



Nimble Clinical Research

AI-powered Clinical Data Solutions

Synergizing Large Language Models (LLMs) and
Human-in-the-Loop (HITL) for Data Noise Reduction
in the RBQM approach

Presented by

Mukesh Babu K

Head of Clinical Analytics
Nimble Clinical Research, India

Vineet Jain

Founder & CEO
Nimble Clinical Research, USA





Agenda



01 AI in RBQM

02 Signal Journey

03 Signal Validation & Noise Example

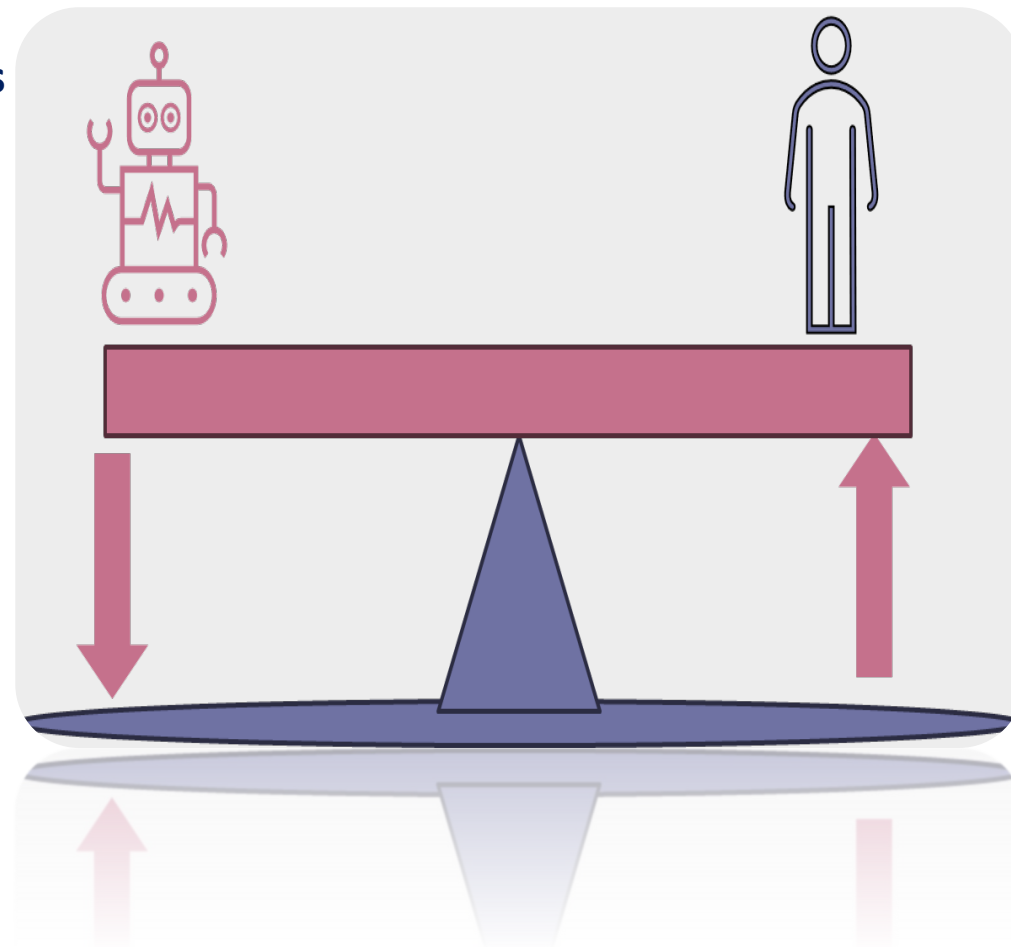
04 Our LLM Approach to Signal Validation

