

Comment on the AI Risk Evaluation Supplement, Version 5.0

Exhibit B Is the Oversight Exhibit

September 26, 2026

Via email to ssobel@naic.org and maromero@naic.org

Commissioner Nathan Houdek, Chair Commissioner Doug Ommen, Co-Vice Chair Mary Block, Co-Vice Chair Big Data and Artificial Intelligence (H) Working Group National Association of Insurance Commissioners

Re: AI Risk Evaluation Supplement, Version 5.0 (exposed August 31, 2026)

Chair Houdek, Co-Vice Chairs Ommen and Block:

The Center for AI Oversight submits these comments on version 5.0 of the AI Risk Evaluation Supplement. A one-page summary follows, then the comment itself, and proposed additions to Exhibit B are attached separately in the exhibit's own format with an editable version for staff use. We would welcome a few minutes on the October 8 call to present the recommendations, and are glad to speak with staff beforehand.

Respectfully submitted,

Brian J. Allen, Esq. Executive Director, Center for AI Oversight info@cfaio.org

Summary

The finding. The Supplement is already built in three layers. Exhibits A, C and D ask what the insurer built, bought and tested. Exhibit B asks who chartered the program and who is accountable for it. Exhibit B is the oversight exhibit. It is distinct in kind from the other three rather than a higher-altitude version of them, and it is the only part of the instrument that will still be true at the next examination, because oversight decisions attach to the company's decision process rather than to the models, use cases and risks that change continuously beneath them.

The recommendation. Complete Exhibit B. Five recommendations, expressed as one amendment to item 1 and seven new Checklist entries, plus one optional item. All are answerable from a document, none requires an examiner to evaluate a model, and none prescribes governance content.

#	Addition	Exhibit	What it fixes
1	What the program governs, its granularity, who approved that as appropriate, and the	B, item 1	A program scoped by decision and one scoped by default answer identically

#	Addition	Exhibit	What it fixes
	standards elected		
2	Accountability at two levels, declared exclusions, and competency	B, 2 series	"Board of Directors or management" can be answered accurately and tell the examiner nothing
3	Ratified AI risk tolerance, with body and dates	B, 3 series	A tolerance that was calculated but never ratified is a working assumption
4	Authority to operate, dated, scoped and expiring	B, 3 series	Automating a decision is a grant of authority, not a design outcome
5	The escalation and disclosure line, with a destination and a record	B, 3 series	Nothing establishes what happens when a system runs outside tolerance

Plus four fields in Exhibit C recording who approved the company's risk criteria, and one change separating what a vendor's officer can attest to from what requires assurance.

The Supplement is already built in three layers

OVERSIGHT GOVERNANCE	What is governed, at what depth, and who answers. Chartered once, then operated.	Still true at the next examination
Exhibit B <i>Five additions proposed</i>		
MANAGEMENT	What the company does: model inventory, use cases, risk criteria, vendor sourcing.	Accurate on the day collected
Exhibits A and C		
TECHNICAL CONTROLS	How the systems are built and tested: testing methodology, data element categories.	Accurate on the day collected
Exhibits C and D		

Exhibits A, C and D ask what the insurer built, bought and tested. Exhibit B asks who chartered the program and who answers for it.

Exhibit B is the oversight exhibit, and it is the only part of the instrument that does not decay between examinations.

Why it matters. This instrument is used inside examinations that run on statutory cycles of up to five years, and in Florida market conduct examinations run "as often as it deems necessary." Exhibit A's model inventory is accurate on the day it is collected and materially stale by the next cycle. The decisions Exhibit B records are stable by nature. When they do change, the program carries the record of the change.

The program is the examinable object. A committee is a governing body. What it governs is a program, and the program does not govern the technology. It governs the company's capacity to make informed decisions about its use of AI. An examiner cannot examine a committee's diligence and should not be asked to examine a model. An examiner can examine a program, with the questions examiners have asked of compliance and anti-fraud programs for decades.

