BECOME A GEN AI LEADER:

Learn the Playbook that Propels Organizations

OBJECT COMPUTING

Please type your questions into the chat at any time. We will answer them near the end.

This webinar is being recorded.

We will send an email with the presentation after the webinar.

Our Team

Andrew Montgomery
Vice President
of Strategy

Dr. Yaqi Chen
AI/ML & Data Insights Area Lead,
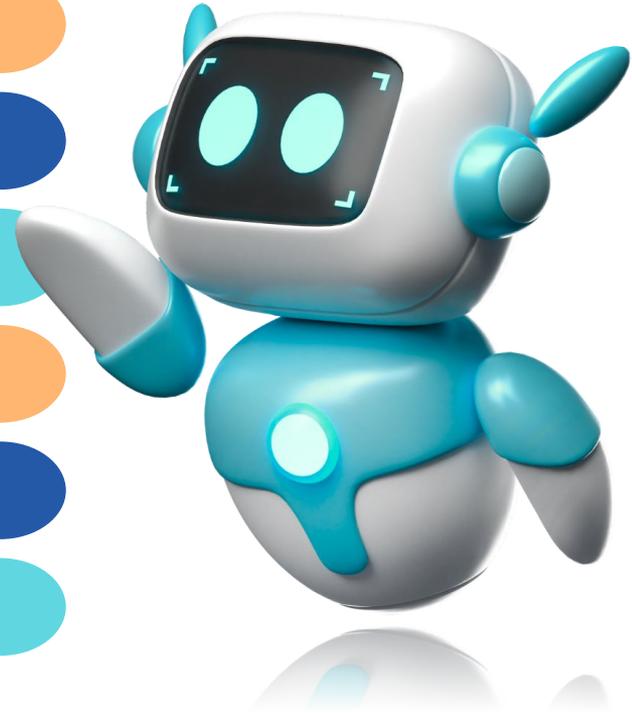Distinguished Scientist

Jimmy Moore
Data Engineering Team Lead,
Principal Data Engineer

Agenda

- Who is Object Computing?

- The Evolution of Gen AI

- Our Playbook Approach to Gen AI

- The Problem and Why a Gen AI Solution

- Challenges, Limitations & Considerations of Gen AI

- The Future Powered by Gen AI

- Q&A

OBJECT COMPUTING
YOUR OUTCOMES ENGINEERED

TECHNOLOGY THAT EMPOWERS

We are a modern consulting company that builds innovative, sustainable, and impactful systems. We specialize in software engineering, AI, machine learning, DLT/blockchain, and applications development.
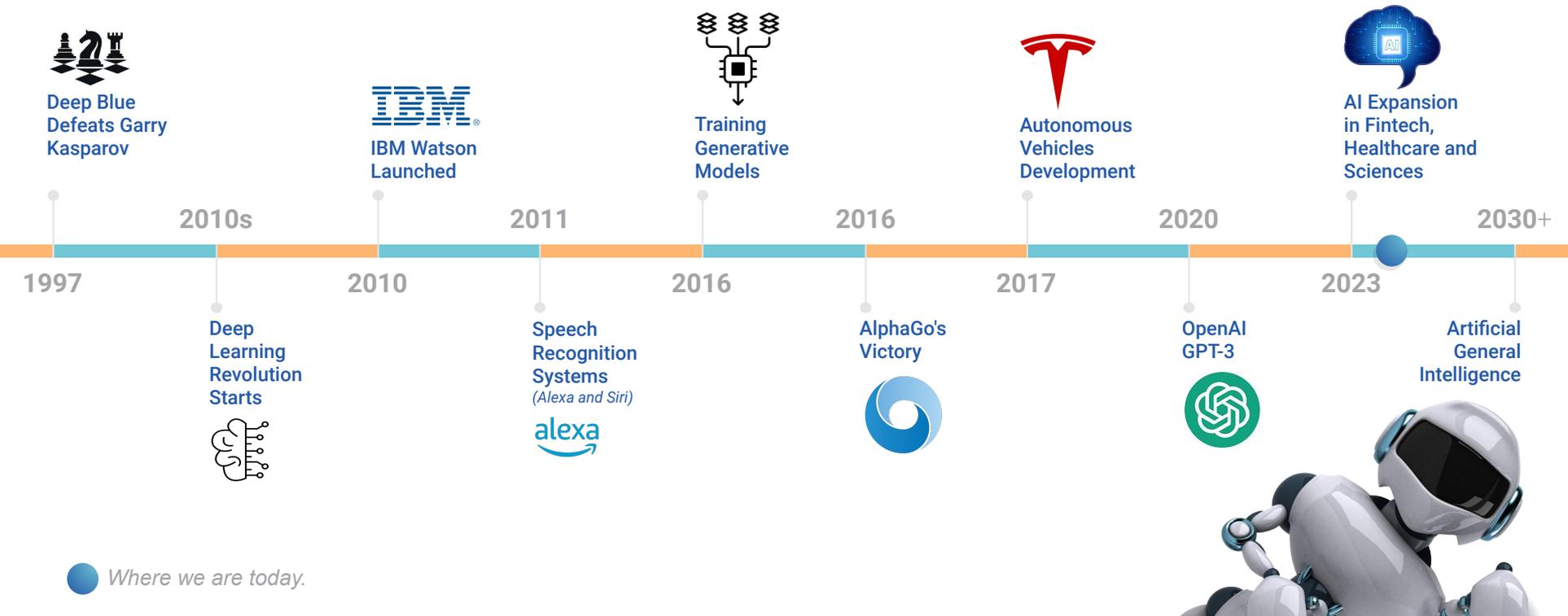
OUR EXPERIENCE

