ChatGPT

Commodity Article Processing – Comprehensive Logic and Decision Guide

Below is a detailed guide for processing commodity-related articles into structured data fields. Each component is explained in natural language, followed by a simplified decision tree outlining the logic flow, and a corresponding if-else pseudocode suitable for automation. This ensures clarity in understanding how each field should be derived from article content.

Commodity Column Logic

Natural Language Explanation

The **Commodity** column identifies the primary commodity (or commodities) affected or discussed in the article. Determine which raw material or resource is at the core of the article's content: - **Single Commodity Focus:** If the article clearly revolves around one commodity (e.g., *copper, crude oil, wheat*), record that as the commodity. For example, an article about *"Copper prices surge due to mine closures"* should be tagged with **Commodity = Copper**. - **Multiple Commodities:** If multiple commodities are mentioned with equal relevance (e.g., an article comparing *steel and aluminum shortages*), either list the primary commodity (if one is dominant) or use a general label like **"Multiple"** to indicate that several commodities are involved. - **No Specific Commodity:** If the article doesn't mention a specific commodity (perhaps it discusses a broad sector or an index), use a neutral value such as **"N/A"** or **"None"** as per guidelines. However, most commodity processing articles will reference at least one commodity explicitly.

The goal is to capture the key commodity that readers or analysts would associate with the event or issue described. Consider article titles, summaries, or repeated mentions to identify the main commodity.

Decision Tree Representation

- **IF** exactly one commodity is explicitly mentioned as central:
  → **Commodity =** that commodity name (e.g., Gold, Corn, Lithium).
- **ELSE IF** multiple commodities are discussed:
- **IF** one commodity is clearly primary (e.g., mentioned in the headline or focus):
  → **Commodity =** primary commodity name.
- **ELSE:** (no single primary, truly equal focus)
  → **Commodity = "Multiple"** (or list the top few commodities as appropriate).
- **ELSE:** (no commodity mentioned)
  → **Commodity = "N/A"** (or use a default category per guidelines).

If-Else Logic Format

```
if article.containsSingleCommodity():
    commodity = extractCommodityName(article)  // e.g., "Copper"
elif article.containsMultipleCommodities():
    primary = identifyPrimaryCommodity(article)
    if primary is not None:
        commodity = primary  // use the main commodity
    else:
        commodity = "Multiple"  // multiple commodities with equal focus
else:
    commodity = "N/A"  // no specific commodity mentioned
```

CAS ID Column Logic

Natural Language Explanation

The **CAS ID** column is meant to capture a unique identifier related to the commodity or event. In many contexts, *CAS ID* refers to a **Chemical Abstracts Service number** for chemical substances. If the commodity is a specific chemical or material with a known CAS registry number, that number should be recorded. If the commodity isn't a chemical or no such ID is applicable, then this field might be left blank or marked **"N/A."** Key considerations: - **Chemical Commodity:** For commodities that are chemical substances (e.g., *benzene, acetone, polyethylene*), lookup or use the standard CAS number (format: **NNNNN-NN-N**) corresponding to that chemical. *Example:* If the article is about **Acetone**, use CAS ID **67-64-1**. - **Non-Chemical Commodity:** For commodities like metals (gold, iron) or agricultural products (wheat, corn), there typically isn't a CAS number. In such cases, enter **"N/A"** or the appropriate placeholder since CAS ID doesn't apply. - **Multiple Commodities or General Topic:** If multiple commodities are involved (and especially if they include non-chemicals) or the article is not about a specific substance, use **"N/A"**. Only provide a CAS ID when a single, identifiable chemical substance is the focus.

In summary, this column is populated only when a precise chemical identifier is relevant, ensuring consistency and utility for chemical commodities.

Decision Tree Representation

- **IF** the identified commodity has a known CAS registry number (typically true for specific chemicals):
  → **CAS ID =** that chemical's CAS number.
- **ELSE** (commodity is not a chemical substance or no single commodity):
  → **CAS ID = "N/A"** (or leave blank if instructed, indicating not applicable).

If-Else Logic Format

```
if commodity is not None and commodity.hasCASNumber():
    cas_id = lookupCAS(commodity)  // e.g., "67-64-1" for Acetone
```

```
else:
    cas_id = "N/A"  // CAS ID not applicable for this commodity or no commodity
```

Activity Column Logic

Natural Language Explanation

The **Activity** column describes the nature of the activity or aspect of the supply chain that the article's event relates to. This categorization captures *what part of the commodity's lifecycle or market is affected*. Determine the primary activity as follows: - **Production-Related:** If the article deals with production changes – such as **mining output, manufacturing, plant operations, capacity expansions or shutdowns** – label the activity as **"Production."** For instance, *"Refinery maintenance cuts output by 20%"* would be a production activity. - **Supply/Distribution-Related:** If the focus is on **supply chain disruptions, inventory levels, or distribution/logistics** issues (e.g., shipping delays, supply shortages, transport problems), use **"Supply"** or **"Logistics"**. *Example: "Port congestion delays copper shipments"* falls under supply/logistics activity. - **Demand-Related:** If the article highlights **demand changes** or consumption trends (like surging demand from an industry, or demand drop due to market conditions), use **"Demand."** *Example: "Electric vehicle boom increases lithium demand"* is a demand activity. - **Market/Price-Related:** If the core of the article is about **price movements, market speculation, or trading** (e.g., price surge, futures market changes) without tying to a specific supply or production event, classify as **"Market"** or **"Price."** - **Policy/Regulatory:** If the event is about **government policy, tariffs, regulations, or political factors** impacting the commodity, you might label it as **"Regulatory"** (or similar term used in your schema). - **Other:** If it doesn't fit the above (for example, a technology development affecting the commodity, or an environmental event not directly production or supply), choose the closest category or an **"Other"** category if available.

