GENINVO™

CodeMagic™

AI-Assisted Code Generation For Statistical Programming

Key Highlights

- SDTM, ADaM and TFL Code Generation using structured specifications and mock shells
- Predictable & controlled code generation vs ad-hoc AI code
- Human-in-the-Loop framework ensuring accuracy, quality, and compliance throughout the process
- Generates code in open-source programming languages like R for flexibility and reproducibility
- Seamlessly integrates with client Statistical Computing Environments (SCEs)
- Ensures GxP-compliant, audit-ready programming code with full documentation
- Customizable configuration to meet specific business requirements and workflows

Strategic Benefits

- Accelerate Time-to-Insight: Shorten development timelines enabling faster data-driven decision making
- Reduce Manual Errors: Intelligent code suggestions and structured prompts enhance data integrity
- Ensure Compliance: CDISC standards and Good Programming Practices are integrated into AI-generated code
- Optimize Resources: Free up programmer capacity for strategic analysis and innovation
- Minimize Training: Help SAS programmers transition to R with built-in learning support
- Codes can be used for production, validation or QC purposes

Current Capabilities

- **Automated ADaM Generation:** Variable-level metadata mapping and derivation logic
- **TFL Code Generation:** Creates TFL code from mock shells and ADaM specs
- **Raw to SDTM:** Automates conversion using detailed specifications
- **Reusable Standards:** Global standard programs for common derivations

Future Roadmap

- **SAS to R Migration:** Convert legacy SAS macros to modern R functions
- **Model Agnostic:** Customizable to work with any LLM model
- **Language Agnostic:** Generate Python, SAS, and other language code
- **Enhanced Capabilities:** Continuous expansion of automation features

GENINVO™

Technical Excellence
Building an AI-Assisted Coding System

Developing **CodeMagic** required solving complex technical challenges that bridge AI capabilities with the rigorous demands of **statistical programming**. Our engineering approach prioritizes accuracy, compliance and practical usability



Context Engineering

Integrating Study metadata (SDTM and ADaM specifications, TFL mock shells), standard programs and user inputs into precise, AI-ready context



Context Management

Managing large amounts of contextual data across workflow states while staying within system limits and preserving critical information



Prompt Engineering

Designing prompts that guide AI through complex, domain-specific tasks with accuracy and consistency



MCP Servers & Tools

Designing specialized MCP tools for diverse tasks with robust error handling, parameter validation and cross-environment compatibility



Retrieval-Augmented Generation

Implementing Retrieval-Augmented Generation with vector databases and semantic search for domain-specific knowledge retrieval



Domain Complexity

Translating deep statistical programming expertise, industry knowledge and best practices into automated, metadata-driven logic



Dynamic File Processing

Supporting diverse file types, structures and naming conventions through flexible parsing and robust error handling



Multi-Phase Code Generation

Breaking complex output generation into structured phases, where each stage builds on the previous one, requires careful coordination and validation throughout development

GENINVO is the go-to partner for our life science industry delivering innovative software solutions and Data Science services. We are focused on solving the toughest and most complex business problems and produce transformative solutions. Our unique combination of deep expertise in life sciences, leading edge technologies and software development, data-engineering practices, cloud expertise, and artificial intelligence research truly help bring life-changing treatments to patients faster