August 12, 2024

Via Electronic Submission

Mr. W. Moses Kim
Director
Office of Financial Institutions Policy
Department of the Treasury

Re: TREAS-DO-2024-0011-0001, Request for Information on Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector.

Dear Mr. Kim:

On behalf of SentiLink, I am pleased to submit the following comments in response to Treasury's Request for Information ("RFI") seeking input on "Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector."

SentiLink provides identity verification and fraud mitigation solutions to US-based financial institutions. Our tools enable institutions and individuals to transact confidently with one another by preventing identity fraud -- synthetic fraud and identity theft, as primary examples -- at the point that a consumer is applying for any type of financial account. SentiLink was also the first company to use the Social Security Administration's Electronic Consent Based SSN Verification service ("eCBSV") to validate account application data. Each day we evaluate over 1,000,000 consumer applications, helping financial institutions open accounts with confidence and consumers obtain the financial products and services they are entitled to.

Underneath the surface of our solutions are statistical machine learning models. To maintain best-in-class precision, these models rely on 1) a limited set of highly vetted external data sources, such as credit header data; 2) information from authoritative data sources, such as the SSA's eCBSV; 3) data provided by our clients (referred to as "partners") in the course of accessing our solutions; and most importantly 4) labels applied to individual cases by our team of expert fraud analysts. The lessons learned through this intensive investigation and labeling process, as assessed and incorporated into our models by SentiLink's staff, are a key factor making it possible for SentiLink to stay abreast of leading edge fraud tactics and criminal activity. Our

models are subject to extensive internal governance processes and data science-led development and review, in addition to rigorous third-party testing and evaluation.

It is from this perspective that we provide responses to select questions contained in the RFI.

Overview

SentiLink's solutions are used by financial institutions to address at least three of the functions cited in the RFI:

- *Risk management* through identity fraud detection and mitigation;
- *Product and service provisioning* through fast, friction-less identity verification;
- *Regulatory compliance* with "Know Your Customer" and "Customer Identification Program" rules.

We support Treasury's efforts to stay abreast of technological developments in AI in support of future policymaking but urge caution in defining “AI” too broadly. The term "artificial intelligence" is often used generically and imprecisely to encompass types of data processing that may seem to superficially fit into a category of AI, but in reality should not. For example, SentiLink and our models as they currently exist appear to fit into many academic and industry definitions of "AI" today, which we fundamentally believe is inaccurate. Beginning with a more precise definition of AI, as we will describe in response to Questions 1 and 2, would help avoid the risk of stifling innovation or over-burdening non-AI or AI-adjacent technologies. It is not simply that the regulation of AI should be risk-based (it should), but that some technologies conventionally considered to fall within a generic "AI" basket are not AI, as understood from a risks and outcomes perspective.

We encourage Treasury to develop a more tailored definition of AI that focuses on technologies that present the most risk and least predictability, and that excludes technologies that may be sophisticated and advanced but do not increase incremental risk for the financial institutions that use them. Furthermore, as frameworks for AI policymaking emerge, we urge Treasury to consider: how the risks introduced by AI will vary by the degree to which the AI is generative and/or predictive; its explainability relative to other data processing techniques; its ability to directly impact consumer outcomes; and the specific production use cases, as examples of important factors.

Consolidated Response to Questions 1 and 2:

Is the definition of AI used in this RFI appropriate for financial institutions? Should the definition be broader or narrower, given the uses of AI by financial institutions in different contexts?

What types of AI models and tools are financial institutions using? To what extent and how do financial institutions expect to use AI in the provision of products and services, risk management, capital markets, internal operations, customer services, regulatory compliance, and marketing?

Treasury's cited definition is overly broad for use in the financial services sector. While some types of machine learning models are integral to generative AI and predictive AI models -- which rightfully warrant scrutiny by Treasury and other regulators due to the increased risks they may pose -- not all machine learning-powered models raise the same risks or concerns. An overly broad definition will hinder the ability of regulators to efficiently examine core AI use cases.