None of it is a new disclosure. Where AI is mission critical, the duty to maintain a board-level monitoring and reporting system already attaches under the Caremark line, applied to mission-critical risk in *Marchand v. Barnhill*. A company that has met it can answer these fields from documents it already keeps.

The mechanism exists, and extending it means naming the object. Insurance regulation names accountable officers twice: the CGAD signature under Model Act #305 Section 3.B and the ORSA Summary Report signature under Part C of the Guidance Manual's Introduction. Neither reaches AI, and neither states what the officer is accountable for with enough precision to examine. Extending them is not a matter of adding a third signature. It is a matter of naming the object: the AI systems program, and what that program governs.

The precedent is recent, and it is cyber. The 2014 Cybersecurity Framework had five functions and no GOVERN. The oversight layer went unbuilt for a decade, was then closed from outside the standards community, and is still not fully closed. This Working Group can put the layer in its own instrument now, on its own terms.

If only two additions are taken, take 2 and 5. With any system that automates a decision, add 4.

What we do not comment on. Exhibit A, Exhibit C's testing methodology fields, and Exhibit D. Those concern how an insurer builds, tests and sources its systems. They belong to the operator and the examiner.

1. Scope

The Center for AI Oversight is an independent nonprofit research and education organization. It defines oversight governance for AI in regulated industries: what must be governed and who is accountable for it. It does not prescribe how AI systems are built, tested, validated or run.

This comment observes that boundary. It addresses Exhibit B, two provisions in Exhibit C, and the attestation in the companion *Regulatory Framework for Third-Party Data and Model Vendors*. The Center takes no positions on matters involving specific institutions.

2. Exhibit B is the oversight exhibit

Three layers, and why the distinction is not academic

Technical controls govern how the system is built, constrained and watched. **Management** governs how the organization pursues its AI objectives and manages the risk of doing so; ISO/IEC 42001 is written for this layer, with the NIST AI Risk Management Framework supplying the risk practice inside it. **Oversight governance** sits above both. It decides what is governed, at what depth, and who answers, and it never decides how.

Management here names an activity, not a rank. Setting objectives, prioritizing, resourcing and treating risk are management wherever performed, including at the top of the company. What makes a decision oversight is not the seniority of the person making it, but that it settles what is governed and who answers for it.

The Supplement already reflects this

Exhibits A, C and D sit at the management and technical layers. Exhibit B sits at the oversight layer, in both of the options version 5.0 offers. The Narrative asks the company to provide its AIS Program and to discuss the governance structure, Board reporting and frequency, how responsibility for governance is assigned, and the integration of AI Systems into the ORSA and ERM assessments. The Checklist asks whether a written AIS Program was adopted, whether the Board of Directors or management was involved in adopting it and what their role in it is, and then runs the 3 series across ERM at 3f, ORSA at 3g, vendor procurement at 3l, and vendor-developed systems at the new 3p. Every one of those is an oversight question.

Version 5.0 moved further in the same direction. The Checklist now carries a Document Name and Page # column, and the instructions direct that "if the responses to the checklist 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." That is the working principle behind every field proposed here: an oversight question should be answerable from a document the company either holds or does not, and the absence of the document is itself the finding.

This comment does not ask the Working Group to adopt an outside framework. It asks the Working Group to finish the exhibit it already wrote.

What the exhibit is examining is a program

Most companies respond to AI by forming a committee. That is not wrong, but it is not governance. A committee is a governing body: it is the WHO. What it governs is a program: the WHAT. And the program does not govern the technology. It governs the company's capacity to make informed decisions about the balance of risk and reward in its use of AI.

That is what makes the subject supervisable. An examiner cannot meaningfully examine a committee's diligence, and should not be asked to examine the technology. An examiner can examine a program. Does it exist. Was it chartered. Is it scoped. Is it granular in proportion to the environment. Did the accountable body approve it. Those are the questions market conduct and financial examiners have asked of compliance programs, anti-fraud programs and risk management programs for decades, and they require no new authority to ask.

A committee that meets without a program beneath it is evidence of attention, not of accountability.

