

**BRIDGING DISCIPLINES
WITH ARTIFICIAL
INTELLIGENCE**

BRIDGING DISCIPLINES WITH ARTIFICIAL INTELLIGENCE

Edited by

Dr. Aanchal Arora

Dr. Khushdip Kaur

Dr. Isha Dudeja

VYOM HANS PUBLICATIONS

© 2026 Editors (**Dr. Aanchal Arora, Dr. Khushdip Kaur, & Dr. Isha Dudeja**)

Date of Publications: 24/06/2026

ISBN (E): 978-81-686211-1-4

ISBN (P): 978-81-686211-0-7

Typesetting: LaTeX

Published By-

Imprint: Vyom Hans Publications

Vyom Hans Vision

967/EH1, Industrial Area, DeraBassi,

SAS Nagar (Mohali), PIN-140507, Punjab, India.

E-Mail: sushil.kumar@vyomhansjournals.com

Website: <https://press.vyomhansjournals.com>

Mob: +91 9781711925

Disclaimer:

The authors, editors, and publisher of this volume do not provide medical or professional advice, nor do they endorse the application of any specific methods or techniques. The content is presented solely for academic, educational, and research purposes. Readers are encouraged to exercise scholarly judgment and consult appropriate experts where necessary. The authors, editors, and publisher shall not be held liable for any consequences arising from the use or interpretation of the material presented in this work.

This edited volume comprises contributions from multiple authors. Responsibility for the accuracy, originality, and integrity of each chapter rests solely with the respective contributors. The views and opinions expressed are those of the individual authors and do not necessarily reflect those of the editors or publisher.

Some contributors may have utilized artificial intelligence (AI) tools to support aspects of writing, editing, or data processing. Such use remains the responsibility of the respective authors, who have ensured appropriate verification, validation, and scholarly oversight of all content. The editors and publisher do not assume responsibility for the use of such tools but expect adherence to academic standards of integrity and transparency.

The publisher affirms that the content of this volume does not knowingly infringe upon any existing copyrights or intellectual property rights. If any source has been inadvertently omitted or improperly attributed, the publisher should be notified in writing so that appropriate corrective action may be taken.

Acknowledgement

The Editors sincerely express their gratitude to all contributors whose dedication and scholarly efforts made this publication possible. We deeply appreciate the authors for their valuable chapters, intellectual insights, timely submissions and cooperation throughout the grounding of this book. Their scholastic contributions have enriched the quality and diversity of this book.

We extend our heartfelt thanks to the academic experts for their insightful comments, constructive suggestions; professional guidance and continuous engagement with emerging technologies inspire meaningful academic discourse and innovation.

The Editors also acknowledge the continuous support, patience and assistance provided by the publishing team in bringing this work to completion with efficiency and professionalism.

Finally, we gratefully recognize the encouragement and support received from our institution, colleagues and family members throughout this endeavor.

Editors:

Dr. Aanchal Arora^a Dr. Khushdip Kaur^b Dr. Isha Dudeja^c

^aAssistant Professor, Department of Chemistry, Khalsa College for Women, Civil Lines, Ludhiana, arora.aanchal511@gmail.com

^bAssociate Professor, Department of Commerce, Khalsa College for Women, Civil Lines, Ludhiana, khushdip1170@gmail.com

^cAssistant Professor, Department of Chemistry, Khalsa College for Women, Civil Lines, Ludhiana, ishadudeja275@gmail.com

Foreword

It is with great pride and satisfaction that I present this edited book titled “Bridging Disciplines with Artificial Intelligence” to the readers from diverse disciplines, professional and intellectual backgrounds. This book has been envisioned with the belief that Artificial Intelligence (AI) is not merely a technological innovation but a life-changing force that is reshaping every human activity. The book reflects the rapidly expanding role of AI in transforming not only scientific and technological domains but also the fields of humanities, commerce and management. AI is no longer confined to computer science laboratories; rather it has become an interdisciplinary force integrating scientific innovation with human understanding, economic systems and managerial practices. The title of this book was deliberately chosen to emphasize the integrative and transformative nature of AI across multiple domains of knowledge. We are living in an era where the boundaries between disciplines are becoming increasingly interconnected. AI stands at the centre of this convergence, offering unprecedented opportunities while also presenting complex challenges that require multidisciplinary understanding.

The chapters included in this volume explore a broad spectrum of themes that demonstrate the multidimensional applications of AI. Several contributions examine the role of artificial intelligence in biological sciences, biotechnology, chemistry, physics, mathematics and environmental sustainability. These chapters highlight how AI-driven technologies are accelerating scientific discovery, improving diagnostics, enhancing computational modeling, and supporting sustainable solutions for global challenges. At the same time, the book also reflects the growing significance of AI in humanities and social sciences. The chapters discussing artificial intelligence in art and creativity, governance, public policy and societal transformation provide meaningful insights into the ethical, cultural and social implications of intelligent technologies. These discussions are especially relevant in an era where technological advancement must remain aligned with human values and responsible innovation. The volume further addresses the increasing integration of AI into commerce, banking, accounting, marketing, and human resource management. Artificial intelligence is reshaping organizational decision-making, consumer engagement, financial systems, cybersecurity and workplace dynamics.

This edited book aims to provide readers with a comprehensive understanding of the opportunities, challenges, ethical concerns, and future

prospects associated with artificial intelligence. As editor, I believe that the future of AI will not be determined solely by technological capability but by our collective wisdom in applying it responsibly, ethically and inclusively. The true promise of artificial intelligence lies not in replacing human intelligence but in augmenting it thus enhancing our ability to solve complex problems, generate new knowledge, foster innovation and improve the quality of life for individuals and communities around the world.

I extend my sincere gratitude to all contributors, reviewers, and collaborators whose dedication and expertise have made this work possible. Their commitment to interdisciplinary scholarship reflects the very spirit of this book. I also thank the readers who have chosen to engage with this important subject and contribute to the ongoing dialogue about the future of intelligent technologies. It is my hope that this volume will inspire thoughtful inquiry, encourage cross-disciplinary collaboration, and promote a deeper understanding of how artificial intelligence can serve humanity while advancing scientific, economic, cultural, and managerial development.

Editors:**Dr. Aanchal Arora^a**

Editor, Bridging Disciplines with Artificial
Intelligence
Khalsa College for Women Civil Lines,
Ludhiana.

Contents

Acknowledgement	i
Foreword	iii
1 AI-DRIVEN BIOINFORMATICS: TRANSFORMING DISEASE DISCOVERY, DIAGNOSIS AND THERAPY	1
1.1 Introduction	2
1.2 Biological and Clinical Data Foundations For AI-Driven Disease Research	3
1.3 Artificial Intelligence Methodologies In Disease Modeling	4
1.4 AI In Disease Discovery	5
1.5 AI In Diagnosis and Patient Stratification	7
1.6 AI Foundation Models and Multimodal Clinical Intelligence	8
1.7 AI In Drug Discovery and Therapeutic Development	10
1.8 Challenges and Limitations of AI In Biomedical Applications	11
1.9 Conclusion and Future Perspectives	12
References	13
2 AI IN BIOLOGICAL SCIENCES	15
2.1 Introduction	16
2.2 Evidence of The Usage of AI In Biological Sciences	16
2.3 Types of AI	18
2.4 Integration of Artificial Intelligence with Biosciences	18
2.5 Applications of AI	19
2.6 Medical Science and Biotechnology Advancements	20
2.7 Challenges and Limitations	24
2.8 Conclusion	25

References	26
3 ROLE OF ARTIFICIAL INTELLIGENCE IN BIOLOGICAL SCIENCES	31
3.1 Introduction	32
3.2 Applications of Artificial Intelligence In The Field of Biology .	33
3.3 Role of AI in Genomics and Bioinformatics	34
3.4 Pros and Cons of AI In Biological Field	38
3.5 Future Prospects of AI In Biology	39
3.6 Conclusion	39
References	40
4 THE PHYSICS OF INTELLIGENCE: COMPUTATIONAL MODELS INSPIRED BY NATURAL LAWS	43
4.1 Introduction	44
4.2 Intelligence as A Physical Phenomenon	44
4.3 Thermodynamics And The Cost of Intelligence	45
4.4 Statistical Mechanics And Learning Dynamics	46
4.5 Nonlinear Dynamics and Self-Organization	47
4.6 Information Theory And Representation	48
4.7 Physics-Inspired Computational Models	49
4.8 Embodied Intelligence And Physical Interaction	50
4.9 Challenges And Open Questions	51
4.10 Future Perspectives	52
4.11 Conclusion	53
References	53
5 ARTIFICIAL INTELLIGENCE IN CHEMICAL SCIENCES: TRANSFORM- ING DISCOVERY, DESIGN AND DATA	57
5.1 Introduction	58
5.2 Literature Review	59
5.3 Molecular Design Using Machine Learning	66

5.4	Limitations	75
5.5	Conclusion	75
	References	76
6	AI IN GREEN CHEMISTRY: PATHWAY TO SUSTAINABLE CHEMICAL PROCESSES	83
6.1	Introduction	84
6.2	AI Techniques Supporting Sustainable Chemistry	86
6.3	AI-Assisted Molecular And Solvent Design	87
6.4	Reaction Optimization And Green Catalysis	89
6.5	Sustainable Separation Processes And Process Intensification .	90
6.6	AI For Low-Carbon Manufacturing And Circular Chemistry . .	90
6.7	Digital Twins And Smart Sustainable Manufacturing	90
6.8	Challenges And Limitations of AI Adoption In Green Chemistry	90
6.9	Future Perspectives Of AI In Green Chemistry	91
6.10	Conclusion	92
	References	93
7	ARTIFICIAL INTELLIGENCE IN MATHEMATICAL SCIENCES: METHODS, APPLICATIONS AND EMERGING TRENDS	97
7.1	Introduction	98
7.2	Mathematical Foundations Supporting AI	99
7.3	Applications In Pure Mathematics	101
7.4	Applications In Applied Mathematics	103
7.5	Interdisciplinary Applications	104
7.6	Emerging Trends In AI-Driven Mathematics	105
7.7	Challenges And Limitations	107
7.8	Future Directions	107
7.9	Conclusion	107
	References	108

8 HARNESSING ARTIFICIAL INTELLIGENCE FOR CYBERCRIME DETECTION AND PREVENTION: THREATS, TECHNIQUES AND CHALLENGES 111

8.1 Introduction 112

8.2 Conventional Cyber Crimes 113

8.3 Role of Artificial Intelligence In Cybercrime Control 120

8.4 Challenges And Limitations of Using AI Based Prevention Methods 122

8.5 Conclusion 123

References 124

9 ARTIFICIAL INTELLIGENCE IN NANOTECHNOLOGY 129

9.1 Introduction 130

9.2 Applications of Nanotechnology And Artificial Intelligence In Healthcare 132

9.3 Applications In Daily Use 136

9.4 Ethical And Regulatory Challenges In The Convergence of Nanotechnology And AI 137

9.5 Conclusion 138

References 139

10 ARTIFICIAL INTELLIGENCE IN ART AND CREATIVITY 143

10.1 Introduction 144

10.2 Creativity As A Unique Human Skill 144

10.3 Role of AI In The Field of Art And Creativity 145

10.4 AI Art Generator Tools 147

10.5 Impacts of AI on Artistic Expression 150

10.6 Potential of AI In Art And Creativity 151

10.7 Recent Trends - AI Is Revolutionizing Art 151

10.8 Some Ethical Challenges In AI-Generated Art 152

10.9 Future Of Artificial Intelligence In Art – Trends And Innovations 153

10.10 Use of AI In Different Areas of Creativity 154

10.11 Conclusion	154
References	155
11 ARTIFICIAL INTELLIGENCE, GOVERNANCE AND PUBLIC POLICY: BRIDGING TECHNOLOGY AND SOCIETY	157
11.1 Introduction	158
11.2 Foundations of AI Governance	159
11.3 Emerging Models of AI Governance	160
11.4 AI And Public Administration	162
11.5 Civil Liberties, Constitutionalism And AI Governance	164
11.6 Public Trust And Democratic Legitimacy	165
11.7 Comparative And Global Perspectives	166
11.8 Policy Implications And Future Directions	170
11.9 Conclusion	171
References	171
12 ADVERSARIAL ATTACKS AND AI SECURITY	173
12.1 Introduction	174
12.2 Artificial Intelligence And Trust	175
12.3 Security As A Fundamental Concern	176
12.4 Evolution of Adversarial Machine Learning	176
12.5 Taxonomy of Adversarial Attacks	178
12.6 Adversarial Attacks As An AI Security And Cybercrime Issue	180
12.7 Defensive Strategies In AI Security	183
12.8 Towards Resilient And Self-Healing AI Systems	186
12.9 Ethical, Social, And Management Implications	188
12.10 Future Directions And Open Challenges	191
12.11 Conclusion	192
References	192

13 EMOTION AI AND “FEELING AI” IN MARKETING: AD OPTIMIZATION, CONSUMER REACTIONS, AND ETHICAL BOUNDARIES	195
13.1 Introduction	196
13.2 Review of Literature	198
13.3 Theoretical Foundations of The Conceptual Model	199
13.4 Conceptual Framework	200
13.5 Application Scenarios And Use Cases	203
13.6 Conclusion	205
References	206
 14 THE END OF THE “MARKET FOR LEMONS”: AI AS A CATALYST FOR INFORMATION ASYMMETRY AND WELFARE OPTIMISATION	 211
14.1 Introduction	212
14.2 Structural Failures of The Problem of Information Asymmetry	215
14.3 AI As An Information Catalyst: The End Of Information Scarcity	217
14.4 Empirical Analysis: AI Applications In Mitigating Information Asymmetry: Ai-Driven Solutions In Different Sectors Of The Economy	219
14.5 AI And India: A Fertile Ground For Symmetry	221
14.6 The Macro Implications For “Viksit Bharat”	222
14.7 Conclusion	223
References	224
 15 ARTIFICIAL INTELLIGENCE AND BANKING SECTOR: A SYSTEMETIC REVIEW	 227
15.1 Introduction	228
15.2 Research Methodology	229
15.3 Data Analysis And Results	231
15.4 Discussions	232
15.5 Conclusion	245
References	245

Chapter 1

AI-DRIVEN BIOINFORMATICS: TRANSFORMING DISEASE DISCOVERY, DIAGNOSIS AND THERAPY

Harshil Dhiman

Assistant Professor, Department of Biotechnology, Khalsa College for Women, Civil lines, Ludhiana, Punjab, India- 141001.

Email: harshilbiotech14@gmail.com

Abstract

The introduction of Artificial Intelligence (AI) in the field of Bioinformatics has completely changed disease discovery, diagnosis and treatments. High-throughput technology generates large amounts of genomic, transcriptomic, proteomic and clinical data that are challenging to analyze using conventional statistical analyses alone. Machine learning and deep learning techniques have enabled scientists and clinicians to quickly identify and extract biologically relevant features from large volumes of complex datasets, thereby allowing them to make quicker and more accurate conclusions regarding disease-causing genes, pathways and molecular fingerprints. We use clustering and predictive modelling methods and classify a disease into very specific biological categories that allows for earlier diagnoses and the ability to predict the risk of developing a disease. This method is especially advantageous in the areas of cancer, neurological disorders and rare inherited disorders. Using AI for applications in bioinformatics has revolutionized the drug discovery process by accelerating many key biochemical processes, including target identification, Protein Structure Prediction, Virtual Screening and Drug-Protein Interaction Modelling. These new capabilities allow researchers to develop therapeutics faster and with lower costs. This chapter reviews recent advances in AI-driven bioinformatics and highlights

how intelligent computational models are transforming disease research from a slow, trial-and-error process into a fast, predictive and data-driven scientific discipline.

Keywords: Artificial intelligence, bioinformatics, disease discovery, drug discovery, precision medicine

1.1 Introduction

Artificial Intelligence (AI)'s integration into bioinformatics constitutes one of the most impressive paradigm shifts in contemporary biomedical research over the last two decades. High-throughput technology advancements have brought about an explosion of biological data through next-generation sequencing, single-cell RNA sequencing, proteomics, metabolomics and large-scale clinical data systems. These tools allow researchers to better explore how molecular mechanisms cause disease; but at the same time, they generate such complex, high-dimensional data that it exceeds the ability of standard statistical methodologies to analyze them. Conventional approaches often find it difficult to facilitate modeling for non-linear relationships and multilayered dependencies that are common in multi-omics and clinical datasets (Wekesa and Achieng, 2023).

Bioinformatics has changed from a discipline concentrated on the alignment and annotation of sequences to an integrated, systems-level science. Disease biology today recognises that the majority of disease states arise due to dynamic interactions between genes, regulatory networks, cellular states and environmental influences rather than as a result of independent genetic mutations. Developing adequate methods to represent this complexity requires computational frameworks that can learn hierarchically from biological data at many scales. Modern algorithms for machine learning and deep learning provide scalable methods for discovering hidden structures in large datasets, while new foundation models will enhance the capability to generalise biological knowledge across various situations (Cui et al., 2024).

AI-based bioinformatics is changing the way of identifying diseases. The implementation of deep neural networks has led to better interpretations of pathogenic variants and improved gene prioritization

in rare inherited disorders, helping to minimize the uncertainty associated with missense variant analysis (Cheng et al., 2023). In addition, systems-level foundational models like Gene former have provided the ability to recreate gene regulatory networks, providing mechanistic insights of cardiovascular diseases, such as cardio myopathy (Theodoris et al., 2023). Integration of multiple omics data via AI-generated approaches has improved classification of subtypes of tumors and identification of biomarkers in cancer, with improved prognostic stratification (Sartori et al., 2025).

In addition to discovery and diagnosis, AI has accelerated the development of new therapies. The ability to utilize advancements in protein structure predictions and protein-protein and protein-DNA interaction modeling through the large-scale implementation of alphaFold has allowed protein structure-based drug development and improved drug target validation (Varadi et al., 2024). The use of a combination of genomic, transcriptomic, proteomic, structural and clinical data within predictive computational systems is changing biomedical research into a much more rapid and data-based process and is being carried out in a precision-based manner. This chapter reviews the foundations of the data, methodological advances and translation of the application of AI as a means of advancing the process of disease discovery, diagnosis and therapy.

1.2 Biological and Clinical Data Foundations For AI-Driven Disease Research

AI-powered bioinformatics relies heavily on high-quality and diverse biological datasets for effective analysis and interpretation. Modern disease research relies on the integration of multiple data sources to understand disease mechanisms, diagnosis and therapeutic development. This combined information supports the development of AI procedures to help researchers and clinicians in exploring human disease. Whole genome and whole exome sequencing provide in excess of millions of mutations in every individual. AI-assisted variant detection can improve the reliability of variant detection across sequencing platforms (Wekesa and Achieng, 2023), enabling the researchers to better understand the pathophysiology and cause of rare diseases (Cheng et al., 2023).

Transcriptomic data provides functional insight into gene activity, while bulk RNA sequencing provides tissue-level expression changes and single-cell RNA sequencing provides high-resolution mapping of cellular heterogeneity. Methods such as scGPT and Geneformer create foundation models that learn generalizable representations of gene expression patterns that aid in cell type annotation and in discovering disease subtypes (Cui et al., 2024). These approaches are particularly beneficial for cancer and immune-related diseases in which the diversity of the cellular populations influences disease progression (Theodoris et al., 2023). Proteomics and structural biology expand the analysis of the molecular level and include the study of protein structure. Recently developed AI-based structural prediction tools, including alphaFold and others, are transforming protein structure modeling and predicting interactions between biomolecules (Varadi et al., 2024). Lastly, clinical datasets, such as electronic health records and medical imaging, provide phenotypic information. Integration of multi-omics can give powerful insight but harmonizing or validating any external data is important to provide replicable evidence and decrease bias in predictions. More recently, multimodal AI models have integrated both molecular and clinical data to facilitate decision-making for discovery, diagnosis and treatment of a disease (Ding et al., 2025).

1.3 Artificial Intelligence Methodologies In Disease Modeling

The main methodologies of AI in disease research today are classified into three areas: machine learning, deep learning and advanced foundation learning models that can identify complex, non-linear relationships within high-dimensional datasets. The conventional statistical machine learning techniques, including support vector machines (SVMs), random forests, and gradient boosting, are still very useful for research involving the life sciences. All these methods work well when feature sets are well-defined, and sample sizes are moderate to large. These methods also have an advantage over other methodologies such as explanatory models because they can provide a degree of interpretability i.e. produce an understandable output from the data and relatively stable performance from test to test.

Deep learning allows for even further automation of feature representation extraction through hierarchical learning on unpro-

cessed materials. Convolutional Neural Networks (CNNs) have been extensively used in the analysis of genomic sequences and medical images, while recurrent and transformer architectures enable the modeling of long-distance interactions in biological sequences and gene expression profiles. Ultimately, these architectures work well for predictive tasks including the effects of variants, single-cell clustering, and molecular interaction modeling (Cheng et al., 2023).

Graph Neural Networks (GNNs) are becoming increasingly significant in drug discovery and network biology because biological systems are naturally represented using graph structures which allows GNNs to predict drug-target interactions and identify modules linked to disease states (Wang et al., 2024). Recently, the emergence of foundation models trained on large biological datasets has led to a paradigm shift in bioinformatics. For instance, transformer-based systems such as Geneformer and scGPT are capable of learning generic representations of gene expression across millions of cells and hence these models can facilitate transfer learning from one biological context to another (Cui et al., 2024). Also, large protein language models can directly infer structural and functional information from amino acid sequences (Lin et al., 2023). These pretrained models minimize dependence on labelled datasets, thereby delivering enhanced performance in data-poor settings. It is important that the development of models include external validation cohorts, transparency and clear understanding of AI methods in order that safe medical treatment of patients can be done based on reliable models.

1.4 AI In Disease Discovery

Advancements in bioinformatics powered by AI have expedited the discovery of diseases by accurately identifying mature mutations found within genes and systems that disrupt molecular pathways or processes. A model called Alpha Missense created by deep learning has increased the accuracy of predicting pathogenic missense mutations from the entire human proteome, aided in finding new genes for rare inherited diseases and providing quicker diagnoses of Mendelian disorders (Cheng et al., 2023). This work decreases the number of unknown significance mutations and enhances reliance on finding causal genes.

AI has also aided in finding new mechanisms in cardiovascular disease processes through the application of a foundational model called Geneformer to develop gene regulatory networks for the cardiomyopathies and identify the gene targets from those networks that, when modified in an experimental manner, would improve the cardiomyocytes (Theodoris et al., 2023). This work converts from purely statistical associations to targeted discovery validated through experimentation. In cancer research, multi-omics deep learning approaches have identified new subtypes of tumors and prognostic biomarkers that apply to many cancers by combining growth profiles from genomic and transcriptomic sources (Sartori et al., 2025). This sort of analysis will help identify molecular heterogeneity, which might be useful for selecting precision therapies. Evidently, through these examples, it is concluded that AI has changed the speed at which scientists analyze and interpret data and will contribute greatly to ongoing efforts in identifying disease mechanisms and new therapeutic targets in at least three major areas: rare genetic conditions, cardiovascular diseases and cancer.

AI-enabled disease discovery increasingly incorporates integrated multi-omics frameworks, assessing system-wide disruptions resulting from the combination of genomic, transcriptomic and proteomic components instead of just assessing isolated mutations. The combining of heterogeneous biological datasets with deep learning models enables better identification of disease-associated networks and prediction of biomarkers across complex multifactorial diseases (Wekesa and Achieng, 2023). Recent advancements using very large protein language models trained on large scale sequence data provide the ability to connect genetic variants with functional interpretation of genetic variants by generating functional correlates between sequences, structural properties and biochemical properties (Lin et al., 2023). Platforms providing structural prediction tools, including AlphaFold, will provide additional mechanistic insight by clarifying how identified mutations impact the structural stability and interactions of proteins (Varadi et al., 2024). Collectively, these advances provide evidence that AI-based bioinformatics can shorten the time required for discovery as well as improve understanding the mechanistic level at both molecular and structural levels of the disease process.

1.5 AI In Diagnosis and Patient Stratification

In addition to discover new diseases, AI bioinformatics is integral to the diagnosis of patients as well as dividing them into groups of similar patients. In modern medicine, the idea that two patients could both be diagnosed with a particular illness yet represent different molecular expressions of the illness (and risk factors and therapeutic response) is well established. AI models can help to identify these subpopulations using predictive modelling, integrating genomic, transcriptomic and clinical attributes.

In the field of oncology, cancer classification based on deep learning multi-omics models has allowed for significant improvement in accurate cancer classification and prognostic prediction through integration of gene expression, mutation profiles and proteomics data (Sartori et al., 2025). This integration of clinical data, molecular characteristics and genetic data supports the identification of new subtypes of cancer based on molecular characteristics and provides insights into survival differences and varied treatment responses. This stratifies treatment so that providers are able to develop individualized treatment plans instead of relying on histopathological classification alone. In the area of rare genetic disorders, AI-assisted tools are increasingly being used in clinical genetics to interpret genetic variants as part of the clinical diagnosis pipeline. Recently published systematic reviews summarize how ML (machine learning) models support rapid prioritization of genetic variants and reduce time to reach a definitive genetic diagnosis for patients with Mendelian disorders (Dias and Torkamani, 2023). These models identify relevant genetic variants faster than traditional methods and identify genetically related disorders or syndromes earlier than previously thought possible. Therefore, ML increases the overall accuracy of diagnosis and supports timely patient intervention. AI shows how important it is to stratify cardiovascular disease. Foundation models that operate on gene expression can effectively identify disease-associated cellular states associated with cardiomyopathy (Theodoris et al., 2023). They can also be used to identify molecular targets that are relevant to disease progression. This information will allow for more precise risk assessment through molecular reassessment instead of solely relying on imaging or symptom evaluations of patients. Multimodal AI systems can also greatly extend the ability to diagnose disease by combining both

molecular data and clinical and imaging records. Recent reviews have noted that the development of multimodal medical foundation models will improve diagnostic performance and assist in selecting therapies to treat complex diseases (Ding et al., 2025). These technologies will assist the healthcare industry in using predictive and integrated data for clinical decision-making in real-time.

In summary, AI-driven bioinformatics will provide enhanced diagnosis through molecularly informed stratification, moving away from symptom logically identified classifications. By recognizing biologically relevant divisions of patients and providing a means to anticipate the outcomes of a particular disease process, AI will aid in delivering precision medicine and individualized treatment plans.

1.6 AI Foundation Models and Multimodal Clinical Intelligence

An important change in AI within the field of Bioinformatics is the move away from Machine Learning (ML) systems that are trained to complete specific tasks toward the use of ML models that serve as foundations for producing general representations across many different forms of biomedical data. While ML models used to be focused on very specific and defined tasks, foundation models have been generated from large-volume datasets and will then be refined for downstream applications related to new disease identification/discovery and new disease diagnosis; and clinical decision support. The importance of this evolution lies in the fact that, to effectively understand diseases, there needs to be a combination of molecular, imaging, pathological and clinical data for analysis, rather than relying solely on a single data type or one data type (Topol, 2023).

Large language models have enhanced the capacity to synthesize biomedical literature, identify gene-disease associations and assist in generating hypotheses through data extraction from texts. One example of this is BioGPT, which serves as an example of how generative models for a specific domain (e.g., biomedical) will support the mining of biomedical texts and accelerate the pipeline for knowledge discovery (Luo et al., 2022). Although these applications of large language models do not replace validation through experimentation, they expand and strengthen the interpretative layer of bioinformatics

workflows through the organization and contextualization of a large amount of scientific evidence.

Digital pathology has revolutionized the use of high-resolution whole slide imaging for processing diagnostic and prognostic tasks in oncology. For example, multimodal foundational models developed in digital pathology have shown strong success in performing classification and representation learning using high-resolution whole slide images. In addition, multimodal foundational models can be utilized as scalable features in pathology and clinical decision support. However, validation across multiple institutions is required for this application. This demonstrates the need for more research into multimodal foundational models as a method to improve diagnostic accuracy and therapeutic decision-making. In dermatology, an example of multimodal integration is seen through the use of vision-based foundational models. These generalizable, large-scale pre-trained systems are effective across various types of skin pathology and provide real-world applications. Integrating visual and contextual clinical information through large-scale foundational models improves accuracy, consistency and may assist in therapeutic decision making for clinicians. Increasingly, oncology is using a multimodal approach to integrate several sources of data in order to better understand tumors and to improve their diagnosis and treatment. The research has shown that integrating image; molecular and clinical data leads to better risk stratification and patient sub-group identification than if using just one source of data, in other words, through unimodal data models (Waqas et al., 2024). These same authors suggest that multi-omics models representing the molecular variability of a tumor; in other words, as multiple associated molecules can be used to link the molecular variability of a tumor to real-world clinical outcomes (Sartori et al., 2025). By combining multiple types of data, multimodal approaches create a stronger relationship between biological processes and clinical decision-making.

Despite the potential of foundation models, they should be developed with caution. There are still a number of challenges such as dataset bias, distribution shifts and lack of interpretability that need to be addressed before deploying foundation models in the clinic. If foundation models are not validated and governed adequately during development, they may incorrectly identify tumours or

produce inconsistent clinical outcomes. Therefore, the effectiveness of multimodal clinical intelligence solutions depends not so much on how sophisticated the algorithms are, but rather on how well those algorithms are validated and how they are used with input from clinical experts.

In summary, foundation models will represent a significant advancement in the ability to link molecular bioinformatics to clinical practice in the real world. They will provide the means to scale up how clinical and biological data are integrated and reinforce that AI will play a transformative role in disease discovery, diagnosis and treatment innovations.

1.7 AI In Drug Discovery and Therapeutic Development

The use of AI for bioinformatics is helping to speed up the process of finding new medications through advances in the ability to find good medication targets, predict the structures of those targets and create computer models of how the drugs will interact with those targets. Drug development has always been an expensive, time-consuming process; however, due to the high rate of failure seen in phased-in clinical trials, wings of drug discovery can now use these AI-based techniques to identify true potential drug targets in an efficient and timely manner.

One area of significant progress is in the use of deep learning neural networks to predict 3D shapes of proteins based solely on their sequences of amino acids. AlphaFold is a neural network-based software application for accurately modeling any given protein's 3D shape and is expanding its potential by beginning to incorporate mechanics of a protein interacting with a small molecule. Therefore, alphaFold can be used for "structure-based" drug design by allowing researchers to make predictions for different mechanisms of how small molecules will bind to the protein, even when no experimental crystal structure exists. The exploration of drug-target interactions, as well as enhancing lead compounds continues to be aided by accelerating the ability of AI-powered framework to undertake virtual drug screening processes through the use of GNNs and other models (Wang et al., 2024) along with the use of AI-enhanced models to track potential therapeutic compounds through various types of diseases (Noor et al., 2024). Target discovery in disease-specific locations has

also seen improved performance due to foundation-based methods within gene expression networks associated with cardiomyopathy, which are aided by identification and validation of potential therapeutic targets (Theodoris et al., 2023). These types of studies demonstrate the integration between AI, disease modeling and then developing treatments.

AI methods/protocols applied from target identification to lead optimization are modifying the overall drug discovery and development process. Ultimately, while experimental validation is still necessary, the overall use of AI provides significant improvements related to the overall cost, time required and inability to complete therapeutic development leading to an efficient bridge between discovery and clinical use.

Recent developments in biomolecular interaction modeling have been supported by advances in AI-supported therapeutic development. Using AlphaFold 3, researchers can predict structure not only for isolated proteins but also for those interacting with other proteins and their ligands (fragments of large molecules like enzymes), thus making drug binding mechanism predictions more accurate and optimizing lead compounds before they are experimentally tested more reliably than ever before (Abramson et al., 2024).

1.8 Challenges and Limitations of AI In Biomedical Applications

Although AI can change the world; however, there are challenges with AI-driven bioinformatics that need to be resolved to apply the technology safely. The challenge is with heterogeneous data as many of the datasets may be made up of multi-omics and clinical data, and they have batch effects, demographic imbalances, or missing data, which can create biases in predictive models. Without appropriate harmonization of data, and proper external validation, the AI systems may create overly optimistic and population-specific results.

Interpreting models adequately is another significant problem as deep learning and foundation models are generally considered "black-box" systems, therefore predicting some outputs may be difficult to validate their rationale. As these predictions are often being made in biomedical settings, which may have significant impact on diagnosis or treatment, a lack of visibility may limit the ability for

the clinician to trust the information and receive regulatory approval. The development of explainable AI frameworks is an ongoing area of investigation.

If there is no separate training and validation datasets strictly, overfitting and data leakage can lead to negative outcomes. Reliable benchmarks and reproducible methodologies are the foundation for scientific credibility. In addition, AI deployment must also be governed by ethical issues around data privacy, informed consent and equal access to healthcare. Algorithms trained mostly on a particular population yield likely to underperform for other groups leading to increased disparities in healthcare.

Lastly, experimental validation and clinical validation are needed to compliment the computational prediction of AI models. AI models generate hypotheses and target prioritization; however, biological confirmation is essential. Future success of AI-based bioinformatics depends on both advancements in algorithms and the successful integration of experimental science and clinical expertise.

1.9 Conclusion and Future Perspectives

Bioinformatics has been radically transformed by AI through the introduction of large-scale, integrated and predictive methods for investigating complex biological systems. Together with disease discovery, AI assists in identifying pathogenic variants, dysregulated pathways and new therapeutic targets. AI has also aided drug development, leading to a faster validation of targets and optimised leads via enhanced development of structure prediction, graph-based modelling and virtual screening methods. Collectively, this represents a shift from descriptive molecular profiling to data-driven and mechanism-based biomedical research.

The future of AI-powered bioinformatics is focused on maximizing the capabilities of multimodal learning. These new advanced models will incorporate genomic, transcriptomic, proteomic, imaging and electronic health record information into single predictive systems that can predict how an individual will develop a given disease. Furthermore, there will also be increased use of Generative AI to create proteins, produce synthetic biology products and develop new drugs through rational design. Finally, improved modelling of the dynamic processes that occur biologically, such as transitions

between cellular states and changes in protein conformation, will allow us to better understand the ways how diseases progress and respond to treatment. It will also be important that we have a consistent approach to interpretability and clinical reliability so that we can effectively translate AI-driven bioinformatics into clinical practice. Overall, we need interdisciplinary collaboration among biologists, computational scientists, clinicians and regulatory bodies to responsibly grow the field.

Bioinformatics driven by AI is creating a new way of approaching diseases, how diseases are diagnosed and treated. Through ongoing methodological development and ethical application, AI has the potential to usher in a new era of predictive, precise and patient-centred biomedical science.

References

- [1] Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O..... et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630, 493-500. <https://doi.org/10.1038/s41586-024-07487-w>
- [2] Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., and Avsec, Ž. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*, 381(6664), eadg7492. <https://doi.org/10.1126/science.adg7492>
- [3] Cui, H., Wang, C., Maan, H., Pang, K., Luo, F., Duan, N., and Wang, B. (2024). ScGPT: Toward building a foundation model for single-cell multi-omics using generative AI. *Nature Methods*, 21(8), 1470–1480. <https://doi.org/10.1038/s41592-024-02201-0>
- [4] Dias, R., and Torkamani, A. (2023). Artificial intelligence in clinical genomics. *Nature Reviews Genetics*, 24, 617–634. <https://doi.org/10.1038/s41576-023-00598-0>
- [5] Ding, T., Wagner, S. J., Song, A. H., et al. (2025). A multimodal whole-slide foundation model for pathology. *Nature Medicine*, 31(11), 3749–3761. <https://doi.org/10.1038/s41591-025-03982-3>
- [6] Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., et al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637), 1123–1130. <https://doi.org/10.1126/science.ade2574>
- [7] Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., and Liu, T. -Y. (2022). BioGPT: Generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, 23(6), bbac409. <https://doi.org/10.1093/bib/bbac409>

- [8] Noor, F., Zaman, N., and Ahmed, S. (2024). Deep learning-driven virtual screening for drug discovery: Recent advances and challenges. *Scientific Reports*, 14, 79799. <https://doi.org/10.1038/s41598-024-79799-w>
- [9] Sartori, F., Codicè, F., Caranzano, I., Rollo, C., Birolo, G., Fariselli, P., and Pancotti, C. (2025). A comprehensive review of deep learning applications with multi-omics data in cancer research. *Genes*, 16(6), 648. <https://doi.org/10.3390/genes16060648>
- [10] Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Alvarado, D. M., et al. (2023). Transfer learning enables predictions in network biology. *Nature*, 618, 616–624. <https://doi.org/10.1038/s41586-023-06139-9>
- [11] Topol, E. J. (2023). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 29, 44–56. <https://doi.org/10.1038/s41591-022-02147-5>
- [12] Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2024). AlphaFold Protein Structure Database in 2024: Expanding coverage and improving structure quality. *Nucleic Acids Research*, 52(D1), D368–D375. <https://doi.org/10.1093/nar/gkad1011>
- [13] Wang, R., Sun, Y., Li, J., and Chen, X. (2024). Graph neural networks in drug discovery: Methods and applications. *Briefings in Bioinformatics*, 25(2), bbac652. <https://doi.org/10.1093/bib/bbac652>
- [14] Waqas, A., Tripathi, A., Ramachandran, R. P., Stewart, P., and Rasool, G. (2024). Multimodal data integration for oncology in the era of deep neural networks: A review. *Frontiers in Artificial Intelligence*, 7, 1408843. <https://doi.org/10.3389/frai.2024.1408843>
- [15] Wekesa, J. S., and Achieng, A. O. (2023). A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment. *Frontiers in Genetics*, 14, 1199087. <https://doi.org/10.3389/fgene.2023.1199087>

Chapter 2

AI IN BIOLOGICAL SCIENCES

Sanjana Sharma

Assistant Professor, Department of Biotechnology, Khalsa College for Women, Civil Lines, Ludhiana, Punjab, India-141001.

Email: sanjanak2203@gmail.com

Abstract

AI has become a vital tool in biological sciences, mainly through machine learning and deep learning techniques. Modern biological research produces abundant data from areas such as genetics, protein studies, medical imaging and laboratory experiments. To examine this complex data using traditional methods is often difficult and slow. AI helps scientists to analyse and understand this data by processing it quickly and to figure out the details that may not be easily noticed by humans. One of the most important contributions of AI is in protein structure prediction. With the help of the tool called ALPHA FOLD, it is easier to predict protein shapes more accurately, which ultimately saves time and effort in laboratory experiments. In case of drug discoveries, AI helps researchers to identify potential drug targets, predict the interaction of a drug with biological molecules and design better medicines at lower cost. AI also supports making medicine according to the patient's requirement by combining the genetic information with patient health records to suggest treatments according to the individual's requirement or say illness. AI also helps in the laboratory by improving efficiency, reducing human error and speeding up research work. Lastly, the use of AI is transferring biology into a more predictive and technology-based science, leading to important advancements in healthcare, agriculture and environmental studies. In short, AI tools speed up biological research by accurately analysing complex data, providing more accurate diagnostic results, drug discovery and better knowledge

about life processes. It also supports faster, smarter and more precise scientific advancements.

Keywords: accuracy, ALPHAFOLD, drug discoveries, improved efficiency, modern biological research, scientific advancement

2.1 Introduction

There is no exact scientific definition of Artificial Intelligence (AI), but broadly we can refer to it as the ability of any machine that can simulate the intelligence of higher organisms. The field of AI plays an important role in almost every branch of research, including philosophy, mathematics, computing, psychology and biology (Chowdhary, 2020). You can refer to an ideal AI system as a self-aware, logical and able to learn from experiences. It would also be able to interpret and react to external environments. Algorithmic assistance is primarily of two types, i.e., machine learning (ML) and deep learning (DL) approaches. With the help of these two approaches, an interactive system can be developed to carry out activities that require human reasoning ability.

In recent years, AI technologies have been widely adopted in various sectors such as healthcare, education, agriculture, finance, transportation and cyber security. These technologies help in improving efficiency, reducing human effort and increasing the accuracy of decision-making processes. Machine learning enables systems to learn patterns from data without being explicitly programmed, whereas deep learning uses artificial neural networks to solve more complex problems such as image recognition, speech processing and natural language understanding. The rapid growth of computational power and availability of large datasets have significantly accelerated the development of AI-based applications. As a result, AI is becoming an essential component in the advancement of smart systems and automation technologies across the world.

2.2 Evidence of The Usage of AI In Biological Sciences

- In the Dartmouth Conference, the term artificial intelligence (AI) was first coined by John McCarthy in 1956 for an intelligent machine system.

- The initial work in AI includes the contribution of mathematician Alan Mathison Turing. He proposed his ideas in a public lecture in London about the concept of self-learning and self-instructed machines that learn from their own experiences in the same way as a human being does (Turing et al., 1950; Buchanan, 2005). Based on his study of facts about smart machines, Alan Turing is also regarded as the father of AI and modern computer science. The statement “The human brain functions essentially like a digital computer” was proposed by him (Akman and Blackburn, 2000). He did an experiment known as “The Turing Test”, which became a historic event in the development of AI. His paper titled ‘Computing Machinery and Intelligence,’ looked into the possibility of non-living computer thinking like a human and was a landmark in this area (Turing et al., 1950).
- In 1973, the first humanoid robot was developed by a group of Japanese engineers, who had proficiency in machine learning at the time, including the ability to walk upright, hold objects and converse in Japanese just like a human. This test was carried out simply to figure out whether a machine has the ability to think.
- Another significant event in the AI timeline was the construction of IBM’s supercomputer, Deep Blue, which was capable of playing chess completely equivalent to humans. It was the first artificial intelligence that defeated Grandmaster Garry Kasparov Russian chess player in a timed match (McCorduck, 2004).
- NASA’s space-based autonomous vehicles have started the impactful use of AI, where they use this technology to steer and move vehicles on their own without human intervention.

To create more advanced, self-aware artificial systems, the two approaches, i.e. DL and ML, are trained to look into the knowledge from data of various sources that are either generated directly or indirectly from the natural intelligence system. The more exposure or training is done, the more self-reliant the system can become (Sarker, 2022).

2.3 Types of AI

AI can be divided into two different categories, i.e. Narrow or Weak AI and Artificial General Intelligence or Strong AI.

Weak AI makes some attempt to copy or mimic human cognitive thought; it enables the automation of the majority of tasks in ways that humans are incapable of (Fjelland, 2020). The most visible examples of weak AI on a daily basis include Tesla's Autopilot feature, facial recognition on our smart phones, Google's search engine, Instagram's AI for understanding user interests, Apple's Siri and Amazon's Alexa.

Strong AI is far more advanced and complex than weak AI. Strong AI is not restricted by human-made laws, and it is able and efficient to think and control the system entirely on its own.

2.4 Integration of Artificial Intelligence with Biosciences

AI has now made its way into the biological field, demonstrating its worth through innovative procedures (Han and Liu, 2019). Additionally, the world has seen a true revolution in the field of information technology (IT). Both information technology and biology have flourished a lot during the past half-century.

According to Moore's law, the computing power can be greatly increased as the number of increases. The numbers of transistors present on the chip will double every two years. Also, the biosciences and biotech industries have tremendously paced their way as there was a marvelous development in sequencing and other high-throughput techniques (Pettersson et al., 2009).

AI-based algorithms have made it possible to build intelligent computer systems that can handle complex decisions, as the raw and unstructured data can be stored and analyzed, which ultimately leads to finding useful data from this (Williams et al., 2018; Sellwood et al., 2018).

The development of a wide range of products and services in different fields, including biosciences and the biotech industry, is possible because of the advancement and improvement in areas of data generation, storage and analysis.

2.5 Applications of AI

Artificial Intelligence is being used in biology, including the precise identification of the 3D geometry of biological molecules such as proteins, which is one of the most critical tasks and is useful in biological research. Moreover, in biological science, AI plays a critical role not only in laboratories scale by promoting innovations and inventions, but also throughout the lifecycle of a medication or chemical product. Furthermore, AI-based tools and applications also helped us to mechanize the complicated production procedures, which ultimately lead to the fast-rising demand for medications, chemicals for use in industry, and food and other bio-based raw materials. ML further helps in the estimation of results by applying large-scale permutations and combinations of datasets available for the drug molecules to determine the best combination without relying on traditional manual methods in the lab (Dara et al., 2022). Although nowadays also, the old and conventional model-driven methods are still useful for analysing biological data, there is an inability to use large amounts of data available, or even big data, to unravel the information, predict data nature and understand complex data linkages. The massive use of large data is leading to an important role in biotechnology and bioinformatics as it continues to grow and remains accessible to academicians and scientists for analysis throughout the world (Chakraborty and Choudhury, 2017). This massive information is measured in terms of multi-omics (study of multiple data types like genomics, proteomics, metabolomics, transcriptomics, etc.) obtained from a variety of biological sources, which further needs to be properly explained and examined to understand complex biological systems. The usage of AI approaches is now widely observed in the field of biological science and is expected to increase further in the near future as this technology matures.

Furthermore, medical imaging and drug responses among the population contribute complex but significant data needs to be analysed efficiently by algorithmic programs. Therefore, ML- and DL-based AI is attaining much attention due to their capabilities for faster processing of huge data and the capability of deriving meaningful information. AI-based digital image processing, drug designing and virtual drug tests might bring a tremendous change in medical science in the near future (Paul et al., 2021, Carpenter and Huang, 2018).

This chapter highlights how Artificial Intelligence and its components could be used in the medical sector to make human life more sustainable.

2.6 Medical Science and Biotechnology Advancements

Medical Science and Biotechnology advancements have paved the way for new ways of developing medications and antibiotics. AI has immense potential for universal applications in the pharmaceutical industry. A branch of AI known as ML is commonly used in disease detection since it utilizes the results of diagnostic testing to improve the reliability of results (Kumar et al., 2023). AI further allows researchers to manage challenging issues, including quantitative and predictive epidemiology, precision-based medicines and host–pathogen interactions (Jiang et al., 2017). AI can help in disease detection and diagnosis, and make computer code more accessible for non-technical individuals (Jiang et al., 2017). Predictive epidemiology uses data, statistical models and computation tools to predict the spread or risk of diseases before or during outbreaks. Individual-based precision medicine and the analysis of host–pathogen interactions are examples of research areas that could benefit from machine and deep learning breakthroughs (Zhang et al., 2017). These methods have further helped in the disease diagnosis and individual case identification, more efficient forecasts and fewer mistakes, faster decision making and better risk analysis. The increasing ratio of tissue biomarkers and the complexity of their evaluations have remarkably increased the use of AI-based techniques.

These AI-based biomarkers are working in a very efficient way to further help physicians in predicting and analysing diseases, about patients' body responses to the treatment and chances of survival of the patient (Lancellotti et al., 2021; Bedi et al., 2015). ML-based methods are also used to help give positive results in the efficiency and reliability of epidemiological models (Pinaire et al., 2017; Vrakas and Vlahavas, 2008). The advancement in ML programs has also helped to invent ten cellular parameters algorithmic program-based models which would precisely differentiate between benign and malignant tumours (Osareh and Shadgar, 2010).

The surgical outcome of pulmonary cancer LCP-CNN, Brock model was achieved accurately. Clinical data, LCP-CNN was able to predict the malignancy of pulmonary nodules with higher accuracy and lower false negative results than the Brock model. It is important to mention the importance of the combination of these algorithms with medical expertise. New pharmaceutical compounds are being developed by involving data analysis using AI, which further decreases the number of clinical trials, which further promotes the market to be flooded with medications without undermining the health of patients (Paul et al., 2021).

Moreover, AI is helpful in the detection of genetically susceptible diseases at earlier stages, which further facilitates patients to be able to prevent and treat the onset of inherited diseases. One of the applications of AI in the pharmaceutical industry is “Open Targets”, which explains the connection between drug targets, diseases and genes, and also explains to us the impact of genes on certain diseases. SPIDER is another AI technique that is used to identify the role of natural products in drug discovery (Paerhati et al., 2026).

Furthermore, quantitative structure–activity relationship (QSAR) studies are beneficial in generating new and more effective medications in a very short span of time using a computer simulation tool. Natural language processing (NLP), ML and robotic process automation are jointly the three key domains for the advancement of AI in the field of medicine. Natural Language Processing has recently been used to enhance the testing of colonoscopy (Harkema et al., 2011). According to medical practitioners, the way a body responds to certain drug action mainly depends upon the metabolic, physical, physiological and genetic makeup of an individual. In spite of knowing this fact, the doctors are practicing an umbrella approach in which the same drug is used to treat all patients for a particular disease, knowing the fact that there can be a slow response towards the recovery or there can be a chance of infection, as that particular drug may not be suitable for all individuals. In short, we can say that doctors are using an umbrella approach regardless of the varying conditions with the same drug. Nowadays, we are on the verge of developing a technology where we can have personalised medications, or we can say a medicine is tailored according to an individual's body needs and adaptability. Despite it seeming to be a very simple procedure, in actuality this method involves

a large amount of data collection, processing, maintenance, and execution (Mathur and Sutton, 2017). Using this technique, physicians and clinicians may better predict which disease treatments and preventative strategies will be most effective for particular patient groups.

Researchers are nowadays working on how to use AI in DNA, RNA and protein studies to understand the effect of drug dosage on living tissue over time and reorganize signaling networks during therapy (Azofeifa and Dowell, 2017, Azofeifa et al., 2018). Based on AI, IBM Watson helped in the formation of the relevant treatment plan for a patient depending on the patient's medical history and personal data, including genetic makeup (Aggarwal et al., 2016). An AI-based system of personalised medicine will not only reduce treatment cost but also minimise the side effects of drugs in the patient (Maceachern and Forkert, 2021). AI helps in saving time and improving patient care, it can also provide ease in gene editing, radiography and drug management planning procedure (Brogi and Calderone, 2021). Furthermore, electronic health records (EHRs) can be maintained in a better way with evidence-based clinical decision support systems (Milano et al., 2014, Gómez-González et al., 2020). This technique contributes to the development of innovative models of disease pathophysiology and can help predict major adverse effects of prolonged medications.

In a recent study, an AI-based approach was found to be very beneficial for the early identification, diagnosis, prognosis and treatment of myopia (Zhang et al., 2022). In cardiology, dermatology and oncology, deep learning algorithms help physicians in the diagnosis of disease (Luo et al., 2016). Computer algorithms can detect metastatic breast cancer in sentinel lymph node biopsies in full slide images with an accuracy rate of more than 91 per cent, and this was raised to 99.5 percent when physician inputs were added (Joshi and Patel, 2026).

One of the proven applications of AI in risk analysis is for diagnosing heart malfunction through cardiovascular imaging. It works by automated monitoring of any deviations from normal conditions based on image processing, myocardial function, detection and analysis of coronary atherosclerotic plaques (Zhang et al., 2022). The YOLOV3 algorithm was used for AI-based medical image

segmentation for 3D printing and naked eye 3D visualisation to detect the prostate in T2-weighted MRI images (AIMIS3D) (Jia et al., 2022). Certain other variables can be analysed through AI, which involves determining the conditions that are resistant to some antibiotics and not to others. Such analysis can help physicians in detecting diseases and considerably decrease unnecessary testing and costs in medical care.

Alzheimer's disease RF, SHAP Clinical and Imaging data: An AI model can detect and predict the progression of Alzheimer's disease with an accuracy of 93.95% in first layer and 87.08% in the second layer. COVID-19 PA Clinical data, with an accuracy of 70–80% was achieved in predicting severe COVID-19 cases (Jiang et al., 2020).

Ovarian cancer ANN Clinical data: An accuracy of 93% was achieved in predicting the survival of ovarian cancer patients, and 77% accuracy was achieved in predicting the surgical outcome (Enshaie et al., 2015).

Pulmonary cancer LCP-CNN, Brock model Clinical data: LCP-CNN was able to predict the malignancy of pulmonary nodules with higher accuracy and lower false negative results than the Brock model (Baldwin et al., 2020).

Influenza IAT-BPNN: CDC data and Twitter dataset IAT-BPNN was able to predict influenza-like illness in a large population size with a high accuracy using a computer simulation tool (Li et al., 2018).

Natural language processing (NLP), ML and robotic process automation are the three key roles playing in the advancement of AI in the field of medicine (Zou et al., 2020). Natural Language Processing has recently been used to enhance colonoscopy analysis, which further improves the detection of adenoma and polyps. An ML approach is used to predict diseases such as a trial fibrillation and urinary tract infections in patients by using models such as a support vector machine (SVM) based on clinical features of the disease. Similar methods have been used to detect and improve heart diseases using a heart-murmur-detecting technology.

The FDA has approved up to 29 AI-based medical devices and algorithms in various fields of medical sciences. The first AI-based model approved by the FDA in the healthcare sector was a diagnostic model based on an autonomous AI system, IDx-DR. The role of

this model is mainly to detect diabetic retinopathy with sensitivity, specificity and imageability of 87.2%, 90.7% and 96.1%, respectively. In this model, a sample dataset consisting of individuals of different ages, races and sexes is used, which minimises the possibility of incorrect results in different groups (Abràmoff et al., 2018).

A number of randomised clinical trials (RCTs) have also been done to test the capability and safety of AI and ML models in clinical practice. The application of AI systems in the medical field for the detection and prediction of diseases at earlier stages has favored the patients in a very positive way by reducing the cost of complete medical treatment, including blood tests, scans and also the medicines. In a study, the use of a procalcitonin-based decision algorithm (PCTDA) for hospitalised sepsis and lower respiratory tract infection has helped the patients to a shortened duration of stay, lower amounts of antibiotics given, lower exposure to artificial ventilation, and also there is a decrease in the number of patients with infections and antibiotic resistance. According to the latest report, PCTDA-based treatment has resulted in a 49% and 23% downfall in overall expenses from conventional treatment for sepsis and lower respiratory tract infections, respectively (Peng et al., 2019).

The pharmaceutical industries are now able to understand genetic information in a much better way with improved AI and ML skills. Robotic process automation, when used with ML and NLP, is much more useful and also has the ability to reshape medical science in the future. Regardless of the inventions we have studied so far, there is still a lot of work to be done before AI-based therapy becomes a regular practice.

