, but producing a formal constraint with precise operational semantics requires a fundamentally different representational commitment. The challenge is compounded in multi-framework regulatory environments where constraints from EU-GMP, ISO standards, and the EU AI Act interact and must be jointly satisfied—a constraint composition problem with no existing solution.

Current state and gaps. Work on legal requirements translation demonstrates the feasibility of extracting formal representations from regulatory text [Singhal and Breaux, 2025]. ShieldAgent [Chen *et al.*, 2025] demonstrates LLM-assisted policy formalization into probabilistic logic rules for agent monitoring, showing that the neuro-to-symbolic pipeline is viable for safety policies. The broader literature on neuro-symbolic knowledge extraction from text [Hitzler *et al.*, 2022; Raedt *et al.*, 2020] provides foundational methods. However, none of these works address the specific challenge of embedding extracted constraints into *executable process models* within *multi-framework regulatory environments* with *formal verifiability* requirements. The gap is not extraction alone but operationalization: producing constraints that are simultaneously faithful to regulatory intent, formally verifiable, and executable at runtime within a process engine.

Research direction. Hybrid pipelines where LLMs generate candidate formalizations from regulatory text and symbolic methods verify their consistency, completeness, and faithfulness to regulatory intent. The proposed project's approach of translating regulations into machine-checkable policies motivates this direction; the systematic methodology for reliably achieving it remains an open research problem.

3.2 Foundational: Grounding Agents in Symbolic Process Structure

The problem. In current agentic architectures (ReAct, AutoGen, CrewAI), process knowledge enters the system through the prompt: the agent is told about the workflow in natural language and expected to follow it. Yao *et al.* [2023] demonstrated the power of interleaving reasoning and acting, but the reasoning is unconstrained, i.e., the LLM decides what to do next based on its interpretation of instructions. In regulated processes, this is structurally insufficient: the process model, not the LLM, must determine the action space.

The neuro-symbolic principle. The process model should function as a formal structure that *constrains* the agent's action space: at each process state, the set of enabled activities is determined by the model, not by the LLM's prompt interpretation. The LLM's role is to ground each enabled activity by interpreting unstructured case artifacts and producing typed outputs that satisfy the activity's contract. The symbolic layer determines *what* to do next; the neural layer determines *how* to do it.

Neuro-symbolic integration is centered around a layer between neural and symbolic systems where symbolic process state meets neural execution. Consider a concrete example: when the deviation management process enables the activity `CLASSIFY_SEVERITY`, the mediation layer constructs a prompt containing the deviation description, relevant SOP excerpts from the process data store, and historical severity distributions retrieved from a Case Graph. It specifies the output schema (`MINOR, MAJOR, CRITICAL`) and the constraint that `CRITICAL` triggers mandatory regulatory notification. The LLM produces its classification; the mediation layer validates the output against the type and checks the notification invariant, and either accepts the result (advancing the process state) or triggers re-prompting with a diagnostic that identifies the specific constraint violation. The process model never enters a state that the LLM chose, but rather only states the model prescribed states.

Current state and gaps. Neuro-symbolic planning research demonstrates the value of binding LLM reasoning to symbolic action specifications. Tantakoun *et al.* [2025] survey the landscape of LLMs as planning formalizers, documenting extensive work on LLM+PDDL integration. Capitanelli and Mastrogiovanni [2024] propose a neurosymbolic (NS) robot planning framework using LLMs with PDDL. Kwon *et al.* [2025] combine LLM-based goal decomposition with symbolic planners for faster, more accurate task planning. However, these approaches target single-agent, goal-directed plan synthesis. Regulated process automation differs in three respects: processes are long-running rather than one-shot, concurrent and multi-party rather than single-agent, and governed by normative constraints rather than goal states. The extension from plan synthesis to ongoing process execution under regulatory constraints is a significant generalization that the community has not yet undertaken.