For 30 years, we've excelled at solving large-scale business problems in nearly every area of software engineering.

1993	1995	2000	2002	2004	2007	2008	2009
Washington University	Iridium	Express Scripts	BAE Systems	Bloomberg	Aerospace	DARPA	Scottrade

2011	2016	2017	2017	2019	2020	2022	2023
ESPN	Monsanto	Block.one	NASA	Kansas City Southern Railway	Pet Care	Tolam Earth	Google Earth Engine

AI/ML · Cloud · DLT/Blockchain · IoT · Security · App Dev

OUR PARTNERS

aws partner network

Google Cloud Partner

Microsoft Solutions Partner

Evolution of AI into Generative AI and the Future

Deep Blue Defeats Garry Kasparov

IBM Watson Launched

Training Generative Models

Autonomous Vehicles Development

AI Expansion in Fintech, Healthcare and Sciences

2010s   **2011**   **2016**   **2020**   **2030+**

1997   **2010**   **2016**   **2017**   **2023**

Deep Learning Revolution Starts

Speech Recognition Systems
(Alexa and Siri)

AlphaGo's Victory

OpenAI GPT-3

Artificial General Intelligence

Where we are today.

objectcomputing.com

Why GenAI Playbook?

A comprehensive guide and recipe that outlines **standardized procedures**, **best practices**, **methodologies**, **tools**, and **frameworks** for developing, deploying, and managing GenAI solutions.