2.7 Challenges and Limitations

Even though AI has immense potential, it is still not used in everyday life in healthcare practices. AI has the capacity to bring a positive change in the field of biology, but at the same time, it can also have a negative impact on the existing data. The reason behind this change is mainly because working of AI majorly depends upon the data used to train it, and that data is labeled. To make sure about the reliability of neural networks, a number of steps are required like collection, organising and cleaning of high-quality data. Identification and correction of errors in real-life applications must be done under

the guidance of field experts to develop and check the AI systems. Even after the development of AI tools and learning models has progressed so far, the only a few of them are used in healthcare practices, and many potential applications are still unwound. The most common reason for not accepting AI in medical practices is the absence of clinical data due to administrative guidelines, inconsistent results in data handling, obtaining results and unresolved issues around responsibility and transparency to patients. An example of non-interpretative results of a clinical AI-based system is DeepGestalt, a tool for analysing facial dysmorphology (Reiter et al., 2024). This tool performed poorly when identifying people with Down syndrome who were of African heritage (36.8%) compared to those who were of European origin (80%). Due to various segregations in training datasets, genetic disease vulnerability modeling is also responsible for differential performance across demographic groups. Thus, it is important to consider that the usage of AI-based approaches will not always lead to better prediction than present methods. AI is a tool that should be used within the proper context procedures, which further helps to sort an important issue. Thus, large-scale data sets and proportional studies are required to develop AI and ML-based models for real-time monitoring and control of bioreactors and bioprocesses.

2.8 Conclusion

One of the great achievements we have seen in this era is the ability of a machine to replicate the capacities of living systems, particularly the intelligence of a human. The ability to recognise objects and make decisions is a crucial characteristic of biological systems. AI is able to recognise objects and make decisions using many of the logical observational abilities of living systems. The power of AI can be used in the biological world, including medical research, agriculture and bio-based industries, for our sustainable way of life. The early prediction and identification of disease and its short time-consuming treatment based on personalised medicine, even while the diseases are in asymptomatic conditions, are examples of how AI can help in medical sciences. This can not only save millions of lives but also reduce medical treatment costs, which ultimately benefit the lower-class people, who lose the lives of their near and dear ones because of high treatment costs.

References

- [1] Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N., and Folk, J. C. (2018). Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NpjDigital Medicine*, 39(1). <https://doi.org/10.1038/s41746-018-0040-6>
- [2] Aggarwal, M., and Madhukar, M. (2016). IBM's Watson analytics for health care: A miracle made true. In *Cloud computing systems and applications in healthcare*, 117–134. IGI Global.
- [3] Akman, V., and Blackburn, P. (2000). Editorial: Alan Turing and artificial intelligence. *Journal of Logic, Language and Information*, 9, 391–395. <https://doi.org/10.1023/A:1008389623883>
- [4] Azofeifa, J. G., Allen, M. A., Hendrix, J. R., Read, T., Rubin, J. D., and Dowell, R. D. (2018). Enhancer RNA profiling predicts transcription factor activity. *Genome Research*, 28, 334–344. [10.1101/gr.225755.117](https://doi.org/10.1101/gr.225755.117)
- [5] Azofeifa, J. G., and Dowell, R. D. (2017). A generative model for the behavior of RNA polymerase. *Bioinformatics*, 33, 227–234. [10.1093/bioinformatics/btw599](https://doi.org/10.1093/bioinformatics/btw599)
- [6] Baldwin, D. R., Gustafson, J., Pickup, L., et al. (2020). External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax*, 75, 306–312. <https://doi.org/10.1136/thoraxjnl-2019-214104>
- [7] Bedi, G., Carrillo, F., Cecchi, G. A., et al. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophrenia*, 1, Article 15030. <https://doi.org/10.1038/nipschz.2015.30>
- [8] Brogi, S., and Calderone, V. (2021). Artificial intelligence in translational medicine. *International Journal of Translational Medicine*, 1, 223–285. <https://doi.org/10.3390/ijtm1030016>
- [9] Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *AI Magazine*, 26, 53.
- [10] Carpenter, K. A., and Huang, X. (2018). Machine learning-based virtual screening and its applications to Alzheimer's drug discovery. *Current Pharmaceutical Design*, 24, 3347–3358. <https://doi.org/10.2174/1381612824666180607124038>
- [11] Chakraborty, I., and Choudhury, A. (2017). Artificial intelligence in biological data. *Journal of Information Technology and Software Engineering*, 7, 207. <https://doi.org/10.4172/2165-7866.1000207>
- [12] Chowdhary, K. R. (2020). Fundamentals of artificial intelligence. Springer.
- [13] Dara, S., Dhamercherla, S., Jadav, S. S., et al. (2022). Machine learning in drug discovery: A review. *Artificial Intelligence Review*, 55, 1947–1999. <https://doi.org/10.1007/s10462-021-10058-4>

- [14] Enshaie, A., Robson, C. N., and Edmondson, R. J. (2015). Artificial Intelligence systems as prognostic and predictive tools in ovarian cancer. *Annals of surgical oncology*, 22(12), 3970-3975. <https://doi.org/10.1245/s10434-015-4475-6>
- [15] Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–118. <https://doi.org/10.1038/nature21056>
- [16] Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, 7, 10. <https://doi.org/10.1057/s41599-020-0494-4>
- [17] Gómez-González, E., Gomez, E., Márquez-Rivas, J., et al. (2020). Artificial intelligence in medicine and healthcare. *arXiv*. <https://doi.org/10.48550/arXiv.2001.09778>
- [18] Han, H., and Liu, W. (2019). The coming era of artificial intelligence in biological data science. *BMC Bioinformatics*, 20, 712. <https://doi.org/10.1186/s12859-019-3225-3>
- [19] Harkema, H., Chapman, W. W., Saul, M., Dellon, E. S., Schoen, R. E., and Mehrotra, A. (2011). Developing a natural Language processing application for measuring the quality of colonoscopy procedures. *Journal of the American Medical Informatics Association*, 18(Supplement_1), i150-i156. <https://doi.org/10.1136/amiajnl-2011-000431>
- [20] Jia, G., Huang, X., Tao, S., Zhang X., Zhao, Y., Wang, H., et al. (2022). Artificial Intelligence based medical image segmentation for 3D printing and naked eye 3D visualization. *Intelligent Medicine*, 2(01), 48-53. <https://doi.org/10.1016/j.imed.2021.04.001>
- [21] Jiang, F., Jiang, Y., Zhi, H., et al. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2, 230–243. <https://doi.org/10.1136/svn-2017-000101>
- [22] Jiang, X., Coffee, M., Bari A., Wang J., Jiang, X. Y. et al. (2020). Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. *Computers, Materials and Continua*, 63(1), 537-551. <https://doi.org/10.32604/cmc.2020.010691>
- [23] Joshi, S., and Patel, U. (2026). Breast Cancer Diagnosis with AI Technologies. *Journal of Multiscale Modelling*. <https://doi.org/10.1142/s1756973726400196>
- [24] Kumar, Y., Koul, A., Singla, R., and Ijaz, M. F. (2023). Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*, 14(7), 8459-8486. <https://doi.org/10.1007/s12652-021-03612-z>

- [25] Lancellotti, C., Cancian, P., Savevski, V., et al. (2021). Artificial intelligence and tissue biomarkers. *Cells*, 10, 787. <https://doi.org/10.3390/cells10040787>
- [26] Li, M., Zhang, H., Chen, B., et al. (2018). Prediction of pKa values for neutral and basic drugs. *Scientific Reports*, 8, 3991. <https://doi.org/10.1038/s41598-018-22332-7>
- [27] Maceachern, S. J., and Forkert, N. D. (2021). Machine learning for precision medicine. *Genome*, 64, 416–425. <https://doi.org/10.1139/gen-2020-0131>
- [28] Mathur, S., and Sutton, J. (2017). Personalized medicine could transform healthcare. *Biomedical Reports*, 7, 3–5. <https://doi.org/10.3892/br.2017.922>
- [29] McCorduck, P. (2004). *Machines who think*. CRC Press.
- [30] Milano, C. E., Hardman, J. A., Plesiu, A., et al. (2014). Simulated electronic health record curriculum: teaching EHR skills and use of the EHR for disease management and prevention. *Academic Medicine*, 89, 399–403. <https://doi.org/10.1097/acm.000000000000149>
- [31] Osareh, A., and Shadgar, B. (2010). Machine learning techniques to diagnose breast cancer.
- [32] Paerhati, Y., Aikebaier A., Dilimulati, D., Baishan, A., Yusufujiang, N., Qiu, X., et al. (2026). Rethinking nature's pharmacy: AI era and natural product drug discover. *Pharmaceuticals*, 19 (2), 301. <https://doi.org/10.3390/ph19020301>
- [33] Paul, D., Sanap, G., Shenoy, S., et al. (2021). Artificial intelligence in drug discovery and development. *Drug Discovery Today*, 26, 80–93. <https://doi.org/10.1016/j.drudis.2020.10.010>
- [34] Peng, F., Chang, W., Xie, J. F., Sun Q., Qiu H. B., Yang, Y. (2019). Ineffectiveness of procalcitonin-guided antibiotic therapy in severely critically ill patients: A meta-analysis. *International Journal of Infectious Diseases*, 85, 158-166. <https://doi.org/10.1016/j.ijid.2019.05.034>.
- [35] Pettersson, E., Lundeberg, J., and Ahmadian, A. (2009). Generations of sequencing technologies. *Genomics*, 93, 105–111. <https://doi.org/10.1016/j.ygeno.2008.10.003>
- [36] Pinaire, J., Azé, J., Bringay, S., and Landais, P. (2017). Patient healthcare trajectory. *Health Information Science and Systems*, 5, 1. <https://doi.org/10.1007/s13755-017-0020-2>
- [37] Reiter, A. M. V., Pantel, J. T., Danyel, M., Horn, D., Ott, C. E., Mensah, M. A. (2024). Validation of 3 Computer-Aided Facial Phenotyping Tools (DeepGestalt, GestaltMatcher, and D-Score): Comparative Diagnostic Accuracy Study. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/42904>.

- [38] Sarker, I. H. (2022). AI-based modeling: Techniques, applications and research issues. *SN Computer Science*, 3, 1–20. <https://doi.org/10.1007/s42979-022-01043-x>
- [39] Sellwood, M. A., Ahmed, M., Segler, M. H. S., and Brown, N. (2018). Artificial intelligence in drug discovery. *Future Medicinal Chemistry*, 10, 2025–2028. <https://doi.org/10.4155/fmc-2018-0212>
- [40] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.
- [41] Vrakas, D., and Vlahavas, I. P. L. (2008). Artificial intelligence for advanced problem solving techniques. IGI Global.
- [42] Williams, A. M., Liu, Y., Regner, K. R., et al. (2018). Artificial intelligence, physiological genomics, and precision medicine. *Physiological Genomics*, 50, 237–243. <https://doi.org/10.1152/physiolgenomics.00119.2017>
- [43] Zhang, C., Zhao, J., Zhu, Z., et al. (2022). Applications of artificial intelligence in myopia. *Frontiers in Medicine*, 9, 840498. <https://doi.org/10.3389/fmed.2022.840498>
- [44] Zhang, J., Han, R., Shao, G., et al. (2022). Artificial intelligence in cardiovascular atherosclerosis imaging. *Journal of Personalized Medicine*, 12, 420. <https://doi.org/10.3390/jpm12030420>
- [45] Zhang, W., Chien, J., Yong, J., and Kuang, R. (2017). Network-based machine learning for precision oncology. *NPJ Precision Oncology*, 1, 1–15. <https://doi.org/10.1038/s41698-017-0029-7>
- [46] Zou, K. H., Li, J. Z., Imperato, J., et al. (2020). Harnessing real-world data for regulatory use. *Journal of Multidisciplinary Healthcare*, 13, 671–679. <https://doi.org/10.2147/JMDH.S262776>

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 3

ROLE OF ARTIFICIAL INTELLIGENCE IN BIOLOGICAL SCIENCES

Manisha

Student, Department of Botany, Government College for Girls, Ludhiana, Punjab, India-141001.

Email: mishasingh920@gmail.com

Abstract

Artificial Intelligence (AI) is causing a revolutionary upswing in biological sciences by making research quicker, more knowledgeable, and more accurate. Contemporary biological research develops massive data through genomics, proteomics, metabolomics etc. Technologies that generate enormous amounts of data are, of course, difficult to analyse using classical methods, but AI gives the right set of tools to handle, analyse and interpret very large datasets comfortably. In biological research, AI assists biologists in elucidating genes, predicting protein structures and understanding the genetic basis of diseases. AI's drug discovery contribution is probably the most significant since it aids in identifying possible drug molecules, predicting drugs, and target interactions. It shortens the research and development process and is cost effective. Besides, AI is beneficial in the area of systems biology, where it helps create models of intricate biological processes as well as interactions inside cells. Within precision medicine, AI constitutes an essential instrument that is used for personalized treatment fabrication and accurate diagnosis by matching genetic data with clinical records. Also, being applied in agriculture, ecology, and microbiology, AI helps to identify diseases, locate species improve crops as well as monitors the environment. Furthermore, graphical interpretations based on AI technologies have made microscopic image analysis and examination of medical scans more legitimate and efficient.

Keywords: genomics, proteomics, metabolomics, intricate, accurate diagnosis

3.1 Introduction

Biological sciences study living organisms and life processes at all levels such as molecular biology to ecology. At the same time, rapid biological data accumulation through genomics, proteomics, imaging, and clinical studies have made traditional methods of data analysis no longer adequate. Artificial Intelligence (AI) is now a solution for handling complicated biological data, recognizing patterns, and performing accurate forecasting. By combining computer science, statistics, and biology, AI helps research, diagnostics, and innovations in life sciences (Mohanani, 2025). Biological sciences are mainly dependent on experimental observations, manual data analyses, and human knowledge to explain the structure and functions of living organisms. In fact, for many years, advances in genetics, molecular biology, physiology, ecology, and medicine were the result of wet, lab experiments, microscopy, statistical calculations, and hypothesis, driven studies. While these traditional methods brought about many famous discoveries like DNA double helix, genetic code, and central metabolism, they were slow, and quite intensive in terms of manual labor and they lacked in their capability to deal with complex biological data (Yetgin, 2025). Before the rise of Artificial Intelligence (AI), the analysis of biological data was mostly a linear and reductionist approach which considered only one gene, protein or pathway at a time. It was considered that genome sequencing took a long time and was very expensive, figuring out the structure of protein would take years of experiment work, and drug discovery was just an extension of trial, and, error. The diagnosis of diseases was mostly a matter of how experienced the doctors were, and usage of agricultural and environmental studies depended heavily on the eye and the past. As biological data became larger and more complicated, these old methods had a hard time keeping up, which often resulted in delays, increased costs, and sometimes even incomplete interpretations. The sudden modernization of technologies such as next, generation sequencing, proteomics, imaging, and various omics, based studies resulted in an enormous amount of biological data (Dai and Shen, 2022).

This was a point after which traditional analytical tools became insufficient. It was at this point that Artificial Intelligence came in as a game changer, providing the capability to handle huge datasets, identify complicated patterns, and come up with predictive models that exceed human potential. AI brought a revolutionary change to biological sciences by combining computational intelligence with biological research (Biswas et al., 2023). Currently, machine learning and deep learning algorithms facilitate automated genome annotation, provide highly accurate protein structure predictions, help in quick drug screening, and allow for early disease diagnosis (Senior et al., 2020; Cheng et al., 2008). Instead of following old methods, AI systems are capable of exploring complex datasets at once and uncovering hidden associations between genes, proteins, metabolites, and environmental factors. The research process has been greatly facilitated, and studies have become more accurate and reproducible. Therefore, the most important function of AI in biological sciences is probably in its power to break through the barriers of the methods used before AI. AI revolutionized biology, which used to be a descriptive and labor-intensive science, into a field driven by data, predicting outcomes, and combining various aspects. Instead of totally replacing human intelligence, AI has been a blessing, a crucial tool in pushing forward modern biological research, medicine, farming, and conserving nature.

3.2 Applications of Artificial Intelligence In The Field of Biology

Artificial Intelligence (AI) is a field of computer science that focuses on giving machines or computer systems the ability to perform activities that typically need human intelligence. Such operations encompass learning from data, reasoning, problem solving, decision making, pattern recognition and language comprehension. AI has infiltrated our daily lives to assist us in making tasks more convenient and speedier. Some of the ways AI has been integrated into smartphones are through face and fingerprint recognition, voice assistants, and camera enhancement features (Ghorbani et al., 2025; Muhammad, 2025). In the world of internet and social media platforms, AI effectively delivers personalized content recommendations and spam filtering. When it comes to online shopping, AI not only suggests products but also supports chatbots and personalizes advertisements

(Waghambare et al., 2024). Navigation applications such as Google Maps harness AI to forecast traffic and come up with the shortest path. The world of entertainment is yet another sector where AI is constantly updating your playlist and movie list as per your taste and also enhances the gaming experience (Sadiku et al., 2025). Besides, AI is infiltrating healthcare through fitness trackers and health apps, banking for fraud detection and chatbots, education for personalized learning, and smart homes for voice, controlled devices. We are unknowingly using AI in many ways throughout our day and it has become such an integral part of our life.

3.3 Role of AI in Genomics and Bioinformatics

The Human Genome Project (HGP) was a worldwide scientific research collaboration, the goal of which was to sequence and map the entire human genome (which consists of about 3 billion base pairs). The project was launched in 1990 and completed in 2003. While the initial phases mostly used conventional computational methods, the final stages and post, HGP analyses were greatly dependent on bioinformatics and AI based techniques (Gannett, 2008).

Artificial Intelligence has significantly contributed to the Human Genome Project mainly by making the analysis and interpretation of the huge volume of genetic data resulting from the project much more efficient. The sequencing of the human genome yielded billions of DNA base pairs in the form of millions of small fragments. The utilization of AI based and sophisticated bioinformatics algorithms targeted at the problem of the accurate matching of these fragments by figuring out the overlaps and subsequently combining them into long, continuous DNA sequences. This assembly by the machine took only a fraction of time and was significantly less prone to errors compared to the manual or conventional computational methods which in fact allowed the project to be completed within the timeline. Moreover, AI helped in locating and annotating the genes in the human genome. By using machine learning models, it became possible to separate the DNA coding regions from the non coding ones and also to locate the genes and regulatory elements (Giassa and Alexiou, 2021). The going, on automated annotation disclosed that the human genome is made up of about 20,000-25,000 genes, thus contradicting the early hypotheses and bringing more

profound understanding of the genome structure. Besides, AI tools assisted in identifying genetic variations like single nucleotide polymorphisms which are crucial for studying genetic diversity and disease susceptibility. AI, powered bioinformatics systems also significantly facilitated the handling, storage, and retrieval of the enormous amount of genomic data produced by the Human Genome Project. These systems assisted in the development of publicly accessible genomic databases which enable scientists globally to inspect and extend the HGP data. To sum up, AI revolutionized the Human Genome Project from merely producing data to a significant scientific discovery by making precise genome assembly, gene identification, and future uses in personalized medicine and biomedical research possible.

3.3.1 Role of AI in field of Proteomics and Structural Biology

Recently, artificial intelligence has found its role in the field of proteomics and structural biology through the AI based protein structure prediction tools like AlphaFold3, CASP13 and related methods. Proteins are the main components of the cell playing an important role in enzyme functioning, structural support and controlling gene expression. Their malfunctioning often leads to diseases thus making them crucial targets for drug discovery. The complex nature of the proteome presents a tremendous challenge and predicting their three-dimensional structure has been the greatest challenge in the field of biology. For a long time, the process of predicting the three-dimensional structure of proteins was incredibly time consuming, costly and involved complex experimental techniques such as X-ray crystallography, nuclear magnetic resonance (NMR) and cryo electron microscopy. Now, AI based deep learning models are being used to predict protein structures from amino acid sequences with high accuracy, thus revolutionizing the field of structural biology (Schauperl and Denny, 2022). The latest versions of these AI tools include AlphaFold3 that is capable of predicting not only the shapes of single proteins but also joint structure of complex molecules such as nucleic acids and small drug candidates, which is a critical aspect of understanding cellular mechanisms and developing new drugs. This model provides improved accuracy as compared to previous models.

3.3.2 Role of AI in Drug Discovery and Development

Artificial Intelligence has primarily changed drug discovery and development by making the process much faster and more accurate, as well as significantly cutting the costs. In the past, drug discovery was a time consuming and costly process that is primarily based on the trial and error and screening of thousands of chemical compounds. From the past few years, drug design has evolved by incorporating computational approaches to modify natural products. However, AI has revolutionized this method by leveraging machine learning and deep learning algorithms to scour vast biological and chemical datasets, uncover drug target interactions, and spot potential new drugs at the very start of the process. The advent of AI has revolutionized the drug development by offering the use of advanced computational tools. The intervention of AI has significantly revolutionized the pharmaceutical industry by rationalizing the drug discovery, optimization of the lead compounds and designing the clinical trials (Abbas et al., 2024). Drug discovery basically focuses on the identification of the target molecules responsible for the specific condition and the next step is the identification of an agent to amend this physiological condition. Besides this, AI can work through the billions of chemical compounds in a matter of seconds and it can also predict their binding affinity, toxicity, and stability, thus significantly cutting down on the laboratory workload and time. It is also implemented in drug design where it assists in inventing new molecules of drugs that have the desired properties using generative models (Liebman, 2022).

In the field of vaccines, AI played an important role in the ultra, fast design of mRNA vaccines through the prediction of antigen structures and immune responses. This greatly reduced the development schedule, thus enabling the creation and approval of COVID-19 vaccines in a time frame unmatched by traditional vaccine development processes. AI also played a role in tracing virus mutations and variants, thus facilitating vaccine strategy updates. In short, AI has dramatically changed drug discovery and development by leading to its evolution into a data, driven and predictive discipline. It is playing a pivotal role during the COVID-19 pandemic has proved the enormous potential of AI in the case of global health emergencies (Mohanti et al., 2020). With the continuous evolution of AI technologies, it is anticipated that they will contribute even more

significantly to the development of safer, faster, and more efficient drugs and vaccines in the years to come.

3.3.3 Role of AI in Medical and Clinical Biology

Artificial Intelligence is playing a significant role in medical and clinical biology, helping to diagnose diseases, plan treatments and improve patient care. In the past, medical decisions were mostly based on clinical experience, laboratory tests, and the manual interpretation of medical images and patient records that is quite expensive and time consuming. But now days with the advent of advanced AI tools, medical imaging using computational biomarkers has cut short the expenses along with the time required to read the radiological data and thus proceed towards new treatment response. Implementation of AI using deep learning and machine learning techniques has automatized several tasks that were earlier performed manually thus also reducing the human based errors in the results (Mayorga-Ruiz et al., 2019). By incorporating genomic data, medical histories and other lifestyle factors, AI can predict the patient's response to a particular treatment and thus allowing physicians to go for the most effective therapy that has minimum side effects. For example, in cancer treatment, AI can give information about the tumor's genetic background and then suggests precisely treatment that significantly reduces the chances of remission (Liao et al., 2023). The application of AI in clinical biology is quite extensive as it helps in the analysis of laboratory data like blood tests, biochemical markers and other microbiological findings. AI based automated systems can swiftly interpret test results, identify infections, and ultimately foretell the reason behind the disease. Furthermore, AI contributes towards the regular monitoring of patients using wearable devices, thus making it possible to track vital signs and detect health problems at a very early stage (Ashraf et al., 2025). In drug development, AI enhances the efficiency of clinical trials by identifying suitable candidates, devising trial protocols and analyzing data. It also warns towards negative drug reactions and other risks related to the course of treatment (Panda and Mahopatra, 2024). Overall we can conclude that AI has changed medical and clinical biology into a discipline that relies more and more on data and predictions. By improving the accuracy of diagnoses, allowing for reliable therapies and opening new avenues for patient surveillance, AI has become indispensable

to healthcare professionals and will continue to revolutionize the field of medicine.

3.4 Pros and Cons of AI In Biological Field

Artificial intelligence has greatly impacted the field of biological sciences in several ways like its ability to analyze data at a much faster pace, improved accuracy and the development of advanced prediction models. Nevertheless, AI has its drawbacks and problems which come alongside the benefits.

3.4.1 Pros of AI in Biological Sciences

AI allows the processing of enormous and complicated biological data such as genome sequences, proteomic data and clinical records that are traditionally hard to process in a very short time and that too with spellbound accuracy. It facilitates biological research by performing automatically repetitive tasks such as genome annotation, protein structure prediction, and drug screening with high speed and efficiency (Tiwari et al., 2023; Hein et al., 2025). AI enhances the precision and uniformity thus minimizing the risk of human involved errors in diagnosis and data interpretation. Besides that, it is a great help for predictive and personalized medicine which means that diagnosis can be done according to the person's genetic constitution (Abdallah et al., 2023). Also, AI is capable of lowering the research costs over time and speeding up breakthroughs in drug development, disease diagnosis, agriculture, and environmental biology, among other fields.

3.4.2 Cons of AI in Biological Sciences

Even though AI has got ample of advantages but its success is generally based on the quality, accuracy, range and comprehensiveness of the data provided to it. If the data is biased, incomplete, or contains wrong information, AI can easily give out incorrect or even misleading results. Moreover, the design, training, and updating of advanced AI models require not only powerful computers but also professionals with deep technical and domain knowledge, which might be beyond the reach of some research institutes. It is also quite common that AI models sometimes serve as black boxes as they don't give any explanation that why or how they reach a particular decision. Most importantly, several other issues are arising concerning ethics and

privacy especially when genetic, medical, or other sensitive data are involved. On top of that, over dependence on AI could limit the contribution of human skills and judgment in biological studies.

3.5 Future Prospects of AI In Biology

Artificial Intelligence holds immense potential to revolutionize biology in the coming years through faster, more accurate and predictive technology. One of the main areas where AI will have a significant impact is genomics and personalized medicine. It will be instrumental in predicting disease risk and supporting advanced gene editing techniques like CRISPR (Patel et al., 2025). Moreover, application of AI in drug discovery and vaccine development will drastically reduce cost and time by accurately identifying drug targets and designing new molecules efficiently. Besides playing role in drug discovery, AI is expected to revolutionize proteomics and structural biology, where it will contribute in the enhancement of protein structure and their interaction studies. In medical and clinical biology, early disease detection and patient monitoring will be greatly improved with the help of AI (Dixit et al., 2024). It will also benefit precision treatment significantly. Furthermore, AI is expected to make a remarkable contribution in the field of agriculture and environmental biology by enhancing crop productivity, conserving biodiversity, and tackling the issue of climate change. Summarizing, AI is going to revolutionize biology turning it into a more data driven and preventive science, thus serving as an assistant to biologists and healthcare professionals.

3.6 Conclusion

Artificial Intelligence has been the most important technologies reshaping the field of biological sciences. It has profoundly changed the way of generating, analyzing and applying the biological data. In the past, biological research is mainly dependent on manual experimentation, traditional statistical methods along with human interpretation, which were not only time consuming but also involved handling of complex and large datasets. With the introduction of AI, the above-mentioned limitations are tremendously reduced as AI based technologies provide rapid data processing, accurate pattern recognition and predictive analysis in different biological disciplines. AI has significantly contributed to the field of genomics

and bioinformatics. This is evidenced by major projects such as the Human Genome Project, where AI based tools were instrumental in genome assembly, gene annotation and identification of genetic variations. In proteomics and structural biology, AI applications are playing their part in predicting protein structures and studying their vast interactions. Similarly, AI has revolutionized drug discovery and development methods by identifying drug targets and designing new molecules. Moreover, in medical and clinical biology, AI has brought about improvements in disease diagnosis, personalized medicine, patient monitoring, and clinical decision, thus facilitating better health outcomes. The technology's utilization in agriculture, environmental biology, and even everyday life demonstrates its wide-ranging impact on biological systems and human welfare. Though AI provides a multitude of benefits such as fastness, correctness, and efficiency, the challenges associated with the quality of data, moral issues, and excessive reliance on the technology have to be dealt thoroughly. Generally, the use of Artificial Intelligence has made a great impact on biology by turning it from a descriptive and experimental science into a data, driven, integrative and predictive discipline. If only enhancing human intellect instead of substituting it, AI is likely to gain more and more significance in the field of biological research, medical services, and the implementation of measures for a sustainable development in the years to come.

References

- [1] Abbas, M. K. G., Rassam, A., Karamshahi, F., Abunora, R., Abouseada, M.(2024). The Role of AI in Drug Discovery. *Chemical Biology*, 25, e202300816. <https://doi.org/10.1002/cbic.202300816>.
- [2] Abdallah S., Sharifa M., I. K. H. Almadhoun M., et al. (2023). The Impact of Artificial Intelligence on Optimizing Diagnosis and Treatment Plans for Rare Genetic Disorders. *Cureus*, 15(10): e46860. <https://doi.org/0.7759/cureus.46860>
- [3] Ashraf, A. A. Rai, S., Alva, S., Alva, P. D., and Naresh, S. (2025). Revolutionizing clinical laboratories: The impact of artificial intelligence in diagnostics and patient care. *Diagnostic Microbiology and Infectious Disease*, 111(4), 116728. <https://doi.org/10.1016/j.diagmicrobio.2025.116728>.
- [4] Biswas, A., Kumari, A., Gaikwad, D. S., and Pandey, D. K. (2023). Revolutionizing biological science: The synergy of genomics in health, bioinformatics, agriculture, and artificial intelligence. *OMICS: A*

- Journal of Integrative Biology*, 27(12), 550-569. <https://doi.org/10.1089/omi.2023.0197>.
- [5] Cheng, J., Tegge, A. N., and Baldi, P. (2008). Machine learning methods for protein structure prediction. *IEEE reviews in biomedical engineering*, 1, 41-49. <https://doi.org/10.1109/rbme.2008.2008239>.
- [6] Dai, X., and Shen, L. (2022). Advances and trends in omics technology development. *Frontiers in medicine*, 9, 911861. <https://doi.org/10.3389/fmed.2022.911861>.
- [7] Dixit, S., Kumar A., Srinivasan K., Vincent P. M. D. R., and Ramu Krishnan N. (2024). Advancing genome editing with artificial intelligence: opportunities, challenges, and future directions. *Frontiers in Bioengineering and Biotechnology*, 11, 1335901.
- [8] Gannett, L. (2008). The human genome project.
- [9] Ghorbani, Z., Safari, F., Sabouri, Z., and Momtaz, A. S. (2025). The AI Revolution in Smartphones: How Devices Learn, Adapt, and Assist. In *Development Engineering Conferences Center Articles Database*, 2(9).
- [10] Giassa, I. C., and Alexiou, P. (2021). Bioinformatics and machine learning approaches to understand the regulation of mobile genetic elements. *Biology*, 10(9), 896. <https://doi.org/10.3390/biology10090896>
- [11] Hein, Z. M., Guruparan, D., Okunsai, B., Che Mohd Nassir, C. M. N., Ramli, M. D. C., and Kumar, S. (2025). AI and Machine Learning in Biology: From Genes to Proteins. *Biology*, 14(10), 1453. <https://doi.org/10.3390/biology14101453>
- [12] Liao, J., Li, X., Gan, Y., Han, S., Rong, P., Wang, W., Li, W., and Zhou, L. (2023). Artificial intelligence assists precision medicine in cancer treatment. *Frontiers in Oncology*, 12. <https://doi.org/10.3389/fonc.2022.998222>
- [13] Liebman, M. (2022). The Role of Artificial Intelligence in Drug Discovery and Development. *Chemistry International*, 44(1), 16-19. <https://doi.org/10.1515/ci-2022-0105>
- [14] Mayorga-Ruiz, I., Jiménez-Pastor, A., Fos-Guarinos, B., Lopez-Gonzalez, R., García-Castro, F., and Alberich-Bayarri, Á. (2019). The Role of AI in Clinical Trials. In: Ranschaert, E., Morozov, S., Algra, P. (eds) *Artificial Intelligence in Medical Imaging*. Springer, Cham. https://doi.org/10.1007/978-3-319-94878-2_16
- [15] Mohanan, P. V. (Ed.). (2025). *Artificial intelligence and biological sciences*. CRC Press, Taylor and Francis Group.
- [16] Mohanty, S., Rashid, M. H. A., Mridul, M., Mohanty, C., and Swayamsiddha, S. (2020). Application of Artificial Intelligence in COVID-19 drug repurposing. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(5), 1027-1031. <https://doi.org/10.1016/j.dsx.2020.06.068>.

- [17] Muhammad, A. (2025). The Multimodal Biometric Authentication: Integrating Fingerprints, Face, and Voice Using AI: An AI-Based Approach to Secure Identity Verification Using Fingerprint, Face, and Voice Biometrics. <https://doi.org/10.31224/4808>
- [18] R. (2024). Revolutionizing Patient Safety: Machine Learning and AI for the early detection of adverse drug reactions and drug-induced toxicity. <https://doi.org/10.2174/0129503752308394241013165241>
- [19] Patel, J., Patel, D., and Raval, A. (2025). Artificial Intelligence for Predictive Modeling in CRISPR/Cas9 Gene Editing: a Survey of Methods and Design Strategies. *The Journal of Gene Medicine*, 27(12), e70061. <https://doi.org/10.1002/jgm.70061>.
- [20] Sadiku, M. N., Ajayi, S. A., and Sadiku, J. O. (2025). Artificial Intelligence in Media and Entertainment. *International Journal of Scientific and Applied Research (IJSAR)*, eISSN: 2583-0279, 5(5), 1-14. <https://doi.org/10.54756/IJSAR.2025.5.1>
- [21] Schauerl, M. and Denn, R. A. (2022). AI-based protein structure prediction in drug discovery: impacts and challenges. *Journal of Chemical Information and Modeling*, 62(13), 3142-3156. <https://doi.org/10.1021/acs.jcim.2c00026>.
- [22] Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T, et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792), 706-710. <https://doi.org/10.1038/s41586-019-1923-7>
- [23] Tiwari, P. C., Pal, R., Chaudhary, M. J., and Nath, R. (2023). Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges. *Drug Development Research*, 84, 1652–1663. <https://doi.org/10.1002/ddr.22115>
- [24] Waghambare, M. A., Prabhu, S., and Ashok, P. (2024). Artificial intelligence (AI)-Powered chatbots for marketing and online shopping. In *Digital Technologies in Modeling and Management: Insights in Education and Industry* (pp. 21-39). IGI Global Scientific Publishing. [10.4018/978-1-6684-9576-6.ch002](https://doi.org/10.4018/978-1-6684-9576-6.ch002)
- [25] Yetgin, A. (2025). Revolutionizing multiomics analysis with artificial intelligence and data processing. *Quantitative Biology*, 13(3), e70002. <https://doi.org/10.1002/qub2.70002>

Chapter 4

THE PHYSICS OF INTELLIGENCE: COMPUTATIONAL MODELS INSPIRED BY NATURAL LAWS

Amrit Singh^a, Pratibha Sharma^b, Ankur Sharma^b, Amandeep Saroa^{b,*}

^aAssistant Professor, Department of Physics, Sri Guru Teg Bahadur Khalsa College, Sri Anandpur Sahib, Punjab, India-140118.

^bAssistant Professor, Department of Chemistry, Sri Guru Teg Bahadur Khalsa College, Sri Anandpur Sahib, Punjab, India-140118.

*Email: amansaroa87@gmail.com

Abstract

Intelligence, whether biological or artificial, is fundamentally constrained and shaped by physical laws which are governed by matter, energy, and information. The artificial intelligence has conventionally been developed within computational and mathematical frameworks. The growing evidence indicates that intelligent behavior arises from physical processes such as; thermodynamic parameter, nonlinear dynamics, statistical physics, and efficient information encoding. This chapter examines intelligence as a physical phenomenon and explores how computational models encouraged by laws of nature provide a unifying framework for understanding learning, adaptation, and decision-making. By integrating concepts from various fields such as; thermodynamics, statistical mechanics, information theory, and embodied cognition, the chapter shows how approaches grounded in physical principles can substitute the development of intelligent systems that are more effective, easier to interpret, and environmentally sustainable. By bridging the basic approach of physics, neuroscience, artificial intelligence, and philosophy, the chapter gives an interdisciplinary viewpoint on the future of

intelligence research.

Keywords: physics of intelligence, thermodynamics, information theory, self-organization, embodied cognition

4.1 Introduction

Intelligence is commonly understood on the basis of ability to learn from experience, process information, and adapt to changing environments. In both biological and artificial systems, intelligence operates in parallel with physical reality. Neural computation in the brain depends on the electrochemical signalling, consumption of metabolic energy and functions under noise and constraints of resources, whereas artificial intelligence systems are similarly limited by hardware, consumption of energy, and costs of data transmission (Landauer, 1961; Shannon, 1948). These observations accentuate that intelligence is not merely an abstract algorithmic property but a physical process governed by fundamental laws of matter, energy, and information.

Recent advances in artificial intelligence have achieved amazing performance across diverse fields, however such systems often require extensive computational resources and energy, in sharp disparity to the efficiency of biological intelligence (Laughlin, 2005; Friston, 2010). This disparity has transformed our interest in understanding intelligence through the lens of physics, based on principles of thermodynamics, statistical mechanics, nonlinear dynamics, and information theory, which provide valuable insights into learning and adaptation (Jaynes, 1957; Bialek, 2012). By using intelligence in natural laws, researchers can better explain how complex cognitive behavior arises from simple physical interactions and can develop computational models that are more efficient, robust, and interpretable. In this chapter, we adopt this physics-based viewpoint to explore the foundations and future directions of intelligent systems.

4.2 Intelligence as A Physical Phenomenon

Information processing is intimated with physical reality. According to information theory and statistical physics, the storage, transmission and erasing information requires energy and is subject to fundamental limits (Shannon, 1948; Cover and Thomas, 2006).

Landauer's principle establishes that even the removal of a single bit of information has a minimum thermodynamic cost, directly linked the computation to energy dissipation (Landauer, 1961). Biological intelligence demonstrates how complex information processing can occur while functioning close to these physical limits. In the case of neural signaling electrochemical processes are involved, whereas learning reflects structural shake up of synaptic connections (Kandel et al., 2021). From this viewpoint, intelligence is not only a purely symbolic abstraction but it arises through continuous interaction between matter, energy, and information.

The figure 4.1 shows intelligence as a growing process that arises from the dynamic interaction between matter, energy, and information. At the center, an "Intelligent System" (biological brain or artificial system) is shown as a layered structure. Surrounding it are three interconnected domains:

- **Matter:** Neural tissue, hardware substrates, synaptic structures, or computational architecture
- **Energy:** Metabolic energy, electrical power, thermodynamic costs of computation
- **Information:** Sensory inputs, internal representations, learning signals

Bidirectional arrows connect all the three domains, highlighting that information processing requires energy, energy operates through physical matter, and matter encodes information. An outer boundary represents physical constraints such as thermodynamics, noise, and finite resources.

4.3 Thermodynamics And The Cost of Intelligence

Thermodynamics plays a major role in shaping intelligent behavior. The second law of thermodynamics suggests that while maintaining internal order of the system, it requires continuous energy exchange with the environment or surroundings. Biological or living systems attain this by transferring entropy, allowing localized order and adaptive behaviour (Schrödinger, 1944; Jaynes, 1957). Learning and decision-making are fundamentally dissipative processes. The principle of free-energy proposes that intelligent agents which acts to minimize disbelief by maintaining internal predictive models of their environment (Friston, 2010; Friston et al., 2006). This

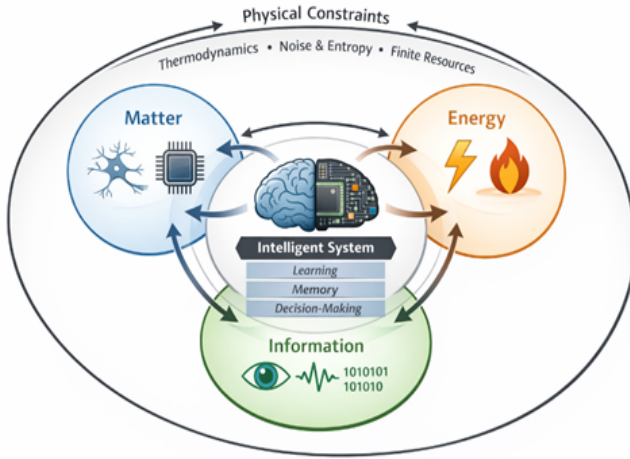


Figure 4.1: Conceptual representation of intelligence as an emergent phenomenon arising from the coupled interaction of matter, energy and information under physical constraints.

framework amalgamates the perception, action, and learning within a single thermodynamic design and has prejudiced computational neuroscience and artificial intelligence. To inspire energy-aware computational approaches, biological systems appear to operate near theoretical efficiency limits (Mead, 2020; Indiveri and Liu, 2015). Further, these approaches object to reduce consumption of energy while maintaining adaptability and performance of the system.

4.4 Statistical Mechanics And Learning Dynamics

Statistical mechanics provides influential tools for understanding the complex collective behavior of the systems. In both biological and artificial neural networks, there are many interacting units whose global behavior cannot be reduced to individual components (Hopfield, 1982; Hinton and Sejnowski, 1986). Learning can be construed as the reshaping of an energy landscape, where attracting states correspond to stored memories or learned solutions. Hopfield networks demonstrated how associative memory can emerge from collective dynamics analogous to spin systems in physics (Hopfield, 1982). These ideas help to explain robustness, generalization, and gradual convergence in learning algorithms (Mehta et al., 2014). The overfitting, generalization, drawing parallels between optimal

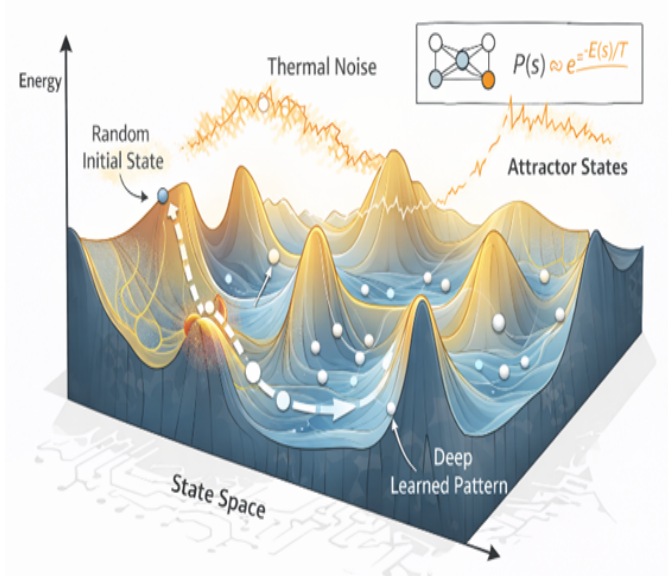


Figure 4.2: Schematic representation of learning dynamics as evolution over an energy landscape.

learning regions and critical phenomena near phase transitions (Bialek, 2012) insights are also provided by statistical mechanics.

To illustrate the concept of statistical mechanics, the figure 4.2 shows learning as motion over a multidimensional energy landscape. The Valleys represent stable attractor states corresponding to memories, whereas hills and barriers represent unstable or high-energy configurations. A trajectory symbolizes that how a learning system evolves from random initial states towards deeper minima. Thermal noise shows transitions between shallow minima, which prevents premature convergence and promoting generalization. Inset indicates analogy to the spin systems, Boltzmann distributions, or Hopfield networks.

4.5 Nonlinear Dynamics and Self-Organization

Intelligent systems mainly operate in nonlinear regions, where small perturbations can produce large-scale effects. Nonlinear dynamics and chaos theory simply explain how complex behavior can emerge from simple interactions (Bak et al., 1987). Experimental evidence suggests that neural systems operate near criticality, which means

it operates in a regime that maximizes the information processing capacity and responsiveness (Beggs and Plenz, 2003). Computational models based on this principle, exploit nonlinear dynamics to process temporal information efficiently (Maass et al., 2002). The principles based on this concept are such as reservoir computing and recurrent neural networks. Further, self-organization demonstrates how intelligence can arise without centralized control. Swarm intelligence and collective robotics demonstrate that simple local interaction rules can lead to complex, adaptive global behavior (Mitchell, 2009; Gershenson, 2013).

4.6 Information Theory And Representation

Information theory provides a rigorous quantitative base for understanding intelligence as a process of efficient representation and prediction under the effect of physical constraints. The intelligence involves reduction of uncertainty about the environment while retaining relevant information necessary for decision-making and action. The fundamental limits on information transmission, storage, and noise tolerance, recognized by Shannon's formulation of information theory demonstrated that all intelligent systems operate within finite capacity bounds (Shannon, 1948; Cover and Thomas, 2006). These limits are mainly significant for biological systems, where energy efficiency and reliability are critical. Biological intelligence attains notable efficiency by accepting compact and sparse representations of sensory inputs. Neural systems encode information to minimize redundancy and retaining features that are important for prediction and survival (Rieke et al., 1997; Bialek, 2012). Such encoding strategies offer valuable design principles for artificial systems and to reduce metabolic cost and enhance robustness against noise. In machine learning, these ideas have inspired representation learning methods, which include autoencoders, sparse coding, and dimensionality reduction techniques, aim to extract useful or meaningful low-dimensional structure from high-dimensional data.

The principle of information bottleneck authenticates this trade-off between compression and significance. This can be done by enclosing learning as an optimization problem which balances preservation of information with predictive accuracy (Tishby and Zaslavsky, 2015). From this viewpoint, intelligent systems learn representations that delete irrelevant details while preserving information essential

for future inference. By grounding representation learning in information-based principles, we can better understand generalization, robustness, and interpretability in both natural and artificial intelligence. This reinforces the view point that intelligence is fundamentally constrained and shaped by physical laws.

4.7 Physics-Inspired Computational Models

To design algorithms and architectures that are efficient, robust, and adaptive the physics-inspired computational models will generate directly from fundamental laws. Many foundational approaches in artificial intelligence are similar to that of physical processes, particularly associated with stability and collective dynamics. For example, simulated annealing is clearly based on thermodynamic cooling, where a system gradually explores its solution space and avoids local optima by probabilistically accepting higher-energy states at raised “temperatures” before converging toward a global minimum (Jaynes, 1957). Similarly, Boltzmann machines and related stochastic neural networks are grounded in statistical mechanics, modelling learning as the minimization of an energy function that encodes probabilistic relationships among variables (Hinton and Sejnowski, 1986; Hopfield, 1982).

The direct combination of physics and intelligence through neural computation in hardware rather than software is represented by neuromorphic computing. These systems check the analog signal processing, parallelism, and local memory. Further, it also checks it with closely resembling biological neural dynamics while attaining significant improvements in energy efficiency (Mead, 2020; Indiveri and Liu, 2015). Such architectures are particularly important for real-time and edge computing applications, where power constraints are critical. Autonomously, quantum-inspired computational models explore how superposition and high-dimensional state spaces can improve optimization and learning even when implemented on classical hardware (Wolfram, 2002). These models inspired by physics accentuate that intelligence can be achieved not only through algorithmic complexity but also through alignment with physical principles. By incorporating computation systems within natural dynamics these approaches propose scalable pathways for more sustainable and interpretable artificial intelligence systems.

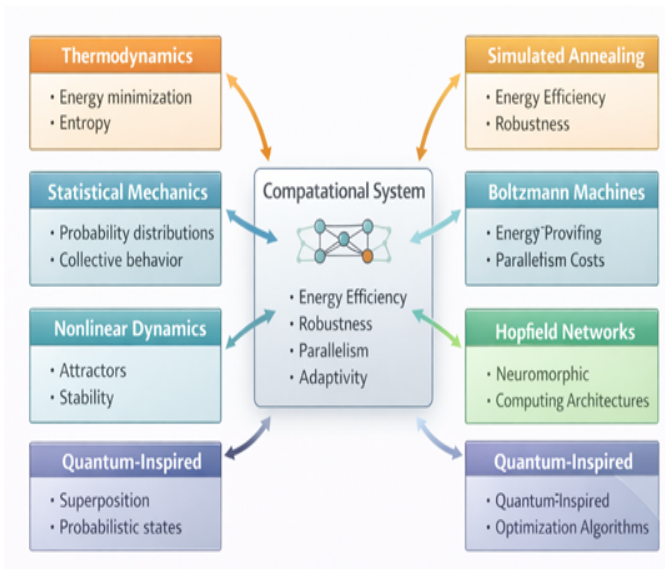


Figure 4.3: Conceptual mapping of physics-inspired principles to computational intelligence models.

The figure 4.3 shows a structured correspondence between fundamental physical principles and related computational intelligence models. On the left, core physics concepts are shown, including: Thermodynamics (energy minimization, entropy), Statistical Mechanics (probability distributions, collective behavior), Nonlinear Dynamics (attractors, stability), Quantum-Inspired Principles (superposition, probabilistic states). On the right, these principles are linked via arrows to computational implementations such as: Simulated Annealing, Boltzmann Machines, Hopfield Networks, Neuromorphic Computing Architectures, Quantum-inspired Optimization Algorithms. At the centre, a schematic “Computational System” highlights shared characteristics such as energy efficiency, robustness, parallelism, and adaptivity. Color coding distinguishes biological inspiration, algorithmic models, and hardware implementations.

4.8 Embodied Intelligence And Physical Interaction

Embodied intelligence highlights that intelligent behavior does not arise from internal computation solely, but arises through continuous interaction between an agent’s body, its environment, and physical laws. The embodied cognition recognizes that perception, action,

and learning are deeply shaped by physical constraints few are gravity, friction, inertia, and energy dissipation which is not present in traditional symbolic or purely algorithmic models (Brooks, 1991; Pfeifer and Bongard, 2007). Cognition is inseparable from sensory motor experience in biological systems, where the body plays an active role in shaping how information is acquired and processed.

Research in robotics provides convincing evidence for the benefits of embodiment. Simple robotic systems that use natural mechanical dynamics can achieve stable and adaptive behavior with very fewer minimal control and computation. By incorporating physical properties of materials and morphology these systems offload some computational burden onto the body itself and further reducing the need for complex algorithms (Pfeifer and Bongard, 2007). This principle referred to as morphological computation and highlights how physical structure can contribute directly to intelligent behavior. Embodied intelligence also aligns with physics-based intelligence by concentrating on energy efficiency and real-world adaptableness. The physically located agents must balance perception, action and energy use in dynamic environments and further reinforcing the role of thermodynamic and informational constraints. Incorporating embodiment into artificial intelligence encourages the development of systems that are more robust, context-aware, and capable of learning through real-world interaction. Ultimately, embodied intelligence joins the physical dynamics and cognitive processes underscoring that intelligence is not only computed but ratified through physical engagement with the world.

4.9 Challenges And Open Questions

Despite the theoretical strength of physics-inspired approaches to intelligence, many challenges remain unresolved. One major difficulty is to translate abstract physical principles into scalable and practical computational models. Further, the concepts such as energy minimization, criticality and information compression offer powerful metaphors. The direct implementation of these in large-scale artificial systems mainly requires simplifying assumptions that may limit realism or performance (Zador, 2019). Balancing physical interpretation ability with computational efficiency remains an open problem. Another challenge concerns about integration with modern data-driven approaches. Deep learning systems

achieve remarkable accuracy but often operate as opaque “black boxes” whereas physics-inspired models emphasize transparency and theoretical grounding. Developing hybrid frameworks that combine physical constraints with flexible learning architectures is an active area of research yet no universally accepted methodology has emerged (Mehta et al., 2014). Further, ensuring robustness across noisy real-world environments continues to pose difficulties particularly for embodied and autonomous systems.

The consumption of energy and sustainability of system also raise critical issues. As artificial intelligence systems grow in large scale and their environmental and economic costs also increase further, underscoring the need for energy-efficient and physically grounded computation (Mead, 2020). Ethical questions complicate the landscape including responsible deployment and societal impact of intelligent systems (Valiant, 2013). Addressing these challenges will require sustained interdisciplinary collaboration across other fields such as; physics, computer science, neuroscience and ethics. Finally, resolving these open questions is necessary for realizing intelligence that is not only powerful but also transparent, sustainable and aligned with physical reality.

4.10 Future Perspectives

The research based on the physics of intelligence is expected to advance through deeper integration of data-based computational models with physical principles. One major direction is to develop physics-informed artificial intelligence where constraints derived from thermodynamics, statistical mechanics and information theory are explicitly embedded into learning algorithms. Such models have the potential to improve interpretation ability, generalization, and robustness while reducing dependence on large datasets. Advances in neuromorphic and unconventional computing architectures are also reshaping the landscape of intelligent systems. By leveraging analog computation, parallelism and local memory inspired by biological neural systems these technologies offer pathways toward low-power adaptive intelligence suitable for edge and autonomous applications. In parallel hybrid approaches combining classical computation with quantum-inspired concepts may open new possibilities for optimization and learning in high-dimensional spaces.

Another key frontier involves embodied and adaptive intelligence, where learning occurs through continuous interaction with dynamic environments. Integrating physical embodiment with learning systems may enable more resilient and context-aware behavior, narrowing the gap between artificial and natural intelligence. Finally, understanding intelligence as a physical phenomenon has broader philosophical and societal implications, challenging traditional distinctions between natural and artificial cognition. As interdisciplinary collaboration deepens, the physics of intelligence is poised to provide a unifying framework that guides the development of transparent, sustainable, and ethically responsible intelligent systems for the future.

4.11 Conclusion

Through the lens of Physics while viewing intelligence, it offers a powerful and unifying viewpoint that connects computation, energy and information within a common framework. By grounding learning, adaptation, and decision-making in thermodynamics, statistical mechanics, nonlinear dynamics, and information theory intelligence can be understood as an emergent property of interacting physical systems rather than a purely algorithmic construct. This perspective helps to explain the remarkable efficiency and robustness of natural intelligence whereas providing valuable direction for the designing artificial systems. Physics-based computational models, embodied intelligence and energy-efficient architectures validate that aligning computation with natural laws can lead to more interpretable and sustainable intelligent technologies. The interdisciplinary research promises to close gaps between physics, neuroscience, and artificial intelligence, despite the fact that there are still major obstacles in the areas of scalability, integration, and ethical deployment. Ultimately, understanding the physics of intelligence not only advances technological innovation but also deepens our understanding into cognition itself and shaping the development of intelligent systems that are efficient, transparent, and responsibly aligned with the physical world.

