INDIUM

Advancing Cancer Immunotherapy Research with RAFT: A Domain-Specialized AI Approach for Oncology

A Whitepaper

Note to Readers

This whitepaper includes several medical and scientific terms related to oncology and cancer immunotherapy. For clarity and ease of understanding, below is a glossary of key terms and acronyms used throughout the document:

Glossary of Terms & Acronyms

Term / Acronym	Full Form / Meaning
RAFT	Rapid Annotation and Fine-Tuning / Research-Augmented Fine-Tuning – A domain-specialized AI model for oncology research
CAR-T cells	Chimeric Antigen Receptor T cells – A type of personalized immunotherapy using modified T cells to attack cancer
Checkpoint Inhibitors	Drugs that block proteins (like PD-1 or CTLA-4) used by cancer cells to evade immune attack
PD-1 / PD-L1	Programmed Cell Death Protein 1 / Programmed Death-Ligand 1 – Immune checkpoint proteins targeted by immunotherapy
TME	Tumor Microenvironment – The surrounding environment of a tumor, including immune cells, blood vessels, and signaling molecules
LLM	Large Language Model – An AI model trained on vast text corpora to understand and generate human language
RNA-seq	RNA Sequencing – A technique for analyzing the expression of genes in a biological sample

Term / Acronym	Full Form / Meaning
TCGA	The Cancer Genome Atlas – A comprehensive database of genetic information from various cancer types
ICGC	International Cancer Genome Consortium – A global initiative to study the genomic changes in cancer
KEGG	Kyoto Encyclopedia of Genes and Genomes – A database for understanding high-level biological functions and pathways
Reactome	A curated database of biological pathways and processes
SNOMED CT	Systematized Nomenclature of Medicine – Clinical Terms – A comprehensive clinical healthcare terminology
UMLS	Unified Medical Language System – A set of biomedical and health-related vocabularies
HIPAA	Health Insurance Portability and Accountability Act – U.S. legislation for protecting patient data privacy
GDPR	General Data Protection Regulation – European regulation on data protection and privacy
EHR	Electronic Health Record – A digital version of a patient's medical history maintained by healthcare providers
Biomarker	A measurable indicator of a biological condition, often used for diagnosis or monitoring of diseases like cancer

Executive Summary

Cancer immunotherapy has emerged as one of the most promising pillars of modern oncology, offering targeted treatment options where traditional therapies often fall short. Yet, researchers face significant hurdles: vast and unstructured biological data, inconsistent patient responses, and the difficulty of identifying effective biomarkers in time.

This is where **RAFT (Research-Augmented Fine-Tuning)** steps in – a domain-specialized AI framework built specifically for cancer research. Unlike generic AI models, RAFT is trained on oncology-specific literature, genomic databases, and real-world clinical data, enabling it to deliver context-aware insights and accelerate discovery.

At **Indium,** we believe that responsible, high-impact AI requires both technological excellence and deep domain knowledge. Our GenAI solutions, like RAFT, are built to empower healthcare teams with faster insights, personalized treatment pathways, and regulatory-ready decision support – all while ensuring privacy, bias mitigation, and scientific rigor

Introduction

Over the past decade, immunotherapy has shifted from a promising concept to a mainstream pillar of cancer treatment, offering patients new options when traditional approaches like chemotherapy or radiation fall short. From checkpoint inhibitors to CAR-T cells (Chimeric Antigen Receptor T cells), the immunotherapy arsenal continues to grow, offering new hope in tackling some of the toughest cancers. Despite its success stories, immunotherapy is far from a one-size-fits-all solution, and that's where the real challenge lies. Tumors can evolve, adapt, and resist even the most advanced treatments; what works for one patient might fail for another.

Did you know?

> According to Global Growth Insights, over **40%** of current oncology clinical trials globally involve combination therapies with immunotherapy agents, particularly checkpoint inhibitors combined with chemo, targeted therapy, or radiation.