Always choose the activity that best represents the *primary focus* of the article. If multiple activities are covered, pick the one that has the most direct impact on the commodity or is most emphasized.

Decision Tree Representation

- **IF** article is about production changes (new capacity, shutdown, maintenance):
  → **Activity = "Production."**
- **ELSE IF** article is about supply chain disruptions or inventory/logistics issues:
  → **Activity = "Supply"** (or **"Logistics,"** depending on terminology).
- **ELSE IF** article discusses changing demand or consumption trends:
  → **Activity = "Demand."**
- **ELSE IF** article focuses on price change or market speculation primarily:
  → **Activity = "Market"** (or **"Price"** related activity).
- **ELSE IF** article covers government/policy decisions affecting the commodity:
  → **Activity = "Regulatory"** (policy-related activity).
- **ELSE:** (for any other type of activity not covered above)
  → **Activity = "Other"** (or best fitting category).

If-Else Logic Format

```
if article.topic == "production_change":
    activity = "Production"
elif article.topic == "supply_disruption" or article.topic == "logistics":
    activity = "Supply"
elif article.topic == "demand_change":
    activity = "Demand"
elif article.topic == "price_market":
    activity = "Market"
elif article.topic == "policy_regulation":
    activity = "Regulatory"
else:
    activity = "Other"
```

Event Category Column Logic

Natural Language Explanation

The **Event Category** column classifies the broad **type of event or cause** described in the article. It answers the question *"What kind of event is this?"* by grouping events into high-level categories. Determine the category by identifying the nature of the incident or situation: - **Natural Disaster:** If the event is driven by natural causes such as **earthquakes, hurricanes, floods, wildfires, droughts** etc., use **"Natural Disaster."** For example, *"Floods in mining region disrupt iron ore output"* falls here. - **Industrial Accident/Disruption:** If it's a man-made accident or operational issue – e.g., **factory fires, explosions, equipment failure, power outage** – categorize as **"Operational Disruption"** or **"Industrial Accident."** - **Labor & Workforce:** Events stemming from human resources issues like **strikes, labor union protests, workforce shortages**, etc., go under **"Labor Disruption."** *Example: "Port workers strike affects copper exports."* - **Political/Regulatory:** If the cause is **government action, policy change, tariff, trade restriction, sanctions, or war/conflict**, use **"Political/Regulatory."** *Example: "New export ban on rare earth metals"* would be a regulatory event. - **Market Dynamics:** Use this for events that are **market-driven** without a single discrete incident, such as **price crashes, demand booms, or speculative bubbles**. Essentially, if the event is a trend or market condition (not a one-off incident), **"Market Dynamics"** fits. *Example: "Global oil oversupply leads to price slump."* - **Logistics/Transportation:** If the event centers on **transportation issues** (shipping, freight, port closures, logistics breakdowns), this can be a separate category like **"Logistics Disruption."** *Example: "Suez Canal blockage disrupts commodity shipments."* - **Other Category:** If none of the above categories capture the event (for instance, a new **technological development, financial crisis, or corporate action** affecting supply), use an **"Other"** or most relevant category available.

By classifying the event at a high level, this column provides a quick sense of what *kind* of disruption or change the article is about. It sets the context for the more detailed sub-category next.

Decision Tree Representation

- **IF** caused by natural hazards (earthquake, weather, etc.):
  → **Event Category = "Natural Disaster."**

- **ELSE IF** caused by accidents or operational failures (fire, outage, etc.):
  → **Event Category = "Operational Disruption"** (or **"Industrial Accident"**).
- **ELSE IF** caused by labor or workforce issues (strike, labor shortage):
  → **Event Category = "Labor Disruption."**
- **ELSE IF** caused by political or regulatory actions (laws, tariffs, war):
  → **Event Category = "Political/Regulatory."**
- **ELSE IF** a market-driven event/trend (price crash, demand surge without singular cause):
  → **Event Category = "Market Dynamics."**
- **ELSE IF** primarily a logistics/transportation issue (shipping delays, port issues):
  → **Event Category = "Logistics Disruption."**
- **ELSE:** (none of the above fits)
  → **Event Category = "Other"** (or closest matching miscellaneous category).

