Ethical Algorithms: AI and Social Media

Grade 9-12 Activity Write Up

actüa

Youth · STEM · Innovation
Jeunesse · STIM · Innovation

Ethical Algorithms: AI and Social Media

Terms of Use

Prior to using this activity or parts thereof, you agree and understand that:

- It is your responsibility to review all aspects of this document and the associated activity write ups, and ensure safety measures are in place for the protection of all involved parties. Any safety precautions contained in the "Safety Considerations" section of the write-ups are not intended as a complete list or to replace your own safety review process.
- Actua shall not be responsible or liable for any damage that may occur due to your use of this content.
- You may adapt the content for your program (remix, transform, and build upon the material), providing appropriate credit to Actua and indicating if changes were made. No sharing of content with third parties without written permission from Actua.

About Actua

Actua is creating a Canada where every child has the skills and confidence they need to achieve their full potential. As a leading science, technology, engineering and mathematics (STEM) outreach organization, Actua includes over 40 universities and colleges, engaging 500,000 youth in 600 communities each year. For 25 years, Actua has focused on identifying and removing the barriers for entry into STEM and now have national programs dedicated to engaging Indigenous youth, girls and young women, Black youth, those facing economic barriers and youth in Northern and remote communities. For more information, please visit us online at www.actua.ca and on social media: Instagram, LinkedIn, Facebook and YouTube! For more information, please visit us online at www.actua.ca and on social media: Instagram, Facebook, LinkedIn, TikTok and YouTube!

Ethical Algorithms: AI and Social Media

Activity Summary

In this activity, participants will examine how AI shapes social media feeds. Through a card-based simulation, they will experience how recommendation algorithms sort users into "clusters" and create filter bubbles. By stepping into the role of the algorithm, they will discover how engagement-focused systems can amplify misinformation and targeted advertisements. The activity culminates in an ethical debate, where participants act as a fictional AI Ethics Board to weigh the conflicting responsibilities of a social media platform and decide on their algorithm's primary goal.

Developed by Actua, 2025.

Delivery Environment	Activity Duration	Intended Audience	Tech
In-Person	1 Hour	Grades 9-12 (*Ages 14-18*)	**Facilitators should have access to a laptop, projector, speakers, and a screen or blank wall to project onto.** • Projector • Speaker • Screen/Blank Wall • Laptops/Tablets

Achievement Goals

Learning Goals

Following this activity, participants will:

- **Understand** how AI uses data clustering to create personalized feeds, filter bubbles, and echo chambers.
- **Analyze** how an algorithm designed to maximize engagement can accelerate the spread of misinformation and targeted advertising.
- **Evaluate** the ethical responsibilities of social media platforms by weighing the impacts of different algorithmic goals on various stakeholders.

Success Criteria

Following this activity, participants can express:

- **I can explain** how AI sorts users into "clusters" and how that creates a filter bubble.
- **I can describe** how an algorithm that prioritizes "likes" and engagement can help fake or misleading information spread.
- **I can identify** different stakeholders (like users and the company) and explain how an algorithm's goals can help or harm each of them.

Logistics (Timing, Group Sizing, Materials)

Section Title	Time	Group Size	Materials
Opening Hook	5 minutes	*Whole Group*	**Facilitators** - AI and Social Media - Activity Slide Deck *(Appendix C)*

Section Title	Time	Group Size	Materials
Section 1: Algorithm Simulation	10 minutes	*Small Groups (3-6)*	**Facilitators** • AI and Social Media - Activity Slide Deck *(Appendix C)* • Interest Cards *(Appendix C)* • Media Cards *(Appendix C)*
Section 2: Algorithm Ethics	15 minutes	*Small Groups (3-6)*	**Facilitators** • AI and Social Media - Activity Slide Deck *(Appendix C)* **Per Group** • Ethical Matrix Worksheet *(Appendix C)* • Writing utensil
Reflection & Debrief	10 minutes	*Whole Group*	**Facilitators** • AI and Social Media - Activity Slide Deck *(Appendix C)*

Safety Considerations

Safety considerations have been provided below to support safety during this activity, however they are not necessarily comprehensive. It is important that you review the activity and your delivery environment to determine any additional safety considerations that you should be implementing for the delivery of these activities.

Emotional Safety

• Facilitators should understand that participants have different lived experiences and prior knowledge about AI safety, AI, and digital citizenship. This activity may involve or lead to discussions of sensitive topics, such as ethical implications of AI. Facilitators should encourage open, respectful discussions and acknowledge all perspectives. Facilitators should always keep

the participants' emotional safety in mind in these discussions, and defer to training from their institution and training received.

- This activity involves discussion on misinformation and disinformation, which may be a sensitive topic. Thinking about these issues might remind participants of past experiences or bring up strong feelings. In addition to the suggestions above, facilitators should:
  - Make it clear that participants can step back from any activity, discussion, or resource that feels too heavy.
  - Build in space for reflection, questions, or quiet processing.
  - Offer the following resource on the topic of gendered misinformation and disinformation if relevant, Tackling Online Gendered Disinformation: Youth Guide, which discusses what it is and how to take action. More resources can be found on actua.ca/misinformation.

Curriculum Links

This activity aligns with these components found in the UNESCO AI Competency Framework for Students:

Human-Centered Mindset: Human Agency

- Learners are expected to be able to recognize that AI is human-led and that the decisions of the AI creators influence how AI systems impact human rights, human-AI interaction, and their own lives and societies (p. 29-30).

