

Big Data and Artificial Intelligence (EX) Working Group Survey Comments

APCIA Comments	1
Arbor Strategies Comments.....	4
Barry Duffy / CalypsoAI Comments	12
InsurTech Coalition Comments.....	13
NAMIC Comments.....	16

July 6, 2021

Commissioner Doug Ommen, Chair
Superintendent Elizabeth Kelleher Dwyer, Vice-Chair
Commissioner Mark Afable, Vice-Chair
Big Data and Artificial Intelligence (EX) Working Group
NAIC Central Office
1100 Walnut, Suite 1500
Kansas City, MO 64106-2197

Attn: Denise Matthews, Data Coordination and Statistical Analysis Director at NAIC
Tim Mullen, Director of Market Regulation

VIA Electronic Mail: dmatthews@naic.org and tmullen@naic.org

RE: Artificial Intelligence/Machine Learning Draft Survey

Dear Commissioner Ommen, Superintendent Dwyer, and Commissioner Afable:

The American Property Casualty Insurance Association (APCIA)¹ appreciates the opportunity to provide feedback to the National Association of Insurance Commissioners' (NAIC) Big Data and Artificial Intelligence (EX) Working Group (Working Group) regarding the draft industry survey on Artificial Intelligence (AI) and Machine Learning (ML) (draft survey). APCIA supports the Working Group's efforts to understand the insurance industry's use of AI/ML in the context of private passenger automobile insurance as a means of informing its next steps and appreciates the thought and work that went into developing the current draft.

As detailed below, APCIA is concerned that the survey, as currently drafted, will not provide companies with a consistent understanding of the information being sought, and, ultimately, may not provide the NAIC with a meaningful and accurate picture of the insurance industry's use of AI/ML. APCIA finds that many of the questions are too generalized and open to different interpretations (e.g. a model governance or ML fairness assessment might mean something different from person to person), while at the same time seek unnecessary, detailed information (e.g. model name and external data providers). Further, even assuming the ambiguities are reconciled, responding to this draft survey will be labor intensive, requiring coordination among multiple departments at the company. It will take time to respond to this draft survey with the level of thought and detail this important topic deserves. For this reason, a minimum of 90 days would be requested for responding in any meaningful way. Also, APCIA understands that this survey applies only to private passenger automobile insurance. If later surveys are issued applicable to additional lines of business, APCIA would request an opportunity to provide comment on those as well.

Included below are initial examples to explain our concerns, but ultimately, we respectfully request the Working Group provide additional time to identify specific edits for the drafters' consideration. APCIA

¹ Representing nearly 60 percent of the U.S. property casualty insurance market, the American Property Casualty Insurance Association (APCIA) promotes and protects the viability of private competition for the benefit of consumers and insurers. APCIA represents the broadest cross-section of home, auto, and business insurers of any national trade association. APCIA members represent all sizes, structures, and regions, protecting families, communities, and businesses in the U.S. and across the globe.

recognizes the timeliness of this issue and the additional time is not intended to delay the process, but rather offer constructive engagement that will assist regulators in achieving their objectives.

Ambiguity

General Tab

The survey should clearly state that if a respondent answers that they do not use AI/ML, then they do not have to fill out the remainder of the survey.

Consider deleting the question seeking information about “other areas not listed above,” since it appears the question does not further the needs of the survey.

The question about a company’s “plan to use or explore to use” in the future is unclear. AI/ML is an emerging tool and plans as to the use/non-use will continually evolve as companies decide how, and if, AI/ML fits into their business.

Definitions

While APCIA appreciates the drafters’ intent to narrow the scope of the definition of AI/ML, there is a consistent lack of understanding to what is and is not in scope. Ambiguity has raised questions as to whether the scope truly is narrow and if different interpretations will skew survey responses. For instance, some interpretations of the definition of AI/ML suggests it could apply to broad categories of algorithms that do not rise to the level of AI/ML. This would increase the scope beyond the survey’s purpose and add to what is already an overly detailed and burdensome survey. The definition of AI/ML also suggests properties that do not currently exist. The system does not just respond to any condition in the external environment, but rather responds to data inputs. Additionally, is the survey intended to only address consumer impacts or are the questions designed to understand the use of ML for analysis, not production purposes or other back office uses? This is not clear and could lead to varying company responses. Our recommendation is to limit it to the consumer impact, but to be clear that is the intent. Finally, AI/ML seem to be discussed as if they are the same thing. Is that the drafters’ intent?

