

**DEVELOPING A REVENUE MODEL FOR
BIOINFORMATICS R&D IN INDIA
USING
MANAGEMENT APPROACH**

SHASHANK SHEKHAR



SCHOOL OF INTERDISCIPLINARY RESEARCH (SIRE)

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2024

©Indian Institute of Technology Delhi (IITD), New Delhi, 2024

**DEVELOPING A REVENUE MODEL FOR
BIOINFORMATICS R&D IN INDIA USING
MANAGEMENT APPROACH**

by

SHASHANK SHEKHAR

School of Interdisciplinary Research (SIRe)

submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy
to the



Indian Institute of Technology Delhi

July 2024

Certificate

This is to certify that the thesis entitled “Developing a revenue model for bioinformatics R&D in India using management approach,” being submitted by **Mr. Shashank Shekhar** to the Indian Institute of Technology, Delhi, for the award of the degree of **Doctor of Philosophy in School of Interdisciplinary Research (SIRe)** is a record of bonafide research work carried out by him. Shashank Shekhar has worked under our guidance and supervision and has fulfilled the requirements for submitting this thesis, which has reached the requisite standard to our knowledge.

The results of this dissertation have not been submitted in part or full to any other University or Institute for the award of any degree or diploma.

Prof. Sanjay Dhir
Department of Management Studies
Science Indian Institute of Technology Delhi
India

Prof. James Gomes
Kusuma School of Biological
Indian Institute of Technology Delhi
India

Prof. B. Jayaram
Department of Chemistry
Indian Institute of Technology Delhi
India

Dated:

Acknowledgments

My Ph.D. research has been an incredible journey in which I have encountered and experienced a range of emotions at various points in time. I wholeheartedly thank my supervisors, Prof. Sanjay Dhir, Prof. James Gomes, and Prof. B. Jayaram, for guiding, inspiring, motivating, and supporting me all through. It has been an honor to work under their mentorship. I am grateful to them for their invaluable advice, unwavering support, and encouragement throughout this exciting passage of research, which enabled me to advance as a researcher. The luxury of receiving ideas and inputs from experts from three different areas has been a unique value proposition for me and my research. I am deeply indebted to my supervisors for showing confidence in my abilities and helping me complete the research part of my Ph.D., followed by thesis writing as a part of the formalities for the award of my Ph.D. degree.

I am grateful to students' research committee (SRC) members Prof. Nalin Pant, Prof. Vivekanandan Perumal, and Prof. Arpan Kar; apart from other faculty and staff members of the Department of Chemistry, School of Biological Sciences, Department of Management Studies, and School of Interdisciplinary Research (SIRe), Indian Institute of Technology Delhi, for their help and support at different stages of my Ph.D. program. I thank the Department of Biotechnology, Department of Electronics & Information Technology (DeitY), Govt. of India, Centre for Development of Advanced Computing (CDAC), Pune, for their financial support during this research.

I thank all my past and current lab members of SCFBio (my workplace) and the School of Interdisciplinary Research (where I have enrolled in this Ph.D. program). I thank my fellow lab mates, especially Akshata, Smriti, Dinesh, and Devendra, for their technical and scientific inputs. Also, A. Mohan Rao and Puneeta were always there to provide administrative support whenever required.

I am also thankful to various groups and individuals, such as the respondents to my survey, the interviewees from academia and industry who shared their views in developing insights on my research topics, and program officers from various funding agencies in India who helped me in understand the financial support mechanism for bioinformatics research and all those who helped and contributed directly or indirectly during the commencement, continuation, and completion of beautiful journey of my Ph.D. and because of whom I have come to this stage where I am representing my work through this thesis.

I thank all my family members, especially my younger brothers (Saumitra and Saharsh), my uncle (Chacha), and my in-laws (especially my brother-in-law, Parijat, who kept asking me that when I am going to submit my thesis) for their love, affection, blessings, and unconditional support throughout the expedition of my doctoral research. I am immensely thankful to my wife, Vandana, for her love, affection, compassion, support, and sacrifices. My two lovely daughters, Shreya and Swara (Strawberry), have always been my motivation and source of inspiration as they more often create a learning environment at home, which pushed me to focus on my research alongside my routine office work, which has many facets. Also, being the senior most member of the research center, I have had the responsibilities and administrative duties required for the smooth functioning of the research center, which is among the best in the country.

