

# SolutionChat: Real-time Moderator Support for Chat-based Structured Discussion

Sung-Chul Lee (KAIST)

Jaeyoon Song (SNU)

Eun-Young Ko (KAIST)

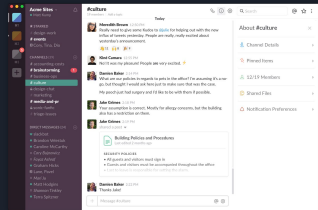
Seongho Park (KAIST)

Jihee Kim (KAIST)

Juho Kim (KAIST)

# Background

- Chat as a channel for problem-solving and decision-making



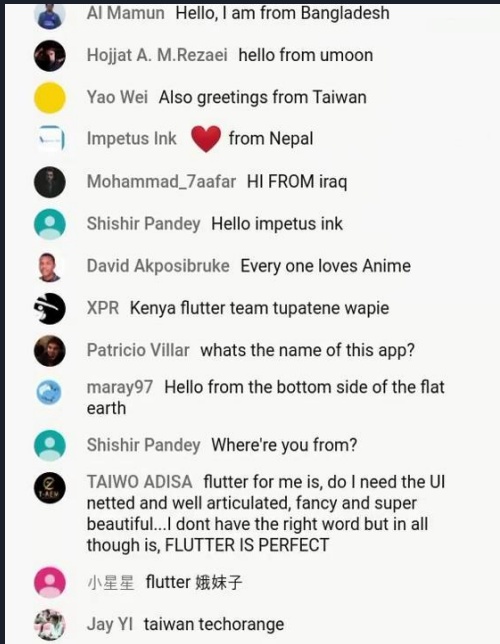
Comcast  
employees in a  
Slack channel



Collectively pulled a protest in a  
self-organized way

<https://twitter.com/SedaGirl/status/902620987602092032>

# Background - Challenges in Online Chat Discussion (1)



Fast message flow and chaotic argument sharing [1]

## Background - Challenges in Online Chat Discussion (2)

For this stage, What was our consensus?



Difficult for participants to keep track of the discussion and follow up a missed conversation

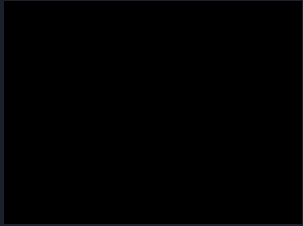
## Background - Challenges in Online Chat Discussion (3)

### Moderator's burden for various supports (examples)



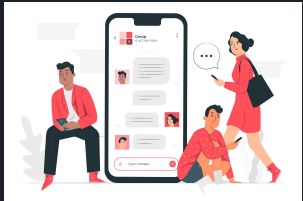
Task management	"For this stage, we will discuss pros of the solution" "Shall we vote?" "We have X minutes left" "Shall we move to the next stage?"
Argument building	"What is the evidence for that?" "Do you have any idea?"
Contribution management	"I think P1 is not talking" → "I want what P1 thinks"
Encouraging	"Thank you for your opinion"

# Background - Challenges & Approach



Fast message flow and chaotic argument sharing [1]

Support structured discussion in chat



Difficult for participants to keep track of the discussion and follow up a missed conversation

Support discussion Stage Tracking



Moderator's burden

Support moderator's tasks



# Video Demo



## Related Work: The Role of Moderators (1)

- Improving the quality of students' discussion for students' learning gain [3].

Assisting discussion [1, 2, 6, 7, 11, 23, 37, 42]

- e.g., Support argument building

Stimulate discussants [3]

- e.g., “Can you add something here?”

[1] Christa SC Asterhan and Baruch B Schwarz. 2007.

[2] Christa SC Asterhan and Baruch B Schwarz. 2009.

[3] Christa SC Asterhan and Baruch B Schwarz. 2010.

[6] Christine Chin and Jonathan Osborne. 2010.

[7] Elaine B Coleman. 1998.

[11] Erica De Vries, Kristine Lund, and Michael Baker. 2002.

[23] Alison King and Barak Rosenshine. 1993.

[37] BB Schwartz, Y Neuman, and S Biezuner. 2000

[42] Carla Van Boxtel, Jos Van der Linden, and Gellof Kanselaar. 2000.