This uncertainty pushes researchers to dig deeper into the tumor microenvironment, study genetic and molecular signals, and decode how individual immune systems respond. The result is an explosion of diverse data, from genomics and proteomics to patient health records and real-world evidence, all of which hold vital clues for improving therapies. But turning this scattered information into practical, testable insights remains a massive hurdle for oncology teams.

Adding to this complexity is the need to test new ideas quickly and cost-effectively, while ensuring patient safety and meeting strict regulatory standards. Researchers and biopharma companies are pressured to find better biomarkers, predict patient responses earlier, and design more precise clinical trials. This demands not just any AI, but intelligent systems that understand the unique language of oncology and can bridge the gap between raw data and real-world impact.

This white paper explores how RAFT revolutionizes cancer immunotherapy research by delivering accurate, context-aware, and biologically informed insights that enable faster discoveries, more reliable hypothesis testing, and highly personalized treatment strategies.

Why Cancer Immunotherapy Needs Specialized AI

Cancer immunotherapy represents a complex confluence of molecular biology, genomics, immunology, and clinical science. The immune system's interaction with tumors is dynamic and varies widely between patients, tumor types, and even regions of the same tumor.

Key challenges that hinder progress include:

Tumor Heterogeneity: Tumors are not monolithic; they consist of various cell populations with differing genetic profiles. This variability complicates the identification of universal biomarkers and treatment responses.

Dynamic Immune Landscapes: The tumor microenvironment (TME) evolves continuously, influenced by cytokines, T cells, dendritic cells, and checkpoint proteins like PD-1/PD-L1 **(Programmed Cell Death Protein 1/ Programmed Death-Ligand 1).** This dynamic nature makes capturing snapshots of immune response through traditional models difficult.

Data Complexity and Silos: From genomics and proteomics to patient health records and imaging data, researchers must analyze large volumes of multimodal data. Often, these datasets are siloed, poorly annotated, or biased toward specific populations.

Long Research Timelines: It can take years from hypothesis to discovery to clinical application. Researchers need AI tools that accelerate this timeline and do so with a deep understanding of the oncology domain.

Large Language Models like GPT or BERT have showcased remarkable capabilities across domains. However, in a specialized and high-stakes field like oncology, general-purpose models face severe limitations:

Lack of Domain Understanding: Most LLMs are trained on generic web-based corpora that include social media, books, news, and Wikipedia, resources lacking the depth needed to understand scientific literature and oncological jargon.

