

# PDF parsing engine using ChatGPT Structured Output

Khunakorn Luyaphan (Gates)

2025-07

Delivering Growth – in Asia and Beyond.

# High Level View



# High Level View



Extract PDF into  
specific excel format

<b>Product Name</b>	<b>Length / Width / Height (m)</b>
<b>Manufacturing site address and country</b> <i>Company who produces, manufactures the product</i>	<b>Volume of a packaging unit (m3)</b>
<b>Manufacturer Article Number</b> <i>(It is the manufacturer internal reference number of the product)</i>	<b>Pallet Type &amp; Material</b>
<b>Physical Location of the Goods (country)</b>	<b>Storage conditions (temperature &amp; humidity)</b>
<b>Substance Name</b>	<b>Transport conditions (temperature)</b>
<b>Role of Substance</b> <i>(Select among additive, carrier, emulsifier, ingredient, impurity including residual solvents), preservative, processing aids, solvent, stabilizer, other (specify).)</i>	<b>Shelf-life</b>
<b>EC Number</b>	<b>Lot/Batch number structure</b> <i>Example: 200101 (2 last digits of year+month+day)</i>
<b>%</b> <i>(Percentage of each ingredient in the formula, not the assay. For pure substances, please indicate 100%)</i>	<b>Tariff code</b> <i>For EU supplier : provide TARIC code (10 digits) – For non EU supplier : provide your local tariff code</i>
<b>Does the product contain substance(s) from animal origin?</b>	<b>Origin (country of)</b> <i>Country of origin = the country in which the goods are manufactured, produced or grown.</i>
<b>Does the product contain substance(s) from vegetal origin?</b>	<b>Customs status</b> <i>"Union" if the product originates in the EU or was released into free circulation in the EU; "Non-Union" otherwise</i>
<b>Does the product contain substance(s) use of palm origin derivatives?</b>	<b>Preferential origin as per free trade agreement with EU</b>
<b>Does the product contain substance(s) from mineral origin?</b>	<b>Preferential origin as per free trade agreement with UK</b>
<b>Does the product contain substance(s) from use of conflict minerals?</b> <i>These conflict minerals are tin, tantalum, tungsten, gold and their derivatives</i>	<b>Preferential origin as per free trade agreement with Switzerland</b>
<b>Does the product contain substance(s) from synthetic origin?</b>	<b>EU preferential origin</b> <i>For EU supplier</i>
<b>If other, please specify:</b>	<b>Estimated Cost in Local Currency</b>
<b>Outer Packaging unit (bag, box, drum, IBC...) (7)</b> <i>In case of more than 2 packagings, please provide a separate document with all the available packagings.</i>	<b>Quantity in 1st PO to Supplier</b>
<b>Outer Packaging material (PE, aluminium, ...)</b>	<b>Estimated Current Year Quantity</b>
<b>UN-homologated Outer Packaging</b> <i>In case of Dangerous Goods, please provide a picture of the UN homologation mark of the packaging</i>	<b>Custom Clearance is done by whom (Us, supplier)</b>
<b>Inner Packaging unit</b>	<b>Country Sold To</b>
<b>Inner Packaging material</b>	<b>Physical location of the goods at the time of PO</b>
<b>Gross weight (kg)</b>	<b>Country where the goods will be custom clear</b>
<b>Net weight (kg)</b>	

