# iCliniq Capabilities and Offerings  - Data Annotation

Every healthcare organization is collecting enormous volumes of data, including patient queries, clinical notes, lab reports, images, videos, and EHR records. But the challenge is that most of this data is unstructured. It is present in free-text documents, scanned files, or fragmented systems, making it almost impossible to use directly for AI, analytics, or regulatory reporting.

This is exactly where annotation becomes critical: turning raw data into structured, standardized, and AI-ready datasets. At iCliniq, we specialize in this transformation, helping partners convert scattered clinical information into actionable knowledge. Our strong foundation as a global telemedicine platform, with millions of real-world patient interactions, uniquely positions us to deliver annotation that is both clinically accurate and context-aware.

**The Need for Annotation-Met by iCliniq:**

- For pharmacovigilance, clinical trial safety monitoring, and regulatory submissions, where every adverse event or indication must be consistently coded and reported. iCliniq's medically trained annotators ensure datasets are aligned with ICD-10 and MedDRA standards, reducing compliance risks and improving regulatory readiness.
- To train models for diagnostics, symptom triaging, and decision support, which requires large, high-quality, labeled datasets. Leveraging our deep expertise in real-world clinical queries, iCliniq delivers annotations that mirror how patients present symptoms and how doctors respond, making training datasets both clinically relevant and AI-ready.
- To make sense of patient records, organize EHRs, and unlock insights that improve outcomes, streamline workflows, and support clinical research. With iCliniq's support, healthcare institutions can extract structured intelligence from years of unstructured records, fueling better care decisions and advancing medical research.

Without annotation, data remains fragmented, inconsistent, and unusable. Annotation transforms raw text, scans, or recordings into structured, coded, and AI-ready assets. For example:

- A free-text complaint of "breast lump" becomes mapped to ICD-10 (N63.x) and MedDRA (Breast lump).
- A DICOM MRI scan becomes segmented and linked to its radiology report for AI training.
- A pathology report becomes tagged with structured findings that can be used for research.

At iCliniq, this transformation is not just theoretical. The exact value we deliver to clients is achieved by combining clinical expertise, compliance-first workflows, and human-in-the-loop quality checks. We ensure annotation that is accurate, scalable, and trusted. This helps in unlocking the full potential of the healthcare data.

**Data Annotation Services at iCliniq:**

At iCliniq, we have developed a doctor-led annotation framework that covers the full spectrum of healthcare data:

- Patient Queries – Authentic patient-physician interactions.
- Case Summaries and Clinical Notes – Structured SOAP format documentation.
- Reports – Laboratory findings, radiology, and pathology reports.
- Medical Images – Case summary linked images and educational resources.
- Videos – Structured condition-wise patient education content.
- Textual Health Articles and Knowledge Content – Multi-level tagged medical information.

What began as our internal framework to structure iCliniq's own data has now evolved into a service we offer to clients. Whether it is a pharma company, hospital, AI/MedTech startup, or research institution, we can apply the same proven framework to the datasets, making them structured, coded, searchable, and AI-ready.

Our strength lies not only in identifying the different modalities of healthcare data but also in applying a consistent annotation strategy across them. That's what makes iCliniq a reliable partner for organizations seeking scalable, compliant, and clinically meaningful annotation services.

**iCliniq's Annotation-as-a-Service Offerings:**

**1. Textual Content Annotation:** We annotate medical articles, educational content, and clinical notes with a multi-level tagging framework:

- Category – Diagnosis, treatment, prevention, lifestyle, etc.
- Specialty – Cardiology, neurology, dermatology, etc.
- Primary Medical Condition – The main condition or disease.
- Secondary Medical Condition – Related medical conditions, risk factors, and comorbidities.
- Signs and Symptoms – Patient complaints explicitly tagged for discoverability.

**For Our Clients:** This ensures large text datasets are structured, easily retrievable, and mapped to global standards (ICD-10, MedDRA), making them usable for training NLP models, powering knowledge systems, or supporting clinical search engines.

**2. Patient Query Annotation -** Over the past decade, iCliniq has built a robust annotation framework, originally developed to organize millions of patient–physician interactions on our own platform. With a network of over 4500 physicians across specialties, we bring deep medical expertise to every annotation task.

Our approach is not just annotation; it is multi-layered labeling, expert labeling, and structuring designed for scalability, licensing, and real-world usability. Our expertise lies in structuring

authentic patient–physician interactions, where unstructured free text is converted into structured data. Our multi-layer annotation framework has been designed to offer the following:

**A. Medical Coding (Standardization and Compliance):** Every patient complaint, symptom, and physician's diagnosis is coded against international standards:

- ICD-10 → Diseases and conditions.
- MedDRA → Symptom and adverse event classification for pharmacovigilance.

