DETECT

A Tool to Improve Loss Prevention from Human Error in High Hazard Industries through the Identification of Error Traps Using Machine Learning

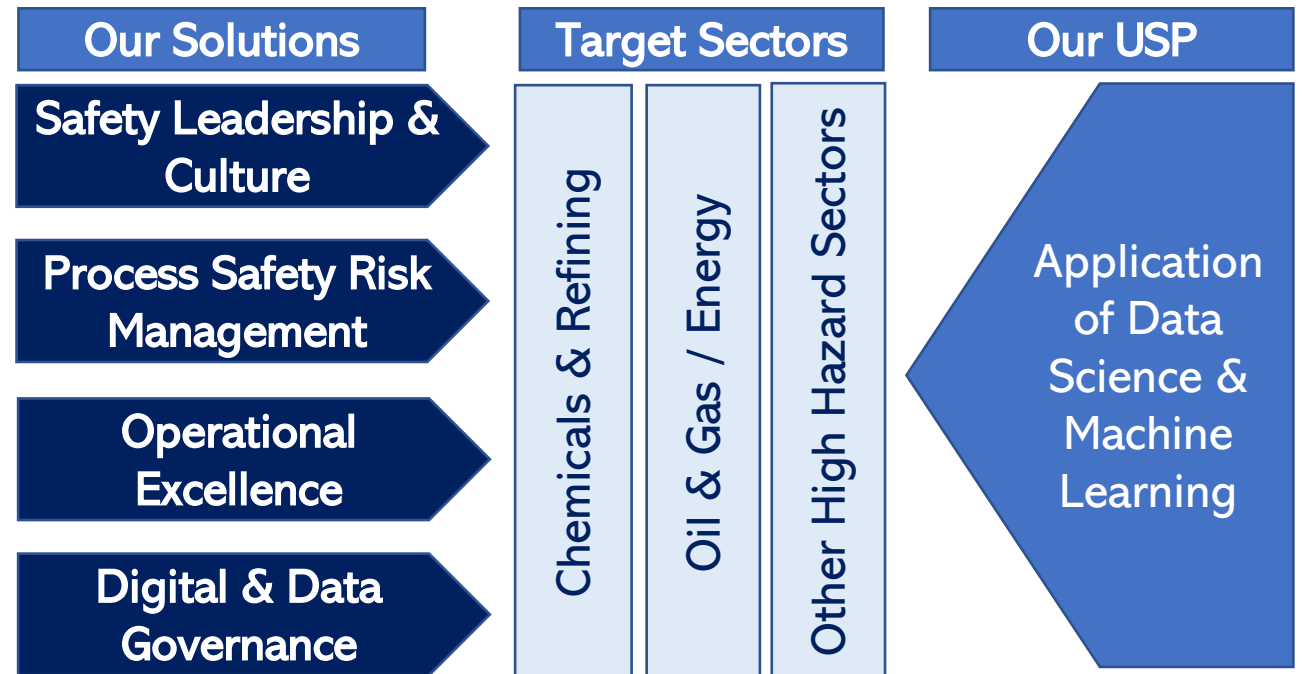**Presenter: Nyala Noë, Data Scientist at Empirisys**

About Empirisys

> **We use engineering & operational expertise and machine learning to create actionable insight to help complex, high hazard organisations manage their risks**

About the Company

Established in October 2020 by two co-founders

- Gus Carroll – former Chief Engineer at Centrica (£30bn end to end energy company)
- Pete Sueref – former Data Science Director at Centrica
- We have a team of highly experienced operational & engineering professionals and PhD level data scientists

What We Do

Our Solutions

- Safety Leadership & Culture
- Process Safety Risk Management
- Operational Excellence
- Digital & Data Governance

Target Sectors

- Chemicals & Refining
- Oil & Gas / Energy
- Other High Hazard Sectors

Our USP

Application of Data Science & Machine Learning

empirisys

Gas from the wellbore erupts through the rig floor which ignites an explosion killing 11 and injuring 17

Deepwater Horizon

Transocean & BP Oil Rig
Gulf of Mexico
April 20, 2010

3

What is DETECT and What Does It Do?

empirisys

- DETECT helps to identify the organisational weaknesses that lay "hidden in plain sight" at every operational asset or facility

- These are often called "Error Traps" or "Performance Influencing Factors"

