

Measure output, outcome, and AI performance

Set targets before the pilot starts. Each selected initiative needs at least one KPI, one OKR, and one efficacy measure, plus exit criteria and re-review triggers. Without efficacy, a tool can hit its usage target and its outcome goal while making poor judgments.

Measure	Answers	Example
KPI (output)	Is it being used and delivered as planned?	80% of advisors use the early-alert dashboard weekly.
OKR (outcome)	Is it moving the goal it was meant to serve?	First-year retention rises by 3 percentage points.
Efficacy (AI performance)	Is the AI accurate, reliable, and fair?	At least 80% of AI risk flags are confirmed by advisors, with no gap across student groups.
Exit criteria	When do we pause or stop early?	Efficacy below 70% for two consecutive months; any confirmed data breach.
Re-review triggers	When must it be scored again?	New data source, expanded population, vendor model change, or a serious complaint.

GOOD MEASURES ARE

- Specific: one number, one population, one time frame.
- Baselined: record where you start before the pilot.
- Owned: one named person reports each measure.
- Proportionate: Low-band pilots need one of each; High and Critical need fuller efficacy and equity checks.
- Checked for equity: break efficacy results down by relevant groups.

REFLECT

At the end of the pilot, compare actuals to targets for all three measures, record lessons, and decide: scale, iterate, pause, or stop.