AI Risk Evaluation Supplement version 6.0 – Comments Received

Table of Contents

	Page
01 KGAgentic_NAIC_AI_Risk_Evaluation_Supplement_v6_Comments	2
02 ReconAI_NAIC_AI_Risk_Evaluation_Supplement_v6_Comment	5
03 Legion - NAIC_v6_Comment_Letter_2026-10-20	8
04 SBJ_Comments_NAIC_AI_Risk_Evaluation_Supplement_v6.0	12
05 EasyEssence_NAIC_Comment_Letter_v6_2026-10-23	19
06 Falsify-OU-comments-AI-Risk-Evaluation-Supplement-v6.0	24
07 Comments on the AI Risk Evaluation Supplement v6.0 - Kaenan Hertz, Insurtech	26
08 Speck-Ward-NAIC-AI-Risk-Evaluation-Supplement-v6.0-Comment	28
09 Ali_Houshangi_NAIC_AI_Risk_Evaluation_Supplement_v6_Comments_2026-10-26	31
10 SecuShift - Comment on AI Risk Evaluation Supplement Version 6.0	32
11 ClearTaskAI, Gang Zhang	33
12 InsurTech Coalition - GLMs and the AI Risk Evaluation Supplement v6.0	34
13 ACLI Comment Letter AIRES v6.0 - 10.30.26	37
14a Verisk Comments on Version 6.0 of the AI Risk Evaluation Supplement	40
14b Verisk AI Risk Evaluation Supplement v6.0 (tracked changes)	42
15 APCIA Comments_AI Risk Evaluation Supplement version 6.0	69
16 NAIC_AI_Risk_Evaluation_Public_Comment_Dapp_Detroit_ARAI	72
17a 2026-1030_NAMIC Comments_NAIC BDAIWG ARES v.6	77
17b 2026-1030_NAMIC Redlines_NAIC BDAIWG ARES v.6	80
18 Execution-Rights-NAIC-AI-Supplement-v6.0-Comments-2026-10-24_1	107
19 BCBSA Comments_NAIC AI Risk Eval Supplement V6	122
20 Mebrat.Hill.Public_Comment_Submission	127

Krishcon Guide Agentic Solutions Inc.

operating as KGAgentic

Cambridge, Ontario, Canada | rakesh.sharma@krishcons.com | www.krishcons.com

October 20, 2026

Nathan Houdek, Chair
Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
Via email: Scott Sobel (ssobel@naic.org) and Miguel Romero (maromero@naic.org)

Re: Comments on the AI Risk Evaluation Supplement, Version 6.0 (exposed October 16, 2026)

Summary of requests: (1) restore a yes/no Exhibit A question on whether Agentic AI can act without prior approval; (2) scale oversight to autonomy and reversibility; (3) define meaningful human review; (4) add an optional Exhibit C item on decision-level records; and (5) extend Third-Party prompts to vendor-supplied and cross-vendor AI agents.

Dear Chair Houdek and Members of the Working Group:

Krishcon Guide Agentic Solutions Inc., operating as KGAgentic ("KGAgentic"), appreciates the opportunity to comment on version 6.0 of the AI Risk Evaluation Supplement (the "Supplement"). These comments are submitted on behalf of KGAgentic by its co-founders. We commend the Working Group and NAIC staff for the pilot-informed revisions and especially support three features of the draft: the clarification in the Intent section that the Supplement supplements the Handbooks and "does not independently create substantive legal duties"; the tiered approach in the Instructions, where Exhibit A is used to decide whether Exhibits B–D are needed; and the growing attention to Agentic AI, including the new stand-alone Agentic AI column in Exhibit A (column 6) and the new Exhibit C Ref 3a, which build on the Agentic AI definition and Exhibit B, Narrative 1g. We use the Supplement's defined terms throughout; "AI agent" means an AI System that meets the Supplement's definition of Agentic AI.

About our perspective. KGAgentic develops KGA ARIA, an Agentic AI orchestration layer for insurers' policy, billing and claims workflows that operates within each insurer's own authority matrix and is designed so that the AI does not bind, cancel, pay, post reserves or deny claims on its own. We are also developing ARIA Assure, a read-only, vendor-neutral oversight tool. We disclose this only to explain our perspective; none of the suggestions below depends on any particular product, vendor or technology.

Our comments share one theme. As insurers begin to use Agentic AI in operational workflows, examiners will need to tell three things apart: (i) whether an AI System *recommends* or *acts*; (ii) who held authority for a given action; and (iii) whether a qualified human actually approved it. Each suggestion is framed as an optional prompt within an existing exhibit. That keeps it consistent with the Supplement's regulator-customizable design and with the statement in Exhibit B that its questions are not intended to create new requirements. These prompts matter mainly where AI can act; consistent with the Supplement's materiality and risk-assessment guidance, regulators could scale them down or leave them out for recommend-only, low-materiality uses.

1. Distinguishing AI that recommends from AI that acts (Exhibit A, column 6; Exhibit C, Refs 3a and 9; Definitions)

Exhibit A's Agentic AI checkbox shows whether Agentic AI is present. It does not show whether an AI agent can act without prior human approval. That is an important factor in deciding whether Exhibits B–D are warranted, alongside materiality and Direct Consumer Impact, which remain relevant for consequential recommendation systems. The tracked-changes version shows the following question as deleted: "is that AI System authorized to execute one or more actions that can directly create or materially alter a consumer,

financial, or financial-reporting outcome without human approval before the action?" We recognize the value of keeping Exhibit A short. Even so, we encourage the Working Group to restore that question as a single yes/no column, which regulators could drop when customizing the exhibit.

- **Exhibit C, Ref 3a:** we suggest expanding "authorizations without human approval" to ask the company to identify the actions the system can execute (including through tools or APIs), which of those require human approval before execution, any actions executed without prior approval and the limits that apply (for example, transaction type or dollar thresholds), and who approves changes to those authorizations.
- **Definitions of "Augmentation" and "Automation" (and Exhibit C, Ref 9):** a single Agentic AI workflow may augment one step and automate the next. We suggest adding: "For Agentic AI, classification may differ by action within a workflow. An action executed before human approval, including one subject only to after-the-fact review or treated as approved when a reviewer does not respond, should be considered Automation."

2. Scaling oversight to autonomy and reversibility (Exhibit B, Narrative 1g and 7; Checklist)

We welcome Narrative 1g ("How does the insurance company assess autonomy, reversibility, and reporting impact risk of AI Systems?"). We suggest adding brief guidance to Narrative 1g along these lines: "Describe how the AIS Program matches the level of human oversight to the autonomy, reversibility and potential impact of the AI System's actions, including which actions (for example, binding, cancellation, claim payment, reserve changes or claim denial) require human approval before execution, whether the system operated in an observation-only ('shadow') mode before going live, and how the company can promptly suspend the system and revert to manual processing." This approach matches Sound Practice 10 of the Financial Stability Board's *Sound Practices for Responsible Adoption of Artificial Intelligence* (October 15, 2026), which Canada's Office of the Superintendent of Financial Institutions (OSFI) has endorsed. That practice calls for "appropriate and effective human oversight relevant to the materiality, risk, autonomy, complexity, and explainability of different AI use cases." Aligning with it would help insurers that operate across jurisdictions.

- **Narrative 7:** the tracked-changes version shows deletions addressing "incidents and near misses," "third-party changes," and "safe retirement or fallback." We suggest including these topics for Agentic AI, because for AI agents, Third-Party model or tool updates and a tested fallback are central controls.
- **Checklist:** the Checklist has no item matching Narrative 1g or Narrative 7. In line with version 6.0's effort to align the two versions of Exhibit B, we suggest adding matching Checklist items.

3. Ensuring meaningful human review (Exhibit B, Narrative 2b; Checklist 3l)

Narrative 2b, which already asks whether there is a human in the loop, now also asks "how the company assures itself the human review is appropriate." We support this addition. We suggest it list the elements regulators would expect a response to address: (a) the reviewer holds authority for the decision under the company's existing delegated authority (for example, underwriting or claims authority limits); (b) the reviewer receives enough information to evaluate the AI output, including key inputs and the stated basis for the recommendation; (c) approval is an affirmative act and is not inferred from silence, inaction or a timeout; (d) edits, overrides and rejections are recorded, with reasons where the action is material or higher-risk; (e) the company monitors for over-reliance on AI outputs; and (f) what happens when an applicable authority rule is missing or unclear, for example whether the workflow stops and escalates rather than proceeding. We suggest similar wording for Checklist 3l ("ensures human oversight is appropriate").

4. Retaining decision-level evidence (Exhibit C, Refs 3a, 10–11 and 13; Narrative 4d)

Exhibit C focuses on testing and validation at the model level. For Agentic AI, much of the risk of Adverse Consumer Impact and Material Financial Impact sits in individual actions within a workflow. We suggest adding an optional item to Exhibit C (for example, as Ref 3b) or a regulator instruction along these lines: "For Agentic AI use cases, describe the records retained for individual actions or recommendations, with the level of detail proportionate to the materiality and risk of the action. For material or higher-risk actions, include the inputs and output, the authority rule or limit applied, whether a human approved, modified or rejected

the action (and who and when), and the reason recorded for any override or rejection. Describe how these records are protected from alteration or deletion and how long they are retained. Regulators may request a sample of such records." This complements Narrative 4d (explainability for models "recommending or making decisions") and Ref 13 (compliance review). It would let examiners test whether the controls described in Exhibit B work in practice. We suggest keeping the wording technology-neutral; append-only or cryptographically chained logs could be listed only as examples of tamper-evident records.

5. Overseeing vendor-supplied and cross-vendor AI agents (Exhibit B, Narrative 4b, 5 and 7; Checklist 3m–3n; Exhibit C, Refs 5–6)

Agentic AI is increasingly built into vendor platforms, including core administration and claims systems, and a single workflow may involve AI agents from more than one Third Party. The Supplement already asks about AI Systems supplied by a Third Party (Narrative 4b and 5; Checklist 3m–3n). The prompts below build on the Third-Party AI Systems and Data section of the NAIC Model Bulletin on the Use of AI Systems by Insurers (Section 3, item 4.0), including its contract terms on audit rights and cooperation with regulatory inquiries. We suggest Narrative 4b (with matching wording in Checklist 3n) add: "For Agentic AI supplied by a Third Party, describe whether the company's own authority limits or the Third Party's defaults govern the AI agent's actions; whether, and on what contractual terms, the company can obtain decision-level records of those actions in a form usable for examination; how oversight is kept consistent when AI agents from more than one Third Party act in the same workflow; and how Third-Party model or tool updates are identified as material changes under Narrative 7."

Closing

We appreciate the Working Group's open, iterative process and the short timeline toward the Fall National Meeting. We would be glad to discuss these comments at a forthcoming public meeting of the Working Group, or to provide suggested redline language for any of the items above. Please direct any questions to Rakesh Sharma at rakesh.sharma@krishcons.com.

Respectfully submitted,

Anita Sharma
Co-founder & Chief Executive Officer
Krishcon Guide Agentic Solutions Inc. (KGAgentic)

Rakesh Sharma
Co-founder & Chief Technology Officer
Krishcon Guide Agentic Solutions Inc. (KGAgentic)

Krishcon Guide Agentic Solutions Inc. (KGAgentic), Cambridge, Ontario, Canada | Contact: rakesh.sharma@krishcons.com

RECON.AI

Evidence-based governance for AI and agentic systems

October 21, 2026

Via Electronic Submission

To: Scott Sobel, AI Policy Advisor; Miguel Romero, Director of Innovation, Cybersecurity and Technology

Re: AI Risk Evaluation Supplement Version 6.0 - Operational Governance Evidence and Risk-Based Examination

Dear Mr. Sobel and Mr. Romero:

Recon.AI appreciates the opportunity to comment on Version 6.0 of the NAIC AI Risk Evaluation Supplement. We support its objective of helping insurance regulators evaluate AI-related consumer and financial risks through existing supervisory and examination processes. We also support its recognition that the Supplement does not independently create new substantive legal duties or expand regulatory authority.

Our recommendations address a narrow examination question:

> **How can an insurer demonstrate, using proportionate and inspectable evidence, that the controls governing consequential AI actions actually operated as intended?**

An insurer may document an AI governance policy, establish a control, and deploy that control without necessarily being able to demonstrate how it operated during a particular consequential decision or examination period.

We recommend distinguishing three evidentiary states:

Evidence category	Supervisory purpose
1. Control design evidence	Policies, procedures, assigned responsibilities, and documented control requirements.
2. Control implementation evidence	Evidence that controls have been deployed in the applicable workflow.
3. Control operation evidence	Records demonstrating whether controls operated as intended, including exceptions and corrective actions.

This distinction can strengthen existing examination inquiries without prescribing a particular technology, logging architecture, or universal record-retention requirement.

1. Preserve examination visibility into consequential agentic authority

Draft reference: Exhibit A, Part Two (proposed deletion); Exhibit C, Question 3a.

The tracked-changes draft proposes deleting Exhibit A, Part Two, including an inquiry into whether agentic AI may execute actions that directly create or materially alter consumer, financial, or financial-reporting outcomes without prior human approval. Exhibit C, Question 3a, separately addresses agentic AI objectives

and authorizations. We recommend retaining the substance of the consequential-action inquiry within an existing, appropriately scoped exhibit rather than preserving an additional questionnaire in its entirety.

Suggested examination language:

> For agentic AI use cases within the scope of examination, describe the categories of consequential actions the AI System is authorized to execute without prior human approval, the applicable authorization limits, and how the company determines whether those limits were observed. Where material exceptions occurred, describe how they were identified, escalated, and resolved.

This helps distinguish technical capability to perform an action from organizational authority to execute it.

2. Preserve material-change and response evidence

Draft reference: Exhibit B, Narrative Question 7 (proposed deletion).

The tracked-changes draft proposes deleting a question addressing material-change criteria, approval, revalidation, enhanced monitoring, incidents, drift, third-party changes, and fallback. We recommend retaining a shorter, risk-based inquiry within Exhibit B.

Suggested examination language:

> For material AI Systems, describe the criteria used to identify changes or events requiring reassessment, approval, enhanced monitoring, restriction, or corrective action. Where applicable, explain how the company documents significant triggers, resulting decisions, and verification of corrective or recovery measures.

This preserves examiner visibility into the relationship between established thresholds and actions taken when those thresholds are reached.

3. Clarify evidence of effective human oversight

Draft reference: Exhibit B, Narrative Question 2(b).

The draft appropriately asks whether human review is present, how the company assures itself that review is appropriate, and how exceptions or errors are identified and resolved. We recommend clarifying that, for material consumer-impacting workflows, examiners may consider evidence of how review operates.

Suggested examination language:

> Where human review is a material risk control, describe how the company evaluates whether required reviews occur at the appropriate point in the workflow, whether reviewers have the authority and information needed to intervene, and how material exceptions or overrides are documented and resolved.

This does not assume human approval is required for every AI-assisted decision; it supports examination of oversight mechanisms appropriate to the risk.

4. Distinguish model-performance monitoring from governance-control monitoring

Draft reference: Exhibit C, ongoing testing and validation; Exhibit B, third-party oversight.

We support the draft's attention to model drift, accuracy, unfair discrimination, performance degradation, material changes, and corrective action. However, model performance may remain stable while operational controls become less effective - for example, through expanded executable permissions, recurring approval

overrides, or changed third-party integrations.

Suggested examination language:

> Where relevant to the AI System's risk profile, describe how the company assesses whether material authorization, human-oversight, escalation, and corrective-action controls continue to operate effectively following deployment or significant changes, including when the AI System relies on third-party components or services.

Supporting evidence may include existing system records, control testing, validation reports, exception reviews, or independent assessments. Examiners should consider the reliability, scope, provenance, and limitations of such evidence.

Implementation and proportionality

Risk-based scope. Focus more detailed operational evidence inquiries on material financial risks, direct consumer impacts, and consequential autonomous actions.

Existing evidence first. Consider reliable records and assessments already maintained by insurers or their service providers.

Technology neutrality. Do not prescribe a governance platform, logging system, or evidence format.

Established authority. Continue to operate within existing examination authorities without independently establishing new legal obligations.

Evidence that a governance control operated is not, by itself, proof that an underlying insurance decision was lawful, accurate, fair, or appropriate. Operational evidence complements, rather than replaces, substantive consumer-protection and financial-risk examination.

Conclusion

Recon.AI supports the NAIC's continued work to provide practical, proportionate examination guidance for AI Systems. We encourage the Working Group to preserve visibility into consequential agentic authorizations and material-change responses while refining existing questions to distinguish control design, implementation, and demonstrated operation.

The central supervisory question is not solely whether an insurer has adopted appropriate AI governance controls, but whether available evidence allows an examiner to determine that those controls functioned when needed.

Respectfully,

Dwayne Clark
Founder & CEO, Recon.AI
https://reconai.net

October 20, 2026

Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
Attention Scott Sobel and Miguel Romero

Comments on the AI Risk Evaluation Supplement Version 6

Dear Members of the Working Group:

On behalf of Legion Modular IP Holdings, Inc., I recommend advancing Version 6.0 with the targeted refinements below. These comments build on Legion's September 10, 2026 submission and are offered in a private capacity. They do not represent any governmental employer or agency.

Version 6 makes important improvements. It identifies the AI System as the primary unit of governance, separates generative and agentic uses in Exhibit A, asks about agentic objectives and authority in Exhibit C, and expands inquiries into material change, testing, and enterprise AI use. We support these changes and the express preservation of existing law and regulatory authority.

The remaining task is to make those improvements consistent and usable. An examiner should be able to determine whether a consequential decision or action stayed within approved authority and whether a harmful outcome can be identified and addressed. The focused amendments below preserve regulator judgment, permit equivalent evidence, and do not ask the NAIC to adopt a private standard or particular technology.

1 Keep the narrative and checklist options equivalent

References: Exhibit B instructions; Narrative questions 7 and 8(a)-(d); Checklist questions 2b, 2c, 3g, 3i, 3l, 3n, and 3p

Exhibit B directs regulators to use either the Narrative or the Checklist, not both. The new Narrative questions expressly address material-change review, unsanctioned AI, nonpublic personal information entered into unapproved tools, enterprise assistants, and workflow-specific risk. The Checklist has related governance prompts but does not present equivalent explicit inquiries. The chosen format should not determine whether these subjects receive attention.

Add corresponding checklist items and a short crosswalk, while retaining the choice of format. Insurers should be able to cross-reference responsive existing records rather than answer both versions.

Proposed language. The Narrative and Checklist versions cover equivalent governance topics. Use one version, tailored to the inquiry. For each applicable topic, the company may provide a response or identify and provide responsive current documentation as requested; a second-format response is not required.

2 Complete the distinction between models and system agency

References: General Guidance on AI System vs AI Model; Exhibit A columns 5 and 6; Exhibit C title and questions 3, 3a, and 9

Exhibit A's separate generative and agentic fields and Exhibit C question 3a improve the distinction between model characteristics and system capability. Exhibit C question 3 still groups generative and agentic AI within a model-type prompt. Complete the correction by asking separately about model technique, content-generating capability, and the system's authority to act. Align Exhibit C's title and purpose with the system-primary guidance, for example by referring to high-risk AI Systems and Models.

For a selected use that warrants detailed review, a concise dependency record should connect the accountable business owner, relevant models and providers, human approval points, and affected insurance

process. An existing diagram, inventory, or narrative should suffice if it answers the inquiry. This does not call for a new enterprise-wide architecture submission.

Proposed language. Identify the model techniques used in the selected AI System or use case. Separately describe whether it generates content, initiates or executes actions, or delegates tasks, and identify the scope of required human approval. Identify the accountable business owner and material dependencies relevant to the risk under review.

3 Test action limits and reconstruct consequential events

References: Exhibit B Narrative questions 1(g) and 2(b); Checklist questions 2b and 3l; Exhibit C questions 3a, 9, 10, and 11

Question 3a's inquiry into agentic objectives and authority is useful. It should also distinguish authority granted from authority enforced. Human involvement can mean observation, review after execution, or approval before an action takes effect. For consequential actions, the response should explain who can intervene, when, and with what practical effect.

For selected material uses, a risk-based sample or significant incident should connect the relevant system version and material inputs to the output, authorization or override, action, known outcome, and remediation. Such evidence helps establish whether the stated controls operated. It need not include every internal computation or require identical reproduction of a nondeterministic output. Record scope and retention should remain proportionate and consistent with applicable law, privacy, and security needs.

Proposed language. Describe authorized, prohibited, and approval-dependent actions; how those boundaries are enforced and tested; and who may override, suspend, or terminate the selected use. Explain the response to attempted unauthorized actions and the available correction or fallback when an action has taken effect. For a risk-based sample or significant event, provide sufficient evidence to reconstruct the material decision or action and the operation of the relevant controls, identifying significant gaps and remediation.

4 Apply change review and testing to the deployed system

References: Exhibit B Narrative questions 4(a)-(d) and 7; Checklist questions 3n and 3p; Exhibit C questions 10, 10a, and 11

We support the new material-change question and expanded predeployment and ongoing testing inquiries. Clarify that their reach includes material changes outside the underlying model. Instructions, permissions, connected services, data sources, or workflow can change the risk of a deployed use even if the model version remains unchanged.

Testing should address the intended use, foreseeable failures, and relevant consumer outcomes, including potential proxy effects where applicable. Question 11 appropriately seeks results and corrective actions. A focused addition should ask whether unresolved findings affect continued use and how remediation is verified. Independent review can be appropriate when risk, complexity, conflicts, or limits of internal expertise justify it; independence need not mean an external assessor for every use.

Proposed language. Apply material-change criteria to the deployed AI System, including changes to instructions, permissions, material dependencies, and workflow. Explain the approvals, testing or revalidation, and monitoring triggered by such changes. For significant findings, describe remediation, verification of its effectiveness, and any limits on continued use. Explain how the scope and independence of validation are determined in light of the use's risk.

5 Clarify consumer risk without equating a signal with a violation

References: Intent; General Guidance on Materiality and Risk Assessment; definitions of Adverse Consumer Impact and Direct Consumer Impact; Exhibit B Narrative question 1(h) and Checklist questions 2c and 3h; Exhibit C questions 7, 10, and 10a

Version 6 replaces Adverse Consumer Outcome with Adverse Consumer Impact and removes the definition's express condition that the impact violate regulatory standards. That can support earlier inquiry into harmful effects, but the guidance should distinguish a risk signal from an established violation or admission. This would reinforce the Intent section's existing limits without reducing attention to consumer harm.

Financial materiality should retain its financial-examination meaning. It should not become the sole filter for consumer-protection inquiry. Severity, frequency, duration, reversibility, potential unlawful discrimination, and complaint patterns may warrant attention even when aggregate financial effects are limited. A back-office label does not by itself establish an absence of consumer consequences. Inherent risk should remain identifiable separately from mitigating controls; those controls can inform the depth of further inquiry.

Proposed language. For purposes of scoping and evaluating an inquiry, identification of a potential or actual Adverse Consumer Impact does not, by itself, establish a violation of applicable law or constitute an admission of such a violation. Assess compliance under applicable law. Consider consumer-risk significance separately from financial materiality, including the nature, severity, recurrence, and reversibility of potential harm, while identifying inherent risk and relevant mitigating controls.

6 Address third party evidence gaps without diluting accountability

References: Exhibit B Narrative questions 4(b), 5, and 5(a); Checklist questions 3m and 3n; Exhibit C questions 5, 6, and 10

The existing third-party inquiries should remain substantive. Outsourcing a model, data source, or service does not remove the insurer's responsibility for its use. The insurer may nevertheless lack control over a component or unrestricted access to a provider's implementation. Review should examine the adequacy of available evidence and the response to limitations, rather than treating either a contract or a vendor assurance as conclusive.

For dependencies relevant to the selected risk, useful evidence may include due diligence, testing results or access arrangements, change and incident notification, known limitations, escalation routes, and continuity options. Equivalent evidence is appropriate when it answers the examination objective, but should not excuse a material assurance gap.

Proposed language. For material third-party dependencies of the selected use, explain how the company obtains sufficient evidence to assess and monitor risk and learns of material changes or incidents. Identify significant evidence limitations, compensating controls, the accountable company owner, and the conditions for restricting or discontinuing use.

7 Match evidence requests to the inquiry and protect lawful access

References: Instructions on prior submissions; General Guidance on Confidentiality; Exhibit A Company and Regulator Instructions

Version 6 supports current prior submissions and existing records, and its confidentiality guidance links protection to the requesting state's authority. Preserve those provisions. Harmonize Exhibit A's company instruction to provide the inventory together with Exhibit A with the regulator instruction allowing the inventory to make a separate Exhibit A unnecessary.

Confidential examination production must also remain distinct from the public comment record. Focused descriptions and existing evidence should ordinarily come first. Where those are insufficient, regulators should retain lawful access to additional materials through secure handling and applicable confidentiality procedures. Neither commercial sensitivity nor a vendor arrangement should create a categorical exemption. Conversely, sensitive implementation materials should not become routine prerequisites when less intrusive evidence adequately resolves the inquiry.

Proposed language. Where a current inventory or prior submission adequately addresses the requested information, the regulator may accept it in place of duplicative responses. Identify any remaining information needed. Tailor production to the examination objective and handle confidential, proprietary, personal, and security-sensitive material under applicable law and authorized procedures. Nothing in this guidance limits a regulator's lawful authority to obtain necessary information.

These refinements can be implemented within the existing exhibits and guidance. They would make the two governance formats consistent, connect system authority to evidence, and preserve consumer-risk visibility without imposing uniform technical requirements. Legion supports moving the Supplement forward on that basis. Thank you for considering these comments.

The Working Group may use these policy recommendations; this submission does not disclose or grant rights in Legion's proprietary implementation technology.

Respectfully,

William C. Biddle
Chief Executive Officer
Legion Modular IP Holdings, Inc.
wcb@legionmodularip.com

Source documents

NAIC AI Risk Evaluation Supplement Version 6.0 clean exposure draft

Legion September 10, 2026 comments in the NAIC published v5.0 comment compilation, PDF pages 127 to 138

COMMENTS ON THE AI RISK EVALUATION SUPPLEMENT, VERSION 6.0

NAIC Big Data and Artificial Intelligence (H) Working Group

Submitted by Solutions by Jewel (SBJ)

Submitter	Jewel Darlington, Chief AI Governance Officer & Founder
Organization	Solutions by Jewel, Global AI Governance Firm
Contact	governance@solutionsbyjewel.com
Website	solutionsbyjewel.com
Submitted to	Scott Sobel, AI Policy Advisor, and Miguel Romero, Director of Innovation, Cybersecurity and Technology, NAIC
Exposure	Artificial Intelligence (AI) Risk Evaluation Supplement, Version 6.0, exposed for a 14-day comment period ending October 30, 2026
Date	October 22, 2026
Reference	SBJ-PRC-039

1. About the Submitter

Solutions by Jewel (SBJ) is a global AI governance firm serving mid-market and enterprise organizations, including organizations in financial services. SBJ builds AI governance frameworks, decision records, and control mappings aligned to the NIST AI Risk Management Framework 1.0 and ISO/IEC 42001. SBJ's work centers on the evidence an organization keeps to show that its AI controls operate as designed, which is the evidence an examiner requests and tests.

2. Summary of Recommendations

SBJ supports the Working Group's development of the Supplement as a resource that lets examiners request AI governance information in a consistent, coordinated form. Version 6.0 [1][2] strengthens the Supplement in several places: it confirms that the Supplement operates within the authority of the existing Handbooks, it adds a tiered path from Exhibit A to Exhibits B through D, it asks for existing AI System and Model inventories, and it adds questions on unsanctioned AI use, nonpublic personal information, and general-purpose tools. SBJ's six comments build on those changes. Each concerns the evidence a company holds and an examiner can test.

1. **Exhibit C, Item 3a:** specify the Agentic AI authorization record, using the outcome-based test from Version 5.0.
2. **Exhibit B:** align the Narrative and Checklist forms so each produces the same coverage.
3. **Exhibit B Narrative, Item 7:** restore an illustrative list of material-change triggers and ask how the resulting decision is recorded.
4. **Exhibit B Narrative, Item 2b:** identify the evidence that shows human review is appropriate.

5. **Exhibit C, Item 6, and Exhibit B, Items 4b and 5:** ask how the company learns of and evaluates changes made by third parties.
6. **Exhibit A and Exhibit C:** link inventory fields to Exhibit C and restore the company's own risk classification of each model.

Section 9 lists editorial corrections. Exhibit and item references follow the numbering of the Version 6.0 clean draft [2]. References to Version 5.0 text come from the tracked-changes draft [1].

3. Comment 1: Agentic AI Authorizations (Exhibit C, Item 3a)

Draft text

> *"Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval."*

Observation

Version 5.0 asked, in Exhibit A, whether each Agentic AI System "is authorized to execute one or more actions that can directly create or materially alter a consumer, financial, or financial-reporting outcome without human approval before the action" [1]. Version 6.0 moves agentic inquiry to Exhibit C, Item 3a, and adds an Agentic AI column to the Exhibit A table. The Version 5.0 wording gives examiners a precise, outcome-based threshold for the agentic actions that matter most to consumers and to financial soundness.

Item 3a asks the company to identify its authorizations. To test the control, an examiner also needs the record behind each authorization: who granted it, what it covers, what the AI System can reach, and what causes it to be reviewed or withdrawn. Exhibit B Narrative, Item 1g, already asks how the company assesses "autonomy, reversibility, and reporting impact risk of AI Systems." Item 3a is the natural place for the model-level evidence of that assessment.

Recommendation

Restore the Version 5.0 outcome test in Item 3a and ask for the elements of the authorization record. SBJ's published AI Decision Record v2.0 [5] is one example of a record organized around these elements, including authorized actions, required human approval, escalation conditions, authorization source and date, access scope, and reassessment triggers.

Suggested text

> 3a. Whether the use case(s) involve Agentic AI. If so, identify: (i) the objectives of the Agentic AI; (ii) each action the AI System is authorized to execute that can directly create or materially alter a consumer, financial, or financial-reporting outcome without human approval before the action; (iii) the function that authorized those actions, the date of authorization, and the systems and data the AI System can access; (iv) the human approval points and escalation conditions that apply; (v) whether the authorized actions can be reversed, and how; and (vi) the conditions that trigger review or revocation of the authorization. Where the company maintains authorization records, provide them.

