

RESPONSIBLE AI (RAI) IN LIFE SCIENCES



Responsible AI (RAI) in life sciences is crucial as AI-driven decisions directly impact patient health, drug development, and medical decisions, where errors can have severe consequences. Ensuring AI is fair, transparent, and reliable is essential for ethical and effective use of AI in the life sciences (healthcare, biotechnology, and pharmaceutical) industry. This POV provides an overview on the need for RAI in the life sciences AI models, and the Infosys RAI framework that helps in identifying the vulnerabilities of AI models in life sciences and applying RAI guardrails to the AI solutions.



AI in Life Sciences

The pharmaceutical industry is progressing in every facet, with the application of artificial intelligence (AI), from drug discovery and clinical trials to supply chain optimization. According to a study, integration of AI in drug design has resulted in faster drug discovery, cost savings, reduced resource and manpower usage, and decreased attrition rates in clinical trials.⁽¹⁾ AI usage in life sciences has risen rapidly in recent years. According to the latest surveys, 80% of pharmaceutical and life sciences professionals use AI for drug discovery.⁽²⁾ According to a report, the usage of AI in Pharma market size will grow to \$5.62 billion in 2028 at a compound annual growth rate (CAGR) of 28.5%.⁽³⁾

Examples of AI Applicability

- **Drug Discovery:** AI is applicable in drug discovery, across drug target identification, accelerated drug design, prediction of 3D structures of target proteins, and prediction of drug-protein interactions.
- **Clinical Trials:** AI can improve patient recruitment and retention, increase the medication adherence rate of patients by 25%⁽⁴⁾, and identify suitable candidates for clinical trials.
- **Sales and Marketing:** AI can be leveraged for augmented sales representative, as well as for content generation and review.
- **Personalized Care:** AI can support doctors in identifying potential patient diagnoses based on their genetics and symptoms and predict how genetic variants impact human cell behavior at the protein level.

The Problem Statement

The potential of AI models used in life sciences is tremendous, however they also bring significant challenges and limitations.

- **Data Privacy and Security:** Handling sensitive health data can lead to breaches and misuse of personal information. For example, a study demonstrated an algorithm capable of re-identifying over 85% of adults in a cohort study, despite data aggregation and the removal of protected health information.⁽⁵⁾