1 Strategic Alignment

2 Standardization & Consistency

3 Efficiency & Speed

4 Scalability

5 Risk Mitigation
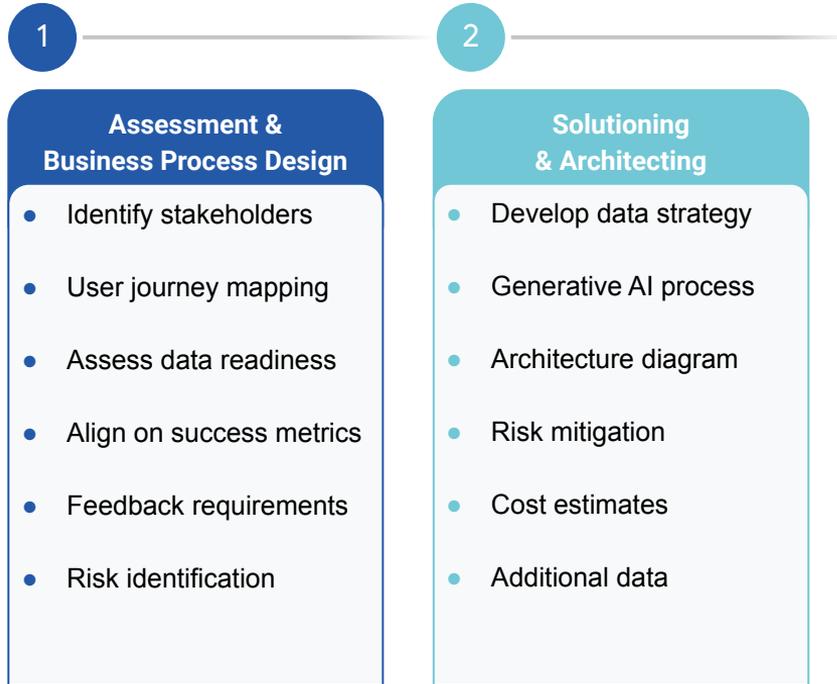
Generative AI Engagement Playbook Approach
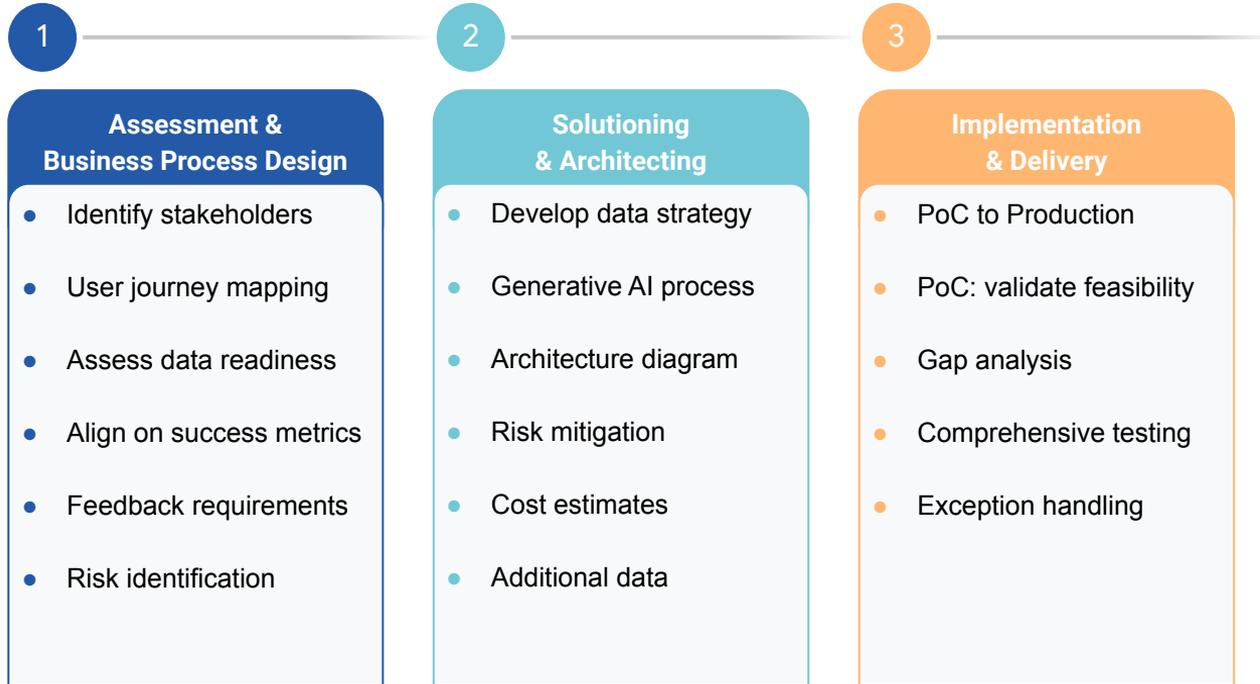
1

