Terence Kok · Framework reference brief · Version 1.1 · First published 5 June 2026 · Revised 12 September 2026 · CC BY-NC 4.0 · terencekok.com/frameworks/measurement/

ARCHITECTURE FRAMEWORK

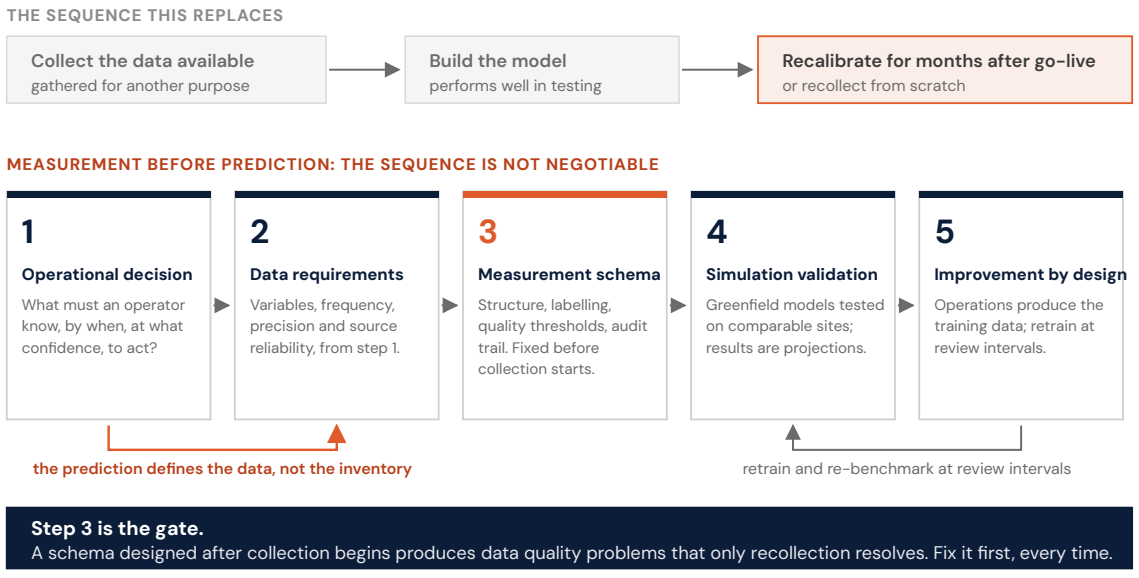
# Measurement-Before-Prediction Architecture

Developed for infrastructure and industrial AI deployments where prediction accuracy depends on a measurement schema designed before data collection begins.

Most infrastructure AI deployments fail not because the models are wrong but because the measurement framework was not designed before the data was collected. When the data schema, collection frequency, and quality standards are designed to serve a different purpose, AI predictions built on that data require months of recalibration after go-live. This framework reverses that sequence.

## REFERENCE DIAGRAM

The conventional collect-then-model sequence and the five-step sequence that replaces it. Step 3, the schema, is the gate: it is fixed before any collection begins.



STEP 1

### Define what the prediction needs to produce

Start with the operational decision, not the data. What does an operator need to know, by when, with what confidence level, to take a meaningful action? This definition determines what the prediction must produce, and therefore what data the prediction requires as input.

STEP 2

### Work backwards to the data requirements

From the prediction output requirements, define the input data: variables, update frequency, precision, and source reliability. This becomes the data collection specification: not an inventory of available data, but a specification of required data derived from the prediction requirements.

STEP 3

### Design the measurement schema before collection begins

Define data structure, labelling conventions, quality thresholds, and audit trail requirements before any collection infrastructure is built. A measurement schema designed retrospectively around collected data produces persistent data quality problems that cannot be resolved without recollection.

STEP 4

### Validate predictions against simulation before go-live

For greenfield deployments with no historical data, validate initial prediction models against simulation data from comparable operational environments. Treat simulation-validated figures as projections, not production metrics, and design the live data collection to confirm or revise them from day one.

STEP 5

### Build model improvement into operational design

The data collected during operations is the asset that improves prediction accuracy over time. Design operational workflows to produce data that improves model accuracy as a by-product of normal operations, not as a separate data collection exercise. Define the review intervals at which prediction models are retrained and accuracy benchmarks are reassessed.

**Application principle:** The sequence is non-negotiable. Changing the measurement schema after data collection begins requires recollection or produces a persistent gap between historical and current data. Building prediction models on data collected for a different purpose produces models that perform well in testing and fail under operational conditions. Design the measurement framework first, every time.

## CITE THIS FRAMEWORK

| VERSION | FIRST PUBLISHED | LAST REVISED | LICENCE |
|---|---|---|---|
| 1.1 | 5 June 2026 | 12 September 2026 | CC BY-NC 4.0 |

DOI
10.5281/zenodo.22722225

APA | Harvard | Chicago | BibTeX

Kok, T. (2026). Measurement-Before-Prediction Architecture (version 1.1) [Framework reference]. terencekok.com. https://doi.org/10.5281/zenodo.22722225

Copy citation