Research direction. Designing the mediation interface: how typed activity specifications are translated into effective prompts, and how LLM outputs are validated against symbolic contracts. The key open problem is the expressiveness-

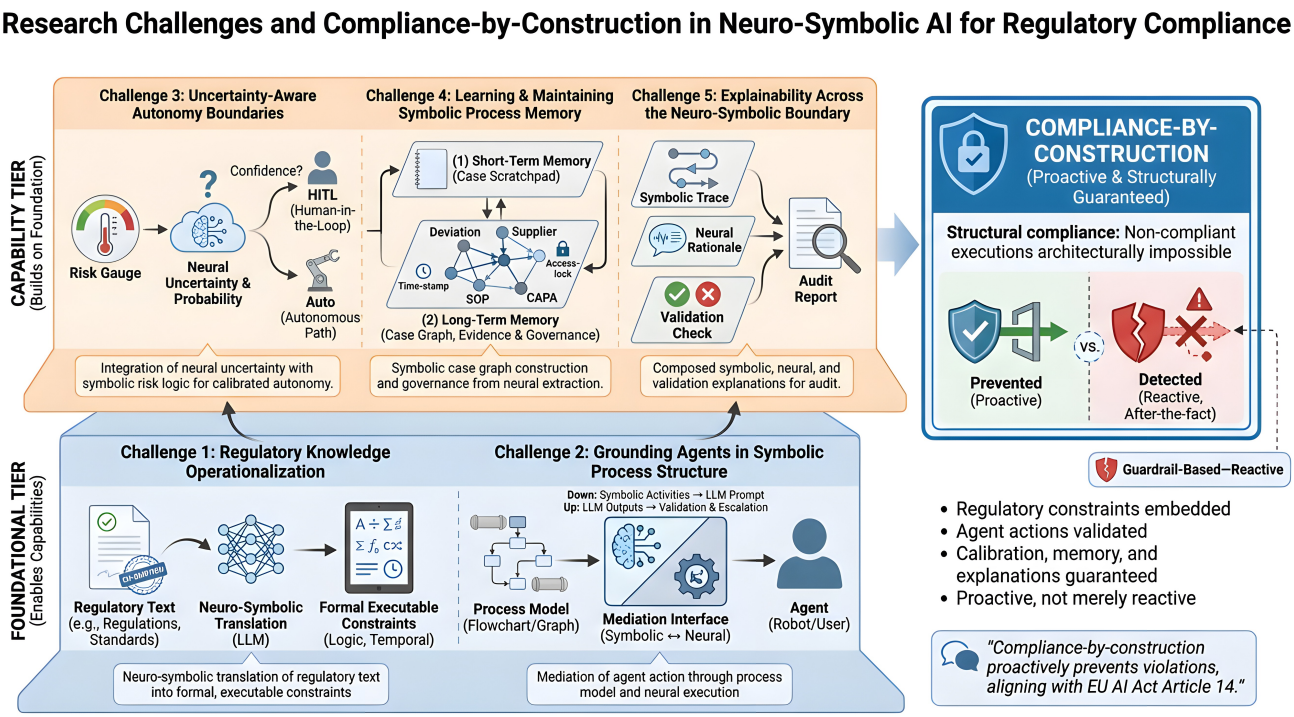

Figure 1: Two-tier challenge architecture. The *foundational tier* (Challenge 1: regulatory operationalization; Challenge 2: symbolic process grounding with mediation interface) defines the structural base. The *capability tier* (Challenge 3: uncertainty-aware autonomy; Challenge 4: symbolic process memory; Challenge 5: cross-boundary explainability) extends it. Compliance-by-construction emerges as a property of addressing both tiers jointly by proactively preventing violations rather than detecting them post-hoc.

verifiability tradeoff: trivial type checks are verifiable but catch little, rich semantic postconditions (e.g., "the CAPA addresses the root cause") are useful but unverifiable by formal methods alone. Identifying the middle ground, perhaps through calibrated confidence scores, lightweight semantic validators, or structured output schemas with partial formal coverage, is a concrete research challenge. We note that the broader perspective of process models as agent world models, with a full formal correspondence between process specifications and world-model components, is a direction we develop separately.

3.3 Capability: Uncertainty-Aware Autonomy Boundaries

The problem. Not every process activity warrants the same level of human oversight. Drafting an initial deviation summary from a structured incident report is low-risk and fully automated. Confirming a severity classification that triggers mandatory regulatory reporting to national competent authorities is high-risk and demands human approval. Between these extremes lies a continuum in which the appropriate autonomy level depends on case-specific factors: the novelty of the deviation type, the completeness and quality of the available evidence, and the confidence in the LLM's analysis. The EU AI Act's Article 14 explicitly requires human oversight mechanisms "proportionate to the risks" making calibrated autonomy a legal requirement [European Union, 2024].

