

REINHARDT PROGRAMME ASSURANCE MODEL

Validation Study

Three Major Infrastructure Programmes

Reinhardt Consulting Ltd | Version 1.0 | March 2026 | Published

Overview

This document presents the results of three validation assessments conducted using the Reinhardt Programme Assurance Model (V1.0). Each assessment applies the model's structured scoring framework to a major infrastructure programme using publicly available evidence, testing whether the model would have produced the correct delivery confidence signal at a defined point in time.

The three programmes were selected to test the model across the full spectrum of delivery outcomes: a successful programme, a programme in early structural difficulty, and a catastrophically failed programme. The assessments were conducted independently by Reinhardt Consulting Ltd.

Validation Summary

| Programme | Period | Model Score | Model Rating | Actual Outcome |
|----------------------------|---------|-------------|--------------|---|
| Heathrow Terminal 5 | 2007 | 86.9 | Green | Delivered on time and within budget |
| Crossrail (Elizabeth Line) | 2016–17 | 59.8 | Amber | Major delay announced Aug 2018; opened May 2022 |
| Berlin Brandenburg Airport | 2013 | 38.3 | Red | 9-year delay; costs trebled; opened Oct 2020 |

Methodology Note

All assessments use publicly available evidence only. Inputs reflect programme conditions at the stated assessment date. Evidence quality is classified as Objective Data, Documented Evidence, Observable Indicators, or Expert Judgement, and a confidence multiplier is applied accordingly. The model produces a Programme Assurance Score (0–100), a Delivery Confidence Rating (Green/Amber/Red), a Programme Failure Risk Score (0–100) based on six structural trigger indicators, and a benchmark comparison against typical megaproject performance ranges. It should be noted that these assessments are retrospective in nature. Inputs were assembled using publicly available evidence with knowledge of each programme's actual outcome. A prospective live deployment against a current programme is required before the validation claim can be considered fully independent.

REINHARDT PROGRAMME ASSURANCE MODEL

Validation Report

Heathrow Terminal 5

Assessment Period: 2007 (delivery/commissioning phase)

Actual Outcome: Delivered on time and within budget; opened March 2008

| Programme Assurance Score | Delivery Confidence Rating | Programme Failure Risk | Evidence Confidence |
|---------------------------|----------------------------|------------------------|-------------------------------|
| 86.9 | Green | Low (0/100) | 0.93 — Moderate–Strong |

Model Output: Successful delivery appears highly likely.

1. Assessment Context

This assessment was conducted using publicly available evidence from BAA project records, parliamentary reports, and academic documentation, representing programme conditions in 2007 during the delivery and commissioning phase. Terminal 5 opened in March 2008, on schedule, and is widely regarded as one of the most successfully delivered major infrastructure projects in UK history.

2. Key Programme Inputs

- Baseline budget: £4.3bn | Forecast final cost: £4.3bn (on budget)
- Schedule: no slip (forecast completion = baseline, December 2008)
- SPI: 0.97 | CPI: 0.98
- High risks: 4 | Risk mitigation coverage: 90%
- Decision turnaround: 4 days | Leadership changes: 0 in 12 months
- Unresolved engineering interfaces: 2

3. Dimension Score Profile

| Dimension | Score | Commentary |
|---------------------------------|------------|--------------------------------|
| Strategic Alignment | 100 | Strong confidence |
| Business Case Credibility | 100 | Strong confidence |
| Governance Effectiveness | 100 | Strong confidence |
| Sponsor & Leadership Capability | 100 | Strong confidence |
| Delivery Model & Organisation | 100 | Strong confidence |
| Schedule Credibility | 80 | Generally positive – some gaps |
| Cost & Contingency Health | 85 | Strong confidence |
| Risk & Issue Management | 87 | Generally positive – some gaps |
| Technical / Design Readiness | 87 | Strong confidence |

| | | |
|------------------------------------|------------|--------------------------|
| Stakeholder & Commercial Readiness | 100 | <i>Strong confidence</i> |
| Front-End Planning Maturity | 100 | <i>Strong confidence</i> |

