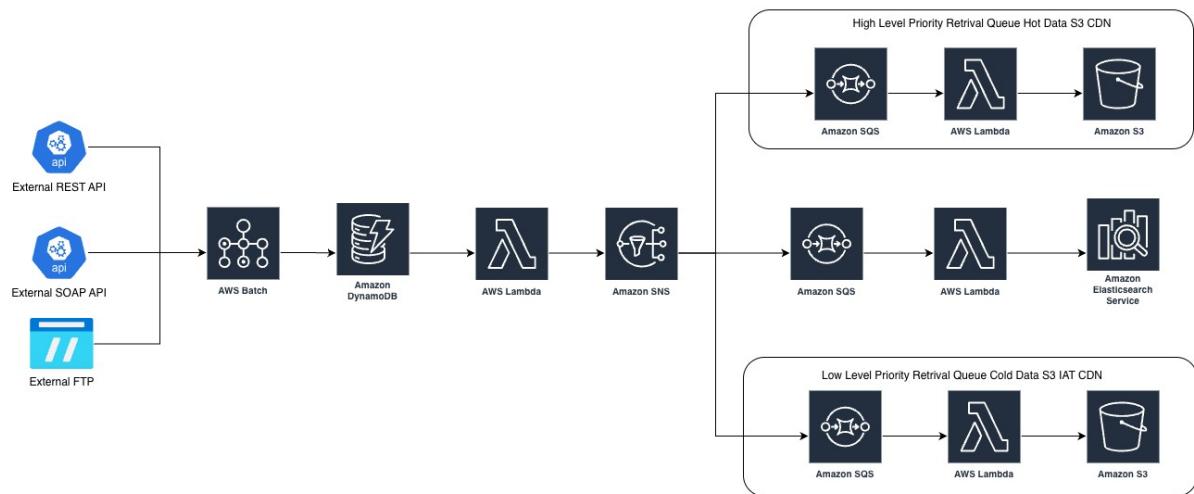


Beyond Microservices

The Architectural Blueprints for Building Business Optimized Data Platforms



Author

Ivan Vukasinovic

Technology Executive and Solutions Architect

About the Author

The author is a seasoned technology executive and solution architect with a proven track record of turning technical infrastructure into significant business value. Three of his four startups were successfully acquired, a feat driven specifically by the advanced capabilities of the proprietary data platforms he engineered. In one notable instance, he designed a high-velocity architecture for a real estate startup capable of processing billions of records in near real-time, achieving this scale at a fraction of the cost of competitor solutions. With over 15 years of experience ranging from early AI adoption to securing critical services for global tech giants, the author specializes in architecting complex distributed systems that serve as the scalable foundation for modern business growth.

Why This Book?

This book serves as a comprehensive guide for building and scaling robust, agile, and data-driven software platforms. It emphasizes the need for architectures that are both flexible and scalable, balancing performance with cost-efficiency to align with modern business needs. Using a car-search platform as a running example, the book illustrates practical applications of concepts, addressing challenges in platform evolution and demonstrating solutions grounded in real-world scenarios.

The journey from a simple system software application such as a monolithic web application to a complex, distributed software platform processing petabytes of data extracted from hundreds of different data sources in real-time is filled with challenges and questions that will be answered in this book. Here are just some of those challenges:

- How do you design for scalability and flexibility from day one?
- How do you balance performance and cost, so you don't burn all your cash on running the software side of your business?
- What technologies and design patterns will ensure your platform remains relevant in an ever-changing technological landscape?
- How do you handle vague business requirements and ever changing business landscape on features and SLAs?

Audience

This book provides a roadmap for architects, engineers, and decision-makers to navigate the complexities of building complex software platforms. Whether you are starting or currently have a system that is not working properly, whether you have with a monolithic architecture, transitioning to microservices, or exploring how to leverage AI within your current system, this book equips you with the insights and tools to succeed. Aside from hands on tech people this book is included for an even wider audience as a decision-making information, this broad audience includes:

- **Founders and Entrepreneurs** learning enough to able to make product decisions
- **Solution Architects** looking to design scalable and maintainable platforms.
- **Data Engineers** focused on building robust data pipelines with AI capabilities.
- **Technology Leaders** aiming to future-proof their data strategies.