Why it is neuro-symbolic. The risk framework that defines HITL gates and autonomy tiers (e.g., classifying activities by regulatory risk) constitutes symbolic knowledge. The confidence estimate is a neural signal. Adaptive autonomy requires the symbolic layer to reason about neural uncertainty: if the LLM's severity classification carries high uncertainty (a deviation underrepresented type), the symbolic layer should escalate to HITL regardless of the default risk tier.

Current state and gaps. Conformal prediction (CP) has emerged as a principled, distribution-free framework for uncertainty quantification in LLM systems [Angelopoulos and Bates, 2022]. Ren *et al.* [2023] apply CP to calibrate LLM planner confidence for robot task planning, triggering human assistance when prediction sets are non-singleton, thereby directly demonstrating the viability of UQ-driven autonomy control for LLM agents. Quach *et al.* [2023] extend CP to open-ended language generation. In the process domain specifically, prior work has developed uncertainty-aware predictive process monitoring using CP as well as calibrated execution of enterprise tasks such as business document processing [Mehdiyev *et al.*, 2023, 2025a,b,c; Majlatow *et al.*, 2025; Rombach and Mehdiyev, 2026]. This establishes methodological foundations for calibrated uncertainty in process-aware systems.

The gap lies in integration: extending these methods from offline process prediction to real-time agent autonomy control. Here, calibrated uncertainty estimates feed into symbolic

risk-classification logic that governs HITL escalation during live process execution. The goal is not necessarily that the agent is always correct, but that its autonomy is always calibrated to its demonstrated competence.

Research direction. Integrating calibrated neural uncertainty (via CP, ensemble disagreement, or epistemic uncertainty decomposition) with symbolic autonomy-control logic in real-time process execution. Such integration allows for the verification of learning-based agents.

3.4 Capability: Learning and Maintaining Symbolic Process Memory

The problem. Over time, an organization accumulates a rich history of deviation cases: which raw materials were involved, which suppliers, what root causes were identified, which corrective actions proved effective, which ones failed and had to be revised. This history is invaluable for handling new cases and tasks such as detecting recurring patterns, identifying systemic supplier issues, and calibrating severity assessments against precedent. However, this knowledge exists primarily in unstructured form: investigation reports written in natural language, email threads with suppliers, free-text fields in deviation records.

Why it is neuro-symbolic. The proposed project models this as a dual-memory architecture: a Short-Term Memory (STM) serving as a case-specific scratchpad for current hypotheses and intermediate evidence, and a Long-Term Memory (LTM) structured as a Case Graph, which represents a knowledge graph (KG) with typed entities (deviation, batch, supplier, SOP, CAPA) connected by domain relations (*caused_by*, *prevented_by*, *linked_to*). Building and maintaining this graph is a paradigmatic instance of learning symbolic abstractions from sub-symbolic data.

The raw inputs are unstructured. Extracting typed entities and relations requires neural processing (named entity recognition, relation extraction, coreference resolution). The resulting graph must be symbolically consistent, meaning that entities must conform to the domain ontology, relations must satisfy type constraints, and the graph must support formal queries that agents rely on at runtime. Moreover, regulated environments impose governance requirements absent from typical KG construction: provenance tracking (which source document did this entity come from?), temporal validity (is this SOP reference still current or has it been superseded?), and access control (which agents can modify which portions of the graph?).

Current state and gaps. Neuro-symbolic KG construction where neural extraction is guided by symbolic ontologies has shown promise [Hitzler *et al.*, 2024]. However, the requirements of regulated environments (e.g., provenance, governance, and the feedback loop where agents query the graph they help build) are largely unaddressed. The Case Graph is not static, but rather an active resource where the root-cause agent searches for precedents, the CAPA agent retrieves prior actions and outcomes, and each agent's outputs feed back in, creating a neuro-symbolic co-learning dynamic.

Research direction. Neuro-symbolic architectures for KG lifecycle management that jointly satisfy extraction accuracy (neural), ontological consistency (symbolic), provenance (regulatory), and runtime queryability (agentic). The feedback loop between agents and the Case Graph is a compelling instance of neuro-symbolic co-learning.

