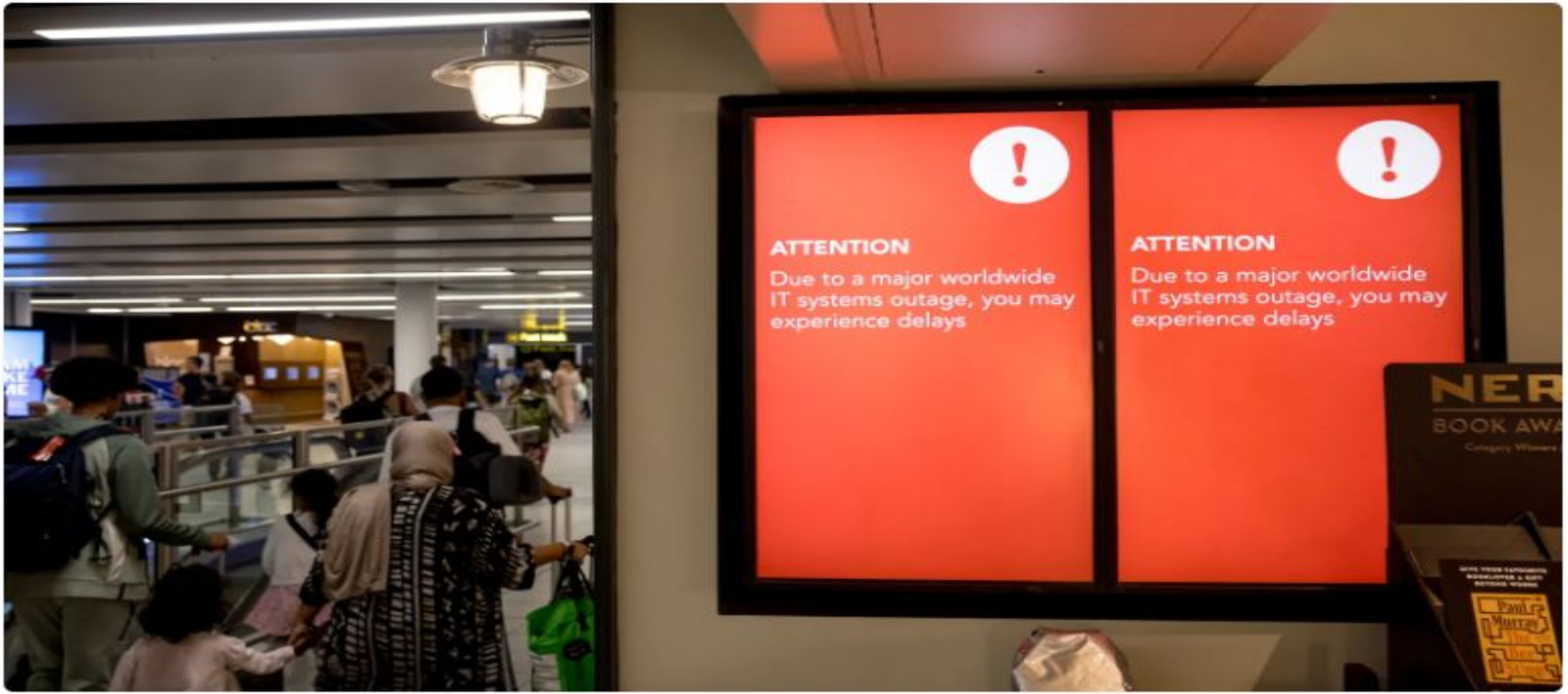
Evaluating Change Requests to Predict and Avoid Future Incidents

Team Name: Snowflakes
John Heisler
Suresh Vadakath
Spencer Ellingson

Where were you?



Problem Statement

Financial institutions face significant operational challenges due to frequent outages and incidents linked to IT modernization efforts. Every week hundreds (maybe thousands??) of change requests are implemented across organizations to enhance the technological infrastructure and improve overall service delivery.

At any given time any one of these change requests can have inadvertent (significant) negative consequences including:

- Operational Disruptions
- Financial Losses
- Reputational Damage
- Increased Recovery Costs
- Data Loss or Corruption
- Regulatory Compliance Risks
- Customer Attrition
- Employee Morale
- Increase Risk Exposure

Solution Overview

Solution Approach: Preventative Measures

Two of the main causes of outages are incremental updates to systems and human errors (1). We aim to mitigate the exposure to these causes, which often combine, to reduce overall outages.

