Platform	Description	Link
Google Dataset Search	A search engine across thousands of open data repositories worldwide.	datasetsearch.research.google.com
UCI Machine Learning Repository	Classic machine learning datasets for research and academic use.	archive.ics.uci.edu/ml
Hugging Face Datasets	Focused on NLP/AI; thousands of datasets for transformers, LLMs, and deep learning.	huggingface.co/datasets
Papers with Code – Datasets	Connects datasets to research papers and state-of-the-art benchmarks.	paperswithcode.com/datasets
Data.gov (USA)	U.S. government open data for health, energy, environment, education, and more.	data.gov
OpenML	Datasets with collaborative tools for machine learning experimentation and benchmarking.	openml.org
Awesome Public Datasets (GitHub)	Curated list of high-quality public datasets across domains (image, text, audio, finance, etc.)	github.com/awesomedata/awesome-public-datasets
FiveThirtyEight	Data-driven journalism with accessible CSVs on politics, sports, culture, and more.	fivethirtyeight.com
Zenodo	Research datasets from CERN and EU Horizon 2020 projects, with DOIs for citation.	zenodo.org
Microsoft Research Open Data	Collection of datasets by Microsoft for research and development purposes.	msropendata.com
AWS Open Data Registry	Cloud-hosted public datasets from satellites, genomics, ML, and more.	registry.opendata.aws