



Session 2 “Online Safety: Harassment and Filter Bubbles” (3 hours)

PURPOSE

- This session will enable students to recognize online harassment, understand how filter bubbles and algorithms intensify online harassment and explore ways to engage against it.

KEYWORDS

- Digital safety, digital security, sexting, cyberbullying, hate speech, filter bubble, algorithm and filter bubble

RESOURCES

- DW Akademie MIL Guidebook, 3rd edition, p. 176 – 200. Especially: 176-179; 193
- **Fact Sheet Algorithm:**
 - A long time ago, social media news feeds showed users updates from people they follow in chronological order. Today, algorithms determine which posts to distribute to which users.
 - The intent of any platform's algorithm is to keep the user on that platform as long as possible. The individual user's behavior and other factors determine when algorithms deliver content to which user (e.g., age, gender, interests, posts they have liked...).
 - Social media algorithms are constantly evolving and being refined. Algorithms will vary depending on the platform, but whatever the formula, the outcome is always the same: getting the user the content they want to see.



MEDIA AND INFORMATION LITERACY

- Algorithms are used in social media to deliver content to the user. Not all social media platforms use algorithms, however, many have adopted news feeds that are delivered via an algorithm in recent years. Newsfeeds use algorithms to determine the content to show you based on your interests, activity, and interactions on the platform.
- Social media algorithms treat engagement (likes, comments and shares) as a sort of snowball effect. The more engagement a piece of content gets, the more likely it is to be rewarded by the algorithm.
- Algorithms filter and display news content which are likely to match their users' preferences (interests, hobbies, political, ideological etc.). Such selective exposure can increase polarization if users receive most of their news from social media.
- The standard definition of an algorithm is ...
An algorithm is a finite series of instructions - well-defined and computer-implementable - to solve a problem, or a class of problems, or to perform a computation or function.
Therefore, in simple terms the meaning of the algorithm is ...
An Algorithm is a set of rules that solve a problem.
- Example Instagram
Three key factors determine the ranking of posts:
 - Interest – How much Instagram thinks you want to see a post. This is based on past content you've viewed, liked and interacted with.
 - Recency – How recently the content was posted, prioritizing timely posts over week-old posts.



MEDIA AND INFORMATION LITERACY

- Relationship – Your relationship with the user who posted the content. Users you've interacted with most in the form of comments, likes, tags and direct messages will appear higher in the news feed.

Furthermore, three secondary factors also determine the ranking of posts in the news feed.

- Frequency – How often you open the app, as the algorithm will show you the best post since your last visit.
- Following – The more people you follow the less frequent you'll see multiple posts from the same people.
- Usage – The time you spent in the app will also determine which posts Instagram choose to show you. The more time you spend in the app, the more variety of posts you will see.

(see <https://www.hopperhq.com/social-media-marketing-glossary/algorithm/>)

- Video: Social media algorithms explained – CBC Kids news

<https://youtu.be/F5tz887wXCY>

- **Fact Sheet Filter Bubble:**

- A filter bubble or ideological frame is a state of intellectual isolation that can result from personalized searches when a website algorithm selectively guesses what information a user would like to see based on information about the user, such as location, past click-behavior and search history. As a result, users become separated from information that disagrees with their viewpoints, effectively isolating them in their own cultural or ideological bubbles. The choices made by



MEDIA AND INFORMATION LITERACY

these algorithms are not transparent. Prime examples include Google Personalized Search results and Facebook's personalized news-stream. (see https://en.wikipedia.org/wiki/Filter_bubble)

- To increase user engagement, social media companies connect users with ideas they are already likely to agree with, thus creating echo chambers of users with very similar beliefs.
- Data is used to “train” the algorithms that determine what we see in our newsfeeds. In the case of social media, our online behaviors are the data. In general, a “like” means positive engagement. Therefore, this becomes a signal that is fed into the training of the machine learning model (the algorithm) and used to update and refine the representation of the user, meaning, the set of preferences that the algorithm has learned for this particular user. The machine learning models are aiming at determining what is the best next thing that a user would be most likely to engage with, for example, “like” or “give 5-star rating.” Given that the model is trained with this as a primary objective, it will favor content that resembles content that the user has already engaged with, and this means that, in the vast ocean of content being shared in a platform, it will most likely rank other content lower. (see <https://news.ucr.edu/articles/2021/02/03/how-burst-your-bubble-broadening-your-social-media-horizons>)
- Video: Flat Earth: How did YouTube help spread a conspiracy theory?
<https://www.bbc.com/news/av/stories-49021903>



