

Rethinking Data Quality

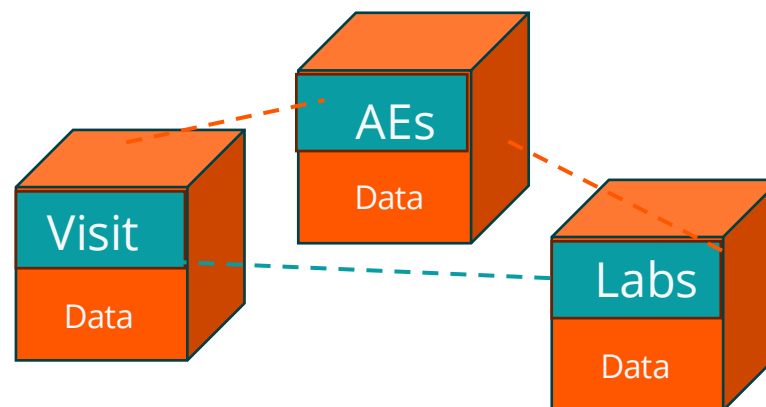
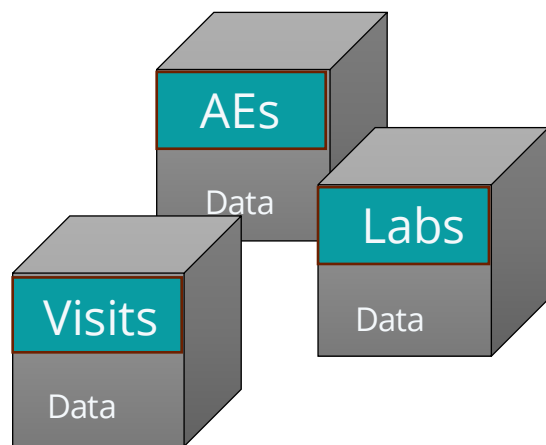
From Fragmented Checks to **Connected** Insights
Moving beyond validation to true clinical intelligence.

Disclaimer

The information presented in this session is for informational purposes only and reflects the current capabilities and intended use of Capish's technology. It does not constitute legal, regulatory, or investment advice. All examples and demonstrations are based on publicly available or simulated data and do not represent real patient or clinical trial data. Any opinions expressed are those of the presenter and do not necessarily reflect the views of Capish or its affiliates. For specific inquiries about the technology, please contact Capish directly.

We've built systems to ensure correctness. Now we must build **understanding**.

In modern clinical trials, data quality depends on seeing **relationships, timing, and context together** — not in isolation.