My parents are in the heavenly abode, and I miss them a lot, but their blessings have been the driving force for me during my Ph.D. journey. Last but not least, I thank the almighty for everything.

Shashank Shekhar

Abstract

In recent years, research in biological sciences has witnessed a significant shift in the research approach while moving from *in vitro* and *in vivo* experimentation to adopting *in silico* methodologies, which fall under the bioinformatics domain. This interdisciplinary area works on ways to understand and analyze biological data, especially when the data sets are extensive and complex in terms of size and characteristics. This prominent inclusion of *in silico* experimentation depends mainly on the development and efficacies of various research methods, computational tools, software, and databases to deal with scientific problems related to various organisms.

The rapid growth in bioinformatics research resulted in advancing scientific methods and developing relevant technologies, culminating in products and services beneficial to society. Considering the significance and relevance of various efforts in the life science area, the bioinformatics and computational biology based approaches attract researchers and professionals belonging to different subject areas, such as computer science, information technology, chemistry, biological sciences, mathematics, and several other related areas to work together through a common platform to create an innovation ecosystem of greater value eventually and to address prevailing and potential agricultural and healthcare issues affecting various stakeholders.

This thesis examines the scenarios and factors that enable academia, where most basic research and development occurs, to embed commercialization aspects to make academic research self-sustainable in the long run. This exploration gives academia new opportunities and enables them to think beyond the traditional approach of developing and disseminating novel knowledge with stakeholders. Most scholars and decision-makers

agree that entrepreneurship is the future and vital to society. In this thesis, bioinformatics research management is examined, and a growth and revenue model is proposed.

The thesis is divided into nine chapters. Chapter 1 discusses the current status of bioinformatics research in the country and how it strengthens the life sciences and healthcare-based research ecosystem with the advancement in information technology-driven applications. Apart from that, management perspectives are introduced and discussed to formalize and streamline the research pipeline. This is followed by adding up the self-sustainability aspects by exploring various revenue generation streams and realizing the vision of having a full-fledged revenue model within the academic boundaries.

Chapter 2 focuses on the systematic literature review related to the thesis topic, considering both the research and management aspects of bioinformatics activities in India and the best practices adopted at the global level. This part of the thesis investigates the empirical studies published in the literature on synchronizing research and management aspects. Based on the findings, an attempt is made to develop a sustainable model in bioinformatics research and development in India.

Chapter 3 has a detailed description of measurement scales, various factors, sub-factors included in this study, and the methodology adopted for analyzing the data collected either through literature review or from the respondents of the questionnaire survey and interviews to conduct the study. This is to help develop a revenue model in sync with the objectives of this thesis.

Chapter 4 discusses how bioinformatics researchers from academia and entrepreneurs in the life science and healthcare industries join hands to develop a conducive environment to complement each other's work. The adoption of management-based methodologies and approaches further bolsters this effort.

Chapter 5 illustrates and discusses the role of developing a knowledge-based ecosystem for bioinformatics R&D by strengthening human resource development programs that help create a pool of trained bioinformatics professionals in the country who can drive and manage this highly specialized sector linked with public health and agricultural productivity.

Chapter 6 highlights the role and significance of a High-Performance Computing (HPC) technology-based revenue model for bioinformatics R&D in academia and how the various factors and their interdependencies play their part in creating a technological ecosystem that paves the way for developing a framework from within the academic boundaries with the help of a management-based approach.

Chapter 7 addresses developing and managing bioinformatics technologies, including user-friendly and interactive software/tools and databases inside academic boundaries where most of these research activities are conducted and which are the genesis of most of the successful commercial ventures when it comes to bioinformatics software suites offered by commercial companies.

Chapter 8 dealt with triangulation and synthesis aspects, which are important research principles that provide authenticity and significance to the academic work. It culminates in multiple viewpoints and data sources to provide comprehensive knowledge. This chapter also includes the implications of the research work undertaken in this thesis.