4. Comment 2: Parity Between the Narrative and Checklist Forms (Exhibit B)

Draft text

The Regulator Instructions for both forms of Exhibit B state that the Exhibits are designed so that regulators use one of the two forms.

Observation

Because a regulator selects one form, a topic that appears in one form and is absent from the other reaches a company when the regulator selects the form that contains it. A comparison of the two Version 6.0 forms shows the following:

Topic	Narrative	Checklist
Autonomy, reversibility, and reporting impact risk	Item 1g	Absent
Uses generating material financial impact, direct consumer impact (including adequacy of human review), financial statement impact, and risk or control assessment impact	Items 2a–2d	Absent
Independence of validation from development	Item 4a	Absent
Explainability and transparency of models recommending or making decisions	Item 4d	Item 3n (vendor systems)
Material change criteria and revalidation triggers	Item 7	Absent
Unsanctioned AI use, nonpublic personal information, enterprise assistants, general-purpose tools	Items 8a–8d	Absent
Quantification of AI System risk levels	Absent	Item 3e
Consumer complaints resulting from AI Systems	Absent	Item 3j
Consumer awareness and disclosure	Absent	Item 3k
Employee training and prohibited practices	Absent	Item 3l

Recommendation

Add Checklist rows for the topics that appear in the Narrative alone, and Narrative sub-items for the topics that appear in the Checklist alone, so that either form produces the same coverage and responses remain comparable across states.

Suggested text

Checklist, new rows under Item 3 (“How the Insurance Company…”):

3q. Assesses the autonomy, reversibility, and reporting impact risk of AI Systems
3r. Assures itself that human review of AI System outputs is appropriate, including for high-volume transactions
3s. Performs validation and testing of AI Systems independently of development
3t. Ensures explainability and transparency of AI models recommending or making decisions
3u. Identifies material changes and triggers approval, revalidation, or enhanced monitoring
3v. Discovers unsanctioned AI use and prevents nonpublic personal information from entering

unapproved AI models
3w. Governs enterprise assistants with AI capabilities and determines when a general-purpose tool becomes a high-risk use because of the workflow in which it is deployed

Narrative, new sub-items under Item 3 (“Discuss how the company:”):

c. Quantifies AI System risk levels
d. Identifies, tracks, and addresses consumer complaints resulting from AI Systems
e. Promotes consumer awareness of the use of AI Systems through disclosures, policies, and procedures for consumer notification, as appropriate
f. Trains employees about AI System use and defines prohibited practices (if any)

5. Comment 3: Material Change Criteria (Exhibit B Narrative, Item 7)

Draft text

“Discuss the criteria for identifying a material change and for triggering approval, revalidation, or enhanced monitoring, including predeployment testing”

Observation

The tracked-changes draft shows an enumerated list struck from this item: “periodic or event-triggered reassessment, changes in operational outcomes, incidents and near misses, drift or degradation, third-party changes, and safe retirement or fallback” [1]. An illustrative list helps companies understand the scope of the question and gives examiners consistent points to compare across companies. The Supplement already confirms that its questions leave AIS Program requirements unchanged, so an illustrative “may include” list leaves the criteria themselves to each company.

The item asks for criteria. The record of what the company decided once a trigger occurred is the evidence that shows the criteria operate.

Recommendation

Restore the list as illustrative and ask how the resulting decision is documented and approved.

Suggested text

7. Discuss the criteria for identifying a material change and for triggering approval, revalidation, or enhanced monitoring, including predeployment testing. Criteria may include, as applicable, periodic or event-triggered reassessment, changes in operational outcomes, incidents and near misses, model drift or performance degradation, changes made by third parties to AI Systems or data, and safe retirement or fallback. Describe how the resulting decision to continue use, revalidate, modify, or retire the AI System is documented and approved.

6. Comment 4: Evidence of Appropriate Human Review (Exhibit B Narrative, Item 2b)

Draft text

“…whether there is a human in the loop, whether there are high volume transactions, how the company assures itself the human review is appropriate, and how any exceptions or errors are

found and resolved."

Observation

The clause added in Version 6.0, "how the company assures itself the human review is appropriate," moves the question from whether a reviewer exists to whether the review functions. Companies that operate human review at scale typically hold evidence on this point: the reviewer's function and authority, the information available at the time of review, review volumes relative to reviewer capacity, and the rate at which reviewers modify or override AI System outputs. A consistently low override rate in a high-volume workflow is a useful indicator for examiner follow-up, since it can signal that review has become confirmation of the AI System output.

Recommendation

Name the elements a company should address when describing its assurance over human review.

Suggested text

> In describing how the company assures itself the human review is appropriate, address, as applicable: the function and authority of the reviewer, including authority to override the AI System output; the information available to the reviewer at the time of review; review volumes relative to reviewer capacity; the frequency with which reviewers modify or override AI System outputs and how that frequency is monitored; and the records retained for each review.

7. Comment 5: Changes Made by Third Parties (Exhibit C, Item 6; Exhibit B, Items 4b and 5)

Draft text

> *"Did the Company purchase the AI model, was the AI model developed specifically for the Company? Is the Company able to modify the AI model as necessary?"*

Observation

Item 6 asks whether the company can modify a third-party model. A third party's update to a model, its training data, or its configuration can change outcomes for consumers independently of any action by the insurer, and the insurer's ability to detect that change depends on notice terms and monitoring. The NAIC Model Bulletin [4] addresses third-party AI Systems in Guideline 4, including due diligence (4.1) and contract terms that "provide audit rights and/or entitle the Insurer to receive audit reports by qualified auditing entities" and "require the third party to cooperate with the Insurer with regard to regulatory inquiries and investigations" (4.2). Exhibit B, Item 4b, asks about validation and testing of vendor-supplied AI Systems, and Item 5 asks about AI Systems used by third-party service providers. A question on the evidence the company holds from the third party would complete both items.

Recommendation

Ask how the company learns of and evaluates third-party changes, and what evidence it holds from the third party, consistent with Model Bulletin Guideline 4.

Suggested text

Exhibit C, Item 6, add:

> Does the third party notify the Company of changes to the AI model, its training data, or its configuration? How does the Company evaluate those changes before they affect production use?

Exhibit B Narrative, Item 4b, add:

> Identify the evidence the Company holds from the third party to support its validation and testing, such as test results, model documentation, and audit reports, and describe any contract terms addressing audit rights, notification of changes, and cooperation with regulatory inquiries.

The same addition can apply to Item 5 and to Checklist Item 3n.

8. Comment 6: Linking the Inventory to Exhibit C (Exhibit A; Exhibit C)

Draft text

> *"If a company has an existing AI System and/or Model inventory, the company should provide the AI System and/or Model inventory together with the completed Exhibit A."*
>
> *"AI model risk criteria are set by the insurance company." (Exhibit C, Purpose)*

Observation

Version 6.0's request for existing inventories strengthens Exhibit A as the tool a regulator uses to select AI Systems for Exhibits B through D. Three fields, where a company's inventory already includes them, would let a regulator move directly from the inventory to targeted requests: whether the AI System or Model is internally developed or supplied by a third party (with vendor name), the accountable owner by function, and the date of last validation or testing. The vendor field also connects the inventory to Exhibit B, Items 4b and 5.

The tracked-changes draft shows Exhibit C, former Item 7, "Model risk classification (high, moderate, low, etc.)," deleted [1]. Because Exhibit C leaves risk criteria to the company, the company's own classification of each model, and the criteria it applied, is the field that connects an Exhibit C response to the "Inherent Risk Level" column the regulator reviews in the inventory.

Recommendation

Add the three inventory fields on a where-maintained basis, and restore a risk classification row in Exhibit C.

Suggested text

Exhibit A, Company Instructions, add to the list of inventory columns:

> Where the inventory includes them: source of the AI System and/or Model (internal, third party with vendor name, or both); accountable owner, by function; and date of last validation or testing.

Exhibit C, new row following Item 2:

> The Company's inherent risk classification of the model and the criteria applied, consistent with the Inherent Risk Level reported in the AI System and/or Model inventory.

9. Editorial Notes

Location	Draft text	Suggested correction
Intent, first paragraph	"…from the use of AI Systems.."	Remove the duplicate period.
Exhibit A heading	"Exhibit A: Part One - Quantify Regulated Entity's Use of AI Systems"	Version 6.0 deletes Part Two. Retitle as "Exhibit A: Quantify Regulated Entity's Use of AI Systems."
Exhibit A, Purpose	"…their AIS Program (Exhibits B)…"	"(Exhibit B)"
Exhibit B Narrative, Item 2b	"…found and resolved.."	Remove the duplicate period.
Exhibit B Narrative, Item 6	Unnumbered paragraph beginning "Discuss the unit(s) responsible for the AIS Program…"	Number as Item 6a.
Exhibit B Narrative, Items 3, 7, and 8	Response fields absent; Item 7 lacks terminal punctuation.	Add "Click or tap here to enter text." response fields and terminal punctuation, consistent with Items 1, 2, 4, and 5.
Exhibit C, Item 9	"Does the AI modelautomate…"	"Does the AI model automate…"
Exhibit D, Purpose	"…data used directlyby an AI model(s)…"	"…data used directly by an AI model(s)…"
Definitions, introduction	"…Use of AI Systems by Insurers ():"	Insert the citation or link to the Model Bulletin in the empty parentheses.
Definitions, "Generalized Linear Model (GLM)"	"…extends linear regression to by relating predictive variables…"	"…extends linear regression by relating predictive variables…"

10. Closing

SBJ appreciates the Working Group's attention to these comments and its work to give state regulators a consistent, coordinated resource for evaluating AI governance at insurers. SBJ welcomes questions on any of these recommendations at governance@solutionsbyjewel.com.

References

1. NAIC Big Data and Artificial Intelligence (H) Working Group, Artificial Intelligence (AI) Risk Evaluation Supplement, Version 6.0 (Tracked Changes). https://content.naic.org/sites/default/files/inline-files/ai-risk-evaluation-supplement-v6.0-tracked-changes_1.docx
2. NAIC Big Data and Artificial Intelligence (H) Working Group, Artificial Intelligence (AI) Risk Evaluation Supplement, Version 6.0 (Clean). https://content.naic.org/sites/default/files/inline-files/ai-risk-evaluation-supplement-v6.0-clean_1.docx
3. NAIC, Big Data and Artificial Intelligence (H) Working Group, Exposure Drafts. https://content.naic.org/committees/h/big-data-artificial-intelligence-wg
4. NAIC, Model Bulletin: Use of Artificial Intelligence Systems by Insurers, adopted by the Executive (EX) Committee and Plenary, December 4, 2023. https://content.naic.org/sites/default/files/cmte-h-big-data-artificial-intelligence-wg-ai-model-bulletin.pdf.pdf
5. Solutions by Jewel, AI Decision Record v2.0. https://solutionsbyjewel.com/ai-decision-record-template

EASYESSENCE
pkangari@easyessence.com | easyessence.com | 813-444-7830

October 23, 2026

VIA ELECTRONIC MAIL (ssobel@naic.org; maromero@naic.org)

Commissioner Nathan Houdek, Chair
Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
1100 Walnut Street, Suite 1500
Kansas City, MO 64106

Re: Comments on the AI Risk Evaluation Supplement, Version 6.0

Dear Commissioner Houdek:

Thank you for the opportunity to comment on Version 6.0 of the AI Risk Evaluation Supplement. I write as the founder of EasyEssence, an independent practice that evaluates insurers' consumer-facing AI outputs against their policy forms. We build none of the systems we review and take no role in tuning them. We commented on Version 5.0 in a letter filed September 8, 2026. Points carried over from that letter are restated here in shorter form so this letter can be read on its own.

Version 6.0 strengthens the Supplement's questions about what an AI System produces, and the changes we support retaining are listed at the end of this letter. Version 6.0 also makes Exhibit A the exhibit regulators will use most often, while removing its Part Two. A shorter Exhibit A now carries more weight, because its answers decide whether a regulator looks further.

When an AI System answers a consumer's coverage question incorrectly, the consumer acts on an answer the policy contradicts. The Supplement's Background states that insurers' decisions and actions impacting consumers that are "made or supported by" advanced analytical and computational technologies, including AI Systems, must comply with all applicable insurance laws and regulations. Our suggestions concern what a regulator learns about those AI Systems at the first step, and whether the Supplement's accuracy questions reach them.

Of the six suggestions below, four extend existing text, one adds a sentence to a definition, and one adds three Checklist rows that mirror existing Narrative questions. Proposed sentences appear as indented blocks where they are introduced, and all suggested language is collected in the appendix.

1. We suggest adding one sentence to column (4) of Exhibit A, asking whether the outputs of systems with a direct consumer impact are tested for accuracy after deployment, and by what method.

Column (4) is now the one place every respondent describes consumer impact. Version 6.0 removed Exhibit A's Part Two, including the Model Inventory request, and asks for inventory information only from companies that already keep one. We understand the aim of a lighter Exhibit A and support it. The Supplement's accuracy questions are in Exhibit C, which a company receives only when a regulator decides to look further, and only for models the company itself rates high-risk. Column (4) is therefore where a regulator decides whether those questions get asked at all.

Column (4), headed "Use Case," currently reads "Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact." We suggest adding one sentence to the column, and the same question to the inventory columns for companies that keep an inventory:

> For systems with a direct consumer impact, state whether their outputs are tested for accuracy after deployment, and if so, by what method.

Under the revised column, one company might answer "Outputs are checked for accuracy against the policy forms in force," and another "Outputs are not separately tested. Representatives review each answer before delivery." The method is what makes the answer usable, because it shows whether a company is measuring its outputs or describing a control around them, and an examiner can see which one is in place without a follow-up request. The addition applies only to rows that include a system with direct consumer impact, and it asks a company to state what it already does rather than to do anything new. A company that does not test can answer in a few words. The regulator keeps full discretion over what the answer means, and the answer provides a second signal alongside the company's own risk rating, at the step where the decision to look further is made.

2. We suggest the Direct Consumer Impact definition state that an AI System has Direct Consumer Impact when its answer reaches a consumer, whether the AI delivers the answer directly, as in a chat window, or a company representative relays it.

Carriers increasingly use AI that drafts the answer to a consumer's coverage question for a company representative to deliver. Under the current definitions, a company could reasonably classify that AI as Support rather than Augmentation, because the representative passes the answer along rather than making a decision with it. Support falls outside Direct Consumer Impact. The consumer still receives the AI's answer, yet the AI System would not be reported as having a direct consumer impact in Exhibit A, and would not be considered for Exhibit C on that basis.

We suggest adding one sentence to the definition, after its list of examples. The sentence does not change how any other AI System is classified, and it settles one configuration the definitions currently leave open. It would read:

> An AI System has Direct Consumer Impact where its output is communicated to a consumer, whether the AI System delivers that output directly or a company representative relays it.

Version 6.0 already recognizes that a tool's risk can depend on how it is used. Exhibit B question 8d asks when a general-purpose tool becomes high-risk "because of the workflow in which it is deployed." The suggested sentence carries that idea into classification, so that an AI System is classified by where its answer ends up. Question 8d addresses a company's governance policy, while the definition determines what a company counts and reports. With the sentence, two insurers running the same configuration would report it the same way, and the totals a regulator uses to scope inquiries would account for AI Systems whose answers reach consumers through a representative.

3. We suggest the Risk Assessment guidance state explicitly that two common safeguards are controls: human review of an AI System's outputs before delivery, and automated checks its platform runs on those outputs.

The Supplement asks companies for an inherent risk assessment, meaning risk before controls, and the company's own rating determines which models it treats as high-risk for Exhibit C. The second safeguard includes vendor features that compare an AI's answer to policy language or block certain answers before the consumer sees them. The guidance does not say where either safeguard belongs. A company that credits either one when setting the inherent rating will rate the same AI System lower than a company that does not.

Exhibit B question 2b asks how the company assures itself that human review is appropriate, which treats that review as a control to be evaluated. We suggest the Risk Assessment guidance say the same where it governs the rating. The Inherent Risk definition would stay as it is, and ratings would become comparable across companies. The added sentence would read:

> Controls include human review of an AI System's outputs before delivery and automated checks that the AI System's platform performs on those outputs before delivery.

4. We suggest Exhibit C's last-testing field also ask when the model's live outputs were last checked for accuracy, and who performed the check, including whether they were independent of the model's development.

We welcome the request in Exhibit C for the results of the last testing and any corrective actions. A model keeps changing after deployment. Where its version is not pinned, it updates on the supplier's schedule. The policy forms it draws on are amended and its instructions are tuned. Version 6.0 recognizes these changes by asking about testing following material changes to the model, data, or use of the model. A testing date alone does not show whether the answers consumers received during the review period were checked. The date of the last live-output check, with the model version in use at the time, does.

Knowing who performed the check, and whether they were independent of the model's development, tells an examiner how much weight a result carries. Exhibit B asks the same independence question at the policy level. The check is sometimes performed by another AI model, and research presented at NeurIPS 2024 found that large language model evaluators give inflated scores to their own outputs.[1] The suggested wording asks a company to name any AI model used. It does not prefer one testing arrangement over another.

The field currently reads "Last date of model testing or validation. Please include the results and corrective actions as a result of the last model testing, if applicable." We suggest adding one sentence:

> Please also include the date of the most recent accuracy check of the model's live outputs, the model version in use at that time, and who performed the check, including whether they were independent of the model's development and any AI model used to evaluate the outputs.

5. We suggest the Exhibit B Checklist include three questions that now appear only in the Narrative.

Version 6.0 asks regulators to use either the Exhibit B Narrative or the Checklist, not both. Several Narrative questions have no Checklist counterpart, and three of them bear directly on how consumer-facing AI is tested and rated. They ask whether validation and testing are independent from development (question 4a), what material changes trigger revalidation or enhanced monitoring (question 7), and when a general-purpose tool becomes high-risk because of its workflow (question 8d). Question 4a is the only place in the Supplement that asks whether testing is independent of development. A regulator who chooses the Checklist will not ask any of the three. Adding three rows to the Checklist would make these inquiries available in either format. The rows would read:

> - Determines whether validation and testing are performed by someone that is independent from development
> - Identifies material changes that trigger approval, revalidation, or enhanced monitoring
> - Decides when a general-purpose tool with AI capabilities becomes a high-risk use because of the workflow in which it is deployed

6. We suggest the Exhibit C purpose note that, for AI Models with Direct Consumer Impact, the number of consumers affected is a consideration alongside the dollar threshold.

Exhibit C asks companies to consider materiality when rating a model's risk, and the Supplement defines materiality as a dollar amount tied to the insurer's financial position. That measure fits financial models well. A consumer-facing model's errors are small in dollars per conversation and repeated across many conversations, so a dollar threshold alone may rate such a model low. Exhibit B already recognizes the difference by asking whether uses with direct consumer impact involve high volume transactions. We suggest adding one sentence to the Exhibit C purpose, leaving the Materiality definition and its alignment with the Financial Condition Examiners Handbook as they are:

> For AI Models with Direct Consumer Impact, the number of consumers affected is also a consideration.

Changes we support retaining

Several changes in Version 6.0 support the aims of these comments, and we ask that they be retained in Version 7.0. Exhibit C distinguishes validation before deployment from ongoing monitoring after deployment, and asks about testing after material changes. Exhibit B asks how a company assures itself that human review is appropriate. Exhibit B also asks how a company governs enterprise assistants and decides when a general-purpose tool becomes high-risk. Exhibit A separates Generative AI from Agentic AI and asks for an existing inventory alongside the completed exhibit.

Our six suggestions keep a shorter Exhibit A pointed at the consumer-facing AI Systems that warrant a closer look, and help the accuracy questions in Exhibits B and C reach those systems. None asks the Working Group to replace company self-assessment, which is the right foundation for this instrument. Each makes an existing question reach the exposure it was written for.

I would welcome the opportunity to discuss any of these points with the Working Group or NAIC staff, including at the November 2 meeting if time for interested-party comment is available.

Respectfully submitted,

Phillip Kangari
Founder, EasyEssence
pkangari@easyessence.com

cc: Commissioner Doug Ommen, Co-Vice Chair; Mary Block, Co-Vice Chair; Scott Sobel, NAIC; Miguel Romero, NAIC

---

[1] Panickssery, Bowman, and Feng, "LLM Evaluators Recognize and Favor Their Own Generations," NeurIPS 2024.

Appendix: Suggested Language

Point 1. Exhibit A, Part One, column (4) "Use Case," add after "whether there is a direct consumer impact or a material financial impact": For systems with a direct consumer impact, state whether their outputs are tested for accuracy after deployment, and if so, by what method.

Exhibit A, Company Instructions, add to the inventory columns regulators will look for: Whether the AI System and/or Model's outputs are tested for accuracy after deployment, and if so, by what method

Point 2. Definitions, "Direct Consumer Impact," add after "claims denial and settlement-related decisions.": An AI System has Direct Consumer Impact where its output is communicated to a consumer, whether the AI System delivers that output directly or a company representative relays it.

Point 3. General Guidance, Risk Assessment, add after "(i.e., risk before the consideration of controls or mitigation measures).": Controls include human review of an AI System's outputs before delivery and automated checks that the AI System's platform performs on those outputs before delivery.

Point 4. Exhibit C, field "Last date of model testing or validation," add after "if applicable.": Please also include the date of the most recent accuracy check of the model's live outputs, the model version in use at that time, and who performed the check, including whether they were independent of the model's development and any AI model used to evaluate the outputs.

Point 5. Exhibit B, Checklist, item 3 ("How the Insurance Company…"), add three rows:

- Determines whether validation and testing are performed by someone that is independent from development
- Identifies material changes that trigger approval, revalidation, or enhanced monitoring
- Decides when a general-purpose tool with AI capabilities becomes a high-risk use because of the workflow in which it is deployed

Point 6. Exhibit C, Purpose, add after "Note that materiality should be a consideration in determining the level of risk.": For AI Models with Direct Consumer Impact, the number of consumers affected is also a consideration.

Falsify OÜ · Reg. 17574308 · Narva mnt 5, 10117 Tallinn, Estonia · cuneyt@falsify.dev

26 October 2026

To: Scott Sobel and Miguel Romero, Big Data and Artificial Intelligence (H) Working Group, NAIC

Re: AI Risk Evaluation Supplement version 6.0

Dear Mr. Sobel and Mr. Romero,

Falsify OÜ (Tallinn, Estonia) submits two focused comments on version 6.0 of the AI Risk Evaluation Supplement, exposed for comment until 30 October 2026. These follow up on our version 5.0 submission, listed as No. 30 in the comment compilation dated 2 October 2026.

Comment 1: Exhibit C, field 10

The discussion of pre-deployment validation would be easier to assess if it identified the acceptance criteria and when they were established. We suggest adding:

> "Identify the acceptance criteria for deployment, including metrics and thresholds where applicable. State when those criteria were established relative to the availability of validation results, and describe any subsequent revisions, their rationale and approval."

Comment 2: Exhibit B (Narrative), item 7

For the criteria used to identify material changes and trigger approval, revalidation or enhanced monitoring, we suggest adding:

> "Describe how the company records when these criteria were established and preserves subsequent changes, including their dates, rationale and approval."

Neither proposal requires a particular format, registry, tool or vendor. Companies could reference existing model governance and document control records where these contain the requested information.

An independent timestamp is one optional way to support a record's existence and integrity. A validated timestamp supports the claim that the committed version existed no later than the stated time and that a subsequently disclosed copy matches it. It does not, by itself, establish when validation began or results became available, whether the criteria were followed, or whether they were adequate.

The proposals allow for justified revisions during development and validation. Their purpose is to make the chronology, rationale and approval of changes available for regulatory review.

Interest statement. Falsify maintains an open specification for tamper-evident records of evaluation criteria and their timing, and offers related paid services. Neither proposal depends on our specification or services.

We consent to this submission being made public.

Kind regards,

Cüneyt Öztürk Founder, Falsify OÜ Tallinn, Estonia

Dr. Kaenan Hertz
Managing Partner, Insurtech Advisors LLC
kh@insurtechadvisors.com | insurtechadvisors.com
October 26, 2026

By email: Scott Sobel (ssobel@naic.org) and Miguel Romero (maromero@naic.org)
NAIC Big Data and Artificial Intelligence (H) Working Group

Re: Comments on the AI Risk Evaluation Supplement, Version 6.0 Exposure

Dear Mr. Sobel and Mr. Romero,

Thank you for the opportunity to comment on version 6.0 of the AI Risk Evaluation Supplement. I write as an advisor to regional and mutual insurance carriers on AI governance, and as someone who has spent three decades inside the insurance industry's judgment processes. Version 6.0 reflects genuine responsiveness to the comments the Working Group received on earlier versions. The points below are offered in that spirit.

1. Support the legal-effect sentence. The statement that the Supplement does not independently create substantive legal duties, standards of conduct, admissions, or expanded regulatory authority is the most important sentence in version 6.0. Please carry it into the final version unchanged. It is what allows a carrier to answer an examiner candidly.

2. Support Exhibit A-first sequencing, and define the escalation path. The design intention that regulators will most often use Exhibit A, escalating to Exhibits B through D for only a subset of a company's use of AI, is right. The final version should say what triggers escalation, even in general terms, so that proportionality does not depend entirely on individual discretion.

3. Support the GLM exclusion, and consider making it the default. For property and casualty carriers, the most consequential models are generalized linear models inside filed and reviewed rating plans. Rate filing is the existing oversight. I would suggest the final version state that GLMs within filed rating plans are presumed out of scope for Exhibits B through D absent specific cause, rather than leaving each regulator to choose. Otherwise every carrier's rating plan risks being re-litigated model by model at examination.

4. Strengthen the acceptance of existing documentation. "May consider accepting" should be "should accept" where a company's existing AI inventory, ERM documentation, or model risk papers are responsive. For a regional carrier, duplicative document production is the single largest cost of this Supplement, and it buys no consumer protection.

5. Extend coordination across states. The new coordination paragraph should expressly cover multi-state coordination, so that a carrier examined by several states is not asked to produce the same Exhibit A inventory in parallel. The Handbook practices the paragraph cites already point this way.

6. Define X in Exhibit A. "Number of AI Systems Implemented in the Past X Months" needs a value. Suggest 24 months, which is long enough to show trajectory and short enough to be answerable from change logs.

7. Make the subject-matter-expert list available on request. Requiring a detailed list of subject matter experts up front assumes a bench that a 50-person carrier does not have. The same information, provided when follow-up inquiries actually arise, achieves the regulatory objective.

8. Thank you for removing the staffing-reductions question. Workforce levels are not a solvency or consumer-protection inquiry, and the replacement question, how the company assures itself that human review is appropriate, is the better one.

9. Proportionality for unsanctioned-AI discovery. Expecting a carrier to discover unsanctioned AI use assumes tooling many regional carriers do not own. Suggest adding "reasonable measures proportionate to the insurer's size and complexity," consistent with the proportionality principle in the NAIC model bulletin that more than half the states have now adopted.

10. Calibrate vendor transparency to what a carrier can actually obtain. Checklist item 3n asks carriers to govern, monitor, test, and ensure transparency of vendor-developed AI. A regional carrier has contractual diligence and inquiry rights, not the ability to compel a vendor's technical disclosure. The final version should tie the carrier's obligation to reasonable inquiry and contractual diligence, or the market will simply stop selling AI to small carriers, which is the opposite of the Supplement's intent.

If it would help the drafting, I am glad to walk any regulator through how these points land at a regional carrier, particularly the proportionality items.

Respectfully,

Dr. Kaenan Hertz
Managing Partner, Insurtech Advisors LLC

SPECK & WARD

GOVERNED DECISION WORKFLOWS FOR INSURANCE OPERATIONS

speckandward@gmail.com

October 22, 2026

Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
Via Scott Sobel and Miguel Romero

> **Re: Comments on the AI Risk Evaluation Supplement v6.0 Exposure — Exhibit B's Question on How a Company "Assures Itself the Human Review Is Appropriate"**

Dear Members of the Working Group:

I am the founder of Speck & Ward, a consultancy that designs governed decision workflows for insurance operations — the systems through which AI-assisted recommendations are reviewed, approved, escalated, and reconstructed after the fact. I write from the builder's side of the table: the controls your exhibits ask about are controls someone must design, instrument, and operate.

I comment on one item only, because I believe it is the highest-leverage sentence in v6.0: Exhibit B's new question asking how a company *assures itself the human review is appropriate*, echoed in the checklist exhibit at item 3l, "ensures human oversight is appropriate."

The question is the right one, because presence is not a control.

Most current practice can answer "Is there a human in the loop?" with a process diagram. That answer is compatible with review that is operationally meaningless: a reviewer who cannot change the outcome; who lacks the expertise the decision requires; who is given seconds per case at queue volumes no human can genuinely evaluate; or whose agreement rate is effectively 100 percent, because disagreement carries a personal cost and agreement carries none.

Commenters on earlier drafts made the same point from the consumer side, noting that review completed in a matter of seconds cannot be meaningful review. "Appropriate" is the word that closes that gap — if companies are asked to demonstrate it, rather than assert it.

Appropriateness has testable components.

In our design work, a human review step is defensible only when a company can evidence four things:

1. **Authority.** The reviewer can reverse, modify, or escalate the AI-assisted decision, and the record shows that reversals actually occur and are documented with reasons. A review step with a zero lifetime reversal rate is not oversight; it is a signature.

2. **Competence matched to the decision.** The reviewer's qualifications correspond to what is being decided — the same principle the Supplement applies to model governance, applied to the human checkpoint.

3. **Capacity proportionate to consequence.** The time, information, and caseload a reviewer is given are sufficient for decisions of that consequence. This is where a tiering principle does real work: the depth of human review should scale with the consequence and materiality of the decision. An adverse coverage or claims decision warrants a different review envelope than a low-impact administrative output. A single, undifferentiated "human in the loop" standard will be simultaneously too weak for consequential decisions and too burdensome for trivial ones.