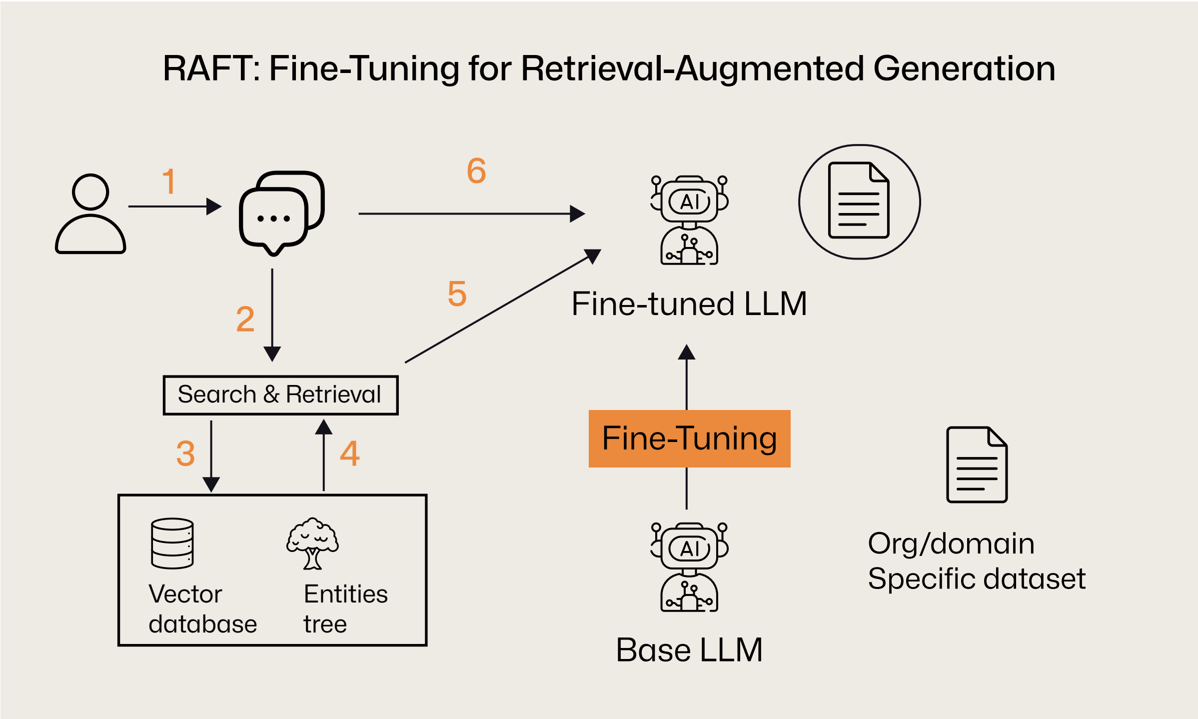
Inaccurate or Hallucinated Outputs: AI hallucination, a phenomenon in which models generate incorrect or fabricated information, can have dire consequences in research, particularly in the life sciences.

Inability to Understand Biological Context: In cancer immunotherapy, understanding context, such as signaling pathways, drug interactions, or immune cell behavior, is critical. Generic models cannot effectively correlate these complexities.

No Training on Experimental Data: Generic models rarely have access to proprietary datasets like patient outcomes, experimental protocols, or gene expression profiles, which are vital for immunotherapy research.

These gaps highlight the urgent need for domain-specific solutions like RAFT, which are built from the ground up with the oncology research ecosystem in mind.

Introducing RAFT: Domain-Specialized AI for Cancer Research



RAFT (Research-Augmented Fine-Tuning) is a transformer-based AI model developed explicitly for oncology applications. It integrates domain-relevant training data, medical ontologies, and research-specific fine-tuning protocols to assist researchers in:

Literature synthesis	Hypothesis generation	Experimental design support	Clinical trial outcome analysis	Biomarker discovery

RAFT's unique selling proposition lies in its training methodology and dataset curation.

Did you know?

> AFT reduces literature review time by up to 85%, turning 6-week research cycles into 5-day insight generation sprints.

Curated Knowledge Base

RAFT is trained on:

- Thousands of peer-reviewed oncology papers from PubMed, Nature, and JCO
- Real-world de-identified clinical trial data from immunotherapy trials
- Genomic databases such as TCGA and ICGC
- Molecular pathway databases (e.g., KEGG, Reactome)
- Structured medical ontologies like SNOMED CT and UMLS

Fine-Tuning with Human-in-the-Loop

RAFT incorporates human-in-the-loop (HITL) feedback loops from oncologists and biomedical researchers to reduce hallucinations and improve interpretability.

Multimodal Support

RAFT can process text and images (like histopathology slides), genomics matrices, and structured tables, enabling holistic insights from cross-domain inputs.

