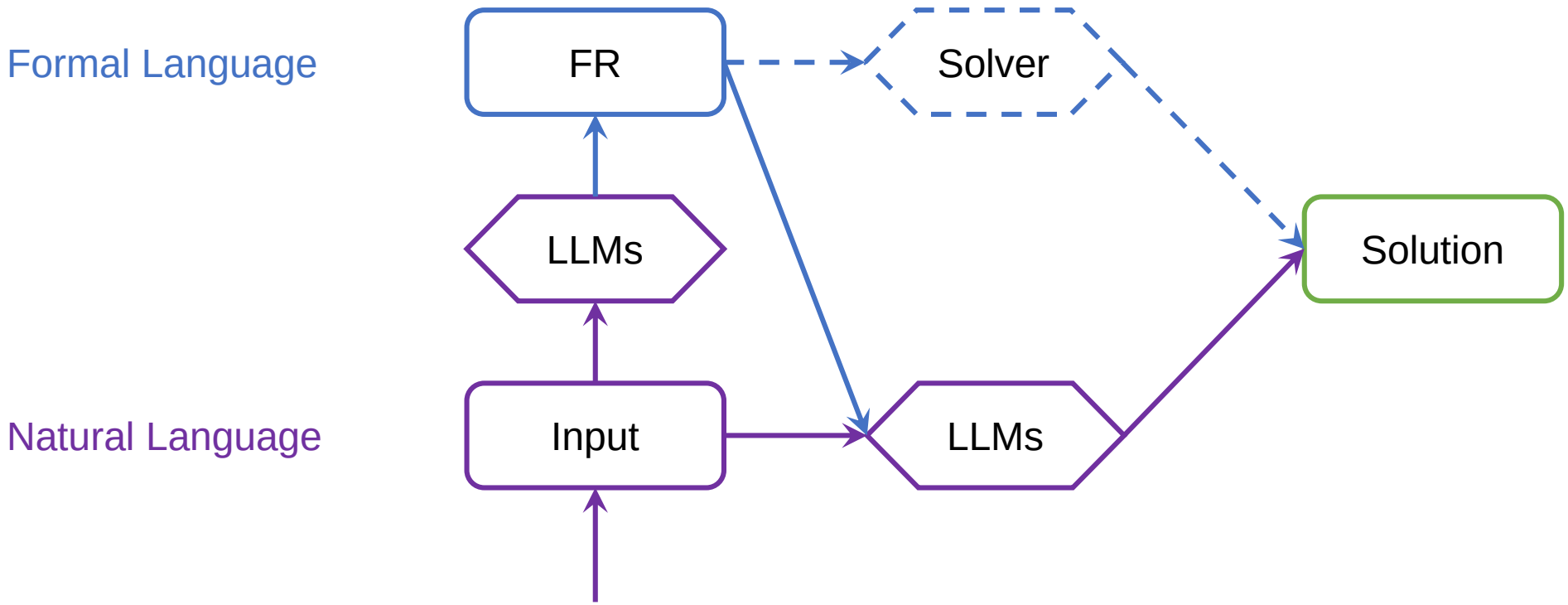
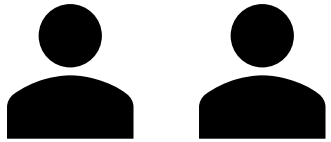

Formal Language
FR
Solver
LLMs
Solution
Natural Language
Input
LLMs

Initiation We build a small dataset, finalize the annotation guidelines and design an AI-assistant for annotation.

Verification We select the annotators from a group of candidates by measuring their performance on the small dataset.



Annotation Each question is annotated by two annotators, then validated by another validator. We randomly check 3% of the data.

Finalization We check the complete dataset with a stronger AI-assistant through cross-validation. Then we finalize the dataset splits.