

# AI Evaluation Insights

Over the past month, I've worked on projects that focus on improving how AI models understand instruction, recognize actions, and interpret visual data. In this project journal, I'll share what I learned about testing AI performance, creating training examples, and ensuring data quality without sharing any private project details.

## Instruction Following & Prompt Testing:

I evaluated how well AI followed natural language instructions, such as “rotate the box” or “zoom in on this object.” My focus was on clarity, accuracy, and consistency – checking whether the AI understood both the action and the context behind it. I also tested multiple variations or prompts to see how phrasing affected results, helping identify which instructions produced the most reliable model responses.

**Skills:** Prompt design, analytical testing, data quality, model behavior analysis

## Video Segmentation & Action Tracking:

I reviewed short video clips to identify specific visual changes or actions (for example, movement, rotation, or item interaction). For each change, I created reference examples so the AI could learn to detect similar actions in other clips. This helped improve the model's ability to recognize sequences and respond accurately to user instructions involving motion or transformation.

**Skills:** Patterns recognition, video data annotation, visual recognition

## Image Set Annotation:

I worked with image sets showing gradual changes in an object's position, angle, or context. My task was to pair written instructions (turn the bottle, open the book, etc.) with the correct image results, ensuring accurate cause and effect examples for model training. This process strengthened the AI's understanding of visual instructions and contextual matching between text and imagery.

**Skills:** Multimodal understanding, labeling accuracy, context matching

## Entity Tagging and Product Recognition:

I tagged products, clothing, and visual entities in short social media videos to help AI models recognize brands, textures, and categories. My annotations supported the development of systems that identify visual elements and connect them to relevant metadata, improving both product detection and contextual recommendations.

**Skills:** Data annotation, entity recognition, categorization building

**Quality Assurance & Consistency Checking:**

I regularly checked datasets for labeling consistency, corrected mismatches, and ensured all examples met project standards. This helped maintain balanced data across categories and improved the overall accuracy of the AI models being trained.

**Skills:** Data validation, quality assurance mindset, attention to detail, problem solving

Through hands-on AI training tasks, including product tagging, object movement, and video sequencing, I developed an understanding of how to create clear, reliable instructions for model learning. By refining task instructions and providing consistent outputs, I contributed to improving the quality and accuracy of AI model training. This experience has strengthened my skills in data operations, workflow optimization, and AI support, demonstrating the impact of careful, precise guidance on model performance.