Human-Centered Mindset: Human Accountability

- Learners are expected to be able to recognize that human accountabilities are the legal obligations of AI creators and AI service providers, and understand what human accountabilities they should assume during the design and use of AI. They should also foster an awareness that human accountability is a legal and social responsibility when using AI to assist decisions that affect humanity and uphold the principle that humans should not cede the determination to AI when making high-stakes decisions. They are also expected to enhance their judgement on, and attitudinal resilience to, the illusive claims on the use

of outputs and as well as predictions that AI can usurp humans' thinking and decision-making (p. 37-38).

Human-Centered Mindset: AI Society Citizenship

- Learners are expected to be able to build critical views on the impact of AI on human societies and expand their human-centred values to promoting the design and use of AI for inclusive and sustainable development. They should be able to solidify their civic values and the sense of social responsibility as a citizen in an AI society. Learners are also expected to be able to reinforce their open-minded attitude and lifelong curiosity about learning and using AI to support self-actualization in the AI era (p. 45-46).

Ethics of AI: Embodied Ethics

- Learners are expected to be able to develop a basic understanding of the ethical issues around AI, and the potential impact of AI on human rights, social justice, inclusion, equity and climate change within their local context and with regard to their personal lives. They will understand, and internalize the following key ethical principles, and will translate these in their reflective practices and uses of AI tools in their lives and learning: Do no harm, Proportionality, Nondiscrimination, Sustainability, Human determination, and Transparency (p. 31-32).

Ethics of AI: Safe and Responsible Use

- Learners are expected to be able to carry out responsible AI practices in compliance with ethical principles and locally applicable regulations. They are expected to be conscious of the risks of disclosing data privacy and take measures to ensure that their data are collected, used, shared, archived and deleted only with their deliberate and informed consent. They are also expected to be conscious of typical AI incidents and the specific risks of certain AI systems, and be able to protect their own safety and that of their peers when using AI (p. 39-41).

AI Techniques and Applications: AI Foundations

- Learners are expected to develop basic knowledge, understanding and skills on AI, particularly with respect to data and algorithms, and understand the importance of the interdisciplinary foundational knowledge required for gradually deepening understanding of data and algorithms. They should also be able to connect conceptual knowledge on AI with their activities in society and daily life, concretizing a human-centred mindset and ethical principles through an understanding of how AI works and how AI interacts with humans (p. 32-34).

AI Systems Design: Problem Scoping

- Learners are expected to be able to understand the importance of 'AI problem scoping' as the starting point for AI innovation. They are also expected to acquire the knowledge and project-planning skills needed in order to conceptualize and construct an AI system (p. 35).

AI Systems Design: Iteration and Feedback

- Learners are expected to enhance and apply their interdisciplinary knowledge and practical methods to evaluate the humanistic appropriateness and methodological robustness of an AI model and its impact on individual users, societies and the environment. They should be able to acquire age-appropriate technical skills to improve the quality of datasets, reconfigure algorithms and enhance architectures in response to results of tests and feedback. They should be able to apply human-centred mindset and ethical principles in simulating decision-making on when an AI system should be shut down and how its negative impact can be mitigated. They are also expected to cultivate their identities as co-creators in the larger AI community (p. 50-51).

This activity can be connected to the following subject areas:

Science

- Investigating systems with specific inputs, processes, and outputs.
- Understanding the role of science and technology in society and daily life.

Language Arts

- Demonstrate an understanding of how some digital media can be created and how they affect the audience.

Mathematics

- Collecting, organizing, and interpreting qualitative and quantitative data.

Psychology

- Analyze how social and digital environments influence individual or group behaviour.
- Investigate how information is processed and how external stimuli (e.g. social media algorithms) impact perception and decision making.
- Evaluate the impact of digital technology and social media on well-being.

Community Connections

Community connections are suggestions from Actua, grounded in our approach, on how facilitators can adapt the activity to reflect the strengths, interests, and priorities of the community where or with whom it is delivered. Consider the following guiding questions to adapt the activity in meaningful ways:

- **Consult with community:** Are there local organizations, Knowledge Keepers, or community members who could contribute insight or context to this topic?
- **Draw on youth experience:** How can you give participants opportunities to share, reflect on, and apply how this learning is relevant to them or their community? Invite participants to identify what knowledge, who, and where they already learn from.

- **Integrate local examples:** How can you tailor this activity to local or regional interests, industries, or community priorities (e.g. land and environment, health, technologies)?

Activity Procedure

To Do in Advance

SECTION	PREPARATION
General	**Think ahead and be ready to adapt:** - Determine your **delivery method** and leverage ideas from the delivery recommendations and adaptations sections. - While **estimated times** are provided, it will be helpful to think about how much time you would like to spend on different activities and discussions. - While **group sizes** (individual, pairs, groups) are suggested, many activities are flexible for whatever will work in your classroom. **Prepare for the content:** - Have **answers in mind** to share with participants for the various reflection questions asked. - Examine the provided materials to determine if they are **suitable** for your participants. **Equipment:** - Ensure device, screen and projector are set up. - Prepare participant devices.
Section 1: Algorithm Simulation	- Print and cut out the *Interest Cards (Appendix C)* - Print, cut out, and shuffle the *Media Cards (Appendix C)*