05 LLM Context Building Examples

06 Integrated LLM with Workflow

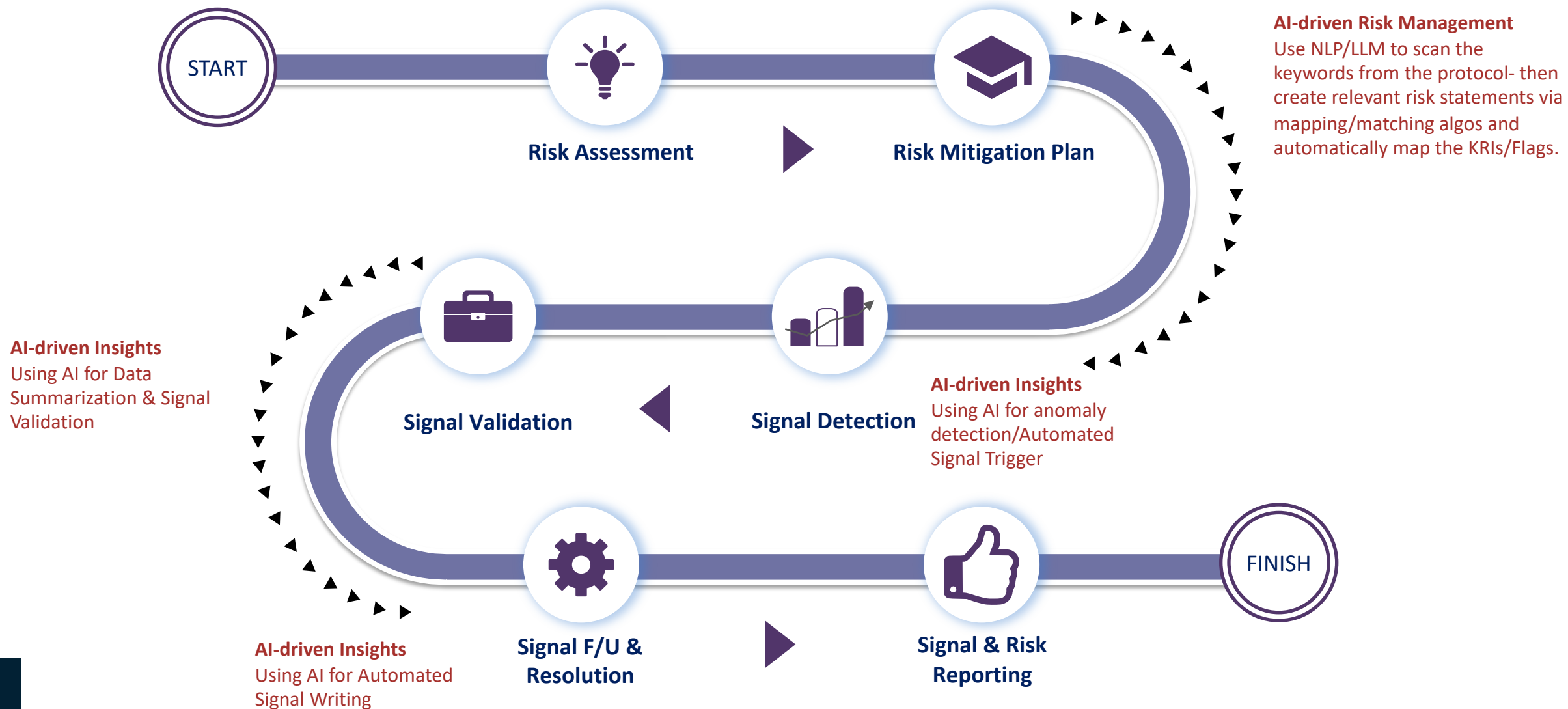
07 Integrated LLM to Analytics Applications- general approach

08 Key Take Aways





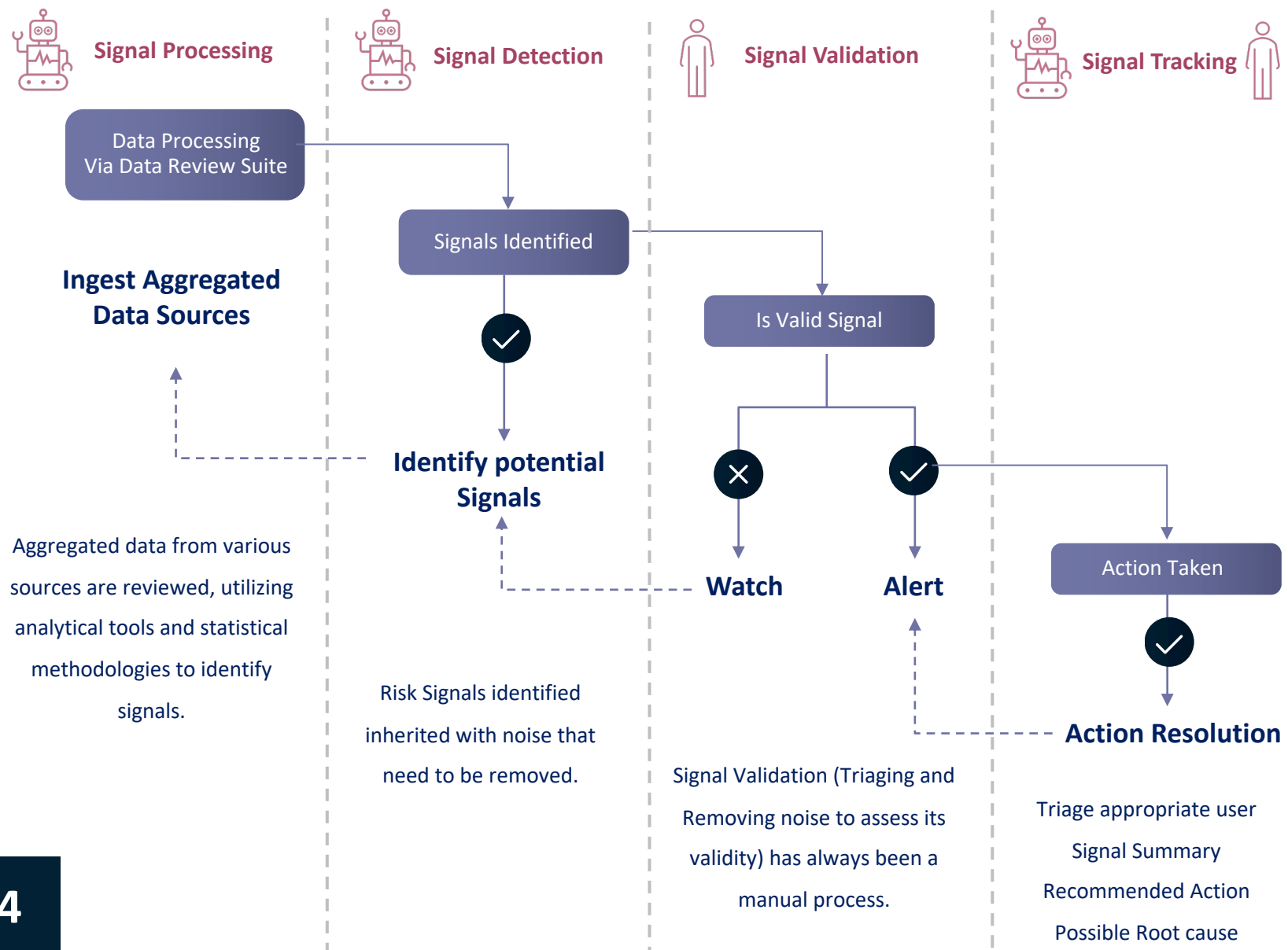
AI in RBQM





Signal Journey

Current Process and Challenges



Signal Validation Key Challenges

1. Noisy Signal Impact on Action

Resources Time & Efforts, Question the Signal Detection Tools

2. Pre-defined Filters for Noise Removal

Emerging risks may be masked and inadvertently remove valid signals

3. Resource-Intensive Context Building

Manual review of multiple data domains to substantiate specific and Tailored actions

4. Subjectivity and Bias

Interpretations of human reviewers may introduce subjectivity into the evaluation process.



