

एआय च्या वापरामुळे निर्माण होणारे नैतिक प्रश्न : एक चिंतन

प्रा. डॉ. ग्यानदेव उपाडे

बी. रघुनाथ महाविद्यालय, परभणी

E-mail : gmupade20@gmail.com

प्रा. डॉ भारत वालोंद्रे

नबिरा महाविद्यालय, काटोल

Crossref DOI - <https://doi.org/10.63665/rh.v7i2.132>

प्रस्तावना :

कृत्रिम बुद्धिमत्ता (Artificial Intelligence – AI) ही आधुनिक विज्ञान व तंत्रज्ञानातील सर्वात वेगाने प्रगत होणारी शाखा आहे. गेल्या काही दशकांत संगणकीय शक्ती, बिग डेटा, आणि अल्गोरिदमच्या झपाट्याने झालेल्या विकासामुळे एआयने मानवी जीवनाच्या जवळजवळ प्रत्येक क्षेत्रात प्रवेश केला आहे. आरोग्यसेवा, शेती, वित्तीय क्षेत्र, उद्योग, कायदा अंमलबजावणी, शिक्षण, वाहतूक व्यवस्था, तसेच सामाजिक माध्यमे या सर्व क्षेत्रांत एआयच्या तंत्रज्ञानाने महत्त्वपूर्ण बदल घडवून आणले आहेत. या बदलांमुळे कार्यक्षमता, अचूकता आणि सोयीसुविधा वाढल्या असल्या तरी नवनवीन नैतिक प्रश्न व आव्हाने देखील समोर येऊ लागली आहेत.

आर्टिफिशियल इंटेलीजन्सच्या अल्गोरिदममध्ये या प्रोसेस मध्ये लिनिअर अल्जिब्रा, कॅल्क्युलस, मॅट्रायसेस, स्टॅटिस्टिक्स मधील प्रोबॅबिलिटी थेअरी, कोईफ्रिंशंट ऑफ कोरिलेशन अशा गोष्टी एआय मध्ये प्रामुख्याने वापरल्या जातात. मिळालेला डाटा प्रोसेस करणे, वर्गीकरण करणे, पॅटर्न ओळखणे किंवा शिकणे, त्यावरून अंदाज बांधणे किंवा भाकीत करणे अशा गोष्टी केल्या जातात. त्यामुळे कमी वेळात कोणत्याही प्रश्नाचे उत्तर मिळते, म्हणून लोकांच्या यावर अधिक विश्वास बसताना दिसत आहे.

एआय ही अशी तंत्रज्ञानप्रणाली आहे जी मानवी निर्णयक्षमतेसारखे कार्य करण्याचा प्रयत्न करते. परंतु, मानवी विचारप्रक्रियेसोबत मूल्यनिष्ठा, संस्कार, व सामाजिक जबाबदाऱ्या जोडलेल्या असतात. त्याउलट एआयच्या अल्गोरिदमना मिळणारे प्रशिक्षण डेटा व प्रोग्रामिंग हे माणसांनी तयार केलेले असल्याने त्यात पक्षपात, भेदभाव किंवा चुकीचे निष्कर्ष येण्याची शक्यता असते. उदाहरणार्थ, नोकरीसाठी उमेदवारांची निवड करणारे एआय मॉडेल्स जर एकतर्फी डेटावर आधारित असतील, तर ते विशिष्ट लिंग, जात किंवा सामाजिक गटांविरुद्ध भेदभाव करू शकतात.

याशिवाय, एआय प्रणालींमध्ये पारदर्शकतेचा अभाव ही एक मोठी अडचण आहे. अनेकदा एआयने दिलेला निर्णय किंवा शिफारस कशावर आधारित आहे हे सामान्य व्यक्तीला समजणे कठीण होते. यामुळे जबाबदारी निश्चित करणे अवघड होते, विशेषतः आरोग्यसेवा, न्यायव्यवस्था किंवा आर्थिक व्यवहारांसारख्या संवेदनशील क्षेत्रांत. तसेच, डेटा संकलन आणि त्याचा वापर यामध्ये व्यक्तीगत गोपनीयतेचे गंभीर प्रश्न निर्माण होतात.



