

# Custom AI Tool Development in Regulated Industries: Why Off-The-Shelf LLM Solutions Fall Short

Gerald T. Rigdon  
Fellow, Software Engineering  
Boston Scientific, Inc.  
[gerald.rigdon@bsci.com](mailto:gerald.rigdon@bsci.com)

## 1. Introduction

When I started working in the medical device industry almost 20 years ago, static analysis tools had captured the spotlight and attention of the medical device industry. This was apparent in a 2007 press article [1], that highlighted the United States Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) substantial investment in a software forensics laboratory. Brian Fitzgerald from the FDA was quoted at the time, saying "We're hoping that by quietly talking about static analysis tools, by encouraging static tool vendors to contact medical device manufacturers, and by medical device manufacturers staying on top of their technology, that we can introduce this up-to-date vision that we have."

I witnessed this outreach firsthand as I fielded numerous sales calls from static analysis tool vendors. Fortunately, I had already been grounded in real-world data and so in 2010 published a paper [2] for the *Embedded Systems Conference* in defense of customized static analysis tool solutions. As a point of interest, the custom solution featured in that paper is still in use today and has discovered a disproportionate number of software defects compared to OTS counterparts used to enforce organizational coding standards. Now, 15 years later this topic has risen in the context of AI tools, and I find myself compelled to speak once again.

Serious interaction with commercial AI platforms and tools such as Cursor, GitHub Copilot, Windsurf, and various enterprise AI web interfaces demonstrate the power and capabilities of this technology and OTS tools. However, riding along the wave of this enthusiasm is a misconception that organizations can simply purchase and deploy these OTS tools and then somehow fully realize the transformative potential of AI. While I believe this is generally the case, I will stay in my lane by addressing the unique challenges faced by medical device manufacturers. Intuition alone would seem sufficient to support the argument that pre-trained LLMs, despite their vast training corpus, lack the domain specificity, regulatory awareness, and data access necessary to provide optimal insights in safety-critical

contexts. However, presenting the case for custom tooling requires the need for conscious reasoning.

## **2. Data Integration**

The most significant limitation of OTS AI solutions is their inability to access and leverage proprietary organizational or domain-specific data. Hence, Retrieval-Augmented Generation (RAG) architectures, as described by [3], address this limitation by combining LLM reasoning capabilities with domain-specific knowledge retrieval. The effectiveness of RAG systems vs pre-trained base model LLMs on domain-specific tasks was documented in [4] which revealed 30-50% improvements in LLM response accuracy. Custom AI tools can uniquely implement RAG systems that:

- Index proprietary domain information using semantic embeddings
- Retrieve contextually relevant information from these embedding data sources
- Ground LLM responses in domain data
- Maintain organizational security boundaries

## **3. Domain-Specific Workflows and Process Integration**

The FDA's Quality System Regulation (QSR) and international standards such as ISO 13485 define specific workflows and defer to other standards such as ISO 14971 for risk management and IEC 62304 for software lifecycle processes which includes verification and validation activities, change control and configuration management, etc. While this information is in the public domain and part of the vast training corpus available to LLMs, each medical device manufacturer has their own unique quality system derived from these standards and principles. What does this mean in practice?

Modern AI tool development increasingly employs multi-agent architectures where specialized LLM agents [5] manage specific workflow stages. For medical device development, this might include:

- Extracting and validating requirements from internal proprietary specifications
- Analyzing designs against regulatory standards and best practices and organizational domain constraints
- Generating compliant code following organizational coding standards
- Creating verification test cases with traceability to documentation that exists outside of immediate LLM context
- Generating documentation with proper formatting such as organizational templates

OTS solutions can only provide this level of sophistication if they have knowledge of organizational processes and their respective quality management systems.

#### **4. Tool Integration and Ecosystem Connectivity**

The research in [6] demonstrates that LLMs perform significantly better with the use of appropriate tools. The Model Context Protocol (MCP), introduced by Anthropic in 2024 [7], is leading the way by providing a universal protocol for connecting LLMs to data sources and tools through a client-server architecture.

Although this is a universal standardization effort, MCP actually reinforces the need for custom tool development instead of eliminating it. Organizations must still build custom MCP servers that understand their domain-specific data structures, implement security access controls, and handle proprietary data file formats. This includes:

- Building connectors to legacy systems
- Reformatting data for MCP resources
- Managing authentication and authorization
- Understanding how to appropriately expose data to MCP resources
- Expertise in MCP tool implementations
- Maintaining MCP servers as requirements change

#### **5. Cost-Effectiveness and ROI**

The information in [8] supports the claim that custom AI solutions outperform OTS options. Hence, organizations achieving significant ROI share common characteristics such as deep integration with core business processes, data-driven approaches leveraging proprietary information, continuous improvement cycles, and custom solutions tailored to specific needs. Moreover, custom tool development, though requiring upfront investment, provides long-term cost advantages such as:

- Unlimited internal usage
- Full control over infrastructure and scaling
- Reusable components across multiple applications

Objections that emphasize an organization's primary product focus and are quick to recommend either OTS only solutions or outsourcing development to consultants or vendors over internal resources risk missing a core understanding of the nature of AI tool development and the strategic value of domain expertise. Given the exposure to problem-solving, understanding algorithms and data structures, etc., it would not be a stretch to conclude that these transferable skills would support the claim that software engineers with

strong fundamentals can achieve proficiency in LLM application development significantly faster than domain experts can acquire deep technical knowledge of complex systems. So, the dream scenario for an organization desirous of maximizing AI utility would be domain experts who are skilled software engineers. The practical challenge is the appropriate allocation of those resources.

## 6. Conclusion

There is substantial evidence to support the need for custom AI tool development in regulated industries like medical device manufacturing. While OTS AI solutions can provide value, the future of AI technology in regulated industries will require building intelligent systems that deeply understand and complement domain-specific expertise. AI is quickly becoming a core engineering capability. Organizations that treat this technology as something to outsource rather than a core competency should recalibrate their strategic awareness or risk losing a competitive advantage.

## References

- [1] Chloe Taft. (2007, October). *CDRH Software Forensics Lab: Applying Rocket Science To Device Analysis*. Medical Devices Today.
- [2] Rigdon, G. (2010, July). *Static Analysis Considerations for Medical Device Firmware*. Embedded Systems Conference Proceedings.
- [3] Lewis, P., et al. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. Advances in Neural Information Processing Systems, 33, 9459-9474.
- [4] Gao, Y., et al. (2023). *Retrieval-Augmented Generation for Large Language Models: A Survey*. arXiv preprint arXiv:2312.10997.
- [5] Park, J. S., et al. (2023). *Generative Agents: Interactive Simulacra of Human Behavior*. arXiv preprint arXiv:2304.03442.
- [6] Schick, T., et al. (2023). *Toolformer: Language Models Can Teach Themselves to Use Tools*. arXiv preprint arXiv:2302.04761.
- [7] Anthropic. (2024). *Introducing the Model Context Protocol*. Anthropic Blog. <https://www.anthropic.com/news/model-context-protocol>
- [8] Markovic, D. (2025). *Why Custom AI Solutions Outperform Off-the-Shelf Options*. Medium. [https://medium.com/@dejanmarkovic\\_53716/why-custom-ai-solutions-outperform-off-the-shelf-options-0b9463b9fefc](https://medium.com/@dejanmarkovic_53716/why-custom-ai-solutions-outperform-off-the-shelf-options-0b9463b9fefc)