Example of Noise and Impact on Action

Dashboard

Patient	N of Visits	Risk Rank	Risk Score	PE Compliance	LB Compliance	VS Compliance	EG Compliance	MH Rate	AE Rate
NIM-55-22-1003-10470	13	4.0	15.0	●	●	●	●	●	●

Subject Rate: 1.6
N of AEs: 21
N of Visits Completed: 13
Risk Level: High
Risk Flag: High Reporting

Signal Validation

Checking for Duplicated AEs



USUBJID	AESEQ	AELT	AESER	AEACN	AEREL	AEOU
NIM-55-22-1003-10470	1	Constipation	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED
NIM-55-22-1003-10470	2	Headache	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED
NIM-55-22-1003-10470	3	Fever	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED
NIM-55-22-1003-10470	4	Constipation	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED
NIM-55-22-1003-10470	7	Dehydration	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	8	Nausea	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	9	Vomiting	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	10	Dehydration	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	11	Dehydration	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	12	Chills	Y	NOT APPLICABLE	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	13	Dizziness	N	DOSE NOT CHANGED	UNRELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	14	Abdominal Pain	N	DOSE NOT CHANGED	UNRELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	15	Dizziness	N	DOSE NOT CHANGED	UNRELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	17	Migraine	N	DOSE NOT CHANGED	UNRELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	19	Nausea	N	DOSE NOT CHANGED	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	23	Fatigue	N	DOSE NOT CHANGED	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	24	Fatigue	N	DOSE NOT CHANGED	RELATED	NOT RECOVERED/NOT RESOLVED
NIM-55-22-1003-10470	25	Hypotension	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED WITH SEQUELAE
NIM-55-22-1003-10470	27	AST increased	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED WITH SEQUELAE
NIM-55-22-1003-10470	29	Allergy	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED WITH SEQUELAE
NIM-55-22-1003-10470	30	Hypotension	N	DOSE NOT CHANGED	UNRELATED	RECOVERED/RESOLVED WITH SEQUELAE

Preexisting Medical Histories



Reported Term for the Medical History	Dictionary-Derived Term	High Level Term
QT PROLONGATION	Electrocardiogram QT prolonged	ECG investigations
SENSORY NEUROPATHY (FEET BILATERALLY)	Peripheral sensory neuropathy	Peripheral neuropathies NEC

Checking for Overlapping Ongoing Medical History reported as AE.



Reported Term for the Medical History	Dictionary-Derived Term	High Level Term
QT PROLONGATION	Electrocardiogram QT prolonged	ECG investigations
SENSORY NEUROPATHY (FEET BILATERALLY)	Peripheral sensory neuropathy	Peripheral neuropathies NEC

Prior/Ongoing Concomitant Medications



Category for Medication	Dose per Administration	Route of Administration	Start Date/Time of Medication	Study Day of Start of Medication	ATC Level 1 Term	CM Indication Identifier	Subject Study Site Identifier for the Study	CMENRF	CMINDC
Concomitant Medications	1.0	ORAL	2018-08-06	-543.0	ALIMENTARY TRACT AND METABOLISM	Digestive disorder	1003.0 1003-10470		ADVERSE EVENT
Concomitant Medications	1.0	ORAL	2019-02-24	-341.0	ALIMENTARY TRACT AND METABOLISM	Digestive disorder	1003.0 1003-10470		MEDICAL HISTORY
Concomitant Medications	10.0	INTRAVENOUS	2017-10-01	-852.0	BLOOD AND BLOOD FORMING ORGANS	Infusion arm pain	1003.0 1003-10470		OTHER

