

---

# 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

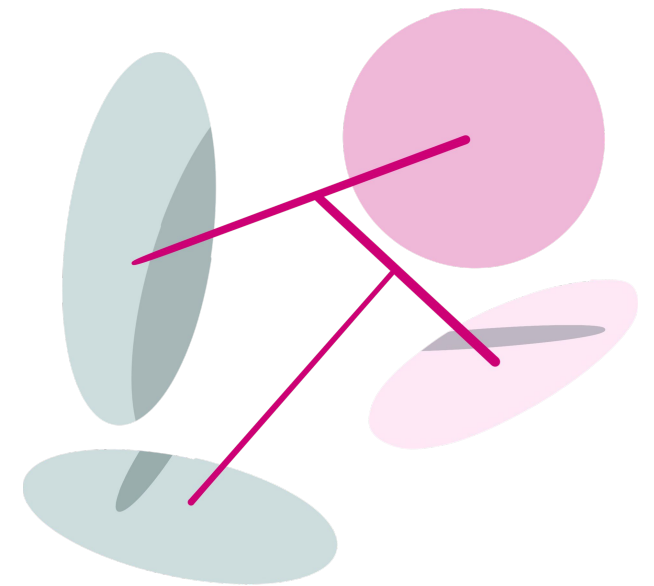
# Introduction

**01 Ask**

**02 Inspect**

**03 Answer**

**04 Audit**



# Priority Matrix

## IN THIS SECTION

If you are unable to decide on your most important target audiences, use meaningful dimensions to set priority.

## INSTRUCTIONS

Decide criteria that makes one agent more or less important or critical, and place agents on a 2×2 matrix based on the criteria.

## OUTCOMES

A prioritized list of agents who are the target audience for whom your Data Cards must work.

## ACTIVITY LEVEL

Intermediate

# What are **important criteria** to help **prioritize** one agent over the other?

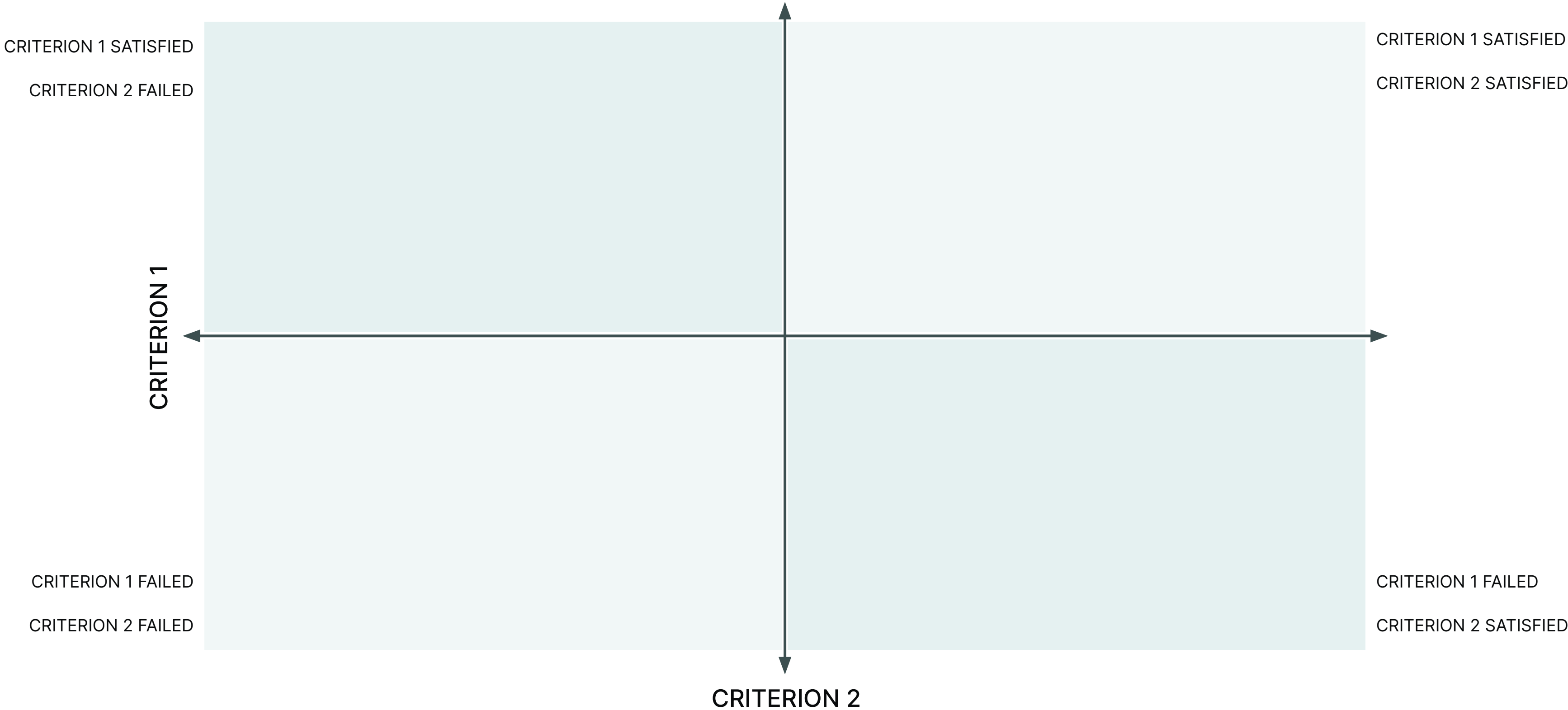
Consider that not all agents engage equally with Datasets and Data Cards. Inversely, the same Data Card may not work equally effectively for everyone.

