



REWRITING THE FUTURE

AI-POWERED MODERNIZATION OF NORTH CAROLINA'S BIOLOGICAL DATABASE

Initiated: January 2023

Completed: April 2025

Agencies: N.C. Department of Information Technology
N.C. Department of Environmental Quality

Contact: Stephanie McInerny
stephanie.mcinerny@deq.nc.gov
(252) 515-5501

Executive Summary

Within the North Carolina Department of Environmental Quality, the Division of Marine Fisheries (DMF) depends on timely, accurate, and comprehensive biological data to manage marine species critical to the state's environment and economy. Since its inception in 1980, the Biological Database (BDB) has grown to house more than 18 million records across over 120 programs, encompassing data dating back to the 1960s. These data are central to species stock assessments, fishery management plans, and regional and federal reporting obligations. However, despite its foundational importance, the BDB's aging technical architecture had become a liability. The 40-year-old legacy system was coded in a patchwork of outdated programming languages including Gawk, C, Visual Basic, and Windows batch files. Many of these scripts were undocumented, poorly standardized, and written by staff no longer with the agency. The tools used daily at the division to access the BDB were built using outdated versions of software that aren't compatible with modern technology, and for many, the source code was unavailable so updates couldn't be made to fix issues or meet current business needs. Some functionality was even hardware-dependent, operated solely on unsupported operating systems such as Windows XP. This posed growing risks to operational continuity, security, and compliance with both state and federal reporting mandates. In addition, it made data querying and analysis difficult, time-consuming, and in some cases impossible for those trying to use the system. The new BDB application launched in July 2024, replacing dozens of legacy tools. Continued development led to the successful deployment of a new data entry solution in April 2025.

Traditional approaches would have required years of manual code review, translation, and debugging. Recognizing the urgency and complexity of the situation, DMF IT adopted an innovative strategy: leveraging Generative Artificial Intelligence (AI) to accelerate and enhance every stage of the process to modernize this mission-critical system. Through the responsible application of OpenAI's ChatGPT using paid licenses costing a total of \$600 annually and deployed in a private, non-training environment, the DMF IT team successfully translated, documented, and refactored over **25,000** lines of legacy code and built a modern web-interface to replace the deprecated applications. The new interface provided a more efficient and optimized experience for the user while strengthening the system's security, resilience, and maintainability. Additionally, a custom GPT assistant—built to guide users through common tasks, answer real-time questions, and resolve errors—surpassed expectations and earned enthusiastic feedback from stakeholders during its pilot phase.

This AI-assisted approach reduced the projected modernization timeline by **50%**, decreased scripting and debugging time by **90%**, saved an estimated **\$340,000** in direct consultant costs, and provided a scalable model for other state agencies struggling with similar legacy system challenges. The new modern architecture provides more robust capabilities including a **10x** speed increase for querying data, enhanced data verification tools that cut editing and verification time by **50%**, increased transparency in data operations by **100%**, and increased security. With Windows XP eliminated and AI-enabled workflows implemented, staff can now securely maintain and update the BDB remotely, an essential capability for operational

continuity and resilience post-COVID. The transformation significantly improved data access and accuracy, significantly enhanced user satisfaction, and established North Carolina as a pioneer in using AI to modernize legacy government IT systems, setting a practical example for other states to follow. The project's success underscores the potential of AI not as a replacement for human expertise, but as a force multiplier that enhances team capabilities, safeguards public data, and delivers measurable improvements in efficiency, engagement, and security.

IDEA

The Biological Database (BDB), created in 1980, is vital to the North Carolina Division of Marine Fisheries' (DMF) core mission and supports the long-term collection, analysis, and reporting of marine fisheries data critical to environmental stewardship and regulatory compliance. It houses more than 18 million records across 120+ fishery-dependent and fishery-independent sampling programs, encompassing data dating back to the 1960s. Prior to modernization, the system relied on fixed-width ASCII text files, where each 128-character line encoded hierarchical structured data across categories like environmental data, gear type, species, individual fish ages, and tag recapture details. These records support stock assessments and fishery management plans used for regulatory compliance at the state and federal level. Data from the BDB feeds into national initiatives, such as the Southeast Area Monitoring and Assessment Program, NOAA's Fisheries' Trip Interview Program, and the Atlantic States Cooperative Statistics Program. Yet the system's aging technical foundation had become a liability.

