

HSE PREDICT

Data Discovery and Prediction Validation Questionnaire

A predictive safety system should be assessed on evidence, not on the confidence of its interface. This questionnaire establishes what will be predicted, how performance will be measured, and whether the data required to validate that performance exists.

It is completed with your HSE and IT leads before any predictive output is relied upon operationally, and it forms the basis of the pilot design.

Organisation	_____	Date	_____
Site / portfolio in scope	_____	Facilitator	_____
HSE lead	_____	IT / OT lead	_____
	_____		_____

How to use this document

The questionnaire is normally completed in a single facilitated session of two to three hours. Sections A to E define the prediction and how it will be judged. Sections F and G establish whether the organisation can supply what validation requires. Section H converts the answers into a pilot design.

Where an answer is not known, record that it is not known rather than estimating. An unknown identified now is a scoping item; the same unknown discovered during a pilot is a failed pilot.

Sections A to E apply to any predictive safety system under evaluation, not only to HSEPredict. They are worth working through even if the evaluation goes no further.

SECTION A**Defining the prediction**

A prediction that is not specified precisely cannot be measured, and anything that cannot be measured cannot be relied upon.

A1 What event is to be predicted?

State the event type in operational language. "Risk is elevated" is not a prediction. "Elevated probability of a gas release at this manifold" is.

A2 What is the unit of prediction?

The entity the prediction attaches to.

☐ Asset ☐ Area ☐ Site ☐ Shift / crew ☐ Portfolio

A2b If other, or if the unit differs by event type, specify.

A3 What is the time window?

The period within which the prediction applies. A useful window leaves sufficient time to intervene — long enough to act, short enough to be actionable.

A4 What action would a warning trigger, and who takes it?

If no action follows a warning, the prediction has no operational value regardless of its accuracy.

A5 Are there event types explicitly out of scope for this pilot?

Naming exclusions prevents scope drift during evaluation.

SECTION B**The base rate**

Serious incidents are rare. A model that consistently predicts nothing will appear highly accurate and provide no value whatsoever. The base rate is what allows genuine performance to be distinguished from arithmetic artefact.

B1 How often has the event in A1 occurred historically?

Give a count and the period it covers, for the scope defined in A2.

B2 Over what period is that history available and reliable?

Note any point at which recording systems, categories or thresholds changed — data before such a change may not be comparable with data after it.

B3 Does the rate vary materially by site, season, shift or operating mode?

A portfolio-level base rate can conceal wide variation. Where it does, the pilot may need to be scoped to a subset.

B4 Are near misses and precursor events recorded separately from incidents?

Precursors are often the more useful prediction target, because they are more frequent and still actionable.

B5 Given the base rate, how many events would be expected during a pilot of the length under consideration?

If the expected count is very low, the pilot cannot demonstrate predictive performance and should be scoped to monitoring value instead. This is worth establishing before committing.

SECTION C**The acceptable error balance**

False positives consume inspection and response capacity. False negatives are missed warnings. The two trade against one another, and the correct balance is a business decision for the client rather than a technical one for the vendor.

C1 What does a single false positive cost you?

In inspection hours, travel, downtime, or diverted response capacity. An approximate figure is sufficient; the purpose is relative weighting.

C2 What is the consequence of a missed warning for the event in A1?

Safety, environmental, regulatory, financial and continuity consequences as applicable.

C3 How many warnings per week can your organisation actually action?

This is a hard operational ceiling. A system that generates more alerts than can be actioned produces alert fatigue, and alert fatigue is worse than no alerts.

C4 Where should the balance sit for this pilot?

☐ Favour sensitivity ☐ Balanced ☐ Favour precision

C4b Record the reasoning, and who owns this decision.

This is a business decision and should be attributable.

C5 Is there a threshold of performance below which the system would not be adopted?

Agreeing this in advance prevents the goalposts moving in either direction after results are seen.

SECTION D**The intervention question**

If a warning prompts an inspection and no incident subsequently occurs, was the prediction incorrect, or was it successful? This has a material effect on measured accuracy and must be resolved before measurement begins.

D1 How will interventions triggered by a warning be recorded?

Without an intervention record, a prevented event is indistinguishable from a false alarm, and the system is penalised for working.