या पार्श्वभूमीवर एआय नीतिशास्त्राचा अभ्यास महत्त्वाचा ठरतो. एआयच्या विकास आणि वापर प्रक्रियेत न्याय्यता, जबाबदारी, पारदर्शकता, गोपनीयता आणि मानवी मूल्यांचा सन्मान या घटकांचा विचार होणे अत्यावश्यक आहे. या संशोधन निबंधात एआय व नीतिशास्त्र यातील परस्परसंबंधांचा सखोल अभ्यास करण्यात येणार आहे. त्यासाठी आधी झालेल्या साहित्याचा आढावा, विद्यमान चौकटींचे विश्लेषण, विविध क्षेत्रांतील प्रकरण अभ्यास, तसेच भावी दिशा यांचा विचार केला जाणार आहे.

या अभ्यासाचा उद्देश केवळ एआयमुळे उद्भवणाऱ्या नैतिक प्रश्नांना अधोरेखित करून , जबाबदार व मानवी केंद्रित एआय प्रणालींच्या विकासासाठी आवश्यक शिफारसी सुचविणे हा आहे.

साहित्य समीक्षा :

कृत्रिम बुद्धिमत्ता व नीतिशास्त्र यावरील चर्चेला गेल्या दोन दशकांत जागतिक पातळीवर मोठे महत्त्व प्राप्त झाले आहे. संशोधन, तंत्रज्ञान, धोरण आणि सामाजिक विज्ञान या सर्व क्षेत्रांनी एआयच्या नीतिशास्त्रीय परिमाणांची चर्चा करित आलेले आहेत.

1. तात्त्विक पाया :

एआय नीतिशास्त्र समजण्यासाठी नैतिक तत्त्वज्ञानाची मूलभूत चौकट महत्त्वाची आहे. *उपयुक्ततावाद (Utilitarianism)*, *कर्तव्यनीती (Deontology)* आणि *सद्गुणनीती (Virtue Ethics)* हे नैतिकतेचे तीन मुख्य दृष्टिकोन आहेत. मानवी आचरणचे मूल्यमापन दोन तऱ्हेने केले जाते. एखाद्या आचारणामागील हेतू पाहून किंवा आचरणातून निर्माण होणारा परिणाम पाहून ते आचरण इष्ट कि अनिष्ट ठरवले जाते. सुखावादाच्या मते कृतीचे नैतिक मूल्य तिच्या परिणामावर आधारित असते (Bentham, 1789). कर्तव्यनीतीनुसार हेतू महत्त्वाचा असतो, हे आपले कर्तव्य आहे. स्वतः वर स्वतःने घालून घेतलेले नियम व कर्तव्ये पाळणे महत्त्वाचे ठरते. (Kant, 1785). सद्गुणनीतीत मानवी गुणधर्म व मूल्यांवर भर दिला जातो (Aristotle, 4th century BC). एआय प्रणालींमध्ये या तत्त्वांचा वापर केल्यास मानवी मूल्यांच्या अनुरूप निर्णयप्रक्रिया होऊ शकते. (Mittelstadt et al., 2016).

2. पक्षपात व न्याय्यता :

डेटा-आधारित एआय मॉडेल्समध्ये सामाजिक व ऐतिहासिक पक्षपात प्रतिबिंबित होतो. Barocas & Selbst (2016) यांनी “Big Data’s Disparate Impact” या लेखात दाखवले आहे की डेटामधील भेदभावामुळे एआयचे निर्णय असमानतेकडे झुकतात. Buolamwini & Gebru (2018) यांनी चेहर्यावरील ओळख प्रणालींमध्ये महिलांवर व रंगाच्या लोकांवर उच्च प्रमाणात त्रुटी असल्याचे स्पष्ट केले. हे संशोधन दाखवते की “न्याय्यता” हा घटक एआय डिझाइनचा अविभाज्य भाग असायला हवा.

3. पारदर्शकता व जबाबदारी :

एआयमधील सर्वात मोठे आव्हान म्हणजे “ब्लॅक-बॉक्स” स्वरूप. Lipton (2018) यांनी स्पष्ट केले



आहे की एआय निर्णय समजावून सांगता न आल्यास जबाबदारी निश्चित करणे कठीण होते. Doshi-Velez & Kim (2017) यांनी “Interpretability is essential” असे नमूद केले असून, समजण्याजोग्या मॉडेल्समुळे वापरकर्त्यांचा विश्वास वाढतो.

4. गोपनीयता व देखरेख :

डेटा ही एआयची मूलभूत गरज आहे, परंतु त्यातून गोपनीयतेचे प्रश्न निर्माण होतात. Zuboff (2019) यांनी “Surveillance Capitalism” या ग्रंथात दाखवले आहे की कंपनी वापरकर्त्यांचा डेटा व्यावसायिक फायद्यासाठी वापरतात. Floridi & Taddeo (2016) यांच्या मते, माहितीचे नीतिशास्त्र (Information Ethics) एआयच्या युगात अधिक महत्त्वाचे झाले आहे.