APCIA is digging into the definition document in an effort to provide specific consensus recommendations.

Inaccurate Interpretation

APCIA is also concerned that the binary nature of some of the responses will not give the regulators an accurate view into how AI/ML is being used. It is certainly our recommendation that closed questions routinely offer more comparable results than open-ended questions. However, APCIA suggests this can be done in a way that also elicits an understanding of how important certain aspects or elements of the use are. A binary response requires the respondent to answer yes/no for the whole category regardless of where or how it is used in the process, thus creating potential false positives of areas of concern for regulators. For example, a respondent may answer “yes,” but in reality, the respondent may only be using the technology to repair data quality or to extract data from historic PDF documents in back-office operations.

One additional thought, could the survey and subsequent report accurately correlate the maturity of a governance model with the maturity level and risks associated with the AI/ML program?

Too Detailed and Unnecessary

Requesting companies to identify model names, AI in research and prototype stages of development, “other” non-traditional data elements, and external data providers are examples of the type of information that could divulge proprietary and intellectual property and/or compromise competitive positions. This information is not even shared broadly within companies during the initial development stages due to the proprietary nature of it. APCIA is also unsure what value these data points provide to inform the regulators’ mission of getting a better understanding of how insurer’s use AI/ML and the governance approaches. Additionally, it

is unlikely that many vendors would share sufficient information to answer this survey, which will eventually be included in a public report.

Respectfully, while there is a significant amount of information requested by this survey, the type of information actually requested is not specific enough as to shed light on key questions like model robustness or model bias mitigation. Additionally, the survey does not take into account that AI/ML use would be subject to existing privacy and cybersecurity law and regulations, and accordingly, it is not clear what additional compliance information is being sought from a security perspective. One option for consideration is whether regulators would consider a narrower survey that would give an initial overview of industry use. This survey could then educate a follow-up survey or be complimented by voluntary individual company conversations.

Process

In addition to the substance, APCIA has the following questions about the process for the data collection and subsequent report:

- (1) How will the survey be distributed – by the NAIC or individual states?
- (2) Is the survey voluntary?
- (3) Will the survey go to all insurers writing business in the identified line of business or will there be a certain premium threshold?
- (4) Can companies report on a group basis?
- (5) How will confidentiality and trade secret protections be implemented?
- (6) What will be the record retention period for the survey responses?
- (7) Will there be protections that the information will not trigger targeted market conduct exams?
- (8) How long will companies have to complete the survey?
- (9) What will the public report look like? Do you anticipate that it will be in narrative form or just an aggregated/anonymous version of the excel spreadsheet? How will meaningful connections be made to avoid inaccurate assessments of the data?

Thank you again for the opportunity to provide high-level feedback. APCIA is happy to answer any questions that you may have and looks forward to working with you on further developing this survey.

Respectfully submitted,

Angela Gleason

Arbor Strategies, LLC

Chris Petersen

Office: 804-916-1728

Mobile: 202-247-0316

cpetersen@arborstrategies.com

July 6, 2021

VIA E-MAIL

Commissioner Doug Ommen, Chair
Big Data and Artificial Intelligence (EX) Working Group
Iowa Insurance Division
Two Ruan Center
601 Locust, 4th Floor
Des Moines, Iowa 50309-3738

Re: Big Data Survey Draft and Definitions

Dear Commissioner Ommen:

I write on behalf of a coalition of health insurers¹ representing some of the country's largest major medical insurers and health maintenance organizations to comment on the recently exposed *Big Data Survey Draft* and the definitions document that accompanies that draft. We appreciate the opportunity to provide input.

We have a number of specific amendments that we suggest to the language of the definitions document. A redline of those proposals, along with some specific questions and comments, is attached along with this letter.

