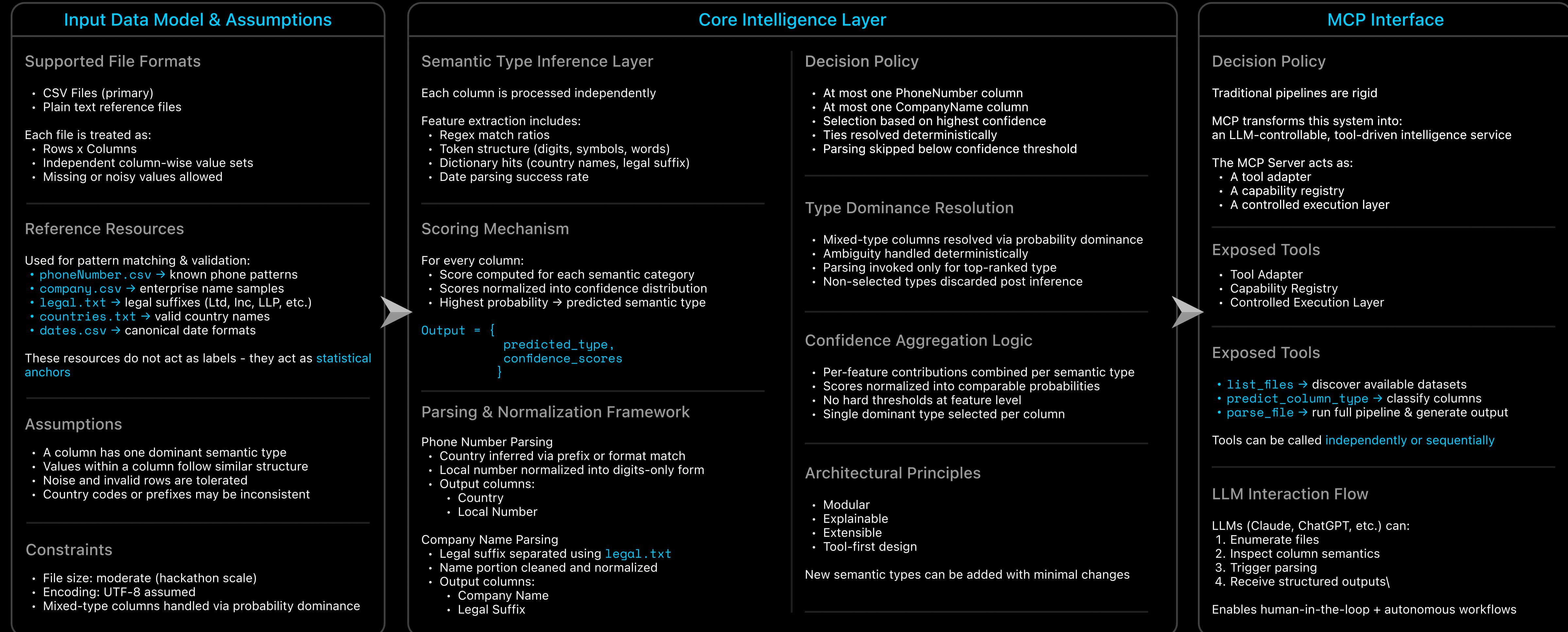


# System Architecture

## Problem Definition & Scope

Automatically infer the semantic meaning of dataset columns using values alone, without relying on headers, and convert unstructured fields into structured, machine-readable formats. The system focuses on identifying Phone Numbers, Company Names, Countries, Dates, and Other, and exposes this capability via an MCP server for LLM-driven workflows.



## Output Schema Specification

- A single output.csv
- Original columns preserved
- Parsed columns appended
- Missing values explicitly represented

## System Invariants

- Input files are never mutated
- Results are deterministic
- Same input → same output
- Operations are idempotent

# Workflow & Execution

## High Level View

....