4. **Evidence of the review itself.** The company audits its reviewers, not only its models: distributions of time-on-review, override and escalation rates, sampled re-review of human decisions, and how exceptions and errors found in review are resolved — which, I note, the same Exhibit B passage already asks about. A company that "assures itself" should be able to produce these artifacts on request.

Two suggestions.

First, consider accompanying guidance — in the Supplement's instructions or in the Working Group's response to comments — that gives examples of evidence demonstrating appropriateness. Without examples, this question risks collapsing into narrative attestation, and the companies with the weakest review will write the smoothest paragraphs.

Second, the appropriateness lens connects naturally to Exhibit C's new item 3a on agentic AI: the objectives and authorizations a system is granted *without* human approval. Where no human reviews each decision, the appropriateness question becomes: What envelope of pre-authorized action has a named human accepted accountability for, what bounds it, and what monitoring evidence shows the system stayed inside it? The Exhibit B question and the Exhibit C question are the same control viewed from two sides, and the final text would be stronger if that connection were visible.

I appreciate the Working Group's work on v6.0 and the care evident in the tracked changes. I am available to discuss the operational detail behind any of the above.

Respectfully submitted,

Daniel Whiteley
Founder, Speck & Ward
speckandward@gmail.com

Ali Houshangi, MSc, CISA, CISSP
Huntington Beach, California | ali.houshangi@gmail.com
October 26, 2026

Scott Sobel, AI Policy Advisor
Miguel Romero, Director of Innovation, Cybersecurity and Technology
National Association of Insurance Commissioners

Re: Comments on AI Risk Evaluation Supplement version 6.0 - Exhibit C item 3a

Individual capacity. These comments are submitted solely in my individual capacity. They do not represent any employer, client, professional association, or other organization.

Dear Mr. Sobel and Mr. Romero:

Thank you for the opportunity to comment on the AI Risk Evaluation Supplement version 6.0 (v6.0), and for considering my September 8, 2026 comments in adding Exhibit C item 3a. These comments are limited to item 3a. I do not propose further changes to Exhibit A or Exhibit B.

1. Clarify "authorizations"

Item 3a reads: "Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval." The word "authorizations" could be read as a request for access permissions or credentials, which my September 8 comments expressly did not seek. The inquiry appears intended to identify what the AI System is permitted to do without prior human approval. Referring to types of action would make that clear and would track the Supplement's definition of Agentic AI ("taking actions (often through tools or APIs)").

2. Extend monitoring to actions taken by the AI System

Item 3a identifies what an agentic AI System is permitted to do. Item 10 addresses validation and ongoing monitoring of model outputs. For an AI System that takes actions, the actions themselves can affect consumers or the company, and Exhibit C does not expressly ask how the company monitors them. Exhibit B Narrative item 2b asks, for uses that generate a direct consumer impact, "how any exceptions or errors are found and resolved." The Checklist version has no corresponding item, and v6.0 provides that regulators would use one version or the other. The wording below applies the same question to the use case identified in item 3a.

Item	Wording
Option A (preferred): item 3a	"Whether the use case(s) involve Agentic AI, and if so, identify its objectives and the types of action it is permitted to take without prior human approval, and describe how the company monitors actions taken by the AI System and how any exceptions or errors are found and resolved."
Option B: item 3a	"Whether the use case(s) involve Agentic AI, and if so, identify its objectives and the types of action it is permitted to take without prior human approval."
Option B: new item 3b	"If the use case(s) involve Agentic AI, describe how the company monitors actions taken by the AI System and how any exceptions or errors are found and resolved."

Neither option prescribes a monitoring method, record format, or technical architecture, or requires human approval of any action. Both ask the company to describe its existing practices, consistent with v6.0's statement that the Supplement does not independently create substantive legal duties. I recognize that either option adds some response effort; it is limited to high-risk models within Exhibit C whose use cases involve Agentic AI.

Thank you for considering these comments. I would be glad to provide further detail if that would be useful.

Respectfully submitted,

Ali Houshangi, MSc, CISA, CISSP
Huntington Beach, California
ali.houshangi@gmail.com

Comment on AI Risk Evaluation Supplement Version 6.0

Submitted to: NAIC Big Data and Artificial Intelligence (H) Working Group
Submitted by: Narayanan Viswanathan, PhD, Founder & CEO, SecuShift, Inc.
Date: October 27, 2026

I appreciate the Big Data and Artificial Intelligence (H) Working Group's continued development of the AI Risk Evaluation Supplement and its focus on effective AI governance, Adverse Consumer Impacts, consumer complaints, testing, monitoring, and corrective action.

I respectfully suggest that the Working Group consider adding a targeted inquiry concerning the **outcome of remediation following an identified Adverse Consumer Impact**.

Version 6.0 appropriately asks how an insurer considers consumer complaints resulting from AI Systems and whether they are identified, tracked, and addressed. Exhibit C also asks insurers to discuss testing for potential Adverse Consumer Impacts and, where testing identifies material differences or potential impacts, the process for evaluating results, identifying contributing factors, and determining whether changes to models, data, methodologies, controls, or other reasonably available alternatives are appropriate. The Supplement further requests results and corrective actions arising from model testing.

These inquiries provide important information concerning whether corrective-action processes are functioning. However, where an AI System has contributed to an identified Adverse Consumer Impact affecting a specific consumer, completion of an organizational corrective process may not by itself demonstrate whether the impact experienced by that consumer was effectively addressed.

The Working Group could therefore consider adding an inquiry such as:

"Where an AI System has contributed to an identified Adverse Consumer Impact, describe how the company determines whether resulting remediation or corrective action has been completed and whether the identified Adverse Consumer Impact has been effectively addressed for the affected consumer. Where applicable, describe how that determination is documented."

This inquiry could complement the Supplement's existing questions without prescribing a particular remediation process or technology or independently establishing a substantive standard of conduct.

The additional information may help distinguish between an insurer's completion of a corrective process and the resulting outcome for the affected consumer.

Thank you for the opportunity to provide comments and for your consideration.

From:

George (Gang) Zhang
ClearTaskAI LLC

Thank you for the chance to comment on version 6.0, exposed on October 16, 2026. We work on Open ADEP, an open profile for evidence of prior-authorization decisions. We ask for two clarifications in the next version. We do not ask the Working Group to adopt that profile.

Name these determinations among the examples. A prior-authorization, utilization-management, or level-of-care determination that grants or withholds a service is already capable of being a Direct Consumer Impact. The definition is an open list, and it reaches a system whose output materially affects a specific consumer. Exhibit A also asks whether systems in that row have a direct consumer impact.

The named examples are underwriting, premium, ratemaking, and claims denial and settlement. They do not name those three determinations, although Exhibit A lists the row. A prior-authorization denial is often made before a claim exists, so it is not the same event as a claims denial. Please add the three determinations to the examples in version 7.0, and keep the words “including but not limited to.”

A decision record is not a finding on clinical correctness or discrimination. A company can retain the recommendation that was signed, whether a human review was recorded, and which evidence classes were used, including classes present, redacted, or not captured. That record shows what was before the decision and what was signed. It does not establish that the determination was clinically appropriate, faithful to a guideline in substance, or free of unfair discrimination. Those questions remain with the company.

Please say this in the instructions for Exhibits B and C. An examiner may ask for the signed recommendation, the record of human review, and the evidence classes used. Producing those records does not, by itself, demonstrate clinical correctness or the absence of unfair discrimination. The questions on explainability, the adequacy of human review, testing, and proxy effects should remain questions to the company.

Thank you for considering this comment. I am glad to answer questions.

October 22, 2026

Via electronic mail
Miguel Romero
Director of Innovation, Cybersecurity and Technology
National Association of Insurance Commissioners
1100 Walnut Street, Suite 1000
Kansas City, MO 64106-2197
maromero@naic.org

Re: AI Risk Evaluation Supplement Version 6.0

Generalized Linear Models Used for Ratemaking and Risk Classification

Dear Mr. Romero and Members of the Big Data and Artificial Intelligence (H) Working Group:

The Insurtech Coalition appreciates the opportunity to comment on version 6.0 of the AI Risk Evaluation Supplement (Supplement). We respectfully request an exclusion throughout the Supplement, including Exhibit A, for generalized linear models (GLMs) used to develop rates, rating factors, risk classifications, or tier or company placement that are filed with, or subject to review by, a state insurance department under applicable rating laws.

Version 6.0 allows regulators to exclude ratemaking GLMs from inquiries under Exhibits B through D. This leaves a static actuarial tool within an AI governance framework unless a regulator determines otherwise. We recommend a clear scope exclusion that recognizes these models' actuarial purpose, operation and existing oversight. The exclusion would preserve existing rating oversight and leave other GLM uses subject to the Supplement where its definitions apply. Proposed revisions appear in Appendix A.

Traditional ratemaking GLMs require actuarial oversight

GLMs are established actuarial tools used in property and casualty pricing. Actuaries exercise professional judgment in selecting variables, distributions and link functions, evaluating model fit and developing rating factors. This work can be complex, but its complexity concerns actuarial assumptions, loss experience and risk classification. Unlike those systems that the AI Model Bulletin addresses, GLMs are long used, well-established actuarial tools that do not continue learning or independently change parameters during use.

Presently, data quality, model performance and unfair discrimination require actuarial review, appropriate controls and regulatory oversight. An additional AI governance framework requires insurers to fit those established disciplines into a second process without a demonstrated gap in existing protections. Those protections exist and have existed to prevent excessive, inadequate or unfairly discriminatory rates, regardless of the methodology used.

The NAIC's Regulatory Review of Predictive Models and Model Review Manual support review of the actuarial characteristics and risks of ratemaking models. They provide tools specifically suited to this work.

Actuarial standards add professional obligations concerning model risk, testing, validation and documentation. These include ASOP No. 56, Modeling; ASOP No. 12, Risk Classification; and ASOP No. 23, Data Quality, where applicable. Allowing regulators to accept previously submitted rate filing materials may reduce repeated submissions, but it does not resolve the mismatch between a static actuarial tool and AI governance.

Forcing actuarial ratemaking into AI governance creates unnecessary burdens

The burden comes from requiring insurers to fit an actuarial process into an AI governance program. The relevant expertise, validation methods and professional standards remain actuarial. Classifying these models as AI systems requires insurers to recast existing controls even though the models' operation and risks have not changed.

Including GLMs requires actuaries to participate in a parallel governance process and adapt their methods and documentation to its structure, in addition to meeting existing actuarial and rating requirements.

The NAIC's ORSA and Annual Financial Reporting model frameworks recognize proportionality. Governance should likewise fit a model's characteristics and existing oversight. A clear exclusion would avoid uncertainty over whether each regulator will require this parallel process. Materiality does not resolve the mismatch. A GLM can have significant consumer and financial effects while remaining a static actuarial tool. Version 6.0's inclusion of ratemaking within Direct Consumer Impact therefore supports addressing the issue through an express scope exclusion.

For small insurers, including many Coalition members, this draws directly on the actuaries responsible for maintaining rating plans. It consumes scarce internal time or requires outside support to repackage oversight of risks already addressed. The statement that the Supplement creates no new requirements does not eliminate that practical burden or establish a corresponding consumer protection benefit.

The proposed exclusion preserves review of other uses

The Working Group has raised potential uses of GLMs outside ratemaking. Our proposed exclusion applies only to the specified rating and risk-classification uses filed with, or subject to review by, the department under applicable rating laws. Other uses remain subject to the Supplement where its definitions apply. The Working Group's 2022–2023 homeowners survey grouped regression, static and pre-2000 models together rather than isolating GLMs. Its findings do not undermine this targeted exclusion, which leaves non-rating uses available for review where otherwise within the Supplement's scope.

Other commenters also support excluding conventional GLMs from the Supplement, recognizing their static operation, established oversight and different governance needs. We respectfully request adoption of the exclusion in Appendix A. It would preserve actuarial accountability and existing consumer protections while avoiding the cost of forcing static ratemaking tools into an AI governance process that does not fit their operation.

Thank you for considering our comments. We would welcome the opportunity to discuss the proposed language with you and the Working Group.

Sincerely,

Jennifer Crutchfield
Insurtech Coalition

Appendix A

Proposed Revisions to Version 6.0

The following revisions would implement the requested exclusion and align the related guidance, inventory heading and definition.

1 General Guidance on statistical and machine learning models

In the first paragraph under Machine Learning (ML) Models, revise the opening and closing sentences to read:

> "Statistical and machine learning models are widely used to estimate the relationship between a dependent variable and one or more independent variables."
>
> "Many statistical and machine learning techniques can be used for this purpose."

2 General Guidance on conventional GLMs

Replace the second paragraph under Machine Learning (ML) Models with:

> "Generalized linear models (GLMs) are a statistical technique in which the modeler specifies the variables, distribution and link function. In a conventional ratemaking application, the fitted model's parameters do not change during use unless the insurer revises the model. GLMs used in ratemaking are subject to oversight under applicable rating laws."

3 Ratemaking and Risk Classification Exclusion

Replace the third paragraph under Machine Learning (ML) Models with:

> "Ratemaking and Risk Classification Exclusion. This Supplement, including Exhibit A, does not apply to a Generalized Linear Model to the extent the model is used to develop rates, rating factors, risk classifications, or tier or company placement that are filed with, or subject to review by, the Department under applicable rating laws. A model is not an AI Model solely because iterative computational techniques were used to estimate its parameters. A Generalized Linear Model used for any other purpose remains subject to this Supplement to the extent it otherwise meets the definition of an AI System or AI Model."

4 Exhibit A column 7

Revise the Generalized Linear Models (GLMs) column heading to read:

> "Generalized Linear Models (GLMs), excluding GLMs within the Ratemaking and Risk Classification Exclusion"

5 Definition of Generalized Linear Model

Correct the phrase "extends linear regression to by relating predictive variables" to read "extends linear regression by relating predictive variables."

October 30, 2026

Commissioner Nathan Houdek, Chair
Commissioner Doug Ommen and Mary Block, Vice Chairs
NAIC Big Data and AI (H) Working Group
National Association of Insurance Commissioners
1100 Walnut Street, Suite 1000
Kansas City, MO 64106-2197

Submitted Electronically

Re: ACLI Comments on AI Risk Evaluation Supplement Version 6.0

Dear Chair Houdek and Vice Chairs Ommen and Block:

The American Council of Life Insurers (ACLI) appreciates the continued work of the NAIC Big Data and AI (H) Working Group in developing the AI Risk Evaluation Supplement (Supplement), and particularly the Working Group's continued engagement with stakeholders as the Supplement approaches completion. Version 6.0 reflects substantial progress and addresses many of the recommendations in ACLI's September 29 comments.

At this stage in the process, ACLI offers a limited number of targeted comments intended to assist the Working Group in resolving remaining implementation issues and producing a final Supplement that is practical, appropriately risk-based, and useful to regulators and insurers alike.

Inherent and Residual Risk
ACLI appreciates Version 6.0's additional discussion of risk assessment and its reliance on the Financial Condition Examiners Handbook framework. The draft, however, focuses its risk-assessment instructions on inherent risk, while separately defining residual risk without explaining how or where residual risk should inform responses or regulatory decisions.

We recommend clarifying that, where risk ratings are used to determine whether further inquiry is warranted, inherent risk should be considered together with relevant controls and residual risk. An elevated inherent-risk rating does not necessarily reflect the risk presented by a system after effective governance, testing, human oversight, monitoring, and other mitigation measures are considered. This clarification would also better align the Supplement with the risk-focused framework incorporated from existing examination guidance.

Traditional Fixed and Transparent Predictive Models
ACLI appreciates the new language recognizing that regulators may exclude some or all GLMs from Exhibits B through D based on existing regulatory oversight through ratemaking. To build on this meaningful progress, we recommend that this not focus solely on GLMs, but more generally on the

The American Council of Life Insurers (ACLI) is the leading trade association driving public policy and advocacy on behalf of the life insurance industry. 90 million American families rely on the life insurance industry for financial protection and retirement security. ACLI's member companies are dedicated to protecting consumers' financial wellbeing through life insurance, annuities, retirement plans, long-term care insurance, disability income insurance, reinsurance, and dental, vision and other supplemental benefits. ACLI's 275 member companies represent 94 percent of industry assets in the United States.

characteristics of the deployed model. Traditional predictive models that result in static formulas or rules do not learn or autonomously change after deployment and present different governance considerations from adaptive or autonomous AI systems.

Accordingly, ACLI suggests refining this guidance on page 3, and in the instructions for Exhibits A through D, as follows:

"Regulators should exclude GLMs and other predictive models that result in the use of rule-based systems, deterministic algorithms, fixed mathematical formulas, and transparent predictive models from the Supplement, especially those which receive other regulatory oversight such as from rate filings."

This approach would preserve appropriate regulatory flexibility while reducing duplicative review of longstanding, transparent models already subject to established regulatory processes.

Exhibit C Question 10a: Proxy Relationships

ACLI recommends deleting new Exhibit C Question 10a concerning evaluation for 'potential proxy relationships or effects.' The term 'proxy relationship' is not defined, and the question could be interpreted as establishing or presupposing a testing obligation independent of otherwise applicable insurance laws and regulatory requirements. Version 6.0 appropriately states elsewhere that the Supplement does not itself create new substantive legal duties or standards of conduct. Question 10a should be consistent with that principle.

Coordination and Reliance on Prior Regulatory Review

ACLI appreciates Version 6.0's strengthened coordination provisions and its recognition that regulators may accept current and applicable information previously submitted to the requesting department or another state department of insurance.

As a final implementation refinement, we encourage the Working Group to strengthen reliance on work already completed by an insurer's domestic regulator. Where substantially the same AI Systems and issues have already been reviewed, other departments should rely on those submissions and regulatory work to the greatest extent practicable. Supplemental requests should generally focus on material changes, information not previously reviewed, or matters specific to the requesting jurisdiction or examination objective.

This approach would reduce duplicative work for regulators and insurers while preserving each department's ability to obtain information necessary for their examination purpose.

Company Contacts and Subject Matter Experts

ACLI appreciates the importance of ensuring that regulators can efficiently reach appropriate personnel for follow-up questions. However, we recommend that Exhibit B provide companies with flexibility to identify an appropriate contact, function, or team to coordinate those inquiries rather than requiring a detailed list of individual subject matter experts. This approach would ensure that regulators connect with the appropriate expertise while allowing companies to manage their internal resources and determine the appropriate personnel to respond based on the nature of a particular inquiry. As such, we recommend the following revised language for Exhibit B:

"For company contact information, please identify an appropriate contact, function, or team that can coordinate responses to follow-up regulatory inquiries and connect regulators with relevant subject matter experts, as appropriate."

Existing Company Inventories and Exhibit A

Finally, ACLI recommends resolving an inconsistency in Exhibit A. The company instructions currently indicate that a company with an existing AI System or Model inventory should provide that inventory together with a completed Exhibit A, while the regulator instructions appropriately recognize that the information in an existing inventory may make a separate Exhibit A submission unnecessary.

To align these provisions and avoid unnecessary duplication, ACLI recommends revising the Company Instructions in Exhibit A to state:

"If a company has an existing AI System and/or AI Model inventory, the company should first provide that inventory. Depending on the information provided, the regulator may determine that completion of some or all of Exhibit A is unnecessary."

This revision would allow existing company documentation to serve its intended purpose while preserving the regulator's ability to request any additional information needed for the review.

Conclusion

ACLI appreciates the significant progress reflected in Version 6.0 and the Working Group's sustained effort to incorporate feedback from regulators, industry, consumer representatives, and other interested parties. The changes made throughout this process have materially improved the Supplement.

As the Working Group moves toward a final version, our remaining comments are intentionally limited and focused on practical implementation. We believe these targeted clarifications will further strengthen the Supplement, reduce unnecessary duplication, and promote a consistent, risk-based examination process while preserving regulators' ability to obtain information relevant to their oversight responsibilities.

ACLI and its members remain committed to working collaboratively toward a final Supplement that is useful and workable for all parties. We greatly appreciate the opportunity to provide these comments.

Respectfully submitted,



Jared Kosky
Assistant Vice President & Counsel
American Council of Life Insurers
jaredkosky@acli.com

cc: Miguel Romero, NAIC
Scott Sobel, NAIC
Dorothy Andrews, NAIC

Insurance Services Office, Inc.
545 Washington Boulevard
Jersey City, NJ 07310-1686
www.iso.com

Laura L. Panesso, CPCU
VP, State Relations and
Regulatory Reporting
Government Relations
T: 201 469 2287
LPanesso@iso.com

October 30, 2026

Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
1100 Walnut Street, Suite 1000
Kansas City, MO 64106-2197

Re: Version 6.0 of the AI Risk Evaluation Supplement

Thank you for the opportunity for Verisk to provide comments on Version 6.0 of the AI Risk Evaluation Supplement ("Supplement Version 6"). To best understand our perspective, please note that Verisk and its subsidiaries provide information (including statistics, underwriting and claims information), actuarial analyses, policy language, and consulting and technical services in connection with multiple lines of property/casualty insurance, as well as catastrophe models and information about specific properties. Our customers include insurers and reinsurers, as well as agents, brokers, self-insureds, risk managers, financial services firms, regulators, and various government and law enforcement agencies.

The proposed language contained in Supplement Version 6 would benefit from additional clarification regarding its scope. To provide that clarification and address some of the concerns identified by stakeholders previously, we propose the following changes:

Machine Learning (ML) Models:

We appreciate regulators' efforts to clarify the scope of the AI supplement with respect to generalized linear models (GLMs). While there was support for completely removing GLMs from the supplement, we believe the recent revisions represent a positive step in the right direction, although further clarification remains necessary. In the corresponding redline document, we suggest adding that "this supplement asks about a company's use of specific models, *when utilized within an AI System,* in Exhibit A to give regulators flexibility in determining their scope of interest." Our proposed edits help clarify that the supplement is intended to focus on how specific models are utilized within an AI System, rather than treating every individual modeling technique as the primary focus. We appreciate regulators' consideration of these concerns and believe these changes help better align the supplement's scope with its intended purpose while preserving regulatory flexibility to identify models of interest.

We also suggest clarifying the AI System/AI Model discussion and Exhibit A so that scope is not determined solely by whether a model uses a GLM. The focus should be on how the model is used and whether it falls within the Supplement's definitions.

Thank you again for the opportunity to provide comments on Version 6.0 of the AI Risk Evaluation Supplement. We appreciate the Working Group's efforts and look forward to ongoing collaboration. Please feel free to contact me should you require additional information concerning our position relative to these matters.

Sincerely,



Laura L. Panesso, CPCU
VP, State Relations and Regulatory Reporting

Artificial Intelligence (AI) Risk Evaluation Supplement

Background, Intent, and Instructions

Background:

The rapid expansion of big data and the adoption of artificial intelligence and machine learning are significantly transforming insurance practices. These technologies can offer substantial benefits to both insurance companies and consumers by facilitating the development of innovative products, improving customer interface and enhancing service, simplifying and automating processes, and promoting efficiency and accuracy. However, without robust governance and effective controls, the use of *AI Systems* ~~may lead~~carries a risk that could lead to Adverse Consumer ~~Outcomes~~ Impacts or compromise the financial soundness of an insurance company. Insurers are responsible for managing the risks associated with the development and implementation of *AI Systems* and must demonstrate to regulators that appropriate risk-based oversight mechanisms are in place and are functioning effectively. Ultimately, insurers that make decisions or perform actions impacting consumers or financial soundness that are made or supported by advanced analytical and computational technologies, including from *AI Systems*, must comply with all applicable insurance laws and regulations.

Intent:

The NAIC's Innovation, Cybersecurity and Technology (H) Committee charged the Big Data and AI Working Group (BDAIWG) to create ~~tool(s)~~resources that would assist regulators to identify and assess both financial and consumer risks from the use of *AI Systems*~~. on an on-going basis~~.

The following exhibits are designed to supplement existing market conduct, financial analysis, and/or financial examination review procedures for reviewing *AI Systems*. As this supplements existing NAIC resources, regulators should continue to consider existing NAIC resources as authoritative but may consider drawing from this ~~supplement~~ Supplement to assist in understanding and assessing a company's use of *AI Systems*. Inquiries and information requests performed related to this ~~supplement~~ Supplement will be coordinated consistent with the guidance provided by the *Market Regulation Handbook (MRH)*, *Financial Condition Examiners Handbook (FCEH)*, and the *Financial Analysis Handbook (FAH)*. Consistent with the Handbooks, this Supplement does not independently create substantive legal duties, standards of conduct, admissions, or expanded regulatory authority.

Moreover, consistent with the coordination practices described in the FCEH, MRH, and FAH, information requests made leveraging the Supplement should be coordinated as much as possible.

With the intention of the document as a supplement acknowledged, there may be some important context that is not specified in this document, particularly related to deciding which companies are in scope of inquiry. This is to avoid the perception of conflicting guidance as many important decisions leading to the full use of this Supplement would be governed by the aforementioned Handbooks.

Commented [MR1]: Edit to address input by Consumer Representatives.

Commented [MR2]: APCIA suggested change.

Commented [MR3]: Added acronyms to address input from Connecticut.

Commented [MR4]: ACLI input.

Commented [MR5]: APCIA and Monitaur related input.

DRAFT

These optional exhibits allow regulators to determine the extent of *AI Systems* usage for an insurance company and whether additional analysis is needed focusing on financial and consumer risk.

This Supplement consists of the following exhibit questionnaires:

- **Exhibit A: Quantify Regulated Entity's Use of AI Systems**
- **Exhibit B: AIS Program (Two Options: Narrative or Checklist)**
- **Exhibit C: High-Risk AI Models**
- **Exhibit D: AI Model Data**

Instructions:

Information obtained from the responses to the following Exhibits may supplement guidance and ~~tools~~ resources used during an existing market conduct, certification, financial analysis, and/or financial examination review, to enhance the regulator's understanding of the extent of *AI Systems* utilization and assessment of potential risk to consumers and the insurance company. Effective assessment requires regulators to maintain a fluent understanding of the applicable laws and regulations, including but not limited to those pertaining to unfair trade practices, unfair claims settlement practices, unfair discrimination protections, corporate governance annual disclosures, confidentiality, privacy protections, financial reporting, and rating.

This Supplement is designed with the intention that regulators using this document will most often use Exhibit A. Based on responses received, regulators may thereafter leverage Exhibit B to inquire about a company's governance practices. Finally, for a subset of a company's use of AI, companies may receive Exhibit C and Exhibit D related inquiries where regulators wish to probe further.

Therefore, ~~R~~regulators using this Supplement may wish to first use Exhibit A to determine if further inquiry using Exhibits B-D is necessary. It is possible that while the company responding indicates usage of *AI Systems*, its use of AI is limited or low in inherent risk such that further inquiry as contemplated by subsequent exhibits is not warranted. For instance, if a regulator is using the Supplement to inquire about a company's claims practices and based on responses to Exhibit A, the regulator determines that the company's use of AI is out of scope of a claims-related targeted exam, the regulator may determine it is appropriate to thereafter rely on the ~~Market Regulation Handbook~~MRH to guide the remaining work.

If information requested through this Supplement has already been provided to this department or any other state department of insurance, the company's response should so state, in which case the regulators may accept prior submissions if the prior response is still current and applicable. Where companies have existing documentation that may satisfy regulatory inquiries, regulators may consider accepting company documentation consistent with guidance described in the FCEH, MRH, and FAH.

Commented [MR6]: Edits to this statement are related to input provided by the consumer representatives.

Formatted: Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border)

Commented [MR7]: ACLI related input.

Commented [MR8]: APCIA, BCBSA, InsurTech Advisors, and Monitaur related input.

DRAFT

The responses to these exhibits will be considered by regulators when identifying the inherent risks of the insurer. They should also affect the planned examination or inquiry approach, as well as the nature, timing and extent of any further procedures performed.

General Guidance

Machine Learning (ML) Models

Machine learning models are widely used to estimate the relationship between a dependent variable and one or more independent variables. For example, an insurer may want to estimate how a driver's age relates to the probability of being involved in a car accident. There are many ML models that can be used for this purpose.

In particular, generalized linear models (GLMs) are one such technique. As with other machine learning techniques, GLMs are not without risk of causing unfair discrimination or other ~~adverse consumer outcomes~~Adverse Consumer Impacts and require effective governance over data quality, model development, validation, implementati~~o~~n, ongoing performance monitoring, and compliance with applicable laws and regulations. This supplement asks about a company's use of specific models, when utiliz~~i~~ed within an AI System, in Exhibit A to give regulators flexibility in determining their scope of interest.

Based on existing regulatory oversight provided via rate making, regulators may choose to exclude some or all GLMs for Exhibit B through D related inquiries if appropriate.

AI System vs AI Model

Definitions are provided for both terms in the Definitions and Appendix section of the Supplement. Generally, this Supplement is designed with a focus on AI Systems. In relation to each other, this Supplement and each of the terms are defined such that an AI System is viewed as the primary unit at which governance and reporting discussions will take place; a model is a component of an AI system; and a stand-alone model may be examined separately when it independently creates direct consumer impact or material financial risk.

However, there may be times where inquiries are directed at an AI System and are better addressed at the AI Model level or vice versa. In those instances, companies may engage with department officials to ensure information provided meets expectations and assists in accomplishing regulatory objectives.

Commented [MR9]: Extensive feedback was provided by various parties on these two terms. Leveraging input from Monitaur, and in lieu of tracking many additional edits, this new overarching guidance was added. This input was also raised by NAMIC and other parties (APCIA and ACLI) had general input this may help address.

Materiality

This Supplement was designed to allow regulators the flexibility to either specify materiality or rely on the company's *materiality* calculations as companies respond to Supplement related inquiries.

The concept of *materiality* leverages the definition otherwise used to drive financial exam-related decisions and as defined in the FCEH. For purposes of completing responses related to this Supplement, regulators

may opt to rely on a company's internal materiality thresholds provided the information supporting a company's *materiality*, including an explanation for why that threshold was considered appropriate.

However, nothing in the guidance provided within this Supplement should be construed to supersede FCEH guidance.

Risk Assessment

This Supplement also, at times, relies on a company's Risk Assessment. The concept of Risk Assessment centers around the discussion of *Inherent Risk* which leverages the definition reflected in the FCEH. For purposes of completing responses to this Supplement, regulators ask that Risk Assessment information provided to regulators include explicit indication of *Inherent Risk* Assessments (i.e., risk before the consideration of controls or mitigation measures).