References

- [1] Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of $1/f$ noise. *Physical Review Letters*, 59(4), 381–384. <https://doi.org/10.1103/PhysRevLett.59.381>

- [2] Beggs, J. M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23(35), 11167–11177.
- [3] Bialek, W. (2012). *Biophysics: Searching for principles*. Princeton University Press.
- [4] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [5] Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1–3), 139–159. [https://doi.org/10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
- [6] Cover, T. M., and Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley.
- [7] Dayan, P., and Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.
- [8] Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- [9] Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology–Paris*, 100(1–3), 70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001>
- [10] Gershenson, C. (2013). The implications of interactions for science and philosophy. *Foundations of Science*, 18(4), 781–790. <https://doi.org/10.1007/s10699-012-9305-8>
- [11] Hinton, G. E., and Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart and J. L. McClelland (Eds.), *Parallel distributed processing*, 1, 282–317. MIT Press.
- [12] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8), 2554–2558. <https://doi.org/10.1073/pnas.79.8.2554>
- [13] Indiveri, G., and Liu, S. -C. (2015). Memory and information processing in neuromorphic systems. *Proceedings of the IEEE*, 103(8), 1379–1397. <https://doi.org/10.1109/JPROC.2015.2444094>
- [14] Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4), 620–630. <https://doi.org/10.1103/PhysRev.106.620>
- [15] Kandel, E. R., Koester, J. D., Mack, S. H., and Siegelbaum, S. A. (2021). *Principles of neural science* (6th ed.). McGraw-Hill.
- [16] Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191. <https://doi.org/10.1147/rd.53.0183>
- [17] Laughlin, R. B. (2005). *A different universe: Reinventing physics from the bottom down*. Basic Books.

- [18] Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation. *Neural Computation*, 14(11), 2531–2560. <https://doi.org/10.1162/089976602760407955>
- [19] Mead, C. (2020). How we created neuromorphic engineering. *Nature Electronics*, 3(7), 434–435. <https://doi.org/10.1038/s41928-020-0448-2>
- [20] Mehta, P., Bukov, M., Wang, C. H., et al. (2014). A high-bias, low-variance introduction to machine learning for physicists. *Physics Reports*, 810, 1–124. <https://doi.org/10.1016/j.physrep.2019.03.001>
- [21] Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.
- [22] Pfeifer, R., and Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. MIT Press.
- [23] Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1997). *Spikes: Exploring the neural code*. MIT Press.
- [24] Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge University Press.
- [25] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- [26] Still, S., Sivak, D. A., Bell, A. J., and Crooks, G. E. (2012). Thermodynamics of prediction. *Physical Review Letters*, 109(12), 120604. <https://doi.org/10.1103/PhysRevLett.109.120604>
- [27] Tishby, N., and Zaslavsky, N. (2015). Deep learning and the information bottleneck principle. In *Proceedings of the IEEE Information Theory Workshop* (pp. 1–5).
- [28] Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215(3), 216–242. <https://doi.org/10.2307/25470707>
- [29] Valiant, L. G. (2013). *Probably approximately correct: Nature's algorithms for learning and prospering in a complex world*. Basic Books.
- [30] Von Neumann, J. (1958). *The computer and the brain*. Yale University Press.
- [31] Wolfram, S. (2002). *A new kind of science*. Wolfram Media.
- [32] Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications*, 10, 4617. <https://doi.org/10.1038/s41467-019-11786-6>

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 5

ARTIFICIAL INTELLIGENCE IN CHEMICAL SCIENCES: TRANSFORMING DISCOVERY, DESIGN AND DATA

Parul Sharma^a & Ramandeep Kaur^b

^aDepartment of Chemistry, Oklahoma State University, Stillwater, Oklahoma 74075, USA.

^bDepartment of Chemistry, College of Basic Sciences and Humanities, Punjab Agricultural University, Ludhiana, India-141004.

Email: ramansaini@pau.edu

Abstract

Artificial Intelligence (AI) has made huge progress in all areas of science. One such area is the field of chemical sciences, where there has been an advent of new data-driven strategies in contrast to experiment-based ones. With the use of artificial intelligence in chemical sciences and computational chemistry, a methodology has been developed to address longstanding issues that required repeated reaction setups and observations of their effects/results. The objective of this chapter is to study and compile the developing AI methodologies in chemical sciences, particularly those concerning machine learning, the effect of AI in pharmaceutical sciences, and the application of deep learning in the various fields of chemistry and biochemistry, such as synthesis route design, drug discovery, workflow, designing materials, and analytical chemistry. This would include a study of the impact of artificial intelligence modeling on reaction optimization and prediction of reaction feasibility, as well as an analysis of its applications and limitations in chemical reactions across agriculture and pharmaceuticals, among others.

Keywords: machine learning, chemistry, drug discovery, chemical synthesis, analytical chemistry

5.1 Introduction

Artificial intelligence (AI) is a field of computer science that deals with creating systems capable of performing human activities like learning, reasoning, decision-making, and problem-solving. Through the imitation of cognition, artificial intelligence has contributed immensely to the progress of technology today (Ahmad et al., 2024). It is used for efficient collection, processing, and analysis of big data, which results in cost-saving solutions and improved precision.

AI is based on several methodologies, such as machine learning (ML) and deep learning (DL). ML involves a learning algorithm that allows an agent to optimize the output function based on the acquired experience (Amodei et al., 2016). This type of learning is used in various applications, including natural language processing, recognition of patterns, planning strategy, and making predictions. Another advantage of AI is that it is highly adaptable, allowing models to continuously learn and make adjustments to outputs based on new input data (Bhide and Patil, 2025).

As one of the main branches of science, chemistry serves as the basis for some vital industries, such as the manufacture of drugs and medical equipment, materials science, energy generation, agriculture, and the management of environmental processes (Bishop and Nasrabadi, 2006). Traditionally, chemical research involved human intuition and extensive experimentation, sometimes taking months and years to produce a result. Even though it was successful at delivering results, this approach was very time-consuming due to the large amount of available data related to chemical compounds and reactions. AI may be considered a tool that can help to solve these issues (Bonastre et al., 2001; Cai et al., 2020).

The recent studies have proven that AI can be effectively used to explore and analyze complex chemical data, revealing structure–property relationships, predicting the outcomes of certain chemical reactions, and designing novel chemical structures (Cardoso Rial, 2024; Coley et al., 2017). Therefore, it can serve to speed up the research process in this field, providing more efficient solutions. At present, artificial intelligence is used in various stages of chemical research, including molecular design, prediction of physicochemical and electronic properties, reaction optimization, retrosynthesis

analysis, catalysis, and materials discovery (Chen et al., 2018; Chen et al., 2024).

One of the areas in which the application of AI is especially valuable is drug discovery, since this process is usually costly and time-consuming, involving many different objectives that should be accomplished (Dana et al., 2018). AI algorithms are currently being developed to identify bioactive compounds, predict drug–target interaction and structure–activity relationships, and generate optimized leads. In addition, it is applied to plan organic syntheses, analyze environmental chemistry, and optimize wastewater treatment and pollution control operations (de Almeida et al., 2019; De Carlo et al., 2024).

In connection with the above-discussed information, it could be stated that the synergy between AI and other digital technologies, such as automated labs, Internet of Things, digital twins, blockchain, and nanotechnology, has significantly transformed the landscape of chemistry and the chemical industry. This chapter will provide a detailed overview of the use of AI in chemistry.

5.2 Literature Review

Artificial Intelligence (AI) in the field of chemistry has evolved through a series of decades. Initial research in the 1970s and 1980s used expert systems based on rules, where manual encoding of chemical knowledge was applied to retrosynthesis and predicting reactions, but such systems were not scalable (de Almeida et al., 2019). However, the significant change that took place in the past two decades is characterized by the development of powerful computers, big chemical datasets, and new ML techniques like deep learning (de Almeida et al., 2019). With the introduction of the latter, AI became capable of transitioning from heuristics to more complex predictive and generative models (Debus et al., 2021). The first ML models that were developed in chemistry concerned Quantitative Structure Activity Relationships (QSAR) and Quantitative Structure Property Relationships (QSPR), using conventional algorithms like support vector machines, decision trees, and random forests (Duan et al., 2023; Floridi et al., 2018). In recent years, more sophisticated neural networks, such as convolutional, recurrent, and graph-based neural

networks, have been developed; out of all, the latter has proven to be highly effective due to representing molecules by graphs.

5.2.1 Artificial Intelligence in Retrosynthesis

One of the most fundamental techniques in the field of organic chemistry is retrosynthesis, which consists of disassembling an objective compound into simpler structures till simpler starting materials are achieved (Goh et al., 2017). This process gets really difficult when one attempts to design reactions for complicated organic compounds comprising various functional groups and stereocenters. The use of artificial intelligence has revolutionized this technique, where algorithms learn from gigantic reaction databases compiled from scientific publications on the subject (Gioiello et al., 2020). While initially, researchers created algorithms that used predefined reaction types encoded into the program, the application range of such algorithms was very limited. In modern approaches based on deep learning, the algorithm learns from large reaction databases (Hautier et al., 2011). For instance, neural networks are used for generating SMILES strings corresponding to the reaction products (Hirani et al., 2024).

5.2.2 AI in Reaction Outcome Prediction

The prediction of reaction outcomes is yet another domain where AI has been performing impressively. Reaction templates are initially used to predict potential products, after which machine learning algorithms score them based on yield probability (Himanen et al., 2019). The identification of dominant products among several possibilities is facilitated by the use of SoftMax scoring. This results in greater efficiency in reactions since less trial and error is required, and more is made out of reaction history data (Hou et al., 2024; Huang et al., 2023).

5.2.3 AI in Molecular Design and Property Prediction

Molecular design through artificial intelligence has become prevalent in material science, organic electronics, and photovoltaic studies (Jin et al., 2025). The machine learning models have the ability to predict the properties of molecules and understand the relationship between the structures and the properties without any knowledge of the chemistry behind them (Jumper et al., 2021). With deep

learning algorithms, features can be extracted from images, graphs, or even descriptors, which help predict the properties of molecules, including their frontier molecular orbitals, solubility, melting point, and electronic behavior (Klambauer et al., 2019).

5.2.4 AI in Environmental and Wastewater Treatment

The application of AI techniques is also useful in wastewater treatment, which helps in modeling, forecasting, and optimizing the complicated biological and chemical processes occurring therein (Klenam et al., 2023). Machine Learning techniques are used to monitor the operational data of treatment facilities in order to predict the effectiveness of pollutants' removal, especially for parameters like chemical oxygen demand (COD), Nitrogen, and Phosphorus concentrations. Some of the methods that can be applied include artificial neural networks, neuro-fuzzy controllers, and reinforcement learning techniques such as Q-learning (Korshunova et al., 2021; Kündig, 2006).

5.2.5 AI in Nanotechnology

The synergy between AI and nanotechnology has brought forth new ways to investigate materials, their characteristics, and predictions (LeCun et al., 2015). AI technology is employed in scanning probe microscopy to understand its data, for calculating dielectric constants, finding the thickness of a thin film, and analyzing the morphology of a nanomaterial (Luong and Singh, 2024). Genetic algorithms, neural networks, and density functional theory have all been used extensively in researching nanoparticle aggregates, carbon nanotubes, and their optical properties (Malashin et al., 2024).

5.2.6 Digital Intelligence and Data-Driven Innovation in Modern Chemistry

The swift adoption of digital technologies has revolutionized the way chemical experiments and interpretations are conducted (Mater and Coote, 2019). Whereas artificial intelligence (AI), including its various subdivisions such as machine learning (ML) and deep learning (DL), alongside sophisticated data analysis techniques, has emerged as the primary driver of innovation in recent years, the impact of digital technologies on chemistry cannot be limited to just one domain alone (Miller, 2019).

The major digital technologies influencing contemporary chemistry can be broadly organized as follows:

- **Chemoinformatics and Chemical Data Analytics:** Computerized methods for the organization, analysis, and interpretation of chemical information.
- **Big Data and Chemical Database Design:** Infrastructure design to harmonize heterogeneous data sources for large-scale analysis and interdisciplinary collaboration.
- **Natural Language Processing in Chemical Text Mining:** Automated methods for extracting insights from chemical scientific literature, patents, and database entries.
- **ML Methods for Prediction of Chemical Properties:** Machine learning-based prediction of physicochemical and biological properties for molecular and material design.
- **Deep Learning Techniques for SAR:** Advanced machine learning for predicting structure-activity relationships for drug design.
- **Artificial Intelligence for Quantum Chemistry and Molecular Modeling:** Enhanced computation and improved prediction of results using artificial intelligence for quantum chemical calculations and molecular modeling.
- **AI-based Chemical Reaction Prediction and Optimization:** Approaches for reaction prediction, optimal conditions, and reaction discovery with AI assistance.
- **Synthetic Planning Using Automated and AI-Assisted Approaches:** Automated synthesis pathway generation and refinement for reducing experimental efforts.
- **High Throughput Experimentation Using Artificial Intelligence:** High throughput experimentation involving design, experiment execution, and data analytics using robots.
- **Robotic and Automation in Laboratory Operations:** Use of robotics and automation in laboratory operations for improving the precision and repeatability of laboratory activities.
- **Smart and AI-Enabled Process Control Systems:** Smart and AI-assisted process control to improve efficiency and adaptability.

- **Digital Twinning for Chemical Processes:** Digital twin-based technologies used in simulating and optimizing chemical processes and equipment.
- **Networks of Laboratory Instruments and IoT Technology:** Networks of instruments used in real-time monitoring and analysis of chemical experiments using IoT technology.
- **AI Assisted Spectroscopic Data Interpretation and Method Development:** Intelligent interpretation of spectroscopic data and development of new analysis methods using AI.
- **Virtual and Augmented Laboratory Environments and Tools:** Virtual environment and augmented reality technologies used in laboratories.

Together, these digitally enabled technologies highlight the transformative role of data, automation, and artificial intelligence in modern chemistry (Peplow, 2014). They expand the boundaries of fundamental research, accelerate innovation, and support more efficient and sustainable chemical practices. The following sections analyze each of these technologies in detail, emphasizing their scientific significance, practical applications (Table 5.1), and future prospects (Roy et al., 2024).

Table 5.1: Summary of the possible influence of AI-based technologies on chemical research, including applications, impact, and relative research activity.

Sr. No.	Digital Chemistry Trajectory	Impact	Application
1	Drug discovery	Accelerates drug discovery and development	AI-assisted identification and optimization of drug candidates
2	Big data and integrated data	Enables comprehensive and data-driven analysis	Consolidation of diverse chemical and experimental data sources

Sr. No.	Digital Chemistry Trajectory	Impact	Application
3	Automated laboratory platforms	Improves precision, reproducibility, and throughput	Highly automated experimental systems
4	Integration of lab instruments and IoT	Enhances laboratory efficiency and data analysis	Real-time data acquisition and monitoring
5	Spectroscopy and analytical method development	Improves analytical accuracy and method development	AI-assisted spectroscopic data interpretation
6	Blockchain in chemical supply chains	Increases traceability and supply chain integrity	Secure and transparent supply chain tracking
7	Digital twins	Optimizes chemical processes and system performance	Virtual replicas for real-time simulation and optimization
8	Virtual laboratories and augmented reality	Enhances learning outcomes and accessibility	Digital platforms for training and education
9	Natural language processing (NLP) in chemical space	Streamlines knowledge discovery and data mining	Extraction of chemical knowledge from literature
10	Predictive toxicology	Improves drug safety and risk assessment	AI models for toxicity and safety prediction

Sr. No.	Digital Chemistry Trajectory	Impact	Application
11	Environmental chemistry and sustainability	Promotes sustainable chemical practices	Digital tools supporting green chemistry
12	Machine learning in molecular design	Accelerates compound and material design	Prediction of molecular and material properties
13	Smart process control	Improves efficiency and sustainability	Intelligent control systems for chemical operations
14	Deep learning in structure–activity relationships (SAR)	Enables advanced drug discovery insights	Modeling complex structure–activity relationships
15	High-throughput experimentation (HTE)	Speeds up experimental research workflows	Automated parallel experimentation with AI analysis
16	Digital materials design and materials informatics	Accelerates materials development	Data-driven materials discovery
17	Data-driven chemical reaction optimization	Improves reaction efficiency and yield	Machine learning for reaction condition optimization
18	Automated synthesis planning	Streamlines chemical synthesis planning	AI-based synthetic route design

Sr. No.	Digital Chemistry Trajectory	Impact	Application
19	Chemoinformatics and chemical data analysis	Supports data-centric chemical research	Management and analysis of large chemical datasets

5.3 Molecular Design Using Machine Learning

Machine learning (ML) technology is gaining traction and is at the forefront of molecular design approaches owing to its ability to accurately predict molecular characteristics and rationally design molecules/materials with desired properties (Russell and Norvig, 2011). In essence, this is achieved by training ML algorithms using large chemical databases with diverse molecules whose features include structure, physicochemical, biological, and reactive properties. After training, ML algorithms have the capability to predict how structural changes affect molecular performance and functionality (Schwaller et al., 2019).

The main advantage of ML in this case is learning from historical data involving both good and bad designs of molecules. This process allows making precise predictions about the properties of new substances without conducting numerous and lengthy experiments (Segler et al., 2018). ML can help in finding suitable molecules with appropriate biological behavior and minimize expenditures on development and time required for it (Shahi and Singh, 2024). In materials science, ML helps to develop materials with improved mechanical, electronic, catalytic, and thermal characteristics. It solves the problems associated with expensive experimental research by allowing rapid screening and selection of promising compounds. ML-based molecular design finds applications in a wide range of industries, including pharmaceutical, energy production, electronics, and ecological purposes (Shen et al., 2021).

5.3.1 Smart Control

The concept of intelligent control within the context of chemical operations and machinery relates to the implementation of intelligent control methodologies, which utilize sophisticated computing algorithms, sensor networks, and automation techniques to enhance operational efficiency, safety, and reliability (Singh et al., 2023). Such control systems are capable of real-time monitoring of essential parameters, decision-making based on information, and adjustment of operation settings for maintaining optimum operations. Furthermore, smart control systems allow for remote monitoring and control, thus minimizing reliance on constant human presence at the site (Tshitoyan et al., 2019).

The computational models employed by smart control systems have become much more advanced in recent times. The latest trends indicate that there is an increase in their use along with AI and ML for the prediction of system behavior using past and current process data (Tripathy and Shahi, 2024). These computational algorithms detect regular patterns of behavior through the processing of sensor data gathered from reactors, separation processes, and other support facilities and also make predictions about any deviation from normal system operation (Vamathevan et al., 2019).

The adoption of smart control technologies directly addresses key challenges in chemical manufacturing by enhancing reaction efficiency, improving product yields, and minimizing energy consumption and waste generation (Venkataraman et al., 2025). From a safety and environmental perspective, these systems enable early identification of abnormal events, including equipment malfunctions, leaks, and unintended emissions, allowing for timely corrective interventions (Wang et al., 2022). Moreover, predictive maintenance enabled by smart control systems supports continuous assessment of equipment health, thereby reducing unplanned downtime and maintenance costs. The growing integration of AI-assisted reaction planning with automated synthesis platforms further underscores the role of smart control in enabling flexible, scalable, and sustainable chemical production systems (Wu et al., 2023).

5.3.2 Structure–Activity Relationships (SAR)

Deep Learning-based SAR Analysis utilizes multi-layered neural network models to represent the relationship between structure and biological activity (Xu et al., 2024). It is very effective and highly popular for drug development in modern times, since quick and precise recognition of active molecules is one of the major problems of the drug industry. Due to big and annotated datasets correlating chemical structure and experimentally measured activity, deep learning offers better insight into complicated interactions (Zeng et al., 2022).

Common architecture includes convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are trained on molecular representations such as graphs or simplified molecular-input line-entry system (SMILES) strings (Zhou et al., 2017). Unlike conventional SAR methods that depend on handcrafted molecular descriptors, deep learning models automatically learn hierarchical and non-linear features directly from raw structural data. This data-driven feature extraction allows more faithful capture of subtle structural motifs and long-range dependencies that influence biological activity (Zhou et al., 2025).

Some of the recent works demonstrate the incorporation of deep learning approaches with the identification of potential molecules based on protein-protein interactions, along with drug-target interaction networks for the discovery of leads, especially in pain therapy involving voltage-gated sodium channels (Jung et al., 2021). This approach includes systematic virtual screening, alongside the use of advanced natural language processing and multi-model machine learning techniques for improving lead generation, drug repurposing, and ADMET prediction purposes. On the same note, the employment of multiscale topological Laplacian and artificial intelligence is used in high-throughput molecular modeling of drug addiction for making correct predictions of their activity and toxicity (Molnár and Tamás, 2024). Throughout the whole process of drug discovery, deep learning-based SAR provides help at every stage, ranging from screening, lead generation, to lead optimization and safety assessment (Chen et al., 2025).

5.3.3 AI-Driven High-Throughput Experimentation (HTE)

High-Throughput Experimentation (HTE) is an advanced approach for conducting experiments that involves automation and AI, together with robotics, for the rapid execution of thousands of chemical experiments at once (Tropsha et al., 2024). Through high throughput, HTE allows one to explore large spaces of chemistry very fast and efficiently, with a level of speed not achievable under normal laboratory conditions. Current HTE tools are capable of conducting several hundred to thousands of experiments per day, producing huge and reliable data sets (Tropsha et al., 2024).

Robotic automation forms the backbone of HTE systems, executing repetitive laboratory operations with exceptional precision and reproducibility. Core tasks such as reagent dispensing, sample preparation, reaction condition control, and automated analysis via spectroscopic and chromatographic techniques are performed with minimal human intervention. This automation reduces experimental variability and human error while allowing researchers to devote greater attention to experimental strategy, hypothesis generation, and data interpretation (Bentéjac et al., 2021; Gunasekaran et al., 2021).

AI plays an important role in augmenting the capabilities of HTE by facilitating the process of experiment design, prediction, and analysis of the collected data (Sheikh et al., 2025). Machine learning algorithms can suggest the best combination of reaction conditions or materials based on previously obtained experimental data. After collecting the information, machine learning can quickly find optimal conditions, trends, and even surprises from multidimensional sets of experimental data (Balram et al., 2019; Jia et al., 2023).

HTE has shown great potential, especially in the areas of materials science, catalysis, and pharmaceutical research, due to the existence of multidimensional sets of experimental parameters. In addition, HTE facilitates the development of a positive feedback loop between the processes of experiment planning and execution on the one hand, and computational analysis on the other (Azevedo et al., 2024).

5.3.4 Digital Materials Design and Materials Informatics

The fields of digital materials design and materials informatics constitute an emerging, fast-paced interdisciplinary paradigm that links together materials science and data science/information technology to allow for the discovery, knowledge, and design of novel materials (Dana et al., 2018; Roy et al., 2024). As an information-driven technique, digital materials design employs large sets of information about materials, which include the chemical makeup, microstructure, processing details, and functionality of a wide range of materials, using machine learning, data analysis, and computer modeling (Dana et al., 2018). Through the systematic identification of complicated relationships among structures, properties, and processing conditions, digital materials design makes it possible to rationally engineer new materials with desired functionalities, such as state-of-the-art catalysts (Rana and Bhushan, 2023).

Predictive modeling is essential in this context, and these are typically generated from data collected from experimental studies, theoretical computation, and scientific publications. Predictive modeling allows scientists to accurately predict various characteristics of materials depending on their compositions and processing conditions. Digital materials design, hence, becomes a faster and more effective process than conventional trial-and-error approaches (Xiao-Yong and Hong-Bin, 2009).

Digital Materials Design has proved versatile and applicable in various industries and technologies. In aerospace and automotive fields, it aids in designing lightweight yet strong alloys that have greater strength and durability while remaining resilient against environmental conditions. In the field of energy, digital materials design helps in designing more efficient, stable, and long-lasting batteries, catalytic systems, and energy conversion materials. Similarly, in electronics and photonics, digital materials design aids in designing new semiconducting and functional materials for electronic and photonic devices of the future (Khadhour et al., 2024).

Apart from aiding in developing innovations, digital materials design is useful for creating sustainable materials and processes by using less resource-intensive material formulations and processing methods (Mroz et al., 2025).

5.3.5 Data-Driven Chemical Reaction Optimization

Machine learning (ML) models can be trained on experimental data and used to predict reaction outcomes, optimize reaction conditions, and discover new reactivity patterns (Peplow, 2014). These models make possible more rational and efficient decision-making in synthetic chemistry by analyzing large datasets generated from literature, databases, and experimental records. This process relies heavily on techniques such as Bayesian optimization, reinforcement learning, and neural networks (Xiao-Yong and Hong-Bin, 2009).

ML models learn complex correlations between reaction variables, e.g., reactant structure, catalysts, solvents, temperature, and pressure, to reliably predict performance. This allows *in silico* scouting of reaction space before experiments, leading to a major decrease in time, cost, and material usage. Early state-of-the-art achievements in autonomous synthesis on platforms integrated with robotic experimentation, ML-based planning, and real-time monitoring have demonstrated self-optimizing chemical system analytics (Khadhour et al., 2024). Applied across synthetic chemistry, materials science, and pharmaceutical manufacturing, data-driven optimization improves yield, selectivity, and sustainability while supporting green chemistry principles.

5.3.6 Synthesis Planning

Automated synthesis planning involves the use of artificial intelligence (AI) to aid chemists in synthesizing molecules through efficient and scalable processes (Miller, 2019). In contrast to conventional techniques that depend heavily on expertise and laborious analysis, these methods provide a computer-guided path towards identifying synthesis routes (Peplow, 2014). Driven by machine learning (ML) algorithms, which draw their knowledge from big reaction databases, these systems identify and prioritize several synthesis routes on the basis of parameters like step efficiency, material availability, costs, feasibility, and ecological impact (Roy et al., 2024). This feature makes the tool particularly useful in complicated cases of organic and pharmaceutical synthesis, in which route selection significantly influences development period and manufacturing costs (Russell and Norvig, 2011).

5.3.7 Chemoinformatics and Chemical Data Analysis

Chemoinformatics and chemical data analysis are a diverse set of computational methods for organizing, analyzing, and interpreting complex datasets in chemistry (Shahi and Singh, 2024). Chemoinformatics leverages algorithms, statistical techniques, and dedicated software applications to convert data on chemical structures, properties, and bioactivity into valuable insights. The overarching aim is to turn large-scale chemical data into actionable knowledge that can aid in understanding, prediction, and speeding up discovery (Shen et al., 2021).

Chemoinformatics is really about using computers to understand chemical structures. This means we can study these structures in a way that uses computers (Tshitoyan et al., 2019). We use something called descriptors to learn more about chemicals. These descriptors are like numbers that tell us about the structure, properties, and shape of a chemical. We use these numbers to learn about lots of chemicals at the same time (Tripathy and Shahi, 2024). By looking at these descriptors, we can figure out how the structure of a chemical is related to its properties and how it works. This is something that can be hard to do by doing experiments in a lab. Chemoinformatics and molecular descriptors are tools for understanding chemicals and how they work (Venkataraman et al., 2025).

Chemoinformatics is really useful when we are dealing with a lot of data that's hard to understand. For example, in the process of discovering drugs, chemoinformatics helps us to look at a lot of compounds and figure out which ones might work well against certain targets. This helps us to find the compounds and make them even better (Xu et al., 2024).

In the field of materials science, chemoinformatics is used to design and improve materials by looking at how the molecules are arranged and how this affects the material's performance. It is also used in regulatory chemistry to predict how toxic a chemical might be, how long it will last in the environment, and what will happen to it over time (Klenam et al., 2023; Korshunova et al., 2021; Kündig, 2006).

As chemoinformatics gets better, it is being used more and more with machine learning and artificial intelligence. This makes it even better at predicting what will happen. All of these developments

are helping us to come up with ideas, make better decisions, and use data to improve chemistry in all areas of science and industry. Chemoinformatics is playing a role in this and will continue to be important as it keeps improving.

5.3.8 AI in Quantum Chemistry and Simulations

Quantum Chemistry, which utilizes quantum mechanical techniques, is a method that predicts and studies chemical structures, properties, and behaviors using atomic-molecular interactions. Despite its reliability, conventional quantum chemistry techniques consume vast amounts of computational power when dealing with larger systems.

The application of artificial intelligence in this field has revolutionized the area through increased efficiencies and scalability in quantum simulations (Korshunova et al., 2021; Venkataraman et al., 2025; Wang et al., 2022). Machine learning algorithms using high-quality training data from quantum chemistry calculations help make approximations to quantum mechanical methods, which makes electronic structure calculations faster and more cost-effective. AI-assisted quantum chemistry has found use in drug design, materials science, and catalysis, among other chemical applications, since it requires electronic structure information. This technique enables scientists to conduct faster virtual screenings in drug discovery, design functional materials, and come up with better catalysts for various chemical reactions (Xu et al., 2024).

5.3.9 Broader AI Themes in Chemistry

The twenty influential AI-driven technologies discussed in this manuscript can be organized into seven broader thematic categories reflecting their primary focus and application domains. These themes provide a structured framework for understanding the collective impact of digital technologies on the evolution of chemistry and its transition into the digital era (Korshunova et al., 2021; Kündig, 2006).

5.3.10 Chemical Biology and Digital Medicine:

This theme includes AI-driven drug discovery, predictive toxicology, and deep learning-based structure-activity relationship analysis (Klenam et al., 2023). These technologies highlight the critical role of AI in accelerating therapeutic development, improving chemical

safety assessment, and elucidating complex biochemical interactions. Globally, this theme supports advances in personalized medicine, reduced development costs, and improved healthcare outcomes.

5.3.11 Environmental Technologies:

This category encompasses AI applications in environmental chemistry, sustainability, and predictive toxicology (Balram et al., 2019). It reflects the growing emphasis on leveraging digital tools to assess environmental impact, support green chemistry practices, and protect ecosystems. Globally, this theme aligns with the urgent need for sustainable development and environmental stewardship.

5.3.12 Data Analysis:

Technologies focused on big data integration, natural language processing, chemical reaction optimization, and knowledge extractions are grouped under this theme. Their shared emphasis on data-driven insight underscores the transformative role of AI in chemical analysis, process optimization, and scientific discovery, enabling more informed and efficient research practices worldwide (Gunasekaran et al., 2021; Rana and Bhushan, 2023).

5.3.13 Laboratory and Process Innovations:

This theme includes laboratory automation, IoT-enabled instrumentation, analytical method development, high-throughput experimentation, automated synthesis planning, and smart control systems (Malashin et al., 2024; Mater and Coote, 2019). These technologies collectively modernize laboratory workflows and chemical manufacturing, improving precision, safety, scalability, and efficiency globally (Balram et al., 2019).

5.3.14 Simulation and Modeling:

Encompassing digital twins, virtual laboratories, and AI-enhanced quantum simulations, this theme highlights the growing importance of predictive modeling and virtual experimentation (Tshitoyan et al., 2019; Tripathy and Shahi, 2024). These tools reduce experimental costs, support remote education and research, and democratize access to advanced chemical knowledge and innovation.

5.4 Limitations

Despite a wide range of analytical approaches and data types used in this project, there are numerous limitations inherent in this investigation due to reliance upon publicly accessible literature and database resources. Unpublished, proprietary, and emergent research developments are not included in the nationally accepted definition of the relevant fields of research, therefore introducing the risk of incomplete representation of this area of research. Rapidly changing technology will likely result in some conclusions becoming outdated before new techniques and applications can be reported on. In addition to these limitations, the combination of interpretative and effortful analysis-based approaches also leads to multiple layers of subjectivity during derivative trend determination and technology valuation, which creates possible bias in both trend determination and technology valuation. The limitations discussed here demonstrate the importance of timely, continued, and regular updating, validation, and revision of analytic practices as digital chemistry continues to advance. Continued research and iterative evaluation will be crucial for a current and complete understanding of a field that is changing quickly.

5.5 Conclusion

The progression of chemistry into the modern age is characterized by the rise of digital chemistry, a transformative approach that combines computational techniques, machine learning, and large-scale data analysis to accelerate and refine scientific discovery. This shift introduces strong predictive capabilities, improves efficiency, and enables innovative strategies for addressing complex chemical problems beyond the limits of traditional experimental approaches. Consequently, research in chemistry is moving toward more targeted, hypothesis-driven methods, reducing dependence on serendipitous findings and fostering systematic, efficient, and sustainable innovation.

It is worth noting that many of the digital research pathways discussed in this work significantly overlap, often tackling similar scientific problems through complementary or partially intersecting approaches. Given the novelty and rapid development of this field, it remains challenging to draw clear distinctions between individual

directions. Instead of merging these pathways, they have been presented as originally identified to maintain the richness and diversity of emerging trends. As the field evolves, more precise classification and conceptual clarity are expected to develop.

Looking ahead, the future of chemistry will depend on the integration of traditional and digital practices. Conventional chemistry continues to play a vital role, particularly in experimental validation, education, and direct interaction with materials. At the same time, digital chemistry is set to grow rapidly, offering powerful tools to explore extensive chemical spaces and stimulate innovation. The categorization of AI-driven technologies highlights a broader transition from labor-intensive, empirical experimentation to predictive, data-centered research. Together, these complementary approaches create a synergistic framework that will guide future research directions, improve sustainability, and broaden the impact of chemistry on science and society.

References

- [1] Ahmad, M. I., Veg, E., Joshi, S., et al. (2024). Thiosemicarbazone derivatives in search of potent medicinal agents: QSAR approach (a review). *Russian Journal of General Chemistry*, 94, 2470–2495. <https://doi.org/10.1134/S1070363224090251>
- [2] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., and Mané, D. (2016). Concrete problems in AI safety (arXiv:1606.06565). *arXiv*. <https://doi.org/10.48550/arXiv.1606.06565>
- [3] Azevedo, B. F., Rocha, A. M. A. C., and Pereira, A. I. (2024). Hybrid approaches to optimization and machine learning methods: A systematic literature review. *Machine Learning*, 113(7), 4055–4097. <https://doi.org/10.1007/s10994-023-06467-x>
- [4] Balram, D., Lian, K. -Y., and Sebastian, N. (2019). Air quality warning system based on a localized PM2.5 soft sensor using a novel approach of Bayesian regularized neural network via forward feature selection. *Ecotoxicology and Environmental Safety*, 182, 109386. <https://doi.org/10.1016/j.ecoenv.2019.109386>
- [5] Bentéjac, C., Csörgő, A., and Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937–1967. <https://doi.org/10.1007/s10462-020-09896-5>
- [6] Bhide, S. S., and Patil, S. R. (2025). Artificial intelligence: Role in healthcare. *Journal of Postgraduate Medicine*, 71(2), 57–60. https://doi.org/10.4103/jpgm.jpgm_95_25

- [7] Bishop, C. M., and Nasrabadi, N. M. (2006). Pattern recognition and machine learning. *Springer*.
- [8] Bonastre, A., Ors, R., and Peris, M. (2001). Distributed expert systems as a new tool in analytical chemistry. *TrAC Trends in Analytical Chemistry*, 20(5), 263–271. [https://doi.org/10.1016/S0165-9936\(01\)00058-9](https://doi.org/10.1016/S0165-9936(01)00058-9)
- [9] Cai, J., Chu, X., Xu, K., Li, H., and Wei, J. (2020). Machine learning-driven new material discovery. *Nanoscale Advances*, 2(8), 3115–3130. <https://doi.org/10.1039/D0NA00388C>
- [10] Cardoso Rial, R. (2024). AI in analytical chemistry: Advancements, challenges, and future directions. *Talanta*, 274, 125949. <https://doi.org/10.1016/j.talanta.2024.125949>
- [11] Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., and Blaschke, T. (2018). The rise of deep learning in drug discovery. *Drug Discovery Today*, 23(6), 1241–1250. <https://doi.org/10.1016/j.drudis.2018.01.039>
- [12] Chen, H., Xiang, Q., Hu, J., Ye, M., Yu, C., Cheng, H., and Zhang, L. (2025). Comprehensive exploration of diffusion models in image generation: A survey. *Artificial Intelligence Review*, 58(4), 99. <https://doi.org/10.1007/s10462-025-11110-3>
- [13] Chen, Y., Huang, Z., Yu, J., Wang, H., Qin, Y., Xing, L., and Du, L. (2024). Research progress of Pt-based catalysts toward cathodic oxygen reduction reactions for proton exchange membrane fuel cells. *Catalysts*, 14(9), 569. <https://doi.org/10.3390/catal14090569>
- [14] Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., and Jensen, K. F. (2017). Prediction of organic reaction outcomes using machine learning. *ACS Central Science*, 3(5), 434–443. <https://doi.org/10.1021/acscentsci.7b00064>
- [15] Dana, D., Gadhiya, S. V., St. Surin, L. G., Li, D., Naaz, F., Ali, Q., et al. (2018). Deep learning in drug discovery and medicine: Scratching the surface. *Molecules*, 23(9), 2384. <https://doi.org/10.3390/molecules23092384>
- [16] de Almeida, A. F., Moreira, R., and Rodrigues, T. (2019). Synthetic organic chemistry driven by artificial intelligence. *Nature Reviews Chemistry*, 3, 589–604. <https://doi.org/10.1038/s41570-019-0124-0>
- [17] De Carlo, A., Ronchi, D., Piastra, M., Tosca, E. M., and Magni, P. (2024). Predicting ADMET properties from molecular SMILES: A bottom-up approach using attention-based graph neural networks. *Pharmaceutics*, 16(6), 776. <https://doi.org/10.3390/pharmaceutics16060776>
- [18] Debus, B., Parastar, H., Harrington, P., and Kirsanov, D. (2021). Deep learning in analytical chemistry. *TrAC Trends in Analytical Chemistry*, 145, 116459. <https://doi.org/10.1016/j.trac.2021.116459>

- [19] Duan, C., Nandy, A., Terrones, G. G., Kastner, D. W., and Kulik, H. J. (2023). Active learning exploration of transition-metal complexes to discover method-insensitive and synthetically accessible chromophores. *JACS Au*, 3(2), 391–401. <https://doi.org/10.1021/jacsau.2c00547>
- [20] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- [21] Gioiello, A., Piccinno, A., Lozza, A. M., and Cerra, B. (2020). Medicinal chemistry in the era of machines and automation: Recent advances in continuous flow technology. *Journal of Medicinal Chemistry*, 63(13), 6624–6647. <https://doi.org/10.1021/acs.jmedchem.9b01956>
- [22] Goh, G. B., Hodas, N. O., and Vishnu, A. (2017). Deep learning for computational chemistry. *Journal of Computational Chemistry*, 38(16), 1291–1307. <https://doi.org/10.1002/jcc.24764>
- [23] Gunasekaran, H., Ramalakshmi, K., Rex Macedo Arokiaraj, A., Deepa Kanmani, S., Venkatesan, C., and Suresh Gnana Dhas, C. (2021). Analysis of DNA sequence classification using CNN and hybrid models. *Computational and Mathematical Methods in Medicine*, 2021, 1835056. <https://doi.org/10.1155/2021/1835056>
- [24] Hautier, G., Fischer, C., Ehrlacher, V., Jain, A., and Ceder, G. (2011). Data-mined ionic substitutions for the discovery of new compounds. *Inorganic Chemistry*, 50(2), 656–663. <https://doi.org/10.1021/ic102031h>
- [25] Himanen, L., Geurts, A., Foster, A. S., and Rinke, P. (2019). Data-driven materials science: Status, challenges, and perspectives. *Advanced Science*, 6(21), 1900808. <https://doi.org/10.1002/advs.201900808>
- [26] Hirani, R., Noruzi, K., Khuram, H., Hussaini, A. S., Aifuwa, E. I., et al. (2024). Artificial intelligence and healthcare: A journey through history, present innovations, and future possibilities. *Life*, 14(5), 557. <https://doi.org/10.3390/life14050557>
- [27] Hou, Y. F., Zhang, L., Zhang, Q., Ge, F., and Dral, P. O. (2024). Physics-informed active learning for accelerating quantum chemical simulations. *Journal of Chemical Theory and Computation*, 20(18), 7744–7754. <https://doi.org/10.1021/acs.jctc.4c00821>
- [28] Huang, G., Guo, Y., Chen, Y., and Nie, Z. (2023). Application of machine learning in material synthesis and property prediction. *Materials*, 16(17), 5977. <https://doi.org/10.3390/ma16175977>
- [29] Jia, J., Liang, W.-C., and Liang, Y. (2023). A review of hybrid and ensemble in deep learning for natural language processing. *arXiv*. <https://doi.org/10.48550/arXiv.2312.05589>

- [30] Jin, D., Liang, Y., Xiong, Z., Yang, X., Wang, H., Zeng, J., and Gu, S. (2025). Application of transformers to chemical synthesis. *Molecules*, 30(3), 493. <https://doi.org/10.3390/molecules30030493>
- [31] Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- [32] Jung, E., Luna, M., and Park, S. H. (2021). Conditional GAN with an attention-based generator and a 3D discriminator for 3D medical image generation. In M. de Bruijne et al., (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2021* (Vol. 12906, pp. 318–328). Springer. https://doi.org/10.1007/978-3-030-87231-1_31
- [33] Katzmann, A., Taubmann, O., Suehling, M., and Kaergel, R. (2024). Method for providing medical imaging decision support data (EP Patent No. 4462445).
- [34] Khadhoury, S., Hramyka, A., Gallagher, K., Light, A., Ippoliti, S., Edison, M., Alexander, C., Kulkarni, M., Zimmermann, E., Nathan, A., et al. (2024). Machine learning and external validation of the IDENTIFY risk calculator for patients with haematuria referred to secondary care for suspected urinary tract cancer. *European Urology Focus*, 10(6), 1034–1042. <https://doi.org/10.1016/j.euf.2024.06.004>
- [35] Klambauer, G., Hochreiter, S., and Rarey, M. (2019). Machine learning in drug discovery. *Journal of Chemical Information and Modeling*, 59(3), 945–946. <https://doi.org/10.1021/acs.jcim.9b00136>
- [36] Klenam, D. E. P., Asumadu, T. K., Vandadi, M., Rahbar, N., McBagonluri, F., and Soboyejo, W. O. (2023). Data science and material informatics in physical metallurgy and materials science. *Results in Materials*, 19, 100455. <https://doi.org/10.1016/j.rinma.2023.100455>
- [37] Korshunova, M., Ginsburg, B., Tropsha, A., and Isayev, O. (2021). OpenChem: A deep learning toolkit for computational chemistry and drug design. *Journal of Chemical Information and Modeling*, 61(1), 7–13. <https://doi.org/10.1021/acs.jcim.0c00971>
- [38] Kündig, P. (2006). Chemistry: The future of organic synthesis. *Science*, 314(5798), 430–431. <https://doi.org/10.1126/science.1134084>
- [39] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- [40] Luong, K. D., and Singh, A. (2024). Application of transformers in cheminformatics. *Journal of Chemical Information and Modeling*, 64(11), 4392–4409. <https://doi.org/10.1021/acs.jcim.3c02070>
- [41] Malashin, I., Tynchenko, V., Gantimurov, A., Nelyub, V., and Borodulin, A. (2024). Applications of long short-term memory (LSTM) networks in polymeric sciences: A review. *Polymers*, 16(18), 2607. <https://doi.org/10.3390/polym16182607>

- [42] Mater, A. C., and Coote, M. L. (2019). Deep learning in chemistry. *Journal of Chemical Information and Modeling*, 59(6), 2545–2559. <https://doi.org/10.1021/acs.jcim.9b00266>
- [43] Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- [44] Molnár, S., and Tamás, L. (2024). Variational autoencoders for 3D data processing. *Artificial Intelligence Review*, 57(2), 42. <https://doi.org/10.1007/s10462-023-10687-x>
- [45] Mroz, A. M., Basford, A. R., Hastedt, F., Jayasekera, I. S., Mosquera-Lois, I., Sedgwick, R., Ballester, P. J., Bocarsly, J. D., del Río Chanona, E. A., Evans, M. L., Frost, J. M., Ganose, A. M., Greenaway, R. L., Hii, K. K. (M.), Li, Y., Misener, R., Walsh, A., Zhang, D., and Jelfs, K. E. (2025). Cross-disciplinary perspectives on the potential for artificial intelligence across chemistry. *Chemical Society Reviews*, 54(11).
- [46] Peplow, M. (2014). Organic synthesis: The robo-chemist. *Nature*, 512, 20–22. <https://doi.org/10.1038/512020a>
- [47] Rana, M., and Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: Diagnosis to detection. *Multimedia Tools and Applications*, 82(17), 26731–26769. <https://doi.org/10.1007/s11042-022-14305-w>
- [48] Roy, S., Kumar, R., Acooli, A., Roy, S., Chatterjee, A., Chattaraj, S., Nayak, J., Jeon, B. H., Basu, A., Banerjee, S., Chakraborty, S., and Tripathy, S. K. (2024). Transforming nanomaterial synthesis through advanced microfluidic approaches. *Journal of Composites Science*, 8(10), 386. <https://doi.org/10.3390/jcs8100386>
- [49] Russell, S., and Norvig, P. (2011). Artificial intelligence: A modern approach (3rd ed.). Pearson.
- [50] Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., and Lee, A. A. (2019). Molecular transformer for chemical reaction prediction. *ACS Central Science*, 5(9), 1572–1583. <https://doi.org/10.1021/acscentsci.9b00576>
- [51] Segler, M. H. S., Preuss, M., and Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 555, 604–610. <https://doi.org/10.1038/nature25978>
- [52] Shahi, S., and Singh, S. K. (2024). Microfluidic devices with nanosensors for infectious disease detection. *Nanotechnology Perceptions*, 20, 3369–3377. <https://doi.org/10.62441/nanontp.vi.3295>
- [53] Sheikh, S., Vela, B., Honarmandi, P., Morcos, P., Shoukr, D., Kotb, A. M., Karaman, I., Elwany, A., and Arróyave, R. (2025). High-throughput alloy and process design for metal additive manufacturing. *npj*

- Computational Materials*, 11(1), 179. <https://doi.org/10.1038/s41524-025-01670-x>
- [54] Shen, Y., Borowski, J. E., Hardy, M. A., et al., (2021). Automation and computer-assisted planning for chemical synthesis. *Nature Reviews Methods Primers*, 1, 23. <https://doi.org/10.1038/s43586-021-00022-5>
- [55] Singh, Y. R., Shah, D. B., Kulkarni, M., Patel, S. R., Maheshwari, D. G., Shah, J. S., and Shah, S. (2023). Current trends in chromatographic prediction using artificial intelligence. *Analytical Methods*, 15(23), 2785–2797. <https://doi.org/10.1039/D3AY00362K>
- [56] Tripathy, P., and Shahi, S. (2024). AI-powered mining of nature's chemical library. *Research Journal of Pharmacology*, 18(3), 32–37. <https://doi.org/10.36478/makrjp.2024.3.32.37>
- [57] Tropsha, A., Isayev, O., Varnek, A., Schneider, G., and Cherkasov, A. (2024). Integrating QSAR modelling and deep learning in drug discovery: The emergence of deep QSAR. *Nature Reviews Drug Discovery*, 23(2), 141–155. <https://doi.org/10.1038/s41573-023-00832-0>
- [58] Tshitoyan, V., Dagdelen, J., Weston, L., et al. (2019). Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571, 95–98. <https://doi.org/10.1038/s41586-019-1335-8>
- [59] Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery. *Nature Reviews Drug Discovery*, 18(6), 463–477. <https://doi.org/10.1038/s41573-019-0024-5>
- [60] Venkataraman, M., Rao, G. C., Madavareddi, J. K., and Maddi, S. R. (2025). Leveraging machine learning models for ADMET evaluation. *ADMET and DMPK*, 13(3), 2772. <https://doi.org/10.5599/admet.2772>
- [61] Wang, J., Bhattarai, A., Do, H. N., Akhter, S., and Miao, Y. (2022). Molecular simulations and drug discovery of adenosine receptors. *Molecules*, 27(7), 2054. <https://doi.org/10.3390/molecules27072054>
- [62] Wu, Z., Wang, J., Du, H., et al. (2023). Chemistry-intuitive explanation of graph neural networks. *Nature Communications*, 14, 2585. <https://doi.org/10.1038/s41467-023-38192-3>
- [63] Xiao-Yong, P., and Hong-Bin, S. (2009). Robust prediction of B-factor profile from sequence using two-stage SVR based on random forest feature selection. *Protein and Peptide Letters*, 16(12), 1447–1454. <https://doi.org/10.2174/092986609789839250>
- [64] Xu, L., Zhu, J., Shen, X., et al. (2024). Machine-learning-accelerated ligand design in Cu(I)-catalyzed hydroxylation. *Angewandte Chemie International Edition*, 63, e202412552. [urlhttps://doi.org/10.1002/ange.202412552](https://doi.org/10.1002/ange.202412552)

- [65] Zeng, X., Wang, F., Luo, Y., et al. (2022). Deep generative molecular design reshapes drug discovery. *Cell Reports Medicine*, 3(12), 100794. <https://doi.org/10.1016/j.xcrm.2022.100794>
- [66] Zhou, W., Ruan, H., Zhu, L., Chen, S., and Yang, M. (2025). Novel glioblastoma deep molecular profiling. *ACS Omega*, 10(10), 10230–10250. <https://doi.org/10.1021/acsomega.4c09586>
- [67] Zhou, Z., Li, X., and Zare, R. N. (2017). Optimizing chemical reactions with deep reinforcement learning. *ACS Central Science*, 3(12), 1337–1344. <https://doi.org/10.1021/acscentsci.7b00492>

Chapter 6

AI IN GREEN CHEMISTRY: PATHWAY TO SUSTAINABLE CHEMICAL PROCESSES

Sanchita

Assistant Professor, Department of Applied Science and Humanities,
The Technological Institute of Textile and Sciences, Bhiwani, India-
127021.

Email: sanchita.gawri@gmail.com

Abstract

The growing need for sustainability has positioned green and sustainable chemistry as a crucial element in present-day industrial and scientific research. With the rise in technology, artificial intelligence (AI) is increasingly being used to create new sustainable systems. This progress enables faster chemical discovery, the design of eco-friendly processes, the use of greener solvents, ultimately contributing to a waste-free environment. Artificial intelligence united with machine learning has significantly paved the way by offering innovative insights into reaction outcomes, solvent selection, catalyst design, and the development of sustainable synthetic routes. The importance of AI-driven tools extends beyond merely predicting reaction outcomes; they also aid in finding environmentally benign solvents, developing efficient and selective catalysts, and lowering harmful by-products, particularly in areas such as organic synthesis, pharmaceutical manufacturing, and industrial chemical production. This book chapter highlights the recent advancements, challenges, and future perspectives of AI-driven green chemistry, emphasizing its potential to foster sustainable chemical industries.

Keywords: artificial intelligence, green chemistry, sustainable chemical processes, low-carbon manufacturing, process optimization

6.1 Introduction

The chemical industry reinforces modern society by providing essential materials, pharmaceuticals, energy carriers, and consumer products. However, conventional chemical processes are often associated with high energy consumption, hazardous waste generation, and significant greenhouse gas emissions. These environmental challenges have intensified global efforts toward sustainable development, prompting the adoption of green chemistry principles and sustainable chemical engineering practices.

Green chemistry emphasizes the prevention of waste, the use of safer chemicals and solvents, energy efficiency, renewable feedstocks, and inherently safer process design (Kreuder et al., 2017). Despite their conceptual clarity, the practical implementation of these principles remains challenging due to the complexity of chemical systems and the vast design space encompassing molecules, reactions, and industrial processes. Traditional experimental and mechanistic modeling approaches are often time-consuming, expensive, and limited in their ability to explore large solution spaces efficiently.

Artificial intelligence has emerged as a transformative tool capable of overcoming these limitations (Leonard et al., 2024). AI techniques, including machine learning (ML) and deep learning (DL), enable the extraction of meaningful patterns from large and heterogeneous datasets generated through experiments, simulations, and industrial operations (Butler et al., 2018). By learning from data, AI models can predict chemical properties, optimize reaction conditions, guide solvent and catalyst selection, and enhance process efficiency with reduced environmental impact (Ali et al., 2024).

AI-driven approaches are increasingly applied across the entire chemical value chain—from molecular-level design of environmentally benign chemicals to plant-level optimization of manufacturing operations (Uwineza and Waskiewicz, 2020). This chapter explores the role of AI in advancing green chemistry and sustainable chemical processes, highlighting methodologies, applications, challenges, and future opportunities.

Artificial Intelligence is making significant contributions to the field of green chemistry, which focuses on developing environmentally sustainable processes and products. The application of AI in green

chemistry enhances efficiency, reduces waste, and facilitates the design of more sustainable and eco-friendly chemical processes (Jiang et al., 2022). AI algorithms can predict the properties of new molecules and recommend synthetic routes with minimal environmental impact. This helps researchers design greener alternatives to existing chemicals and pharmaceuticals. AI is used to optimize chemical reactions by analysing reaction parameters and suggesting conditions that minimize waste. AI assists in the design of catalysts for chemical reactions that are more selective and efficient, reducing the need for harsh conditions and minimizing the generation of by-products. AI models can recommend environmentally friendly solvents by analysing the properties of different solvents and their impact on the overall environmental footprint of chemical processes (Ananikov et al., 2024). AI is employed to predict and optimize reaction pathways to minimize by-products and waste generation during chemical synthesis. This contributes to more sustainable and efficient manufacturing processes. Furthermore, to develop metrics and models for assessing the environmental impact of chemical processes AI is used (Hassija et al., 2024). This includes evaluating factors such as energy consumption, carbon footprint, and toxicity. It helps identify sustainable and environmentally friendly materials by analysing large datasets. In addition, it monitors and control chemical processes in real-time, adjusting parameters to optimize efficiency and reduce environmental impact. This includes the use of sensors and feedback mechanisms. AI tools contribute to educational initiatives by providing virtual labs, simulations, and interactive platforms to teach green chemistry principles and methodologies. AI assists in ensuring compliance with environmental regulations by automating the analysis of chemical processes to meet green chemistry standards. Kevin C. Leonard reported in his paper that we now have the power to use artificial intelligence to logically guide sustainable chemistry research by uncovering complex performance relationships. The integration of AI in green chemistry not only accelerates research and development but also provides a systematic approach to designing and optimizing chemical processes with reduced environmental impact (Leonard et al., 2021). By leveraging AI, researchers can make informed decisions to advance the principles of sustainability in the chemical industry.