One consequence for the drafting

Version 5.0 defines an AIS Program as "the controls and processes that an insurer adopts and implements to ensure the responsible use of AI Systems," and says such a program "may be comprised of general guidelines, governance framework, risk management and internal controls, and third-party oversight."

That is a management definition. Controls and processes are what an organization does. The definition places governance inside the program as one component among several rather than above it as the layer that charters it, and it is the reason Exhibit B cannot presently distinguish a company that has built an oversight program from one that has built controls and called them governance.

The Center does not propose that the Working Group redefine the term. The narrower fix is to ask the company. A program whose stated object is the company's AI systems is a management program, and Exhibits A, C and D already reach that ground. A program whose stated object is the company's capacity to

decide about those systems is an oversight program, and it is the only one Exhibit B can test without duplicating the other exhibits. Recommendation 6.1 records it in the opening item, in one clause.

The Working Group's own data supports the emphasis

In the NAIC health insurance AI/ML survey released in May 2025, 93 companies responded and 84 percent said they apply AI or machine learning to health insurance operations, which is 78 companies. Among those 78, 55 percent develop internally using third-party components and 15 percent rely entirely on third-party vendors, with a further 13 percent using a combination. On the two clearest categories alone, seven in ten AI-using health insurers run systems built wholly or partly by someone else. For those companies the model is not theirs to explain. The decision to deploy it, under whose authority and with what withdrawal criteria, is.

3. The answers already exist in a well-governed company

The duty attaches before the exhibit asks

Nothing proposed here is a new disclosure. Where AI is mission critical to underwriting, pricing or claims handling, the obligation to maintain a board-level system of monitoring and reporting already attaches. The Delaware courts developed that duty from *In re Caremark Int'l Inc. Derivative Litig.*, 698 A.2d 959 (Del. Ch. 1996), and applied it to mission-critical risk in *Marchand v. Barnhill*, 212 A.3d 805 (Del. 2019), where the claim survived dismissal because the board had no committee charged with the company's central safety risk, no protocol requiring management to report on it, and no record of having discussed it.

A company that has met that duty can answer every field proposed in this comment from documents it already keeps. A company that cannot has that gap whether or not this exhibit asks.

The record is protective as well as exposing, and this sector holds the exposure twice

In *Firemen's Ret. Sys. of St. Louis v. Sorenson*, C.A. No. 2019-0965-LWW (Del. Ch. Oct. 5, 2021), a documented and monitored program earned dismissal of the oversight claim after a breach of roughly 500 million guest records. The artifacts that answer an examiner's question are the same artifacts that defend the board.

Insurers in scope of this Supplement hold that exposure from both sides. They deploy AI in their own underwriting, pricing and claims operations. Many of them also write directors and officers liability, where the oversight failures of their insureds arrive as claims on the same theory that has already reached cybersecurity.

That second exposure is not theoretical. Cornerstone Research counted sixteen AI-related securities class actions in 2025 and fifteen in the first half of 2026 alone, on pace to nearly double the 2025 total. Those fifteen were thirteen percent of core filings and seventy-three percent of the Disclosure Dollar Loss index, and five of the seventeen mega filings in the period were AI-related. A smaller and growing set of derivative suits pleads board oversight failure directly.

No court has yet decided an AI oversight claim on the merits, and the Delaware pleading bar remains high. That distinction matters less to a D&O writer than it might appear. Defense costs are incurred from the complaint forward whatever the outcome, and settlements in this line have been reached on governance reforms without any merits ruling at all. The loss experience is already running.

This Supplement is written for financial analysis and financial examination as well as market conduct, and that is where the two sides meet. Exhibit B asks the insurer about its own AIS Program. The D&O book is an underwriting, pricing and reserving question this instrument does not reach. But the document set is the same one on both sides, and the examiner reading Exhibit B is reading the record the insurer would need to produce as a defendant and would want to see from a D&O applicant.

This comment does not ask the Working Group to adopt, interpret or enforce any corporate law standard. Model Act #305 Section 1.B reflects the same limit in the corporate governance context, providing that nothing in that act shall be construed to prescribe corporate governance standards beyond applicable state corporate law. The point is practical: these fields impose a collection cost, not a governance cost.