Consistent with the discussion of *materiality*, nothing in the guidance provided within this Supplement should be construed to supersede FCEH guidance as it relates to discussions of risk assessment.

Confidentiality

The information requested via these Exhibits will be held confidential, consistent with the requesting state's relevant authority.

Regulators using any of these exhibits should cite examination or other authority as appropriate when requesting information from insurers. Regulators should cite all relevant confidentiality statutes or other specific protections related to documents, materials or other information in the possession or control of regulators that are obtained by or disclosed to the regulators or any other person in the course of a market conduct inquiry and all information reported or provided to the regulator pursuant to cited examination or other authority.

DRAFT

Exhibit A: Part One - Quantify Regulated Entity's Use of AI Systems

Purpose: To obtain information pertaining to the number of *AI Systems* and *AI Models* used by a company to help facilitate risk assessment. Based on the responses from the company, regulators may ask for additional information related to their *AIS Program* (Exhibits B), high-risk models (Exhibit C), and the types of data (Exhibit D~~).) where there is risk for *Adverse Consumer Outcomes* or material adverse financial impact.~~

Company Instructions:

- Provide the most current counts and use cases of the following, as requested.
- ~~For purposes of responding to information requests related to this Exhibit, models that *augment* or *automate* decision-making related to consumers are considered to have direct consumer impact.~~
- ~~*"AI System"*, *"Adverse Consumer Outcome"*, and *"Use Case"* are defined in the "Definitions and Appendix" section of this Supplement.~~
- Include all lines of business. If the governance differs by company, line of business, or state, work with your regulator to determine if multiple submissions are needed.
- If a company has an existing *AI ~~Systems~~* System and/or *Model* inventory, ~~it may suggest submission of the inventory in lieu of completing Exhibit A~~the company should provide the AI System and/or Model inventory together with the completed Exhibit A. Regulators will look for columns with information on:
    - AI System and/or Model Name
    - Description of how the AI System and/or Model is used and the broader AI system in which it is contained
    - Operation or Program Area(s) the AI System and/or Model is Used Within
    - Inherent Risk Level Based Upon the Company's Risk Assessment
    - What direct consumer impact does this AI System and/or Model have?
    - What financial impact does this AI System and/or Model have?
- If a company has questions about any of the fields, please submit them in writing to <regulator contact email address>.

Regulator Instructions:

- Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.
- Regulators should either specify a materiality threshold to guide company responses to the Supplement or request that the company provide a threshold that it used in completing this Exhibit.
- Depending on the information provided in a company's AI System and/or Model inventory, regulators may consider if a separate Exhibit A submission is necessary.
- For column 3, regulators should specify the number of months.

Formatted

DRAFT

Company Legal Name and Group Name: _____

NAIC CoCode and Group Code: _____

Company Contact Name: _____ Email: _____

Line(s) of Business for Which This Response Applies: _____

Materiality Threshold <regulator to provide or specify if the company will provide>: _____

Date Form Completed: _____

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	(4) Use Case *Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact*	Do any of the AI Systems utilize any of the following (check all that apply)			
				(5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Marketing							
Premium Quotes & Discounts							
Underwriting/ Eligibility							
Ratemaking/ Rate Classification/ Schedule Rating/ Premium Audits							

Formatted Table

Commented [SS10]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

DRAFT

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	(4) Use Case *Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact*	Do any of the AI Systems utilize any of the following (check all that apply)			
				(5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Claims/ Adjudication*							
Customer Service							
Utilization Management/ Utilization Review/ Prior Authorization/ Level of Care Determination							
Fraud/Waste & Abuse							
Investment/ Capital Management							
Reserves/ Valuations							
Catastrophe Triage							
Reinsurance							
Other Insurance							

DRAFT

Formatted Table

Commented [SS10]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	(4) Use Case *Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact*	Do any of the AI Systems utilize any of the following (check all that apply)			
				(5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Practices** (if applicable)							
** Includes Salvage/Subrogation, adjuster assignments, etc.*							
*** Any additional operation or program area with the potential to result in either Direct Consumer Impact or Material Financial Impact.*							

Formatted Table

Commented [SS10]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

DRAFT

Exhibit A: Part Two – AI Systems Request List

In Support of the regulatory inquiries related to this Supplement, the regulators request information or documentation related to the following:

Regulator Instructions: Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.

1. Materiality – Explanation related to the company's *materiality* calculation (if applicable) used in responding to inquiries related to this Supplement.
2. Risk Assessment – Documentation explaining the company's risk assessment process for purposes of its use of *AI Systems*.
3. Model Inventory – A listing of the company's models, that specifies risk for the following program areas <regulator to specify program areas in scope of the request>. Additionally, regulators request the model inventory or accompanying documentation include:
    a. *AI Model* Name
    b. Description of how the *AI Model* is used
    c. Description of the broader *Use Case* in which each *AI Model* is used
    d. Operation or Program Area(s) the *AI Model* is Used Within
    e. *Inherent Risk* Level Based Upon the Company's Risk Assessment
    f. What consumer impact does this *AI Model* have?
    g. What financial impact does this *AI Model* have?
g. Indicate whether each use case involves Agentic AI? If yes, is that AI System authorized to execute one or more actions that can directly create or materially alter a consumer, financial, or financial-reporting outcome without human approval before the action?

Formatted

Commented [SS11]: Suggestion from Ali Houshangi

DRAFT

Exhibit B: (Narrative) AIS Program

Purpose: To obtain the Company's *AIS Program*, including how the company handles governance, risk identification, mitigation, and management framework and internal controls, and third-party *AI Systems* and data for *AI Systems*; and the process for acquiring, using, or relying on third-party *AI Systems* and data. Market and financial regulators should coordinate to gain access to the relevant section of the policies governing the use of *AI Systems*.

Company Instructions: Below are both Checklist and Narrative sections of Exhibit B of the NAIC's AI Risk Evaluation Supplement. If the responses to the ~~checklist~~ narrative indicate that an explanation of a policy, procedure or process can be found on a given page of a document, indicate the page number and the document title and include the document referenced in the submission. If other sections of the document are irrelevant, the relevant sections of the document may be submitted in lieu of the entire document. Please see definitions at the end of this document.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. The Exhibits are designed so that regulators would use either the Narrative or the Checklist Exhibit but not both.

The references to, and questions about, elements of a company's *AIS Program* are not intended to be interpreted as creating new requirements for *AIS Program*.

Group and Company Legal Name: _____

NAIC Group and Company CoCodes: _____

Company Contact Name: _____ Email: _____

Date Form Completed: _____

1. Provide your *AIS Program*. Click or tap here to enter text.
    a. What role maintains the *AIS Program*? Click or tap here to enter text.
    b. Discuss the governance structure, Board reporting and frequency. Click or tap here to enter text.
    c. Discuss the process by which the *AIS Program* is integrated throughout the organization, assessed and remediated. Click or tap here to enter text.
    d. Discuss the process by which the effectiveness of the *AIS Program*, *AI Systems*, and individual models are assessed and modified. Click or tap here to enter text.

Commented [MR12]: Pending - CAI says there are inconsistencies between the two versions of Exhibit B.

Commented [SS13R12]: Changes were made to the two versions of Exhibit B for consistency based on feedback from CAI

Commented [MR14]: Edit to address input by ACLI and NAMIC.

Formatted: Indent: Left: 0.5", No bullets or numbering

Commented [MR15]: Based on input from Monitaur

e. Discuss how responsibility for governance within the organization is assigned and how the organization ensures consistency and alignment. Please include a detailed list of subject matter experts who will be able to answer follow up regulatory inquiries. Click or tap here to enter text.

f. Discuss the integration of the *AI Systems* in the Own Risk and Solvency Assessment (ORSA) and Enterprise Risk Management (ERM) assessments. Click or tap here to enter text.

g. How does the insurance company assess autonomy, reversibility, and reporting impact risk of *AI Systems*? Click or tap here to enter text.

~~g.~~h. Discuss the company's materiality threshold and risk assessment process in evaluating its use of AI. Click or tap here to enter text.

2. Discuss the uses of *AI Systems* that:
   a. Generate a *material financial impact*. Click or tap here to enter text.
   b. Generate a *direct consumer impact*. The company's response should include, but not be limited to, whether there is a human in the loop, whether there are high volume transactions, how the company assures itself the human review is appropriate, and how any exceptions or errors are found and resolved~~, and whether there were any staffing reductions as part of the *AI System* implementation~~. Click or tap here to enter text.
   c. Generate or impact material information reported in financial statements. Click or tap here to enter text.
   d. Generate or impact risk or control assessment. Click or tap here to enter text.
   e. Discuss the development, testing, and implementation of material *AI Systems* that the Company has implemented. If appropriate, include details regarding where any systems differ from established IT systems and data handling protocols. Discuss the basis for deviation from established practices. Click or tap here to enter text.

3. Discuss how the company:
   a. Assesses, mitigates, and evaluates inherent risks related to the use of AI Systems including the possibility of the company engaging in unfair trade practices or that it may violate non-discrimination protections.
   b. Considers *AI System* risks within the software development lifecycle (SDLC) or within the broader *AIS Program*

~~3.~~4. Provide the policy for, and discuss the use and oversight of, material *AI System* vendors, model design and testing:
   a. Discuss the validation and testing procedures performed on internally-developed *AI Systems*, including whether validation and testing are performed by someone that is independent from development. Click or tap here to enter text.
   b. Discuss the validation and testing procedures performed on third-party vendor-supplied *AI Systems*. Click or tap here to enter text.
   c. Discuss the testing and verification that has occurred including frequency, scope and methodology. Click or tap here to enter text.

Commented [MR16]: Carrying over Exhibit A part 2

Commented [MR17]: Edit to address input by consumer representatives.

Commented [MR18]: ACLI and BCBSA related input.

Formatted: Indent: Left: 1", No bullets or numbering

Commented [MR19]: Edits to create more alignment between the two versions of Exhibit B.

DRAFT

d. Discuss how the company ensures explainability and transparency into any *AI models* recommending or making decisions. Click or tap here to enter text.

~~4.~~5. Provide the policy for, and discuss the use and oversight of, material *AI Systems* by third-party providers including actuarial, claim, MGA, audit, and/or other professional services. Click or tap here to enter text.

a. Discuss the testing and verification that has occurred, frequency, scope, and methodology. Click or tap here to enter text.

~~5.~~6. Discuss additional aspects of the framework design and evaluation pertaining to *AI Systems*. Click or tap here to enter text.

~~a.~~ Discuss the unit(s) responsible for the *AIS Program*, assessment approach and frequency, and involvement with the program area to the extent it differs from that discussed previously. Click or tap here to enter text.

7. Discuss the criteria for identifying a material change and for triggering approval, revalidation, or enhanced monitoring, including predeployment testing~~, periodic or event-triggered reassessment, changes in operational outcomes, incidents and near misses, drift or degradation, third-party changes, and safe retirement or fallback.~~

8. ~~For generative AI, h~~How does your company:
   a. Discover unsanctioned AI use
   b. Prevent nonpublic personal information from entering unapproved AI ~~tools~~models
   c. Govern enterprise assistants that have AI capabilities and support many low- and medium-risk uses
   d. Decide when a general-purpose tool with AI capabilities becomes a high-risk use because of the workflow in which it is deployed.

Formatted: Indent: Left: 1", No bullets or numbering

Commented [SS20]: Based on feedback from Legion

Formatted: Indent: Left: 0.5", No bullets or numbering

Formatted

Commented [SS21]: Based on suggestion from InsurTech Advisors to address enterprise generative AI and shadow AI as operating risks

DRAFT

Exhibit B: (Checklist) AIS Program

Purpose: To obtain information about the Company's *AIS Program* including information about how the company handles governance, risk management and internal controls, and third-party *AI Systems* and data for *AI Systems*; and the process for acquiring, using, or relying on third-party *AI Systems* and data. The references to, and questions about, elements of a company's *AI Systems* Governance Risk Assessments are not intended to be interpreted as creating new requirements for *AI Systems* Governance Risk Assessments.

Company Instructions: Provide responses to the questions regarding how the governance of *AI Systems* fits within your company's system of supervision or Enterprise Risk Management program. When referencing documents that are responsive to this exhibit, please include the documents referenced in the submission. Include all companies and lines of business. If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below.

For company contact information, please also include a detailed list of subject matter experts who will be able to answer follow-up regulatory inquiries.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. The Exhibits are designed so that regulators would use either the Narrative or the Checklist Exhibit but not both.

Group and Company Legal Name(s): _____

NAIC Group and Company Code(s): _____

Company Contact Name: _____ Email: _____

Date Form Completed: _____

Ref	AI Systems Use Questions for Company	Company Response
1	Has the company adopted a written *AIS Program*? If yes, when was it adopted, what does it govern, who approved the program, and what is the frequency of review for updating?	

Commented [MR22]: Edit to address input by ACLI and NAMIC.

Commented [MR23]: Edit to address input by the Center for AI Oversight.

DRAFT

2	Was the Board of Directors or management involved in the adoption of an *AIS Program*?		
	2a. What is the role of the Board of Directors or management in the *AIS Program*?		
	2b. How is accountability assigned within the AIS Program and what training has been provided for those charged with oversight of the AIS Program?		
	2c. Discuss the company's materiality threshold and risk assessment process in evaluating its use of AI.		
Ref	**AI Systems Use Questions for Company**		**Company Response**
3	**Reference the processes and procedures of the Company *AIS Program* that addresses the following:**		
	How the Insurance Company…	**Document Name and Page #**	**If not specified in the *AIS Program*, provide details below:**
	3a. Considers *AI System* risks within its Enterprise Risk Management (ERM)		
	3b. Considers *AI System* risks within the Own Risk and Solvency Assessment Report (ORSA), as applicable.		
	3~~a~~c. Assesses, mitigates, and evaluates *inherent risks* related to the use of *AI System* including the possibility of the company engaging in unfair trade practices		
	3d. Determined "*materiality*" and what threshold was used for assessing *AI Systems* risk and determined the company's risk assessment processes for its use of AI Systems.		
	3e. Quantifies *AI System* risk levels		
	3f. Considers *AI System* risk impact on financial reporting		

Commented [MR24]: Edit to address input by Center for AI Oversight.

Commented [MR25]: Carrying over Exhibit A part 2

Commented [SS26]: Suggestion from Yash Lundia: Also ask for version # and date of AIS Program document

Commented [MR27]: To carry over former Exhibit A part 2

DRAFT

	3~~b~~g. Ensures *AI Systems* are compliant with applicable state and federal laws and regulations, including non-discrimination protections		
	3~~c~~h. Evaluates the risk of *Adverse Consumer ~~Outcomes~~Impact*		
	3~~d~~i. Considers data privacy and protection of consumer data used in *AI Systems*		
	3j. Considers consumer complaints resulting from *AI Systems* and whether they are identified, tracked, and addressed		
	3k. Promotes consumer awareness of the use of *AI Systems* through disclosures, policies, and procedures for consumer notification, as appropriate		
	3l. Trains employees about *AI System* use, ensures human oversight is appropriate, and defines prohibited practices (if any)		
	3m. Provides standards and guidance for procuring and engaging *AI System* vendors		
	3n. Governs, monitors, tests and ensures transparency of *AI Systems* developed by vendors.		
	3o. Considers *AI System* risks within the software development lifecycle (SDLC) or within the broader *AIS Program*		
	3~~e~~p. Evaluates whether *AI Systems* are suitable for their intended use and should continue to be used as designed		

Commented [MR28]: Edit to address input by Consumer Representatives

DRAFT

Exhibit C: High-Risk AI Models

Purpose: To obtain detailed information on in-production high-risk *AI models*, such as models making automated decisions that could cause ~~*Adverse Consumer Outcomes,*~~ *material financial impact*~~,~~ or direct consumer impact ~~material consumer impact, or material financial reporting impact~~. *AI model* risk criteria are set by the insurance company. Note that *materiality* should be a consideration in determining the level of risk. ~~To assist in identifying models for which this information is requested, regulators may request information on the company's risk assessment and a model inventory if such information has not otherwise already been provided.~~

Company Instructions: Fill in the details for each of the *AI model(s)* requested. ~~Include all companies and lines of business.~~ If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below. The template below refers to both *AI Systems* and *AI Models* depending on the information being requested. There may be some instances where a company feels information should be provided in relation to the *AI System* and not the *AI Model* or vice-versa. This should be discussed with regulators as part of the submission process to avoid misunderstanding.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam.

Group and Company Legal Name(s): _____

NAIC Group and Company Code(s): _____

Company Contact Name: _____ Email: _____

Provide who prepared and who reviewed the responses, by function. _____

Date Form Completed: _____

Ref	AI Model Information Requests	Company Response
1	*AI model* name and version number.	
2	Use cases and purpose(s) of model	
3	Model type used in the *AI System* (*GLM, generative/agentic AI, or other AI/ML algorithm*)	

Commented [MR29]: Suggestion from Consumer Representatives, ACLI, and NAMIC.

Commented [SS30]: Suggested by Jeffrey Sharer

DRAFT

3a	Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval.	
4	Model Implementation Date	
5	Model development (internal, ~~"owned" third party,~~ or ~~true~~ third party, or a mix of both – include vendor name if applicable)	
6	Did the Company purchase the *AI model*, was the *AI model* developed specifically for the Company? Is the Company able to modify the *AI model* as necessary?	
~~7~~	~~Model risk classification (high, moderate, low, etc.)~~	
~~8~~7	Model risk(s) (i.e., describe the potential for *adverse consumer impact*)	
~~9~~8	Model limitations (i.e., limitations on *uses cases* or data inputs outside the domain of the training data, or other guardrails to manage model related risk)	
~~10~~9	*Does the AI model* ~~type (~~automate, augment, or support decision-making?~~)~~	
~~11~~10	Discuss testing and validation of model outputs, as applicable, ~~(e.g.~~ testing for model drift, accuracy, unfair trade practices, unfair discrimination, performance degradation, and other potential Adverse Consumer Impacts, ~~etc.).~~ Discuss validation performed prior to deployment and ongoing monitoring after deployment, including testing following material changes to the model, data, or use of the model. Where testing identifies material differences or potential Adverse Consumer Impacts, discuss the company's process for evaluating the results, identifying contributing factors, and determining whether changes to the model, data, methodology, controls, or other reasonably available alternatives are appropriate. ~~and how the model was validated prior to being deployed as well as how its performance is monitored on an ongoing basis.~~	

Commented [SS31]: Suggestion from Ali Houshangi

Commented [MR32]: To address input provided by NAMIC.

Commented [MR33]: Edit to address input by NAMIC and other parties.

Formatted: English (United States)

Commented [MR34]: Edits to incorporate CT's input.

DRAFT

~~10~~1~~a~~	Where applicable, discuss whether the company evaluates input variables, derived variables, or model outputs for potential proxy relationships or effects that may contribute to Adverse Consumer Impacts or unfair discrimination, including the methodology used and any resulting action taken.	
~~12~~11	Last date of model testing or validation. Please include the results and corrective actions as a result of the last model testing, if applicable.	
~~13~~12	Discuss how the model affects the financial statements, risk assessment or controls.	
~~14~~13	Discuss how the model is reviewed for compliance with applicable state and federal laws, including but not limited to the unfair trade practices act and unfair claims settlement laws.	
~~15~~14	To the extent permitted by law, discuss if the company has had any actions taken against them for use of this model. Actions may include, but are not limited to, informal agreements, voluntary compliance plans, administrative complaints, ongoing third-party monitoring, cease and desist, remediation, restitution, fines, penalties, investigations, consent orders or other regulatory agency actions.	

Commented [MR35]: New question proposed by CT.

Commented [SS36]: Suggestion from Yash Lundia: Connect testing with the resulting decision/action

DRAFT

DRAFT

Exhibit D: AI Model Data

Purpose: To obtain detailed information of the source(s) and type(s) of data used directly ~~in~~by an *AI model(s)* to identify risk of *adverse consumer impact*, unfair trade practices, material financial impact, or material financial reporting impact.

Company Instructions: Provide details below for the data used in *AI model(s)*. If any of the data elements listed are used in the training or test data as part of the development of an *AI model*, provide information on whether the data element is sourced internally or whether the data element is sourced from a third party, in which case provide the name of the third-party vendor. Leave blank if a data source is not used in the development of *AI model* for the insurance operation. Include all companies and lines of business. If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. Regulators may also wish to customize this supplement for the line of business as some data types may or may not be pertinent based on the responding company's lines of business. Depending on responses received, a regulator may wish to ask for a data dictionary to gain a more complete understanding of a data used for a given model.

Group and Company Legal Name(s): _____

NAIC Group and Company CoCode(s): _____

Company Contact Name: _____ Email: _____

List the *AI System* or *AI models* which rely on the data described in response to this submission:

_____

Describe the Line of Business for Which This Response Applies (complete one for each line of business):

_____

Date Form Completed: _____

Commented [MR37]: APCIA related input

DRAFT

DRAFT

Ref	(1) Type of Data Element Used in AI Model	(2) Describe how the AI Model uses this data including how the output is used within operations; or input data is used for training and testing	(3) Internal Data Source	(4) Third Party Data Source / Vendor Name
1	Aerial Imagery			
2	Age, Gender, Ethnicity/Race			
3	Consumer or Other Type of Insurance/Risk Score			
4	Crime Statistics			
5	Criminal Convictions (Exclude Auto-Related Convictions)			
6	Driving Behavior			
7	Education Level (Including school aptitude scores, etc.)			
8	Facial or Body Detection / Recognition / Analysis			
9	Geocoding (including address, city, county, state, ZIP code, lat/long, MSA/CSA, etc.)			
10	Geo-Demographics (including ZIP/county-based demographic characteristics)			
11	Household Composition			
12	Image/video Analysis			
13	Income			
14	Job History			
15	Loss Experience			
16	Medical, including Biometrics, genetic information, pre-existing conditions, diagnostic data, etc.			
17	Natural Catastrophe Hazard (Fire, Wind, Hail, Earthquake, Severe Convective Storms)			

DRAFT

18	Online social media, including characteristics for targeted advertising			
19	Personal Financial Information			
20	Reasonable Accommodations or Policy Modifications (granted or requested)			
21	Telematics/Usage-based insurance			
22	Vehicle-Specific Data including VIN characteristics			
23	Voice Analysis			
24	Weather			
25	Other: Non-Traditional Data Elements (Regulator to specify)			

DRAFT

DEFINITIONS AND APPENDIX

Where available, for the purposes of this evaluation terms are defined in accordance with the NAIC Model Bulletin on the Use of *AI Systems* by Insurers (https://content.naic.org/sites/default/files/2023-12-4%252520Model%252520Bulletin_Adopted_0.pdf):

"Adverse Consumer ~~Outcome~~Impact" refers to an *AI System* decision (output) by an insurance company that is subject to insurance regulatory standards enforced by the Department that adversely impacts the consumer ~~in a manner that violates those standards~~.

"Agentic AI" refers to *AI Systems* that can pursue objectives across multiple steps without requiring human input at each stage, which could involve perceiving the environment, deciding what to do next, taking actions (often through tools or APIs), observing/analyzing the results, and/or adjusting the action accordingly. The difference is in agency – the system does not just respond to a single prompt or instruction, but actively executes a process.

"Algorithm" means a clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result.

"AI Model" refers to an integral component of an *AI System* that learns from data to perform tasks via computational, statistical and/or machine learning techniques. It combines algorithms, training data and learned parameters to produce outputs from a given set of inputs, such as making predictions or decisions on data it has never seen before.

"AI System" is a machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, content (such as text, images, videos, or sounds), or other output influencing decisions made in real or virtual environments. *AI Systems* are designed to operate with varying levels of autonomy.

"**AIS Program**" refers to the controls and processes that an insurer adopts and implements to ensure the responsible use of *AI Systems*. An *AIS program* may be comprised of general guidelines, governance framework, risk management and internal controls, and third-party oversight.

"Artificial Intelligence (AI)" refers to a branch of computer science that uses data processing systems that perform functions normally associated with human intelligence, such as reasoning, learning, and self-improvement, or the capability of a device to perform functions that are normally associated with human intelligence such as reasoning, learning, and self-improvement. This definition considers machine learning to be a subset of artificial intelligence.

DRAFT

“Augmentation” refers to ~~an~~ the use of an *AI System* that suggests an answer and/or advises a human who is making a decision.

“Automation” refers to the use of an *AI System* that does not involve human intervention.

~~**“Consumer Impact”** refers to a decision by an Insurer that is subject to insurance regulatory standards enforced by the Department.~~

“Direct Consumer Impact” refers to an *AI System* that *augments* or *automates* human decision-making whose output has the potential to create final consumer outcomes including but not limited to underwriting including tier and coverage eligibility determinations, premium determinations, ratemaking, and claims denial and settlement-related decisions. This has less applicability to *AI Systems* used for internal analytics, model monitoring, and back-office tooling with no consumer-visible output or other internal operational or administrative functions that do not determine or materially affect an outcome for a specific consumer.

“Externally Trained Models” refers to transferred learnings from pre-trained models developed by a third party on external reference datasets.

“Generative Artificial Intelligence (Generative AI)” refers to a class of *AI Systems* that generate content in the form of data, text, images, sounds, or video, that is similar to, but not a direct copy of, pre-existing data or content.

“Generalized Linear Model (GLM)” refers to a statistical modeling technique that extends linear regression to by relating predictive variables to an outcome variable using an assumed additive or multiplicative relationship through a link function and an appropriate probability distribution.

“Inherent Risk” refers to the risk of economic loss or inaccurate financial reporting before considering internal controls.

“Internally Trained Models” refers to models developed from data internally obtained by the company.

“Machine Learning (ML)” refers to a field within artificial intelligence that focuses on the ability of computers to learn from provided data without being explicitly programmed.

“**Material Financial Impact**” should be determined based on either the company’s materiality threshold which should be disclosed when responding to Supplement related inquiries or based on the threshold to be provided by regulators.

“**Materiality**” refers to the dollar amount above which the examiner’s perspective of an insurer’s financial position will be influenced.

Commented [MR38]: Edits to Augment, Automate, and Support were intended to address NAMIC’s and Monitaur’s feedback.

DRAFT

“Model Drift” refers to the decay of a model’s performance over time arising from underlying changes such as the definitions, distributions, and/or statistical properties between the data used to train the model and the data on which it is deployed.

“Neural Network Model” refers to a machine learning model that mimics the complex functions of the human brain. These types of models consist of interconnected nodes or neurons that process data, learn patterns and enable tasks such as pattern recognition and decision-making, including but not limited to: Single/multi-layer perceptrons/fully connected networks (MLPs/FCs), Deep Learning (DL), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Neural Networks (LSTMs), Sequence Models, Large Language Models (LLMs), and Reinforcement Learning Models (RLs).

“Predictive Model” refers to the mining of historic data using algorithms and/or machine learning to identify patterns and predict outcomes that can be used to make or support the making of decisions.

“Residual Risk” refers to an assessment of risk after considering risk-mitigation strategies or controls.

“Support” refers to the use of an *AI System* that provides information but does not suggest a decision or action to a human.

“Third Party” for purposes of this Supplement means an organization other than the insurance company that provides services, data, or other resources related to AI.

“Validation Method” refers to the source of the reference data used for validation, whether Internal, External, or Both.

“Use Case” refers to a description of a specific function in which a product or service is used.

Operations

Claims - Examples: Claim assignment, triage/fast-tracking, individual/bulk claim reserving including loss estimation, imaging/video analysis, fraud detection, litigation, estimation of closure rates, salvage/subrogation, examination/report ordering.

Customer Service - Examples: Agent/broker/internet/customer service interaction (chatbots), online/smart phone apps, loss prevention/risk mitigation advice, payment plans, complaints.

Marketing - Examples: market research, target advertising, market/coverage expansion, customer segment target marketing, demand modeling, agent/broker incentive plans, up/cross-selling.

Other: Cyber Security, Strategic Operations, Reserving, Investments, Capital Management, Financial Reporting, Reinsurance, Legal, Legal Exposure, Reputation Risk.

Ratemaking/Pricing - Examples: Development of overall/base rates, expense/loss loadings, estimation of trends and loss development, development of manual rating factors, tiering criteria, insurance credit scoring, territory boundary definitions, numeric/categorical level groupings and interactions, individual risk rating, telematics/UBI, price optimization, schedule rating factors.

Underwriting - Examples: Policy/coverage acceptance or eligibility, company placement/tiering, schedule rating, decisions based on telematics/UBI, report ordering, retention modeling, inspections, anomaly detection.

DRAFT



October 30, 2026

NAIC Big Data and Artificial Intelligence (H) Working Group
1100 Walnut Street
Suite 1500
Kansas City, MO 64106
Attn: Scott Sobel (ssobel@naic.org) and Miguel Romero (maromero@naic.org)

RE: *APCIA Comments on the AI Risk Evaluation Supplement version 6.0*

Chair Houdek and Members of the Big Data and Artificial Intelligence (H) Working Group:

APCIA appreciates the opportunity to comment on Version 6.0 of the AI Risk Evaluation Supplement. We support the Working Group's efforts to develop a practical, risk-based tool to help regulators understand insurers' use of AI Systems and assess potential consumer and solvency risks, and we appreciate its ongoing engagement with stakeholders throughout this process.

Version 6.0 reflects meaningful progress. We appreciate the Working Group's emphasis on risk-based supervision, the clarification that the Supplement does not create new legal obligations or regulatory authority, the recognition that Exhibit A should serve as the primary examination tool, and the flexibility to rely on existing documentation and prior submissions. These changes improve the Supplement's usability and alignment with existing examination frameworks.

While the Supplement has improved considerably, several significant issues remain. Please find below several recommendations from our members to help address the remaining concerns.

Preserve a Risk-Based Approach

Version 6.0 appropriately recognizes that insurers establish materiality thresholds and identify high-risk AI uses through their governance frameworks. That principle should be applied consistently throughout the Supplement. Exhibits C and D should be reserved for a limited subset of high-risk uses rather than becoming default requests applicable to all AI systems. Otherwise, a risk-based supervisory tool could become an enterprise-wide inventory and reporting exercise rather than a targeted supervisory tool. APCIA also encourages the Working Group to clarify that insurers may rely on existing documentation, model inventories, model cards, risk assessments, and similar materials when reasonably responsive to the inquiry.

