1. Consider Adoption of its Summer National Meeting Minutes

Attachment 1

–Commissioner Nathan Houdek (WI)

Draft: 8/23/26

Big Data and Artificial Intelligence (H) Working Group
Columbus, Ohio
August 13, 2026

The Big Data and Artificial Intelligence (H) Working Group of the Innovation, Cybersecurity, and Technology (H) Committee met in Columbus, OH, Aug. 13, 2026. The following Working Group members participated: Nathan Houdek, Chair, and Coral Manning (WI); Doug Ommen, Co-Vice Chair (IA); Mary Block, Co-Vice Chair (VT); Richard Fiore (AL); Alex Romero (AK); Lori Munn (AZ); Chris Erwin (AR); Ken Allen (CA); Jason Lapham (CO); Joshua Hershman (CT); Nicole Crockett (FL); Matt Kilgallen (GA); Weston Trexler (ID); Ryan Gillespie (IL); Jake Vermeulen (IN); Saima Sanjida Shila (LA); Sandra Darby (ME); Mary Kwei (MD); Caleb Huntington (MA); Kate Stojsih (MI); Phil Vigliaturo (MN); Brad Gerling (MO); Connie Van Slyke (NE); Gennady Stolyarov II (NV); Sean Rosene (NH); Dave Wolf (NJ); Vanessa DeJesus (NM); Angela Hatchell (NC); Colton Schulz (ND); Matt Walsh (OH); Nicole Nash (OK); Gary Jones and Michael McKinney (PA); Andreea Savu (SC); Tony Dorschner (SD); Michael Schulz (TN); Marianne Baker (TX); Eric Lowe (VA); Ellen Potter (WV); and Lela D. Ladd (WY).

1. Adopted its July 22 Minutes

The Working Group met July 22 and took the following action: 1) received an update on the Artificial Intelligence (AI) Risk Evaluation Supplement pilot process and 2) heard a panel discussion on perspectives from actuarial organizations on generalized linear models (GLMs) and AI governance.

Stolyarov made a motion, seconded by Trexler, to adopt the Working Group's July 22 minutes (Attachment Two-A). The motion passed unanimously.

2. Received an Update on the AI Risk Evaluation Supplement Pilot

Commissioner Houdek said the 12 states participating in the pilot have been engaged since March and have met nearly weekly to discuss their individual state experiences implementing the pilot. He said that, based on feedback received during the pilot and input from subject matter experts, the supplement will be revised with an updated version ready for a public exposure comment period in early September.

He said each participating state has been engaged in the work over the spring and summer, and that some states used the supplement as part of a market conduct or financial examination that was already planned, while others used it as an ad hoc questionnaire outside the typical examination process. Each state has the flexibility to modify or add questions to meet its specific needs. The pilot will continue through September, and feedback has been collected from companies involved in the pilot, with each participating state also reaching out to its pilot companies to solicit input through a centralized survey of the experience. Written feedback has been received, which will help inform revisions to the supplement. Opportunities for verbal feedback will be provided during the public meetings following the two exposure periods planned for this fall, with the goal of finalizing the supplement and adopting the final version at the Fall National Meeting.

Crockett commented that the pilot process has been effective due to the continued engagement and strong collaboration between the states.

3. Heard a Presentation from AM Best on AI Governance in Insurance

Edin Imsirovic (AM Best) presented on AI governance in insurance, stating that the presentation was for discussion purposes and does not change AM Best's criteria, methodology, or rating guidance, nor does it create a formal AI governance assessment framework. He said the discussion on AI governance draws on public regulatory and supervisory standards, model risk, and AI security measures to address how governance expectations may change as AI systems move from predictive to generative to action-oriented workflows. He said the NAIC *Model Bulletin on the Use of Artificial Intelligence Systems by Insurers* and the AI Risk Evaluation Supplement pilot provide a foundation for a company's AI governance policy toward how governance operates.

Imsirovic defined three working terms: 1) predictive AI, which uses data to score, rank, classify, or flag something, often used in pricing models or claims triage, usually supporting a decision inside a well-defined process; 2) generative AI, which creates new content, such as summarizing documents, drafting language, extracting details, or turning unstructured data into something more useful, but its output can sound confident when it is wrong; and 3) agentic AI, which goes a step further and can work through a series of steps toward a goal, use tools, look up information, update a system, route work, or move a workflow forward. This system may affect processes directly rather than support a human decision. Each creates a different governance problem.

Imsirovic said AI does not create a separate risk category, as the exposures already flow through existing categories, such as product and underwriting, operational, and regulatory and legal risk, and that the enterprise risk management (ERM) question remains whether risk management capability is commensurate with the risk profile.

He said that well-governed AI deployment may support outcomes such as expense and processing improvement, underwriting quality, claims efficiency and consistency, and fraud control, while risks of execution, control breakdowns, cyber and data exposure, vendor dependency, and regulatory and explainability exposure can arise from deployment. He said what matters is whether governance mitigates the risks.

He said that if AI affects pricing, risk selection, segmentation, or portfolio mix, that points to product and underwriting risk. Data quality, claims-handling errors, cyber exposure, workflow breakdowns, or weak oversight of third-party models fall under operational risk and consumer protection. Explainability, litigation, or regulatory scrutiny can flow through the legislative, regulatory, judicial, and economic categories. One use case can span several categories; for example, a claims AI tool that begins as an operational issue becomes a regulatory issue if consumers are treated inconsistently and creates financial consequences if errors affect payments or reserve information. He said the categories are familiar, but the path into them is changing rapidly, as carriers move from predictive models to generative models to agentic actions. As a result, risks can emerge faster, become harder to understand, and move further without human intervention.