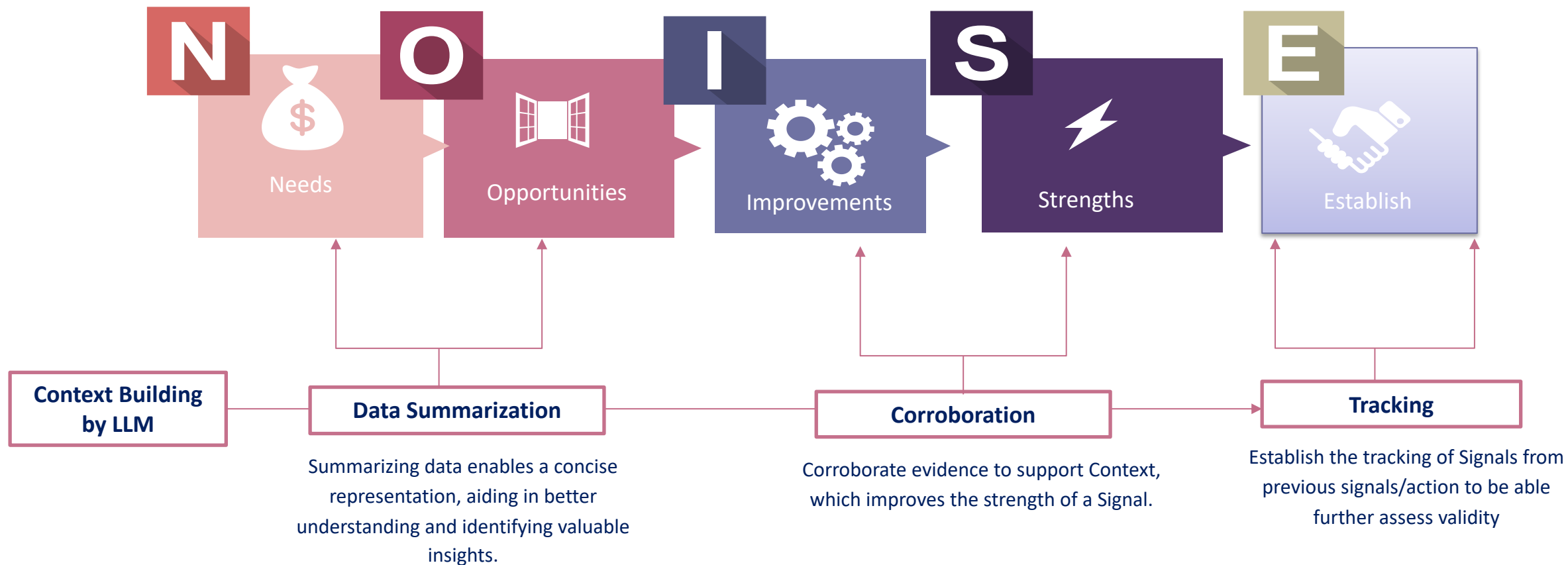
Direct escalation (without Validation)

Lacking the Signal Validation Step would result in stakeholders expending "unnecessary" time and effort deciphering "What needs to be done."

None of the specified criteria support the assertion of AE High Reporting. It appears to be a potential "false positive" stemming from the distribution of AEs. Upon investigation, it is noted that the Study-level AE rate is <0.5. Consequently, numerous patients do not experience sufficient AEs, leading to the classification of some patients as High reporting without necessitating further action.



Our Approach: Context building by LLMs for Noise Reduction



Uncovered unexplored territory in the field by leveraging LLMs, specifically exploring context building through data summarization for signal validation. This innovative approach aims to identify substantial evidence around signals, providing a promising avenue for improvement.



Context Building by Data Summarization



LLM prompt

Create a summary of the following site-level demographics and Adverse event aggregations.

Subject Demographics: Age Groups: 18-30: 25% 31-45: 35% 46-60: 20% 60+: 20% Gender Ratio: Male: 40% Female: 60%.

Pre-existing Conditions: Diabetes: 15% Hypertension: 20% cardiovascular diseases: 10%

Frequency of AEs: Nausea: 30% Headache: 25% Dizziness: 20% Fatigue: 15% Rash: 10%

Severity of AEs: Mild: 40% Moderate: 35% Severe: 25%

AE Rate at This Site: 5 Average (SD) AE Rate Across Other Sites: 3 (0.5)

CM Rate at This Site: 11 Average (SD) AE Rate Across Other Sites: 8 (1.4)

MH Rate

Create one summary paragraph for Subject Demographics, Adverse Event Patterns, and site-specific Rate Comparison.

Signal Validation through Summarization

- Concise summarization of data from the high volume of data
- Summarize multi-source data from diverse critical data points such as demographics, AE patterns, and medications together

LLM Response

Subject Demographics: The demographic profile at this site is predominantly young and female, with the majority being under 45 years of age. Notably, a significant portion of subjects have pre-existing conditions like hypertension and diabetes, which are common in about one-fifth and one-sixth of the subjects, respectively.

Adverse Event Patterns: Adverse events reported predominantly include nausea and headache, with other significant occurrences like dizziness and fatigue. The severity of these events varies, but there is a considerable incidence of moderate to severe cases.

Site-Specific Rate Comparison: The site exhibits a higher AE rate than the average across other sites. Similarly, the Concomitant Modification Rate is elevated here, indicating a higher frequency and potential severity of clinical modifications in response to AEs at this site than others in the study.



Context Building by Corroboration



LLM prompt

Determine if the high Adverse Event (AE) rate at a specific site is valid or noise by analysing the following key co-related metrics and comparing them to averages across other sites:

AE Rate: Site: 4.5 Other Sites Average: 3.0 (SD: 0.7)

MH Rate: Site: 3.8 Other Sites Average: 2.5 (SD: 0.6)

CM Rate: Site: 2.6 Other Sites Average: 1.9 (SD: 0.5)

AE Duplication Rate: Site: 0.5 Other Sites Average: 0.3 (SD: 0.1)

Data Entry Delay (days): Site: 2 Other Sites Average: 1.5 (SD: 0.4)

Average Exposure (days): Site: 80 Other Sites Average: 115 (SD: 8)

Usually, a high AE rate, accompanied by a high CM rate, MH rate, AE duplication rate, and. exposure, and low data entry delay, justifies a high AE rate. Assess if the site's high AE rate is an outlier considering these co-related metrics. Present the conclusion in 50-100 words, followed by a detailed analysis.

