**Transcript for Research Proposal Presentation**

**Slide 1: Title Slide**

Hello, my name is Mauricio Lozano, and this presentation outlines my proposed research on automating network security compliance. Specifically, I aim to use a declarative modeling framework in combination with Large Language Models (LLMs) to improve how Risk Management Framework (RMF) documentation is generated, maintained, and evaluated within Department of Defense (DoD) network environments. This work builds on existing modernization initiatives like SWIFT while addressing the unique challenges associated with network architecture documentation, which remains heavily manual and interpretation-driven.

**Slide 2: Introduction & Context**

Let’s begin with some context. The Department of Defense relies heavily on the Risk Management Framework, or RMF, as a structured process to manage cybersecurity risks across its vast and complex network infrastructure. The RMF is composed of six steps and is grounded in guidance such as NIST SP 800-37 and SP 800-53. It mandates the documentation and review of more than 1,200 individual security and privacy controls for each system, tailored to that system's operational context. While this framework offers a standardized, policy-driven approach to achieving cybersecurity assurance, it also suffers from well-documented shortcomings, especially when it comes to efficiency.

A recurring issue voiced by compliance professionals, system owners, and authorizing officials is that the documentation burden placed by RMF is excessive. Generating the System Security Plan (SSP), the Plan of Actions and Milestones (POA&M), and supporting risk acceptance narratives requires hundreds of labor hours. This is particularly true for medium- to high-impact systems. As noted by researchers like Scanlon and auditors from the Government Accountability Office (GAO), most of this effort doesn't go into implementing technical controls or collecting evidence. Instead, it is consumed by interpreting raw artifacts and translating them into context-aware, narrative documentation that is defensible during an ATO—or Authority to Operate—review.

**Slide 3: Research Problem & Significance**

One GAO study examined 19 weapon-system authorizations and found that in several cases, risk decisions were based on incomplete or outdated documentation. The review warned that checklist-driven approvals often missed context-specific risks, such as electromagnetic emission hazards aboard Navy vessels. The takeaway from that finding is not just that more review is needed, but that better, more adaptive documentation methods could reduce human error while improving assurance.

To modernize the RMF process, the DoD has recently piloted initiatives such as the Secure Workflow for Institutional FastTracking, or SWIFT. Led by Acting CIO Katie Arrington, SWIFT represents a pivot away from one-time ATO decisions toward continuous authorization, especially for software systems. It leverages Software Bills of Materials—SBOMs—and feeds them into automated AI pipelines to assess compliance. This has allowed software teams to automate much of the authorization process and move toward real-time security assurance. However, SWIFT’s current scope is limited to software supply chains. Network security, which has more complex topologies, mutable trust boundaries, and hybrid deployment models, has largely been left behind in terms of automation.

**Slide 4: Research Question**

That’s where this research comes in. My central question is this: How can declarative network architecture modeling—combined with LLMs—be used to automate RMF documentation in a way that maintains trustworthiness and contextual accuracy? Can we close the gap between structured automation formats like OSCAL or Infrastructure as Code, and the rich, interpretive documentation needed for risk acceptance? This question sits at the intersection of cybersecurity, AI ethics, and public sector modernization and has implications for how future ATOs might be conducted across the federal enterprise.

**Slide 5: Aims & Objectives**

To answer this question, my research is structured around four key objectives. The first is to prototype a declarative, machine-readable language and system for defining, validating, and enforcing network architecture and policy. This framework will allow security architects to describe devices, roles, zones, and policies in a structured YAML-based syntax that supports inheritance, tagging, and policy logic. The language will act as a source of truth that can be validated, simulated, and transformed.

Second, I plan to use this structured model as input for Large Language Models. The goal here is to develop prompt templates that translate model data into first-pass SSP narratives. This would include descriptions of the network environment, inheritance of controls, trust boundary justifications, and residual risk statements. The LLM will not replace the human authorizer but will instead accelerate the drafting process, freeing human reviewers to focus on higher-risk areas.

Third, I will benchmark this AI-assisted documentation workflow against the traditional manual process. Metrics will include time to completion, edit count, word count, and coverage of applicable controls. I will use multiple synthetic network environments with varying complexity levels to ensure generalizability. The aim is to establish whether this approach leads to measurable gains in efficiency without compromising the completeness or quality of documentation.

Finally, I will address the ethical and governance implications of using AI to support RMF documentation. Issues such as hallucinated outputs, lack of explainability, bias in training data, and over-reliance on automation must be studied closely. I will ensure that all generated outputs are subject to human-in-the-loop review and incorporate mechanisms for provenance tracking and citation enforcement.

**Slide 6: Related Work – RMF & Compliance Burden**

To ground this proposal further in existing literature: NIST SP 800-37 outlines RMF steps with the expectation that humans will author all documentation. SP 800-53 includes over 1,200 security controls but offers no implementation guidance on how to scale the narrative burden. OSCAL introduces a machine-readable schema for security documentation but does not produce contextual text or risk interpretation. Similarly, Infrastructure as Code platforms like Terraform and Ansible are excellent at automating infrastructure provisioning but cannot represent architectural intent or compliance logic in a way that maps directly to RMF.