SECTION	PREPARATION
Section 2: Algorithm Ethics	• Print out *Ethical Matrix Worksheets (Appendix C)*

Opening Hook

1. Using AI and Social Media - Activity Slide Deck *(Appendix C)*, show participants two distinct mood boards.
   a. Mood Board #1: Technology, games, photography, art, movies.
   b. Mood Board #2: Travel, animals, nature, adventure.
2. Ask participants: "Who is interested in the first mood board? Who is interested in the second?"
   a. Some participants may say that they are interested in both, or neither. That's okay! There are some many hobbies and interests, these mood boards just have a few of them!
   b. Now imagine that every click, like, watch, search, and/or follow you make online adds a piece to your own digital mood board.
3. Ask participants: "If you use social media, do you notice any patterns in the kinds of posts that show up in your feed?"
   a. Ask participants:
      i. "How do platforms like TikTok, YouTube, or Instagram get so good at showing you exactly what you are interested in?"
      ii. "How might this be helpful? How might this be limiting?"
4. Explain to participants that the "brains" behind your feeds are recommendation algorithms powered by **Artificial Intelligence (AI)**. Today, we're going to simulate how these systems work.

Section 1: Algorithm Simulation

Part 1: Cluster Creation

1. Using AI and Social Media - Activity Slide Deck *(Appendix C)*, explain to participants that the first thing that these algorithms do is sort large groups of users in **data clusters**.

2. This is a process where users are automatically sorted into groups (clusters) based on shared patterns in their data, like videos they've watched or posts they've liked. This allows the AI to predict that you will like something because other people in your cluster also liked it.

    a. This is not like being added to a group chat - data clusters are constantly adjusting and shifting. A user could be in the same cluster for a week, and then because of liking one post be moved to a different one.

3. To simulate simplified data clusters, participants will be given Interest Cards (*Appendix C*) at random. Once they have received their interest card, they will find the other people in the room who have the same card.

4. Distribute the interest cards to participants and have them gather into their matching groups.

5. Once they have done so, explain the concept of **filter bubbles.**

    a. These groups are a physical representation of data clusters online, and they filter content based on what your interests are. This often leads to something called an **echo chamber**, where everyone shares the exact same opinion/interests.

    b. On a more personal level, these bubbles use **Confirmation Bias** to keep you on their platform. People have a natural tendency to like things that confirm what we already believe. Therefore, when our Social Media is filled with content that affirms our beliefs, we are more inclined to keep scrolling! Their goal is to keep users engaged.

6. Ask participants: "How might this be helpful? How might this be limiting?"

    a. Share this example: "Do you ever notice posts to buy things or advertisements on your feed?"

      i. Companies will pay Social Media platforms to host sponsored posts. The platforms will then often push those ads to the clusters that they think will be most likely to click them due to aligning interests!

Part 2: Feed Simulation

1. Explain that these groups will now receive twelve randomized Media Cards (*Appendix C*), simulating the algorithm feeding them content.
2. Participants will act as their clusters algorithm by reviewing each card and deciding if it is relevant to their cluster's interest.
   a. If it does, they keep the card in the group. If it doesn't, they can discard it at a designated discard pile.
   b. Once they have done this, the remaining Media Cards will be shuffled and redistributed evenly. This will continue until there are no more cards left to distribute.
3. Once complete, ask participants: "Do you think all of the media in your group is factual/accurate?"
4. Explain mis/disinformation:
   a. Misinformation: False or inaccurate information that is spread by **mistake** or unwittingly.
   b. Disinformation: False or inaccurate information that is **deliberately** intended to deceive, manipulate or confuse.
5. Using the AI and Social Media - Activity Slide Deck *(Appendix C)*, go through each interest and highlight which Media Cards are mis/disinformation.
6. Discuss with participants why these stayed in their bubble, despite being inaccurate.
   a. Because it matched the interest card!
7. Explain that they didn't do anything wrong, they performed the task given to them! The task was to "keep relevant media", not "keep relevant media but remove potential mis/disinformation"!
8. Ask participants: "Did you notice any advertisements or "sponsored" posts in your group?

9. Using the AI and Social Media - Activity Slide Deck *(Appendix C)*„ highlight the "sponsored posts" in the Media Cards.

    a. All of these share one key part - a **call to action**. This usually comes in the form of a prompt to click a link and/or buy something. On social media, they often grab your attention through a discount code or "limited time offer".

10. Most Social Media algorithms function the same way your bubbles did: **prioritizing engagement.** They have little to do with managing accuracy and/or safety. This is how mis/disinformation can go viral without any disclaimer to the viewer!

11. Explain to participants that it has become increasingly common to find 100% AI generated content going viral on platforms like Instagram, YouTube, TikTok, and Facebook - especially with short-form content.

    a. Remember, the algorithm's goal is engagement, not accuracy. This means the majority of social media algorithms aren't able to detect whether or not a piece of content is generated by AI.

12. With advances in technology, it has never been easier to create fully AI generated video content. These videos usually come with a small watermark that serves as their only "disclaimer" that they are generated.

    a. However, with the rise in popularity of these tools there have been entire applications created with the sole purpose of removing these watermarks, making it even more difficult to determine whether or not they are AI generated.

    b. And here is the twist. Those watermark removal apps? Also powered by artificial intelligence.