Improve Consistency and Clarity Throughout the Supplement

Several key terms influence reporting obligations and examination scope without sufficient clarity. Terms such as *materiality*, *material change*, *financial impact*, *direct consumer impact, effectiveness*, and *transaction* warrant additional clarification to promote greater consistency across jurisdictions and examinations.

While Version 6.0 improves the distinction between *AI Systems* and *AI Models*, that distinction should be

applied consistently throughout the Supplement, including references to *generative AI* and *agentic AI*. Exhibits C and D should function as drill-down inquiries within an AI System governance framework rather than creating a separate model-governance framework. Finally, APCIA encourages the Working Group to maintain consistency with the NAIC AI Model Bulletin's concept of "*adverse consumer impact*." Insurance decisions may result in unfavorable outcomes for individual applicants or policyholders, but an unfavorable outcome does not, by itself, mean that a decision is unfair, discriminatory, or inconsistent with applicable law. References throughout the Supplement should therefore focus on *adverse* consumer impacts rather than consumer impacts generally.

Generalized Linear Models Should Be Excluded

APCIA remains concerned that the Supplement treats conventional generalized linear models (GLMs) as potential AI models subject to further inquiry. Although Version 6.0 permits regulators to exclude GLMs, that decision remains discretionary. Because GLMs used in ratemaking and underwriting are already subject to longstanding actuarial standards, existing regulatory review frameworks, and specific NAIC guidance, APCIA recommends excluding them from Exhibits B, C, and D. Regulatory attention should remain focused on technologies and use cases that present heightened or novel AI-related risks rather than conventional statistical models that have been reviewed by regulators for decades.

Clarify Testing Expectations for High-Risk AI Models

APCIA recognizes that Exhibit C is intended to provide additional information regarding a limited subset of high-risk AI models. However, Questions 10 and 10a read more like model governance requirements than examination inquiries. Question 10 requests extensive information regarding testing, validation, monitoring, adverse consumer impacts, and potential model changes, while Question 10a requests information regarding evaluations of potential proxy relationships or effects.

While inquiries into governance and risk-management practices may be appropriate for high-risk AI models, these questions could be interpreted as establishing expectations regarding specific testing methodologies or analyses. APCIA is also concerned that the reference to "*other reasonably available alternatives*" could be interpreted as creating a new substantive expectation rather than seeking information on existing governance practices.

Consistent with the Supplement's statement that it creates no new legal obligations or regulatory standards, Questions 10 and 10a should focus on understanding an insurer's governance framework and risk-management approach rather than implying uniform testing processes or methodologies. Any inquiry regarding proxy relationships should be tied to applicable law, the specific use case, available data, and the risks presented by the AI model.

Narrow and Clarify AI Governance Questions

Exhibit B includes questions regarding unsanctioned AI use, enterprise AI assistants, and general-purpose AI tools that could be interpreted broadly enough to encompass routine productivity tools with little connection to regulated insurance activities.

The Supplement should remain focused on material and higher-risk AI uses that affect regulated insurance functions, consumers, or financial outcomes, rather than routine low-risk applications or operational business processes that do not directly affect regulated insurance activities. It should also provide greater clarity regarding the distinction between AI that automates or augments decision-making and AI tools that merely

provide support or general productivity assistance, including enterprise productivity tools that have AI-enabled features but are not used to make or materially influence insurance decisions.

Calibrate Third-Party Oversight Expectations

We appreciate the improvements made regarding third-party AI oversight. However, oversight expectations should remain risk-based and limited to information reasonably available to the insurer. Insurers often lack access to proprietary vendor information, and the Supplement should not create an expectation that insurers obtain information that vendors are unwilling or unable to provide.

Limit Exhibit D to Targeted, High-Risk Inquiries

As drafted, Exhibit D could require extensive information across dozens of data categories and resemble an enterprise-wide data inventory. The information requested should instead be limited to data elements used in production high-risk models that are the subject of a specific regulatory inquiry.

Reporting should focus on material production inputs rather than all data sources considered during development or testing. Requiring disclosure of data evaluated but not ultimately incorporated into production models provides little supervisory value while raising operational and intellectual property concerns. The Supplement should also avoid requests that could effectively require reconstruction of model-development processes or production of information that may be interpreted inconsistently across companies and examiners.

Promote Consistent and Coordinated Implementation

APCIA encourages the NAIC to emphasize coordinated, risk-based implementation across regulatory functions and jurisdictions. Pilot participants reported overlapping AI-related information requests from multiple states and regulatory disciplines, underscoring the importance of avoiding duplicative or inconsistent requests. As the Supplement moves toward adoption, clear guidance on implementation and examination coordination will be important to reducing duplication and promoting consistent insurer responses.

Conclusion

APCIA appreciates the progress reflected in Version 6.0 and the Working Group's engagement with industry. We respectfully encourage the Working Group to consider the recommendations outlined above as it moves toward final adoption and future implementation of the Supplement. Addressing these remaining concerns will help promote consistent, risk-based oversight while avoiding unnecessary burden and unintended regulatory expectations.

We appreciate your consideration of these comments.

Thank You,

Kristin Abbott
Sr. Director and Counsel, Cyber & Privacy

Dave Snyder
Vice President, International & Counsel

Public Comment on the NAIC AI Risk Evaluation Supplement

Version 6.0 | October 30, 2026

To: NAIC Big Data and Artificial Intelligence (H) Working Group

From: Dapp Detroit — Enterprise AI Systems Architecture and Governance

In collaboration with: ARAI Global — AI Governance Education and Certification Development

Subject: Technical Comments on AI System Risk Evaluation, Agentic Workflows, Third-Party Dependencies, and Insurance Governance Evidence

Dear Members of the Working Group:

Thank you for the opportunity to provide comments on Version 6.0 of the Artificial Intelligence Risk Evaluation Supplement.

I submit these comments from the perspective of an enterprise technology architect and AI governance subject-matter expert, with professional experience spanning complex systems, financial and wealth-sector technology, workflow automation, and the design of governed AI infrastructure.

Through Dapp Detroit, my work focuses on enterprise AI architecture, system authorization, risk evaluation, workflow governance, and auditable implementation. In parallel, my work with ARAI Global includes the development of professional AI governance competencies and insurance-specific certification frameworks intended to improve the ability of technology, compliance, and operational professionals to assess AI-enabled systems.

These comments are offered as independent technical and industry observations, rather than as representations of a regulated insurance carrier or insurance operator.

I support the Working Group's risk-based approach and its recognition that AI risk evaluation must consider both direct consumer impact and material financial risk. I also appreciate the distinction between an AI system and its underlying models, the inclusion of agentic AI, and the proposed coordination with existing regulatory examination procedures.

My recommendations focus on several implementation environments where AI governance can become difficult to demonstrate, particularly as insurers increasingly rely on interconnected technology systems and external solution providers.

1. Proof-of-Concept, Startup, and Sandbox Risk

Relevant draft sections: Exhibits A and B

A significant implementation risk arises during the transition from a proof-of-concept or experimental environment into operational use.

Insurance organizations may evaluate AI solutions through internal development teams, startup technology providers, InsurTech platforms, innovation laboratories, and controlled sandboxes. These systems may demonstrate strong functional performance before the organization has established sufficient evidence

regarding security, model provenance, data protection, integration dependencies, or operational accountability.

I recommend that the Supplement encourage regulators to consider whether insurers maintain documented authorization and review procedures for AI systems progressing through distinct implementation stages, including exploratory testing, controlled sandbox deployment, data and cybersecurity assessment, vendor due diligence, limited production, full production approval, and ongoing monitoring.

A successful technical demonstration should not, by itself, establish that an AI system is appropriate for production within a regulated insurance workflow. The depth of review should remain proportionate to inherent risk, information sensitivity, and the consequences of potential actions.

2. Agentic AI, System Permissions, and Decision Authority

Relevant draft sections: Exhibits A, B, and C

I support the Supplement's explicit identification of agentic AI. However, agentic systems introduce risks that may not be fully understood through model classification or output validation alone.

An AI agent may retrieve information, initiate transactions, modify records, invoke external tools, communicate with consumers, or trigger actions across multiple systems. Risk evaluation should consider not only what an AI system produces, but also what it is authorized to do.

I recommend additional inquiry into agent and service-account identity; accessible data and systems; distinctions between recommendation, modification, approval, and execution; human approval and override; logging of tool calls and consequential changes; and mechanisms for suspending or containing autonomous actions.

These considerations are especially important in claims administration, underwriting, policy servicing, financial transactions, customer communications, and distribution operations. A human-in-the-loop designation should be supported by evidence that the reviewer has sufficient authority, information, and practical opportunity to intervene.

3. Third-Party AI and Distributed Insurance Technology Ecosystems

Relevant draft sections: Exhibits B and C

Insurance AI systems increasingly operate through combinations of internal software, third-party platforms, external models, application programming interfaces, and embedded service providers.

In addition to traditional vendor relationships, regulators may encounter AI capabilities supplied through managing general agents, managing general underwriters, third-party administrators, insurance marketing organizations, field marketing organizations, distribution technology providers, and other intermediaries.

I recommend that the Supplement encourage evaluation of material downstream dependencies, including fourth-party providers, external model services, embedded AI features, and delegated operational authority.

The central issue is whether the insurer can identify and assess consequential AI actions and obtain appropriate evidence of controls even when another organization supplies portions of the technology or workflow. This need not assume that every intermediary has the same regulatory obligations as the insurer.

4. Open-Source AI, Software Supply Chains, and Cybersecurity Exposure

Relevant draft sections: Exhibit B — vendor oversight, system development, and material changes

Open-source and open-weight AI technologies can offer innovation, cost efficiency, local deployment, and improved control over sensitive data. Their use does not eliminate cybersecurity or governance obligations.

Risks may arise from unverified dependencies, compromised repositories, malicious packages, model or dataset provenance, prompt injection, unauthorized tool execution, insecure integrations, and subsequent model or environment changes.

I recommend that the Supplement encourage insurers to consider whether existing technology risk processes address provenance, security testing, dependency management, access restrictions, and change monitoring of material AI components. The relevant distinction is actual system exposure and operational controls, not whether technology is proprietary or open source.

5. Insurance Distribution, Life Products, and Wealth-Related Workflows

Relevant draft sections: Exhibit A — consumer and financial impact

AI-enabled risk may emerge through producer-facing and intermediary workflows supporting insurance marketing, suitability-related documentation, policy illustrations, product comparisons, customer communications, and policy servicing.

These concerns are particularly relevant to life insurance, annuities, and adjacent wealth-sector systems where information and recommendations may move through multiple organizations before reaching a consumer.

For example, an AI assistant used by an insurance distribution organization may generate product comparisons or customer-facing explanations from carrier materials, external datasets, or embedded models. Evaluation should consider information currency and traceability, authorized roles, meaningful supervision, and records sufficient to reconstruct consequential interactions.

I recommend that regulators consider these distributed workflows when determining whether an AI system creates direct consumer impact or material financial exposure.

6. Risk Classification, Governance Evidence, and Continuous Assurance

Relevant draft sections: Exhibits A, B, C, and D

I support the Supplement's reliance on inherent risk, materiality, and existing documentation. From a systems governance perspective, risk evaluation can be strengthened by examining the consequence of incorrect or unauthorized decisions, data sensitivity and provenance, autonomy, connected-system reach, and maturity of relevant providers.

These characteristics can support proportionate risk classification without prescribing a single scoring methodology.

I recommend consideration of evidence demonstrating how a consequential AI-assisted decision was reached, what authority applied, which controls operated, and whether exceptions were identified and resolved. Relevant artifacts may include inventories, risk assessments, authorization records, model and integration documentation, override records, vendor assessments, change histories, incidents, and remediation tracking.

The objective should be to evaluate whether governance controls operate effectively—not simply whether written policies exist.

7. Workforce Competency and Governance Readiness

Relevant draft sections: Exhibit B — accountability, training, and oversight

AI governance is not exclusively a technical or compliance responsibility. Effective implementation requires coordination among technology architects, cybersecurity personnel, legal and compliance teams, insurance operations, risk management, and executive leadership.

A practical consideration for regulatory inquiry is whether personnel responsible for approving, supervising, or reviewing material AI systems have the competencies necessary to understand system behavior, authorization boundaries, escalation requirements, and supporting evidence.

Closing — Industry Competency, Risk Assessment, and Continued Collaboration

The AI Risk Evaluation Supplement provides an important foundation for consistent, risk-based regulatory inquiry. I encourage the Working Group to preserve its flexible, supplemental approach while strengthening consideration of emerging AI implementation risks, particularly those involving agentic systems, experimental deployments, external technology dependencies, and distributed insurance workflows.

Recognizing the growing need for industry-specific governance competencies, we have developed a proposed insurance specialization to build upon ARAI Global's existing Certified AI Governance & Compliance Specialist (C-AIGCS) credential.

The proposed C-AIGCS-I (Certified AI Governance & Compliance Specialist: Insurance) pathway is designed to equip professionals with practical competencies to evaluate, authorize, govern, and monitor AI-enabled systems within regulated insurance environments. It addresses enterprise architecture, AI agents, third-party providers, startup and sandbox readiness, cybersecurity, human decision authority, and audit evidence.

Complementing the individual credential, we have also developed a proprietary risk-scoring and readiness-assessment methodology incorporated into the proposed ARAI Insurance AI Governance Readiness Assessment (I-AIGRA) framework.

This applied framework is designed to evaluate organizational and technology-provider readiness, classify AI implementation risks, assess control maturity, and produce structured evidence registers, risk findings, and remediation recommendations. Its approach considers decision consequences, data sensitivity, system autonomy, integration exposure, and provider maturity.

These are two distinct but complementary initiatives. The C-AIGCS-I credential is intended to validate individual professional competencies; it does not certify an employer or AI product. I-AIGRA is intended to support governance, procurement, risk assessment, and examination readiness; it does not constitute an independent audit, regulatory approval, legal opinion, or guarantee of compliance. Both initiatives remain subject to applicable development, validation, and quality-review processes.

Our objective is to help bridge the gap between regulatory expectations and the technical realities of deploying and governing AI across the insurance industry.

We would welcome opportunities to contribute further technical perspectives, participate in relevant industry discussions, or share implementation insights that may assist the Working Group as its AI evaluation resources evolve. Additionally, we would welcome the opportunity to demonstrate our proprietary **Insurance AI Governance Readiness Assessment (I-AIGRA)** framework, currently in development, as a potential industry resource for evaluating AI readiness, classifying implementation risks, assessing governance maturity, and producing structured evidence and remediation recommendations. We believe this approach could complement the NAIC's risk-based evaluation objectives by helping translate governance expectations into measurable, auditable operational practices.

Thank you for your consideration and for the Working Group’s continued efforts to advance responsible, evidence-based AI governance in insurance.

Respectfully,

Ramona Fellmy

Founder | Dapp Detroit

Enterprise AI Systems Architect & AI Governance Advisor

U.S. Training & Certification Partner | ARAI Global

www.dappdetroit.com | www.araiglobal.org

NAMIC®
NATIONAL ASSOCIATION OF MUTUAL INSURANCE COMPANIES

317.875.5250 | [F] 317.879.8408
10 W. 106th St., Suite 100 | Indianapolis, Indiana 46290
202.628.1558 | [F] 202.628.1601
20 F Street N.W., Suite 510 | Washington, D.C. 20001

October 30, 2026

Commissioner Houdek (WI), Chair
NAIC Big Data and Artificial Intelligence (H) Working Group
c/o Miguel Romero, Director of Innovation, Cybersecurity and Technology
via email maromero@naic.org

Re: Big Data and AI (H) Working Group Version 6 of the AI Risk Evaluation Supplement

Chair Houdek and Members of the Working Group:

On behalf of the National Association of Mutual Insurance Companies (NAMIC)[1], we are writing in response to the exposure of the AI Risk Evaluation Supplement, Version 6 (hereinafter "Supplement"). Below we detail our remaining substantive and thematic concerns, and attached we include suggested redlines for the Working Group to consider.

[1] New Substantive Additions to the Supplement Should be Removed
In our attached redlines, the Working Group will find areas that NAMIC has identified as new substantive additions to the Supplement that have not undergone proper vetting, discussion, nor pilot. Given the abbreviated exposure period and anticipated vote for adoption at the Fall National Meeting, NAMIC does not believe it is procedurally adequate to introduce new substantive aspects to the Supplement this late in the process, after 18 months of drafting and discussion.

One of the new substantive additions we highlight is the "Use case" column in Exhibit A. While we agree with the removal of the unpiloted Part II of Exhibit A, NAMIC remains concerned that the substance of that Part II is now showing up in the table through this "Use cases" section. The original purpose of Exhibit A was to be used as an introductory inquiry providing background information for a regulator to help determine if additional inquiry using the remaining exhibits is necessary. This new and unpiloted "Use case" column expands the scope of Exhibit A and the burden associated with providing this information, and can lead to differences in how insurers approach the quantification without clear definitions or further regulatory guidance in the Supplement. Moreover, Exhibit A's Column 1 already requires insurers to classify their uses of AI systems. As such, the proposed "Use case" column is only a repetitive, yet more burdensome, recycling of Column 1. NAMIC strongly recommends removal of this new column, from both a procedural and substantive perspective.

[1] The National Association of Mutual Insurance Companies is the foremost trade association representing the property/casualty insurance industry. Serving more than 1,300 member companies, including local and regional insurers as well as some of the nation's largest carriers, NAMIC members collectively write $453 billion in annual premiums, representing 58 percent of the homeowners and 52 percent of the automobile insurance markets. For more than 130 years, NAMIC has been the leading voice advancing public policy solutions and regulatory frameworks that promote a strong, competitive market and protect member companies and their policyholders.

Another new substantive addition is Ref. 10 in Exhibit C, which reads "Where applicable, discuss whether the company evaluates input variables, derived variables, or model outputs for potential proxy relationships or effects that may contribute to Adverse Consumer Impacts or unfair discrimination, including the methodology used and any resulting action taken." The discussion over "proxy relationships" has received robust debate outside of the NAIC. It is not a standard in insurance law, and it is not one that has been piloted or discussed by this Working Group over the past year and a half. Further, the Working Group is exposing this version for an abbreviated 14-day comment period and is looking to adopt the Supplement at the Fall National Meeting. As such, NAMIC strongly suggests the removal of Ref. 10, as it has not been rigorously vetted the way other parts of the Supplement have.

The Working Group has stated that its intent is for the Supplement to evolve alongside the technology, so NAMIC does not see a pressing need to include new substantive sections that have not yet been piloted or discussed in the upcoming vote for adoption. We therefore strongly encourage the removal of the noted new substantive additions in our redlines.

[2] GLMs Used in Ratemaking Should be Explicitly Removed from Scope
NAMIC's position continues to be that GLMs and other predictive models should be explicitly removed from scope, for the reasons we've set forth in our previous comments. For sake of brevity, we have not repeated those arguments here, but we encourage Working Group members to reference them in our prior comment letters.

If the Working Group intends for GLMs and other predictive models to remain in scope of the Supplement despite there already being adequate regulatory oversight, NAMIC strongly suggests that the Supplement clarify that GLMs and other predictive models used in Ratemaking are explicitly out of scope. It appears the Working Group might be attempting to do this through its inclusion in the "General Guidance" of: "Based on existing regulatory oversight provided via rate making, regulators may choose to exclude some or all GLMs for Exhibit B through D related inquiries if appropriate." NAMIC agrees that existing regulatory oversight covers GLM and predictive model use in ratemaking, and that GLMs or predictive models used in ratemaking should be outside the scope of the Supplement as a whole. GLMs and other predictive models used in ratemaking have oversight, are required to adhere to regulatory requirements (including prohibitions on unfair discrimination), and already receive regulatory scrutiny through the state filing process. We therefore include redlines to this effect. We also believe that a more explicit exclusion of GLMs in ratemaking from the scope of the entire Supplement will help aid consistency and uniformity of use across jurisdictions, as opposed to some departments choosing to include them in the ratemaking context and others not.

[3] Incorporate the Pilot's Successful State Coordination
One of the reasons the pilot of this Supplement was successful is because the states committed to focusing on their domestic insurers and coordinating to avoid repetitive requests. NAMIC members remain concerned over the potential for multiple states to request completion of the Supplement, even simultaneously. Such regulatory duplication was avoided during the pilot because the states coordinated

and committed to a "focus on using the Tool with domestic insurers."[2] NAMIC encourages the NAIC to incorporate this vetted and successful aspect of the pilot into the supplement by memorializing that states should focus on using the Supplement with their domestic insurers. Our redline suggestions include language to accomplish this shared objective.

[4] Adverse Consumer Impact
Version 6 of the Supplement contains a revised definition of "Adverse Consumer Impact" which removes "in a manner that violates those standards." NAMIC is concerned that this change means that legitimate, legal actions (such as cancellation, non-renewal, premium increase, etc.) will be considered "adverse consumer impact." The use of this phrase throughout the Supplement suggests a regulatory violation, not a permissible action, and the definition should be narrowly tailored to match. Therefore, NAMIC asks the WG to revise the definition to ensure that legitimate and legal actions are not considered an adverse consumer impact, as we do not believe that is the intent of the WG. Such revision should include "in a manner that violates those standards," and a material and cause-based connection between the AI system and the challenged decision or outcome. NAMIC strongly suggests that "Adverse Consumer Impact" and its use throughout the Supplement be carefully limited and tied to applicable law.

Further, from a consistency and scope standpoint, some references in version 6 of the Supplement maintain the prior adverse impact threshold, and some do not. All references to consumer impact should continue to focus on adverse impact to consumers. This is consistent with insurance AI governance regulatory requirements, and the corresponding risk assessment frameworks that insurers have developed. Further, the focus of AI governance regulatory requirements is on adverse consumer impacts, and the Supplement should follow this approach. In Exhibit B, for example, the inclusion of and focus on AI capabilities that support low and medium risk uses should be removed, as they are acknowledged to be lower risk.

[5] Redline Suggestions Attached
NAMIC has attached redline suggestions to go along with our substantive comments. We look forward to continued collaboration with the Working Group to arrive at a product that meets the needs of the market.

Sincerely,

Lindsey Stephani

Lindsey Stephani
Senior Policy Vice President
Data Science, Artificial Intelligence, and Cybersecurity
NAMIC

[2] https://content.naic.org/sites/default/files/inline-files/Pilot%20Project%20Summary_1.pdf

Artificial Intelligence (AI) Risk Evaluation Supplement

Background, Intent, and Instructions

Background:

The rapid expansion of big data and the adoption of artificial intelligence and machine learning are significantly transforming insurance practices. These technologies can offer substantial benefits to both insurance companies and consumers by facilitating the development of innovative products, improving customer interface and enhancing service, simplifying and automating processes, and promoting efficiency and accuracy. However, without robust governance and effective controls, the use of *AI Systems* ~~may lead~~carries a risk that could lead to Adverse Consumer ~~Outcomes~~ Impacts or compromise the financial soundness of an insurance company. Insurers are responsible for managing the risks associated with the development and implementation of *AI Systems* and must demonstrate to regulators that appropriate risk-based oversight mechanisms are in place and are functioning effectively. Ultimately, insurers that make decisions or perform actions impacting consumers or financial soundness that are made or supported by advanced analytical and computational technologies, including from *AI Systems*, must comply with all applicable insurance laws and regulations.

Intent:

The NAIC's Innovation, Cybersecurity and Technology (H) Committee charged the Big Data and AI Working Group (BDAIWG) to create ~~tool(s)~~resources that would assist regulators to identify and assess both financial and consumer risks from the use of *AI Systems*~~ on an on-going basis~~.

The following exhibits are designed to supplement existing market conduct, financial analysis, and/or financial examination review procedures for reviewing *AI Systems*. As this supplements existing NAIC resources, regulators should continue to consider existing NAIC resources as authoritative but may consider drawing from this ~~supplement~~ Supplement to assist in understanding and assessing a company's use of *AI Systems*. Inquiries and information requests performed related to this ~~supplement~~ Supplement will be coordinated consistent with the guidance provided by the *Market Regulation Handbook (MRH)*, *Financial Condition Examiners Handbook (FCEH)*, and the *Financial Analysis Handbook (FAH)*. Consistent with the Handbooks, this Supplement does not independently create substantive legal duties, standards of conduct, admissions, or expanded regulatory authority.

Moreover, consistent with the coordination practices described in the FCEH, MRH, and FAH, information requests made leveraging the Supplement should be coordinated as much as possible.

With the intention of the document as a supplement acknowledged, there may be some important context that is not specified in this document, particularly related to deciding which companies are in scope of inquiry. This is to avoid the perception of conflicting guidance as many important decisions leading to the full use of this Supplement would be governed by the aforementioned Handbooks.

Commented [MR1]: Edit to address input by Consumer Representatives.

Commented [MR2]: APCIA suggested change.

Commented [MR3]: Added acronyms to address input from Connecticut.

Commented [MR4]: ACLI input.

Commented [MR5]: APCIA and Monitaur related input.

DRAFT

These optional exhibits allow regulators to determine the extent of *AI Systems* usage for an insurance company and whether additional analysis is needed focusing on financial and consumer risk.

This Supplement consists of the following exhibit questionnaires:

- **Exhibit A: Quantify Regulated Entity's Use of AI Systems**
- **Exhibit B: AIS Program (Two Options: Narrative or Checklist)**
- **Exhibit C: High-Risk AI Models**
- **Exhibit D: AI Model Data**

Instructions:

Information obtained from the responses to the following Exhibits may supplement guidance and ~~tools~~ resources used during an existing market conduct, certification, financial analysis, and/or financial examination review, to enhance the regulator's understanding of the extent of *AI Systems* utilization and assessment of potential risk to consumers and the insurance company. Effective assessment requires regulators to maintain a fluent understanding of the applicable laws and regulations, including but not limited to those pertaining to unfair trade practices, unfair claims settlement practices, unfair discrimination protections, corporate governance annual disclosures, confidentiality, privacy protections, financial reporting, and rating.

This Supplement is designed with the intention that regulators using this document will most often use Exhibit A. Based on responses received, regulators may thereafter leverage Exhibit B to inquire about a company's governance practices. Finally, for a subset of a company's use of AI, companies may receive Exhibit C and Exhibit D related inquiries where regulators wish to probe further.

Therefore, ~~R~~regulators using this Supplement may wish to first use Exhibit A to determine if further inquiry using Exhibits B-D is necessary. It is possible that while the company responding indicates usage of *AI Systems*, its use of AI is limited or low in inherent risk such that further inquiry as contemplated by subsequent exhibits is not warranted. For instance, if a regulator is using the Supplement to inquire about a company's claims practices and based on responses to Exhibit A, the regulator determines that the company's use of AI is out of scope of a claims-related targeted exam, the regulator may determine it is appropriate to thereafter rely on the ~~Market Regulation Handbook~~MRH to guide the remaining work.

If information requested through this Supplement has already been provided to this department or any other state department of insurance, the company's response should so state, in which case the regulators may accept prior submissions if the prior response is still current and applicable. If the company's domestic regulator has evaluated its use of AI Systems, then other departments should refrain from repetitive requests. Where companies have existing documentation that may satisfy regulatory inquiries, regulators may consider accepting company documentation consistent with guidance described in the FCEH, MRH, and FAH.

Commented [MR6]: Edits to this statement are related to input provided by the consumer representatives.

Commented [MR7]: ACLI related input.

Commented [LS8]: NAMIC recommends this inclusion to memorialize the approach taken during the pilot - that states work together to avoid duplicative requests of companies.

Commented [MR9]: APCIA, BCBSA, InsurTech Advisors, and Monitaur related input.

The responses to these exhibits will be considered by regulators when identifying the inherent risks of the insurer. They should also affect the planned examination or inquiry approach, as well as the nature, timing and extent of any further procedures performed.

General Guidance

Machine Learning (ML) Models

Machine learning models are widely used to estimate the relationship between a dependent variable and one or more independent variables. For example, an insurer may want to estimate how a driver's age relates to the probability of being involved in a car accident. There are many ML models that can be used for this purpose.

In particular, generalized linear models (GLMs) are one such technique. As with other machine learning techniques, GLMs are not without risk of causing unfair discrimination or other ~~adverse consumer outcomes~~Adverse Consumer Impacts and require effective governance over data quality, model development, validation, implementation, ongoing performance monitoring, and compliance with applicable laws and regulations. This supplement asks about a company's use of specific models in Exhibit A to give regulators flexibility in determining their scope of interest.

Based on existing regulatory oversight provided via rate making, GLMs and other predictive models used in the rate making context are not in scope of this Supplement. Additionally, regulators may choose to exclude some or all GLMs for Exhibit B through D related inquiries if appropriate.

Commented [LS10]: NAMIC's position continues to be that GLMs and other predictive models should be explicitly removed from scope, for the reasons we've set forth in previous comments.

If the Working Group intends for GLMs and other predictive models to remain in scope of the Supplement, despite adequate regulatory oversight, NAMIC suggests these changes to clarify that GLMs used in Ratemaking are explicitly out of scope. We provide greater detail on this point in our substantive comments.

AI System vs AI Model

Definitions are provided for both terms in the Definitions and Appendix section of the Supplement. Generally, this Supplement is designed with a focus on AI Systems. In relation to each other, this Supplement and each of the terms are defined such that an AI System is viewed as the primary unit at which governance and reporting discussions will take place; a model is a component of an AI system; and a stand-alone model may be examined separately when it independently creates direct consumer impact or material financial risk.

However, there may be times where inquiries are directed at an AI System and are better addressed at the AI Model level or vice versa. In those instances, companies may engage with department officials to ensure information provided meets expectations and assists in accomplishing regulatory objectives.

Commented [MR11]: Extensive feedback was provided by various parties on these two terms. Leveraging input from Monitaur, and in lieu of tracking many additional edits, this new overarching guidance was added. This input was also raised by NAMIC and other parties (APCIA and ACLI) had general input this may help address.

Materiality

This Supplement was designed to allow regulators the flexibility to either specify materiality or rely on the company's *materiality* calculations as companies respond to Supplement related inquiries.

