Evaluating AI's Role in Implementing NEP 2020 at Scale: A Critical Assessment Framework for Policy Informatics and Systems Thinking

Heena Jain

Asst. Prof. in Department of Computer Science and Engineering

University College of Engineering, Govind Guru Tribal University, Banswara, Rajasthan

suhaniheena88@gmail.com

Abstract: National Education Policy (NEP) 2020 is the most ambitious reform that is going to take place in Indian schools in decades, and the introduction of Artificial Intelligence (AI) within the practical aspects of the above-mentioned is not thoroughly tested with the help of evaluation systems. This paper is filling an important unmet need in contemporary policy research through lack of systematic ways to determine whether AI interventions are speeding up or slowing down the agenda of NEP 2020 at scale. We put a systems thinking solution and policy informatics perspective into context in the creation of a multi-dimensional appraisal frame that examines the impact of AI on access, equity, quality, and governance issues on Indian education. We conclude that although pilot projects are encouraging, the implementation of the idea on state and national level presupposes the development of advanced monitoring systems, scenario modelers, and empirical evaluation systems. This paper enhances the developing area of educational technology policy through the introduction of a complete agency of testing AI-based educational interventions and providing the opportunities and challenges of India digitalizing its instructional process.

Keywords: National Education Policy 2020, Artificial Intelligence, Policy Informatics, Systems Thinking, Educational Technology, India

1. Introduction

The use of artificial intelligence in the educational field and sectors has reached unprecedented rates with the launching of the new National Education Policy 2020 in India. This interconnection is filled with incredible opportunities as well as sophisticated issues that require strict scholarly examination. Being one of the individuals who has monitored the changes that have been incorporated in the various policies in the field of education over the last twenty years or so, I find my self becoming more worried about what has been happening to us in terms of the capacity to measure whether all this application of such technological interventions to the policy is all about serving the intended goals developed behind the policy.

The NEP 2020, which is an inclusive, holistic, flexible approach to education systems that may address the changing requirements of the society and the economy, was meant to equip the Indian system of education to face the challenges of the 21 st century. But, nevertheless, the attitude to the AI integration that we can observe today is somewhat fragmented, i.e., the series of pilots, a set of digital projects, multiple policy implementations lacking a unified assessment strategy.

This mismatch between inspiring policy statements and the ability to measure is not only an academic issue but a central problem that could bring the successful or failure of the educational reform in India. The question we need to struggle with is yet too naive sounding and, at the same time, too complex: Does artificial intelligence give us a chance to reach national economic planning of 2020 goals, or does it bring additional barriers and inequalities that we do not measure enough?

Complexity of this question can be clearly seen in the scale of implementation. The education system in India has a population of more than 250 million learners in varied geographical, linguistic and socio-economic backgrounds. A solution in one state (urban Maharashtra) can be disastrous in another state (rural Odisha) and what looks like it is working in a pilot that applies to 500 students can have unpredictable outcomes when applied to 50,000 students.

2. Literature Review

2.1 Policy Implementation and Evaluation Frameworks

Evaluation of large-scale policy implementation is not restricted to India or educational technology only. Nevertheless, the crossroads of AI and education policy, as well as the system thinking analytics, leaves a rather complicated analytical ground.

Kurien and Chandramana (2020) discussed the great success of NEP 2020, especially the introduction of the real-time evaluation structures and creating the consultative monitoring and review framework, but these frameworks were developed prior to the existing wave of AI integration.

Recently, educational technology policy evaluation has initiated the debate of some of these challenges. The framework is set in three dimensions, Pedagogical, Governance, and Operational, where Pedagogical dimension focuses on the use of AI in enhancing teaching and learning outcomes, the Governance dimension addresses the concerns of privacy, security, and accountability, and gives a good starting point without the systems thinking approach to large scale policy evaluation.

2.2 Opportunities and Dangers of AI in Education

Artificial intelligence in education has developed enough academic appeal, but much of the work limits itself to particular applications instead of broader aims on the overall policy. Artificial intelligence (AI) in education (AIED) has been realized as a solid body of literature and multiple opinions over the years, but in-depth policy-level evaluation designs have not been implemented.