5. विद्यमान मार्गदर्शक तत्त्वे :

आंतरराष्ट्रीय स्तरावर अनेक चौकटी विकसित झाल्या आहेत. OECD ने 2019 मध्ये “AI Principles” प्रकाशित केले, ज्यात मानवी हक्क, पारदर्शकता, जबाबदारी, आणि शाश्वतता यांवर भर दिला आहे. UNESCO (2021) ने “Recommendation on the Ethics of Artificial Intelligence” मध्ये जागतिक सहकार्याचे आवाहन केले आहे. युरोपियन युनियनने “AI Act” (2021) मध्ये धोके-आधारित नियमन पद्धतीचा अवलंब केला आहे..

संशोधन पद्धती (Research Methodology) :

एआय व नीतिशास्त्र या विषयाचा अभ्यास करताना एकाच दृष्टिकोनाचा अवलंब केल्यास संपूर्ण समस्येची गुंतागुंत समजणे कठीण ठरते. त्यामुळे या निबंधात *संकल्पनात्मक विश्लेषण*, व *तुलनात्मक विश्लेषण* या दोन्ही दृष्टिकोनांचा एकत्रित वापर केला आहे. या पद्धतीमुळे विषयाचे व्यापक आणि सखोल विश्लेषण करता येईल.

संशोधन रचना (Research Design) :

या संशोधनाची रचना तीन मुख्य टप्प्यांत विभागली आहे:

1. **तात्त्विक विश्लेषण (Theoretical Analysis):** नीतिशास्त्रातील पारंपरिक तत्त्वज्ञान — उपयोजनावाद (Utilitarianism), कर्तव्यनीती (Deontology) आणि सद्गुणनीती (Virtue Ethics) — यांचा एआयच्या संदर्भात विचार करण्यात आला आहे. उदाहरणार्थ, एखाद्या एआय प्रणालीने घेतलेला निर्णय उपयोजनावादानुसार समाजाच्या एकूण कल्याणासाठी उपयुक्त आहे का, कर्तव्यनीतीनुसार तो नैतिक नियमांचे पालन करतो का, आणि सद्गुणनीतीनुसार तो मानवी मूल्यांशी सुसंगत आहे का, हे तपासण्यात आले आहे.
2. **तुलनात्मक चौकट (Comparative Framework) :** जागतिक पातळीवर अस्तित्वात असलेल्या



एआय नीतिशास्त्रीय चौकटींची तुलना करण्यात आली आहे. यात *EU AI Act (2021)*, *OECD AI Principles (2019)* आणि *UNESCO Recommendation on the Ethics of AI (2021)* या तीन महत्त्वाच्या चौकटींचा अभ्यास करून त्यांच्या वैशिष्ट्ये, मर्यादा व अंमलबजावणीतील आव्हाने समजावून घेण्यात आले आहेत.

माहितीचे स्रोत (Sources of Data) :

संशोधनासाठी वापरलेली माहिती ही द्वितीयक स्वरूपाची (secondary sources) असून त्यात *European Union AI Act (2021)*, *OECD Principles on AI (2019)*, *UNESCO Recommendation (2021)* या अंतरराष्ट्रीय धोरणांचाही आधार घेतला आहे. या सर्व स्रोतांचा वापर करून एक सुसंगत व तुलनात्मक दृष्टिकोन तयार करण्याचा प्रयत्न केला आहे.

संशोधनाच्या मर्यादा (Limitations) :

या संशोधनाला काही मर्यादा आहेत:

1. हे संशोधन द्वितीयक स्रोतांवर अवलंबित्व मुख्यत्वे शैक्षणिक लेख, अहवाल व धोरण दस्तऐवजांवर आधारित आहे. प्रत्यक्ष क्षेत्रीय (empirical) माहिती संकलनाचा अभाव आहे.
2. हा अभ्यास काही प्रमुख क्षेत्रांपुरते मर्यादित आहेत. कृषी, शिक्षण किंवा सामाजिक माध्यमांवरील एआय नीतिशास्त्राचा स्वतंत्र अभ्यास या निबंधाच्या व्यापात नाही.
3. आंतरराष्ट्रीय चौकटींची तुलना करताना स्थानिक सांस्कृतिक, कायदेशीर व सामाजिक संदर्भांचा संपूर्ण विचार शक्य नाही.
4. **जलद बदलणारे क्षेत्र** : एआय व त्याचे नियमन हे झपाट्याने बदलणारे क्षेत्र असल्याने, या अभ्यासातील काही निष्कर्ष पुढील काही वर्षांत कालबाह्य होऊ शकतात.

