Data Product Canvas

Motivation

Problem

What problem are we solving?

Who are the target users and beneficiaries?

Why is the current solution insufficient?

Business Value

How will this reduce costs (e.g., speed, fewer errors)?

How will this increase revenue (e.g., new product feature, retention)?

What is the scale of impact (number of users, volume)?

Does this enable future projects (e.g., reusing data infrastructure)?

Does this create a competitive advantage or is needed to just keep up?

Success Criteria

Definition of done and measurable KPIs.

Guardrail metrics that must not degrade.

Current baselines for comparison.

Approach

Deliverables

Software system to make predictions continuously or insights from a one-off analysis?

Integration

How does the solution fit into existing workflows?

What input data is required and where does it come from?

What outputs are produced and how are they utilized?

Build or Buy?

Does this require unique domain knowledge?

Required Resources & Biggest Risks

Data

How much data is available (incl. rare events & labels)?

Do subject matter experts believe the data is sufficient, or are key variables missing?

How easy is it to access and consolidate the data?

What effort is required to build a pipeline for continuous data collection?

Deployment & Maintenance

Where should the model run (cloud, on-prem, edge)?

What infrastructure or hardware (e.g., GPUs) is needed?

How will the model be monitored, retrained, and maintained over time?

Will the solution scale as data and usage grow?

Risks & Governance

What regulations apply (e.g., GDPR, HIPAA, EU AI Act)?

Are there risks around data privacy, fairness, or transparency?

What are the consequences if the model makes an error?

Should humans remain in the loop for critical decisions?

Could users try to game or exploit the system?

What barriers might prevent user trust or adoption?

Under what conditions should the project be stopped (kill switch criteria)?

ML

Is there a simpler solution without ML?

Has a similar problem been solved with ML before?

People

Who owns the solution and its outcomes?

What expertise is required, and do we need support from other teams or external partners?

Will users need training to adopt it effectively?