4. Oversight is the stable half, and cyber shows the cost of leaving it out

The lower layers move at the speed of the technology

A model deployed this quarter behaves differently next quarter. Use cases are added and retired, and risks are reclassified as the environment changes. The oversight layer does not move at that speed, because the obligation attaches to the company's decision process rather than to any generation of the technology. That matters here because of the cycle this instrument runs on.

Among the pilot states, Virginia, Pennsylvania and California each examine licensed insurers at least once every five years (Va. Code Section 38.2-1317(B); 40 P.S. Section 323.3(b); Cal. Ins. Code Section 730(b)). Florida conducts financial examinations on a risk-tiered basis of three to five years (Fla. Stat. Section 624.316) and market conduct examinations "as often as it deems necessary" (Fla. Stat. Section 624.3161).

Set that against Exhibit A, which asks for the number of AI Systems currently in use, the number implemented in a regulator-specified lookback window, and counts of AI Models with direct consumer impact and with material financial impact. Those figures are accurate on the day they are collected. Three years later a substantial share of the inventory will have been retrained, replaced, retired or reclassified. An examiner reading the prior examination's Exhibit A is reading a photograph of a moving object.

Exhibit B does not decay that way. A charter, a ratified tolerance, a named accountable officer and an escalation line change only when the company decides to change them, and the program carries the record of that change.

Two regulators have already acted on this

Connecticut, a pilot state, requires its domestic insurers to complete an annual Artificial Intelligence Certification under Insurance Department Bulletin No. MC-25 of February 26, 2024, due September 1, 2024 and annually thereafter. The bulletin rests entirely on pre-existing authority, principally the Unfair Insurance Practices Act (Conn. Gen. Stat. Sections 38a-815 to 38a-819), market conduct examination authority (Sections 38a-15 and 38a-16), the corporate governance annual disclosure statute (Section 38a-142a) and the rating laws. A recurring AI oversight filing, with no AI statute behind it. NYDFS Insurance Circular Letter No. 7 (2024), at paragraph 26, states that an insurer's board or other governing body, its committees, or senior

management under delegated authority "should review and approve" the insurer's AI-related policies and procedures at least annually. It is guidance rather than a rule, and it places the annual approval at the governing level.

Why existing practice does not supply the layer

Governance, risk and compliance practice produces control inventories, risk registers, policy sets and testing programs, each with an owner. It rarely states which layer a responsibility belongs to, or who is answerable at the top of each. The layers collapse into one undifferentiated function, and a company can then describe in detail how its AI risk is managed without being able to say who is accountable for a given deployment. An examiner reading that record sees activity everywhere and answerability nowhere.

The management instruments do not close it, and they are right not to try. They leave scope, granularity and reporting depth to the organization: testing proportionate to consequence, review calibrated to criticality and risk tolerance. Each is a calibration they are correct to leave to the operator. None says who holds it, or on what record. A calibration nobody ratified is a working assumption, and an examiner cannot tell the two apart.

That is the shape of the gap. It is also the shape cybersecurity left open for a decade.

The last time this layer went unbuilt, the gap took a decade to close

This is not the first control discipline to mature from the bottom up. The 2014 Cybersecurity Framework organized its core around five functions, Identify, Protect, Detect, Respond and Recover, and contained no GOVERN function. The control layers matured first and the oversight layer went largely unbuilt for a decade.

It was filled eventually, and mostly from outside the framework. The Securities and Exchange Commission adopted cybersecurity disclosure rules on July 26, 2023, requiring registrants to describe "the board of directors' oversight of risks from cybersecurity threats" under Item 106 of Regulation S-K. The courts applied the oversight duty to cybersecurity on their own schedule. NIST then added GOVERN as a sixth function, including an Oversight category, in Cybersecurity Framework 2.0 on February 26, 2024.

The sequence is the point. For ten years the artifact everyone reached for had nothing to say about who answers, and the answer arrived from a disclosure rule and a courtroom before the standard caught up.