Rather than mitigate already existing problems (which is another approach–see future work), we aim to prevent changes that would cause an outage from ever entering the environment.

Our Solution: Project Harbinger - AI-Enabled Outage Prevention

To this end, we leverage AI to identify patterns in data and oversights in development that may have otherwise gone unnoticed. With that insight, we flag the risky changes for for further review to significantly reduce outages due to change requests.

Key Features: Elegance, Speed to Market, Flexibility, Minimal Overhead

Collapsing up to dozens of run times, our data warehouse, data engineering, model inference, and front end are all in a single Snowflake runtime. Our solution can be integrated into any change release process with minimal customization and near zero maintenance overhead. Lastly, the solution allows users to switch between dozens of LLMs at the click of a button to optimize outcomes.

Organizational and Industry Impact

Asymmetrical & Scalable Positive Impact

The solution has the potential to reduce financial impact of IT outages by millions of dollars for IT organizations annually, can scale to any company with access to LLMs, and extremely low costs at a tiny fraction of the savings.

Hypothetical Single Use Case

260 changes per year (5 changes per week)
20 outages caused by changes
$5,500/min Cost of each minute of downtime
90 Minutes Average outage duration
50% Reduction in outages

Annual Cost

Outage Time: 1,800m *(90m * 20 outages)*
Cost: $9,900,000 *(1,800 * 5,500)*

Annual Savings (50% reduction)

$4,950,000 Savings

Industry Impact, Collaboration and Production Viability

Potential Path to Production Granted the user has access to change request data and Snowflake, **this solution can be deployed immediately** as is in their environment.

Scalability The scalability of this solution is predicated on the context window of the models used for inference. When run in Snowflake, the TCO will remain very low and access to models is out of the box.

Industry Partner Collaboration Our solution leverages a language models from a wide variety of industry leaders like Anthropic, Meta, Mistral, RekaAI, and Jamba. This “model garden” affords end users ultimate flexibility in output with no overhead.

Challenges and Limitations we see two potential challenges

- Should the enterprise not have access to Snowflake, the users would need to alter the inference to call LLM end points (minimal effort).
- If the users do not have access to their change request data or the data is not clean, users would need to perform the data engineering to make it accessible in a clean state.