Using SentiLink as an example: Our models primarily use a form of "supervised machine learning"[1] (regression algorithms and ensemble trees) to make probabilistic conclusions about fraud. Our models generate a risk score output indicating a confidence level between 1-999 whether an applicant for a financial product is who they claim to be, or is a case of identity fraud.

Contrasting the specific language of the definition of "artificial intelligence" adopted by the RFI with SentiLink's approach to machine learning helps illustrate important distinctions and the need for more precise definitional language. Here, italicized words are drawn from the codified language:[2]

- Unlike a credit score, the risk score outputs of our machine learning models do not *predict* or *perceive* the likelihood of some future condition: they use analysis of historical data to make conclusions about a given set of PII at a specific moment in time.
- In addition, our models derive their outputs based on raw observed and recorded data, not *perceptions* that become layered onto or *abstracted from* that raw data. This approach ensures our models remain highly explainable

[1] See: Hastie, Trevor, et al. "The Elements of Statistical Learning." 2nd ed., Springer, 2009.
[2] For reference: *The term "artificial intelligence" or "AI" has the meaning set forth in 15 U.S.C. 9401(3): a machine-based system that can, for a given set of human-defined objectives, make* ***predictions****,* ***recommendations****, or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human–based inputs to* ***perceive*** *real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to* ***formulate options*** *for information or action.* (emphasis added)

and predictable because the analytic functions are performed directly on relevant observed and documented data.

- Our models do not *recommend* or *formulate options* for any course of action for a financial institution to take. A SentiLink risk score is one of often many inputs that a financial institution considers when evaluating an application. Our models do not attempt to replicate or substitute for any *decision*-oriented human judgment. Instead, our models provide valuable analytic data points that are utilized in a financial institution's subsequent formulation of fraud prevention options, decisions, and judgments.

Outputs of AI models that "predict," "perceive," "decision" or any of the other emphasized words in the bullet points above are foundational hallmarks of what distinguishes AI from the broader category of models and algorithmic tools designed and managed by humans to perform analytic tasks. SentiLink's fraud detection models do not bear these hallmarks -- since they are constrained by human-applied labels, with no ability to learn or make decisions independently -- and in their current form should not be regarded as "AI."

Taking this a step further, see Treasury's analysis of the adopted definition from the RFI:

> Treasury interprets this definition to describe a wide range of models and tools that utilize data, patterns, and other informational inputs to generate outputs – including statistical relationships, forecasts, content, and recommendations – for a given set of objectives. For the purposes of this RFI, Treasury is seeking comment on the latest developments in AI technologies and applications, including but not limited to advancements in existing AI (e.g., machine learning models that learn from data and automatically adapt and improve with minimal human interference, rather than relying on explicit programming) and emerging AI technologies including deep learning neutral network such as generative AI and large language models (LLMs).

Regarding the green highlighted text, we agree that the codified definition encompasses a wide range of models and tools, which is also why we believe it must be narrowed for financial regulatory purposes. Using data and other inputs to generate outputs has been done since the early days of PCs running DOS, and also describes a core function of an Excel spreadsheet -- neither of which any reasonable person would conflate with modern "artificial intelligence."

The yellow highlighted text does not describe SentiLink's machine learning models. However, it does accurately describe different types of machine learning models used for generative and predictive AI use cases. We believe this should be the guiding perspective for how to narrow the definition of AI for applications in the financial sector.

In sum, application of an overly broad definition of "artificial intelligence" for financial regulatory purposes will:

1. Improperly classify specific types of technology as "AI;"
2. Add costs onto technologies that are not introducing incremental risk and inhibit their effectiveness at advancing important economic and consumer protection outcomes; and
3. Frustrate Treasury's goal of supporting responsible innovation and competition in the financial sector.