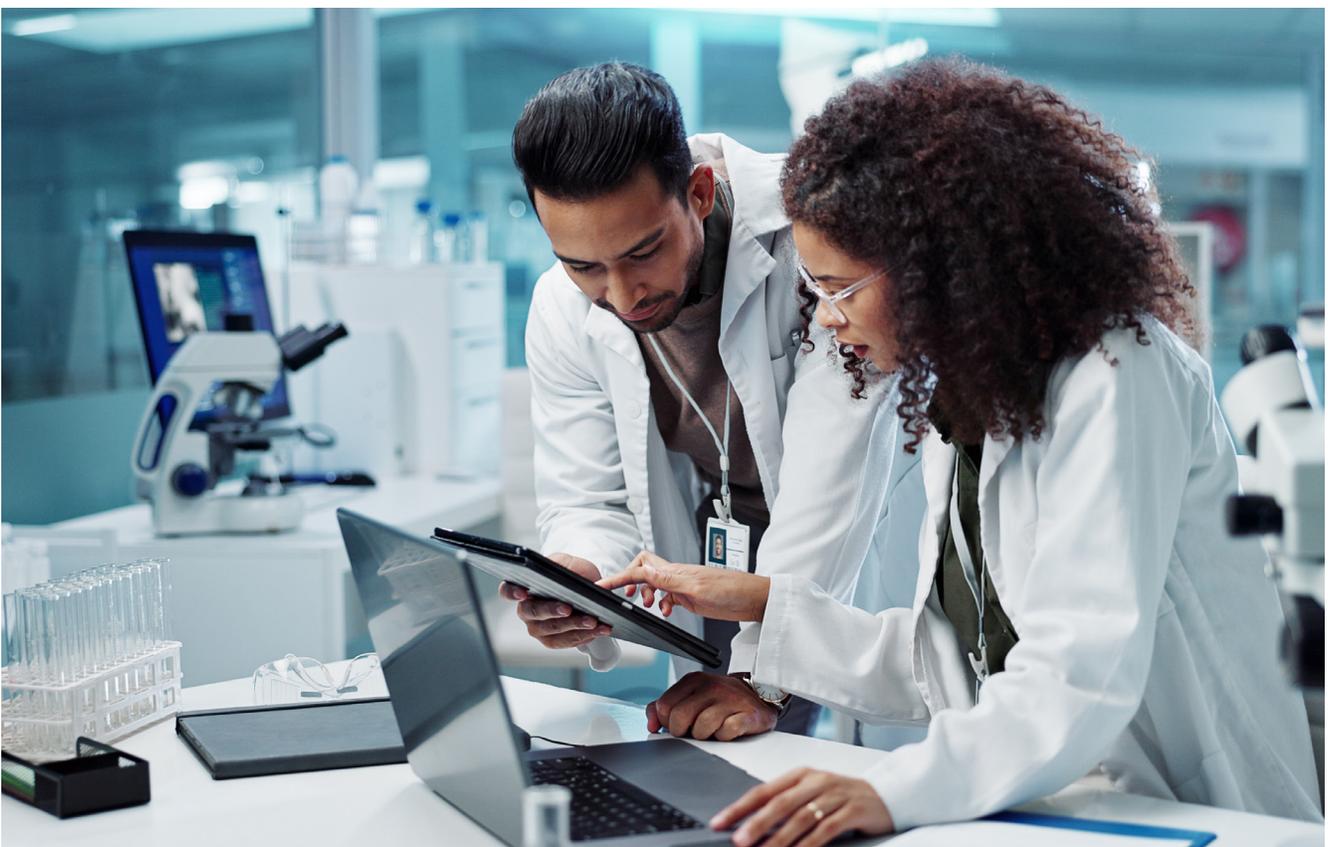
- **AI Algorithm Bias:** AI models may inherit biases from training data, leading to unfair or inaccurate outcomes. For instance, a review found that none of the AI models used to detect and prognosticate COVID-19 using chest radiographs and CT scans were of potential clinical use due to methodological flaws and/or underlying biases. ⁽⁶⁾
- **Lack of Transparency:** Many AI models, especially deep learning ones, are black boxes, making it hard to understand their decision-making process. The challenge of transparency is exemplified by applications like IBM Watson for Oncology, where understanding the basis for diagnosis and treatment recommendations can be difficult. ⁽⁷⁾
- **Overreliance on AI:** Overreliance on AI can lead to reduced human oversight, potentially missing errors, or nuances that a human expert might catch. Introducing a human-in-the-loop

approach, such as with AI-assisted detection, can help mitigate this risk. ⁽⁸⁾

- **Accuracy and Reliability:** If not properly validated, AI models can produce inaccurate results, leading to misdiagnoses or inappropriate treatments.
- **Consideration of Ethics:** The use of AI in life sciences raises ethical issues, such as the potential for AI to replace human jobs or make decisions about life and death.

EU AI Act and RAI

AI systems used in parts of the life sciences value chain, like drug discovery, clinical trials, and medical devices, are classified as high-risk if they affect human health, safety, and fundamental rights. EU AI Act enforces RAI principles through strict regulations particularly for high-risk AI applications in the sectors like life sciences, finance, and law enforcement. ⁽⁹⁾



Responsible AI to Overcome the Challenges in Using AI

Responsible AI defines an approach to developing, deploying, and using AI systems that align with ethical principles and societal values.

