

API/MODELS/FRAMEWORK Security WG/RG/CG 1st Meeting

By: Ronaldson Bellande
PhD Student
Founder/CEO/CTO/COO Bellande
Technologies Corporation Inc
Founder of Bellande Research
Organizations

Meeting Agenda

- Integrity Protection for Data
 - Integrity Protection for Model
 - Integrity Protection for Inference
 - Confidentiality for Data
 - Confidentiality for Model
 - Availability
 - Framework Security
 - Secure Coding Practices
 - Dependency Management
 - Configuration Management
 - Testing and Validation
 - Adversarial Attack Defense
 - Privacy-Preserving Machine Learning
 - Model Watermarking and Fingerprinting
 - Secure Model Deployment
 - API Security
 - Authentication and Authorization
 - Data Encryption
 - Rate Limiting and Throttling
 - Logging and Monitoring
 - Input Validation
-
- Importance of Embedded Security
 - Challenges in ML/AI Security
 - Trends in ML/AI Security
 - Benefits of Secure ML/AI
 - Implementing ML/AI Security

Integrity Protection for Data

- Data integrity ensures training and input data hasn't been tampered with.
- It involves implementing robust data validation and verification processes.



Integrity Protection for Model

- Model integrity prevents unauthorized modifications to the AI/ML model.
- It requires implementing secure model versioning and change detection systems.



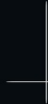
Integrity Protection for Inference

- Inference integrity ensures model outputs are accurate and unmanipulated.
- It involves implementing output validation and anomaly detection mechanisms.



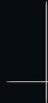
Confidentiality for Data

- Data confidentiality protects sensitive information used in training and inference.
- It requires robust encryption and access control measures.



Confidentiality for Model

- Model confidentiality secures proprietary models from theft or reverse engineering.
- It involves implementing model obfuscation and secure storage techniques.



Availability

- Availability ensures ML/AI services are resilient against Denial of Service attacks.
- It requires implementing load balancing and redundancy strategies.



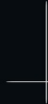
Framework Security

- Framework security incorporates security measures into ML/AI development tools.
- It covers areas like secure coding, dependency management, and configuration.



Secure Coding Practices

- Secure coding practices help avoid common vulnerabilities in ML/AI systems.
- They include preventing injection attacks, buffer overflows, and insecure deserialization.



Dependency Management

- Dependency management involves regularly updating third-party libraries.
- It helps mitigate risks from known vulnerabilities in external components.



Configuration Management

- Secure configuration management ensures safe default settings in ML/AI systems.
- It makes implementing security configurations easier for developers.



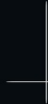
Testing and Validation

- Thorough security testing identifies potential vulnerabilities in ML/AI systems.
- It includes techniques like fuzz testing and adversarial testing.



Adversarial Attack Defense

- Adversarial attack defense protects against manipulated inputs designed to fool models.
- It involves developing models resilient to adversarial examples.



Privacy-Preserving Machine Learning

- Privacy-preserving ML protects user data during training and inference.
- It includes techniques like federated learning and differential privacy.



Model Watermarking and Fingerprinting

- Model watermarking embeds unique signatures within ML/AI models.
- It helps identify ownership and detect unauthorized use of models.



Secure Model Deployment

- Secure deployment protects ML/AI models in production environments.
- It utilizes technologies like containerization and trusted execution environments.



API Security

- APIs play a crucial role in integrating and securing ML/AI models.
- They require robust security measures to protect against various threats.



Authentication and Authorization

- Strong authentication ensures only authorized users can access ML/AI services.
- Role-based access control limits user permissions based on their role.



Data Encryption

- Encryption protects data transmitted to and from ML/AI APIs.
- It involves using protocols like HTTPS/TLS and encrypting data at rest.



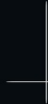
Rate Limiting and Throttling

- Rate limiting prevents abuse of ML/AI APIs.
- It helps protect against Denial of Service attacks.



Logging and Monitoring

- Detailed logging of API activities helps detect suspicious behavior.
- Real-time monitoring enables quick response to potential security incidents.



Input Validation

- Input validation prevents injection attacks in ML/AI APIs.
- It ensures the integrity of data processed by the model.



Importance of Embedded Security

- Embedded security in ML/AI systems provides proactive protection.
- It's more effective than adding security measures after development.



Challenges in ML/AI Security

- The ML/AI threat landscape is constantly evolving.
- Balancing security with performance and usability is an ongoing challenge.



Future Trends in ML/AI Security

- Emerging technologies will reshape ML/AI security practices.
- AI itself will play a growing role in enhancing security measures.



Benefits of Secure ML/AI

- Secure ML/AI systems protect sensitive data and intellectual property.
- They help maintain user trust and comply with regulatory requirements.



Implementing ML/AI Security

- Implementing ML/AI security requires a holistic, multi-layered approach.
- It involves collaboration between data scientists, developers, and security experts.



Collaboration Opportunities & Next Steps & Networking & Resources

- GitHub Working Group Repository Information: <https://github.com/Artificial-Intelligence-Computer-Vision/BAI-CVRI-Machine-Learning-Artificial-Intelligence-Models-Framework-Security-Community-Group>
- Presentation-Notes: <https://github.com/Artificial-Intelligence-Computer-Vision/BAI-CVRI-Machine-Learning-Artificial-Intelligence-Models-Framework-Security-Powerpoint-Notes>
- GitHub Organization: <https://github.com/Artificial-Intelligence-Computer-Vision>
- Website: <https://artificial-intelligence-computer-vision.github.io>
- Discord Group: <https://discord.gg/KBwqcPdx2H>
- ~~Github Profile: <https://github.com/RonaldsonBellande>~~