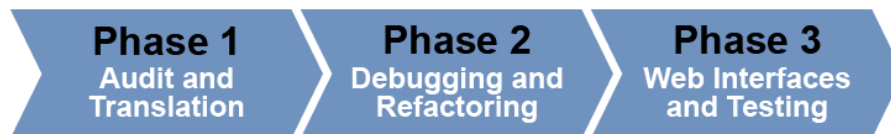
The BDB relied on antiquated codebases lacking proper documentation. Many scripts were written in programming languages that had fallen out of use in state government IT environments, including Gawk, C-based data engines, and layers of batch files with unknown dependencies. With minimal knowledge transfer, the unforeseen departure of the database administrator left behind opaque, unmaintainable systems that functioned without being fully understood. The essential data entry tool operated solely on Windows XP, introducing significant operational and security vulnerabilities. Additionally, the absence of version control further hindered maintainability and oversight.

The fragile BDB architecture was experiencing system failures regularly, but it reached a breaking point in 2023 when a system crash revealed severe gaps in maintainability as new data were unable to be entered, existing data were unreachable within the system, and data analyses weren't reproducible. Notably, DMF IT had already initiated development of a new web-based solution at the time, recognizing the aging system's risks and proactively working toward a sustainable future. However, the crash accelerated the urgency of these efforts. To compound matters, in early 2025, the XP machines that ran the data entry software also failed and were unrecoverable. Without swift action, DMF faced not only the risk of noncompliance with mandatory data requirements—which could have triggered regulatory penalties and loss of federal funding—but also broader consequences including operational paralysis, compromised decision-making due to inaccessible data, and erosion of stakeholder confidence in the agency's ability to fulfill its mission.

Traditional solutions were either prohibitively expensive or infeasible. Contracting consultants with the requisite experience in obsolete languages would have cost hundreds of thousands of dollars and potentially taken years to fully restore the system’s functionality. Training existing staff to become proficient in multiple outdated coding languages was neither realistic nor scalable. DMF needed a solution that could scale expertise, move fast, and build internal capacity—not just replace one fragile system with another. Generative AI presented a viable path forward, not only addressing this specific challenge but also serving as proof of concept for a broader application of Generative AI that aligns with several priorities outlined by state CIOs. This solution effectively targets three of NASCIO’s Top Ten Priorities for State CIOs in 2025: 2–AI/Machine Learning/Robotic Process Automation, 4–Data Management and Analytics, and 5–Legacy Modernization.

IMPLEMENTATION

The use of AI in the modernization of the BDB began in early 2023. DMF selected OpenAI’s ChatGPT—deployed in a private, state-managed environment that opted out of training on user data. Scripts were sanitized to remove credentials and uploaded in stages. What followed was a transformative partnership between human developers and AI to build new and enhanced functionality for the BDB as a web-based application with a SQL Server backend. This project encompasses the work done over the past two years and the launch of the new BDB web application. Incorporating AI into the software development process occurred in a three-phase approach. This approach was repeated for each issue that occurred and each part of the BDB that needed to be replaced.



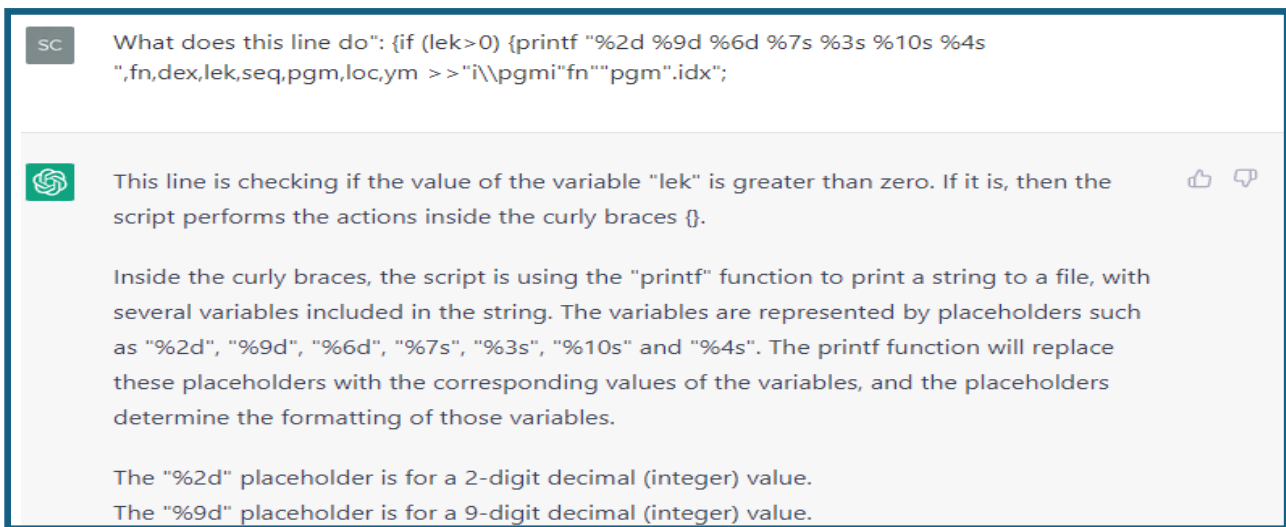
Phase 1: Audit and Translation

DMF’s IT staff conducted a comprehensive audit of the existing codebase. They sanitized scripts by removing sensitive information, ensuring that no secure credentials or proprietary server addresses were exposed. The team then uploaded these scripts into a private ChatGPT workspace that complied with strict non-training protocols, ensuring that no data provided to the AI would be used for broader model training or accessible beyond the project scope.

