ChatGPT and the impact on education

OpenAI has trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. In this cycle we explore the impact of ChatGTP on society and specifically on (higher) education.

Created by: FHICT
Created on: January 5, 2023 9:26 AM
Changed on: January 12, 2023 7:49 AM

Context of use: Other

Technology Impact Cycle Tool

ChatGPT and the impact on education

Impact on society
What impact is expected from your technology?

What is exactly the problem? Is it really a problem? Are you sure?
Untill now people could go to Google to ask a question and get a lot of results to choose from. But ChatGPT will change this concept forever by giving a direct and automated answer to any question (including coding questions!) directly, based on a large language model that the model is trained upon. This is better than Google because it limits choices, it curates the answer and it is a system that is natural to interact with. But does it really solve a problem?

Is it really a problem that people get too many choices when they type something into Google. Isn't it good to get multiple choices? Wasn't it already dubious that Google increasingly gave an immediate answer? The big question that remains with ChatGPT is: what problem does this actually solve? It is not clear about that.

A few ideas:
The world is too linguistic and therefore unfair. People who can talk well (who can not bullshit) already have the advantage. By giving the people who can't do that a fully automatic bullshitting machine, the world becomes fairer. Then the non-linguistic person can also write that nice advisory report, objection or letter of motivation;

Education revolves around certification and that is almost always linguistic, so CHATGPT forces you to explore other ways of education;

We spend too much time on things that don't matter, such as writing nonsensical code or writing reports, or answering emails, we could use a help.

- in education students are dependent on the coach/experts and their whim's/ habits of explaining. CHATGPT can provide a personalized explanation service. for example, if you can learn about a theory or technology by using metaphors in your subject of interest, that can improve your learning abilities.

Or something else. One of the fundamental findings of CHATGPT is that on first sight it does not immediately solve a problem. It's, let's say, the Gorillaz or Getir of language models: find something that isn't a problem and fix it!

Are you sure that this technology is solving the RIGHT problem?
This is a hard question if it is not clear which problem CHATGPT aims to solve.

It is more likely that CHATGPT is one of the cases where the developers first wanted to try if this is possible, before reverse-engineering the possible use cases for it.

On the other hand, if you 'unleash the lion', there might be no way back!

How is this technology going to solve the problem?

Again, if it is not clear what problem CHATGPT aims to solve, it is hard to determine it CHATGPT will be succesfull. Take for example a custom on-demand delivery service which guarantees that your groceries will be delivered within 10 minutes. Such a service may not solve the right problem, but it will solve the problem they indicate they want to solve (sometimes you need groceries quickly). With CHATGPT this is much more diffuse, because it is difficult to determine exactly which problem needs to be solved.

What negative effects do you expect from this technology?

There are a lot of negative effects possible. For example:

The bag of tea effect: ChatGPT trains itself with input it has generated and thus everything it generates becomes increasingly weak;

Education has been blindsighted, and certificates are becoming worthless, because it becomes so easy to cheat;

There will be a plug-in for e-mail, so that chatGPT can automatically reply to your e-mail and other people do the same on the other side;

Information is becoming less and less valuable;

A plagiarism rat race ensues;

ChatGPT determines the answer to questions, but on what basis and with what moral values? The anwers are censured and curated.

People uses it to flood the web with automatic blogging;

People use it to automatically write malware;

- Answers are random and can be personalized. So, it is possible that people are more and more being drawn into filetr bubbles.

- By creating another tool which allow us to avoid human interaction, we destroy basic human learning contexts. Similar to automatic scanning in supermarkets: not necessary anymore to talk to a cashier tand people have more difficulty with smalltalk. for children this used to be a good learning context where they had to confront fear (talk to strangers)

And so on.

They say that in 2025 90% of digital content will be AI-generated. If that is the case, we can ask ourselves what the value of information is still worth by then and if anybody will be able to read it all?

In what way is this technology contributing to a world you want to live in?

A short time ago there was no CHATGPT. There was a lot wrong with the world, and I don't immediately see solutions offered by CHATGPT for all sorts of problems. It is therefore difficult to determine whether and in what way CHATGPT contributes to a world in which you would like to live.

Perhaps it will lead to a democratization of the art of writing, so that people with less sense of language also get a fair chance. Maybe everyone can have their own 'personal assistent' to guide them and help them with whatever they want. It can help write them essays, poems, give answers to questions, helps debugging code, etc, etc.Maybe code is being written faster, so other problems are solved faster. Perhaps it is a driver for educational innovation. Maybe not all.