We believe tightening the codified definition can avoid these consequences by delineating more clearly that "AI" in the financial sector should encompass predictive and decisioning technologies that substitute for human judgment and cannot be fully explained by model development principles. Specifically, we recommend the following changes be made to a definition of "artificial intelligence" as it applies to the financial sector:

> [A] machine-based system that can, for a given set of human-defined objectives, make predictions~~,~~ ~~recommendations,~~ or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human–based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to generate predictive or decisioning outputs ~~formulate options for information or action~~.

As Treasury continues its monitoring of developments in this space, we again emphasize the importance of certain risk-based aspects in the evaluation of different data processing technologies, such as the degree to which the AI is generative and/or predictive; its explainability relative to other data processing techniques; its ability to directly impact consumer outcomes; and the specific production use cases, as examples of important factors.

Note: For the following questions we are offering perspective with the above response in mind, i.e. as a non-AI company today that builds and maintains sophisticated and human supervised machine learning-based models.

Question 5: What are the actual and expected benefits from the use of AI to any of the following stakeholders: financial institutions, financial regulators, consumers, researchers, advocacy groups, or others? Please describe specific benefits with supporting data and examples. How has the use of AI provided specific benefits to low-to-moderate income consumers and/or underserved individuals and communities (e.g., communities of color, women, rural, tribal, or disadvantaged communities)?

SentiLink's models review more than 1,000,000 financial applications across our entire partner base each day. Of those, more than 20,000 are high-risk attempts to leverage the stolen identities of consumers. Among those cases, on average 24% are victims over the age of 60, and 16% are thin-file consumers. Based in part on the fraud risk signals offered by our models, financial institutions are able to take appropriate steps to prevent further harm to these victims.

Of the 1,000,000 cases we review each day, approximately 13,000 are high-risk attempts to leverage synthetic identities. While a smaller portion of overall identity fraud attempts, synthetic fraud drives outsized losses for financial institutions: synthetic tradelines for credit cards are 16.5X more likely to default within the first year compared to typical consumers, and 13.7X more likely to default for auto loans. Given the nature of this type of identity crime, it is impossible for a financial institution to recoup any of these losses. Financial institutions that use SentiLink's models targeting synthetic identities -- along with our pioneering use of the eCBSV -- often see reductions in fraud losses totalling millions of dollars each year.

Overall, use of SentiLink's solutions help millions of people gain access to the financial products and services they are entitled to, without adding friction to their application experience. Further, our ability to validate the identities of thin-file consumers and match information with authoritative data sources such as the eCBSV often helps consumers in vulnerable populations that would otherwise be considered high-risk.

Question 10: How are financial institutions addressing any increase in fair lending and other consumer-related risks, including identifying and addressing possible discrimination, related to the use of AI, particularly emerging AI technologies? What governance approaches throughout the development, validation, implementation, and deployment phases do financial institutions expect to establish to ensure compliance with fair lending and other

consumer-related laws for AI models and tools prior to deployment and application?

Protecting against model bias and disparate impact are table stakes for SentiLink and our partners. We invest heavily in people, process, and technology and rely on reputable data sources, proven human-led model development techniques, and cross-functional reviews to prevent disparate outcomes. While we do not collect information about race, ethnicity, or gender, SentiLink relies on rigorous third-party audits of our models to make sure they are accurate and fair for all consumers.

SentiLink maintains a robust internal governance process. This process establishes a framework by which our data scientists evaluate current and prospective model features and data inputs. This is critical as our partners -- many of which are heavily regulated depository institutions -- subject our models to their own internal governance standards as well as third-party vendor management and cybersecurity risk regulatory obligations.[3] Our partners will routinely request a feature-by-feature justification and review -- a process only made possible by the explainability of our models. Given the non-AI nature of our current models, existing regulatory guidance is more than sufficient to ensure that the services SentiLink provides support a safe and sound banking system. That said, regulatory alignment for use in the financial sector, as we've suggested above, would significantly help smooth out inconsistencies that currently exist on an institution-by-institution basis.

