

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

This document provides survey filing guidance and data definitions. It is organized by survey sections.

This survey aims to aid regulators in understanding what artificial intelligence/machine learning (AI/ML) techniques life insurers are using and where in their insurance operations they are using them. This survey does not aim to have insurers provide details of trade secret components of the various techniques they are implementing. This Life Insurance survey is restricted to interests in three operational areas: 1) pricing and underwriting; 2) marketing; and 3) risk management. Within each operational area, interest is restricted to a limited number of use cases. It is expected that leaders and data scientists in these operational areas will have significant involvement in responding to the survey items.

Any company licensed to write life insurance in one of the 14 participating states (Colorado, Connecticut, Illinois, Iowa, Louisiana, Minnesota, Nebraska, North Dakota, Oregon, Pennsylvania, Rhode Island, Vermont, Virginia, and Wisconsin) and meeting one or more of the criteria below is required to complete the survey.

**Criteria for Selecting Participating Insurance Companies:**

1. A selected InsurTech company.
2. A company with more than \$250 million in premiums on all individual policies in 2021.
3. A term life insurance writer that has issued policies on more than 10,000 lives.

A state market share analysis must show that the selected companies adequately represent the 14 states.

**Definition of Artificial Intelligence/Machine Learning for This Survey – Applicable to All Sections**

For purposes of this survey, AI is defined as models that can simulate learning in performing tasks. ML is a subset of algorithms that facilitate learning without being explicitly programmed to achieve a predetermined result. Models that are considered AI and built using ML include robotics, natural language processing, and sentiment analysis.

Standard features of AI systems adopted for purposes of this survey include:

- Systems that adapt and adjust to new data and experiences without manual human intervention.
- 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 use deep-learning neural network algorithms.
- Systems that learn under a supervised, semi-supervised, unsupervised, or reinforcement learning style.
- 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), or any other approach that does not require either preprogramming or human intervention in every instance of an action or decision.
- Systems that automatically generate adaptive responses based on interactions with a consumer

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITONS**

or third party.

- Systems that determine which data elements to rely upon, in a non-preprogrammed fashion, among various possible alternatives.
- Generalized linear modeling (GLM) or generalized additive modeling (GAM) are considered ML. See the Appendix for an expanded list of ML techniques.

For purposes of this survey, the following AI systems are excluded:

- Scorecards that deterministically map consumer or other risk characteristics to treatments or decisions.
- Tables of point or factor assignments in risk classes.
- Deterministic “phone trees” that navigate consumers through prerecorded voice prompts.

**Components of Survey**

This Life Insurance survey is made up of the following sections:

1. Respondent Information
2. General Information
3. Underwriting and Pricing Product Matrix
4. Pricing and Underwriting
5. Marketing
6. Risk Management
7. Governance

The remainder of this document will describe each of the above sections to help respondents understand the intent of the questions and expected responses.

**Section 1: Respondent Information**

This section requests basic demographic information on the respondent, as depicted below. Sample responses have been populated.

RESPONDENT INFORMATION		
1	NAIC Company Code	28207
2	Company Name	Best
3	Contact Name	JC
4	Contact Title	Mr
5	Contact Phone Number	555-555-5555
6	Contact Email Address	j@naic.org
7	Comments:	Yes

Comments are optional but are encouraged if any of the respondent’s responses need to be clarified.

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

**Section 2: General Information**

This section attempts to understand to what extent the respondent uses AI/ML in three key operational areas: 1) pricing and underwriting; 2) marketing; and 3) risk management. This survey primarily focuses on consumer impact models used in these operational areas. However, the respondent can include other operational areas in the “Other” line in Question 3.

The respondent is referred to the definition of AI/ML above to respond to the questions in this section. There are six basic questions in this section of the survey. They are summarized as follows:

1. Does your company use, plan to use, or is currently exploring the use of AI/ML?
2. Provide the implementation status of AI/ML in each of the operational areas identified above. The status options are:
  - a. N/A
  - b. < 1 Year
  - c. 1–3 Years
  - d. > 3 Years
3. List areas other than pricing and underwriting, marketing, and risk management for which your company uses, plans to use, or is currently exploring the use of AI/ML.
4. Indicate whether your company has developed or plans to develop AI/ML systems internally, use vendors, or both internally with input from vendors.
5. Choose from the reasons given why your company does not use, plan to use, or explore AI/ML.
6. Indicate whether your company is using automated ML tools like DataRobot or Salesforce Einstein.

