Training Guide for CharXIV Rework (Client Feedback)

1. Why We're Resetting

A significant number of tasks were rejected due to **low quality, lack of detail, or incorrect reasoning**. The client specifically flagged **logic gaps and incomplete CoTs** as the main weaknesses. However, 80% of approved tasks required no edits, which shows the potential is strong if we stick to consistent quality.

The goal now is **consistent clarity, accuracy, and logical flow across all tasks.**

---

2. Common Errors to Avoid

*Work items reveal an **overall 46% pass rate** (115 total errors), below thresholds. Primary driver of errors is incorrect or inaccurate information within chain-of-thought across work items. 49.06% showed a lack of detail (comprehensiveness), specifically missing steps or explanation to support CoT and 41.51% contained incorrect chart value readings and data point misinterpretations (factuality) that compound into work item rejections. Other errors included question misalignment and lack of derivation processes, which contribute to logical gaps and unexplained conclusions for 54% of the work items.*

Error Category	Count of Appearance	Percent of Total Errors
Lack of detail in CoT	26	49.06%
Factuality issues	22	41.51%
Incorrect relevance - external information	17	32.08%
Incorrect specificity	16	30.19%
Complexity issues	10	18.87%
Formatting errors	10	18.87%
Redundancy issues	7	13.21%
Lack of relevance	7	13.21%
TOTAL	**115**	**100.00%**

👉 Biggest issue = **CoTs that skip reasoning steps or leave gaps**.

---

3. Golden Rules for Chain of Thought (CoT)

A strong CoT must **teach the user how to solve the task**, not just give an answer. The following steps can be undertaken by **QAs** to verify the comprehensiveness of any CoT.

Step 1: Check that the goal has been clearly stated

- Ensure the prompt has been restated clearly in the task.
- Check that the goal is explicitly defined (e.g., compare values, identify trends, compute a ratio).
- Look for clarification of any ambiguities (e.g., “minimum value” is explained as “minimum along the y-axis” or “minimum within a range”).
- Verify the CoT does not jump directly to calculations; it should first establish why the question matters.

Step 2: Check that the visuals have been described in detail

- Confirm the figure is described in a structured way:
    - Chart/Plot title included.
    - Axes labels and units mentioned (and special scaling, if any).
    - Legend referenced by method/label names, not just colors.
    - Subplots identified clearly (top-left, bottom-right, etc.).
- Watch for vague descriptors (“the blue line”); the CoT must always anchor to legend terms and subplot positions.
- Ensure the description makes it possible for the reader to understand the figure before reasoning begins.

Step 3: Check that the solution strategy has been mentioned before the actual reasoning steps

- Verify there is a clear plan of action before any calculation.
- Check that the CoT explains why data points are being extracted.
- Confirm that formulas are introduced and explained before values are plugged in.
- Ensure variable assignments are justified and comparisons have a stated purpose (e.g., “to check if Method A consistently outperforms Method B”).
- Strategy should be concise but purposeful, not skipped.

Step 4: Check that the reasoning is sequential, accurate, and comprehensive

- Check that reasoning is broken into numbered sequential steps.
- Each step should contain only one atomic action: read → interpret → calculate.
- Ensure every step is anchored to chart elements (axes, points, ranges, subplots).
- Always **double-check extracted values**.
- Ensure related points are grouped into **comprehensive steps** (not fragmented mini-steps).
- Always reference graphs properly:
    - ✅ "Top-right subplot named [NAME], y-axis shows accuracy (%)"
    - ❌ "second graph" or "the line in red."
- Keep **formatting** consistent:
    - LaTeX for math.
    - **Units** included in reasoning.
    - **Same labels** used throughout.
- Explain the **relevance** of each part of the analysis.
    - **Don't assume** the user understands why certain graphs or data points are being analyzed.
    - If the prompt or image references external information (e.g., European date notation, specific technical terms not in the legend), **explicitly state that this information is being used and where it comes from.**

Numerical reasoning checks:

- Approximations use fractions or percentages when between ticks.
- Rounding is consistent (default = 2 decimals, use ≈\approx≈ with "approximately").
- Exact values use =.
- Language avoids vagueness ("slightly higher" must be expressed as something like "approximately one-third between 0.2 and 0.3 (≈0.23\approx 0.23≈0.23)").
- Logical flow is intact: each step must show **why** it leads to the next.

Step 5: Check that the final answer is stated clearly

- Confirm there is one definitive final answer (no ambiguity).
- Ensure the conclusion explicitly links back to the goal stated in Step 1.