Purchase Order labeling Instructions
Labelling Guidance

Getting Started
● Start by annotating a small number of documents
● Review them, correct any mistakes, update the labeling guidance to make sure data is labeled consistently.

Choosing the right data/model
● A general rule is to up-train with customer production traffic. The closer the training distribution is to production traffic, the better the performance.
● Depending on your customer’s requirements, consider utilizing more than one processor in order to support multiple languages
● One model for every region (group of untrained languages) vs a universal model

Dataset Size
● Training - at least 1000 per language
● Test - at least 200 per language
Multi-Occurrence entities

There are 2 occurrence types that we are relevant to purchase order labeling:

- **OPTIONAL_ONCE**: This entity is expected to have no more than one ground truth value.
- **OPTIONAL_MULTIPLE**: This entity can have any number of ground truth values.

For Purchase Order, `line_item` entities are **OPTIONAL_MULTIPLE**, while all other entities are **OPTIONAL_ONCE**.

Note that while **OPTIONAL_ONCE** entities can still appear multiple times with the same value. In this case, we recommend to label all appearances, except for “easy” entities (in the current entity list, only `currency`).
Optical Character Recognition (OCR) Errors

Drawing bounding boxes around entities will automatically populate a text captured through OCR. In some cases, the captured text is not accurate and doesn’t reflect what is on the page, especially for documents that are blurry or of low resolution. These values should be corrected manually through “Value” input text.

If a value doesn’t exist in the document, there is no need to label it.
Purchase Order Entity Definitions
<table>
<thead>
<tr>
<th>Entity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>currency</td>
<td></td>
</tr>
<tr>
<td>delivery_date</td>
<td></td>
</tr>
<tr>
<td>payment_terms</td>
<td></td>
</tr>
<tr>
<td>purchase_order_date</td>
<td></td>
</tr>
<tr>
<td>purchase_order_id</td>
<td></td>
</tr>
<tr>
<td>receiver_name</td>
<td></td>
</tr>
<tr>
<td>ship_to_address</td>
<td></td>
</tr>
<tr>
<td>ship_to_name</td>
<td></td>
</tr>
<tr>
<td>total_amount</td>
<td></td>
</tr>
<tr>
<td>line_item/amount</td>
<td></td>
</tr>
<tr>
<td>line_item/description</td>
<td></td>
</tr>
<tr>
<td>line_item/product_code</td>
<td></td>
</tr>
<tr>
<td>line_item/purchase_order_id</td>
<td></td>
</tr>
<tr>
<td>line_item/quantity</td>
<td></td>
</tr>
<tr>
<td>line_item/receiver_reference</td>
<td></td>
</tr>
<tr>
<td>line_item/unit_of_measure</td>
<td></td>
</tr>
<tr>
<td>line_item/unit_price</td>
<td></td>
</tr>
</tbody>
</table>
Receiver entities

**receiver_name**: Name of the organization/person/entity to which the purchase order is made or billed. Examples: “Zinfra”.

**ship_to_address**: These labels correspond to the entity/organization address to whom the items are being shipped to. The key for this field might be the same as the ship_to_address field. Example: ship to, deliver to.

**ship_to_name**: Actual name of the entity/person, to whom the goods are being shipped to. Include full organization name including prefixes such as “LLC”, “Pvt Ltd”. Example: Google LLC
Line item entities

**line_item/amount:** The actual amount billed for a particular item in the items table. Usually, one value per line item in the items table. Examples: “20”, “65.4”.

**line_item/description:** It is actual description of the item. There are usually many values in a single image. Examples: “Laptop”, “Paper bundle”, “HP Headphones”

**line_item/product_code:** It refers to items describing the product code/ID. There are usually many values in a single image, one corresponding to each item. Examples: “BN-150”, “75”, “B-108”.

**line_item/purchase_order_id:** Actual purchase order identifier/number. Include any prefix/suffix associated with the purchase order id. Example- “PO128”.

**line_item/quantity:** It is the actual number of units ordered. There are usually many values in a single image, one corresponding to each item. Examples: “10”, “20” in "20 lbs", “30” in "30 bags”.

**line_item/receiver_reference:** It refers to customer code assigned by the receiver to the product or service that is being bought. For example, a supplier may have a product code XYZ, but the receiver has defined the same product in their system with code ABC.

**line_item/unit_of_measure:** It is the measurement unit of the item being ordered. There are usually many values in a single image, one corresponding to each item. Examples: “lb”, “pounds”, “kg”, “hours”

**line_item/unit_price:** It is the actual price per unit of the item being ordered. There are usually many values in a single image, one corresponding to each item. Examples: “70”, “50 $/hr”, “20 $/pound”.