6.2 AI Techniques Supporting Sustainable Chemistry

With advances in computer science, artificial intelligence (AI) now provides powerful tools to guide sustainable chemistry research by uncovering complex relationships that influence chemical performance. Although AI has already revolutionized fields like bioinformatics and drug discovery, its adoption in catalysis and sustainable chemistry has been limited, leaving substantial untapped potential. Catalysis development, for instance, has traditionally relied on trial-and-error experimentation or empirical heuristics, which are time-consuming and often produce unpredictable outcomes. Catalyst performance depends on a multitude of factors, including operating conditions, elemental composition (metals, supports, impurities), morphological characteristics (phase, porosity, surface area, conductivity), and reactor configuration. The vast number of variable combinations makes systematic exploration using conventional methods extremely challenging (Mohamadpour et al., 2024).

AI technologies, particularly machine learning (ML) and natural language processing (NLP), offer effective solutions to these challenges. ML builds predictive models from large datasets, enabling researchers to identify patterns and relationships that are difficult to discern manually. Although chemistry and catalysis literature contain extensive experimental and computational data, this information is often scattered across journals in non-uniform formats, complicating its direct use (Makarov et al., 2023). NLP can extract, organize, and standardize this textual information, allowing AI models to understand and leverage complex experimental details, including catalysts, reactants, products, and reaction conditions (Burkov, 2019).

A promising framework, CataLST, combines AI with human expertise to accelerate catalyst discovery. The approach involves four stages: cataloging literature and experimental data using NLP, learning insights through ML, designing new catalysts or reaction conditions guided by these insights alongside computational models and human intuition, and experimentally testing the predicted outcomes. This iterative process can be applied at both small and large scales, potentially giving rise to an “Internet of Catalysis” (IoC) that archives, analyzes, and exploits catalyst and reaction data to rationally design new chemical processes. (Leonard et al., 2021)

Deep learning, a rapidly growing subset of ML, further enhances catalyst and reaction design by automatically extracting hierarchical features from raw data without manual input (Strubell et al., 2020). Techniques such as convolutional neural networks, recurrent neural networks, and generative models (e.g., variational autoencoders and generative adversarial networks) have shown remarkable capabilities in predicting molecular properties and generating new catalyst candidates with high precision—often comparable to computationally expensive density functional theory (DFT) calculations.

Beyond catalysis, ML applications extend to chemical engineering and process optimization. Surrogate models can approximate complex nonlinear chemical phenomena for multiscale process systems, improving tractability in large-scale optimization problems (Grimstad and Andersson, 2019). In carbon capture, utilization, and storage (CCUS), ML aids in screening materials, predicting physicochemical properties, optimizing process flows, and analyzing supply chain logistics (Hasan et al., 2015). Similarly, AI facilitates real-time plant monitoring, assessing safety metrics, environmental impact, and operational efficiency. In drug discovery, AI accelerates compound design while reducing chemical waste, supporting both sustainability and efficiency.

Overall, integrating AI with mechanistic understanding offers a transformative approach to sustainable chemistry and catalysis, enabling systematic exploration of chemical spaces, rational design of catalysts, process optimization, and environmentally responsible chemical innovation (Taylor et al., 2023).

6.3 AI-Assisted Molecular And Solvent Design

6.3.1 Sustainable Molecular Design

AI-driven approaches have emerged as powerful tools for sustainable molecular design by enabling rapid, accurate, and cost-effective prediction of molecular properties. Advanced machine learning models can simultaneously evaluate a wide range of physicochemical, toxicological, and environmental parameters—such as solubility, stability, bioaccumulation potential, and ecotoxicity—at the early stages of molecular development. This predictive capability significantly reduces dependence on time-consuming experimental trial-and-error

methods, thereby conserving resources and minimizing chemical waste (Coley et al., 2017).

By employing quantitative structure–property relationship (QSPR) and quantitative structure–activity relationship (QSAR) modeling, AI systems can identify key structural features that govern molecular behavior and environmental impact. These insights guide the rational design of molecules that are biodegradable, exhibit low toxicity to humans and ecosystems, and maintain high functional performance. In addition, generative AI models can propose novel molecular structures optimized for multiple sustainability objectives, balancing efficiency, safety, and environmental compatibility (Jha et al., 2020).

Furthermore, AI-assisted molecular screening supports compliance with regulatory frameworks and green chemistry metrics by proactively eliminating hazardous candidates before synthesis. As a result, sustainable molecular design driven by AI not only accelerates innovation but also embeds environmental responsibility into the core of chemical discovery. This integration directly advances green chemistry principles by promoting safer chemicals, reduced environmental footprints, and more sustainable industrial practices.

6.3.2 Green Solvent Selection and Design

Solvents play a crucial role in determining the environmental impact of chemical reactions, as many traditionally used solvents such as benzene and dichloromethane are highly toxic, non-biodegradable, and pose serious challenges in terms of safe handling and disposal. In response to these concerns, green chemistry strongly advocates the transition toward safer, environmentally benign, and sustainable solvent systems (Diorazio et al., 2016). In recent years, AI has emerged as a powerful and reliable tool in this pursuit by enabling detailed analysis of solvent properties and accurately predicting their behavior under specific reaction conditions. Machine learning (ML) algorithms are capable of screening large solvent databases and recommending non-toxic, biodegradable, and recyclable alternatives, thereby aligning solvent selection with the core principles of green chemistry (Rodrigues, 2019).

A notable contribution in this area was reported by Sels et al., who developed an AI-driven framework for solvent selection and substitution. Their approach employed a neural network based on

a Kohonen Self-Organizing Map to cluster a comprehensive solvent database according to key physicochemical properties. These solvent clusters were subsequently validated using both chemical insight and statistical analysis, and the results were presented through a user-friendly interface in the form of the Sustainable Solvents Selection and Substitution Software (SUSSOL) tool. SUSSOL enables users to systematically explore solvent options and generates ranked greener alternatives for a given solvent based on safety, health, and environmental performance criteria. The reliability of this platform was demonstrated through several case studies, where its recommendations closely matched those reported in existing green chemistry literature (Sels et al., 2020).

6.4 Reaction Optimization And Green Catalysis

Optimizing reaction conditions in organic chemistry—temperatures, catalysts, and solvents—has traditionally relied on labor-intensive trial-and-error experimentation. Machine learning offers a more efficient alternative by analyzing large datasets of historical reactions to predict key parameters such as yield, reaction time, and selectivity. This approach reduces resource consumption and minimizes environmental impact (Winterton, 2021).

AI-assisted reaction optimization improves yields, selectivity, and energy efficiency. ML models guide catalyst discovery and optimization, enabling reactions to proceed under milder conditions while minimizing energy use (Weber et al., 2021). AI-based retrosynthesis planning supports the selection of sustainable pathways, reducing reaction steps and hazardous by-products. For example:

- Kariofillis et al., (2022) applied data science techniques to analyze substrate scope in Ni/photoredox-catalyzed cross-coupling reactions, combining DFT featurization, dimensionality reduction, and hierarchical clustering to identify high-efficiency reaction conditions.
- Jorayev et al., (2022) optimized a biowaste-derived chemical route to p-cymene using a Bayesian two-stage evolutionary multi-objective optimization (TS-EMO) algorithm, achieving maximal conversion and selectivity without a detailed kinetic model.

AI also supports flow chemistry, enabling real-time optimization of continuous reactions to reduce waste and energy consumption. Platforms such as Synthia provide AI-driven retrosynthesis for greener synthetic routes, avoiding toxic by-products.

6.5 Sustainable Separation Processes And Process Intensification

Separation processes are among the most energy-intensive operations in the chemical industry. AI-driven optimization of distillation, absorption, extraction, and membrane-based separations leads to significant energy savings. Process intensification strategies guided by AI enhance mass and heat transfer while reducing equipment size and environmental footprint (Zhang et al., 2023).

6.6 AI For Low-Carbon Manufacturing And Circular Chemistry

AI plays a critical role in reducing carbon emissions through optimized energy integration, predictive maintenance, and real-time process control. In circular chemistry, AI supports waste valorization, recycling process optimization, and the development of closed-loop manufacturing systems. AI-driven life-cycle assessment enables data-informed decision-making for sustainable process design (Pimenow, 2024)

6.7 Digital Twins And Smart Sustainable Manufacturing

Digital twins integrate AI models with real-time plant data to create dynamic virtual replicas of chemical processes. These systems enable predictive optimization, safety enhancement, and emission reduction, thereby supporting sustainable manufacturing and regulatory compliance (Mane et al., 2024).

6.8 Challenges And Limitations of AI Adoption In Green Chemistry

Despite its potential, AI adoption in green chemistry faces several challenges:

- **Data Quality and Availability:** High-quality, curated datasets are essential for AI model accuracy, yet many green chemistry datasets are limited in size or diversity. Insufficient data can lead to biased or unreliable predictions, especially in emerging areas like eco-friendly solvent design. Collaborative data-sharing initiatives and standardized datasets are needed (Weber et al., 2021).
- **Complexity of Chemical Systems:** Organic reactions involve multiple variables (solvents, catalysts, reagents) and non-linear relationships, which can challenge ML algorithms. Dynamic systems affected by temperature, pH, and pressure further complicate predictive modeling (Zhang et al., 2023).
- **Interdisciplinary Knowledge:** Effective AI integration requires collaboration between chemists, data scientists, and environmental experts. Misaligned terminologies and methodologies can hinder progress. Interdisciplinary education and shared knowledge platforms are critical (Pimenow, 2024).
- **Scalability and Generalization:** AI models trained on specific reactions may not generalize to other reactions or greener alternatives. Techniques such as transfer learning can help improve model adaptability across chemical systems (Smith, 2023).
- **Ethical and Regulatory Concerns:** AI raises issues related to data privacy, intellectual property, transparency, and environmental impact. Regulatory frameworks and clear ethical guidelines are necessary for responsible AI deployment in green chemistry (Hermann et al., 2021).
- **Energy Consumption of AI Models:** Training large AI models is computationally intensive and can offset environmental benefits. Developing energy-efficient algorithms, hardware, and renewable-powered AI infrastructure is essential for sustainable implementation. (Kamkar et al., 2024).

6.9 Future Perspectives Of AI In Green Chemistry

Future AI-driven green chemistry will focus on:

- Autonomous laboratories for high-throughput experimentation
- Explainable and trustworthy AI models for regulatory compliance

- Secure, open-access data-sharing platforms
- Integration of AI with chemical theory, robotics, and quantum computing

Accessible AI platforms will allow chemists without advanced data science expertise to leverage AI tools. Cloud-based systems and standardized, high-quality datasets will democratize AI adoption globally. Balancing energy consumption with environmental benefits and establishing regulatory oversight will be critical to responsible deployment. Through these advances, AI is poised to accelerate sustainable chemical innovation, optimize resource utilization, reduce waste, and enable environmentally responsible chemical manufacturing on a global scale.

In the coming years, AI is also expected to enable localized and context-specific sustainable solutions by integrating regional climate conditions, resource availability, and policy constraints. AI-driven life cycle assessment, supply chain optimization, and waste minimization strategies will further enhance sustainability across chemical industries. Moreover, the integration of AI with autonomous laboratories, Industry 4.0 technologies, and circular economy principles will create a more interconnected and resource-efficient chemical ecosystem, where discovery, production, and recycling are seamlessly aligned.

Overall, while challenges remain, the future of AI in green chemistry is highly promising. With careful attention to data quality, energy efficiency, ethical governance, and interdisciplinary collaboration, AI can serve as a powerful catalyst for achieving long-term environmental sustainability and societal benefit.

6.10 Conclusion

Artificial intelligence is increasingly recognized as a key driver of green chemistry and sustainable chemical processes, enabling faster discovery, improved resource efficiency, and reduced environmental impact. By supporting the design of eco-friendly materials, optimizing chemical reactions, and advancing low-carbon and circular chemical industries, AI contributes significantly to sustainability goals. This analysis highlights that green chemistry is fundamental to building a sustainable society, with green solvents, sustainable synthesis, catalysis, and optimized reaction conditions playing central roles. While AI and machine learning offer immense potential to transform

green chemistry, their impact depends on responsible implementation, high-quality data, and strong multidisciplinary collaboration. Ultimately, AI should be viewed as a powerful enabling tool that, when combined with human expertise and cooperative frameworks, can help realize a cleaner, greener, and more resilient future for coming generations.

References

- [1] Ali, R. S. A. E., Meng, J., Khan, M. E. I., and Jiang, X. (2024). Machine learning advancements in organic synthesis: A focused exploration of artificial intelligence applications in chemistry. *Artificial Intelligence in Chemistry*, 2, 100049. <https://doi.org/10.1016/j.aichem.2024.100049>
- [2] Ananikov, V. P. (2024). Classification of artificial intelligence applications in chemistry: From automation to digital scientific thinking. *Artificial Intelligence in Chemistry*, 2, 100075. <https://doi.org/10.1016/j.aichem.2024.100075>
- [3] Burkov, A. (2019). The hundred-page machine learning book. *Journal of Information Technology Case and Application Research*, 22,136. <https://doi.org/10.1080/15228053.2020.1766224>
- [4] Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., and Walsh, A.(2018). Machine learning for molecular and materials science. *Nature*, 559, 547–555. <https://doi.org/10.1038/s41586-018-0337-2>
- [5] Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., and Jensen, K. F. (2017). Prediction of organic reaction outcomes using machine learning. *ACS Central Science*, 3, 434–443. <https://doi.org/10.1021/acscentsci.7b00064>
- [6] Diorazio, L. J., Hose, D. R. J., and Adlington, N. K. (2016). Sustainable synthesis design: Development of a green route via continuous flow bromination. *Organic Process Research and Development*, 20, 760–773. <https://doi.org/10.1021/acs.oprd.6b00015>
- [7] Grimstad, B., and Andersson, H. (2019). ReLU networks as surrogate models in mixed-integer linear programs. *Computers and Chemical Engineering*, 131, 106580. <https://doi.org/10.1016/j.compchemeng.2019.106580>
- [8] Hasan, M. F., First, E. L., Boukouvala, F., and Floudas, C. A. (2015). A multiscale framework for CO₂ capture, utilization, and sequestration: CCUS and CCU. *Computers and Chemical Engineering*, 81, 2–21. <https://doi.org/10.1016/j.compchemeng.2015.04.034>
- [9] Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Huang, K., Scardapane, S., Goel, D., Spinelli, I., Mahmud, M., and Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial

- intelligence. *Cognitive Computation*, 16, 45–78. <https://doi.org/10.1007/s12559-023-10179-8>
- [10] He, C., Zhang, C., Bian, T., Jiao, K., Su, W., Wu, K. -J., and Su, A. (2023). Artificial intelligence enabled design, synthesis, and process optimization of chemical products for Industry 4.0. *Processes*, 11(2), 330. <https://doi.org/10.3390/pr11020330>
- [11] Hermann, E., Hermann, G., and Tremblay, J. C. (2021a). Ethical artificial intelligence in chemical research and development: A dual advantage for sustainability. *Nature Communications*, 12, 6830. <https://doi.org/10.1038/s41467-021-27048-8>
- [12] Hermann, E., Hermann, G., and Tremblay, J. C. (2021b). Ethics and governance of artificial intelligence in scientific research. *Science and Engineering Ethics*, 27, 45. <https://doi.org/10.1007/s11948-021-00325-6>
- [13] Jha, D., Choudhary, K., Tavazza, F., Liao, W.-K., Choudhary, A., Campbell, C., and Agrawal, A. (2019). Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning. *Nature Communications*, 10, 1–12. <https://doi.org/10.1038/s41467-019-10343-0>
- [14] Jiang, P., Doan, H. A., Madireddy, S., Assary, R. S., and Balaprakash, P. (2022). Value-added chemical discovery using reinforcement learning. *Digital Discovery*, 1(1), 79–89. <https://doi.org/10.1039/D1DD00022A>
- [15] Jorayev, P., Russo, D., Tibbetts, J. D., Schweidtmann, A. M., Deutsch, P., Bull, S. D., and Lapkin, A. (2022). Autonomous chemical process optimization using machine learning. *Chemical Engineering Science*, 247, 116938. <https://doi.org/10.1016/j.ces.2021.116938>
- [16] Kamkar, M., Leonard, K. C., Ferrer, I., et al. (2024). Artificial intelligence and automation for sustainable chemistry. *ACS Sustainable Chemistry and Engineering*, 12, 2924–2945. <https://doi.org/10.1021/acssuschemeng.4c01004>
- [17] Kapustina, O., Burmakina, P., Gubina, N., Serov, N., and Vinogradov, V. (2024). User-friendly and industry-integrated AI for medicinal chemists and pharmaceuticals. *Artificial Intelligence in Chemistry*, 2, 100072. <https://doi.org/10.1016/j.aichem.2024.100072>
- [18] Kariofillis, S. K., Jiang, S., Zurański, A. M., Gandhi, S. S., Martinez Alvarado, J. I., and Doyle, A. G. (2022). Predicting and improving reaction outcomes via data-driven catalyst design. *Journal of the American Chemical Society*, 144, 1045–1055. <https://doi.org/10.1021/jacs.1c12203>
- [19] Kreuder, D. V., Ashley, T.-K., Whitford, J., et al. (2017). A method for assessing greener alternatives between chemical products following the 12 principles of green chemistry. *ACS Sustainable Chemistry and*

- Engineering*, 5, 2927–2935. <https://doi.org/10.1021/acssuschemeng.6b02399>
- [20] Leonard, K. C. (2021). Can artificial intelligence and machine learning be used to accelerate sustainable chemistry and engineering? *ACS Sustainable Chemistry and Engineering*, 9(18), 6126–6129. <https://doi.org/10.1021/acssuschemeng.1c02741>
- [21] Makarov, D. M., Lukanov, M. M., Rusanov, A. I., Mamardashvili, N. Z., and Ksenofontov, A. A. (2023). Machine learning approach for predicting the yield of pyrroles and dipyrromethanes condensation reactions with aldehydes. *Journal of Computational Science*, 74, 102173. <https://doi.org/10.1016/j.jocs.2023.102173>
- [22] Mane, S., Dhote, R. R., Sinha, A., and Thirumalaiswamy, R. (2024). Digital twin in the chemical industry: A review. *Digital Twins and Applications*, 1(2), 118–130. <https://doi.org/10.1049/dgt2.12019>
- [23] Mohamadpour, F., and Amani, A. M. (2024). Photocatalytic systems: Reactions, mechanism, and applications. *RSC Advances*, 14, 20609–20624. <https://doi.org/10.1039/D4RA03259D>
- [24] Pimenow, S., Pimenowa, O., and Prus, P. (2024). Challenges of artificial intelligence development in the context of energy consumption and impact on climate change. *Energies*, 17, 5965. <https://doi.org/10.3390/en17235965>
- [25] Rodrigues, T. (2019). The good, the bad, and the ugly in chemical and biological data for machine learning. *Drug Discovery Today: Technologies*, 32–33, 3–8. <https://doi.org/10.1016/j.ddtec.2020.07.001>
- [26] Sels, H., De Smet, H., and Geuens, J. (2020). Recent advances in catalytic molecular transformations in green chemistry applications. *Molecules*, 25, 3037. <https://doi.org/10.3390/molecules25133037>
- [27] Smith, E. K. (2023). Artificial intelligence for sustainable chemical innovation. *Chemical Science*, 15, 5143–5154. <https://doi.org/10.1039/D3SC04928K>
- [28] Strubell, E., Ganesh, A., and McCallum, A. (2020). Energy and policy considerations for deep learning in NLP. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 13693–13696. <https://doi.org/10.1609/aaai.v34i09.7123>
- [29] Taylor, C. J., Pomberger, A., Felton, K. C., et al. (2023). A brief introduction to chemical reaction optimization. *Chemical Reviews*, 123(6), 3089–3128. <https://doi.org/10.1021/acs.chemrev.2c00798>
- [30] Uwineza, P. A., and Waskiewicz, A. (2020). Recent advances in supercritical fluid extraction of natural bioactive compounds from natural plant materials. *Molecules*, 25(17), 3847. <https://doi.org/10.3390/molecules25173847>

- [31] Weber, J. M., Guo, Z., Zhang, C., Schweidtmann, A. M., and Lapkin, A. (2021). Data-driven methods for chemical reaction and process optimization. *Chemical Society Reviews*, 50, 12013–12041. <https://doi.org/10.1039/D1CS00477H>
- [32] Winterton, N. (2021). A critical review of clean production and environmental policy in industrial practice. *Clean Technologies and Environmental Policy*, 23, 2499–2510. <https://doi.org/10.1007/s10098-021-02188-8>
- [33] Zhang, S. Q., Xu, L. C., Li, S. W., Oliveira, J. C. A., Li, X., Ackermann, L., and Hong, X. (2023). Machine learning-assisted reaction mechanism elucidation. *Chemistry – A European Journal*, 29, e202202834. <https://doi.org/10.1002/chem.202202834>

Chapter 7

ARTIFICIAL INTELLIGENCE IN MATHEMATICAL SCIENCES: METHODS, APPLICATIONS AND EMERGING TRENDS

Vandana Verma

Assistant Professor, P. G. Department of Mathematics, Khalsa College for Women, Civil Lines, Ludhiana, Punjab- 141001.

Email: vandanaverma1945@gmail.com

Abstract

Artificial Intelligence (AI) has gradually become part of the computational landscape of the Mathematical sciences. While AI is often discussed as a modern technological tool, its development relies heavily on a long-established area of mathematics. Concepts from linear algebra, calculus, probability theory, and optimization continue to shape the structure of learning algorithms and provide the theoretical basis needed to understand their behavior and limitations. In pure mathematics, the use of computational help has become increasingly common over the years. The techniques such as automated proof verification and computer-assisted exploration are now applied in areas like number theory and combinatorics, where many problems are too complicated to be handled entirely by hand. Applied mathematics has benefited from AI through methods like reinforcement learning in stochastic control, graph-based neural networks for discrete optimization, and simulation on high-dimensional manifolds. In addition, intelligent systems for proof verification and algorithmic reasoning are gradually being integrated into standard mathematical workflows, complementing traditional analytical approaches. These methods are intended to support, not replace, traditional mathematical reasoning. In addition, AI-assisted tools are increasingly used across fields such as physics, biology, economics, and the social

sciences to analyze complex mathematical models more. This abstract highlights AI as a supportive framework that strengthens both theoretical inquiry and practical problem-solving. By combining computational efficiency with mathematical rigor, application-oriented methodologies in AI continue to expand the scope and relevance of mathematical sciences in contemporary research.

Keywords: automated theorem, machine learning, mathematical modeling, probability theory, and reinforcement learning

7.1 Introduction

Artificial Intelligence (AI) has slowly developed into an important part of contemporary mathematical sciences. Over the past few years, its use has increased significantly, especially as a computational support system for mathematical investigation, analytical work, and optimization tasks. Problems that previously demanded lengthy manual computation can now be approached more efficiently using AI-based techniques. As a result, AI is beginning to shape both theoretical research and a practical application within mathematics (Goodfellow et al., 2016).

It is equally necessary to understand that mathematics itself provides the groundwork on which AI is built. AI systems do not operate in isolation; they function through mathematical logic and structured computation. Areas such as linear algebra, calculus, probability, statistics, and optimization supply the essential tools required for algorithm construction, data analysis, and performance assessment. These mathematical foundations help ensure that AI models remain consistent, accurate, and computationally feasible (Bishop, 2006)

In general terms, Artificial Intelligence refers to the design of computer systems that are capable of carrying out tasks commonly associated with human intelligence. These tasks include reasoning, problem-solving, language comprehension, perception, learning, and decision-making. Despite their apparent complexity, each of these capabilities depends heavily on mathematical formulation and algorithmic processing (Mitchell, 1997; Poole and Mackworth, 2010).

This chapter focuses on examining how AI contributes to both pure and applied mathematics. It discusses the use of computational systems in theorem proving, identification of mathematical patterns, symbolic manipulation, and optimization modeling. In addition, it reflects on existing challenges and the future scope for deeper integration of AI into mathematical research, while maintaining strict logical and analytical standards.

7.2 Mathematical Foundations Supporting AI

Mathematics forms the structural base of Artificial Intelligence. Rather than acting as a secondary support, it functions as the core framework that allows AI systems to operate effectively. Every stage of AI processing — from data handling to predictive output — involves mathematical computation. Through mathematical modeling, AI systems are able to recognize patterns, address uncertainty, and refine learning processes over time (Deisenroth et al., 2020; Murphy, 2012).

7.2.1 Linear Algebra: The Backbone of Data Representation and Manipulation

Artificial Intelligence systems primarily process numerical data. Linear algebra provides systematic methods for organizing this information using vectors, matrices, and higher-dimensional arrays. For instance, digital images can be represented as matrices containing pixel intensities, while textual information may be converted into vector form so that machines can analyze linguistic patterns numerically (Lay, 2012).

Matrix operations play a central role in machine learning. Neural networks depend on matrix multiplication to transmit and transform information across computational layers. In the same way, regression techniques and deep learning structures rely on linear algebraic transformations to identify relationships embedded in datasets (Goodfellow et al., 2016; Aggarwal, 2018).

Another significant contribution of linear algebra is dimensionality reduction. When data contains a large number of variables, computation becomes complex and time-consuming. Techniques such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) help reduce dimensional scale while preserving

essential features. This leads to faster computation and reduces the likelihood of over fitting (Jolliffe, 2025; Golub and Van Loan, 2013).

7.2.2 Calculus: The Engine of Optimization and Learning

Calculus plays an essential role in enabling AI systems to learn from data. During training, model parameters must be adjusted so that prediction errors are minimized. Differential calculus provides the mathematical mechanism for measuring how such adjustments influence outcomes (Stewart, 2015; Apostol, 2024).

Gradient descent is among the most widely used optimization procedures in machine learning. By determining the gradient of a loss function, the method identifies how parameters should be modified to reduce error. Through repeated iterations, the model gradually approaches an optimal state (Ruder, 2016; Bottou, 2010).

Back propagation represents another major application of calculus within AI. Using the chain rule, it calculates how prediction errors move backward through neural network layers, enabling efficient weight updates. Calculus also supports sensitivity analysis by showing how variations in input variables influence outputs. This understanding is important for improving model stability and predictive performance (Rumelhart et al., 1986; LeCun et al., 2015).

7.2.3 Probability and Statistics: Handling Uncertainty and Extracting Insights from Data

Probability and statistics are fundamental when dealing with uncertain or incomplete data. Probability theory allows AI systems to estimate event likelihoods and operate effectively even when outcomes cannot be determined with certainty (Ross, 2014; Grimmett and Stirzaker, 2001).

Statistical methods assist in organizing, summarizing, and interpreting large datasets. Through statistical analysis, AI systems can recognize trends, identify correlations, and extract meaningful insights. Applications such as recommendation platforms and spam filtering systems rely heavily on probabilistic reasoning (Casella and Berger, 2002; Hastie et al., 2009).

Approaches including Bayesian inference and hypothesis testing enable AI systems to make informed decisions grounded in available

evidence. Model evaluation also depends on statistical indicators such as accuracy, precision, recall, and F1-score. Validation procedures like cross-validation help ensure that models perform reliably beyond their training datasets (Wasserman, 2004; Powers, 2011).

7.2.4 Discrete Mathematics: The Foundation for Algorithms and Logical Reasoning

Discrete mathematics provides the theoretical structure required for algorithm design and logical reasoning within AI. It deals with finite mathematical systems such as sets, graphs, trees, and combinatorial arrangements, all of which are essential for computational modeling (Rosen, 2012; Cormen et al., 2009).

Graph theory is especially relevant because many real-world systems can be represented as networks. Applications in social media analysis, communication infrastructures, and recommendation engines rely on graph-based representations to study relationships and interactions (West, 2001; Enderton, 2001).

Logic forms the basis of reasoning in AI. Knowledge representation systems and rule-based models depend on formal logical structures to derive conclusions from given information. Combinatorial optimization further supports AI by helping identify the most effective solution from a finite range of possibilities. This is particularly useful in scheduling, routing, and resource allocation problems. Mathematics, therefore, serves as the universal language through which AI systems interpret, process, and respond to data.

7.3 Applications In Pure Mathematics

7.3.1 Automated Theorem Proving (ATP):

It is also known as automated deduction and is a subfield of automated reasoning and mathematical logic dealing with proving mathematical theorems by computer programs (Harrison, 2009). ATP system transforms mathematical reasoning into a mechanical process through several steps:

a) Formula Preprocessing: User-inputted first order formulae undergo lexical and syntax analysis to generate a parse tree.

b) Normal Form Conversion: Formulae are often converted into Conjunctive Normal Form (CNF) or Skolem Normal Form (SNF) to simplify the search for contradictions.

c) Resolution and Refutation: Most provers use refutation, attempting to prove a statement by showing that its negation is unsatisfiable. The Robinson Resolution Principle is a standard method for deducing new data from these clauses.

d) Search Guidance: Because the space of potential proof steps grows exponentially, provers use heuristics to select which paths to explore first (Robinson, 1965; Harrisosn, 2009).

7.3.2 Conjecture Generation and Pattern Recognition:

AI systems can analyze large mathematical datasets to identify recurring patterns and relationships. In number theory, algebra, and combinatorics, machine learning models assist in generating plausible conjectures that can later be examined through rigorous proof.

Such tools do not replace creativity or intuition but serve as exploratory instruments that expand the range of hypothesis available for mathematical investigation (Colton, 2001; Urban et al., 2008).

7.3.3 Symbolic Computation:

Symbolic Computation systems use AI-inspired algorithms to manipulate algebraic expressions, solve equations, and simplify mathematical structures. These tools are widely used in algebraic geometry, differential equations, and mathematical physics (Geddes et al., 1992; Cox et al., 2015).

7.3.4 Algebra and Number theory:

In this Machine learning models analyze integrate sequences, prime distributions, and modular structures to detect hidden relationships. AI tools assist mathematicians in testing conjectures by generating large datasets of algebraic structures and identifying counterexamples (Hardy and Wright, 2008; Tao, 2016).

7.4 Applications In Applied Mathematics

7.4.1 Reinforcement learning and Stochastic Control:

Reinforcement learning (RL) is a computational framework in which agents learn strategies by interacting with dynamic environments. Mathematical models of Markov decision processes and stochastic control underpin RL algorithms.

Applications include optimal control problems, robotics, financial modeling, and operations research. Mathematical analysis ensures stability, convergence, and robustness of learning policies (Bertsekas, 2019).

7.4.2 Numerical Analysis and Approximation:

AI techniques enhance numerical methods by providing adaptive approximation strategies. Neural networks are increasingly used to approximate solutions to differential equations, especially in high dimensional settings where traditional numerical methods face limitations.

Such approaches support efficient simulation and modeling in engineering, physics, and applied sciences (Burden and Faires, 2011; Raissi et al., 2019).

7.4.3 Graph theory and Network Optimization:

Graph neural networks extend classical graph theory by enabling learning on structured data. These models are applied to transportation networks, communication systems and social networks, where discrete optimization problems are prevalent (Kipf and Welling, 2017; Hamilton, 2020).

Mathematical analysis of graph structures ensures interpretability and scalability of these models.

7.4.4 Differential Equations and Dynamical Systems:

Differential equations constitute a central pillar of applied mathematics, modeling phenomena across physics, biology, and engineering. Traditional numerical solvers rely on discretization techniques such as finite difference and finite element methods. While effective, these

approaches become computationally expensive for nonlinear or high dimensional systems (Raissi et al., 2019; Strogatz, 2018).

AI introduces neural differential equation solves that learn solution manifolds directly from data. Physics informed neural networks integrate governing equations into the learning process, ensuring that predicted solutions obey physical laws. This fusion of analytical structure with machine learning flexibility allows efficient modeling of turbulent flows, heat transfer systems, and biological oscillations.

Dynamical systems theory also benefits from AI-based phase space reconstruction, attractor detection, and chaos prediction. Intelligent models assist in identifying stability regions and bifurcation patterns that are difficult to compute analytically.

7.4.5 Financial Mathematics:

AI contributes to stochastic modeling, risk assessment, derivative pricing, and fraud detection. Deep learning models process high-frequency financial data to identify patterns beyond traditional stochastic calculus frameworks (Hull, 2018; Shreve, 2004).

7.5 Interdisciplinary Applications

The integration of AI and mathematics has led to substantial interdisciplinary impact. In physics; AI supports simulations of complex systems such as fluid dynamics and quantum interactions. In biology, mathematical AI models assist in understanding gene regulation, protein interactions, and population dynamics.

Economic and social sciences benefit from AI-driven mathematical models for forecasting, resource allocation, and policy analysis. These applications demonstrate the broad relevance of mathematical AI across diverse domains. A defining feature of AI in mathematical sciences is its role as a bridge between pure theory and applied practice (Mehta et al., 2019; Varian, 2019).

For example:

- Algebraic structures inform cryptographic algorithm design.
- Spectral graph theory underpins graph neural network architecture.

- Category theory concepts influence compositional deep learning models.

Such interdisciplinary synthesis demonstrates that AI is not only a computational accelerator but also a conceptual framework unifying diverse mathematical disciplines.

7.6 Emerging Trends In AI-Driven Mathematics

Artificial Intelligence is not only enhancing existing mathematical methodologies but is also reshaping the future direction of mathematical inquiry. Emerging trends in AI-driven mathematics reflect a shift from purely computational assistance toward collaborative discovery, automated reasoning and self-improving mathematical systems. These developments are redefining how; mathematical knowledge is generated, verified and applied.

7.6.1 Neural-Guided Symbolic Reasoning:

One of the most significant emerging trends is the integration of neural networks with symbolic mathematics. Classical computer algebra systems rely on deterministic symbolic manipulation, whereas neural-guided frameworks learn heuristic strategies for simplification, integration, and transformation of expressions. By combining pattern recognition with rule-based reasoning, these hybrid systems improve the efficiency of symbolic computation and reduce combinatorial search complexity (Marcus, 2018).

Such approaches are particularly valuable in integral calculus, differential operator algebra, and tensor simplification, where multiple transformation pathways exist. Neural guidance helps prioritize optimal proof or simplification routes, thereby accelerating symbolic discovery.

7.6.2 Hybrid Symbolic-Numeric Approaches:

Emerging research focuses on combining symbolic reasoning with numerical computation. These hybrid models aim to preserve mathematical rigor while benefiting from data-driven adaptability.

7.6.3 Explainable and Interpretable AI:

Interpretability remains a central concern in AI-driven mathematics. Current research emphasizes transparent models that align with mathematical principles, enabling meaningful theoretical insights.

7.6.4 Data-efficient Learning:

Mathematical datasets are often limited or abstract. Developing AI methods that function effectively with minimal data is an active area of research, guided by mathematical constraints and prior knowledge.

7.6.5 AI in High-Dimensional Geometry:

The study of high-dimensional spaces is central to modern mathematics and physics. AI tools now assist in visualizing, approximating, and classifying high-dimensional manifolds that are otherwise intractable through classical visualization.

Deep learning supports:

- Dimensionality reduction of geometric datasets.
- Manifold learning
- Curvature estimation
- Shape Classification

This trend is influencing string theory, cosmological modeling, and complex systems geometry.

7.6.6 Self-Improving Mathematical Systems:

An emerging frontier is the development of self-refining AI systems capable of improving their own mathematical reasoning. Through meta-learning and recursive training, these systems evaluate prior solutions, identify inefficiencies, and refine their algorithms.

Such systems demonstrate early forms of autonomous mathematical exploration, where AI iteratively enhances proof strategies, optimization heuristics, and approximation schemes without direct human intervention (Pearl and Mackenzie, 2018)

7.7 Challenges And Limitations

Despite its promise, AI presents challenge within mathematical sciences. Computational cost, lack of interpretability, and dependence on data quality pose limitations. Ensuring mathematical validity and avoiding over-reliance on heuristic methods remains critical concerns.

Ethical considerations also arise in applications involving decision-making and societal impact, emphasizing the need for responsible mathematical modeling (Doshi-Velez and Kim, 2017; Rudin, 2019).

7.8 Future Directions

Future research is likely to focus on deeper integration between AI and formal mathematical reasoning. Collaborative systems that support proof discovery, verification, and explanation may redefine mathematical workflows. Advances in computational mathematics, optimization theory, and probabilistic modeling will further enhance AI's role as a supportive tool rather than replacing it for human insight. Several promising directions are likely to shape the next phase of AI-driven mathematical sciences.

- Integration of Quantum computing with AI optimization.
- Development of fully automated proof discovery systems.
- Energy –efficient AI algorithms grounded in mathematical sparsity.
- Topological data analysis for high dimensional learning.
- Formal verification of neural network safety.

These research avenues illustrate the expanding frontier where mathematical rigor and computational intelligence converge (Preskill, 2018; Aaronson, 2013).

7.9 Conclusion

Artificial Intelligence has emerged as a powerful complementary framework within the mathematical sciences. Grounded in rigorous mathematical foundations, AI enhances both theoretical exploration and applied problem solving. Its application and interdisciplinary research highlight its transformative potential. Mathematics isn't just a supporting player; it's the fundamental language and operational

engine that powers everything AI systems do. Without a strong mathematical foundation, AI simply wouldn't exist as we know it.

While challenges related to interpretability and rigor persist, continued collaboration between mathematical and AI researchers promises to strengthen the relevance and applicability of mathematical sciences in addressing complex, real world problems.

References

- [1] Bertsekas, D. (2019). Reinforcement learning and Optimal Control (Vol.1). Athena Scientific.
- [2] Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer. 4(4), 738.
- [3] Bronstein, M. M., Bruna, J., Lecun, Y., Szlam, A., and Vandergheynst, P.(2017). Geometric deep learning: Going Beyond Euclidean data, *IEEE signal processing Magazine*, 34(4), 18-42.
- [4] Casella, G., and Berger, R. (2002). *Statistical Inference* Duxbury Press. Pacific Grove, CA.
- [5] Colton, S. (2001). *Automated Theory Formation in Pure Mathematics*. Springer.
- [6] Cormen,T. H., Leiserson,C. E., Rivest, R. L., and Stein, C. (2009). Introduction to Algorithms (3rd ed.). MIT Press and McGraw-Hill.
- [7] Gadavidis, G. (2017). Artificial intelligence, Computational thinking and Mathematics Education. *The International Journal of Information and Learning Technology*, 34(2), 133-139. <https://doi.org/10.1108/IJILT-09-2016-0048>
- [8] Goodfellow, I., Bengio, Y., and Courville, A. (2016). Regularization for Deep learning, 7, 216-261.
- [9] Hamilton, W. (2020). Graph Representation Learning. Morgan and Claypool.
- [10] Hardy, G. H., and Wright, E. M. (2008). An Introduction to the Theory of Numbers. Oxford University Press.
- [11] Harrison. (2009). Handbook of Practical Logic and Automated Reasoning. Cambridge University Press.
- [12] Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning.
- [13] Jolliffe, I. (2025). Principal Component Analysis. In International Encyclopedia of Statistical Science, Berlin, Heidelberg: Springer Berlin Heidelberg.
- [14] McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133. <https://doi.org/10.1007/BF02478259>

- [15] Mitchell, T. M. (1997). Artificial Neural Networks. *Machine Learning*, 45(81), 127.
- [16] Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT Press.
- [17] Panstar, M. (2024). Theorem proving in artificial neural networks: New frontiers in mathematical AI. *European Journal for Philosophy of Science*, 14(1), 4. <https://doi.org/10.1007/s13194-024-00569-6>
- [18] Pearl, J., and Mackenzie. (2018). *The book of why: The new science of cause and effect*. Basic books.
- [19] Poole, D. L., and Mackworth, A. K. (2010). *Artificial Intelligence: Foundations of Computational Agents*. Cambridge University Press.
- [20] Powers, D. (2011). Evaluation: Precision, recall and F-measure. *International Journal of Machine Learning Technologies*, 2(1), 37-63. <https://doi.org/10.48550/arXiv.2010.16061>
- [21] Robinson, J. A. (1965). A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1), 23-41.
- [22] Rosen, K. H. (2012). *Discrete Mathematics and its Applications*, McGraw-Hill. United States.
- [23] Ross, S. M. (2014). *Introduction to Probability Models*. Academic Press.
- [24] Ruder, S. (2016). An overview of gradient descent optimization algorithms, arXiv preprint arXiv: 1609.04747.
- [25] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323 (6088), 533-536. <https://doi.org/10.1038/323533a0>
- [26] Silverman, J. H. (2009). *The Arithmetic of Elliptic Curves* (Vol. 106, No. 2). New York: Springer.
- [27] Stewart, J. (2015). *Calculus: Early Transcendentals*. Cengage Learning.
- [28] West, D. B. (2001). *Introduction to Graph Theory* (Vol. 2, pp. 1-150). Upper Saddle River: Prentice Hall.
- [29] Wolfram, S. (2002). *A New Kind of Science: Notes from the Book*: Wolfram Media.

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 8

HARNESSING ARTIFICIAL INTELLIGENCE FOR CYBERCRIME DETECTION AND PREVENTION: THREATS, TECHNIQUES AND CHALLENGES

Inderpreet Kaur

Assistant Professor, Department of Computer Science, Khalsa College for Women, Civil Lines, Ludhiana, Punjab, India-141001.

Email: inderpreet.kcwldh@gmail.com

Abstract

As technology has become an essential part of everyday life, our dependence on digital systems has increased and so has the number of cybercrimes. In the past, cybercrime mainly involved activities such as identity theft, phishing scams, malware, spam and other common online attacks. However, with the rapid growth of Artificial Intelligence (AI) and Machine Learning (ML), cybercriminal activities have become far more advanced. These technologies enable attackers to automate, adapt and scale their attacks, making cybercrime more dangerous and increasingly difficult to detect and prevent. Cyber Security is vital for protecting our digital world. AI has transformed the cyber security industry. On one hand cybercriminals are using AI to conduct more sophisticated and targeted cyber-attacks; on the other hand, AI is driving significant advancements in cyber security defences, enabling security teams to identify and respond to attacks with greater speed and precision than ever before. Traditional cyber security systems use established techniques like firewalls, antivirus software, and intrusion detection systems to protect against threats. These approaches depend on set rules and human monitoring to recognize and block harmful activities. But by using AI technologies

individuals and organizations can detect and respond to threats with greater accuracy and speed than traditional measures. In this chapter, an attempt has been made to discuss various traditional as well as advanced cybercrimes and throw light on the role of AI in cybercrime control. The challenges and limitations of using AI based prevention methods have also been discussed.

Keywords: artificial intelligence, cybercrime, cybercriminals, machine learning, online attacks, technology

8.1 Introduction

Data has become a critical asset for individuals as well as organizations in the digital era. With increasing reliance on online platforms, a lot of data gets generated every second. Protecting this huge data from unauthorized access has become a major challenge. Even a minor data breach can lead to significant personal, financial and organizational losses. Hence, cybersecurity has emerged as an essential mechanism for safeguarding data and ensuring trust in digital systems. According to the Cyber Threat Report 2026 released by Seqrite, more than 265 million cyberattacks were recorded in India in 2025. It has also stated that Trojans and File Infectioners account for 70% of all attacks and regions like Maharashtra, Gujarat and Delhi are the most affected regions (Anonymous, 2025). Globally too, cyberattacks has increased sharply, with organizations experiencing nearly 1,925 attacks per week in early 2025—a 47% rise compared to the previous year (Suresh, 2025). The use of artificial intelligence (AI) and Machine Learning (ML) techniques in digital forensics for evidence analysis and cybercrime detection was covered by Akeiber (2025). The article highlighted how advanced methods like ransomware, phishing and deep fake technologies are increasingly being used by cybercriminals. According to Dilek et al., (2015), AI provides adaptive, intelligent and real-time answers to escalating dangers, while traditional defense methods are inadequate. It offers a thorough examination of the ways in which AI techniques—including data mining, expert systems, ML and neural networks are used to identify, stop and address cybercrime. Ghillani (2022) proposes a deep learning-based AI framework to strengthen cyber security. Deep learning uses neural network architectures such as CNNs, RNNs, auto encoders and GANs

to detect, predict, and reduce cyber threats more effectively than traditional rule-based systems.

Dash et al., (2022) emphasizes that hybrid approaches combining AI with human oversight are essential to balance opportunities with emerging risks in cybersecurity. The authors stressed that AI-based intrusion detection systems offer enhanced accuracy, adaptability and real-time threat detection compared to traditional methods. They also highlighted that AI-based intrusion detection systems also face challenges such as adversarial attacks, data dependency and interpretability issues. Treleaven et al., (2023) highlighted that AI and emerging technologies led to a rapid increase in cybercrime, creating unprecedented risks for individuals, businesses and governments. It emphasizes the urgent need for new regulatory, technical and collaborative frameworks to counter this “cybercrime tsunami.”

Adewale (2022) discussed how cybercriminals are exploiting AI technologies to create sophisticated, scalable and hard to detect digital threats. It emphasizes the urgent need for proactive defence strategies and regulatory frameworks to counter this new frontier of cybercrime. Velasco (2022) provides an overview of how international organizations address the intersection of cybercrime and AI within criminal justice systems. The authors examine various applicable international legal instruments and highlight the need for coordinated global frameworks to regulate AI-driven cyber threats. Shetty et al., (2024) examines how AI is reshaping the landscape of cybercrime by enabling more sophisticated, scalable and evasive attacks. It also reviews emerging trends and proposes countermeasures, emphasizing the need for proactive defense strategies, regulatory oversight and hybrid approaches that combine AI with human expertise. The objectives of this chapter are:

- To study various types of traditional and advanced cyber crimes
- To examine the role of AI in cybercrime control
- To review the challenges and limitations of using AI based prevention methods

8.2 Conventional Cyber Crimes

Traditional cybercrimes fall under three major categories, under which there are other subcategories. They are:

8.2.1 Attacks on individuals

These are malicious and deliberate attempts targeted at specific individuals. It is done by a person or group to exploit vulnerabilities in a victim's digital life for harm or gain. These attacks often leverage human psychology and technological weaknesses to steal data, money or disrupt personal life. It includes:

Identity theft

It is a crime where one person, an "identity thief" or "impostor," secretly takes another individual's personal information and uses it to impersonate them, usually for financial gain (Arshad et al., 2021).

Phishing

It is a type of cybercrime where cybercriminals use fake emails or websites to steal sensitive information from individuals. These emails or websites are designed to look like legitimate ones and often include requests for personal information such as passwords or credit card details. Once the information is obtained, the cybercriminals can use it for fraudulent purposes (Arshad et al., 2021).

Cyber stalking

It refers to a perpetrator who uses electronic means to repeatedly harass, intimidate, or monitor a specific victim, causing fear or substantial emotional distress (Kaur et al., 2021).

Cyber bullying

It refers to the use of technology to intimidate, harass or embarrass an individual. Cyber bullies use various platforms such as social media, online chat rooms and instant messaging to target their victims (Macaulay et al., 2022).

Financial fraud

It is done when a person uses another person's credit card information or the physical card itself to make purchases or withdraw money without his/her permission. (Arshad et al., 2021)

Sex tortion

It is a type of blackmail where an individual is threatened with the release of intimate photos or videos unless they pay a ransom. This can occur through online dating scams or other means (Macaulay et al., 2022).

Attacks on property

In this type of cyber-crime, a hacker or malicious user makes use of technology to interfere with an individual's digital or physical property rights. The attackers manipulate human behavior to achieve that goal (Goni et al., 2022).

Distributed Denial-of-Service (DDoS) attacks

It is a cyber attack that overwhelms a target server, website, or network with a massive flood of malicious internet traffic from multiple compromised devices or a botnet, making the service unavailable to legitimate users by over utilizing its resources and bandwidth (Singh and Gupta, 2022).

Malware and Virus Attacks

This involves installing malicious software or viruses on computers or networks to corrupt data, spy on users, or gain unauthorized access (Furnell, 2013).

Copyright Infringement

It includes unauthorized copying, distribution, or use of copyrighted digital content such as software, music, movies, or books (Wall, 2024).

8.2.2 Attacks on institutions and government

Attacks on government systems are considered highly serious cybercrimes as they threaten national security, public safety and critical infrastructure. These crimes are often politically motivated.

Data Breaches

It is a type of cybercrime where a criminal gains unauthorized access to an organization's computer systems and steals sensitive information. It can include personal data, financial information and

intellectual property. Data breaches can result in significant financial losses for organizations as well as harm to their reputation and loss of trust from their customers. There are several ways criminals can carry out a data breach, including exploiting vulnerabilities in software or using stolen login credentials etc. Organizations can protect themselves from data breaches by implementing strong access controls, regularly updating software and systems and training employees on cyber security best practices (Goni et al., 2022).

Business Email Compromise

Business Email Compromise is a sophisticated form of targeted spear phishing. It is a type of cybercrime where criminals use social engineering tactics to trick an organization's high-level employees into transferring money or sensitive information to them. These attacks can be highly sophisticated and convincing and can result in significant financial losses for organizations. Business Email Compromise attacks can be carried through emails or by exploiting vulnerabilities in email systems. Organizations can protect themselves from BEC attacks by implementing strong access controls, training employees on cyber security best practices and using multi factor authentication to prevent unauthorized access (Wassermann et al., 2023).

Insider threats

These are types of cybercrimes that are carried out by individuals who have authorized access to an organization's systems, networks or data. These individuals can be employees, contractors or even third-party vendors who have access to sensitive information. These attacks can include stealing sensitive information, sabotaging computer systems or conducting fraud. Insider threats can be difficult to detect and prevent as the attackers have legitimate access to the organization's system and data. The impact of insider threats can be significant including financial losses, damage to reputation and legal liabilities. Organizations must take proactive measures to protect themselves from insider threats including implementing strong access controls, monitoring employee activity and providing security awareness training to employees (Goni et al., 2022).

Intellectual Property theft

Intellectual Property theft is when a criminal steals an organization's intellectual property such as patents, trade secrets or copyrighted materials (Wall, 2024). This information can be used to gain a competitive advantage or sold to competitors. Intellectual property theft can be carried out through hacking, social engineering or physical theft. Organizations can protect themselves from intellectual property theft by implementing strong access controls, regularly monitoring for unauthorized access and using encryption to protect sensitive data.

Hacking

It is the most common form of cybercrime against the state. Hacking refers to unauthorized access to government networks, databases, or systems to steal sensitive information causing damage or disruption to government operations or stealing money or other assets (Furnell, 2013).

Cyber terrorism

It is another form of cybercrime against the state. It involves the use of internet or computer networks to create fear, panic, or large-scale disruption. It often targets critical infrastructure such as power plants, financial systems, transportation systems, defense systems and communication networks etc. Cyber terrorism can be carried out by individuals or organized groups and it poses a significant threat to a nation's security and stability (Singer and Friedman, 2013).

Cyber Espionage

It involves the use of internet or computer networks to gain access to confidential information or data belonging to a government or private organization. Cyber espionage can be conducted by individuals or state sponsored groups to steal sensitive information such as military secrets and trade secrets. The stolen information can be used for financial gain or to gain a strategic advantage in international relations. (Singer and Friedman, 2013)

Spreading Propaganda

It is relatively new form of cybercrime against the state. It includes the use of digital platforms such as social media platforms to spread false information, fake news, extremist ideologies, or propaganda. The main aim is to manipulate public opinion or creating social unrest. This type of cybercrime can be used to destabilize governments and create chaos.

Cyber Sabotage

It involves the use of the internet or computer networks to damage or destroy a government or private organization's data systems or equipment. Cyber saboteurs may use viruses, worms or other malicious software to cause physical damage to critical infrastructure such as dams, transportation systems, power grids etc. (Singer and Friedman, 2013). Cyber sabotage can also be used to disrupt the normal functioning of government agencies, financial systems and businesses.

8.2.3 Advanced Cyber crimes

Advanced cybercrimes increasingly employ generative AI for automated phishing kit creation, credential harvesting workflows, and large-scale deployment of deceptive websites (Begou et al., 2023). It uses modern technologies such as AI, ML, Deepfakes, Advanced Persistent Threats (APTs) etc. to commit cybercrime. It includes:

AI-powered phishing

AI-powered phishing attacks are becoming highly personalized and scalable (Schmitt and Flechais, 2024). In AI phishing schemes, hackers employ different AI learning techniques, such as behavioral analysis, voice replication, and natural language processing (NLP) models, to generate highly realistic AI-created cyber representations (Muzammil, 2024). Utilizing these technologies, attackers can realistically imitate the looks, voice, or even writing and speaking styles of friends and coworkers, thereby facilitating deception of their targets and enhancing their success rates (Aggrey et al., 2024).

Deep fake fraud

Another common form of cybercrime done using AI is deep fake fraud in which voice/video cloning is used for financial scams. In this, multimodalities such as videos, images, and audio are used to create media deep fake that appear genuine but are totally manipulated. The rise of deep fake technology in cybercrime is a growing threat, and the consequences of falling victim to a deep fake phishing scam can be devastating (Kemp et al., 2025).

APT attacks

Advanced persistent threats (APT) increasingly exploit AI-driven zero-day vulnerability discovery and stealth intrusion techniques (Alansary et al., 2025). An APT attack is a targeted, stealthy cyberattack in which an attacker gains long-term unauthorized access to a system to steal sensitive information. Attacker uses sophisticated techniques such as custom malware, zero-day exploits, and social engineering and does not leave after one attack. They maintain access for weeks, months, or even years without being noticed. The attacker is usually a well-funded group aiming for long-term, stealthy attacks on governments.

AI-based malware

AI based malware is malicious software that can alter its operation, code, and attack vectors in real-time. It is also known as Adaptive Malware as these types of malwares adapt to detection systems. Unlike traditional malware, which follows pre-set instructions, malware powered by AI can learn and change over time to get beyond standard protection mechanisms (Ansarullah et al., 2024). These threats are able to build their own code, change in real time, avoid detection, and even imitate the actions of trustworthy users.

Automated credential stuffing

Credential stuffing is one of the most common and successful forms of cyber-attacks. It is a ML driven password attack that aims to accomplish an account takeover (ATO) for fraudulent misuse. It operates by entering stolen usernames and passwords into the login fields of the system. Some people use the same usernames and passwords to log in on numerous systems, which is why it functions

so well. A significant credential stuffing assault occurred at HSBC in 2018, endangering the financial information of their customers. Begou et al., (2023) demonstrate how ChatGPT can support large-scale phishing deployment by generating obfuscated code, reverse proxy integration, and automated hosting setup.

