

Form2Seq : A Framework for Higher-Order Form Structure Extraction (Supplementary)

Milan Aggarwal¹, Hiresht Gupta², Mausoom Sarkar¹, Balaji Krishnamurthy¹
Media and Data Science Research Labs, Adobe¹
Adobe Experience Cloud²

1 Form Structure Extraction Application

Document structure extraction has been performed for digitising documents to make them re-flowable which is useful in web based services. Organisations across various domains, such as finance, administration, healthcare etc., which have been using paper forms or flat PDF forms would want to digitize them by converting them into an appropriate digitised version (such as an HTML). Once these forms are made re-flowable, they can be used on many devices with different form factors so that whole form layout can be rendered dynamically as shown in figure 1. This availability across devices automatically increases the ease of doing business or provide services since people can interact with them easily. Form digitisation also enables other capabilities such as better handling of data filled in digitised version, applying validation checks on data filled in fields, consistent form design control, auto-filling similar fragments in a form etc.



Figure 1: Form digitisation to re-flow a form (left) on multiple devices (right).

2 Forms Dataset

The Forms Dataset comprises of 23k form documents from various governmental and private organizations. We studied different form elements and structures in order to identify the constructs. Post identification, we collected a diverse set of example cases for each structure to design a tagging schema to tag the complete hierarchical structure of Form Documents comprising of Fields, Choice-Groups and Lists. We created a Forms Tagging

Tool (as shown in Figure 2) capable of grouping Text elements, associating Text Captions with Widgets (marking Fields) and grouping elements into ChoiceGroups and Lists to generate rich hierarchical annotations for our Forms Dataset. We employed multiple annotators to tag the form images using our Forms Tagging Tool, and instructed them using the tagging scheme designed through various demonstrations. There were multiple rounds of tagging comprising where we reviewed their annotations and suggested possible improvements iteratively.

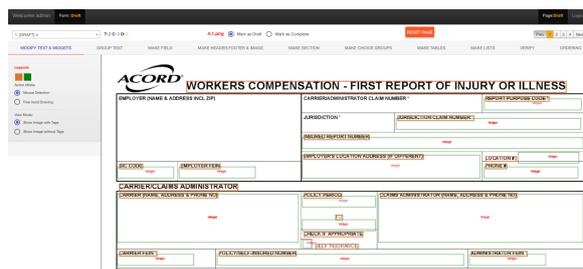


Figure 2: Figure illustrating the designed Tagging Tool created for tagging Forms Dataset.

3 Additional Visualisations

Figure 3 and 4 below illustrate Form2Seq model prediction comparisons between with and (without text) variants. As can be seen, having text helps in better grouping of choice group elements.

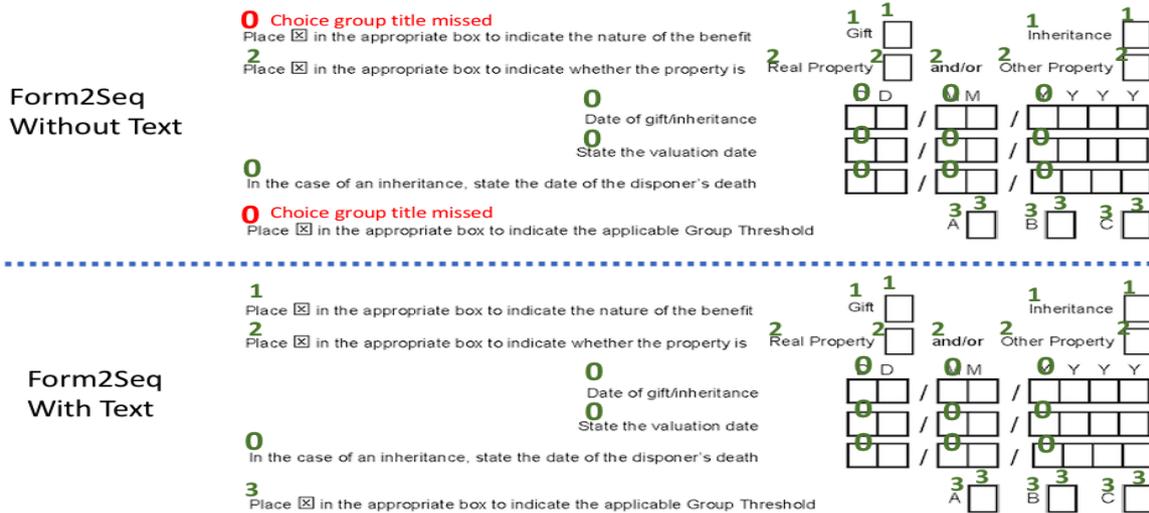


Figure 3: Visualization comparing Forms2Seq model predictions for Choice Group Identification task. It can be seen that the Form2Seq w/o Text model misses to group two Choice group titles with their corresponding choice fields (highlighted in red).

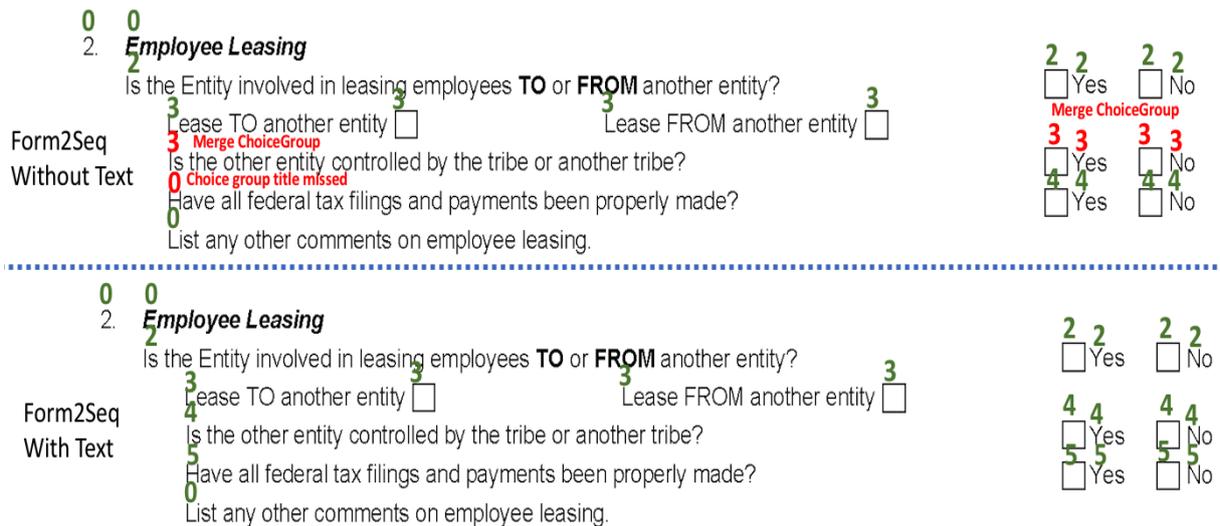


Figure 4: Visualization showing Forms2Seq models predictions for Choice Group Identification task. It can be seen that the Form2Seq without Text model merges two choicegroups (highlighted in red) because of missing textual context.