4. Programme Failure Risk Triggers

| Failure Risk Indicator | Trigger Condition | Triggered | Current Value |
|-----------------------------|-------------------------------|-----------|-----------------|
| Front-End Planning Weakness | Score < 60 | No | 90.0 (adjusted) |
| Schedule Credibility Risk | Score < 65 | No | 80 |
| Contingency Depletion | Remaining contingency < 20% | No | 50% remaining |
| Leadership Instability | Leadership changes > 1 (12m) | No | 0 changes |
| Risk Exposure Trend | Trend = Deteriorating | No | Stable |
| Governance Delay | Decision turnaround > 10 days | No | 4 days |

5. Megaproject Benchmark Comparison

| Metric | Current Value | Benchmark Range | Assessment |
|---------------------------------------|---------------|-----------------|------------------------------|
| Design Maturity Before Delivery | 95% | 75%–90% | Above Benchmark Range |
| Schedule Performance Index (SPI) | 0.97 | 0.90–1.05 | Within Typical Range |
| Cost Performance Index (CPI) | 0.98 | 0.90–1.05 | Within Typical Range |
| Contingency Remaining During Delivery | 50% | 20%–40% | Above Benchmark Range |
| Decision Turnaround Time (days) | 4 | 3–7 | Within Typical Range |
| High Risk Count | 4 | 3–7 | Within Typical Range |
| Risk Mitigation Coverage | 90% | 75%–95% | Within Typical Range |
| Unresolved Engineering Interfaces | 2 | 0–5 | Within Typical Range |

6. Validation Commentary

The model correctly returned a Green rating with no failure risk triggers active. All six failure risk indicators were within safe thresholds. Schedule Credibility was the lowest-scoring dimension at 80, reflecting the known tight margins on the schedule in the final delivery phase — an accurate signal. The result is consistent with the programme’s actual outcome and demonstrates that the model does not generate false positives for well-managed programmes. Two metrics — design maturity and remaining contingency — appeared above the typical benchmark range. In both cases this reflected exceptionally strong programme preparation rather than a genuine risk signal, which is correctly reflected in the overall Green rating.

REINHARDT PROGRAMME ASSURANCE MODEL

Validation Report

Crossrail (Elizabeth Line)

Assessment Period: 2016–2017

Actual Outcome: Programme delayed August 2018; Elizabeth Line opened May 2022

| Programme Assurance Score | Delivery Confidence Rating | Programme Failure Risk | Evidence Confidence |
|---------------------------|----------------------------|--------------------------|------------------------|
| 59.8 | Amber | Elevated (60/100) | 0.89 — Moderate |

Model Output: Successful delivery is feasible, but significant issues exist.

1. Assessment Context

This assessment was conducted using publicly available evidence from NAO reports (including the NAO’s 2019 report “Crossrail: A Progress Update”), Public Accounts Committee publications, and Crossrail Learning Legacy documentation, representing programme conditions in 2016–17 — approximately 18–24 months before the August 2018 delay announcement. The model produced an Amber rating with Elevated Failure Risk, consistent with what is now documented about the programme’s structural condition during that period.

2. Key Programme Inputs

- Baseline budget: £14.8bn | Forecast final cost: £17.0bn (>10% overrun threshold)
- Schedule slip: >90 days (forecast completion 2019 vs baseline 2018)
- SPI: 0.86 | CPI: 0.91
- High risks: 11 | Risk mitigation coverage: 70%
- Decision turnaround: 11 days | Leadership changes: 2 in 12 months
- Unresolved engineering interfaces: 7

3. Dimension Score Profile

| Dimension | Score | Commentary |
|---------------------------------|------------|-----------------------------------|
| Strategic Alignment | 100 | Strong confidence |
| Business Case Credibility | 87 | Generally positive – some gaps |
| Governance Effectiveness | 60 | Major delivery risk |
| Sponsor & Leadership Capability | 67 | Material concerns – action needed |
| Delivery Model & Organisation | 65 | Material concerns – action needed |
| Schedule Credibility | 48 | Major delivery risk |
| Cost & Contingency Health | 65 | Material concerns – action needed |
| Risk & Issue Management | 57 | Major delivery risk |

| | | |
|------------------------------------|-----------|--|
| Technical / Design Readiness | 60 | <i>Major delivery risk</i> |
| Stakeholder & Commercial Readiness | 73 | <i>Material concerns – action needed</i> |
| Front-End Planning Maturity | 73 | <i>Material concerns – action needed</i> |