Question 18: What actions are necessary to promote responsible innovation and competition with respect to the use of AI in financial services? What actions do you recommend Treasury take, and what actions do you recommend others take? What, if any, further actions are needed to protect impacted entities, including consumers, from potential risks and harms?

As we have stated, we believe SentiLink's existing human-supervised machine learning models should not be considered "AI." However, there may be other identity verification service providers that do use AI in their customer-facing deployments, such as for document verification and some biometric-based solutions. The viability of both AI and non-AI technologies used in fraud mitigation and identity verification may be impaired by an anticipated Consumer Financial Protection Bureau ("CFPB") rulemaking.

---

[3] *See, for example*: "Interagency Guidelines Establishing Information Security Standards." Board of Governors of the Federal Reserve System, accessed at: https://www.federalreserve.gov/supervisionreg/interagencyguidelines.htm and "Interagency Guidance on Third-Party Relationships: Risk Management." Federal Deposit Insurance Corporation, accessed at: https://www.govinfo.gov/content/pkg/FR-2023-06-09/pdf/2023-12340.pdf

The CFPB is expected to soon propose a rule to expand the scope of the Fair Credit Reporting Act ("FCRA") to encompass CIP/KYC and fraud prevention activities, as well as the data needed to perform these tasks and fulfill regulatory obligations important to Treasury and FinCEN.[4] In particular, it is possible the CFPB's proposal will classify credit header data as a "consumer report;" classify firms that sell datasets used for identity verification and fraud reduction as "data brokers" that are "consumer reporting agencies" under the FCRA; and classify firms that use credit header and other data acquired from data brokers for fraud prevention and identity verification as "consumer reporting agencies."

Fundamentally, we believe the FCRA was never intended to govern financial crime prevention and identity verification activities in the financial industry. The FCRA implicitly assumes that the consumers addressed under that Act are the consumers they purport to be and is intended to provide consumers and businesses with the information they need to transact honestly with each other (for example, by determining eligibility for credit). To achieve the FCRA's purposes, there is a necessary bright line regulatory distinction between 1) ex-ante identity verification and identity fraud reduction efforts (subject to the GLBA and all of the BSA-derived CIP, KYC and AML authorities under the jurisdiction of FinCEN) and 2) the ex-post regulation of credit reporting and evaluation of consumer credit-worthiness (under the FCRA) for a consumer whose identity has been verified.

Subjecting fraud mitigation and identity verification providers -- and the data they rely upon -- to the FCRA would compromise the delivery of the services that make our models an important component of CIP and other BSA/AML processes. If the CFPB's rulemaking ultimately collapses identity verification into the FCRA framework, companies like SentiLink will need to reconstitute the analytic elements of their fraud prevention tools, and that effort would be hindered by FCRA requirements. For example, fraudsters could exploit FCRA consumer rights to undermine fraud detection and reveal fraud prevention details to then hone strategies for slipping through identity verification defenses. Further, limiting the use of credit header data -- which is often the foundation for identity verification and fraud mitigation efforts employed by banks -- would severely compromise the ability of banks to see patterns in identity use and misuse.

The likely effect of such a change would be to inhibit consumer access to financial services and increase costs as well as safety and soundness risks to regulated

[4] *See*: "Small Business Advisory Review Panel for Consumer Reporting Rulemaking: Outline of Proposals and Alternatives Under Consideration," September 15, 2023. Accessed at https://files.consumerfinance.gov/f/documents/cfpb_consumer-reporting-rule-sbrefa_outline-of-proposals.pdf

institutions. In short, the proposal would obstruct our ability to provide a vital fraud prevention and CIP function that protects consumers and enables our financial institution partners to focus their subsequent credit-worthiness assessments on genuine applicants. We urge Treasury to engage with the CFPB to understand the implications of this rulemaking and prevent a regulatory outcome that directly conflicts with the authorities granted by Congress to FinCEN and the prudential regulators.

Thank you for this opportunity to comment and for your consideration of our views.

Sincerely,

/s/

Jason Kratovil
Head of Policy