Data poisoning attacks

It is an adversarial assault that occurs when an attacker purposefully modifies the training data needed to create an AI or ML model. The objective of data poisoning attack is to modify the model's behavior during training so that it continues to behave that way after deployment. This may result in the model exhibiting unpredictable behavior, making inaccurate predictions, or introducing hidden weaknesses. There are various methods for data poisoning such as intentionally adding inaccurate or deceptive data to the training dataset, changing the current dataset, removing a section of the dataset etc. According to experimental data from Kure et al., (2025), poisoning can cause fraud detection models to lose accuracy by up to 22% and image recognition tasks by up to 27%.

Ransomware-as-a-Service (RaaS)

It is a subscription-based ransomware system that makes it possible for even novice cybercriminals to initiate ransomware assaults. RaaS programs remove the necessity for malicious code to be written by attackers. It is therefore widely used by cybercriminals who lack the technological know-how to create ransomware themselves. RaaS is available to anybody by only purchasing a subscription (Akinsuli, 2024).

8.3 Role of Artificial Intelligence In Cybercrime Control

In the rapidly evolving cyber security landscape, AI has become an indispensable asset. Organizations automate the critical processes of identifying, analysing, and pre-emptively mitigating cyber security threats by employing ML and sophisticated AI algorithms (Dash et al., 2022).

8.3.1 AI-Based Cybercrime Prevention Mechanisms

Predictive threat modelling

Predictive threat modelling is a cyber security strategy that aims to foresee assaults before they occur. It analyzes trends, signals and patterns that identify possible hazards using AI and behavioral analytics. By identifying these early warning signs, security teams can prepare in advance, lower risk, and thwart attacks before they can cause any harm (Salem et al., 2024). The fact that predictive models are always learning makes them more appropriate for contemporary, rapidly evolving threat environments.

Early warning systems

Early warning systems alter reactive safety measures into proactive defense strategies (Salem et al., 2024). These systems integrate threat intelligence, behavioral analytics, and automated response capabilities to detect new threats in hybrid cloud environments. It works as an organization's digital radar, looking for new threats and operational anomalies before they can affect business operations. In contrast to typical monitoring solutions that send out alerts after issues happen, EWS platforms detects odd activities such as suspicious authentication attempts, unusual network patterns or configuration changes that indicates possible compromise.

Automated response systems

Automated incident response makes use of software and algorithms to monitor, identify, and address security problems without the need for human involvement. It enables quick reaction to threats by automating repetitive operations. AIR systems guarantee consistent and prompt responses to security problems by utilizing specified protocols and actions (Narayan et al., 2025). Platforms like Security Orchestration, Automation, and Response (SOAR) can be used as AIR systems by the organizations to instantly isolate compromised endpoints, block malicious IP addresses, and kill suspicious processes without human intervention.

AI-driven access control and authentication

AI is revolutionizing the way businesses manage mobility and secure their premises (Balasubramanian et al., 2025). Intelligent

authentication and predictive threat detection are just two examples of how AI-powered access management is rapidly becoming essential across businesses. AI-powered authentication and access control make sure the right people have the correct access at the right time with the least amount of hassle.

Behavioral biometrics

Behavioral biometrics can detect fraud and identity theft by analyzing a user's digital, physical, and cognitive behavior. It uses ML to examine patterns in human behavior. As authenticated users and cyber criminals engage with digital platforms in different ways, behavioral biometrics are used to differentiate between authenticated users and cybercriminals. Using ML, it determines whether an individual is who they say they are when interacting online and whether their actions are the result of an automated assault or human intervention.

8.4 Challenges And Limitations of Using AI Based Prevention Methods

8.4.1 Adversarial attacks on AI models

A developing danger that targets the fundamental principles of ML model operation is adversarial AI attacks. An adversarial AI attack is a malevolent method that manipulates ML models by purposefully giving them false information in order to produce inaccurate or unexpected behavior (Malatji and Tolah, 2025). They can frequently get past traditional security measures and are made to be undetectable. Different adversarial AI attacks such as evasion, poisoning, model extraction and inversion attacks are used to target ML systems.

8.4.2 Data bias and fairness issues

One key challenge is the reliance on accurate data. Incomplete or outdated information can lead to incorrect permissions or missed detection of unauthorized access. Regular audits and updates to your infrastructure and permission data can help ensure your AI works effectively (Grgic-Hlaca et al., 2024).

8.4.3 Model interpretability (black-box problem)

Many AI/ML models, sometimes known as "black boxes," are opaque, which presents serious problems for interpretability, transparency, and trust. This lack of explainability can be particularly problematic in fields like malware analysis and network intrusion detection, where comprehending the logic underlying model predictions is essential for efficient decision-making and system enhancement (Reynaud and Roxin, 2025).

8.4.4 Scalability and real-time constraints

Scalability and real time constraints are two of the biggest, most interconnected problems in cyber security. Security tools must handle enormous data volumes and react to threats instantaneously as IT environments develop into sophisticated, fast-moving, and dispersed ecosystems (Boateng et al., 2026). Failure to do so results in performance bottlenecks, elevated susceptibility, and high false-positive rates.

8.4.5 Ethical and legal challenges

Despite the establishment of privacy rules by legal frameworks like the CCPA and GDPR, there are still difficulties in upholding ethical standards in cyber security and AI (Ilari et al., 2026). While the Facebook-Cambridge Analytica case emphasizes the dangers of uncontrolled data usage, growing AI adoption calls for ethical standards to guard against bias and discrimination. Addressing legal and ethical issues in the digital age requires stronger regulatory frameworks, corporate social responsibility, and digital literacy.

8.5 Conclusion

Businesses need to keep changing and using the latest cyber security tools to be safe. By using AI-enhanced cyber security frameworks and leveraging its knowledge, one may keep ahead of the ever-evolving cyber threat landscape and ensure comprehensive protection against advanced cyber threats. Businesses must remain ahead of the curve as cyber dangers continue to change. Protecting against the upcoming generation of cyber threats will require implementing AI-driven cyber security protection, utilizing sophisticated threat intelligence, and implementing behavior-based security models. Through the

fusion of traditional approaches and AI techniques, businesses that make investments in scalable, AI-ready infrastructure now will be in the best position to securely and smoothly implement tomorrow's advancements, resulting in environments that are safer and more secure.

References

- [1] Adewale, T. (2022). Artificial Intelligence in Cybercrime: Unveiling the Emerging Landscape of Intelligent Threats.
- [2] Aggrey, R., Adjei, B. A., Afoduo, K. O., Dsane, N. A. K., Anim, L., and Ababio, M. A. (2024). Understanding and Mitigating AI-powered Cyber-Attacks. *International Journal for Multidisciplinary Research*, 13(1), 1–21. https://www.computerscijournal.org/pdf/vol13no1/OJCST_Vol13_N1_p_01-21.pdf
- [3] Akeiber, H. J. (2025). A comprehensive study of cybercrime and digital forensics through machine learning and AI. *Al-Rafidain Journal of Engineering Sciences*, 3, 369–395. <https://doi.org/10.61268/hff1pp49>
- [4] Akinsuli, O. (2024). The rise of AI-enhanced ransomware-as-a-service (RaaS): A new threat frontier. *World Journal of Advanced Engineering Technology and Sciences*, 1(2), 85–97. <https://doi.org/10.30574/wjaets.2021.1.2.0019>
- [5] Alansary, S. A., Ayyad, S. M., Talaat, F. M., and Saafan, M. M. (2025). Emerging AI threats in cybercrime: A review of zero-day attacks via machine, deep, and federated learning. *Knowledge and Information Systems*, 67, 10951–10987. <https://doi.org/10.1007/s10115-025-02556-6>
- [6] Anonymous (2025). India records 265 million cyber attacks in 2025: Report. TOI Tech Desk. Times of India. <https://timesofindia.indiatimes.com/technology/tech-news/india-records-265-million-cyber-attacks-in-2025-report/articleshow/125772984.cms>
- [7] Ansarullah, S. I., Wali, A. W., Rasheed, I., and Rayees, P. Z. (2024). AI-powered strategies for advanced malware detection and prevention. In *The art of cyber defense* (pp. 3-24). CRC Press.
- [8] Arshad, A., Rehman, A. U., Javaid, S., Ali, T. M., Sheikh, J. A., and Azeem, M. (2021). A systematic literature review on phishing and anti-phishing techniques. *arXiv*, 13(3), 1-30. <https://doi.org/10.48550/arXiv.2104.01255>
- [9] Balasubramanian, P., Liyana, S., Sankaran, H. et al. (2025). Generative AI for cyber threat intelligence: applications, challenges, and analysis of real-world case studies. *Artificial Intelligence Review*, 58, 336. <https://doi.org/10.1007/s10462-025-11338-z>
- [10] Begou, N., Vinoy, J., Duda, A., and Korczyński, M. (2023). Exploring the dark side of ai: Advanced phishing attack design and deployment using

- chatgpt. In *2023 IEEE Conference on Communications and Network Security (CNS)* (pp. 1-6). IEEE. [10.1109/CNS59707.2023.10288940](https://doi.org/10.1109/CNS59707.2023.10288940)
- [11] Boateng, Y. J., Mim, N. J., Akhter, N., Naha, R., Mahanti, A., and Barros, A. (2026). Application of AI in Cyberattack Detection: A Review. *Sensors*, 26(5), 1518. <https://doi.org/10.3390/s26051518>
- [12] Dash, B., Ansari, M. F., Sharma, P., and Ali, A. (2022). Threats and opportunities with AI-based cyber security intrusion detection: A review. *International Journal of Software Engineering and Applications (IJSEA)*, 13(5), 13-21. <https://doi.org/10.5121/ijsea.2022.13502>
- [13] Dilek, S., Çakır, H., and Aydın, M. (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. *International Journal of Artificial Intelligence and Applications*, 6, 21-39. <https://doi.org/10.48550/arXiv.1502.03552>
- [14] Furnell, S. (2013). Hackers, viruses and malicious software. In: *Handbook of internet crime* (pp. 173-193). Willan, London. <https://doi.org/10.4324/9781843929338>
- [15] Ghillani, D. (2022). Deep learning and artificial intelligence framework to improve the cyber security. *Authorea Preprints*. <https://www.authorea.com/users/506161/articles/587142-deep-learning-and-artificial-intelligence-framework-to-improve-the-cyber-security>
- [16] Goni, O., Ali, M. H., Showrov, M. M. A., and Shameem, M. A. (2022). The basic concept of cyber crime. *Journal of Technology Innovations and Energy*, 1(2), 16-24. <https://doi.org/10.5281/zenodo.6499991>
- [17] Grgić-Hlača, N., Ali, J., Gumadi, K. P., and Wortman Vaughan, J. (2024). (De) Noise: Moderating the Inconsistency Between Human Decision-Makers. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW2), 1-38. <https://doi.org/10.1145/3686987>
- [18] Ilari, L., Tiribelli, S., and Caruso, F. (2026). From AI security to ethical AI security: A comparative risk-mitigation framework for classical and hybrid AI governance. *AI and Ethics*, 6, 91. <https://doi.org/10.1007/s43681-025-00935-x>
- [19] Kattinig, M., Angerschmid, A., Reichel, T., and Kern, R. (2024). Assessing trustworthy AI: Technical and legal perspectives of fairness in AI. *Computer Law and Security Review*, 53, 106053. <https://doi.org/10.1016/j.clsr.2024.106053>
- [20] Kaur, P., Dhir, A., Tandon, A., Alzeiby, E. A., and Abohassan, A. A. (2021). A systematic literature review on cyberstalking. An analysis of past achievements and future promises. *Technological Forecasting and Social Change*, 163, 120426. <https://doi.org/10.1016/j.techfore.2020.120426>

- [21] Kemp, M., Kalutarage, H., and Al-Kadri, M. O. (2025). AI-Powered Spearphishing Cyber Attacks: Fact or Fiction?. *arXiv preprint arXiv:2502.00961*. <https://arxiv.org/pdf/2502.00961>
- [22] Kure, H. I., Sarkar, P., Ndanusa, A. B., and Nwajana, A. O. (2025, May). Detecting and preventing data poisoning attacks on AI models. In *2025 Photonics and Electromagnetics Research Symposium-Spring (PIERS-Spring)* (pp. 01-12). IEEE. <https://doi.org/10.1109/PIERS-Spring66516.2025.11276197>
- [23] Macaulay, P. J., Betts, L. R., Stiller, J., and Kellezi, B. (2022). An introduction to cyberbullying. In *Cybersecurity and cognitive science*, 197-213. <https://doi.org/10.1016/B978-0-323-90570-1.00002-4>
- [24] Malatji, M., Tolah, A. (2025) Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. *AI Ethics*, 5, 883–910. <https://doi.org/10.1007/s43681-024-00427-4>
- [25] Muzammil, K. (2024). Aalpha Information Systems. (n.d.). Artificial intelligence and machine learning in cyber security. <https://www.aalphanet/articles/artificial-intelligence-and-machine-learning-in-cyber-security/>
- [26] Narayan, Y., Gupta, A., Charan, P., Lakshmi, T. H., Ghumman, M. K., and Saghra, H. S. (2025, September). Strengthening Network Security through the Implementation of AI-Driven Automated Incident Response Systems. In *2025 IEEE International Conference on Compute, Control, Network and Photonics (ICCCNP)* (pp. 1-5). <https://doi.org/10.1109/ICCCNP63914.2025.11233253>
- [27] Reynaud, S., and Roxin, A. (2025). Review of eXplainable artificial intelligence for cybersecurity systems. *Discover Artificial Intelligence*, 5(1), 78.
- [28] Salem, A. H., Azzam, S. M., Emam, O. E., and Abohany, A. A. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. *Journal of Big Data*, 11, 128. <https://doi.org/10.1186/s40537-024-00957-y>
- [29] Schmitt, M., and Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. *Artificial Intelligence Review*, 57, 324. <https://doi.org/10.1007/s10462-024-10973-2>
- [30] Shetty, S., Choi, K. S., and Park, I. (2024). Investigating the intersection of AI and cybercrime: Risks, trends, and countermeasures. *International Journal of Cybersecurity Intelligence and Cybercrime*, 7(2), 28–53. <https://doi.org/10.52306/2578-3289.1187>
- [31] Singer, P. W., and Friedman, A. (2013). *Cybersecurity and Cyberwar: What Everyone Needs to Know®*. Oxford University Press.

- [32] Singh, A., and Gupta, B. B. (2022). Distributed denial-of-service (DDoS) attacks and defense mechanisms in various web-enabled computing platforms: Issues, challenges, and future research directions. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 18(1), 1-43. <https://doi.org/10.4018/IJSWIS.297143>
- [33] Suresh, L. (2025). Global cyberattacks surge sharply; new data shows unprecedented rise in threats worldwide. *The Week*. <https://www.theweek.in/news/defence/2025/12/04/global-cyberattacks-surge-sharply-new-data-shows-unprecedented-rise-in-threats-worldwide.html>
- [34] Treleaven, P., Barnett, J., Brown, D., Bud, A., Fenoglio, E., Kerrigan, C., Koshiyama, A., Sfeir-Tait, S. and Schoernig, M. (2023). The Future of Cybercrime: AI and Emerging Technologies Are Creating a Cybercrime Tsunami. Available at SSRN: <https://ssrn.com/abstract=4507244> or <http://dx.doi.org/10.2139/ssrn.4507244>
- [35] Velasco, C. (2022) Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, 23, 109–126. <https://doi.org/10.1007/s12027-022-00702-z>
- [36] Wall, D. S. (2024). *Cybercrime: The transformation of crime in the information age*. John Wiley and Sons.
- [37] Wassermann, S., Meyer, M., Goutal, S., and Riquet, D. (2023). Targeted attacks: Redefining spear phishing and business email compromise. *arXiv preprint arXiv:2309.14166*. <https://doi.org/10.48550/arXiv.2309.14166>

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 9

ARTIFICIAL INTELLIGENCE IN NANOTECHNOLOGY

Shraddha S. Khamkar* & Jayashri Naphade

Assistant Professor, Department of Chemistry, VIVA College of Arts, Science and Commerce, Virar (West), Palghar, India-400303.

Email*: shraddha.khamkar@vivacollege.org

Abstract

Artificial Intelligence (AI) is emerging as a powerful enabler of nanotechnology's advancement. It offers a transformative approach to the design, characterization and application of nanoscale systems. This chapter explores the convergence of AI techniques – specifically, machine learning, deep learning and data-driven modeling with nanoscale science and engineering. It mainly focuses on how AI facilitates the prediction of material properties, accelerates the discovery of novel nanomaterials and optimizes synthesis and fabrication processes with improved accuracy and efficiency. The chapter further examines the role of AI in interpreting complex experimental data, like microscopy and spectroscopy outputs, which enable the enhanced visualization and understanding of nanoscale phenomena. Applications across key domains, including nanomedicine, nanoelectronics, energy storage and environmental remediation, are also discussed to explain the broad impact of this interdisciplinary integration. The chapter also explains the current challenges, which include the limited high-quality datasets, module interpretability and the need for robust integration between the computational and experimental workflows. Overall, this chapter will provide a comprehensive overview of how AI is reshaping nanotechnology, which offers insights into the current capabilities, while identifying the pathways for future innovation at the atomic and molecular levels.

Keywords: nanoscale, nanomaterials, artificial intelligence, disease diagnosis

9.1 Introduction

Recently, artificial intelligence (AI) has developed so much among other transformative fields. AI is a technology that develops systems capable of thinking and performing tasks just like humans. AI and nanotechnology have come up as the most significant fields in the 21st century. Their correlation is always a fascinating topic for researchers. AI technology works by training on existing data and restructuring data processing. It solves complex problems and performs automation in diverse areas. On the other hand, nanotechnology enables the transformation of materials at the atomic level, thereby forming materials with distinct characteristics and functions (Li et al., 2025).

Nanotechnology is a branch of science that deals with the engineering and designing of materials and devices with the size ranging between 1 and 100 nanometers (nm)—where one nanometer is one billionth of a meter. Nanomaterials are novel materials with features offering better performance. These kinds of materials being at nanoscale have better physical, chemical, biological, and mechanical properties ideal for their usage in multiple fields. The science behind nanotechnology uses these properties to optimize solutions for issues in the field of medicine, electronics, environmental science, etc. (Sohail, 2025). Nanotechnology works in interdisciplinary fields with its application in physics, chemistry, biology, engineering, materials science, and computer science.

AI refers to a group of technologies that enables computers to learn and make decisions, doing things that were previously done by humans, like processing language, analyzing information, and making useful recommendations. AI is an innovative technology with the power to create substantial and constructive transformations in people's lives and in the world at large. It cuts across a wide range of fields, from computer science to data analysis and statistics, software engineering, linguistics, neuroscience, and philosophy, among others. AI is about teaching computers to do the amazing things our own brains can do (Golec et al., 2025). For instance, AI helps

in optical character recognition (OCR), where text/data is extracted from a variety of images/documents. It converts unstructured data into structured data useful for business purposes. This helps in gaining valuable insights. AI may not be one particular technological application; however, it covers a wide range of areas such as machine learning (ML), deep learning (DL), natural language processing (NLP), and computer vision.

Both nanotechnology and AI are gaining a lot of focus nowadays and have the ability to transform several facets substantially. The combination of nanotechnology and artificial intelligence has emerged as a powerful tool in revamping industries and disclosing new potentials in science, technology, and medicine. The unique features of nanotechnology with the power of AI are helping scientists and researchers achieve groundbreaking results in medical care, energy, technology, and other fields (Chakroborty et al., 2025). The field of computer science that focuses on the creation of intelligent machines that are capable of performing activities that often call for human intelligence is known as artificial intelligence, or simply AI (Kaushik et al., 2023). When AI and nanotechnology converge, the potential for a breakthrough is vast (Rawat et al., 2023). Cascini et al. (2022) reported that the combination of AI with nanotechnologies has the potential to help in the advancement of several fields, like energy-efficient products, personalized medicine, diagnosis and better treatment of disease, and smart environmental monitoring systems (Cascini et al., 2022). Though nearly all AI-based successful paradigms like artificial neural networks and algorithmic evolution are biology-based, nanotechnology combines the basics of chemistry, physics, and engineering. Therefore, the developments are bridging the gaps between nanoscience and AI. It is predicted from decades that AI-incorporated nanoscience will be great to build new communication and technologies. The approaches in scientific and engineering fields that develop, manufacture, and evaluate devices and materials having functioning structure at nanoscale are collectively referred to as "nanotechnology." It affects the basic molecular structure of a material and resizes it to nano-dimensions, which results in chemical and physical properties. It is substantially critical to find and observe the behavior of a single molecule or group of interacting molecules to infer the major properties of the material or devices at the nano level (Saini et al., 2010). Nanotechnology

is an interdisciplinary branch with multiple roles in biomolecular detection, therapeutics, microelectronics, DNA sequencing, optics, and sensors. In recent times, it has gained huge importance in the field of pharmaceutical sciences.

Nanotechnology has grown swiftly in previous years, impacting sectors, be it electronics, material science, or medical science. At the same time, AI has emerged greatly, powered by developments in machine learning, computational power, and data analytics. Both branches individually play their vital role in revolutionizing the research in specific domains; their merger possesses even greater power, especially in healthcare sectors, where both precision and efficiency are foremost (Jaber et al., 2024).

The combination of AI and nanotechnology even suggests unending chances in the early detection and handling of diseases, particularly at the molecular level. A major instance of AI and nanotechnology is always healthcare, since in this sector AI helps in the analysis of complex biological data that are generated by nano-devices, thus enabling the early diagnosis and treatment of probable fatal diseases such as cancer, cardiovascular disorders, and neurological conditions. Further, reports suggest nanotechnology enhances the precision and efficacy of drug delivery, minimization of the side effects, and maximization of therapeutic outcomes. The role of AI and nanotechnology in the creation of efficient environmental monitoring systems and sustainable energy solutions has also been reported by Malik et al. (2023). But as it is well understood, both technologies may have cons as well, which are not to be forgotten.

9.2 Applications of Nanotechnology And Artificial Intelligence In Healthcare

The integration of superior features of nanomaterials and analytical capabilities of artificial intelligence is steering in a new era of innovation in health care. Several applications have been enlisted below:

9.2.1 Detection of Early Disease and Diagnosis

The disease detection is the most crucial step for its cure. The delay in diagnosis can impact the efficacy of standard treatments. The timeline plays a pivotal role in healthcare. Here is the revolutionizing

role of AI based nanotechnology in healthcare, it helps in early disease detection. The conventional methods are laborious and time consuming which alter the detection times significantly, leading to delayed treatment and poor outcomes. Even if there are no symptoms nano devices like nano-sensors helps in detection of the molecular changes at the cellular level and AI algorithms, can easily analyse large data generated by these devices in real time. Hence, by identifying patterns, the diseases like neuron disorders, cardiovascular problems and even cancer like diseases can be caught at early stages. For example, AI-powered nano sensors can be used to detect cancer biomarkers in blood samples with high specificity, enabling earlier and more accurate diagnosis. Additionally, AI models can predict disease progression, which is based on nanoscale data, thus allowing healthcare providers to meet the treatment plans for individual patients (Kaur et al., 2022).

9.2.2 AI-assisted Surgical Tools and Nano Robotics

Nano robots are the tiny machines capable of navigating through the bloodstream. They are being developed to perform precise medical interventions at the cellular level. The Nano robots can be programmed to repair the damaged tissues and also to clear blockages present in the blood vessels. When these are guided by the AI algorithm, the Nano robots make autonomous decisions that are based on real-time data. Their actions are adjusted so as to optimize the outcomes of the patients. In surgery, the AI-assisted tools that incorporate the nanoscale technology are being developed to enhance the precision of the procedures. They also minimize invasiveness and reduce the recovery time. Surgeons can use these AI-guided Nano devices for real-time imaging and navigation, which allows more accurate and less traumatic interventions (Ansari et al., 2025).

9.2.3 Technological Innovations in the Convergence of Nanotechnology and AI

The convergence of nanotechnology and artificial intelligence has led to creative innovations that are reshaping various sectors particularly healthcare. The advancements arise from the ability of nanotechnology to manipulate matter at the atomic level and the capability of AI to process and also to analyze the vast datasets, pattern recognition and other important decision-making tasks.

By combining the precision offered by nanotechnology and the predictive authority of AI, healthcare sector can achieve exceptional accuracy in the diagnosis of several diseases and their treatment. Besides healthcare, blending of these two technologies can offer abundant applications in the field of environment and manufacturing of smart materials thus improving the quality of life globally (Elqassas et al., 2024).

9.2.4 Nano-scale imaging and AI-driven analysis

This is one of the most important areas where innovation lies in nano-scale imaging technologies. Nanotechnology allows the creation of highly detailed images at the molecular level. But, here, the analysis of the images has been labor-intensive historically and also time-consuming. AI plays a very important role in the interpretation and processing of complex data from images at a higher speed and accuracy. Machine learning algorithms are also being trained to detect nanoscale images for many applications. They include identifying cancer cells, detecting the presence of pathogens and also observing the biological processes at a molecular level, e.g., AI-powered nanoscale imaging systems can detect the abnormalities in the cells before there are visible symptoms in a patient, which facilitates the early diagnosis and therefore more specific and effective treatments can be given (Egwuatu et al., 2024).

9.2.5 AI Optimized Nanomaterials for Drug Discovery

In the field of drug discovery, researchers are developing nanoparticles for targeted treatment of diseases. AI has significantly enhanced the development of these nanomaterials by optimizing the design for a specific therapeutic purpose. By machine learning, researchers can now predict how the different nanoparticles will interact with the biological systems. This enables the development of more effective and safer drug delivery methods, e.g., the AI models are used to simulate the nanoparticles to travel through the bloodstream and get attached to the target cells to release their therapeutic effect. The optimization process reduces the time and the cost associated with the trial-and-error approaches. It also fastens the creation of nanomaterials that can target the site of disease, such as tumors, without harming the adjoining healthy tissues (Gholap et al., 2024).

9.2.6 Nano Robots and AI Driven Autonomy

The integration of artificial intelligence and micro/ Nano robotics is fundamentally reshaping bio-sensing by enabling autonomous, high-resolution and adaptive biological analysis. One of the most futuristic innovations emerging from this convergence is the development of nano-robots. These smart Nano robots have the potential to revolutionize minimal invasive medical treatments, thus improving the outcomes and reducing the time required for the recovery (Selvi et al., 2025). AI enhances these Nano robots by providing them with autonomous decision-making capabilities. This allows them to navigate complex biological environments and also to respond to the real-time conditions. These miniaturized robotic systems, fabricated by using advanced techniques such as soft lithography, nanoimprinting and self-assembly, can navigate complex biological environment to perform the targeted sensing and therapeutic delivery (Sharma et al., 2025). For example, AI-powered nano robots can monitor changes in a patient's physiological condition and adjust their actions. In the treatment of cancer, nano-robots can be programmed to identify the malignant cells and release drugs directly into them. This minimizes the side effects and also increases the treatment efficacy.

9.2.7 AI-guided Nano Manufacturing and Material Design

Nanotechnology depends on the precision manufacturing at an atomic scale. AI is playing a transformative role in this area. AI-guided nano manufacturing processes enable the design of new nanomaterials with the properties for some specified applications. This includes the development of materials with enhanced conductivity, strength and reactivity, which can be used in various industries, from healthcare to electronics. AI is also speeding up the discovery of new nanomaterials by analyzing the datasets of molecular structures. It also predicts the behavior under different conditions. This capability allows for the fast prototyping of nanomaterials. These nanomaterials are suitable for particular functions, such as lightweight nanocomposites for energy-efficient catalysts for industrial processes (Nandipati et al., 2024).

9.2.8 Real-time Health Monitoring with Nano sensors

The integration of nanotechnology in health care, with its unique and versatile properties, has brought a revolution, such as disease detection and management. This involves nano-scale sensors which can continuously monitor the physiological parameters such as heart rate, metabolic activity, blood glucose, etc. (Gulati et al., 2025). When combined with AI, these sensors can provide real-time insights for a patient's health. This enables personalized and predictive healthcare. The unique physical and chemical properties of the nanoparticles provide a basic structural framework on which to build the successive chemical and biological detection systems.

9.3 Applications In Daily Use

Nano-scale additives in polymer composites are used for making tennis rackets, bicycles, motorcycle helmets, automobile parts, luggage, tennis balls and power tool housings. These additives increase the physical properties of the products by making them stiff, durable, lightweight and resilient (Poluektova et al., 2017). These are very fine films present on laptops, mobile screen displays, facades and eyeglasses. When used as surfaces, these additives can make them water and residue repellent, resistant to ultraviolet light, anti-fog, antimicrobial, scratch resistant, etc. (Li et al., 2019). Nanomaterials in Automotive products – These include the high-performance rechargeable battery systems. They are used for temperature control, as a fuel additive for cleaning the exhaust and for extended range. Nano structured ceramic coatings exhibit much greater toughness than conventional wear-resistant coatings for machine parts (Zois et al., 2009). Nano structured smart fabrics are equipped with flexible nanoscale sensors and electronics with features for health monitoring, energy harvesting through movement, etc. (Castano and Flatau, 2014). Personal Care Products synthesized using Nanotechnology includes titanium dioxide and zinc oxide, which have been used in sunscreens to protect from the sun while appearing invisible on the skin (Adawi et al., 2018).

9.4 Ethical And Regulatory Challenges In The Convergence of Nanotechnology And AI

Nanotechnology and Artificial Intelligence are being integrated extensively from past few years but along with offering abundant benefits their amalgamation has resulted in significant ethical and regulatory challenges. For the long term success and acceptability of these technologies, it is important that these technologies must be used responsibly and ensuring absolute ethical safety. The blend of these technologies unlocks plentiful applications but it demands strict vigilance for the privacy and accountability of the data provided. In healthcare sector, patient's privacy and AI driven treatment and decision making plays a vital role. AI systems can pose a serious threat to human safety if they are not carefully designed. AI could eventually become so intelligent that it surpasses human intelligence, thus leading to a scenario in which AI systems control our lives (Bostrom, 2014).

9.4.1 Ethical Concerns in AI-Driven Medical Decisions

AI based technologies are being used to predict medical condition and decide the course of action in combination with nano science-based technologies but in all these cases, the question arises regarding the accountability and transparency. AI based results can sometimes provide vague algorithms that are difficult to understand (Radanliev et al., 2025). Patients and healthcare providers have to believe that AI driven decisions are accurate and impartial. AI systems used in combination with nanotechnology must be completely explainable with clear guidelines on human oversight and decision-making authority.

9.4.2 Regulatory Framework and Global Standards

To develop a comprehensive regulatory framework that will govern a safe and ethical use of these technologies is very important. These regulations must include issues of safety, transparency, accountability and efficacy. The rapid speed of innovation in nanotechnology and AI presents challenges for the regulatory bodies, which often struggle to keep up with technological advances. One more challenge is about harmonizing regulatory standards across countries (Firoozi et al., 2025). Ensuring that the regulatory frameworks are consistent and

compatible worldwide is very important so as to prevent the gaps that could lead to unsafe or unethical applications of these technologies.

9.4.3 Environmental Impact and Safety

Nanoparticles, due to their small size, can potentially accumulate in organs or also the ecosystems, which can lead to dangerous consequences. Nanotechnology is increasingly combined with AI-powered systems, specifically in the areas of healthcare and environmental applications. It is hence essential to assess the safety of all these materials. Moreover, the environmental impact of nanotechnology especially its potential toxicity to wildlife and ecosystems, must be taken care of. AI can also help in predicting these impacts by replicating the interactions of the nanomaterials with biological and environmental systems. Therefore, a regulatory framework is required to ensure that nanomaterials are properly tested before they are deployed (Mosa et al., 2024).

9.4.4 Risk of Nanotechnology Misuse

The ability of nanotechnology to manipulate the matter at the atomic level offers a remarkable potential, but it also has a concern about its misuse. For example, if in the wrong hands, nanotechnology could be weaponized to create harmful agents, such as nanobots, which are capable of causing damage to biological systems. In healthcare, nanotechnology can be used for drug delivery and diagnostic purposes, but if not properly regulated, then there is also a risk of unsafe nanomaterials entering the market. Strict regulations are required here to prevent the misuse and also to ensure that nanotechnology and AI are applied only for beneficial and safe purposes (Ray et al., 2009).

9.5 Conclusion

AI has the potential to revolutionize patient care, improve public health surveillance, personalize treatment programs and increase diagnostic accuracy. Combining AI with nanotechnology represents a fundamental paradigm shift with wide-ranging implications for businesses and scientific domains. The convergence of nanotechnology and artificial intelligence (AI) holds immense promise across various sectors, with healthcare at the forefront. AI-driven nanotechnology is revolutionizing healthcare and illness management by fastening the

creation of customized drugs and targeted therapies. The integration of AI's computational process with the precision of nanotechnology is revolutionizing diagnostics, personalized medicine, drug delivery and even real-time health monitoring. These creations have the potential to improve patient outcomes and also make healthcare more predictive and effective. Beyond healthcare, the application of these technologies also extends to industries such as materials science and manufacturing, offering sustainable and efficient solutions for global challenges. Furthermore, global regulatory standards must also evolve to keep pace with technological advances to ensure safety, accountability and transparency. Collaboration between governments, researchers and industry leaders will be the key to achieving this goal, thus promoting fairness in the deployment of these technologies. Ultimately, the convergence of nanotechnology and AI is poised to revolutionize healthcare and many more. The success of this will depend on the ability to navigate the ethical and regulatory challenges it presents. By balancing innovation with responsibility, we can definitely harness the transformative technologies for the benefit of society as a whole.

References

- [1] Adawi, H. I., Newbold, M. A., Reed, J. M., Vance, M. E., Feitshans, I. L., Bickford, L. R., and Lewinski, N. A. (2018). Nano-enabled personal care products: Current developments in consumer safety. *Nano Impact*, 11, 170-179. [10.1016/j.impact.2018.08.002](https://doi.org/10.1016/j.impact.2018.08.002)
- [2] Ansari, Z. J., Aher, A., and Thitame, S. N. (2025). Advancements in robotics and AI transforming surgery and rehabilitation. *Journal of Pharmacy Bioallied Science*, 17(1), S46–S48. [10.4103/jpbs.jpbs_1937_24](https://doi.org/10.4103/jpbs.jpbs_1937_24).
- [3] Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press, 352, ISBN 978-0-19-967811-2.
- [4] Cascini, F., Beccia, F., Causio, F. A., Melnyk, A., Zaino, A., and Ricciardi, W. (2022). Scoping review of the current landscape of AI based applications in clinical trials, *Frontiers in Public Health*, 10, 949377. <https://doi.org/10.3389/fpubh.2022.949377>.
- [5] Castano, L. M., and Flatau, A. B. (2014). Smart fabric sensors and e-textile technologies: a review. *Smart Materials and structures*, 23(5), 053001. [10.1088/0964-1726/23/5/053001](https://doi.org/10.1088/0964-1726/23/5/053001)
- [6] Chakraborty, S., Nath, N., Sahoo, S., Singh, B. P., Bal, T., Tiwari, K., Hailu, Y. K., Singh, S., Kumar, P. and Chakraborty, C. (2025). A review of emerging trends in nanomaterial-driven AI for biomedical applications.

- Nanoscale Advances*, 7(12), 3619–3630. <https://doi.org/10.1039/d5na00032g>.
- [7] Egwuatu, O. P., Ori, M. O., Samuel, H. S. and Ekpan, F. M. (2024). AI-enabled Diagnostics and Monitoring in Nanomedicine. *Eurasian Journal Science and Technology*, 4(3), 208-229. <https://doi.org/10.48309/EJST.2024.42>.
- [8] Elqassas, R., Alrakhawi, H. A., Elsobeihi, M. M., Habil, B., Abu-Nasser, B. S., and Abu-Naser, S. S. (2024). Convergence of Nanotechnology and Artificial Intelligence: Revolutionizing Healthcare and Beyond. *International Journal of Academic Engineering Research (IJAER)*, 8(10): 25-30.
- [9] Firoozi, A. A., Firoozi, A. A., and Maghami, M. R. (2025). Transformative Impacts of Nanotechnology on Sustainable Construction: A Comprehensive Review, *Results in Engineering*, 26, 104973. <https://doi.org/10.1016/j.rineng.2025.104973>.
- [10] Gholap, A. D., Uddin, M. J., Faiyazuddin, M., Omri, A., Gowri, S., and Khalid, M. (2024). Advances in artificial intelligence in drug delivery and development: a comprehensive review. *Computational Biology Medicine*, 178, 108702, <https://doi.org/10.1016/j.compbiomed.2024.108702>.
- [11] Golec, M., Hatay, E. S., Gill, S. S., and Buyya, R. (2025). Artificial Intelligence (AI): Foundations, trends and future directions, *Telematics and Informatics Report*, 20. <https://doi.org/10.1016/j.teler.2025.100265>.
- [12] Gulati, S., Yadav, R., Kumari, V., Nair, S., Gupta, C., and Aishwari, M. (2025). Nanosensors in healthcare: transforming real-time monitoring and disease management with cutting-edge nanotechnology, *RSC Pharmacy*, 2, 1003-1018. <https://doi.org/10.1039/D5PM00125K>.
- [13] Jaber, H. A., Saleh, Z., Jabber, W., Badran, A., and Nasser, H. (2024). Artificial Intelligence Integration With Nanotechnology. *Artificial Intelligence in the Age of Nanotechnology*. IGI Global, 25-41. <https://doi.org/10.4018/979-8-3693-0368-9>.
- [14] Kaur, D., Uslu, S., Rittichier, K. J., and Durrezi, A. (2022). Trustworthy, Artificial intelligence: a review, *ACM Computing Surveys*, 55, 1–38. <https://doi.org/10.1145/3491209>
- [15] Kaushik, S. and Suhag, D. (2023). Nanotechnology for AI in Healthcare. *Integrated Nanomaterials and their Applications*. Singapore: Springer Nature Singapore, 291-310. https://doi.org/10.1007/978-981-99-6105-4_14.
- [16] Li, F., Xu, J. (2025). Revolutionizing AI-Enabled Information Systems Using Integrated Big Data Analytics and Multi-Modal Data Fusion. *IEEE Access*. <https://doi.org/10.1109/2025.355203>.

- [17] Li, X., Shi, T., Li, B., Chen, X., Zhang, C., Guo, Z., and Zhang, Q. (2019). Subtractive manufacturing of stable hierarchical micro-nano structures on AA5052 sheet with enhanced water repellence and durable corrosion resistance. *Materials & Design*, 183, 108152. [10.1016/j.matdes.2019.108152](https://doi.org/10.1016/j.matdes.2019.108152)
- [18] Malik, S., Muhammad, K., and Waaheed, Y. (2023). Emerging Applications of Nanotechnology in Healthcare and Medicine, *Molecules*, 28(18), 6624. <https://doi.org/10.3390/molecules28186624>.
- [19] Mosa, M. J., barhoom, A. M., Alhabbash, M. I., Harara, F. E., Abu-Nasser, B. S., and Abu-Naser, S. S. (2024). AI and Ethics in Surveillance: Balancing Security and Privacy in a Digital World. *International Journal of Academic Engineering Research (IJAER)*, 8(10), 10-17.
- [20] Nandipati, M, Fatoki, O., and Desai, S. (2024). Bridging Nanomanufacturing and Artificial Intelligence—A Comprehensive Review. *Materials*, 17(7), 1621. <https://doi.org/10.3390/ma1707162>.
- [21] Poluektova, V. A., Shapovalov, N. A., and Vladykin, V. N. (2017). From nano-scale to macrostructure in composite for additive technologies. In *International Conference Actual Issues of Mechanical Engineering* (pp. 614-619). Atlantis Press.
- [22] Radanliev, P. (2025). AI Ethics: Integrating Transparency, Fairness and Privacy in AI Development, *Applied Artificial Intelligence*, 39(1). <https://doi.org/10.1080/08839514.2025.2463722>.
- [23] Rawat, B., Bist, A. S., Supriyanti, D., Elmanda, V., and Sari, S. N. (2023). AI and nanotechnology for healthcare: A survey. *APTISI Transactions on Management*, 7(1), 86-91. <https://doi.org/10.33050/atm.v7i1.1819>.
- [24] Ray, P. C., Yu, H., and Fu, P. P. (2009). Toxicity and environmental risks of nanomaterials: challenges and future needs. *Journal of Environmental Science and Health Part C*, 27(1), 1–35. <https://doi.org/10.1080/10590500802708267>.
- [25] Saini, R., Saini, S., and Sharma, S. (2010). Nanotechnology: the future medicine. *Journal of Cutaneous and Aesthetic Surgery*, 3(1), 32–33. <https://doi.org/10.4103/0974-2077.63301>
- [26] Selvi, S. K., Kumar, A., and Verma, R. (2025). Nano Robots Promising Advancements and Challenges in Healthcare: Nanobots in Healthcare. In: Kumar A, editor. *Exploring the Micro World of Robotics Through Insect Robots*. Wiley; p. 1-16. <https://doi.org/10.4018/979-8-3693-6150-4.ch00>.
- [27] Sharma, P. K., and Yuan-Chen, C. (2025). AI-Integrated Micro/Nanorobots for Biomedical Applications: Recent Advances in Design, Fabrication and Functions. *Biosensors*, 15(12):793. <https://doi.org/10.3390/bios15120793>.

- [28] Sohail, A. (2025). Progress in nanomaterials: Synthesis, characterization and applications, *Nest Nanotechnology*, 8. <https://doi.org/10.1016/j.nxnano.2025.100263>.
- [29] Zois, D., Lekatou, A., and Vardavoulias, M. (2009). A microstructure and mechanical property investigation on thermally sprayed nanostructured ceramic coatings before and after a sintering treatment. *Surface and Coatings Technology*, 204(1-2), 15-27. <https://doi.org/10.1016/j.surfcoat.2009.06.013>

Chapter 10

ARTIFICIAL INTELLIGENCE IN ART AND CREATIVITY

Sukhjeet Kaur Grewal

Assistant Professor, Department of Fine Arts, Khalsa College for Women, Civil Lines, Ludhiana, Punjab, India-141001.

Email: sgrewalsk60@gmail.com

Abstract

Artificial intelligence (AI) has played an important role in human creativity and artistic practices these days. AI introduces new methodologies for artistic production. As AI technologies become a direct threat that are transforming traditional approaches to creativity and redefining the relationship between humans and digital tools. AI as a collaborative element enhances human creative capabilities rather than replacing them. The incorporation of AI into the creative field of work can enhance innovation, efficiency and the overall quality of artistic output. Artificial intelligence refers to the development of systems having capability of performing many diverse tasks that typically require human intelligence. In the context of art and design, AI-based applications have the potential to transform both creation and interpretation. It can work on creating new forms, styles, gradations and compositional elements; expressions of ideas and addressing complex creative challenges. This chapter examines how AI helps the creative industry grow globally by addressing challenges in idea expression and investigating novel design opportunities. The chapter also analyzes current AI applications in artistic practices and discusses their implications for the future of creativity.

Keywords: AI in contemporary creativity, future of AI, hybrid workforce, innovation in expression, art

10.1 Introduction

Art is a form of creative expression; artists use their imagination and creativity to produce visual and audio works. Artists typically create works of art using a wide variety of materials such as paint, clay, glass, wood, paper, fabric, metal and so on. The art they create includes paintings, sculptures, drawings, photographs, ornaments, engraved items and more. "Arts" refers to the theory, human application and physical expression of creativity found in human cultures and societies through skills and imaginations in order to produce objects, environments and experiences. It includes visual arts, architecture, music, theater, film, dance and literature. Although the creation of art has been a human activity from thousands of years ago, now new technologies make it feasible for artists to create art using Artificial Intelligence (AI). This process typically involves algorithms that can learn from data inputs and generate new, original art pieces; it can take various forms, including digital images, paintings, sculptures, music and poetry. Modern artists often use a hybrid of traditional and AI techniques, creating their artworks. AI refers to the development of computer systems capable of performing cognitive tasks (Zhang et al., 2021). It consists of both hardware and software and can be used for a wide range of applications, including cars, phones, home appliances and intelligent robots. In today's scenario all of us have seen brands, news channels and celebrities flood the internet with Studio Ghibli versions of their photos. Everyone is participating in the Art vs. Algorithm discourse. In this chapter we will understand the role of AI in the field of art and creativity and learn more about creative technologies forming artistic expressions.

10.2 Creativity As A Unique Human Skill

Creativity is the ability to form novel and valuable ideas or works using one's imagination. It is often considered as an "intuition" or "talent" and can't be easily interpreted in a logical way (Wu et al., 2020). The creative fields often refer to fine arts, graphic design, music, fashion, advertising, media or entertainment industries, related to the extraordinary executions by supreme creative individuals. However, creativity actually lies in all creative activities, from the painting to music, from everyday life activities to industry production. Boden (2007) who pioneered the field of cognitive science, offers the definition "Creativity is the ability to come up with ideas or artifacts

that are new, surprising and valuable”. People used the creativity in art and technology and making it one of the most striking features of the human species. Creativity is becoming an important part forming various fields including education, art and design, entertainment and so on. Creativity is the driving force behind all artistic processes. Creativity helps you to wake up your inner artist and harness your full creative potential, because of this, creativity in the artistic process is still important to many artists. AI tools into creative workspace can affect the ability of human creators (Zylinska, 2020). Creativity involves in a process of exploration and discovery that makes us to generate new ideas and artifacts. In order to understand creativity; we need to examine the processes by which new ideas are generated. A key component of creativity involves abstraction, the process of learning how to make sense of information by identifying the conceptual components which are relevant for the construction of art. This leads to the development of different forms and structures and is the way which we explore and develop new ideas.

10.3 Role of AI In The Field of Art And Creativity

A major development is seen in deep-learning tools, such as Midjourney and Dall-E, that raised some questions in minds about creativity and AI's role in fine arts. Can AI be creative? This question arises in the mind every time because computational creativity becomes very important these days. Boden (1998) made a distinction between combinational creativity, exploratory creativity and transformational creativity. AI could be used in multiple ways to enhance creativity, the production of artistic events; it could also be highly beneficial for art exhibitions and museums. AI's role would not necessarily challenge human creativity in the fields of social sciences, arts and humanities, especially in the field of fine arts (Graham, 2005). There are certain ways in which AI is playing role in art and creativity and are discussed in following sections

10.3.1 To perform cognitive tasks

AI is increasingly used to do cognitive tasks humans previously completed and the most mechanical and routine tasks are the ones in which humans could be replaced first. With the development of AI technology, AI systems could take over higher-level cognitive tasks in various fields with superior efficiency and precision (Gerlich,

2025). In the long run, AI could take over tasks in the field of fine arts such as recognizing patterns in a complex visual composition; making decisions about color, form, composition and texture; abstract concepts; and creating art that involves intricate hand-eye coordination and control; and in writing, choosing precise language and constructing coherent arguments or narratives, etc., that humans previously completed. Because of the recent developments of AI and its capabilities, people continue to be captivated, sparking their curiosity about its possibilities (Oksanen et al., 2023).

10.3.2 AI transforms society and culture

AI has the potential to bring about numerous positive changes in society, as it can help in solving complex problems and make our daily lives easier and more convenient. Its usage and ongoing advancements have become unavoidable for everyone. AI is integrated into most of everyday devices, systems and technologies we rely on. So, this transformation involves not only a twist in what already exists but also combining different ways. In music, it can compose music for a particular genre but also can create a totally new genre or a subgenre. New movements in art inspired by AI fit into society and culture in the 21st century (Dwivedi et al., 2021). New trends show that AI is revolutionizing art, opening the way for new forms of creativity and opening up art to new audiences.

10.3.3 Evolution of Artistic and Creative Expression

Artistic progression is a continuous, decade-by-decade, complex transformation brought about by cultural, technological and societal revolutions. Visual arts like painting, sculpture and music have existed for ages and people attributed them to displaying sentiments, narrating histories and depicting the human experience. Yet, with the help of AI in the creative field, the novel history of art has begun. AI has given the brush of a new approach to creativity and provided artists/creators with new options to push the boundaries of creativity. AI algorithms and machine learning help artists create pieces of art made half by the artist and half by the computer. Thus, the scope of the creative world appears endless, from creating paintings with the help of AI to composing music pieces and designing sculptures and installations. The entwining of AI in art has led to a discussion about creativity and the artist's position in the virtual world. It's

still confusion whether AI is an instrument to stimulate creativity and enrich the array of creative practices or a threat that diminishes the essence of art. Nonetheless, with the development of AI and its penetration into the field of art, changes are observed in the development of art and its new landmarks are progressing in an unknown direction (Mazzone and Elgammal, 2019).

10.4 AI Art Generator Tools

Nowadays, the world of AI technologies is highly developed and we are currently in a space of transformative societal and cultural change. Current developments in AI involve creativity in fine arts, encompassing the visual arts, architecture, music, theater, film, dance and literature. There are numerous AI art generator tools listed in Table 10.1 given below:

Table 10.1: List of AI art generator tools

Sr. No.	AI art generator tools	Working area	Website
1.	Stable-Diffusion	It is very easy for anyone to create AI images without having any advanced technical knowledge. It is an open-source tool that can be installed locally; you can use it offline and have greater control over your creative processes. This enables you to adjust image prompts and control settings for better-quality outputs. There are multiple platforms that include this tool: Stable Assistant, Dream Studio and Hugging Face.	https://stability.ai/stable-image

Sr. No.	AI art generator tools	Working area	Website
2.	Canva AI	It is a widely used graphic design platform that has the feature of a built-in AI art generator known as Magic Media. With the help of this art generator too, users can create customized images for various purposes, including social media posts, marketing materials and educational content. You can explore different image styles, color combinations and different backgrounds to customize your personal designs.	https://www.canva.com
3.	DALL-E3	It allows users to create detailed images for numerous applications. It can produce concept-based art, social media illustrations and customized graphics. DALL-E 3 has an advanced system, generating images that work according to your instructions with high accuracy.	https://openai.com/dall-e-3
4.	Starryai	It converts simple text prompts into highly customizable visual artworks. You can use many styles, including anime, pixel art, photorealistic and abstract styles that allow for significant creative expression according to your needs.	https://starryai.com

Sr. No.	AI art generator tools	Working area	Website
		It also lets you adjust image parameters to tailor the generated images to fit your vision. Aside from its Text-to-Art feature, the platform supports blending AI-generated images with your own digital creations through its Image Fusion capability. This integration with the user is another functionality that distinguishes Starryai from other AI art generators.	
5.	Midjourney	It is best for imaginative artwork; you can upload multiple photos and blend all of them to make a single image. It also has advanced rendering techniques, allowing you to make images having rich textures and intricate details.	https://www.midjourney.com
6.	Leap AI	Best for personal design projects, it can quickly produce images from keywords and offers different art styles, including abstract, sketch and realistic, allowing you to craft content that matches your project's theme and personal preference in seconds.	https://www.tryleap.ai
7.	Prodia	Best for building content libraries, Prodia has a modern and uncluttered layout with just a few but important image customization options; its interface is easy to navigate.	https://prodia.com

Sr. No.	AI art generator tools	Working area	Website
8.	Jasper Art	It is ideal for businesses, marketing and professionals. This tool can be used to produce images based on textual and visual input, eliminating the hassle of searching through stock photos or paying for design services.	https://www.jasper.ai/art

10.5 Impacts of AI on Artistic Expression

To comprehend the benefits, obstacles and application of AI in art, there is a need to consider the current scope of the phenomenon. A primary benefit of incorporating AI as part of art is that AI assists and strengthens people’s imagination. The AI algorithms can also analyze big data and generate patterns and ideas that may not be easy or impossible to develop through everyday thinking. This creates new opportunities for the growth of artistic expression, as artists can experiment with ideas and create work that can be seen as provocative to the normative society. Moreover, AI can enrich creative practices and enhance people’s experience by enabling artists to try more different styles, techniques and media variants quickly. Thus, AI tools like image generators, music creation and text analyzers led artists to concentrate on creative intent. However, democratic artworks, with the help of AI, are problematic with topics such as copyright, ownership and cultural appropriation. In massive cases of creations such as art with or by the use of AI, the parameters within which authorship and, hence, the rights of ownership are determined become a contentious subject due to the convergence of AI and art in the bigger picture of tried and tested tools of authorship. This shows AI’s effects on art, which identifies the advantages of technology as a form of creation for the further development of art. When art is created using AI, artists, technologists and viewers are forced to consider conventional creativity, authorship and aesthetic practices while introducing new ways of practicing and imagining art in the postmodern digital context (Rammal et al., 2025).

10.6 Potential of AI In Art And Creativity

AI's potential is significant, for example, helping people in organizing art events, finding events and marketing art. However, the use of AI also raises concerns about the production and authorship of art and the replacement of people working in various fine arts fields. AI has become a collaborative agent to achieve specific goals. It can automatically analyze paintings of various artists and schools of art; it works as a classifier that not only identifies a painting's style but also provides explanations for its classification. As another example, Song (2021) used an AI-based improved support-vector machine algorithm for the analysis of paintings. Gualandi (2021) developed two machine-learning tools for the automatic recognition and classification of pottery, utilizing shape and decorative features. Manitsaris (2014) developed a methodology for capturing and modeling the gestural knowledge and skills used in wheel-throwing pottery. They created a technological prototype called ArtOrasis to validate the methodology, which provided continuous feedback in real time by measuring the differences between an expert's and a learner's gestures. Such AI systems can be used to learn the gestures. Mazzone et al., (2019) described an AI process developed for making images and proposed a partnership between human and machine creativity to overcome the challenges of AI creativity in art and ensure it complements rather than challenges.

10.7 Recent Trends - AI Is Revolutionizing Art

- **Perfect Brainstorming Assistant:** Artists might encounter periods of creative blocks when they can't generate actionable ideas, which slows down their processes. AI can be the perfect assistant in these situations. It can help generate ideas for mediums, composition and style; provide references; and refine ideas, allowing artists to overcome creative obstacles and accelerate creation (Veale and Pérez, 2020).
- **Enhanced Workflow:** AI is known for its ability to enhance the ability to reduce stunning visuals with minimal effort, decision-making, automating repetitive tasks and providing more room for the artists to explore their creativity. Even digital artists can use AI-driven tools for color blocking in elaborate works, add elements using generative features and allow them to create multiple versions of their work in different color schemes and perspectives, presenting distinct POVs and refining artworks (McCormack et al., 2020).
- **Personalized Styles:** AI-driven tools can inspire new ideas, offer fresh perspectives and explore creative directions, introducing new aesthetic possibilities. It can lead to more diverse and new styles, forms and innovative outputs. Artists can now use AI tools to generate

ideas that resemble their personal style, saving time and increasing efficiency and productivity (Rammal et al., 2025).