The concept of *materiality* leverages the definition otherwise used to drive financial exam-related decisions and as defined in the FCEH. For purposes of completing responses related to this Supplement, regulators

may opt to rely on a company's internal materiality thresholds provided the information supporting a company's *materiality*, including an explanation for why that threshold was considered appropriate.

However, nothing in the guidance provided within this Supplement should be construed to supersede FCEH guidance.

Risk Assessment

This Supplement also, at times, relies on a company's Risk Assessment. The concept of Risk Assessment centers around the discussion of *Inherent Risk* which leverages the definition reflected in the FCEH. For purposes of completing responses to this Supplement, regulators ask that Risk Assessment information provided to regulators include explicit indication of *Inherent Risk* Assessments (i.e., risk before the consideration of controls or mitigation measures).

Consistent with the discussion of *materiality*, nothing in the guidance provided within this Supplement should be construed to supersede FCEH guidance as it relates to discussions of risk assessment.

Confidentiality

The information requested via these Exhibits will be held confidential, consistent with the requesting state's relevant authority.

Regulators using any of these exhibits should cite examination or other authority as appropriate when requesting information from insurers. Regulators should cite all relevant confidentiality statutes or other specific protections related to documents, materials or other information in the possession or control of regulators that are obtained by or disclosed to the regulators or any other person in the course of a market conduct inquiry and all information reported or provided to the regulator pursuant to cited examination or other authority.

DRAFT

Exhibit A: Part One – Quantify Regulated Entity's Use of AI Systems

Purpose: To obtain information pertaining to the number of *AI Systems* and *AI Models* used by a company to help facilitate risk assessment. Based on the responses from the company, regulators may ask for additional information related to their *AIS Program* (Exhibits B), high-risk models (Exhibit C), and the types of data (Exhibit D~~).) where there is risk for *Adverse Consumer Outcomes* or material adverse financial impact.~~

Company Instructions:

- Provide the most current counts and use cases of the following, as requested.
- ~~For purposes of responding to information requests related to this Exhibit, models that *augment* or *automate* decision-making related to consumers are considered to have direct consumer impact.~~
- ~~*"AI System", "Adverse Consumer Outcome",* and *"Use Case"* are defined in the "Definitions and Appendix" section of this Supplement.~~
- Include all lines of business. If the governance differs by company, line of business, or state, work with your regulator to determine if multiple submissions are needed.
- If a company has an existing *AI ~~Systems~~* System and/or *Model* inventory, ~~it may suggest submission of the inventory in lieu of completing Exhibit A~~the company should provide the AI System and/or Model inventory ~~together with the completed Exhibit A~~. Regulators will look for columns with information on:
  - AI System and/or Model Name
  - Description of how the AI System and/or Model is used and the broader AI system in which it is contained
  - Operation or Program Area(s) the AI System and/or Model is Used Within
  - Inherent Risk Level Based Upon the Company's Risk Assessment
  - What direct consumer impact does this AI System and/or Model have?
  - What financial impact does this AI System and/or Model have?
- If a company has questions about any of the fields, please submit them in writing to <regulator contact email address>.

Regulator Instructions:

- Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.
- Regulators should either specify a materiality threshold to guide company responses to the Supplement or request that the company provide a threshold that it used in completing this Exhibit.
- Depending on the information provided in a company's AI System and/or Model inventory, regulators may consider if a separate Exhibit A submission is necessary.
- For column 3, regulators should specify the number of months.

Commented [LS12]: NAMIC suggests this deletion, as it does not align with the added bullet point under "regulator instructions," which states a company can first provide its own inventory, and regulators can then determine if Exhibit A is necessary. NAMIC agrees that the added bullet point under "Regulator Instructions" appropriately describes how the process should work. Companies that have an existing inventory should only need to complete Exhibit A if their existing inventory is deemed insufficient.

Formatted

DRAFT

Company Legal Name and Group Name: _____

NAIC CoCode and Group Code: _____

Company Contact Name: _____ Email: _____

Line(s) of Business for Which This Response Applies: _____

Materiality Threshold <regulator to provide or specify if the company will provide>: _____

Date Form Completed: _____

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	~~(4) Use Case~~ *~~Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact~~*	Do any of the AI Systems utilize any of the following (check all that apply) (5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Marketing							
Premium Quotes & Discounts							
Underwriting/ Eligibility							
Ratemaking/ Rate Classification/ Schedule Rating/ Premium Audits						Not Applicable	

Formatted Table

Commented [SS15]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

Commented [LS13]: NAMIC continues to suggest that the column "Number of AI Systems Implemented in the Past X Months" be time limited to no longer than 12 months. Alternatively, NAMIC suggests removal of this column altogether, as this should be sufficiently covered by capturing "Number of AI Systems Currently in Use."

Commented [LS14]: While the Working Group has removed Part II of Exhibit A, NAMIC remains concerned that the substance of that Part II is now showing up in the table through this "Use cases" section.

The original purpose of Exhibit A was to be used as an introductory inquiry providing background information for a regulator to help determine if additional inquiry using the remaining exhibits is necessary. This new "Use case" column expands the scope of Exhibit A and the burden associated with providing this information. Further, this portion was never piloted.

As such, NAMIC strongly recommends removal of this new column, from both a procedural and substantive perspective.

Commented [LS16]: In line with NAMIC's suggestion that GLMs used in ratemaking be explicitly removed from scope, we recommend clarifying that GLMs used in ratemaking are not expected to be noted on the Exhibit A table.

DRAFT

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	~~(4) Use Case~~ *~~Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact~~*	Do any of the AI Systems utilize any of the following (check all that apply)			
				(5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Claims/ Adjudication*							
Customer Service							
Utilization Management/ Utilization Review/ Prior Authorization/ Level of Care Determination							
Fraud/Waste & Abuse							
Investment/ Capital Management							
Reserves/ Valuations							
Catastrophe Triage							
Reinsurance							
Other Insurance							

DRAFT

Formatted Table

Commented [SS15]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

Commented [LS13]: NAMIC continues to suggest that the column "Number of AI Systems Implemented in the Past X Months" be time limited to no longer than 12 months. Alternatively, NAMIC suggests removal of this column altogether, as this should be sufficiently covered by capturing "Number of AI Systems Currently in Use."

Commented [LS14]: While the Working Group has removed Part II of Exhibit A, NAMIC remains concerned that the substance of that Part II is now showing up in the table through this "Use cases" section.

The original purpose of Exhibit A was to be used as an introductory inquiry providing background information for a regulator to help determine if additional inquiry using the remaining exhibits is necessary. This new "Use case" column expands the scope of Exhibit A and the burden associated with providing this information. Further, this portion was never piloted.

As such, NAMIC strongly recommends removal of this new column, from both a procedural and substantive perspective.

(1) Use of *AI System* in Operations or Program Area	(2) Number of *AI Systems* Currently in Use	(3) Number of *AI Systems* Implemented in the Past X Months	~~(4) Use Case~~ *~~Describe how the systems are used in your business and whether there is a direct consumer impact or a material financial impact~~*	Do any of the AI Systems utilize any of the following (check all that apply) (5) *Generative AI*	(6) *Agentic AI*	(7) Generalized Linear Models (GLMs)	(8) *Other AI/ML Models*
Practices** (if applicable)							

** Includes Salvage/Subrogation, adjuster assignments, etc.*

*** Any additional operation or program area with the potential to result in either Direct Consumer Impact or Material Financial Impact.*

Formatted Table

Commented [SS15]: Edits address inputs provided by the ACLI, BCBSA, CAI, NAMIC, InsurTech Advisors, and Shadow Dynamic Solutions. This was modified to provide more guidance on use cases, to remove most of the quantification and to allow for checkboxes to share whether certain types of machine learning are in place. A new table is shown here and changes are not shown in tracking mode because it becomes confusing.

Commented [LS13]: NAMIC continues to suggest that the column "Number of AI Systems Implemented in the Past X Months" be time limited to no longer than 12 months. Alternatively, NAMIC suggests removal of this column altogether, as this should be sufficiently covered by capturing "Number of AI Systems Currently in Use."

Commented [LS14]: While the Working Group has removed Part II of Exhibit A, NAMIC remains concerned that the substance of that Part II is now showing up in the table through this "Use cases" section.

The original purpose of Exhibit A was to be used as an introductory inquiry providing background information for a regulator to help determine if additional inquiry using the remaining exhibits is necessary. This new "Use case" column expands the scope of Exhibit A and the burden associated with providing this information. Further, this portion was never piloted.

As such, NAMIC strongly recommends removal of this new column, from both a procedural and substantive perspective.

DRAFT

Exhibit A: Part Two – AI Systems Request List

In Support of the regulatory inquiries related to this Supplement, the regulators request information or documentation related to the following:

Regulator Instructions: Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.

1. Materiality – Explanation related to the company's *materiality* calculation (if applicable) used in responding to inquiries related to this Supplement.
2. Risk Assessment – Documentation explaining the company's risk assessment process for purposes of its use of *AI Systems*.
3. Model Inventory – A listing of the company's models, that specifies risk for the following program areas <regulator to specify program areas in scope of the request>. Additionally, regulators request the model inventory or accompanying documentation include:
    a. *AI Model* Name
    b. Description of how the *AI Model* is used
    c. Description of the broader *Use Case* in which each *AI Model* is used
    d. Operation or Program Area(s) the *AI Model* is Used Within
    e. *Inherent Risk* Level Based Upon the Company's Risk Assessment
    f. What consumer impact does this *AI Model* have?
    g. What financial impact does this *AI Model* have?
4. Indicate whether each use case involves Agentic AI? If yes, is that AI System authorized to execute one or more actions that can directly create or materially alter a consumer, financial, or financial-reporting outcome without human approval before the action?

Commented [SS17]: Suggestion from Ali Houshangi

DRAFT

Exhibit B: (Narrative) AIS Program

Purpose: To obtain the Company's *AIS Program*, including how the company handles governance, risk identification, mitigation, and management framework and internal controls, and third-party *AI Systems* and data for *AI Systems*; and the process for acquiring, using, or relying on third-party *AI Systems* and data. Market and financial regulators should coordinate to gain access to the relevant section of the policies governing the use of *AI Systems*.

Company Instructions: Below are both Checklist and Narrative sections of Exhibit B of the NAIC's AI Risk Evaluation Supplement. If the responses to the ~~checklist~~ narrative indicate that an explanation of a policy, procedure or process can be found on a given page of a document, indicate the page number and the document title and include the document referenced in the submission. If other sections of the document are irrelevant, the relevant sections of the document may be submitted in lieu of the entire document. Please see definitions at the end of this document.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. The Exhibits are designed so that regulators would use either the Narrative or the Checklist Exhibit but not both.

The references to, and questions about, elements of a company's *AIS Program* are not intended to be interpreted as creating new requirements for *AIS Program*.

Group and Company Legal Name: _____

NAIC Group and Company CoCodes: _____

Company Contact Name: _____ Email: _____

Date Form Completed: _____

1. Provide your *AIS Program*. Click or tap here to enter text.
    a. What role maintains the *AIS Program*? Click or tap here to enter text.
    b. Discuss the governance structure, Board reporting and frequency. Click or tap here to enter text.
    c. Discuss the process by which the *AIS Program* is integrated throughout the organization, assessed and remediated. Click or tap here to enter text.
    d. Discuss the process by which the effectiveness of the *AIS Program*, *AI Systems*, and individual models are assessed and modified. Click or tap here to enter text.

Commented [MR18]: Pending - CAI says there are inconsistencies between the two versions of Exhibit B.

Commented [SS18R2]: Changes were made to the two versions of Exhibit B for consistency based on feedback from CAI

Commented [MR19]: Edit to address input by ACLI and NAMIC.

Formatted: Indent: Left: 0.5", No bullets or numbering

Commented [MR20]: Based on input from Monitaur

e. Discuss how responsibility for governance within the organization is assigned and how the organization ensures consistency and alignment. Please include a detailed list of subject matter experts who will be able to answer follow up regulatory inquiries. Click or tap here to enter text.

f. Discuss the integration of the *AI Systems* in the Own Risk and Solvency Assessment (ORSA) and Enterprise Risk Management (ERM) assessments. Click or tap here to enter text.

g. How does the insurance company assess autonomy, reversibility, and reporting impact risk of *AI Systems*? Click or tap here to enter text.

~~g.~~h. Discuss the company's materiality threshold and risk assessment process in evaluating its use of AI. Click or tap here to enter text.

2. Discuss the uses of *AI Systems* that:
   a. Generate a *material financial impact*. Click or tap here to enter text.
   b. Generate a *direct consumer impact*. The company's response should include, but not be limited to, whether there is a human in the loop, whether there are high volume transactions, how the company assures itself that the level of human review is appropriate, and how any exceptions or errors are found and resolved~~, and whether there were any staffing reductions as part of the *AI System* implementation~~.. Click or tap here to enter text.
   c. Generate or impact material information reported in financial statements. Click or tap here to enter text.
   d. Generate or impact risk or control assessment. Click or tap here to enter text.
   e. Discuss the development, testing, and implementation of material *AI Systems* that the Company has implemented. If appropriate, include details regarding where any systems differ from established IT systems and data handling protocols. Discuss the basis for deviation from established practices. Click or tap here to enter text.

3. Discuss how the company:
   a. Assesses, mitigates, and evaluates inherent risks related to the use of AI Systems including the possibility of the company engaging in unfair trade practices or that it may violate non-discrimination protections.
   b. Considers *AI System* risks within the software development lifecycle (SDLC) or within the broader *AIS Program*

~~3.~~4. Provide the policy for, and discuss the use and oversight of, material *AI System* vendors, model design and testing:
   a. Discuss the validation and testing procedures performed on internally-developed *AI Systems*, including whether validation and testing are performed by someone that is independent from development. Click or tap here to enter text.
   b. Discuss the validation and testing procedures performed on third-party vendor-supplied *AI Systems*. Click or tap here to enter text.
   c. Discuss the testing and verification that has occurred including frequency, scope and methodology. Click or tap here to enter text.

Commented [MR21]: Carrying over Exhibit A part 2

Commented [LS22]: NAMIC suggests including a reference to the level of human review that might be appropriate. The first part of the sentence asks "whether there is a human," so the later part should be consistent and not default to such presumption.

Commented [MR23]: Edit to address input by consumer representatives.

Commented [MR24]: ACLI and BCBSA related input.

Formatted: Indent: Left: 1", No bullets or numbering

Commented [LS25]: As written, the later part of the sentence reads as an assumption that a violation of unfair trade practices is an inherent risk of using AI. NAMIC suggests either removing this part of the sentence, or editing the verbiage to say something to the effect of "Assesses, mitigates, and evaluates inherent risks related to the use of AI Systems, including compliance with existing insurance law."

Commented [MR26]: Edits to create more alignment between the two versions of Exhibit B.

DRAFT

d. Discuss how the company ensures explainability and transparency into any *AI models* recommending or making decisions. Click or tap here to enter text.

~~4.~~5. Provide the policy for, and discuss the use and oversight of, material *AI Systems* by third-party providers including actuarial, claim, MGA, audit, and/or other professional services. Click or tap here to enter text.

a. Discuss the testing and verification that has occurred, frequency, scope, and methodology. Click or tap here to enter text.

~~5.~~6. Discuss additional aspects of the framework design and evaluation pertaining to *AI Systems*. Click or tap here to enter text.

~~a.~~ Discuss the unit(s) responsible for the *AIS Program*, assessment approach and frequency, and involvement with the program area to the extent it differs from that discussed previously. Click or tap here to enter text.

7. Discuss the criteria for identifying a material change and for triggering approval, revalidation, or enhanced monitoring, including predeployment testing~~: periodic or event-triggered reassessment, changes in operational outcomes, incidents and near misses, drift or degradation, third-party changes, and safe retirement or fallback.~~

8. ~~For generative AI, h~~How does your company:

a. Discover unsanctioned AI use

b. Prevent nonpublic personal information from entering unapproved AI ~~tools~~models

c. Govern enterprise assistants that have AI capabilities and support many low- and medium-risk uses

d. Decide when a general-purpose tool with AI capabilities becomes a high-risk use because of the workflow in which it is deployed.

Formatted: Indent: Left: 1", No bullets or numbering

Commented [SS27]: Based on feedback from Legion

Formatted: Indent: Left: 0.5", No bullets or numbering

Commented [LS28]: NAMIC suggests removal of this provision, as the focus of AI governance regulatory requirements is on adverse consumer impacts.

Commented [SS29]: Based on suggestion from InsurTech Advisors to address enterprise generative AI and shadow AI as operating risks

DRAFT

Exhibit B: (Checklist) AIS Program

Purpose: To obtain information about the Company's *AIS Program* including information about how the company handles governance, risk management and internal controls, and third-party *AI Systems* and data for *AI Systems*; and the process for acquiring, using, or relying on third-party *AI Systems* and data. The references to, and questions about, elements of a company's *AI Systems* Governance Risk Assessments are not intended to be interpreted as creating new requirements for *AI Systems* Governance Risk Assessments.

Company Instructions: Provide responses to the questions regarding how the governance of *AI Systems* fits within your company's system of supervision or Enterprise Risk Management program. When referencing documents that are responsive to this exhibit, please include the documents referenced in the submission. Include all companies and lines of business. If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below.

For company contact information, please also include a detailed list of subject matter experts who will be able to answer follow-up regulatory inquiries.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. The Exhibits are designed so that regulators would use either the Narrative or the Checklist Exhibit but not both.

Group and Company Legal Name(s): _____

NAIC Group and Company Code(s): _____

Company Contact Name: _____ Email: _____

Date Form Completed: _____

Ref	AI Systems Use Questions for Company	Company Response
1	Has the company adopted a written *AIS Program*? If yes, when was it adopted, what does it govern, who approved the program, and what is the frequency of review for updating?	

Commented [MR30]: Edit to address input by ACLI and NAMIC.

Commented [MR31]: Edit to address input by the Center for AI Oversight.

DRAFT

2	Was the Board of Directors or management involved in the adoption of an *AIS Program*?		
	2a. What is the role of the Board of Directors or management in the *AIS Program*?		
	2b. How is accountability assigned within the AIS Program and what training has been provided for those charged with oversight of the AIS Program?		
	2c. Discuss the company's materiality threshold and risk assessment process in evaluating its use of AI.		
Ref	**AI Systems Use Questions for Company**		**Company Response**
3	**Reference the processes and procedures of the Company *AIS Program* that addresses the following:**		
	How the Insurance Company…	**Document Name and Page #**	**If not specified in the *AIS Program*, provide details below:**
	3a. Considers *AI System* risks within its Enterprise Risk Management (ERM)		
	3b. Considers *AI System* risks within the Own Risk and Solvency Assessment Report (ORSA), as applicable.		
	3~~a~~c. Assesses, mitigates, and evaluates *inherent risks* related to the use of *AI System* including the possibility of the company engaging in unfair trade practices		
	3d. Determined "*materiality*" and what threshold was used for assessing *AI Systems* risk and determined the company's risk assessment processes for its use of AI Systems.		
	3e. Quantifies *AI System* risk levels		
	3f. Considers *AI System* risk impact on financial reporting		

Commented [MR32]: Edit to address input by Center for AI Oversight.

Commented [MR33]: Carrying over Exhibit A part 2

Commented [SS34]: Suggestion from Yash Lundia: Also ask for version # and date of AIS Program document

Commented [MR35]: To carry over former Exhibit A part 2

DRAFT

	3~~b~~g. Ensures *AI Systems* are compliant with applicable state and federal laws and regulations, including non-discrimination protections		
	3~~c~~h. Evaluates the risk of *Adverse Consumer ~~Outcomes~~Impact*		
	3~~d~~i. Considers data privacy and protection of consumer data used in *AI Systems*		
	~~3j. Considers consumer complaints resulting from *AI Systems* and whether they are identified, tracked, and addressed~~		
	3k. Promotes consumer awareness of the use of *AI Systems* through disclosures, policies, and procedures for consumer notification, as appropriate		
	3l. Trains employees about *AI System* use, ensures human oversight is appropriate, and defines prohibited practices (if any)		
	3m. Provides standards and guidance for procuring and engaging *AI System* vendors		
	3n. Governs, monitors, tests and ensures transparency of *AI Systems* developed by vendors.		
	3o. Considers *AI System* risks within the software development lifecycle (SDLC) or within the broader *AIS Program*		
	3~~e~~p. Evaluates whether *AI Systems* are suitable for their intended use and should continue to be used as designed		

Commented [LS36]: State law does not require the tracking of AI complaints and it may not be clear that AI has contributed to a complaint. As such, NAMIC suggests removal of this provision.

Commented [MR37]: Edit to address input by Consumer Representatives

DRAFT

Exhibit C: High-Risk AI Models

Purpose: To obtain detailed information on in-production high-risk *AI models*, such as models making automated decisions that could cause ~~*Adverse Consumer Outcomes,*~~ *material financial impact*~~;~~ or direct consumer impact ~~*material consumer impact*, or material financial reporting impact~~. *AI model* risk criteria are set by the insurance company. Note that *materiality* should be a consideration in determining the level of risk. ~~To assist in identifying models for which this information is requested, regulators may request information on the company's risk assessment and a model inventory if such information has not otherwise already been provided.~~

Company Instructions: Fill in the details for each of the *AI model(s)* requested. ~~Include all companies and lines of business.~~ If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below. The template below refers to both *AI Systems* and *AI Models* depending on the information being requested. There may be some instances where a company feels information should be provided in relation to the *AI System* and not the *AI Model* or vice-versa. This should be discussed with regulators as part of the submission process to avoid misunderstanding.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam.

Group and Company Legal Name(s): _____

NAIC Group and Company Code(s): _____

Company Contact Name: _____ Email: _____

Provide who prepared and who reviewed the responses, by function. _____

Date Form Completed: _____

Ref	AI Model Information Requests	Company Response
1	*AI model* name and version number.	
2	Use cases and purpose(s) of model	
3	Model type used in the *AI System* (*GLM, generative/agentic AI, or other AI/ML algorithm*)	

Commented [MR38]: Suggestion from Consumer Representatives, ACLI, and NAMIC.

Commented [SS39]: Suggested by Jeffrey Sharer

DRAFT

3a	Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval.	
4	Model Implementation Date	
5	Model development (internal, ~~"owned" third party,~~ or ~~true~~ third party, or a mix of both – include vendor name if applicable)	
6	Did the Company purchase the *AI model*, was the *AI model* developed specifically for the Company? Is the Company able to modify the *AI model* as necessary?	
~~7~~	~~Model risk classification (high, moderate, low, etc.)~~	
87	Model risk(s) (i.e., describe the potential for *adverse consumer impact*)	
98	Model limitations (i.e., limitations on *uses cases* or data inputs outside the domain of the training data, or other guardrails to manage model related risk)	
~~10~~9	*Does the AI model* ~~type (~~automate, augment, or support decision-making?~~)~~	
~~11~~10	Discuss testing and validation of model outputs, as applicable, ~~(e.g.~~ testing for model drift, accuracy, unfair trade practices, unfair discrimination, performance degradation, and other potential Adverse Consumer Impacts. ~~etc.).~~ Discuss validation performed prior to deployment and ongoing monitoring after deployment, including testing following material changes to the model, data, or use of the model. Where testing identifies material differences or potential Adverse Consumer Impacts, discuss the company's process for evaluating the results, identifying contributing factors, and determining whether changes to the model, data, methodology, controls, or other reasonably available alternatives are appropriate. ~~and how the model was validated prior to being deployed as well as how its performance is monitored on an ongoing basis.~~	

Commented [SS40]: Suggestion from Ali Houshangi

Commented [MR41]: To address input provided by NAMIC.

Commented [MR42]: Edit to address input by NAMIC and other parties.

Commented [MR43]: Edits to incorporate CT's input.

DRAFT

~~10~~1~~a~~	~~Where applicable, discuss whether the company evaluates input variables, derived variables, or model outputs for potential proxy relationships or effects that may contribute to Adverse Consumer Impacts or unfair discrimination, including the methodology used and any resulting action taken.~~	
~~12~~11	Last date of model testing or validation. Please include the results and corrective actions as a result of the last model testing, if applicable.	
~~13~~12	Discuss how the model affects the financial statements, risk assessment or controls.	
~~14~~13	Discuss how the model is reviewed for compliance with applicable state and federal laws, including but not limited to the unfair trade practices act and unfair claims settlement laws.	
~~15~~14	To the extent permitted by law, discuss if the company has had any actions taken against them for use of this model. Actions may include, but are not limited to, informal agreements, voluntary compliance plans, administrative complaints, ongoing third-party monitoring, cease and desist, remediation, restitution, fines, penalties, investigations, consent orders or other regulatory agency actions.	

Commented [MR44]: New question proposed by CT.

Commented [LS45]: The discussion over "proxy relationships" has received robust debate outside of the NAIC, it not currently a standard in insurance law, and it is not one that has been piloted or discussed by this Working Group over the past year. Further, the Working Group is exposing this version for an abbreviated 14 day comment period and is looking to adopt the Supplement at the Fall National Meeting. As such, NAMIC strongly suggests the removal of Ref. 10, as it has not been rigorously vetted the way other parts of the Supplement have.

Commented [SS46]: Suggestion from Yash Lundia: Connect testing with the resulting decision/action

DRAFT

DRAFT

Exhibit D: AI Model Data

Purpose: To obtain detailed information of the source(s) and type(s) of data used directly ~~in~~by an *AI model(s)* to identify risk of *adverse consumer impact*, unfair trade practices, material financial impact, or material financial reporting impact.

Company Instructions: Provide details below for the data used in *AI model(s)*. If any of the data elements listed are used in the training or test data as part of the development of an *AI model*, provide information on whether the data element is sourced internally or whether the data element is sourced from a third party, in which case provide the name of the third-party vendor. Leave blank if a data source is not used in the development of *AI model* for the insurance operation. Include all companies and lines of business. If the governance differs by entity, line of business, or state, work with your domestic regulator to determine if multiple submissions are needed. See definitions below.

Regulator Instructions: Regulators should customize this supplement to limit information requested to more targeted inquiries for use in a limited scope exam. Regulators may also wish to customize this supplement for the line of business as some data types may or may not be pertinent based on the responding company's lines of business. Depending on responses received, a regulator may wish to ask for a data dictionary to gain a more complete understanding of a data used for a given model.

Group and Company Legal Name(s): _____

NAIC Group and Company CoCode(s): _____

Company Contact Name: _____ Email: _____

List the *AI System* or *AI models* which rely on the data described in response to this submission:

_____

Describe the Line of Business for Which This Response Applies (complete one for each line of business):

_____

Date Form Completed: _____

Commented [MR47]: APCIA related input

DRAFT

DRAFT

Ref	(1) Type of Data Element Used in AI Model	(2) Describe how the AI Model uses this data including how the output is used within operations; or input data is used for training and testing	(3) Internal Data Source	(4) Third Party Data Source / Vendor Name
1	Aerial Imagery			
2	Age, Gender, Ethnicity/Race			
3	Consumer or Other Type of Insurance/Risk Score			
4	Crime Statistics			
5	Criminal Convictions (Exclude Auto-Related Convictions)			
6	Driving Behavior			
7	Education Level (Including school aptitude scores, etc.)			
8	Facial or Body Detection / Recognition / Analysis			
9	Geocoding (including address, city, county, state, ZIP code, lat/long, MSA/CSA, etc.)			
10	Geo-Demographics (including ZIP/county-based demographic characteristics)			
11	Household Composition			
12	Image/video Analysis			
13	Income			
14	Job History			
15	Loss Experience			
16	Medical, including Biometrics, genetic information, pre-existing conditions, diagnostic data, etc.			
17	Natural Catastrophe Hazard (Fire, Wind, Hail, Earthquake, Severe Convective Storms)			

DRAFT

18	Online social media, including characteristics for targeted advertising			
19	Personal Financial Information			
20	Reasonable Accommodations or Policy Modifications (granted or requested)			
21	Telematics/Usage-based insurance			
22	Vehicle-Specific Data including VIN characteristics			
23	Voice Analysis			
24	Weather			
25	Other: Non-Traditional Data Elements (Regulator to specify)			

DRAFT

DEFINITIONS AND APPENDIX

Where available, for the purposes of this evaluation terms are defined in accordance with the NAIC Model Bulletin on the Use of *AI Systems* by Insurers (https://content.naic.org/sites/default/files/2023-12-4%252520Model%252520Bulletin_Adopted_0.pdf):

"Adverse Consumer ~~Outcome~~Impact" refers to an *AI System* decision (output) by an insurance company that is subject to insurance regulatory standards enforced by the Department that adversely impacts the consumer in a manner that violates those standards, and in a manner that results in a material and cause-based connection between the AI System and the challenged decision or outcome. ~~in a manner that violates those standards~~.

"Agentic AI" refers to *AI Systems* that can pursue objectives across multiple steps without requiring human input at each stage, which could involve perceiving the environment, deciding what to do next, taking actions (often through tools or APIs), observing/analyzing the results, and/or adjusting the action accordingly. The difference is in agency – the system does not just respond to a single prompt or instruction, but actively executes a process.

"Algorithm" means a clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result.

"AI Model" refers to an integral component of an *AI System* that learns from data to perform tasks via computational, statistical and/or machine learning techniques. It combines algorithms, training data and learned parameters to produce outputs from a given set of inputs, such as making predictions or decisions on data it has never seen before.

"AI System" is a machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, content (such as text, images, videos, or sounds), or other output influencing decisions made in real or virtual environments. *AI Systems* are designed to operate with varying levels of autonomy.

"**AIS Program**" refers to the controls and processes that an insurer adopts and implements to ensure the responsible use of *AI Systems*. An *AIS program* may be comprised of general guidelines, governance framework, risk management and internal controls, and third-party oversight.

"Artificial Intelligence (AI)" refers to a branch of computer science that uses data processing systems that perform functions normally associated with human intelligence, such as reasoning, learning, and self-improvement, or the capability of a device to perform

Commented [LS48]: NAMIC is concerned that this change means that legitimate, legal actions (such as cancellation, non-renewal, premium increase, etc.) will be considered "adverse consumer impact." The use of this phrase throughout the Supplement suggests a regulatory violation, not a permissible action, and the definition should be narrowly tailored to match.