1. 🖊️

2. 🖊️

Tie breaker: 🖊️

# Plot your agents in the appropriate quadrants reflecting your criteria



# What are **important criteria** to help **prioritize** one agent over the other?

Consider that not all agents engage equally with Datasets and Data Cards. Inversely, the same Data Card may not work equally effectively for everyone.

## **Frequency of Use**

How frequently will the agent use your dataset in their work?

## **Importance of Review**

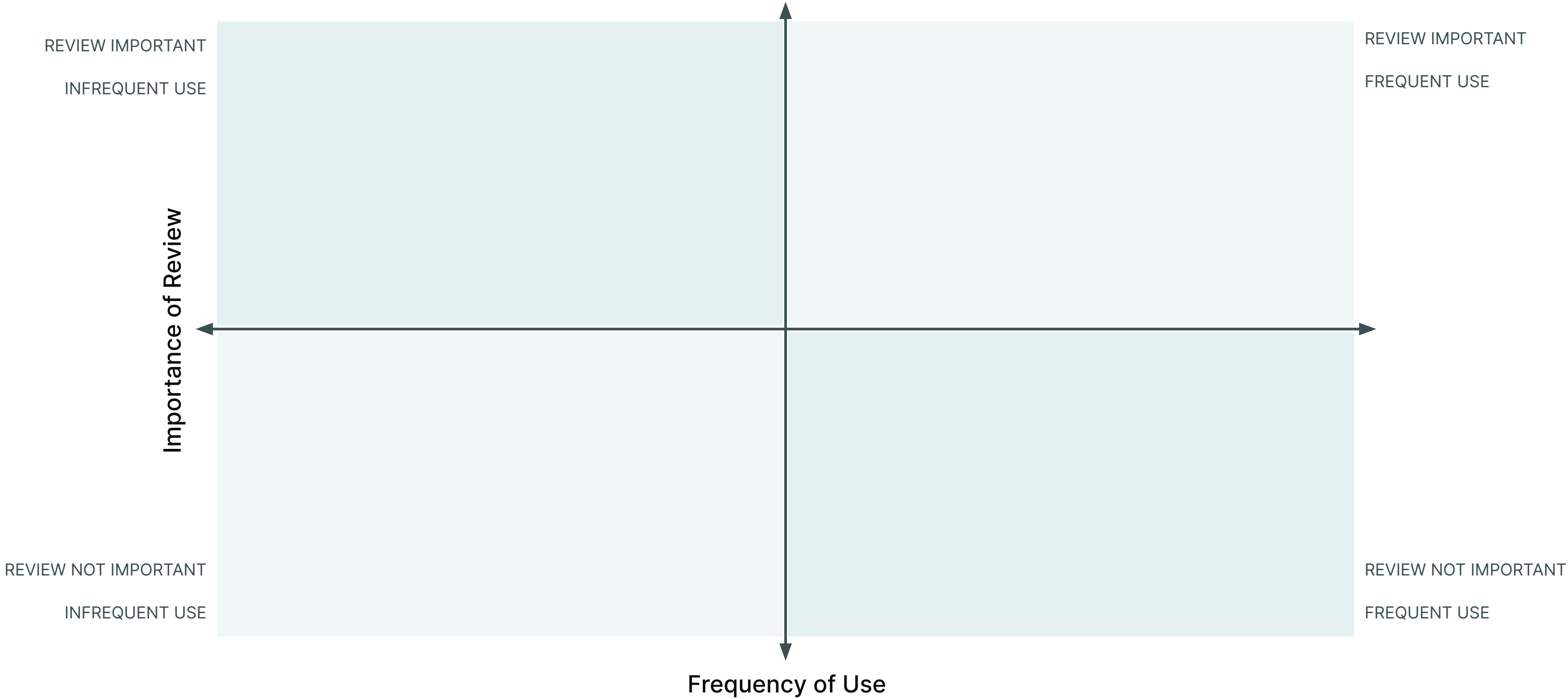
How important is the agent's review, approval, or feedback for the dataset's success?