## Related Work: The Role of Moderators (2)

- Moderators should provide various support during the discussion

Used as a discourse taxonomy	Managerial Support (Task management)	"For this stage, we will discuss pros of the solution" "Shall we vote?" "We have X minutes left" "Shall we move to the next stage?"
	Pedagogical Support (Argument building)	"What is the evidence for that?" "Do you have any idea?"
	Interaction Support (Contribution management)	"I think P1 is talking" → P1, Do you have any idea?
	Social Support	"Thank you for your Idea"



## Related Work: Moderation in online communities

- Online communities often moderate content and user behaviors
- Content moderation
  - Automated moderation [21]
  - Hybrid (human + automation) moderation [20]
- Process moderation
  - Automated repetitive process helpers [41]
- Challenges
  - Algorithmic moderation in response to the dynamics of group discussion
    - limited NLP performance [25]
    - high cost of failed interactions

[20] Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. 2019a.

[21] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2019b

[25] Lorenz Cuno Klopfenstein, Saverio Delpriori, Silvia Malatini, and Alessandro Bogliolo. 2017.

[41] Niels van Berkel, Jorge Goncalves, Danula Hettiachchi, Senuri Wijenayake, Ryan M. Kelly, and Vassilis Kostakos. 2019.

# Related Work: Discussion Summarization and Real-time Message Recommendation



Deliberatorium

- Topic summarization
  - Deliberatorium [24] - a tree-structured network
  - Wikum [47] - multi-level and recursive summarization workflow
  - Tilda [46] - chat message markup



ConsensusUs

- Consensus summarization
  - ConsensusUs [29] - visualizes participants' consensus for multi-criteria decision
  - ConsiderIt [26] visualizes the level of agreement

[24] Mark Klein. 2011.

[26] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012.

[29] Weichen Liu, Sijia Xiao, Jacob T Browne, Ming Yang, and Steven P Dow. 2018.

[46] Amy X Zhang and Justin Cranshaw. 2018

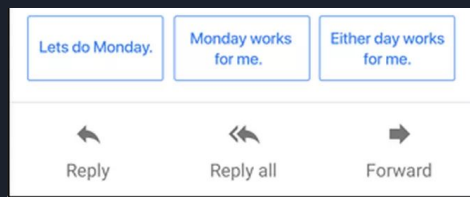
[47] Amy X Zhang, Lea Verou, and David Karger. 2017.

# Related Work: Discussion Summarization and Real-time Message Recommendation



Keyboard  
Prediction

- Message recommendation
  - Keyboard applications (Emojis, Words) [17, 39, 32, 40]
  - Gmail's Smart Reply [18]
  - Gmail's Smart Compose [5] suggests words and phrases as the user



Gmail's Smart Reply

- [5] Mia Xu Chen, Benjamin N Lee, Gagan Bansal, Yuan Cao, Shuyuan Zhang, Justin Lu, Jackie Tsay, Yinan Wang, Andrew M Dai, Zhifeng Chen, and others. 2019.
- [17] Google. 2016. GBoard - the Google Keyboard. (May 2016). <https://play.google.com/store/apps/details?id=com.google.android.inputmethod.latin> Accessed: 2019-09-20.
- [18] Matthew Henderson, Rami Al-Rfou, Brian Strope, Yun-hsuan Sung, László Lukács, Ruiqi Guo, Sanjiv Kumar, Balint Miklos, and Ray Kurzweil. 2017.
- [32] Microsoft. 2017. Word Flow Keyboard. (2017). <https://www.microsoft.com/en-us/garage/profiles/word-flow-keyboard/> Accessed: 2019-09-20.
- [39] TouchPal. 2008. TouchPal Keyboard. (2008). <http://www.touchpal.com/> Accessed: 2019-09-20.
- [40] TouchType. 2010. SwiftKey Keyboard. (2010). <https://swiftkey.com/en> Accessed: 2019-09-20.



# Formative Study

Participants	18 participants
Group	6 groups (3 members per group)
Compensation	12.5 USD for a hour

Condition	Pre-selected structure	Moderator
Structure only	To all discussants	No
Moderator only	No	Yes
Moderator+Structure	To moderator only	Yes

Chat (40min)

Survey (20min)  
& Interview

# Formative Study Observations (1)

Structure

Covered most of the structures

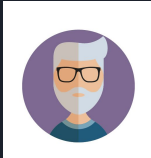
More aspects

No structure

Made their own structure  
(e.g. pros and cons)