# High Level View



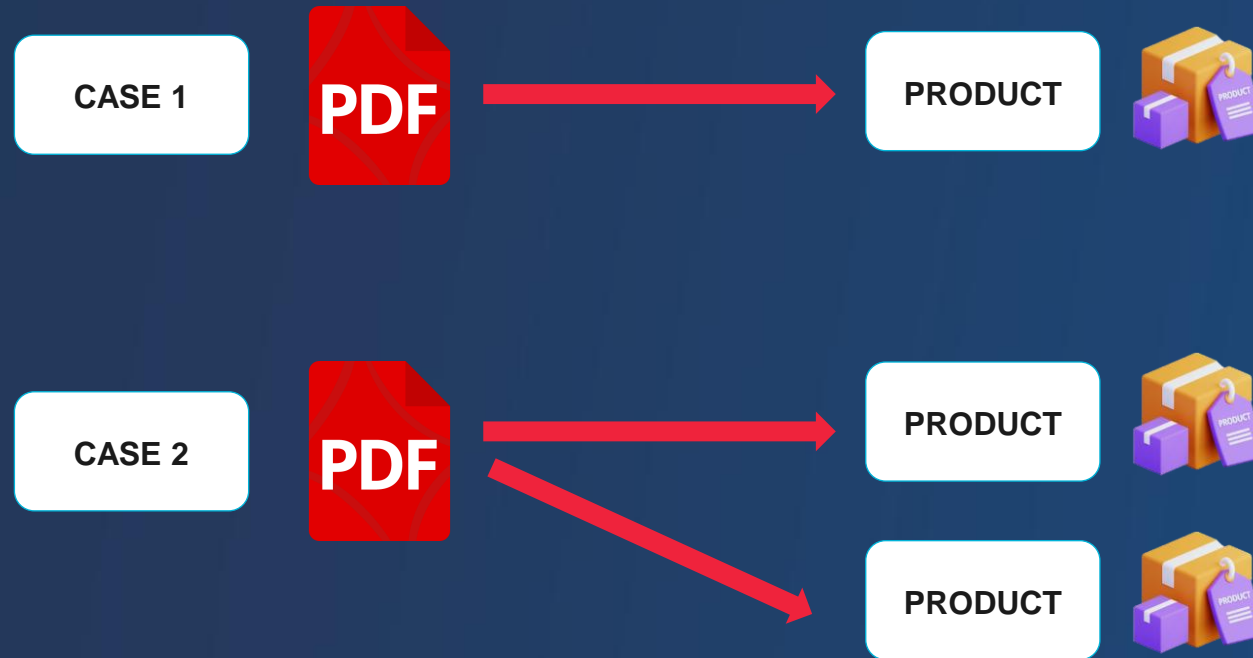
43 Fields



Extract PDF into  
specific excel format

<b>Product Name</b>	<b>Length / Width / Height (m)</b>
<b>Manufacturing site address and country</b> <i>Company who produces, manufactures the product</i>	<b>Volume of a packaging unit (m3)</b>
<b>Manufacturer Article Number</b> <i>(It is the manufacturer internal reference number of the product)</i>	<b>Pallet Type &amp; Material</b>
<b>Physical Location of the Goods (country)</b>	<b>Storage conditions (temperature &amp; humidity)</b>
<b>Substance Name</b>	<b>Transport conditions (temperature)</b>
<b>Role of Substance</b> <i>(Select among additive, carrier, emulsifier, ingredient, impurity including residual solvents), preservative, processing aids, solvent, stabilizer, other (specify).)</i>	<b>Shelf-life</b>
<b>EC Number</b>	<b>Lot/Batch number structure</b> <i>Example: 200101 (2 last digits of year+month+day)</i>
<b>%</b> <i>(Percentage of each ingredient in the formula, not the assay. For pure substances, please indicate 100%)</i>	<b>Tariff code</b> <i>For EU supplier : provide TARIC code (10 digits) – For non EU supplier : provide your local tariff code</i>
<b>Does the product contain substance(s) from animal origin?</b>	<b>Origin (country of)</b> <i>Country of origin = the country in which the goods are manufactured, produced or grown.</i>
<b>Does the product contain substance(s) from vegetal origin?</b>	<b>Customs status</b> <i>"Union" if the product originates in the EU or was released into free circulation in the EU; "Non-Union" otherwise</i>
<b>Does the product contain substance(s) use of palm origin derivatives?</b>	<b>Preferential origin as per free trade agreement with EU</b>
<b>Does the product contain substance(s) from mineral origin?</b>	<b>Preferential origin as per free trade agreement with UK</b>
<b>Does the product contain substance(s) from use of conflict minerals?</b> <i>These conflict minerals are tin, tantalum, tungsten, gold and their derivatives</i>	<b>Preferential origin as per free trade agreement with Switzerland</b>
<b>Does the product contain substance(s) from synthetic origin?</b>	<b>EU preferential origin</b> <i>For EU supplier</i>
<b>If other, please specify:</b>	<b>Estimated Cost in Local Currency</b>
<b>Outer Packaging unit (bag, box, drum, IBC...) (7)</b> <i>In case of more than 2 packagings, please provide a separate document with all the available packagings.</i>	<b>Quantity in 1st PO to Supplier</b>
<b>Outer Packaging material (PE, aluminium, ...)</b>	<b>Estimated Current Year Quantity</b>
<b>UN-homologated Outer Packaging</b> <i>In case of Dangerous Goods, please provide a picture of the UN homologation mark of the packaging</i>	<b>Custom Clearance is done by whom (Us, supplier)</b>
<b>Inner Packaging unit</b>	<b>Country Sold To</b>
<b>Inner Packaging material</b>	<b>Physical location of the goods at the time of PO</b>
<b>Gross weight (kg)</b>	<b>Country where the goods will be custom clear</b>
<b>Net weight (kg)</b>	

