

AI Skills,  
Powered  
by Values



## Module 8: Final Capstone Module

# The AI-assisted apology message



[www.valuesai.eu](http://www.valuesai.eu)



Co-funded by  
the European Union

Module 8 © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)



WHY THIS MATTERS

# AI has moved into our most personal conversations.

## THE SHIFT

A decade ago, AI helped us spell-check emails. Today it drafts apologies, break-up messages, condolence notes and answers to our parents. The technology is the same - the stakes are very different.

When a generator writes on our behalf in an emotional situation, it is not just producing text. It is shaping a relationship.

## Speed

AI drafts in seconds. Repair takes time we no longer give it.

## Polish

Smooth wording can hide blame, deflection or factual errors.

## Distance

It feels safer to let a machine speak - but the other person still receives the words.

# Anatomy of a sincere apology

A framework from conflict research (Lazare; Lewicki) - four moves that make repair feel real.

01

## Acknowledge

Name what happened, clearly, with the other person's experience in mind.

### AVOID

*Vague: "if anyone was upset..."*

02

## Accept responsibility

Own the part you are responsible for. Do not shift blame to the upset person.

### AVOID

*"You took it the wrong way"*

03

## Repair

Address the impact. Offer something concrete where it is possible and honest.

### AVOID

*Empty gesture or deflection*

04

## Change

Describe what will be different. Only promise what you can actually keep.

### AVOID

*"We will never do this again"*

**Takeaway** A short, honest message that hits all four moves usually lands better than a long, polished one that skips any of them.

# Where AI helps - and where it fails

*Language models are strong at form and weak at truth. That matters more in conflict than almost anywhere else.*

## WHERE IT HELPS

- ✓ **Structure**  
Suggests a clear order - acknowledge, impact, next step.
- ✓ **Register**  
Calms a hot draft into a calmer tone you can then own.
- ✓ **Options**  
Offers alternative phrasings when the right word is stuck.
- ✓ **Accessibility**  
Helps non-native speakers and people with dyslexia produce readable text.

## WHERE IT FAILS

- ✗ **Facts**  
Invents details (“hallucination”). In a conflict, invented facts re-open the argument.
- ✗ **Blame-shifting**  
Default apology templates often imply it was the other person’s reaction, not the act.
- ✗ **Overpromising**  
Confident phrasing (“this will never happen again”) lands as dishonest when it can't be kept.
- ✗ **Privacy**  
Draws in names, screenshots or context that should not be published in a public chat.

# Four guardrails for AI-assisted messages

Before you send anything, a machine helped you write, check it against these four.



## Verify

Every factual claim about what happened must be checked with the people involved - not with the AI.

**Ask yourself:** *Would I stake my reputation on this sentence being accurate?*



## Fairness

Consider who is visible, who is missing and who could be harmed if this message is screenshot and forwarded.

**Ask yourself:** *Would the person I'm writing about feel this is a fair summary?*



## Privacy

Names, identifying details and third-party information should be removed unless there is a clear reason to keep them.

**Ask yourself:** *Does this message share something that is not mine to share?*



## Disclosure

In a conflict, being honest about AI assistance builds trust; hiding it often breaks trust if it later comes out.

**Ask yourself:** *Would it feel wrong if the recipient learned AI wrote part of this?*

HOW TO PLAY

# The AI-assisted apology message

*An interactive decision game. Read a scenario, apply your tools, make a choice and see what happens.*

**1**

**Scenario**

A group-chat apology

**3**

**Paths**

A · B · C

**9**

**Outcomes**

Discussion + reflection

# The scenario & the golden rule

*Set the stage before you make a single click.*

## THE SCENARIO

**A disagreement in your youth group has spilled into a group chat.**

A message was posted that upset someone. A leader asks your group to help draft an apology that will be shared in the chat.

Your task: decide how to handle the apology: with AI, without AI, or somewhere in between.

## GOLDEN RULE

**Use at least two tools before every decision.**

## YOUR TOOLKIT - pick any two

- STOP - THINK - CHECK - CHOOSE
- THINK: True · Helpful · Intent · Neutral · Kind
- Fairness check - who is included, missing, harmed?
- Verification habits - pause, check details, check source
- Pause · Verify · Protect · Support · Report

# How a round works - five steps

*Read → apply tools → choose → read outcome → reflect.*

1

**Read the start card**

Look at the scenario on card B0. Notice the decision prompts on the right.