Chapter 9 discusses the summary, a few perspectives emerging from the thesis work, limitations, and future scope.

सार

हाल के वर्षों में, जैविक विज्ञान में अनुसंधान ने इन विट्रो और विवो प्रयोग से लेकर सिलिको पद्धति को अपनाने के प्रतिमान में बदलाव देखा है, जो जैव सूचना विज्ञान डोमेन के अंतर्गत आता है। यह अंतःविषय क्षेत्र जैविक डेटा को समझने और उसका विश्लेषण करने के तरीकों पर काम करता है, खासकर जब डेटा सेट व्यापक और जटिल होते हैं। यह प्रमुख समावेशन मुख्य रूप से विभिन्न जीवों से संबंधित वैज्ञानिक समस्याओं से निपटने के लिए विभिन्न कम्प्यूटेशनल उपकरणों और सॉफ्टवेयर के विकास और प्रभावकारिता पर निर्भर करता है।

जैव सूचना विज्ञान अनुसंधान में तेजी से वृद्धि के परिणामस्वरूप वैज्ञानिक तरीकों को आगे बढ़ाने और प्रासंगिक प्रौद्योगिकियों को विकसित करने में मदद मिली, जिसका परिणाम समाज के लिए लाभकारी उत्पाद और सेवाएं हैं। जीवन विज्ञान क्षेत्र में विभिन्न प्रयासों के महत्व और प्रासंगिकता को ध्यान में रखते हुए, जैव सूचना विज्ञान और कम्प्यूटेशनल जीव विज्ञान आधारित दृष्टिकोण विभिन्न विषय क्षेत्रों, जैसे कम्प्यूटर विज्ञान और सूचना प्रौद्योगिकी, रसायन विज्ञान, जैविक विज्ञान, गणित और अन्य संबंधित क्षेत्रों से संबंधित शोधकर्ताओं और पेशेवरों को आकर्षित करते हैं। प्रचलित स्वास्थ्य देखभाल मुद्दों से जुड़ी जैविक अनुसंधान समस्याओं के समाधान के लिए अंततः अधिक मूल्य का एक नवाचार पारिस्थितिकी तंत्र बनाने के लिए एक आम मंच के माध्यम से हाथ मिलाना और एक साथ काम करना।

यह थीसिस उन परिदृश्यों और कारकों की जांच करने और समझने की कोशिश करती है जो शिक्षा जगत को, जहां सबसे बुनियादी अनुसंधान और विकास गतिविधियां होती हैं, आगे सोचने और इन गतिविधियों के साथ व्यावसायीकरण पहलुओं को शामिल करने और लंबे समय में खुद को आत्मनिर्भर बनाने में सक्षम बनाती हैं। यह परिदृश्य शिक्षा जगत के लिए कई अप्रयुक्त अवसरों को खोलता है और उन्हें हितधारकों के लिए नए ज्ञान को विकसित करने और प्रसारित करने के पारंपरिक दृष्टिकोण से परे सोचने में सक्षम बनाता है। अधिकांश शिक्षाविद और निर्णय-निर्माता इस बात से सहमत हैं कि उद्यमिता एक नया रास्ता है और बड़े पैमाने पर समाज की उन्नति और कल्याण के लिए आवश्यक है। यह थीसिस जैव सूचना विज्ञान में अनुसंधान के प्रबंधन पहलू का पता लगाने और उस दिशा में एक व्यवहार्य विकास और राजस्व आधारित मॉडल विकसित करने का प्रयास करती है।

थीसिस नौ अध्यायों में विभाजित है। अध्याय 1 देश में जैव सूचना विज्ञान अनुसंधान की वर्तमान स्थिति पर चर्चा करता है, और यह सूचना प्रौद्योगिकी-संचालित अनुप्रयोगों में

