empirisys

Making Sense of the Noise

AI and the Value Hidden in Operational Data

Gus Carroll | CEO & Co-Founder, Empirisys

Hazards Forum | 2026

The Challenge: A Sea of Unstructured Data

Process industries generate vast volumes of workforce and operational data — but that data arrives fragmented, inconsistent, and largely unanalysed.

Fragmented

Data scattered across sites, systems, and formats — no single coherent picture.

Inconsistent Quality

56% of Life-Saving Rule fields left empty or marked 'N/A' in pre-AI industry data.

Volume Overload

Tens of thousands of data points: beyond what any analyst can manually review.

Latent Signal

60–70% of HSE inspection findings already visible in prior workforce observation data — undetected.

The result: critical risk signals hidden in plain sight.

The Five-Year Warning

Our core finding — grounded in industry data

60-70%

of HSE inspection findings were already visible in prior workforce observation data; they just weren't being read.

The Opportunity

If AI can surface latent signals from observations, organisations gain years of warning before a major event.

The Proof Point

Demonstrated across UKCS operators via Step Change in Safety E-Obs programme, now incorporating IOGP 544 barrier taxonomy.

The Stakes

In process safety, the cost of missing a leading indicator can be catastrophic and irreversible.

Source: Empirisys 'Five-Year Warning', 2025 Edition

How AI Helps: Classification Models

Each data point can be enriched in real time through an AI classification pipeline — turning raw text into structured, analytics-ready data.

Classification Outputs

- Safe / Unsafe observation
- Act (behaviour) or Condition (plant)
- Life-Saving Rule (LSR) — 9 categories
- Hazard & Sub-hazard (14/80+ categories)
- IOGP 544 Barrier assignment
- Suggested remedial Action
- Potential Consequence

How It Works

Binary ML models trained on real industrial data for Safe/Unsafe and Act/Condition

Industry-infused LLMs for LSR, Hazard and Barrier classification

Text-generative LLM for action & consequence suggestion

Not 'ChatGPT + prompt' — purpose-built, validated pipeline

Accuracy v Human SME

Accuracy
>97% | ~40%

Recall
>85% | ~30%

Drift
>70% | n/a

How AI Helps: Thematic Analysis

Once classified, observations are analysed collectively — revealing patterns no individual could detect across thousands of records.

01

Embed & Cluster

Text embedding + dimensionality reduction + clustering identifies statistically coherent groups of similar observations.

02

Name & Summarise

AI names, summarises and extracts key points from each theme — turning mathematical clusters into human-readable findings.

03

Evolve Over Time

The evolving model tracks themes longitudinally — new themes emerge; existing themes are monitored for prevalence shifts.

Example in practice

In practice, this approach identified identified a 'Door Integrity' theme across offshore operator data — revealing both a behavioural pattern (doors left wedged open) and a mechanical failure pattern. This triggered a targeted maintenance review AND a behavioural intervention. Theme prevalence tracking then confirmed whether the intervention was working.

Why Not Use Co-Pilot

Copilot is a general-purpose language model interface built to augment productivity across generic enterprise workflows

01

No domain training data

02

No validated classification architecture

03

Cannot meet COMAH / ONR regulatory requirements

03

No explainability or audit trail

From Boardroom to Control Room: Value-Adding Insights

The same underlying data speaks differently at different levels of the organisation.

Board & Executive

Portfolio-level barrier health status. Are our major hazard controls degrading? Are we generating leading indicators ahead of regulatory inspection?

Process Safety & HSE Management

Hazard & barrier breakdown by site, asset, LSR category. Theme prevalence trends. Early warning of systemic weaknesses requiring targeted intervention.

Operations & Site Management

Real-time observation quality feedback. Reviewer-assisted action assignment. Theme analysis by shift / area for localised risk focus.

Control Room & Field Workforce

Inline AI assistance at point of observation submission: category suggestions, action recommendations — faster, higher-quality reporting.

Grounded in Industry Standards: Safety Critical Frameworks

Boost maps observation data to the IOGP barrier taxonomies — turning anecdotal reports into structured Major Accident Hazard and Severe Injury & Fatality intelligence.

Why Barrier Classification Matters

- Each observation becomes an indicator of barrier health — not just an event record
- Challenges to SC-PPP controls can be overlaid on bow-tie / barrier diagrams alongside other process safety / SIF KPIs
- Degradation trends visible before inspection findings emerge
- Provides structured comparison across sites, operators, and time periods
- Aligns with COMAH competent authority expectations for leading indicator monitoring

IOGP Frameworks

The IOGP process safety framework defines hardware and non-hardware barrier categories enabling systematic cross-industry comparison of barrier performance

IOGP Life Saving Rules provide common presursors to Severe Injury & Fatality (SIF) events

Both used as Boost's classification backbone.

