Pryon's Guide to Responsible AI and the RAI SHIELD Assessment

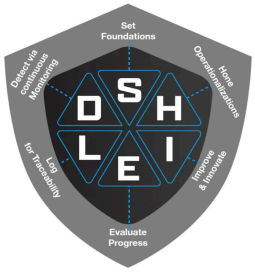
What is Responsible AI?

Responsible Artificial Intelligence (RAI) is a dynamic approach to the design, development, deployment, and use of trustworthy AI capabilities. DoD agencies must ensure their AI projects adhere to the RAI framework, which integrates DoD AI Ethical Principles while emphasizing the necessity for technical maturity to build effective, resilient, robust, reliable, and explainable AI. The RAI framework is intended to support AI development at the speed necessary to meet the National Defense Strategy.

The RAI SHIELD Assessment

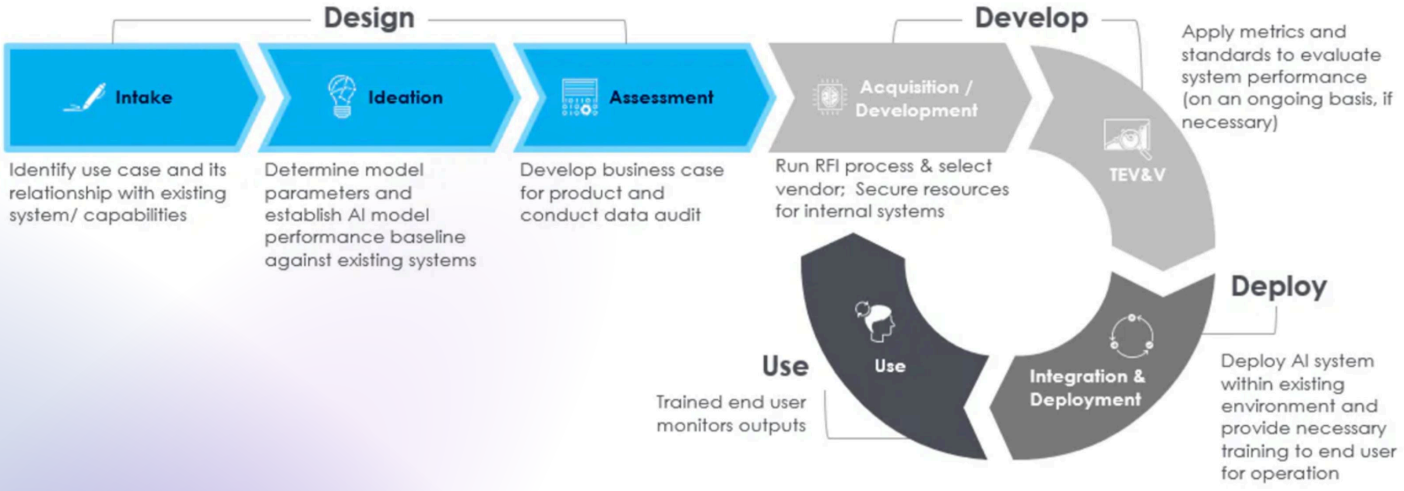
The SHIELD Assessment, part of the RAI Toolkit provided by the DoD, ensures each project identifies and addresses RAI-related risks and opportunities at all stages of the product lifecycle. SHIELD breaks these activities down into six sequential categories:

- **S**et Foundations
- **H**one Operationalizations
- **I**mprove & Innovate
- **E**valuate Progress
- **L**og for Traceability
- **D**etect via Continuous Monitoring



How Pryon's Deployment Methodology Adheres to RAI SHIELD

The SHIELD framework delineates four stages of the lifecycle of an AI project: Design, Develop, Deploy, and Use. As shown in the diagram below, these four stages can be divided into a total of seven phases. Pryon's solutions team will work directly with agency stakeholders throughout the entire lifecycle of an AI project and can flexibly adapt to meet client requirements, be they technical or procedural.



PRYON

www.pryon.com

V25_0106

Intake
The Intake phase typically occurs during the final negotiations of a Pryon RAG Suite procurement. Having identified and engaged stakeholders, the Pryon team will confirm the AI use case(s). Use cases will be confirmed and finalized per the stated requirements (i.e., workflow mapping, impact goals, and AI's role in the solution) for which Pryon's AI solution will be implemented. Upon finalization of procurement, the Pryon solutions team, in partnership with client stakeholders, will decide to proceed to the Ideation phase and agree to how the joint team will engage and track RAI-related issues. This includes having conducted feasibility and data/content availability analysis.

Ideation
The Ideation phase occurs upon procurement of Pryon software and starts with a refinement of requirements as preliminarily understood during Intake, including any necessary translation to functional requirements and design specifications. During this phase, any risks or statements of concern, by either the Pryon solutions team or client stakeholders, are identified and incorporated into the ultimate delivery plan to ensure they are tracked and managed throughout the project lifecycle.

Assessment
During the Assessment phase, the Pryon team ensures all requirements and statements of concern have success/failure criteria that have been mutually agreed upon with stakeholders. Moreover, all applicable content and data to be included in the solution is collected and labelled or tagged to ensure ground truth and eliminate bias. Prior to the Assessment stage, the Pryon team will have ensured AI is suitable for the solution given the explicit requirements. These elements of the Assessment phase are captured by updating existing project documents.

Acquisition/Development
Having developed requirements in accordance with DoD AI Ethical Principles, the Pryon team partners with customers to ensure the following are adhered to: documentation requirements, permissions, data access, roles and responsibilities, traceability, continuous monitoring, stakeholder engagement, and appropriate documentation procedures. Moreover, the Pryon team partners with customers to develop user acceptance testing and ongoing maintenance plans, including agreed-upon processes for updating the system and adding new content.

TEVV (Test & Evaluation, Validation & Verification)
Pryon RAG Suite is engineered to avoid the inherent risks of Generative AI seen in other products. Prior to the TEVV phase, the Pryon development team will already have tested the deployed Pryon solution to mitigate the risk of any adversarial attacks or model drift. During the TEVV phase, the Pryon solutions team partners with customers to test for performance, security, and adherence to requirements. As the system evolves through initial deployment, user acceptance testing, and ongoing updates per user requirements and content changes, documentation is updated accordingly.

Integration and Deployment
Pryon partners with customers to thoroughly test the deployed solution, train users, conduct user acceptance testing, educate customers on how to report support incidents, and update documentation as necessary.

Use
Once a Pryon AI solution has been deployed, the Pryon team conducts quarterly reviews with clients. This includes reviewing usage reports and metrics to determine how the system may be (re)configured and/or what content should be updated or removed. Through this partnership, customers can evolve their Pryon solution per the usage patterns and requirements of the user community, as these often become clearer with real-world usage. Moreover, if additional use cases are identified, the Pryon team can facilitate necessary updates, collection creation, and/or necessary configurations in support of these.

Conclusion

The Responsible AI Toolkit, DoD Ethical AI Ethical Principles, and SHIELD framework provide DoD agencies with a critical foundation to understand and assess how to implement and deploy AI in a government context. Pryon, a leader in retrieval-augmented generative (RAG) AI, stands ready to help DoD agencies complete the SHIELD framework and quickly achieve rapid results with trustworthy AI.