He described the development of AI capabilities as a compounding progression rather than a timeline. He said that using AI for fraud detection and pricing, or for computer vision and language processing used in claims and document work, was foundational, and that the governance problem at this stage remained closer to traditional data and model governance. He said that generative AI changed that, as those systems work with less structured information, produce less predictable output, and may be harder to explain, presenting a different governance problem rather than a harder version of an old one. He said that agentic AI changed it again, as those systems can act rather than only recommend, and one error can trigger others and move through a workflow before a person can intervene; therefore, a single governance approach is not enough.

He noted that predictive AI raises familiar questions about whether the model was properly tested, is fair, is documented, and whether its performance is changing over time. Generative AI raises questions about whether the output can be trusted, whether sensitive information stays where it should, and what happens when the underlying model is updated. Agentic AI raises questions about the system's access, what gets recorded, where a person steps in, and whether an action can be stopped or reversed. The risks build, such that an agentic system may still carry predictive and generative risks in addition to new permission and action risks, and that the comparison should be read as a comparison of governance problems rather than as a simple risk ranking.

He said that for predictive AI, monitoring should lead to action, so that when a model drifts or produces a weak outcome, the natural thing to look for is whether the company adjusted a threshold, limited the model's use, or added human review, and a dashboard that never changes anything is weak evidence. For generative AI, the question is whether controls hold up through the live workflow, including whether the level of review increases when the output matters more, whether the company tests the system again when a vendor updates the model, and whether the company can show that sensitive information is secured. For agentic AI, the evidence has to be more concrete, including whether permissions are tested, whether actions are logged for company reconstruction, whether there is a kill switch to roll the process back to a safe state, and whether a person has the authority to override actions. He said two companies can have the same policy but very different governance in practice.

He said the predictive, generative, and agentic labels do not indicate how much is at stake. He identified five factors: 1) how important the decision is, including whether the output directly affects a consumer or a financial outcome; 2) how much the system can do on its own; 3) whether the company can explain why the system produced a particular result or took a particular action; 4) how quickly the system can change, whether gradually through drift or suddenly through a vendor update; and 5) how dependent the insurer is on a third party, including whether the company can independently test, challenge, or override the system. Underlying these factors is proportionality, in that governance around a system should become stronger as stakes, autonomy, lack of transparency, speed of change, or third-party dependence increase, and he said this is consistent with the NAIC *Model Bulletin on the Use of Artificial Intelligence Systems by Insurers* and with recent work from the European Insurance and Occupational Pensions Authority (EIOPA), which also takes a risk-based and proportionate approach to AI governance in insurance. Thus, predictive AI is not inherently lower risk, and agentic AI is not automatically unacceptable. The label indicates the type of system, while the use case indicates more about the level of risk and the governance it may need.

He said lower-impact tools that mainly assist a person, such as an underwriting research assistant, still need governance, but the likely consequences of an error are more limited, and a person remains close to the process. Where the system is still assisting a person but the output matters more, such as extracting information from medical records for underwriting or preparing claim summaries for adjusters, the review has to be meaningful. This raises the question of whether the reviewer has enough information, time, and authority to question the result rather than accept it. He said that where the system has more autonomy but the impact is operational, such as narrowly defined automated endorsement processing, the questions become whether the boundaries are clear, whether actions are logged, and whether errors can be contained and reversed. He said the most demanding governance zone combines greater autonomy with more important consumer or financial outcomes, such as an agentic system routing and paying claims or making underwriting decisions. Controls such as human override, action logging, rollback, and clear limits on the system's actions become especially important.

He said controls are built in layers, with earlier controls staying in place as each layer is added. Predictive AI governance begins with model risk controls, including independent validation. Revalidation triggers mean deciding in advance which kinds of changes should lead to another review or testing, such as changes to a predictive model

or its data. For generative AI, a company should know what prompts are used; otherwise, the output will be harder to explain. For agentic AI, the level of approval should depend on the action, as looking up information is not the same as changing coverage or issuing payment. Some controls depend on vendor contracts, such as testing rights and notice of important updates, which can be much harder to obtain.

He said that where governance lives can make a difference, as two companies can use the same AI tool and have similar controls but end up with very different levels of risk in how the tool is integrated into the business. Governance should be built into the process. Fragile tool layering occurs when a tool is added to an existing process without changing the workflow or its accountability. He said companies do not choose that setup deliberately but arrive at it gradually. As a pilot works, more people begin using it, the use case spreads, and only later does the company realize that the process around the tool never caught up. A well-designed predictive model in a weak operating environment can create more risk than a more advanced AI system that is properly integrated and controlled.

He said shared dependency is broader than reliance on a single vendor, as many carriers may depend on the same vendors, models, data sources, cloud providers, or other infrastructure, and noted that the Financial Stability Board (FSB) has highlighted third-party dependency and provider concentration as potential financial stability concerns. For a single carrier the issue comes down to three questions: 1) visibility, whether the company understands how the system behaves; 2) control, whether the company can test and challenge the system, be notified when something important changes, and step in when needed; and 3) portability, whether it is difficult to replace or unwind a tool once it becomes part of an important workflow. A vendor change can affect how a system behaves, even if the carrier makes no changes, and risks can compound if the company does not review it. The larger point is at the market level, as each carrier may see only its own dependency, but if many carriers rely on the same underlying provider, a single weakness can affect many companies at the same time. This does not mean third-party use is a problem, as vendors can offer useful tools, but the key is whether insurers can understand what the system is doing and have the authority to oversee it and step in when something changes. The issue is whether oversight works effectively.