If-Else Logic Format

```
if event.cause_type == "natural":
    event_category = "Natural Disaster"
elif event.cause_type == "accident" or event.cause_type == "operational":
    event_category = "Operational Disruption"
elif event.cause_type == "labor":
    event_category = "Labor Disruption"
elif event.cause_type == "political" or event.cause_type == "regulatory":
    event_category = "Political/Regulatory"
elif event.cause_type == "market_trend":
    event_category = "Market Dynamics"
elif event.cause_type == "logistics":
    event_category = "Logistics Disruption"
else:
    event_category = "Other"
```

Event Sub-category Column Logic

Natural Language Explanation

The **Event Sub-category** provides a more **specific description of the event** within the broader category. After determining the general category, identify the precise type of event: - **Natural Disaster Sub-categories:** If *Event Category* is Natural Disaster, choose the specific disaster: e.g., **"Earthquake," "Flood," "Hurricane," "Drought," "Wildfire,"** etc., based on the article. *Example: "Earthquake"* for an earthquake in a mining region. - **Operational Disruption Sub-categories:** If *Category* is an accident or operational issue, specify the event like **"Fire," "Explosion," "Power Outage," "Equipment Failure," "Cyber Attack,"** etc., as described. *Example:* for a factory fire, sub-category = **"Fire."** - **Labor Disruption Sub-categories:** Under labor issues, identify **"Strike," "Lockout," "Labor Shortage," "Protest,"** etc. *Example:* a truckers' strike would be sub-category **"Strike."** - **Political/Regulatory Sub-categories:** If political/regulatory, specify the type: **"Tariff," "Export Ban," "Sanction," "New Regulation," "Political Conflict,"** etc. Use the article's details to choose. *Example:* a new tariff on steel imports → sub-category **"Tariff."** - **Market Dynamics Sub-categories:** For market-driven events, pick terms like **"Price Surge," "Price Drop," "Demand Surge,"**

“Demand Decline,” “Oversupply,” “Shortage,” etc., whichever best fits the situation. *Example:* if prices are skyrocketing due to speculation, sub-category = **“Price Surge.”** - **Logistics Disruption Sub-categories:** If logistics, specify the issue: **“Port Closure,” “Shipping Delay,” “Container Shortage,” “Transportation Cost Spike,”** etc. *Example:* a closed border crossing causing delays → **“Border Closure.”** - **Other:** If the event is unusual or doesn’t fall neatly into predefined sub-categories, use a descriptive term from the article. For instance, a **“Technological Breakthrough”** or **“Financial Collapse”** if that’s relevant.

Essentially, the sub-category should be a short descriptor that pinpoints what exactly happened, complementing the broader category with detail. It’s often directly gleaned from a keyword in the article describing the event.

Decision Tree Representation

- **IF** Category = Natural Disaster:
- Look for specific type (earthquake, flood, hurricane, etc.) in article.
  → **Event Sub-category =** that disaster type.
- **IF** Category = Operational Disruption/Accident:
- Identify the incident (fire, explosion, outage, etc.).
  → **Event Sub-category =** that incident type.
- **IF** Category = Labor Disruption:
- Determine nature of labor issue (strike, protest, shortage…).
  → **Event Sub-category =** that specific labor event.
- **IF** Category = Political/Regulatory:
- Identify policy action (tariff, ban, sanction, conflict…).
  → **Event Sub-category =** that specific action.
- **IF** Category = Market Dynamics:
- Identify market condition (price surge/drop, oversupply, demand spike…).
  → **Event Sub-category =** that condition.
- **IF** Category = Logistics Disruption:
- Identify logistics issue (port closure, shipment delay, etc.).
  → **Event Sub-category =** that issue.
- **ELSE:** (Category = Other or needs custom value)
- Use the article’s description for a fitting label.
  → **Event Sub-category =** descriptive term (e.g., “Technological Upgrade,” “Financial Crisis”).

If-Else Logic Format

```
switch(event_category):
    case "Natural Disaster":
        if article.mentions("earthquake"): event_subcat = "Earthquake"
        elif article.mentions("hurricane"): event_subcat = "Hurricane"
        elif article.mentions("flood"): event_subcat = "Flood"
        ... // other disaster types
        else: event_subcat = "Other Natural"
        break
    case "Operational Disruption":
        if article.mentions("fire"): event_subcat = "Fire"
```

```
        elif article.mentions("explosion"): event_subcat = "Explosion"
        elif article.mentions("power outage"): event_subcat = "Power Outage"
        ... // other operational issues
        else: event_subcat = "Other Accident"
        break
    case "Labor Disruption":
        if article.mentions("strike"): event_subcat = "Strike"
        elif article.mentions("shortage"): event_subcat = "Labor Shortage"
        elif article.mentions("protest"): event_subcat = "Protest"
        ... // other labor issues
        else: event_subcat = "Other Labor"
        break
    case "Political/Regulatory":
        if article.mentions("tariff"): event_subcat = "Tariff"
        elif article.mentions("sanction"): event_subcat = "Sanction"
        elif article.mentions("ban"): event_subcat = "Export Ban"
        elif article.mentions("war"): event_subcat = "Conflict"
        ... // other policy issues
        else: event_subcat = "Other Policy"
        break
    case "Market Dynamics":
        if article.mentions("record high") or article.mentions("price surge"):
event_subcat = "Price Surge"
        elif article.mentions("price drop") or article.mentions("plunge"):
event_subcat = "Price Drop"
        elif article.mentions("demand surge") or article.mentions("boom"):
event_subcat = "Demand Surge"
        elif article.mentions("oversupply") or article.mentions("glut"):
event_subcat = "Oversupply"
        ... // other market trends
        else: event_subcat = "Other Market"
        break
    case "Logistics Disruption":
        if article.mentions("port closure"): event_subcat = "Port Closure"
        elif article.mentions("delay"): event_subcat = "Shipping Delay"
        elif article.mentions("shortage of trucks") or
article.mentions("container shortage"): event_subcat = "Transport Shortage"
        ... // other logistics issues
        else: event_subcat = "Other Logistics"
        break
    default:
        // Category is Other or unspecified
        event_subcat = deriveCustomSubcategory(article)  // fallback to a
descriptive term from article
```

