AI Evaluation Methodology Lead

Do you want to define the evaluation methodology and technical requirements that underpin effective AI evaluation and monitoring within the business at one of the leading financial institutions in the Nordics?
Join Danske Bank's newly established AI Model Assessment team in Business Controls & Product Management. As AI Evaluation Methodology & Platform Requirements Lead, you will hold a 1st Line of Defence role with clear ownership of the technical methodologies used to evaluate AI models and AI-enabled solutions, and of the technical requirements for the platforms that support their evaluation and ongoing monitoring. You will translate business-owned risk and control needs into clear methods, metrics, data, evidence, traceability and platform capabilities that can be implemented consistently across the Bank's technology landscape.
This is a unique opportunity to shape a consistent, scalable and evidence-based capability through which the business identifies, evaluates, monitors and manages AI model risk.

This role can be based in Copenhagen, Denmark, Vilnius, Lithuania, Warsaw, Poland, Stockholm, Sweden or Helsinki, Finland.

What you will be doing
The AI Model Assessment team is a business-owned control capability that performs evidence-based assessments of AI models and AI-enabled solutions across the Bank. We work with business owners, developers, engineers and platform teams to understand how solutions operate, define the evidence required for assessment, and establish the evaluation and monitoring capabilities needed to identify, assess and manage model risk where it arises.
Our mission is to enable responsible innovation by providing the business with a consistent and scalable framework for evaluating AI systems, monitoring their performance and risk, and demonstrating that appropriate controls are operating effectively.

As AI Evaluation Methodology & Platform Requirements Lead, you will own the evaluation methodology and define the technical requirements for the platforms, tooling and data capabilities used by the 1st Line of Defence to evaluate and monitor AI models and AI-enabled solutions. Your mandate is to ensure that the business has robust, implementable and auditable means to assess its AI risks and controls. Your focus will be on specifying what the evaluation and monitoring capability must deliver rather than owning software development or platform operations.

You will work closely with the independent 2nd Line AI model risk function, which sets the overarching risk framework and provides oversight and constructive challenge. You will translate that framework into practical evaluation methodologies, technical requirements and business-owned controls, provide transparent evidence of how requirements are implemented, and support the resolution of identified gaps.

Your responsibilities will include:
• Owning, developing and maintaining the methodology for evaluating AI models and AI-enabled solutions, including evaluation objectives, test approaches, acceptance criteria and evidence standards.
• Defining the functional and technical requirements for evaluation and monitoring platforms that enable risk management, including data ingestion, test execution, metric calculation, observability, evidence capture, reporting, access controls and auditability.
• Translating model risk, business and assessment requirements into reusable evaluation methods, representative datasets, metrics, thresholds, monitoring indicators and structured evidence requirements.
• Determining how AI system architectures, components, dependencies and data flows affect evaluation scope, test design, monitoring coverage and the technical capabilities required from supporting platforms.
• Defining platform-agnostic requirements that support evaluation and monitoring across environments such as AWS Bedrock and AgentCore, Databricks, Azure OpenAI and embedded third-party AI platforms.
• Setting requirements for reproducibility and end-to-end traceability across models, prompts, datasets, configurations, system components, evaluation results, monitoring signals and changes over time.
• Prioritising opportunities to standardise and automate evaluation and monitoring activities, and working with evaluators, platform owners and engineers to convert approved control requirements into practical platform capabilities.
• Engaging with the Second Line model risk function to interpret applicable requirements, demonstrate how they are operationalised, respond to constructive challenge, and ensure that evaluation and monitoring evidence supports effective independent oversight.

About you
An experienced AI evaluation, model risk, data, architecture or platform professional who can establish rigorous evaluation methodologies and turn business risk and control needs into clear, implementable technical requirements for enterprise platforms and tooling.

You have:
• Proven experience designing or governing AI/ML evaluation methodologies, including test objectives, datasets, metrics, thresholds, acceptance criteria and evidence standards.
• Experience translating evaluation, monitoring, risk and control needs into functional and technical requirements for platforms, tooling, data pipelines and reporting capabilities.
• Experience with AI/ML evaluation, business-owned model risk management and controls, technical assurance, MLOps, LLMOps, AI platforms or AI observability.
• Familiarity with cloud AI services or enterprise AI platforms; experience with AWS Bedrock or AgentCore, Databricks or Mosaic AI, Azure OpenAI, or comparable platforms is an advantage.

• Strong understanding of evaluation and monitoring principles, including representative datasets, experiment configuration, metric design, regression testing, monitoring indicators, traceability and evidence capture.
• Ability to distinguish methodology and requirements ownership from platform implementation, while collaborating effectively with engineers and platform owners. Python or similar tools for analysis, prototyping or investigation are an advantage, but software development is not the primary focus of the role.
• Excellent collaboration and communication skills, with the ability to articulate evaluation, monitoring, risk and control requirements across business, assessment, platform, engineering, architecture, data and risk functions.
Experience with AI evaluation or observability technologies, for example Phoenix or comparable tools, is an advantage but is not a prerequisite.

What we offer you
You will join the newly established AI Model Assessment team in Business Controls & Product Management, helping build the Bank's consistent and scalable business-owned capability for AI evaluation, monitoring and model risk management.
The team combines technical AI evaluation expertise with model risk management and works across the Bank's AI initiatives. In this role, you will connect evaluators with platform, engineering, architecture and data teams by defining the methods and technical requirements their solutions must support, ensuring that risk ownership and control evidence remain anchored in the business.

If hired in Lithuania, your monthly salary would range from **4320 EUR** to **6480 EUR** gross (based on your competencies relevant for the job).
Additionally, each Danske Bank employee receives employee benefits package which includes:
• **Growth opportunities:** professional & supportive team, e-learnings, numerous development programs; (incl. professional certificates); 100+ professions for internal mobility opportunities.
• **Health & Well-being:** a diverse, inclusive, work & life balance work environment; additional health insurance; mental well-being practices; partial psychologist counselling compensation; silence and sleep zones at the office; game rooms.
• **Hybrid working conditions:** Work from home up to two days a week; home office budget (after the probation period); modern Danske Campus workplace developed with anthropologist for the best employee experience.
• **Additional days of leave:** for rest, health, volunteering, exams in higher education institutions, and other important activities. Moreover, for seniority with Danske Bank.
• **Monetary compensation package:** accidents & critical diseases insurance; financial support in case of unfortunate events, travel insurance; IIIrd Pillar Pension Fund contribution.
See all the benefits HERE.

Your job function (position) in the job contract will be AI Science

Are you ready for the next step?
If you are passionate about defining robust AI evaluation methodologies and the technical requirements for scalable evaluation and monitoring platforms that support effective business ownership of AI risk, we would love to hear from you.
For questions about the role, please contact **Khaled Alizai**, Head of AI Model Assessment at khali@danskebank.dk. We are screening the applications on an ongoing basis, so we encourage you to apply.
Join us and help build the business-owned foundation for effective AI model evaluation, monitoring and risk management at Danske Bank.

Danske Bank København K, 1300 København K danskebank.dk	**Ansøgningsfrist** Løbende jobsamtaler AI Evaluation Methodology Lead	**Kontaktperson** kontakt@finanskandidaten.dk