The respondent may further elaborate on any response in the “Additional Comments” column provided. The respondent is not required to add additional commentary.

**Section 3: Underwriting and Pricing Product Matrix**

This section of the survey attempts to understand any product attribute thresholds that may be in place that trigger the use of AI/ML by the respondent. The two threshold attributes of interest are age and face amount for the following products:

1. Term life
2. Whole life
3. Universal life (including variable universal life)

For each product, the respondent is asked to specify the following:

1. The product types subject to AI/ML processes.
2. Minimum and maximum age across products that trigger AI/ML use.

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

3. Minimum and maximum face amounts that trigger AI/ML use.

There are dropdown boxes to indicate no minimum and no maximum age and face amount thresholds for each product. The respondent may select “Applicable” or “Not Applicable” as appropriate from response options.

The table below depicts the product types listed for term, whole life, universal life, and variable life that appear in the product response options.

<b>Product Response Options</b>			
<b>Term</b>	<b>Whole Life</b>	<b>UL</b>	<b>VL</b>
Level	Single Premium	Universal Life	Variable Life
Convertible	Simplified Issue	Variable Universal Life	
Annual	Guaranteed Issue		
Decreasing	Periodic Premium		
Group	Group Life		
Other			

For the term product, “Other” may be selected, and the respondent is expected to enter the corresponding product type(s) in free-form text.

**Operational Areas**

The following discusses the instructions that are common to each operational area. Instructions unique to each operational area will follow this section.

Each operational area has specific “uses” listed for AI/ML. The first question in each operation area is: Does your company use, plan to use, or is currently exploring the use of AI/ML? This is the same question from the “General Information” section but is being asked for each specified use case.

The follow-up question requests the respondent to identify the level of deployment of AI/ML for the use case. The levels are: research; proof of concept; prototype; or implemented in production. For example, “Setting Pricing Assumptions” is a use case listed under the “Pricing and Underwriting” section. The respondent should select the highest level of deployment of AI. The deployment levels from the lowest level of use to the highest level of use are:

- **Research:** The investigation into and study of materials and sources to establish facts and reach new conclusions. This is the collection of information about a particular subject and is the lowest level of use.
- **Proof of Concept (POC):** 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 can be achieved in development. It is testing the model for functional

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITONS**

viability to be sure it runs and delivers a result.

- **Prototype:** Prototyping provides the opportunity to visualize how the product will function. It is a working interactive model of the end product that gives an idea of the design, navigation, and layout. Prototyping involves testing the model with actual data, in a limited, controlled environment. A prototype brings the POC idea to life.
- **Implemented in Production:** The model is being used in a live, production environment using real data. This is the highest level of use.

The third common question asks the respondent to list each AI/ML model that has been developed for each use case either internally, by a third party, or both. For each model listed, there is a grid (see below) that requests the respondent to indicate which AI/ML techniques are used for each named model. The respondent is required to indicate each of the techniques that a model uses by selecting the “Yes” option for the applicable technique. A sample grid is below. The red triangles in the heading row can be hovered over to view an expansion of the associated acronym for convenience. Just hover over the red triangle to see the full name.

AI/ML Model Category Types												
Select All That Apply												
DL	ENS	NN	REG	RS	RGS	BAY	DT	DR	IB	CLU	COX	Other - Specify

The column headings are defined as follows:

1. DL – Deep Learning
2. ENS – Ensemble
3. NN – Neural Network
4. REG – Regularization
5. RS – Rule System
6. RGS – Regression
7. BAY – Bayesian Methods
8. DT – Decision Trees
9. DR – Dimensionality Reduction
10. IB – Instance-Based
11. CLU – Clustering
12. Cox Proportional Hazard Modeling

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

13. Other

For example, if a named model used principal component analysis (PCA), a dimensionality reduction (DR) technique, to reduce a data set to independent principal components and then used those principal components as independent variables in a regression model (RGS), the respondent would indicate “Yes” for both the DR and RGS columns. If the list of techniques above does not reflect the techniques(s) used by the named model, the respondent is to select the “Other” option and fill in the name of the technique. The respondent is allowed to use a combination of the listed techniques and the “Other” category. The company should indicate all categories that apply.

