

How LAMS should be designed in order to conform with the principles of Trustworthy AI?

By Salman Khan.
Due Date: 26th Oct 2020

What is Trustworthy AI?

Artificial intelligence is a field of computer science in which machines or software are coded in such a way that they can predict or perform tasks without the intervention of humans. AI designed tools use previous data for its operations. Trustworthy AI is an AI which fulfill the following conditions throughout the system's life cycle:

- (1) it should be lawful, complying with all applicable laws and regulations
- (2) it should be ethical, ensuring adherence to ethical principles and values and
- (3) it should be robust, both from a technical and social perspective since.

Since individually the aforementioned conditions are necessary for Trustworthy operations of AI but not self-sufficient. We need a system where all three components are fulfilled and make one harmonic whole.

Each of these components are discussed below in details.

Is AI unethical?

Like any other piece of modern technology-AI's ethical or unethical behavior is dependent on its usage. Though it is commonly perceived that AI must be biased since it is dependent/trained on data that is limited and may be biased. But unless system is used to achieve some unethical goal; AI systems by their nature are not unethical even if they are biased. The end application and situations created if intentionally are not unethical; AI system in their own self are never unethical.

Are AI ethics important?

Yes, AI ethics are crucially important given enormous power of AI systems. There are many cases where unethical usage of AI has resulted in enormous real-life consequences. Use of AI systems for surveillance

on minority/persecuted communities in some states resembles fictitious Orwellian societies. Future of the modern civilization is dependent on ethical usage of AI and as such it is necessary to develop enforceable ethical codes to protect individual, collective and businesses interests.

What is LAMS?

Learning Activity Management System is an online teaching platform used by most of the institutes in Europe. It is a free as well as open source platform and one can design and manage it as per their needs.

It consists of two main part for end users:

- 1) Teaching Part, who create tasks and other class activities.
- 2) And a student side, where they can access the tasks and activities created by their teachers.

How LAMS should be designed according to Trustworthy AI?

Learning activity management system that we are currently using is not an AI tool. It is an easy tool or website mostly used by younger students. We need to make it available for the use of higher studies students, but it will make the system complex and hard to handle. So, the solution to make it consistently good in quality or performance we need to use AI and it must be trustworthy AI.

1 - Legitimate AI Based LAMS:

AI is the future but some of the past consequences of AI system have not boded well with public. And as such public is very skeptical about the usage of AI. In some cases, people have faced gender and racial discrimination as result of unsound AI algorithms. In other cases, AI systems were weaponized by political campaigns and adverse states to influence elections and exploit racial/communal divides. Therefore, we

must be wary of limitations of AI systems and keep data protection etc. laws in mind when applying AI to LAMS.

European Union is actively extending legal frame works to AI systems to regulate them. Whereas United States is more lenient when it comes to regulating AI.

2 - Ethical LAMS:

Ethical LAMS is necessary to preserve academic honesty and fairness. To apply AI to LAMS system is to provide a degree of autonomy in LAMS operation and it vastly increases the responsibility of LAMS to be ethically sound. Major ethical challenge for an AI LAMS is its impact on end users. They should be treated in fair and objectively unbiased way. Their data should be preserved and protected at any cost. And there must be an inbuilt accountability process to check any unfair practice or consequence.

- The principle of respect for human autonomy: Learning process is by nature a very interactive process. Rigidity in learning systems can kill creativity and innovation. AI LAMS by its very nature will be limited by its algorithm and data. It is crucial to respect student's agency of autonomy and not rigidly restrict them. AI LAMS should be flexible and give a very high degree of autonomy to students in learning process and methods of assessment.
- The principle of prevention of harm: Since AI LAMS will have access to student data-it is very necessary that their data must not fall into harmful hands. AI LAMS will have a high degree of autonomy in grading students and assessing their performances. Any glitch in AI LAMS can have far consequences on the career of students and therefore it is necessary that they are prevented from any bias in system. Grades and exam papers are very private for students and

as such it is very necessary that their data is encrypted and secure from hackers.

- The principle of fairness: It is important for platform like LAMS to perform its operations fairly. We have seen some of the examples in past where AI was not working fine and was biased. There was a case, where AI image processing was not of sound quality and was identifying some innocent people as criminal based on their appearance. LAMS will be using the exam papers of the student for making the result and it may use camera to keep check on students during exams. So, it is very important to use high quality data for training AI for LAMS to avoid such incidents in future.
- The principle of explicability: AI LAMS will owe a vast responsibility to students. Students are normally used to challenging their grades and other assessment results to their teacher. Teachers abiding to fairness code are transparent to students and explain everything. AI LAMS can in no way be a “black-box Algorithm” since explicability is the foundation of fair assessment. AI LAMS has to explain assessment criteria, mechanism etc. to students in advance and bring them all on a fair footing. And in the end must be auditable to satisfy any queries of the students.

3 - Requirements of Trustworthy LAMS:

A Trustworthy AI LAMS has to fulfill a number of general criteria and above all must meet all of them in one harmonic operation.

Transparent, Just and Privacy.

AI LAMS will owe it to students to be accountable and transparent to them. It should treat them justly and not violate any privacy or their human autonomy/dignity laws. Their data must be protected at every cost.

Respecting Diversity, non-discrimination, and fairness.

Since diverse body of students will make AI LAMS end users. Students must not be discriminated due to their gender/race/beliefs/orientation etc. There have been cases in past where AI system profiled people of particular skin color or discriminated based on gender or first name. The designer of AI LAMS must be very careful about this aspect of AI LAMS.

Human agency and oversight:

AI LAMS should respect human agency. It should be flexible enough to accommodate student needs and diverse demands. Above all there should be teacher oversight on AI LAMS to check any limitations of its operation or assess its performance.

Accountability:

AI LAMS will be affecting future of students and as such they have fundamental right of inquiring about the way they have been judged. AI LAMS therefore will be accountable to students and also to regulatory bodies to see it is not violating any laws.

Technical Robustness:

Since AI LAMS is expected to handle very large swath of data and be used by hundreds of thousands of students over a very vast geographical divide; it is crucial that AI LAMS is technically very robust. It must be able to handle all the data without crushing and cater for the needs of all the users-including those who may not have access to latest devices or very fast internet connections. To equally treat everyone AI LAMS should be designed in a way that it caters for the need of least advantaged student.

Conclusion:

AI LAMS is needed to free teachers/scholars from repetitive and unrewarding student assessment process and focus their energies on teaching and solving new problems. But since AI LAMS will have serious consequences for the future of all students; it is very necessary that it fulfills all the requirements of a technically robust and ethically sound AI system. It should be fair, secure, accountable, and technically robust enough to equally serve everyone.

References:

<https://dataflog.com/read/we-need-ethical-ai-5-initiatives-ensure-ethical-ai/7571>

https://www.researchgate.net/figure/Pre-service-teachers-perceptions-of-the-limitations-of-LAMS_fig3_256757434

<https://www.forbes.com/sites/cognitiveworld/2020/02/20/ai-laws-are-coming/#afd6e17a2b48>