Maybe it widen the gap between the 'have and the have-nots', the 'rich and the poor', the 'digitals and the not-sodigitals' and the 'higly educated and low litterates'... Maybe it will be a paid service in the future, only available to people who are already in charge anyway. Perhaps the negative effects above will all come true.

Perhaps in a world where everything could be faked or AI-generated, there will be an new wave of trusted content or true news. People will get less and less gullible.

And maybe hardly anything will change.

Now that you have thought hard about the impact of this technology on society (by filling out the questions above), what improvements would you like to make to the technology? List them below.

At the moment CHATGPT has been trained with data up to and including September 2021. The data is one and a half years old. That's wise. It does not seem advisable for CHATGPT to get involved in current discussions.

CHATGPT is censored and has not been trained with all kinds of erotic material (say 90% of the internet). But it is not clear when it is and when it is not censored. That could be clearer.

CHATGPT honestly indicates what the shortcomings are. It tells that answers are sometimes incorrect and nonsensical, but it would be good if that were stated with every answer given instead of in the fine print. Output should be labeled as 'AI-generated' or 'human-generated'.

CHATGPT should remain free. Some say a tool like this should be banned, but we can also look for a way to guide it in a better direction.

Hateful and criminal actors
What can bad actors do with your technology?

In which way can the technology be used to break the law or avoid the consequences of breaking the law?
In an educational context, someone who cheats is a bad actor. CHATGPT is quite capable of writing essays or open book articles, so that's cheating. However, this 'problem' has been around for much longer with tools like Google, Translate and Grammarly. So the question is whether something will change in essence or whether it is simply an extra tool for cheaters (and thus a stimulus for educational innovation).

There are however other ways to use CHATGPT to break the law or avoid the consequences. For example:

- Write malicious code with CHATGPT (malware);
- Covering an agency with appeals written by CHATGPT;
- Indentity theft by posing (using CHATGPT) as someone else;
- Creating fake science;
- Using it for creating SPAM.
- And so on.

"Give a bad actor a personal assistant and a world is opening up."

Can fakers, thieves or scammers abuse the technology?
Yes, of course, as seen above CHATGPT is perfect for creating fake science, fake articles and other fake media outlets. It is nothing new, but CHATGPT can help to create more volume and be more convincing.

CHATGPT is a bullshitter, like Harry Frankfurt coined in his book. A bullshitter does not lie, a bullshitter just does not care about the truth. A bullshitter wants to sound convincing to reach its goals. The goal of CHATGPT is to sound convincing which makes it the ultimate bullshitters - assistant.

There is also an incentive for fakers and people who crave chaos to fill the web with incorrect data. The more incorrect data CHATGPT will see online, the more incorrect output it will produce, as it cannot distuingish between correct or incorrect. So, misinformation and misinterpretation of data can become a big problem here.

Can the technology be used against certain (ethnic) groups or (social) classes?
Depending on your agenda bad actors can try to use CHATGPT to produce output that can be used against certain groups of people. However this is

hard because the language model is not transparent and very large and CHATGPT is heavily curated, censored and shys away from sensitive statements.

In which way can bad actors use this technology to pit certain groups against each other? These groups can be, but are not constrained to, ethnic, social, political or religious groups.
We already have examples from the past where user input generated a hatefull, holocaust-deneing, women-unfriendly and racist enviroment (e.g. Tay from Microsoft).

This is probably harder with CHATGPT (see remarks above). However, never underestimate the creativity of a bad actor.

How could bad actors use this technology to subvert or attack the truth?
See earlier remarks on using CHATGPT to create (fake) content, bullshitting and trying to generate (fake) content on which CHATGPT is trained.

There is also a fundamental issue and that is that CHATGPT is trying to sound convincing, so it just answers the question. It doesn't always mention the sources of that answer, people can tend to think that it is a 'truth', while in real life, multiple viewpoint might arise and people have to decide for themselves who to believe or not.

Now that you have thought hard about how bad actors can impact this technology, what improvements would you like to make? List them below.
I think the tool has to explain where the answers come from, so people can make up their own minds. Also people should get the same kind of access to it, so that everyone can benefit from it in the same amount. Co-Founder Elon Musk says the best defense to people creating things we are concerned about is "to empower as many people as possible to have AI. If everyone has AI powers, then there's not any one person or a small set of individuals who can have AI superpower."