## **Tie breaker: Impact on Agent**

How might the agent be impacted once your dataset is used in production?



# Importance of Review & Frequency of Use





# Summarize: Who are the target readers (agents) of your Data Card?

Top Priority:

1. 🖊️

Medium Priority:

2. 🖊️

3. 🖊️

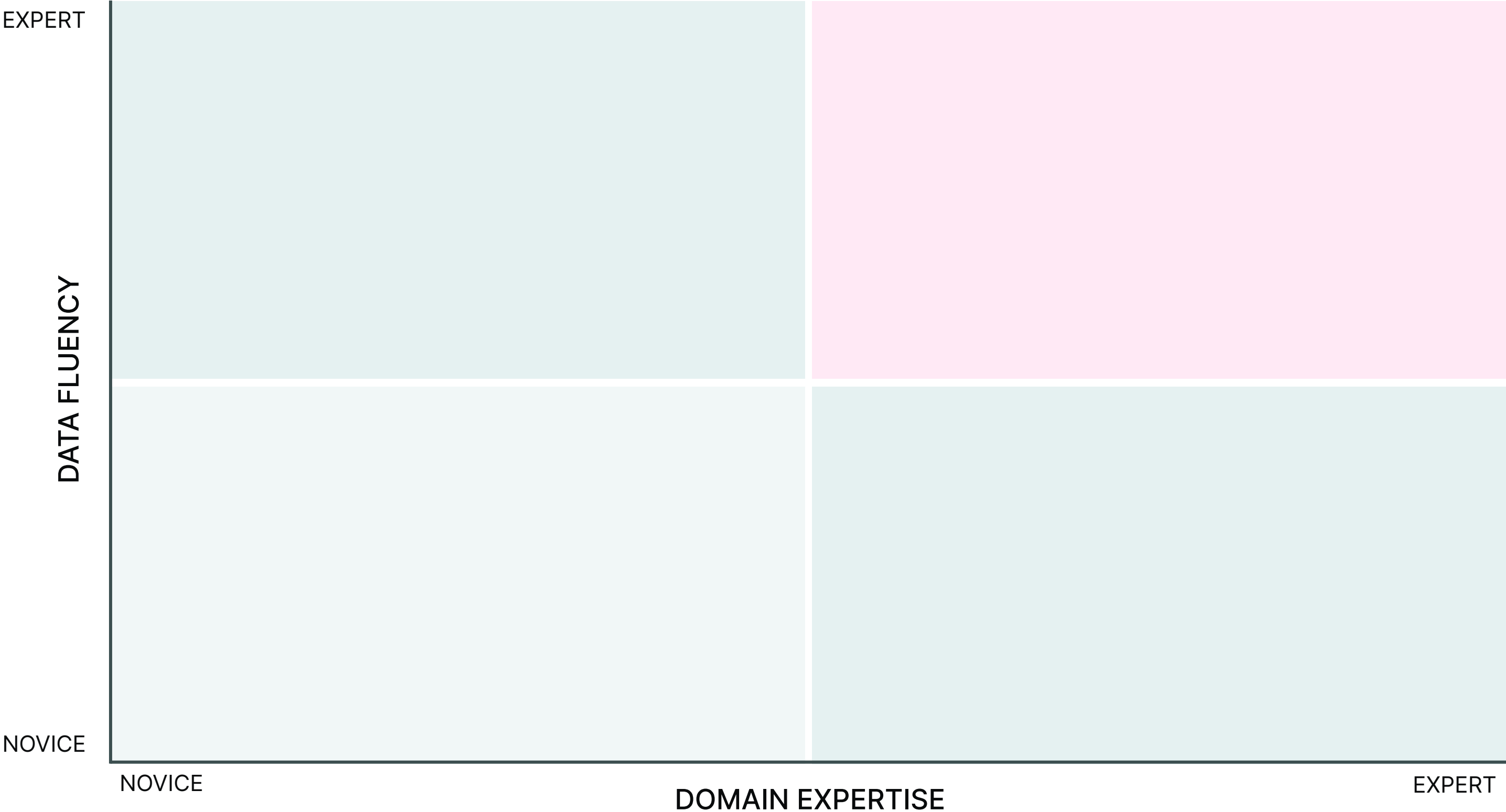
4. 🖊️

5. 🖊️

6. 🖊️



# Agent Characteristics: Data Fluency and Domain Expertise



# Checklist

YOU SHOULD NOW HAVE

—

- ✔ Agreed on your highest priority agents
- ✔ Defined the importance of each agent to your dataset
- ✔ A clear agreement of the data fluency of the agents
- ✔ A clear agreement of the domain expertise of the agents



#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>.