↑ Earlier Detection of Issues Vs Errors

↑ Confidence Analysis & Decisions

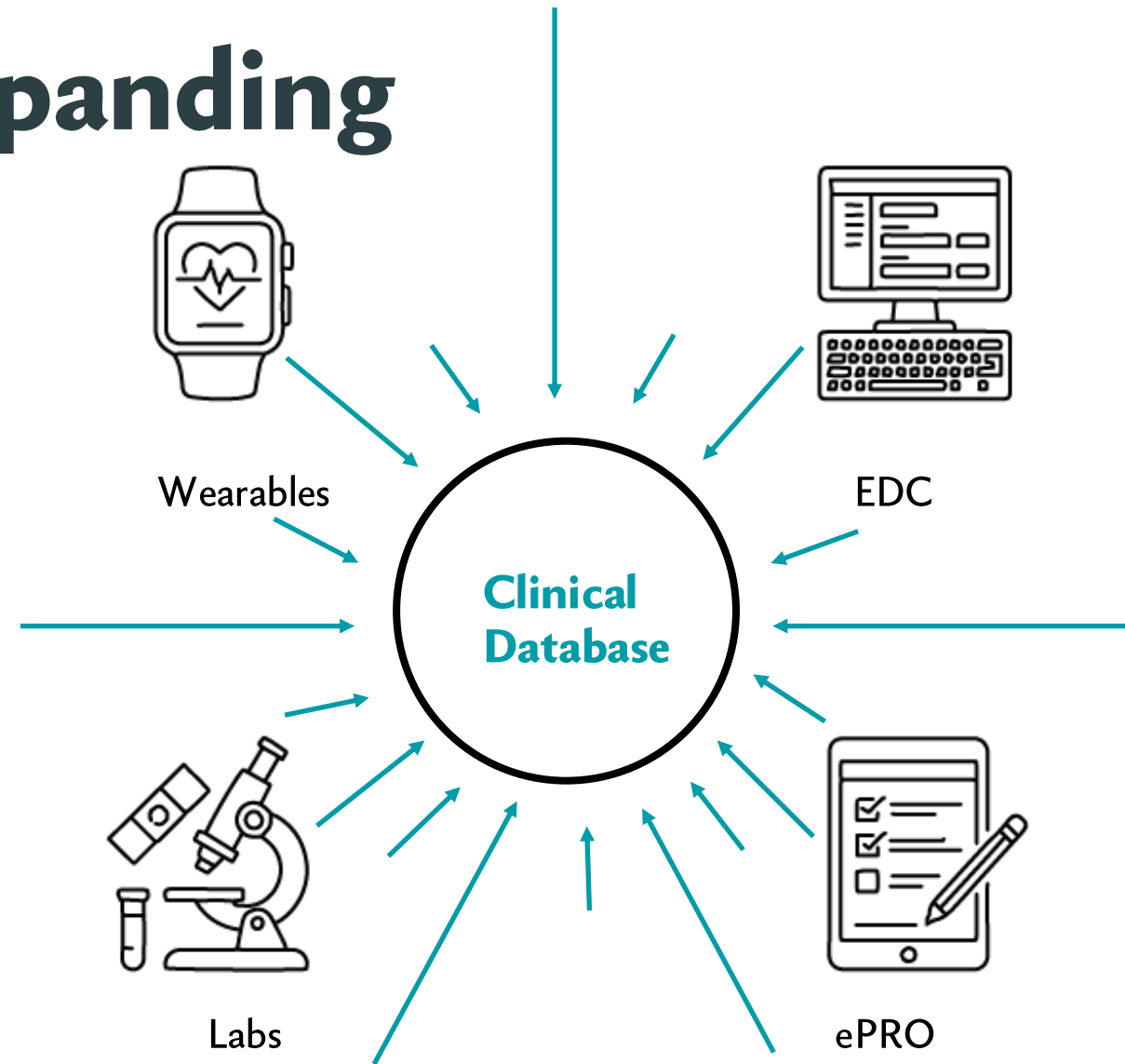
↑ Reuse of Submission Data

The Data Landscape is **Expanding**

Trials have moved beyond the clinic.

Decentralized models, wearables, and real-world inputs are reshaping how data is generated. We have access to richer, more timely information than ever before.

Data quality is no longer just a back-end task — it now shapes timelines, safety interpretation, and patient outcomes.



Correct data alone is **no longer** enough.

The Trap: Traditional systems focus on validation rules and compliance.
A data point can be "valid" (within range, correct format) but still be misleading if the context is missing.

Limited Visibility: Review typically happens across separate domains (silos) and static summaries.

The Result: Data is complete, but relationships and timing remain difficult to see.

120/80	45mg	125	37.5 °C
70bpm	500 ml	1.2g	60mg
•	140/90	50mg	130
38.0 °C	75bpm	400ml	1.5g

Traditional Table View

The Blind Spot: Visit-Based Views Hide Reality

The Abstraction



Administrative View: Looks like a single point in time.

The Reality



Real World View: Events are disconnected by time.

Building **Understanding**, Not Just Correctness

Data quality in complex trials depends on seeing three things together:

01

Relationships

How data across domains connects

02

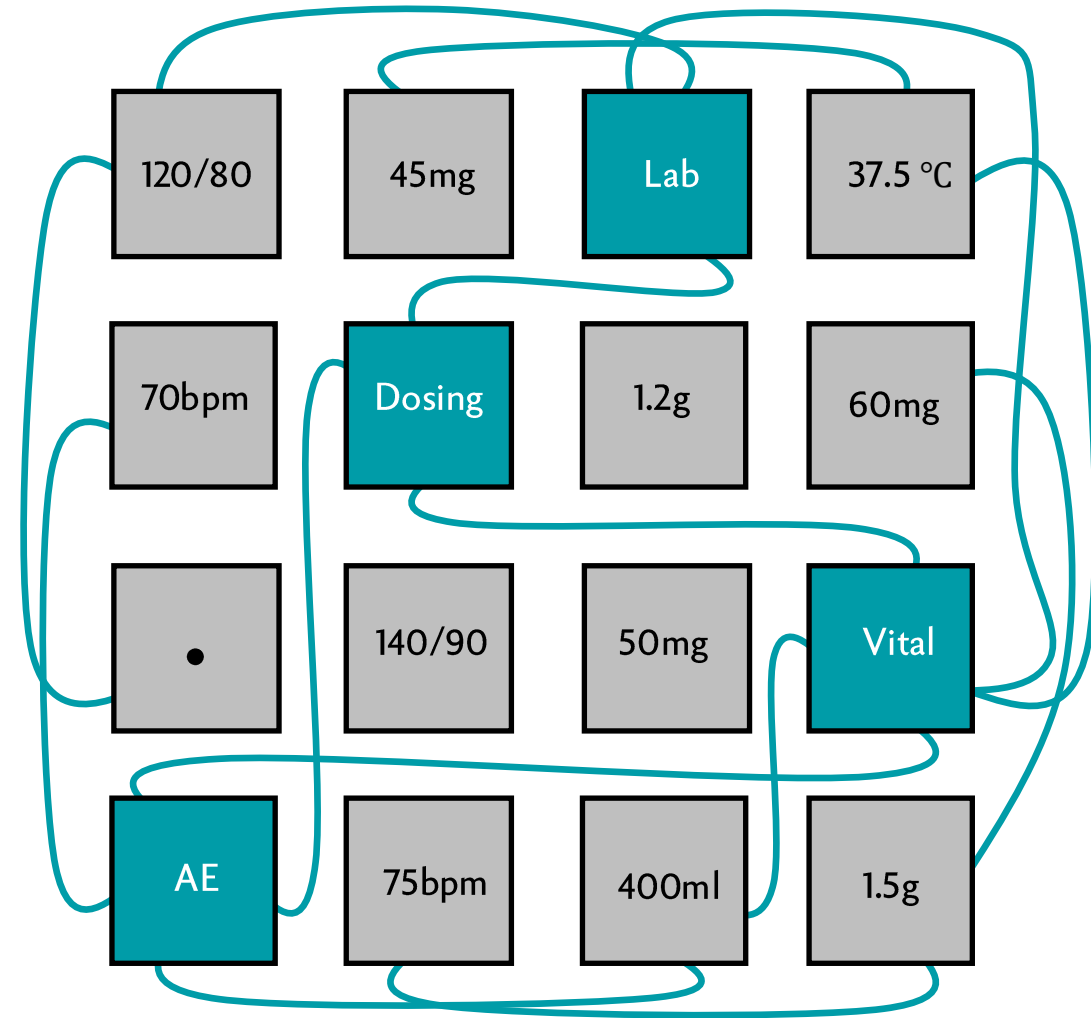
Timing

The true chronological sequence

03

Context

The operational environment



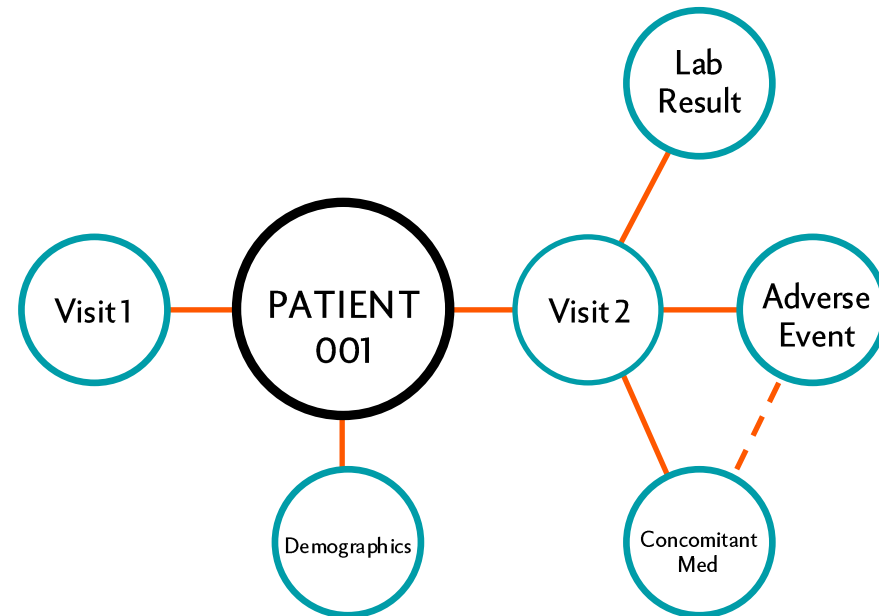
From Silos to Connected Graph

Transforming Tables into a Connected **Graph**

Technical Differentiator

Capish uses a graph-based data model:

- **Entities, Not Rows:** Clinical data as connected nodes (Patient, Visit, Dosing) rather than isolated rows
- **Cross-Domain Reasoning:** Link adverse events directly to exposure and labs without complex SQL
- **No Pre-Defined Queries:** Exploration is flexible, mirroring how clinicians think



Temporal Fidelity: **Bridging** Visits and Days

Dual Prospective

Visit-Based View

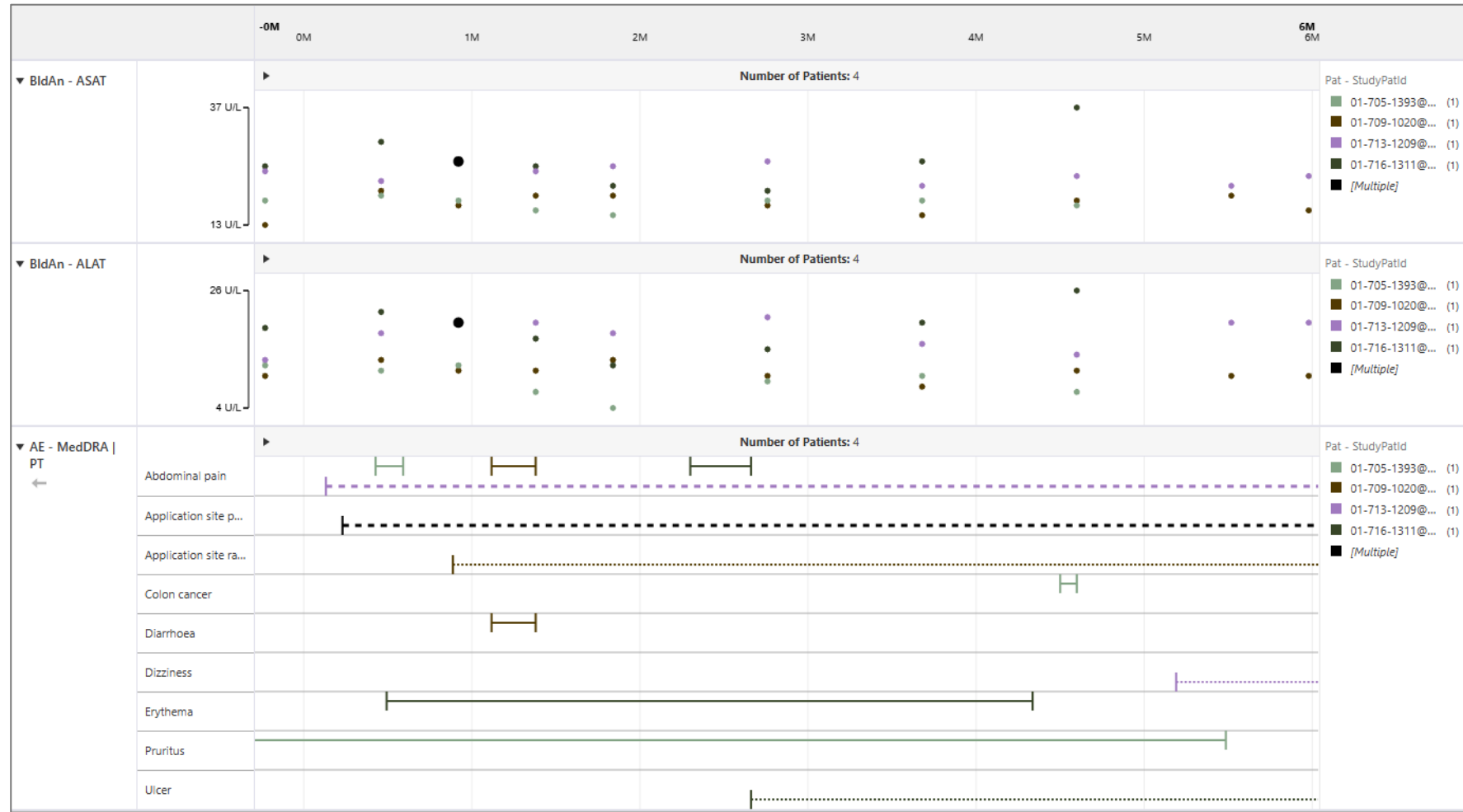
Essential for standardization, reporting, and compliance