4. Programme Failure Risk Triggers

| Failure Risk Indicator | Trigger Condition | Triggered | Current Value |
|-----------------------------|-------------------------------|------------|-----------------|
| Front-End Planning Weakness | Score < 60 | No | 65.7 (adjusted) |
| Schedule Credibility Risk | Score < 65 | YES | 48 |
| Contingency Depletion | Remaining contingency < 20% | No | 40% remaining |
| Leadership Instability | Leadership changes > 1 (12m) | YES | 2 changes |
| Risk Exposure Trend | Trend = Deteriorating | YES | Deteriorating |
| Governance Delay | Decision turnaround > 10 days | YES | 11 days |

5. Megaproject Benchmark Comparison

| Metric | Current Value | Benchmark Range | Assessment |
|---------------------------------------|---------------|-----------------|------------------------------|
| Design Maturity Before Delivery | 75% | 75%–90% | Within Typical Range |
| Schedule Performance Index (SPI) | 0.86 | 0.90–1.05 | Below Benchmark Range |
| Cost Performance Index (CPI) | 0.91 | 0.90–1.05 | Within Typical Range |
| Contingency Remaining During Delivery | 40% | 20%–40% | Within Typical Range |
| Decision Turnaround Time (days) | 11 | 3–7 | Above Benchmark Range |
| High Risk Count | 11 | 3–7 | Above Benchmark Range |
| Risk Mitigation Coverage | 70% | 75%–95% | Below Benchmark Range |
| Unresolved Engineering Interfaces | 7 | 0–5 | Above Benchmark Range |

6. Validation Commentary

The model correctly identified Schedule Credibility, Governance Effectiveness and Risk & Issue Management as the three lowest-scoring dimensions — precisely the areas identified in subsequent post-mortems as the primary drivers of delay. Four of six failure risk triggers were active and the Programme Failure Risk Score reached 60/100 — Elevated — approximately 18–24 months before the delay became public. This would have constituted an actionable early warning signal at a point when programme reporting remained broadly positive.

REINHARDT PROGRAMME ASSURANCE MODEL

Validation Report

Berlin Brandenburg Airport (BER)

Assessment Period: 2013 (mid-construction, post-first failed opening)

Actual Outcome: 9-year delay; opened October 2020; costs more than trebled

| Programme Assurance Score | Delivery Confidence Rating | Programme Failure Risk | Evidence Confidence |
|---------------------------|----------------------------|---------------------------|------------------------|
| 38.3 | Red | Critical (100/100) | 0.89 — Moderate |

Model Output: *Successful delivery appears highly unlikely without fundamental intervention.*

1. Assessment Context

This assessment was conducted using publicly available evidence from German parliamentary inquiries, Flughafen Berlin Brandenburg GmbH reports, and academic case studies, representing programme conditions in 2013 — after the May 2012 opening had been cancelled days before it was due to take place. The airport eventually opened in October 2020, nine years late. Final costs exceeded €7 billion against an original budget of approximately €2 billion.

2. Key Programme Inputs

- Original budget: ~€2bn | Estimated final cost: >€7bn (>250% overrun)
- Schedule slip: >3,000 days (opened October 2020 vs original May 2012)
- SPI: 0.70 | CPI: 0.75
- High risks: 18 | Risk mitigation coverage: 50%
- Decision turnaround: 18 days | Leadership changes: 3 in 12 months
- Unresolved engineering interfaces: 15

3. Dimension Score Profile

| Dimension | Score | Commentary |
|---------------------------------|-----------|--|
| Strategic Alignment | 80 | <i>Generally positive – some gaps</i> |
| Business Case Credibility | 67 | <i>Material concerns – action needed</i> |
| Governance Effectiveness | 40 | <i>Critical – immediate intervention</i> |
| Sponsor & Leadership Capability | 40 | <i>Critical – immediate intervention</i> |
| Delivery Model & Organisation | 30 | <i>Critical – immediate intervention</i> |
| Schedule Credibility | 30 | <i>Critical – immediate intervention</i> |
| Cost & Contingency Health | 40 | <i>Major delivery risk</i> |
| Risk & Issue Management | 43 | <i>Critical – immediate intervention</i> |
| Technical / Design Readiness | 33 | <i>Critical – immediate intervention</i> |

| | | |
|------------------------------------|-----------|--|
| Stakeholder & Commercial Readiness | 40 | <i>Critical – immediate intervention</i> |
| Front-End Planning Maturity | 43 | <i>Critical – immediate intervention</i> |

