Research Note
BY RICHARD WINTER • AUGUST 2025

"You're Gonna Need a Bigger Boat"

A Metaphor — and a Recommendation — for AI and the Data Platform

AN HOUR INTO THE 1975 MOVIE *"JAWS,"* the main characters finally see the shark for the first time. The size and true menace of the shark is shocking to them, and to us in the movie audience. Even though everyone already had heard that a shark was killing people, no one is prepared for the enormity, and ferocity, of the animal as it first reveals itself. Shaken by his first look, our hero, the local police chief (played by Roy Scheider) says to the shark hunter (played by Robert Shaw), "You're gonna need a bigger boat." This line is widely quoted as one of the best in the history of movies. You can watch the clip here: *Jaws* (1975) - You're Gonna Need a Bigger Boat Scene (4/10) | Movieclips (YouTube).

AI is like that giant shark during the first hour of the movie — it is making its presence felt but most of us still don't realize how massive it is — or perhaps how hungrily AI will consume enterprise data.

This research note is about the demands about to be generated by AI and about the "bigger boat" — the data platform(s) or data architecture that many data intensive customers are going to need — to handle those enormous and hungry AI requirements. I'm going to recommend a crucial first step that I believe every data leader should take. Like the characters in *Jaws*, those who don't prepare for this change are likely to find that their projects, if not their entire companies, will soon be in mortal peril.

WinterCorp
www.wintercorp.com

THE COMING AI WORKLOADS

I HAVE BEEN ANALYZING and projecting analytic data workloads for over 30 years to help customers select data platforms, develop solutions and evaluate architectural options. In my view this is one of the most crucial steps in selecting the data platform and developing the data architecture. It makes a great deal of difference whether your workloads are heavy or light; whether they are dynamic or steady; and, what the mix of queries and analytics looks like. A relatively small number of complex operations can put as much stress on a data platform as a large number of simple queries. Analytics that touch a lot of data are fundamentally different than those access a single record. Large, complex joins are often a particularly challenging component of workloads — as are many advanced statistical and machine learning operations.

The point is that the nature and scale of workloads can profoundly affect both architectural requirements and the cost of operations.

So, what is going to happen as the usage of enterprise AI builds up? Most companies are pursuing strategies in which AI:

- Makes access to enterprise data ***available to a much larger audience*****.** In my opinion, many companies will find that **the population of users directly accessing the data via AI will be *ten or more times larger* than in the mostly BI (Business Intelligence)-driven era of the last ten years**;
- **Provides *natural language query* against enterprise data and documents, thereby making it *much easier to ask complex business questions* than in the past**; in the past, a business knowledge worker or decision maker with such a question either had to: (a) find an analyst who could create and submit a complex query and who was available to do so; or, (b) struggle through doing so himself or herself, often using tools in which it is not easy to create and manage complex queries;
- **Enables *autonomous agents* (think: digital robots) to perform all sorts of tasks that we used have to do ourselves**; many of these will involve interaction with enterprise data; many people working in business functions will have multiple such agents working on their behalf, often accessing enterprise data on their own;
- **Supports *semantic search* (also called similarity search) on large volumes of enterprise data**; this means that ***many less precise questions will be asked*** ("show me products similar to that one"); these are more complicated for the query engine to process than most traditional database queries;
- Enables ordinary business users to **easily generate analytics that range over a wide variety of data, *including documents, audio, images, video, sensor data*, in addition to the structured data that has been the focus of the BI era**.

These are just some of the data intensive strategies that are being widely pursued now. As AI advances, surely it will result in making more use of enterprise data, not less.

AI will increase the user population actually using enterprise data by a large factor in most organizations. If you have 1000 people today — business analysts and data analysts generating queries, dashboards and reports — and perhaps conducting the occasional citizen data science project — my guess is that you will between 10,000 and 100,000 doing that as your AI strategies take hold. And, many of those people will have agents acting on their behalf, also generating their own interactions with the data. And, who will these people be? Initially, employees, but within a few years the user population will expand to include the entire extended enterprise: suppliers, partners, customers… anyone with a relationship to the business. External users will often be prevented from accessing sensitive data, but there is a great deal of useful data that is not sensitive — or that pertains to the user's particular relationship with the enterprise — and their use of it will likely increase by a large factor.

SO WHAT?

SO, WHY SHOULD YOU CARE that AI will bring massive new workloads to your data architecture? In short, here is the reason:

> **In most cases, the data platform has been selected, and the data architecture has been created, for the world of business intelligence (BI) — NOT artificial intelligence (AI). Sometimes the existing data platform may be able to handle all the required work in the brave new world of AI, but…**
>
> **I believe that *most companies* are going to "need a bigger boat": they will need to re-assess their data architecture with AI workloads in mind.**

In fact, I think it is not only about the workloads. The nature of business is that it naturally invites people to combine data from different sources to answer business questions. This was one of the key insights that led to the emergence of the data warehouse and the modern era of competing on analytics. In the world of BI, it wasn't as easy for non-specialists to formulate the queries that bring together many different kinds of information. A lot of companies are living with a lot of siloed data in part because nothing forces them to integrate it. In the world of AI, there is likely to be much more pressure to integrate data from disparate sources.