The point for this Working Group is narrow. The oversight layer arrives eventually, either inside the examination instrument on terms the regulator sets, or afterward through a disclosure rule or a courtroom on terms set elsewhere. The Working Group holds an advantage NIST did not: it is an examination body rather than a standards body, and it can simply ask the question. This exposure is the moment to ask it.

5. The naming mechanism already exists, and extending it means naming the object

Insurance regulation already requires a named individual to attest, annually, on a record filed with the department. It does so twice, and neither instrument reaches AI.

Corporate Governance Annual Disclosure

Model Act #305 Section 3.B requires "a signature of the insurer or insurance group's chief executive officer or corporate secretary attesting to the best of that individual's belief and knowledge that the insurer has implemented the corporate governance practices."

The signer is named. The practices are not defined, and they are not broken out by layer. An insurer can satisfy that attestation by describing management processes and calling the result governance, and an examiner reading a CGAD filing cannot tell whether AI oversight decisions are covered by it, covered in part, or not addressed at all.

That is not a defect this Working Group is charged with repairing, and the Center does not propose that it try. It is the reason Exhibit B should ask the question directly rather than rely on a company's corporate governance filing to have answered it. One note for whichever NAIC group holds that charge. Defining the attested practices and breaking them out by layer expectation, so that an attestation distinguishes the decisions made at the oversight layer from the processes operated beneath it, would make both instruments more examinable.

Own Risk and Solvency Assessment

The Introduction to the ORSA Guidance Manual, at Part C (General Guidance), requires "a signature of the insurer's chief risk officer or other executive having responsibility for the oversight of the insurer's ERM process attesting to the best of his/her belief and knowledge that the insurer applies the ERM process described in the ORSA Summary Report and that a copy of the ORSA Summary Report has been provided to the insurer's board of directors or the appropriate committee."

The attestation stops at confirming that the report was *provided* to the board. Provision is receipt. It does not establish what the board did with it, and that is where AI oversight most often fails. Recommendation 6.5 closes that half step by asking for the recorded response.

What extending the mechanism actually requires

The Working Group does not need to invent an accountability mechanism for AI. The sector has one, in two forms, filed by every domestic insurer. But extending it is not a matter of adding a third signature to a third form, because an attestation is only as examinable as its object. CGAD leaves the practices undefined. ORSA defines the object as the ERM process, which is a management process. Neither states that anyone is accountable for an AI oversight program, or what such a program governs.

Extending the mechanism therefore means three things, and Exhibit B can do all three: name the committee that holds the program by charter and the officer who answers to it; state what that officer is accountable for, which is the AIS Program; and state what the program governs, which is the company's capacity to make informed decisions about its use of AI, as distinct from the AI systems it operates. A signature with no stated object attests to nothing in particular.

6. Five additions to Exhibit B

Each is a question about the program. None requires the examiner to evaluate a model. None imposes a substantive governance standard.

6.1 Ask what the program governs, at what granularity, who approved that as appropriate, and which standards the company elected

Checklist item 1 asks whether a written AIS Program has been adopted, when, and how often it is reviewed for updating. Narrative item 1 asks the company to provide the program. Neither asks what the program governs, at what granularity, whether anyone decided that scope and granularity were right for this company, which external standards were elected, at what depth, or who decided.

Two gaps sit inside that, and the first is the more consequential.

Scope and granularity are decisions, and the exhibit treats them as givens. A rural domestic carrier and a multinational group perform the same oversight acts at very different depth. That is correct, and it is why a national instrument should not set one depth for both. But a program that covers the right ground because someone decided it should and a program that covers it because nobody looked further produce identical answers to item 1 as written. Reviewing a document for updating is not the same as testing whether its scope is still appropriate to a company whose use of AI has changed since the last examination. This is the field that distinguishes a chartered program from an inherited one.

An elected standard is a commitment. Adopting a voluntary framework at a chosen depth creates an obligation the company set for itself, and it is the commitment, not the framework, that assurance later tests. The exhibit does not record what was elected or how far.