Key Themes and Insights:

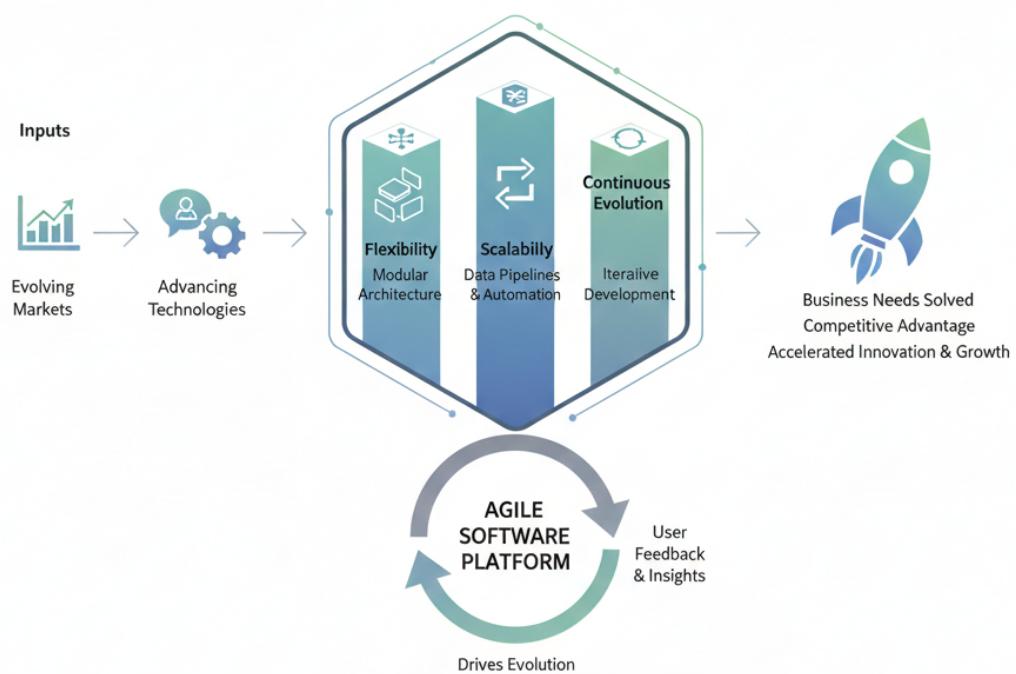
- **Foundation and Evolution:**
 - Explains what software platforms are and the need to evolve from monolithic architectures to Service-Oriented Architectures (SOA) and eventually to microservices.
 - Focuses on principles like modularity, scalability, and guided evolution to ensure adaptability in dynamic business environments.
- **Data Ingestion and Processing:**
 - Covers strategies for ingesting diverse data from APIs, file uploads, and other sources.
 - Highlights real-time and batch processing approaches, including trade-offs between cost, performance, and SLAs.
- **Analytics and AI Integration:**
 - Details the importance of splitting operational and analytical storage to ensure efficient data handling and enable advanced analytics.
 - Explores AI integration, such as validating car models in photos and ensuring image compliance using tools like AWS SageMaker and Rekognition.
- **Operational Excellence:**
 - Provides best practices for CI/CD, monitoring, and disaster recovery.
 - Discusses managing environments (alpha, beta, production) to streamline development, testing, and deployment.
- **Scaling the Platform:**
 - Identifies key challenges of scaling, including bottlenecks in file retrieval, analytics costs, and database constraints.
 - Advocates transitioning to microservices for modularity and scalability, employing distributed storage (e.g., DynamoDB) and tiered storage for cost efficiency.
- **Performance Optimization:**
 - Explores techniques for optimizing Time-To-Live (TTL) in data pipelines, balancing event-driven architectures with batch processing.
 - Stresses alignment between engineering and business goals to optimize costs while meeting SLAs.

Chapter 1: Introduction to Agile Software Platforms

1.1. What is an Agile Software Platform?

In today's data-driven world, software platforms are the backbone of innovation. They enable businesses to solve business needs, automate processes, harness data, generate insights, and drive decisions at scale. But as markets evolve, user expectations grow, and technologies advance, platforms must adapt to stay relevant. This adaptability—the ability to respond quickly to change—is what defines an **Agile Software Platform**.

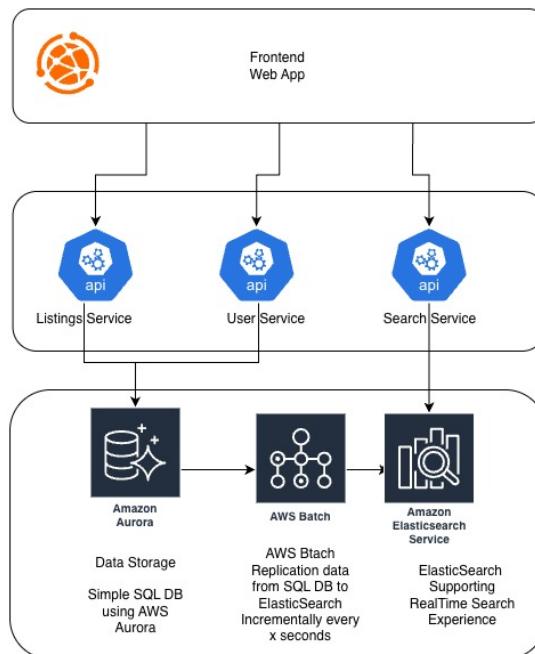
An Agile Software Platform is designed with flexibility, scalability, and evolution at its core. It's not just about handling large volumes of data; it's about doing so efficiently, reliably, and in a way that supports continuous improvement. Unlike traditional platforms that are rigid and slow to adapt, Agile Software Platforms prioritize iterative development, modular architecture, and a user-focused approach. They evolve in response to feedback, allowing businesses to stay ahead in competitive landscapes.