He identified four areas where a company should be able to show more than a policy or framework: 1) AI inventory, where knowing which decisions each system touches and how much it can do on its own shows which systems need the most attention; 2) validation and monitoring, where the question is what happens when a trigger is hit, including any tests and approvals performed before the system went live and what the company did when monitoring later found a problem; 3) human oversight, where a company should be able to show an example in which a person challenged the result, changed it, or stopped the process; 4) vendor and agentic controls, where the question is whether the company can see a vendor change, test its effect, stop the system, or return to a safe version. Policy alone is not enough; rather, operational governance is necessary.

He said the question is whether the use of AI stays bounded, monitorable, governable, and resilient as systems change, take on more autonomy, and become harder for people to oversee. There is still no single AI rulebook, but regulators want to know what AI is being used for, what decisions it affects, how it was tested, who is accountable, and evidence showing that controls actually work. He said that the last point represents the real shift, as much of the earlier work was principles-based, while newer work is moving toward implementation and examination.

He said modern AI systems may include the model, the prompt, the information it retrieves, memory, tools, and permissions, so the control boundary becomes much larger than the model itself. He said the UK AI Security Institute (AISI), working with the Bank of England, examined what developers are building rather than only testing models in a laboratory, reviewing more than 177,000 agent tools built over the preceding 16 months. Two findings

stood out: 1) tools that take action by changing something in an external system grew from about one-quarter of usage to 65% of usage over that period; and 2) while most action tools support medium-stakes work, such as editing a file (finance was an exception), with high-stakes financial roles having more tools that can change things than would be expected, including growth in the number of servers that can execute payments from about 46 to more than 1,200 in a single year. He said memory fits within that same inventory question, raising data governance questions about what is kept when a system carries information from one task to the next, how long it is kept, and whether it can be corrected or deleted when a policyholder asks, with the difference being that the data now sits inside the system rather than beside it, so identity, delegated authority, the ability to revoke access, and what the system remembers may all become core governance questions rather than technical details.

He said that capability can change even when the model does not. For example, the AISI tested frontier models on multistep cyberattack scenarios and increased the compute budget from 10 million tokens to 100 million tokens, which improved performance by 59% with no retraining, additional fine-tuning, or special expertise. This means a company could keep the same model and still materially change the risk by changing what surrounds it, whether through more time, more attempts, better tools, or wider permissions, such that approving the model may not be the same as approving the deployment.

Commissioner Houdek asked where companies struggle the most in governance. Imsirovic said that one of the biggest areas for improvement is building or revising the operating model around the use of AI, as too many companies use various tools without redesigning the operating model.

Block commented that agentic AI seems like a transition from human to AI models, back to another human being doing the work. Because agentic AI operates independently, it feels like a continuum in which supervision of a nonhuman employee is required. She appreciated the discussion of having a plan for a quick kill switch, for what happens when a system fails, and for the process to shut a system down, and that she views this as part of the ERM disaster recovery process to be treated the same way as planning for when a key employee disappears. Imsirovic agreed and said that agentic AI is not as widespread as generative AI at the moment, in part because companies are cautious about deploying agentic systems due to operational and regulatory concerns. He said most of the deployments he sees are in lower-stakes roles, such as a frequently asked questions (FAQs) chatbot or research assistance, and that the controls companies most often discuss involve keeping a human in the loop throughout the process and ensuring the AI system is monitored and tested, with the caveat that agentic AI remains mostly in early pilot stages.

Commissioner Ommen said a number of Working Group members are receiving different responses through that structured effort. He asked whether Imsirovic is beginning to see policies and procedures standardized in writing or whether there is still a great deal of communication through meetings. Imsirovic responded that the amount of capability an AI system has draws a different amount of scrutiny within insurance companies, noting that companies typically tell him they do not use AI for underwriting or, if they do, that it is monitored, and that in underwriting, it is usually used to turn unstructured data into structured data to enrich the submission process. Another area is in the claims process, helping adjusters with claims summaries, and that the most consistent answer he receives to governance questions is that a human is in the loop. Commissioner Ommen asked whether that response comes in conversation or whether written procedures are beginning to appear. Imsirovic responded that most of it comes through conversations, as AM Best does not have an official process for gathering data on AI governance, although that may change. He said some companies volunteer more information than others and that most are keen to describe how AI will improve their processes, but he has not seen a great deal of that documentation.

Stolyarov asked whether Imsirovic is seeing insurers adopt a consequential approach to AI governance in which the impact of a particular failure is considered not only for a consumer but for the enterprise as a whole. He illustrated the distinction with a chatbot that interacts with consumers and hallucinates information, which could mislead consumers and draw regulatory concern but is fundamentally correctable because a person at the insurer could intercede and provide the correct information, compared with an AI agent that, in an effort to be helpful, proactively and irreversibly deletes the company's claims information to free storage space, which is a failure from which the enterprise as a whole could not recover. The member cited the July 2024 CrowdStrike incident, in which a new employee implemented a systemwide update that disabled the systems of airlines and other enterprise clients, as the kind of catastrophic failure governance needs to prevent, and asked whether insurers are prioritizing those systemwide consequences. Imsirovic responded that the level of sophistication depends in part on the sophistication of a company's use of AI systems, as companies that entered the space earlier had more time to develop proper governance frameworks for those tools and tend to be higher-rated with well-established ERM frameworks. He said his concern is with companies that are newer to the technology and lack that history, which may implement it without having thought it through.