- **Newer Mediums:** The prompt-based generative AI can fill in the details as required, acting on the instructions provided and create entirely new styles, fuse different artistic traditions and reinterpret traditional techniques and art forms. It offers ample opportunities for artists to experiment with their art and find their unique personal style (Mazzone and Elgammal, 2019).
- **Immersive Artworks:** AI supports AR/VR technologies that allow artists to create interactive and immersive 3D art installations that break the boundaries of traditional art forms and change the perception of art in the modern world. AI interactive installation art generally follows a similar pattern. First, information is captured through sensors. Second, the information is sent to the program equipped with an AI algorithm for analysis and processing. Finally, the processing results are fed back to the participants through the output device (Grau, 2003).

10.8 Some Ethical Challenges In AI-Generated Art

- **Concerns ownership:** Anyone can provide prompts to AI-based software and generate impressive artworks. However, AI models are fed existing art, which allows them to process and generate newer work. This raises some questions, like, who owns AI-generated art? Can a machine own artworks? Does feeding prompts to AI-based software count as creation? The ambiguous nature of ownership standards and unclear regulatory laws remain a looming ethical dilemma for artists worldwide (Lovato et al., 2024).
- **Copyrights:** AI models are trained with large datasets. There are many copyrighted artworks by renowned studios and artists who have worked for years to develop a unique style or form. Experts have thus raised concerns about copyright infringement, protecting someone's art. The rise of infringement lawsuits by copyright owners against AI companies can be seen (Xie and Xie, 2026).
- **Lack of Depth and Originality:** While AI can generate artworks inspired by existing traditional and modern styles and forms, fusing different techniques and using magnificent colors and strokes, it cannot process art as the human brain does. There is emotional depth and variations possessed by artists, making AI-generated art appear unessential, soulless and disingenuous. The existing AI models can't create original art, raising questions about their value, ethics and use of artistic styles (Elgammal et al., 2017).

- **The security threats associated with AI-generated pictures:** The major threat related to AI-generated images that can be posed to cybersecurity is the threat of data breaches. When the users post their pictures on an AI platform, they are authorizing the platform's ability to access facial and biometric information. In the case when the platform does not ensure sufficient security, their data can get compromised and this may lead to such crimes as identity theft, stalking and other forms of harassment (Mirsky and Lee, 2021).
- **Environmental Concerns:** Data centers that power these AI models consume large amounts of electricity and water to function. Rapid expansion of AI also leads to the generation of harmful electronic waste and an increase in mining activities to extract rare earth metals. While the former can have long long-lasting impact on biodiversity, human rights abuses and destruction of the environment in the nearby regions (Strubell et al., 2019).

10.9 Future Of Artificial Intelligence In Art – Trends And Innovations

According to a Forbes report, AI has a rather bright future in art and related fields. As more artists use tools, hybrid art is ready to become a new way of art, driving the market and shifting how art is created, taught, sold and perceived. There is a rise in human-AI collaboration as many artists are using AI tools to generate ideas, add creativity and even produce artworks. This collaboration between AI and human creativity is creating a new world within the art field, with its own new genre, style and immersive qualities (McCormack et al., 2020). While AI becomes an important part of the process, artists are adopting approaches that address ethical concerns, paving the way for a future full of harmonious innovations. From ensuring transparency and informing the viewers about the extent to which AI was used and seeking consent from artists before using their work to train models, to making sure the data fed was free of bias, avoiding hyper-dependence on AI and determining the ownership—we are looking forward to balanced strategies in art and design. Emerging AI-based tools might increase the workflow for freelancers. Many platforms like Upwork and Fiverr, a multinational online marketplace, have already begun to work together on projects, with AI being able to create promotional art and design for brands, video and animation and music and audio. However, new opportunities will emerge for people who are acquainted with various AI tools and know how to navigate the landscape by creating the perfect balance between AI and human creativity. With the continuous growth of AI and creative technology, the curriculum and methods of teaching need to evolve with AI as well. Many institutes are updating their syllabi to include art and AI training. Mark Leibert, an

artist and professor of the practice at Georgia Tech. His expertise is in art, AI and digital technologies. He taught students fundamentals of various artistic media, including charcoal drawing and oil painting. Students used these principles to create art using AI ethically and creatively. AI entered our lives so rapidly; aim to empower students (Mazzone and Elgammal, 2019; McCormack et al., 2020).

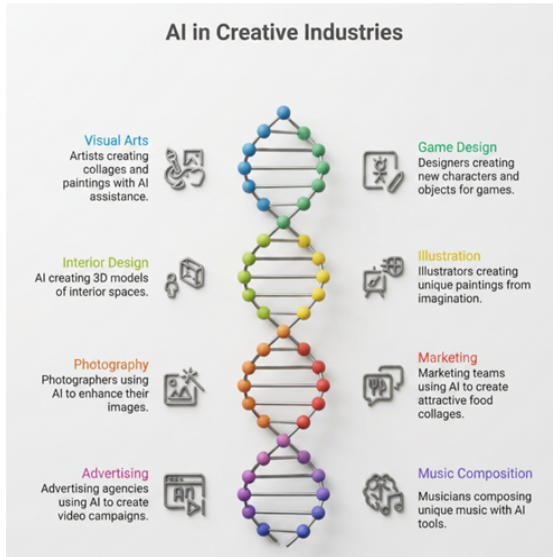


Figure 10.1: Applications of AI in art and creativity (Made using Napkin AI <https://app.napkin.ai/>)

10.10 Use of AI In Different Areas of Creativity

The uses of AI in art and creativity are enlisted in figure 10.1.

10.11 Conclusion

AI-generated art is advancing rapidly, open exciting pathways for creative expression in many fields. By blending innovation of AI with creativity, it bridges the past and future and offering new lens to view the world of creativity. It brings new possibilities to creative work. By recognizing present deficiencies and encouraging teamwork between humans and machines, we can achieve a future. Today's generative AI makes creativity more available and diverse. It is a powerful tool of inspiration can generate new content, enhance creativity and overcome difficulties. Companies and professionals in every field know how to use generative AI in creative work to achieve more.

References

- [1] Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, 103(1–2), 347–356. [https://doi.org/10.1016/S0004-3702\(98\)0055-1](https://doi.org/10.1016/S0004-3702(98)0055-1)
- [2] Boden, M. A. (2007). Creativity in a nutshell. *Think*, 5(15), 83-96. <https://doi.org/10.1017/S147717560000230X>
- [3] Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., et al. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- [4] Elgammal, A., Liu, B., Elhoseiny, M. and Mazzone, M. (2017). *CAN: Creative Adversarial Networks, generating “art” by learning about styles and deviating from style norms*. In Proceedings of the 8th International Conference on Computational Creativity (ICCC 2017), 96–103. Association for Computational Creativity. <https://arxiv.org/abs/1706.07068>
- [5] Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- [6] Graham, G. (2005). *Philosophy of the arts: An introduction to aesthetics*. Routledge. <https://www.routledge.com/Philosophy-of-the-Arts-An-Introduction-to-Aesthetics/Graham/p/book/9780415327243>
- [7] Grau, O. (2003). *Virtual art: From illusion to immersion*. The MIT Press, Cambridge, Massachusetts, London, England.
- [8] Gualandi, M. L., Gattiglia, G., and Anichini, F. (2021). An open system for collection and automatic recognition of pottery through neural network algorithms. *Heritage*, 4(1), 140-159. <https://doi.org/10.3390/heritage4010008>
- [9] Lovato, J., Zimmerman, J. W., Smith, I., Dodds, P., and Karson, J. L. (2024). Foregrounding artist opinions: A survey study on transparency, ownership, and fairness in AI generative art. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 7(1), 905-916. <https://ojs.aaai.org/index.php/AIES/article/view/31691>
- [10] Manitsaris, S., Glushkova, A., Bevilacqua, F., and Moutarde, F. (2014). Capture, modeling, and recognition of expert technical gestures in wheel-throwing art of pottery. *Journal on Computing and Cultural Heritage (JOCCH)*, 7(2), 1-15. <https://doi.org/10.1145/2627729>
- [11] Mazzone, M., and Elgammal, A. (2019). Art, creativity and the potential of artificial intelligence. *Arts*, 8(1), 26. <https://doi.org/10.3390/arts8010026>

- [12] McCormack, J., Gifford, T., Hutchings, P., Llano, M. T., Yee-King, M., and d'Inverno, M. (2020). Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound*, 25(1), 41–52. <https://doi.org/10.1017/S1355771819000451>
- [13] Mirsky, Y., and Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), 1–41. <https://doi.org/10.1145/3425780>
- [14] Oksanen, A., Cvetkovic, A., Akin, N., Latikka, R., Bergdahl, J., Chen, Y., and Savela, N. (2023). Artificial intelligence in fine arts: A systematic review of empirical research. *Computers in human behavior: Artificial humans*, 1(2), 100004. <https://doi.org/10.1016/j.chbah.2023.100004>
- [15] Rammal, H., Hejase, H. J., and El Takach, A. (2025). Artificial intelligence in art: Bridging the gap between automation and human expression. *Asian Journal of Social Science Studies*, 10(1), 1-9. <https://doi.org/10.20849/ajsss.v10i1.1445>
- [16] Song, B. (2021). Social-cultural perspective of fine arts policies using AI model. *International Journal of Technology Management*, 86(2-4), 167-182. <https://doi.org/10.1504/IJTM.2021.118312>
- [17] Strubell, E., Ganesh, A., and McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP*. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 3645–3650. <https://doi.org/10.18653/v1/P19-1355>
- [18] Veale, T. and Pérez y Pérez, R. (2020). Leaps and Bounds: An Introduction to the Field of Computational Creativity. *New Generation Computing*, 38, 551–563. <https://doi.org/10.1007/s00354-020-00116-w>
- [19] Wu, Y., Mou, Y., Li, Z., and Xu, K. (2020). Investigating American and Chinese subjects' explicit and implicit perceptions of AI-generated artistic work. *Computers in human behavior*, 104, 106186. <https://doi.org/10.1016/j.chb.2019.106186>
- [20] Xie, Y., and Xie, Y. (2026). When artificial intelligence makes everything similar: The risks of content homogenization. *Chinese Journal of Sociology*, <https://doi.org/10.1177/2057150X261419573>
- [21] Zhang, R., Zhao, Y., Kong, J., Cheng, C., Liu, X., and Zhang, C. (2021). Intelligent recognition method of decorative openwork windows with sustainable application for Suzhou traditional private gardens in China. *Sustainability*, 13(15), 8439. <https://doi.org/10.1177/2057150X261419573>
- [22] Zylinska, J. (2020). *AI art: Machine visions and warped dreams*. Open humanities press. <https://www.openhumnaitiespress.org/books/titles/ai-art/>

Chapter 11

ARTIFICIAL INTELLIGENCE, GOVERNANCE AND PUBLIC POLICY: BRIDGING TECHNOLOGY AND SOCIETY

Ruchi Dhiman

Assistant Professor, Department of Political Science, A.S. College, Khanna, Punjab, India-141401.

Email: ruchidhiman994@gmail.com

Abstract

Artificial Intelligence is continuously changing how government authorities are exerted; the way in which governments distribute their resources; and, the ways in which citizens interact with their government. AI is an evolving agent of change that will continue to influence government, political authority, and the policymaking process; positioned within the context of Political Science. The chapter begins with an overview of AI Governance and includes a description of various forms of AI Governance developed from many of the Governance Models created today. Examples include the establishment of AI Governance Systems using Regulatory Frameworks, Administrative Oversight Tools, and Responsible AI Principles. The growing need for democratic control of AI with increased accountability and transparency led to the emergence of these structures. Government has started using it to save its time and finance but besides this, it is also required that it must be used on democratic principles i.e. must protect data, provide security and safeguard human rights by ensuring fair justices that it may be actually proved helpful in uplifting society and not diminishing it. Many of the public or private organizations are using AI for increasing the efficiency of the agencies of government. It is an urgent need that there must be a mechanism to establish good governance for AI which can be achieved by making artificial

intelligence mechanisms approachable to the citizens as well in a participatory format and ensuring its adaptiveness according to the changing times and circumstances so that AI can better be utilized not only for an individual or government but for all members of the society as a whole.

Keywords: democratic control, governance, policy making, principles, regulatory frameworks

11.1 Introduction

Artificial Intelligence or AI has become a major agent of change in the modern forms of governance. It is transforming the way of connectivity between the government and citizens; formulation of policies and administration of services. Therefore, the governments are also trying to use basic algorithms to make better decisions and providing services in an effective way along with to improve their own efficiency. This modern era shows a drastic shift from traditional rule based bureaucratic model to digital governance which can be better termed as ‘*Smartocracy*’. It is taking administration away from the slow, paper heavy and rigid system towards the quicker, flexible and data driven procedure.

Ahmad et al., (2024) explain that AI driven tools are increasingly assimilating into the policy making and execution which aid in real time decision making and predictive governance.

Traditionally governance was based on the constitutional values and human behavior in the events such as elections but now a days, development of AI technology is putting a question mark on moving on the lines of constitution as well as maintaining democratic values, i.e. It has become a major challenge to maintain democratic values as well as adopting the vast techno changes. In other words, AI is not only making technological innovations but also transforming political scenario in terms of governance and implementation of schemes and policies. It can be seen as a medium of changing public- private relationship and also providing a new pathway to link government with the common masses. It predicts a new way of engaging with the government authorities. Table 11.1 highlights transition from traditional to AI-mediated governance.

Table 11.1: Transition from Traditional to AI-Mediated Governance

Feature	Traditional Bureaucracy	AI-Driven Gover- nance
Opera- tional Pace	Slow and massive re- liance on paper work	Flexible, fast, real- time
Decision Basis	Rigid rules, static guide- lines and precedents	Data-driven and predictive
Interaction Model	Reactive (You ask, they answer)	Proactive (they predict, you will want/need something)
Primary Tool	Human administrative oversight	Algorithmic and automated systems

Aswal and Joshi (2025) contend that the growing reliance on AI in governance raises ethical and political concerns, including algorithmic bias, reduced democratic oversight, and the concentration of decision-making power.

This chapter presents the ways in which adaptive public policy frameworks can be developed to facilitate communication between technological change and the governance of democracies.

11.2 Foundations of AI Governance

11.2.1 Definition of AI Governance

In simple terms, “Artificial Intelligence refers to a machine- based system that has the ability to learn from data, make decisions, and carry out cognitive functions” (Aarab et al., 2025). The use of artificial intelligence is in governance, such as welfare targeting, monitoring regulatory compliance and managing public health, along with urban planning.

Aarab et al., (2025) conceptualise AI governance as the institutional, legal, and ethical arrangements that guide the development and use of artificial intelligence within public governance systems. Politically, the idea of AI governance expands on reforms to governance systems that will address the state’s abilities and how to protect democracy.

11.2.2 Algorithmic Power and Political Power

In a democracy, the people are the ultimate source of power, and they give "sovereign authority" to elected officials to make decisions. But the use of AI in Governance is changing how authorities (those that have political power) govern. The decision-making authority of public officials is being transitioned to automated systems and technical professionals. This transition presents challenges in relation to accountability, transparency, and democratic control.

Chen and Wang (2025) demonstrate that the relation between the ruling governance and AI is continuously modifying the structure of cooperate and public administration. They asserted that human must take the responsibility as AI is adopting the responsibility of tasks to be done in a way fast quicker than them. There is a need to evolve new laws as the traditional methods are not convenient to use, manage and control machines.

11.3 Emerging Models of AI Governance

11.3.1 Regulatory Governance

Governments are using many types of regulatory frameworks to control the dangers of AI out of which the major three components are given as follows-

- a) Data privacy,
- b) Algorithmic accountability and
- c) AI Specific Regulation.

The main objective of such laws is to ensure the use of AI to be responsible and ethical.

The World Bank (2023) highlights that various countries have adopted and implemented laws governing the use of artificial intelligence in administration based on their political systems and socio- economic perspectives. However, there is always a continuous struggle experienced between governance and rapid technological developments which often leads to uncertainty.

11.3.2 Administrative and Institutional Oversight

We cannot just prepare AI tools and forget them. It is important to have some mechanisms as well that can prevent wrong use of public data and keep its functioning error-free. Administrative governance doesn't mean to stop innovation rather it must be able to build trust using AI technology. The government or any organisation or company must keep a watch over the outcome of using digital tools. It is also essential to note that AI must match with human values and demands.

There are some core safeguards that are required in the public sector for effective governance such as continuous capacity building, organisational readiness and oversight mechanisms, without which AI can be proved fatal for the administration because of serious risks.

11.3.3 Principles of Ethical and Responsible AI

AI ethics means to enable the use of AI technology for the upliftment and betterment of the society. Therefore, it is required to make AI fulfilled with common principles such as transparency and accountability. Data must be protected and under the democratic values to establish governments on the basis of ethical as well as responsible AI. The following principles provide the foundation for this new framework:

- **Accurate and reliable:** To provide results those are correct and align with real-world data or intended outcomes. System to perform accurately according to environment and time.
- **Accountable and transparent:** Final authority or ownership to remain with the humans, so that users must acknowledge when and how they are interacting with AI.
- **Fairness:** Must be free from any type of harmful biasness that could lead to discrimination based on race, gender, or other sensitive traits.
- **Human-centric:** To support human well-being, assist communities rather than replacing human agency and also to respect human rights.
- **Safe:** Must be safe from any physical or psychological point of view, to individuals, society, or the environment.

- **Ethical:** Development must be linked to the international norms or laws and to avoid any type of unauthorized surveillance.
- **Secure:** There must not be any unauthorised access and system attacks such as data poisoning.
- **Resilient:** Must be able to recover quickly from technical failures.
- **Interpretable:** Stakeholders should be able to understand the "why" behind a specific AI prediction or decision.
- **Documented:** Recording of the overall process or cycle of AI for auditing and governance.
- **Privacy-enhanced:** To prevent identity leaks by respecting data privacy rights.
- **Data governed:** Data must be gathered with permission and making sure to use it ethically.

According to Papagiannidis (2025), the principles of responsible AI governance should be incorporated into the regulatory and institutional framework of a country's institution, not just through a voluntary set of guidelines.

11.4 AI And Public Administration

11.4.1 State Capacity Building

AI has the ability to enhance and develop state governance efficiency through many ways such as state capacity building, decreasing the cost of structure formulation and improving administrative efficiency. Governments, around the world have started utilising AI technology or tools for transforming their governance and rule from traditional into the modern updated aspects without the need of extensive use of manpower in the common tasks such as tax collection or urban decentralised governance.

- **Tax and Revenue:** AI technologies can analyse massive amounts of data to search for evidence to support potential evaders and mistakes. This streamlines the collection of taxes and makes the collection system more accurate.
- **Welfare functions:** With the help of this method, governments can make better use of resources for the effective execution of

welfare schemes and programmes. It can help in identifying the good utilisation by the right people of society.

- Health and Hospitals: AI tools are using with greater power in the health sector as they are the most reliable source of getting knowledge of the health condition of the patients and improving the chances of the better treatment by acquiring the medical historical data of any person within no time.
- AI can also help to manage urban infrastructure.
- Helpful in running business enterprises successfully by evaluating the profit and loss timely. It increases the prospects of good opportunities as well.
- Smart TVs, smart phones, smart watches or smart doorbell are the variant examples of AI tools that help in day- to- day life making it easier.
- Finance and banking are other areas of successfulness of AI technology making human beings comfortable with their transactions and mapping their lifestyle.
- In the form of much other advancement such as in the field of automobiles and language learning or translations providing greater satisfaction on the part of government also.
- During Crisis: Using data from AI, States can determine how to respond to crises before they become a bigger problem.

Ahmad et al., (2024) observe that AI-driven administrative systems can enhance policy effectiveness by optimising resource allocation and improving service delivery outcomes. The development of these initiatives has resulted in better overall government performance and greater citizen satisfaction.

11.4.2 Challenges to Fair Justice

Although, AI in government may bring about operational efficiency, but there are very real concerns about how AI-based governance affects citizens' acumen of fair process (procedural justice).

Specific challenges to procedural justice due to AI include:

- **Lack of transparency or 'Black-box':** It is a problem that often make decisions unexplainable to citizens as Artificial Intelligence (AI) often draws conclusions via billions of calculations that even its creators cannot fully articulate. In situations where

an individual is denied a permit or flagged by law enforcement, the answer of 'the computer said so' is not an acceptable legal reason, therefore eliminating any chance of verifying whether AI's reasoning behind the decision was logical.

- **Algorithmic bias and discrimination:** AI derives its learning from historical data, and if there are biases exhibited within the data, such as through previous human decision making, such as decisions about a certain neighbourhood or member of a particular race, the AI will learn those biases and perpetuate them. This results in systemic biasness.
- **Erosion of due process and right to contest:** If an AI renders a decision against an individual, one cannot contest the AI's decision as all such 'evidence' will be trapped in complicated computer code.
- **Lack of accountability and oversight:** In case of an inaccurate conclusion by AI, it can be difficult to determine whether the error rests with the individual who developed the software, the entity that bought the software or the government entity utilizing the software.
- **Risks of data security and privacy:** Large amounts of personal information create a large number of risks including identity theft and the acquisition or unethical use of sensitive information, which would violate the data privacy rights.
- **Undermine the role of human judgements:** According to Aswal and Joshi (2025), if automated decisions are not transparent, contestable, or overseen by humans, they can undermine procedural justice through algorithms. Thus, the establishment of procedural fairness should be a key goal of democratic AI governance.

11.5 Civil Liberties, Constitutionalism And AI Governance

Integrating AI into state enforcing and state surveillant arms of government is one of the major threats to modern constitutionalism. AI applications used in surveillance, policing and data analysis through such technology as facial recognition and predictive policing have large implications for civil liberties and constitutional rights. These applications could violate rights to privacy, equal protection, and due process if there are not sufficient protections in place.

Joshi (2024) argues that AI governance in India must be guided by constitutional principles to prevent the erosion of fundamental rights and democratic accountability. Aligning AI policy with constitutional norms is essential for preserving democratic legitimacy.

11.5.1 Digitalizing the Loss of Privacy

The right to privacy has historically been a central principle to liberty within almost all democratic societies, however, the way AI usually operates requires that large format datasets be collected about citizens and their behaviors which are a breach of their expected anonymity in communal spaces. Using AI as the basis of their authorities for surveillance, the state is now moving from being a surveillance model where they survey individuals who they have already monitored and confirmed to be of interest to them, to an indiscriminate comprehensive approach whereby all individuals' behaviors will be analysed via multiple datasets around them to identify what might be considered abnormal. This transition could encroach on individuals' access to equal protection and/or their due process rights unless there are strong safeguards implemented.

11.5.2 Constitutional Safeguards and India

The tension between Governance and AI in India is especially intense, as current governance must operate under constitutional restrictions to avoid infringing on any citizen's right. It is required to check that AI policies are actually in accordance with the existing constitutional laws and frameworks. These AI laws must ensure the existence of the right to fair trial and freedom from any type of discrimination. Otherwise, there can be the probability of the ignorance of balance and judicial control that can possibly harm civic rights.

11.6 Public Trust And Democratic Legitimacy

People's trust in an institution and the way they perceive AI as fair will determine if the public supports AI governance. Therefore, creating a governance framework for AI that emphasizes elements of transparency, accountability, and ethical safeguards provide the public greater confidence in those systems and increase the chances of their support to them. Public trust in AI governance is strongly influenced by transparency, ethical oversight, and inclusive decision-making processes.

If there is no significant public involvement in shaping how we govern artificial intelligence technology, then there's a real chance that governance will be purely technical with no connection to the real world.

11.6.1 The Social Contract is Eroding

Democratic legitimacy traditionally comes from the "sovereign authority" that people give their elected representatives through the electoral process. However, as the decision-making power shifts to a non-human (automated) entity, this may create a "legitimacy gap".

- **De-humanisation of the Organisation:** If algorithms are put in charge of distributing social welfare payments or determining whether an individual should be subjected to a tax audit, there will be an increased likelihood that the person will consider the organisation as a "cold" machine or entity, rather than as an emotionally responsive entity.
- **Trust as Predictive versus Reactive:** Trust is shifting from a reactive basis, where the person forms their judgement of the organisation on past actions/service levels, to a proactive basis, where the person's trust in the organisation is built on how accurately the organisation is expected to fulfill forecasts based on the individual.
- **The "Black Box" Barrier:** If it is impossible for a citizen to know (or understand) "why" a specific AI prediction was made, the citizen will not develop trust in the organisation that created the prediction, but instead will develop distrust towards it.

11.7 Comparative And Global Perspectives

There are many different examples of how countries are governing AI technology. It is because of the diversity in their systems of government and the differences in how they regulate according to their socio-economic conditions. Cross-cultural distinctions show that local conditions must always govern the design of a particular AI governance framework.

Guenduez et al., (2025) demonstrate that national political and institutional contexts play a decisive role in shaping AI policy formulation and implementation. International organisations increasingly

advocate adaptive governance frameworks that balance global norms with local realities.

11.7.1 Diversity in Regulatory Philosophy

AI governance can't be done universally; it's based on the political and legal characteristics of each individual country. Comparative research suggests that although technology is global, regulation of the technology is local and varies greatly from country to country.

There is diversity between countries in their regulatory philosophy.

- **The European Approach**—The EU AI Act is one example of this model, which is based on a risk-based approach/framework and the promotion/protection of fundamental human rights.
- **The Technical Standards Approach**—This model is provided by technical standards organizations (e.g., NIST-National Institute of Standards and Technology, in the USA, ISO/IEC), which seek to create standards that focus on the technical aspects of safety and risk management as the basis for delivery of capable, secure, and resilient AI technologies. These organizations create blueprints for AI safety.
- **Global South**—Three main areas of focus for developing countries are how AI can increase the capacity of the state to provide services in healthcare and urban development while overcoming constraints on the administrative system/resources.

11.7.2 Global Norms and Local Reality are bridged

Global organizations are increasingly promoting the use of adaptive governance. It is defined as reconciling universal ethical standards with the cultural realities and legal realities of a region. The extent to which a national AI policy (implementation of the AI policy) will occur in those nations will be determined by the nature of the national political context. A nation with a higher degree of administrative capacity can implement AI regulation more effectively than a nation with lower level of administrative capacity.

11.7.3 Case Study: Competing Layers of Embedded Strategies (India and The Eu)

An analysis of AI governance has revealed that there is a critical difference between the structure of The European Union (the "EU AI Act", 2024) which is 'Rights -Based and India's 'Innovation-First/Adaptive' model, as both are unique in how they define technology based upon the overall political will of the given population.

The European Union: AI Legislation (Hard Law)

The European Union AI Act (2024) represents a new direction in AI governance. It focuses on how AI will be developed and distributed, both as an extension of basic human rights (for instance, to privacy and freedom from discrimination) and as a condition of product safety. A new risk-based classification system exists in the EU AI Act which classifies all AI applications into the following four risk categories:

Unacceptable Risk: (i.e., governmental social scoring) - All forms of unacceptable risk in AI are prohibited under law.

High-Risk: (i.e., AI used in recruitment and law enforcement) - All high-risk applications will be subject to an independent, mandatory performance audit, and required to comply with the following two new requirements - high-quality data and human oversight.

Limited Risk: (i.e., chatbots) - All limited risk applications will be required to comply only with minimal levels of transparency.

Minimal Risk: (i.e., AI-enabled video games) - All minimal risk applications will remain exempt from any applicable regulations.

The existence of 'Hard Law' will ensure that AI governance structures exist within the 'EU AI Act' when developing or implementing compliance programs; however, there are also ongoing criticisms by those in the tech industry that Hard Law may limit innovation via excessive or alternative compliance cost structures have yet to be satisfied.

11.7.4 India: Flexible and Principles-Based Approach

India has adopted a flexible, principles-based approach to AI regulation instead of a prescriptive, comprehensive AI regulation law

as is the case with the EU. Rather than develop a comprehensive AI regulatory framework, India has chosen to work within the framework of existing laws, such as the Digital Personal Data Protection (DPDP) Act of 2023, as well as sector-specific guidelines issued by the Ministry of Electronics and Information Technology (*MeitY*).

India’s AI strategy, referred to as #AIforAll, is designed to promote social inclusion and economic scale. The Indian approach recognizes that not all risks associated with the use of AI technologies can be known ahead of time. It can be said that Indian approach is grounded on ‘soft law’. It consists of ethical regulations, structures and standards that can be reshaped with the development of technology. This soft law helps India to be innovative and creative in developing its start-up ecosystems. The use of a soft law framework allows India’s rapidly growing entrepreneurial hub to be innovative free from legal burdens. But along with this, there is need to examine biasness also in the delivery of services (Aswal and Joshi, 2025).

The following table presents the points of key differences between the above stated regions of the world:

Feature	European Union (EU)	India (Bharat)
Primary Legislation	EU AI Act (2024)	DPDP Act (2023) and Sectoral Guidelines
Regulatory Philosophy	Rights-Based / Pre-cautionary	Innovation-Focused / Adaptive
Enforcement Model	Strict Risk Categories and Penalties	Principles-based and "Light-touch"
Objective	Privacy of Individuals and Market Safety	Social Inclusion and State Efficiency

11.8 Policy Implications And Future Directions

To achieve effective governance of artificial intelligence (AI), there must be a multi layered approach to balance between clear organisational structure and standardized risk frameworks. There is a need to implement proactive policies that create regulatory frameworks, administrative oversight, ethical standards, and opportunity for public participation. As AI evolves faster than traditional regulations, so an effective model must be adaptive rather than static.

The following major implications can be suggested:

- Adopt Recognized Frameworks such as the NIST AI Risk Management Framework (AI RMF), the EU AI Act, or the ISO/IEC 42001 standard for AI management systems. The EU AI Act (Regulation 2024/1689) is the first comprehensive law of its kind in the world. It came into effect on 1 August 2024. Its regulatory approach is known as “risk-based” which means the greater the potential risk to people posed by an AI system, the stricter and firmer the regulations will be for it.
- Encourage active participation and involvement of different segments of society such as researchers, users or employees to keep an eye over its efficiency.
- Adoption of the various methods to ensure safety, accountability, transparency, and privacy of the users.
- With the help of AI tools, deciding clear organisational goals and objectives to achieve and ensuring democratic governance through AI.
- Implementation of a trial project first in order to check the reliability of the policy.
- Proper training and education must be provided to those who are managing AI tools so that technology can be going long way in accordance with the demand of the administrators.
- Continuous observance and maintenance of control over AI technology is required.
- Proper management of any types of risks by following already existing structures.
- Confirm the guarantees of urgent intervene by humans or actual owners in case of high risk as after all human are the real manager of AI tools.

- Employment of well educated, dedicated and trained officials for instance from IT, legal or business firms only so that they will be able to frame policies taking care of all the dangerous online prospects.
- Different segments of society must work together for the betterment of society as a whole and be the masters of AI technology or tools rather than letting AI to be their master.
- Categorise the levels of risks that can be occurred like low risk, medium risk, high risk or infinite risk to put successful control.

11.9 Conclusion

It can be stated that although, AI provides the chances for better administration and an effective implementation of governance policies but also raises the questions of fundamental principles such as transparency, protection of human rights and civil liberties and assurance of accountability. It is high time that AI must be used in a responsible manner rather than just accepted as a tool of governance. AI must be properly aligned with human rights so that citizens' safety and security are not to be endangered in any possible ways. There must be increasing role of common masses, marginalised communities, researchers or civil society in the effective policy formulation using AI tools so that the way towards achieving Democratic Governance can be paved forward.

References

- [1] Aarab, A., El Marzouki, A., Boubker, O., and El Moutaqi, B. (2025). Integrating AI in public governance: A systematic review. *Digital*, 5(4), 59. <https://doi.org/10.3390/digital5040059>
- [2] Ahmad, M. A. Y., Hamdan, M., Anshari, M., Fitriyani, N. L., and Syafrudin, M. (2024). Governing with intelligence: The impact of artificial intelligence on policy development. *Information*, 15(9), 556. <https://doi.org/10.3390/info15090556>
- [3] Aswal, N., and Joshi, P. (2025). Artificial intelligence and the future of governance: Ethical and political challenges. *International Journal for Multidisciplinary Research (IJFMR)*, 7(1). <https://www.ijfmr.com/research-paper.php?id=62727>
- [4] Cordeiro, C. M., and Adomaitis, L., and Huang, L. (2025). The AI-policy-governance nexus: How regulation and AI shift corporate governance. *Technology in Society*, 84, 103117. [10.1016/j.techsoc.2025.103117](https://doi.org/10.1016/j.techsoc.2025.103117)

- [5] Digital Personal Data Protection Act, (2023). *The Gazette of India: Extraordinary* (No. 22 of 2023). Ministry of Law and Justice, Government of India. <https://www.meity.gov.in/content/digital-personal-data-protection-act-2023>
- [6] European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*. <https://eurlex.europa.eu/eli/reg/2024/1689/oj>
- [7] Guenduez, A. A., Fuchs, S., and Demircioglu, M. A. (2025). AI in a comparative perspective: Linking policy formulation and implementation. *Journal of Comparative Policy Analysis: Research and Practice*, 27(5-6), 647–670. <https://doi.org/10.1080/13876988.2025.2543984>
- [8] International Organization for Standardization. (2023). *Information technology — Artificial intelligence — Management system* (ISO/IEC Standard No. 42001:2023). <https://www.iso.org/standard/81230.html>
- [9] Joshi, D. (2024). AI governance in India – Law, policy and political economy. *Communication Research and Practice*. <https://doi.org/10.1080/22041451.2024.2346428>
- [10] National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- [11] NITI Aayog. (2018). *National Strategy for Artificial Intelligence: #AIforAll*. Government of India. <https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf>
- [12] Papagiannidis, E., Mikalef, P., and Conboy, P. K. (2025). Responsible artificial intelligence governance: A review and research agenda. *The Journal of Strategic Information Systems*, 34(2), 101885. <https://doi.org/10.1016/j.jsis.2024.101885>
- [13] Schwalbe, N., and Wahl, B. (2020). Artificial intelligence and the future of global health. *The Lancet*, 395(10236), 1579–1586. 10.1016/S0140-6736(20)30226-9.
- [14] World Bank. (2023). *Global trends in AI governance: Evolving country approaches*. World Bank Publications. openknowledge.worldbank.org

Chapter 12

ADVERSARIAL ATTACKS AND AI SECURITY

Jayal Shah^a, Kshitij Dandekar^{a,*}, Sakshi Giglani^b, Sakshi Makwana^a

^aSchool of Computational Technology, Institute of Advanced Research, Gandhinagar, Gujarat, India - 382426.

^bFaculty of Engineering and Technology, Gujarat Law Society University, Ahmedabad, Gujarat, India- 380006.

Email: kshitijdandekar2005@gmail.com

Abstract

Artificial intelligence (AI) is now everywhere, from self-driving cars to X-ray machines, yet a major security flaw is ignored: it is incredibly easy to trick. These susceptibilities, called adversarial attacks, work by making small tweaks, like changing just a few pixels that are unseen to humans. This is not a technical glitch; it is a form of cybercrime that puts most critical systems at risk. In this chapter, we explore how these "hidden" attacks weaken the trust we place in AI. As a response to these concerns, Latent Uncertainty Guided Adversary Intervention (LUGAI) is discussed as a mechanism for identifying and recovering from adversarial attacks. Unlike basic security that just shuts down when it senses trouble, LUGAI is designed to "self-heal." The system acts like a digital immune system. It watches for two main red flags: first, when the AI becomes unusually confused, and second, when data appears distorted during a double check. If an attack is detected, the system doesn't quit; it clears the smudges, converting fake or blurry inputs back into clear, original images. This allows the AI to see reality and make the right decision. The tests show that the system performs exceptionally well in real-world scenarios. Observations indicate that the system has the ability to maintain accuracy while autonomously correcting adversarial distortions. Ultimately, we argue that AI needs to be more than just "aware" of

problems; it must be a self-healing machine that can fix its own wounds instantly.

Keywords: adversarial machine learning, ai security, cyber-crime, robustness, self-healing

12.1 Introduction

Artificial Intelligence is now part of the foundation of many systems and processes. For example, AI has been used in hospital diagnostic tools, banks to detect fraudulent transactions, optimizing logistics, and analysing government-provided data to help support evidence-based decision-making. Because of all these uses, AI has become an integral part of many aspects of people's lives, which creates tremendous benefits (Goodfellow et al., 2015). However, there are also a lot of risks to deploying AI systems, such as failures in completed systems that can lead to life-altering failures or losses.

Machine learning is quite different from the traditional type of software; there's a primary difference between machine learning and traditional programming. Machine learning does not use explicitly defined instructions. It uses actual data to learn how to behave appropriately. The flexibility and power provided by machine learning will create many new "failure modes" (ways that ML systems will fail). An error could arise from a subtle change in the input data, from a biased training set, or from a pattern that the model has never seen before. This means that new types of security and operational processes will need to be developed to account for the possibility of failure from machine learning (Biggio and Roli, 2018; and Ren et al., 2020).

Adversarial attacks take advantage of this very gap in how we analyze neural network-based predictions and inferencing capabilities. Attacks exploit the inputs the model relies on to make incorrect predictions with confidence due to very small, sometimes invisible alterations of real-world input data (Goodfellow et al., 2015; and Szegedy et al., 2014). Many times, adversarial attacks can be classified as either targeted or untargeted, with motivation behind attacks ranging from financial gain to malicious destruction to politically motivated interference. Because the original infrastructure remains functionally operational, detection, attribution, and legal

liability for these types of attacks are challenging, creating a systemic issue across engineering, operations, and governance domains (ENISA, 2021).

This chapter outlines the problem and provides guidance for building the systems needed to respond. We will provide a definition of adversarial attack models and explain how the construction and validation of oppressors and publications occur. We will also discuss the defence models that are available, such as adversarial training, detection, transformation, and architectural hardening, while also detailing the trade-offs and possible failure modes of each defence mechanism. We will then discuss the operational and regulatory issues that exist related to the response and/or remediation of these incidents, as well as the potential liabilities for organizations that are involved in high-risk business sectors or high-risk technologies. Finally, we provide practical recommendations to organizations about creating a layered defence approach, integrating security throughout the entire ML life cycle, developing formalized continuous testing and monitoring mechanisms, and creating an alignment between their technical and legal/risk teams. All of the recommendations are intended to assist organizations in creating resilient systems that can respond to emerging threats and sustain public assurance in the use of artificial intelligence.

12.2 Artificial Intelligence And Trust

Trust is the main reason behind the people who use AI in different fields like medicine and finance. However, the accuracy score of these AI systems is high; users tend to blindly trust them without knowing how the decisions are made (European Commission, 2019; and OECD, 2023). Where we stop questioning the AI and make it easy for adversarial attacks to cause a problem or disaster without anyone noticing it.

These adversaries are a major problem, which proves that AI can be easily tricked in such a way that they are ‘hidden’ and are not visible to human eyes, making them difficult to detect and more harmful. If the problem is not fixed, people will start losing faith in AI (Brundage et al., 2018; European Commission, 2019). Organizations may implement strict rules or become reluctant to use AI-based solutions.

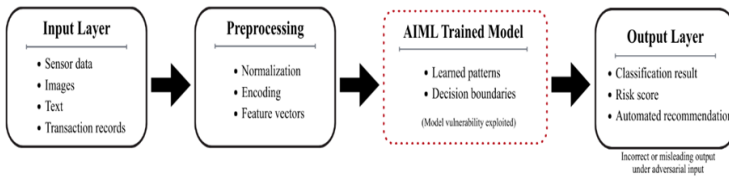


Figure 12.1: High-level AI system pipeline showing input, model, and decision stages with adversarial attack surfaces highlighted

12.3 Security As A Fundamental Concern

Security issues in AI are different than a normal or a regular bug. While most of the errors are accidental or unintentional, caused by coding mistakes, these adversarial attacks are deliberately made and adaptive ‘digital crimes’ that target the internal logic of AI (ENISA, 2021). Because the attackers can change their ways to attack the systems and use different tactics to bypass defence, they are more dangerous than normal glitches.

These adversarial attacks are much closer or related to cybersecurity threats rather than a normal technical failure. The failure is just not a normal computer error; it also becomes a major threat to financial and human stability. This raises serious ethical and legal problems: if AI makes a mistake, who is responsible (Brundage et al., 2018; and European Commission, 2019; and NIST, 2023)? We cannot treat this as a minor problem that can be fixed later. The need to treat AI as a fundamental component from the beginning ensures it stays safe and trustworthy for everyone.

12.4 Evolution of Adversarial Machine Learning

The development of AI systems led to increased interest in adversarial machine learning as a way of addressing security vulnerabilities in these systems. The fact that we can generate significant variations in the performance of an AI model through small changes to its input highlights the importance of this research. AI systems will be increasingly susceptible to manipulation in the future; therefore, we need to move away from a focus on quick fixes and towards developing a systematic approach for addressing these vulnerabilities.

12.4.1 Early Discovery of Adversarial Examples

Early studies used many experimental variations that humans would find nearly impossible to detect to misguide neural networks into providing results that would seem reliable yet were completely wrong (Szegedy et al., 2014).

This was accomplished because neural networks utilize decision-making rules that are completely different than the rules people use when making decisions.

Because small adjustments within many input dimensions can cause a cumulative change in the model's internal representation, creating completely different internal representations for the model will create completely different results. After conducting follow-up studies, researchers found that the phenomenon was not caused by coding errors or unusual characteristics of the original neural network design but, rather, is actually an intrinsic property of how statistical models are "banked" on data (Biggio and Roli, 2018; and Goodfellow et al., 2015). Additionally, the potential attack surfaces on high-resolution inputs (e.g., imagery, audio, and/or sensor inputs) are so broad that many small, geographically diverse input variations, made up of low-magnitude input changes, can result in major changes in outputs.

12.4.2 From Academic Curiosity to Security Threat

The adversarial examples, which were designed for one type of machine learning model, also have similar effects on other types of machine learning models (Figure 12.2). Additionally, many of the alterations made to an image as part of generating an adversarial example will persist after being subjected to real-world variations (e.g., shadowing or different camera angles) (Eykholt et al., 2018), making the practical application of adversarial examples even more dangerous. In summary, falsified traffic signs, adversarial clothing items, and modified medical scans can result in harm due to an adversarial attack on a person.

Rather than targeting bugs in the AI's code itself, adversarial attacks concentrate on the data flow within a machine learning system. Traditional defences against security breaches and incident response that utilize software code inspection, for example, cannot provide

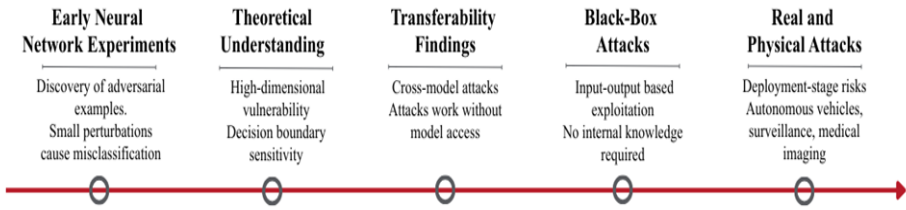


Figure 12.2: Evolution of adversarial machine learning

adequate protection against this form of attack (Carlini and Wagner, 2017; Carlini et al., 2019).

The use of adversarial examples to illustrate the findings in this research suggests that we should consider the adversarial risk of AI as being a systemic risk rather than one that is limited to unintentional or random events. To ensure the ongoing safety of AI in production, we must build multilayered technical protections into AI systems, continuously test those AI systems against adaptive adversaries, and adopt a governance framework and operational process that incorporates adversarial resilience as one of the core aspects of the entire machine learning lifecycle.

12.5 Taxonomy of Adversarial Attacks

There are three distinct ways to organize a methodology of performing adversarial attacks (Figure 12.3); they are based on: (i) how much information the attacker has about the target system, (ii) where the attack occurs in the machine learning system, and (iii) the data the attacker uses in performing the attack. These organizations provide clarity and form the basis for forthcoming work. We define each adversarial attack type and provide an example to clarify it.

12.5.1 Attacker Knowledge

There are three distinct levels of resources available to an adversary in their attempts to misclassify information using a machine learning model. These include white-box, grey-box, and black-box resources.

White-Box Resources: The adversary knows everything about the model, including its architecture, weight values, and how it was trained. Because of this, they can create very precise noise to specifically deceive the model, such as by having a model that

is supposed to recognize medical images incorrectly classify these images.

Gray-Box Resources: The adversary does not have all the same information but does know some contextual information, such as where the training data came from, what kind of model it is, etc., or they can create a 'shadow' copy of the adversary's model using their own data. With a 'shadow' model, the adversary can use it to create transferable attacks that will work on an actual deployed system (Papernot et al., 2017; Goodfellow et al., 2015).

Black-Box Resources: The adversary only has input/output access to the machine learning model, so they will have to use the input/output response of the machine learning model to try to deduce how it works. Even though they do not have access to the internals of the model (Papernot et al., 2017), the adversary may be able to identify weaknesses with the model by making repeated requests and modifying the input parameters after receiving their respective output responses.

12.5.2 Attack Stage

Inferences made away from the system during inference time may be compromised by a live input change designed to immediately warp the model into making an error.

One example is modifying a driver's perspective of a simple traffic sign or using a device that can mislead a driver at the moment of judging whether to stop or go.

Injecting jeopardizing information into a training dataset during training is also known as training time. As a consequence of injecting damaging information into a dataset (Akhtar and Mian, 2018; Ren et al., 2020), it creates opportunities for hackers to insert "Trojans" into that dataset that can be activated later with a predefined command, causing an eventual loss of trust in the supply chain and third-party model assurance.

Finally, adversaries can use the output of an adversarial third-party model to extract functionality from the model or to speculate on the private training data, which competitors could potentially use to reassemble the original proprietary model and/or assemble a competitive model of that proprietary model.

12.5.3 Data Modality

Vision: Small pixel changes will change the classification of images.

Implication: These small pixel changes, even when distributed around the surface of the image, could result in the incorrect classification of the image when used for medical or security purposes (Eykholt et al., 2018; Finlayson et al., 2019).

Original Text: Prompts and token substitutions that are done well will enable an actor that intends to do bad things, whether that actor is a human or a machine, to manipulate the outcomes created by language processing systems when those systems are used as customer service agents.

Embedded Commands in Audio and Tabular Data: Embedded commands in audio recordings may take the form of commands during the recording process. These commands may be in a very small amount of time or may be in numerical form, i.e., in a spreadsheet or in a log entry.

Implication: If an adversarial example is placed in audio, it can wake up an AI-driven voice assistant. In addition, if a financial entry is manipulated, it could affect the input of an automated risk assessment model, leading to erroneous results.

Adversarial vulnerabilities recur across knowledge levels, stages, and modalities, so defences must be layered and tailored to the specific threat profile.

12.6 Adversarial Attacks As An AI Security And Cyber-crime Issue

Adversarial attacks directed against artificial intelligence (AI) systems threaten to have security implications similar to traditional cyber-crime but with the added potential of catastrophic consequences due to their placement at the confluence of the cybercrime business model and the critical infrastructure industry. By contrast to unintended failures or accidents, adversarial attacks are planned and executed with malicious intent. This section will provide an overview of how adversarial attacks pose threats to critical infrastructure systems, the relationship of adversarial attacks to established cybercriminal

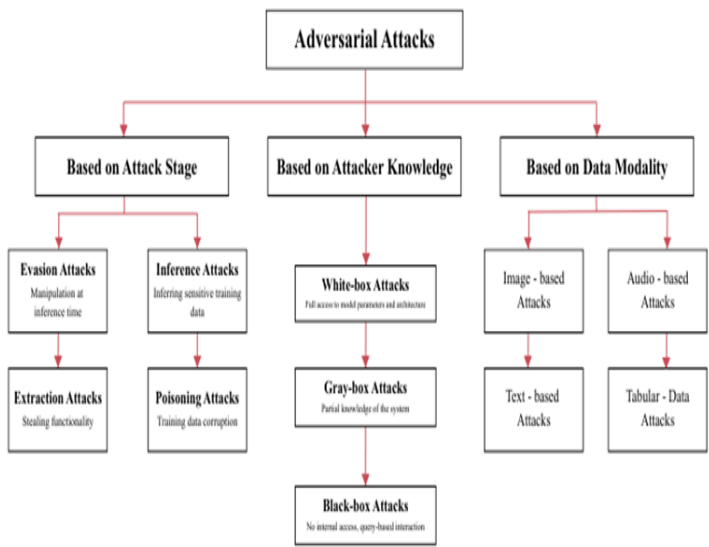


Figure 12.3: Hierarchical taxonomy of adversarial attacks in machine learning

models, and certain unresolved legal and accountability issues associated with adversarial attacks.

12.6.1 Threats to Critical Systems

In healthcare, adversarial manipulation of AI-based diagnostic and decision-support systems can have severe consequences. Small, carefully crafted changes to medical images or patient data may lead to incorrect diagnoses, delayed treatment, or inappropriate clinical decisions. Because clinicians often trust AI outputs as supportive evidence, such failures can directly affect patient safety (Finlayson et al., 2019).

Autonomous and semi-autonomous systems are also particularly vulnerable. In applications such as self-driving vehicles, drones, and smart transportation systems, adversarial inputs can distort the system’s perception of its environment. Manipulated traffic signs, altered road markings, or modified sensor signals can cause misinterpretation of critical cues, increasing the risk of accidents and system failure (Eykholt et al., 2018).

Then there is the financial sector, where the goal is usually fraud or market manipulation. Adversarial attacks trigger financial fraud, market manipulation, or ignoring regulatory constraints, leading to economic losses and loss of stakeholder confidence in automated financial technologies. Given the scale and speed at which financial AI systems operate, even small vulnerabilities can be amplified into significant systemic risks (ENISA, 2021).

12.6.2 Adversarial Attacks and Cybercrime

Attacks through adversarial means appear and function similarly to traditional cybercrimes, as they are intentional, targeted, and quick to adapt as defences improve. These attacks do not typically occur in the conventional manner. Rather than exploiting software vulnerability or a lingering misconfiguration, adversarial attackers will modify the training data that the model relies upon and use their own input that exploits the learned assumption of the model to result in a failure of the system while maintaining the same infrastructure (Carlini and Wagner, 2017).

Seeing these incidents through a cybercrime lens moves them from academic curiosities to real operational threats. Whether driven by financial motives, competitive sabotage, or political goals, adversarial machine learning is no longer confined to research settings; it is a real-world risk for deployed systems (Brundage et al., 2018; ENISA, 2021).

Reducing these risks goes beyond minor model adjustments. To manage adversarial threats, developers must implement existing cybersecurity best practices (threat modeling, structured risk assessments, continuous monitoring, and incident response playbooks) for AI systems. The defence will fail by having AI developers and security teams working independently. Coordinated efforts will be required across technical and organizational areas to transform AI from a potential vulnerability into an actionable resilience capability.

12.6.3 Legal and Accountability Challenges

Adversarial attacks are often stealthy and intentional, which makes them a poor fit for existing legal frameworks. It's rarely clear who should be held accountable when an AI system produces incorrect results because someone manipulated its inputs. Investigation,

enforcement, and remedying an issue become more complex when the responsibility for an event is shared by several entities, such as when these entities are referred to as multiple actors: system developers, organizations that deploy AI, data providers, and ultimately the attacker(s), who may also be held accountable for their actions (Brundage et al., 2018; European Commission, 2019; NIST, 2023).

Sectors like healthcare, finance, and transportation are especially vulnerable because their rules were written for software bugs or human mistakes, not to address failures triggered by adversarial interference. Current compliance standards don't always say how far organizations must go to anticipate or defend against novel adversarial techniques, leaving open questions about liability and reasonable due diligence when attacks exploit previously unknown weaknesses (NIST, 2023; OECD, 2023).

Addressing this requires a dual approach. Technical defences paired with institutional reform are necessary. Even though we need strong technical protections (such as comprehensive testing), there must also exist clear legal accountability that establishes how organizations will be held liable when they experience an attack on their technology, what types of technical protection should be employed, what evidence will be collected in order to establish whether or not an organization met its defensive obligations, and so on. These reforms create a public trust in AI researchers, developers, and users, as well as help promote responsible use and collaboration in the technology.

12.7 Defensive Strategies In AI Security

As awareness of adversarial threats has increased, a wide range of defensive strategies have been proposed to improve the robustness of machine learning systems. Methods try to shrink a model's attack surface, flag suspicious inputs before they're processed, or blunt the impact of malicious data once it's inside the system, but the problem is far from solved. This section breaks down the most prominent defences, critically examining not just how they work but the specific limitations that attackers continue to exploit (Ren et al., 2020; Abomakhelb et al., 2025).

12.7.1 Adversarial Training

Adversarial training is among the most extensively explored techniques for defending machine learning models against adversarial attacks (Goodfellow et al., 2015). It operates on a "training-by-fire" principle: you force the model to confront adversarial examples during its learning phase. It's like putting the model through stress tests during training, forcing it to face adversarial examples so it learns how to spot and resist them. This process toughens its decision boundaries, making the system more resilient when similar manipulations show up later in real-world use.

Adversarial training presents notable practical challenges, despite its effectiveness against certain attack types. Generating adversarial examples on the fly is computationally intensive and lengthens training cycles, increasing resource use and engineering complexity. It can also affect performance on clean data if not carefully balanced. A deeper limitation is limited generalization (Carlini et al., 2019; Abomakhelb et al., 2025). The robustness achieved through this method is typically closely tied to the specific attacks used during training. A model trained to resist a certain type of attack may fail against others. Because adversaries continually change tactics, a defence built around a fixed threat model can become outdated. As adversarial techniques continue to evolve, this dependence on known attack models limits the long-term reliability of adversarial training as a standalone defence.

12.7.2 Detection-Based Defence Mechanisms

Detection-based defence Mechanisms identify adversarial inputs before they can influence the decisions of a model. Detection methods typically rely on abnormal shifts in confidence, feature activation, and inconsistencies in latent representations as detected signals identifying potential adversarial input (Metzen et al., 2017). If an input is abnormal when compared to a prior behavior or expected normal behavior, it may be filtered before the user or model comes into contact with it, decreasing the risk of creating harmful or misleading outputs. Therefore, the purpose of detection methods is to filter data streams, allowing only legitimate input to be processed by the model.