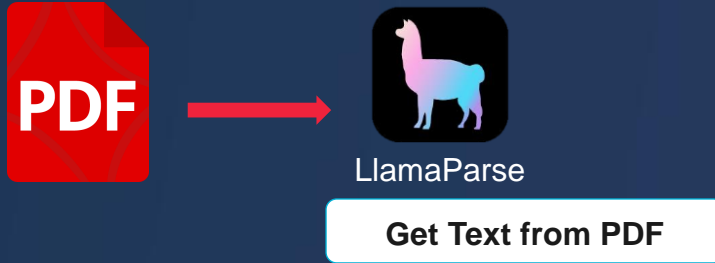
# High Level View



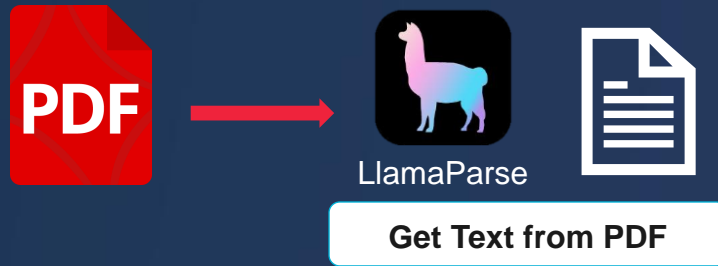