Context Building by Corroboration

- Site Level: Interpret complex data relationships across several metrics like AE rate, CM rate, EX duration, and MH rates, using detailed prompts with cross-relationships defined
- Subject Level: Link records across medical history, medications, and adverse events to identify patterns

Conclusion: The high AE rate at the site appears justified and not an outlier when considering the correlated metrics. Elevated MH, CM, and AE duplication rates and a slightly increased data entry delay align with the higher AE rate. However, this site's notably lower average exposure diverges from the typical pattern, suggesting that other factors might influence the high AE rate.

Detailed Analysis:

AE Rate: The site's AE rate of 4.5 is significantly above the average of 3.0 at other sites. The deviation is notable, considering the standard deviation 0.7 at other sites.

MH Rate: The MH rate at the site is 3.8, much higher than the 2.5 average at other sites. Given its standard deviation of 0.6, this elevated rate suggests a higher prevalence of medical history issues, which could contribute to the increased AE rate.

CM Rate: With a CM rate of 2.6 at the site versus an average of 1.9 at other sites, and considering the standard deviation of 0.5, there is a clear increase in clinical modifications, likely in response to more frequent AEs.

AE Duplication Rate: The site's AE duplication rate of 0.5, higher than the 0.3 average at other sites (standard deviation of 0.1), indicates more instances of repeated AE reporting, contributing to the higher overall AE rate.

Data Entry Delay: The site's average data entry delay of 2 days, slightly above the 1.5-day average at other sites (standard deviation of 0.4), points to marginally slower reporting, which could affect AE rate calculations.

Average Exposure (Days): The average exposure of 80 days at the site is significantly lower than the 115-day average at other sites (standard deviation of 8). This lower exposure duration contradicts the expectation that a higher AE rate correlates with longer exposure.

LLM Response



Context building by Tracking



LLM prompt

Review the progress of the following signal identified four weeks ago as part of RBQM central monitoring process based on the new data compared with previous data values and provide a conclusion in 50-100 words.

Previous Signal:

Given the recent analysis indicating an anomaly in your site's Adverse Event (AE) rate, we request an immediate investigation into potential causes. The AE rate at your site is 4.5, notably higher than the average of 3.0 across other sites. However, your site shows lower rates in Medical History (MH), Concomitant Modification (CM), and a similar AE Duplication rate compared to other sites, which contradicts the high AE rate. Please review your data collection and reporting processes.

Following are the current KRI values for the site, previous KRI values for the site, and the current average. KRI value across sites after a month:

AE Rate: 3.5, 4.5, 3

MH Rate: 1.8, 1.8, 2.5

CM Rate: 1, 0.6, 1.2

AE Duplication Rate: 0.3, 0.5, 0.3

Data Entry Delay: 2, 2, 3

Average Exposure: 110, 85, 135

Note: Usually, a high AE rate, when accompanied by a high CM rate, MH rate, AE duplication rate, and exposure, and low data entry delay, justifies a high AE rate.

Context Building by Tracking

Tracks and compares updated data over time, highlighting improvements and areas for ongoing monitoring in clinical sites.

LLM Response

The reduction in the AE rate and improvements in the CM rate and the AE duplication rate suggest that the actions taken in response to the initial signal have been effective. The site's performance in these KRIs is more in line with the average across other sites, indicating a successful resolution of the initial anomaly. However, continued monitoring of the MH rate and average exposure is recommended to ensure alignment with broader trends and identify any emerging concerns.



Integrating LLM with Signal Validation Workflow

Level 1: Convenient Listing Dashboard



Patient	N of Visits	Risk Rank	Risk Score	PE Compliance	LB Compliance	VS Compliance	EG Compliance	MH Rate	AE Rate
NIM-55-22-1003-10470	13	4.0	15.0						

Subject Rate: 1.6
N of AEs: 21
N of Visits Completed: 13
Risk Level: High
Risk Flag: High Reporting

Related Data Listings
Patient Profile
AI Summary & Analysis

Dashboards

Private

Safety Graphics

Population Filter

Select existing filters

Add

Reset

Apply

USUBJID

NIM-55-22-1001-10010

Reports

Section 16.2

Listing 16.2.1 Disposition

Listing 16.2.2 Demograp...

Listing 16.2.3 Medical Hi

Untitled

Listing 16.2.2 Demographics

Listing 16.2.2: Demographic Characteristics
Safety Population

	Subject Number	Age	Year of Birth	Gender	Ethnicity	Race
2	1001-10010	42	1978	Male	Not Hispanic or Latino	White

Listing 16.2.6 - Adverse Events

Listing 16.2.6: Adverse Events (AE)
Safety Population

	System Organ Class/ Preferred Term/ Verbatim Term	Start Date/Time	End Date/Time	SAE ?	Severity
2	/ Alveolitis / Alveolitis	2020-03-16 11:00	2020-03-25 15:15	N	Grade 1 (Mild)
3	/ Anaemia / Anaemia	2020-02-22 11:48	2020-02-28 12:02	N	Grade 1 (Mild)
4	/ Anaemia / Anaemia	2020-03-12 01:01	2020-03-25 19:19	N	Grade 2 (Moderate)
	/ Atelectasis / Atelectasis	2020-02-29 15:15	2020-03-16	N	Grade 2 (Moderate)

NOTE: SAE = Serious Adverse Events

Listing 16.2.3 Medical History

Listing 16.2.3: Medical History
Safety Population

	Subject ID	System Organ Class/ Preferred Term/ Verbatim Term	Subject has Pre-specified condition	Start Date	End Date
2	1001-10010	/ Loss Of Consciousness / LOSS OF CONSCIOUSNESS	Y	2020-01-19	2020-01-24
3		/ Hypertension / HYPERTENSION	Y	2020-02-02	2020-02-10
4		/ Focal Deficit / FOCAL DEFICIT	Y	2020-01-26	2020-02-05