Fewer aspects

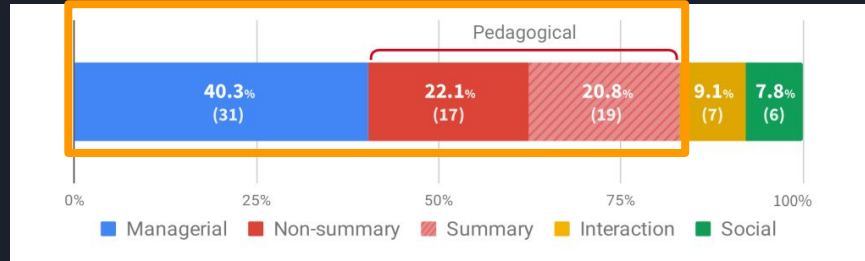
Condition	Pre-selected structure	Moderator
Moderator+Structure	To moderator only	Yes



Discussants' wanted to see the discussion structure

## Formative Study Observations (2)

- Moderator Messages (counts and ratios)



- The managerial support and pedagogical support take the major share of moderator messages.



# Design goals

G1. Assist discussion stage management by exposing the discussion structure and highlighting the current stage to all participants.

Structure

More aspects

G2. Reduce moderators' constant burden in summarizing throughout the discussion.

G3. Facilitate moderators' managerial support by assisting with repetitive managerial messages.

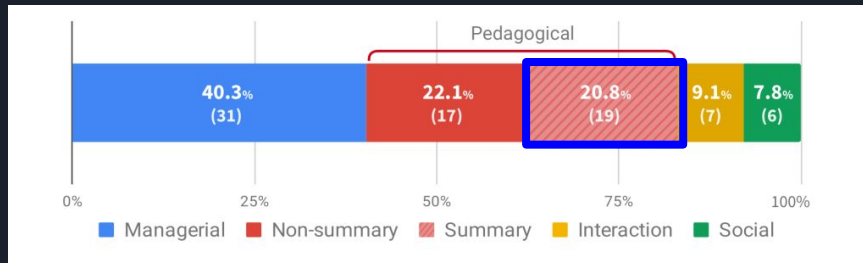
G4. Facilitate moderators' pedagogical support by assisting with repetitive pedagogical messages.



# Design goals

G1. Assist discussion stage management by exposing the discussion structure and highlighting the current stage to all participants.

G2. Reduce moderators' constant burden in summarizing throughout the discussion.



G3. Facilitate moderators' managerial support by assisting with repetitive managerial messages.

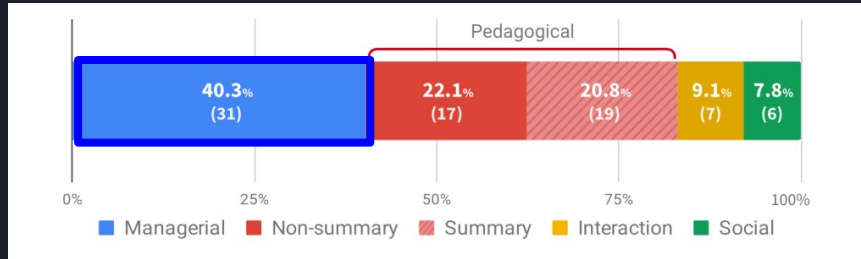
G4. Facilitate moderators' pedagogical support by assisting with repetitive pedagogical messages.

# Design goals

G1. Assist discussion stage management by exposing the discussion structure and highlighting the current stage to all participants.

G2. Reduce moderators' constant burden in summarizing throughout the discussion.

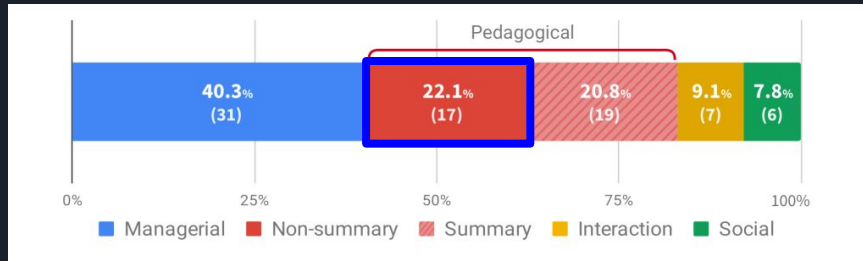
G3. Facilitate moderators' managerial support by assisting with repetitive managerial messages.



G4. Facilitate moderators' pedagogical support by assisting with repetitive pedagogical messages.