# Plan Of Attack



# Plan Of Attack

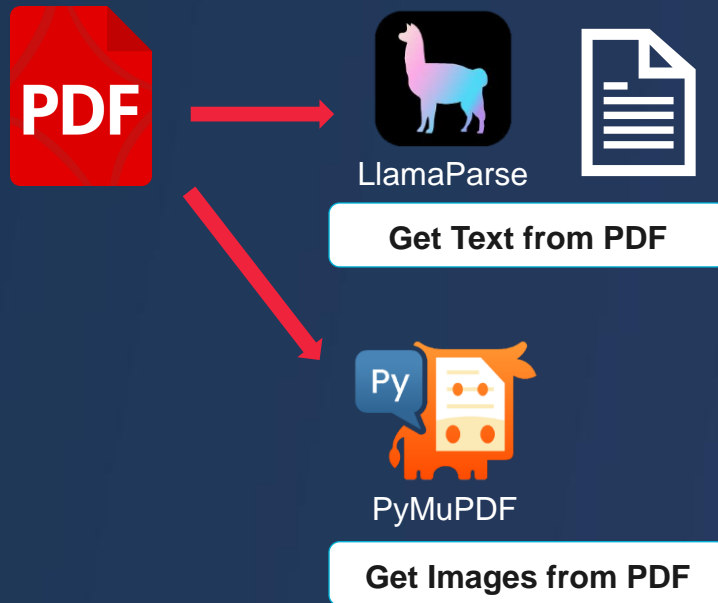


# Plan Of Attack

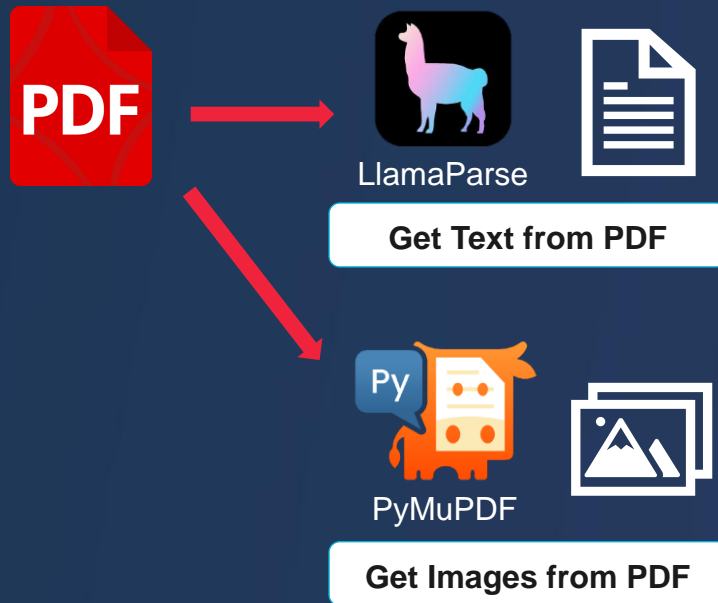