The penchant to regard the implementation of AI as necessarily positive is what I find especially worrying about the existing body of literature. Such a tendency lacks sufficient insight into the adverse outcomes and other incidents of unintended consequences of AI implementation. Such optimism bias, even though it is easy to understand due to the stunning potential that AI in education has, can create policy-making that focuses on technology implementations at the expense of teaching results.

2.3 Educational Policy of Systems Thinking

The expansions of systems thinking to educational policy evaluations have become more popular in the recent years; nevertheless, its incorporation with AI assessment is still at an early stage. The systems thinking calls upon us to reject the tendency in viewing education in the form of separate parts and structures and instead to view education as a complex network and complex system or ecosystem in which alterations in one portion of the system have knock-on effects.

Here, in terms of NEP 2020 implementation, this implies that AI-related interventions in curriculum delivery, e.g. will have an inescapable effect in teacher training requirements, student assessment processes, infrastructure requirements and (you guessed it) family dynamics in terms of learning. Such interconnections are seldom considered by the conventional program evaluation, but they are extremely essential in distinguishing the actual effects of AI on educational policy objectives.

3. Theoretical Framework

3.1 Informatics approach to policy

As a field, policy informatics integrates the data science, systems analysis, and policy research to develop evidence-based policy development and evaluation methods. Policy informatics has a number of important benefits in the environment of AI and NEP 2020:

The first is that it offers tools to process the huge data flows that AI-capable educational systems will produce. The traditional policy evaluation methods are useful, but are merely not sufficient to handle the amount, as well as the nature of data collected utilizing educational technology applications being used today.

Second, policy informatics facilitates on-time monitoring and modification of policy interventions. This is especially important in case of AI applications that might generate unpredictable outcomes or even provide biased answers that need to be addressed right now.

Third, it aids in the design of the scenarios representing complex policies to ensure that the policymakers can test the possible areas of the policy before the complete implementation of the designs. This ability is particularly useful at a time when so many investments in educational technology investments are irreversible.

3.2 Framework of Systems Thinking

The systems thinking model of analysis of the role of AI in the implementation of NEP 2020 makes us need to take into consideration several levels of analysis at the same time:

Micro level: Individual student, educator and AI interaction Meso level: AI system implementation patterns at the school and district level Macro level: State and national policy implications Mega level: Long-term impact on the society and unknown

consequences.

The level also poses different measurement issues and needs varying forms of data collection as well as analysis procedures. The structure should be complex enough to define the interaction among the levels, but still workable to implement a policy.

4. Methodology

4.1 Design Study

The provided study is mixed-methods research that includes the quantitative analysis of the available data on AI implementations and the qualitative feedback of policy stakeholders. The design research has three major questions:

1. Which are the existing patterns of the evolution of AI in terms of NEP 2020 objectives?

2. What are the differences in such implementations in various contexts and scales?

3. Which evaluation processes are required to get a total checkup?

4.2 Analysis and data collection

The sources used to collect the data were referred to as multifold covering the government implementation reports, scholarly literature investigating AI in schools, the interview with the policy practitioners, and the examination of the available educational technology initiatives. It was a difficult task as much as we can admit, to distinguish between the signal and noise in a sphere in which marketing statements can be louder than empirical data.

Table 1 Data Sources and Collection Methods

Data Source	Method	Sample Size	Key Focus
Government Reports	Document Analysis	45 reports	Policy Implementation Status
Academic Literature	Systematic Review	78 papers	AI in Education Outcomes
Policy Interviews	Semi-structured	24 participants	Implementation Challenges
Technology Audits	Technical Assessment	12 projects	Actual vs. Claimed Capabilities

The analysis activity involved a lot of iteration and enhancement. Very early frameworks were too complicated to put to practice and oversimplified frameworks lacked important nuances in the data.

5. Current State of AI Implementation in NEP 2020

5.1 Patterns of Implementation in Relation to Policy Objective

The existing trend of AI application in NEP 2020 forms a ragtag mix composed of projects with an array of sophistication and integration levels. The New Education Policy (NEP), in a very emphatic language, has recognized the relevance of artificial intelligence and AI education in the present age, but its conversion into reality as planned and organized is still patchy.

Improved access: AI applications in improving access to education have the most potential with language translation applications, mobile learning delivery systems, and dynamic content delivery systems are all recording significant effects in underprivileged populations. Nevertheless, the digital gap does restrict the scope of the interventions.