Commissioner Hershman noted Imsirovic had identified three types of AI, that those systems are leveraged for different use cases carrying different risks to the consumer or the company, and that different governance structures should surround those different types of systems and risks. He asked how the Working Group should think about the different levels of governance and how governance tools should be distinguished, such as by high risk versus low risk or by the type of AI being used. Imsirovic responded he would look for a pattern rather than one document, including a useful AI inventory, clear testing and approval standards, monitoring that leads to action, examples in which human review changed the outcome, evidence that the company can respond when vendor systems change, and whether the company is changing its operating model around the technology. He said the answer returns to the impact chart and the five factors: The higher the impact of a particular tool, the more governance that tool will need. He said an agentic tool used for narrowly defined endorsement processing may not require as much governance as the same type of tool used to pay claims or to automate underwriting, so the use case and what is at stake should determine the level of governance around a particular technology.

4. Discussed Other Matters

Commissioner Houdek reminded participants that an updated version of the supplement will be exposed for a public comment period at the beginning of September.

Having no further business, the Big Data and Artificial Intelligence (H) Working Group adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/H Cmte/2026_Summer/WG-BDAI/Minutes-BDAIWG081326-Final.docx

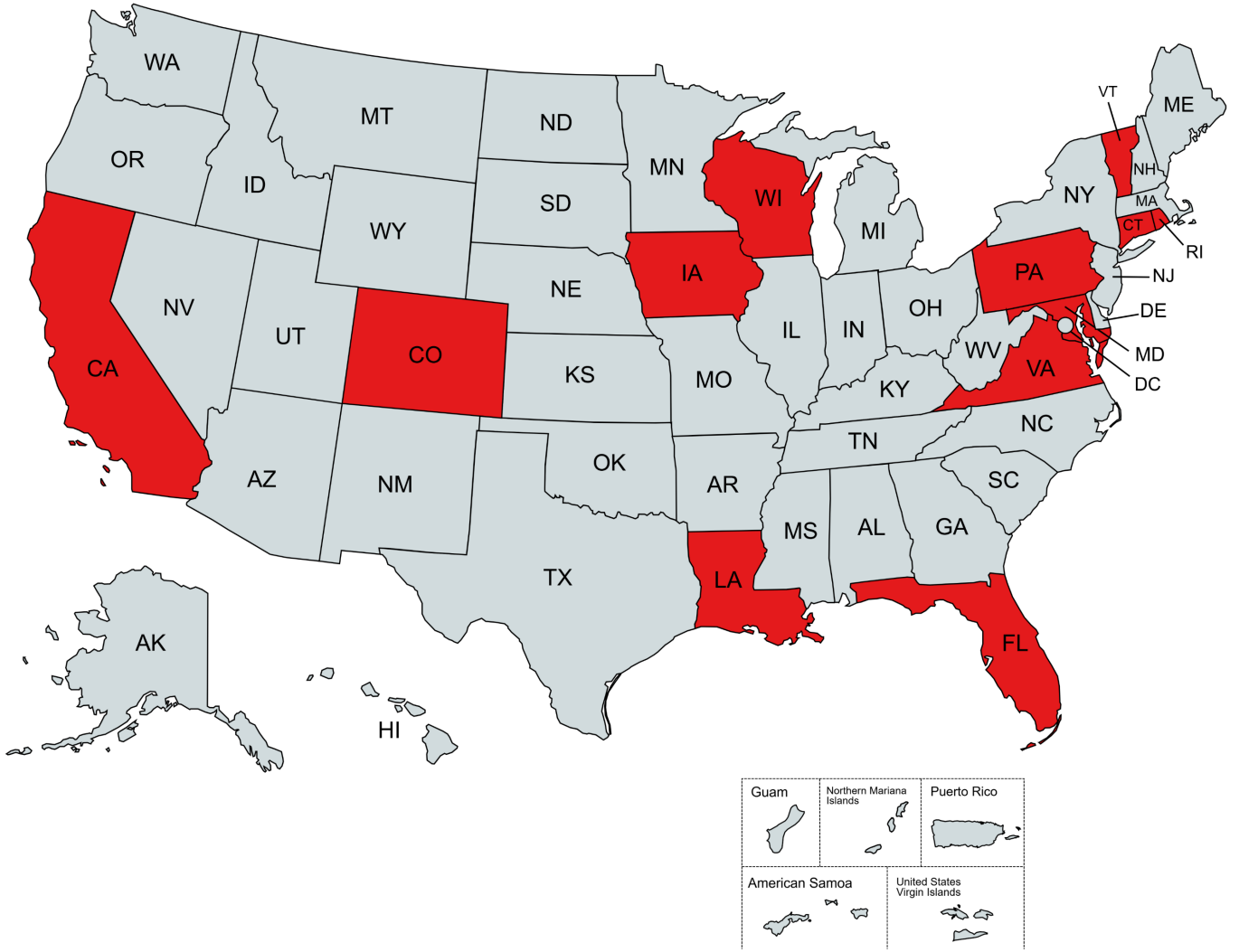
2. Hear an Update on the Proposed Changes to the AI Risk Evaluation Supplement