प्रगति के साथ जीवन विज्ञान और स्वास्थ्य देखभाल-आधारित अनुसंधान पारिस्थितिकी तंत्र को कैसे मजबूत कर रहा है। इसके अलावा, अनुसंधान पाइपलाइन को औपचारिक बनाने और सुव्यवस्थित करने के लिए प्रबंधन दृष्टिकोण पेश किए जाते हैं और उन पर चर्चा की जाती है। इसके बाद विभिन्न राजस्व सृजन धाराओं की खोज करके आत्मनिर्भरता पहलुओं को जोड़ा जाता है और अकादमिक सीमाओं के भीतर से ही एक पूर्ण राजस्व मॉडल रखने की दृष्टि को साकार किया जाता है।

अध्याय 2 भारत में जैव सूचना विज्ञान गतिविधियों के अनुसंधान और प्रबंधन दोनों पहलुओं और वैश्विक स्तर पर अपनाई गई सर्वोत्तम प्रथाओं पर विचार करते हुए, थीसिस विषय से संबंधित व्यवस्थित साहित्य समीक्षा पर केंद्रित है। थीसिस का यह हिस्सा मूल रूप से अनुसंधान और प्रबंधन पहलुओं के सिंक्रनाइज़ेशन पर साहित्य में प्रकाशित अनुभवजन्य अध्ययनों की जांच करता है, और निष्कर्षों के आधार पर, भारत में जैव सूचना विज्ञान अनुसंधान और विकास में एक स्थायी मॉडल विकसित करने का प्रयास किया जा रहा है।

अध्याय 3 में इस अध्ययन में शामिल माप के पैमाने, विभिन्न कारकों, उप-कारकों और साहित्य समीक्षा के माध्यम से या प्रश्नावली सर्वेक्षण और साक्षात्कार के उत्तरदाताओं से अध्ययन के संचालन और अंततः एकत्र किए गए डेटा का विश्लेषण करने के लिए अपनाई गई पद्धति का विस्तृत विवरण है। इस थीसिस के उद्देश्यों के अनुरूप मॉडल विकसित करने में मदद करता है।

अध्याय 4 में चर्चा की गई है कि कैसे जैव सूचना विज्ञान शोधकर्ता शिक्षा जगत से संबंधित हैं और जीवन विज्ञान और स्वास्थ्य सेवा उद्योगों के उद्यमी एक-दूसरे के काम के पूरक के लिए अनुकूल वातावरण विकसित करने के लिए हाथ मिलाते हैं। प्रबंधन-आधारित पद्धतियों और दृष्टिकोणों को अपनाने से इस प्रयास को और बल मिलता है।

अध्याय 5 मानव संसाधन विकास कार्यक्रमों को मजबूत करके जैव सूचना विज्ञान अनुसंधान एवं विकास के लिए एक ज्ञान-आधारित पारिस्थितिकी तंत्र विकसित करने की भूमिका को दर्शाता है और चर्चा करता है जो देश में प्रशिक्षित जैव सूचना विज्ञान पेशेवरों का एक पूल बनाने में मदद करता है जो सार्वजनिक स्वास्थ्य से जुड़े इस अत्यधिक विशिष्ट क्षेत्र को चला और प्रबंधित कर सकते हैं।

अध्याय 6 शिक्षा जगत में जैव सूचना विज्ञान अनुसंधान एवं विकास के लिए उच्च-प्रदर्शन कंप्यूटिंग (एचपीसी) प्रौद्योगिकी उन्मुख राजस्व मॉडल की भूमिका और महत्व पर प्रकाश डालता है और कैसे विभिन्न कारक और उनकी अन्योन्याश्रयताएं एक तकनीकी पारिस्थितिकी तंत्र बनाने में अपनी भूमिका निभाती हैं जो एक रूपरेखा विकसित करने का

मार्ग प्रशस्त करती है। प्रबंधन-आधारित दृष्टिकोण की सहायता से शैक्षणिक सीमाओं के भीतर से।

अध्याय 7 शैक्षणिक सीमाओं से उपयोगकर्ता-अनुकूल और इंटरैक्टिव सॉफ्टवेयर/टूल्स और डेटाबेस सहित जैव सूचना विज्ञान प्रौद्योगिकियों के विकास और प्रबंधन पर केंद्रित है, जहां इनमें से अधिकतर शोध गतिविधियां की जा रही हैं और जो मूल रूप से अधिकांश सफल वाणिज्यिक उद्यमों की उत्पत्ति हैं। वाणिज्यिक कंपनियों द्वारा पेश किए गए जैव सूचना विज्ञान सॉफ्टवेयर सुइट्स में आता है।