13. Discuss with participants:

    a. Who do you think should be responsible for labeling or checking AI-generated content? The creator, the platform, the viewer?

    b. What are some actions you can take to verify content before trusting and/or sharing it?

Section 2: Algorithm Ethics

1. Discuss with participants: "Who decides the algorithm's behaviour?"
   a. The company that owns the Social Media platform does! But what leads them to making that decision?
2. Introduce the term **stakeholders**.
   a. Everyone who is affected by a company's choices is a stakeholder. Some examples include users of the platform, employees of the company, or financial investors!
3. Explain that participants are now going to act as an "AI Ethics Board" for a Social Media company. The company is fictional, but their discussion reflects real questions companies might face. **Their job is to decide the algorithm's primary goal.**
4. The company has narrowed down the goal to two possible options:
   a. **Goal 1: Maximize Engagement**
      i. The priority is to keep users on the app as long as possible. It will prioritize exciting, emotional, and tailored content, even if it's unverified or potentially misleading.
      ii. This makes the most money.
   b. **Goal 2: Prioritize User Safety**
      i. The priority is to ensure accuracy and user well-being. It will slow down, flag, or remove unverified content.
      ii. This costs more money and might make the feed "less exciting," but makes the platform overall a more trustworthy source of information.
5. Distribute an Ethical Matrix (*Appendix C)* worksheet for each group.
6. Each group will discuss the pros and cons of each goal to help determine what the algorithm's primary goal will be. For each stakeholder, they must list the potential positive (pros) and negative (cons) impacts of choosing Goal A versus Goal B from that perspective.
7. After the discussion, each group will share their decision - Will their algorithm prioritize Goal #1, or Goal #2? Why?

8. Remind participants that although the company is fictional, this exercise supports deeper understanding of similar real-world tools (for example, understanding why companies may more heavily prioritize one goal over the other), leading to helping the participants make more informed decisions.

Reflection & Debrief

1. Ask participants: "What are the biggest trade-offs between "Maximizing Engagement" and "Prioritizing Safety & Truth" for the Users, Platform, and Society?"
2. How can algorithms that focus mainly on likes and interests make it easier for misinformation to spread?
    a. These algorithms do not detect whether or not content / information is coming from a legitimate source. The only thing the algorithm wants is to keep the user on the platform - no matter what the content is, or where it comes from!
3. What are some things you can do to counteract the spread of misinformation on your Social Media feed?
    a. Seek multiple/external sources.
        i. If you see something unbelievable, research the topic, fact, statement, or image on a separate platform. Whenever possible, search for sources of statements, and read up on those sources - do not just take them at face value!
    b. Think critically of content that seems too good to be true, or surprising.
    c. Try to diversify your "feed" through different forms of media and creators.

Delivery Adaptations

Modifications are ways to make the activity more accessible, **extensions** are ways to make the activity last longer or more challenging.

Modifications

SECTIONS 1: ALGORITHM SIMULATION

- For younger participants, it may be better to use a simple analogy for filter bubbles/data clusters, like "friend groups". Focus on the core idea that the algorithm only shows you what it thinks your group will like!
- Rather than participants forming groups and filtering Media Cards independently, this can be done with the whole class. The facilitator holds up a media card one at a time, and participants vote on which of the four categories they think it belongs in.

SECTION 2: ALGORITHM ETHICS

- Reduce the stakeholders in the Ethical Matrix to just two, **The Users** and **The Platform.** This simplifies the discussion to a more direct conflict of interest.

Extensions

SECTION 2: ALGORITHM ETHICS

- Have participants brainstorm on ways platforms could compromise between Goal A and Goal B that benefits both users and financial stakeholders, instead of choosing one or the other.

REFLECTION AND DEBRIEF

- Continue the conversation with participants:
  - How can being more intentional about your feed help you feel more informed and empowered online?
  - What are some things you can do to improve your sense of well-being online?

References & Gratitude

Actua. (2025). *Misinformation*. actua.ca/misinformation

Amazon Web Services. (n.d.). *What is Machine Learning?*. Amazon.
aws.amazon.com/what-is/machine-learning/

Beyari, H., & Hashem, T. (2025). *The Role of Artificial Intelligence in Personalizing Social Media Marketing Strategies for Enhanced Customer Experience.*
https://doi.org/10.3390/bs15050700

Inspirit AI. (2023). *AI in Social Media: The Benefits and Risks.*
https://www.inspiritai.com/blogs/ai-student-blog/ai-in-social-media

Government of Canada. (2024, May). H*ow to identify misinformation, disinformation, and malinformation.*
cyber.gc.ca/en/guidance/how-identify-misinformation-disinformation-and-malinformation-itsap00300

Morrison, S. (2023). *The tricky truth about how generative AI uses your data.* Vox.
vox.com/technology/2023/7/27/23808499/ai-openai-google-meta-data-privacy-nope

NVIDIA. (n.d.). *Generative AI – What is it and How Does it Work? .*
nvidia.com/en-us/glossary/generative-ai/

Proofpoint. (2025). *What Is a Deepfake? Definition & Technology.*
proofpoint.com/us/threat-reference/deepfake

Stewart, J. (2025). *How AI is Changing the Social Media Landscape and Its Impact on Self-Esteem.* The Center for Integrative Counseling and Psychology.
https://thecentercounseling.org/how-ai-is-changing-the-social-media-landscape-and-its-impact-on-self-esteem/