At a high level, we suggest that the NAIC draw the appropriate distinction between “artificial intelligence” (AI) and “machine learning” (ML). Artificial intelligence covers a broad spectrum of issues, many with different considerations than machine learning. The definitions proposed by the NAIC in the current document make clear that it is actually machine learning –

¹ Coalition members include CVS Health/Aetna, Anthem, Cigna and UnitedHealthcare, who together provide health insurance and health maintenance organization coverage to more than 200 million members nationwide.

the ability of software to be trained to recognize patterns and draw conclusions, rather than simply

perform rules-based activities, that is at the core of the conversation. Our suggested amendments to the first paragraph are intended to make this clarification.

Similarly, we provide edits to the bullets in the section labeled “*Artificial Intelligence/Machine Learning Systems include*” to attempt to clarify that ML is not intended to cover every operation that may occur without human intervention. We are clear that the NAIC does not intend to bring general rules-based systems into scope (which could include things such as thermostats and thermometers); our amendments are intended to make that clarification. We suggest that further clarification to ensure that only true “machine learning” is included in the scope of the definitions is likely warranted.

In addition to those comments, we make the following suggestions:

There is no opportunity in the survey instrument to discuss the notions of “explainability” or accuracy. That is, given that this is an evolving concept at its core, there should be some narrative opportunity for carriers to explain how their systems reliably reach conclusions and are becoming more transparent.

It appears that the view of ML is, at present, quite over-simplified and lacks important nuance. As a general principle, there are three types of ML techniques – supervised learning, unsupervised learning, and semi-supervised learning. The examples in the work group materials focus almost entirely on unsupervised systems that can learn without guidance from humans, using data that is unlabeled and organized in free form. Most systems today, however still work within constraints of requiring pre-labeled data and strict limits on what aspects of the algorithm the machine can alter by itself (supervised learning) or systems that combine teaching from using labeled data to inform how to look at the unorganized (unlabeled) data. These supervised and semi-supervised systems are less susceptible to uncontrolled learning, have increased opportunities for gate checks, and should be given appropriate focus and discussion within the survey.

We also note that the survey document itself uses language and requires information that is overbroad. For example, multiple demands to list “all data vendors” is burdensome, but more importantly, unnecessary. Companies may utilize hundreds of sources for data, and certain health carriers may leverage multiple sources, including public ones, from which data can be gathered. It is more important, at this point in the evolution of ML, to understand the categories of data sources rather than the identity of any particular “vendor.”

As a last point, we note that it is unclear what is intended to be covered by the term “vendor” in the context of the survey. If this is only intended to cover data that is directly purchased from a third-party that collects data and then resells it, that should be defined.

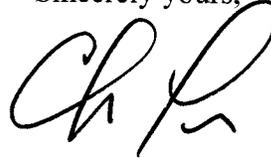
Arbor Strategies, LLC

July 6, 2021

Page | 3

We thank you for the opportunity to provide our input and suggestions. We look forward to discussing them with you in the near future. Please contact me at 202-247-0316 or at cpetersen@arborstrategies.com with questions regarding our comments.

Sincerely yours,

A handwritten signature in black ink, appearing to read 'Chris Petersen', written in a cursive style.

Chris Petersen
Arbor Strategies, LLC

cc: Tim Mullen
Denise Matthews
CVS Health/Aetna
Anthem
Health Care Services Corporation (HCSC)
UnitedHealthcare

**Data Survey Definitions
DRAFT**

Artificial Intelligence/Machine Learning (AI/ML)

AI/ML Machine learning describes an automated process a set of methods in which a system software can be trained to recognize begins recognizing patterns and draw conclusions without explicit instruction about how to do so without being specifically programmed to achieve a pre-determined result. This is different from a heuristic or rules based algorithm or process that is explicitly developed to solve a problem standard algorithm in that an algorithm is a process or set of rules executed to solve an equation or problem in a pre-determined fashion. There is no model created by the software. Evolving algorithms are considered a subset of AI/ML.