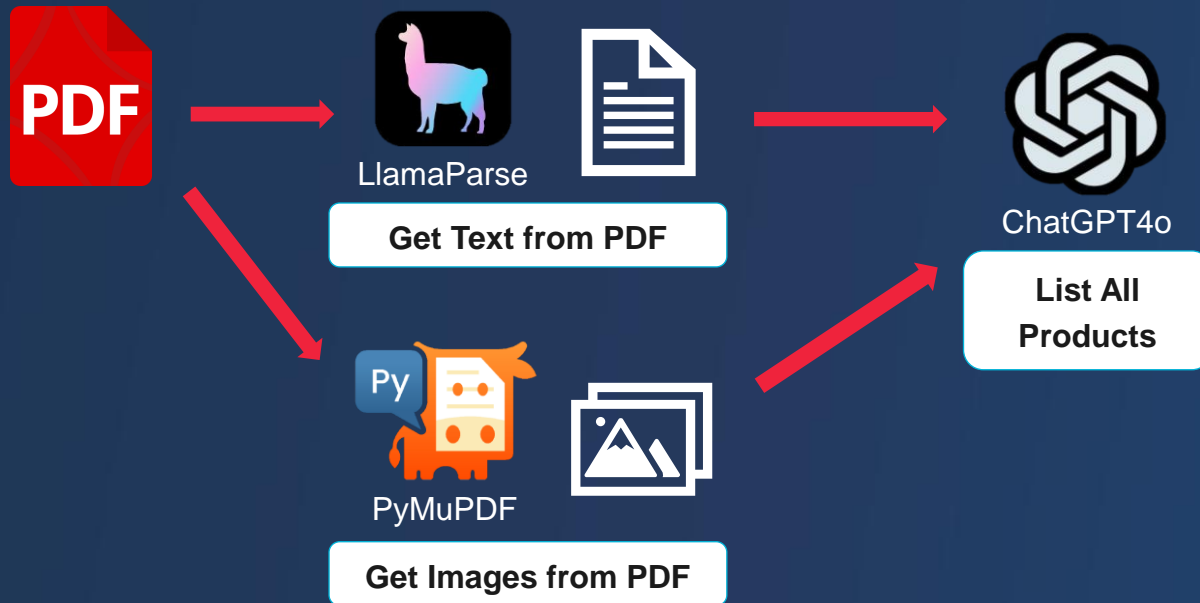
# Plan Of Attack



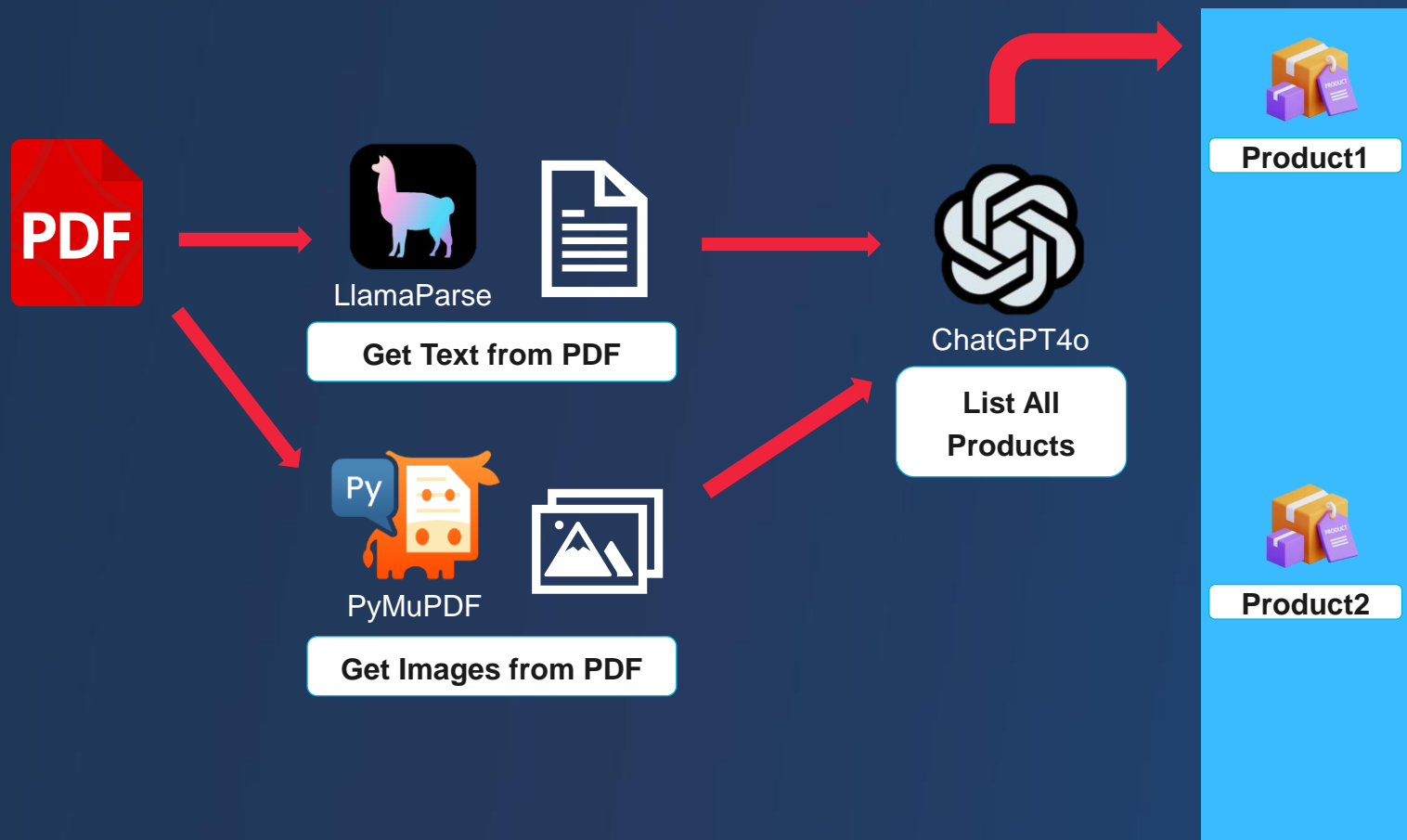