If this is true, remains to be seen.

Technology Impact Cycle Tool

ChatGPT and the impact on education

Privacy

Are you considering the privacy & personal data of the users of your technology?

Does the technology register personal data? If yes, what personal data?

Yes. There is a lot of data collected and it can be used in all kinds of ways. OpenAI is pretty clear and honest about that.
See: https://openai.com/privacy/

Do you think the technology invades the privacy of the stakeholders? If yes, in what way?

Yes, it does, but I do not think people will feel that their privacy is invaded in an unjust way. After all, you agree to share your data when you sign up and you get to use the (for now) free services of CHATGPT in return.

Is the technology is compliant with prevailing privacy and data protection law? Can you indicate why?

For now it is as far as we can see.

However, the technology is so new that probably laws and regulations have to be written. As many times the case innovation comes first and then laws will follow.

Does the technology mitigate privacy and data protection risks/ concerns (privacy by design)? Please indicate how.

No, it does not. If you read the privacy policy, you see that a lot of data is collected and shared. There is clearly a business model connected to the collection of data.

Most privacy by design guidelines are ignored.

In which way can you imagine a future impact of the collection of personal data?

Yes, for sure. It is the same with Google. The things we search for are very personal. If CHATGPT becomes a kind of Google or a trusted advisor you can imagine that people will have very personal conversations with CHATGPT.

These conversations are stored and it is easy to imagine a future in which they are misused.

Now that you have thought hard about privacy and data protection, what improvements would you like to make? List them below.
There are some easy improvements here.

- Ask for a suscription;
- In return allow people to manage their own data;
- Only use data to improve the system;
- Do not personalize except when absolutely necessary;
- Do not share personal data.
But, I am afraid the data is too valuable to do so.

Human values
How does the technology affect your human values?

How is the identity of the (intended) users affected by the technology?
If information can be synthetically generated, what is then the worth of it and what does it say about us as humans if even those highly-humanized things like creating (text or images) can be outsourced?

CHATGPT is a bullshitter (it creates Bullshit as meant by Harry Frankfurt). We humans have a hard time distinguishing human-generated bullshit from AI-bullshit. This says a lot about us, of course.

Some possible effects:

- People who are really good at bullshitting can lose their mojo;

- People who have great competences but a hard time putting their thoughts on paper are empowered;

- People who are great at writing are no longer valued as much;

If something is valuable to do (like writing an article or motivation letter) but it can also be done by an AI, will people still do it. And if they do not, are they robbed or rewarded?

How does the technology influence the users' autonomy?
It can help us but it can also make us dependent on it.

What is the effect of the technology on the health and/or well-being of users?
It can get people stuck in 'information bubbles'. As the tool doesn't offer multiple answers and sources are often not good refered to, people are not challenged to make up their mind for themselves. That can make them feel less capable.

Now that you have thought hard about the impact of your technology on human values, what improvements would you like to make to the technology? List them below.
It might be an idea to let the tool give multiple answers and references, so people can dive into information from others also, instead of listening to a system 'mansplaining' everything from one perspective.

Technology Impact Cycle Tool

ChatGPT and the impact on education

Stakeholders
Have you considered all stakeholders?

This category is only partial filled.

Who are the main users/targetgroups/stakeholders for this technology? Think about the intended context by answering these questions.

Name of the stakeholder
Students

How is this stakeholder affected?
Students can use the tool, to make their life easier or get some help writing essays or writing or debugging code. But, they can also be affected if other students use the tool while they haven't. In this way an unequal playing field might occur.

Did you consult the stakeholder?
Yes

Are you going to take this stakeholder into account?
Yes

Name of the stakeholder
Teachers

How is this stakeholder affected?
Teachers can use the tool themselves, but for student work they might not be able to detect which work is made by a student or writen by an AI. This might become a problem if schools depent their grades based on written output.

Did you consult the stakeholder?
Yes

Are you going to take this stakeholder into account?
Yes

Name of the stakeholder
Authorities

How is this stakeholder affected?
Governments are many times responsible for funding (higher) education, but how can they tell if education is effective and students learned something if

their input is generated by an AI?

Did you consult the stakeholder?
Yes

Are you going to take this stakeholder into account?
Yes

Name of the stakeholder
Parents

How is this stakeholder affected?
Parents are many times co-funders of a student education. How can they tell if their money is 'well spend' if a students passes his/her exams but hasn't learned something from it and fails in their professional life afterwards?