Artificial Intelligence /Based Machine Learning Systems include:

- Systems that adapt and adjust to new data and experience without manual human intervention are trained with examples of data to either classify or predict an outcome or event based on new or current data.
- Systems that arrive at results for which the outcomes and the stepwise approach toward the outcomes were not configured in advance by a human programmer.
- Systems that dynamically respond to conditions in the external environment without the specific nature of such responses being known in advance to the designers of the systems.
- Systems that utilize leverage neural networks and/or deep learning algorithms, such as supervised, semi-supervised, and unsupervised learning algorithms.
- Systems that engage in automatic speech recognition, facial recognition, image recognition, text recognition, natural language processing, generation of customer-specific recommendations, automated customer communications (e.g., chatbots with non-preprogrammed prompts), autonomous or semi-autonomous vehicle operation or data gathering, or any other approach that does not require either preprogramming or a manual human intervention in every instance of an action or decision.
- Systems that automatically generate adaptive responses based on interactions with a consumer or third party.
- Systems that determine which data elements to rely upon, in a non-preprogrammed fashion, among a variety of possible alternatives.

Artificial Intelligence / Machine Learning Systems are not:

- Static "scorecards" that deterministically map consumer or other risk characteristics to treatments or decisions. (However, an AI/ML system may use the output of such static "scorecards" as input data for the AI/ML system to consider.)
- Systems with solely preprogrammed decision rules (e.g., "If A, then B" applied invariably in all situations).
- Tables of point or factor assignments in rating plans.

Comment [A1]: We are unclear which algorithms are being referenced.

Comment [A2]: Algorithms are trained on historical data to classify or predict an event in the future.

Comment [A3]: This does not specifically define ML and is over-broad, as it could, in fact, cover things such as thermostats.

Comment [A4]: While most of these descriptors may include ML, there are other systems that have automated responses based on rules that do not use ML. This language should be more focused and tailored to specifically define ML and clarify what the goal is to specify.

Comment [A5]: "Static" is not the appropriate adjective. It would be helpful to have a clarification of the goals for this provision.

- **Static ratemaking and/or predictive-modeling methodologies, including linear regression, generalized linear modeling (GLM), or generalized additive modeling (GAM).**
Purely informational static databases, such as databases used to obtain reference amounts for claim settlements, or static databases pertaining to consumer characteristics or experience, regardless of the amount of information in the database. However, if AI/ML is used to create a static predictive model, that AI/ML system is considered within the scope of this survey.
- **Deterministic “phone trees” that navigate consumers through pre-recorded voice prompts.**
- **Any approach that an insurer could have realistically utilized in the year 2000 or prior.**

Comment [A6]: Same comment.

Comment [A7]: Why this date?

AI/ML Use Descriptions and/or Explanations

- **Rating: AI/ML Uses**
 - **Rating Class Determination:**
 - **Price Optimization: As defined in the NAIC Casualty and Actuarial Statistical (C) Task Force white paper:**
https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf
 - **Retention Modeling:**
 - **Numerical Relativity Determination:**
- **Underwriting: AI/ML Uses**
 - **Automated Approval:**
 - **Automated Denial:**
 - **Underwriting Tier Determination:**
 - **Company Placement:**
 - **Input into Non-Automated Approval Decision:**
 - **Input into Non-Automated Denial Decision:**
 - **Automate Processing Thru the Agency Channel:**
- **Claims: AI/ML Uses**
 - **Claim Approval:**
 - **Claim Denial:**
 - **Determine Settlement Amount:**
 - **Claim Assignment Decisions:**
 - **Informational Resource for Adjusters:**
 - **Evaluation of Images of the Loss:**
- **Fraud Detection: AI/ML Uses**
 - **Fast Tracking of Likely Non-Fraudulent Claims:**
 - **Referral of Claims for Further Investigation:**
 - **Detect Medical Provider Fraud:**
 - **Detect First-Party Liability:**
 - **Detect Third-Party Liability:**

Comment [A8]: We suggest making clear that the following lists are exclusively addressed to the property/casualty industry. Many are inappropriate, for example, for the health insurance industry.