Assessment & Business Process Design

- Identify stakeholders

- User journey mapping

- Assess data readiness

- Align on success metrics

- Feedback requirements

- Risk identification

objectcomputing.com

Generative AI Engagement Playbook Approach

1

Assessment & Business Process Design

- Identify stakeholders
- User journey mapping
- Assess data readiness
- Align on success metrics
- Feedback requirements
- Risk identification

2

Solutioning & Architecting

- Develop data strategy
- Generative AI process
- Architecture diagram
- Risk mitigation
- Cost estimates
- Additional data

Generative AI Engagement Playbook Approach

1

Assessment & Business Process Design

- Identify stakeholders
- User journey mapping
- Assess data readiness
- Align on success metrics
- Feedback requirements
- Risk identification

2

Solutioning & Architecting

- Develop data strategy
- Generative AI process
- Architecture diagram
- Risk mitigation
- Cost estimates
- Additional data

3

Implementation & Delivery

- PoC to Production
- PoC: validate feasibility
- Gap analysis
- Comprehensive testing
- Exception handling

objectcomputing.com

Use Case:
Manufacturer/Distributor

Data Challenges in Manufacturer Decision-making

Data Heterogeneity

Varying Formats, Structures & Types

Data Volume and Velocity

High Volume and High Velocity

Data Quality and Consistency

Errors, Missings and Duplications

Security and Compliance

Data Privacy and Regulations

objectcomputing.com

Current Solutions Rely Heavily on Data Platforms and Middleware

Data Lake/Data Warehouse

Middleware and APIs

Diverse Data Sources

Master Data Management

objectcomputing.com

Current Solutions Rely Heavily on Data Platforms and Middleware

Bespoke ETL

Data Lake/Data Warehouse

Brittle Parsers

Middleware and APIs

Cloud Agnostic Strategies

…

Diverse Data Sources

Master Data Management

Current Solutions Rely Heavily on Data Platforms and Middleware

Bespoke ETL

Data Lake/Data Warehouse

Brittle Parsers

Middleware and APIs

Cloud Agnostic Strategies

…

Diverse Data Sources

Master Data Management

Why LLMs?

Broad Question Coverage

RAGs, Agent and Tooling

Handle Diverse Data

Intrinsic semantic capabilities

Cost Efficiency

Robust and Built-in Solutions

Enhanced Maintenance

Security, Clear Errors, Updating

objectcomputing.com

Challenges: Security Considerations & Importance

Considerations from OWASP Top 10 for LLM Applications

LLM01
Prompt Injection

LLM04
Model Denial of Service

LLM06
Sensitive Data Exposure

LLM08
Excessive Agency

Our security webinars are a great place to learn more:

objectcomputing.com

Challenge: Diverse Data Formats and Data Structures Across Manufacturers

Require a Unified Approach to Extract Error Information out of Different Structures and Key Names
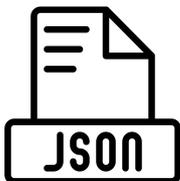
```
{"errors": [{"E50": "POWER_FAILURE"}, ...]}
{"errorList": [{"POW": "CRITICAL", ...]}
{"machine_errors": {"ERROR_A": 3, "ERROR_B": 2, ...}}
```

```
<errors><error id=1><error id=3></errors>
```

```
2024-06-25T16:00:00.0Z [ERROR] ID=53 POWERFAIL
```

How to Address: 4 Approaches

JSON Agent

RAG

Retrieval Augmented Generation (RAG)

Full Context

Generative Tools

objectcomputing.com

JSON Agent

Solution

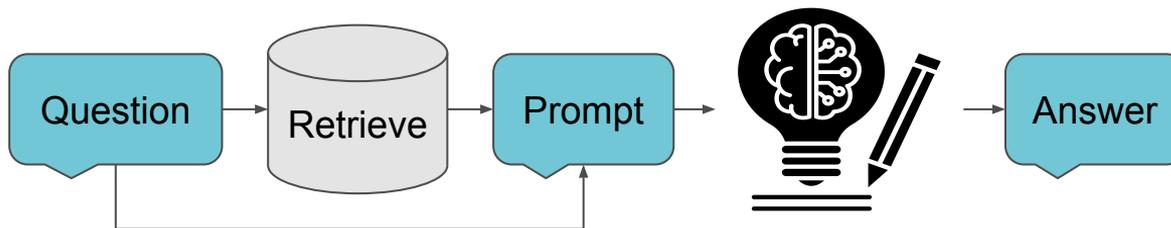Provides the ability to iteratively traverse JSON files without the entire file being in context

Con

◆ Each search is "nearsighted"
◆ Limited applications

"What are the required parameters in the request body to the /completions endpoint?"