Key Considerations of RAI in Life Sciences

- **Patient Health and Safety:** Life Sciences AI models are often used in high-stakes areas such as disease diagnosis and drug discovery, where errors could lead to misdiagnosis or inappropriate treatment. Hence it is essential to ensure that AI models are validated and safe for clinical and research purposes.
- **Ethical Considerations:** Enhancing the AI models with ethical principles would ensure patient rights, privacy, and dignity by adhering to fair practices.
- **Trust and Adoption:** Having clear explainable AI models ensures that the public perceives AI as trustworthy and increases acceptance in the healthcare ecosystems.
- **Regulatory Compliance:** Life sciences AI models complying with the strict regulations like EU AI Act, FDA/EMA, GDPR, and HIPAA ensure legal and ethical adherence.
- **Avoiding Unintended Consequences:** Having humans in the loop as part of the AI model would prevent blind reliance on technology and hence avoid unintended consequences.

To overcome different challenges from the AI models, the below responsible AI principles could be used:

	Fairness and Bias	Ensuring AI systems do not discriminate against individuals or groups based on race, gender, age, or other characteristics. This involves addressing biases in data and algorithms.
	Privacy	Protecting an individual's data and ensuring it is used responsibly. This involves adhering to data protection laws and implementing robust data security measures.
	Security	Protecting AI systems from potential security threats, such as adversarial attacks or unauthorized access. This includes safeguarding against malicious parties who may attempt to manipulate or compromise AI models.
	Safety	Ensuring AI systems operate safely and do not cause harm. This includes rigorous testing and monitoring to prevent malfunctions or misuse.
	Explainability	Making AI systems and their decision-making processes understandable and explainable to users and stakeholders. This includes clear communication about how AI works and what data it uses.

Example: An AI Model for Oncology Diagnosis and Treatment

Purpose: The AI model assists Oncologists by diagnosing cancer and recommending treatment plans. The model utilizes medical imaging, pathology reports, genetic reports, and clinical history. It uses deep learning, natural language processing (NLP), and reinforcement learning for the recommendation.

Responsible AI Framework and Implementation Strategy

There are various applicable RAI principles to overcome challenges in AI implementations and enterprises must follow a systematic approach towards RAI implementation.

 AI Vulnerability	 Applicable RAI Principle	 Mitigation Strategy
Bias in training data (such as underrepresentation of minority populations)	Fairness and Bias	<ul style="list-style-type: none"> • Diverse and inclusive datasets • Fairness-aware ML algorithms
Explainability limitation (such as black box decisions in treatment recommendations)	Transparency and Explainability	<ul style="list-style-type: none"> • Interpretable AI models
Inaccurate prediction (such as false positives/false negatives in diagnosis)	Safety and Reliability	<ul style="list-style-type: none"> • Rigorous validation on multiple data sets
Privacy risks (such as leakage of patient health records)	Privacy and Data Governance	<ul style="list-style-type: none"> • Strong encryption and access control
Regulatory compliance gaps (such as not complying with regulations like, GxP, GDPR, EU AI Act, and guidelines from FDA and EMA)	Compliance and Accountability	<ul style="list-style-type: none"> • Adherence to healthcare regulations



Three Main Components of Responsible AI3S Framework

Scan

Continuous AI Model Risk Assessment: From project inception, continuous audits and risk assessments are conducted to identify potential AI model vulnerabilities. This process includes a detailed questionnaire covering the AI's impact on individuals, the specific model used, input and output data, and effects on business functions. The resulting risk assessment identifies associated risks and vulnerabilities.

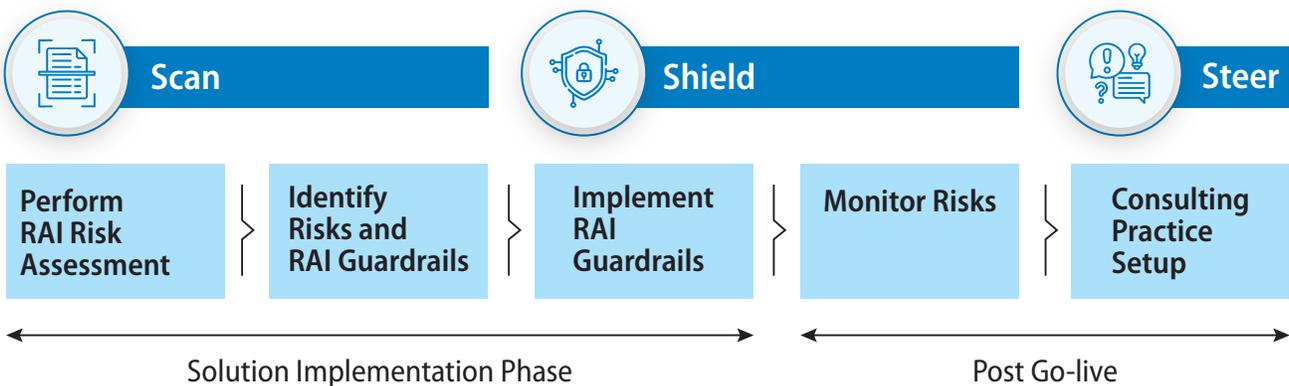
- **Regulation Readiness Prep:** Reviewing contracts for AI-related clauses to safeguard against risks