TRAINING METHODS

- Topic game
- Guided discussion
- Group work and presentation
- Active media work (video production)
- Feedback discussion

POSSIBLE “TRAINING OF TRAINERS” ELEMENTS

- A student conducts a topic game
- A student visualizes a discussion outcome
- A student moderates a Feedback session after the presentation of Media products
- A Student conducts the conclusion (“Wall of Love”)

HOW TO CONNECT TO PREVIOUS MEET-UP “VLOG”

- Social media is a good platform to learn, to interact with others, and to access information, but it is also risky in terms of online harassment. Sharing online content can make a user vulnerable to online harassment. Social media algorithms may increase the visibility of these attacks or even cause them to go viral, which can lead to even more or more severe online attacks. Algorithms lead to social media users only seeing what the algorithm wants them to see and interact with. In the long run, users mostly get to see posts from people who hold similar opinions as they do and that reconfirm their beliefs. Algorithms prevent users from accessing diverse media content. This creates filter bubbles and echo chambers.



LEARNING OUTCOMES – THE HEAD (KNOWLEDGE)

- Students recognize online harassment, explain how it emerges and engage against it.
 - Students recognize how algorithms and filter bubbles can intensify online harassment.
-

LEARNING OUTCOMES – THE HAND (SKILLS)

- Students counteract online harassment. They circumvent algorithms and prevent filter bubbles by diversifying the sources and contents they consume on social media.
-

LEARNING OUTCOMES – THE HEART (ATTITUDE)

- Students behave carefully and respectfully on social media and help put an end to online harassment.



INTRODUCTION (15 MIN)

Introduction: Topic game: Snowball fight of ideas

| | | |
|------------------|--|----------------|
| Activity | <ul style="list-style-type: none"> ● Topic game “Snowball fight of ideas”: (1) online harassment, cyber-bullying and hate speech (2) ideas to counter them | 15 ” |
| Objective | <ul style="list-style-type: none"> ● Students think about an example of online harassment, cyber-bullying or hate speech as a risk on social media and are encouraged to develop ideas to counter it | |
| Steps | <ul style="list-style-type: none"> ● Put the three flipcharts with definitions of online harassment, cyber-bullying, and hate speech up for everyone to see ● Trainers ask students to read the definitions and make sure there is a basic understanding before you play the topic game. The terms were covered in the basic MIL Club (Unit 2, Session 7) ● Trainers ask students to take a piece of paper and write down one example of online harassment, cyber-bullying or hate speech on social media ● Students crunch their pieces of paper like a ball and throw them at one another a few times, like in a snowball fight. After a while, trainers ask students to stop ● Trainers ask each student to pick up the nearest paper ball, unfold it, read what was written on the paper and think of a reaction to counter it. Each student adds their idea on their paper ● Students crumple the papers up again and start the next round of snowball fighting | |



MEDIA AND INFORMATION LITERACY

| | |
|-----------------|---|
| | <ul style="list-style-type: none"> • Trainers again ask each student to pick up the paper ball closest to them, unfold it and read what was written on the paper. If that student has a different strategy against the mentioned example or harassment, cyberbullying or hate speech, they should add it to their paper • Play three to four rounds before stopping the snowball fight • Ask students to unfold their papers and hang them up on a pin board. You could let them cluster the examples for online harassment, cyberbullying and hate speech. Give everyone a few minutes to read all the risks and suggestions for counteractions • Trainers ask students to reflect on the game. What do they think about the problems and the ideas to counter them, and what was new to them? |
| Material | <ul style="list-style-type: none"> • Pieces of paper, pens, flipchart • Markers |
| Notes | <ul style="list-style-type: none"> • Prepare three flipcharts with short definitions of “Online harassment”, “Cyberbullying” and “Hate speech”. Leave room on the flipcharts for additional key words and ideas that will be developed in the group discussion after the topic game • Visualize instructions on a flip chart • Carefully explain how to play the game to the students. Make sure everyone knows what a snowball looks like and how it works • Before starting to play, give students an opportunity to ask clarifying questions if they do not fully understand the instructions or how to play the game |



MEDIA AND INFORMATION LITERACY

| | |
|-------------------------------|---|
| Possible ToT - element | <ul style="list-style-type: none"> Reflect on the game with the students afterwards and help them link it to the following session |
| Possible ToT - element | <ul style="list-style-type: none"> Let a volunteering student conduct this topic game. The volunteer will need instructions on how to conduct the game before the session begins. Make sure the volunteer has understood the objective of the game, how to play it and how to reflect on it and connect it to the following session. <p>After this ToT-element, the volunteer should get feedback from the group and from the MIL Club trainers on what they did well and how they could improve as trainers</p> |