# Plan Of Attack



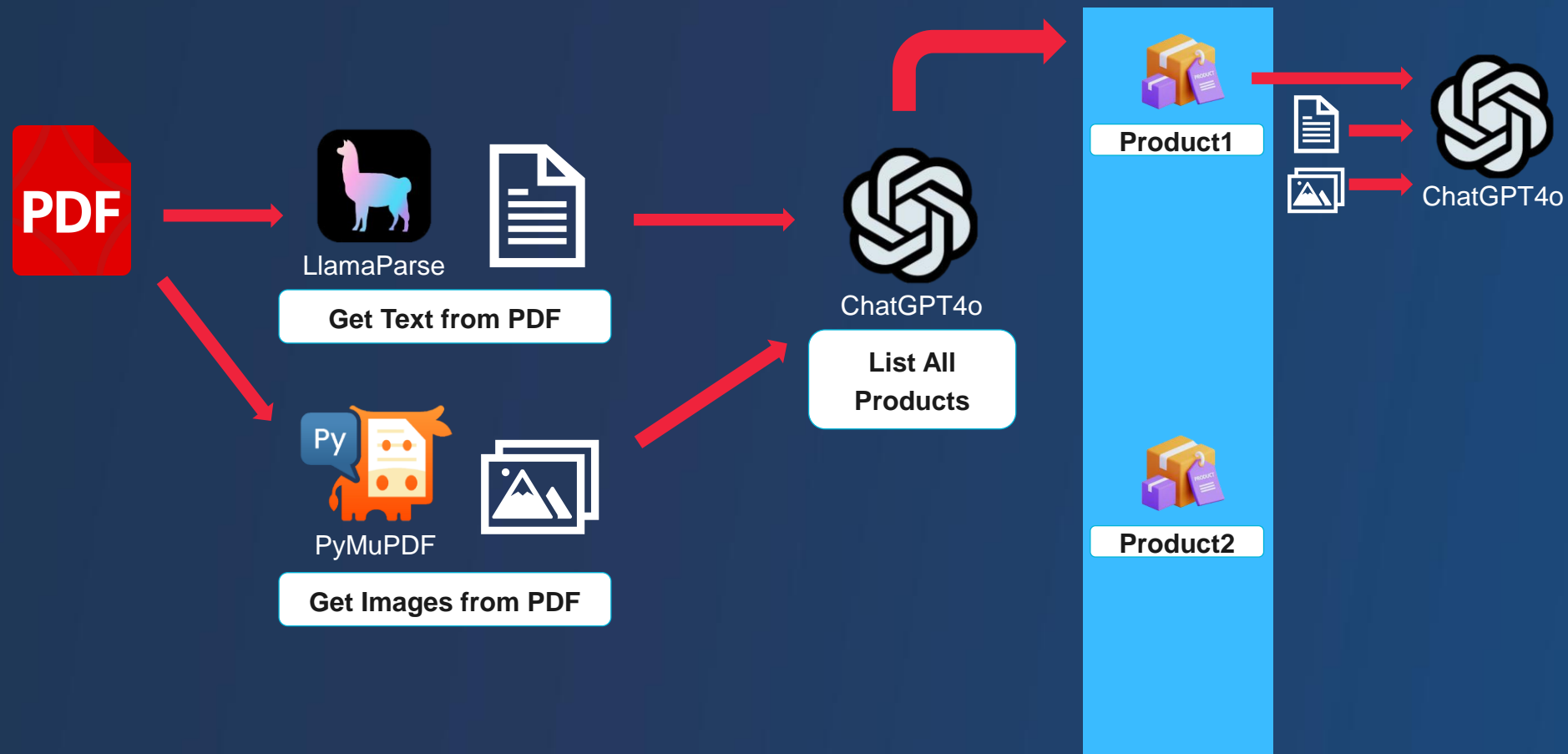
# Plan Of Attack



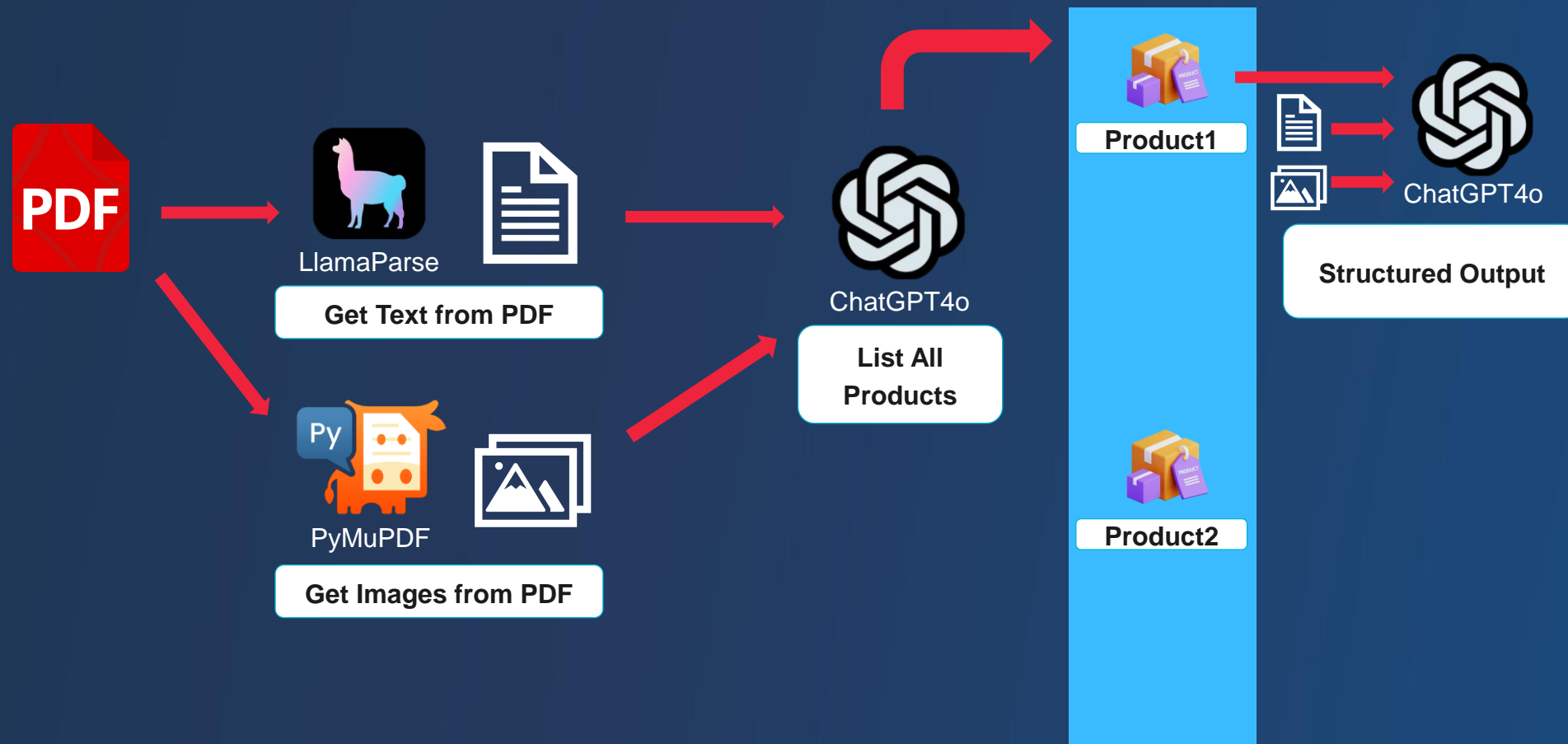