2

**Pick a path: A, B or C**

A = use AI, send it.  
B = AI draft + edit.  
C = write it yourselves.

3

**Apply two tools**

Before clicking, name the two toolkit items you are using and why.

4

**Read the outcome**

See what happens in the chat and discuss the three open questions.

5

**Capture reflection**

On the reflection slide, record your path, your tools and a safer follow-up action.

**TIP** Different paths lead to very different outcomes - compare choices with your group before moving on to the reflection slide.

# Module 8 Capstone

AI Decision Game

## The AI-assisted apology message

---

[Start](#)

*Use at least two tools before each decision*

## How to play

# Navigate the scenario like a digital decision game

This deck is set up for classroom discussion, self-study or live facilitation.

### Flow

1. Read the start card
2. Pick a route: A, B or C
3. Apply at least two tools
4. Click into the next card
5. Read the outcome and discuss the open questions
6. Capture evidence on the reflection slide

### Suggested facilitation moves

- Ask learners to name the two tools they used before clicking
- Pause after each outcome and compare different paths
- Invite learners to edit the wording live and explain why
- Use the reflection slide as an evidence record or exit task

### Deck map

**B0 Start card**



**B-A, B-B or B-C**



**Outcome cards**



**Reflection**

Jump into the toolkit first if you want the prompts visible before starting.

[Open toolkit](#)

[Title](#)

[Start card](#)

[Toolkit](#)

[Reflection](#)

*Use at least two tools per decision*

# Decision tools to apply before each choice

Use at least two prompts every time you make a decision in the game.

## Core prompts

- STOP–THINK–CHECK–CHOOSE
- THINK = True, Helpful, Intent, Neutral, Kind
- Fairness check = who is included, who is missing, who could be harmed
- Verification habits = pause, check details, check source

## Extended prompts

- Pause–Verify–Protect–Support–Report
- Create → Verify → Fairness check → Disclose
- NSL, Consent Gate, Fair Data Check, SAFE Workflow Ladder



[Start the scenario](#)

# The apology draft

Choose a path. Each button links to the next scenario or outcome card.

## Scenario

A disagreement in your youth group has spilled into a group chat. A message was posted that upset someone. A leader asks your group to help draft an apology message that will be shared in the chat.

### Decision prompts

Pause before posting  
Check what is accurate  
Consider privacy and fairness

## Choose one

### Choice A

Use AI to write the apology and send it straight away.

### Choice B

Use AI to draft, then edit it in your own words and check the details first.

### Choice C

Do not use AI. Write it yourselves to avoid mistakes.

# The AI message looks polished

Choose a path. Each button links to the next scenario or outcome card.

## Scenario

The AI draft sounds calm and confident. It includes one line about what happened that you are not sure is accurate. It also names the person who was upset.

### Decision prompts

Verification first  
Do not name people casually  
Shared facts matter

## Choose one

Choice A1  
Send it as written.

Choice A2  
Remove the name and the disputed line,  
then send it.

Choice A3  
Pause and ask the group to confirm what  
actually happened before sending  
anything.

# Editing the draft

Choose a path. Each button links to the next scenario or outcome card.

## Scenario

You use AI to draft, then you start editing. The AI draft includes a phrase that sounds like it is blaming the upset person (“if you took it the wrong way...”). It also includes a promise (“we will never do this again”) that you cannot guarantee.

### Decision prompts

Check tone for blame  
Do not promise what you cannot keep  
Choose a calmer process

## Choose one

### Choice B1

Keep the phrase and promise because it sounds professional.

### Choice B2

Rewrite the message to focus on impact and remove promises you cannot keep.

### Choice B3

Share the draft with the whole chat first and ask people to vote on it.

# Writing without AI

Choose a path. Each button links to the next scenario or outcome card.

---

## Scenario

You write your own apology. It is short and honest, but you are worried it might sound too blunt. Someone suggests adding an AI-generated line to “smooth it out”.

### Decision prompts

Clarity matters  
Tone still needs checking  
AI should not replace accuracy

## Choose one

### Choice C1

Keep it short and send it as-is.

