EthicsHub: Integrating Ethical Questions into AI/ML Projects

Thomas Zembowicz

Advisors: Elisa Celis (S&DS) and Nisheeth Vishnoi (CS)

Goal

Create an interactive GitHub plugin that helps users create documentation files detailing the provenance, proper use, and ethical considerations of models, datasets, and code hosted on public depositories.

Background

Data do not speak for themselves. As Gitelman [4] explains, "raw data" are not naturally-occurring but instead created, organized, and used in specific contexts. When models and datasets are published and reused online (especially on code-sharing sites such as GitHub, where files from personal projects and academic papers are often publicly accessible), these contexts are often missing. This is a problem because reusing components originally created for different purposes can create ethical problems and exacerbate social biases, as Bolukbasi et al. [1] show in their study of gender-biased word embeddings used in settings such as hiring.

There have been several recent proposals for standardizing how to provide context for models and datasets. These proposals draw widely on practices from other fields and industries. Mitchell et al. [6] propose model cards that disclose the proper use, benchmarked performance metrics, and other useful information for trained models. Gebru et al. [3] suggest "datasheets for datasets" that call on dataset creators to answer questions about how the data were gathered and what purpose they serve, in a format based on standards in automobile safety, clinical trials, and electronics. Holland et al. [5] propose flexible "nutrition labels" that show metadata as well as more domain-specific information inspired by labels on food packaging. Deon [2] is a command line tool that helps users create an ethics checklist for data science projects. These projects are by no means final or mutually exclusive; rather, they intend to spark discussions about what kinds of questions and disclosures are necessary for using published work in new projects or complex AI systems.
However, few tools integrate these ethical questions by default into the workflows of data scientists, software engineers, and others who handle data. I think this presents a real opportunity to start important conversations about the ethics of reuse in computing projects, and the role that the open source community can take in keeping each other accountable.

Description

My project is to build a GitHub plugin that encourages users to document the provenance and proper use of datasets included in their repositories. This plugin, packaged as a Google Chrome extension, will add a panel to the repository view that allows users to review an ethics checklist and add necessary documentation in the form of an ETHICS.md file using an interactive wizard. I plan on synthesizing existing research to create a modified collection of questions that should be considered when making models and data public. I’ll also draw on my design experience to build a user interface that carefully poses questions without being overwhelming. The tool could also dynamically show and hide options in response to specific features, like the mention of sensitive categories such as race or gender.

My plugin will also be relevant for GitHub users who aren’t repository maintainers. Users will see a button when browsing other projects that allows users to generate a public GitHub issue in the repository that poses specific questions about a model or dataset and gives the maintainer instructions for how to better document their work. By creating two-way communication between GitHub users and repository maintainers, I hope to leverage the power of the open source community to encourage helpful documentation.

I’m going to focus on GitHub because it’s a familiar platform used by the open source community, hosts millions of software and data science projects, and offers a robust set of APIs and webhooks that would let me integrate my plugin with existing workflows. GitHub users are already used to working with ZenHub (agile development), Circle CI (continuous integration) and other plugins offered through GitHub’s Marketplace that alter the appearance and functionality of GitHub—but none are currently devoted to ethics. Time permitting, I will adapt my plugin to work on a standalone website that allows people to create documentation outside of GitHub.

I’m interested in pursuing this project because I believe that ethical questions raised by computing demand greater visibility in the everyday activities of programmers and data scientists. I think that the open source community will play a huge role in creating standards and demanding accountability from each other as the ethical consequences of data misuse become clearer, and I want to explore how thinking about these ethical concerns can become second nature. I’m also excited for the opportunity to practice my design and implementation skills on a project I can see through from start to finish.
Deliverables

1. GitHub plugin. This is the tool described above. I also hope to publish it on GitHub’s Marketplace for greater visibility.

2. Final report. A document synthesizing the literature relevant to “model cards,” “datasheets” and “nutrition labels,” describing how my plugin works and the design choices, and highlighting areas for future work.

Timeline

<table>
<thead>
<tr>
<th>Week</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Submission of project proposal (2/7)</td>
</tr>
<tr>
<td>1-2</td>
<td>Research and synthesis of techniques for data documentation, learning about GitHub APIs/ecosystem</td>
</tr>
<tr>
<td>3-4</td>
<td>Development work, mockups of user interactions and functionality</td>
</tr>
<tr>
<td>5-8</td>
<td>Continued development work</td>
</tr>
<tr>
<td>9-11</td>
<td>Implementing additional functionality and testing, work on final report</td>
</tr>
<tr>
<td>12</td>
<td>Submission of all project deliverables (5/2), publishing plugin</td>
</tr>
</tbody>
</table>

References