Therefore, NAMIC asks the WG to revise the definition to ensure that legitimate and legal actions are not considered an adverse consumer impact, as we do not believe that is the intent of the WG. Such revision should include "in a manner that violates those standards," and a material and cause-based connection between the AI system and the challenged decision or outcome.

functions that are normally associated with human intelligence such as reasoning, learning, and self-improvement. This definition considers machine learning to be a subset of artificial intelligence.

"Augmentation" refers to ~~to an~~ the use of an *AI System* that suggests an answer and/or advises a human who is making a decision.

"Automation" refers to the use of an *AI System* that does not involve human intervention.

~~**"Consumer Impact"** refers to a decision by an Insurer that is subject to insurance regulatory standards enforced by the Department.~~

"Direct Consumer Impact" refers to an *AI System* that *augments* or *automates* human decision-making whose output has the potential to create final consumer outcomes including but not limited to underwriting including tier and coverage eligibility determinations, premium determinations, ratemaking, and claims denial and settlement-related decisions. This has less applicability to *AI Systems* used for internal analytics, model monitoring, and back-office tooling with no consumer-visible output or other internal operational or administrative functions that do not determine or materially affect an outcome for a specific consumer.

"Externally Trained Models" refers to transferred learnings from pre-trained models developed by a third party on external reference datasets.

"Generative Artificial Intelligence (Generative AI)" refers to a class of *AI Systems* that generate content in the form of data, text, images, sounds, or video, that is similar to, but not a direct copy of, pre-existing data or content.

"Generalized Linear Model (GLM)" refers to a statistical modeling technique that extends linear regression to by relating predictive variables to an outcome variable using an assumed additive or multiplicative relationship through a link function and an appropriate probability distribution.

"Inherent Risk" refers to the risk of economic loss or inaccurate financial reporting before considering internal controls.

"Internally Trained Models" refers to models developed from data internally obtained by the company.

"Machine Learning (ML)" refers to a field within artificial intelligence that focuses on the ability of computers to learn from provided data without being explicitly programmed.

"**Material Financial Impact**" should be determined based on either the company's materiality threshold which should be disclosed when responding to Supplement related inquiries or based on the threshold to be provided by regulators.

Commented [MR49]: Edits to Augment, Automate, and Support were intended to address NAMIC's and Monitaur's feedback.

DRAFT

"**Materiality**" refers to the dollar amount above which the examiner's perspective of an insurer's financial position will be influenced.

"Model Drift" refers to the decay of a model's performance over time arising from underlying changes such as the definitions, distributions, and/or statistical properties between the data used to train the model and the data on which it is deployed.

"Neural Network Model" refers to a machine learning model that mimics the complex functions of the human brain. These types of models consist of interconnected nodes or neurons that process data, learn patterns and enable tasks such as pattern recognition and decision-making, including but not limited to: Single/multi-layer perceptrons/fully connected networks (MLPs/FCs), Deep Learning (DL), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Neural Networks (LSTMs), Sequence Models, Large Language Models (LLMs), and Reinforcement Learning Models (RLs).

"Predictive Model" refers to the mining of historic data using algorithms and/or machine learning to identify patterns and predict outcomes that can be used to make or support the making of decisions.

"Residual Risk" refers to an assessment of risk after considering risk-mitigation strategies or controls.

"Support" refers to the use of an *AI System* that provides information but does not suggest a decision or action to a human.

"Third Party" for purposes of this Supplement means an organization other than the insurance company that provides services, data, or other resources related to AI.

"Validation Method" refers to the source of the reference data used for validation, whether Internal, External, or Both.

"Use Case" refers to a description of a specific function in which a product or service is used.

Operations

Claims - Examples: Claim assignment, triage/fast-tracking, individual/bulk claim reserving including loss estimation, imaging/video analysis, fraud detection, litigation, estimation of closure rates, salvage/subrogation, examination/report ordering.

Customer Service - Examples: Agent/broker/internet/customer service interaction (chatbots), online/smart phone apps, loss prevention/risk mitigation advice, payment plans, complaints.

Marketing - Examples: market research, target advertising, market/coverage expansion, customer segment target marketing, demand modeling, agent/broker incentive plans, up/cross-selling.

Other: Cyber Security, Strategic Operations, Reserving, Investments, Capital Management, Financial Reporting, Reinsurance, Legal, Legal Exposure, Reputation Risk.

Ratemaking/Pricing - Examples: Development of overall/base rates, expense/loss loadings, estimation of trends and loss development, development of manual rating factors, tiering criteria, insurance credit scoring, territory boundary definitions, numeric/categorical level groupings and interactions, individual risk rating, telematics/UBI, price optimization, schedule rating factors.

Underwriting - Examples: Policy/coverage acceptance or eligibility, company placement/tiering, schedule rating, decisions based on telematics/UBI, report ordering, retention modeling, inspections, anomaly detection.

DRAFT

October 24, 2026

Scott Sobel, AI Policy Advisor
Miguel Romero, Director of Innovation, Cybersecurity and Technology
National Association of Insurance Commissioners

Submitted by email to ssobel@naic.org and maromero@naic.org

Re: Comments on the AI Risk Evaluation Supplement, version 6.0, Big Data and Artificial Intelligence (H) Working Group

Dear Mr. Sobel, Mr. Romero, and Working Group members,

Thank you for the opportunity to comment on version 6.0 of the AI Risk Evaluation Supplement. We appreciate the Working Group's careful work across successive versions and its openness to public input.

We especially welcome how version 6.0 recognizes that AI Systems now act, not only advise. It flags Agentic AI in the inventory[1] and asks companies to identify authorizations made without human approval.[2] It also asks how companies assess autonomy and reversibility,[3] how they make sure human review is appropriate,[4] what triggers new approval after a material change,[5] and when a workflow turns a tool into a high-risk use.[6]

One question is still open. When an AI System pays a claim, changes a policy, or sends a coverage decision, was that specific action authorized before it took effect? The supplement examines the program, the model, and the vendor. For a system that acts, the harm happens at the action. A well-governed program can still let one unauthorized payment through. The supplement asks how errors are found and resolved after the fact,[7] but no existing question asks whether that payment was authorized before it took effect.

Our recommendations close that gap without changing the supplement's design. They build on version 6.0's own questions about Agentic AI and turn them into evidence an examiner can test. They apply only to AI Systems that take Consequential Actions, a term we propose to define. It covers actions that move money, change coverage or a claim, send a coverage or claim decision outside the company, or disclose nonpublic information. The proposed text names no product or method, and a documented manual process can answer it.

Execution Rights develops controls for the moment before an AI System's action takes effect, and that work informs these comments.

Footnotes give the location of existing version 6.0 text. “Proposed” refers to text we recommend adding. Attachment A gives the exact proposed text and where it goes in version 6.0, and each reference to a

[1] Exhibit A, Part One table, column (6), “Agentic AI.”
[2] Exhibit C, Ref 3a: “Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval.” This Ref is new in version 6.0.
[3] Exhibit B (Narrative), item 1g: “How does the insurance company assess autonomy, reversibility, and reporting impact risk of AI Systems?”
[4] Exhibit B (Narrative), item 2b, which asks “how the company assures itself the human review is appropriate.”
[5] Exhibit B (Narrative), item 7: “Discuss the criteria for identifying a material change and for triggering approval, revalidation, or enhanced monitoring, including predeployment testing.”
[6] Exhibit B (Narrative), item 8d: “Decide when a general-purpose tool with AI capabilities becomes a high-risk use because of the workflow in which it is deployed.”
[7] Exhibit B (Narrative), item 2b, which asks “how any exceptions or errors are found and resolved.”

proposed change links to it. Attachment B gives example examination procedures that use the proposed questions.

We respectfully recommend six changes:

#	Recommendation	Proposed text
1	Use the Exhibit A inventory to identify which AI Systems take Consequential Actions and whether those actions are authorized before they take effect	Changes 1 to 4
2	Add one core examination item to Exhibit B. Before each Consequential Action takes effect, is it cleared, declined, or held, and is there a record?	Changes 5 and 6
3	For high-risk AI Systems, ask for evidence and test results in Exhibits B and C	Changes 5 to 9 and 11
4	Ask for the records of those decisions in a common format in Exhibit D	Change 12
5	Ask how amounts from Consequential Actions reach the financial statements	Changes 5, 6, and 10
6	Define the key terms so that every company and examiner reads them the same way	Changes 13 to 16

Sections 1 through 6 explain each recommendation in turn. Section 7 addresses scope and burden. Each recommendation can be adopted on its own, together with the definitions it uses.

1. Identify the AI Systems that take Consequential Actions

The inventory asks whether a system uses Agentic AI. That describes the technology. It does not tell a regulator whether the system can move money or change coverage. A predictive model connected to a payment system can pay claims without Agentic AI. An AI System that only drafts text takes no Consequential Action, even if it uses Agentic AI.

We recommend four small proposed changes:

Proposed change	What it does
Two inventory questions	Ask whether the AI System can take Consequential Actions, and whether each one is cleared, declined, or held before it takes effect
Inventory columns (9) and (10)	Count the AI Systems that can take Consequential Actions, and how many of those have Consequence Authorization in place. A gap between the two columns shows a regulator where to look.
One sentence in the Exhibit B instructions	Lets a regulator use the proposed Exhibit B, C, and D items together as a limited-scope inquiry when the columns show a gap
A clause in the item on workflow risk	Connects workflow risk to Consequential Actions

2. Ask whether each action is authorized before it takes effect

This is our core recommendation. It proposes one examination item with four questions:

1. Before a Consequential Action takes effect, is it cleared, declined, or held? Cleared means it may proceed. Declined means it may not. Held means it waits for a person with authority to decide.
2. Can a declined or held action take effect anyway, through a retry, another tool, or another route?
3. What happens when the check cannot be completed? The safe answer is that the action does not take effect.
4. What record shows each decision and what then occurred? We call this a Decision Record.

The item would appear in Exhibit B as new narrative items 9a to 9d and checklist rows 4a to 4d. It builds directly on version 6.0's question about authorizations made without human approval.[8] It asks how the company knows that each authorization was valid.

An example shows why this matters. An AI Agent that handles claims may pay approved claims up to $10,000 to the payee on file. Company policy requires separate verification before payment details change. An email that appears to come from the claimant gives a new bank account. The AI Agent changes the payee and pays the claim without that verification.

A check of the payment amount would pass, because the amount was within the AI Agent's limit. Proposed items 9a and 9h would catch it. The new account came from content the AI Agent read, not from anyone with authority to change it.

This is a familiar control. Funds-transfer procedures use measures such as separate initiation and approval, authentication, and callback verification. Under many such procedures, when a required callback cannot be completed, the wire waits. An AI System that pays a claim or changes coverage raises the same control question, and the same kind of control answers it.

3. Ask for evidence on high-risk AI Systems

For the AI Systems a regulator selects for closer review, we recommend proposed questions that produce evidence an examiner can test. Each can be dropped without changing the core item.

Proposed question	What it asks for	Builds on version 6.0's question about
Ref 3b	Each type of action, each route to it, the credentials that carry it out, and whether a separate control holds those credentials until the action is cleared	Agentic AI authorizations[9]
Ref 3c	How each action is cleared, declined, or held, and who resolves held actions	Agentic AI authorizations
Ref 3d	A sample of Decision Records that the regulator selects, reconciled with what the company's systems actually did	Agentic AI authorizations
Ref 3e and the example tests	The latest authorization test results, including nine example tests	Testing and validation[10]
Ref 3f	Changes since the last review and the approval that followed each one	Material change[11]
Ref 3g	The latest test of stopping further actions, and which effects can be reversed	Reversibility[12]
Ref 3h	The company's measurement, if it has one, of what share of actions took effect without authorization	No existing question
Ref 3i	What each clearance covers, so that a later step outside it is held	No existing question
Ref 3j	Whether the company confirms authorization itself when a vendor carries out an action	Third-party providers[13]
Addition to Ref 1	An identifier that changes when an AI Agent's model, instructions, tools, or authority change	Model name and version[14]
Ref 11a	An independent check that the control is in the path of every Consequential Action	Date of testing or validation[15]

[8] Exhibit C, Ref 3a.
[9] Exhibit C, Ref 3a.
[10] Exhibit C, Ref 10, on testing and validation of model outputs.
[11] Exhibit B (Narrative), item 7.
[12] Exhibit B (Narrative), item 1g.
[13] Exhibit B (Narrative), item 5, and Exhibit B (Checklist), row 3n.
[14] Exhibit C, Ref 1, “AI model name and version number.”
[15] Exhibit C, Ref 11, on the last date of model testing or validation.

Proposed question	What it asks for	Builds on version 6.0's question about
Items 9e to 9o and 9q, rows 4e to 4o and 4q	The narrative and checklist detail behind these questions, including human review (9j) and stopping actions (9n)	Autonomy, reversibility, and human review[16]

These answers also give regulators a view no single company has. Several companies may name the same model provider or vendor component in proposed Ref 3b. A regulator can then see which companies a change or failure in that component would reach. Attachment B shows how an examiner could use these answers in a limited-scope exam.

4. Ask for Decision Records in a common format

Version 6.0 already requests data about high-risk AI models.[17] For an AI System that acts, the most important data is the record of each authorization decision. We recommend a proposed Exhibit D table of twelve Decision Record fields. They cover:

- the action as proposed, and where each payment or coverage value came from
- the rules that applied and the conditions evaluated
- the determination, and who or what decided
- what then occurred, including any payment or financial entry

A common format lets examiners test records the same way across companies. It also tells companies what to keep before an exam begins.

5. Connect financial reporting to authorization

Version 6.0 already asks how AI affects the financial statements.[18] When an AI System pays a claim or changes a premium, the amount reaches the financial records whether or not the action was authorized.

We recommend proposed item 9p, row 4p, and Ref 12a. They ask whether amounts from Consequential Actions are recognized only when linked to a valid authorization determination. They also ask how the company identifies and corrects amounts from actions later found to be unauthorized. Financial-condition examiners can then trace a sample of financial entries to the Decision Records behind them.

6. Define the key terms

Without shared definitions, each company answers a slightly different question, and answers cannot be compared. We recommend nine proposed definitions.

Proposed term	In plain words
Consequential Action	An action that moves money, changes coverage or a claim, sends a decision outside the company, or discloses nonpublic information, whatever the technology
AI Agent	An AI System that can take Consequential Actions
Consequence Authorization	The control that clears, declines, or holds each Consequential Action before it takes effect
Decision Record	The record of each of those decisions, made before the action takes effect
Authorization Evidence	Records that let someone outside the company verify that actions were authorized
Clearance Scope	What a clearance covers
Composition Identifier	An identifier for the AI Agent's model, instructions, tools, and authority
Control Verification Statement	An independent check that the control is in place

[16] Exhibit B (Narrative), items 1g and 2b, and Exhibit B (Checklist), row 3l.
[17] Exhibit D, AI Model Data.
[18] Exhibit B (Narrative), item 2c; Exhibit B (Checklist), row 3f; and Exhibit C, Ref 12.

Proposed term	In plain words
Exposure Measurement	A company's statement of what share of its actions took effect without authorization

7. Scope and burden

The proposed questions reach only AI Systems that take Consequential Actions. A company with no such system states that they do not apply. Version 6.0 asks regulators to customize each exhibit for targeted inquiries.[19] A regulator can therefore use proposed items 9a to 9d or rows 4a to 4d alone, with the core definitions.

Each proposed change is an examination question or supports one. Consistent with the supplement's Intent section,[20] none independently creates a substantive legal duty or standard of conduct. Any implementation that produces the evidence can answer the questions, including a documented manual process or a control the company builds itself.

Conclusion

Version 6.0 has already asked the right question about AI that acts. These proposed changes let examiners confirm the answer, action by action. We respectfully recommend that the Working Group include them in version 7.0, starting with the core item in Recommendation 2.

On request, we would be glad to provide a tracked-changes copy of version 6.0 with these proposed changes, a reference measurement method, and an examiner workpaper form. The Working Group may adopt or adapt any of this text as it sees fit. We would also welcome the chance to summarize these comments at the November 2 meeting. Thank you for considering them.

Respectfully submitted,

Tim Fox
Founder, Execution Rights
executionrights.com
tim@executionrights.com

[19] Each exhibit's instructions state: “Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.”
[20] The Intent section states that the supplement “does not independently create substantive legal duties, standards of conduct, admissions, or expanded regulatory authority.”

ATTACHMENT A

Proposed Changes to Version 6.0

Each proposed change below has two parts. “Where in version 6.0” names the existing location and quotes existing NAIC text in quotation marks. “Proposed text” is the text we recommend adding. Changes are listed in the order they would appear in version 6.0. Core changes, together with the core definitions, stand on their own. Build-out changes add detail and can be dropped one at a time. No existing item is renumbered.

#	Where in version 6.0	Proposed change	Tier	Rec.
1	Exhibit A, Part One, Company Instructions, inventory list	Add two items	Core	1
2	Exhibit A, Part One table	Add columns (9) and (10)	Core	1
3	Exhibit B (Narrative), instructions	Add one sentence on targeted inquiries	Build-out	1
4	Exhibit B (Narrative), item 8d	Add a closing clause	Build-out	1
5	Exhibit B (Narrative), after item 8d	New item 9	Core: lead-in and 9a to 9d. Build-out: 9e to 9q	2, 3, and 5
6	Exhibit B (Checklist), after row 3p	New row 4	Core: lead-in and 4a to 4d. Build-out: 4e to 4q	2, 3, and 5
7	Exhibit C, Ref 1	Add a closing clause	Build-out	3
8	Exhibit C, after Ref 3a	New Refs 3b to 3j	Build-out	3
9	Exhibit C, after Ref 11	New Ref 11a	Build-out	3
10	Exhibit C, after Ref 12	New Ref 12a	Build-out	5
11	After the Exhibit C table, before Exhibit D	New table of example authorization tests	Build-out	3
12	End of Exhibit D, before Definitions and Appendix	New table of Decision Record data	Build-out	4
13	Definitions, after “Algorithm”	“AI Agent”	Core	6
14	Definitions, after “Augmentation”	“Authorization Evidence”	Build-out	6
15	Definitions, after “Automation”	Six entries, from “Clearance Scope” to “Decision Record”	Core: Consequence Authorization, Consequential Action, and Decision Record. Build-out: the other three	6
16	Definitions, after “Direct Consumer Impact”	“Exposure Measurement”	Build-out	6

Proposed change 1. Add two inventory questions to Exhibit A (core)

Where in version 6.0: Exhibit A, Part One, Company Instructions, in the existing list that follows “Regulators will look for columns with information on:”. Add as the last two items, after “What financial impact does this AI System and/or Model have?”

Proposed text:

> Can this AI System and/or Model take Consequential Actions (see Definitions), and of what types?
>
> Is each of those Consequential Actions cleared, declined, or held before it takes effect (see Consequence Authorization in Definitions)?

Proposed change 2. Add two inventory columns to Exhibit A (core)

Where in version 6.0: Exhibit A, Part One table. Add two new columns after existing column (8), “Other AI/ML Models,” outside the existing group header “Do any of the AI Systems utilize any of the following (check all that apply).”

Proposed text:

> (9) Number of AI Systems That Can Take Consequential Actions (see Definitions)
>
> (10) Number of the AI Systems in Column (9) Whose Consequential Actions Are Cleared, Declined, or Held Before They Take Effect (see Consequence Authorization in Definitions)

Proposed change 3. Add a sentence on targeted inquiries to the Exhibit B instructions (build-out)

Where in version 6.0: Exhibit B (Narrative), in the existing instructions, after the existing sentence “Regulators should customize this Exhibit to limit information requested to more targeted inquiries for use in a limited scope exam.”

Proposed text:

> Where Exhibit A, column (10) is lower than column (9), some AI Systems can take Consequential Actions without Consequence Authorization in place. Regulators may then use narrative item 9, checklist row 4, Exhibit C, Refs 3b to 3j, and Exhibit D, Decision Record Data, together as a limited-scope inquiry.

Proposed change 4. Extend Exhibit B narrative item 8d (build-out)

Where in version 6.0: Exhibit B (Narrative), existing item 8d, which ends “because of the workflow in which it is deployed.” Replace the final period with the proposed text.

Proposed text:

> , including when the workflow lets it take Consequential Actions.

Proposed change 5. Add Exhibit B narrative item 9 (core for the lead-in and 9a to 9d)

Where in version 6.0: Exhibit B (Narrative), a new item 9 after existing item 8d, at the end of the narrative and before the existing heading “Exhibit B: (Checklist) AIS Program.”

Proposed text:

> 9. For AI Systems that can take Consequential Actions, discuss whether and how Consequence Authorization (see Definitions) is in place, including how the company:
>
> a. Clears, declines, or holds each Consequential Action before it takes effect, separately from the limits of any credential, account, or payment instrument the AI System uses.
>
> b. Prevents a declined or held action from taking effect through any other tool, route, retry, credential, or change to the workflow.
>
> c. Declines or holds the action when the authorization determination cannot be completed, or when the information it needs is missing or conflicting.
>
> d. Keeps a Decision Record of each authorization determination, made before the action takes effect, and links it to what then occurred.
>
> e. Identifies which internal policies, limits, contract terms, regulatory requirements, delegated authorities, and approval routes apply to each action; records which were considered, which applied, and which were set aside under a stated order of precedence; holds the action when it cannot be

> determined whether a source applies or when sources conflict without resolution; and re-evaluates affected actions when one of those sources changes.
>
> f. Clears an action only when every applicable condition is met for the actual action, including its target, its values, and how they combine, and holds any later stage of the action that falls outside its Clearance Scope (see Definitions) or what was actually approved.
>
> g. Confirms that authority is current when the action takes effect, including when an approval has expired, been revoked, or been superseded.
>
> h. Requires authority for values that come from content the AI System reads, such as an email, an uploaded document, or a tool response, before those values set a payee, an account, an amount, or a coverage term.
>
> i. Prevents a declined or held action from being achieved by rewording or splitting it, and prevents a retry from repeating an effect.
>
> j. Shows, for each held action, that the reviewer was designated for that action and that what the reviewer saw matches the action that took effect.
>
> k. Requires a new approval before the AI System takes further Consequential Actions after a change to its model or provider, instructions, tools, workflow, delegated authority, or level of human supervision, including any change to its Composition Identifier (see Definitions), and traces a change to a shared component to every AI System that relies on it.
>
> l. Checks a payment again when the authority behind it changes while the payment can still be stopped or amended, reconciles what happens after release, and links each completed payment to the authorization determination that cleared it.
>
> m. Sets the evidence a vendor of a third-party AI System must produce and how often, treats evidence not produced on schedule as a departure from the last evaluated configuration, and allocates by contract the responsibility for each kind of adverse outcome, recognizing that the company remains responsible for the action.
>
> n. Stops pending and further Consequential Actions for one AI System, one type of action, or every AI System that relies on a shared component; returns affected processes to a safe state; identifies which completed effects can be reversed; and tests that procedure.
>
> o. Measures, if it does, the share of its Consequential Actions that took effect without a prior authorization determination or a matching Decision Record, under a versioned method identified by name and version, and reconciles that measurement with its answers to this item.
>
> p. Recognizes amounts from Consequential Actions in its financial records only when they are linked to a valid authorization determination, and withholds or reverses that recognition when an authorization is later found invalid, even if the action was carried out.
>
> q. Obtains a Control Verification Statement (see Definitions) before the AI System takes Consequential Actions, at least annually, and after a change to its Composition Identifier.

Proposed change 6. Add Exhibit B checklist row 4 (core for the lead-in and 4a to 4d)

Where in version 6.0: Exhibit B (Checklist), a new row 4 after existing row 3p, with the same three columns as existing row 3 (“How the Insurance Company…”, “Document Name and Page #”, and “If not specified in the AIS Program, provide details below:”). Proposed rows 4a to 4q follow proposed narrative items 9a to 9q.

Proposed text:

> 4. For AI Systems that can take Consequential Actions, reference the processes and procedures of the Company AIS Program that address the following:

4a. Clears, declines, or holds each Consequential Action before it takes effect, separately from the limits of any credential, account, or payment instrument the AI System uses

4b. Prevents a declined or held action from taking effect through any other tool, route, retry, credential, or change to the workflow

4c. Declines or holds the action when the authorization determination cannot be completed, or when the information it needs is missing or conflicting

4d. Keeps a Decision Record of each authorization determination, made before the action takes effect, and links it to what then occurred

4e. Identifies which internal policies, limits, contract terms, regulatory requirements, delegated authorities, and approval routes apply to each action; records which were considered, applied, and set aside under a stated order of precedence; holds the action when applicability is unclear or unresolved; and re-evaluates affected actions when a source changes

4f. Clears an action only when every applicable condition is met for the actual action, and holds any later stage outside its Clearance Scope or what was actually approved

4g. Confirms that authority is current when the action takes effect, including when an approval has expired, been revoked, or been superseded

4h. Requires authority for values that come from content the AI System reads, such as an email or uploaded document, before those values set a payee, account, amount, or coverage term

4i. Prevents a declined or held action from being achieved by rewording or splitting it, and prevents a retry from repeating an effect

4j. Shows, for each held action, that the reviewer was designated for that action and that what the reviewer saw matches the action that took effect

4k. Requires a new approval before the AI System takes further Consequential Actions after a change to its model or provider, instructions, tools, workflow, delegated authority, or level of human supervision, including any change to its Composition Identifier, and traces a change to a shared component to every AI System that relies on it

4l. Checks a payment again when the authority behind it changes while the payment can still be stopped or amended, reconciles what happens after release, and links each completed payment to the authorization determination that cleared it

4m. Sets the evidence a vendor of a third-party AI System must produce and how often, treats evidence not produced on schedule as a departure from the last evaluated configuration, and allocates by contract the responsibility for each kind of adverse outcome, recognizing that the company remains responsible for the action

4n. Stops pending and further Consequential Actions for one AI System, one type of action, or every AI System that relies on a shared component; returns affected processes to a safe state; identifies which completed effects can be reversed; and tests that procedure

4o. Measures, if it does, the share of its Consequential Actions that took effect without a prior authorization determination or a matching Decision Record, under a versioned method identified by name and version, and reconciles that measurement with its answers to this row

4p. Recognizes amounts from Consequential Actions in its financial records only when linked to a valid authorization determination, and withholds or reverses recognition when an authorization is later found invalid

4q. Obtains a Control Verification Statement before the AI System takes Consequential Actions, at least annually, and after a change to its Composition Identifier

Proposed change 7. Extend Exhibit C, Ref 1 (build-out)

Where in version 6.0: Exhibit C, existing Ref 1, “AI model name and version number.” Replace the final period with the proposed text.

Proposed text:

> ; for an AI Agent (see Definitions), also its current Composition Identifier (see Definitions).

Proposed change 8. Add Exhibit C, Refs 3b to 3j (build-out)

Where in version 6.0: Exhibit C, new Refs 3b to 3j after existing Ref 3a, “Whether the use case(s) involve Agentic AI, and if so, identify its objectives, and authorizations without human approval.”

Proposed text:

> 3b. If the AI System can take Consequential Actions, list each type it can take or cause, each route to each type, the types of credentials or access it holds that can carry them out, and the model provider and other third-party components each route relies on. State how the company classifies an action that fits no listed type. State whether the AI System holds the credentials that carry out its Consequential Actions itself, or whether a separate control holds them and releases them only after the action is cleared.
>
> 3c. Describe how each Consequential Action is cleared, declined, or held before it takes effect, which rules apply to it, and who resolves held actions.
>
> 3d. Describe how the company reconciles its Decision Records with the attempted and completed Consequential Actions recorded by the systems that received or carried them out. Then provide Authorization Evidence (see Definitions) for a sample of Consequential Actions from actual use. The regulator selects the sample, or it is drawn from a population of Decision Records fixed before selection, and it includes at least one declined or held action. Link each Decision Record to what the system that carried out the action accepted or performed. If no action was declined or held in the review period, provide a control test and its result.
>
> 3e. Provide the date and results of the latest authorization tests, including the example authorization tests in the table that follows this one.
>
> 3f. List changes since the last review to the model or its provider, instructions, tools, workflow, delegated authority, or level of human supervision, including changes to a component shared with other AI Systems, and the approval that followed each change.
>
> 3g. Provide the date and findings of the latest test of stopping pending and further actions, and describe which completed effects can be reversed.
>
> 3h. Provide the latest Exposure Measurement (see Definitions) for the AI System, by type of Consequential Action, and state whether it and the sampled Decision Records match the company's other answers about this AI System. State separately whether any untested route could let an action take effect without an authorization determination. If no measurement is available, identify the records and tests that support the company's answers and describe their limits.
>
> 3i. For each type of Consequential Action, describe what its Clearance Scope (see Definitions) covers, including the values, combinations of values, and later stages it permits, and how a later stage that falls outside it is held.
>
> 3j. If a vendor or other third party carries out Consequential Actions for the company, state whether the company confirms authorization itself before the action takes effect, rather than relying on the vendor's statement that it was authorized. State what evidence the vendor must produce and how often, and what the company does when that evidence is not produced, such as suspending the vendor's ability to take Consequential Actions.

Proposed change 9. Add Exhibit C, Ref 11a (build-out)

Where in version 6.0: Exhibit C, a new Ref 11a after existing Ref 11, which asks for the last date of model testing or validation.

Proposed text:

> 11a. If the AI System can take Consequential Actions, provide the latest Control Verification Statement (see Definitions), including its date, who performed it, its scope, and its findings. State whether verification is repeated at least annually and after a change to the AI System's Composition Identifier.

Proposed change 10. Add Exhibit C, Ref 12a (build-out)

Where in version 6.0: Exhibit C, a new Ref 12a after existing Ref 12, “Discuss how the model affects the financial statements, risk assessment or controls.”

Proposed text:

> 12a. If the AI System can take Consequential Actions, describe how amounts from those actions reach the financial statements, whether each amount is linked to its Decision Record, and how amounts from actions later found to be unauthorized are identified and corrected.

Proposed change 11. Add a table of example authorization tests (build-out)

Where in version 6.0: a new table after the existing Exhibit C table and before Exhibit D, titled “Exhibit C: Example Authorization Tests.” Proposed Ref 3e refers to it.