- **Marketing: AI/ML Uses**
 - Targeted Online Advertising:
 - Identification of Recipients of Mail or Phone Advertising:
 - Provision of Offers to Existing Customers:
 - Identification of Potential Customer Groups:
 - Demand Modeling:
 - Direct Online Sales:

- **Loss Prevention: AI/ML Uses:** With advances in technology, the loss-prevention function is likely to grow in importance as insurers strive to improve their results not only through efforts at risk selection and matching the rate to the risk, but also through reducing the probability that the underlying insured losses would occur in the first place. Possible uses of artificial intelligence in loss prevention for private passenger automobile insurance may include, but are not limited to, the following:
 - **Identification of High-Risk Customers:** The goal of such identification in a loss-prevention context is not to make an underwriting or rating decision, but rather to recognize which specific customers may benefit most from loss-prevention advice and mitigation techniques that the insurer may be able to provide, thereby reducing such customers' frequency and/or severity of losses. For example, an AI/ML system might determine that certain households with youthful drivers are more likely to benefit from risk-mitigation advice and other approaches.
 - **Risk-Mitigation Advice to Consumers:** Artificial intelligence systems might be used to target messaging to consumers based on specific risks identified for a given policy. For example, in a household with youthful drivers, AI/ML-targeted messaging and incentives could focus on ways those drivers could gain experience in a low-risk manner and drive more carefully in day-to-day context. For households in mountainous areas, AI/ML systems could provide targeted advice about safe driving in rugged terrain.
 - **Determination of Advance Payments:** In many situations, small payments issued at or shortly after the time of loss, prior to the full adjustment of the claim, can help the insured or third-party claimant prevent much larger amounts of damage that would otherwise greatly raise the costs of the claim for the insurer. In a private passenger automobile context, examples could include, but are not limited to, (i) making a payment for minor repairs that restore the vehicle to a drivable condition, whereas the insured and/or insurer would have otherwise needed to spend much more money to rent another vehicle or to pay for storage of a non-functional vehicle; (ii) making a payment for prompt, inexpensive medical treatment of a claimant, which could prevent the emergence of a longer-term, chronic, and much more costly health condition; or (iii) making a payment for expenses related to towing an insured's or claimant's vehicle away from the scene of the accident and reasonable costs of storage for the vehicle until the insurer or vehicle owner are able to gain possession of the vehicle. In the absence of such prompt payments, vehicles at towing-company storage yards may accumulate significant charges for which the insurer may ultimately become responsible.

Comment [A9]: Similar comment. It would be helpful to explicitly state that these "AI/ML Uses" are exclusive to the property/casualty industry.

Data Use Definitions

1. Criminal Convictions (exclude auto-related convictions)
2. Demographic (age, gender, address, marital status, other non-behavioral attributes of a consumer or population attributes of an area)
3. Driving Behavior (tickets, years of driving experience, annual miles driven)
4. Education (level of education, GPA)
5. Vehicle-Specific Data (type of vehicle(s) driven or owned, history of the vehicle(s), value of contents inside the car)
6. Facial Detection / Recognition / Analysis (picture to confirm identity, estimate biological age or gender of the consumer)
7. Geocoding (latitude and longitude coordinates of a physical address)
8. Natural Catastrophe Hazard (frequency and severity of natural hazards)
9. Job Stability (current employment, length of employment at prior employers, unemployment)
10. Income (annual income, income source)
11. Occupation
12. Personal Financial Information (net worth, type of bank account or credit account, number of bank accounts or credit accounts, available credit, payment history data)
13. Loss Experience (claim history for PPA, claims from other lines of insurance)
14. Medical (medical history, medical condition, prescription data, lab data)
15. Online Media (web searches, online purchases, social media activities)
16. Telematics (time-of-day driving data, location of driving data, braking data, acceleration data, maximum spend, turn speed)
17. Voice Analysis (speed, pitch, volume)]
18. Consumer or Other Type of "Score": A numeric value generated based on a combination of any underlying attributes or behaviors of the consumer, insured risk, or any items considered by the insurer to be relevant to the consumer or insured risk. Scores are computed using deterministic algorithms or models which are not themselves considered to be AI / ML systems. Inquiries in this survey regarding such scores seek to understand whether these scores are used as input data elements within AI / ML systems.
19. Other