Ensures datasets are regulatory-grade, interoperable, and ready for global healthcare use.

**How iCliniq Uses MedDRA — Our Capabilities**

**a. Patient Queries and Answer Annotation -** Tag patient questions and physician answers with MedDRA terms.

- Example: Patient says, "I feel chest tightness and breathlessness" → mapped to MedDRA "Dyspnoea" (code: 10013968).
- Benefit: Creates searchable, structured content for research, AI, and clinical analytics.

**b. Content Standardization -** Map diverse expressions of the same condition to a single MedDRA concept.

- Example: "Gastric ulcer," "Peptic ulcer," "Stomach sore" → unified under one MedDRA term.
- Benefit: Prevents duplication, improves search accuracy, and enhances data consistency across clinical tags.

**c. Adverse Drug Reaction (ADR) Reporting -** Tag patient-reported side effects using MedDRA for pharmacovigilance.

- Example: "I got swelling after taking amlodipine" → MedDRA: "Oedema peripheral."
- Benefit: Supports regulatory reporting and safety monitoring for pharma companies and health authorities.

**d. Analytics and Insights -** Generate structured reports on patient-reported symptoms and trends.

- Example: "Top 50 symptoms reported this quarter."
- Benefit: Identifies emerging health trends, comparable to WHO or FDA safety signal monitoring.

**e. AI Training Data -** Annotate patient–physician interactions with MedDRA to create high-quality training datasets for ML/NLP models.

- Use cases: Chatbot triage, symptom checkers, and AI-driven decision support.
- Benefit: Ensures AI speaks standardized medical language aligned with pharma and regulatory standards.

iCliniq is already implementing all of these MedDRA-based annotation workflows. We deliver clinically accurate, structured, and standardized datasets from both text and imaging sources, enabling research, AI development, pharmacovigilance, and regulatory readiness.

**B. SOAP Structuring (Clinical Documentation) -** We convert free-text conversations into SOAP notes (Subjective, Objective, Assessment, Plan).

This transforms fragmented exchanges into clinically recognizable documentation aligned with global EHR standards.

**C. Symptom and Condition Tagging (Granularity) -** Every interaction is explicitly tagged with:

- Primary medical conditions (diagnosis).
- Secondary medical conditions (comorbidities).
- Signs and Symptoms (complaints).
- Risk factors (lifestyle, sexual history, exposures).

Enables fine-grained discoverability for clinical and AI applications.

**D. Contextual Annotation (Layered Intelligence) -** Queries and responses are segmented into two core blocks:

- Patient's Concern or Worry → The raw voice of the patient.
- Physician's Case Summary/Response → The structured medical evaluation and advice.

Each block is further annotated by:
- Category: Diagnostic, therapeutic, preventive, lifestyle.
- Specialty: Dermatology, Cardiology, Endocrinology, etc.
- Voice Separation: Distinguishing patient from physician for conversational AI.
- Primary vs. Secondary tags: Core condition vs. associated elements.

In the data licensing market, this is delivered as labeled and expert-labeled datasets, ready for syndication, licensing, or downstream AI use.

**E. Treatment and Medication Annotation (Therapeutic Mapping)** - Mentions of medications, interventions, and management strategies are annotated for:

- Drug name.
- Drug class.

- Clinical purpose and observed outcome.

Creates real-world treatment datasets for pharma companies and AI-driven research.

**3. Image Annotation -** iCliniq provides end-to-end annotation of patient records (images, scanned reports), radiology DICOMs, pathology slides, and other imaging data using single-expert and multi-expert labeling workflows. All image and DICOM findings can be coded to MedDRA.

**iCliniq Image and DICOM Annotation Offerings:**
**A. Annotation of Patient Records**

- Scope: Includes images/photographs, scanned lab reports, prescriptions, and other patient-submitted images.
- Value: Converts real-world patient uploads into structured, clinically meaningful datasets, enabling hospitals and researchers to gain actionable insights from everyday clinical interactions.

**B. Radiology DICOM Annotation**

- Scope: Annotation of CT, MRI, X-ray, and ultrasound DICOM series, capturing lesions, organs, and abnormal findings.
- Value: Supports AI model training, clinical research, and regulatory submissions with accurate, high-quality, expert-reviewed imaging data.

**C. Pathology Slide Annotation**

- Scope: High-resolution pathology slides annotated for tissue types, cellular abnormalities, and disease markers.
- Value: Facilitates precise research, clinical trial endpoints, and AI-driven diagnostic applications by creating structured, reproducible datasets from complex histopathology data.

**D. Single-Expert Labeling**

- Scope: Each report or DICOM file is reviewed and labeled by a single clinician or specialist.
- Value: Provides a fast, cost-effective solution for straightforward or well-documented cases, while maintaining clinical accuracy and consistency.