**Slide 7: Related Work – Automation Tools**

I’m assessing hree major automation tools or frameworks that inform the foundation for this research, while also exposing their limitations in addressing interpretive RMF documentation.

**OSCAL** – The Open Security Controls Assessment Language, developed by NIST, is a machine-readable format that aims to streamline RMF documentation. It standardizes the way security controls, system implementation, and assessment results are expressed. However, OSCAL does not address how those controls apply in specific operational contexts. It provides structure, but not interpretation. Human analysts are still required to produce contextual narratives and risk justifications.

**Infrastructure as Code (IaC)** – Tools like Terraform and Ansible allow for declarative provisioning of infrastructure in a repeatable and scalable manner. These tools are crucial for DevSecOps workflows, ensuring infrastructure consistency and compliance alignment. Yet, they are not compliance tools themselves. They lack native support for mapping infrastructure configurations to RMF controls or for generating SSP narratives. IaC describes what is deployed, but not why it meets a security requirement.

**SWIFT** – The Secure Workflow for Institutional FastTracking initiative by the DoD focuses on continuous ATO for software systems. It uses SBOMs and automated analytics to assess software compliance in near-real-time. While SWIFT represents a significant leap forward for software pipelines, its scope is limited to applications and code. It does not extend to network infrastructure, where trust zones, boundary controls, and hardware interactions require different forms of evaluation.

Together, these tools show progress toward RMF modernization, but none fully address the interpretive, policy-aware needs of network RMF documentation. This research seeks to fill that gap by combining structured architecture modeling with AI-generated narrative support.

**Slide 8: Related Work – AI and LLMs**

This slide focuses on the intersection of AI, ethics, and the emerging role of Large Language Models in documentation generation.

**AI Ethics** – Corrêa et al. (2024) conducted a review of over 200 global AI ethics frameworks. They identified transparency, accountability, fairness, and bias as recurring themes. These principles are highly relevant in defense contexts, where automation must meet strict assurance requirements. Yet, Corrêa’s review also reveals a major shortfall: fewer than 10% of the reviewed policies specify practical mechanisms for auditing model behavior or tracing data provenance—capabilities that are critical for RMF applications.

**LLMs and Narrative Generation** – Recent advances in natural language processing have enabled LLMs like GPT-4 to generate context-aware, coherent technical narratives. When properly prompted, these models can synthesize structured inputs—such as network topology models—into textual explanations, justifications, or summaries that align with compliance documentation needs. This unlocks the possibility of using LLMs to automate first-draft RMF narratives, particularly the SSP sections that describe system context, control applicability, and residual risks.

**Opportunity and Gap** – Despite these capabilities, no prior studies have combined LLMs with network architecture models specifically for RMF documentation. While examples exist in software code generation and document summarization, there is a gap in research exploring how LLMs can interpret network-focused structured data to produce tailored SSP content. This project aims to fill that gap by integrating LLMs with a machine-readable domain-specific language, governed by ethical controls to ensure transparency and traceability.

**Slide 9: Proposed Artefact**

The artefact proposed in this research is a Structured Network Modeling Framework designed to bridge the gap between architectural design and compliance documentation. At its core is YAML-based Domain-Specific Language, which allows engineers and security architects to declaratively model the topology, device roles, trust boundaries, and communication rules within a network. This language forms the basis for the rest of the automation pipeline.

Supporting this language is a validation engine that interprets the DSL files, applies inheritance logic and tagging semantics, and verifies whether defined rules align with Zero Trust principles and NIST control mappings. The engine ensures internal-to-external traffic paths are flagged, trust zones are respected, and architectural intent is preserved.

Building on this, the framework integrates with a Large Language Model via a structured prompt layer. This layer translates the validated architecture into tailored narrative outputs resembling the content typically required in System Security Plans. The generated text covers elements such as boundary definitions, control applicability justifications, and residual risk narratives—core components of RMF documentation.

Overall, this artefact operationalizes the idea that security architecture and compliance need not be siloed processes. By merging modeling and narrative generation into a single system, the research proposes a scalable and auditable way to modernize RMF workflows in networked environments

**Slide 10: Methodology**

Methodologically, my approach is mixed. The quantitative phase will involve constructing SSP documentation for the same test environments using both traditional and AI-assisted workflows. I will log time to completion, word count, reviewer edits, and control coverage. I also plan to assess how many iterations were required to reach a reviewable draft. Additional testing may involve LLM fine-tuning experiments to compare general-purpose models like GPT-4 with domain-adapted variants.