Shield

1. **Infosys Responsible AI Gateway:** Embedding automated RAI workflows in internal processes and systems
2. **Infosys GenAI Guardrails:** Identifying the applicable RAI guardrails (Fairness and Bias, Privacy, Security, Safety, Explainability) to protect against the vulnerabilities of LLMs and foundation models.
3. **Infosys Responsible AI Toolkit:** Providing a collection of automated RAI pipelines and APIs that can be embedded into MLOPS pipelines.
4. **AI Security:** Simulating and identifying attacks on AI models to strengthen our AI models and build counter defense models

Steer

1. **RAI Advisory Services:** Providing oversight for an RAI program and formulating overall strategy.
2. **RAI Practice Setup:** Developing end-to-end RAI practices across the organization, with a mix of frameworks, advisory services, tailored best practices, playbooks, and technical guardrails to institute strong AI governance
3. **AI Crisis Management:** Engaging in swift remediation via crisis management protocols or instituting disaster recovery mechanisms in the event of unforeseen AI-related incidents.





RACI Matrix for RAI Implementation

Responsible	Infosys RAI Scan Team	Infosys RAI Scan Team Infosys RAI Shield Team	Infosys RAI Shield Team	Support and Maintenance Team	Infosys RAI Steer Team
Accountable	Project Delivery Team	Project Delivery Team	Project Delivery Team	Application Owner	Infosys RAI Steer Team
Consulted**	Legal Regulatory Compliance Data Privacy Office Domain SME Cybersecurity	Legal Regulatory Compliance Data Privacy Office Domain SME Cybersecurity			Legal Regulatory Compliance Data Privacy Office Domain SME Cybersecurity
Informed	Project Sponsor PMO	Project Sponsor PMO	Project Sponsor PMO	Project Sponsor	Project Sponsor

Conclusion

In summary, the integration of RAI in life sciences is essential to ensure ethical, transparent, and safe application of AI. RAI is not just a necessity but an imperative for AI-driven advancements in life sciences. By implementing strong governance, transparency, and ethical AI principles, the

industry can harness AI's full potential. The Infosys Responsible AI Toolkit, as part of the Infosys Responsible AI Framework, is one such solution that enterprises can utilize to ensure responsible AI across different AI models in life sciences.



References

1. Role of Artificial Intelligence in Revolutionizing Drug Discovery - ScienceDirect
2. Recent Research Shows That Artificial Intelligence Has the Potential to Make the Pharma Sector More Effective
3. AI In Pharma Market Overview: Market Size, Major Drivers And Trends - Latest Global Market Insights
4. Artificial intelligence in drug discovery and development - PMC
5. Privacy and artificial intelligence: challenges for protecting health information in a new era
6. <https://www.nature.com/articles/s42256-021-00307-0>
7. Case Study 20: The \$4 Billion AI Failure of IBM Watson for Oncology - Henrico Dolfing
8. View of Review on AI-Driven Innovations in Stroke Care: Enhancing Diagnostic Accuracy, Treatment Efficacy, and Rehabilitation Outcomes
9. EU AI Act: first regulation on artificial intelligence | Topics | European Parliament

Author

Pooja Durgad



For more information, contact askus@infosys.com

© 2025 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.