#### This label is a draft

The below is not a final, just a preview

What is this label?



# Description

SimCoPilot is a benchmark for evaluating LLMs to perform as a "copilot"-style, interactive coding assistant, testing their ability to add and complete code in complex real-world software environments and analyzing how LLMs manage different code dependencies and logic complexities.

### Keywords

Benchmark Dataset G

Gen-Al LLM

Program Synthesis

## How to use it?



## ⚠ Restrictions on Use

CC BY-NC-ND 4.0

### ✓ Known Uses

SimCoPilot is a benchmark for evaluating LLMs to perform as a "copilot"-style, interactive coding assistant, testing their ability to add and complete code in complex real-world software environments and analyzing how LLMs manage different code dependencies and logic complexities..

https://github.com/mj33rice/SimCoPilot

## ① Do Not Use

- Domain. N/A
- commercial purposes.
- infer personal information . privacy and ethical violations

## About the dataset

#### People

Owned by

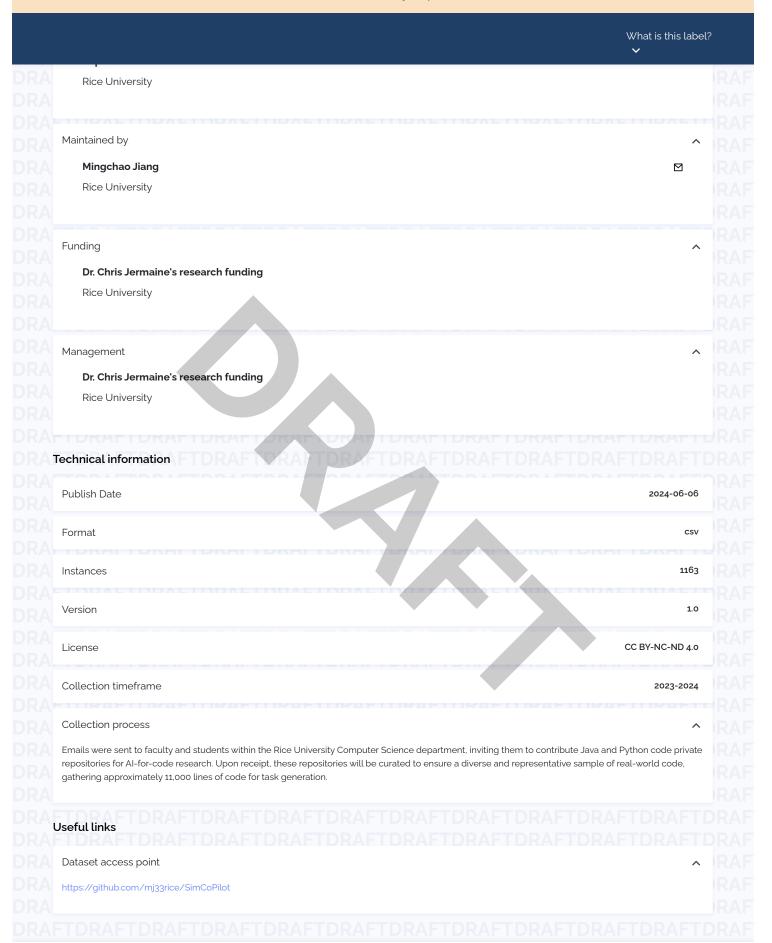
Created by

#### Mingchao Jiang

Rice University

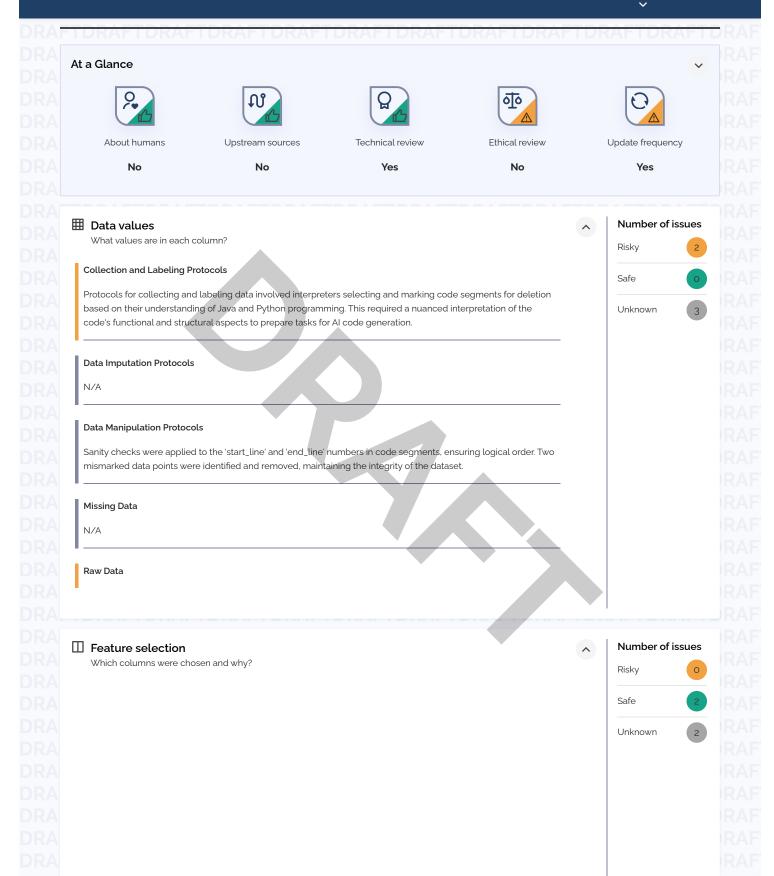
https://datanutrition.org/labels/v3/?id=aa73c968-4f70-47bc-8c65-96f920b44ccc

#### This label is a draft



#### This label is a draft





### This label is a draft

			What is this label? ✔	
DRA	Proxy Characteristics			RAF
	Planning Representation			
	Contributors to the dataset not only provided their source code but also identified ideal code blocks to serve as test points.			
	Domain Knowledge			
	Understanding the field names in the dataset requires familiarity with basic programming concepts in Java and			
	Python, as well as knowledge of Al-for-code methodologies. Additionally, awareness of software development practices and annotation specifics used in generating infill and completion tasks is essential.			
	I and the second			
	% Representation		Number of issues	
	Which rows were included and why?	^		
	Subpopulation Information		Risky	
			Safe 2	
	Not Applicable		Unknown 4	
	Representation			
	The dataset includes Java and Python code contributions primarily from students and faculty at Rice University's			
	Computer Science department, representing a community of academic programmers and developers.			
	Individual Inferences			
	While SIMCOPILOT promotes AI-for-code precision, enhancing development efficiency, developers might over-rely			
	on these tools, diminishing human oversight in coding.			
	Individual Inferences - Mitigation			
	To address the concern of over-reliance on AI-for-code tools like SIMCOPILOT, it's essential to maintain a balance			
	between automation and human oversight. Implementing human-in-the-loop systems ensures critical decisions			
	and code reviews are still performed by developers, preserving essential skills and preventing error propagation.			
	Collection Representation			
	Contributors to the dataset not only provided their source code but also identified ideal code blocks to serve as test points.			
	Other Representation Issues			
	•			
			·	