If there are 10 or less models, the respondent is required to provide: the names of the models; the ML techniques used; additional information on whether each model was developed internally, by a third party, or both; its level of influence; and whether there is a model governance framework in place for the model. This information can be listed in parentheses after the model’s name.

If there are more than 10 models, the models with the highest level of deployment (i.e., most impactful) should be listed in the first 10 lines, and only the models with the lowest level of deployment (i.e., least impactful) should be listed on line “K” of each named model section.

If only researching the “use” of an AI model for an operational area, it is understood a model may not have a name associated with it. If that is the case, the respondent should indicate “Not Yet Named” as the AI/ML model name and complete the remaining columns. If better information can be provided through a brief description of the model, as opposed to a model name, that is acceptable. However, if a third-party vendor has given its model a specific name, then that name should be provided.

For each named model, the survey asks the respondent to indicate whether a named model was developed internally, by a third party, or both.

If the named model was developed by a third party, the respondent must enter the name(s) of the third parties involved.

The survey next asks the respondent to indicate the level of decisions influenced by the named model. The options to choose from are:

1. Automation (no human intervention on execution)
2. Augmentation (model advises human who makes the decision; model suggests the answer)
3. Support (model provides information but does not suggest a decision or action)
4. Other

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

The respondent need not specify an option for the “Other” category.

The survey next asks whether there is a model governance structure in place for each named model. This response expected is either “Yes” or “No” for the named model.

Finally, there is space provided for the respondent to add additional commentary on their responses for each named model. The respondent is not required to add additional commentary.

**Data Use Table Definitions**

For each operational area, the survey seeks to understand the extent to which third-party data is used in the development of AI/ML in the operational areas of pricing and underwriting; marketing; and risk management. The third-party data elements are listed and defined below:

1. **Credit-Based Insurance Scores:** Scores developed by the insurer that are separate and distinct from a score issued by a credit rating agency such as TransUnion but may use similar credit rating variables.
2. **Financial Credit Score:** Scores issued by a credit rating agency such as TransUnion, Equifax, or Experian.
3. **Other Type of Non-Credit “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 systems. Inquiries in this survey regarding such scores seek to understand whether these scores are used as input data elements within AI systems.
4. **Public Records:** Such as criminal convictions (excluding auto-related convictions), bankruptcy records, mortgage information, driving records, an officer in a skydiving club, etc.
5. **Demographic:** Age, gender, address, marital status, other non-behavioral attributes of a consumer or population attributes of an area; Education: Level of education, GPA; Stability: current employment, length of employment at prior employers, unemployment; Income: Annual income, income source; Occupation: Primary profession, service, or trade for which a person is paid.
6. **Telematics Type Data:** Time-of-day driving data, location of driving data, braking data, acceleration data, maximum speed, turn speed, and other telematics-related variables.
7. **Driving Behavior:** Tickets, years of driving experience, and annual miles driven.
8. **Biometrics:** Facial Detection/Recognition/Analysis: Picture to confirm identity and estimate

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

biological age or gender of the consumer; Voice Analysis: speed, pitch, and volume.

9. **Medical:** Medical history, medical condition, prescription data, and lab data.
10. **Online Media:** Web searches, online purchases, and social media activities.
11. **Other:** Any data element category not defined above.

The three common survey questions with respect to the data use table are:

1. Does your company use the following data elements within an AI/ML system? Respond “Yes” or “No” to each data element type.
2. If “Yes,” then indicate if the data source is “Internal,” “External,” or “Both.”
3. If “External” or “Both” in the above question, list each data vendor unless you do not use external vendors for any of the AI/ML scoring models identified above that provide the data element indicated.

The expected response to the first question is “Yes” or “No.” The response to the second question is “Internal,” “External,” or “Both.” The response to the third question is a little more complex. The diagram below will facilitate the explanation.