Moving Factor Column Logic

Natural Language Explanation

The **Moving Factor** column identifies whether the primary force at play is related to **supply or demand** (or other) regarding the commodity's change. It essentially asks: *What fundamental factor is "moving" the situation?* Determine this by examining the nature of the event: - **Supply Side Factor:** If the article's event deals with the **availability, production, or distribution** of the commodity (i.e., affecting supply), then the moving factor is **"Supply."** This includes events like production cuts, mine closures, export bans, logistic bottlenecks – all of which constrain or boost supply. *Example:* A mine shutdown reducing output → supply factor. - **Demand Side Factor:** If the event is driven by **changes in consumption or demand** for the commodity, choose **"Demand."** Examples include a surge in demand due to new applications or a drop in demand due to economic slowdown. - **Mixed/Other Factors:** Sometimes events have both supply and demand aspects or fall outside either category: - If both supply and demand elements are significant (e.g., one event causes supply drop and simultaneously demand spike), you might label **"Both"** or prioritize the dominant factor if one clearly outweighs the other. - If the factor is external or doesn't neatly fit (like a **regulatory change** or **macroeconomic factor** that indirectly influences supply/demand), decide which side it primarily affects. For instance, a tariff on imports is primarily a supply constraint domestically, so consider it a supply factor. - Some frameworks may allow an "Other" category for factors like technology or currency, but generally in commodity context it boils down to supply vs demand.

By specifying **Supply** or **Demand**, this column highlights the basic market force impacted by the event, which in turn usually signals the price direction (handled in the prediction field).

Decision Tree Representation

- **IF** event primarily affects **supply** (production/output/logistics of the commodity):
  → **Moving Factor = "Supply."**
- **ELSE IF** event primarily affects **demand** (consumption/need for the commodity):
  → **Moving Factor = "Demand."**
- **ELSE IF** both supply and demand are affected significantly:
  → **Moving Factor = "Both"** (if such a label is allowed; otherwise, choose the predominant factor).
- **ELSE:** (if factor is external/unclear but leans more to one side)
  → **Moving Factor =** whichever side is more impacted (or **"Other"** if using a separate classification).

If-Else Logic Format

```
if event.impact_area == "supply":
    moving_factor = "Supply"
elif event.impact_area == "demand":
    moving_factor = "Demand"
elif event.impact_area == "both":
    moving_factor = "Both"
else:
    // If not explicitly supply or demand, infer based on context or mark as
other
```

```
    moving_factor = inferFactorFromContext(event)  // e.g., could default to
"Supply" or "Other"
```

Type Column Logic

Natural Language Explanation

The **Type** column indicates the **nature of the article/event in terms of time frame or certainty**. It helps distinguish whether the content is about an actual event or a forecast/prediction, among possibly other distinctions. Common type designations might include: - **Actual Event (Real/Current):** If the article describes something that has already happened or is currently happening (a concrete event or change), classify it as **"Actual"** or **"Real Event."** For example, news of a factory fire or a current strike is an actual event. - **Forecast/Outlook:** If the article is projecting something that *might* happen or giving an analysis of future conditions (e.g., an analyst predicting a price rise, or a warning of potential shortage), use **"Forecast"** or **"Predictive."** *Example: "Report forecasts copper deficit in next year."* - **Ongoing/Developing:** Some articles describe an ongoing situation or evolving story (e.g., *"Strike continues for second week"*). In some schemas this might still be "Actual," but you could label it as **"Ongoing"** if that distinction is needed. - **Analysis/ Report:** If the article is more of a general analysis or a summary of trends (not news of a single event), the type might be **"Analysis"** or **"Review."** However, this is less about the event nature and more about article style; use only if your schema distinguishes it. - **Other Types:** Depending on guidelines, there might be other types like **"Announcement"** (for planned events or company press releases about future actions) or **"Policy Update"**. Ensure to follow any predefined type categories.

In summary, **Type** tells us whether we're dealing with a factual occurrence or a speculative/analytical piece. This guides how we interpret the information (e.g., actual events versus warnings of possible events).

Decision Tree Representation

- **IF** the article is reporting an event that has occurred or is happening now:
  → **Type = "Actual Event."**
- **ELSE IF** the article is primarily a prediction or outlook (future possibility):
  → **Type = "Forecast/Predicted."**
- **ELSE IF** the article discusses a situation that is ongoing or developing:
  → **Type = "Ongoing"** (if such a category is used, otherwise keep as Actual).
- **ELSE IF** the article is an analytical piece or retrospective report:
  → **Type = "Analysis/Report."**
- **ELSE:** (if none of the above apply or different classification exists)
  → **Type = "Other"** (e.g., "Announcement" for future planned events, etc., if defined).