During this phase, the team used AI to translate over **25,000** lines of legacy code across multiple programming languages. ChatGPT provided plain-language explanations for obscure syntax and flagged potential errors or inefficiencies. Below is an example of just **one** line of GAWK code that was translated using AI.

Phase 2: Debugging and Refactoring

As staff juggled SAS, R, SQL, Python, Gawk, VB, C#, and C, the risk of errors grew. One recurring issue came from an R script failing to correctly assign a dataframe. ChatGPT identified the scope mismatch and rewrote the block, resolving the bug. These fixes were not just faster; they were teachable moments. The AI not only resolved issues but explained why, upskilling staff in real time. AI played an instrumental role in identifying logic flaws and security vulnerabilities hidden deep within the legacy scripts. It guided the team through error resolution in real time, often providing insights that would have taken human analysts days or even weeks to uncover. Errors that previously required two to three days to diagnose were resolved within two to three hours with AI assistance.



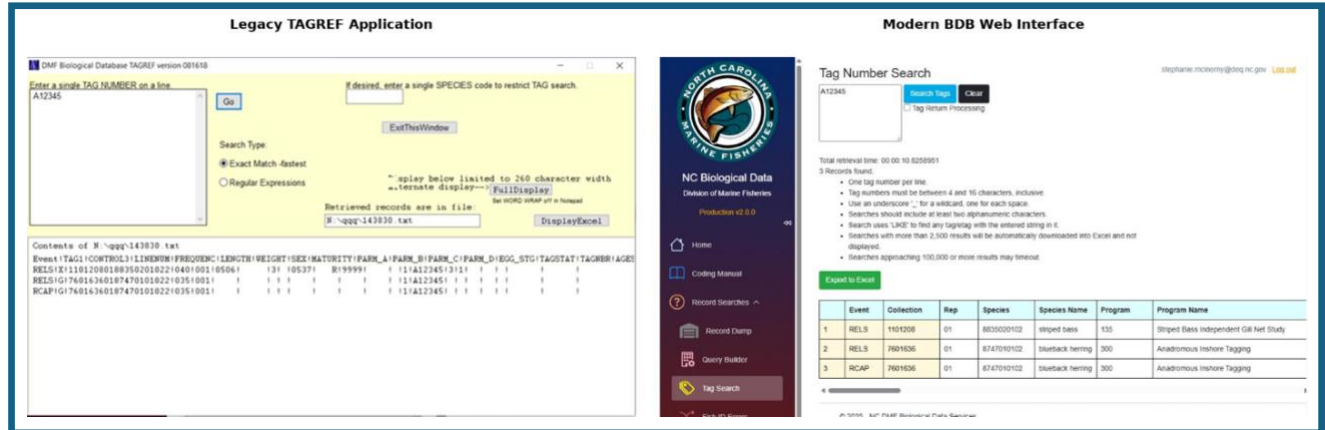
The screenshot shows a chat interface with a user question and an AI response. The user asks: "What does this line do": {if (lek>0) {printf "%2d %9d %6d %7s %3s %10s %4s",fn,dex,lek,seq,pgm,loc,ym >>"i\pgmi"fn""pgm".idx};". The AI response explains that the line checks if the variable "lek" is greater than zero. If true, it uses the printf function to print a string to a file, with variables represented by placeholders like "%2d", "%9d", "%6d", "%7s", "%3s", "%10s", and "%4s". The response also clarifies that "%2d" is for a 2-digit decimal (integer) value and "%9d" is for a 9-digit decimal (integer) value.

AI provided scaffolding for new code in modern languages such as C# and Python, accelerating development of the new web application and ensuring adherence to best practices. Modular services replaced one-off legacy solutions, improving both maintainability and scalability.