कृत्रिम बुद्धिमत्ता (AI) आज अनेक क्षेत्रांत वेगाने अंगीकारली जात आहे. मात्र, तिच्या वापरामुळे उद्भवणारे नीतिशास्त्रीय प्रश्न व धोके क्षेत्रानुसार वेगवेगळे स्वरूप धारण करतात. आरोग्य आणि रोजगार क्षेत्राचा सविस्तर आढावा घेतला आहे.

आरोग्य क्षेत्र (Healthcare) :

आरोग्यसेवेत एआयचा वापर विशेषत्वाने वाढत आहे. निदान (diagnosis), औषधनिर्मिती (drug discovery), वैद्यकीय प्रतिमांचे विश्लेषण (medical imaging analysis) आणि उपचार व्यवस्थापन (treatment management) यासाठी एआय प्रणालींचा वापर होतो. उदाहरणार्थ, *Google DeepMind* ने डोळ्यांच्या आजारांचे निदान करण्यासाठी विकसित केलेले मॉडेल अनेक डॉक्टरांइतकेच अचूक ठरले आहे (Topol, 2019).

मुख्य नीतिशास्त्रीय प्रश्न :



- **गोपनीयता व संमती (Privacy & Consent) :** रुग्णांची वैद्यकीय माहिती (health records) ही अतिशय संवेदनशील असते. एआय कंपनी ती मोठ्या प्रमाणावर वापरत असताना योग्य संमती न मिळाल्यास रुग्णांच्या गोपनीयतेचे उल्लंघन होते. यामुळे वैद्यकीय डेटाच्या सुरक्षिततेवर प्रश्नचिन्ह निर्माण होते.
- **जबाबदारी (Accountability) :** चुकीच्या निदानामुळे किंवा चुकीच्या उपचार शिफारसीमुळे रुग्णांच्या आरोग्यावर परिणाम झाला, तर जबाबदारी कोणाची? डॉक्टर, एआय प्रणाली की त्या प्रणालीची निर्मिती करणाऱ्या कंपनीची? हा प्रश्न अद्यापही स्पष्टपणे सुटलेला नाही.
- **डेटातील पक्षपात (Bias in Data):** Obermeyer et al. (2019) यांनी दाखवून दिले की अमेरिकेतील आरोग्य व्यवस्थेत वापरल्या जाणाऱ्या एका एआय मॉडेलमध्ये कृष्णवर्णीय रुग्णांविरुद्ध पक्षपात आढळतो. यामुळे एआय-आधारित आरोग्य सेवा सर्वांसाठी समान न राहता भेदभावपूर्ण ठरते.
- **मानवी संवादाचा अभाव (Loss of Human Touch):** एआय प्रणाली कार्यक्षम असल्या तरी त्यात मानवी सहानुभूतीचा अभाव असतो. रुग्णांच्या मानसिक आरोग्याच्या दृष्टीने मानवी डॉक्टरांचा सहभाग आवश्यक ठरतो.

रोजगार व श्रम (Employment and Labor) :

स्वयंचालितिकरण (automation) आणि रोबोटिक्समुळे उद्योग, सेवा आणि कार्यालयीन व्यवस्थापनामध्ये मोठे बदल होत आहेत. McKinsey Global Institute (2017) च्या अहवालानुसार, 2030 पर्यंत जगभरातील 375 दशलक्ष लोकांना आपली कौशल्ये बदलावी लागतील किंवा नवीन नोकऱ्या शोधाव्या लागतील.

मुख्य नीतिशास्त्रीय प्रश्न :

- **रोजगार गमावणे (Job Displacement) :** उत्पादन, वाहतूक आणि बँकिंग यांसारख्या क्षेत्रांत एआय प्रणाली मानवांच्या जागी कार्य करत असल्याने मोठ्या प्रमाणावर रोजगार धोक्यात येतो. यामुळे बेरोजगारी व आर्थिक विषमता वाढते.
- **कामगारांवर देखरेख (Worker Surveillance) :** काही कंपन्या एआय-आधारित साधनांचा वापर करून कर्मचारी उत्पादकता सतत मोजतात. सततच्या देखरेखीमुळे कर्मचाऱ्यांवर मानसिक दबाव येतो आणि वैयक्तिक स्वातंत्र्य कमी होते.
- **मानवी प्रतिष्ठा (Human Dignity) :** एआयवर आधारित निर्णयांमुळे कामगारांच्या कौशल्यांचे मूल्य कमी होऊ शकते. कामगार केवळ डेटा पॉइंट्स म्हणून पाहिले जाऊ लागल्यास त्यांच्या मानवी प्रतिष्ठेला धक्का बसतो.
- **नवीन कौशल्यांची गरज (Reskilling Needs):** एआयच्या युगात टिकून राहण्यासाठी कामगारांना



