



A Dharma Perspective on Artificial Intelligence

by Rajiv Malhotra | Apr 8, 2021 | Insights

Artificial Intelligence has become the foundation for what is called the Fourth Industrial Revolution. No doubt, this will bring amazing benefits to humanity, just as the previous industrial revolutions did, and much has been written on this aspect, almost portraying AI as a sort of Holy Grail. However, one must balance such utopian views with serious issues and challenges that the new technology brings in its wake.

Many of these challenges are generic and applicable to all traditions and faiths across the board – such as the loss of jobs in several categories. But there are also faith-specific challenges that need to be on the table for conversations with mutual respect. At the same time, there are also likely to be faith-specific *positive* consequences of AI.

Rather than the populist “all religions are the same” approach, it is far more honest to acknowledge the diversity of faiths and relate to each other with mutual respect. The worldview of each tradition is the result of its subjective experience, and none can claim the universalism that would justify its imposition on others. This article adopts the posture of *difference with mutual respect* among faiths.

With this spirit, I am adopting the Dharma lens for discussing the social impact of AI. I will show that Artificial Intelligence is far from being neutral in its treatment of all faiths. There are two major points that make the impact on Dharma fundamentally different from the corresponding impact on the Abrahamic religions: the absence of a single historical book given the same kind of absolute status as the Bible or Qur’an, and the deep impact of colonization.

The Absence of a Single Book

Dharma is not based on one book that is historically unique, but on the human experiences of exemplars in higher states of consciousness. This has both advantages and disadvantages. A benefit is that such a tradition does not get frozen in time and boxed into the rigid walls of some dogma that is historically fixated. Since history cannot be altered, any absolute truth ground in a historical singularity becomes inflexible. Dharma is based not on one book but a library of works from individuals who attained higher states – like the Buddha and the rishis. This gives it much more room to evolve, adapt and negotiate its worldview in each new epoch.

However, a serious pragmatic disadvantage is that in the absence of the one book-based absolutism, Dharma can become fragmented into numerous groups with divergent claims of truth. This disadvantage turns into vulnerability in the face of AI, because algorithms can be trained by their owners depending on the social-political preferences they wish to proliferate. I am often asked, “Which version of Hindu Dharma do you want to advocate?” Even though I have explained what unifies all Dharma in considerable detail in my books, this can challenge the masses and AI can be used to create a wedge among the followers of Hindu Dharma by training the AI algorithms using different texts and interpretations.

The Deep Impact of Colonialization on Dharma Training Texts

The second issue concerns the impact of a thousand years of colonization, and this deserves detailed treatment. Algorithms are fundamentally biased because of the way they are trained. An analogy with the education of a child is useful. The training of a child depends on the type of exposure from parents, school, media and the type of reading it does. What the child is taught as true or false includes many subjective aspects and these shape its belief systems and worldviews. Likewise, in the case of artificial intelligence, the data used to train is especially important in shaping what the algorithm considers as right and wrong, and a worthy outcome. African Americans have complained that there is bias against them in many of the AI systems because these were trained on Whites considered normal and Blacks as something of an aberration from the norm. Similarly, feminists and LGBTQA people have complained that their members are considered inferior to males and straight people, respectively. This is because the datasets used for training algorithms have been dominated by people of certain types and excluded others.

It is important to examine the same issues of bias concerning faith-based groups. For instance, I am unaware of any training of algorithms using texts and interpretations from the Dharma traditions. Wikipedia is a common dataset for training algorithms because it is assumed to be the result of a democratic process in which all members of the public participate equally. But this is simply not true. Wikipedia's editors have a hierarchy of privileges, and many of its entries concerning Dharma are outright biased because they are based on Western Indology from the colonial era. When a machine is trained using Wikipedia as a primary dataset, it contains the same biases as Wikipedia.