United Nations Educational, Scientific and Cultural Organization. (2024). *UNESCO AI framework: Competency levels.*
unesdoc.unesco.org/ark:/48223/pf0000391105

University of Toronto Libraries. (2025). *Artificial Intelligence for Image Research: How Generative AI Models Work.*
guides.library.utoronto.ca/c.php?g=735513&p=5297039

U.S. GAO. (2020). *Deconstructing Deepfakes—How do they work and what are the risks?*

gao.gov/blog/deconstructing-deepfakes-how-do-they-work-and-what-are-risk

Zewe, A. (2023). *Explained: Generative AI.* Massachusetts Institute of Technology.

news.mit.edu/2023/explained-generative-ai-1109

Appendices

Appendix A: Career & Mentor Connections

AI/MACHINE LEARNING ENGINEER

- Builds and trains artificial intelligence systems that can learn from data to make predictions or perform complex actions. They support the machine learning researcher.

COMPUTER PROGRAMMER

- A computer programmer is a person who creates computer software. They write code to build websites, computer games, financial analysis and many more.

MACHINE LEARNING RESEARCHER / DATA SCIENTIST

- Machine learning researchers or data scientists clean and interpret data while building models using a combination of that data and machine learning algorithms.

Appendix B: Background Information

ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is a branch of Computer Science that deals with a machine's ability to simulate intelligent behaviour. This includes cognitive functions we associate with human minds, such as perceiving, reasoning, learning, and adapting.

AI is becoming increasingly vital in our lives. From digital assistants, GPS navigation, and autonomous vehicles to tools like Siri/Google Home and generative AI tools (e.g., OpenAI's Chat GPT), its impact on our daily lives is growing. AI plays a crucial role in various aspects of work, enhancing efficiency, and taking on hazardous or monotonous tasks. As AI applications grow, discussions on AI ethics and responsible practices are increasingly important.

MACHINE LEARNING

Machine learning (ML) is a type of artificial intelligence where computers learn from data, improve at tasks over time, and adapt without being programmed step by step. Instead of giving the computer every possible instruction, we give it ways to learn from experience, similar to how people practice and improve.

For example, you might teach a computer to play checkers. Rather than listing every possible move, you could program it to learn from playing many games. Eventually, it could get so good that it plays better than the person who wrote the program

The learning process usually includes a few steps:

1. **Give it Data:** First, the computer is given a lot of data to process. This could be pictures, numbers, text, or sounds.
2. **Find the Patterns:** The computer then analyzes all this data and starts to find patterns, building its own "rules" based on experience.
3. **Make Predictions or Actions:** Using these patterns, the computer can make decisions or predictions when it encounters new data.

4. **Get Better and Better:** As the computer gets more data, it can adjust its patterns to become more accurate. The more data and computing power it has, the better its predictions become

Machine learning is used across different fields and in our everyday lives. In your daily life, ML works quietly behind the scenes in many of the tools and apps you use. It powers autocorrect and predictive text on your phone, suggests videos to watch on streaming platforms, and runs the filters on social media apps that recognize and track your face. It also helps your phone unlock through facial recognition by learning your unique features, and it keeps your email inbox clean by filtering out junk mail.

In addition, ML is making an impact in the wider world. In healthcare, it helps doctors diagnose illnesses more accurately, such as spotting broken bones in X-rays. It is also the "brain" behind self-driving cars, enabling them to observe and respond to the world around them. In industries like agriculture and manufacturing, ML supports smarter ways to grow food and produce goods more efficiently.
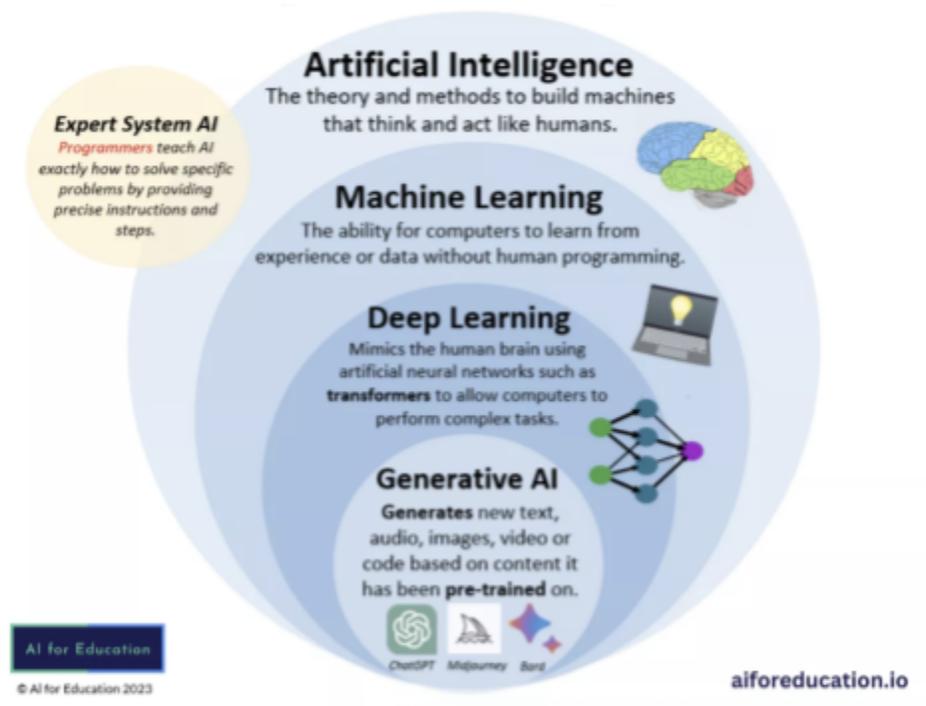
AI for Education. (2023). *Generative AI Explainer.*
aiforeducation.io/ai-resources/generative-ai-explainer

Types of Machine Learning

Machine learning methods are usually grouped into three main types, depending on the goal and the kind of data they use.