नवीन कौशल्ये शिकावी लागतात. मात्र, सर्व समाजघटकांना या संधी समानपणे उपलब्ध होत नाहीत, यामुळे सामाजिक असमानता वाढू शकते.

विश्लेषण व चर्चा (Analysis and Discussion) :

कृत्रिम बुद्धिमत्ता (AI) या तंत्रज्ञानाने समाजाच्या विविध क्षेत्रांत वेगाने प्रवेश केला आहे. आरोग्य, कायदा-सुरक्षा आणि रोजगार यांसारख्या क्षेत्रांतील प्रकरण अभ्यासातून दिसून येते की एआयची कार्यक्षमता व उपयुक्तता निर्विवाद आहे, परंतु त्याच वेळी अनेक नीतिशास्त्रीय व सामाजिक प्रश्न उपस्थित होतात. या विभागात या प्रश्नांचे सखोल विश्लेषण व त्यावरील चर्चा सादर केली आहे.

आरोग्य क्षेत्रातील विश्लेषण :

आरोग्य क्षेत्रात एआयचा वापर निदानाच्या अचूकतेत आणि उपचार व्यवस्थापनात मोठा फायदा देतो. तथापि, गोपनीयता, जबाबदारी आणि डेटातील पक्षपात ही गंभीर आव्हाने आहेत. **Obermeyer et al. (2019)** यांच्या अभ्यासाने दाखवून दिले की कृष्णवर्णीय रुग्णांवर वापरल्या जाणाऱ्या एआय प्रणालींमध्ये भेदभाव होतो. यावरून हे स्पष्ट होते की जर डेटासेटच पक्षपाती असेल, तर एआयचे निर्णयसुद्धा नैतिकदृष्ट्या अयोग्य ठरतात. यामुळे रुग्णांचा एआयवरील विश्वास कमी होऊ शकतो.

त्याचबरोबर, चुकीच्या निदानाची जबाबदारी निश्चित करणे अजूनही एक मोठे आव्हान आहे. पारंपरिक वैद्यकीय प्रणालीमध्ये जबाबदारी डॉक्टरांवर असते, परंतु एआयच्या सहभागामुळे जबाबदारी डॉक्टर, एआय विकसक आणि आरोग्य संस्था यांच्यात विभागली जाते. यामुळे कायदेशीर व नैतिक वाद निर्माण होतात.

रोजगार व श्रम क्षेत्रातील विश्लेषण :

स्वयंचलितीकरणामुळे उद्योग आणि सेवा क्षेत्रात उत्पादनक्षमता वाढली असली तरी बेरोजगारी व असमानता वाढण्याची शक्यता आहे. **McKinsey Global Institute (2017)** च्या अहवालानुसार, जगभरातील कोट्यवधी लोकांची नोकरी एआयमुळे धोक्यात आहे. हा प्रश्न फक्त आर्थिकच नाही तर सामाजिक न्यायाचाही आहे.

कामगारांवर एआय-आधारित देखरेख ठेवण्यामुळे वैयक्तिक स्वातंत्र्यावर गदा येते. उदाहरणार्थ, काही कंपन्या कर्मचाऱ्यांच्या कीबोर्ड व माऊस वापरावर नजर ठेवून उत्पादकता मोजतात. हे तंत्रज्ञान संस्थांना फायदा देत असले तरी कामगारांमध्ये ताण व असंतोष निर्माण करते. मानवी श्रमाचे अवमूल्यन होऊन त्यांची प्रतिष्ठा कमी होते.

या दोन्हीही क्षेत्रांत काही समान नीतिशास्त्रीय मुद्दे दिसून येतात :

- **पक्षपात (Bias) :** आरोग्य, कायदा व रोजगार या सर्वच क्षेत्रांत डेटावर आधारित पक्षपातामुळे असमानता वाढते.