For each data type, the respondent is required to list the vendors that provide third-party data for that data type. The vendor names should be typed in the white spaces provided below each data type, as depicted in the diagram below. If there are more than 10 third-party data vendors for a given data element, the respondent should list them on line “17K” separated by commas.

PRICING & UNDERWRITING DATA USE		Data Element Types:	
		Credit Based Insurance Score	
15	Does your company use the following data elements WITHIN AN AI/ML SYSTEM - Respond Yes or No to each Data Element Type		
16	If yes to #15, then indicate if data source is Internal, External, or Both		
17	If External OR Both in the above question, list each data vendor unless you do not use external vendors for any of the AI/ML scoring models identified above that provide the data element indicated.	Credit Based Insurance Score	
17A		Vendor 1	Experian
17B		Vendor 2	Equifax
17C		Vendor 3	Transunion
17D		Vendor 4	
17E		Vendor 5	
17F		Vendor 6	
17G		Vendor 7	
17H		Vendor 8	
17I		Vendor 9	
17J		Vendor 10	
17K	If more than 10 AI/ML data vendors are used for the respective data element type, please list them here: (For the OTHER category, also list the OTHER data type being used)		

The respondent may further elaborate on any response in the “Additional Comments” column provided. The respondent is not required to add additional commentary.



**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

These are the end of the common sections for the three operational areas. Questions unique to each operational area are discussed in the next three sections of this guidance document.

**Section 4: Pricing and Underwriting**

Definitions and Explanations Specific to Pricing and Underwriting

- **Setting Assumptions:** Used to understand the deviation of actual experience from current assumptions and to understand mortality when no prior knowledge exists.
- **Speed and Accuracy:** Used to verify application and claims data and speed approval processes. Speed up the delivery of policy pricing rates online.
- **Specialty Products for Certain Conditions, Such as Diabetes:** Studying and tracking behaviors that coincide with improved outcomes on certain diseases.
- **Automated Premium Rates:** Differentiated pricing tailored to pricing and market conditions with transparency on the impact of a feature on the price.
- **Automated Approval:** Approving an application without human intervention on that particular application.
- **Automated Denial:** Denying an application without human intervention on that particular application.
- **Underwriting Tier Determination:** Decisions regarding the criteria to use to establish specific named or numbered categories (called tiers) that use combinations of attributes that affect an insurer's underwriting decision.
- **Input Into Non-Automated Approval Decision:** Providing data, analysis, or recommendations regarding a decision to approve an application in a situation where a human decision-maker still has the ability and responsibility to affirmatively consider this information and make a decision independently of the AI system. In this situation, the AI system cannot automatically approve the application, and protocols exist that ensure that each recommendation from the AI system is actively reviewed and not adopted by default.
- **Input Into Non-Automated Denial Decision:** Providing data, analysis, or recommendations regarding a decision to deny an application in a situation where a human decision-maker still has the ability and responsibility to affirmatively consider this information and make a decision independently of the AI system. In this situation, the AI system cannot automatically deny the application, and protocols exist that ensure that each recommendation from the AI system is actively reviewed and not adopted by default.
- **Automate Processing Through the Agency Channel:** Enabling agencies to receive certain information about applicants automatically without specifically requesting that information and/or to provide quotes to the applicants and/or recommend a decision regarding the application to the agent without being based on preprogrammed decision rules.

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

In addition to the common survey questions, the Pricing and Underwriting section adds an extra request in question 5A. Question 5 relates to specialty products for conditions such as diabetes. Question 5A requests respondents to list the top three conditions for which they have a specialty product. The respondent is required to write the names of the conditions in the text boxes provided.

Under data use, the respondent is asked to respond to the following questions unique to pricing and underwriting for each data element listed:

1. How do you determine the data elements are accurate and reliable?
2. For each data element, how do you determine whether there is a reasonable relation to mortality?
3. Do you disclose to applicants for life insurance what external data is accessed and used in your underwriting program?
4. Do you have a process for applicants for life insurance to dispute mistakes in this data element?

For each of these questions, the respondent is required to provide a response in the text boxes provided.

