

Configuring the Job: Determining Which Files To Read

- By default, `FileInputFormat.setInputPaths()` will read all files from a specified directory and send them to Mappers
 - Exceptions: items whose names begin with a period (`.`) or underscore (`_`)
 - Globbs can be specified to restrict input
 - For example, `/2010/*/01/*`
- Alternatively, `FileInputFormat.addInputPath()` can be called multiple times, specifying a single file or directory each time
- More advanced filtering can be performed by implementing a `PathFilter`
 - Interface with a method named `accept`
 - Takes a path to a file, returns `true` or `false` depending on whether or not the file should be processed