If-Else Logic Format

```
if article.nature == "news" and article.timeframe == "current":
    type = "Actual"
elif article.nature == "forecast" or article.containsPredictions():
    type = "Forecast"
elif article.nature == "news" and article.timeframe == "ongoing":
```

```
    type = "Ongoing"
elif article.nature == "analysis":
    type = "Analysis"
else:
    type = "Other"
```

Repetitive Nature Column Logic

Natural Language Explanation

The **Repetitive Nature** column indicates whether the event described is **recurring or a one-off occurrence**. This helps in understanding if similar events are expected periodically or if this is an unusual incident: - **Repetitive/Recurring:** Mark this if the event is part of a cycle or a known repeated pattern. For example: - Seasonal or annual events (**seasonal floods, hurricane season impacts, yearly maintenance shutdowns**). - Regular strikes (if a particular industry strikes every contract cycle, etc.). - Cyclical market phenomena (like a price spike that happens every few years due to known cycles). In such cases, you might indicate **“Yes”** (it is repetitive) or use a term like **“Recurring”** or **“Seasonal”** if provided. - **Non-Repetitive/ One-off:** If the event is unusual, unexpected, or happens very rarely (no known cycle), then mark **“No”** (not repetitive) or **“One-time.”** For instance, a **major accident** or an **unexpected geopolitical event** (like an unanticipated coup affecting supply) would typically be non-repetitive. - **Uncertain:** If it’s not clear (for example, a shortage that could be part of a trend but not enough history to tell), you might note **“Unknown”** or make a judgment call. Generally, lean on historical precedent: if a similar event happened before regularly, it's repetitive.

This field often boils down to a **Yes/No** flag (or similar) indicating recurring nature. Use context: does the article or past knowledge hint that this is part of a pattern? If yes, mark it as repetitive.

Decision Tree Representation

- **IF** the event is part of a known cycle or pattern (seasonal, periodic, historically repeated):
  → **Repetitive Nature = “Yes”** (recurring event).
- **ELSE IF** the event is a unique occurrence with no regular precedent:
  → **Repetitive Nature = “No”** (one-time event).
- **ELSE:** (unclear or not enough info)
  → **Repetitive Nature =** make best assessment (default to “No” if uncertain, or use “Unknown” if allowed by schema).

If-Else Logic Format

```
if event.isRecurringPattern():
    repetitive = "Yes"
elif event.isOneOffIncident():
    repetitive = "No"
else:
```

```
    repetitive = "Unknown"  // or default to "No" if unsure and no category for
unknown
```

Resilinc Prediction Column Logic

Natural Language Explanation

The **Resilinc Prediction** column captures the **predicted outcome or consequence** of the event on the commodity's market, as assessed by Resilinc's analytics or your analysis. Essentially, after reading the article, what do we expect will happen to the commodity's supply, demand, or price? Key possibilities include: - **Price Increase (Bullish impact):** If the event is likely to tighten supply or boost demand, predict a **price increase** or **upward pressure** on prices. For example, a supply disruption (mine closure, export ban) or a surge in demand suggests prices will rise. - **Price Decrease (Bearish impact):** If the event could loosen supply or dampen demand, predict a **price decrease** or **downward pressure**. For example, discovery of new reserves (more supply) or a demand drop implies prices may fall. - **Supply Shortage or Constraint:** Sometimes phrased as an outcome, you might predict a **"shortage"** or **"supply crunch"** if the event significantly reduces available supply. - **Supply Surplus or Improvement:** Conversely, an event like a major production boost might lead to **"oversupply"** or improved availability, pushing prices down. - **No Major Impact / Stable:** If the event is not likely to significantly move the market (perhaps it's minor or already anticipated), the prediction could be **"No significant impact"** or **"Stable."** - **Other Predictions:** In some cases, the prediction might be about **lead times or delays** (e.g., *"likely delivery delays for 2-3 weeks"* if logistics are affected) or **substitution effects** (e.g., *"buyers may switch to alternative materials"*). Use whatever prediction format is standard (Resilinc might have specific labels).

The prediction should be a concise statement or label reflecting what **you expect to happen next** in the commodity market due to the event. It's often informed by supply-demand logic: - Supply down or demand up → upward price pressure (bullish). - Supply up or demand down → downward price pressure (bearish).

Decision Tree Representation

- **IF** event causes supply to decrease or demand to increase:
  → **Resilinc Prediction = "Price Increase"** (or **"Upward price pressure / Shortage"** as appropriate).
- **ELSE IF** event causes supply to increase or demand to decrease:
  → **Resilinc Prediction = "Price Decrease"** (or **"Downward price pressure / Surplus"**).
- **ELSE IF** event causes logistic delays (short-term disruption):
  → **Resilinc Prediction =** something like **"Delivery Delays"** or **"Short-term supply disruption"**.
- **ELSE IF** no clear market-moving impact:
  → **Resilinc Prediction = "No Significant Impact"** (market remains stable).
- **ELSE:** (other specific outcome as per context)
  → **Resilinc Prediction =** a custom prediction (e.g., **"Shift to alternate suppliers"** if that is the expected outcome).