Because of the colonial past, a large number of Indian texts were translated and interpreted under the supervision of rulers who were outside the Dharma tradition. Initially, the influence of Muslims was felt in the public discourse in India. But the foreign influence was much heavier during the colonial times when Europeans (who called themselves Orientalists), started the systematic interpretation of Sanskrit texts, partly because they assumed Sanskrit was the heritage of European languages by the so-called Aryans. Therefore, from the late 1700s and throughout the 1800s, the study of Sanskrit and its seminal texts became a thriving academic discipline across Europe. The effect of this was lasting: Almost all present day works on Dharma texts in English are based on frameworks established by scholars who were not practitioners but

outsiders. Sometimes they had vested interests and motives, and at other times their biases became incorporated innocently and unconsciously in their works. Regardless of the causes, the effect has been that the corpus of Western Indological works today appear unacceptable to most experts from within the Dharma tradition. In fact, decolonizing India studies and Indology is a major movement of Hindus today (for which they often get branded as nationalists).

Thus, whether you use Wikipedia or any Indological work for training an algorithm, the effect can be biased. This can range from outright nonsensical and blatant falsehoods to more subtle and nuanced prejudices.

A good example of such academically generated biases would be the interpretation of the word *swastika*. This is an old term from classical Sanskrit that has nothing to do with its modern distortion resulting from the unfortunate misappropriation by the Nazis. In the Dharma tradition, swastika is a highly revered term found in Hindu, Buddhist, Jain, and Sikh traditions and found in houses of worship in all these traditions. Swastika refers to auspiciousness, good luck, positive wishes especially for a new beginning, and has nothing to do with the horrible interpretation given by the Nazis. But when an algorithm is trained using Western writings, the swastika becomes associated with violence and fascism. Posts containing the word or image of swastika often get banned. However, if an algorithm was trained in India, it would notice that the swastika is publicly displayed on shops, trucks, billboards, and used as a brand on products from matchboxes to other consumer goods.

Related to the word swastika is the word Aryan. It too has been misinterpreted by western Indology. *There are no Aryan people as a race*. In fact, there is no such word in Sanskrit. The correct Sanskrit word is *arya*, without the 'n'. *Arya* is an adjective for nobility, and describes people who are courageous, possess good character and high standards. Anyone can be an *arya* regardless of race or faith. The error is equivalent to interpreting 'tennis player' as a race, instead of a description of anyone who plays that game. The mistranslation of the term *arya* (noble quality) to *Aryan* (race) was to solve serious identity problems in Germany. Thus, a false theory was formulated that Sanskrit started somewhere close to Europe, presumably somewhere north of the Caspian Sea. The people who developed Sanskrit, according to this theory, were Aryans because the word *Arya* appeared in Sanskrit texts. The theory became that these Aryans were a race and they spread Sanskrit and it became adapted into German and other European languages. Though

the pure form went to India, it was argued that the people of India were inferior because the Aryan invaders made the mistake of marrying the dark-skinned natives.

None of this makes any sense, nor is supported by any scientific evidence. There has never been any Aryan invasion into India contrary to the claims of Western Indology that culminated in Nazi scholarship. The British adopted this Aryan invasion theory because it suited their vested interests to divide Indians into those of the north (declared of Aryan origin) and a different race in the south of India called Dravidians. This helped their own divide-and-rule policies. Indian scholars have never accepted these theories. The problem of Aryan seen as racism has entered AI because algorithms are trained on texts and material based on the Aryan invasion theory. This continues to exacerbate the Aryan/Dravidian divisiveness as the racial substratum of Indian society.

The point I am making can be generalized to literally hundreds of Sanskrit words which have been misinterpreted. This is why I wrote a book called *Sanskrit Non Translatable*, with the purpose of explaining that certain Sanskrit words cannot be replaced in English because of the bias inherent in their translation into English.

While it appears that all languages and faiths suffer in the same way, the problem is particularly acute with respect to the Dharma traditions because of the colonial past. The Jews, despite long periods of persecution, always managed to control the scholarship of their tradition because it remained securely in the hands of their own authorities and exemplars. The Jewish diaspora took their own internal version to every corner of the world where Jews were practicing. Likewise, the Christians were never colonized to the extent of losing control of their own scholarship, in the same sense as the Dharma traditions were. The Muslims have also enjoyed the continuity of scholarship of their own tradition under their own rulers.

The same cannot be said of Hinduism, Buddhism, Sikhism or Jainism. These traditions suffered at the hands of invasions and conquests from foreign armies and empires that represented alien faiths. For the past 1000 years, large parts of India, including its main sacred sites and learning centers, were under the rule of non-Dharmic rulers, either Islamic or Christian. Only in the past 75 years, India has gained independence and been able to appoint its own rulers from within. This means that the custodians of the houses of

worship, schools of learning, practices and rituals, and interpretations of texts were largely in the hands of foreigners for a period of centuries.