The respondent may further elaborate on any response in the “Additional Comments” column provided. The respondent is not required to add additional commentary.

### **Section 5: Marketing**

Definitions and Explanations Specific to Marketing

- **Geographic Marketing:** Determination of which individuals by geographic area to send advertisements from the insurer.
- **Targeted Online Advertising:** Determination of which individuals on the internet should receive or see advertisements from the insurer.
- **Identification of Recipients of Mail or Phone Advertising:** Determination of which individuals would be desirable recipients of an insurer’s advertisements via the telephone or physical mail.
- **Provision of Offers to Existing Customers:** Determination of which customers should be notified of new insurance products, discounts, options to be written in a different book of business, or any other benefit or favorable treatment that the insurer seeks to extend.
- **Identification of Potential Customer Groups:** Determination regarding which consumer sub-populations could likely become additional customers of the insurer and/or benefit from the insurer’s products and services.
- **Demand Modeling:** Identification of consumers’ needs for and interest in specific types of insurance and insurance products that the insurer is offering or whose development or sale the insurer may be considering or exploring.
- **Direct Online Sales:** Selling insurance policies to consumers through a direct internet-based

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

channel in a manner that does not rely solely on preprogrammed decision rules.

Under data use, the respondent is asked to respond to the following question related to data use in marketing:

1. How do you determine the data elements are accurate and reliable?

The respondent is required to write a response in the text box provided.

The respondent may further elaborate on any response in the “Additional Comments” column provided. The respondent is not required to add additional commentary.

### **Section 6: Risk Management**

Definitions and Explanations Specific to Risk Management

- **Wearable Devices:** Wearable devices refer to smart electronic devices with sensors that collect and deliver biometric information. The technology includes devices that are worn on the wrist and other forms such as jewelry, glasses, clothing, shoes, and implanted devices. The main category of wearables in the market are fitness trackers and smartwatches, which gather metrics associated with physical activity: step count, activity minutes, distance traveled, floors climbed, calories burned, heart rate, and sleep patterns.
- **Wellness Initiatives:** A program intended to improve and promote health and fitness that's usually offered through the workplace, although insurance plans can offer them directly to their enrollees. The program allows employers or plans to offer premium discounts, cash rewards, gym memberships, and other incentives to participate. Some examples of wellness programs include smoking cessation, diabetes management programs, weight loss programs, and preventative health screenings.
- **Discount Medical Programs:** Programs that provide a discount on medical services or prescription drugs.
- **Technology to Detect Smoking:** Wearable sensor technology used to detect and help people quit smoking.
- **Disease Detection:** AI and ML programs designed to diagnose disease using training data—such as the patient’s history, lab results, scans, symptoms, and images of confirmed and susceptible cases—or real-time facial recognition images.

Under data use, the respondent is asked to respond to the following question related to data use in risk management:

1. How do you determine the data elements are accurate and reliable?

The respondent is required to write a response in the text box provided.

As with previous sections of the survey, the respondent may further elaborate on any response

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

in the “Additional Comments” column provided. The respondent is not required to add additional commentary.

**Section 7: Governance**

The purpose of the questions related to AI/ML governance is to obtain a better understanding of a company’s awareness of specific risk areas tied to the National Association of Insurance Commissioners’ (NAIC’s) Artificial Intelligence Principles. The questions in this section apply to the three operational areas: 1) pricing and underwriting; 2) marketing; and 3) risk management.

The NAIC recommends that insurance companies and all persons or entities facilitating the business of insurance that play an active role in the AI system life cycle—including third parties such as rating, data providers, and advisory organizations (hereafter referred to as “AI actors”)—promote, consider, monitor, and uphold the NAIC AI Principles according to their respective roles.

The principles are intended to establish consistent high-level guiding principles for AI actors. These principles are guidance and do not carry the weight of law or impose any legal liability. This guidance can serve to inform and establish general expectations for AI actors and systems emphasizing the importance of: accountability; compliance; transparency; and safe, secure, fair, and robust outputs.