BODY PART 1 (MIXTURE OF TRAINER AND TRAINEE CENTERED) (75 MIN)

Trainee Centered: Group discussion: Reasons for online harassment, cyber-bullying and hate speech

| | | |
|------------------|---|----------------|
| Activity | <ul style="list-style-type: none"> Group discussion: Students share their knowledge and ideas about reasons for online harassment, cyberbullying and hate speech | 20 ” |
| Objective | <ul style="list-style-type: none"> Students recognize various forms and reasons of online harassment, cyberbullying and hate speech on social media | |
| Steps | <ul style="list-style-type: none"> Trainers ask students to share their experience with online harassment, cyberbullying or hate speech. It can be either their own experience or what they heard or read about it | |



MEDIA AND INFORMATION LITERACY

| | |
|-----------------------------|---|
| | <ul style="list-style-type: none"> ● Trainers engage the trainees in a discussion by asking questions such as <ul style="list-style-type: none"> - what kind of harassment, cyberbullying or hate speech have you seen? - who did what to whom? Who was harassed? By whom? - Why did it happen? - What dynamics did it develop? - Why does harassment, cyberbullying or hate speech on social media take on a different quality than in real life / in person? ● Trainers add characteristics of online harassment, cyberbullying and hate speech to the three flipcharts so everyone is clear on these terms again |
| Material | <ul style="list-style-type: none"> ● Flipchart / whiteboard to visualize key points of the discussion ● Markers |
| Notes | <ul style="list-style-type: none"> ● This training segment is just a reminder of topics covered in the basic MIL Club. That's why trainers can keep it short here. ● Coordinate the discussion ● Encourage students to share ideas ● Trainers need to prepare examples to make sure the students are not confused |
| Possible ToT element | <ul style="list-style-type: none"> ● Let a volunteering student visualize this discussion on flipcharts or a whiteboard <p>Make sure the volunteer has all the key points and writes clearly so everyone can see</p> <p>After this ToT-element, the volunteer should get short feedback</p> |



from the group and from the MIL Club trainers on what they did well and how they could improve as trainers

Give additional hints on how to make visualizations easy to understand and interesting (colors, shapes, icons, planning ahead what space you will need...)

Trainer Centered: How the online disinhibition effect promotes online harassments, cyber-bullying and hate speech

| | | |
|------------------|---|----------------|
| Activity | <ul style="list-style-type: none"> Guided discussion (see DW Akademie MIL Guidebook, p. 176 – 177) | 15 ” |
| Objective | <ul style="list-style-type: none"> Students identify the “online disinhibition effect” and understand the reasons why it intensifies online harassment, cyberbullying and hate speech Students understand how online harassment and hate speech affect social media users in Cambodia. | |
| Steps | <ul style="list-style-type: none"> Trainers add to the previous discussion and help the students discover why online harassment, cyberbullying and hate speech happen. Trainers differentiate between personal reasons (posts or comments directed against people one knows in online-harassment or cyberbullying) and ideological reasons (posts or comments directed against groups or ideas in hate speech). Trainers supply additional information Trainers explain the “online disinhibition effect” as extreme or radical social harassment because people do not see each other. | |



| | |
|-----------------|--|
| | <p>Within this context, the attackers seem not to feel guilty of hurting anyone, even though their actions generate huge negative impacts on the victims and society (also see DW Akademie MIL Guidebook, p. 177)</p> |
| Material | <ul style="list-style-type: none"> • Flipcharts “Online harassment”, “Cyberbullying” and “Hate speech” (from previous activity) • Flipchart to explain the online disinhibition effect • Makers and cards |
| Notes | <ul style="list-style-type: none"> • Visualize instructions on a flip chart • Research (e.g., DW Akademie MIL Guidebook p. 176 – 179) |

Trainee Centered: Group work Train to explain: Effects of algorithms and filter bubbles

| | | |
|------------------|---|----------------|
| Activity | <ul style="list-style-type: none"> • Group work “Train to explain”: Effects of algorithms and filter bubbles | 40 ” |
| Objective | <ul style="list-style-type: none"> • Students read, understand, and evaluate definitions of algorithms and filter bubbles. They summarize the key points • The students explain how algorithms and filter bubbles influence the contents of social media news feeds and how they often intensify online harassment, cyberbullying and hate speech | |



