

Supplementary Material

1. CORD dataset prompts

Function	Prompt Template
Role/Task	Claude, you are a seasoned invoice annotator. Given a document, your task is to extract the following information in JSON format:
Format	{ "line_item":
Format	[{ "discount_price": "",
Format	"identifier": "",
Format	"name": "",
Format	"other": "", "quantity": "",
Format	"sub": [{"sub_name": "",
Format	"sub_price": "",
Format	"sub_quantity": ""}],
Format	"subtotal_price": "",
Format	"total_price": "", "unit_price": "" }],
Format	"subtotal": { "discount_price": "",
Format	"other": [],
Format	"service_price": "", "subtotal_price": [],
Format	"tax_price": [] },
Format	"total": { "cash_price": [],
Format	"change_price": "",
Format	"credit_card_price": "",
Format	"emoney_price": "",
Format	"line_item_quantity_count": "",
Format	"line_item_type_count": "",
Format	"other": "",
Format	"total_price": [] } }
Rules	Please be mindful of the following rules:
Rules	-The extraction must respect the JSON schema.
Rules	-Only extract entities specified in the schema. Do not skip any entity types.
Rules	-The values must only include text found in the document.
Rules	-Use null or [] for missing entity types.
Rules	-Do not indent the json you produce.
Rules	-Examples of valid string value format: "\$ 1234.50", "John Do", null.
Rules	-Examples of valid list value format: ["\$ 1234.50", "John Do"], [].
Rules	-Report the entire text present for each field. For example, if the price mentioned is usd 123.50, mention usd 123.500 and not 123,500. If unit price mentioned is @12, mention @12 and not 12. Respect the format of the text.
Rules	- For price, double check whether the number format uses comma (eg 124,000) or dot (eg 124.000).
Rules	If 124.000 is mentioned, use that as price.
Rules	- Place special attention on the newline character \n. Strings not separated by it are likely part of the same entity.
Rules	- Return the json, double check the json format, do not write anything extra
Rules	Return output for this image
image data	<i>image</i>

Table 1. TAIL prompt for CORD

For TAIL labels, we modified the attributes keys to meaningful names for better zero-shot extraction ([1]). For example,

”menu” was changed to ”line.item”, ”nm” was changed to ”name”, and so on. The full prompt is given in Table 1:

References

- [1] Vincent Perot, Kai Kang, Florian Luisier, Guolong Su, Xiaoyu Sun, Ramya Sree Boppana, Zilong Wang, Zifeng Wang, Jiaqi Mu, Hao Zhang, Chen-Yu Lee, and Nan Hua. Lmdx: Language model-based document information extraction and localization, 2024. 1