

---

# The Data Cards Playbook

A toolkit for purposeful and people-centric dataset documentation for transparency in AI systems.

<https://pair-code.github.io/datacardsplaybook/>

#datacardsplaybook



THE DATA CARDS PLAYBOOK

# Introduction

01 Ask

**02 Inspect**

03 Answer

04 Audit



# Dimensions Evaluation

## IN THIS SECTION

Evaluate your Data Card or template on the basis of five dimensions. Use these to identify action items to refine your Data Card template



## INSTRUCTIONS

Use the Dimensions to uncover directional insights using qualitative approximations. Find immediate opportunities to start refining your Data Card or template.

## OUTCOMES

A first pass of refining your Data Cards or templates. Repeat as needed!

## ACTIVITY LEVEL

Moderate

**Dimensions** help us ascertain the usefulness of Data Card for arriving at acceptable conclusions about the dataset.

Can the Data Card template sufficiently **represent the dataset**? Does it provides **additional context** that a reader might require for responsible decision making?

**Consider:** Some information themes in your Data Card will contribute to multiple dimensions, while others will satisfy a single dimension.



# 5 Dimensions for Transparency

## ACCOUNTABILITY

Demonstrates adequate ownership, reflection, reasoning, and systematic decision making surrounding the dataset and its use by the people who own and maintain it. producers.

## UTILITY OR USE

Provides details that satisfy the information needs of the readers for a responsible decision-making process that establishes the suitability of dataset for their tasks and goals.

## QUALITY

Summarizes the rigor, integrity and completeness of the dataset, often communicated in a manner that is accessible and understandable to readers of different backgrounds.

## IMPACT OR CONSEQUENCES OF USE

Sets expectations for outcomes when using and managing the dataset.

Acknowledges any first or second-order consequences that could negatively impact the readers' goals.

## RISK AND RECOMMENDATIONS

Makes readers aware of known and potential risks and limitations, that stem from the provenance, representation, use, or context of use.

Provides enough information and alternatives to help readers make responsible trade-offs.

## Accountability

Demonstrates adequate ownership, reflection, reasoning, and systematic decision making by its producers.

**Consider:** Authorship, Maintenance, Responsibility, Intentionality.

## Proficiency

How much fluency and expertise might be needed to arrive at an informed conclusion about this Dimension from the Data Card?

### Data Fluency:

1 2 3 4 5

No Fluency

Expert Fluency

### Domain Expertise:

1 2 3 4 5

No Expertise

Expert

**Rationale** How well is this dimension currently supported and why?

- Add rationale and reasoning in or against favor here.

**Evidence** What questions or sections in support this dimension and how?

- Add examples from sections or questions here

**Action Items** What steps must be taken to refine or address shortcomings?

- Action items

## Utility or Use

Provides details that satisfy the needs of the readers' decision-making process responsibly to establish the suitability of system for their tasks and goals.

**Consider:** Feedback, Usability Context, Downstream Applications, Societal Needs.

## Proficiency

How much fluency and expertise might be needed to arrive at an informed conclusion about this Dimension from the Data Card?

### Data Fluency:

1 2 3 4 5

No Fluency

Expert Fluency

### Domain Expertise:

1 2 3 4 5

No Expertise

Expert

**Rationale** How well is this dimension currently supported and why?

- Add rationale and reasoning in or against favor here.

**Evidence** What questions or sections in support this dimension and how?

- Add examples from sections or questions here

**Action Items** What steps must be taken to refine or address shortcomings?

- Action items



## Quality

Summarizes the rigor, integrity and completeness of the system, often communicated in a manner that is accessible and understandable to many readers.

**Consider:** Validity, Reliability, Integrity, Reproducibility.

## Proficiency

How much fluency and expertise might be needed to arrive at an informed conclusion about this Dimension from the Data Card?

### Data Fluency:

1

2

3

4

5

No Fluency

Expert Fluency

### Domain Expertise:

1

2

3

4

5

No Expertise

Expert

**Rationale** How well is this dimension currently supported and why?

- Add rationale and reasoning in or against favor here.

**Evidence** What questions or sections in support this dimension and how?

- Add examples from sections or questions here

**Action Items** What steps must be taken to refine or address shortcomings?

- Action items

## Impact and Consequences

Sets expectations for outcomes when using and managing the system. Acknowledges any first or second-order consequences that could negatively impact the readers' goals.

**Consider:** Efficacy in Use, Group Benefit, Relevance of Outcomes, Implications of Deviations.

## Proficiency

How much fluency and expertise might be needed to arrive at an informed conclusion about this Dimension from the Data Card?

### Data Fluency:

1 2 3 4 5

No Fluency

Expert Fluency

### Domain Expertise:

1 2 3 4 5

No Expertise

Expert

**Rationale** How well is this dimension currently supported and why?

- Add rationale and reasoning in or against favor here.

**Evidence** What questions or sections in support this dimension and how?

- Add examples from sections or questions here

**Action Items** What steps must be taken to refine or address shortcomings?

- Action items

## Risk and Recommendations

Makes readers aware of known potential risks and limitations, stemming from provenance, representation, use, or context of use. Provides enough information and alternatives to help readers make responsible trade-offs.

**Consider:** Risk Magnitude, Mitigations, Recommendations, Group Harm.

## Proficiency

How much fluency and expertise might be needed to arrive at an informed conclusion about this Dimension from the Data Card?

### Data Fluency:

1

2

3

4

5

No Fluency

Expert Fluency

### Domain Expertise:

1

2

3

4

5

No Expertise

Expert

**Rationale** How well is this dimension currently supported and why?

- Add rationale and reasoning in or against favor here.

**Evidence** What questions or sections in support this dimension and how?

- Add examples from sections or questions here

**Action Items** What steps must be taken to refine or address shortcomings?

- Action items

# Final checks

YOU SHOULD NOW HAVE

—

- ✔ Evaluated your Data Card Template for five dimensions of transparency
- ✔ Uncovered directional insights about content quality, readability, and utility
- ✔ Identified action items for a more robust and refined template
- ✔ A concrete plan to refine your Data Card Template





#datacardsplaybook



[The Data Cards Playbook](#) is an adaptable toolkit of participatory activities, conceptual frameworks, and guidance that support Responsible AI practices for transparency in dataset documentation.

If you've adapted, implemented, or have feedback for this guidance, we'd love to hear from you at <https://github.com/pair-code/datacardsplaybook>.

Find the complete playbook at  
<https://pair-code.github.io/datacardsplaybook>



The [Data Cards Playbook](#) <sup>↗</sup> by [the People + AI Research Initiative](#) <sup>↗</sup> at [Google Research](#) <sup>↗</sup> is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. You are free to share and adapt this work under the [appropriate license terms](#) <sup>↗</sup>.