Data Use Tab Example

Example:

- **ROW 3: Rating Section:** Put a "Y" in the box for every category of data used in an AI/ML system for a Rating use.
- **ROW 3:** For columns S and T, put a "Y" in the box IF a consumer score for a category not specified in columns B – R is being used and put a "Y" in the box if some type of "Other: Non-Traditional Data Elements" are being used not specified in columns B – R.
- **ROW 4:** Put a Y in the box for every category of data used in an AI/ML system for Rating use that is a derived "score" representing that data category (for example, a consumer "education" score or a driving behavior "telematics" score).
- **ROW 5:** For each category of data element used in an AI/ML system (where there is a "Y" in the box), indicate whether it is sourced internally (I), externally (E) or both internally and externally (B). This pertains to data elements, not "scores".

- **ROW 6:** If Row 5 indicates “E” for externally sources or “B” for sources internally and externally, list the vendor.
- **ROW 6:** In column S, if in Row 5 for column S there is either an “E” or “B”, indicating consumer or other “scores” used are sourced externally or both internally and externally, list the score data providers.

Other AI/ML Use Descriptions:

- What is the difference between a Proof of Concept (POC) and a Prototype?

A Proof of Concept (POC) is a small exercise to test the design idea or assumption. The main purpose of developing a POC is to demonstrate the functionality and to verify a certain concept or theory that can be achieved in development. Prototyping is a valuable exercise that allows the innovator to visualize how the product will function. A prototype is a working interactive model of the end-product that gives an idea of the design, navigation, and layout. While a POC shows that a product or feature can be developed, a prototype shows how it will be developed.

DRAFT

From: [Matthews, Denise](#)
To: [Matthews, Denise](#)
Subject: FW: NAIC Call Notice - Big Data and Artificial Intelligence (EX) Working Group
Date: Wednesday, July 7, 2021 10:27:18 AM

From: Barry Duffy
Sent: Wednesday, July 7, 2021 5:28 AM
To: Mullen, Timothy B.
Subject: Re: NAIC Call Notice - Big Data and Artificial Intelligence (EX) Working Group

CAUTION: This email originated from outside of the organization. Do not click links or open attachments unless you recognize the sender and know the content is safe.

I've registered for Friday's call and am looking forward to it. I also think the survey draft is excellent. It looks to be pitched just right for the subject matter.

Feedback

1. For the listing of AI/ML models, it might be good to ask the model's purpose and type (Vision, Tabular, Time Series, etc.). Model names, particularly internal ones, are unlikely to be indicative of purpose. Also, a single job could be done by a group of models. For example, vehicle damage estimation could be the combined outcome of 3 models

- A vision model for damage location

- A different vision model for damage severity

- A wrapper tabular model taking in the above and adding vehicle location, vehicle cost, etc.

2. The NAIC AI Principles of Robustness isn't included in the Model Governance section. I'd like to know if insurers are able to ensure real-world performance of models. I believe a lot of models get stuck in the Proof of Concept stage because insurers don't have the necessary data/confidence to push the button to go live.

I'll join in on Friday and will see how this develops.

Thanks,
Barry Duffy, Head of Product
CalypsoAI



July 6th, 2021

Hon. Doug Ommen, Chair
Hon. Elizabeth Kelleher Dwyer, Vice Chair
Big Data and Artificial Intelligence (EX) Working Group
National Association of Insurance Commissioners
110 Walnut Street, Suite 150
Kansas City, MO 64106-2197

Dear Commissioner Ommen and Superintendent Dwyer,

The InsurTech Coalition, a group of property and casualty InsurTech companies, welcomes the opportunity to collaborate with the NAIC on the Big Data & AI Working Group's draft survey. We appreciate the opportunity to lend our unique perspective to this important conversation. As leaders in both technology and insurance, we strongly endorse the NAIC's work to ensure that consumer protections are in place around Artificial Intelligence (AI)/Machine Learning (ML)/Predictive Models, and we offer our assistance, education, and partnership to your Working Group and to the NAIC. Below, please find specific comments and responses regarding the exposures of both the Big Data Survey Draft and the Definitions Data Use List Draft from the Big Data and Artificial Intelligence (EX) Working Group. We look forward to further discussions with the shared goal of encouraging a vital and robust insurance market in each state, while protecting and educating consumers.