The qualitative phase will include semi-structured interviews with Information System Security Managers and compliance professionals. I’ll ask them to evaluate the clarity, trustworthiness, and usefulness of LLM-generated text, and whether they would accept or reject such content in a real-world ATO context. These interviews will help assess human perceptions of AI reliability and will be thematically coded to extract actionable insights. Interviews will also explore whether reviewers believe automation affects risk posture.

For tooling, the research will leverage Git for version-controlled storage of DSL files, Open Policy Agent (OPA) for potential policy testing, GPT-4 or similar models for narrative generation, and a Python-based validation engine for interpreting and simulating the DSL content. Visual modeling and export features may be implemented through open-source graph libraries or custom plugins for VS Code. The test environments will be created in isolated, synthetic conditions to simulate realistic enclave configurations.

**Slide 11: Evaluation Approach**

The evaluation approach will compare traditional RMF documentation practices with the AI-assisted method developed in this research. Two parallel documentation workflows will be executed: one using conventional manual SSP drafting and one leveraging the DSL framework and LLM-generated narrative output. Each will address the same synthetic network environments.

Quantitative metrics to be collected include time-to-completion, total word count, number of reviewer edits, and percentage of applicable controls covered. Qualitative analysis will stem from structured reviewer feedback and observations of usability or interpretability issues. The study aims to determine whether the AI-assisted workflow materially reduces effort while maintaining compliance accuracy and reviewer trust.

Multiple test runs with varying levels of network complexity will be conducted to ensure generalizability. Reviewer feedback will be categorized using a combination of thematic coding and Likert-scale responses to evaluate perceptions of reliability, transparency, and overall documentation quality.

**Slide 12: Ethical Considerations & Risks**  
On the ethical front, I will ensure full compliance with university research ethics policies and applicable legal requirements. No classified or sensitive data will be used—only synthetic test environments created explicitly for research purposes. All AI-generated content will undergo human review, and citations or sources will be explicitly tracked to mitigate the risk of hallucinated or unverifiable statements.

Additional safeguards include the use of adversarial testing to assess how the LLM responds to malformed prompts or biased input data. Bias audits will be conducted by varying prompt structures and assessing consistency across outputs. Where possible, models will be configured to produce outputs with confidence indicators or justification references.

Ethical oversight will include pre-study approval by an institutional review board (IRB), documentation of all prompt-response pairs, and traceability mechanisms embedded in the pipeline. These measures ensure transparency, accountability, and the safe integration of LLMs into a compliance-driven domain like RMF.

**Slide 13: Timeline – 4-Month Plan**  
The timeline for this research has been carefully planned to span a four-month period, allowing for focused development, testing, and analysis. Each phase is time-boxed to ensure meaningful progress while maintaining quality and rigor.

* **Month 1**: Finalize the YAML-based DSL schema and create two to three representative synthetic network environments. These environments will be modeled to capture a range of common architectural patterns and trust boundary configurations.
* **Month 2**: Develop the DSL validation engine, implement inheritance and tagging logic, and begin constructing the LLM prompt layer. This month will also include preliminary output testing for accuracy and control mapping.
* **Month 3**: Run side-by-side comparisons between traditional and AI-assisted SSP workflows. Begin conducting qualitative interviews with reviewers. Capture time metrics and feedback.
* **Month 4**: Analyze quantitative results, synthesize qualitative findings, and prepare final deliverables including the written report, supporting documentation, and the narrated presentation.

**Slide 14: Expected Contributions**  
Expected contributions of this research are fourfold and span both technical and governance dimensions:

1. **A machine-readable language model and system** for defining, validating, and enforcing network architecture and policy. This artefact can reduce ambiguity and support compliance automation at scale.
2. **Empirical benchmarking data** comparing traditional RMF documentation workflows to AI-assisted alternatives. This includes time savings, control coverage, and reviewer trust metrics.
3. **An integration strategy for LLMs in compliance-driven domains**, including best practices for prompt design, error handling, and output validation—all critical for DoD and other high-assurance environments.
4. **A practical framework for ethical AI use in cybersecurity workflows**, offering concrete recommendations for citation enforcement, bias mitigation, and human-in-the-loop oversight.

**Slide 15: Conclusion**  
To conclude this presentation, I would like to reiterate the transformative potential of this research. By leveraging a machine-readable architecture modeling framework in tandem with AI-assisted narrative generation, this work seeks to reduce the compliance burden that has long impeded agile and secure system deployments within the Department of Defense. The research directly responds to long-standing pain points identified in both practitioner commentary and official audits—chief among them the inefficiencies of narrative-heavy RMF processes.

More than just a technological proof of concept, this project aims to provide a governance-aware model that preserves the trust relationships inherent to the ATO process. It acknowledges the human element—not as something to replace, but to augment. The goal is not to automate judgment, but to automate preparation, enabling security professionals to focus their efforts where human discernment is most needed.

Ultimately, this proposal lays the groundwork for future exploration into how declarative modeling, AI, and secure automation can converge to support mission readiness, risk transparency, and operational resilience across government networks.

Thank you for your attention and have a great day.