3.5 Capability: Explainability Across the Neuro-Symbolic Boundary

The problem. When a process decision is audited, the explanation must span both layers. The auditor needs to know: (a) *Why was this activity executed at this point?*—a symbolic question answered by the process model's state and transition logic; (b) *What evidence did the agent consider and how did it reach its judgment?*—a neural question answered by the LLM's reasoning trace and evidence attributions; and (c) *Were the formal constraints satisfied?*—a neuro-symbolic question answered by the mediation layer's validation record. No single explainability technique covers all three.

The specific technical challenge. A symbolic process trace is faithful by construction. An LLM rationale is a post-hoc reconstruction that may not faithfully represent the model's internal computation [Turpin *et al.*, 2023]. A cross-boundary framework must reconcile these fidelity levels by indicating which parts are structurally guaranteed and which are best-effort neural attributions. Colelough and Regli [2025] find explainability in only 28% of NS publications. Prior work on XAI for process mining [Mehdiyev and Fettke, 2021; Mehdiyev *et al.*, 2024] provides a starting point, but extension to real-time multi-layer agent explanations with differentiated fidelity guarantees is open.

Research direction. Layered explanation architectures that compose symbolic process traces, neural evidence attributions, and contract validation records into audit-ready artifacts with explicit fidelity annotations. Such frameworks would directly address the EU AI Act's transparency requirements (Article 13) for high-risk systems.

4 Compliance-by-Construction

Addressing the foundational and capability challenges jointly yields a structural property we call *compliance-by-construction*: an architectural foundation that, by design, makes structural control-flow violations (wrong sequencing, missing approvals, skipped mandatory steps) impossible. If regulatory requirements are operationalized as process-model invariants (§3.1), and the agent's action space is structurally bound to the process state with validated typed contracts (§3.2), every transition is sanctioned by the model. Calibrated autonomy (§3.3) ensures human oversight at the right moments; governed process memory (§3.4) ensures trustworthy knowledge; cross-boundary explainability (§3.5) ensures auditability.

Complementarity with guardrails. Compliance-by-construction and guardrails address different failure modes and are complementary. The structural layer prevents deterministic, fully preventable control-flow violations through architectural means. Guardrails (including sophisticated

symbolic approaches like ShieldAgent [Chen *et al.*, 2025] and QuadSentinel [Yang *et al.*, 2025]) remain essential for catching semantic errors: an LLM misjudging severity, a poorly reasoned CAPA, or an unforeseen edge case that falls outside the contract specification. Together, they provide defense in depth: the architecture handles what can be guaranteed, and guardrails handle what can only be monitored. The EU AI Act's requirement that human oversight be designed to "prevent or minimize the risks" (Article 14(2)) supports this layered approach [European Union, 2024].

Compliance-by-construction guarantees *structural* compliance, not semantic correctness of individual judgments. But it ensures every judgment is made in the right context, validated against formal constraints, documented with provenance, and when confidence is insufficient, routed to human expertise. Guardrails then provide the additional monitoring layer that catches what structural guarantees cannot.

5 Timeliness

Three developments make this agenda urgent. First, **agentic AI is entering regulated production now**. Projects are being built today, and their architectural decisions (guardrails alone vs. guardrails plus structural compliance, prompt-based control vs. formal grounding) will determine trustworthiness. Second, **the EU AI Act creates legal demand**: high-risk obligations (Articles 9–17) take binding effect on August 2, 2026 [European Union, 2024], imposing risk management, technically embedded human oversight, and documented quality management on AI systems in pharmaceutical and medical device manufacturing. Neuro-symbolic architectures with formal guarantees are becoming legally necessary. Third, **foundational components exist but are not integrated**: neuro-symbolic planning [Tantakoun *et al.*, 2025], CP for LLM calibration [Ren *et al.*, 2023; Angelopoulos and Bates, 2022], process mining with conformance checking [van der Aalst, 2016], uncertainty quantification for process monitoring [Mehdiyev *et al.*, 2025c], regulatory NLP [Singhal and Breaux, 2025], symbolic guardrails [Chen *et al.*, 2025], and KG construction [Hitzler *et al.*, 2024] are all mature. But the integrative program connecting them is missing.

6 Discussion and Limitations

Our proposal has several limitations. First, it assumes formal process models exist and are maintained as authoritative artifacts. This holds for GMP-regulated manufacturing, not universally. For processes without preexisting models, a process discovery step would precede our foundational tier, introducing its own neuro-symbolic challenges [van der Aalst, 2016].