अध्याय 8 त्रिकोणासन और संश्लेषण पहलुओं से संबंधित है, जो महत्वपूर्ण शोध सिद्धांत हैं जो अकादमिक कार्यों को प्रामाणिकता और महत्व प्रदान करते हैं। यह व्यापक ज्ञान प्रदान करने के लिए कई दृष्टिकोण और डेटा स्रोतों का समापन करता है। इस अध्याय में इस थीसिस में किए गए शोध कार्य के निहितार्थ भी शामिल हैं।

अध्याय 9 सारांश, थीसिस कार्य से उभरे कुछ दृष्टिकोण, कुछ सीमाएँ और भविष्य के दायरे पर चर्चा करता है।

Table of Contents

CERTIFICATE.....	I
ACKNOWLEDGEMENTS.....	II
ABSTRACT.....	IV
संर.....	VII
TABLE OF CONTENTS.....	X
LIST OF FIGURES.....	XVI
LIST OF TABLES.....	XVII
LIST OF ABBREVIATIONS	XIX
CHAPTER 1: INTRODUCTION.....	1
1.1 Introduction to the study.....	2
1.2 Research Background.....	5
1.3 Motivation for the study and need for a revenue model	6
1.4 Research Aim and Objectives.....	9
1.5 Research Questions.....	10
CHAPTER 2: LITERATURE REVIEW.....	12
2.1 Introduction.....	13
2.2 Scope of Literature review.....	13
2.3 Academic Research and Entrepreneurship.....	17
2.3.1 Understanding the importance of IPR in Entrepreneurship.....	28
2.4 Human resource development at the core of the bioinformatics ecosystem.....	29
2.5 Bioinformatics research and the need for computing infrastructures.....	31
2.6 Bioinformatics research illustration through user-friendly Software and Apps....	41
2.7 Initiatives so far in India in the field of bioinformatics.....	44
2.7.1 Initiatives taken to promote entrepreneurship in life science domain in India....	46
2.7.2 Entrepreneurship in experimental biology using data analysis in India.....	48
2.8 Synthesis and Research gaps.....	50
2.8 Concluding Remarks.....	50
CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY.....	56
3.1 Introduction.....	57
3.1.1 Participatory Research.....	57

3.1.2	Action Research.....	58
3.2	Research philosophy and its adoption in the current study.....	58
3.3	Research Approach.....	62
3.4	Research Design.....	63
3.4.1	Research Method.....	63
3.4.2	Sampling Decisions.....	65
3.4.3	Data Collection.....	66
3.4.4	Data Analysis.....	68
3.4.4.1	SWOT Analysis.....	69
3.4.5	Time Horizon.....	70
3.5	Credibility of Research Findings.....	71
3.5.1	Validity.....	71
3.5.2	Reliability	71
3.6	Research Direction and Research Flow.....	72
3.7	Concluding Remarks.....	75
	CHAPTER 4: ACADEMIC BIOINFORMATICS COLLABORATION WITH ENTREPRENEURIAL VENTURES: A FUTURISTIC STRATEGY IN LIFE SCIENCE RESEARCH (CASE STUDIES FROM INDIAN ACADEMICS, START- UPS, AND THE LIFE SCIENCE SECTOR.....	76
4.1	Introduction.....	77
4.1.1	Bioentrepreneurship.....	78
4.1.2	Bioinformatics Research Activities in the Indian Context.....	79
4.1.3	Detailed analysis of BTISNet.....	83
4.1.4	Academic institutions from India and their linked start-ups.....	92
4.2	Methodology.....	99
4.3	Key observations and outcomes.....	100
4.4	SAP-LAP Analysis of the Case Study.....	104
4.4.1	Situation.....	104
4.4.1.1	Industry Perspective.....	104
4.4.1.2	Academic Perspective.....	105
4.4.2	Actor.....	105
4.4.3	Process.....	105
4.4.4	Learning.....	106