# Plan Of Attack



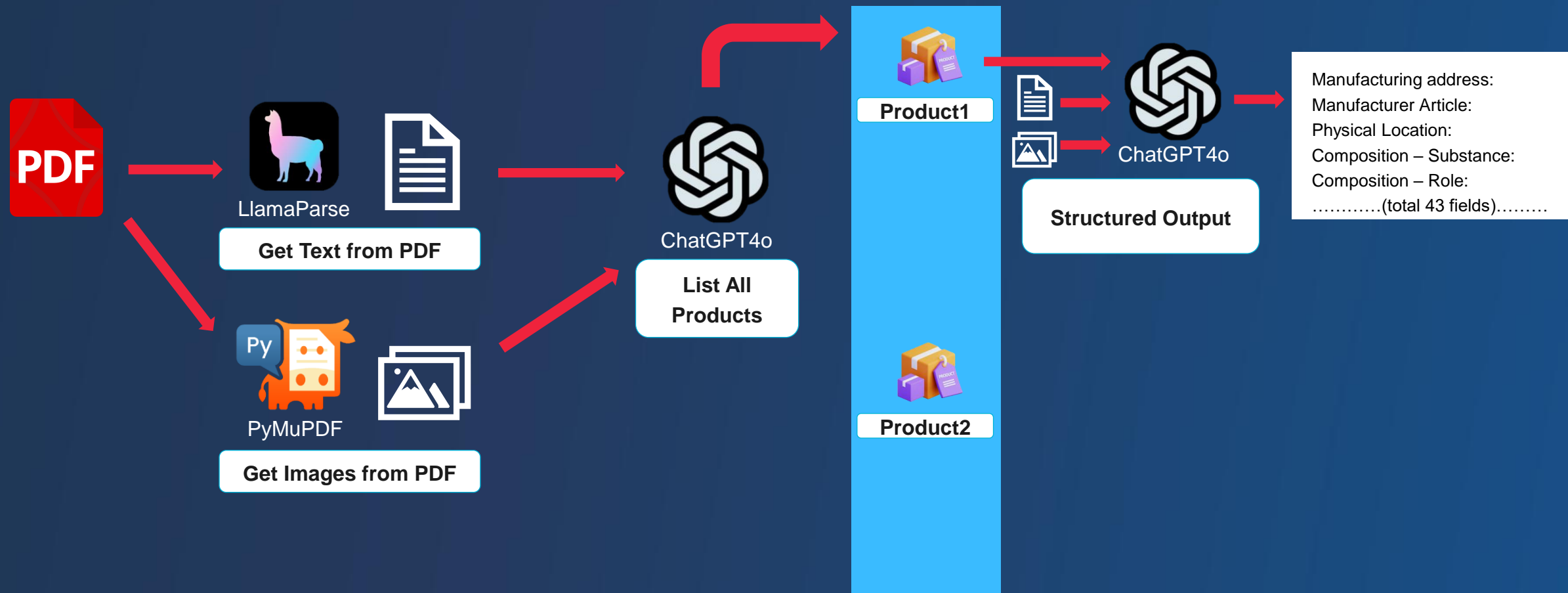
# Plan Of Attack