Attachment 2

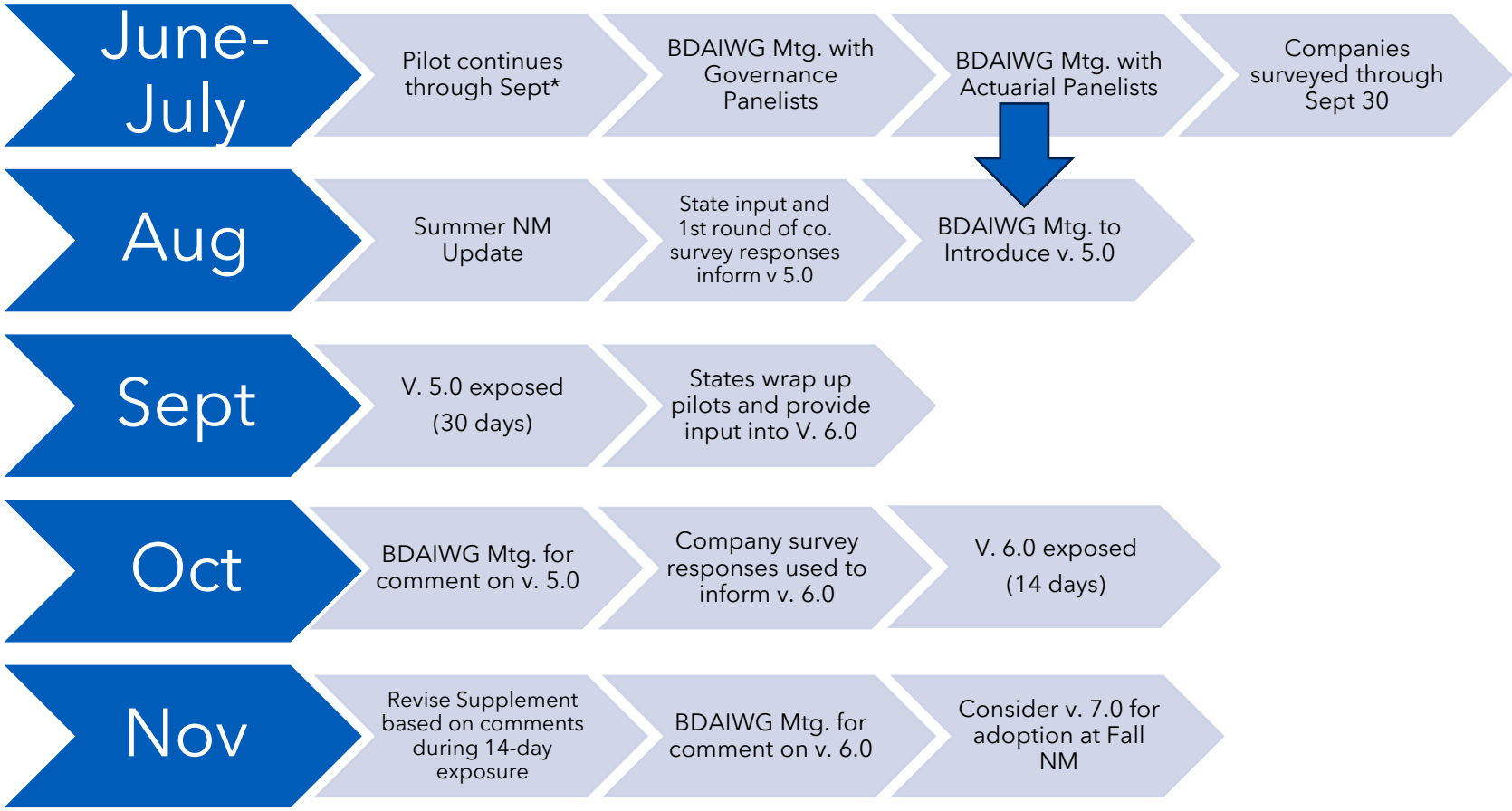
—Coral Manning (WI)

Pilot Participation States

1. California
2. Colorado
3. Connecticut
4. Florida
5. Iowa
6. Louisiana
7. Maryland
8. Pennsylvania
9. Rhode Island
10. Vermont
11. Virginia
12. Wisconsin



AI Risk Evaluation Supplement Pilot–Timeline



*Timing for each state's pilot will vary and every pilot may not be completed by 9/30/26. Information gleaned from the pilot experience will inform the Supplement drafts through adoption and future iterations.

AI Risk Evaluation Supplement Pilot– Summary of Changes from v4.0 to v5.0

- Clarification of Definitions
  - Materiality
  - Direct Consumer Impact
  - Risk Assessment
  - AI System, AI Model, AI Program
- Handling GLMs
- Exhibit Edits

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 to Adverse Consumer Outcomes 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) 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 to assist in understanding and assessing a company's use of *AI Systems*. Inquiries and information requests performed related to this supplement will be coordinated consistent with the guidance provided by the *Market Regulation Handbook, Financial Condition Examiners Handbook*, and the *Financial Analysis Handbook*.

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.

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 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, corporate governance annual disclosures, confidentiality, financial reporting, and rating.

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 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.

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 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.

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 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.

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.

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* inventory, it may suggest submission of the inventory in lieu of completing Exhibit A.
- 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.
- For column 3, regulators should specify the number of months.