Proposed text:

> Tests 1 to 8 check that an action without current authority does not take effect. Test 9 checks that a permitted action is cleared.

Test	Condition	Expected result
1	Authority is revoked or expires after the action is planned but before it takes effect, such as an adjuster's payment authority withdrawn during a claim	The action does not take effect on the revoked or expired authority. A new authorization determination may clear it.
2	The action or its values change after approval	Not taken on the earlier approval. A new authorization determination is required.
3	An approval the action needs is missing, or comes from someone without authority for that action	Declined or held
4	The action exceeds its authorized scope or a limit, including values that are each allowed but were never authorized together	Declined or held
5	A retry, another tool, a direct call, or another route attempts a declined or held action while the reason for the decline or hold is unresolved and no new authorization exists	Still declined or held
6	Content the AI System read, such as an email with new bank details, is used in place of the authority or verification the action requires to set a payee, account, amount, or coverage term	Declined or held until the required authority or verification is obtained

Test	Condition	Expected result
7	The authorization determination cannot be completed because of a timeout, missing evidence, a missing rule, or conflicting rules	Declined or held. The action does not take effect.
8	Split, simultaneous, or repeated actions together exceed a limit, or a retry would repeat an effect that was authorized once	The excess or duplicate action is declined or held.
9	An action within valid, current authority	Cleared by the authorization control, subject to other applicable controls. What then occurs is recorded separately.

Proposed change 12. Add a table of Decision Record data to Exhibit D (build-out)

Where in version 6.0: the end of existing Exhibit D, AI Model Data, before the existing Definitions and Appendix. Title the new table “Exhibit D: Decision Record Data.”

Proposed text:

> For AI Systems that can take Consequential Actions, provide the Decision Records for the sample in Exhibit C, Ref 3d, with the fields below. State which fields the company does not keep.

#	Field	What it shows
1	Record identifier, date, and time	When the authorization determination was made, before the action took effect
2	AI System and Composition Identifier	Which AI System proposed the action, and its model, instructions, tools, and delegated authority at that time
3	Action as proposed	The type of Consequential Action, its target, and its values, such as the payee, account, amount, or coverage term
4	Source of each value	For each value that sets a payee, account, amount, or coverage term, the input it came from and whether that input carried authority for this use
5	Governing sources	The policies, limits, contract terms, regulatory requirements, delegated authorities, and approval routes considered; which applied; which were set aside; and the version of the order of precedence used
6	Conditions evaluated	Each condition evaluated and its result, including any condition that could not be evaluated
7	Determination	Cleared, declined, or held
8	Decided by	The rule, or the human approver and the authority under which the approver acted
9	Clearance Scope	For a cleared action, the values and later stages the clearance covers
10	Review of a held action	The designated reviewer, what the reviewer was shown, the decision, and its date and time
11	What then occurred	What the receiving system accepted or performed, including the payment or settlement record, or that nothing occurred
12	Financial entry	Where applicable, the financial-record entry linked to the action

Proposed changes 13 to 16, the definitions

Execution Rights defined the term consequence authorization. The proposed definitions of Authorization Evidence, Consequence Authorization, and Decision Record reproduce Execution Rights' wording exactly after the lead-in “refers to.” In those definitions, “agent” and “AI agent” mean an AI Agent, and “consequential action” means a Consequential Action. Allow, deny, and hold correspond to cleared, declined, and held. A Decision Record is completed with the outcome once the action takes effect or does not.

Proposed change 13 (core). Where in version 6.0: Definitions and Appendix, after existing "Algorithm" and before existing "AI Model."

Proposed text:

> **"AI Agent"** refers to an AI System that can take Consequential Actions, whether or not it uses Agentic AI.

Proposed change 14 (build-out). Where in version 6.0: Definitions and Appendix, after existing "Augmentation" and before existing "Automation."

Proposed text:

> **"Authorization Evidence"** refers to the body of decision records and related artifacts sufficient for a party other than the operator to verify, after the fact, that consequential actions were authorized before execution: complete (a record exists for every in-scope action, including declines and holds), correspondent (each executed action matches its record, and declined actions produced no execution), and reconstructable (the decision, its basis, and its authority can be re-derived from the record alone). A tag, flag, or metadata assertion of authorized status is not authorization evidence; authorized status derives only from an authorization determination rendered against current authority, as reflected in the decision record.

Proposed change 15 (core for Consequence Authorization, Consequential Action, and Decision Record; build-out for the others). Where in version 6.0: Definitions and Appendix, after existing "Automation" and before existing "Direct Consumer Impact." Six entries, in this order.

Proposed text:

> **"Clearance Scope"** refers to the values, and the combinations of values, evaluated when a Consequential Action is cleared, together with the later stages of the action that the clearance permits. A value, combination, or stage outside the Clearance Scope is not cleared and is held until a new authorization determination is made.
>
> **"Composition Identifier"** refers to an identifier computed over an AI Agent's model, instructions, tools, and delegated authority, so that a change to any of them produces a different identifier.
>
> **"Consequence Authorization"** refers to the pre-execution layer that decides whether an AI agent's action is allowed to become real, and the agent cannot bypass the layer itself: no pathway reaches the consequence except through the authorization. Before a consequential action executes, it runs a composite authorization across spending authority, contractual rights, policy, regulation, delegated authority, and human-approval routing, and returns one of three decisions: allow (cleared to execute), deny (declined, and the agent cannot route around it), or hold (routed to a human for approval). Every decision produces a record of what was authorized and the basis for it.
>
> **"Consequential Action"** refers to an action an AI System takes or causes, directly or through any tool, route, or other system, that (a) moves money or commits the company to a payment; (b) changes coverage or the terms of a policy, the status or amount of a claim, a premium, or a financial record; (c) sends a consumer, claimant, or other party outside the company a communication that grants, denies, or changes coverage, a claim, a payment, or a deadline; or (d) discloses or transfers nonpublic information outside the company, or deletes it. An action is a Consequential Action whether or not a person approved it, and whatever the AI System or its function is called. An action that is attempted but declined or held remains a Consequential Action.
>
> **"Control Verification Statement"** refers to a dated record of an independent test of whether Consequence Authorization is in the path of every Consequential Action of an AI System. It identifies the controls tested, the evidence examined, the findings, and the conditions for continued reliance.

> **"Decision Record"** refers to the per-action evidence artifact of consequence authorization. Created at decision time — before the consequence becomes real — it captures the action as proposed, the authorization decision rendered (allow, deny, or hold), the dimensions evaluated, the rule or human approver that decided, the authority it was decided under, and the outcome. It exists for every decision, including declines and holds, so refusals are as provable as approvals.

Proposed change 16 (build-out). Where in version 6.0: Definitions and Appendix, after existing "Direct Consumer Impact" and before existing "Externally Trained Models."

Proposed text:

> **"Exposure Measurement"** refers to a statement of the share of an AI System's Consequential Actions of a given type that took effect without a prior Consequence Authorization determination, without a matching Decision Record, or both. It is computed under a versioned method that the company identifies by name and version, so that measurements made under the same version, over comparable scopes and periods, can be compared. It states its result, the method and its version, how it was produced, which inputs were observed and which were declared, the AI Systems it covers, and its date.

ATTACHMENT B

Example Examination Procedures

These example procedures show how an examiner could use the proposed questions in a limited-scope exam of AI Systems that take Consequential Actions. The Working Group may adapt them for examiner guidance or a workpaper. References are to the proposed changes in Attachment A.

1. **Set the scope.** Compare proposed Exhibit A columns (9) and (10). Select AI Systems that can take Consequential Actions, giving priority to any without Consequence Authorization in place.
2. **Fix the population.** Obtain the full population of Decision Records for the review period, with a count or other identifier fixed before any sample is selected (proposed Ref 3d).
3. **Select the sample.** Select Decision Records from that population. Include at least one declined or held action.
4. **Reconcile.** Compare the actions recorded by the systems that carried them out, such as payment and policy systems, with the Decision Records. Treat any completed action without a matching Decision Record as an exception.
5. **Re-perform.** For each sampled record, rebuild the determination from the record alone, using the fields in proposed Exhibit D. Confirm the conditions, rules, and authority in effect support the recorded result.
6. **Test routes around the control.** With the company, attempt a declined or held action again through a retry, another tool, or a direct call. Expect it to remain declined or held (proposed example tests 5 and 8).
7. **Test failure.** Make the authorization determination unable to complete. Expect the action not to take effect (proposed example test 7).
8. **Test changes.** Compare the history of each AI Agent's Composition Identifier with the approvals the company recorded (existing Ref 1 as proposed to be extended, and proposed Ref 3f).
9. **Review verification and stop testing.** Review the latest Control Verification Statement and the latest stop test (proposed Refs 11a and 3g).
10. **Trace to the financial statements.** Trace a sample of financial entries arising from Consequential Actions to their Decision Records (proposed Ref 12a).
11. **Report.** Summarize exceptions by type, AI System, and type of Consequential Action.

BlueCross BlueShield Association
750 9th Street, NW
Washington, DC 20001
www.bcbs.com

October 30, 2026

Commissioner Nathan Houdek, Chair
Commissioner Doug Ommen, Co-Vice Chair
Deputy Commissioner Mary Block, Co-Vice Chair
Big Data and Artificial Intelligence (H) Working Group
National Association of Insurance Commissioners
1100 Walnut Street, Suite 1000
Kansas City, MO 64106-2197

Submitted via email to Scott Sobel (ssobel@naic.org) and Miguel Romero (maromero@naic.org)

RE: Artificial Intelligence Risk Evaluation Supplement Version 6.0

Dear Chair Houdek and Vice Chairs Ommen and Block:

The Blue Cross Blue Shield Association (BCBSA) appreciates the opportunity to provide comments on the National Association of Insurance Commissioners (NAIC)'s Artificial Intelligence (AI) Risk Evaluation Supplement version 6.0.

BCBSA is a national federation of independent, community-based and locally operated Blue Cross and Blue Shield companies (Plans) that collectively provide health care coverage for one in three Americans. BCBS Plans contract with 97% of hospitals and 83% of doctors across the country and serve those who are covered through Medicare, Medicaid, an employer or purchase coverage on their own.

BCBSA shares the Working Group's commitment to responsible AI use that serves consumers. Patients should be able to trust that the AI tools used to serve them are fair, accurate and well governed. Used responsibly, AI can improve health outcomes, relieve administrative burdens and lower health care costs. Focusing the Supplement on the AI posing the greatest risk to consumers will help regulators and Plans advance these shared goals.

We commend the collaborative process the Working Group has led throughout the development of the Supplement and for the ongoing consideration of stakeholder feedback. Version 6.0 reflects that thoughtful engagement and is a stronger and more focused resource. We thank the Working Group for including several of the considerations we – and other stakeholders – raised on Version 5.0.

Among these improvements, Version 6.0 removes the separate counting of generalized linear models in Exhibit A and allows regulators to exclude some or all generalized linear models from Exhibits B through D where appropriate. These changes keep Exhibit A more focused on the generative and agentic applications that warrant regulatory attention. We also appreciate that Exhibit C remains directed at high-risk models and no longer instructs companies to include all

companies and lines of business, which keeps its detailed, model-by-model reporting focused on the models that present the greatest risk. We welcome the new General Guidance distinguishing AI systems from AI models, which responds to our request for clarity and helps companies and examiners apply the exhibits more consistently. Finally, the removal of the staffing-reduction reference from the Exhibit B narrative question on direct consumer impact keeps that question focused on the factors most relevant to whether an AI system's output affects a specific consumer.

BCBSA also supports two new additions in Version 6.0 and encourages the Working Group to retain them: the statement that, consistent with the Handbooks, the Supplement does not independently create substantive legal duties, standards of conduct, admissions, or expanded regulatory authority; and the clarification that regulators would use either the Narrative or the Checklist version of Exhibit B, but not both, which avoids duplicative reporting without reducing the information available to regulators.

Recommendations on New Content in Version 6.0

BCBSA offers two recommendations on content introduced in Version 6.0, intended to keep the Supplement's new terms and questions precise and appropriately scoped so that companies respond consistently and regulators receive more useful information.

I. Refine the definition of "Adverse Consumer Impact."

Version 6.0 revised this definition to describe an AI System output that "adversely impacts the consumer," removing the prior qualifier that the impact violate the applicable regulatory standards. Because the term now anchors several of the expanded testing questions in Exhibit C, including new questions 10 and 10a, it could be read to capture routine, appropriate outcomes. For consumers, a risk evaluation tool is valuable when it surfaces the outcomes that warrant attention. Removing the qualifier may inadvertently bring more outcomes into scope which could limit the supplement's ability to identify risks that could harm consumers. At the same time, limiting the definition to outcomes that already violate regulatory standards could narrow the Supplement's ability to identify risks before they result in harm.

BCBSA recommends that the Working Group revise the definition so it reaches an AI System output that adversely impacts a consumer and is inconsistent with applicable regulatory standards or with the established criteria the company applies to that decision. Like the NAIC Model Bulletin on the Use of Artificial Intelligence Systems by Insurers (Model Bulletin), which defines adverse outcomes by reference to the regulatory standards a department enforces, this approach would keep the definition grounded in those standards, while also allowing the Supplement to surface errors and emerging risks before they result in harm. It would also exclude appropriate outcomes that do not signal risk, keeping companies' responses and regulators' attention focused on the outcomes that matter most.

II. Remove new Exhibit C question 3a on Agentic AI.

BCBSA shares the Working Group's interest in understanding the distinct considerations that agentic systems present. However, we recommend that the Working Group remove question 3a, because the agentic AI considerations it addresses are already captured by Exhibit C questions 2, 3 and 9 and by the Exhibit B question on autonomy and

reversibility. The autonomy that question 3a's phrase "authorizations without human approval" is reaching for is the automation concept the Supplement already defines and asks about in question 9 and the Exhibit B autonomy question. Question 3a adds complexity and invites inconsistent responses but is not likely to produce information regulators would not already obtain.

Priority Recommendations Carried Forward from Version 5.0

BCBSA's Version 5.0 comments identified areas where further refinement would strengthen the Supplement and better serve its purpose of helping regulators identify and assess the risks arising from the use of AI systems. We offer the most significant below, grouped by objective. These recommendations are intended to focus the Supplement on the AI that most affects consumers, protect the information it collects, promote consistent application and reinforce its grounding in the NAIC's existing processes.

III. Further focus the Supplement on AI with material consumer impact.

In its Model Bulletin on the Use of Artificial Intelligence Systems by Insurers, NAIC promotes oversight that is commensurate with the potential for harm, which we agree is an optimal approach to managing risk effectively. Similarly, concentrating the Supplement on AI to uses that materially affect consumer outcomes directs regulatory attention where it matters most and produces more useful information for regulators. Accordingly, BCBSA recommends that the Working Group:

- **Replace "Direct Consumer Impact" with "Material Consumer Impact"** throughout the Supplement, so the term better matches the Supplement's own definition and parallels the existing "Material Financial Impact" phrase. If the current term is retained, make the materiality concept explicit and operative in the definition rather than expressing it only as a matter of "less applicability."
- **Confirm in the Instructions that the Supplement is intended to evaluate the governance and risk management of AI systems rather than collect or compare enterprise AI inventories across companies**. Making this explicit would help ensure the Supplement is applied consistently with that intent, and it would reflect the practical reality that AI inventories are dynamic and depend heavily on how each organization defines, groups and tracks its AI assets, which makes them appropriate for a risk assessment but of limited value for direct comparison across companies. This clarification is especially useful now that Version 6.0 asks companies to submit an inventory together with Exhibit A.

IV. Protect the sensitive and proprietary information the Supplement draws in.

The Supplement collects sensitive governance, model and vendor information, much of it supplied by third-party vendors that retain proprietary rights over their models and impose confidentiality obligations. Clear protections and realistic expectations on what can be shared and how the information will be used is necessary to enable effective reporting and maintain trust between regulators and companies. To support these goals, BCBSA recommends that the Working Group:

- **Strengthen the Confidentiality section to specify how submitted information will be maintained, whether and how it may be shared across examiners and states**

and confirm that it will not be made public. This is increasingly important given the new requests to identify subject matter experts in Exhibit B and to indicate who prepared and reviewed the responses in Exhibit C.

- **Make the disclosure of third-party data sources and vendor names in Exhibit D optional or accommodate contractual restrictions**, because some vendor contracts bar disclosure of the vendor's identity outside confidential examination settings. These vendors often supply specialized data and capabilities that Plans rely on. Exhibit D already captures the type of data used, how the model uses it and whether it is sourced from a third party, so regulators can assess the data's risk without the vendor's name in every response. A mandatory requirement could deter companies from engaging vendors best suited for a given purpose, narrowing the pool of available vendors and limiting innovation that is in the best interest of members. Making vendor identification optional, or expressly accommodating contractual restrictions, would avoid that outcome while still allowing regulators to obtain the information through confidential examination channels where needed.
- **Confirm that validation and testing of third-party, vendor-supplied AI may be satisfied by either the vendor or the insurer**, including where AI is supplied by a fourth party such as a subcontractor. The insurer remains accountable for the AI it uses, and this clarification concerns how that responsibility may be met, not whether it applies. Because vendors retain proprietary rights over their models and insurers often cannot access the technical detail needed to test every element, insurers should be able to satisfy their oversight obligations by relying on and evaluating vendor testing rather than duplicate work they cannot reasonably perform. This reflects how third-party AI oversight works in practice while keeping responsibility for the AI with the insurer.

V. Promote a consistent, baseline application of the Supplement across states.

BCBSA recognizes that individual states need the flexibility to tailor the Supplement to their specific regulatory needs. At the same time, a shared baseline lowers the burden of implementing the Supplement for states while leaving each free to adapt it while also providing responding companies with more consistency, driving more meaningful reporting states can act on. We appreciate the new language encouraging regulators to coordinate information requests and encourage the Working Group to build on it.

Accordingly, BCBSA recommends that the Working Group establish a baseline response timeframe of at least 30 business days, with extensions available upon request. A single response can require input from governance, compliance, risk management, actuarial, IT, vendor management and finance teams, followed by review before it is submitted. During the pilot, some states required responses in fewer than two weeks, which leaves little time for that coordination.

VI. Coordinate the Supplement's implementation through existing NAIC regulatory frameworks.

The Supplement will be most effective when implemented through existing market conduct and financial regulatory processes rather than as a parallel framework. Version 6.0's reinforced grounding in the Handbooks appropriately reflects that objective.

BCBSA recommends that the Working Group refer the Supplement jointly to the Market Regulation and Consumer Affairs (D) Committee and the Financial Condition (E) Committee for timely, coordinated implementation through the relevant Handbooks. The Handbooks should address when use of the Supplement is appropriate, how inquiries should be scoped based on risk and materiality and how its use should be coordinated across regulatory functions and jurisdictions. This approach follows NAIC precedent, including the referral of the Own Risk and Solvency Assessment (ORSA) framework to the groups responsible for the financial Handbooks. A joint referral would promote consistent application of materiality, inherent risk, examination, and accreditation principles; reduce the potential for conflicting interpretations or duplicative oversight; advance the NAIC Collaboration Forum's objectives; and preserve state flexibility. BCBSA is not recommending that the Working Group delay adoption of the Supplement, but rather that it coordinate implementation within existing regulatory frameworks through a timely referral.

Thank you again for your continued engagement with stakeholders and for the opportunity to comment on Version 6.0 of the AI Risk Evaluation Supplement. If you have any questions, please contact Randi Chapman at Randi.Chapman@bcbsa.com.

Sincerely,



Clay S. McClure
Senior Director, State Affairs
Blue Cross Blue Shield Association

PUBLIC COMMENT SUBMISSION

THE EVIDENTIARY VALIDATION AND CONSISTENCY FRAMEWORK (EVAF) v2.0

A Risk-Based Evidence Framework for the AI Risk Evaluation Supplement Version 6.0

Submitted To:	NAIC Big Data and Artificial Intelligence (H) Working Group
Submitted By:	Mebrat Hill Premium Audit Consultant & Insurance Assurance Practitioner Oakland, California
Date:	October 30, 2026

COVER LETTER

October 30, 2026

Members of the Big Data and Artificial Intelligence (H) Working Group:
NAIC Big Data and Artificial Intelligence (H) Working Group

RE: Public Comment Submission – AI Risk Evaluation Supplement Version 6.0 Exposure Draft

Thank you for the opportunity to comment on the AI Risk Evaluation Supplement Version 6.0 Exposure Draft.

I submit these comments as an insurance risk and assurance practitioner with more than twenty years of experience evaluating insurance operations, governance controls, regulatory requirements, evidentiary standards, risk management practices, and audit defensibility.

I appreciate the Working Group's continued development of a practical, risk-based framework for evaluating insurer use of AI Systems. The Version 6.0 Exposure Draft appropriately emphasizes materiality, proportionality, governance, and supervisory flexibility while recognizing the wide range of insurer sizes, operational models, and AI use cases throughout the industry.

Attached for consideration is the Evidentiary Validation and Consistency Framework (EVAF) v2.0, a technology-neutral governance and examination framework intended to complement the existing structure of the Supplement. The recommendations focus on strengthening examination consistency, improving traceability across exhibits, reinforcing evidence-based governance, and supporting proportionate supervision through existing insurer records whenever appropriate.

The framework is organized around several core concepts:

- Evidence-Based Governance (Policy → Control → Evidence)
- Decision Lineage & Auditability
- Human Authority and Effective Oversight
- Cross-Exhibit Traceability
- Third-Party Accountability
- Risk-Based Proportionality

The recommendations are not intended to prescribe any specific technology, architecture, software platform, governance system, or implementation methodology. Rather, they are intended to help regulators and insurers consistently evaluate a fundamental supervisory question:

> **"Can the insurer demonstrate how a material AI-supported decision was governed, reviewed, controlled, and supported by evidence?"**

I respectfully submit these recommendations for the Working Group's consideration and appreciate the significant effort invested in the continued evolution of the AI Risk Evaluation Supplement.

Respectfully submitted,

Mebrat Hill
Premium Audit Consultant & Insurance Assurance Practitioner

TABLE OF CONTENTS

1. Executive Summary

2. Technology-Neutrality and Implementation Statement

3. Formal Recommendations to the Working Group

4. Governance Should Be Demonstrable Through Evidence

5. Decision Lineage

6. Human Authority and Effective Oversight

7. Cross-Exhibit Traceability

8. Third-Party Accountability

9. Risk-Based Proportionality

10. EVAF Examiner Worksheet

11. Closing Observation

1. Executive Summary

I appreciate the Working Group's continued effort to develop a practical, risk-based framework for evaluating insurer use of AI Systems while preserving flexibility across varying company sizes, use cases, and regulatory contexts. The Version 6.0 draft appropriately emphasizes materiality, inherent risk, proportionality, and coordination across regulatory functions.

This submission recommends a technology-neutral evidentiary framework designed to strengthen regulatory review without creating new substantive obligations or prescribing specific technical architectures.

Governance Should Be Demonstrable, Not Merely Documented

For a material AI-supported decision, an examiner should be able to determine:

- What governed the decision;
- What controls applied;
- Whether those controls operated;
- What authority existed to intervene; and
- What evidence demonstrates what actually occurred.

The proposed EVAF framework is built around three operational concepts:

POLICY → CONTROL → EVIDENCE

- **Policy** establishes what should happen.
- **Control** establishes what is designed to happen.
- **Evidence** establishes what actually happened.

2. Technology-Neutrality and Implementation Statement

The concepts presented in this submission are intended as technology-neutral governance and examination principles. The recommendations are designed to describe supervisory outcomes and evidentiary considerations rather than prescribe any particular implementation method, software architecture, workflow design, governance platform, validation methodology, or technical solution.

Organizations may implement these concepts using approaches appropriate to their size, complexity, risk profile, and operational environment.

> **NOTICE REGARDING IMPLEMENTATION METHODS**
> The concepts described in this submission are intended as governance and examination principles. Any implementation methods, architectures, workflows, algorithms, validation mechanisms, or technical approaches that may support such principles are outside the scope of this submission and are not disclosed herein.

3. Formal Recommendations to the Working Group

Recommendation 0: Adopt an Evidence-Based Governance Principle

For material AI-supported decisions, examination should focus not only on the existence of policies, controls, models, and governance structures, but also on the evidence available to demonstrate how relevant controls operated in practice and how the resulting decision was governed, reviewed, and supported.

Recommendation 1: Distinguish Control Design, Control Operation, and Evidence

Examinations should distinguish between: (1) Whether a control exists; (2) Whether it is appropriately designed; and (3) Whether evidence demonstrates it operated when required.

Recommendation 2: Recognize Decision Lineage as an Examination Consideration

For material AI-supported decisions, regulators may consider whether sufficient records exist to reconstruct the basis for the decision, including applicable requirements, material data, contributing systems or models, required reviews, exceptions, escalations, and final disposition.

Recommendation 3: Clarify That Human Oversight Means Human Authority

Human review should be evaluated based on the reviewer's authority, available information, ability to challenge or reject outputs, and escalation responsibilities rather than merely the existence of human involvement.

Recommendation 4: Establish a Cross-Exhibit Traceability Principle

AI System → Model → Data → Governing Requirement → Control → Human Authority → Decision Basis → Evidence

Recommendation 5: Encourage the Use of Existing Corporate Records

Insurers should be permitted to leverage existing model inventories, Model Risk Management documentation, internal audit records, ORSA/ERM materials, governance documentation, validation records, and change-management records.

Recommendation 6: Apply Proportionality Consistently Across All Exhibits

Examination depth should be proportionate to materiality, operational scale, degree of automation, reversibility, consumer impact, and financial impact.

Recommendation 7: Preserve Accountability for Third-Party AI Systems

The use of third-party tools should not eliminate accountability for resulting insurance decisions.

Recommendation 8: Consider Failure, Exception, and Change Evaluation

Examiners should have the flexibility to evaluate exceptions, overrides, data quality issues, model changes, third-party disruptions, and remediation procedures.

4. Governance Should Be Demonstrable Through Evidence

The existence of a governance framework, model inventory, review requirement, or documented control does not by itself establish that governance operated effectively when a material decision was made.

• Policy = What the organization says should happen.
• Control = What systems and personnel are designed to permit, prevent, require, or escalate.
• Evidence = What records demonstrate what actually happened.

Recommended Examination Question:
Can the insurer demonstrate, using evidence proportionate to risk and materiality, that relevant governance controls operated as intended?

5. Decision Lineage

For a material AI-supported decision, an insurer should be able to identify:
• Applicable governing requirements;
• Relevant policy version or effective period;
• Material data relied upon;
• Material limitations or exceptions;
• AI systems and models contributing to the outcome;
• Required human reviews or approvals;
• Escalations and interventions;
• Final disposition; and
• Records sufficient to reconstruct the basis for the decision.

The insurer should be able to demonstrate the basis upon which a material AI-supported decision was reached and governed.

6. Human Authority and Effective Oversight

For material use cases, examination may benefit from evaluating reviewer authority, the ability to challenge outputs, bypass restrictions, defined escalation tracks, and fallback controls.

Recommended Examination Question:
What authority did the reviewer possess, what evidence was available, and what control operated when required review did not occur?

7. Cross-Exhibit Traceability

EVAF TRACEABILITY PRINCIPLE:
AI System → Model → Data → Governing Requirement → Control → Human Authority → Decision Basis → Evidence

This framework allows examiners to follow the basis of a decision from origin through outcome while preserving flexibility regarding tools, platforms, databases, and documentation formats.

8. Third-Party Accountability

Where reasonably available and proportionate to risk, examination evidence may identify third-party functions involved, applicable oversight mechanisms, known limitations, accountability structures, material changes, and governance over resulting decisions.

9. Risk-Based Proportionality

Examination depth should scale with:
• Materiality: Magnitude of potential consumer, financial, or regulatory impact.
• Operational Scale: Volume and frequency of use.
• Automation Severity: Extent to which automation influences outcomes.
• Reversibility: Ease with which decisions may be corrected.

Higher-risk combinations warrant greater evidence, while lower-risk applications require lighter expectations.

10. EVAF Examiner Worksheet

This functional template serves as a practical assessment tool for examiners to map policy intents to verifiable evidence footprints during an AI risk evaluation.

EVAF Pillar	Core Regulatory Focus	Examiner Evaluation Criteria	Verifiable Evidence Footprint
1. Evidence-Based Governance	Policy -> Control -> Evidence operation.	Verify that explicit corporate policies align directly with automated enforcement triggers and generate persistent system logs.	System run-time logs, audit trail exports, exception tracking dashboards.
2. Decision Lineage	Reconstructability of AI decisions.	Assess if historical inputs, specific model versions, and environmental parameters can recreate a distinct automated outcome.	Version-controlled data schemas, model frozen states, step-by-step transaction traces.
3. Human Authority	Effective human veto & intervention power.	Determine whether human overseers have the practical technical tools and explicit mandate to overrule or reject AI recommendations.	Override authorization logs, formal escalation tickets, documented human rejection histories.
4. Traceability	Cross-exhibit linking & visibility.	Track data transformations seamlessly from initial data ingestion up to the final business output across regulatory disclosures.	Data lineage maps, systemic API mapping records, cross-referenced exhibit IDs.
5. Third-Party Governance	Vendor system accountability.	Review vendor SLAs, validation disclosures, independent audit reports, and local edge monitoring controls over external systems.	SOC 2 Type II reports, third-party model performance validations, API failure alerts.

11. Closing Observation

The objective of effective AI governance is not merely to document policies, inventories, and oversight structures. The objective is to demonstrate that governance operates effectively in practice.

A model may be explainable, a governance program may be documented, and a control may be well designed—yet regulators may still be unable to determine whether governance operated effectively for a material AI-supported decision.

EVAF seeks to provide a practical, technology-neutral, and proportionate framework for answering a fundamental supervisory question: 'Can the insurer demonstrate how a material AI-supported decision was governed, reviewed, controlled, and supported by evidence?'

If that question can be answered consistently and proportionately using records insurers already maintain, the AI Risk Evaluation Supplement will be better positioned to provide meaningful oversight while preserving innovation, flexibility, and risk-based supervision.