Current Challenges Vs Future Possibilities with RAFT

Current Challenges in Immunotherapy Research	How RAFT Transforms the Landscape
Siloed, unstructured data across EHR, genomics, and literature	Harmonizes multimodal data for unified analysis (text, genomics, images)
Long research cycles (6+ weeks for literature review, 12 months for validation)	Literature review in 5 days; biomarker validation reduced by **66%**
Limited interpretability of AI outputs	Built-in attention mapping and layerwise explainability
Generic LLMs hallucinate or miss biological context	Domain-trained with oncology datasets and ontologies (TCGA, KEGG, UMLS)
Inconsistent biomarker discovery across patient groups	**31%** higher accuracy in biomarker prediction; bias mitigation for diverse populations
Slow trial adaptation and fixed protocols	Real-time clinical trial monitoring and adaptive trial design support
Difficulty generating regulatory reports and evidence	Auto-compiled insights and report-ready knowledge maps for FDA/EMA submission

What Sets RAFT Apart from Other Biomedical AI Models

While several industry leaders – such as Tempus, Foundation Medicine, and NVIDIA Clara – have made strides in applying AI to oncology, RAFT stands apart in its narrow focus on cancer immunotherapy research, its multimodal fusion of clinical, genomic, and textual data, and its human-in-the-loop fine-tuning architecture. Unlike closed diagnostic platforms or general-purpose drug discovery models, RAFT was built from the ground up to serve oncology researchers with interpretability, flexibility, and real-time insight generation, making it uniquely suited to the complexity and pace of immunotherapy innovation.

Understanding Other Fine-Tuning Methods in AI for Oncology

To make AI truly useful for cancer research, general models must be adapted to oncology's unique language, data, and questions. This is where fine-tuning comes in: customizing pre-trained AI models for specific tasks or fields.

Common fine-tuning methods include:

Domain Pretraining:

Training a general model further on extensive collections of oncology-specific literature to build deeper biomedical context.

Supervised Fine-Tuning:

Teaching the model with labeled datasets, such as patient data linked to treatment outcomes.

Instruction Tuning:

Training AI to follow specialized prompts, such as those to summarize trial results or flag anomalies.

Reinforcement Learning from Human Feedback (RLHF):

Using expert feedback to help the AI learn which outputs are most accurate.

Federated Fine-Tuning:

Improving models across multiple hospitals or labs without moving sensitive patient data.

These approaches ensure AI tools are not generic black boxes, but trustworthy partners for complex research. RAFT builds on these methods with a domain-specialized pipeline explicitly designed to meet the demands of cancer immunotherapy.

RAFT Architecture and Technical Deep Dive

RAFT is designed on a transformer-based architecture, leveraging recent advances in natural language processing and multi-modal AI to specialize in oncology data.

1

Core Model Architecture

RAFT utilizes a transformer backbone similar to BERT or GPT but augmented with domain-specific modules. Key architectural features include:

- **Domain Embedding Layer:** Integrates specialized embeddings for biomedical terminology and gene/protein symbols.
- **Multi-Head Attention:** This allows RAFT to capture complex relationships between biological entities, such as the interaction between checkpoint inhibitors and immune cells.
- **Cross-Modal Fusion Layer:** Think of this as RAFT's ability to "see the full picture." It connects data from multiple formats – like scientific.
- **Interpretability Modules:** RAFT doesn't just give answers – it shows its work. With features like attention maps and explanation layers, researchers can understand why the AI made a recommendation, ensuring transparency and trust.

2

Training Pipeline

- **Pretraining:** RAFT undergoes pretraining on vast biomedical corpora, emphasizing language patterns and scientific syntax unique to oncology.
- **Fine-Tuning:** The model is then fine-tuned on immunotherapy-specific datasets, including clinical trial reports and experimental data.
- **Human Feedback Integration:** Oncologists annotate output for correctness and relevance, and these annotations are fed back to refine the model iteratively.

3

Data Handling and Privacy

To ensure compliance with healthcare data regulations (HIPAA, GDPR), RAFT incorporates strict de-identification protocols and federated learning capabilities, allowing training on sensitive data without compromising privacy.