Plant Condition 20%	Hardware 50%	10%
	Software 50%	10%
Accident, Event or Near Miss	Violations 10%	8%
Human Error 80%	Lapses 20%	16%
	Organisational Factors 70%	56%

>50% of accidents caused or influenced by organisation factors or error traps
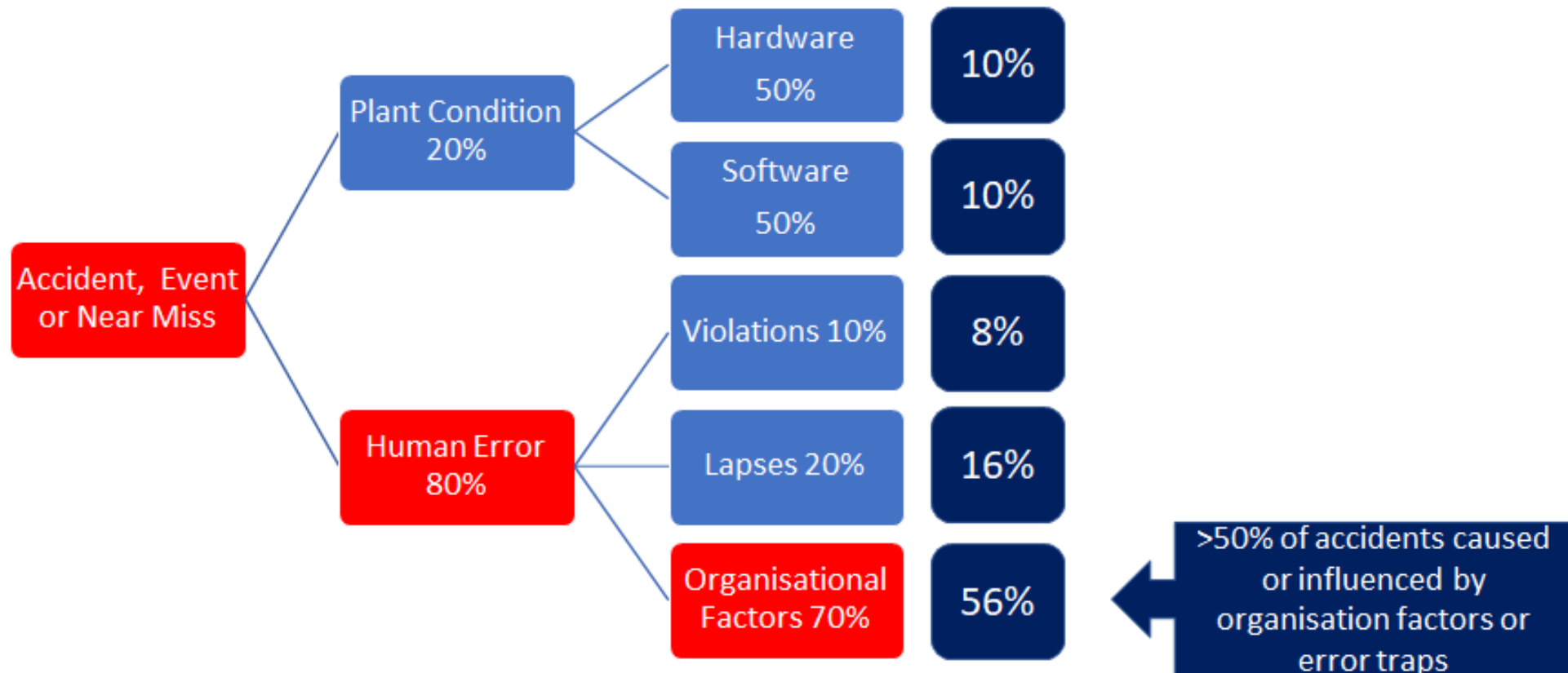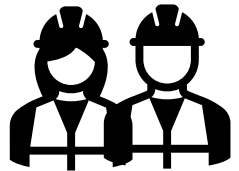
What is an Error Trap?

empirisys

- Result of organisational factors and latent defects that increase the likelihood of human error

- In safety critical settings, this can result in catastrophic consequences

- Error traps are often "hidden in plain sight" but are extremely difficult to spot