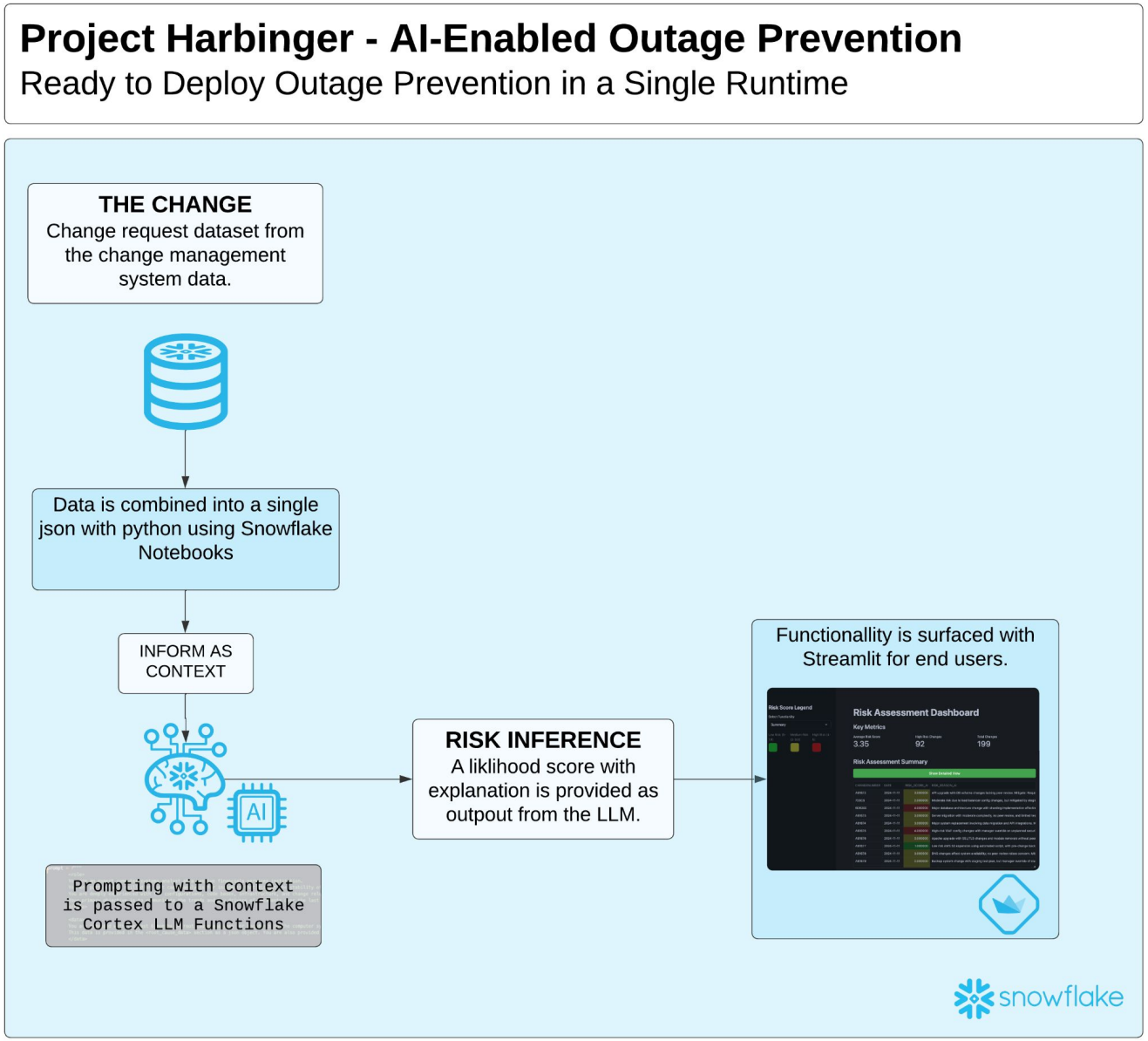
Technical Implementation

Tech Stack Overview: We used Python and SQL only to build our solution. All development was done