Google
Amount entities

**currency**: It refers to text specifically declaring the currency used in the Purchase order. Examples: “USD”, “GBP”, “$”, “€”, “US $”

**total_amount**: The actual text indicating the monetary amount (total amount after taxes are added). Examples: “343.44”, “8,372.81”.
Miscellaneous entities

**delivery_date**: This field refers to the delivery date for the products and services being ordered. Example: “Delivery date”

**payment_terms**: The actual text of the payment terms. Include full text including # of days, due by date, etc (if present). Examples- “NET 30 days”, “NET 45”.

**purchase_order_date**: This field refers to the purchase order date. In some cases, only month & year may be mentioned, instead of the full date. Example “Dec 2019”. Please consider such values as valid and annotate them. Don’t include words like “Date” or “P/O Date”.

**purchase_order_id**: Purchase order is also an identifier of the transaction. Include any prefix/suffix associated with the purchase order id. Example- “PO128”.

# PO Acknowledgment

**Sales Order:** 1281982888  
**PO#:** 234532  
**Freight Order:** 2838723  
**Order Date:** 12/02/21  
**Pricing Date:** 12/02/21  
**Ship Date:** 12/5/21  
**Receiver Name:** Abcxyz Traders  
**Currency:** USD  
**Primary Broker:** 1299188882 - NO BROKER

<table>
<thead>
<tr>
<th>Item</th>
<th>Material</th>
<th>Material Number</th>
<th>GTIN</th>
<th>Product Description</th>
<th>Order Qty</th>
<th>Order Net Weight</th>
<th>Pricing Units</th>
<th>Price/Unit</th>
<th>Cust Exp Price</th>
<th>Price/Allowances/Charges</th>
<th>Est. Extended Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0000003323</td>
<td>021212124</td>
<td>0023732736746</td>
<td>4/4 FT BEEF CHKN KRM MAIN KOSL</td>
<td>160</td>
<td>1,600.00 LB</td>
<td>160</td>
<td>$44,400.00 CS</td>
<td>Lot Price DEL-CB</td>
<td>$7,428.80</td>
<td>$7,428.80</td>
</tr>
<tr>
<td>20</td>
<td>0000001235</td>
<td>0003728327373</td>
<td>0023732736746</td>
<td>7/4 PC BEEF MARINADED</td>
<td>12</td>
<td>216.00 LB</td>
<td>12</td>
<td>$115,150.00 CS</td>
<td>Published Price DL-CB</td>
<td>$1,381.80</td>
<td>$1,381.80</td>
</tr>
<tr>
<td>30</td>
<td>0000001234</td>
<td>0002732732737</td>
<td>0003728327373</td>
<td>7/4 PC BEEF MARINADED</td>
<td>45</td>
<td>506.25 LB</td>
<td>45</td>
<td>$48,710.00 CS</td>
<td>Published Price DL-CB</td>
<td>$2,191.95</td>
<td>$2,191.95</td>
</tr>
<tr>
<td>40</td>
<td>0000001236</td>
<td>0000287272723</td>
<td>0003728327373</td>
<td>7/4 PC FRGRN BEEF MARINADED</td>
<td>100</td>
<td>1,000.00 LB</td>
<td>100</td>
<td>$69,100.00 CS</td>
<td>Price Of Net - DEL</td>
<td>$6,910.00</td>
<td>$6,910.00</td>
</tr>
<tr>
<td>50</td>
<td>0000004839</td>
<td>0002732732737</td>
<td>000328832923</td>
<td>7/4 PC BEEF &amp; LAMB</td>
<td>30</td>
<td>300.00 LB</td>
<td>30</td>
<td>$92,210.00 CS</td>
<td>Published Price DL-CB</td>
<td>$2,768.70</td>
<td>$2,768.70</td>
</tr>
<tr>
<td>60</td>
<td>0000002365</td>
<td>0009289552654</td>
<td>000728121278</td>
<td>7/4 PC BEEF</td>
<td>100</td>
<td>1,200.00 LB</td>
<td>100</td>
<td>$89,900.00 CS</td>
<td>Published Price DL-CB</td>
<td>$8,908.60</td>
<td>$8,908.60</td>
</tr>
<tr>
<td>70</td>
<td>0000002367</td>
<td>0002383277373</td>
<td>001212871277</td>
<td>7/4 PC BEEF</td>
<td>20</td>
<td>200.00 LB</td>
<td>20</td>
<td>$45,200.00 CS</td>
<td>Price Of Net - DEL</td>
<td>$958.00</td>
<td>$958.00</td>
</tr>
<tr>
<td>80</td>
<td>0000002962</td>
<td>0006732673234</td>
<td>001326871127</td>
<td>4/4 FT BEEF CHKN KRM MAIN BEEF</td>
<td>20</td>
<td>200.00 LB</td>
<td>20</td>
<td>$70,500.00 CS</td>
<td>Published Price DL-CB</td>
<td>$1,415.00</td>
<td>$1,415.00</td>
</tr>
<tr>
<td>90</td>
<td>0000004528</td>
<td>000325435344</td>
<td>012926129282</td>
<td>4/4 FT BEEF CHKN KRM MAIN BEEF</td>
<td>10</td>
<td>100.00 LB</td>
<td>10</td>
<td>$45,500.00 CS</td>
<td>Lot Price DEL-CB</td>
<td>$455.00</td>
<td>$455.00</td>
</tr>
</tbody>
</table>