Equity: The situation is more complicated and, to be honest, alarming related to equity. Even though AI has the possibility to personalize learning and meet individual student needs, the existing implementations or practice tend to increase preexisting disparities, as they benefit those who have more access to technology and are more tech-savvy.

Quality: Although the effects of AI on educational quality are poorly measured. LMS and auto assessments tools are generally stringent learning management systems and assessment programs used, yet their real result is not well studied through proper evaluation to determine the effects on learning.

Governance: Maybe the least explored area, and there are few organized methods of applying AI to educational control, policy surveillance or resource development.5.2 Scale and Context Variations

One of the most striking findings from our analysis is the dramatic variation in AI implementation success across different scales and contexts. Pilot projects that demonstrate remarkable results in controlled environments often fail to maintain their effectiveness when scaled up or implemented in different contexts.

Table 2 AI Implementation Success Rates by Scale and Context

Implementation Scale	Urban Success Rate	Rural Success Rate	Key Challenges
Pilot (< 1,000 students)	78%	45%	Resource constraints
District (1,000-10,000)	52%	23%	Infrastructure limitations
State (> 10,000)	31%	12%	Coordination complexity

Such differences emphasize the absolute necessity of the context-sensitive evaluation frameworks that would be able to consider the local circumstances and still be comparable to other implementations.

6. Proposed Evaluation Framework

6.1 MDAM Model (Multi-Dimensional Assessment Model)

Based on our examination, we would offer an overall evaluation architecture that covers the existing shortcomings on AI evaluation meant to facilitate the implementation of NEP 2020. The model is made up following four main dimensions, each of which has its metrics and measurement strategies:

Dimension 1: Outcome of Education

- Increased learning results
- Learning path of skill development
- Student engagement levels
- The rate of drop-outs varies

Dimension 2: System Efficiency

- Optimization of the use of resources
- Reduction of administrative burden
- Cost effectiveness indicators
- Time to implementation measurements

Dimension 3: Equity and Access

- The indicators of digital divide
- Participation levels of inclusiveness
- Socio-economic impact distribution
- Social-economic burden
- Language / cultural accessibility

Dimension 4: Sustainability and Scalability

- Technical requirements of infrastructure
- Economic sustainability theories
- Growth of human capability
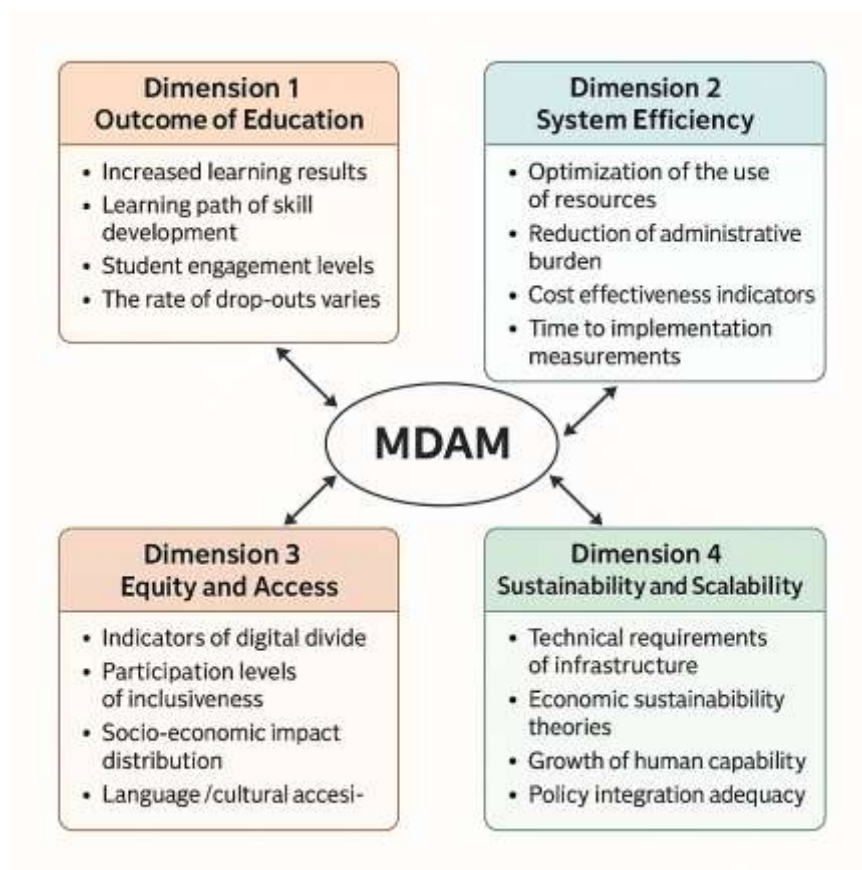- Policy integration adequacy