Second, the added architectural complexity may not be justified for low-risk processes where guardrail-based approaches suffice. Compliance-by-construction is most valuable where the stakes are highest (e.g., patient safety, regulatory reporting).

Third, the expressiveness-verifiability tradeoff (§3.2) remains genuinely unsolved. The most useful postconditions ("the CAPA addresses the root cause") resist formal verification because they require semantic judgment. The best we can offer architecturally is ensuring semantic judgments occur in the right structural context and escalate to humans when neural confidence is insufficient.

Finally, our challenge framing is deliberately at the research-agenda level, where we identify *what* needs to be solved, not *how*. This is intentional for a position paper aimed at catalyzing community engagement, but the practical feasibility of compliance-by-construction remains to be demonstrated through future work.

7 Related Work

Neuro-symbolic AI foundations. Garcez and Lamb [2023] position NS AI as the third wave, emphasizing the integration of learning and reasoning for trust, safety, and explainability. Hitzler *et al.* [2022] survey neuro-symbolic approaches with a focus on knowledge representation. Raedt *et al.* [2020] provide a comprehensive survey bridging statistical relational and NS AI. Our proposal applies this paradigm to a domain in which its core themes are exercised under real-world constraints.

LLM-based agent frameworks and symbolic guardrails. ReAct [Yao *et al.*, 2023], AutoGen [Wu *et al.*, 2023], and CrewAI represent the prevailing LLM-as-orchestrator paradigm, providing flexible tool use and multi-agent coordination but no formal compliance guarantees. ShieldAgent [Chen *et al.*, 2025] extracts safety policies into probabilistic logic rule circuits for action trajectory verification. GuardAgent [Xiang *et al.*, 2025] synthesizes guardrail code via LLM-driven planning. QuadSentinel [Yang *et al.*, 2025] addresses multi-agent safety via sequent-based supervisory control. These provide the monitoring layer that remains essential even under compliance-by-construction. Our contribution is the structural foundation that reduces the surface area guardrails must cover.

Neuro-symbolic planning. Tantakoun *et al.* [2025] survey LLMs as planning formalizers, documenting extensive work on LLM+PDDL integration. Capitanelli and Mastrogiovanni [2024] and Kwon *et al.* [2025] demonstrate concrete neurosymbolic planning architectures. These target single-agent, goal-directed plan synthesis, which extends the symbolic-grounding principle from planning to long-running, multi-party process execution under normative constraints is the generalization we advocate.

LLMs for BPM and trustworthy process AI. LLMs have been applied to process modeling, task classification, and conversational process modeling [Grohs *et al.*, 2024; Lauer *et al.*, 2025]. Agentic BPM emphasizes governance [Vu *et al.*, 2025; Kirchdorfer *et al.*, 2025], and AI-assisted process mining enables context-sensitive analysis [Brützke *et al.*, 2025]. Pfeiffer *et al.* [2025] demonstrate trustworthy LLM-driven process automation. Pery *et al.* [2022] identify transparency and conformance as prerequisites for trustworthy process mining. Prior work on XAI and UQ for process monitoring [Mehdiyev and Fettke, 2021; Mehdiyev *et al.*, 2023, 2024, 2025a,b,c; Majlatow *et al.*, 2025] provides methodological foundations for Challenge 3. These address individual capabilities; we advocate the integrated architecture connecting them.

8 Conclusion

We have argued that regulated process automation structurally demands neuro-symbolic integration as an architectural necessity. The thesis that symbolic structures should co-constitute the agent's reasoning infrastructure leads to a structured research agenda: foundational challenges of regulatory operationalization and process grounding, supported by capability challenges of calibrated autonomy, symbolic memory, and cross-boundary explainability. Addressing them jointly yields compliance-by-construction, a property that is both academically significant and, with the EU AI Act entering enforcement, increasingly legally required. The neuro-symbolic community has built the foundations for exactly this kind of integration. Regulated process automation is a domain where those foundations are urgently needed and directly consequential. We invite the community to engage.

References

Anastasios N. Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*, 2022.

Auria Consulting Group. Calculating the true cost of quality: A GMP perspective for manufacturing professionals. ACG Technical Report, 2022.

