# How to write a killer README

Emily Omier

# About Me

Positioning / product strategy consultant for open source companies

Host of The Business of Open Source podcast

Open Source Founders Summit co-organizer

Your README is more important than you think.

# What is the purpose of your README?

1. Tell people who read it how to install your software.

2. Help people understand if your project will provide the outcome / solve the problem that they want it to.

3. Help people quickly figure out if your project is a poor fit for them.

4. Make people give you a star.

5. Help people see what others are doing with your project.

6. All of the above

A README is like a homepage for your project.

## Installation

- Container
- Binary/OS packages
- Helm Chart

# A README should

- Establish what your project is, ie your market category
- Communicate the value someone gets from using it
- Communicate how it’s different from the other options
- Make it clear the use cases the project is ideal for
- Ideally, give examples of how people use the project
- … and then give them a way to easily get started

# And if you've got a company…

- The ReadMe is where you establish a connection between your project and your products and/or services
- Tell people that your company exists!

# Argilla

## Build high quality datasets for your AI models



Argilla is a collaboration tool for AI engineers and domain experts who need to build high-quality datasets for their projects.

If you just want to get started, deploy Argilla on Hugging Face Spaces. Curious, and want to know more? Read our documentation.

Or, play with the Argilla UI by signing in with your Hugging Face account:

Logo

Outcome you get from using Argilla

High-level positioning

Argilla was acquired by Hugging Face...

## Why use Argilla?

Argilla can be used for collecting human feedback for a wide variety of AI projects like traditional NLP (text classification, NER, etc.), LLMs (RAG, preference tuning, etc.), or multimodal models (text to image, etc.). Argilla's programmatic approach lets you build workflows for continuous evaluation and model improvement. The goal of Argilla is to ensure your data work pays off by quickly iterating on the right data and models.

### Improve your AI output quality through data quality

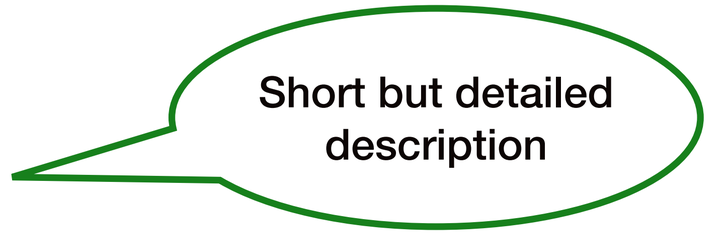
Compute is expensive and output quality is important. We help you focus on data, which tackles the root cause of both of these problems at once. Argilla helps you to **achieve and keep high-quality standards** for your data. This means you can improve the quality of your AI output.

### Take control of your data and models

Most AI tools are black boxes. Argilla is different. We believe that you should be the owner of both your data and your models. That's why we provide you with all the tools your team needs to **manage your data and models in a way that suits you best**.

### Improve efficiency by quickly iterating on the right data and models

Gathering data is a time-consuming process. Argilla helps by providing a tool that allows you to **interact with your data in a more engaging way**. This means you can quickly and easily label your data with filters, AI feedback suggestions and semantic search. So you can focus on training your models and monitoring their performance.




Mix of social proof and example use case

## What do people build with Argilla?

### Open-source datasets and models

The community uses Argilla to create amazing open-source datasets and models.

- Cleaned UltraFeedback dataset used to fine-tune the Notus and Notux models. The original UltraFeedback dataset was curated using Argilla UI filters to find and report a bug in the original data generation code. Based on this data curation process, Argilla built this new version of the UltraFeedback dataset and fine-tuned Notus, outperforming Zephyr on several benchmarks.
- distilabel Intel Orca DPO dataset used to fine-tune the improved OpenHermes model. This dataset was built by combining human curation in Argilla with AI feedback from distilabel, leading to an improved version of the Intel Orca dataset and outperforming models fine-tuned on the original dataset.

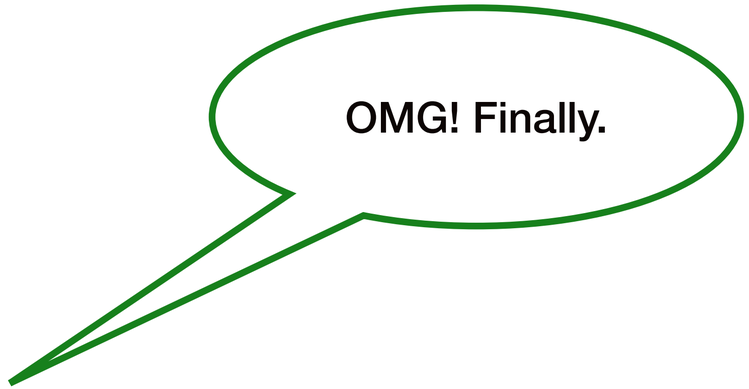
## Examples Use cases

AI teams from organizations such as the Red Cross, Loris.ai and Prolific use Argilla to improve the quality and efficiency of AI projects. They shared their experiences in our AI community meetup.

- AI for good: the Red Cross presentation showcases how the Red Cross domain experts and AI team collaborated by classifying and redirecting requests from refugees of the Ukrainian crisis to streamline the support processes of the Red Cross.
- Customer support: during the Loris meetup they showed how their AI team uses unsupervised and few-shot contrastive learning to help them quickly validate and gain labeled samples for a huge amount of multi-label classifiers.
- Research studies: the showcase from Prolific announced their integration with our platform. They use it to actively distribute data collection projects among their annotating workforce. This allows Prolific to quickly and efficiently collect high-quality data for research studies.

Mix of social proof and example use case

The only thing “missing” is a section on who should not use Argilla




## 👨‍💻 Getting started

### Installation

First things first! You can install the SDK with pip as follows:

```
pip install argilla
```

# Write a Killer README

- Tell me why I should dedicate 30 minutes of my life to “getting started” with your project
- Do everyone a favor and make it easy for people to understand situations where the project is not a good fit
- Make the connection with your company if applicable

Thank you!!
Podcast: The Business of Open Source
Open Source Founders Summit: www.05f5.com
Website: www.emilyomier.com
Email: emily@emilyomier.com