How RAFT Transforms Cancer Immunotherapy Research

RAFT's domain specialization leads to transformative capabilities across multiple facets of immunotherapy research:

	Faster Literature Review: RAFT quickly reads and summarizes thousands of research papers, saving scientists months of manual work.
	New Research Ideas: It finds links in data to suggest new ideas, like potential drug targets or reasons why some treatments stop working.
	Better Experiment Design: RAFT recommends setting up lab tests, such as which doses, cell lines, or biomarkers to use, to get better results.
	Smarter Clinical Trials: It analyzes trial data as it arrives, spotting early signs of success, side effects, or which patients respond best.
	Support for Personalized Care: RAFT matches patient DNA and immune data with the latest therapies, helping doctors choose the best treatment.
	Live Data Insights: It can work with real-time data from lab tests or patient devices to monitor how treatments are working.

Easy Collaboration: RAFT makes it easier for hospitals, labs, and partners to share insights without sharing raw patient data.

Automatic Data Tagging: It auto-labels huge datasets, highlights essential biomarkers, and spots errors, speeding up data analysis.

Finding New Uses for Drugs: RAFT can uncover old drugs that might work for new cancer types, especially rare or hard-to-treat cancers.

Better Patient Grouping: It helps researchers group patients more precisely, making trials more targeted and effective.

Faster Regulatory Reports: RAFT helps teams quickly gather evidence and results to support approvals and submissions to bodies like the FDA (Food and Drug Administration) or EMA (European Medicines Agency).

Clear Knowledge Maps: It builds easy-to-read maps that show how genes, drugs, and outcomes are connected, helping teams see the bigger picture.

The Industry's Path to Better Cancer Immunotherapy: Powered by RAFT

The industry is embracing RAFT as a transformative solution to enhance cancer immunotherapy research.

1

Dataset Creation and Diversification

Organizations are building robust datasets by:

- Sourcing standardized, high-quality clinical protocols from platforms like PubMed, MedLine, and ClinicalTrials.
- Selecting diverse cases across cancer types, treatment approaches, and clinical trial phases.
- Structuring datasets to mimic retrieval-augmented generation (RAG) workflows, facilitating the AI's ability to generalize knowledge across the cancer immunotherapy space, not just memorize protocol specifics.

2 Fine-Tuning and Model Customization

- RAFT models are fine-tuned to understand medical language, terminologies, and reasoning structures unique to oncology, ensuring the model's output is clinically meaningful.
- The industry ensures the RAFT process introduces realistic information retrieval scenarios similar to those faced in clinical or research settings.
- Meeting regulatory and collaborative demands in biomedical research.

3 Rapid, Privacy-Preserving Insights

- RAFT enables organizations to keep sensitive patient data within secure, internal environments while applying advanced AI methods.
- Rapid, accurate answer generation is possible for highly specialized clinical questions, helping drive decision-making in precision immunotherapy.

Industry Benefits and Applications

Accelerated Research: This approach shortens the time from data acquisition to actionable insight by streamlining protocol analysis and experimental planning.

Improved Reproducibility: Fosters reproducible research outcomes, which are essential for clinical translation and regulatory approvals.

Tailored Implementation: Organizations can adapt the RAFT framework for various immuno-oncology tasks, from biomarker discovery to treatment planning.

Integration into Research Pipelines and Hospitals

RAFT is designed for seamless integration into existing workflows:

Research Labs

RAFT provides APIs and user interfaces for literature mining, hypothesis testing, and data analytics compatible with standard bioinformatics tools (e.g., Bioconductor, cBioPortal).

Clinical Settings

RAFT integrates with electronic health records (EHR) systems and clinical decision support tools, delivering personalized immunotherapy recommendations at the point of care.

Cloud and On-Premise Deployment

Flexible deployment options enable institutions with strict data sovereignty requirements to run RAFT locally or leverage cloud-based scalability for compute-intensive tasks.

A Look into Real-World Case Studies

RAFT builds on proven, real-world approaches similar to those of leaders like Tempus, Foundation Medicine, and NVIDIA Clara Discovery. These companies show how domain-focused AI is already transforming how researchers harmonize data, find new biomarkers, and design better immunotherapy strategies, which is precisely what RAFT takes further with its specialized, fine-tuned framework for oncology.