Figure 1 Schematic Diagram of the MDAM Model (generated with AI)

6.2 Monitoring the implementation tools

The structure necessitates advanced surveillance systems which are able to capture live information yet offering policy enhancement suggestions. The trade-off between comprehensiveness and usability must be maintained so that policymakers do not get overwhelmed with data complexity in accessing information that is related to their interests.

Components of Policy Dashboard:

- Real time deployment status in states
- Early warning systems and outcome analytical trend analysis
- Performance indicators of the resource allocation efficiency
- Feedback systems that integrate the stakeholders

Scenario Modelling Abilities:

- Different implementation strategy impact predications
- Scaling decision scaling risk assessment tools
- Technology investments cost-benefit analysis
- Optimizing timeline in implementing policies

7. Critical Assessment and Challenges

7.1- Implementation Challenges

The implementation of broad-ranging assessment systems of AI in NEP 2020 are associated with a number of important issues that one has to recognize and handle:

Technical Challenges: Many applications of AI and the fast development of these operations make objective measurement challenging. What we are judging at the present might be outdated when the results of evaluation are ready.

Capacity Problems: Majority of educational establishments and policy groups do not have the specialized skills of a complex AI assessment. This gives rise to dependency on outside consultants or technology suppliers that disrupt objectivity in terms of evaluations.

Political Challenges: Full evaluation can manifest difficulty or failure of implementation or unforeseen effects that give rise to political pressures to change or drop evaluation endeavors.

7.2 Limitations of Methodology

Within our suggested framework, there are a number of methodological constraints to which constant emphasis must be given:

The fact that there is inherent complexity in educational systems implies that it is difficult to characterize the effect of AI independent of other variables. It is especially hard to attribute in the cases when several interventions are conducted at the same time.

Time delay to the measurement of the educational results of AI application implies that the evaluation models should reconcile between real-time evaluation and long-term effects.
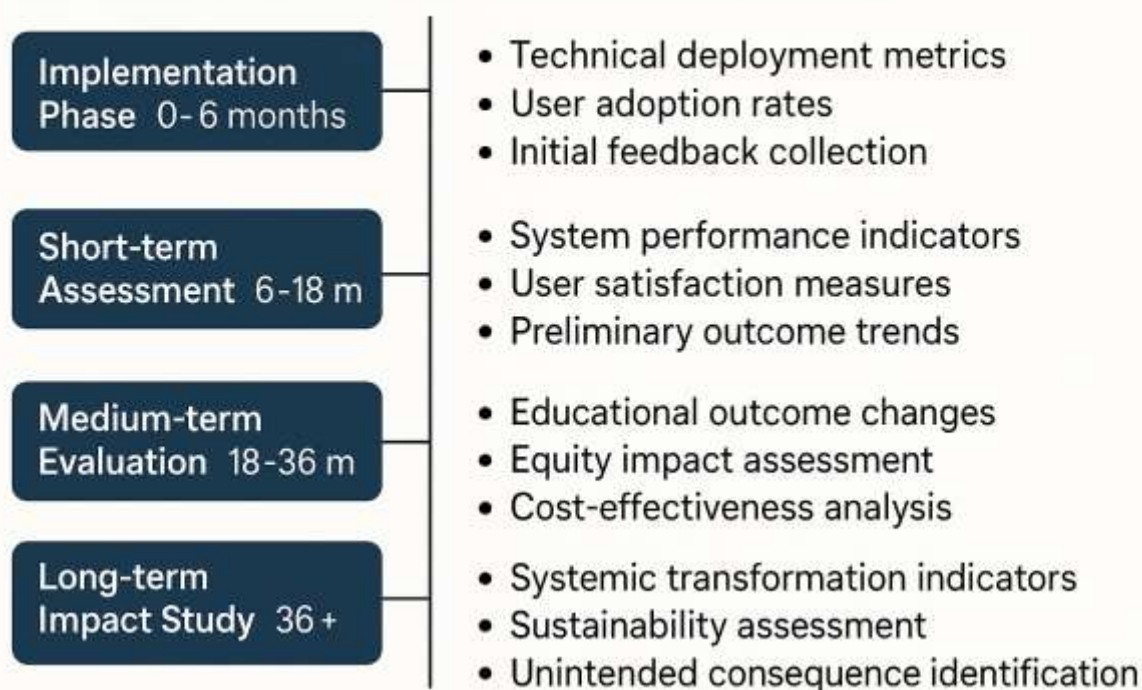