Proposed: state what the AIS Program governs and the granularity at which it operates; identify the individual or body that approved that scope and granularity as appropriate to the company's size, complexity and use of AI Systems, with the date the scope and granularity are next reviewed for continued appropriateness; and identify the external frameworks or standards elected, the depth of application elected, and the individual or body that approved the election, with date.

A program that states it governs the company's AI systems has described a management program. One that states it governs the company's capacity to decide about those systems has described an oversight program. The exhibit cannot currently tell which it received.

6.2 Ask who is accountable for the program, at two levels

Four questions come close and none establishes accountability for the program. Narrative item 1a asks what role maintains the AIS Program. Narrative item 1e asks how responsibility for governance is assigned and how the organization ensures consistency. Checklist item 2 asks whether the Board of Directors or management was involved in adopting the program, and item 2a asks what the role of the Board of Directors or management in it is.

A company can answer all four accurately and still not have said who is accountable for the AIS Program itself. Maintaining a program is an administrative function. Assigning responsibility describes an organizational chart. Involvement in adoption is a historical fact. The disjunctive running through items 2 and 2a permits a response that is true and uninformative.

The object is what matters. Accountability for AI in general is a portfolio description. Accountability for the AIS Program is a position that can be examined, because the program has a stated scope, a stated granularity and required outputs, and someone either answers for them or does not.

That accountability sits at two levels, and the two are not interchangeable. A company that names its Audit Committee as holding oversight of the program by charter has answered the first level, and for many insurers that is the right body. Adding the committee chair's name does not answer the second, because the chair sits inside that body rather than beneath it. The question that settles it is short: when the committee asks whether the program did what it was chartered to do, who answers? If the answer is the committee, the exhibit has recorded a self-report. If a named officer answers to a named committee, the examiner can test both sides.

Nothing in this requires a new role or a new committee. The officer is whoever already owns it, and insurance regulation already names accountable officers twice, as section 5 sets out.

Proposed: identify the committee holding oversight of the AIS Program by charter; identify by name and role the officer accountable for the program, with any delegation and its date; identify any AI systems declared out of scope, the named individual who accepted each exclusion, and the review date; and describe how the company establishes that the accountable officer and the committee hold the competency to interrogate what the program reports to them.

The out-of-scope field is the one most likely to produce a finding. Systems fall outside a program because they are low risk, legacy or acquired with a book of business, and sometimes because no one decided. The first three are decisions, the fourth is not, and the exhibit cannot currently distinguish them.

6.3 Ask who ratified the risk tolerance and when it is next reviewed

Checklist item 3k asks how the company "Quantifies AI System risk levels," and the new item 3o asks how it "Determined 'materiality' and what threshold was used for assessing AI Systems risk." Both ask how the assessment was made. Neither asks who ratified it. Assessment is an act performed by a modeling or risk function. Ratification is an act performed by a body that can be held to it, and a tolerance calculated but never ratified is a working assumption the company is not committed to.

Proposed: identify the stated AI risk tolerance, the body that ratified it, the date of ratification, and the date of next scheduled review.

6.4 Ask whether the program treats removing the human as a grant of authority

Narrative item 1g asks how the company "assess[es] autonomy, reversibility, and reporting impact risk of AI Systems." Assessing the risk of autonomy is not the same as authorizing it, and a company that has assessed the risk well may still be unable to say who granted the system permission to act.

The oversight question is not what any one system was permitted to do. It is whether the program treats the removal of a person from a decision as a grant of authority at all, or as something that happens when a model is deployed. Those produce different records and different examinations.

Proposed, as an Exhibit B item: identify how the company grants, bounds and periodically re-ratifies authority to operate for AI Systems recorded as automating a decision, recording the authorizing individual or body, the date of authorization, its scope, and its expiration or re-ratification date.

Section 7.2 proposes the per-model complement in Exhibit C, where the automate, augment and support classification is already recorded. That field is the record the program produces. This one is the program.