Case Study 1: Biomarker Discovery in Melanoma

Using RAFT, researchers at a leading cancer center identified a novel gene signature predictive of anti-PD-1 therapy response. RAFT's multimodal analysis combined RNA-seq data and clinical records, reducing validation time from 12 months to 4 months.

Case Study 2: Optimizing CAR-T Cell Therapy

RAFT helped design a new experimental protocol for CAR-T (Chimeric Antigen Receptor T-cell) therapy targeting solid tumors by suggesting modifications in antigen-targeting strategies based on immune escape mechanisms identified in the literature and patient data.

Case Study 3: Adaptive Clinical Trial Monitoring

A pharmaceutical company integrated RAFT into its clinical trial analytics pipeline. RAFT provided real-time insights into patient stratification and adverse event trends, allowing dynamic trial modifications that improved patient safety and efficacy endpoints.

Ethical Considerations and Bias Mitigation in RAFT

AI in oncology research must adhere to strict ethical standards:

Bias Detection and Correction

RAFT incorporates bias detection algorithms to identify skewed data distributions that may affect minority groups or rare cancer subtypes, ensuring equitable model performance.

Transparency and Explainability

Through attention mapping and output rationales, RAFT supports explainability, allowing researchers and clinicians to trust its recommendations.

Patient Privacy and Data Security

RAFT employs advanced encryption, anonymization, and federated learning to safeguard patient data during model training and deployment.

Continuous Ethical Oversight

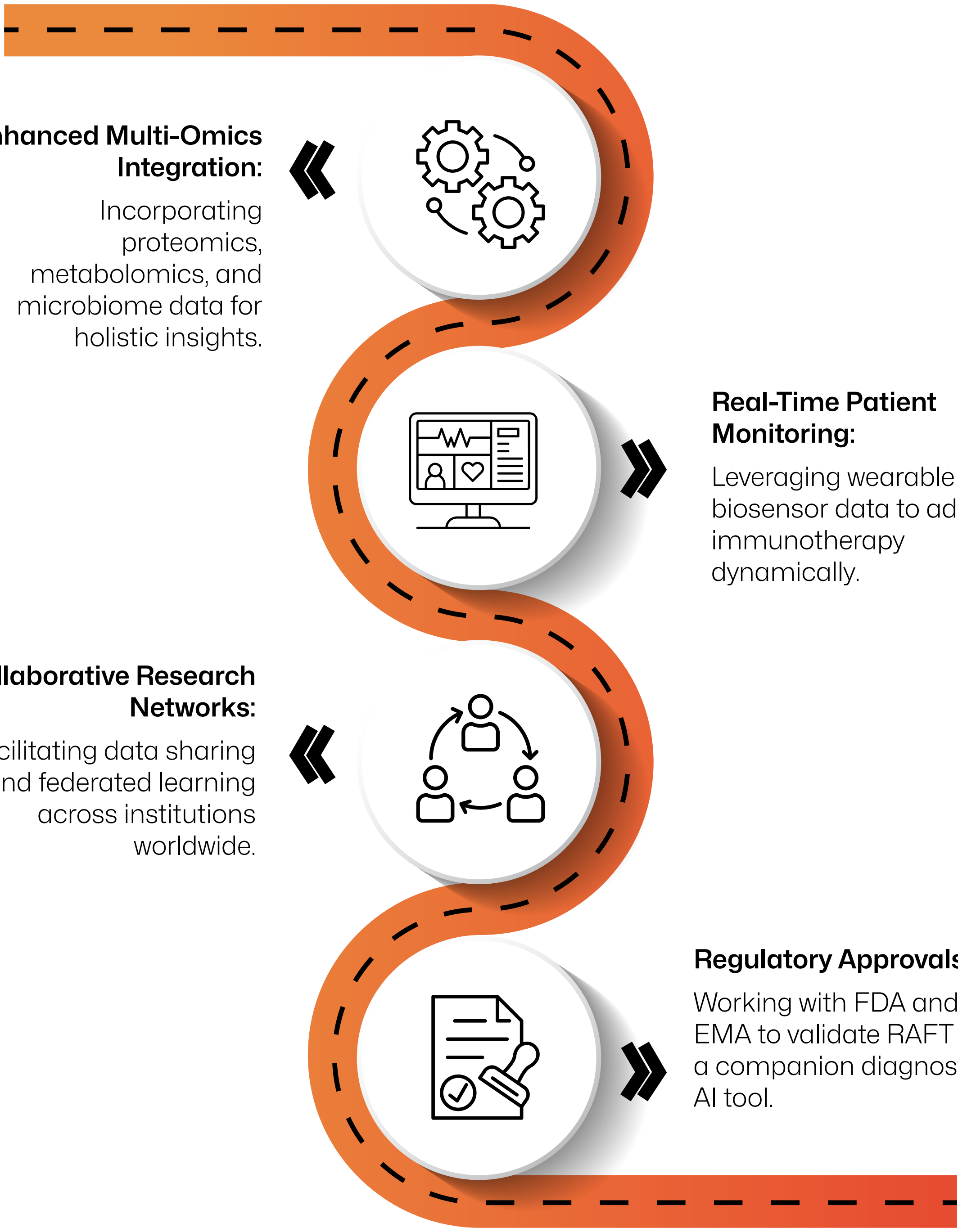
A dedicated ethics board regularly reviews RAFT's outputs and applications to monitor compliance with evolving healthcare and AI ethics guidelines.