If-Else Logic Format

```
if event.effect == "supply_down" or event.effect == "demand_up":
    prediction = "Price Increase"  // expecting upward price movement or
```

```
shortage
elif event.effect == "supply_up" or event.effect == "demand_down":
    prediction = "Price Decrease"  // expecting downward price movement or
surplus
elif event.effect == "logistics_delay":
    prediction = "Delivery Delays"  // supply impact in form of delays
elif event.effect == "no_change":
    prediction = "No Significant Impact"
else:
    prediction = formulateCustomPrediction(event)  // any other specific
predicted outcome
```

Prediction Maturity Period Column Logic

Natural Language Explanation

The **Prediction Maturity Period** column specifies the **time frame over which the predicted impact is expected to materialize or last**. It gives context to the Resilinc Prediction by indicating whether the effect is short-term or long-term: - **Short-Term:** If the impact is expected in the immediate or near future (often defined as within **0 to 3 months**). Use labels like **“Short-Term”** or a specific range (e.g., **“0-3 months”**). This is for events that have an immediate effect or resolve quickly (e.g., a short strike, a temporary outage). - **Medium-Term:** If the effect will take a bit longer to unfold or last, say **3 to 12 months** (up to a year). Label as **“Medium-Term”** or **“Mid-Term (within a year)”**. Example: a policy change that takes effect next quarter or a demand trend building over months. - **Long-Term:** If the consequences are expected to be sustained or only become significant in **over a year**. Use **“Long-Term”** (e.g., **“12+ months”**). This could apply to structural changes like new mines (which take years to develop) or prolonged geopolitical shifts. - **Immediate** (if used): Some systems might specify **“Immediate”** for things happening right now or within weeks, separate from short-term. Often though, “Short-Term” covers this. - **Ongoing/Indefinite:** If an event’s impact period is uncertain or continuous until further notice (for example, an indefinite ban or a permanent change), you might leave as **“Unknown”** or describe as **“Ongoing”** if that fits better than a fixed term.

Choose the period that best fits how the prediction is framed: - If the article or analysis suggests a timeframe (“prices will rise in the next few weeks”), match that. - If not explicitly stated, infer based on the nature of the event (e.g., a brief event likely short-term impact, a structural change likely long-term).

Decision Tree Representation

- **IF** predicted impact is expected in the near future (weeks to a couple of months):
  → **Prediction Maturity = “Short-Term”** (e.g., within 3 months).
- **ELSE IF** impact is expected in a moderate timeframe (several months, up to about a year):
  → **Prediction Maturity = “Medium-Term”** (3–12 months).
- **ELSE IF** impact will be realized over a long period (beyond a year or very sustained effect):
  → **Prediction Maturity = “Long-Term”** (12+ months).
- **ELSE IF** impact timing is immediate or happening now (if distinguished from short-term):
  → **Prediction Maturity = “Immediate”** (if such a category is used).

- **ELSE:** (if uncertain or continuous)
  → **Prediction Maturity = “Unknown/Indefinite”** or **“Ongoing”** (as per allowed values).

If-Else Logic Format

```
if expectedImpact.timeframe <= 3 months:
    maturity = "Short-Term"
elif expectedImpact.timeframe <= 12 months:
    maturity = "Medium-Term"
elif expectedImpact.timeframe > 12 months:
    maturity = "Long-Term"
else:
    maturity = "Unknown"  // or "Immediate"/"Ongoing" depending on context if
timeframe not clearly defined
```

Potential Impact Column Logic

Natural Language Explanation

The **Potential Impact** column gauges **how severe or significant the event’s impact could be on the commodity or supply chain**. It usually is a qualitative rating such as *High, Medium, Low* (or numerical levels) based on scope and intensity: - **High Impact:** The event likely has a **major effect** on the commodity’s availability, price, or supply chain. Criteria for High might include: - A large percentage of supply affected (e.g., a key producer nation’s output is halted). - A critical material with few alternatives experiencing disruption. - Expectation of noticeable price spikes or global ripple effects. *Example:* A war in a country that supplies 30% of a mineral would be High impact. - **Medium Impact:** A **moderate effect** is expected. Perhaps regional or limited in scope: - A noticeable portion of supply or demand is affected but there are buffers. - The commodity market can adjust in time or alternative sources exist. - Prices might move modestly or for a short duration. *Example:* A temporary factory outage causing a short-term dip in output (recoverable in a month) might be Medium. - **Low Impact: Minor effects** or easily mitigated: - The event affects a small producer or small region, or the commodity has ample stock/reserves. - Any price or supply changes will likely be marginal or short-lived. - Also use Low if the situation is contained or market had anticipated it. *Example:* A brief transport delay where inventory is sufficient to cover needs. - **Unknown/To Be Determined:** If you cannot assess impact from the information, or it could range widely, you might label it as **“Unknown”** or leave a placeholder. Typically, though, you make an educated assessment.

Assess the scale (global vs local), duration, and commodity’s market conditions (tight vs surplus) to decide impact. Always align with any defined criteria your organization uses for High/Med/Low.

Decision Tree Representation

- **IF** event affects a **large portion of supply/demand** or likely to cause **major price change**:
  → **Potential Impact = “High.”**
- **ELSE IF** event has a **moderate, contained effect** (noticeable but not disastrous):
  → **Potential Impact = “Medium.”**

- **ELSE IF** event has **minimal effect** or is easily offset/mitigated:
  → **Potential Impact = "Low."**
- **ELSE:** (if impact is unclear or data insufficient)
  → **Potential Impact = "Unknown"** (or choose a cautious Medium if forced to pick).

