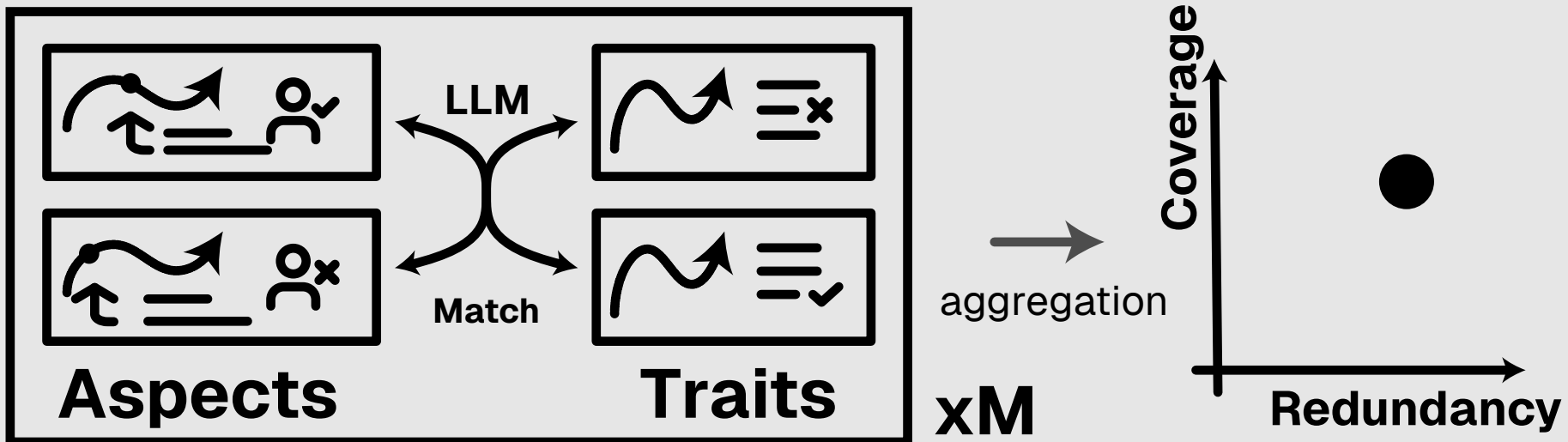


EVALUATION STEP 2

Meta Evaluation

Evaluate how much the detected traits from LLM-as-a-Judge could cover the human feedback aspects. A high **coverage** indicates the completeness of the induced metrics.



Coverage is the overall ratio of aspects matched with traits;
redundancy is the ratio of traits not matched any aspects.