4.4.5	Action.....	106
4.4.6	Performance.....	107
4.5	Challenges, Issues, and Implications.....	107
4.6	Discussion and Suggestions.....	112
4.7	Recommendations leading to a Business Model.....	115
4.8	Concluding remarks, inferences and way forward.....	115
CHAPTER 5: DEVELOPING A KNOWLEDGE-BASED ECOSYSTEM FOR BIOINFORMATICS R&D BY STRENGTHENING HUMAN RESOURCE DEVELOPMENT PROGRAMS: A CASE STUDY AND ANALYSIS IN INDIAN CONTEXT USING SAP-LAP BASED APPROACH.....		117
5.1	Introduction.....	118
5.2	Research Objectives... ..	121
5.3	Research Methodology.....	121
5.4	Survey details, analysis, and critical observations.....	123
5.5	Data Preparation and Tools.....	124
5.6	Exploratory Factor Analysis.....	126
5.6.1	Parallel Scree Plot.....	127
5.6.2	Analysis and Interpretation.....	128
5.6.3	Adequacy Test.....	129
5.6.4	Nomenclature of the factors.....	130
5.7	Reliability and validity of data.....	131
5.7.1	Factor-wise reliability test.....	131
5.7.2	Validity for the questionnaire data.....	132
5.8	SAP-LAP: An overview.....	133
5.9	SAP-LAP linkages.....	134
5.10	SAP Model for Human Resource Development for Bioinformatics R&D.....	135
5.10.1	Situation.....	135
5.10.2	Actor.....	136
5.10.3	Cross-interaction matrix 'Situation × Actor'.....	136
5.10.4	Process.....	139
5.11	LAP Synthesis.....	140
5.11.1	Learning.....	140

5.11.2 Action.....	141
5.11.3 Performance.....	141
5.12 Learning Issues.....	142
5.13 Suggested actions based on the literature and survey analysis.....	143
5.14 Performance Metrics.....	146
5.15 Synthesis of LAP Framework.....	148
5.16 Cross-interaction matrix ('Action × Performance')	149
5.16.1 Binary matrix (Action*Performance)	149
5.16.2 Interpretive cross-interaction matrix (Action*Performance)	149
5.17 Discussion, and Managerial Implications.....	151
5.18 Conclusion and Recommendations.....	153
5.19 Limitations and Future Directions.....	155
CHAPTER 6: ANALYSIS OF THE CRITICAL FACTORS FOR THE DEVELOPMENT OF HIGH-PERFORMANCE COMPUTING TECHNOLOGY- ORIENTED REVENUE MODEL FOR BIOINFORMATICS R&D IN ACADEMIA.....	156
6.1 Introduction.....	157
6.1.2 Details of the supercomputers from India since PARAM 8000.....	159
6.1.3 Purpose and expectations from other series of PARAM.....	161
6.2 Literature Review	163
6.2.1 Identification of critical factors delineating the impact of HPC in Bioinformatics R&D in an academic set up.....	163
6.2.2 List of factors deduced from the literature review.....	164
6.3 Methodology.....	164
6.4 Sample Characteristics (details of the participants of the survey)	167
6.5 Differentiation of factors with respect to the impact of HPC in Bioinformatics.....	169
6.6 TISM Analysis.....	171
6.7 Modified-TISM (m-TISM) process.....	174
6.8 MICMAC analysis.....	179
6.9 Results.....	182
6.10 Discussions on the development of an HPC oriented Revenue model in bioinformatics for growth and sustainability.....	186

