**Typical Directory Structure**

The following table outlines a typical directory file tree for a MatterGen project. This structure helps organize raw data, configuration files, source code, scripts, and output results.

| Folder/File | Description | Example Contents |
|---|---|---|
| **/mattergen/** | Root directory of the MatterGen repository. | README.md, MODEL_CARD.md, LICENSE |
| **/checkpoints/** | Contains model checkpoints for both the base model and fine-tuned conditional models. | mattergen_base/, dft_band_gap/, chemical_system_energy_above_hull/ |
| **/datasets/** | Raw and preprocessed datasets used for training and fine-tuning. | mp_20/ (raw CSVs), alex_mp_20/, datasets/cache/ |
| **/conf/** | Configuration files for training, fine-tuning, and generation. | trainer/default.yaml, diffusion_model.yaml, property_configs/ |
| **/src/** | Source code for MatterGen, including modules for diffusion, data handling, evaluation, etc. | modules/, models/, utils/, **init**.py |
| **/scripts/** | Standalone scripts to run training, generation, and evaluation workflows. | mattergen-train, mattergen-generate, mattergen-evaluate |
| **/results/** | Output directory for generated structures, evaluation metrics, and logs. | results/2025-03-05/, generated_crystals_cif.zip, metrics.json |
| **/logs/** | Log files generated during training and evaluation. | training.log, evaluation.log |
| **target_properties.json** | Example dictionary file for multi-property conditioning | defines a set of target properties to guide conditional generation |