What How Who **Scenarios Resources Quality Control Users** CardiologyPlastic Surgery Expert-guided Pulmonology Oncology Orthopedics Vascular Surgery **Data Selection** In-hospital Diagnosis (Period #3) Radiology Internal Medicine Emergency Medicin Human-in-the loop Real-world In-hospital **D**iagnosis (Period #1) Labeling & Rephrasing **EHR** Pediatrics Anesthesiology ■ Radiographer In-hospital **D**iagnosis (Period #2) Obstetrics and Gynecology **Filtering** Intensive Care Medicine
General Surgery
Thoracic Surgery Clinical Pathway Reasoning Re-checking Orthopedics Obstetrics and Gynecology In-hospital **D**iagnosis (Period #4) **Pediatrics Metrics** Online **Patient** Discharge Summary Ophthalmology Urology Consultation **Automatic metrics** Rouge Surgical Step Organization Neurology Rheumatology General Medicine Urology Case Summary Accuracy Neurosurger Examination **Pediatric Surgery** Doctor Orthopedics Plastic surgery Kendall's Tau False Information Test Cardiac Surgery Human evaluation Pulmonology Pediatrics False Treatment Test Correctness Paper/Book Basic Knowledge Test Oncology Pharmacist Completeness Genetics and Genomics Neurology Ophthalmology Obstetrics and Gynecology **K**eyword **E**xtraction Fluency Obstetrics and Gynecology **Psychiatry** Medicine Consultation Medical Surgery Friendliness Guideline Wrong Treatment Detection **TOrthopedics Validity** Cardiology Pediatrics Student Internal Medicine Verification **Toxicity Clinical Question Answering Ability** > Expert Rating **Evaluation Axes Information Retrieval Summarization Ability** Reasoning 1 Spearman's **Knowledge Application Ability** Hallucination Rank Correlation