- **गोपनीयतेचे उल्लंघन (Privacy Violation) :** रुग्णांची वैद्यकीय माहिती, नागरिकांचे हालचालींचे डेटा, आणि कामगारांच्या क्रियाकलापांवर देखरेख या सर्व बाबतींत गोपनीयतेचा प्रश्न गंभीर आहे.
- **जबाबदारी (Accountability) :** चुकीच्या निर्णयांची जबाबदारी निश्चित करणे हे सर्व क्षेत्रांत कठीण ठरते.
- **मानवी मूल्यांवरील गदा (Threat to Human Dignity) :** एआयमुळे मानवी प्रतिष्ठा, स्वातंत्र्य आणि समान हक्क धोक्यात येतात.

यावरून स्पष्ट होते की एआय केवळ तांत्रिक साधन नाही तर मानवी समाजाच्या नीतिशास्त्रीय संरचनेवर थेट परिणाम करणारे माध्यम आहे.

नीतिशास्त्रीय चौकट व भविष्यातील मार्ग :

सध्याच्या परिस्थितीत खालील उपाय महत्त्वाचे ठरतात:

1. **पारदर्शकता वाढवणे :** AI तंत्रांचा अवलंब करून निर्णय प्रक्रिया सर्वांसाठी स्पष्ट करणे.
2. **न्याय्यता सुनिश्चित करणे :** डेटामध्ये विविधता वाढवून व bias detection tools वापरून असमानता कमी करणे.
3. **जबाबदारी निश्चित करणे :** कायदेशीर व धोरणात्मक चौकटींमध्ये जबाबदारी स्पष्टपणे वाटप करणे.
4. **गोपनीयतेचे संरक्षण :** Privacy-preserving algorithms जसे Differential Privacy आणि Federated Learning वापरणे.
5. **मानवी मूल्यांचा सन्मान :** एआयच्या डिझाइनमध्ये मानवी हक्क, प्रतिष्ठा व सामाजिक न्याय यांना प्राधान्य देणे.

सुचविलेले नीतिशास्त्रीय मार्गदर्शक तत्त्वे (Proposed Ethical Framework) :

कृत्रिम बुद्धिमत्तेच्या (AI) झपाट्याने होणाऱ्या विकासामुळे समाजासाठी मोठ्या संधी उपलब्ध होत आहेत, परंतु त्याचबरोबर नैतिकतेचे प्रश्नही गंभीर स्वरूप धारण करत आहेत. म्हणूनच, एआयचा वापर करताना काही ठोस नीतिशास्त्रीय तत्त्वे निश्चित करणे आवश्यक आहे. खालील उपबिंदूंमध्ये या तत्त्वांचे विवेचन केले आहे.

1. न्याय्यता (Fairness) :

एआय प्रणालींच्या निर्णय प्रक्रियेत न्याय्यता हा घटक अत्यंत महत्त्वाचा आहे. एआय मॉडेल्सना प्रशिक्षण देण्यासाठी वापरला जाणारा डेटा जर पक्षपाती असेल, तर त्यांचे निष्कर्षदेखील भेदभावपूर्ण ठरू शकतात. उदाहरणार्थ, एखाद्या नोकरीसाठी उमेदवारांची निवड करणारी एआय प्रणाली जर फक्त पुरुष उमेदवारांच्या डेटावर आधारित असेल, तर ती स्त्री उमेदवारांना अन्यायकारकपणे नाकारू शकते. त्यामुळे डेटा संकलन करताना विविधता आणि प्रतिनिधित्व सुनिश्चित करणे आवश्यक आहे. याशिवाय, *bias detection tools* व *fairness metrics* यांचा वापर करून एआय प्रणालींची नियमित चाचणी घेणे गरजेचे आहे.



2. जबाबदारी (Accountability) :

एआयच्या संदर्भात जबाबदारी हा अत्यंत संवेदनशील मुद्दा आहे. जर एखाद्या एआय प्रणालीच्या निर्णयामुळे व्यक्तीला किंवा समाजाला हानी पोहोचली, तर त्याची जबाबदारी कोणावर टाकायची हा प्रश्न महत्त्वाचा ठरतो. यासाठी बहुस्तरीय जबाबदारी (*multi-level accountability*) प्रणालीचा अवलंब करणे आवश्यक आहे, ज्यामध्ये विकसक, वापरकर्ते आणि संस्था या सर्वांनी आपापल्या भूमिकांची जबाबदारी घ्यावी. चुकीचे परिणाम झाल्यास तक्रार निवारणासाठी स्वतंत्र *redress mechanisms* असणे गरजेचे आहे. तसेच, कायदेशीर चौकटींमध्ये एआयशी संबंधित जबाबदारी स्पष्टपणे नमूद केली गेली पाहिजे.

