iCliniq Dataset Overview

AI experts caution against using synthetic data—generated by large language models to mimic real-world data—in AI healthcare training, as LLMs often produce inaccurate information. Training on such data can amplify these inaccuracies

iCliniq offers a premium medical dataset, meticulously curated for AI development. Our dataset is a rich repository of real-world clinical interactions, expertly crafted by practicing physicians and rigorously reviewed by our in-house medical team. This ensures unparalleled accuracy, relevance and clinical depth.

Dataset Highlights:

- **Expert-authored content**: Our dataset is exclusively authored by practicing physicians and specialists, guaranteeing clinical precision.
- **Real-world patient interactions**: The data accurately reflects genuine patient concerns, providing invaluable insights for AI model development.
- **Time-saving efficiency**: Pre-labeled, curated and annotated, our dataset accelerates AI development by eliminating data preprocessing.
- **AI training optimized format**: Our structured Q&A format is tailored to enhance reinforcement learning with human feedback (RLHF) models.

Data Richness for Advanced AI Model Training:

Our dataset is meticulously structured to provide essential elements for robust AI model development:

- **Comprehensive patient profiles**: Detailed patient demographics, medical history, and symptoms offer a rich context for AI algorithms.
- **Actionable medical insights:** Clear and structured medical recommendations, including diagnostic tests, treatment plans, and follow-up care, provide valuable training data.
- **Precise medical terminology**: Consistent use of accurate medical language ensures optimal AI performance and understanding.
- **Enhanced model performance**: Our dataset, comprising over 60 million tokens, offers extensive data for optimizing AI model weights and biases, leading to improved accuracy and efficiency.

Unmatched Data Quality and Value:

iCliniq's dataset stands out in the market due to its exceptional quality, curation, and depth. Our data is:

- **Expert verified**: Rigorous review by medical professionals guarantees accuracy and reliability.

- **AI optimized**: Structured format and detailed information enhance model training efficiency.
- **Proven effective**: Our dataset has contributed to numerous academic research projects, demonstrating its value.
- **Trusted Brand:** We partner with renowned health systems like Memorial Sloan Kettering Cancer Center, where we power their second opinion program, and major e-commerce platforms like Amazon Pharmacy, which use our health content to engage their patients.

Dataset Scale and Investment:

Our dataset comprises approximately 70,000 meticulously curated examples, each representing over 30 minutes of specialist input and rigorous quality control. This substantial investment in data curation guarantees exceptional quality and depth.

Benefits for AI Development:

By leveraging our dataset, you gain:

- A rich source of labeled, high-quality medical data
- A foundation for building trustable and effective AI applications
- A competitive edge in the rapidly evolving healthcare AI landscape

Example Q&A on Our Platform:

Example 1:

Question: "I am a 20-year-old female. I hit my elbow at work the night before last, and it hurts to straighten my arm. I can do it, but it is painful. And randomly, I suddenly have shooting pain in my arm. It is swollen, and there is a small lump where I hit it. I also fractured my elbow in the same place when I was a child. I do not know what to do to make it better. But I am having trouble sleeping and I have to go to work tomorrow, and I need to be able to work without pain."

Answer: "From your history, it looks like a blunt injury with a hematoma at the site of the injury. I suggest getting an X-ray of the elbow to rule out a fresh fracture. If there is no fracture, then use an arm pouch and keep the elbow at 90 degrees flexion in the arm pouch for three weeks. I also suggest taking Chymoral forte (Chymotrypsin and Trypsin - enzymes to break down hematoma and swelling) and analgesics like Tramadol-Paracetamol combination for acute pain. Consult a specialist doctor, discuss with him or her, and start taking the medicines with their consent. Give cold compression at the site of injury and swelling."

Example 2:

Question: "I felt the lump seven months back but did the test four months ago and had the surgery three months ago. I have been diagnosed with early-stage breast cancer and had an excisional biopsy procedure where the tumor, safety margin, and three lymph nodes have been dissected. My doctor told me that I need to undergo chemotherapy followed by 20 sessions of radiotherapy, then I will start hormonal therapy. I have started chemotherapy with 130 milliliters

of Taxol 6 mg/L weekly, took five doses till now, and there are seven more. My hair started to fall, and I feel very weak. I cannot lift my arm. Sometimes I feel my heart pounding very fast. And I am starting to think, did I really need all of this even after removing the lump completely? I am attaching the biopsy report that revealed ER, PR positive and HER2 negative. Please advise me."

Answer: "I have gone through your query and attachments (attachments removed to protect the patient's identity). Being an oncologist involved in the care of breast cancer patients, I can understand your concerns. Based on the query, I can gather a few things:

1. You are 29, which is quite early to develop breast cancer.
2. You have undergone breast conservative surgery (BCS) along with sentinel lymph node surgery of the axilla (SLNB). This is the current standard of care (SOC).
3. Your breast cancer is hormone-positive, meaning it is ER (estrogen) or PR (progesterone) positive and HER2 (human epidermal growth factor receptor) negative.

Treatment for this will need:

4. Surgery and SLNB, which you have undergone.
5. Radiation is needed as you have undergone breast conservative surgery. Radiation reduces the risk of the tumor coming back to the same site. If you have undergone removal of the complete breast, radiation may have been avoided depending on final histopathology.
6. Hormonal treatment as your tumor is hormone-positive, it will respond to hormonal treatment. This will be started once you complete your chemotherapy and will be continued for a minimum of five years. It may be extended for ten years, depending on tolerance. I want to tell you that it is very easy to take. These three things are needed in every case.
7. Chemotherapy need is decided on HPE (histopathology) report. For this, I would like to see the entire report. Also, will need to know if you have undergone any additional testing to see the risk of tumor recurrence. Tumors can have different risks of recurrence based on histopathology and genetic features. If the risk is high, chemotherapy is definitely needed to reduce risk by an absolute 10%. Considering all this and your age, most oncologists will be inclined to give you chemotherapy as 29 is too early for CA (carcinoma) breast. I would also like to know if there is any family history of CA breast? And would consider doing genetic testing as every breast cancer below 50 has an almost 15% chance of being genetic. I would be able to give you a clearer answer regarding chemotherapy if you can provide me with histopathology details after surgery and reports of any additional tests. If you have any other questions, please revert here."

Invest in data that drives results. Contact us today to discover how iCliniq's dataset can elevate your AI initiatives.