directly in Snowflake using notebooks.

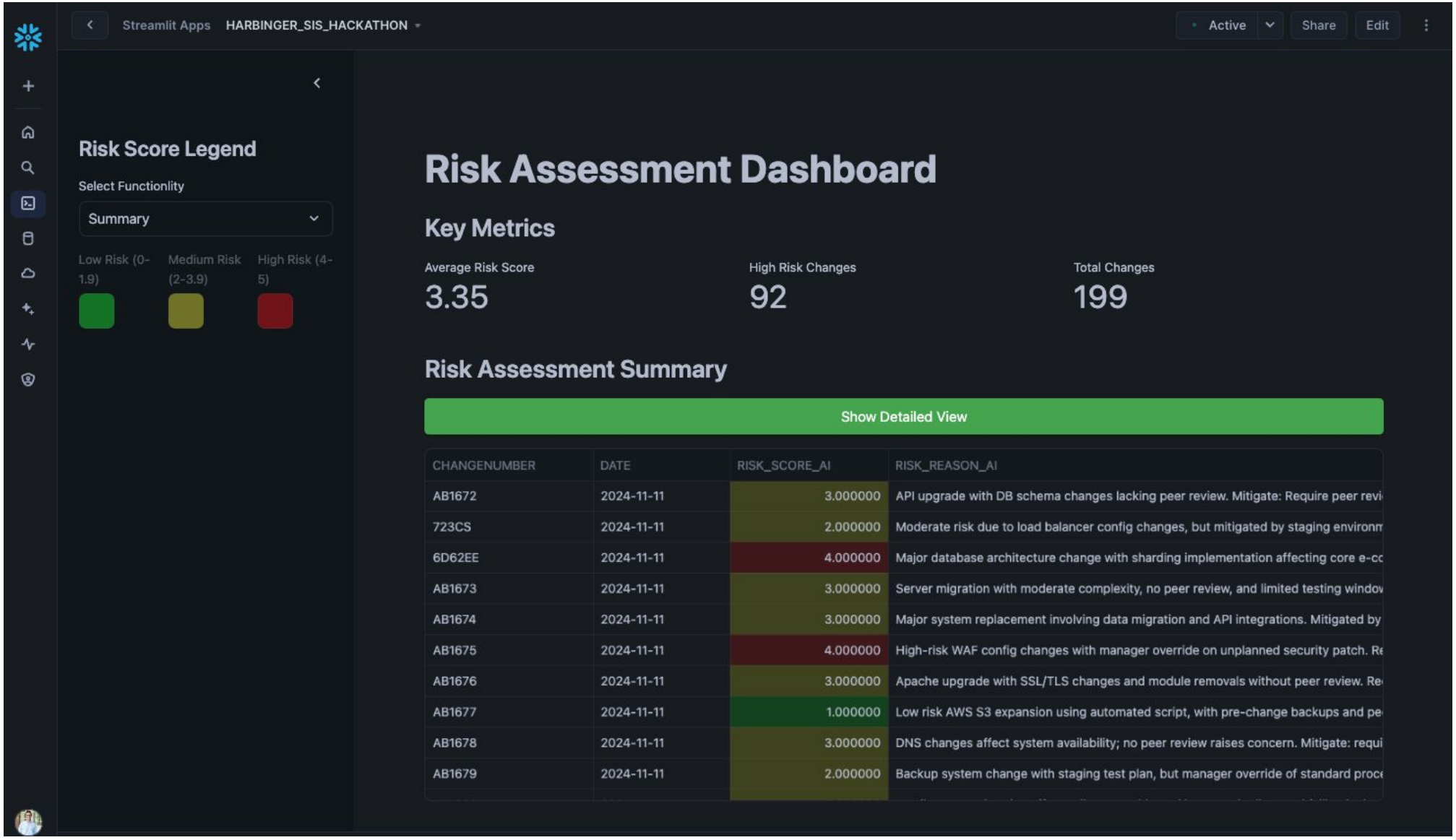
Data Usage: We used Synthetic Data generated with methods shared in our solution.