Frances Bruttin and Doug Dean. Managing the cost of compliance in pharmaceutical manufacturing operations. *Pharmaceutical Engineering*, 24(6), 2004.

Paul Brützke, Robin Killewald, Sandro Franzoi, and Jan vom Brocke. Ai-assisted process mining for context-sensitive analysis support, 2025.

BSI Group. Cost of quality in the pharmaceutical sector. BSI White Paper, 2021.

Alessio Capitanelli and Fulvio Mastrogiovanni. A framework for neurosymbolic robot action planning using large language models. *Frontiers in Neurorobotics*, 18:1342786, 2024.

Zhaorun Chen, Mintong Kang, and Bo Li. ShieldAgent: Shielding agents via verifiable safety policy reasoning. *arXiv preprint arXiv:2503.22738*, 2025.

Brandon C. Colelough and William Regli. Neuro-symbolic AI in 2024: A systematic review. *arXiv preprint arXiv:2501.05435*, 2025.

Hana Derouiche, Zaki Brahmi, and Haithem Mazeni. Agentic AI frameworks: Architectures, protocols, and design challenges. *arXiv preprint arXiv:2508.10146*, 2025.

European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (artificial intelligence act). Official Journal of the European Union, 2024.

Artur d'Avila Garcez and Luís C. Lamb. Neurosymbolic AI: The 3rd wave. *Artificial Intelligence Review*, 56(11):12387–12406, 2023.

Michael Grohs, Luka Abb, Nourhan Elsayed, and Jana-Rebecca Rehse. Large language models can accomplish business process management tasks. In Jochen De Weerdt and Luise Pufahl, editors, *Business Process Management Workshops*, pages 453–465, Cham, 2024. Springer Nature Switzerland.

Pascal Hitzler, Aaron Eberhart, Monireh Ebrahimi, Md Kamruzzaman Sarker, and Lu Zhou. Neuro-symbolic approaches in artificial intelligence. *National Science Review*, 9(6):nwac035, 06 2022.

Pascal Hitzler, Monireh Ebrahimi, Md K. Sarker, and Daria Stepanova. Neuro-symbolic AI and the semantic web. *Semantic Web*, 11(1-3), 2024.

Lukas Kirchdorfer, Robert Blümel, Timotheus Kampik, Han van der Aa, and Heiner Stuckenschmidt. Discovering multi-agent systems for resource-centric business process simulation. *Process Science*, 2(1):4, 2025.

Minseo Kwon, Yaesol Kim, and Young J. Kim. Fast and accurate task planning using neuro-symbolic language models and multi-level goal decomposition. *arXiv preprint arXiv:2409.19250*, 2025.

Chantale Lauer, Peter Pfeiffer, Alexander Rombach, and Nijat Mehdiyev. Conversational business process modeling using llms: Initial results and challenges. In *EMISA 2025*, pages P3+2–O4+5. Gesellschaft für Informatik e.V., Bonn, 2025.

Maxim Majlatow, Fahim A. Shakil, Andreas Emrich, and Nijat Mehdiyev. Uncertainty-aware predictive process monitoring in healthcare. *Applied Sciences*, 15(14), 2025.

Nijat Mehdiyev and Peter Fettke. Explainable artificial intelligence for process mining: A general overview and application of a novel local explanation approach for predictive process monitoring. In Witold Pedrycz and Shyi-Ming Chen, editors, *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, pages 1–28. Springer International Publishing, Cham, 2021.

Nijat Mehdiyev, Maxim Majlatow, and Peter Fettke. Explainable artificial intelligence meets uncertainty quantification for predictive process monitoring. In *PMAI 2023 Workshop at IJCAI*, pages 29–32. CEUR-WS, 2023.

Nijat Mehdiyev, Maxim Majlatow, and Peter Fettke. Counterfactual explanations in the big picture: An approach for process prediction-driven job-shop scheduling optimization. *Cognitive Computation*, 16(5):2674–2700, 2024.

Nijat Mehdiyev, Maxim Majlatow, and Peter Fettke. Augmenting post-hoc explanations for predictive process monitoring with uncertainty quantification via conformalized monte carlo dropout. *Data & Knowledge Engineering*, 156:102402, 2025.

Nijat Mehdiyev, Maxim Majlatow, and Peter Fettke. Integrating permutation feature importance with conformal prediction for robust XAI. *Engineering Applications of Artificial Intelligence*, 149:110363, 2025.