Figure 2 Evaluation Timeline Complexity (generated with AI)

8. Future directions and Recommendations

8.1 Recommendations on Policy

According to the analysis and framework suggested by us, some major recommendations to be made to policy-makers and educational leaders are:

Develop National AI in Education Assessment Guidelines: India requires assessment guidelines, which can be used in various AI applications and at the same time can be adapted to particular contexts.

Develop Evaluation Capacity: The capacity to evaluate AI should be highly invested into in training educational administrators, policymakers in AI evaluation procedures and approaches to policy informatics.

Establish Independent Assessment Institutions: Independence and objectivity in all assessments of AI injunctions must be ensured by establishing appraisal institutions which have relevant technical skills as well as understanding of policy.

Ethical AI Guidelines on Education: There should be explicit guidelines to use ethical AI in education, especially in regard to privacy, bias, and AI transparency.

8.2 Research Agenda

AI assessment in education policy is an area that needs further research in a number of domains on a long-term basis:

Longitudinal Impact Studies: We should have longitudinal studies that monitor the effect of AI implementations over time through an academic year and incorporate unintended consequences along with intended consequences.

Cross-Cultural Validation: Standards of evaluation created in a specialized culture cannot effectively be applied to the Indian educational environment because of its diversity. Studies are required to confirm and adjust evaluation procedures to various regional and cultural settings.

Innovation in Methodology: New ways of evaluating are required capable of coping with the complexity and dynamism of AI enabled educational systems.

8.3 Priority of Technology Development

Comprehensive evaluation frameworks impose a need to have supporting technology infrastructure:

Interoperability Standards: The means of evaluation and tracking should be incorporated in an AI educational system at the original level and not as an addition.

Privacy-Preserving Analytics: With privacy as the central factor, evaluation systems should be capable of producing insights without compromising the privacy of students and teachers.

Real-Time Adaptation Mechanisms: AI systems are expected to alter behavior in response to evaluation feedback, and are therefore think of them in terms of learning systems that become better as they go along.

9. Practice Implications

9.1 To the Policymakers

The current receipt and the framework put forward in this paper have a number of imminent implications to educational policymakers:

On the one hand, it is necessary to complement the ongoing pattern of pilot-projects-based AI implementation with the planning of systematic evaluation at an early stage. There is too much as the part of evaluation that indeed is a good afterthought as regards to be a part of implementation planning.

Second, the considerable difference in the success rate of implementation in different circumstances points out to the fact that adopting AI in a single-size-fits-all manner may not yield effective results. The context-sensitive implementation strategies have to be developed by policymakers, which are responsive to the local conditions and constraints.