AI's Impact on Human Agency

A serious AI related issue affecting all faiths is that as machines are getting smarter, people are getting relatively dumber because they are more dependent on machines. People are putting themselves on autopilot and relying on machines to tell them what is right and true. Wikipedia, Google search and the algorithms driving social media are defining the worldviews and ideas of truth. This amplifies their biases because knowledge is being crowd-sourced, and the relative popularity of opinions has become the criteria for deciding what to believe. And social media controlled by algorithms determines the popularity of different views. At the end of the day, artificial intelligence drives the algorithms, which in turn drive social media, and this ends up shaping popular beliefs, “facts”, and so on. From outright fake news all the way to so-called objective claims, there is nothing that has escaped the power of algorithms in shaping our minds.

When you combine all this with the trend of reduced attention span, superficial learning and the obsession to chase popularity, the result has been an impact on human agency. People are less grounded in knowledge based on the rigor of studies from authoritative sources and more likely to follow the bandwagons on social media. Traditional texts and exemplars including from various faiths, have been side-lined by popularity and power plays. Those faiths like the Dharma traditions that have fragmented structures of traditional authority and weak institutions are far more vulnerable to this onslaught of AI.

Furthermore, a lot of money is being made by modelling human psychology and decision making and selling these models for the benefit of advertisers. In fact, the multi trillion-dollar digital economy runs in large parts on advertising revenue, which looks for algorithms that can predict the behavior of different kinds of individuals and develops messages to achieve the desired outcomes. Therefore, human agency is being replaced by algorithmic power over decision making and this is resulting in a breakdown of faith and other systems that traditionally defined how people see the world and make choices. I have called this dumbing-down syndrome the ‘moronization’ of the masses. The figure below from my book, *Artificial Intelligence and the Future of Power*, illustrates this.