**E. Multi-Expert Labeling**

- Scope: Multiple experts review each report or DICOM file, and final labels are determined by consensus or majority vote.

- Value: Ensures the highest reliability and reproducibility for ambiguous or high-stakes tasks, such as cancer detection or complex disease annotations, making it suitable for regulatory-grade datasets and AI validation.

**6. MedDRA-Coded Annotation**

- Scope: All annotations are standardized and mapped to MedDRA terms for clinical and regulatory use.
- Value: Delivers structured, universally interpretable datasets, enabling seamless integration into research, AI pipelines, pharmacovigilance, and regulatory submissions.

**4. Video Annotation: iCliniq's Annotation Offerings for Healthcare Videos**

Healthcare video libraries, whether patient education resources, training modules, or recorded consultations, often remain unstructured and hard to navigate. iCliniq's video annotation framework transforms them into structured, searchable, and AI-ready datasets.

**A. Condition-Level Annotation -** Each video is anchored to a primary medical condition (e.g., Migraine, Diabetes, Asthma). Ensures every piece of content has a clear clinical identity, making large video repositories easy to segment by disease.

**B. Subtopic Annotation -** Videos are broken down into structured medical subtopics, which are universally relevant across conditions:

- Symptoms – Key patient complaints discussed.
- Diagnosis – Tests, imaging, and criteria mentioned.
- Treatment Options – Medications, therapies, lifestyle advice.

**C. Lifestyle Modifications – Diet, exercise, stress management, prevention.**

**Prognosis/Outcome –** Long-term outlook and disease course. This converts long-form videos into modular, taggable content units.

**D. Multi-Layer Tagging (Contextual Data Labeling) -** Videos are tagged using iCliniq's multi-layer annotation framework:

- Category: Educational, diagnostic, preventive, lifestyle.
- Specialty: Neurology, endocrinology, cardiology, etc.
- Primary Tags: Core condition.
- Secondary Tags: Symptoms, risk factors, and comorbidities.

Delivered as labeled and expert-labeled datasets that clients can license, syndicate, or plug into AI pipelines.

**E. Cross-Linking With Other Data Modalities - Videos can be linked to related:**

- Articles or knowledge resources.
- Patient queries or case summaries.
- Medical records (e.g., DICOM, X-rays, pathology slides).

Builds multi-modal datasets where one condition is represented consistently across text, images, and video.

**F. AI and Research Enablement - Annotated video datasets become ready for:**

- Machine learning training (NLP and video analysis models).
- Condition-wise knowledge hubs for hospitals and research institutes.
- Patient-facing education platforms with structured navigation by medical condition or symptom.

**Why Does This Matter?**
With iCliniq's video annotation framework, any healthcare video dataset can be transformed into:

- Condition-wise learning hubs → Videos grouped by disease and clinical aspect.
- Searchable knowledge bases → Easily indexed by condition, symptom, or specialty.
- Licensing-ready video assets → Delivered as labeled/export-labeled datasets for syndication.
- AI-training datasets → Optimized for ML pipelines in healthcare education and predictive analytics.

And because all annotations at iCliniq are physician-led and quality-checked, clients receive clinically accurate, compliant, and scalable video datasets, far beyond what generic tagging vendors can deliver.

**Clinician-Led Annotation Across All Modalities -** At iCliniq, every annotation is reviewed and validated by experienced medical professionals, including clinicians, radiologists, pathologists, and domain specialists.

- Single-Expert Labeling: Fast, accurate annotation of well-documented or routine data.
- Multi-Expert Labeling: Consensus-based review for complex, high-stakes, or ambiguous cases, ensuring regulatory-grade reliability.

All annotations, whether text, reports, patient queries, images, DICOMs, pathology slides, or videos, are MedDRA-coded and cross-referenced to ICD-10, delivering standardized, globally interpretable, and compliant datasets.

**Key Differentiators That Set iCliniq Apart**

- **Expert-Led Annotation:** Experts like Clinicians, radiologists, and pathologists validate all the datasets.
- **Multi-Modal Coverage:** End-to-end annotation of text, queries, reports, images, DICOMs, slides, and videos.
- **Standardized and Compliant:** MedDRA and ICD-10 mapping ensure global usability.
- **Flexible, Scalable Workflows:** Single- or multi-expert workflows tailored to requirements.
- **Proven Experience:** Millions of structured patient interactions processed and delivered successfully.
- **HIPAA-compliant**, secure, AI-Ready Datasets: De-identified, export-ready, and scalable for AI pipelines, research, and licensing.

iCliniq is more than an annotation service provider; we are a strategic partner in clinical data transformation. By turning unstructured healthcare data into structured, standardized, and actionable intelligence, we empower organizations to make informed decisions, accelerate AI and research initiatives, and confidently meet regulatory and operational goals.