D2 Who records the intervention, and in which system?

If this adds burden to a frontline role, it will decay. Prefer a system already in use.

D3 What will be captured about the intervention outcome?

For example: what was inspected, what was found, what was done, whether a defect or condition was confirmed.

D4 How will a confirmed finding be treated in measurement?

☐ As a successful prediction ☐ As indeterminate ☐ To be decided — see notes

D4b Notes and agreed rule.

D5 Who adjudicates disputed cases?

Name a role. Disputed cases are certain to arise, and resolving them ad hoc erodes trust in the measurement.

SECTION E**Reporting behaviour**

In safety data specifically, reporting culture confounds analysis. An increase in near-miss reports may indicate rising risk, or improving reporting behaviour. These have opposite implications and look identical in the data.

E1 Have there been material changes to reporting policy, categories or thresholds in the period covered by B2?

Include system migrations, category redefinitions, and changes to what must be reported.

E2 Do reporting rates vary materially between sites, shifts or supervisors?

Where they do, signal volume cannot be treated as a risk indicator without adjustment.

E3 Has any campaign, incentive or disciplinary change plausibly affected reporting volume?

Safety campaigns raise reported near misses without raising risk. Disciplinary tightening lowers them without lowering risk.

E4 Is reporting anonymous, attributed, or mixed?

☐ Anonymous ☐ Attributed ☐ Mixed

E5 Are frontline personnel confident that reporting does not attract scrutiny of their own team?

This is the question that determines whether the input data can be trusted at all. An honest answer here is worth more than a favourable one.

SECTION F**Data readiness**

Prediction is a joint undertaking. These are the conditions under which validation is possible.

F1 Which operational and safety systems can be connected, and by what means?

Telemetry, SCADA, asset management, incident records, inspection findings, access control, CCTV metadata, environmental sensors. Note API, export, or database access for each.

F2 Are assets and locations identified consistently across those systems?

If the same physical asset carries different identifiers in different systems, records cannot be joined and correlation is not possible. This is the most common blocker.

F3 How much retained history is available, and in what form?

F4 Can outcomes be established and linked to the assets and areas involved?

Without outcome records, validation cannot be performed at all — only monitoring value can be delivered.

F5 What is the data residency or sovereignty requirement?

☐ On-premise ☐ In-country cloud ☐ Private cloud ☐ No constraint

F6 Which deployment model is required?

☐ Edge ☐ Hybrid cloud ☐ Industrial private cloud ☐ On-premise

SECTION G**Ownership, action and governance****G1 Who receives intelligence alerts, and do they have authority to act?**

Name roles rather than individuals. Intelligence delivered to someone without authority to act produces frustration, not safety.

G2 Who owns the relationship between HSE, engineering and IT for this pilot?

Predictive safety sits across all three. Unowned, it stalls between them.

G3 Which HSE and engineering personnel are available to assess whether a discovered pattern is operationally meaningful?

Pattern validation cannot be done by the vendor alone. A statistically real pattern may be operationally trivial, and only domain experts can tell the difference.

G4 What data protection, retention and consultation requirements apply?

Include works council or union consultation where applicable.

G5 Is it agreed that predictive output will not be used in individual performance management?

☐ Agreed ☐ Requires escalation ☐ Not agreed

G5b Where worker monitoring is in scope — computer vision, movement or access data — record the agreed scope, purpose and consultation route.

If personnel come to believe that reporting increases scrutiny of their team, reporting rates decline and the data degrades. A predictive safety system can undermine its own inputs.

SECTION H**Pilot design summary**

Completed at the end of the session, from the answers above.

Predicted event

Unit and window

Scope of pilot

sites, assets, duration

Expected event count

Agreed error balance

Intervention rule

Known data gaps

Monitoring-only fallbackif validation is not yet
possible

A note on outcome

Where this questionnaire establishes that the data required for validation is not yet available, the recommended outcome is a monitoring-only deployment. That delivers operational value from integration, builds the structured evidence base that prediction requires, and defers predictive claims until they can be supported.

This is a legitimate and often preferable starting position. It is not a failed evaluation.

HSE lead

Date

IT / OT lead

Date

Facilitator

Date