MEDIA AND INFORMATION LITERACY

| | |
|--------------|---|
| | <ul style="list-style-type: none">• Students apply the “train to explain” method and gain training experience |
| Steps | <ul style="list-style-type: none">• Ask students to form two groups and explain the task<ul style="list-style-type: none">– One group will work on the term “algorithm”, the other will research “filter bubble”. Their task will be to understand, summarize and explain their term to the other group– As a starting point for their research, each group will get a worksheet with selected explanations, facts, and links to explainer videos– At the end of the group work, each group should be able to explain to the other group:<ul style="list-style-type: none">→ The meaning of its term (i.e., algorithm or filter bubble)→ What effect it has on a social media newsfeed→ What effect it has on the spread of online harassment, cyberbullying and hate speech• One volunteer from each group will be the “lead trainer”, who has two minutes to explain the group’s term and its effect. They can prepare a visualization if they like with a “co-trainer”. (Task and preparation: 15 minutes)• After each two-minute presentation, the other group has five minutes to ask the “lead trainer” and the “co-trainer” questions. (presentation: 2 x 10 min = 20 min) |



MEDIA AND INFORMATION LITERACY

| | |
|-----------------------------|--|
| | <ul style="list-style-type: none"> • Correct the group presentations and add information where necessary. Have a short feedback round for the lead-trainer and the co-trainer at the end (10 min) |
| Material | <ul style="list-style-type: none"> • Prepared worksheets with facts / definitions / explanations / links to explainer videos. Try to find Khmer definitions or additional explainer videos. • Flipcharts • Markers and cards |
| Possible ToT element | <ul style="list-style-type: none"> • Let a volunteering student visualize this discussion on flipcharts or a whiteboard • Make sure the volunteer has all the key points and writes clearly so everyone can see • After this “Train to explain” segment, the presenters should get feedback from everyone and from the MIL Club trainers on what they did well and how they could improve as trainers • Give additional hints on how to make visualizations easy to understand and interesting (colors, shapes, icons, planning ahead what space you will need...) |

----- BREAK (15 MIN) -----



BODY PART 2 (MIXTURE OF TRAINER AND TRAINEE CENTERED) (75 MIN)

Trainee Centered: Topic Game Post scramble – Counter hate

| | | |
|------------------|--|----------------|
| Activity | <ul style="list-style-type: none"> • A variation of the topic game “Post Scramble” – DW Akademie MIL Guidebook p. 184. | 15 ” |
| Objective | <ul style="list-style-type: none"> • Students practice how to react against online harassment • Students understand how algorithms intensify online harassment | |
| Steps | <ul style="list-style-type: none"> • Ask students to sit in a circle that represents a big group of social media users. Ask one volunteer to stand in the middle. There should be no empty chair in the circle • Explain that the person in the middle is the only one who can “post” something (a message, a piece of information, or a comment) by saying it out loud • For this game, these “posts” should be examples of online harassment or hate-speech • Examples could be: <ul style="list-style-type: none"> ○ “Men should not be allowed to drive cars.” ○ “Women only care about fashion.” ○ “People who eat Durian stink.” • Explain to the students that for the duration of this role-play, they can pretend to be mean. Strong statements and reactions will increase the dynamics of the game • If trainers think the game might get out of hand, they can suggest posting about fruits, objects, or colors instead of people (e.g., | |



MEDIA AND INFORMATION LITERACY

“Durian are dumbest fruits in the world” or “Pink is such an ugly, color, it should be banned from the earth.”

- When the first volunteer has said their “post” out loud, the other social media users can react to it. They can “like” or “dislike” (thumbs up or down) and “post” a comment or counter-speech to the original post. The others react again, and so on
- For the sake of this game, students don’t have to be “politically correct”, they can be bad and “post” mean things, just like many people would on social media. Remind them that this is just a role play, trying to re-enact how algorithms work
- Everyone who wants to react must get up from their chair and go to the middle of the circle. Here, they can “post” by saying their comment out loud. Then they should find a new chair that someone else has vacated because they also want to “post” or react
- If no one wants to react anymore, one person is left in the middle without a seat. They can then formulate a new “post”
- The trainer or the person in the middle can stop a running “conversation” at any time and replace it with a new “post.” End the game if you think trainees are getting bored or if the game is getting out of hand
- One volunteer or co trainer counts the number of interactions each post gets before it is replaced by a new post (i.e., how long is the thread of a post)
- At the end of the game, this scorekeeper reminds everyone how many reactions each post got before it was replaced by a new