**Example 1:**

**PO Type:** Manual  
**Date/Time:** 2021-10-21 13:10:49  
**Sales Order:** 1281982888  
**Sold To:** 12717272772  
**Ship To:** 87287237378  
**ZXCV INDUSTRIES**  
**45 LIGHTNING ROAD**  
**ARIZONA AZ 88776-8328**  
**Country:** US  
**Currency:** USD  

*PRICES MAY BE SUBJECT TO APPLICABLE DISCOUNTS, ALLOWANCES OR RELATES WHICH ARE NOT SHOWN, EXCEPT AS MAY OTHERWISE BE GOVERNED BY A DEFINITIVE AGREEMENT SIGNED BY BOTH PARTIES. ALL SALES ARE SUBJECT TO THE TERMS AND CONDITIONS OF SALE AT WWW.THISWEBSITE.COM/Legal/Terms-And-Conditions-Of-Sale.aspx. ANY ADDITIONAL OR DIFFERENT TERMS PROPOSED BY CUSTOMER ARE REJECTED UNLESS EXPRESSLY ASSENTED TO IN WRITING BY SELLER.*
### Example 1:

**ABCxyz Traders**

**PO Acknowledgment**

**Date/Time:** 2021-10-21 13:10:49  
**Sales Order:** 1291982888

<table>
<thead>
<tr>
<th>Item</th>
<th>Material</th>
<th>GTIN</th>
<th>Product Description</th>
<th>Order Qty</th>
<th>Est Order Net Weight</th>
<th>Pricing Units</th>
<th>Price/Unit</th>
<th>Cust Exp Price</th>
<th>Price/Allowances/Charges</th>
<th>Est Extended Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>1982983389</td>
<td>121982988812</td>
<td>305/3.33 GLD RB PORK STEAK PFK</td>
<td>24 CS</td>
<td>240.00 LB</td>
<td>24</td>
<td>$42.4500</td>
<td>$1,018.80</td>
<td>$1,018.80</td>
<td>ESTIMATED TAX: $0.00</td>
</tr>
</tbody>
</table>

**EST TOTAL CUBE:** 223.339

**LINE ITEM/PRODUCT CODE:** 1982983389  
**LINE ITEM/DESCRIPTION:** 305/3.33 GLD RB PORK STEAK PFK  
**LINE ITEM/QUANTITY:** 24 CS  
**LINE ITEM/UNIT PRICE:** $42.4500  
**LINE ITEM/UNIT OF MEASURE:** 240.00 LB  
**LINE ITEM/AMOUNT:** $1,018.80  
**LINE ITEM/RECEIVER REFERENCE:** PORK STEAK PFK  
**LINE_ITEM/AMOUNT:** $1,018.80  
**TOTAL AMOUNT:** $1,018.80  
**CURRENCY:**  

 Prices may be subject to applicable discounts, allowances or rebates which are not shown, except as may otherwise be governed by a definitive agreement signed by both parties, all sales are subject to the terms and conditions of sale at www.zacosfoods.com/Legal/Terms-and-Conditions-of-Sale.aspx. Any additional or different terms proposed by customer are rejected unless expressly assented to in writing by seller.
Feedback

1. Should Sales Order be invoice_id?
2. Receiver_name should be the “ZXCV INDUSTRIES” under “Sold-To”, not the title.
3. GTIN should be line_item/product_code, not Material.
4. All the comments on page 1 also applies to page 2.