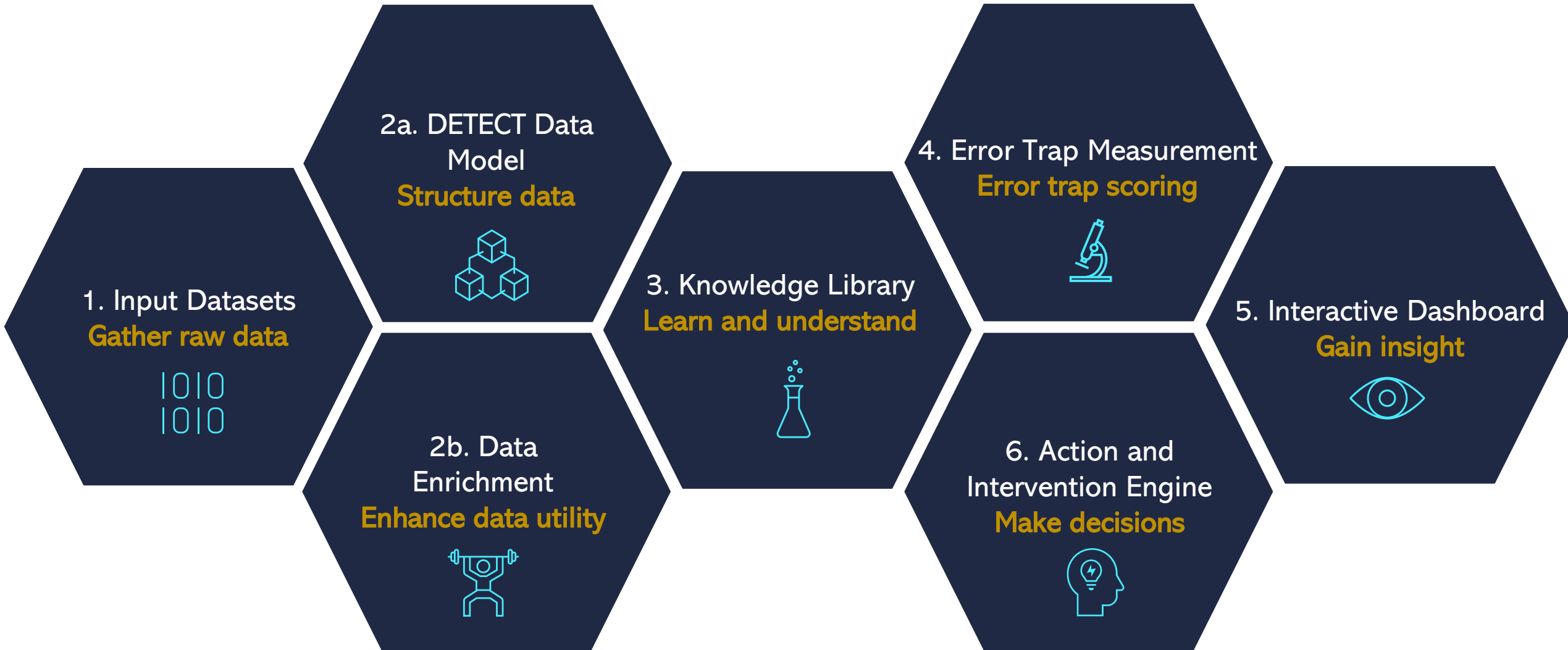| JOB	• Inappropriate procedures • Complex task instructions • Working environment
PERSON	• Stress & morale • Fatigue • Physical capability & condition • …
ORGANISATION	• Work pressures (production vs safety) • Psychological safety • Communication • …

HSE Health & Safety Executive

DETECT

empirisys

1. Input Datasets
Gather raw data

2a. DETECT Data Model
Structure data

2b. Data Enrichment
Enhance data utility

3. Knowledge Library
Learn and understand

4. Error Trap Measurement
Error trap scoring

6. Action and Intervention Engine
Make decisions

5. Interactive Dashboard
Gain insight

Main Challenges

- Is it possible to extract, cleanse and transfer unstructured data from operational technology sources such as maintenance systems, HSE reporting systems etc? **Data Feasibility**

- Is it possible to detect error traps automatically from data? **Machine Learning Feasibility**