Phase 3: Web Interfaces and Testing

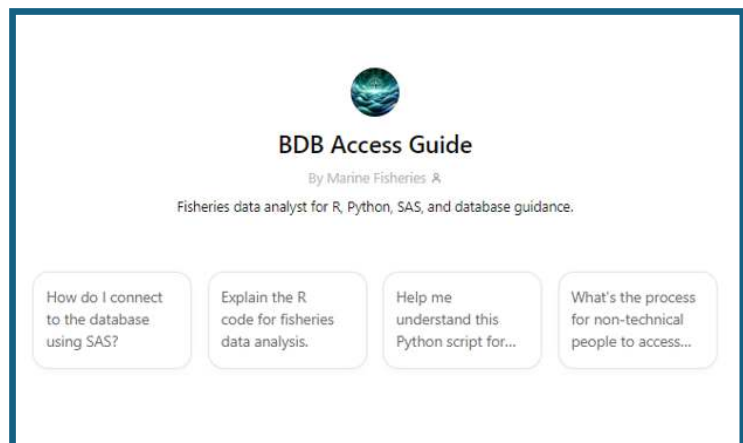
The DMF IT team created a modern web-based application using Microsoft Blazor to interact with the new SQL Server version of the BDB database. AI-assisted scripting allowed staff to more quickly develop a dynamic, user-friendly web application containing custom search functionality and dashboards that present fisheries data in intuitive formats. This **one** modern application replaced **a dozen** piecemeal applications that all failed when the legacy BDB crashed. AI was also leveraged to generate custom unit test scripts, allowing the team to validate system functionality quickly and ensure new features met both technical requirements and user expectations. The new BDB application was released to users in July 2024, delivering features that replaced dozens of legacy systems. Development efforts continued to integrate remaining functionality, including data entry and data tracking components.

With the help of AI, a new solution was developed to allow the existing 16-bit data entry software to run on current operating systems (Windows 11). AI was instrumental in configuring the boot files and in building an easy-to-use interface so existing users would experience a seamless transition to the new application. This solution was deployed in April 2025. Without this, DMF would have again been at risk of missing mandatory state and federal data deadlines.



Custom GPT Assistant

During the project, DMF IT pioneered the development of a custom GPT assistant. This AI tool was trained in sanitized versions of the BDB's user guides and standard operating procedures. The custom GPT provided step-by-step instructions for common tasks, answered staff queries in real time, and even explained error messages to non-technical users. In initial pilots, the custom GPT performed well and was said to be a potential game-changer for training and onboarding. This initiative hasn't been released to the user base yet, but plans for incorporating this into the BDB Web application are forthcoming.



IMPACT

The BDB modernization delivered measurable improvements in cost, speed, security, usability, and staff empowerment.

- **Time Savings:** With the help of AI, script refactoring and debugging timelines decreased by **90%**. Script conversions dropped from 120 hours to just 10, and code turnaround went from 2-3 days to 2-3 hours.
- **Cost Avoidance:** The AI-driven approach saved approximately **\$340,000** in estimated consultant costs (1 contractor at \$70.00/hour for 4,853 hours over 2 years and 4 months).
- **Cost Efficiency:** The license cost for ChatGPT is **\$600 per year**.
- **Security, Resilience, and Transparency:** **100%** of unknown executables were replaced with clean, documented code. The new SQL architecture is fully compliant with state cybersecurity standards, and the modular architecture substantially improved disaster recovery capabilities.
- **User satisfaction:** Stable user-friendly interfaces in the new architecture resulted in a significant increase in user satisfaction by providing a **10x** speed increase in data querying and cutting data editing and verification times by **50%**.
- **Workforce Enablement:** AI removed the bottleneck of relying solely on legacy experts, enabling a broader range of staff to understand, modify, and enhance critical systems. This shift transformed development into a collaborative, inclusive process, accelerating progress and reducing risk.
- **Remote Readiness:** With Windows XP eliminated and AI-enabled workflows in place, staff can now maintain and update the BDB remotely, a critical capability in a post-COVID world.

The DMF BDB modernization exemplifies how AI can responsibly and powerfully transform state government IT. Rather than replacing humans, AI became an expert teammate scaling capacity, accelerating timelines, reducing costs, and enriching institutional knowledge. The result is a secure, modern, maintainable system that has preserved one of the most important fisheries datasets on the East Coast.

This approach has been extended to other development teams within the department where similar results are expected. The solution has been presented in about a dozen conferences and panels targeted at North Carolina state agencies.

Next Steps

Development of new and enhanced functionality for the BDB will continue with AI as a collaborative partner, supporting DMF as it adapts to evolving needs and improves data collection methods to ensure ongoing compliance with state and federal conservation mandates.

In an era where all states face legacy technical debt, workforce turnover, and cybersecurity mandates, this project shows what's possible. North Carolina has not only safeguarded its biological data but also created a roadmap for modernization that others can follow—not in theory, but in practice.