The Data Cards Playbook

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

#datacardsplaybook

Introduction
01 Ask
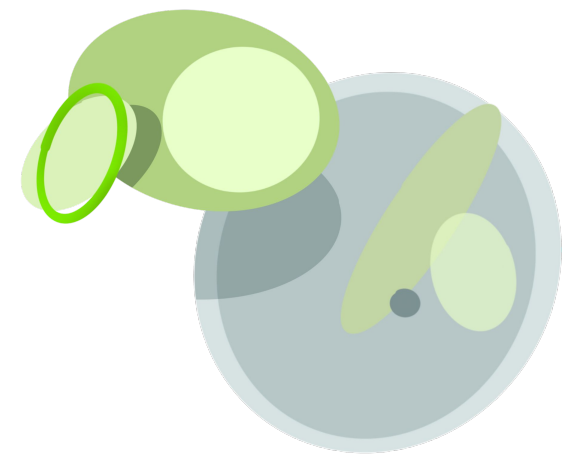# 02 Inspect
03 Answer
04 Audit

Dimensions Rubric

IN THIS SECTION

Have two reviews complete an evaluation of a completed Data Card to help dataset producers improve their Data Cards.

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 a 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.

Overall Rating

Pick one:

Poor

Borderline

Average

Solid

Outstanding

Feedback + Evidence Provide rationale and evidence that supports your rating of the Data Card.

- Add rationale and reasoning in or against favor here, along with corresponding evidence from the Data Card.

Recommendations + Comments Any additional comments or recommended actions. What can publishers do to take their Data Card to the next level?

- Provide your recommendations, action items, and other comments.

5 min

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.

Overall Rating

Pick one:

Poor

Borderline

Average

Solid

Outstanding

Feedback + Evidence Provide rationale and evidence that supports your rating of the Data Card.

- Add rationale and reasoning in or against favor here, along with corresponding evidence from the Data Card.

Recommendations + Comments Any additional comments or recommended actions. What can publishers do to take their Data Card to the next level?

- Provide your recommendations, action items, and other comments.

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.

Overall Rating

Pick one:

Poor

Borderline

Average

Solid

Outstanding

Feedback + Evidence Provide rationale and evidence that supports your rating of the Data Card.

- Add rationale and reasoning in or against favor here, along with corresponding evidence from the Data Card.

Recommendations + Comments Any additional comments or recommended actions. What can publishers do to take their Data Card to the next level?

- Provide your recommendations, action items, and other comments.

5
min

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.

Overall Rating

Pick one:

Poor

Borderline

Average

Solid

Outstanding

Feedback + Evidence Provide rationale and evidence that supports your rating of the Data Card.

- Add rationale and reasoning in or against favor here, along with corresponding evidence from the Data Card.

Recommendations + Comments Any additional comments or recommended actions. What can publishers do to take their Data Card to the next level?

- Provide your recommendations, action items, and other comments.

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.

Overall Rating

Pick one:

Poor

Borderline

Average

Solid

Outstanding

Feedback + Evidence Provide rationale and evidence that supports your rating of the Data Card.

- Add rationale and reasoning in or against favor here, along with corresponding evidence from the Data Card.

Recommendations + Comments Any additional comments or recommended actions. What can publishers do to take their Data Card to the next level?

- Provide your recommendations, action items, and other comments.

Final checks

–

✔ Evaluated a Data Card Template for five dimensions of transparency

✔ Provided actionable feedback for dataset producers for a more robust and refined Data Card

✔ The ability to gauge the overall quality of Data Cards produced by your organization over time

#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 ↗](https://github.com/pair-code/datacardsplaybook).

Find the complete playbook at
[https://pair-code.github.io/datacardsplaybook ↗](https://pair-code.github.io/datacardsplaybook)