- **Supervised learning** happens when the computer is trained with examples that already have the right answers (called labels). The computer learns the link and patterns between the input and output so it can predict answers for new data. This is used in things like email filters that sort *spam vs. not spam*, or in predicting numbers, like house prices.
- **Unsupervised learning** is when the computer is given data without any labels and has to find patterns on its own. It might group similar things together (clustering), or uncover hidden connections in the data. This is used for things like grouping customers into types or making product recommendations.
- **Reinforcement learning** is like learning by trial and error. A computer "agent" makes decisions in an environment, gets rewards or penalties for its actions, and learns strategies to do better over time. This approach is often used in robotics and training AI to play video games.

DATA CLUSTERS

Data clustering is an unsupervised machine learning technique that automatically groups a set of data points based on their similarities. In the context of social media, the "data points" are users, and the "similarities" are derived from their behavior: posts they've liked, videos they've watched, accounts they've followed, and even the time of day they are most active. The AI algorithm analyzes this vast amount of data to identify patterns and sort users into groups, or **"clusters."**

For example, an algorithm might identify a cluster of users who frequently interact with content related to "vintage cameras," "film photography," and "darkroom techniques." A different cluster might form around users who engage with "esports," "PC building," and "tech reviews." The AI does not need to be told what these categories are; it discovers them by finding users who share common patterns of behavior. This process allows platforms to efficiently predict that a user will likely

enjoy a piece of content because other users in their cluster have already engaged with it.

Filter Bubbles and Echo Chambers

While clustering is an efficient way to personalize content, it is also the mechanism that creates **filter bubbles** and **echo chambers**.

- A **filter bubble** is the personalized universe of information that an individual user lives in online. It is created when an algorithm selectively guesses what a user would like to see based on their past behavior and information about them. The algorithm's goal is to maximize engagement, so it prioritizes showing content that it predicts the user will like, while filtering out content it predicts they will dislike or ignore.
- An **echo chamber** is a result of living within a filter bubble. It is an environment where a person only encounters beliefs or opinions that coincide with their own, so their existing views are constantly reinforced and alternative ideas are not presented. The algorithm learns that content which confirms a user's existing biases (**confirmation bias**) is highly engaging, and therefore prioritizes it.

The danger of this process is that it can isolate users from diverse perspectives and create a skewed perception of reality, making them more susceptible to targeted misinformation that is designed to appeal to their specific cluster. An algorithm designed solely to maximize engagement does not inherently distinguish between true and false information; it only distinguishes between engaging and unengaging content.

GENERATIVE AI

Generative AI is a type of artificial intelligence (AI) designed to create new content, such as text, images, music, or code, by learning patterns from existing data. As a subset of AI, which broadly refers to machines performing tasks that normally require human intelligence, generative AI specifically focuses on producing original outputs rather than just analyzing or recognizing information.

Generative AI models generate new data that resembles the examples they were trained on by understanding underlying patterns and structures. Instead of simply responding to inputs with predefined answers, generative AI can create novel and creative content.

What can generative AI create?

- **Text:** Stories, poems, essays, summaries, chat responses, reports, and even programming code.
- **Images:** Drawings, paintings, photorealistic pictures, designs, and digital art from descriptions or sketches.
- **Audio:** Music compositions, sound effects, voice synthesis, and speech generation.
- **Video:** Short animations, deepfake videos, or video sequences from textual prompts (in emerging applications).
- **3D Models:** Shapes and objects for games, simulations, or design prototypes.

MISINFORMATION

Information and Data

Information and data permeate our world at every turn. Information comes in many forms such as words, numbers, symbols, and images. The ability to accurately judge and interpret information helps us predict the weather, build cities, keep a personal budget and even explore outer space. The validation of information is critical to working with information and can be divided into 4 types:

- **Valid:** Factually correct, based on data and not misleading.
- **Inaccurate:** Incomplete or manipulated to change the narrative.
- **False:** Incorrect and can be disproven with other data.
- **Unsustainable:** Cannot be proved or disproved based on available data.

Misinformation, Disinformation and Malinformation

As discussed above, information is a critical part of how we navigate our world. However, as the world becomes more and more connected, the ability to spread false

information has become more prevalent and easy to access. There are 3 types of false information often abbreviated as MDM.

- **Misinformation:** Misinformation refers to false information that is not intended to cause harm. It tends to be accidental and stems from miscommunication. It can be widespread or focused.
- **Disinformation:** Disinformation refers to false information that is intended to manipulate, cause damage and/or guide people in the wrong direction. It is often focused on disparaging a particular group of people or to sway the opinion of a particular group of people.
- **Malinformation:** Malinformation refers to information that stems from the truth but is often exaggerated in a way that misleads and causes potential harm. This is often intentional and intended to mislead or sway the opinions of others.