Listing 16.2.8 - Meds

Listing 16.2.8: Prior and Concomitant Medications
Safety Population

	Subject ID	Prior or Concomitant Medication	Reported Term	Start Date (Day)	End Date (Day)
2	1001-10010	Concomitant Medications	POTTASIIUM SUPPLEMENTS	2020-02-13 (-5)	(NA)
3		Concomitant Medications	ACETAMINOPHEN	2020-02-20 (2)	2020-02-26 (8)
4		Concomitant Medications	FUROSEMIDE	2020-02-25 (7)	2020-03-04 (15)
5		Concomitant Medications	LORAZEPAM	2020-02-16 (-2)	2020-02-26 (8)
6		Concomitant	MAGNESIUM	2020-02-23 (5)	2020-02-29 (11)



Integrating LLM with Signal Validation Workflow

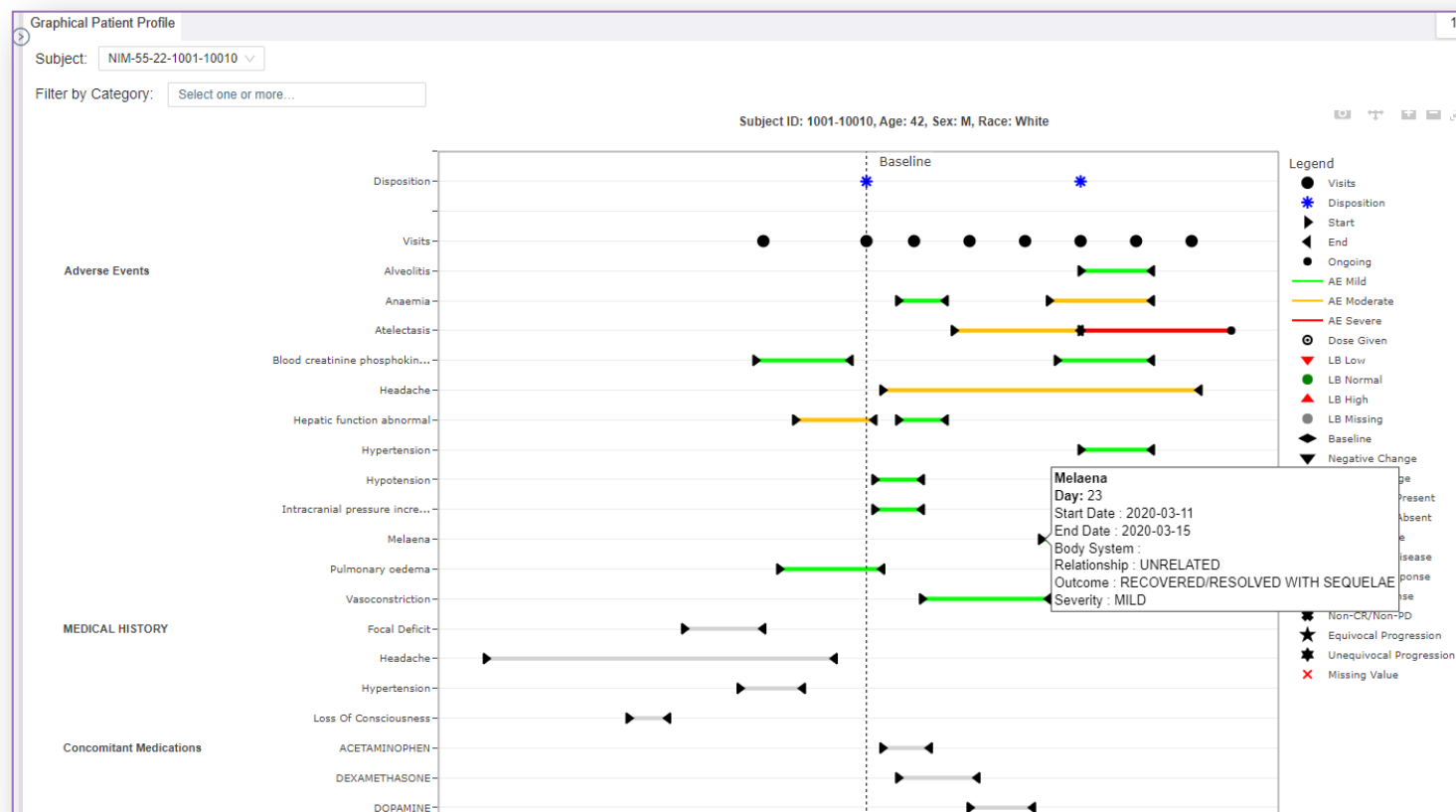
Level 2: Graphical Patient Profile



Patient	N of Visits	Risk Rank	Risk Score	PE Compliance	LB Compliance	VS Compliance	EG Compliance	MH Rate	AE Rate
NIM-55-22-1003-10470	13	4.0	15.0	●	●	●	●	●	●

Subject Rate: 1.6
N of AEs: 21
N of Visits Completed: 13
Risk Level: High
Risk Flag: High Reporting

Related Data Listings
Patient Profile
AI Summary & Analysis





Integrating LLM with Signal Validation Workflow

Level 3: AI Summary & Analysis



Patient	N of Visits	Risk Rank	Risk Score	PE Compliance	LB Compliance	VS Compliance	EG Compliance	MH Rate	AE Rate
NIM-55-22-1003-10470	... 13	4.0	15.0	● ...	● ...	● ...	● ...	● ...	●