- Do users need our product? **Desirability**

- Will user believe the output of the product? **Credibility**

Data Feasibility

empirisys

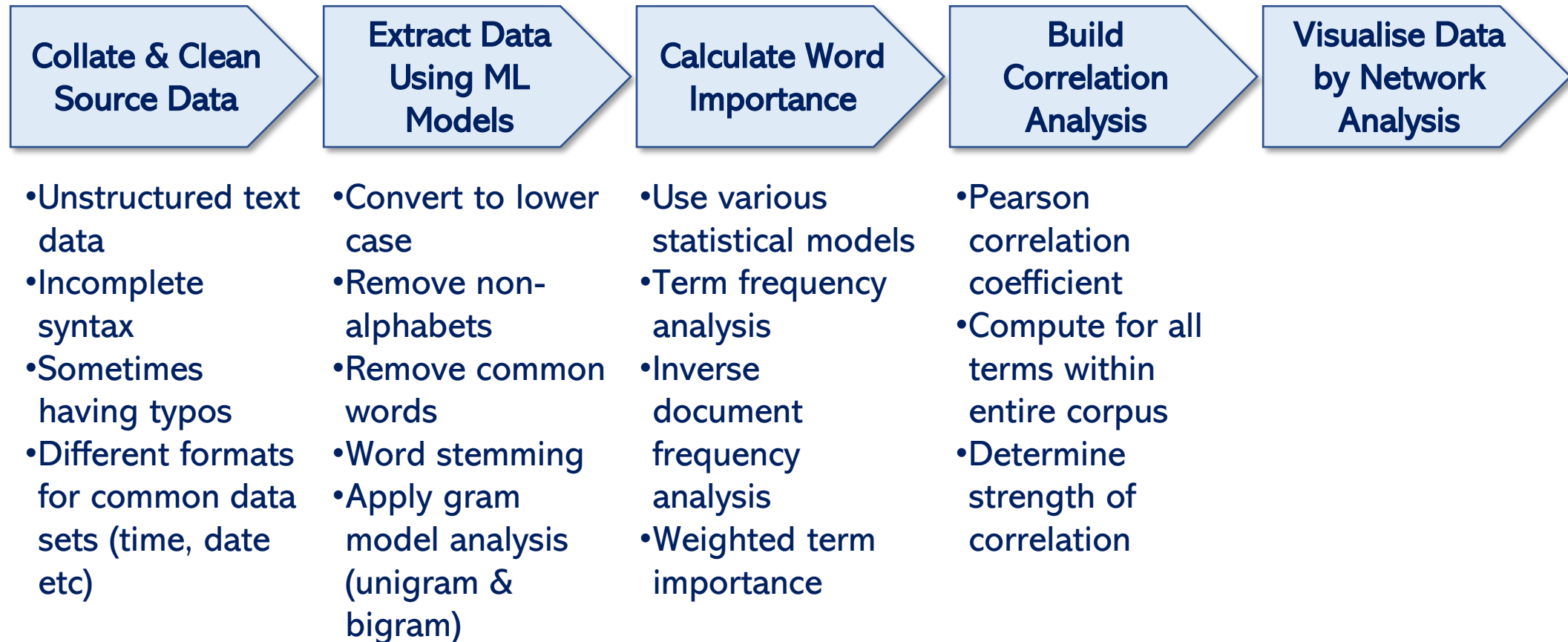|	REFLECTS…	EXAMPLES…	WHAT MACHINE LEARNING OFFERS
STRUCTURED DATA	• Formal structure & process • The way we say things get done • How the world sees us	• DCS/PI • Alarms & alerts • Action closures • Risk ratings • Financial KPIs etc	• Correlate trends with future outcomes • Early warning • Feedback / action loops
UNSTRUCTURED DATA	• Informal structure & process • The way things really get done • The stories we create • The unexplored way we are • The true culture of the organisation	• Worker logs • Observations • Interviews • Surveys • Meeting notes • Audit reports etc • Emails • Social networks • Conversations etc	• Themes & topics • Sentiment • Relationships • Networks • Clusters

- Unstructured data held in different formats on IS platforms that do not easily connect
- Often use different rulesets for key metadata such as timestamps, tags and location.