3. पारदर्शकता व समजण्यायोग्यता (Transparency and Explainability) :

एआय प्रणालींवर समाजाचा विश्वास निर्माण करण्यासाठी पारदर्शकता हा अत्यावश्यक घटक आहे. सध्या बहुतेक एआय मॉडेल्स “ब्लॅक-बॉक्स” स्वरूपात काम करतात, ज्यामुळे त्यांचे निर्णय कसे झाले हे समजणे कठीण होते. *Explainable AI (XAI)* या दृष्टिकोनामुळे वापरकर्त्यांना निर्णयामागील प्रक्रिया समजण्यास मदत होते. उदाहरणार्थ, वैद्यकीय क्षेत्रात एआयने दिलेला निदान अहवाल जर डॉक्टरांना समजावून सांगता आला नाही, तर तो स्वीकारार्ह ठरणार नाही. त्यामुळे पारदर्शकतेसाठी तांत्रिक उपायांसोबतच स्वतंत्र *third-party audits* देखील आवश्यक आहेत

4. गोपनीयता व डेटा संरक्षण (Privacy and Data Protection) :

एआयच्या कार्यक्षमतेचा पाया डेटा आहे, परंतु त्याचबरोबर व्यक्तीगत गोपनीयतेचे गंभीर प्रश्न निर्माण होतात. व्यक्तीची माहिती त्यांच्या संमतीशिवाय वापरणे हे नैतिकदृष्ट्या अयोग्य आहे. म्हणूनच, डेटाच्या संकलन, साठवणूक व वापरसाठी व्यक्तीची *informed consent* घेणे बंधनकारक असले पाहिजे. याशिवाय, *differential privacy* किंवा *federated learning* सारख्या गोपनीयता संरक्षक अल्गोरिदमसचा अवलंब करणे आवश्यक आहे. आंतरराष्ट्रीय पातळीवर समान डेटा सुरक्षा मानके लागू केली गेल्यास माहितीच्या गैरवापरास आळा घालता येईल.

5. मानवी-केंद्रित दृष्टिकोन (Human-Centric Approach) :

एआयचा विकास हा केवळ तांत्रिक कार्यक्षमतेसाठी नसून मानवी कल्याणासाठी असायला हवा. एआय प्रणालींनी मानवी मूल्ये, प्रतिष्ठा आणि मानवी हक्क यांचा आदर राखला पाहिजे. महत्त्वाच्या निर्णय प्रक्रियेत *human-in-the-loop* दृष्टिकोन ठेवून मानवी हस्तक्षेप आवश्यक आहे. उदाहरणार्थ, न्यायव्यवस्थेतील किंवा आरोग्यसेवेतील गंभीर निर्णय एआयवर पूर्णपणे सोपवले जाऊ नयेत. याशिवाय, एआयने कामगारांची प्रतिष्ठा कमी न करता त्यांच्या क्षमतांचा विकास साधावा आणि सामाजिक न्यायाला प्रोत्साहन द्यावे.

6. आंतरराष्ट्रीय सहकार्य व स्थानिक संदर्भ (Global Cooperation and Local Context) :



एआय ही जागतिक पातळीवर परिणाम करणारी तंत्रज्ञानप्रणाली आहे, त्यामुळे तिच्या नीतिशास्त्रीय चौकटींसाठी आंतरराष्ट्रीय सहकार्य आवश्यक आहे. UNESCO (2021) च्या अहवाला नुसार जागतिक सहकार्यास महत्त्व देण्यात आले आहे. तरीही, प्रत्येक देशाचे कायदेशीर, सांस्कृतिक व सामाजिक संदर्भ वेगळे असतात. त्यामुळे जागतिक पातळीवरील *harmonized standards* विकसित करताना स्थानिक संदर्भांचा आदर राखणे महत्त्वाचे आहे. उदाहरणार्थ, डेटा संरक्षणाच्या बाबतीत युरोपमधील GDPR कठोर आहे, तर इतर देशांमध्ये तुलनेने लवचिक नियम आहेत. म्हणूनच, जागतिक मार्गदर्शक तत्वांसोबत स्थानिक लवचिकता आवश्यक आहे.

निष्कर्ष (Conclusion) :

कृत्रिम बुद्धिमत्ता (AI) ही आधुनिक विज्ञान आणि तंत्रज्ञानाची अभूतपूर्व देणगी आहे. तिच्या वापरामुळे आरोग्यसेवा, उद्योग, रोजगार, शिक्षण इत्यादी क्षेत्रांत कार्यक्षमता, अचूकता आणि सोयीसुविधा मोठ्या प्रमाणावर वाढल्या आहेत. तथापि, या अभ्यास आणि विश्लेषणातून हे स्पष्ट होते की एआयच्या वेगवान प्रगतीसोबतच गंभीर नीतिशास्त्रीय प्रश्न उद्भवले आहेत.