Subject Rate: 1.6
N of AEs: 21
N of Visits Completed: 13
Risk Level: High
Risk Flag: High Reporting

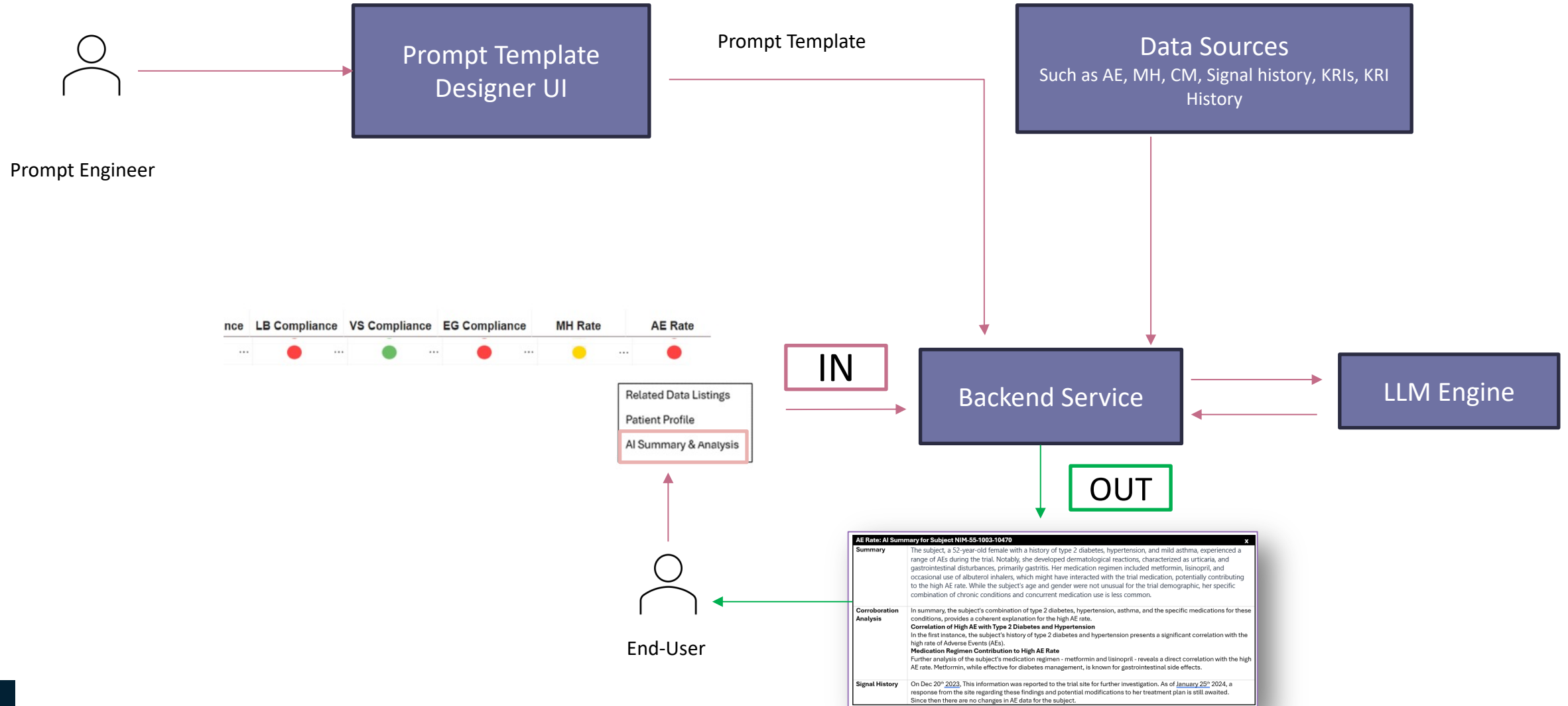
Related Data Listings
Patient Profile
AI Summary & Analysis

AE Rate: AI Summary for Subject NIM-55-1003-10470		x
Summary	The subject, a 52-year-old female with a history of type 2 diabetes, hypertension, and mild asthma, experienced a range of AEs during the trial. Notably, she developed dermatological reactions, characterized as urticaria, and gastrointestinal disturbances, primarily gastritis. Her medication regimen included metformin, lisinopril, and occasional use of albuterol inhalers, which might have interacted with the trial medication, potentially contributing to the high AE rate. While the subject's age and gender were not unusual for the trial demographic, her specific combination of chronic conditions and concurrent medication use is less common.	
Corroboration Analysis	In summary, the subject's combination of type 2 diabetes, hypertension, asthma, and the specific medications for these conditions, provides a coherent explanation for the high AE rate. Correlation of High AE with Type 2 Diabetes and Hypertension In the first instance, the subject's history of type 2 diabetes and hypertension presents a significant correlation with the high rate of Adverse Events (AEs). Medication Regimen Contribution to High AE Rate Further analysis of the subject's medication regimen - metformin and lisinopril - reveals a direct correlation with the high AE rate. Metformin, while effective for diabetes management, is known for gastrointestinal side effects.	
Signal History	On Dec 20 th 2023, This information was reported to the trial site for further investigation. As of January 25 th 2024, a response from the site regarding these findings and potential modifications to her treatment plan is still awaited. Since then there are no changes in AE data for the subject.	