Did you consult the stakeholder?
Yes

Are you going to take this stakeholder into account?
Yes

Did you consider all stakeholders, even the ones that might not be a user or target group, but still might be of interest?
-

Now that you have thought hard about all stakeholders, what improvements would you like to make? List them below.
It might be a good idea to make it transparent to all stakeholders if input is generated by an AI or not.
But, it might also be an option to not assess students on knowlegde alone (because this can easily be generated or replicated), but on the process and the application of that knowlegde. It that way tools like this do not have to be a tread to education, but can be a helpfull tool.

Data
Is data in your technology properly used?

Are you familiar with the fundamental shortcomings and pitfalls of data and do you take this sufficiently into account in the technology?
As we are not the developers of this tool we cannot know everything about the underlying data. But, from what we know now, it looks like each users gets his/her own answer from the AI, because it is also building upon previous questions and answers. This might be a problem, because people can get stuck in their own bubbles because of that.

How does the technology organize continuous improvement when it comes to the use of data?
The underlying model is a Large Language Model with Reinforcement Learning from Human Feedback (RLHF). Large language models have their limitations and the answers can sometimes be false or misleading, but most of the times the model gives a quite good answer to questions. It also seems to know when to answer a question and when not to answer it, or answer it by providing a biased answer with less confidence. That's a good start!

How will the technology keep the insights that it identifies with data sustainable over time?
Microsoft released an AI bot called Tay in 2016, which was shut down in a matter of hours due to misuse. Meta AI took down the Galactica model demo in a matter of two days.
These examples illustrate the difficulties of putting online large language models for anybody to use. We think OpenAI made a smart move by requiring OpenAI credentials for trying the research preview.

In what way do you consider the fact that data is collected from the users?
We don't have prove that user generated data are used and/or sold by now, but that can happen in the future.

For now the user gets far more in return than he/she is giving.

Now that you have thought hard about the impact of data on this technology, what improvements would you like to make? List them below.
We think the model could be better if open-sourced.

Inclusivity
Is your technology fair for everyone?

Will everyone have access to the technology?
For now the technology is free for everyone, but you have to be digital savvy and have to have access to a computer.

Furthermore some people will be more able to give good prompts then other. As always, some people are jumping on the bandwagon right away while other will stay behind. It is unsure if those people will catch up eventually or not.

Does this technology have a built-in bias?
The tool seems to be trained on a very large, but specific set of data. We have already seen some examples where the tool is favoring white males over females or people with color.

Most training material is biased on English source data. for example, when using it for outdoor questions on security of climbing, German is a dominant source language, but ChatGPT uses secondary, derived English sources.

Does this technology make automatic decisions and how do you account for them?
In contrary to Google, which gives multiple options in answer to a question, ChatGTP gives normally one answer to a question. It is not always clear on what basis the system came to that one answer. When not enough source data is available for a certain prompt, the technology starts to "hallucinate" because it always tries to come up with an answer

Is everyone benefitting from the technology or only a a small group? Do you see this as a problem? Why/why not?
Smart people will soon discover how to use the tool to their benefits, while other might be left behind. Passing an exam in a school where only knowlegde is assessed might become easier, while passing an assessment with a portfolio might be close to as hard as it already was. This might make the differences between people and institutions bigger.

the more skillfull you are in writing good prompts and in assessing information, the more useful this tool is. These skills als more common for higher educated people and IT savvy people.

Does the team that creates the technology represent the diversity of

our society?

OpenAI LP is governed by the board of the OpenAI nonprofit, comprised of OpenAI LP employees Greg Brockman (Chairman & President), Ilya Sutskever (Chief Scientist), and Sam Altman (CEO), and non-employees Adam DAngelo, Reid Hoffman, Will Hurd, Tasha McCauley, Helen Toner, and Shivon Zilis.

Our investors include Microsoft, Reid Hoffmans charitable foundation, and Khosla Ventures.

Now that you have thought hard about the inclusivity of the technology, what improvements would you like to make? List them below.

We don't know the cultural backgrounds of the people at OpenAI exactly, but they seem to come from a silicon valley incrowd. IT might be a good idea to make this team broader.

Technology Impact Cycle Tool

ChatGPT and the impact on education

Transparency
Are you transparent about how your technology works?

This category is only partial filled.

