

Data Product Heuristics - Checklist

Data product: _____ Archetype: ☐ source-aligned ☐ aggregate ☐ consumer-aligned Date: _____

Start with the fundamentals, then add the heuristics for the archetype. Heuristics are guiding questions, not rigid rules.

Fundamentals

1. Clearly defined consumer & use cases

Describe its main purpose in one sentence.

- ☐ Can you describe the main purpose in one sentence?
- ☐ Are there any specific teams or roles that want to use this data product right now?

2. Stable ownership

One domain, one team.

- ☐ Is one specific domain or team accountable for semantics, quality, and operations?
- ☐ Would the owner credibly handle future changes?

3. Consistent data quality

Quality attributes consistent across all output ports.

- ☐ Are data quality attributes consistent across output ports?

4. Low integration burden

Smallest useful standalone unit.

- ☐ Is this the smallest useful standalone unit that does not force consumers to stitch data products together?
- ☐ Can a typical consumer immediately start using this data product meaningfully on their own?

5. Bounded scope

Only what is needed, not what might be useful someday.

- ☐ Does the data product include only what is needed for its purpose?
- ☐ Is it limited to only including things that are useful in the present?

Source-aligned

1. Semantic coherence

All data required for interpretation.

- ☐ Does the data product make sense on its own, or does it require the other parts of the source data?
- ☐ Does it feel like a cohesive, integrated whole rather than a random collection of related items?

2. Single business domain

One module. Not all data.

- ☐ Does the decomposition follow meaningful domain modules rather than whole systems?
- ☐ Does the data contain only internal or also cross-domain context?

3. Source data blast radius

Different source, different data product.

- ☐ Do changes on the data source impact only this data product directly?

Aggregate

1. Value & reuse

Build it when 3+ teams need it.

- ☐ Are there more than two teams that need the same derived view with identical meaning?
- ☐ Would teams repeatedly build the same integration or calculation without the aggregate?
- ☐ Does value emerge only after combining sources?

2. Cost & complexity

- ☐ Is the derivation expensive (feature engineering, entity matching, deduplication, cross-source joins)?
- ☐ Is there someone in the company willing to bear the costs of this data product?

3. Scope & governance discipline

General-purpose. Strictly scoped.

- ☐ Is the scope tight enough so the data product is not drifting toward a mini data warehouse?
- ☐ Is the outcome valuable enough to justify the required strong governance?
- ☐ Can the owning team maintain the integrated semantics, despite spanning multiple sources?

Consumer-aligned

1. Clear purpose

One sentence. Verb plus object.

- ☐ Can this purpose be expressed as a verb + object sentence?

2. Natural, focused size

Sized by the job, not the artifact.

- ☐ Does the data product support a single decision context that results in one or more related dashboards, reports, or outputs?

3. Meaningful boundaries

Process lines, not system lines.

- ☐ Does the boundary follow a process and not a system boundary?
- ☐ Does the decomposition reflect how a consumer acts or decides, not how data happens to be stored?

4. Business consumers

Name the team. Then the data product.

- ☐ Will your data product be used by business users, data analysts, data scientists, or applications? Who exactly, and for what?