LLM → ["model"]

Answer found!

Get Value

List Keys

Retrieval Augmented Generation (RAG)

Question → Retrieve → Prompt → 🧠 → Answer
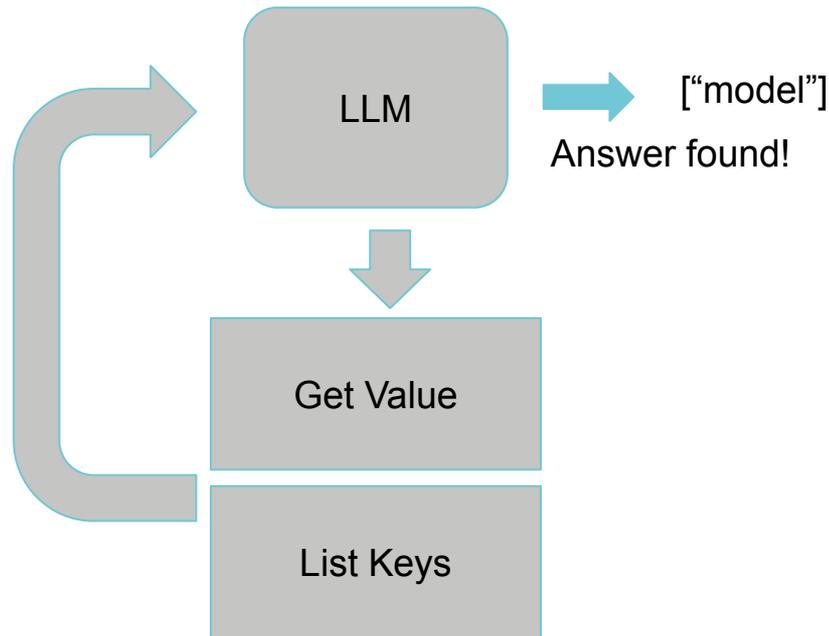
Solution

Retrieve data chunks based on similarity to query

Con

◆ Requires data preprocessing steps (chunking, embedding, etc)
◆ Hallucinations while minimized may still occur
◆ Chunk and overlap sizing needs tuning

objectcomputing.com

Full Context

Solution
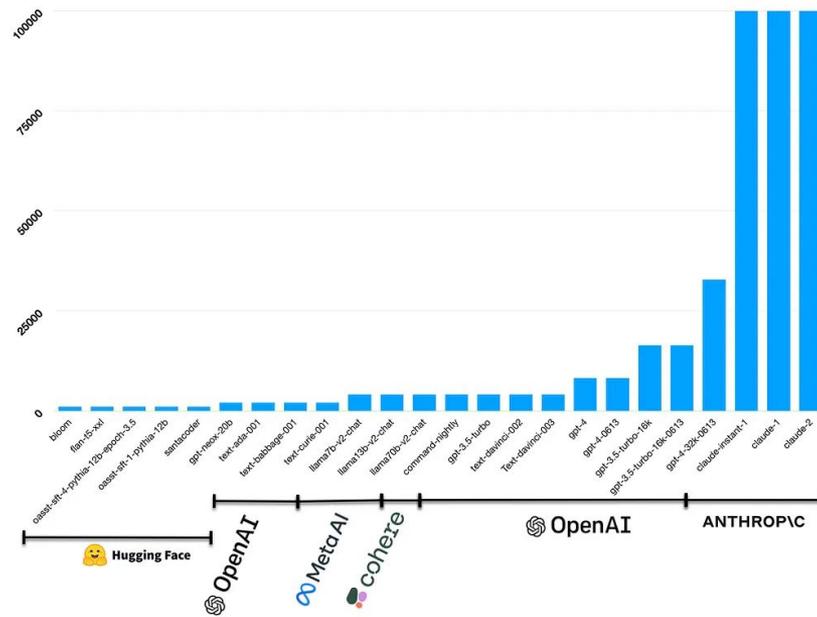Put the entire document in context