Detection-based defence mechanisms can face significant challenges in a practical sense; attackers can strategically devise inputs that mimic a benign data set's statistical profile, bypassing detection methods based on assumption. Conversely, detection systems may present false positives, where a normal input is incorrectly classified as being adversarial. False positives interfere with model functionality and user experience (Carlini et al., 2019). For these reasons, detection mechanisms should not be treated as standalone solutions; detection mechanisms will have the greatest effectiveness and impact when combined with a broader security framework, allowing detection to serve as an early warning mechanism.

12.7.3 Input Transformation and Purification

Input transformation and purification methods aim to neutralize adversarial perturbations before data reaches the model. The process of adjusting input information using one or a combination of various methods (i.e., image smoothing, feature reduction, image compression, or image synthesis using generated or automatically generated models, etc.) aims to eliminate "adversarial" noise while keeping the required essential information for making accurate predictions (Meng and Chen, 2017). All three methods operate on the premise that "adversarial" manipulations have a limited life cycle and that small changes to an input image will affect an attack's pattern and/or implementation but will not change the meaningful representations contained within an input image.

Despite their promise, these methods face clear limitations. Transformations designed to remove adversarial noise can also distort clean inputs, reducing accuracy even when no attack is present. More critically, attackers can adapt by designing perturbations that survive the purification process or exploit its weaknesses, turning the defence into another target (Carlini et al., 2019; Abomakhelb, 2025). As a result, input transformation and purification are rarely sufficient in isolation. They are most effective when deployed as part of a broader, layered defence strategy that combines multiple protective mechanisms to address evolving adversarial threats.

12.7.4 Limitations of Existing Defences

Adversarial defence research is characterized by repeated themes of the inability to develop scalable and generalisable defences (Biggio

and Roli, 2018; Ren et al., 2020). Most defences are based on a very particular threat model and only appear effective under those very constrained circumstances. When an adaptive attacker looks at a defence system's mechanism(s), they will typically exploit this narrow view of defence and therefore defeat that defence. Because of this aspect of generalization, the ability to build a defence that will withstand the changing strategies of an adversary is extremely challenging; it will therefore remain fragile.

Robustness comes at a cost as well. Defensive strategies will generally degrade the accuracy of clear inputs, provide an increase in computational costs, and prohibit large-scale scalability. A defence strategy must be reliable, fast enough for practical applications, and not too slow to meet the demands of applications in the real world. Therefore, instead of seeking a single solution to adversarial defence, the focus on adversarial security is beginning to evolve into one of defence-in-depth; we now have to develop a defence that addresses multiple layers of defence to provide continuous evaluation and adaptation in order to meet new threats as they arise.

12.8 Towards Resilient And Self-Healing AI Systems

Self-healing resilient artificial intelligence systems poised to assume a very utilitarian view, which retains an understanding of the fact that there are a number of problems in artificial intelligence systems, which can never be completely solved but concentrate more on ensuring, in any given scenario, everything works when there is an issue recognized than on striving hard at making everything work properly in the first place. Such artificial intelligence systems shall be more likely utilized in tackling critical issues in the real world when considered from an operational view rather than being considered from an engineering point of view.

This includes where, practically speaking, there are aspects incorporated into a system of self-healing where there is a process of ongoing learning, monitoring, and correction incorporated into a method of the self-healing method (NIST, 2023). The chapter will also highlight the importance of this aspect of self-healing, where there is a process of ongoing learning, monitoring, and correction, and what this incorporates, a system of detecting and purifying, such

as a system known as LUGAI (Meng and Chen, 2017; Metzen et al., 2017).

12.8.1 Need For Resilience in AI Security

Ultimately, the ability of certain types of AI systems (models) to effectively operate in predictable environments differs widely from their effectiveness in unpredictable or hostile environments, where prior successful predictions may now turn out incorrect simply because an alternate method was used in the creation of the prediction itself; all of these prior assumptions made with the benefit of past experience with respect to reducing failure occurrences, detecting expected situations, reducing failures' impact, and recovering from failures no longer apply, as there has been a shift resulting from the resilient effect produced by changing variables.

In important applications such as healthcare or autonomous driving, system resiliency is used in a way that helps eliminate the actual damage of a cyberattack, such that a safe response can be provided to unusual inputs by the system itself, rather than by, or even in anticipation of, human action. This approach to AI security, called operational security, is a good fit in fault-tolerant system development models.

12.8.2 Self-Healing AI as a Design Philosophy

The AI can identify deviations due to the inclusion of self-healing functions into its regular work activities. An AI will be able to monitor for deviations in its behaviour by using several self-healing attributes such as determining the confidence levels associated with its actions and identifying uncertainty indicators. When a deviation occurs, an AI has to use its self-healing attributes to establish the type of deviation that occurred and what kind of corrective action is required. For example, it may use both input quality and fallback methods to correct for the deviation.

Building self-healing systems is essentially like performing trade-offs. Additional checks increase systems with both extra latency and extra price tags; highly sensitive detectors lead to false alarms, while extra pieces increase system complexity. This was achieved by having well-defined boundaries of what constitutes the system itself, what constitutes the high-risk decisions that need humans

in the loop, and constant evaluations so that mechanisms for self-healing improve with increasing robustness. The figure 12.4 shows conceptual self-healing ai loop illustrating detection, correction, and continuous operation.

12.8.3 Case Illustrative Detection and Purification

The detection and purification models may play an important function in the interpretation of self-healing mechanism implementations, including the classification stage after the detection of flagships and then the restoration. For example, LUGAI uses latent uncertainty and reconstruction deviation as supplementary classifiers. Here, data with values higher than a particular value will be conducted and implemented for reconstruction with a denoising autoencoder so as to create a new version of clean data and then report their status with corresponding values before and after classification.

This has the potential to minimize the blind spots we usually have as a result of the singular signal, in addition to the purification cycle and the procedure by which the signal is returned to the training manifold. This can create a more reliable prediction. Even though it is not foolproof because we are working with the element of an adapting adversary and potential latency, at least we have at least one familiar workable procedure.

12.9 Ethical, Social, And Management Implications

High-stakes applications of AI bring technical failures out of the shadows and into the public eye. When systems can cause harm, such as medical misdiagnoses, financial mistakes, or transportation accidents, the issue of safety is no longer merely an engineering problem but has become both a social and governance issue as well. Although technologies like LUGAI, which enhance risk management through uncertainty-driven detection and self-healing reconstruction, provide some mitigation against future harm, they also present new challenges for decision-makers in terms of their management, ethical considerations, and policy implications that need to be resolved.

12.9.1 Trust and Public Perception

When we have predictable behaviours and safeguards that we can see, we can trust them again. By identifying inconsistencies in a

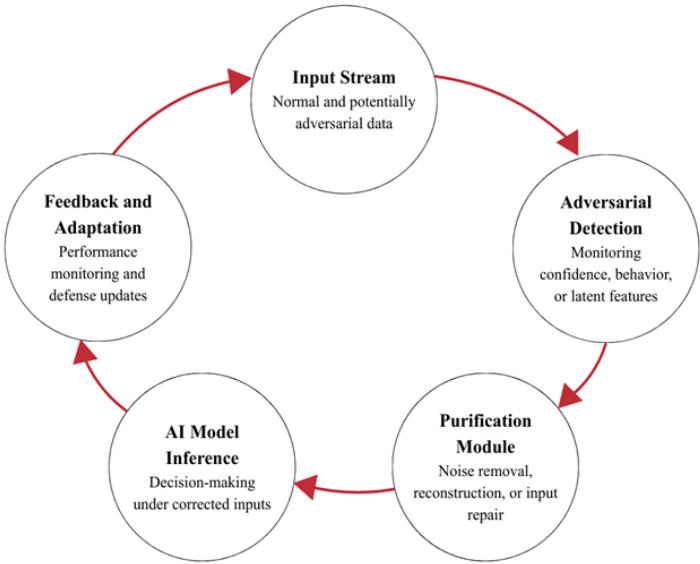


Figure 12.4: Conceptual self-healing AI loop illustrating detection, correction, and continuous operation

system and taking steps to rectify them, we can create new trust in the system by making failures visible or correctable rather than hidden. Providing proof of the technical controls to regulators or auditors will provide a way to support public trust through verified assurances of the controls provided by a system (European Commission, 2019; OECD, 2023).

12.9.2 Management and Organizational Difficulty

Resilience to adversarial threats is an ongoing commitment, not a one-time endeavor; due to the addition of threat detection and mitigations (including operational costs along with additional complexity) to the use of machine learning (ML), which also necessitates the application of multidisciplinary skill sets across ML, security, and legal fields. Some actionable steps for implementing practical solutions are to: -Establish a budget to support continuing efforts to monitor, update and red-team ML systems; -Establish cross-training processes that bring together engineering teams and security professionals with those who have legal backgrounds; -Develop playbooks that contain procedures for responding to alerts, triaging ML systems, and escalating required remedial actions; and -Provide operator training

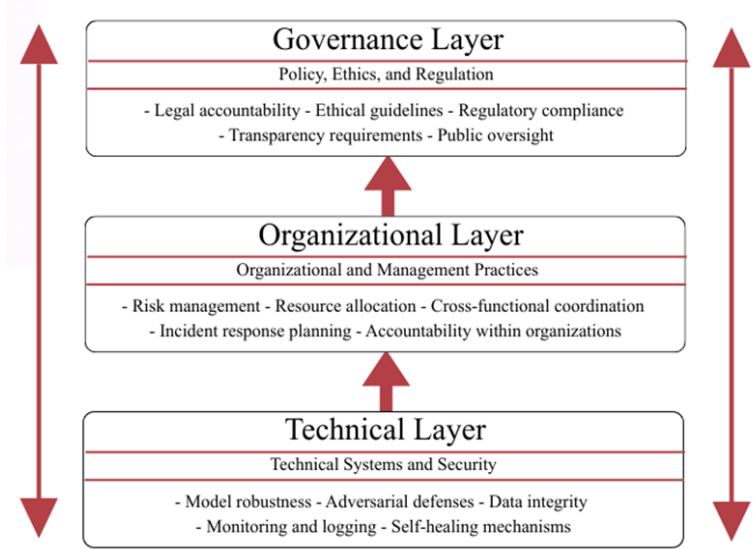


Figure 12.5: Layered framework illustrating technical, organizational, and policy dimensions of AI governance.

in how to read and respond to signals from ML systems. By addressing resilience from a product lifecycle management perspective, organizations can build sustainable and effective defences against threats to their ML systems (ENISA, 2021; NIST, 2023).

12.9.3 Policy and Governance Issues

While regulations that exist today focus on transparency, privacy, and fairness, they do not generally contain explicit requirements for adversarial robustness. When setting policy priorities, it is important to develop guidelines for protecting the entire lifecycle (through monitoring and logging) of safety-critical systems with mandated adversarial testing, develop standardized purification audit trails, and enhance the clarity of liability rules related to responsible sharing among developers, deployers, and operators. By being quantifiable in nature, tools such as LUGAI can help identify the measurable expression of the detection and remediation capabilities of systems, therefore assisting in the establishment of these guidelines (NIST, 2023; OECD, 2023). The figure 12.5 shows layered framework illustrating technical, organizational, and policy dimensions of AI governance.

12.10 Future Directions And Open Challenges

Adversarial defence is a dynamic game between the attackers and defenders. There are still many open questions down these promising paths.

Adaptable adversaries will change their attack methods in response to defences implemented. Adaptive learning must be part of future systems so the purification techniques and detection thresholds adapt to emerging threats (Abomakhelb et al., 2025).

Scaling from small image classifiers to large language models and multimodal systems, self-healing techniques have to struggle with computational issues and latency and work out appropriate definitions of trustworthy uncertainty signals across data types.

Standardized evaluation: The community requires benchmarks under realistic threat models that evaluate end-to-end latency, detection reliability, purification efficacy, and classification accuracy.

Human-in-the-loop design: For important issues, you can't just trust the system with trying to figure everything out by itself. The way we get around this is by including information so that both fixes and explanations are easy to read and easy to understand so that a human can review what action the system took and possibly even correct things if needed. The main challenge here revolves around displaying everything so that a human can make good decisions (European Commission, 2019; NIST, 2023).

Edge deployment, hardware: Simply put, real-time systems can't wait, so they need high response times, whereas the best of all is hardware that allows everything to run faster, including any threat detection algorithm that can run in concert, thereby executing faster, efficiently, like having threat response "out at the edge" of the network where you need it most!

The social and behavioral implications: As they appear to be more effective, they may generate more trust in the system. But, on the other hand, this could have a converse effect too. As they become more comfortable with the system, they may not pay enough attention to what they are actually doing. More studies must be done to understand their behavior with self-healing systems.

12.11 Conclusion

Adversarial attacks really shine a light on a tough reality: even the strongest systems can unexpectedly fall apart. So, just aiming for high accuracy isn't enough; we need to weave resilience into the very fabric of our architecture from the start. That's where the LUGAI framework comes into play, combining latent uncertainty detection with reconstruction-based purification (Meng and Chen, 2017; Metzen et al., 2017). In simpler terms, it's about spotting those strange, suspicious inputs and trying to fix them before the system makes any decisions. But let's be real, having cutting-edge algorithms isn't the whole story. For these systems to work effectively in the real world, we need more than just technology. We need people who are ready to embrace the changes, encourage open and honest reporting, stay alert, and put smart policies in place that set clear guidelines. Self-healing features can do wonders to keep things running smoothly, but they only shine when paired with genuine oversight and accountability. When we shift our mindset and stop seeing resilience as an afterthought, we can build AI that's not just smart, dependable, safe, and truly in tune with what people need (NIST, 2023; OECD, 2023).

References

- [1] Abomakhelb, A., Jalil, K. A., Buja, A. G., Alhammadi, A., and Alenezi, A. M. (2025). A comprehensive review of adversarial attacks and defense strategies in deep neural networks. *Technologies*, 13(5), 202. <https://doi.org/10.3390/technologies13050202>
- [2] Akhtar, N., and Mian, A. (2018) Threat of adversarial attacks on deep learning in computer vision: A survey. *IEEE Access*, 6, 14410–14430. <https://doi.org/10.1109/ACCESS.2018.2807385>
- [3] Biggio, B., and Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, 84, 317–331. <https://doi.org/10.1016/j.patcog.2018.07.023>
- [4] Brundage, M., Avin, S., Clark, J., et al. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute. <https://arxiv.org/abs/1802.07228>
- [5] Carlini, N., and Wagner, D. (2017). Towards evaluating the robustness of neural networks. *Proceedings of the IEEE Symposium on Security and Privacy*, 39–57. <https://doi.org/10.1109/SP.2017.49>

- [6] Carlini, N., Athalye, A., Papernot, N., et al. (2019). On evaluating adversarial robustness. *arXivpreprint*. <https://arxiv.org/abs/1902.06705>
- [7] ENISA. (2021). *Securing machine learning algorithms*. European Union Agency for Cybersecurity. <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>
- [8] European Commission. (2019). *Ethics guidelines for trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- [9] Eykholt, K., Evtimov, I., Fernandes, E., et al. (2018). Robust physical-world attacks on deep learning models. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1625–1634. <https://doi.org/10.1109/CVPR.2018.00175>
- [10] Finlayson, S. G., Bowers, J. D., Ito, J., et al. (2019). Adversarial attacks on medical machine learning. *Science*, 363(6433), 1287–1289. <https://doi.org/10.1126/science.aaw4399>
- [11] Goodfellow, I. J., Shlens, J., and Szegedy, C. (2015). Explaining and harnessing adversarial examples. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1412.6572>
- [12] Meng, D., and Chen, H. (2017). MagNet: A two-pronged defence against adversarial examples. *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, 135–147. <https://doi.org/10.1145/3133956.3134057>
- [13] Metzen, J. H., Genewein, T., Fischer, V., and Bischoff, B. (2017). On detecting adversarial perturbations. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1702.04267>
- [14] NIST. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. National Institute of Standards and Technology. <https://www.nist.gov/itl/ai-risk-management-framework>
- [15] OECD. (2023). *AI risk management and governance*. <https://www.oecd.org/digital/ai/>
- [16] Papernot, N., McDaniel, P., Goodfellow, I., et al. (2017). Practical black-box attacks against machine learning. *Proceedings of the ACM Asia Conference on Computer and Communications Security (AsiaCCS)*, 506–519. <https://doi.org/10.1145/3052973.3053009>
- [17] Ren, K., Zheng, T., Qin, Z., and Liu, X. (2020). Adversarial attacks and defences in deep learning. *Engineering*, 6(3), 346–360. <https://doi.org/10.1016/j.eng.2019.12.012>
- [18] Szegedy, C., Zaremba, W., Sutskever, I., et al. (2014). Intriguing properties of neural networks. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1312.6199>

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 13

EMOTION AI AND “FEELING AI” IN MARKETING: AD OPTIMIZATION, CONSUMER REACTIONS, AND ETHICAL BOUNDARIES

Seema Devi^{a,*} & Sheena^b

^{a,*}Department of Commerce, Ch. Dhajja Ram Janta Mahavidyalaya, Butana, Haryana, India- 131302.

^bDepartment of Commerce, B.K. Birla Institute of Higher Education, Pilani, Rajasthan, India-333031.

Email: sdphougatmd@gmail.com

Abstract

Emotion AI and “Feeling AI” represent a new frontier in marketing, where artificial intelligence (AI) systems are designed to detect, interpret and respond to human emotions in real time. The chapter develops a conceptual framework to examine how these technologies influence advertising optimization, consumer reactions and ethical boundaries. By integrating insights from psychology, computer science and marketing theory, the study highlights how Emotion AI can enhance personalization, improve message resonance and increase consumer engagement. At the same time, it raises concerns about manipulation, privacy invasion, algorithmic bias and the commodification of human feelings. The chapter argues that while Emotion AI offers marketers powerful tools to refine advertising strategies, its deployment must be critically assessed against ethical standards and consumer rights. The theoretical discussion emphasizes the dual role of Emotion AI: as a driver of innovation and efficiency and as a potential source of ethical dilemmas that challenge trust and transparency in digital markets. The study contributes to marketing scholarship by clarifying the conceptual boundaries between technological advancement and ethical responsibility, offering a foundation

for future empirical studies and policy debates. Ultimately, the chapter calls for responsible innovation, regulatory safeguards and interdisciplinary dialogue to ensure that Emotion AI serves both commercial interests and consumer well-being.

Keywords: emotion AI, feeling AI, advertising optimization, consumer reactions, ethical boundaries

13.1 Introduction

Artificial intelligence (AI) has rapidly transformed marketing practices, reshaping how businesses interact with consumers and optimize advertising strategies. Among the most recent developments is Emotion AI, sometimes referred to as “Feeling AI,” which focuses on detecting, interpreting and responding to human emotions through computational systems. Unlike traditional AI that relies primarily on rational data inputs, Emotion AI integrates affective computing, psychology and machine learning to capture subtle cues such as facial expressions, voice tone, physiological signals and online behavior patterns (Calvo and D’Mello, 2010; Peng et al., 2023). This shift represents a profound evolution in marketing, where emotional resonance becomes as critical as rational persuasion. Marketing has always relied on understanding consumer emotions, but Emotion AI introduces unprecedented precision and scalability. Advertisers can now tailor messages in real time, adjusting content based on detected emotional states. For example, emotion recognition tools embedded in digital platforms can identify whether a consumer is bored, excited, or skeptical, and then deliver advertisements that align with their mood (Koelstra et al., 2012; Hossain and Muhammad, 2019). This personalization promises higher engagement and conversion rates, making emotion. At the same time, the integration of Emotion AI raises ethical boundaries that cannot be ignored. Scholars have warned that emotional data is deeply personal and its misuse could lead to manipulation, privacy violations, and reinforcement of social biases (Crawford, 2021). Unlike demographic or behavioral data, emotional signals reveal intimate aspects of human experience, making their commodification ethically sensitive. The challenge for marketers and policymakers is to balance innovation with responsibility, ensuring that Emotion AI enhances consumer well-being rather than exploiting vulnerabilities.

The chapter develops a conceptual framework to examine three interconnected dimensions: (1) advertising optimization through Emotion AI, (2) consumer reactions to emotionally adaptive marketing, and (3) ethical boundaries that define acceptable practices. By synthesizing insights from interdisciplinary research, the study contributes to marketing scholarship by clarifying how Emotion AI can be harnessed responsibly. Emotion AI builds on the foundations of affective computing, a field pioneered by Picard in 1997, which argued that emotions are integral to human intelligence and should be incorporated into computational systems. In marketing, this translates into technologies that analyze consumer emotions to refine targeting, message delivery, and brand positioning (Hudlicka, 2016; Martinez-Miranda and Aldea, 2005). For instance, emotion recognition in video advertisements can predict audience engagement more accurately than traditional metrics such as click-through rates (McDuff et al., 2015). Recent studies show that emotionally adaptive advertising enhances consumer trust and brand loyalty when used transparently (Bagozzi et al., 1999; Vinciarelli et. al., 2008). However, when consumers perceive manipulation, their reactions can be negative, leading to resistance or backlash (Mallika and Siyad, 2025). This duality underscores the importance of ethical boundaries in deploying Emotion AI.

Consumer responses to Emotion AI are complex and context-dependent. Positive reactions often emerge when emotional personalization aligns with consumer needs, creating a sense of empathy and relevance (Kim and Biocca, 2018). Conversely, negative reactions occur when consumers feel surveilled or manipulated, particularly if emotional data is collected without consent (Cowie et al., 2011; Liu, 2025). Research suggests that cultural differences also shape consumer reactions. In collectivist societies, consumers may be more accepting of emotion-based personalization, while in individualist cultures, privacy concerns dominate (Hofstede, 2001). Understanding these nuances is essential for global marketing strategies.

The ethical implications of Emotion AI are profound. Emotional data, unlike transactional data, reveals inner states that individuals may not consciously express. Its collection and use raise questions about autonomy, consent, and fairness (Crawford, 2021). Scholars argue that ethical frameworks must be established to prevent exploitation, including guidelines for transparency, accountability,

and consumer control (Floridi et al., 2018; Mittelstadt et al., 2016). Bias in Emotion AI systems is another critical concern. Studies show that emotion recognition algorithms often misinterpret expressions across different ethnic groups, leading to discriminatory outcomes (Buolamwini and Gebru, 2018). In marketing, such biases could reinforce stereotypes and marginalize certain consumer segments. Addressing these issues requires interdisciplinary collaboration between technologists, marketers, ethicists, and regulators.

The study contributes to the literature by offering a conceptual framework that integrates advertising optimization, consumer reactions and ethical boundaries in the context of Emotion AI. While empirical studies are emerging, theoretical clarity is needed to guide responsible innovation. By situating Emotion AI within broader debates on digital ethics and consumer rights, the paper provides a foundation for future research and policy development.

13.2 Review of Literature

The growing body of scholarship on Emotion AI and “Feeling AI” in marketing reflects the convergence of affective computing, consumer psychology and digital ethics. Early work on emotional computing emphasized the role of affect in human–computer interaction, laying the foundation for marketing applications (Schuller et al., 2011; Zeng et al., 2009). Subsequent studies demonstrated that emotion recognition technologies could enhance advertising effectiveness by predicting consumer engagement more accurately than traditional behavioral metrics (Lewinski et al., 2014; Soleymani et al., 2012). Research in neuromarketing further highlighted how emotional cues influence consumer decision-making, showing that affective responses often precede rational evaluation (Plassmann et al., 2012; Yoon et al., 2006). Scholars have also examined the role of multimodal emotion detection, combining facial analysis, voice recognition and physiological signals to improve personalization in marketing campaigns (Baltrušaitis et al., 2018; Gunes and Schuller, 2013).

Consumer reactions to Emotion AI have been explored through the lens of trust, privacy and perceived manipulation. Studies reveal that while emotionally adaptive advertising can foster empathy and brand loyalty, it may also trigger resistance when consumers feel their emotions are being exploited (Coviello et al., 2014; Sundar

and Marathe, 2010). Cultural and demographic factors further complicate these reactions, with evidence suggesting that acceptance of emotion-based personalization varies across societies and age groups (Matsumoto et al., 2008; Jack et al., 2012). Ethical concerns dominate recent literature, particularly regarding bias in emotion recognition algorithms and the commodification of affective data (Barrett et al., 2019; Stark, 2018). Scholars argue that Emotion AI systems often misinterpret emotions across diverse populations, raising risks of discrimination and reinforcing stereotypes (Barker et al., 2025). Theoretical contributions emphasize the need for frameworks that balance innovation with responsibility. Researchers propose guidelines for transparency, accountability and consumer autonomy in Emotion AI applications (Jobin et al., 2019; Cath et al., 2018). Others highlight the importance of interdisciplinary collaboration, suggesting that ethical marketing with Emotion AI requires input from psychology, computer science, law, and cultural studies (Serrano-Puche, 2015; Stahl and Wright, 2018). Overall, the literature underscores both the promise and peril of Emotion AI in marketing, pointing to a dual trajectory where technological advancement must be tempered by ethical safeguards (Mittelstadt, 2019).

13.3 Theoretical Foundations of The Conceptual Model

The conceptual framework proposed in this study is supported by four interrelated theories: Affective Events Theory (AET), Technology Acceptance Model (TAM), Experiential Marketing Theory and Ethical Theories of Justice and Fairness. Each theory contributes to a different layer of the model, ensuring that the framework is both interdisciplinary and theoretically robust.

(i) Affective Events Theory (AET)

Affective Events Theory (Weiss and Cropanzano, 1996) argues that emotional experiences significantly shape individual attitudes and behaviors. In marketing, Emotion AI captures affective cues such as facial expressions, voice tone and physiological signals, which act as “affective events” influencing consumer decision-making.

(ii) Technology Acceptance Model (TAM)

The Technology Acceptance Model (Davis, 1989) explains how perceived usefulness and ease of use influence user acceptance of new

technologies. Applied to Emotion AI, TAM helps explain consumer reactions to emotionally adaptive advertising. If consumers perceive Emotion AI as enhancing personalization and relevance, they are more likely to accept it; if they perceive it as intrusive, resistance emerges.

Resultantly, acceptance or resistance reflects TAM's constructs of perceived usefulness and perceived intrusiveness.

(iii) Experiential Marketing Theory (EMT)

The Experiential Marketing Theory was developed by Prof. Bernd Schmitt from Columbia Business School in 1999. The theory firstly discussed in book entitled "*Experiential Marketing: How to Get Customers to Sense, Feel, Think, Act and Related to Your Company and Brands*" Schmitt highlight the five strategic experiential modules I.e. sense, feel, think, act and relate that can be used to manage purchaser's experiences. As per the theory the consumer is a rational and emotional human being who can be manipulated, entertained and immersed rather only seeker of information. Experiential Marketing Theory emphasizes that consumer decisions are shaped not only by functional benefits but also by emotional and sensory experiences (Holbrook and Hirschman, 1982; Brakus et al., 2009). Emotion AI enhances experiential marketing by tailoring advertisements to evoke specific emotions, thereby deepening consumer engagement.

(iv) Ethical Theories of Justice and Fairness

Normative ethical theories, particularly Rawls' Theory of Justice (Rawls, 1999) and Floridi's ethics of information (Floridi, 2016), provide the foundation for the Boundary Layer (Ethical Safeguards) in the framework. These theories argue that technologies must respect autonomy, fairness, and transparency. Emotion AI raises ethical dilemmas such as privacy invasion, emotional commodification, and algorithmic bias. Embedding justice and fairness theories ensures that ethical safeguards are integral to responsible innovation.

13.4 Conceptual Framework

The conceptual framework of this study integrates four interdependent layers—Emotion AI inputs, advertising optimization, consumer reactions, and ethical boundaries—to explain how Emotion AI and "Feeling AI" reshape marketing practices. Each layer is grounded in

a distinct theoretical foundation, offering a robust interdisciplinary lens through which to examine the dual trajectory of innovation and responsibility. Emotion AI systems operate by detecting and interpreting consumer emotions through multimodal inputs such as facial expressions, voice tone, and physiological signals. This process aligns with Affective Events Theory (AET), which posits those emotional experiences act as triggers for behavioral responses (Weiss and Cropanzano, 1996). These emotional insights form the Input Layer of the framework and are then applied to advertising optimization, enabling marketers to personalize messages, adjust timing, and refine content for greater resonance.

The Process Layer—advertising optimization—is supported by both the Technology Acceptance Model (TAM) and Experiential Marketing Theory. TAM explains how consumers evaluate the usefulness and intrusiveness of emotion-driven personalization (Davis, 1989), while Experiential Marketing Theory emphasizes the role of emotional and sensory engagement in shaping consumer experiences (Holbrook and Hirschman, 1982; Brakus et al., 2009). Together, these theories illuminate how Emotion AI enhances message relevance and emotional resonance, potentially increasing consumer engagement, trust, and brand loyalty. However, consumer reactions are not uniformly positive. While some individuals appreciate emotionally adaptive advertising for its empathy and relevance, others perceive it as intrusive or manipulative. These reactions—captured in the Outcome Layer—are shaped by cultural norms, demographic factors, and individual privacy concerns. Here again, AET and TAM provide explanatory power: AET accounts for emotional variability in consumer responses, while TAM helps interpret acceptance or resistance based on perceived value and ethical comfort. The fourth dimension—ethical boundaries—functions as a critical filter in the framework. Issues such as privacy invasion, emotional commodification, and algorithmic bias challenge the legitimacy of Emotion AI in marketing. This Boundary Layer is grounded in Ethical Theories of Justice and Fairness, particularly Rawls' principles of autonomy and equity (Rawls, 1999) and Floridi's ethics of information (Floridi, 2016). Ethical safeguards—including transparency, accountability, and regulatory oversight—are necessary to ensure that technological innovation aligns with consumer rights and societal values. Together, these layers form a cyclical framework:

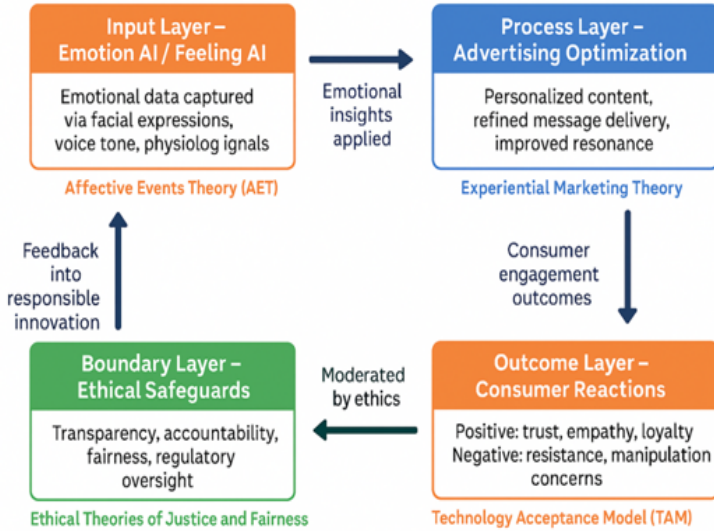


Figure 13.1: Conceptual Framework of Emotion AI and “Feeling AI” in Marketing

Emotion AI inputs □ *Advertising optimization* □ *Consumer reactions* □ *Ethical boundaries* □ *Feedback into responsible innovation*.

This cycle emphasizes that ethical considerations are not external constraints but integral to sustainable marketing practices. The framework can be visualized as a four-layered model, with each layer theoretically anchored and dynamically connected through feedback loops. The figure 13.1 illustrates conceptual framework of Emotion AI and “Feeling AI” in marketing.

This diagram illustrates a four-layered conceptual framework linking Emotion AI technologies to advertising optimization, consumer reactions, and ethical safeguards. Each layer is grounded in a distinct theoretical foundation:

- **Input Layer – Emotion AI / Feeling AI:** Emotional data captured via facial expressions, voice tone, and physiological signals. This layer is anchored in *Affective Events Theory (AET)*, which explains how emotional triggers influence behavior.
- **Process Layer – Advertising Optimization:** Emotionally adaptive advertising strategies that personalize content and refine message delivery. This layer draws on both the

Technology Acceptance Model (TAM) and *Experiential Marketing Theory*, which explain how perceived usefulness and emotional resonance drive consumer engagement.

- **Outcome Layer – Consumer Reactions:** Emotional responses to advertising, ranging from trust and empathy to resistance and perceived manipulation. This layer reflects both *AET* and *TAM*, emphasizing the role of affect and technology perception in shaping consumer behavior.
- **Boundary Layer – Ethical Safeguards:** Principles of transparency, fairness, accountability, and regulatory oversight. This layer is informed by *Ethical Theories of Justice and Fairness*, which ensures responsible innovation and protect consumer rights.

Arrows connect the layers in a clockwise flow:

“Emotional insights applied” → “Consumer engagement outcomes” → “Moderated by ethics” → “Feedback into responsible innovation.”

This cyclical structure emphasizes that ethical considerations are not external constraints but integral to sustainable marketing practices.

13.5 Application Scenarios And Use Cases

Emotion AI and “Feeling AI” are increasingly being deployed across diverse marketing environments, offering practical insights into consumer behavior and enabling emotionally intelligent engagement strategies. This section explores key application scenarios where the conceptual framework developed in this paper can be operationalized, demonstrating its relevance to both industry practice and academic inquiry.

13.5.1 Retail and E-Commerce Personalization

In retail and e-commerce, Emotion AI is used to personalize product recommendations, interface design, and promotional messaging based on real-time emotional feedback. For example, platforms like Alibaba and Amazon have experimented with emotion-aware interfaces that adapt product suggestions based on facial expressions and browsing behavior (Alshehri, 2026). Emotion AI can also enhance virtual try-on experiences by gauging consumer satisfaction and adjusting visual cues accordingly. These applications align with the

advertising optimization layer of the framework, where emotional insights are used to refine message delivery and improve resonance.

13.5.2 Digital Advertising and Campaign Management

Emotion AI is transforming digital advertising by enabling dynamic content adaptation. Advertisers can now test multiple versions of an ad and deploy the one that elicits the strongest emotional response from target audiences (Rehman, 2022). Real-time emotion tracking allows for programmatic ad placement that responds to user mood, increasing engagement and reducing ad fatigue. This use case illustrates the feedback loop between consumer reactions and advertising strategy, as outlined in the framework.

13.5.3 Political and Social Campaigns

Emotion AI has also been applied in political marketing and social awareness campaigns. By analyzing emotional responses to political speeches, social media posts, and campaign videos, strategists can tailor messaging to evoke empathy, urgency, or solidarity (Jain and Agrawal, 2022). However, this raises ethical concerns about emotional manipulation and consent, reinforcing the importance of the boundary layer in the framework. Studies have shown that emotionally charged political content can polarize audiences if not ethically moderated.

13.5.4 Healthcare Marketing and Patient Engagement

In healthcare marketing, Emotion AI is used to improve patient communication and engagement. Emotion-aware chatbots and virtual assistants can detect anxiety or confusion in patient interactions and adjust tone or content accordingly (Lima et al., 2022). Pharmaceutical companies have also used emotion recognition to test public reactions to health campaigns and adjust messaging for clarity and compassion (Stoumpos et al., 2023). These applications demonstrate how ethical safeguards must be embedded in emotionally sensitive domains to protect consumer well-being.

13.5.5 Entertainment and Media Platforms

Streaming services and media platforms use Emotion AI to curate content based on viewer mood and emotional preferences. Netflix

and Spotify, for instance, have explored emotion-based recommendation engines that adjust playlists or show suggestions based on facial cues or voice tone (Gupta et al., 2021). This enhances user satisfaction but also raises questions about emotional profiling and data privacy, reinforcing the need for transparent data governance.

13.5.6 Automotive and Smart Environments

Emotion AI is increasingly integrated into smart vehicles and environments to enhance user experience. In automotive marketing, emotion-aware dashboards can adjust lighting, music, and navigation prompts based on driver mood (Giri et al., 2023). These features are marketed as safety and comfort enhancements, but they also serve as data collection tools for behavioral targeting. The ethical implications of such ambient emotion tracking are significant and must be addressed through regulatory safeguards.

13.5.7 Cross-Cultural and Inclusive Marketing

Emotion AI applications must be adapted for cultural sensitivity and inclusivity. Studies show that emotion recognition algorithms often misinterpret expressions across ethnic and cultural groups, leading to biased marketing outcomes (Das and Roy, 2022). Inclusive marketing strategies require training Emotion AI systems on diverse datasets and incorporating cultural norms into emotional interpretation models. This aligns with the ethical boundary layer of the framework, emphasizing fairness and accountability.

13.6 Conclusion

Emotion AI and “Feeling AI” represent a transformative frontier in marketing, AI systems are designed to detect, interpret, and respond to human emotions in real time. This chapter has developed a conceptual framework that integrates advertising optimization, consumer reactions, and ethical boundaries, offering a structured lens through which to understand both the opportunities and challenges of these technologies. The framework emphasizes that Emotion AI is not merely a technical innovation but a socio-technical system with profound implications for consumer trust, cultural sensitivity and ethical governance. On one hand, Emotion AI enables advertisers to refine personalization, enhance message resonance, and foster deeper consumer engagement. On the other hand,

it raises critical concerns about manipulation, privacy invasion, algorithmic bias, and the commodification of human feelings. These dual trajectories highlight the need for responsible innovation, where commercial benefits are balanced against consumer rights and societal values. The theoretical contributions of this study lie in bridging marketing scholarship with insights from psychology, affective computing, and digital ethics. By situating Emotion AI within interdisciplinary debates, the framework clarifies conceptual boundaries and provides a foundation for empirical testing. Practical applications across retail, healthcare, political campaigns, and entertainment demonstrate the relevance of Emotion AI, while also underscoring the importance of ethical safeguards and regulatory oversight. Ultimately, the paper calls for future research that is methodologically rigorous, culturally inclusive, and ethically grounded. Interdisciplinary collaboration will be essential to ensure that Emotion AI evolves as a tool for innovation that respects human dignity and enhances consumer well-being. By advancing both theoretical clarity and practical guidance, this study contributes to ongoing discourse in marketing scholarship and provides a roadmap for navigating the ethical boundaries of Emotion AI in the digital marketplace.

References

- [1] Alshehri, A. (2026). Emotion-Aware Adaptive User Interfaces for Enhanced User Experience Using Multi-Modal Deep Learning. *International Journal of Advances in Soft Computing and its Applications*, 18,138-155. <https://doi.org/10.15849/ijasca.v18i1.31>
- [2] Bagozzi, R. P., Gopinath, M., and Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science*, 46(2), 184-206. <https://doi.org/10.1177/0092070399272005>
- [3] Baltrušaitis, T., Ahuja, C., and Morency, L. P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>
- [4] Barker, D., Tippireddy, M., Farhan, A. and Bilal, A. (2025). Ethical Considerations in Emotion Recognition Research. *Psychology International*,7(43),<https://doi.org/10.3390/psycholint7020043>
- [5] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., and Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>

- [6] Brakus, J. J., Schmitt, B. H., and Zarantonello, L. (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of Marketing*, 73(3), 52–68. <https://doi.org/10.1509/jmkg.73.3.052>
- [7] Buolamwini, J., and Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15. Available from <https://proceedings.mlr.press/v81/buolamwini18a.html>
- [8] Calvo, R. A., and D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18–37. <https://doi.org/10.1109/T-AFCC.2010.1>
- [9] Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., and Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), 505–528. <https://doi.org/10.1007/s11948-017-9901-7>
- [10] Coviello, L., et al. (2014). Detecting emotional contagion in massive social networks. *PLoS ONE*, 9(3), e90315. <https://doi.org/10.1371/journal.pone.0090315>
- [11] Cowie, R., et al. (2011). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18(1), 32–80. <https://doi.org/10.1109/79.911197>
- [12] Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. *Yale University Press*.
- [13] Das, S., and Roy, A. (2022). Cultural bias in emotion recognition systems: Implications for inclusive marketing. *International Journal of Human-Computer Studies*, 165, 102838.
- [14] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- [15] Floridi, L. (2016). The ethics of information. *Oxford University Press*.
- [16] Floridi, L., Cowls, J., Beltrametti, M., et al. (2018) AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds & Machines*, 28, 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- [17] Giri, M., Bansal, M., Ramesh, A., Satvik, D., and Damotharan, U. (2023). Enhancing Safety in Vehicles using Emotion Recognition with Artificial Intelligence. *IEEE 8th International Conference for Convergence in Technology*, 1-10. <https://doi.org/10.1109/I2CT57861.2023.10126274>
- [18] Gunes, H., and Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions.

- Image and Vision Computing*, 31(2), 120–136. <https://doi.org/10.1016/j.imavis.2012.06.016>
- [19] Gupta, R., Mehta, S., and Bansal, A. (2021). Emotion-based content recommendation in streaming platforms. *Multimedia Tools and Applications*, 80(17), 25987–26005.
- [20] Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. Sage. [https://doi.org/10.1016/S0005-7967\(02\)00184-5](https://doi.org/10.1016/S0005-7967(02)00184-5)
- [21] Holbrook, M. B., and Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, 9(2), 132–140. <https://www.jstor.org/stable/2489122>
- [22] Hossain, M. S., and Muhammad, G. (2019). Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion*, 49, 69–78, ISSN 1566-2535. <https://doi.org/10.1016/j.inffus.2018.09.008>
- [23] Hudlicka, E. (2016). Affective computing for AI-based marketing. *International Journal of Human-Computer Studies*, 95, 27–38.
- [24] Jack, R. E., Garrod, O. G., and Schyns, P. G. (2012). Dynamic facial expressions of emotion reveal cross-cultural differences in recognition. *Journal of Experimental Psychology: General*, 141(4), 684–696.
- [25] Jain, P., and Agrawal, R. (2022). Emotion AI in political marketing: Opportunities and ethical challenges. *Journal of Political Marketing*, 21(3), 245–263.
- [26] Jobin, A., Ienca, M., and Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- [27] Kim, J., and Biocca, F. (2018). Telepresence via television: Two dimensions of telepresence. *Media Psychology*, 21(4), 582–610.
- [28] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., and Patras, I. (2012). DEAP: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31. <https://doi.org/10.1109/T-AFFC.2011.15>
- [29] Lewinski, P., Fransen, M. L., and Tan, E. S. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics*, 7(1), 1–14. <https://doi.org/10.1037/npe0000012>
- [30] Lima, E. S., et al. (2025). Emotional recognition technologies applied to health: review and challenges. *Brazilian Journal of Psychiatry*, 47, e20243963. <https://doi.org/10.47626/1516-4446-2024-3963>

- [31] Liu, H. (2025). The Application Prospects of Consumer Emotion Recognition Technology in Marketing Strategies. *Modern Economics & Management Forum*, 6, 334. <https://doi.org/10.32629/memf.v6i3.4002>
- [32] Mallika, and Siyad, P. K. (2025). The Impact of AI-Driven Personalization on Consumer Purchase Intentions. *International Journal of Integrated Research and Practice*. 101-109. <https://doi.org/10.65579/31075037.092>
- [33] Martinez-Miranda, J., and Aldea, A. (2005). Emotions in human and artificial intelligence. *Computers in Human Behavior*, 21(2), 323–341. <https://doi.org/10.1016/j.chb.2004.02.010>
- [34] Matsumoto, D., Yoo, S. H., and Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology*, 94(6), 925–937. <https://doi.org/10.1037/0022-3514.94.6.925>
- [35] McDuff, D., Kaliouby, R., Cohn, J., and Picard, R. (2015). Predicting Ad Liking and Purchase Intent: Large-Scale Analysis of Facial Responses to Ads. *IEEE Transactions on Affective Computing*, 6(3), 223-235. <https://doi.org/10.1109/TAFFC.2014.2384198>
- [36] Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- [37] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- [38] Peltier, J. W., Dahl, A. J., Drury, L., Khan, T., and Wang, C. (2026). Here's the interactive marketing in the interactive marketing literature? A state of the art review and research agenda. *Journal of Research in Interactive marketing*, 20, 397-439. <https://doi.org/10.1108/JRIM-03-2025-0135>
- [39] Peng, Z., Desmet, P. M. A., and Xue, H. (2023). Mood in Experience Design: A Scoping Review, *She Ji: The Journal of Design, Economics, and Innovation*, 9(3), 330-378, ISSN 2405-8726. <https://doi.org/10.1016/j.sheji.2023.09.001>
- [40] Plassmann, H., Ramsøy, T. Z., and Milosavljevic, M. (2012). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology*, 22(1), 18–36. <https://doi.org/10.1016/j.jcps.2011.11.010>
- [41] Rawls, J. (1999). A theory of justice. *Harvard University Press*.
- [42] Rehman, A. (2022). Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges. *Information*, 13, 268. <https://doi.org/10.3390/info13060268>

- [43] Schuller, B., Rigoll, G., and Lang, M. (2011). Hidden Markov model-based speech emotion recognition. *Speech Communication*, 53(9–10), 1162–1171. <https://doi.org/10.1109/ICME.2003.1220939>
- [44] Serrano-Puche, Javier. (2015). Emotions and Digital Technologies: Mapping the Field of *Research in Media Studies*. <https://doi.org/10.13140/RG.2.2.31240.62729>
- [45] Soleymani, M., Lichtenauer, J., Pun, T., and Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 3(1), 42–55. <https://doi.org/10.1109/T-AFFC.2011.25>
- [46] Stahl, B. C., and Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *Technology in Society*, 54, 35–45. <https://doi.org/10.1109/MSP.2018.2701164>
- [47] Stark, L. (2018). Algorithmic psychometrics and the scalable subject. *Social Studies of Science*, 48(2), 204–231. <https://doi.org/10.1177/0306312718772094>
- [48] Stoumpos, A. I., Kitsios, F., and Talias, M. A. (2023) Digital Transformation in Healthcare: Technology Acceptance and Its Applications. *International Journal of Environmental Research and Public Health*, 20(4), 3407. <https://doi.org/10.3390/ijerph20043407>
- [49] Sundar, S. S., and Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power. *Human Communication Research*, 36(3), 298–322.
- [50] Vinciarelli, A., Pantic, M., Bourlard, H., and Pentland, A. (2008). Social Signals, their Function, and Automatic Analysis: A Survey. *ICMI'08: Proceedings of the 10th International Conference on Multimodal Interfaces*. <https://doi.org/10.1145/1452392.1452405>
- [51] Weiss, H. M., and Cropanzano, R. (1996). Affective Events Theory: A theoretical discussion of the structure, causes, and consequences of affective experiences at work. *Research in Organizational Behavior*, 18, 1–74.
- [52] Yoon, C., Gutchess, A. H., Feinberg, F., and Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, 33(1), 31–40. <https://doi.org/10.1086/504132>
- [53] Zeng, Z., Pantic, M., Roisman, G. I., and Huang, T. S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 39–58. <https://doi.org/10.1109/TPAMI.2008.52>

Chapter 14

THE END OF THE “MARKET FOR LEMONS”: AI AS A CATALYST FOR INFORMATION ASYM- METRY AND WELFARE OPTIMISATION

Rishika Sharma* & Nitika

Assistant Professor, Department of Economics, Khalsa College for Women, Civil Lines, Ludhiana, Punjab, India-141001.

Email*: rishikasharma301997@gmail.com

Abstract

Information is the foundational currency of efficient markets, yet its unequal distribution often leads to systematic collapse and poor decision making. In today's incredibly data-driven economy, where the traditional challenges of asymmetric information are going through a radical transformation, classical economic challenges continue to explain why information causes market failure leading to contractual inefficiencies. This chapter explores Akerlof's (1970) *Market for Lemons* model posits asymmetric information as a fundamental cause of market failure, leading to adverse selection, inefficient pricing, and the collapse of otherwise welfare-enhancing markets. Subsequent work by many economists, particularly Stiglitz (1981), shows the broader implications of information asymmetry in the context of moral hazard and contractual inefficiency. This chapter revisits the lemons market framework in the context of recent advances in artificial intelligence (AI), arguing that AI represents a structural shift in how information asymmetries are generated and processed in modern markets. We conceptualize AI as a digital encyclopedia of information and guesstimate technology that solves the core problems prominent in the lemons market through three mechanisms: (i) algorithmic screening helps in improving the gap of information between buyers and sellers; (ii) dynamic signaling via AI-mediation helps

in analysing huge datasets to understand buyers needs and identify problems by reducing uncertainty; and (iii) reductions in moral hazard enabled by continuous monitoring and real-time data analytics. These mechanisms address the problem of adverse selection between peaches and lemons without relying heavily on costly signaling models. This will further reduce deadweight loss and improve allocative efficiency by enhancing the mechanism. While there persists a concern regarding the concentration of informational power and the emergence of new forms of intelligence asymmetry, the chapter suggests that even if this situation is plausible, AI still functions as a welfare catalyst rather than a source of market failure.

Keywords: artificial intelligence, asymmetric information, market for lemons, moral hazard, welfare optimisation

14.1 Introduction

The information revolution is a revolution in the way we think about the economy... Information has become the most important factor of production, more important than land, labor, and capital.

Joseph Stiglitz, Nobel Laureate (2002)

In the classical Newtonian view of economics, markets were assumed to be stable environments in which buyers and sellers have perfect knowledge. However, the reality of the 21st-century global economy is defined by dynamic fluctuations. Markets are characterized by imperfect knowledge, leading to information asymmetry in which one party in a transaction possesses more information than the other. This silent friction is what leads to market failure, inefficient pricing, and the systemic exclusion of the marginalized and underprivileged. As we stand on the cliff edge of the age of intelligence and robotics, there has been a significant shift in paradigm regime from AI being a mere tool of automation to becoming the definitive corrective force for these market failures. This metamorphosis from a data-poor state to a data-rich economy uses AI to bridge the gap between the formal and informal sectors.

To understand the magnitude of AI intervention globally, we examine the theoretical foundations of market failures, including

information asymmetry, market signaling, and moral hazard. In the Indian context, markets are characterized by linguistic fragmentation, institutional voids, dominance of the informal sector, and a lack of a proper trust mechanism. These factors, together, created an environment of extreme asymmetry and may hamper the functioning of demand and supply chains. India's AI Impact Summit 2026 highlighted the significance of AI as the symmetry engine. It is emphasized that, for the Global South, AI's primary role is not to replace humans, but to fill gaps left by missing institutions and bridge the gap between the known and the unknown. Whether it is a coffee farmer in Ethiopia or a weaver in Varanasi, the problem remains the same: How do I prove my quality to a world that cannot see me? The solution requires combining open-source AI with public digital rails, which act as the new global blueprint for market formalization.

This chapter aims to unify fifty years of Nobel-winning economic theories with the lived experience of 21st-century India. This chapter will explore how AI is addressing Adverse Selection in the used-goods, tech-goods, and Agri-markets. It centers on alleviating Moral Hazard through continuous AI-driven monitoring in the FinTech sector, as well as resolving the principal-agent problem via Monitoring.

Now, let's dive into the theoretical aspect of various types of market failures:

14.1.1 The Ghost of the Used Car Lot: Akerlof's Persistent Shadow of Uncertainty

The paper on "*Market for lemons*" by George Akerlof (1970) provides the fundamental framework of Modern Economics. His thesis established a profound premise: in markets where the seller knows more about a product's quality than the buyer does.

According to his framework, when a seller has more information about a product than a buyer, the seller is punished. Using the used-car market as his primary example in his paper, he exhibited that if a buyer is unable to differentiate between a high-quality car known as a "peach" and a low-quality car as a "lemon", the buyer will only be willing to pay an average price. This occurs due to a lack of correct information. This creates a situation in which sellers of high-quality goods don't get fair prices for their peaches and exit the market, leaving only the lemons. This leads to a phenomenon known

as Adverse Selection. This ultimately causes market failure due to information asymmetry across diverse sectors, ranging from health insurance fraud to the second-hand car market, to bad credit markets, to the skilled and unskilled labor market.

In the Indian context, reliance on informal intermediaries like Dalals leads to the failure of Indian Bazaars due to unverifiable data and a lack of systematic data management. However, India's transformation into a global powerhouse calls for a more organized, algorithmic approach to trust mechanisms in the fields of borrowing, the labor market, and quality producers, who are currently trapped in a spiral of misinformation and high costs.

14.1.2 Adverse Selection to Screening through the Stiglitz Ian Pivot

While Akerlof (1970) focused on quality uncertainty and Spence (1978) on Signaling, Joseph Stiglitz (1975, 1981) introduced the concept of “*Screening*” under “*Adverse Selection*”. Stiglitz argued that in the world of information asymmetry, the buyer, or the uninformed agent, plays an active role in designing mechanisms to extract hidden information from the seller. As a solution to this problem, Stiglitz proposed a framework for contract design—such as offering varying insurance deductibles to force high- and low-risk individuals to reveal their true nature through their choices. For decades, the Indian market suffered from credit rationing. Stiglitz and Weiss (1981) proved that banks cannot raise interest rates to compensate for the risk posed by lemons, as doing so drives away “Peaches,” who are honest borrowers.

14.1.3 Michael Spence and the Evolution of the “Special Signaling.”

While Akerlof focused on market failure due to information asymmetry, Michael Spence (1978) investigated the burden of the seller or the informed agent. He answered the query about proving a high-quality seller as innocent and not a lemon. Spence used the model of “*Signaling*” to kick-start a trust mechanism. In his foundational example of the job market, a university degree serves as a signal not because of the knowledge gained, but because the “cost” (effort,

time and intellect) of obtaining it is lower for high-ability individuals than for low-ability ones, creating a Separating Equilibrium.

”Special Signals” were socially gated in the Indian Economy, as a small merchant could signal their creditworthiness through physical collateral, such as gold or property, or through historically lasting family relations and reputations. This led to a high signaling cost, further trapping millions of India’s informal sector in a Pooling equilibrium in which all types of sellers take the same action, making it difficult to distinguish between peaches and lemons. Therefore, all the sellers are treated as Lemons by default.