Confidentiality

With response to anticipated company specific responses to the data survey, we respectfully request that the NAIC ensure the protection of confidential and proprietary data. Our companies are dedicated to providing the necessary data to have meaningful conversations about rate and insurance but want to ensure the protection of said data is sound. Additionally, the data we collect, and the manner in which it is used in rating and underwriting - while completely within the purview of state insurance regulators to examine - may involve the disclosure of trade secret or proprietary information that has substantial competitive value to the market. Accordingly, we have taken significant steps to protect its secrecy, both internally and externally.

It is our recommendation that any responses to the survey containing proprietary or trade secret information be provided in a manner that ensures the confidentiality of the response. This may be achieved by confidentiality agreements with the states and NAIC, the collection of each response by a domestic regulator pursuant to confidential examination procedures or any other process created by the NAIC and state regulators that affords protection to trade secret or proprietary responses.

Data Use

Although the definitions surrounding AI/ML are referenced, the coalition requests more specificity in questions surrounding data use. The difference in models may not provide the NAIC accurate information since companies may define models differently. Broadly worded definitions will cause different and inconsistent interpretations by companies which diminish the quality of the responses. A more specific wording will give companies in all stages better direction in the survey.

Model Governance

The coalition also requests clarity on the question of *Model Governance*, cited below:

If model governance in place, please check (Yes/No) if the following areas are documented in your governance program. Refer to the NAIC AI Principles for a description of each. PLEASE NOTE: It is expected for the governance levels and maturity to be commensurate with the maturity level and risk associated with the program.

Model Governance has a variety of components, including documentation, measurement, and more. Without clarifying the definition of model governance and what it includes, it leaves what constitutes governance open to interpretation. In order to eliminate the ambiguity, we recommend:

- Defining what is included in Model Governance.
- Providing specific levels of maturity with definitions and examples, as well as examples of how these map to maturity levels of the programs themselves.

Furthermore, the options provided in the Model Governance section are ambiguous and overlapping, which may lead to confusion when answering.

- For example, the 3rd category, “Laws Related to Unfair Discrimination” is a type of fairness concern. This would also be covered by the first broader category, “Fairness and Ethics”.
- Additionally, the phrase “Fairness and Ethics” may mean different things to different individuals, because *Ethics* can encompass a broad range of risks and concerns.

The specific questions about compliance with laws are clear and it may be beneficial to provide similar specificity across the set of categories.

Marketing

We request clarity on two questions which appear in almost every category of the survey. The questions are:

- Are consumers provided information regarding the data elements being used and the purposes for which they are being used? (*Answer should be no if not disclosing any information other than what is required by law*)
- Outside of processes required because of FCRA, do consumers have an opportunity to challenge or correct their specific data?

Both questions ask whether carriers have disclosures and processes around the use of consumer data that go beyond existing statutory requirements. The binary nature of these questions means that a broad, indeterminate spectrum of carrier behavior will be captured in each response. The questions leave room for ambiguity as to the robustness the consumer protections each carrier has in place. Additionally, the reading of the questions will hinge on carrier interpretation of current legal requirements. In order to eliminate ambiguity and meaningfully gauge whether carriers have additional consumer protections in place, the Coalition suggests that the survey be amended to inquire as to specific categories of disclosures or specific processes contemplated by the Working Group.

The advancements in technology are an evolution and progression of insurance that will benefit consumers with fairer and better products. The InsurTech Coalition appreciates your attention to these changes and welcomes further conversations. If you have questions regarding our comments, please feel free to contact Dave Luketic at dave.luketic@joinroot.com.