Model Performance and Results: While we see extremely encouraging results based on the synthetic data. Without true historical data to test upon, performance metrics are difficult to quantify. However, in early pilots of this exact functionality on real historical data, ***we have reliably seen 80%+ accuracy in identifying change requests that had caused incidents.***

Architecture Diagram



Demo



Ubiquitous Industry and Cross-Industry Value

At the Spoke

When deployed at individual institutions our solution maximizes uptime allowing the business to deliver their products and services.

At the Hub

When deployed at central hubs of the financial services industry, like DTCC, the positive impact is shared across the entire industry– ***a compounding positive impact.***

Minimal TCO

On top of the universal value, the solution's elegance makes it extremely cost effective to deploy and own.

Next Steps

Early Viability Signs and Interest

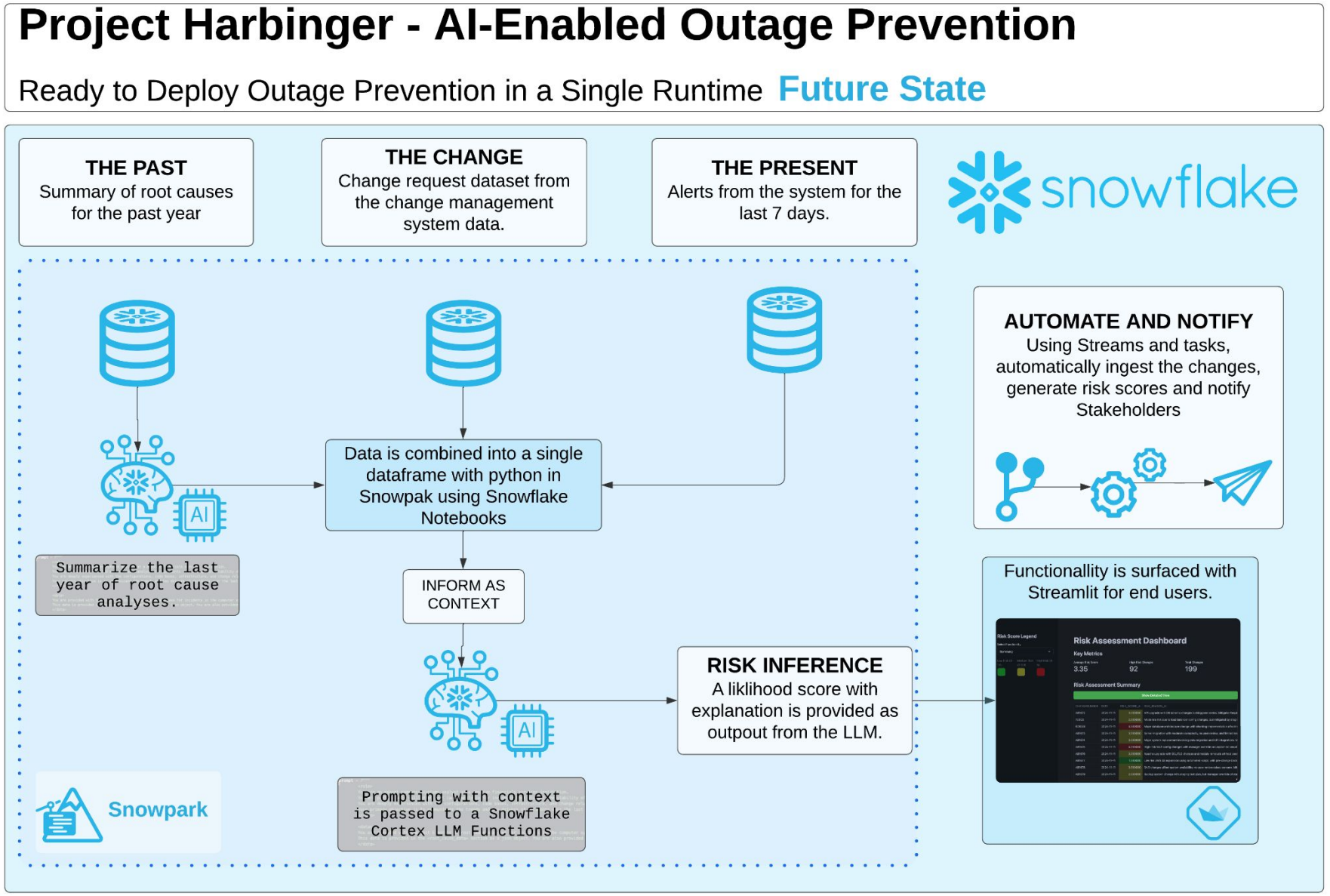
In pilots on real historical data, we've received enthusiastic executive engagement as we have ***seen 80%+ accuracy identifying changes that had resulted in outages.***

Future State & Active Development

We have detailed a future state architecture which proposes automation and broadened data inputs to further enhance this solution which we are actively working on.

We are also excited to begin work accelerating outage root cause analysis with similar AI techniques to attack this problems from both sides.

Architecture Diagram - Future State



Appendix

Judging Criteria and Scoring

	Metric	Title	Maximum Score
Technical Innovation 	**Innovation and Originality of Solution Design**	Does the solution have a creative approach to solve the problem? Was AI used effectively to derive the solution? Judges will evaluate the uniqueness of the solution, the creative approach to solving the use case.	**15**
	Technical Implementation	The quality and sophistication of the solution's architecture, design, tech stack, industry frameworks and efficiency of the solution.	**10**
	Solution Viability	Solution's potential for operationalization and enterprise use.	**10**
	Data Usage and Management	The effective use of provided datasets or external data sources, data preprocessing and feature engineering techniques, and handling of data privacy and security concerns will be evaluated.	**10**
	Ease of Use	Does it have a user-friendly interface? Can a user use the product with minimum instructions.	**5**
Value Proposition 	**Industry Impact and Viability**	Potential value to the Capital Markets industry broadly, market scalability, and revenue generation or cost-saving potential of the solution.	**20**
	Industry Collaboration *(Bonus for "super teams")*	Does the implementation (or future phases of the implementation) facilitate/require collaboration across at least 2 or more industry participants? (e.g., through data sharing or other methodologies).	**15**
Final Presentation and Demo 	**Clear Presentation of Problem and Solution**	A clear explanation of the problem and solution, the quality of the pitch or demo, and technical design documentation of the project will be important factors.	**15**
	Effective handling of Q&A		
	Technical Design Documentation		
	Bonus- visual appeal of presentation material		