1. Comprehensive Road Accident Datasets:

\* Global Road Accidents Dataset: Available on Kaggle, this dataset encompasses over 132,000 records with 30 diverse features capturing accident severity, weather conditions, vehicle involvement, driver characteristics, and geographic regions. It's well-suited for exploratory data analysis, feature engineering, and training machine learning models for predicting accident severity and related outcomes.

\* Road Safety Data (UK): Provided by data.gov.uk, this dataset contains detailed information about personal injury road collisions in Great Britain from 1979. It includes circumstances of the accidents, types of vehicles involved, and resulting casualties. Data is updated regularly, with provisional mid-year data released in November and final annual data in September.

\* Road Accidents in India: Available on the Open Government Data (OGD) Platform India, this dataset provides statistics on the total number of road accidents, fatalities, and injuries in India. It also includes metrics like accidents per lakh population, per ten thousand vehicles, and per ten thousand kilometers of roads. Data is collected annually from police departments across India.

2. Traffic Accident Datasets with Diverse Features:

\* Traffic Accident Prediction Dataset: Hosted on Kaggle, this dataset is designed to predict the occurrence and severity of traffic accidents based on various factors. Features include weather conditions (clear, rainy, foggy, snowy, stormy), road type (highway, city road, mountain road), number of vehicles involved, and road condition (dry, wet, icy, under construction).

\* Road Traffic Accident Dataset of Addis Ababa City (2017-2020): Available on Mendeley Data and Kaggle, this dataset contains records of road traffic accidents in Addis Ababa, Ethiopia, with 32 features and over 12,000 instances. It includes details on the time, day, driver demographics, vehicle type, and accident severity. The dataset has both raw and pre-processed versions.

3. Video-Based Accident Detection Datasets:

\* DoTA (Detection of Traffic Anomaly): This dataset comprises over 4,600 videos collected from dashcams, focusing on traffic anomalies, including accidents.

\* CCD (Car Crash Dataset): This dataset contains real traffic accident videos captured by dashcams, annotated with environmental attributes (day/night, weather), ego-vehicle involvement, accident participants, and reasons. It's valuable for developing safety-guaranteed self-driving systems.

\* Traffic Accident Detection Video Dataset for AI-Driven Computer Vision Systems in Smart City Transportation: Available on IEEE DataPort, this dataset includes approximately 5,700 video files from traffic/surveillance cameras and dashcams, categorized into various accident and normal traffic scenarios. Videos are segmented into 5-second clips.

4. Datasets Combining Traffic and Environmental Data:

\* A Tagged Traffic Accident Dataset for Machine Learning: This dataset on OpenAIRE fuses traffic data from radar detection sensors with weather and light condition data. It includes time series of traffic speeds, flows, and occupancies from sensors near accident locations, along with accident details.

\* A Traffic Accident Dataset for Chattanooga, Tennessee: This dataset, described in a NREL publication, combines traffic data from radar sensors with weather and light conditions for accidents in the Chattanooga, TN area. It provides time series data from sensors upstream and downstream of accident locations.

Key Considerations for Using These Datasets:

\* Data Cleaning and Preprocessing: Most real-world datasets require cleaning, handling missing values, and feature engineering to be suitable for AI model training.

\* Feature Selection: Identifying the most relevant features for accident analysis and prediction is crucial for building effective models.

\* Model Selection: The choice of AI model (e.g., classification, regression, deep learning) depends on the specific prediction task (e.g., accident occurrence, severity level).

\* Ethical Implications: When using accident data, it's essential to consider privacy concerns and ensure responsible data handling.