6.5 Ask where an out-of-tolerance system escalates to, and what was done about it

Narrative item 1b comes closest, asking the company to discuss "the governance structure, Board reporting and frequency." That establishes that reports move and how often. Checklist items 3m and 3n are consumer-facing, addressing complaints and consumer awareness.

None of them establishes what the report carries, what triggers one outside the cadence, who can act on it, or what was done. That gap matters most for the system operating within specification and outside tolerance, which is the harder case and the one that produces supervisory surprise.

Proposed, as two items rather than one, because eight elements in a single checklist cell will not be answered well.

Escalation: the cadence on which the program reports to the accountable committee independent of events; whether that reporting carries adverse information, incidents, findings, inquiries and deviations from tolerance and not only progress; the escalation triggers and materiality thresholds fixed in advance; the party with authority to continue, constrain or withdraw a system; whether at least one reporting pathway to the governing body does not depend on management's discretion; and whether the response to an escalated matter is recorded.

Disclosure: the company's disclosure position, including the criteria for disclosure to the department, to consumers and to the public, and any clock that statute attaches to it.

Two elements carry most of the weight. A routing label without a party who can withdraw the system does not resolve to an action, and a red flag received without a recorded response is, on the record, a red flag ignored. This is the field an examiner can test directly from minutes.

Optional sixth: ask whether anyone independent has tested the program

Narrative item 1d asks how "the effectiveness of the AIS Program and individual models are assessed and modified," and Narrative item 3a asks whether validation and testing of internally developed systems are "performed by someone that is independent from development." The Working Group has therefore already applied the independence principle, but to model validation rather than to the program.

Proposed: identify the provider of any independent assurance over the AIS Program, the scope of the review and the date of the most recent one. The useful object of such a review is whether the company applied what it said it would apply, at the depth it recorded in the opening item, rather than conformance to a standard in the abstract. ISO/IEC 42006:2025, published July 7, 2025, gives this market a published accreditation standard. A blank answer is information, not a deficiency.

Proportionality

Recommendations 6.2 and 6.5 carry most of the value for a small or resource-constrained insurer, with 6.4 added wherever a system automates a decision. That minimum viable form is why an oversight instrument scales from a small domestic carrier to a global group without a separate small-entity version.

7. Two provisions in Exhibit C

7.1 Keep company-set risk criteria, and record who approved them

Exhibit C states that "AI model risk criteria are set by the insurance company." The Center supports that provision and asks the Working Group not to retreat from it. A national instrument cannot set one threshold right for both a small domestic carrier and a multinational group. Allowing the company to set its own criteria is what lets the Supplement scale, and it is consistent with the principle that the obligation is not the right decision but an informed one, reached through a process defensible to the scale of the risk.

The gap is not that the company sets the criteria. It is that setting them is itself a decision, and the exhibit does not treat it as one.

Proposed, four fields alongside the existing statement: the individual or body that approved the risk criteria; the level at which approval was given; the date of approval; and the date the criteria are scheduled for review.

7.2 Treat automate, augment and support as a grant of authority, not a property of the system

Exhibit C item 10 records AI model type as automate, augment or support. The Supplement defines Automation as "an AI System that does not involve human intervention" and Augmentation as one that "suggests an answer and/or advises a human who is making a decision." Those definitions are precise and the Center does not propose changing them.

The field built on them is the most consequential in the exhibit, and it is recorded as a property of the model. It is better understood as a grant of authority. Removing the human from a decision is a decision, and the difference between the two definitions above is the difference between a recommendation and an action taken in the company's name.

Proposed: for any model recorded as automating a decision, identify the individual or body that authorized operation at that level, the date, the scope of the authorization, and its expiration or next re-ratification date.

Authority to operate that does not expire is not authority. It is inertia.