6.11	Conclusion.....	189
6.12	Implications.....	191
6.13	Limitations and Future Scope.....	193
CHAPTER 7: DEVELOPMENT AND MANAGEMENT OF BIOINFORMATICS SOFTWARE & TOOLS: A CASE STUDY ON USER-FRIENDLY AND INTERACTIVE INDIGENOUS SOFTWARE/TOOLS AND DATABASES.....		194
7.1	Introduction.....	195
7.2	Background of this study.....	195
7.3	Literature Review.....	199
7.3.1	Popular Genomics Software.....	202
7.3.2	Popular Protein Structure Prediction Software.....	204
7.3.3	Popular Drug Design Software.....	206
7.4	Methodology.....	208
7.4.1	Research Questions.....	211
7.4.2	Hypotheses.....	211
7.5	Backcasting Model Exploration.....	212
7.6	Description of factors using backcasting approach.....	217
7.7	Elaboration of user-friendly and interactive software/tools developed as a part of the study.....	222
7.8	Discussion.....	225
7.9	Conclusion.....	233
CHAPTER 8: TRIANGULATION AND SYNTHESIS OF RESULTS TOWARDS DEVELOPING A REVENUE MODEL.....		235
8.1	Research Triangulation & Synthesis.....	236
8.2	Business & Revenue Model Perspectives.....	237
8.3	Thesis Triangulation & Synthesis.....	239
8.4	Implication of results and findings.....	242
8.4.1	<i>Implication of Study 1 (Chapter 4 of this thesis)</i>	243
8.4.2	<i>Implication of Study 2 (Chapter 5 of this thesis)</i>	244
8.4.3	<i>Implication of Study 3 (Chapter 6 of this thesis)</i>	245
8.4.4	<i>Implication of Study 4 (Chapter 7 of this thesis)</i>	248
8.5	Distinction between bioinformatics R&D scenario in India and abroad.....	250

8.6	Past & present scenarios and futuristic plans and strategy for India.....	252
8.7	Few business models in bioinformatics in Indian context.....	255
8.8	Development of a Business and Revenue model.....	258
8.9	Cost analysis for the revenue model.....	260
8.9.1	Features of the model ensure sustainability for the stakeholders.....	263
8.10	Schematic representation of the proposed revenue model.....	265
CHAPTER 9: CONCLUSION, LIMITATIONS & FUTURE SCOPE.....		266
9.1	Conclusion & limitations of present research.....	267
9.2	stakeholder's driven policy change leading to attainment of the objectives.....	270
9.3	Directions for future research.....	275
References.....		277
Appendix.....		309
A	Format of the questionnaire/survey.....	310
A.1	Interview questions related to Academic bioinformatics activities joining hands with Entrepreneurial Ventures.....	310
A.2	Questionnaire on creating a pool of bioinformatics professionals in India.....	313
A.3	Questionnaire on developing a revenue model around High-Performance Computing (HPC) technology in bioinformatics.....	316
A.4	Comparison of top pharma companies (global vs Indian) and their revenues.....	320
B.1	Plots for illustrating the impact of human resource training.....	321
B.2	Table for testing validity using Pearson's correlation.....	329
B.3	Factor analysis data and reliability test.....	331
C.1	TISM factors considered for comparison and interpretation.....	334
C.2	Interpretation of TISM factor linkages.....	336
C.3	Segregation of unique factors for an optimized Revenue model.....	342
List of publications.....		345
Curriculum Vitae: Shashank Shekhar.....		347

LIST OF FIGURES

Figure 1: Major activities and milestones achieved by BTISNet since inception.....	87
Figure 2: Summary of publications in bioinformatics in India since 2001.....	88
Figure 3: Correlation plot of the questionnaire.....	125
Figure 4: Correlation plot influencing the industrial exposure and commercialization of Bioinformatics research in academia.....	126
Figure 5: Scree Plot Parallel analysis to find the number of factors.....	127
Figure 6: Factors and correlation with 25 parameters.....	128
Figure 7: Relevant parameters when cut off is 0.3 considering five factors.....	128
Figure 8: Factors and correlation with 25 parameters.....	129
Figure 9: Relevant parameters when cut off is 0.3 considering four factors.....	129
Figure 10: Analysis of 4 latent factors with 25 parameters in questionnaire.....	130
Figure 11: Steps involved in TISM method.....	171
Figure 12: m-TISM: paired comparisons and simultaneous transitivity checks.....	175
Figure 13: TISM model of critical factors depicting the role and impact of HPC in Bioinformatics R&D.....	178
Figure 14: MICMAC analysis of factors showing the impact of HPC in Bioinformatics R&D.....	181
Figure 15: Illustration of an HPC based revenue model component in an academic environment.....	189
Figure 16: Drug discovery flow chart illustrating the significance of Bioinformatics....	214
Figure 17: In silico Drug discovery assembly line developed at SCFBio, IIT Delhi.....	217
Figure 18: Futuristic scenario for personalized medicine using backcasting approach...	222
Figure 19: Windows OS-based workbench, android app for Drug Design modules, and set of basic tools and software.....	224
Figure 20: Revenue model framework for Bioinformatics R&D in academia.....	265
Figure 21-28: Plots representing the views of respondents of the survey for developing a strong pool of bioinformaticians through application-based training and industry exposure.....	321-328