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**	**Number of *AI Models* with Direct Consumer Impact**			**Number of *AI Models* with Material Financial Impact**			(10) ***Use Case(s)***
			(4) ***Generalized Linear Models (GLMs)***	(5) ***Generative/ Agentic AI Models***	(6) ***Other AI/ML Models***	(7) ***Generalized Linear Models (GLMs)***	(8) ***Generative/ Agentic AI Models***	(9) ***Other AI/ML Models***	
Marketing									E.g., UC1: Identify potential consumers interested in product.
Premium Quotes & Discounts									
Underwriting/ Eligibility									
Ratemaking/ Rate Classification/ Schedule Rating/ Premium Audits									
Claims/ Adjudication*									
Customer Service									
Utilization Management/ Utilization Review/ Prior									

(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**	**Number of *AI Models* with Direct Consumer Impact**			**Number of *AI Models* with Material Financial Impact**			(10) ***Use Case(s)***
			(4) ***Generalized Linear Models (GLMs)***	(5) ***Generative/ Agentic AI Models***	(6) ***Other AI/ML Models***	(7) ***Generalized Linear Models (GLMs)***	(8) ***Generative/ Agentic AI Models***	(9) ***Other AI/ML Models***	
Authorization/ Level of Care Determination									
Fraud/Waste & Abuse									
Investment/ Capital Management									
Reserves/ Valuations									
Catastrophe Triage									
Reinsurance									
Other Insurance 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.*									

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?

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 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 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* and individual models are assessed and modified. Click or tap here to enter text.

e. Discuss how responsibility for governance within the organization is assigned and how the organization ensures consistency and alignment. 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.

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 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. 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.
   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. 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. 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.

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. 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.

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 and what is the frequency of review for updating?	
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*?	

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. Assesses, mitigates, and evaluates *inherent risks* related to the use of *AI System* including the possibility of the company engaging in unfair trade practices		
	3b. Ensures *AI Systems* are compliant with applicable state and federal laws and regulations		
	3c. Evaluates the risk of *Adverse Consumer Outcomes*		
	3d. Considers data privacy and protection of consumer data used in *AI Systems*		
	3e. Evaluates whether *AI Systems* are suitable for their intended use and should continue to be used as designed		
	3f. Considers *AI System* risks within its Enterprise Risk Management (ERM)		
	3g. Considers *AI System* risks within the Own Risk and Solvency Assessment Report (ORSA), as applicable.		
	3h. Considers *AI System* risks within the software development lifecycle (SDLC) or within the broader *AIS Program*		
	3i. Considers *AI System* risk impact on financial reporting		
	3j. Trains employees about *AI System* use and defines prohibited practices (if any)		
	3k. Quantifies *AI System* risk levels		
	3l. Provides standards and guidance for procuring and engaging *AI System* vendors		

	3m. Considers consumer complaints resulting from *AI Systems* and whether they are identified, tracked, and addressed		
	3n. Promotes consumer awareness of the use of *AI Systems* through disclosures, policies, and procedures for consumer notification, as appropriate		
	3o. Determined "*materiality*" and what threshold was used for assessing *AI Systems* risk.		
	3p. Governs, monitors, tests and ensures transparency of *AI Systems* developed by vendors.		

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*, 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: _____

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*)	
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	Model risk(s) (i.e., describe the potential for *adverse consumer impact*)	
9	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	*AI model* type (automate, augment, support)	
11	Discuss testing model outputs (e.g. model drift, accuracy, unfair trade practices, unfair discrimination, performance degradation, etc.) and how the model was validated prior to being deployed as well as how its performance is monitored on an ongoing basis.	
12	Last date of model testing	
13	Discuss how the model affects the financial statements, risk assessment or controls.	
14	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	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.	

Exhibit D: AI Model Data

Purpose: To obtain detailed information of the source(s) and type(s) of data used in 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: _____

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)			

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)			

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" 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.

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

“Automation” refers to 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.

“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 mimic 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 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.

NAIC NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS

AI Risk Evaluation Supplement – Summary of Changes

To support stakeholder review of the updated document, NAIC staff drafted this summary of the most impactful changes including rationale where appropriate:

- Intent – This document was designed as a supplement to existing handbooks. During prior comment periods, feedback was received on the scope of the use of the document. Additional text was added to emphasize the document as a supplement. Accordingly, decisions on who will receive Supplement related inquiries are expected to be addressed relying on Handbook guidance and not via the Supplement which is now expressed in the Supplement.
- Instructions – During the pilot, companies often received several parts of the Supplement as part of the field test of the document. However, in practice, companies may receive more limited inquiries. An example of when a regulator may deem it appropriate to limit inquiry was added.
- Machine Learning – Added new guidance explaining Machine Learning and as an example GLMs and offering regulators the opportunity to limit inquiry after Exhibit A responses are received if the inquiring regulator deems it appropriate.
- Materiality – Added additional language to align the definition of materiality to the *Financial Condition Examiners Handbook* and to explain the use of the concept within the Supplement.
- Risk Assessment – Added additional language to explain the use of the concept within the Supplement. Regulators are focused on the consideration of inherent risk, which is the risk before the consideration of mitigating controls.
- Exhibit A:
    - Instructions are now organized via bullets to provide clarity.
    - Exhibit A is now clarified where regulators are seeking information on AI Systems vs AI Models.
    - Added guidance indicating that either the regulator will specify a materiality threshold for responding to Exhibit A related inquiries or will ask the company to specify the threshold the company used to scope its responses.
    - Revised columns to gather more information on the model types. This information is also now requested separately for AI Models that have a Direct Consumer Impact and then also for AI Models with Material Financial Impact. This will allow regulators to understand if a company's use of AI is limited to a specific type of AI which may thus limit further Supplement related inquiries.
    - There were several implied requests in the prior version of the Supplement. These are now explicitly listed. Most notably, regulators now ask for a Model Inventory.
- Exhibit B
    - Following feedback on the AIS Program vs Governance Framework related language, the Exhibit is now called clarified to focus on a company's AIS Program.
    - Instructions have also been clarified and questions posed have been updated leveraging feedback from the pilot.
    - Narrative Version
        - 2b. Added guidance for companies when responding to inquiries related to models with direct consumer impact.
        - 3d. (new) Added an inquiry about explainability and transparency.
    - Checklist Version
        - Added request to provide document name and page # when companies are responding to inquiries. This will allow regulators to connect the company responses to documents provided.