In addition, the survey seeks information to understand if guidelines or best practices are documented. Specifically, if the company is involved in using AI models, does it have a documented process in place to ensure that AI actors are:

**Fair and Ethical**

AI actors should respect the rule of law throughout the AI life cycle. This includes, but is not limited to, insurance laws and regulations, such as those relating to trade practices, unfair discrimination, access to insurance, underwriting, privacy, consumer protection and eligibility practices, ratemaking standards, advertising decisions, claims practices, and solvency.

Consistent with the risk-based foundation of insurance, AI actors should proactively engage in responsible stewardship of trustworthy AI in pursuit of beneficial outcomes for consumers and to avoid proxy discrimination against protected classes. AI systems should not be designed to harm or deceive people. They should be implemented in a manner that avoids harmful or unintended consequences and corrects and remediates for such consequences when they occur.

**Accountable**

AI actors should be accountable for ensuring that AI systems comply with these

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

principles consistent with the actors' roles, within the appropriate context and evolving technologies. Any AI system should comply with legal requirements governing its use of data and algorithms during its phase of the insurance life cycle. Data supporting the final outcome of an AI application should be retained and be able to be produced in accordance with applicable insurance laws and regulations in each jurisdiction. AI actors should be responsible for any AI system's creation, implementation, and impacts, even if the impacts are unintended. AI actors should implement mechanisms and safeguards consistent with the degree and nature of the risks posed by AI to ensure all applicable laws and regulations are followed, including ongoing (human or otherwise) monitoring and, when appropriate, human intervention.

**Compliant**

AI actors must have the knowledge and resources to comply with all applicable insurance laws and regulations. AI actors must recognize that insurance is primarily regulated by the individual states and territories of the U.S., as well as by the federal government, and that AI systems must comply with the insurance laws and regulations within each jurisdiction. Compliance is required whether the violation is intentional or unintentional. Compliance with legal requirements is an ongoing process. Thus, any deployed AI system must be consistent with applicable laws and safeguards against outcomes that are either unfairly discriminatory or otherwise violate legal standards, including privacy and data security laws and regulations.

**Transparent**

To improve the public's confidence in AI, AI actors should commit to transparency and responsible disclosures regarding AI systems to relevant stakeholders. AI actors must be able to protect the confidentiality of proprietary algorithms, provided adherence to individual state law and regulations in all states where AI is deployed can be demonstrated. These proactive disclosures include revealing the data being used, the purpose of the data in the AI system and the consequences for all stakeholders.

Consistent with applicable laws and regulations, stakeholders (which includes regulators and consumers) should have a way to inquire about, review, and seek recourse for AI-driven insurance decisions. This information should be easy-to-understand and describe the factors that lead to the prediction, recommendation, or decision. This information may be presented differently and should be appropriate for applicable stakeholders.

**Secure, Safe, and Robust**

AI systems should be robust, secure, and safe throughout the entire life cycle so

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

that in conditions of normal or reasonably foreseeable use, or adverse conditions, they can function in compliance with applicable laws and regulations. To this end, AI actors should ensure a reasonable level of traceability in relation to datasets, processes, and decisions made during the AI system life cycle. AI actors should enable analysis of the AI system's outcomes, responses, and other insurance-related inquiries, as appropriate in keeping with applicable industry best practices and legal requirements.

AI actors should, based on their roles, the situational context, and their ability to act, apply a systematic risk management approach to each phase of the AI system life cycle continuously to address risks related to AI systems, including privacy, digital security, and unfair discrimination as defined by applicable laws and regulations.

It is understood that governance models vary in terms of components and terms used to describe these risk areas. However, there is a common thread across most governance models, and this language was specifically used in this survey as it ties directly to the NAIC's AI Principles. Where there may be concerns about overlap, the intention is for this additional information to clarify the unique intent of each. The company should reply to each component as precisely as possible.

**Governance Question 1** seeks to understand whether the respondent has AI/ML-related contracts that include any conditions that would limit disclosure or otherwise limit transparency to regulators. The expected response to this question is "Yes" or "No."

**Governance Question 2A** seeks to understand whether the respondent has governance frameworks in place to address the five fairness and ethics considerations defined in the NAIC AI Principles for each operational area. Include expected response type.