LIST OF TABLES

Table 1: TRL levels and brief detail.....	21
Table 2: Entrepreneurial case studies.....	22
Table 3: Factors affecting the entrepreneurial initiatives.....	24
Table 4: Current scenario, prevalent Issues and suggestions in Indian context.....	51
Table 5: Research philosophies and their corresponding attributes.....	60
Table 6: Research objectives and methods adopted.....	72
Table 7: Institutes linked with bioinformatics start-ups and their work domain.....	92
Table 8: Facilities similar to IBDC in USA, Europe, Japan and China.....	93
Table 9: Details of the respondents of the survey.....	123
Table 10: Reliability index.....	131
Table 11: Dynamic SAP framework.....	134
Table 12: Dynamic LAP framework.....	135
Table 13: Binary Matrix ('Situation x Actor')	137
Table 14: Interpretive matrix ('Situation × Actor')	137
Table 15: Summary of LAP framework.....	148
Table 16: Binary matrix ('Action × Performance')	149
Table 17: Interpretive matrix ('Action × Performance')	150
Table 18: India's first Supercomputer details.....	159
Table 19: Details of leading supercomputers from India.....	160
Table 20: Details of the respondents of the survey.....	168
Table 21: List of factors based on the responses in percentage terms.....	169
Table 22: List of factors concerning the impact of HPC in Bioinformatics.....	170
Table 23: Reachability matrix.....	176
Table 24: Reachability matrix with transitivity.....	176
Table 25: Partitioning of the final reachability matrix into different levels (Iteration 1)	176
Table 26: Iteration 2.....	176
Table 27: Iteration 3.....	177
Table 28: Iteration 4.....	177

Table 29: Iteration 5.....	177
Table 30: Iteration 6.....	177
Table 31: Partitioning of the final reachability matrix into different levels.....	180
Table 32: MICMAC analysis.....	180
Table 33: Driving and Dependence Power.....	181
Table 34: Illustration of factors as per MICMAC analysis.....	182
Table 35: Leading genomics software.....	202
Table 36: Leading proteomics software.....	204
Table 37: Leading drug design software.....	206
Table 38: Leading institutes/centers involved in indigenous software development in India.....	209
Table 39: Comparative bioinformatics scenario in India and elsewhere.....	250
Table 40: Business Model Canvas for Bioinformatics R&D Ecosystem.....	258
Table 41: Impact of policy change on the objective of research.....	270

LIST OF ABBREVIATIONS

Abbreviations	Definitions
BTISNet	Biotechnology Information System Network
DBT	Department of Biotechnology
CDAC	Centre for Development of Advanced Computing
DeitY	Department of Electronics and Information Technology
DST	Department of Science and Technology
E7 Countries	A group of emerging market economies from Brazil, China, India, Indonesia, Mexico, Russia and Turkey.
EU	European Union
G7 Countries	A group of advanced economies of Canada, France, Germany, Italy, Japan, the UK, and the US
HPC	High Performance Computing
IIT	Indian Institute of Technology
IISc.	Indian Institute of Science, Bangalore
IoT	Internet of Things
MD/VP	Managing Director/ Vice President
MICMAC	Matrice d'impacts croisés multiplication appliquée à un classment
M-TISM	Modified Total Interpretive Structural Modeling
PI	Principal Investigator
R&D	Research and Development
R&T	Research and Technology
SAP-LAP	Situation, Actor, Process–Learning, Action, Performance
SCFBio	Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi
TRL	Technology Readiness Level