CNL: A DSL for human-in-the-loop data transformation GUIs

Background

- Domain experts in the social sciences often find themselves performing repetitive tasks restructuring large amounts of data.
- They often don’t have the technical expertise to automate these tasks.
- Writing bespoke tools for these tasks from scratch is hard.

System

- End User
- Data + Interactions
- CNL Tool
- CNL Engine
- CNL Program
- Programmer

Tool

Stage 1: Upload Dataset

Stage 2: Issue Matcher Query

Programs in CNL define **Tools**, each containing multiple **Stages**.

Ilustrative Example using CNL:

```python
def dataset_upload_stage():
    file_upload_input = FileUploadComponent('csv', "Upload new data sets here:"
    dataset_name_input = TextInputComponent(str, "Enter a name for the dataset you’re uploading:"

def save_dataset():
    saved_table = db_utils.create_table_from_csv(dataset_name_input, file_upload_input)
    return saved_table

processor = LambdaProcessor(save_dataset)
my_result = processor.result("Here is the saved dataset: ")
result.show_results([my_result])
```

Stages use **Components** to describe GUIs, use **Processors** to describe data transformations, and can show **Results**.