Step Change in Safety E-Obs Programme

This approach has been validated across UKCS operators through the Step Change in Safety E-Obs programme — providing the industry's most comprehensive observation dataset mapped to IOGP PSM barriers and LSRs

Organisational Considerations: Cognitive Enablement

Done well, AI doesn't replace human safety judgment — it enhances it. The goal is cognitive enablement, not cognitive replacement.

AI Provides

- Rapid classification of large observation volumes
- Consistency in category assignment across all reporters
- Pattern detection across tens of thousands of records
- Preliminary risk screening against barrier framework
- Inline prompts that surface the right question at submission

Humans Provide

- Contextual interpretation of plant-specific conditions
- Challenge and override AI classification where needed
- Decision authority on remedial actions and escalation
- Validation through safety studies and engineering judgment
- Maintained knowledge of plant configuration and history

The design principle: AI as a structured prompt, not an authority — engineers remain the decision-makers.

The Other Side: Managing Cognitive Offloading Risk

The HSE ThinkSHARP research programme (AISI-funded, 2026) has identified a parallel risk we must take seriously.

⚠ The Automation Drift Risk

- AI introduced as productivity tool — not safety-critical system — so no formal validation
- AI outputs treated as authoritative; engineers stop checking primary data
- Expertise shifts from understanding the plant to operating the AI system
- Junior staff assume AI already incorporates standards; critical thinking declines
- Safety studies gradually rely on AI-generated summaries; regulators lose traceability
- Result: more 'holes' in the Swiss cheese — incidents become more probable

✓ Designing for Cognitive Resilience

- Treat AI as safety-relevant decision support — require technical validation before deployment
- AI outputs must cite data sources, assumptions and confidence levels
- Safety studies (HAZOP/LOPA) performed independently of AI outputs
- Engineers actively challenge AI outputs; culture rewards critical thinking
- Regulators and internal assurance can review AI decision logs and traceability
- Engineering standards and industry codes always override AI recommendations

Key Takeaways

01

The signal is there

60–70% of inspection findings already exist in workforce & operatinal data. The challenge is surfacing them before the event.

02

Classification unlocks structure

AI converts unstructured text into barrier-mapped, analytics-ready data — turning noise into leading indicators.

03

Thematic analysis reveals systemic risk

Pattern detection across large datasets shows what concerns the workforce — enabling targeted, trackable interventions.

04

Value spans the whole organisation

From real-time field prompts to boardroom barrier health dashboards — the same data serves every level.

05

Cognitive enablement, not offloading

AI should enhance human judgment, not replace it. Governance, culture and traceability are non-negotiable safeguards.

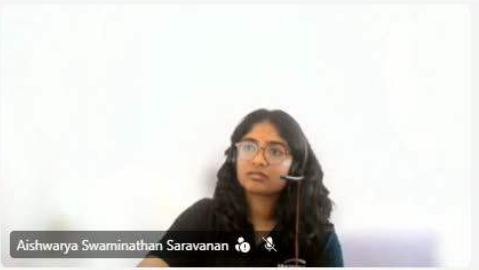
Questions & Discussion

gus.carroll@empirisys.com | empirisys.com



Anna McNeil







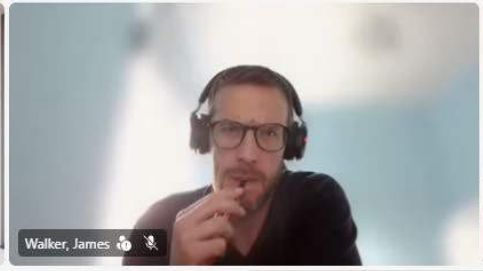
Carroll Alex

Threlfall, David

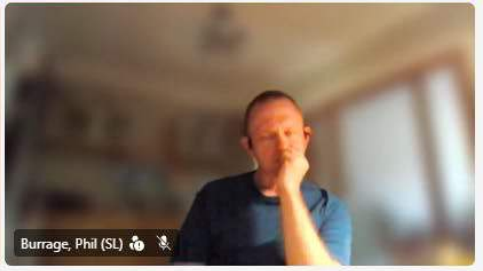
1

Helen Berry









Aaron Vincent