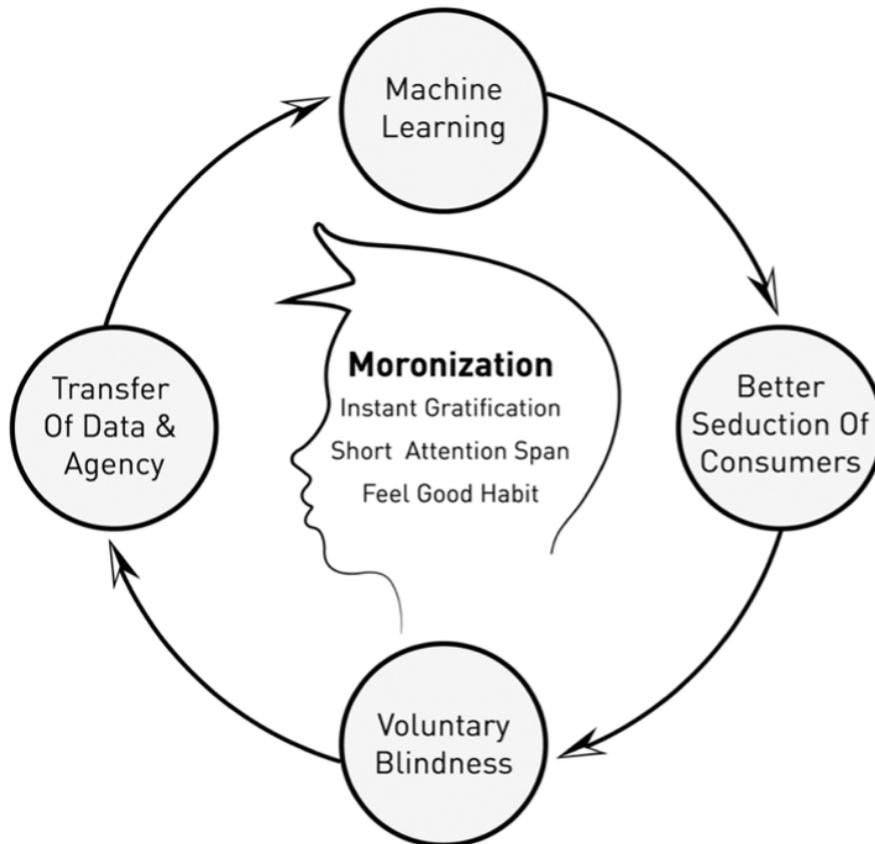


Figure 1: Moronization of the Masses

AI's Impact on Selfhood

But the problem does not end here. I claim that there is a deeper battle for the very notion and experience of selfhood. In all the Dharma traditions, the ultimate Self is divine, but this divinity is masked by the ego masquerading as if it were the Self. The spiritual process and purpose of faith is to rediscover the divinity within oneself, which is called sat-chit-ananda. Many practices, including those of meditation, rituals, lifestyle choices, diet, worship etc. focus on this, but, unlike the Abrahamic traditions, the quest is not for an external god in the sky but for the presence of divinity within oneself. This means that the ultimate bliss one is seeking lies inside and one is advised to give up external sensory pleasures and gratifications based on the body.

Artificial Intelligence takes a person in the opposite direction from the Dharma traditional practices. It offers seductive sensations, gratifications, pleasures using augmented reality and virtual reality, different kinds of wearables and devices and eventually even implants which produce the right kind of hormones, neurological experiences etc. It is true that these advancements

will also solve many mental and emotional health problems by diverting people from negative thoughts such as suicide and depression and replacing them with positive thoughts and memories artificially. However, such AI applications will not stop at merely solving mental health issues.

These very technologies are also the future of the entertainment industry, where virtual realities will take one to imaginary experiences, replacing real life. This means that, rather than the quest for selfhood within as advocated by the Dharma traditions, the push being given by these AI products and services will take the person away from inner quest and make them more extroverted towards externally supplied artificial pleasures. Worse still, some of these AI services will be controlled by others and not the individual. The lure will be too much to resist, and the result will be a two-tiered humanity: colonizers and colonized. The figure below from my book captures this changing face of identity.