Did you know?

> A 2025 report by the AI in Medicine Ethics Board stated that AI models trained on biomedical corpora failed bias audits for underrepresented cancer populations, particularly in Black and Asian cohorts, an issue RAFT explicitly addresses.

What's Next For RAFT?

The roadmap for RAFT includes:



Enhanced Multi-Omics Integration:
Incorporating proteomics, metabolomics, and microbiome data for holistic insights.

Real-Time Patient Monitoring:
Leveraging wearable and biosensor data to adapt immunotherapy dynamically.

Collaborative Research Networks:
Facilitating data sharing and federated learning across institutions worldwide.

Regulatory Approvals:
Working with FDA and EMA to validate RAFT as a companion diagnostic AI tool.

Indium's Commitment to Smarter Healthcare with AI

At Indium, we understand that innovation in healthcare, especially in oncology, requires more than just technology; it demands deep domain expertise, proven frameworks, and an agile mindset. Our Gen AI solutions for the healthcare industry are designed with this very principle. Whether it's enabling AI-assisted research in cancer immunotherapy, accelerating drug discovery with large language models, or transforming clinical decision-making with intelligent analytics, we offer tailor-made solutions built for real-world impact. Our healthcare analytics, predictive modeling, and AI-driven diagnostics work reflects our commitment to precision, performance, and purpose.

What truly sets us apart is our team. Backed by data scientists, healthcare SMEs, and MLOps engineers, Indium ensures that your Gen AI initiatives are innovative but also responsible, scalable, and regulation-ready. We don't just consult, we collaborate, co-create, and deliver. From proof of concept to production-grade deployment, our end-to-end AI services empower healthcare organizations to unlock value at every stage of the patient journey.

With a portfolio of success stories and a passion for impact, Indium is your trusted partner in shaping the future of AI in healthcare.

Discover how Indium is leading in Gen AI – Click Here.

Success Story

Transforming Appeals Management with Gen AI for a Leading U.S. Health Benefits Provider

Success Story Overview

A leading U.S. health benefits provider partnered with Indium to transform its manual, error-prone appeals management process using a GenAI-powered automation platform.

Faced with regulatory complexity and operational inefficiencies, the client needed to accelerate appeals processing, ensure consistent decision-making, automate document generation and translation, and maintain stringent data privacy.