        - 3o. (new) Added question on materiality.
        - 3p. (new) Added question on how a company oversees third party models.
- Exhibit C
    - Instructions have been clarified.
    - Requested information has been updated leveraging feedback from the pilot.
    - Field 1 – Clarified to focus on AI Model
    - Field 2 – Moved from later in the Exhibit and now asks on the use case and purpose of a given model for which regulators are requesting information.
    - Field 3 – Added examples of model type responses
    - Field 4 – Added to the inquiry on how the model was developed
    - Field 7 & 8 – Split so now field 6 asks for model risks and field 7 asks for model limitations; added "(high, moderate, low, etc.)" to field 7.
- Exhibit D
    - Added a field to indicate the AI Models which rely on the data sets described within the company's response to Exhibit D which allows the deletion of a column which asked for the same information for each data field
    - Clarified the column asking for information on how the information is used
    - Added the statement in Regulator Instructions that "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.", recognizing that "Ethnicity/Race" is requested on row #2.
    - Added the statement in Regulator Instructions that "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."
- Definitions
    - Added a definition for agentic AI
    - Added a definition for AI Model leveraging the NIST/WH Executive Order language
    - Added a definition for AIS Program leveraging language from the AI Model Bulletin
    - Removed "Degree of Potential Harm to Consumers"
    - Added a definition for "Direct Consumer Impact"
    - Added a definition for GLMs
    - Clarified the definition of inherent risk to better align with the *Financial Condition Examiners Handbook*
    - Added a definition for materiality. Corresponding changes to Exhibit A result in a dynamic where a company may be asked to complete Exhibit A responses based on a materiality threshold the company specifies but also that it discloses to regulators
    - Added a definition for "Material Financial Impact"
    - Added a definition for third parties leveraging the language in the Third Party Registration Framework, as appropriate.
    - Reordered the operations related definitions provided.

SharePoint/NAIC Support Staff Hub/Member Meetings/H Cmte/2026_Fall/WG-BDAI/Summary-of-Changes.docx

Brian Bayerle
Chief Life Actuary
202-624-2169

David Leifer
VP & Sr. Associate General Counsel
202-624-2089

Colin Masterson
Sr. Policy Analyst
202-624-2463

July 21, 2026

Commissioner Nathan Houdek
Chair, NAIC Big Data and AI (H) Working Group

Re: Feedback on AI Systems Evaluation Tool v4.0

Dear Chair Houdek:

The American Council of Life Insurers (ACLI) appreciates the substantial work that regulators and NAIC staff have devoted to the development of the AI Systems Evaluation Tool. Since the initial exposure, the Working Group has made meaningful progress in refining the Tool, incorporating stakeholder feedback, and advancing a pilot process designed to better understand how the framework may function in practice. As the pilot concludes and the Working Group considers future revisions, ACLI welcomes the opportunity to provide additional feedback informed by member experiences and discussions regarding the version of the Tool used in the pilot (v4.0). Our comments are intended to help clarify key definitions, better align the Tool with its intended examination uses, and ensure that information requests remain focused on areas of meaningful regulatory value.

We hope our comments help inform future exposures and encourage continued dialogue as the Working Group moves toward final adoption of the Tool later this year.

General Comments:

ACLI recommends the Working Group clearly distinguish which portions of the Tool are intended for market conduct examinations and which are intended for financial examinations. These examination functions operate under different authorities, processes, and objectives, and clearer identification of the intended examination context would improve implementation and promote greater consistency across jurisdictions.

Regulatory consistency would be further bolstered by the inclusion of language into the Pilot explicitly allowing companies to produce the same submission to multiple states if the requests were similar and made within a reasonable timeframe. Such coordination among market conduct examinations will minimize duplicative and burdensome requests from multiple states.

Language in the Tool explicitly clarifying confidentiality protections for participating companies would also be prudent.

Definitions:

Several key definitions remain ambiguous and may be interpreted differently by companies and regulators. As suggested below, greater precision is necessary to promote consistent implementation and ensure the Tool is applied only to those systems intended to be within scope. ACLI also recommends reviewing the relationships among the defined terms to ensure there is a clear hierarchy in the framework. As drafted, the interaction among some concepts is unclear and could lead to inconsistent interpretation.

The definition of "AI System" remains overly broad and could be interpreted to encompass traditional deterministic algorithms, scoring methodologies, pricing formulas, and other fixed models that operate according to predefined rules and are not capable of autonomously adapting or modifying their behavior. Such systems are transparent, subject to established governance processes, and may already be reviewed by insurance regulators through existing rate, form, or examination frameworks. ACLI recommends clarifying that traditional rule-based models and fixed mathematical formulas are not AI Systems for purposes of the Tool.