**Governance Questions 2B and 2C** seek to understand whether the respondent has practices in place that ensure accountability for data algorithms' compliance with laws. The respondent is expected to answer "Yes" to indicate they have practices in place that ensure accountability for data algorithms' compliance with laws or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**Governance Questions 2D and 2E** seek to understand whether the respondent has practices in place that ensure accountability for data algorithms' intended impacts. The respondent is expected to answer "Yes" to indicate they have practices in place that ensure accountability for data algorithms' intended impacts or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

**Governance Questions 2F and 2G** seek to understand whether the respondent has practices in place that ensure accountability for data algorithms' unintended impacts. The respondent is expected to answer "Yes" to indicate they have practices in place that ensure accountability for data algorithms' unintended impacts or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**Governance Questions 2H and 2I** seek to understand whether the respondent has practices in place that ensure appropriate resources and knowledge to ensure compliance with laws, including those related to unfair discrimination. The respondent is expected to answer "Yes" to indicate they have practices in place that provide appropriate resources and knowledge to ensure compliance with laws, including those related to unfair discrimination, or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**Governance Questions 2J and 2K** seek to understand whether the respondent has practices in place that ensure transparency with appropriate disclosures, including notice to consumers specific to data being used and methods for appeal and recourse related to inaccurate data. The respondent is expected to answer "Yes" to indicate they have practices in place that ensure transparency with appropriate disclosures, including notice to consumers specific to data being used and methods for appeal and recourse related to inaccurate data, or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**Governance Questions 2L and 2M** seek to understand whether the respondent has practices in place that ensure AI systems are secure, safe, and robust, including decision traceability and security and privacy risk protections. The respondent is expected to answer "Yes" to indicate they have practices in place that ensure AI systems are secure, safe, and robust, including decision traceability and security and privacy risk protections, or "No" if they do not have any practices in place. If the respondent has practices in place, the respondent is expected to briefly describe those practices in the space provided.

**Governance Questions 3 and 3A** seek to understand whether the respondent follows existing guidance that was provided by a third party or a hybrid of an internal governance structure and a third party. The expected responses to questions 3A and 3B are "Yes" or "No."

**Governance Questions 4, 4A, and 4B** request the respondent to identify the third party solicited to help develop their AI/ML governance framework. The respondent is expected to supply the information in the spaces provided.

**Governance Questions 5 and 5A** seek to understand the processes the respondent has in place to

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

provide consumers with the data elements used in AI/ML models that affect them in the operational areas that are the focus of this survey. The follow-up question seeks to understand how the respondent ensures compliance with regulatory frameworks, such as the *Unfair Trade Practices Act* (#880), when using non-Fair Credit Reporting Act (FCRA) data. The respondent is expected to supply the information in the spaces provided.

**Governance Questions 6 and 6A** seek to understand whether the respondent makes consumers aware of the non-FCRA data collected on them and used in AI/ML models. This is a compound question. It further asks the company whether the respondent informs consumers when non-FCRA data is collected and how it is used. If the respondent can answer “Yes” to any part of this question, then the respondent is expected to explain in the space provided the process employed to make consumers aware of the non-FCRA data collected on them, when non-FCRA data is collected, or how it is used.

**Governance Question 7** seeks to understand whether the respondent allows consumers to correct non-FCRA data collected on them and used in respondent AI/ML models. The expected response to this question is “Yes” or “No.”

**Governance Question 8** seeks to understand the extent to which a human is involved in algorithmic decision-making. Does the AI/ML inform decision-making (i.e., the human is in the loop)? Does the AI/ML model replace decision-making (i.e., the human is out of the loop)? Or is it a combination of the two (i.e., the human is on the loop)? The follow-up question seeks to understand how much intervention is involved. The respondent is expected to provide a response in the space provided.

**Governance Question 9** seeks to understand how AI/ML model results are monitored. For example, one type of monitoring could involve the metrics collected to ensure the model is working as designed. The respondent could also monitor the deployment of the AI/ML model to ensure end users are deploying the model as intended. The respondent is expected to provide a response in the space provided.