### Choice C2

Add a single AI-suggested sentence, then edit it to match your voice.

### Choice C3

Ask AI to rewrite the whole message to sound more persuasive.

# Sent quickly, challenged publicly

Outcome card and debrief prompts

## What happens

The message is posted quickly. Some people respond positively. Others reply asking about the disputed line. The named person reacts in the chat and the conversation becomes active again.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. Which details would you want to verify before sending a message like this?
2. What is the risk of naming someone in a public group chat?
3. What would you do if someone replies with a different version of events?

# Edited for privacy, but still prompts questions

Outcome card and debrief prompts

## What happens

The message is posted without naming anyone. Some people accept it. A few ask for more clarity about what happened. The person who was upset reads the message but does not respond immediately.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What information is necessary in an apology and what is optional?
2. How can you acknowledge impact without restarting the argument?
3. What would you do next if the upset person messages privately?

# Shared facts before posting

Outcome card and debrief prompts

## What happens

You do not post yet. You ask for a quick check of what happened. People share different recollections. The group agrees on a short, shared summary before drafting the apology.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What is a fair way to agree a shared summary when people disagree?
2. How can you keep the process calm and respectful?
3. When is it better to move the conversation out of the group chat?

# Professional tone, unsafe wording

Outcome card and debrief prompts

## What happens

The edited message is posted with the blaming phrase and a strong promise. Some people react with thumbs-up. One person replies pointing out the blaming wording. Another says the promise sounds unrealistic.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. Which part of the message could be read as blame and why?
2. What promises are safe to make and which are risky?
3. How could you rewrite the message to focus on impact and repair?

# Impact-focused repair

Outcome card and debrief prompts

## What happens

The final message focuses on impact and avoids blame. It includes one clear next step, for example “we will keep the chat respectful and check before we share”. Several people respond positively and the chat slows down.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What makes a repair message feel sincere rather than scripted?
2. Which tools helped you decide what to change?
3. What is one follow-up action that supports the words in the apology?

# Crowdsourcing brings the argument back

Outcome card and debrief prompts

## What happens

You share the draft to the whole chat for a vote. People comment on wording, tone and details. Some suggestions conflict. The conversation becomes long and the original issue reappears in replies.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What are the risks of crowdsourcing wording in a live group chat?
2. What alternative process could reduce conflict and still include voices?
3. What would you do if someone shares screenshots outside the group?

# Honest but too brief for some readers

Outcome card and debrief prompts

## What happens

You send the short message. Some people appreciate it. A few people interpret it as cold or too brief. The person who was upset replies asking for a clearer acknowledgement of impact.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What information or wording might make this feel more caring?
2. How can you be clear and kind at the same time?
3. What would you add without turning it into a long explanation?

# Light AI support, voice retained

Outcome card and debrief prompts

## What happens

You add one AI-suggested sentence, then edit it. The final message stays short but includes a clearer acknowledgement of impact. The person who was upset responds with a short reply and the chat quiets.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What is one sign that an AI sentence needs editing before sending?
2. What would you disclose, if anything, about AI assistance in a group message?
3. How would you check that the tone is fair to everyone involved?

# Persuasive rewrite, shaky facts

Outcome card and debrief prompts

## What happens

You ask AI to rewrite the whole message. It becomes longer and more persuasive. It includes a phrase that is not accurate to what happened. Some people challenge the message in replies.

Discuss, annotate or jump to the reflection slide to capture evidence.

## Open questions

1. What should you verify when AI rewrites a personal situation?
2. How can persuasive wording create problems in a conflict situation?
3. What would you do next if people accuse the message of being “fake” or “scripted”?

# Quick reflection and portfolio prompts

Use this slide as a learner record, plenary prompt or screenshot template.

## Record your reasoning

Which path did you choose?

---

Which two tools did you use?

---

What did you verify?

---

Who could be affected or excluded?

---

What would you disclose about AI use, if anything?

---

What is one safer follow-up action?

---

## Facilitator prompts

- What made the message feel sincere or unsafe?
- Where did accuracy matter most?
- What changed after the fairness check?
- When should the conversation move out of the public group chat?
- What evidence would show responsible AI use here?