Sincerely,

InsurTech Coalition

July 2, 2021

Commissioner Doug Ommen, Chair
NAIC Big Data and AI Working Group
c/o Denise Matthews at dmatthews@naic.org and
c/o Tim Mullen at tmullen@naic.org
444 North Capitol Street NW, Suite 700
Washington, D.C. 20001

Re: NAMIC comments to initial draft of industry survey on AI/ML and large data sets

Dear Chair Ommen, Co-Vice Chair Dwyer, Co-Vice Chair Afable, and Members of the Working Group,

Please find included herein comments on behalf of the National Association of Mutual Insurance Companies (NAMIC)¹ concerning the NAIC Big Data and Artificial Intelligence Working Group initial draft of a survey to better understand the role of Artificial Intelligence (AI), Machine Learning (ML), and large data sets in insurer operations. We want to thank the Working Group and the leadership of the Chair and Co-Vice Chairs in exploring this important topic concerning AI. NAMIC believes AI/ML creates an outstanding and positive result for insurance consumers who demand innovative and technological advancements in product usage in insurance marketplaces. NAMIC shares the goals of the Working Group in ascertaining appropriate usage of AI/ML and protection of consumers from illegal conduct or harm.

We want to start by thanking the Working Group for exposing the draft survey and the draft definitions. We believe as mentioned in the exposure email that this is a first step and hopefully there will be additional opportunities to provide material and pertinent comments in regard to this matter before a final product is determined. Generally, we appreciate the Working Group's desire to learn how prevalent the use of AI/ML is in the industry and better understand its use before offering further regulatory frameworks, models laws, regulations, or other guidance. This activity seems to be a logical next step in the evolutionary process of reviewing this issue in the regulatory environment.

NAMIC would inquire as to the intent of the Working Group concerning implementing the survey. Is this survey designed to be voluntary for industry or mandatory? If mandatory, what authority do you intend to utilize for collection, and will this data be collected/housed at NAIC? Is this a national project, regional in nature, or only targeted at certain domiciliary states? Is

¹ The National Association of Mutual Insurance Companies is the largest property/casualty insurance trade group with a diverse membership of more than 1,400 local, regional, and national member companies, including seven of the top 10 property/casualty insurers in the United States. NAMIC members lead the personal lines sector representing 66 percent of the homeowner's insurance market and 53 percent of the auto market. Through our advocacy programs we promote public policy solutions that benefit NAMIC member companies and the policyholders they serve and foster greater understanding and recognition of the unique alignment of interests between management and policyholders of mutual companies.

the focus of intent that the survey be an NAIC-level survey or serve as a model for states to utilize within their authority, if applicable?

Further, we are encouraged by the mention of confidentiality but would like to better understand how this information will be kept confidential and whom within the regulatory community it will be shared? There are concerns with exposing sensitive internal operational dynamics that could cause competitive harm if not protected. What assurance can the Working Group provide that these responses will be kept confidential and if they feel they have adequate authority to prevent improper disclosures to third parties?

Additionally, with the understanding that this is only a snapshot in time, changes in certain responses by the industry at a later date should not be a cause for regulatory concern regarding prior responses. That is, industry responses to this survey should not be used in later examinations as evidence of anything other than a timely response to the survey. It is also important to understand if there will be any regulatory compliance issues that flow from responses as the Artificial Intelligence Principles have a number of broad subjective concepts that are not currently defined in most, if not all, insurance state codes.

In regard to the actual survey and definitions, we think that at least initially, there is concern whether there is a stable and uniform understanding and definition of AI/ML so that there will not be regulatory issues with responses or misunderstandings as to future intent. There should be an opportunity to provide explanations where deemed appropriate.

Finally, the initial draft is very thorough in scope. However, NAMIC would query whether there has been any thought in possibly implementing the survey in stages in order to gain feedback from early responses to initial tabs and ascertain where improvements can be made throughout the process and the efficacy of further data collection?

In sum, we thank the Working Group for the ability to respond to this early draft, look forward to working with the Group on further drafts, and appreciate consideration of our comments.

Sincerely,



Andrew Pauley, CPCU
Government Affairs Counsel
National Association of Mutual Insurance Companies (NAMIC)