**Governance Question 10** seeks to understand whether the respondent has a process in place for consumers to contest an adverse underwriting decision. This question only applies to pricing and underwriting. The expected response to this question is “Yes” or “No.”

**Governance Question 11** seeks to understand whether the respondent keeps a log of contested underwriting decisions. The expected response to this question is “Yes” or “No.”



**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

**Appendix: Machine Learning Taxonomy**

The machine learning (ML) taxonomy provides a listing of ML algorithms that fall under each category that makes up an ML taxonomy. The list is representative but not exhaustive. This is one of many existing ML taxonomies. Some include more or fewer techniques than depicted below, more out of convenience than structure.

**Machine Learning Taxonomy Techniques**

Generally, the most common techniques are reflected in each category below. If the technique you are using is not represented by one of the categories below, select the other category and enter the ML type in the indicated column. You may enter more than one ML type separated by commas. Each category below is represented in the spreadsheet by the abbreviations in parentheses.

1. Deep Learning (DL)
  - a. Deep Boltzmann Machine (DBM)
  - b. Deep Belief Network (DBN)
  - c. Convolutional Neural Network (CNN)
  - d. Stacked Auto-Encoder
2. Ensemble (ENS)
  - a. Random Forest
  - b. Gradient Boosting Machine (GBM)
  - c. Bootstrapped Aggregation (Bagging)
  - d. AdaBoost
  - e. Stacked Generalization (Blending)
  - f. Gradient Boosted Regression Trees
3. Neural Networks (NN)
  - a. Radial Basis Function Network (RBFN)
  - b. Perceptron
  - c. Back-Propagation
  - d. Hopfield Network
4. Regularization (REG)
  - a. Ridge Regression
  - b. Least Absolute Shrinkage and Selection Operator (LASSO)
  - c. Elastic Net
  - d. Least Angle Regression (LARS)
5. Rule System (RS)
  - a. Cubist
  - b. One Rule (OneR)
  - c. Zero Rule (ZeroR)
  - d. Repeated Incremental Pruning to Produce Error Reduction (RIPPER)
6. Regression (REG)

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

- a. Linear Regression
  - b. Ordinary Least Squares Regression (OLSR)
  - c. Stepwise Regression
  - d. Multivariate Adaptive Regression Splines (MARS)
  - e. Locally Estimated Scatterplot Smoothing (LOESS)
  - f. Logistic Regression
7. Bayesian (BAY)
- a. Naïve Bayes
  - b. Averaged One-Dependence Estimators (AODE)
  - c. Bayesian Belief Network (BBN)
  - d. Gaussian Naïve Bayes
  - e. Multinomial Naïve Bayes
  - f. Bayesian Network (BN)
8. Decision Tree (DT)
- a. Classification and Regression Tree (CART)
  - b. Iterative Dichotomiser (ID3)
  - c. C4.5
  - d. C5.0
  - e. Chi-Square Automatic Interaction Detection (CHAID)
  - f. Decision Stump
  - g. Conditional Decision Trees
  - h. M5
9. Dimensionality Reduction (DR)
- a. Principal Component Analysis (PCA)
  - b. Partial Least Squares Regression (PLSR)
  - c. Sammon Mapping
  - d. Multidimensional Scaling (MDS)
  - e. Project Pursuit
  - f. Principal Component Regression (PCR)
  - g. Partial Least Squares Discriminant Analysis
  - h. Mixture Discriminant Analysis (MDA)
  - i. Quadratic Discriminant Analysis (QDA)
  - j. Regularized Discriminant Analysis (RDA)
  - k. Flexible Discriminate Analysis (FDA)
  - l. Linear Discriminant Analysis (LDA)
10. Instance-Based (IB)
- a. k-Nearest Neighbor (KNN)
  - b. Learning Vector Quantization (LVQ)
  - c. Self-Organizing Map (SOM)
  - d. Locally Weighted Learning (LWL)
11. Clustering (CLU)
- a. k-Means

**ARTIFICIAL INTELLIGENCE/MACHINE LEARNING  
LIFE INSURANCE SURVEY  
FILING GUIDANCE AND DATA DEFINITIONS**

- b. k-Medians
- c. Expectation Maximization
- d. Hierarchical Clustering