4. Programme Failure Risk Triggers

| Failure Risk Indicator | Trigger Condition | Triggered | Current Value |
|-----------------------------|-------------------------------|------------|-----------------|
| Front-End Planning Weakness | Score < 60 | YES | 38.7 (adjusted) |
| Schedule Credibility Risk | Score < 65 | YES | 30 |
| Contingency Depletion | Remaining contingency < 20% | YES | 10% remaining |
| Leadership Instability | Leadership changes > 1 (12m) | YES | 3 changes |
| Risk Exposure Trend | Trend = Deteriorating | YES | Deteriorating |
| Governance Delay | Decision turnaround > 10 days | YES | 18 days |

5. Megaproject Benchmark Comparison

| Metric | Current Value | Benchmark Range | Assessment |
|---------------------------------------|---------------|-----------------|------------------------------|
| Design Maturity Before Delivery | 65% | 75%–90% | Below Benchmark Range |
| Schedule Performance Index (SPI) | 0.70 | 0.90–1.05 | Below Benchmark Range |
| Cost Performance Index (CPI) | 0.75 | 0.90–1.05 | Below Benchmark Range |
| Contingency Remaining During Delivery | 10% | 20%–40% | Below Benchmark Range |
| Decision Turnaround Time (days) | 18 | 3–7 | Above Benchmark Range |
| High Risk Count | 18 | 3–7 | Above Benchmark Range |
| Risk Mitigation Coverage | 50% | 75%–95% | Below Benchmark Range |
| Unresolved Engineering Interfaces | 15 | 0–5 | Above Benchmark Range |

6. Validation Commentary

The model correctly returned a Red rating with a maximum Programme Failure Risk Score of 100. All six failure risk triggers fired simultaneously. Every benchmark metric was outside the normal range — most significantly: design maturity at 65% (below the 75% threshold), SPI at 0.70, CPI at 0.75, contingency at just 10% remaining, decision turnaround at 18 days, 18 high risks, mitigation coverage at only 50%, and 15 unresolved engineering interfaces. The result is consistent with the programme's actual condition in 2013 — one of the most comprehensively failed infrastructure projects in European history.

Conclusions

Across three programmes spanning successful delivery, early structural difficulty, and catastrophic failure, the Reinhardt Programme Assurance Model produced directionally correct results in each case:

- Heathrow T5 (Green, 86.9): No failure triggers fired. All dimensions in the healthy range. The result is consistent with a programme that delivered on time and on budget — demonstrating the model does not generate false positives for well-managed programmes.
- Crossrail (Amber, 59.8): Four of six failure triggers fired and the Programme Failure Risk Score reached 60/100 — Elevated — approximately 18–24 months before the delay became public. Schedule Credibility, Governance Effectiveness and Risk & Issue Management were identified as the three weakest dimensions — consistent with the drivers of failure documented in subsequent post-mortems. This would have constituted an actionable early warning signal at a point when programme reporting remained broadly positive.
- Berlin Brandenburg Airport (Red, 38.3): All six failure triggers fired at maximum severity. Programme Failure Risk Score reached 100. Every benchmark metric was outside the normal range. The result is consistent with the actual condition of a programme that became one of the most comprehensively documented infrastructure failures in European history.

The validation demonstrates that the model’s structured evidence-weighting approach, failure risk trigger system, and benchmark comparison layer each contribute meaningfully to the overall assessment. No single dimension or metric produces the result; it is the combination of signals across the framework that generates the rating — consistent with the multi-causal nature of programme failure in practice. A prospective pilot deployment is currently being sought. Enquiries regarding the application of the Reinhardt Programme Assurance Model to live programmes are welcome via reinhardtconsulting.co.uk.