The term "AI System Model" is not defined, although it is used throughout the Tool. This creates ambiguity as to whether the reporting obligation is intended to apply to AI Models, AI Systems, or both. If the intent is to focus on AI Models, ACLI recommends replacing references to "AI System Model" with "AI Model" and incorporating the National Institute of Standards and Technology (NIST) definition:

> "**AI Model**" is a component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs."

The definitions of "Adverse Consumer Outcome" and "Consumer Impact" use the term "insurance regulatory standards" without articulating what exactly is intended by such standards. To avoid this concept potentially being interpreted to apply to virtually every process within an insurance company, we suggest defining this term to align with the regulatory models captured by the NAIC AI Model Bulletin and the Model Law on Examinations:

> "**Insurance Regulatory Standards**" means the following laws and regulations enforced by the Department relating to insurance practices: The *Unfair Trade Practices Act* [*insert citation to state statute or regulation corresponding to Model #880*]; The *Unfair Claims Settlement Practices Act* [*insert citation to state statute or regulation corresponding to Model #900*]; The *Corporate Governance Annual Disclosure Act* [*insert citation to state statute or regulation corresponding to Model #305*]; The *Property and Casualty Model Rating Law* [*insert citation to state statute or regulation corresponding to the Model #1780*];The *Market Conduct Surveillance Model Law* [*insert citation to state statute or regulation corresponding to Model #693*]; and The *Model Law on Examinations Act* [*insert citation to state statute or regulation corresponding to Model #390*].

Similarly, "Consumer Impact" as a concept could be interpreted to apply to virtually any process, so we suggest narrowing the definition to clarify that it applies to cases which may directly and adversely impact the consumer:

> "**Consumer Impact**" refers to an AI System or AI System output that is consumer-facing or that informs an insurer decision that directly and adversely affects a consumer and is subject to Insurance Regulatory Standards enforced by the Department.

The term "Material Financial Impact" created significant confusion for companies. We suggest refining the definition to make it more straightforward and less likely to produce inconsistent interpretations. In addition, it should be clarified that a Material Financial Impact classification does not result in an error or failure that would meet a financial statement materiality threshold.

> "**Material Financial Impact**" refers to an AI System with the potential to result in a significant operational loss or direct impact on an insurer's financial statements, assessed on an inherent risk basis without consideration of controls or governance measures.

The category of "Other" under the Operations definitions is overly broad and should align with considerations appropriate for a market conduct or financial examination. We suggest aligning this to the terms "Consumer Impact" and "Material Financial Impact," which would be the focus of those examinations:

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

We note the term "Degree of Potential Harm to Consumers" is not used in the Tool and should be removed from the definitions.

Exhibits A-D:

ACLI recommends narrowing the scope of the exhibits to exclude predictive models with transparent and explainable structures. The Tool should focus on AI Systems that involve more complex decision-making processes or where it is more difficult to connect inputs to outputs. Traditional predictive models with transparent methodologies are well understood and subject to existing governance and oversight practices. Including such models would substantially expand reporting obligations without providing commensurate regulatory value. At minimum, ACLI recommends excluding predictive models with transparent structures, such as Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs), from the scope of the exhibits. This approach is consistent with recent analysis by EIOPA, which concluded that risks associated with GLMs and GAMs arise primarily from governance and business practices rather than the model architecture itself.

Exhibit A:

The operations or program areas in scope of Exhibit A should align with the intended use of the Tool in market conduct or financial examinations. Accordingly, the scope should be limited to operations or

program areas that may result in either Consumer Impact or Material Financial Impact. The exhibit should not apply to enterprise productivity tools or other internal business applications that fall outside the business of insurance or that do not directly affect consumers or insurance decision-making. For these reasons, ACLI recommends striking "Legal/Compliance" from the list of operations or program areas. With respect to "Producer Services," insurers frequently lack visibility into and control over AI Systems independently deployed by producers, and producer conduct affecting consumers is already addressed through existing market conduct oversight; ACLI therefore recommends either striking "Producer Services" or clarifying that the category is limited to AI Systems deployed or controlled by the insurer. Finally, the terms "Waste" and "Abuse" in the "Fraud/Waste & Abuse" category derive from health program-integrity frameworks (such as those administered by the federal Centers for Medicare & Medicaid Services) and lack a settled meaning across all lines of insurance. While insurers' anti-fraud functions are appropriately within examination scope, ACLI requests that the Working Group clarify that "Waste & Abuse" is limited to the health program-integrity context in which those concepts are defined.

The concept of "currently in use" is vague, and we would suggest incorporating a clarification into the company instructions at the beginning of the Exhibit. We would suggest the inclusion of the following sentence:

> For purposes of this Exhibit, "currently in use" means used within the past 12 months.

Similarly, "implemented" in the fifth column of Exhibit A should be clarified as meaning system model(s) that are in full production to avoid including models that are in a limited pilot or testing phase prior to full implementation.

Exhibit D:

ACLI continues to recommend the elimination of Exhibit D. As previously noted, the exhibit requires the collection and manual reporting of extensive information that is overly broad, operationally burdensome, and of limited regulatory value. Moreover, the focus of Exhibit D extends beyond the oversight of AI Systems and into areas related to third-party data and data governance that are being addressed through other NAIC initiatives. In the alternative, if Exhibit D is retained, its scope should be limited to AI Systems determined to be "high risk."

We appreciate the consideration of our feedback as the Working Group continues refining the Tool for eventual adoption.

Sincerely,

B Bonfanti David M. Leifer Colin Masterson

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