Nijat Mehdiyev, Maxim Majlatow, and Peter Fettke. Quantifying and explaining machine learning uncertainty in predictive process monitoring: An operations research per-

spective. *Annals of Operations Research*, 347:991–1030, 2025.

Andrew Pery, Majid Rafiei, Michael Simon, and Wil M. P. van der Aalst. Trustworthy artificial intelligence and process mining: Challenges and opportunities. In Jorge Munoz-Gama and Xixi Lu, editors, *Process Mining Workshops*, pages 395–407, Cham, 2022. Springer International Publishing.

Peter Pfeiffer, Alexander Rombach, Maxim Majlatow, and Nijat Mehdiyev. From theory to practice: Real-world use cases on trustworthy LLM-driven process modeling, prediction and automation. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2025.

Victor Quach, Adam Fisch, Tal Schuster, Adam Yala, Jae Ho Sohn, T. Jaakkola, and Regina Barzilay. Conformal language modeling. volume abs/2306.10193, 2023.

Luc de Raedt, Sebastijan Dumančić, Robin Manhaeve, and Giuseppe Marra. From statistical relational to neuro-symbolic artificial intelligence. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 4943–4950. International Joint Conferences on Artificial Intelligence Organization, 7 2020. Survey track.

Allen Z. Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, Peng Xu, Leila Takayama, Fei Xia, Jake Varley, Zhenjia Xu, Dorsa Sadigh, Andy Zeng, and Anirudha Majumdar. Robots that ask for help: Uncertainty alignment for large language model planners. 2023.

Alexander Rombach and Nijat Mehdiyev. Beyond accuracy: Understanding model confidence in key information extraction with conformal prediction. *International Journal on Document Analysis and Recognition (IJDAR)*, 2026.

Anmol Singhal and Travis Breaux. Legal requirements translation from law. In *2025 IEEE 33rd International Requirements Engineering Conference (RE)*, pages 205–217, 2025.

Marcus Tantakoun, Christian Muise, and Xiao-Dan Zhu. Llms as planning formalizers: A survey for leveraging large language models to construct automated planning models. In *Annual Meeting of the Association for Computational Linguistics*, 2025.

Miles Turpin, Julian Michael, Ethan Perez, and Samuel R. Bowman. Language models don't always say what they think: unfaithful explanations in chain-of-thought prompting. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Wil M. P. van der Aalst. *Process Mining: Data Science in Action*. Springer Texts in Business and Economics. Springer, Berlin, Heidelberg, 2 edition, 2016.

Hoang Vu, Nataliia Klievtsova, Henrik Leopold, Stefanie Rinderle-Ma, and Timotheus Kampik. Agentic BPM: Practitioner perspectives on agent governance. In *BPM: Responsible BPM Forum*, pages 29–43. Springer, 2025.

Zishen Wan, Che-Kai Liu, Hanchen Yang, Chaojian Li, Haoran You, Yonggan Fu, Cheng Wan, Tushar Krishna, Yingyan Lin, and Arijit Raychowdhury. Towards cognitive AI systems: A survey and prospective on neuro-symbolic AI. *arXiv preprint arXiv:2401.01040*, 2024.

Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Beibin Li, Erkang Zhu, Li Jiang, Xiaoyun Zhang, Shaokun Zhang, Jiale Liu, Ahmed H. Awadallah, Ryen W. White, Doug Burger, and Chi Wang. AutoGen: Enabling next-gen LLM applications via multi-agent conversation. *arXiv preprint arXiv:2308.08155*, 2023.

Zhen Xiang, Linzhi Zheng, Yanjie Li, Junyuan Hong, Qinbin Li, Han Xie, Jiawei Zhang, Zidi Xiong, Chulin Xie, Carl Yang, Dawn Song, and Bo Li. Guardagent: Safeguard llm agents by a guard agent via knowledge-enabled reasoning. 2025.

Yiliu Yang, Yilei Jiang, Qunzhong Wang, Yingshui Tan, Xiaoyong Zhu, Sherman S. M. Chow, Bo Zheng, and Xiangyu Yue. QuadSentinel: Sequent safety for machine-checkable control in multi-agent systems. *arXiv preprint arXiv:2512.16279*, 2025.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2023.