8

Machine Learning Feasibility
High Level Process

empirisys

Collate & Clean Source Data >	Extract Data Using ML Models >	Calculate Word Importance >	Build Correlation Analysis >	Visualise Data by Network Analysis >

- Unstructured text data
- Incomplete syntax
- Sometimes having typos
- Different formats for common data sets (time, date etc)

- Convert to lower case
- Remove non-alphabets
- Remove common words
- Word stemming
- Apply gram model analysis (unigram & bigram)

- Use various statistical models
- Term frequency analysis
- Inverse document frequency analysis
- Weighted term importance

- Pearson correlation coefficient
- Compute for all terms within entire corpus
- Determine strength of correlation

Machine Learning Feasibility - Results

empirisys

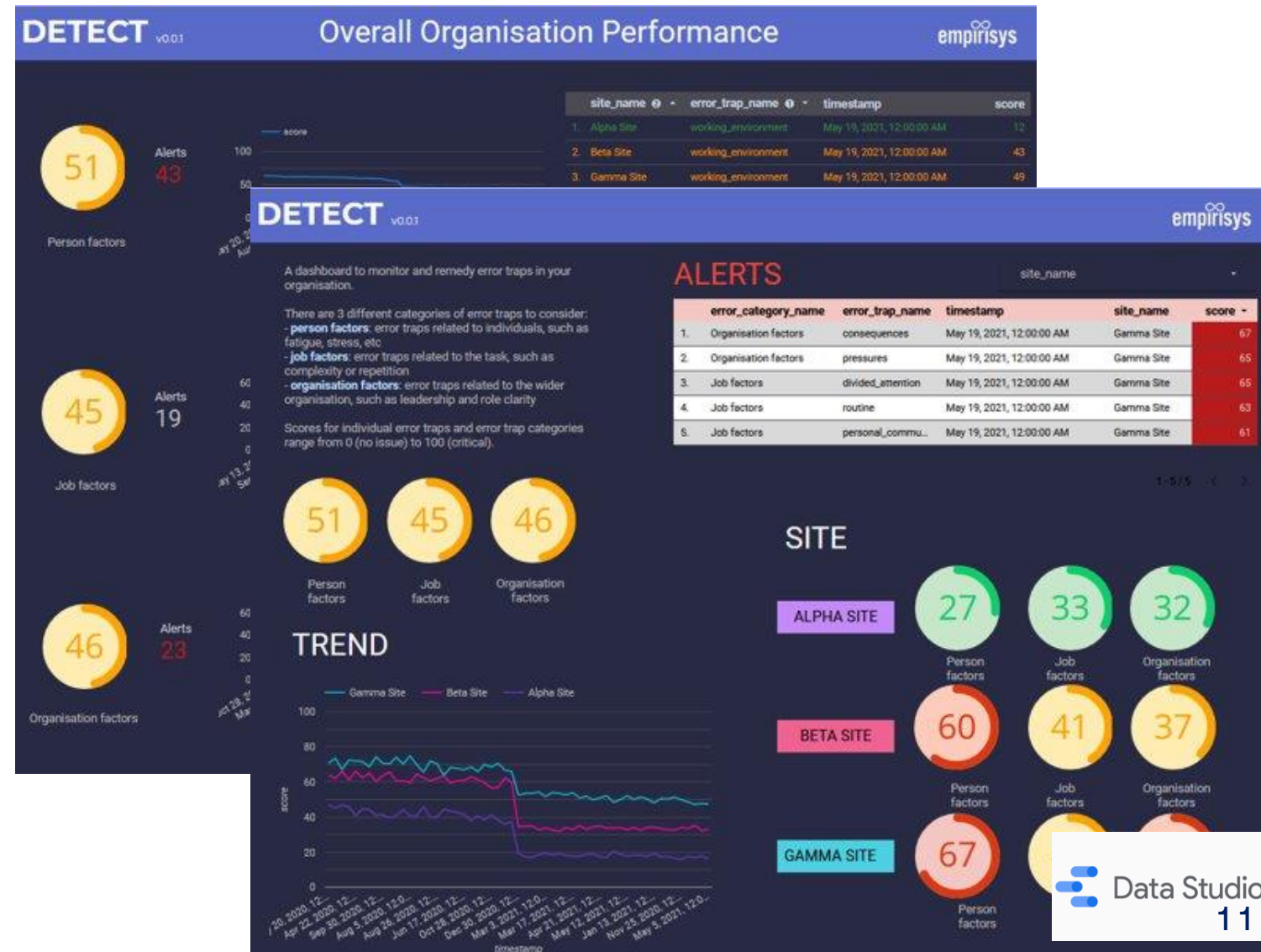**Event Reports & Observation Logs**

Observations leading to 'Electrical Contact' events
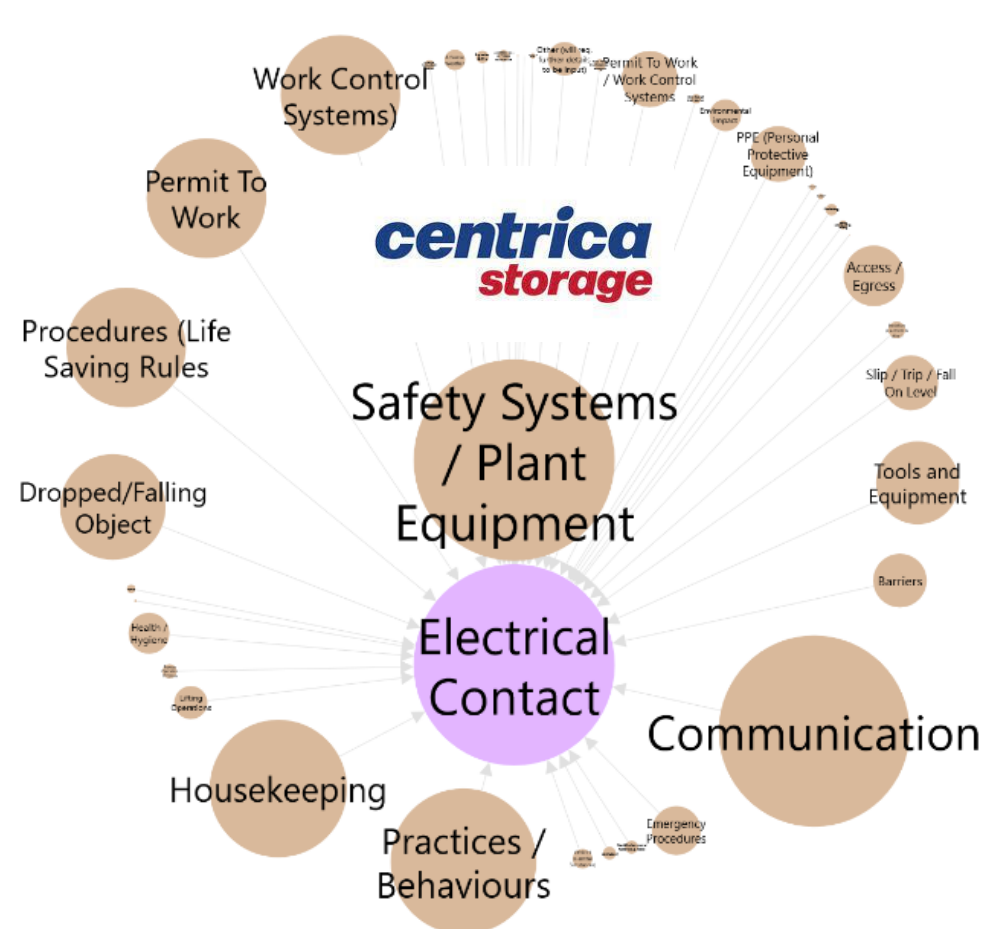
Desirability Testing

- 5 subject-matter experts in high-hazard industries and process safety

- 1-hour interview with interactive prototype (Google Data Studio dashboard)

- User experience and desirability questions

Usability Study - Results

empirisys

All 5 experts interviewed felt there was a need for a product like DETECT

Key Observations

- Tool requires a minimum level of data maturity from the organisation

- Most useful when embedded into existing risk management and process safety tools

- Data evidence for issues is invaluable

Key Recommendations

- Favour simplicity

- Be able to take action and lead to interventions

- Need for transparency

DETECT is Currently in Development

empirisys

| Successful initial user desirability testing | Error trap-marker links established | Machine learning feasibility tested |
| | Operational data available | |

Can we do this at scale?

| DETECT Prototype development – feasibility tested | DETECT MVP launched with partner companies – extensive testing phase | Production-ready version of DETECT launched | Expand to new customers within target sector | Expand to other high-hazard industry sectors |
| Q1 – Q3 2022 | Q4 2022 – Q1 2023 | 2023 - 2024 | 2024-2025 | 2026+ |

Thank You

Website: https://empirisys.io

Presenter contact (Nyala Noe): nyala@empirisys.io

General enquiries: info@empirisys.io