8. The attestation in the third-party vendor framework

*A note on routing. The companion * Regulatory Framework for Third-Party Data and Model Vendors * is before the Third-Party Data and Models (H) Working Group, and its comment period closed August 5, 2026. This point is raised here because the two documents interlock and because Exhibit B addresses vendor procurement at item 3l and, as of version 5.0, vendor-developed AI Systems at the new item 3p. The Center offers the material to that Working Group directly if it is useful.*

The framework provides for an annual attestation by a senior officer that the vendor's governance program is implemented and effective for the data and models in scope. The Center has worked from the version exposed to that Working Group and does not quote it verbatim here; the point below turns on the pairing of implemented with effective rather than on the precise wording.

Most of that is sound. Requiring the attesting person to hold the approval authority rather than merely sign is a deliberate and correct choice, and *implemented* is an existence and operation claim that leader is positioned to make.

Effective is different in kind. Effectiveness is an assurance conclusion, and the framework asks the vendor to assert it about the vendor's own program with no independent party testing the assertion. Sarbanes-Oxley pairs the two deliberately: Section 302 is management's certification, Section 404(a) is management's assessment of internal control, and Section 404(b) is the independent auditor's attestation on it. The framework has written the 302 and not the 404(b).

The Center is not asking the Working Group to mandate an audit. Two narrower changes close the gap.

8.1 Separate what the officer attests to from what requires assurance

Ask the attesting leader to attest to implementation and scope, which that leader knows. Ask separately whether independent assurance has been obtained, from whom, on what scope and as of what date, treating no answer as information rather than as a deficiency.

8.2 Name the object of assurance as the commitment, not the framework

Where assurance has been obtained, the question is whether the vendor applied what it said it would apply, at the depth it recorded, rather than whether the vendor's program conforms to an external standard in the abstract.

This carries weight beyond insurance. OCC Bulletin 2026-13 of April 17, 2026, which replaced the 2011 model risk management guidance, states that generative and agentic AI models "are not within the scope of this guidance." The instrument financial supervisors have relied on for model risk does not reach the systems this Supplement is about.

9. What this comment does not ask for

It does not ask for a new requirement on any insurer. Every field proposed is a question, answerable from a document the company either holds or does not, and the absence of the document is itself the finding. Version 5.0's own addition of a Document Name and Page # column is the same principle, and the Supplement's own instruction that "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" applies to these items exactly as it applies to the existing ones.

It does not ask the Working Group to prescribe governance content. Asking who ratified a decision is not prescribing what the decision should be, which is the distinction Model Act #305 Section 1.B draws in the corporate governance context, as discussed in section 3. Nor does it ask the Working Group to abandon the optional and supplemental design, or the provision that each state may modify or add questions.

And it does not ask for anything the model bulletin has not established. The NAIC Model Bulletin on the Use of Artificial Intelligence Systems by Insurers, adopted December 4, 2023, expects a written program, vests responsibility with senior management accountable to the board, and outlines the information regulators may request. The Supplement is the operational expression of that bulletin, and these recommendations complete the expression rather than extending its reach.

10. About the Center

The Center for AI Oversight is an independent nonprofit research and education organization. Its tagline is *Defining oversight governance for AI in regulated industries*. It works with corporate boards and senior executives in regulated industries, with the professional bodies serving governance practitioners, and with regulators at the federal and state level. Center materials are educational, and the Center takes no positions on matters involving specific institutions.

Its advisory board includes Jim Cunha, former Executive Vice President, Federal Reserve Bank of Boston; Chuck Senatore, former Regional Director, U.S. Securities and Exchange Commission; Rear Admiral David Simpson, U.S. Navy (retired), former Chief of the Public Safety and Homeland Security Bureau, Federal Communications Commission; and Brian Peretti, former Chief Technology Officer, U.S. Department of the Treasury. Its Senior Fellows include Ankur Singhal, former Head of Technology Risk at PayPal and at The Hartford.

The Center's Executive Director, Brian J. Allen, Esq., was appointed by the Chairman of the Federal Communications Commission to represent the communications sector during development of the Cybersecurity Framework, and represented the banking sector through the Bank Policy Institute during development of the NIST AI Risk Management Framework.

This comment applies the Center's general position on the oversight layer, *Distinguishing the Oversight Layer of AI Governance*, available at cfaio.org.

Contact: info@cfaio.org