Day-Based View

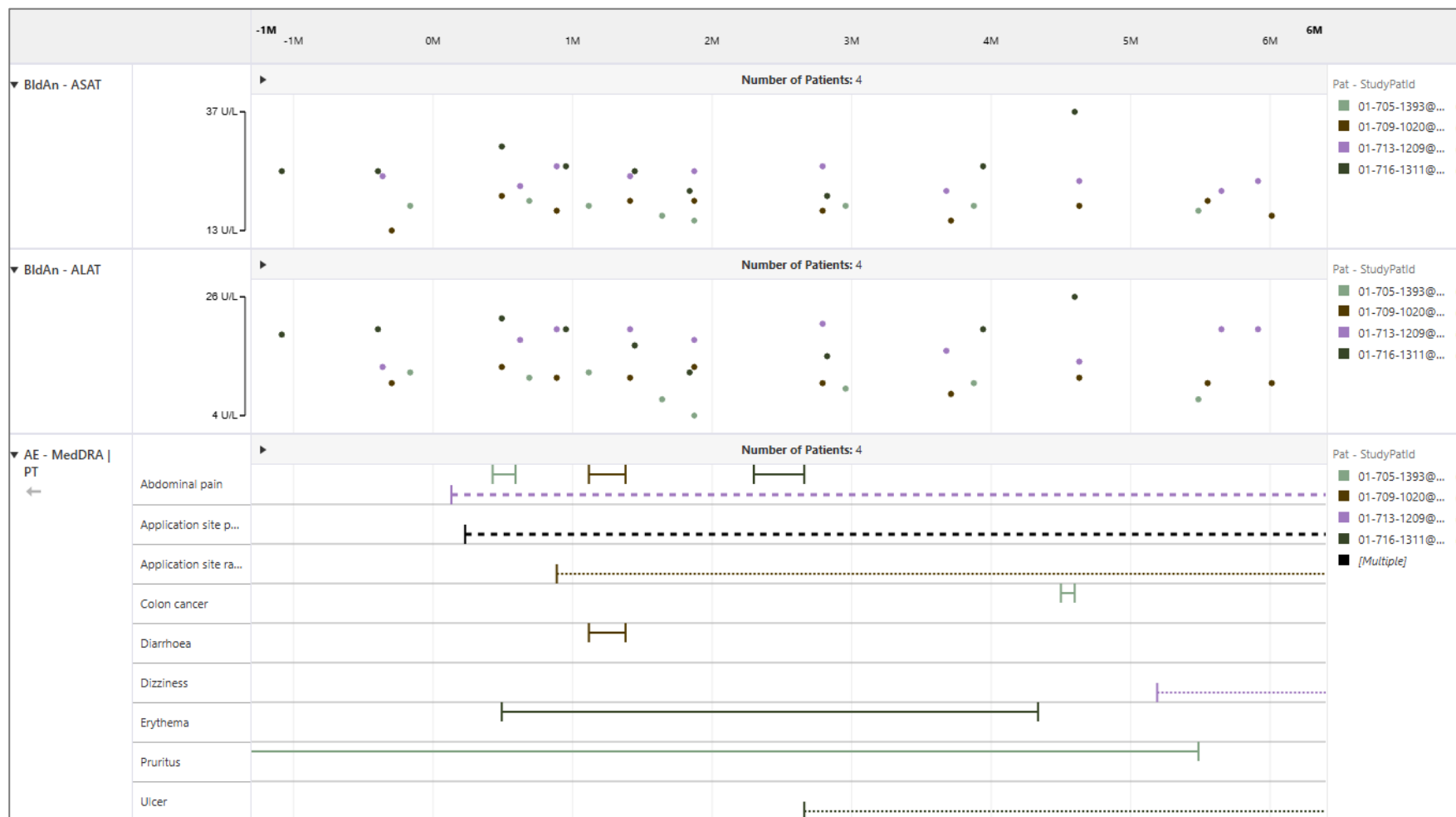
Essential for safety, sequence, and operational reality



Visit



Days



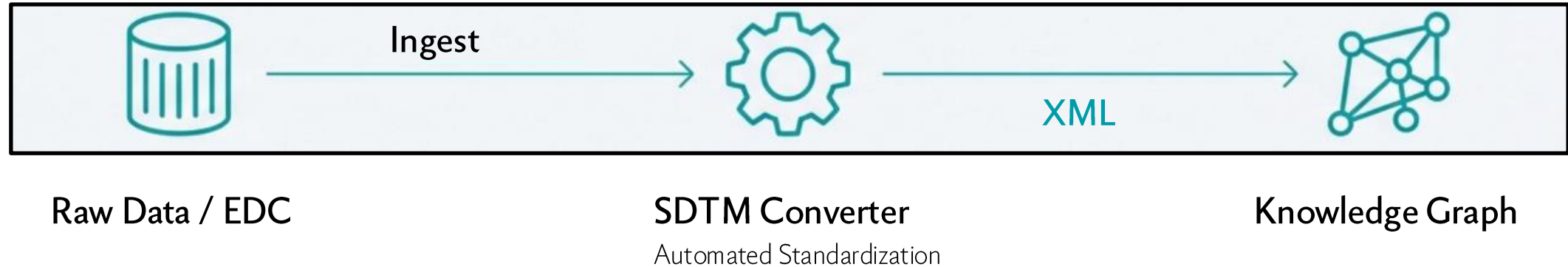
From Static Listings to **Interactive Exploration**

