

# Project: TransportationBench



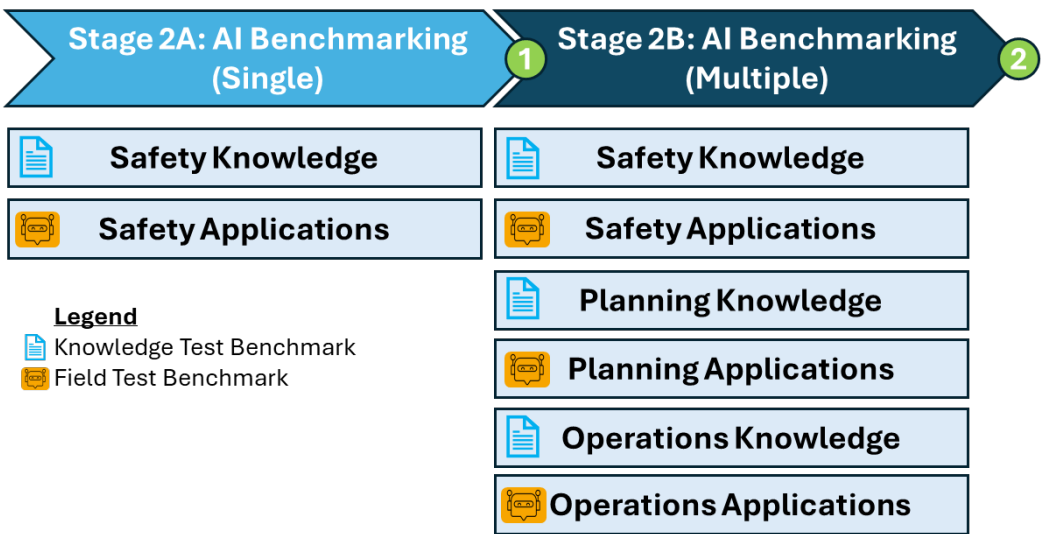
TransportationBench



Knowledge Test (e.g. RSP1)



Field Test (e.g. RSA)



- 1 **Mid-term Evaluation:** Benchmark tool reliably assesses AI model competency on single domain (e.g. RSP1)
- 2 **Final Evaluation:** Benchmark tool reliably assesses AI model competency on multiple domains (e.g. PTP, PTOE) via corridor study

• **Proposed R&D Project and Objectives:** First standardized benchmark for measuring AI models in their knowledge and application of transportation concepts. Objectives: a) build agency trust in AI, b) reduce evaluation costs, and c) establish adaptable testing framework.

• **Technical Landscape and Problem Addressed:** Agencies cannot confidently deploy AI due to lack of trust and domain-specific evaluation tools. This results in limited use for administrative tasks and missed innovation opportunities.

• **Proposed Technical Approach and Plan:** Two-phase validation testing AI against certified professional competency standards (RSP1, PTP, PTOE). 12-18 month R&D plan tests AI knowledge and practical application. Stage 2A focuses on one domain, Stage 2B expands to multi-domain with evaluating AI agent team on supporting a multi-disciplinary corridor study.

• **Commercialization and Transition Approach:** Dual pathway through integration with USDOT programs and direct partnerships with state and local transportation agencies demonstrating practical value through real world deployments.

• **Potential Impact:** Benchmarking helps unlock breakthrough capabilities for AI-augmented studies (RSA, Traffic Signal Studies, Feasibility Studies). Democratizes access of advanced tools for resource constrained smaller / rural agencies (95% of market).