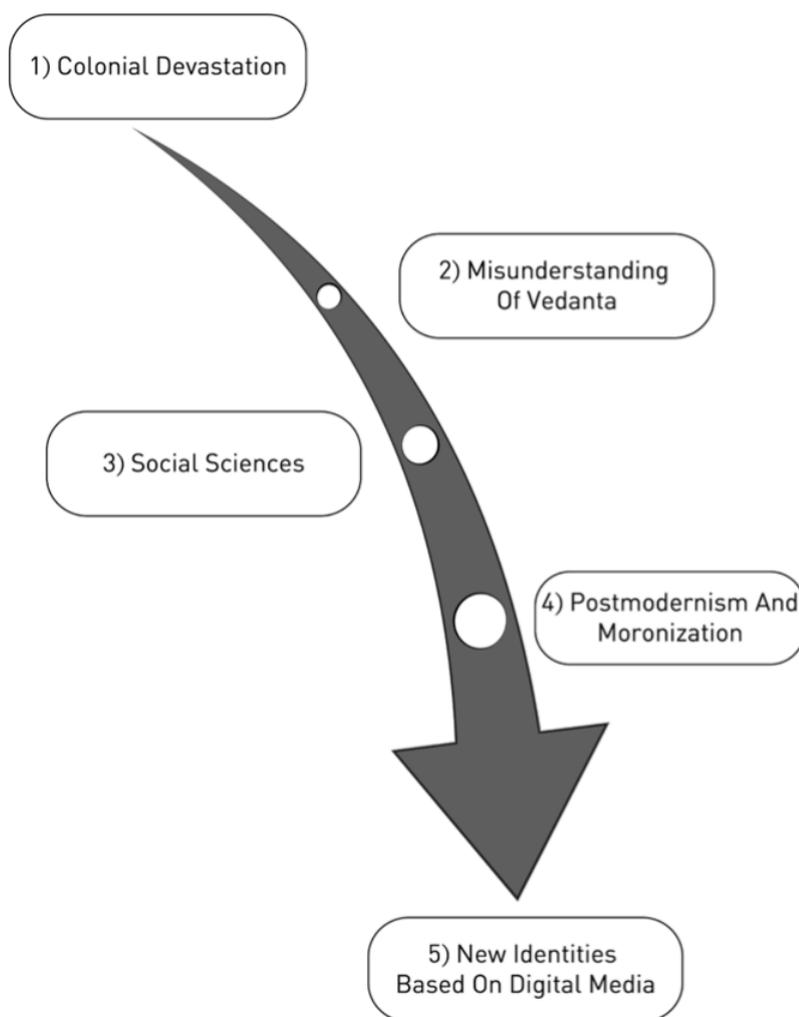


Figure 2: Digital Identity

Thanks to the success of AI, the human being is being modelled by biology as a collection of algorithms. Each organ is a complex network of algorithms and parts can be broken into smaller units which are also mere algorithms. There is thus a web of algorithms that comprise the living creature. This means that, artificially one can create new algorithms to replace the older ones or interact and modify the older algorithms. This is the cutting edge of AI, where biology and neuroscience meet computer science. I refer to this battleground as Algorithm v/s Being.

The Dharma tradition considers the supreme being as having manifested as the living cosmos including all its creatures. Each of us is part of the supreme being and nothing is different in essence or separable from the supreme being. However, the success of artificial intelligence is a boost to the biological, materialistic, mechanical model of the human. The metaphysical debate between the side considering consciousness as primary and the opposing side of biological materialism as the ultimate reality of living creatures, is now being tilted in favor of biological materialism.

Prior to AI's recent successes, there was a thriving consciousness movement that had arrived in the West since the 1960s mainly from Hindu and Buddhist yogis and gurus, including various Hindu teachers and the Dalai Lama. They brought the meditation tradition and other affiliated lifestyles as a part of what has now become a well-established consciousness evolution movement in the West. This movement is all about the Being evolving through us.

Competing against this is the algorithm marching forward, which is a mechanistic, silicon-based intelligence that can augment, enhance and ultimately substitute the being. This algorithm v/s the being is the battle for self. While it affects all faith traditions because they support the idea of soul as opposed to mere biology being the basis for the body, it is particularly acute in the case of the Dharma traditions because the quest for self-realization as opposed to God realization is given considerable emphasis in these traditions.

Issues in Common with Other Faiths

There are many other issues concerning AI and the faith traditions, which are already part of the conversation in most circles. For instance: Will AI take away jobs of those who are vulnerable? Will it make the rich richer and the poor poorer? Will certain weaker economies and societies that are unable to jump on the AI bandwagon become colonized by countries such as the USA

and China in the same manner as the industrial revolution empowered England and France to become colonizers and reduced many countries to colonies? These are serious issues which bring all faiths together in making sure that the less empowered people have a seat at the table.

There are other issues challenging the faith leaders: Will AI based algorithms shape adjudication of disputes and legal proceedings, and if so, how would fairness be guaranteed towards those who are not technologically sophisticated or wealthy or powerful enough to be able to understand and much less negotiate the adverse impact of AI on them?

One partial solution is to learn from other areas of success in attaining fairness. This is about the pharma industry's use of plants from poor countries to discover new drugs. Plant products sourced from places like India, Africa and Latin America became the basis for patents claimed by the world's richest pharmaceutical companies, and similarly the diversity of data found in various parts of the world is the new source of wealth feeding the inexhaustible appetite of the AI giants today. In the case of the pharma industry, after decades of exploitation of the biological resources of the developing world, there were finally some UN laws passed that give a semblance of protection to the countries whose biological products and resources are harvested for modern medicine.

The same kind of discussion has not even started in the case of harvesting of big data. While there is a lot of discussion on issues like privacy and data rights, what has not been given special attention is the way data from poor countries is being sucked into the giant AI machines of the rich countries. This data is being turned into patented products which are sold back to the developing countries, whose data was used to create these products in the first place. Also, the control over algorithms lies with the powerful nations because the poor countries don't have the knowledge or the hardware power to take control of AI in their own hands even if they had the wherewithal to do so.

I fear a world in which the AI enabled societies, both as nations and individuals, will control the rest of humanity and turn them into low level workers and consumers, by keeping them intoxicated and addicted to these new technologies. The time has come for people of all faiths to discuss their separate and shared concerns. We should forge an alliance of faiths to build faith-based principles that must be incorporated in any AI ethics going forward.



Rajiv Malhotra

is the founder of Princeton, New Jersey-based Infinity Foundation and a public intellectual who combines his background in physics and computer science with a quarter century of research on India's historical and future place in the world. His work covers a broad spectrum including mind sciences, civilizations, and geopolitics.