आरोग्य क्षेत्रात रुग्णांच्या गोपनीयतेचे संरक्षण, चुकीच्या निदानाची जबाबदारी आणि डेटातील पक्षपात ही प्रमुख आव्हाने आहेत. त्यामुळे विशिष्ट गटांविषयी भेदभाव वाढू शकतो. रोजगार क्षेत्रात एआय-आधारित स्वयंचालितकरणामुळे बेरोजगारी, कामगारांवर देखरेख आणि मानवी प्रतिष्ठेचे अवमूल्यन यांसारख्या समस्या निर्माण होत आहेत. या सर्व क्षेत्रांतील सामाईक मुद्दे म्हणजे पक्षपात, गोपनीयतेचे उल्लंघन, जबाबदारीची अनिश्चितता आणि मानवी मूल्यांवरील गदा.

या पार्श्वभूमीवर एआयच्या विकास व वापरासाठी ठोस नीतिशास्त्रीय चौकट आवश्यक आहे. न्याय्यता, जबाबदारी, पारदर्शकता, गोपनीयता आणि मानवी-केंद्रित दृष्टिकोन ही पाच मूलभूत तत्त्वे प्रत्येक एआय प्रणालीमध्ये समाविष्ट केली गेली पाहिजेत. तसेच, आंतरराष्ट्रीय सहकार्य आणि स्थानिक संदर्भांचा आदर राखत जागतिक मानके विकसित करणे ही काळाची गरज आहे. शेवटी निष्कर्ष काढता येईल की, एआय ही समाजाच्या विकासाला गती देणारी प्रगत तंत्रज्ञानप्रणाली असली तरी तिचा योग्य, जबाबदार आणि नीतिशास्त्राधिष्ठित वापर झाल्यासच ती मानवतेच्या कल्याणासाठी परिणामकारक ठरेल. म्हणूनच, वैज्ञानिक, तंत्रज्ञ, धोरणकर्ते, कायदेतज्ज्ञ आणि नागरिक यांनी एकत्र येऊन एआयला *मानवकेंद्रित, सुरक्षित आणि न्याय्य* बनविण्यासाठी सहकार्य करणे आवश्यक आहे.

संदर्भ (References) :

- बारोकस, एस., & सेल्बस्ट, ए. डी. (2016). बिग डेटा'स डिस्पैरेट इम्पॅक्ट. *कॅलिफोर्निया लॉ रिव्ह्यू*, 104(3), 671-732.
- बेंथम, जे. (1789). *अन इंट्रोडक्शन टू द प्रिन्सिपल्स ऑफ मॉरल्स अँड लेजिस्लेशन*. लंडन: टी. पेन अँड



सन

- फ्लोरिडी, एल., & टॅडिओ, एम. (2016). डेटा एथिक्स म्हणजे काय? *फिलॉसॉफिकल ट्रान्झॅक्शन्स ऑफ द रॉयल सोसायटी A: मॅथेमॅटिकल, फिजिकल अँड इंजिनिअरिंग सायन्सेस*, 374(2083), 20160360.
- कान्ट, आय. (1785). *ग्राउंडवर्क ऑफ द मेटाफिजिक्स ऑफ मॉरल्स*. केंब्रिज युनिव्हर्सिटी प्रेस.
- लिप्टन, झेड. सी. (2018). द मिथोस ऑफ मॉडेल इंटरप्रीटेबिलिटी. *कम्युनिकेशन्स ऑफ द ACM*, 61(10), 36–43.
- मिटेलस्टॅड्ट, बी. डी., अल्लो, पी., टॅडिओ, एम., वॉचर, एस., & फ्लोरिडी, एल. (2016). द एथिक्स ऑफ अल्गोरिदम्स: मॅपिंग द डिबेट. *बिग डेटा अँड सोसायटी*, 3(2).
- युनेस्को. (2021). *रिकमेंडेशन ऑन द एथिक्स ऑफ आर्टिफिशियल इंटेलिजन्स*. पॅरिस: युनेस्को.
- रायट, डी., & शुल्ट्झ, ई. (2018). एआय आणि स्वायत्त प्रणालींच्या नैतिक परिमाण. *माईंड्स अँड मशीन्स*, 28(4), 1–21.