Pro
The model has everything in its sights to utilize in answering your questions.
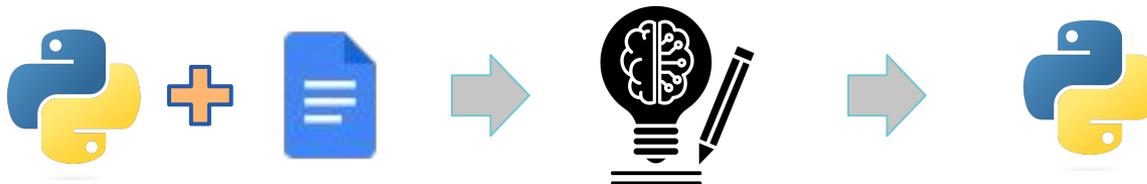
Con
Large token usage == $$ requires many iterations if comparing multiple JSON files (1 JSON file in context at a time).

LLMs Context Size

Generative Tools

Solution
Provide a standard abstract base class and have an LLM write the child class for you for each new file format
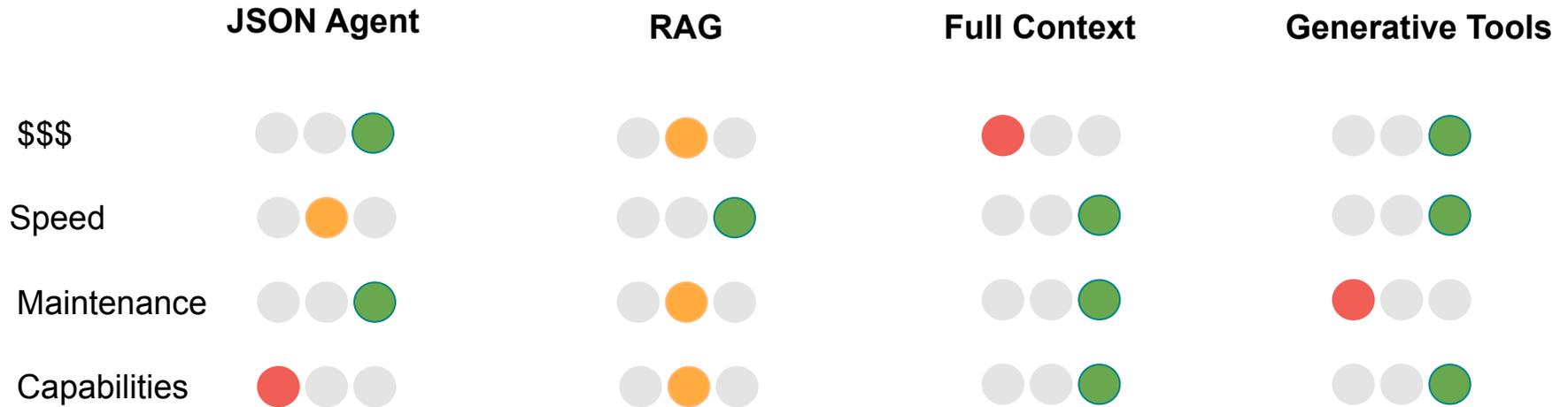
Pro
Flexibility of different data formats/audit types

Con
Would require a strong test suite to be successful, more code to maintain

Approach Comparison Overview

	JSON Agent	RAG	Full Context	Generative Tools
$$$	🟢 (high)	🟠 (mid)	🔴 (low)	🟢 (high)
Speed	🟠 (mid)	🟢 (high)	🟢 (high)	🟢 (high)
Maintenance	🟢 (high)	🟠 (mid)	🟢 (high)	🔴 (low)
Capabilities	🔴 (low)	🟠 (mid)	🟢 (high)	🟢 (high)

