ISO/PAS 8800:2024 — 安全与人工智能培训

Agenda Vector Academy

课程形式: 线下
课时: 2 天
适用人员: 质量保证经理，系统工程师，安全工程师，AI ML 工程师，设计工程师，合规专员，项目经理，审核员
目标: 以 ISO/PAS 8800 为主线，协同 ISO 26262 与 ISO 21448，建立可审计的 AI 安全生命周期与安全案例，让 AI 风险可控 合规可证 量产可落地。

1. 培训内容

> 第 1 天 — 从理念到需求与措施

- 安全概念和内容概述：统一术语与适用范围，明确与 ISO 26262、ISO 21448 的协同关系
- AI 安全管理：角色职责、度量与里程碑设计，形成可审计的管理闭环
- AI 系统安全分析：围绕场景与风险，覆盖数据、架构、需求三个层面的系统化分析
- AI 安全需求的开发：自上而下分解到可验证条款，纳入数据与模型的约束与接口
- AI 安全措施：冗余、多样性、监控与降级策略嵌入架构与流程

> 第 2 天 — 从技术到验证与运行

- AI 技术选择，架构与开发措施：依据功能与约束做技术与架构取舍，将安全机制前置到设计开发
- AI 数据安全设计：建立数据全生命周期治理，覆盖采集、标注、版本、控制
- AI 系统的验证与确认：制定数据集与模型验证、系统级确认、场景覆盖与失效注入策略
- AI 系统运行阶段的措施：在线监测、置信度门控、漂移检测、处置与 OTA 再训练控制
- AI 安全开发框架和软件工具：需求与证据管理、数据治理等

2. 日程概览

第一天	第二天
• 安全概念和内容概述 • AI 安全管理 • AI 系统安全分析 • AI 安全需求的开发 • AI 安全措施	• AI 技术选择 架构与开发措施 • AI 数据安全设计 • AI 系统的验证与确认 • AI 系统运行阶段的措施 • AI 安全开发框架和软件工具

3. 培训概要

> AI 安全的必要性

- AI 进入安全关键系统，必须证明风险可控与剩余风险可接受
- 课程聚焦 ISO/PAS 8800 的核心理念 方法 证据

ISO/PAS 8800:2024 – 安全与人工智能培训

Agenda Vector Academy

- 全球项目实践分享
  - 用真实案例示范：场景定义、数据治理、架构与模型 V&V 的落地做法
- 产品责任与能力要求
  - 强化数据全生命周期治理、能力边界定义与运行期监测处置
- ISO/PAS 8800 关键增强
  - 补齐 AI 特有要求：数据与模型治理，训练，验证，部署和监控
  - 明确 AI 安全生命周期 度量与可信论证，形成可审计安全案例

4. 培训目标：

- 明确场景与风险：掌握 AI 在汽车安全关键系统中的应用边界与主要风险点
- 建立标准协同：理解 ISO/PAS 8800 与 ISO 26262、ISO 21448 的关系与裁剪思路
- 构建安全生命周期：将 AI 安全活动嵌入现有研发生命周期与组织流程
- 定义可验证的安全需求：结合数据生命周期，落地到系统与子系统可检验条款
- 开展系统化安全分析：覆盖数据，架构，需求三层的分析方法与结果闭环
- 实施 V&V 与可信论证：明确验证与确认方法，形成可审计的安全保证与合规证据
- 运行期保障与改进：建立监测，警戒，应急与持续改进机制，闭环全生命周期 ，熟悉常用 AI 框架与平台的可信使用要点

AI lifecycle – Reference AI Safety Lifecycle