6.2. Case Study: Enhanced Search on the Car-Search Platform

The car-search platform needed to enable autocomplete, fast search, and filtering on millions of car listings while ensuring data accuracy and consistency. Solution Architecture is described below as well as depicted in a diagram:

- **Data Syncing:** A replication service extracted relevant fields from RDS (e.g., make, model, price, mileage) and transformed them for indexing in OpenSearch.
- **Autocomplete:** OpenSearch's completion suggestions were configured for car make and model, delivering real-time suggestions.
- **Fast Filtering and Sorting:** Filters and sorting options were implemented using range queries and aggregations in OpenSearch.



Splitting data into Stored Data in Amazon RDS and Indexed Data in OpenSearch enables advanced search capabilities while ensuring performance, scalability, and consistency. By leveraging the strengths of both systems, the platform delivers an exceptional user experience and supports complex business requirements effectively. In our next chapter we will start working on scaling the platform to solve bottlenecks and optimize operational excellence by migrating the architecture from SOA to microservices. Outcome achieved in Car-Search Platform with integration of OpenSearch as depicted in the Solution Architecture Diagram are:

- Achieved sub-200ms search response times.
- Enhanced user experience with autocomplete and dynamic filters.
- Maintained synchronization and data consistency across both storage layers.

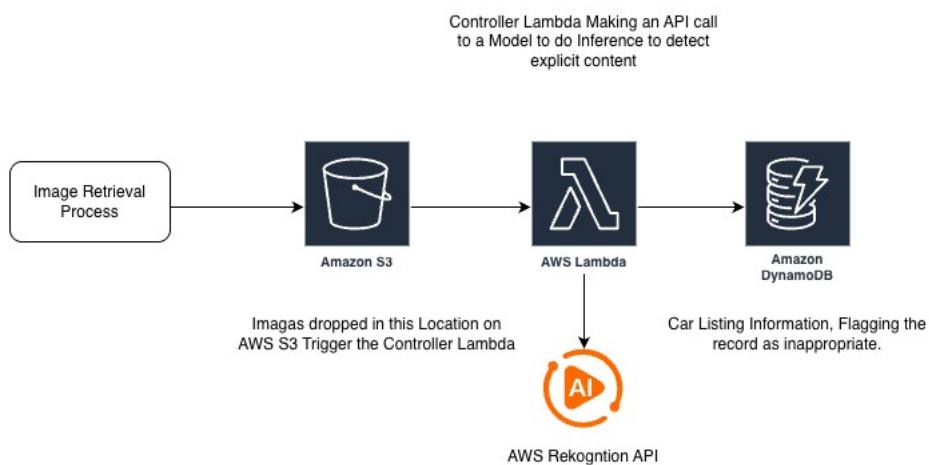
Chapter 10: Integrating AI into the Platform

The integration of artificial intelligence (AI) into data platforms is a transformative step in enhancing platform capabilities, improving user experience, and automating manual operations to keep costs under control as the platform operations scale up due to increased usage. An agile data platform provides a crucial foundation for this integration by enabling modularity, scalability, and flexibility. With its ability to rapidly adapt to new requirements and incorporate advanced components, an agile platform ensures seamless deployment of AI solutions. By automating workflows, personalizing user experiences, and optimizing operations, these platforms empower businesses to stay ahead of evolving demands.

This chapter delves into practical methods for integrating AI, so let's start with a simple use case on a car-search platform by ensuring Image Compliance for Explicit Graphic Content. Let's focus on the first objective, detect and filter out images containing explicit or inappropriate content to ensure compliance with platform policies. Implementation Steps to achieve this are:

- Upload newly added images to an S3 bucket that triggers a Lambda
- The triggered AWS Lambda makes a Call to a LLM Model via API to flag content
- Flagged images prevent a car record to be visible to end users
- Additionally, we can send an alert if this listing came from a dealership

The desired outcome of this improvement is to create a safe and professional environment for the users by removing inappropriate images. Agile platforms also facilitate iterative development, allowing AI systems to be refined over time to address emerging challenges and opportunities. Once above are implemented your architecture should look like the one depicted in the image below.



Integrating AI into the car-search platform enhances user trust, compliance, and operational efficiency. By combining custom-trained models and pre-built solutions like AWS Rekognition, the platform achieves a robust, scalable, and intelligent system. This approach exemplifies how AI can be seamlessly woven into data platforms to meet both technical and business objectives.