Dataset Facts

Dataset Example
Instances Per Dataset 123

Motivation

Original Authors	Alice, Bob
Original Funding	Charlie

Composition

Sample or Complete	Sample
Missing Data	Probably

Collection

Ethical Review	Hopefully
Author Consent	Hopefully

Cleaning and Labeling

Cleaning Done	Hopefully
Labeling Done	Maybe

Uses and Distribution

Notable Uses	This Template
Other Uses	Probably None

Maintenance and Evolution

Corrections or Erratum	Ideal
Methods to Extend	Maybe

Breakdown	% of Example*
Orange 40 items	32.5%
Purple 30 items	24.4%

Green 3.1%	• Blue 2.9%
Red 1.5%	

Not a significant source of real data.

* This dataset is a fabricated example.

Figure 1: A data card styled after nutrition labels. Created for the retrospective BookCorpus datasheet, with help from Stack Exchange. See Gebru et al. for guidance on creating complete datasheets for datasets.