To give one example, in the world of AI, it is easy to ask questions that relate data about markets and customers with data about supply chains. There is going to be much more pressure to break down silos and make data from different sources usable together, meaning yet more complex queries hitting the data platform. This also will add to the workloads and the stresses on data architecture.

WHAT SHOULD YOU DO?

Okay, so there is a giant shark out there called AI and you need to prepare for it.

I recommend that you assess your data architecture *with AI in mind*. Define what you are likely to need and take a look at what you have got. If you see major gaps, start exploring options for addressing them.

If your company is adopting AI strategies that will require changes to your data architecture, *these changes should be built into the AI implementation plan*. You don't want to wait until the AI rollout is failing because the data architecture is unable to support it.

And here is the crucial point: you want a *quantitative basis* for any of the key findings or outcomes.

What are the most likely outcomes of this assessment? One of the following:

a. **No Changes Needed.** Your current data architecture is fine, no changes are needed for the AI requirements that can be anticipated at present; in this case, you continue with your current direction, but plan to reassess this question annually as AI strategies progress; regrettably, I expect that most companies will be facing one of the other outcomes below;

b. **Optimization of the Current Architecture.** Your current data architecture is challenged, but some optimization and/or configuration changes will position you well for any currently known or expected AI requirements; in this case, you count your lucky stars; schedule the optimization projects; and, plan to reassess annually as AI strategies progress;

c. **An AI Engine is Needed.** Your current data architecture is not close to being able to handle the AI workloads and requirements that are clearly coming; in this case you want to launch a project to evaluate your options for bringing in an additional data engine and making the necessary data available to that data engine, possibly via open table formats (which will allow your existing data platform to access that data as well).

d. **Full Migration is Needed.** Your current data architecture is not close to being able to handle the AI workloads and requirements that are clearly coming and there are such serious problems with your existing data platform (e.g., it is too expensive, or it does not meet other key requirements for data and analytics) that you want to migrate your entire data estate to a new architecture or platform that will meet all data and AI requirements.

Each of these outcomes is a major finding and you don't want to be guessing. You want such a finding to have a sound quantitative basis, rooted in facts, measurements, tests and, if appropriate, models. Note that you cannot simply ask your vendor for the answers to these questions: the process must be conducted by a party that is not influenced by any vendor.

However, if you are able to conduct an appropriate assessment and develop a recommendation grounded in the AI strategies and based on informed projections,

About WinterCorp

WinterCorp is an independent consulting firm expert in the strategy, architecture and scalability of the modern analytic data ecosystem, including the application of AI/ML to enterprise data.

Since our founding in 1992, we have architected and engineered solutions to some of the toughest and most demanding analytic data challenges, worldwide.

We help customers define their data-related business interests and vision; develop their data strategies and architectures; select their data platforms; and, engineer their solutions to optimize business value. AI is an increasingly central aspect of the challenges, architectures and platforms that we deal with.

Our customers get business results with analytics and AI in which their return is often ten or more times their investment. Our unique expertise is grounded in our extensive history with large scale realistic testing, implementation and enterprise class operation.

When needed, we create and conduct benchmarks, proofs-of-concept, pilot programs and system engineering studies that help our clients manage profound technical risks, control costs and reach business goals.

With our in-depth knowledge and experience, we deliver unmatched insight into the issues that impede scalability and into the technologies and practices that enable business success.

WinterCorp
www.wintercorp.com
TYNGSBORO, MA
617-695-1800

facts and measurements, then you will have a sound basis for a revision of your data architecture. This can make the difference between failure and success with AI strategies across your enterprise.

RECOMMENDATION

IF YOUR COMPANY IS PLANNING OR IMPLEMENTING major AI strategies that will rely on enterprise data or documents, then you must take a fresh look at your data architecture.

Chances are that your data architecture has been developed for BI, not the coming world of AI. AI will stress your data architecture in ways it has not been stressed before. Applying a metaphor from the movie, *Jaws*, it is likely that you "are gonna need a bigger boat."

Even though all cloud data platforms are *elastic* — and therefore can add capacity — they are not all *scalable* in the presence of large, complex and dynamic workloads of the sort we will see with AI. Your platform and your database architecture may or may not be up to the demands just ahead, as AI is made available in progressively larger programs throughout your enterprise.

Don't wait. AI is almost certainly in the works at your company. Before you find yourself with AI programs failing because the data architecture cannot support them, *I urge you to conduct an assessment such as I describe in this note*. For reference material, resources and/or services to help you complete that assessment, contact WinterCorp.