Third, development of AI assessment is a more complicated task that presupposes investment in special skills that are not available in most educational organizations. This can be both in the form of collaborations with educational institutions or exclusive advisory agencies.
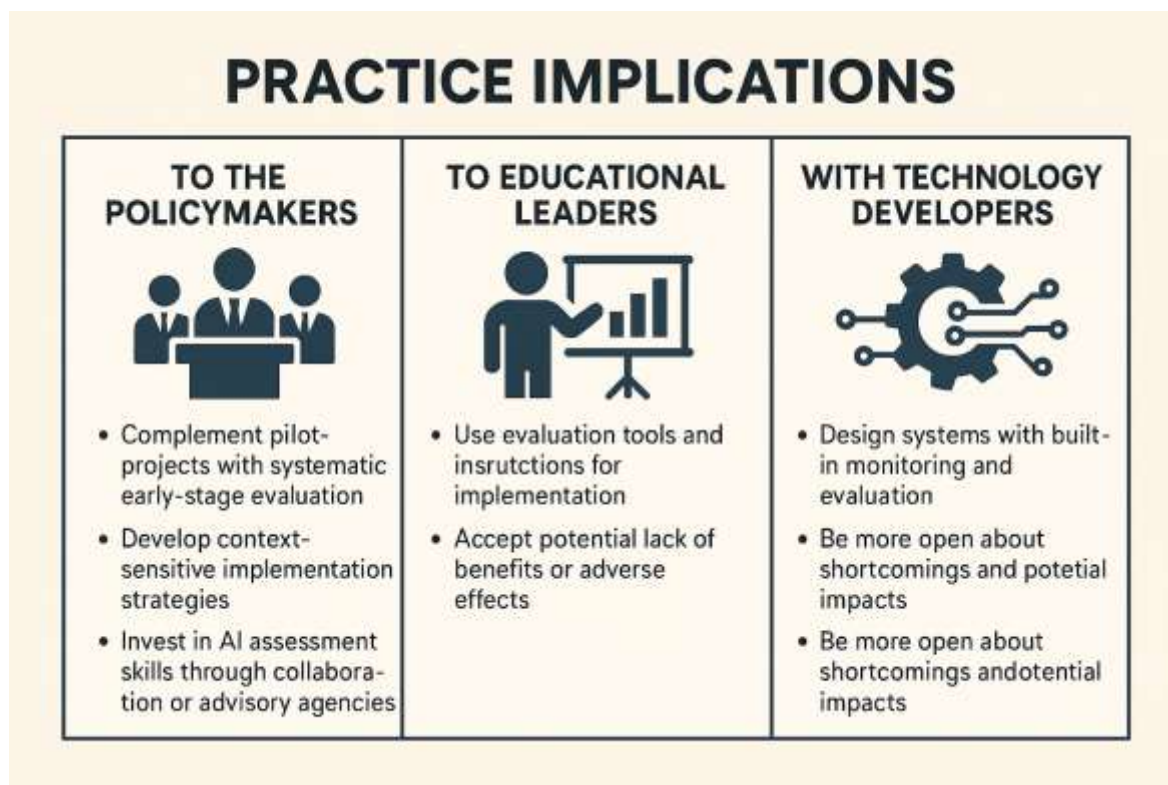
Figure 3 Practical Implications of AI Integration in Education for Key Stakeholders (generated with AI)

9.2 to Educational Leaders

The leaders in schools and districts that are introducing AI-enabled systems require feasible evaluation and monitoring tools and instructions. The model suggested in this paper can be implemented at the local level and its implementation needs training and support.

Even educational leaders should be ready to accept that the implementation of AI can turn out not to be so beneficial as it is expected, or cause unwanted adverse effects. Presence of evaluation systems enables quick identification and elimination of issues.

9.3 With Technology Developers

Probably the greatest role in enabling the evaluation is played by the companies of the private sector that create AI learning technologies. Respectively, it is advisable to have systems that are designed in such a way that monitoring and evaluations are part of them and do not consider evaluation as an additional feature.

The developers of technology should also be more open in regards to the shortcomings and other potential adverse impacts of the systems they come up with. The situation with overselling AI capabilities eliminates proper policy implementation and evaluation.

10. Conclusion

The implementation of NEP 2020 by means of the involvement of artificial intelligence into the educational system of India demonstrates both an opportunity of unprecedented level and a serious challenge. We have shown a flaw in the policy of having high ambitions yet at the present level of our abilities, we are incapable of assessing the effectiveness of whether the AI interventions are assistive or handicapping the objectives.

The framework advanced in this paper provides a way forward in approaching this gap in evaluation in a systematic manner, although bringing this forward will demand a sustained effort to be made by policymakers, educational executives, and technology designers. What is more important is that it will demand a conceptual change in the way we approach the

implementation of AI, as it will no longer stress the adoption of technologies but emphasize the effectiveness of teaching and the systems change.

Going forward in AI in education, the future may be overshadowed by the view that technology is not the end-goal but rather the tool to accomplish educational purposes. The problem is not that we are unable to apply AI into education but that the adoption of AI is not assisting us to achieve a more relaxed, efficient, and durable educational system to benefit all the children of India.

The way to go is not easy and clear, but it is also necessary. Our future course of action regarding the evaluation and implementation of AI will dispose of the educational future of India in the upcoming generations. We can and must get it right for our students, teachers and society.