Indium's solution introduced an AI-driven case review engine, automated approval/denial letter generation, integrated multilingual translation, and seamless workflow integration - all built on a secure, scalable infrastructure.

Driving Measurable Outcomes in Appeals Management

2X Faster Appeal Resolution: Drastically reduced decision-making time, accelerating the appeals lifecycle and improving member turnaround time.

90% Accuracy in AI-Supported Decisions: Consistent and explainable decisions across cases, reduced variability, and ensured compliance.

40% Increase in Nurse Productivity: Automation allowed clinical teams to focus on higher-value tasks, improving operational efficiency.

50% Workflow Streamlining: Enhanced team collaboration and communication, reduced bottlenecks, and improved end-to-end process performance.

See the bigger picture of this GenAI-powered automation platform - Click Here!

Conclusion

RAFT represents more than just a technological advancement; it signals a paradigm shift in how the oncology community approaches cancer immunotherapy research. By merging domain-specialized AI with deep biological context and rigorous scientific validation, RAFT empowers researchers and clinicians to uncover novel insights that were previously obscured by complexity and data silos. With its context-aware reasoning and explainable outputs, RAFT accelerates discovery, optimizes treatment pathways, and enhances confidence in AI-driven decisions, a critical factor in life-or-death scenarios like cancer care.

As precision oncology continues to evolve, RAFT sets a new benchmark for how artificial intelligence can be responsibly and powerfully applied in complex, high-stakes biomedical domains. It paves the way for faster breakthroughs, targeted therapies, and ultimately, more hopeful patient and family outcomes. By bridging the gap between cutting-edge computational power and the nuanced realities of cancer biology, RAFT stands poised to transform the future of cancer immunotherapy, turning today's possibilities into tomorrow's standard of care.

References

https://stayrelevant.globant.com/en/technology/data-ai/medical-technology-smarter-ai-raft-cancer-immunotherapy-research/

RAFT-based Application for Breakthroughs in Cancer Immunotherapy

https://pubmed.ncbi.nlm.nih.gov/34900525

https://pubmed.ncbi.nlm.nih.gov/38597966

https://www.nature.com/articles/s41568-024-00694-7

https://www.frontiersin.org/articles/10.3389/fimmu.2022.1076883/full

https://www.sciencedirect.com/science/article/abs/pii/S1044579X22002309

https://www.mdpi.com/2079-9721/13/1/24

https://www.nature.com/articles/s41416-021-01633-1

https://www.aacr.org/blog/2025/04/28/improving-cancer-care-with-artificial-intelligence

https://www.cancerresearch.org/blog/january-2025/ai-cancer

https://www.ft.com/content/0a8f2c61-77f4-43ce-87d2-a7b421bbda85

https://www.theguardian.com/research-to-reality/2025/jan/24/prostate-cancer-and-ai-the-exciting-advances-that-could-transform-treatment

About Indium

Indium is an AI-driven digital engineering company that helps enterprises build, scale, and innovate with cutting-edge technology. We specialize in custom solutions, ensuring every engagement is tailored to business needs with a relentless customer-first approach. Our expertise spans Generative AI, Product Engineering, Intelligent Automation, Data & AI, Quality Engineering, and Gaming, delivering high-impact solutions that drive real business impact.

With 5,000+ associates globally, we partner with Fortune 500, Global 2000, and leading technology firms across Financial Services, Healthcare, Manufacturing, Retail, and Technology—driving impact in North America, India, the UK, Singapore, Australia, and Japan to keep businesses ahead in an AI-first world.

USA	INDIA	UK	SINGAPORE
Cupertino | Princeton Toll-free: +1-888-207-5969	Chennai | Bengaluru | Mumbai Hyderabad | Pune Toll-free: 1800-123-1191	London Ph: +44 1420 300014	Singapore Ph: +65 6812 7888

www.indium.tech

For Sales Inquiries
sales@indium.tech

For General Inquiries
info@indium.tech