If-Else Logic Format

```
if event.scope == "global" or event.percentage_supply_affected > HIGH_THRESHOLD:
    impact = "High"
elif event.scope == "regional" or event.percentage_supply_affected >
MEDIUM_THRESHOLD:
    impact = "Medium"
elif event.scope == "localized" or event.percentage_supply_affected <=
MEDIUM_THRESHOLD:
    impact = "Low"
else:
    impact = "Unknown"
```

(In the pseudocode above, `HIGH_THRESHOLD` *and* `MEDIUM_THRESHOLD` *represent whatever cut-offs are defined for High/Medium impact in terms of supply share or other metrics.)*

Article URL Column Logic

Natural Language Explanation

The **Article URL** column should contain the **direct web link (URL) to the source article** being processed. This allows others to reference the original source for more details. Ensure the URL is captured and formatted correctly: - It should be the **specific URL of the article** (not just a homepage). For example, *"https://www.newswebsite.com/commodity/2025/05/article-title"*. - **Copy exactly** from the source to avoid broken links. Include the `http://` or `https://` prefix as required. - If multiple sources were used, typically choose the **primary or original source**. If your system allows multiple URLs, you might list them, but generally one main URL is expected. - If the article is not from the web (e.g., a PDF report or a physical publication), use the best available reference: - For a PDF or file, provide the file link or location. - For print-only info, you might put **"N/A"** or a description (though ideally every entry has a URL or source). - Quality check: make sure the URL is accessible and not behind a login (if it is, perhaps note it, but usually commodity articles are publicly accessible news).

This field doesn't require logical branching beyond ensuring the presence of a valid link. It's mostly about correctness and completeness of the reference.

Decision Tree Representation

- **IF** a single source article URL is available:
  → **Article URL =** *that URL* (ensure it's complete and direct).
- **IF** multiple sources were used:
- **IF** only one URL can be listed: use the **primary source's URL**.

- **ELSE IF** multiple URLs are allowed: list all relevant URLs (each separated as per format).
- **ELSE:** (no URL available, e.g., info came from print or internal source)
  → **Article URL =** *"N/A"* or a placeholder (and possibly provide source info in summary or notes).

If-Else Logic Format

```
if article.url is not None:
    url_field = article.url  // primary source URL as a string
elif article.multiple_sources:
    url_field = selectPrimaryURL(article.sources)  // choose main URL from
sources
else:
    url_field = "N/A"  // no URL available
```

Summary Column Logic

Natural Language Explanation

The **Summary** column contains a **brief summary of the article**, highlighting the key points relevant to the commodity and event. The goal is to condense the article into a few clear sentences: - **Include the Essentials:** A good summary should mention **what happened, where, and to whom/what**, as relevant: - Identify the **event** (e.g., factory fire, new policy, record demand). - Identify the **commodity** involved. - If applicable, mention the **location or company** (e.g., "in Chile" or "at XYZ Corporation's plant"). - Include the **impact** or consequence in brief (e.g., "causing a 10% output loss" or "prompting price rise concerns"). - **Be Concise:** Typically 1-3 sentences. It should be comprehensive enough to understand the situation without reading the full article, but not overly verbose. - **Neutral and Factual:** Present the summary objectively, in your own words, without adding analysis beyond the article. Avoid copying long phrases directly (unless necessary for clarity) and avoid subjective language not in the article. - **Example:** For an article titled *"Flood hits copper mines in Peru, supply at risk,"* a summary might be: *"Severe flooding in Peru has hit several major copper mines, disrupting operations. The sudden outage is expected to tighten copper supply and may lead to higher prices in the short term."* This covers the event, location, commodity, and potential impact. - **Multiple Points:** If the article has multiple points (say it's a monthly report covering different commodities), focus on the part relevant to your commodity field entries, or split into separate entries if needed. Usually, each row in processing corresponds to one main story.

Always double-check that the summary is accurate and reflects the article content, since it will be used for quick reference and context.

Decision Tree Representation

- **Start with the Event:** Identify the main event or change described.
- e.g., "**Factory fire**, **New law**, **Surge in demand**," etc.
- **Identify Subject and Commodity:** Who/what is impacted and which commodity is involved.
- e.g., "...at **Company X's plant** affecting **aluminum output**..."
- **Mention Outcome/Impact:** Include the immediate consequence or expected result.
- e.g., "..., resulting in a **10% production cut** and **higher prices**."

- **Assemble into 1-3 sentences:** Combine the above elements logically in a brief narrative.
- **IF** multiple key details (where, what, impact) are present, consider splitting into two shorter sentences for clarity.
- **Review for Clarity:** Ensure the summary stands alone in explaining the situation clearly and correctly.

(The above can be visualized as: Event → Commodity/Subject → Impact, all in a concise narrative.)

If-Else Logic Format

```
summary = ""
// Pseudocode outline for building summary:
event = identifyMainEvent(article)        // e.g., "Flood", "Fire", "Policy
change"
subject = identifySubject(article)        // e.g., location or company involved
commodity = commodity  // from commodity field logic
impact = identifyImpact(article)          // e.g., output loss, price risk

if event and commodity:
    summary = f"{event} affecting {commodity}"
    if subject:
        summary += f" at {subject}"
    if impact:
        summary += f", resulting in {impact}."
    else:
        summary += "."
else:
    summary = article.generateBriefSummary()  // fallback to a default
summarization method
```

The pseudocode constructs a sentence using identified pieces. In practice, natural language generation might be more complex, but it follows this general logic.