The S.P.O.T. Method

The **S.P.O.T.** method provides a structured framework for verifying the source and credibility of the content itself.

- **S:** Is this a credible **Source**?
  - Where does it come from? Is it a random social media account, or shared on the news? Have you come across any other reliable content from this source before?
- **P:** Is the **Perspective** biased?
  - Does the person sharing this have a reason they would want you to believe whatever it is?
- **O:** Are **Other** sources reporting the same story?
  - Do you recognize this photo, video, or information anywhere else? If you research it, is anyone else discussing it?
- **T:** Is the story **Timely**?
  - When was it posted? Are they referencing old information, or claiming to be referencing a brand new study?

Identifying Technical Artifacts

Before analyzing technical details of a piece of content, the easiest way to start is by asking some key questions about the content's context and intent:

- **Does this make sense?**
    - Does the situation seem too good to be true, strange, or "uncanny"? If the content features a known person, does their behavior or speech align with what is publicly known about them? Content that feels out of character or context is a primary red flag.
- **Does this want something from me?**
    - Is it aggressively trying to persuade you to purchase a product, sign up for a service, or adopt a specific viewpoint? Content with a strong persuasive or urgent call to action warrants higher scrutiny.
- **Does this have signs of being AI-generated?**
    - This involves looking for specific technical flaws or artifacts left behind by the AI generation process. See the next section for more information on artifacts to look for.

ETHICS AND AI

Artificial intelligence offers powerful tools and new possibilities. As these systems learn from data, make decisions, and shape our world, it is important to consider their ethical impacts.

Actua has developed a resource (*Appendix C*) to support facilitators in leading discussions with youth about ethics and responsible AI use. Facilitators are encouraged to engage youth in meaningful conversations that empower them to think critically about how AI is designed, used, and experienced in the world around them. This resource emphasizes human agency and responsibility, supports values-based reflection, and creates space for curiosity, dialogue, and informed decision-making as digital citizens.

Appendix C: Additional Resources

GENERAL

Activity Slide Deck

- AI and Social Media - Activity Slide Deck
  - ○ ***Note:*** This link will automatically download to your device.

Supporting Resource

- AI in Context: Responsibility and Ethics in Artificial Intelligence

SECTION 1: ALGORITHM SIMULATION

Activity Page(s)

- Interest Cards (refer below)
- Media Cards (refer below)

SECTION 2: ALGORITHM ETHICS

Activity Page(s)

- Ethical Matrix Worksheet (refer below)

Ethical Algorithms: AI and Social Media

Interest Cards

Interests

SCIENCE *Interest Card*	**SCIENCE** *Interest Card*
SPORTS *Interest Card*	**SPORTS** *Interest Card*
ART *Interest Card*	**ART** *Interest Card*
TECHNOLOGY *Interest Card*	**TECHNOLOGY** *Interest Card*

Ethical Algorithms: AI and Social Media

Media Cards

Science

Vampire bats reported in Canada, how to protect yourself!

@thedailyorbit

A HISTORY OF NASA ORION MISSION DESIGN, COPERNICUS SOFTWARE DEVELOPMENT, AND THE ARTEMIS I TRAJECTORY

@inst_astro

Waves, no matter the type (mechanical, fluid, electric, magnetic, gravitational) all fundamentally behave in the same way.

@mindblown

Diamond filled meteorite lands in Manitoba, government covering it up. What aren't they telling us?

@ecoworldweekly

Turritopsis dohrnii, a jellyfish, can transform back into its juvenile stage, which in turn grows back into an adult over and over, Making this jellyfish functionally immortal.

@dinodynamite

Girl who feels no pain, doesn't sleep, and immune to all diseases, a unique case that shocked science.

@naturewizsolutions

Vertical farms use up to 10 times less water, achieving a 98% reduction in water usage and a 99% decrease in land requirements.

@nationalscirev

Seismic solutions: how Canada's Parliament will be prepared to withstand earthquakes.

@morningbriefing

Scientists disprove climate change, claim atmospheric carbon is let off from oceans, unrelated to human activities

@thetruthseek_88

For 45 Winters, mountain hermit and reluctant climate scientist, Billy Barr, has been quietly at work recording the weather—twice a day, every day—all winter long

@omg_facts

There are 8 x 10^67 ways to shuffle a standard deck of 52 cards. This number is so large that it's statistically more likely that the order of cards in a shuffled deck has never existed before in the history of the universe.

@triviatime

Learn how to make a color changing slime using THERMOCHROMIC PIGMENT! 50% off now, Our Biggest Sale EVER!

@slimetime

Sports

Olympic Games in Antarctica? International Olympic Committee pitches globally neutral permanent location for Olympic Games.

@globalsportspost

In NHL games, the puck is kept frozen ahead of the game to keep them from bouncing. Warm pucks are more likely to go airborne and injure players or fans.

@fastskatefacts

You won't believe this secret super-food to help put on muscle and maximize gains: Star Fruit.

@hiddenfaxproof

The Balancing Act: Negotiating Athleticism and Artistry in Rhythmic Gymnastics and Other Aesthetic Sports by Christine Mazumdar

@fitalyticshub

Dancers are ranked number one most physically demanding job, above hard labourers, as well as Athletes and other Sport Competitors.

@occuinfonetwork

Major League Baseball looking to replace all umpires with Artificial Intelligence.