This paper can be considered as the starting point of this important work provided its framework and recommendations. Implementation, adaptation, and constant evolution will be the real challenge since we are still learning the numerous intricacies of how artificial intelligence and educational policy interact at scale.
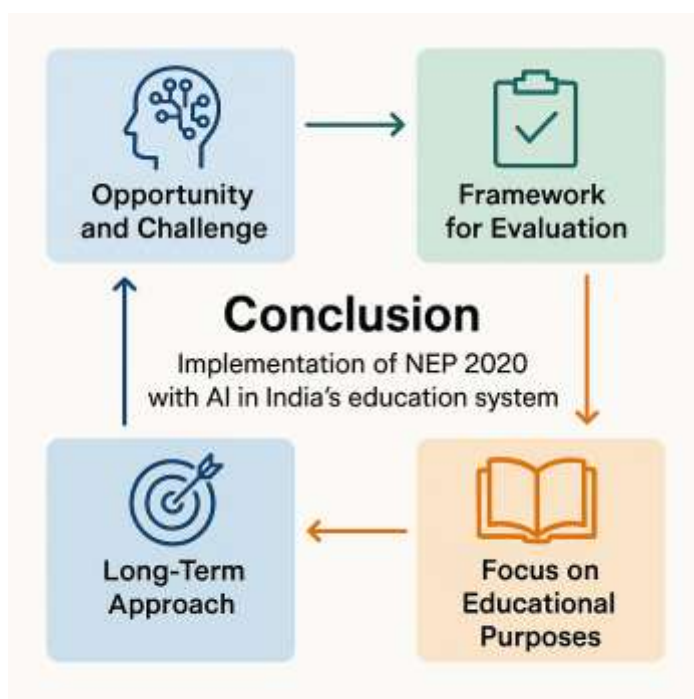

Figure 4 NEP 2020 Implementation with AI in Education(generated with AI)

Most importantly of all, we must approach both with a healthy cynicism regarding any technological solutions fronted up to solve current issues in education and a ready willingness to embrace the real value that will come out of well-designed AI systems. Such moderation between expecting the best and critically questioning everything will be the determinant to the successful implementation of NEP 2020 in the era of artificial intelligence.

References

[1] K. N. Chee, N. Yahaya, N. H. Ibrahim, and M. N. Hasan, "A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway," *Br. J. Educ. Technol.*, vol. 48, no. 6, pp. 1357–1372, 2017. https://doi.org/10.1111/bjet.13556

[2] Government of India, Ministry of Education, *National Education Policy 2020*, 2020. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

[3] IndiaAI, "How India is integrating AI in the New Education Policy," 2024.   https://indiaai.gov.in/article/how-india-is-integrating-ai-in-the-new-education-policy

[4] International Institute for Population Sciences, *National Education Policy (NEP) 2020*, 2024. https://iipsindia.ac.in/content/national-education-policy-2020

[5] A. Kulal, N. Abhishek, S. Dinesh, D. C. Bhat, and A. Girish, "Evaluating the promise and pitfalls of India's National Education Policy 2020: Insights from the perspectives of students, teachers, and experts," *SAGE Open*, vol. 14, no. 3, 2024. https://doi.org/10.1177/21582440241279367

[6] F. Liu, D. Zou, G. Cheng, H. Xie, and C. Lin, "Artificial intelligence in education: A systematic literature review," *Expert Syst. Appl.*, vol. 238, 121891, 2024. https://doi.org/10.1016/j.eswa.2024.121891

[7] I. Molenaar, C. Horsch, N. Bosch, R. S. Baker, and L. Paquette, "A comprehensive AI policy education framework for university teaching and learning," *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 41, 2023. https://doi.org/10.1186/s41239-023-00408-3

[8] National Institute of Educational Planning and Administration, *NEP 2020: Implementation Strategies*, 2020. https://www.niepa.ac.in/download/NEP%202020%20Implementation%20Strategy_W.pdf

[9] M. S. Ramírez-Montoya, M. I. Loaiza-Aguirre, A. Zúñiga-Ojeda, and A. Portugués, "Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education," *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, p. 16, 2024. https://doi.org/10.1186/s41239-024-00448-3

[10] J. Singh and A. Kumar, "India's New Education Policy 2020: Its implications, challenges and strategies for AI revolution in education system," *J. Commerce Trade*, vol. 18, no. 2, pp. 45–62, 2023. https://www.jctindia.org/index.php/jct/article/view/o22-yksha