DreamStruct: Understanding Slides and User Interfaces via Synthetic Data Generation

Yi-Hao Peng[1], Faria Huq[1], Yue Jiang[3], Jason Wu[1], Xin Yue Li[1], Jeffrey P. Bigham[1], and Amy Pavel[2]

[1] Carnegie Mellon University, Pittsburgh PA 15213, USA
[2] University of Texas Austin, Austin, TX, 78712, USA
[3] Aalto University, Otakaari 24, 02150 Espoo, Finland
{yihaop,fhuq,jasonwu2,xal,jbigham}@cs.cmu.edu, yue.jiang@aalto.fi, apavel@cs.utexas.edu

1 Synthetic samples for slides and UIs

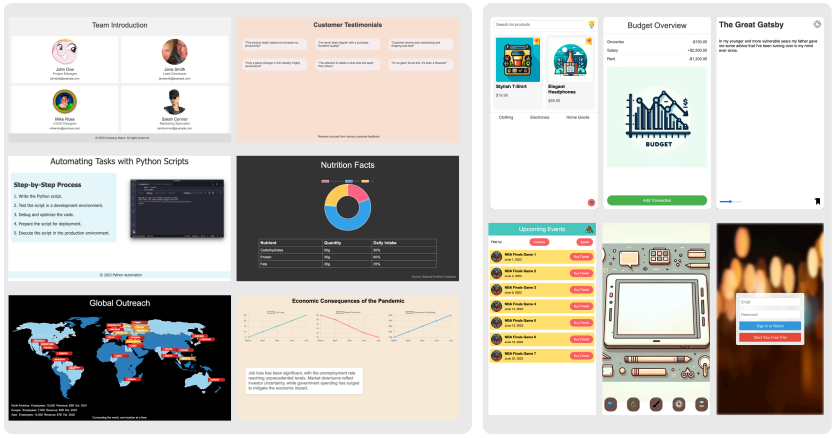

Fig. 1: Examples of generated slides *(left)* and UI *(right)* in DreamStruct.

2 Heuristic-based post-processing for generated semantics

– *Element annotation adjustments*
  - If the `HTML` contains both a background image and a background fill color, remove the fill color so that it is visible.
  - If an item is of `<input type= 'checkbox'>` class, assign 'data-type' = 'checkedview' to its parent.

  - If any child text element under a button is labelled with 'data-type' = 'text', it should be removed for visual and label consistency with human-annotated samples.
- *Element annotation augmentations*
  - Add width and height to `<img>` elements so that it fits only the screen properly. For example, `<img src="https://via.placeholder.com/280x400" ...>` should be updated to `<img .. style="height: 400px; width: 280px">`
  - If an element is a font awesome icon with `<i>` tag, label it with 'data-type' = 'icon'.
  - If the screen contains a sliding menu, call a `js` function to keep it open when the page is being loaded. This makes sure that the sliding menu elements are visible on the screen.

3 Example Model Input and Output

Please refer to `https://github.com/yihaop/dreamstruct` for more detailed formatted prompts as well as example model outputs (*e.g.,* image captions) and corresponding model performance.