14.1.4 Moral Hazard: The Post-transactional Friction

Moral Hazard is a post-contractual problem arising from hidden intent or a lapse in morality on the part of the buyer. This concept has historical origins and was developed by Kenneth Arrow (1978), in which someone else pays for another person's risky behavior. Behavioral changes, such as reckless spending after a credit card is issued or speeding and rash driving after an automobile’s insured becomes the primary driver, can lead to market collapse. This creates an information gap similar to the Lemons market, due to the insurer's lack of monitoring, as the insurer's intentions are difficult to judge. This eventually leads to the principal-agent problem, which arises when the employer, the principal, hires an employee, the agent. The principal is unable to monitor the agent’s actions due to information asymmetry. This type of issue is more prominent in contractual relationships, banking, and corporate settings, where the principal and agent work under different incentives. The agent’s goal is to maximize his personal utility through incentives such as job security, prestige, and working less for a higher salary. On the other hand, a principal’s goal is to focus on higher profits and stock prices. This creates an imbalance leading to conflict of interest and high agency costs.

14.2 Structural Failures of The Problem of Information Asymmetry

Lemon’s market is a description of systematic failures where trust becomes an expensive commodity. After investigating the

fundamental theories of market failures, the following core problems are outlined that are faced on account of Information asymmetry:

- **Adverse Selection:** It refers to the problem faced by sellers when buyers offer an average price due to pooling in equilibrium. This results in a loss for high-quality sellers and an adverse selection of participants.
- **Market Unraveling:** As high-quality sellers exit, the average quality declines further. Buyers respond by lowering their bids again, triggering a feedback loop that continues until the market shrinks to a negligible size or collapses entirely.
- **The Dishonesty Premium:** Honest sellers face unfair competition, where their honesty is taxed while deception is rewarded due to the lack of verification for high-quality product sellers.
- **Cost Inefficiency:** The cost of generating a signal to prove the quality of a peach product is very high. This generates a massive deadweight loss in the economy, in which much of the resources are spent proving quality through expensive certifications rather than enhancing productivity.
- **Counterfeiting and Noise:** Lemons often generate noise by copying Peaches' signals, reducing the creditworthiness of all signals and brutally impacting the honesty of high-quality sellers—for instance, fake resumes in the labor market or superficial aesthetic repairs on a second-hand car.
- **Barriers to Entry for the Peach:** If a signal is purely financial, high-quality individuals who lack resources and initial capital are locked out of the market.
- **Moral Hazard:** Once a transaction is completed, the informed party may change their behavior because the uninformed party cannot monitor the informed party's behavior or assess their credibility. This kind of issue is commonly seen in insurance and credit markets.
- **The Stiglitz-Weiss Paradox:** There is a lack of distinction between risky and safe borrowers in an economic setting. One cannot solve the problem of the trust mechanism by simply raising interest rates to weed out poor-quality suppliers, as this would not address the lack of screening. Consequently, this limits credit supply, leaving productive businesses poorly funded.

- **Middleman's Monopoly:** Informed agents, such as brokers, moneylenders, or specialized agents, may exploit uninformed agents and extract rents or charge high prices. These intermediaries increase the total transaction costs for both buyers and sellers.
- **The Sorting Failure:** Job Markets often fail to sort employees and applicants into the right tiers. A highly skilled worker may end up in a low-skill job because the employer lacks a screening mechanism to verify their true capabilities, leading to a massive misallocation of human capital and underemployment.

14.3 AI As An Information Catalyst: The End Of Information Scarcity

The concept of a machine operating independently dates back thousands of years, to the era of ancient philosophers who questioned the world around them. The earliest records of automatons referred to by Plato and later made by Leonardo Da Vinci around 1495 prove that the history of AI as a catalyst began long before the current era of large language models. The late 1950s to the 1970s were a time of creation and innovation. It was during this time period that the term Artificial Intelligence (AI) was coined. We see the development of programming languages, literature and art that explored the idea of robots. Thus, the idea of AI became main stream during the 1980s, leading to a period of rapid growth of AI, now known as the "AI boom." "Currently, the captivating human creativity of creating artificial sentient beings is on the rise. There is a scientific renaissance thanks to this unwavering quest, where the development of AI holds the key to unlocking a bright future, acting as a catalyst for global wealth and a beacon of enlightenment. Our task is to responsibly and sustainably harness AI's extraordinary powers.

Historically, information management was handled through centralized institutions, including state and private regulators, which were slow, expensive, and prone to corruption and bias. Then came a time when Akerlof, Spence, and Stiglitz laid the foundation for market failure. Mathematical proof showed that "hidden information" could destroy markets. Since then, digitalization in economics has been steadily advancing towards the "Catalytic Age." First came the Internet, which reduced the cost of moving information, followed by Big Data, which increased the volume and hampered the quality of

information. Today, AI represents the third big stage: the ability to verify and interpret information autonomously.

In economics, AI acts as a catalyst by establishing itself as a trust mechanism in a market. In developing countries like India, it is difficult for small business owners to obtain loans due to skyrocketing interest rates, which are fixed to deter lemons in the market. Traditional Signaling makes it difficult for new firms to enter the market, thereby giving rise to monopolies on account of historical significance or a long-lasting family name. AI can therefore help in creating digital meritocracy where data serves as the signal. Information asymmetry further hinders achieving Pareto efficiency by diverting resources from the poor. AI can help prevent this deadweight loss through inclusivity, formal participation, and resource optimization. Thus, the implications of this paradigm shift are profound on the markets and society as a whole.

14.3.1 AI as a Re-balancer

In the classical "Market for Lemons," the fundamental tragedy is of Information Scarcity. AI serves as a universal solution that addresses a range of problems, including quality uncertainty, screening, costing, resource utilization, and even moral hazard. Artificial Intelligence (AI) represents a structural shift from Information Asymmetry to Information symmetry.

- **Solving the signaling trap:** AI by moving the signal from Intent-based to Action-based, reduces "Social Waste" in the form of money and resources spent by putting these back into the production function. Meanwhile, AI replaces massive expenditure signals with simple computational signals through the digital footprints we leave every day.
- **Solving the screening paradox:** AI helps solve problems by providing a dynamic environment rather than a static screen, predicting your future self -based on history. This expands the market's Pareto frontier and helps differentiate the peach from a lemon without the need for paper audits or frictional interviews.
- **Solving the problem of moral hazard:** Through IoT Telemetry, computer vision and predictive behavioral analysis, AI provides continuous and low-cost monitoring by acting as a Digital Eye.

14.3.2 The AI Correction Framework

To move from a collapsed "Lemons Market" to an efficient, symmetric market, a structured process can be followed:

Step 1: Data Granularization

Before AI can solve asymmetry, it must investigate the type of microdata previously hidden. This helps in finding the small units of missing data that lead to Akerlof's pooling equilibrium.

Step 2: Pattern Recognition and Extraction

After the data is captured, AI, being a machine learning model, helps identify the objective truth required to bypass Special Signal Noise.

Step 3: Automated Truth Synthesis

AI synthesizes these patterns into an actionable unit output, such as credit Scores, a health rating, or a quality check certificate. This is a low-Cost signal, as an AI-generated trust score can be generated in just milliseconds, benefiting the honest agent.

Step 4: Algorithmic Screening (Sorting the Market)

Lenders or buyers use AI agents to screen the entire market simultaneously, eliminating credit rationing and allowing the market to offer personalized prices rather than "Average" prices by placing different people into different risk brackets.

Step 5: Continuous Monitoring and Moral Hazard Mitigation

The "Lemons" problem often arises after the transaction is complete, leading to moral hazard. AI-linked IoT sensors ensure that a borrower is using a loan for its intended purpose, preventing the market from sliding back into a death spiral.

14.4 Empirical Analysis: AI Applications In Mitigating Information Asymmetry: Ai-Driven Solutions In Different Sectors Of The Economy

Case Study 1: Solving Signaling Gaps

Platform: DeHaat

The Problem: Small farmers in Bihar and Karnataka are forced to sell their harvests at exploitative rates to intermediaries. There's no way for them to signal the quality of their product.

The AI Solution: Platforms like DeHaat in India use AI-driven quality grading. This app connects suppliers to buyers on a single platform. A farmer has to take a picture of their product and upload it to the app or website. The AI instantly provides a quality grade for the grains based on their color, shape, and size. Further, the AI produces information signals that indicate the product's high quality to potential buyers, making this process low-cost and free of middlemen's interference, and increasing the farmer's income by 30%. The platform's operation is documented in The NITI Aayog Frontier Tech Repository (2025) as a full-stack agricultural solution that bridges the gap between pre- and post-harvest quality grading.

Case Study 2: Adverse Selection Pickle

Platform: Cars 24

The Problem: The second-hand Indian car market is one of the ultimate markets for Lemons. The uninformed agents' paranoia about car quality led to a decline in the prices of high-quality cars, further pushing the Peaches out of the market.

The AI Solution: CARS24 introduced an AI-driven pricing engine backed by a 300-point automated inspection, which analyses millions of historical data points to set a true value based on actual evidence. This has helped in removing the information gap by providing certified signals that highlight the quality of a product and its seller.

Case Study 3: Screening the Credit Rationing

Platform: KreditBee

The Problem: Banks refuse to lend to MSMEs, which are generally credit rationed because they lack collateral and credit history, thereby creating an information gap.

The AI Solution: Screening borrowers with AI-generated alternative signals helps identify reliable borrowers who look like Lemons to banks but are peaches in reality. Information about signals is generated based on factors such as the velocity of UPI transactions, Mobile payments, GST filings, etc. This has helped in achieving billions in credit for small businesses. Recent empirical studies

published in 2025 provide a clear benchmark for the effectiveness of these mechanisms. Findings of the research show that AI-based models have significantly reduced loan default rates to 7%, compared with 13% observed with traditional statistical models.

Case Study 4: Containing Moral Hazard

Platform: ICICI Lombard

The Problem: Post-contractual moral conduct is the core issue driving market failure. Once a car or health policy is issued, there is a moral hazard: the user might become a reckless driver and start speeding just because they are insured.

The AI Solution: Platforms like ICICI Lombard use a real-time mechanism in which customers upload photos of their cars in the app, making damage verification easier through AI-conducted Break-in Inspections, Telematics and AI Monitoring. In motor insurance, AI tracks driving behavior through usage-based insurance. Incentives like low premiums for safe driving attract customers, making it profitable for them to stay safe even after they are insured. Their InstaSpect tool uses live video and AI to approve motor claims in minutes. This helps in reducing information lag and fraudulent claims by the wrong customers.

14.5 AI And India: A Fertile Ground For Symmetry

India is bifurcated into formal and informal spheres due to extreme information asymmetry. This dynamic socio-economic landscape is not due merely to a lack of data, but to a structural by-product of a highly fragmented market composed of 22 official languages, diverse regions, and a dominant, robust informal-sector labor force. Data scarcity has been a concern in the Indian market for the past several decades. As a significant barrier, problems such as a lack of credible measures and morally challenged customers have continuously hampered the growth of the Indian economy. India can be termed as a perennial market for lemons, where the high cost of verifying quality and mere intent led to systemic exclusion and the persistence of the credit gap.

Historically, India's economy has been characterised by a trust deficit in which the public is generally trapped in an Akerlofian lemons market. The invisibility of reliable intentions and high-quality

sellers is evident in the agriculture and labor markets. In the agricultural sector, farmers face adverse selection, in which honest farmers are driven out of the market by dishonest, low-quality producers. In India, traditional credit markets involve middlemen or local moneylenders who charge 30-40% interest rates based solely on intent and long-standing relationships. It is almost impossible for the large rural population in India to escape the cycle of costly signaling to obtain a formal loan or a university degree in the job market. AI has the power to address all these obstacles by developing computational vision, alternative data signals, and solutions to moral hazard problems via consent-linked access to a borrower's historical behavioral patterns.

14.6 The Macro Implications For “Viksit Bharat”

The synthesis of AI and Indian Digital Public Infrastructure (DPI) is acting as a corrective force through a symmetrical revolution, driven by verification and alternative data signals. Platforms like cars24, DeHaat, and ICICI Lombard, through their computational capabilities, have demonstrated the ability to improve economic welfare in India. It ensures that India's economic growth is not concentrated solely in urban hubs of trust but is also distributed across the informal sector. By addressing the problems identified by Akerlof, Spence, and Stiglitz, AI is transforming India into a merit-based market. The Impact of AI in India is most evident in the resolution of adverse selection. Other AI-driven platforms like Bhashini have helped expand access to national markets by removing the obstacle of linguistic asymmetry. As India continues to achieve Viksit Bharat goals, AI is not just about technological advancement; it has now become the infrastructure of trust. It has helped transform intent into information by converting socio-economic barriers into data-driven bridges and ensuring that the Indian market finally achieves the symmetry required for truly inclusive growth.

The AI summit 2026 findings released in the State of AI in India 2026 Report highlight a fundamental shift in the Indian economic fabric: the move from mere gut feeling to Predictive Certainty.

a. The quandary of Mandi asymmetry

One of the most noticeable findings of the 2026 Summit was the documented 28% reduction in post-harvest losses across Punjab,

Haryana, and Andhra Pradesh. This happened due to the integration between Generative AI and the Open Network for Digital Commerce (ONDC). AI models are processing unstructured data regarding weather and sudden market pattern shifts to provide farmers with a 72-hour window of certainty. This has effectively dismantled intermediaries' monopoly over information, ensuring that value is captured at the farm gate rather than on the trading floor.

b. The "Formalisation without Friction" Milestone

A core theme of the summit covered sachetization of credit. The findings outlined that over 45 million MSMEs, previously considered lemons, have been brought under the formal credit umbrella through flow-based lending. The 2026 report specifically highlights that default rates are 40% lower than in traditional collateral-backed lending, proving that high-frequency data is a reliable predictor of intent.

c. Bhashini's Economic Multiplier

The Summit also introduced the concept of linguistic asymmetry, which means non-English speakers must pay a hidden cost to intermediaries to survive in a digital world. The 2026 findings showed that the Bhashini AI layer has added an estimated \$120 billion to the rural GDP by enabling voice-first commerce. The Summit concluded that AI has successfully decoupled digital capability from English proficiency, solving long-standing socio-economic asymmetry.

Finally, the 2026 findings emphasized the success of the India AI Compute Capacity, which prevents the formation of a new asymmetry in which global tech giants own the prediction models of Indian consumers.

The summit concluded with a focus on the uniqueness of India's solution, which includes a design to address market problems.

14.7 Conclusion

This chapter has illustrated that the asymmetry that has existed throughout history is not an enigma to be solved but a technological gap that can be bridged through modern technology and AI. AI has effectively dismantled the barriers seen in the lemon market through a sophisticated layer of digital trust and methodology. Empirical evidence from previous studies shows that AI is no longer

a hypothetical phenomenon but a practical solution that allows the informal sector to signal its value with unprecedented clarity. These are not merely efficiency gains; they are the foundational blocks of a "Market Symmetry" that empowers the bottom of the pyramid.

In the Indian context alone, Viksit Bharat is set to branch out in this field via the India-AI Impact Summit 2026, focusing on the development of seven chakras, including human capital, economic development for social goods, and resilience, innovation, and efficiency through AI. The Indian models, as discussed above, via case studies, prove that when technology is grounded in economic reality and social responsiveness, it does more than optimize markets. As AI begins to initiate transactions and manage autonomous supply chains, the challenge for policymakers around the world will be to maintain the human-centric guardrails that ensure this symmetry leads to equitable growth rather than new forms of algorithmic exclusions. In summary, AI has served as the definitive corrective force against dynamic socio-economic distortions in the market. By converting hidden information into public signals, the world has pioneered a global model for using AI not just for profit, but for the democratization of opportunity.

References

- [1] Agrawal, A., Gans, J., and Goldfarb, A. (2019). Prediction, judgment, and complexity: a theory of decision-making and artificial intelligence. *The economics of artificial intelligence: An agenda*, (pp. 89-110). University of Chicago Press.
- [2] Akavaram, S. (2025). Responsible AI in Financial Decision Systems: Fairness and explainability in credit scoring models. *Journal of International Crisis and Risk Communication Research*, 571–580. <https://doi.org/10.63278/jicrcr.vi.3624>
- [3] Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Uncertainty in Economics*, 235-251. <https://doi.org/10.1016/B978-0-12-214850-7.50022-X>
- [4] Arrow, K. J. (1978). Uncertainty and the welfare economics of medical care. *Uncertainty in Economics*, 345-375. <https://doi.org/10.1016/B978-0-12-214850-7.50028-0>
- [5] Bhashini Division, MeitY. (2025). *Breaking the language barrier: AI-driven multilingual commerce in rural India*. Ministry of Electronics and Information Technology, Government of India.
- [6] Choudhari, A. A., Meshram, K. A., Das, N. J., Mahankal, M. D., Ghadage, A. H., and Shah, A. H. (2026). AI-Driven Financial Management for MSMEs:

- Transforming Decision-Making and Access to Finance in Developing Economies. *Enterprise Development and Microfinance*, 35(2), 74-90. <https://doi.org/10.3362/edm.v35i2.560>
- [7] Digital India Corporation. (2026). *Democratising AI in India*. Ministry of Electronics and information (MeitY), Government of India. URL: [Democratising AI in India - PIB](#)
- [8] EY India. (2025). Is India ready for Agentic AI? *The Aldea of India: Outlook 2026*. Ernst and Young Global Limited. https://www.ey.com/en_in/insights/ai/agentic-ai-india
- [9] G20 India Presidency. (2024). Framework for Digital Public Infrastructure (DPI) and AI integration in emerging markets. *G20 Secretariat*.
- [10] India AI Impact Summit. (2026). The Delhi Declaration: Scaling responsible AI for the Global South. *Proceedings of the India AI Impact Summit 2026, New Delhi*.
- [11] Ministry of Finance. (2026). Economic Survey of India 2025-26: Digital formalisation and the rise of information symmetry. *Department of Economic Affairs, Government of India*.
- [12] NITI Aayog. (2025). Artificial Intelligence for all: A strategy for inclusive growth in the informal sector. *Government of India*.
- [13] Reserve Bank of India (RBI). (2025). Framework for Responsible and Ethical Enablement of AI (FREE-AI) in the financial sector. *RBI Department of Supervision*.
- [14] Spence, M. (1978). Job market signalling. *Uncertainty in Economics*, 281-306. <https://doi.org/10.1016/B978-0-12-214850-7.50025-5>
- [15] Stiglitz, J. E. (2002). Information and the change in the paradigm in economics. *American Economic Review*, 92(3), 460-501. ISSN 0002-8282 (Print)
- [16] Stiglitz, J. E., and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3), 393–410. <https://www.jstor.org/stable/1802787>
- [17] Subramanya, S. V., and Anil, N. (2026). AI-based Credit Scoring and Financial Access for MSMEs: Evidence from Secondary Data. *International Journal of Research in Commerce and Management Studies (IJRCMS)*, 8 (2): 12-17 Article No. 664 Sub Id 1139
- [18] UNCTAD. (2025). Digital economy report 2025: Bypassing linguistic and institutional barriers through AI. United Nations Conference on Trade and Development.
- [19] World Economic Forum (WEF). (2025). The rise of the intelligent stack in emerging economies: A 2030 blueprint. World Economic Forum.

How to cite**Any chapter of this book should be cited as:**

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.

Chapter 15

ARTIFICIAL INTELLIGENCE AND BANKING SECTOR: A SYSTEMATIC REVIEW

Rimpi Kaur

Assistant Professor, Department of Commerce, Sri Guru Gobind Singh College, Sector-26, Chandigarh-160019.

Email: rimpikr24@gmail.com

Abstract

Artificial Intelligence (AI) is revolutionizing every field in the world and banking sector is the most affected area amongst all the sectors. As per global AI adoption rate in industry, the banking sector leads with a 68 % adoption rate mainly due to AI-driven risk management, automation and fraud detection. It is promoting business in terms of enhancing the operational efficiency, decision-making, customer relationships, risk management. AI-driven technologies such as machine learning and robotic process automation are redefining banking operations in India because of the growing demand for automation, fraud detection, personalized customer engagement and real-time decision making. This chapter is an attempt to analyze the current literature to find out the areas covered under the various studies conducted in this field. Research chapters published during the last year have been downloaded from the social sciences research network (SSRN) and studied to analyze the status of the research on AI in the banking sector. AI application is observed in various fields like personalized financial services, cyber security, risk management, back-end BPM, customer engagement through chatbots, etc. The chapter also concludes that AI has been broadly functional in areas such as risk management, fraud detection, customer service, algorithmic trading, and personalized marketing. Among these, fraud detection and credit risk management appear as the most extensively

explored areas, mainly due to the accessibility of structured financial data and regulatory pressure on banks to augment risk controls. The chapter contributes to the understanding of AI adoption in the various areas of operation in the banking sector that will help the policymakers and banks to navigate the future of AI in the banking sector.

Keywords: artificial intelligence, AI adoption, AI led banking revolution, Indian banking sector, fraud detection

15.1 Introduction

Earlier the banks were providing all the services in brick-and-mortar system but now days there is a drastic change when almost all the services are provided on just a click of mouse. The need is for fast internet connection and the customers can visit the branch virtually to get all the services from the banks. After the globalization, then IT revolution the digital world has entered in the banking sector that has made all the banking services faster, accessible and available on 24/7 basis. After the covid-19 the banking touched a new phase where online payments increased to the larger extent because of only option to survive. Then the Fintech companies flourished with exponential increase in number of transactions for online payments. Fintech companies have played important role in providing online payment solutions especially in collaboration with the banks. Fintech funding flow has positive impact on stock returns of the banks (Pisaniuc, 2024) which is a positive sign to use fintech services that enhances the online services with exponential growth.

Now the banks have entered into a new phase where AI is assisting the banks to provide customised banking services faster and efficiently that further enhance the operational efficiency and returns of the banks. AI is used in various forms like blockchain technology, chatbots, predictive analytics, gamification, etc. in almost all the areas of the banks. It further helped the banks to improve operational efficiency, better manage the banking services, customer relationship management, risk management, fraud detection and many other fields like loan processing, retail banking, digital payments, etc.

Artificial intelligence has transformed every aspect of business and banking sector is the most sentimental sector leading with a 68

% adoption rate. Because AI is making the banking process faster and efficient like fund transfer, fraud detection, back-end operational management. It is also serving the banks in personalized customer interactions, especially through chatbots. The chatbots play a leading role in customer interactions that further help to know about their desires to provide customized services (Parthiban et al., 2024). AI has affected all the banking operations related to customer service delivery and operational management. AI with predictive analytics helps the banks to know about the needs and preferences of the customers that enhances the loyalty and transparency. Predictive analytics help the banks in market segmentation, targeting, positioning, which further provide the competitive advantage to the banks with customized services, which helps banks guess what customers will need and prefer, boosting loyalty and satisfaction. Blockchain technology is also a better solution to enhance security and transparency, as these two are crucial to create a trust in digital payments (Kshetri et al., 2018).

Banks are spending huge money on AI especially in developed countries to gain the competitive advantage. This is a crucial time to examine the need and impact of AI to explore the opportunities and meet the challenges efficiently. An attempt is made in this direction to examine the areas where AI is used in its most to provide banking services. Through the review of existing literature, the study will highlight the areas where the AI is used in the most and what is the impact of AI on different performance areas of banks.

15.2 Research Methodology

Present systematic review was conducted by using existing literature available about artificial intelligence in banking sector. The current study is a review study of artificial intelligence in the banking industry.

15.2.1 Data Extraction

In this study, a comprehensive search strategy was developed to identify relevant literature on SSRN website which included specific keywords and phrases related to AI and banking, ensuring that a broad gamut of studies was considered. The keywords "Artificial Intelligence in banking" and "AI in Indian Banking Sector" were used



Figure 15.1: No of Articles Component wise

to maximize the effectiveness of the search strategy. The data was extracted from social sciences and research network website.

15.2.2 Time Period

The chapters downloaded for review were published in the last year (2025-2026) to ensure that the review focused on the most current advancements in AI technologies and their applications in banking. The data was extracted on Jan 14, 2026.

15.2.3 Inclusion and Exclusion Criteria

To ensure the relevance and quality of the literature included in the review, specific inclusion and exclusion criteria were established. The studies published during 2025-26, peer-reviewed, relevant to the theme were included in the review. The studies not directly related to AI technologies and banking sector and published before and after the time period of study are excluded. Total 165 chapters were observed on the database for the selected period of the study. Out of these, 96 chapters either were not available in full text form or not downloadable. 21 chapters were irrelevant to the study. So finally, 48 chapters were included in the study.

The extracted data was analyzed properly and the themes emerged from the analysis were AI technologies in bank management, operational efficiency, fraud detection, ethical considerations, risk management, CRM, HRM, digital payments, credit access, etc. The review process adhered to ethical research standards, as all sources

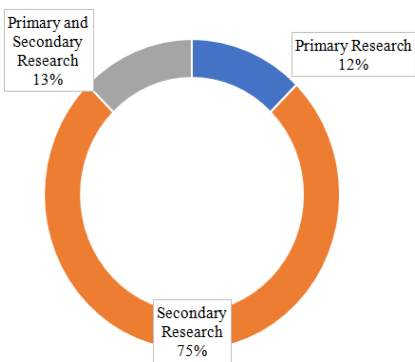


Figure 15.2: Categories of Research on the Basis of Data Used

were appropriately cited, and the review did not involve any original data collection from human subjects and efforts were made to minimize bias in the selection of studies and data extraction, ensuring that the review represents a balanced perspective on the theme.

15.3 Data Analysis And Results

Total 48 chapters were included in this systematic review that included AI in banking sector. The chapters were segregated according to the various components of the banking operations where the AI tools are applied. From Figure 15.1, it is clear that the most chapters are on general implementation, adoption of AI in banking sector where overall performance is examined. These are 15 chapters while 10 chapters are in the field of risk management which is a crucial area of the banks. Fraud detection is another important area where various tools of AI are used to detect frauds in time. Digital payments followed by operational efficiency are the next preferred areas of banking to apply AI technologies. While other areas like CRM, HRM, Market positioning, retail banking microfinance are not much explored areas.

Figure 15.2 highlights that only 12% chapters are based on primary data 13% included both primary and secondary data while 75% chapters are based on secondary data only that revealed the future area of research in this field because primary data helps to know about the qualitative factors like behavior, likes, etc. about the customers, employee problems and prospects and many more.

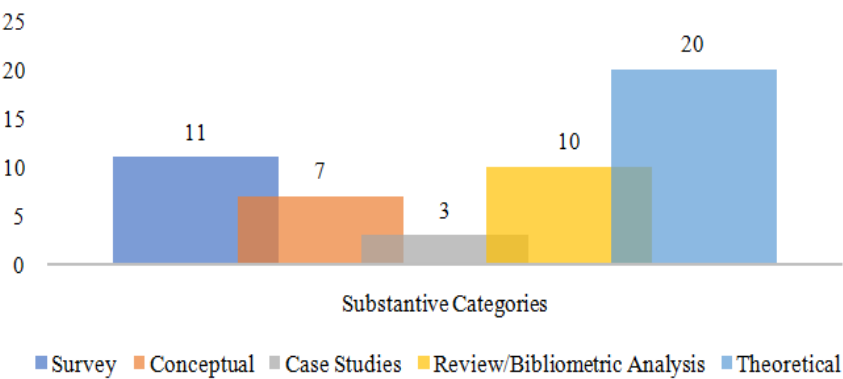


Figure 15.3: Substantive Categories of Research Articles

Figure 15.3 reveals the substantive categories of research article where the most chapters are theoretical that explore the application of AI technologies in banking areas to further understand the concepts in detail. 11 and 10 studies are survey based and review chapters respectively. There are only 25% chapters based on primary data and among those most are survey based that explore the different aspects of the banking with AI while all others are either theoretical, conceptual or case studies. Case studies are very less in number i.e. 3 articles.

15.4 Discussions

15.4.1 AI and Bank Management (General):

There are 15 chapters on application and adoption of AI technologies in banking sector and revealed that AI has positive impact on efficiency, financial inclusion, green finance, usability and accessibility. Challenges and opportunities are also discussed in many chapters. This part includes mixture of theoretical, conceptual and review chapters only that contributes to the existing literature on AI in banking areas with positive impact on performance of the banks.

The systematic review is presented below in table 15.1.

Table 15.1: Literature Review Summary

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source (Publications followed)
AI and Bank Management (General)				
Bholane (2023)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed adoption of AI in Indian Banking, covering the application, challenges and suggestions	11
Jun and Paulson (2025)	SR- Theoretical	Data collection, analysis and interpretation	Discussed application of AI and data analytics in bank transformation in terms of efficiency and innovations	36
Walker et al., (2025)	SR-Review	Data Collection, Analysis and Interpretation	Study presented positive outlook of AI with improved efficiency, customer experience and better risk management along with challenges and opportunities	46
Periasamy et al., (2024)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed the positive impact of AI on Indian banking system in terms of customer experiences, security, efficiency, etc.	16
Odi et al., (2025)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed positive impact of AI, Recommendations and Regulatory and Legal Accountability in Nigeria	22

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Salvi (2025)	SR-Conceptual	Data Collection, Analysis and Interpretation	Discussed the application, Impact, challenges, and comparative adoption of AI in Public and Private banks that can promote financial inclusion, enhance competitiveness, and foster a more resilient and customer-centric banking ecosystem	16
Jameaba (2025)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed the implications of the adoption of BCT and AI on operations and performance of banking in Indonesia	98
Prarthana (2025)	SR-Conceptual	Data Collection, Analysis and Interpretation	Concluded significant impact of AI on banks performance	38
Bakkar (2025)	SR-Review	Data Collection, Analysis and Interpretation	Discussed the role of AI in banking sector through bibliometric analysis	43
Sadaram et al., (2023)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed that AI, ML and Blockchain technologies help to improve value delivery system though having security and other ethical concerns	52
Odi et al., (2025)	SR-Theoretical	Data collection, analysis and interpretation	Discussed AI impact, legal and ethical concerns, regulatory framework and recommendations in Nigerian banking	22
Panigrahi et al., (2025)	SR- Theoretical	Data collection, analysis and interpretation	Discussed role of AI in banking, ethical concerns and its future	5

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Wei le (2025)	SR-Review	Data collection, analysis and interpretation	Discussed AI applications in digital banking platforms that has significantly enhancing accessibility, usability, and engagement	12
Periasamy et al., (2025)	SR- Theoretical	Data collection, analysis and interpretation	Discussed innovative practices along with challenges and opportunities in green finance, digital banking and AI	8
Yilmaz (2025)	SR-Theoretical	Data collection, analysis and interpretation	Revealed positive impact of AI that enables innovations like biometric authentication, AI-powered decision-making and blockchain-based financial services	14
AI and Fraud Detection				
Ahmadi (2022)	SR-Conceptual	Data Collection, Analysis and Interpretation	Discussed application of XAI models in fraud detection, key challenges and future trends	33
Chitraju (2024)	SR-Review	Data Collection, Analysis and Interpretation	Study proposed an AI-driven detection and response framework that offers faster detection, accurate identification and quicker recovery	16
Chenoori (2025)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed positive impact of AI in terms of enhanced fraud detection, customers services and risk management	30
Kothapalli et al., (2022)	SR-Conceptual	Data Collection, Analysis and Interpretation	Discussed trends and behavior in credit cards fraud attacks and transactions	37

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Autade (2021)	PR/SR-Case studies	Data collection, analysis and interpretation	Discussed the potential of the blockchain to transform secure, efficient, and transparent banking transactions	40
Shadreck (2024)	SR-Conceptual	Data collection, analysis and interpretation	Introduced Secure Bank AI, a Capsule Network-Based fraud detection Model that provide higher accuracy and reduces false positives	29
AI and Risk Management				
Chougala (2025)	SR-Conceptual	Data collection, analysis and interpretation	Developed a conceptual model to use AI as intermediary between users and core banking system, operating through perception, action and learning	5
Gates (2025)	PR/SR-Survey/Case Studies/Review	Interview, Analysis and Interpretation	Concluded that automating back-office functions can enhance efficiency and better management of risks	48
Robert and Will (2024)	SR- Theoretical	Data Collection, Analysis and Interpretation	Discussed the positive role of AI in risk management through case studies, challenges and future scope	36
Paleti (2023)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed various aspects of AI implications in risk management along with some examples from case studies	56

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Goswami (2025)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed EU AI Act provisions, GDPR Article 22 and EBA model risk management guidelines through various case studies	77
Hendrix et al., (2025)	SR-Review	Data Collection, Analysis and Interpretation	Discussed best practices in AI implementation where banks prioritizing transparency and accountability through XAI	46
Gates and Will (2025)	SR-Review	Data Collection, Analysis and Interpretation	Discussed the integration of AI into credit risk assessment frameworks and their impact on predictive performance and decision-making processes	34
Hassan (2025)	PR-Empirical	Data collected through Questionnaire, analysis and interpretation	Concluded that integration of AI, ML, NLP and augmented analytics is enhancing financial forecasting, risk management, and investment strategies in terms of improved ROI and decision-making efficiency and risk score reductions though several barriers, like data integration challenges, skill gaps, and high implementation costs are reported	80
Soumya et al., (2025)	PR-Survey	Questionnaire based data collection, analysis and interpretation	Concluded significant but weak impact of AI adoption on credit risk accuracy but negligible impact on reducing loan approval times indicating that traditional factors and constraints, still play a crucial role	20

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
AI and Fiduciary Duties of Banking Board				
Tambun et al., (2025)	SR-Review	Data Collection, Analysis and Interpretation	Discussed the role of AI in discharge of fiduciary duties and constructs a five-dimensional fiduciary oversight framework for AI Governance in Indonesian banks	57
AI and Digital payments/Fintech				
Azuikpe et al., (2024)	PR-Survey	Questionnaire, Analysis and Interpretation	Concluded the requirement of AI to meet the demands with enhanced accuracy, improved efficiency and adaption to changes, also discussed challenges and future prospects	24
Kumar and Bung (2024)	PR-Survey	Questionnaire (Survey and internet poll), Analysis and Interpretation	Examined the impact of AI in terms of Chatbots, automation, fraud detection, decision making and digital payment system with positive outcomes	40
Wu et al., (2025)	PR-Survey	Online survey, analysis and interpretation	Concluded positive impact of Chatbots' service quality on satisfaction but not significant impact on trust means users are more concerned about functional value and use chatbots mainly for functional needs	146

Authors/Ref- erence	Research Type	Research Methodology	Key Findings/Results	Information Source
Pisaniuc (2024)	SR-Case Study	Data collection, analy- sis and interpretation	From monthly stock return data of 56 banks, matching it with monthly Fintech funding data, the study concluded a significant positive relationship between the growth in Fintech funding and the stock returns of EU and the UK incumbent retail banks	32
AI and Operational Efficiency				
Racordon (2025)	SR-Conceptual	Data Collection, Analy- sis and Interpretation	The study introduced the GAB/BAG AI integrated models to enhance efficiency, governance and compliance in Swiss private banking	21
Shaikh et al., (2025)	SR-Review	Data collection, analy- sis and interpretation	Discussed the impact of new technologies on operational efficiency and employment	18
Orjinta and John (2025)	SR-Theoretical	Data collection, analy- sis and interpretation	Concluded that AI strategies enhance sus- tainable performance of deposit money banks in Nigeria	35
AI and CRM				
Tharian (2025)	SR-Review	Data Collection, analy- sis and interpretation	Bibliometric analysis of 209 chapters con- cluded that integration of AI and e-CRM have significantly transformed customer engagement, satisfaction and loyalty	50

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Rani et al., (2025)	PR-Empirical	Questionnaire based data collection, analysis and interpretation	Based on occupation, the study revealed significant difference in customer's adoption factors of personal innovativeness not in other factors and behavioural intention to use AIFA also differentiates among different groups	25
AI and HRM				
Akerele and Ajiyon (2025)	SR-Review	Data collection, analysis and interpretation	Revealed that AI is increasingly adopted in recruitment to automate candidate secret-ing, improve efficiency and reduce bias but not much used in selection and retention process in US banking	44
AI and Ethical considerations				
Pope and Sony (2025)	SR-Theoretical	Data Collection, Analysis and Interpretation	Discussed the ethical considerations in big data through case studies and their consequences if not followed the regulatory framework	18

Authors/Ref- erence	Research Type	Research Methodology	Key Findings/Results	Information Source
Dullo (2025)	PR-Survey	Interviews based data collection, analysis and interpretation	Findings revealed that ethical AI is not fully institutionalizing yet but it is gradually transitioning from being a compliance obligation to a strategic differentiator	11
AI and Strategic forecasting and Market positioning				
Deshmukh and Lilah (2024)	SR-Theoretical	Data collection, analysis and interpretation	Revealed that the strategic integration of AI and predictive analytics enhances operational efficiency of the banks and also redefines their competitive background in terms of market positioning	32
AI and Credit Access/Microfinance				
Chen et al (2026)	SR-Descriptive	Data Collection, Analysis and Interpretation	Concluded that adoption of AI and big data significantly improves credit rating accuracy and loan performance in China	49
Hasanov (2025)	SR- Theoretical	Data collection, analysis and interpretation	Explained that use of AI technologies expanded financial inclusion, reduced operational costs and streamlined lending processes in developing economies	10
AI and Retail Banking				

Authors/Reference	Research Type	Research Methodology	Key Findings/Results	Information Source
Homayssi (2025)	PR-Survey	Interview, analysis and interpretation	Concluded positive impact of AI in employees job roles, efficiency, but negative in terms of job insecurity, ethical concerns	182
Owen et al., (2025)	SR-Theoretical	Data collection, analysis and interpretation	Examined the impact of digital transformation on retail banking and the growing importance of customer-centric approaches	46

15.4.2 AI and Fraud Detection:

6 chapters are found in this area where the researchers revealed that AI technologies especially blockchain are helping in faster detection, accurate identification and quicker recovery of frauds in banking sector. All chapters except one are based on secondary sources and are theoretical, conceptual and review chapters. Shadreck (2024) has introduced a Capsule Network based fraud detection model that provides higher accuracy and reduces false positives.

15.4.3 AI and Risk Management:

The review observed 9 chapters in this field including 3 chapters based on primary data and others are secondary data based. Maximum studies are theoretical while few are conceptual, survey based and review chapters. The studies explored the positive role of AI in risk management that enhance efficiency and better management of risks. Some studies discussed challenges and future scope of AI in this area while some suggested models like ML, NLP, augmented analytics help to improve efficiency in terms of ROI by managing the risks involved in banking area (Hassan, 2025).

15.4.4 AI and Fiduciary Duties of Banking Board:

There is only one study based on secondary data. Tambun et al., (2025) reviewed the role of AI in discharge of fiduciary duties and construct a five-dimensional fiduciary oversight framework for AI Governance in Indonesian banks. AI can support in performing various fiduciary duties like care, loyalty through algorithmic domains but by within the dimensions of regulatory framework.

15.4.5 AI and Digital Payments:

This part includes 4 studies 3 are based on primary data includes survey through questionnaire or internet poll. The studies highlighted the requirement of AI technologies to meet the demands with accuracy and efficiency. Wu et al., revealed positive impact of Chatbots' service quality on satisfaction but not significant impact on trust means users are more concerned about functional value and use chatbots mainly for functional needs.

15.4.6 AI and Operational Efficiency:

All 3 studies included in this area are based on secondary data and are theoretical, conceptual and review studies only. The studies discussed positive impact of AI technologies on operational efficiency. Racordon

(2025) suggested a model GAB/BAG to enhance efficiency, governance and compliance in Swiss private banking.

15.4.7 AI and Customer Relationship Management (CRM)

: There are 2 studies in this area one is secondary data based and another is primary data based. Tharian and Biju (2025) analyzed 209 chapters in bibliometric analysis and concluded that integration of AI and e-CRM has significantly transformed customer engagement, satisfaction and loyalty. Rani et al., (2025) in their primary research revealed a significant difference in customer's adoption factors of personal innovativeness.

15.4.8 AI and Human Resource Management (HRM):

Only 1 chapter by Akerele and Ajiyon (2025) was found in this area which is a secondary data-based review chapter. It is revealed that AI is increasingly adopted in recruitment to automate candidate secreting, improve efficiency and reduce bias but not much used in selection and retention process in US banking. Hence, AI can enhance the efficiency of HRM departments to get the benefits of utilizing the saved time for other productive purposes.

15.4.9 AI and Ethical considerations:

There are 2 chapters included in this area one is primary data based and other one is secondary data-based theoretical chapter. Though ethical consideration is taken care of but Dullo (2025) revealed that ethical AI is not fully institutionalizing yet but it is gradually transitioning from being a compliance obligation to a strategic differentiator.

15.4.10 AI and Strategic forecasting and Market positioning:

There is only 1 chapter by Deshmukh and Lilah (2024) based on secondary data. The study revealed that the strategic integration of AI and predictive analytics enhances operational efficiency of the banks and also redefines their competitive background in terms of market positioning. AI helps to identify the customer preferences and their behavior that help the banks to make strategies to provide tailored customer services than enhances the loyalty and security among the customers.

15.4.11 AI and Credit Access/Microfinance:

There are 2 studies based on secondary data and explored that AI improved the credit rating accuracy and loan performance. Hasanov (2025) concluded that use of AI technologies expanded financial inclusion, reduced operational costs and streamlined lending processes in developing economies.

15.4.12 AI and Retail Banking:

This area is explored by two studies one through primary source and second through secondary sources. Homaissi (2025) revealed positive impact of AI in employees job roles, efficiency, but negative in terms of job insecurity, ethical concerns. Owen, Xavier et al., (2025) examined the positive impact of digital transformation on retail banking and the growing importance of customer-centric approaches.

The overall findings from all the studies can be drawn in positive sense means most of the studies revealed positive impact of AI technologies on operational efficiency, risk management, fraud detection, CRM, HRM, credit access, market positioning, and many more. But few studies highlighted that AI is gradually transitioning from being a compliance obligation to a strategic differentiator but not fully committing yet. Ethical consideration, regulatory frameworks are yet to be institutionalized areas that hit the trust of the customers.

15.5 Conclusion

AI is hitting all the areas of every industry where the financial sector is the most leading area in use of AI. Banking sector is the crucial sector where AI technologies like blockchain, Machine learning, big data, chatbots, predictive analytics, etc. are providing a striking platform to improve upon operational efficiency, risk management, fraud detection and enhancing customer satisfaction by providing customized services. Though AI is entering in other areas too like HRM, credit accessibility and loan processing, microfinance, green finance, regulatory areas and Fintech have also played an important role in enhancing efficiency in digital payments. But these areas are not much explored by the researchers may be due to difficulty in collection of related data and implementation of AI technologies in qualitative aspects. This is the right time to best utilize AI technologies to get maximum output but without taking care of ethical considerations, the goal can't be achieved. The banking organization should take help from AI technologies not as an alternate but as an assistant so that the customers and employees satisfaction level can be achieved.

References

- [1] Ahmadi, S. (2022). Advancing Fraud Detection in Banking: Real-Time Applications of Explainable AI (XAI). *Journal of Electrical Systems*, 18 (4), 141-150. <https://dx.doi.org/10.2139/ssrn.5094556>
- [2] Akerele, O., and Ajiyon, S. T. (2025). Effect of Artificial Intelligence on Employee's Recruitment, Selection and Retention in the US Banking Sector: A Systematic Review. *Effect of Artificial Intelligence on*

- Employee's Recruitment, Selection and Retention in the US banking Sector: A Systematic Review*, 1, <http://doi.org/10.17632/bk79gkn2yb>
- [3] Anetoh, J. (2025). Achieving Sustainable Performance of Deposit Money Banks in Nigeria through Artificial Intelligence Strategies. <https://dx.doi.org/10.2139/ssrn.5204490>
 - [4] Autade, R. (2021). Financial Security and Transparency with Blockchain Solutions. *Turkish Online Journal of Qualitative Inquiry*. <http://dx.doi.org/10.53555/w60q8320>
 - [5] Bakkar, S. A. (2025). How Do ChatGPT/Chatbot Transform the Banking Industry? Prospects and Challenges. <http://dx.doi.org/10.2139/ssrn.5112874>
 - [6] Bholane, K. P. (2023). Adoption, Applications and Challenges of Artificial Intelligence in Indian Banking Sector. *SSRN Electronic Journal*, 17(2), 146-151. <http://dx.doi.org/10.2139/ssrn.5194225>
 - [7] Chenoori, S. (2025). AI-Driven Transformation in Banking: From Fraud Detection to Personalized Services. *Available at SSRN 5403433*. <http://dx.doi.org/10.2139/ssrn.5403433>
 - [8] Chitraju, S. (2024). AI-Driven Detection and Response Strategies for Cloud-Based Ransomware in Banking Systems. *Available at SSRN 5383844*. <http://dx.doi.org/10.2139/ssrn.5383844>
 - [9] Chougala, M. K. (2025). The Emergence of Agentic Artificial Intelligence: A Paradigm Shift for Operational Efficiency and Risk Management in Digital Banking. *Available at SSRN 5784424*. <http://dx.doi.org/10.2139/ssrn.5784424>
 - [10] Deshmukh, A., and Boyd, L. (2024). Predictive Analytics in Banking: Leveraging AI for Strategic Forecasting and Market Positioning. *Available at SSRN 5393028*. <http://dx.doi.org/10.2139/ssrn.5393028>
 - [11] El Homayssi, B. (2025). AI Adoption in Retail Banking: Employees Talent Innovation. <http://dx.doi.org/10.2139/ssrn.5372797>
 - [12] Fabuyi, J., and Balogun, A. Y. (2024). The Necessity of Artificial Intelligence in Fintech for SupTech and RegTech Supervisory in Banks and Financial Organizations. *International Journal of Science and Research Archive*, 12(2). <http://dx.doi.org/10.30574/ijrsra.2024.12.2.1614>
 - [13] Gates, N., and Ellison, G. (2025). Artificial Intelligence for Operational Risk Management: Automating Back-Office Banking Functions. *Available at SSRN 5441714*. <http://dx.doi.org/10.2139/ssrn.5441714>
 - [14] Gates, N., and Flynn, W. (2025). A Comprehensive Review of Artificial Intelligence-Driven Credit Risk Modeling: From Traditional Scoring to Explainable Machine Learning in Modern Banking Systems. <http://dx.doi.org/10.2139/ssrn.5414214>

- [15] Goswami, A. (2025). AI Governance for Risk Modelling in European Banks: A Compliance-First Approach. <http://dx.doi.org/10.2139/ssrn.5360444>
- [16] Hasanov, E. (2025). The Role of Fintech in Enhancing Microfinance Accessibility and Efficiency in Developing Economies. <http://dx.doi.org/10.2139/ssrn.5198027>
- [17] Hassan, T. (2025). Augmented Analytics in Financial Forecasting Enhancing Risk Management and Investment Strategies through AI-Driven Insights. <http://dx.doi.org/10.2139/ssrn.5179431>
- [18] Hendrix, D., Holt, S., and Howe, N. (2025). The Role of Explainable AI in Improving Transparency and Accountability in Banking Decisions. <http://dx.doi.org/10.2139/ssrn.5440794>
- [19] Jameaba, M. (2025). Blockchain Technology, Artificial Intelligence and Financial Stability: Opportunities and Challenges for Financial Institutions. <http://dx.doi.org/10.2139/ssrn.5973614>
- [20] Jun, C., and Paulson, N. (2025). Artificial Intelligence and Data-Driven Banking Transformation: A Review of Efficiency and Innovations <http://dx.doi.org/10.2139/ssrn.5413924>
- [21] Kaul, D., Aditya. (2025). How can Indian Financial Institutions Develop Ethical AI Frameworks that Drive Innovation while Ensuring Compliance with Financial Services Regulations? <http://dx.doi.org/10.2139/ssrn.5349181>
- [22] Kothapalli Sondinti, L. R., and Yasmeen, Z. (2022). Analyzing Behavioral Trends in Credit Card Fraud Patterns: Leveraging Federated Learning and Privacy Preserving Artificial Intelligence Frameworks. *Universal Journal of Business and Management*, 2(1), 1224. Retrieved from <https://www.scipublications.com/journal/index.php/ujbm/article/view/1224>
- [23] Kshetri, Nir B., NC DOCKS at the University of North Carolina at Greensboro (2018). The Indian Blockchain Landscape: Regulations and Policy Measures. Retrieved From: https://libres.uncg.edu/ir/uncg/f/N_Kshetri_Indian_2018.pdf
- [24] Kumar, P., and Bung, P. (2024). AI Implementation in Digital Payment: An Empirical Analysis on Banking Sector in India. <http://dx.doi.org/10.2139/ssrn.5223749>
- [25] Maia, P. (2024). The Impact of Fintech on Bank Management and Valuation: UK Case Study. <http://dx.doi.org/10.2139/ssrn.5110004>
- [26] Odion, A. O., Farinola, T., Okeke, V., Sunmboye, W., Echefula, W., and Oyewolu, O. (2025). Navigating the Double-Edged Sword: The Impact and Challenges of Artificial Intelligence Integration in Modern Banking. <http://dx.doi.org/10.2139/ssrn.5376175>

- [27] Owen, O., Walter, Z., and Huff, C. (2025). Transforming Customer Experience in Retail Banking Through AI-Powered Personalization. <http://dx.doi.org/10.2139/ssrn.5440636>
- [28] Paleti, S. (2023). AI-Driven Innovations in Banking: Enhancing Risk Compliance through Advanced Data Engineering. <http://dx.doi.org/10.2139/ssrn.5244840>
- [29] Pandey, P. (2025). Sustainable and Innovative Practices in the Indian BFSI Sector. <http://dx.doi.org/10.2139/ssrn.5328215>
- [30] Panigrahi, A. (2025). Reengineering Banking Performance through Convergence of AI Technologies. <http://dx.doi.org/10.2139/ssrn.5200084>
- [31] Parthiban, E. S., and Adil, M. (2024). Seeing Through the Mystique Surrounding the Task-Technology-Organization Fit: A Mixed-Method Exploration of Consumers' Intention to Banking Conversational Agents. *Communications of the Association for Information Systems*, 55, 395-433. <https://doi.org/10.17705/1CAIS.05516>
- [32] Periasamy, P., Dinesh, N., and Padmanabhan, S. (2024). Impact of AI in Indian BFSI Sector. *Optimization*, 16(1): 13-17.
- [33] Pope, B., and Vincent, S. (2025). Big Data Ethics in Banking: Balancing Strategic Insights with Privacy, Transparency, and Responsible AI. <http://dx.doi.org/10.2139/ssrn.5395295>
- [34] Prarthana, M. S. (2025). A Study on the Boon and Bane of Artificial Intelligence in the Banking and Financial Sector Industry. <http://dx.doi.org/10.2139/ssrn.5636990>
- [35] Racordon, H. (2025). An Integrated AI-Blockchain Governance Framework for Operational Resilience in Private Banking (Working Chapter Version 1.3). Zenodo / SSRN. <https://orcid.org/0009-0008-9382-3191>
- [36] Rani, N., Sethi, T., and Gupta, P. (2025). Impact of Occupation of Banking Customers on Artificial Intelligence-driven Financial Assistant Adoption Factors and Behavioural Intention to use AIFA. <http://dx.doi.org/10.2139/ssrn.5426734>
- [37] Robert, E. and McGee, W. (2024). Artificial Intelligence for Strategic Risk Management in Banking: Current Practices and Research Directions. <http://dx.doi.org/10.2139/ssrn.5395050>
- [38] Sadaram, G., Sakuru, M., Jha, K. M., Bodepudi, V., Katnapally, N., Maka, S. R., and Karaka, L. M. (2023). Understanding the Fundamentals of Digital Transformation in Financial Services: Drivers and Strategic Insights. *Journal of Artificial Intelligence and Big Data*, 3(1), 72-83. <https://www.scipublications.com/journal/index.php/jaibd/article/view/12>

- [39] Salvi, P. (2025). Analytical Study of Artificial Intelligence Integration in Indian Banks and Its Implications for the Future of Banking and Financial Services in India. <http://dx.doi.org/10.2139/ssrn.5705943>
- [40] Shadreck, M. (2024). Secure Bank AI: A Capsule Network-Based Model for Fraud Detection in Digital Banking. *World Journal of Advance Pharmaceutical Sciences*, 1(1), 31-38.
- [41] Shaikh, A., Khatri, M., and Lamsoge, S. (2025). The Impact of Digitalisation on Operational Efficiency and Employment in the Indian Banking Sector. <http://dx.doi.org/10.2139/ssrn.5337272>
- [42] Soumya, K. G., Suriya, S. V., Vinoth, S., and Gopalakrishnan, C. (2025). Enhancing Lending Decisions and Risk Management in Banking Through Artificial Intelligence. <http://dx.doi.org/10.2139/ssrn.5216138>
- [43] Tambun, T., Yudoko, G., and Aldianto, L. (2025). Strategy and Innovation in AI Oversight: Fiduciary Duties of Banking Boards. *Gatot and Aldianto, Leo, Strategy and Innovation in AI Oversight: Fiduciary Duties of Banking Boards*. <http://dx.doi.org/10.2139/ssrn.5403675>
- [44] Tharian, K. (2025). The Role of CRM and e-CRM in Banking Sector: A bibliometric Review on Customer Loyalty and Performance. <http://dx.doi.org/10.2139/ssrn.5662810>
- [45] Walker, J., Williams, T., and Douglas, F. (2025). Artificial Intelligence Applications in Strategic Bank Management: A Systematic Review. <http://dx.doi.org/10.2139/ssrn.5442535>
- [46] Wei, L. (2025). A Review of AI Applications in Digital Banking Platforms: Enhancing Accessibility, Usability, and Engagement. <http://dx.doi.org/10.2139/ssrn.5406826>
- [47] Wu, J., Hou, J., and Xu, X. (2025), AI Anxiety as a Positive Moderator in Banking Chatbot Continuance Usage Intention: Pls-Sem and Ann Analysis. <http://dx.doi.org/10.2139/ssrn.5244532>
- [48] Yilmaz, Y. (2025). Technology Fusion in the Banking Sector: A Patent Network Analysis Approach. <http://dx.doi.org/10.2139/ssrn.5431466>

NOTES

[illegible]



A blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. There are 21 lines in total, starting from the top and ending near the bottom. The lines are thin and black.

How to cite

Any chapter of this book should be cited as:

Author, A. A. (2026). Title of the cited chapter. In Dr. Aanchal Arora, et al. (Ed.), Bridging Disciplines with Artificial Intelligence. (pp. 1xx–1xx). Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

All chapters of this book published under the CC-BY license at <https://press.vyomhansjournals.com> under Vyom Hans Publications. ISBN (E): 978-81-686211-1-4; ISBN (P): 978-81-686211-0-7

Book is available online at <https://press.vyomhansjournals.com>; Amazon and Flipkart marketplaces.