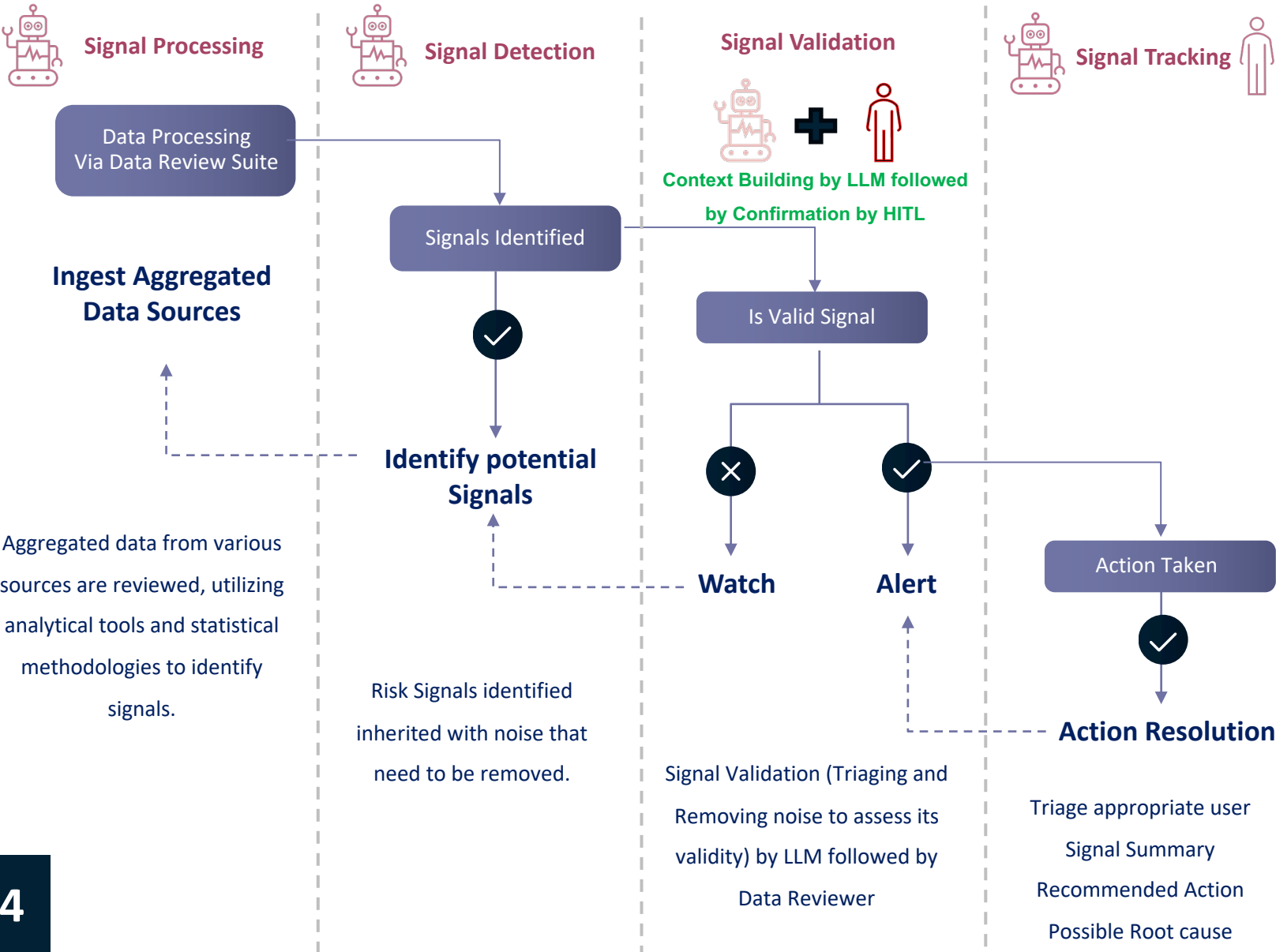
Integrating LLM with Analytics Application

General Approach





Conglomerate of LLM+ HITL in Signal Journey



Key Merits of LLM + HITL Signal Validation

- Significant **Reduction of time spent** on signal validation
- Looking beyond few data to **more insights** through refined prompts
- Action owners receive more actionable, insightful signals, **optimizing everyone's time** and effort.
- This can support **RBQM's overall success** beyond



Key Takeaways & Forward-Looking Statements



01

Success

- Enhanced Insights
- Quicker Decision Making
- Integrated AI Workflow
- Human-LLM Synergy is key

02

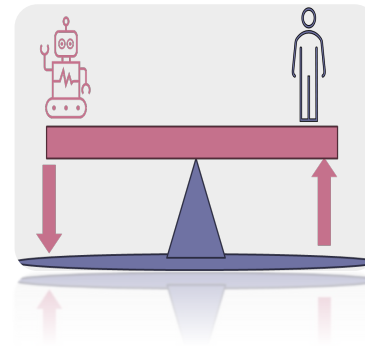
Challenges

- Cautiously Optimistic and Small Wins of AI
- Prompt Engineering
- Robust Data Platform
- Training

03

Similar Opportunities

- Targeted eligibility for Medical Monitories
- Signal Detection in Safety Analytics
- Data Quality Index & Data Surveillance
- Patient Narratives
- Risk Management



We advocate for the iterative, human-linked, process-aligned and judicious utilization of AI!

Thank YOU!
We are happy to answer your questions!



Nimble Clinical Research

AI-powered Clinical Data Solutions

*Your Clinical Data * Our AI-based Solutions * Endless Possibilities*

Collaborate with us:

Vineet Jain

Founder & CEO

Nimble Clinical Research, USA

991 US Highway 22, Suite 200

Bridgewater, New Jersey 08807 USA

Email: info@nimble-cr.com

Website: <https://nimble-cr.com/>