Is it explained to the users/stakeholders how the technology works and how the business model works?
Because the tool is trained on a very large dataset based on reinforcement learning and part of a capped-profit (formally nonprofit) company, total transparancy about how the technology works cannot be expected. During a 2016 conversation about the technological singularity, Altman said that "we don't plan to release all of our source code" and mentioned a plan to "allow wide swaths of the world to elect representatives to a new governance board".

If the technology makes an (algorithmic) decision, is it explained to the users/stakeholders how the decision was reached?
No.

Is it possible to file a complaint or ask questions/get answers about this technology?
Yes. They have a discord and say 'your feeedback will help us improve'.

Is the technology (company) clear about possible negative consequences or shortcomings of the technology?
When you log in, the system tells you more and more about restrictions and limitations.

Now that you have thought hard about the transparency of this technology, what improvements would you like to make? List them below.
This question has not been answered yet.

Technology Impact Cycle Tool
ChatGPT and the impact on education

Sustainability
Is your technology environmentally sustainable?

In what way is the direct and indirect energy use of this technology taken into account?
How much energy does AI consume? A paper from the University of Massachusetts Amherst stated that training a single AI model can emit as much carbon as five cars in their lifetimes. ChatGPT is a big one, consuming probably more.

On the other hand, artificial intelligence (AI) can also be part of the solution, by calculating energy consumption solutions.

Do you think alternative materials could have been considered in the technology?
No, but in the future techniques like this might become more efficient, consuming less energy then they do now. And using a pre-trained model can help to 'not invent the wheel yourself'.

Do you think the lifespan of the technology is realistic?
We don't know. They say that AI tools like this will probably help to get a solution to climate change, but as these kind of tools consume a lot of energy to do so, the question will be what will come first: extinction or a solution to the problem.

What is the hidden impact of the technology in the whole chain?
Chat-GPT (3.5) is based on GPT-3, that is based on.... and so on. GPT-4 will soon follow.

Now that you have thought hard about the sustainability of this technology, what improvements would you like to make? List them below.
Promote the use of pre-trained models and keep the energy consumption as low as possible

Future
Did you consider future impact?

What could possibly happen with this technology in the future?
The whole concept of education can change because of this tool. People have to re-think education in many ways.

Sketch a or some future scenario (s) (20-50 years up front) regarding the technology with the help of storytelling. Start with at least one utopian scenario.
ChatGPT and technologies like it can make life easier and more efficient. People can do what they are good in and outsource things that they find difficult to a system. It can also help them when they are in a writers block. All people can be equally smart and society becomes more equal.

Sketch a or some future scenario (s) (20-50 years up front) regarding the technology with the help of storytelling. Start with at least one dystopian scenario.
"If you have a button that could do bad things to the world, you don't want to give it to everyone" is said by philosophers who are concerned with existential risk from artificial intelligence. And that is exactly what is done here: unleashing the lion before knowing what potential harm it could bring to the world. It could spread misinformation, keep people believing what that already believed, thus keeping them in informationbubbles and demolishing democracies, etc.

Would you like to live in one of this scenario's? Why? Why not?
The utopian scenario sounds best.

What happens if the technology (which you have thought of as ethically well-considered) is bought or taken over by another party?
The transition from a nonprofit to a capped-profit company was viewed with skepticism by Oren Etzioni of the nonprofit Allen Institute for AI, who agreed that wooing top researchers to a nonprofit is difficult, but stated "I disagree with the notion that a nonprofit can't compete" and pointed to successful lowbudget projects by OpenAI and others. "If bigger and better funded was always better, then IBM would still be number one." Following the transition, public disclosure of the compensation of top employees at OpenAI LP is no longer legally required. The nonprofit, OpenAI Inc., is the sole controlling shareholder of OpenAI LP. OpenAI LP, despite being a for-profit company, retains a formal fiduciary responsibility to OpenAI's Inc.'s nonprofit charter. A majority of OpenAI Inc.'s board is barred from having financial stakes

inOpenAI LP.[33] In addition, minority members with a stake in OpenAI LP are barred from certain votes due to conflict of interest.[34] Some researchers have argued that OpenAI LP's switch to for-profit status is inconsistent with OpenAI's claims to be "democratizing" AI. A journalist at Vice News wrote
that "generally, we've never been able to rely on venture capitalists to better humanity".

Impact Improvement: Now that you have thought hard about the future impact of the technology, what improvements would you like to make? List them below.
Make sure that people have equal access to the tool and let the tool propose multiple answers where people can choose from, so they can make up their own minds, but still profit from the options the tool delivers.