Issues surface through exploration rather than just escalation.



Move fluidly from population signals down to individual patient details.

Built on Standards to Unlock **Value**



CDISC SDTM & PHUSE Collaboration

Reuse of submission data for cross-study analysis and AI preparation, aligned with community standards.

Automated standardization transforms raw data into connected intelligence.

The Shift: From Checks **to Insights**

FROM: Validation & Cleanup

TO: **Insight & Intelligence**

FROM: Visit-Based Abstraction

TO: Chronological + Visit Context

FROM: Static Listings

TO: Explorable Views

FROM: Isolated Domains

TO: **Cross-Domain Reasoning**

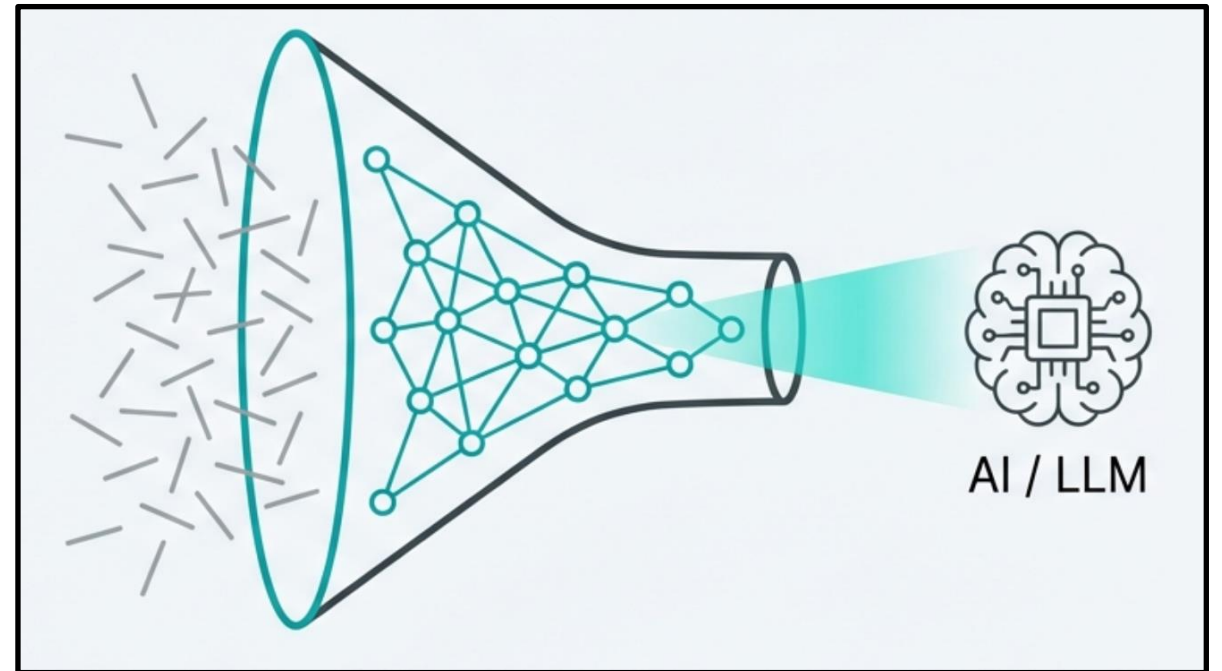
Data quality is no longer a back-end task—it is a decision driver.

A **foundation** for AI and Machine Learning.

Data needs context, not just clean rows. Large Language Models (LLMs) struggle with flat, disconnected tables.

The Advantage:

- Graph models provide the semantic structure AI needs to understand relationships.
- We structure the data so the machine can understand the story.



Data quality is no longer a Back-end task – it is a decision driver.

Let's connect your data.

Eva Kelty, Capish (Malmö, Sweden)
Matthew Kramer, Capish (Boston/San Francisco, USA)