When to use

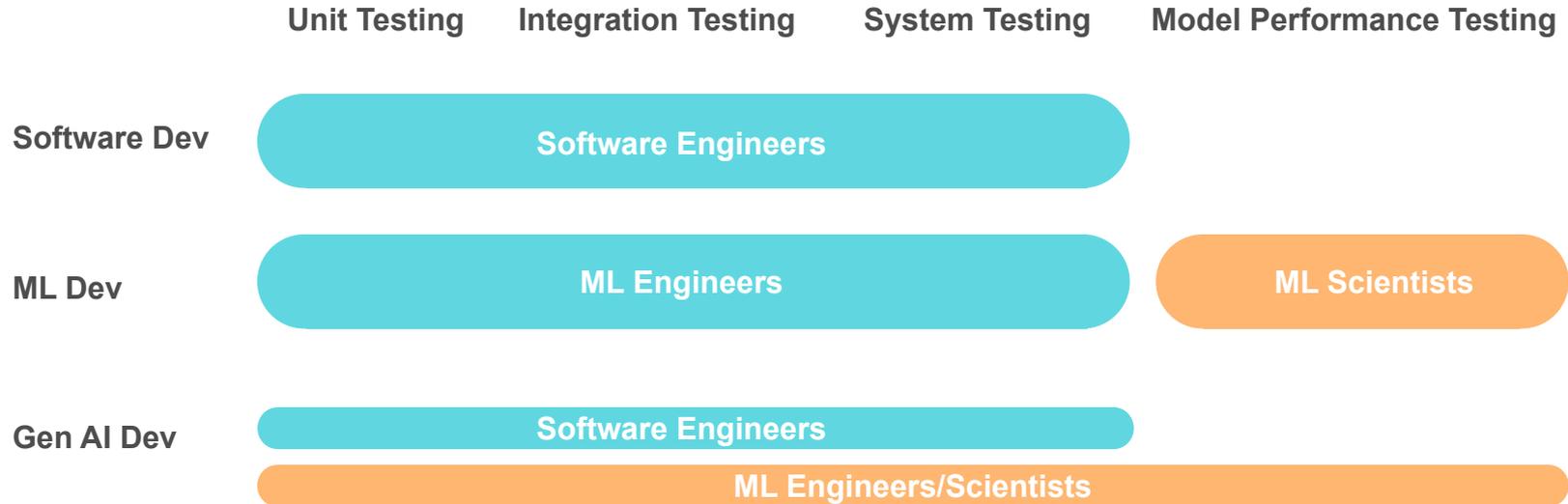|	
JSON Agent	Semantically relevant keys with a relatively shallow structure to retrieve specific information
RAG	The information is within a segment of data (find the needle in the haystack) and you need faster results than JSON agent
Full Context	Quick start to see what's possible or complex questions with low volume
Generative Tools	Consistent outputs, summarizations/calculations/transformations, LLM as an orchestrator that chains together tools and organizes the results

objectcomputing.com

Challenges - Extensive Testing/Test Coverage

Gen AI Dev requires more extensive testing across software engs, ML engs and scientists

	Unit Testing	Integration Testing	System Testing	Model Performance Testing
Software Dev	Software Engineers			
ML Dev	ML Engineers			ML Scientists
Gen AI Dev	Software Engineers / ML Engineers/Scientists			

objectcomputing.com

From Today to Future: Hyper-automation in Data Powered by LLMs

Data Access

- Structured Method
- Schema Dependency

- Natural Language Queries
- Contextual Understanding

ETL Process

- Manual Scripting
- Fixed Pipeline
- Complex Transformations

- Automated Extraction
- Dynamic Transformation
- Adaptive Pipeline

Maintenance

- Maintenance overhead

- Reduced Overhead

Today's State-of-the-Art

Future's Hyper-automation

Q&A

objectcomputing.com

THANK YOU!

LET'S CONNECT

Matt Bremehr
Sales Director
Object Computing Inc.

BremehrM@objectcomputing.com

ObjectComputing.com

28

objectcomputing.com