*@mlbinsider_official_*

A Series of Studies—The Physiological Basis for Strength Training in American Football: Fact Over Philosophy by William J. Kraemer

@dunkchampviral

Lacrosse can be traced back to the Indigenous peoples of North America, who played variations of the game as early as the 12th century.

@historicpreserveorg

St. John's, Newfoundland announces pitch for new CFL team, called the Newfoundland Puffins.

@newsbreaking_247

Cardiovascular Effects and Benefits of Exercise, by Matthew A. Nystoriak and Aruni Bhatnagar in Frontiers of Cardiovascular Metabolism, 2018

@sidelinescoop

Soccer is the most watched sport in the world, with an estimated 3.5 billion fans globally.

@therealfootball

Looking for performance gains, without the stress of figuring out a meal plan? Buy our pre-made high-protein meal kits today. Save 28% when you pay annually!

@weight4it

Tech

Archaeologists discover Ancient Computing Technology in Pharaoh tombs.
Were Egyptians actually the most advanced civilization in history?

@real_grhmhnck

Coding as another language: a pedagogical approach for teaching computer science in early childhood by Marina Umaschi Bers in Journal of Computers in Education 2019.

@cybersecurityinst

The QWERTY keyboard was designed to separate commonly used letter pairs to prevent mechanical typewriters from jamming.

@mechanicalfacts

AI water usage grossly overestimated by environmental groups: the actual benefits of AI on the environment.

@dontwantu2kno

Researchers have unveiled a neural network that can accurately chart the expanse of large Antarctic icebergs in satellite images, working 10,000 times quicker than human researchers can chart.

@digitalagemag

Born Electric, Buried Toxic: The Life Cycle of Generative AI and Its Environmental Impact. By Sarah Mangelsdorf and Lauren Behan in Brief, 2025

@ai_safety_board

World's most accurate smart watch, Guardian Loop, can diagnose over 1000 common diseases and illnesses, before symptoms even appear.

@pro_techloot

Past decade of supercapacitor research – Lessons learned for future innovations Molahalli et al. 2023 in Journal of Energy Storage

@cybersecurityinst

Computer screens only use red, green, and blue lights (called subpixels) to create every single color you see.

@digitalage

The first computer bug was a real bug! In 1947, a team of engineers found a moth causing issues in their computer at Harvard University. They logged this as the first case of a "bug" and "debugging."

@futurefound

An introduction to quantum computing for non-physicists by Eleanor Rieffel and Wolfgang Polak, in ACM Computing Surveys, 2000

@gadgetgazette

You don't need a study-buddy.
You need a study assistant that caters to you. Use code **study4me** to get 20% off your own AI powered study companion.

@thattechbro

Arts

The word "ukulele" is Hawaiian for "jumping flea," which was inspired by the way the player's fingers move across the strings. It also has a poetic meaning as "the gift that came from Portugal"

@dailyculture

The oldest discovered instrument was a recorder-like whistle made from a bear's femur dating from 60,000 years ago

@museumofhistory

Taylor Swift retires from music to become yoga teacher to the stars.

@swifty_swap_x

The language of graphic design revised and updated: An illustrated handbook for understanding fundamental design principles by Richard Poulin

@classicalassoc

$1 billion dollar NFT sale replaces previous most expensive artwork ever sold - Leonardo da Vinci, Salvator Mundi - which sold for $450.3 million at Christie's in New York in 2017

@thegrndset

Another Louvre heist?
Souvenir shop looted.

@pixelperfectpix

Teaching and Assessing Visual Storytelling Techniques for Multimodal Composing With a Digital Camera or Phone. By Ewa McGrail; J. Patrick McGrail

@canvaschronicle

Before You Write Your Novel, Essential Skills for the First-time Novelist by James McCreet

@sketchypencils

Jonas Brothers shock the world with reveal that they are not actually related.

@shocking_truths

"Hearing music as . . .": Metaphors evoked by the sound of classical music. Schaerlaeken et al. 2019

@colorpopdaily

Disney announces cancellation of Marvel Cinematic Universe, states "We've done enough."

@artistsalley

Learning to draw anatomy?
Grab the **Anatomy 101 Guidebook** for just $14.99, out biggest sale ever.

@aiartisart

Ethical Algorithms: AI and Social Media

Ethical Matrix Worksheet

Ethical Matrix

Stakeholders	Prioritize Safety		Prioritize Engagement	
	Pros	Cons	Pros	Cons
Users				
Platform				
Investors				

Ethical Matrix (Example)

Stakeholders	Prioritize Safety		Prioritize Engagement	
	Pros	**Cons**	**Pros**	**Cons**
Users	The platform is more trustworthy, users don't have to wonder if everything they see is fake or mis/disinformation.	Content may not be as interesting or engaging.	Content may be more exciting and relevant to interests.	Content may not be trustworthy or legitimate.
Platform	More trust established with users, which may lead to more word-of-mouth growth.	Higher costs to ensure safety. Users may spend less time on the platform.	More engagement with users, leading to them spending more time on the platform.	May develop a reputation for being untrustworthy and/or unethical.
Investors	More trust with users may lead to them being willing to purchase platform subscriptions or from advertisers.	Users may not spend as much time on the platform, leading to lower sales.	More time on the platform means more users, and lower safety barriers means targeted advertising is easier.	Higher risk for controversy, leading to lower financial performance or boycotting.