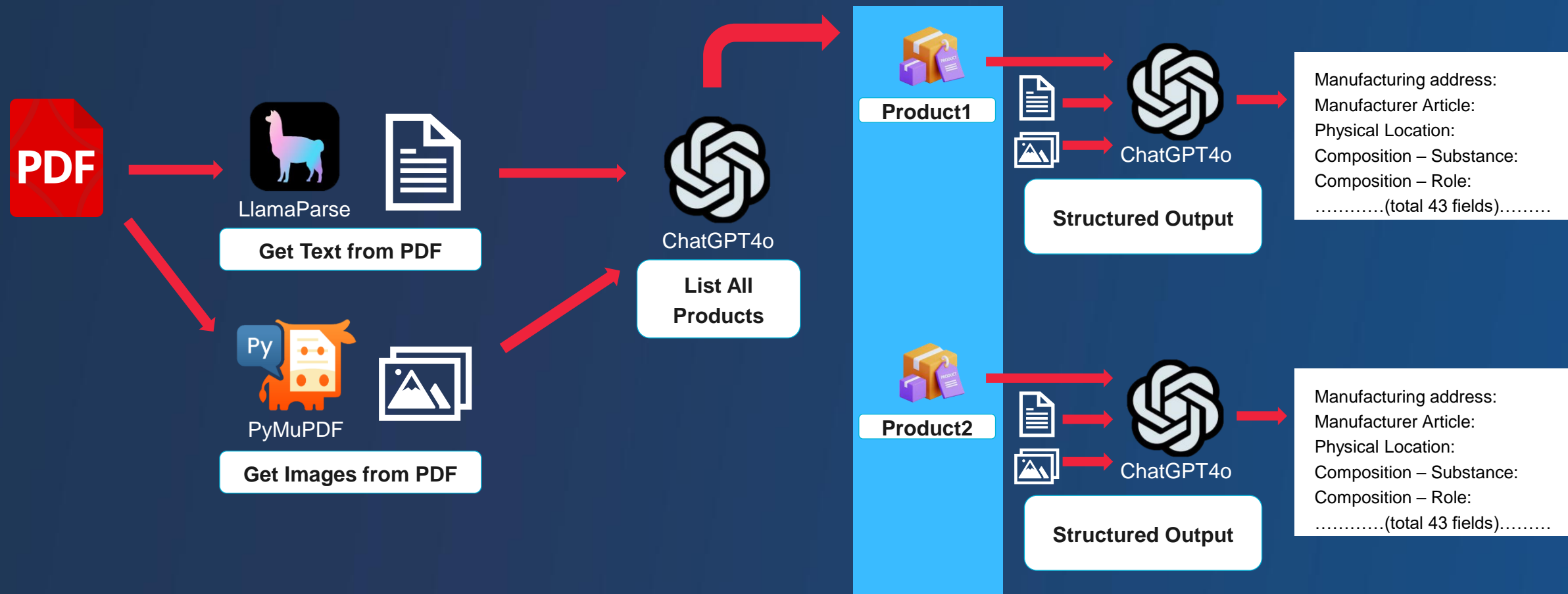
# Plan Of Attack



# Plan Of Attack



# Plan Of Attack





**Thank you for your attention**