Confidence Rating Column Logic

Natural Language Explanation

The **Confidence Rating** column is a measure of **how confident you are in the accuracy and reliability of the information and the above assessments**. Often expressed as **High, Medium, Low** (or a numeric scale), it depends on source credibility and clarity: - **High Confidence:** Use this when the information is from very reliable, verifiable sources and the facts are clear: - Article from a reputable news outlet, official press release, or an industry report. - Multiple sources confirm the same information. - The event details (what, where, when) are well-established and leave little doubt. - You have no reason to suspect major inaccuracies. - **Medium Confidence:** This is a middle ground: - The source is generally reliable but maybe it's a single source without secondary confirmation. - Or the information is plausible but not fully confirmed (e.g., an industry insider report, or early news that might change). - Some minor uncertainties exist, but

overall it's believable. - **Low Confidence:** Mark this if the information has notable doubt: - Source is questionable (a rumor, social media post, or an obscure publication with potential bias). - The article contains conflicting information or lacks detail. - It's a speculative piece or the situation is rapidly changing (so facts may be outdated or incorrect). - Essentially, you're not very sure the reality is as described. - **Use of Numeric Scales:** If your system uses numbers (e.g., 1-5 or 1-10), map High/Med/Low to those (like High = 5, Medium = 3, Low = 1, etc.), or use percentage confidence if required. But the logic remains the same: evaluate source and content certainty.

The confidence rating helps users of the data know how much trust to place in the information. It is subjective but should be justified by the evidence available.

Decision Tree Representation

- **IF** source is highly credible **AND** details are confirmed/clear:
  → **Confidence Rating = High** (information is very likely true and complete).
- **ELSE IF** source is reasonably credible **BUT** some details lack confirmation or clarity:
  → **Confidence Rating = Medium** (probable but with some uncertainty).
- **ELSE IF** source is dubious or information is unverified/unclear:
  → **Confidence Rating = Low** (significant uncertainty in accuracy).
- **ELSE:** (for other cases, e.g., conflicting info)
  → **Confidence Rating = Low** (default to cautious stance if in doubt; or use an intermediate value if partial conflict, but typically err on the side of lower confidence if not sure).

If-Else Logic Format

```
if article.sourceReliability == "high" and article.detailsConfirmed:
    confidence_rating = "High"
elif article.sourceReliability == "medium" or article.detailsPartial:
    confidence_rating = "Medium"
else:
    confidence_rating = "Low"
```

(In practice, you might have a more nuanced check, but this captures the essence: high confidence only if sure, medium for moderate certainty, low for doubtful cases.)

Confidence Reason Column Logic

Natural Language Explanation

The **Confidence Reason** column is a short explanation of **why you assigned the above confidence rating**. It provides context to the rating by citing source credibility or information quality: - If **High Confidence**, explain what gives you certainty: - e.g., "**Trusted source** (Reuters) and **multiple confirmations** of the event," or "Official company press release – data is direct from source." - You could say: *"High – Data confirmed by official press release and reported in several major news outlets."* - If **Medium Confidence**, clarify the uncertainties: - e.g., "Based on a **single source (Trade journal)**, no secondary confirmation yet," or "Information is plausible but **not officially confirmed**." - For example: *"Medium – Report from one industry*

publication, seems credible but not yet confirmed by other sources." - If **Low Confidence**, state the doubts: - e.g., "Source is a **rumor on social media** with no verification," or "Article has **conflicting reports** and unclear data." - For example: *"Low – Information comes from an unverified local blog, details unconfirmed."* - Be concise but specific: one sentence or fragment usually suffices. Mention either the source or the nature of info that influenced your confidence: - **Source credibility:** trusted, known, reputable vs anonymous, unverified. - **Confirmation:** multiple reports vs single source. - **Clarity:** clear facts vs speculative or contradictory info. - This reason helps anyone reviewing the entry to quickly grasp why they should trust or be cautious about the data.

Decision Tree Representation

- **IF** Confidence = High:
- Cite strong source/confirmation.
  → e.g., **Reason =** "Multiple reputable sources confirm this information."
- **IF** Confidence = Medium:
- Note the partial reliability or single source.
  → e.g., **Reason =** "Single source (moderately reliable), not yet confirmed elsewhere."
- **IF** Confidence = Low:
- State the cause for doubt.
  → e.g., **Reason =** "Unverified source or conflicting information, reliability in question."
- **ELSE:** (if using a numeric or other scale, adapt similarly)
- Provide analogous reasoning for the given level.

If-Else Logic Format

```
if confidence_rating == "High":
    confidence_reason = "Confirmed by reliable sources (official or multiple
independent reports)."
elif confidence_rating == "Medium":
    confidence_reason = "Based on limited/single source; plausible but awaiting
confirmation."
elif confidence_rating == "Low":
    confidence_reason = "Source credibility is low or info is unconfirmed/
contradictory."
else:
    confidence_reason = "Confidence level rationale not specified."
```

This completes the comprehensive guide. Each section provided the full logic in words, a decision-tree breakdown for quick reference, and pseudocode illustrating how one might implement the logic programmatically. Use these guidelines to ensure consistent and thorough processing of commodity-related articles into structured data fields.