MEDIA AND INFORMATION LITERACY

| | |
|-----------------|---|
| | <p>post. The more people reacted to the original post or statement, the more this thread would have shown up on other newsfeeds and “gone viral”. This shows how algorithms can intensify forms of online harassment</p> <ul style="list-style-type: none"> • Trainers do a short reflection and remind everyone that this was just a role play. In real life, the students shouldn’t be mean and reinforce online harassment, hate speech or cyber-bullying when they encounter it in social media. Instead, they should react with counter speech, de-escalate and report the offenders • Trainers link to the next activity -- how algorithms and filter bubbles intensify online harassment and what forms of counter speech are successful |
| Material | |
| Notes | <ul style="list-style-type: none"> • Visualize instructions on a flipchart • Carefully explain how to play the game to the students • Before starting to play, give students an opportunity to ask clarifying questions if they do not fully understand the instructions or how to play the game • Reflect on the game with the students afterwards and help them link it to the following session |

Active media work: Tips and tutorials on how to fight online harassment, cyber-bullying, hate speech; explainer videos on algorithms and filter bubbles



MEDIA AND INFORMATION LITERACY

| | | |
|------------------|---|----------------|
| Activity | <ul style="list-style-type: none"> ● Group work, active media work ● Groups create videos Topics: <ul style="list-style-type: none"> - Practical tips what to do when you encounter online harassment - Practical tips what to do when you encounter cyberbullying - Practical tips what to do when you encounter hate speech - What are algorithms and how can we override them? - What are filter-bubbles and how can we get out of them? | 50 ” |
| Objective | <ul style="list-style-type: none"> ● Students offer solutions to online harassment, cyberbullying, hate speech, algorithms and filter bubbles ● Students plan and produce videos with tips ● Students give each other constructive feedback on their media products | |
| Steps | <ul style="list-style-type: none"> ● Trainers divide students into groups (3-4 groups) ● Trainers ask each group to create videos with tips on one of the above-mentioned topics ● Students can do some research and planning. They should spend about 25 minutes on producing their videos ● Students upload their videos to the Facebook group and do a short presentation for 1-2 minutes ● Students and trainers reflect/give feedback on the media products. What was done well, what could be improved? How clear is the | |



| | |
|----------------------|--|
| | message? How valid / important is it for other youths in Cambodia? |
| Material | <ul style="list-style-type: none"> • Flipchart for task explanation • Markers/cards • Computer/smartphones/LCD projector |
| Notes | <ul style="list-style-type: none"> • Visualize instructions on a flipchart • Carefully explain the group work task to the students • Before going into the groups, give students an opportunity to ask clarifying questions if they do not fully understand the instructions |
| TOT - Element | <ul style="list-style-type: none"> • Let a trainee moderate the feedback session after the presentation of the videos <p>After this ToT-element, the volunteer should get feedback from the group and from the MIL Club trainers on what they did well and how they could improve as trainers</p> |

CONCLUSION (10’)

Conclusion Wall of love

| | | |
|------------------|---|----------------|
| Activity | <ul style="list-style-type: none"> • Students write a compliment on a card and stick on the wall of love | 10 ” |
| Objective | <ul style="list-style-type: none"> • Students show their appreciation for each other • Students realize what others appreciate them for or what others value about them | |



MEDIA AND INFORMATION LITERACY

| | |
|--------------------------|--|
| Steps | <ul style="list-style-type: none"> • Every student writes their name on a card and folds it. The trainers collect all folded cards in a hat, bag, or container • Each trainee draws a card with a name on it. If someone draws their own name, they should put the card back in the container and draw another card • Trainers ask each student to think of one thing that impressed them about this classmate in today's training. The students write this on the name card they drew • It may be a good idea if the compliments are anonymous and no one knows who wrote about them • After everyone has finished writing, collect the cards and stick them on the wall of love |
| Material | <ul style="list-style-type: none"> • Cards/markers • Sticky tape |
| Notes | <ul style="list-style-type: none"> • Visualize instructions on a flipchart • Carefully explain the task to the students • Before starting the activity, give students an opportunity to ask clarifying questions if they do not fully understand the instructions |
| ToT - Element | <ul style="list-style-type: none"> • Let a volunteer be the trainer for this game. Make sure the volunteer has understood the objective of the game, how to play it, knows how to explain it and how to reflect on it at the end. After this ToT-element, the volunteer should get feedback from the group and from the MIL Club trainers on what they did well and how they could improve as trainers |