# Design goals

- G1. Assist discussion stage management by exposing the discussion structure and highlighting the current stage to all participants.
- G2. Reduce moderators' constant burden in summarizing throughout the discussion.
- G3. Facilitate moderators' managerial support by assisting with repetitive managerial messages.
- G4. Facilitate moderators' pedagogical support by assisting with repetitive pedagogical messages.



# What is SolutionChat?

- A chat platform that
  - Visualizes multi-stage discussion structure and its related featured opinions
  - Supports moderators in real-time by providing moderator message recommendation

The screenshot displays the SolutionChat interface for a discussion titled "Subjectivity of the grading system". The interface is divided into several panels and sections:

- Panel A (Agenda Panel):** Located on the left, it contains a sidebar with "Problem Representation" and "Solution Generation & Evaluation".
- Panel B (Current Stage and Featured Opinions):** The main content area, showing a sequence of questions and answers. The current stage is "What can be the solution?", with a "Next >" button. Below it, two featured opinions are listed: "Making the whole grading process public (1/1)" and "Make an absolute grading scale (0/1)". A text input field with the placeholder "Press enter to give your answer" is at the bottom.
- Panel C (Stage Divider):** A horizontal bar at the top right of the main content area, indicating the current stage.
- Panel D (Add To Featured Opinions):** A button labeled "Add to list" next to a message.
- Panel E (Message Recommendation):** A section showing recommended messages, including "Now let's talk about possible solutions" and "Making the whole grading process public".
- Panel F (Message Recommendation):** A section showing recommended messages, including "I wonder what P3 thinks. Can you tell us your opinion?", "Any more ideas?", and "Should we vote now?".

At the bottom, there is a text input field with the placeholder "Press enter to send message".

A: Agenda Panel

B: Current Stage and Featured Opinions

C: Stage Divider

D: Add To Featured Opinions

E-F: Message Recommendation

# Overview of SolutionChat

## Subjectivity of the grading system



Problem  
Represent  
-ation

**Solution  
Generation &  
Evaluation**

Previous: What evidence do you have to justify the cause of the problem?

What can be the solution?

Next >

-  Making the whole grading process public (1/1)
-  Make an absolute grading scale (0/1)

Press enter to give your answer

What are the advantages of the solution?

What are the disadvantages of the solution?

What makes this better compared to the other ones?

Discussion structure  
with a current stage indicator  
(Agenda Panel - AP)

Participants



P1

You've discussed for total 12 mins

What can be the solution?



Now let's talk about possible solutions



Okay

suggest make public the process of grading to as many

for this?

P2

I agree

I wonder what P0 thinks. Can you tell us your opinion?

● Facilitate active participation

Any more ideas?

● Facilitate active participation

Should we vote now?

● Encourage voting to make consensus

Press enter to send message

# Overview of SolutionChat

## Subjectivity of the grading system



Problem  
Represent  
-ation

Previous: What evidence do you have to justify the cause of the problem?

**Solution  
Generation &  
Evaluation**

What can be the solution?

Next >

-  Making the whole grading process public (1/1)
-  Make an absolute grading scale (0/1)

Press enter to give your answer

What are the advantages of the solution?

What are the disadvantages of the solution?

What makes this better compared to the other ones?

**Featured opinions  
for the current stage**

Participants



You've discussed for total 12 mins

What can be the solution?



Now let's talk about possible solutions



Okay



Since transparency is the key issue here, I suggest make public the process of grading to as many



Can you elaborate?



I agree

for this?

I wonder what P0 thinks. Can you tell us your opinion?

● Facilitate active participation

Any more ideas?

● Facilitate active participation

Should we vote now?

● Encourage voting to make consensus

Press enter to send message

# Moderator message support

## Subjectivity of the grading system

Problem  
Represent-  
ation

Previous: What evidence do you have to justify the cause of the problem?

Solution  
Generation &  
Evaluation

What can be the solution? Next >

✖ Making the whole grading process public (1/1)

✖ Make an absolute grading scale (0/1)

Press enter to give your answer

What are the advantages of the solution?

What are the disadvantages of the solution?

What makes this better compared to the other ones?

Participants

M me P1

You've discussed for total 12 mins

What can be the solution?

M me

Now let's talk about possible solutions

P1

Okay

P1

Since transparency is the key issue here, I suggest make public the process of grading to as many people as possible.

I see

Thanks for your idea

Any evidence for this?

↳

M me

Can you elaborate?

P2

I agree

Moderator message recommendation (Inline MR)

Moderator message recommendation (Block MR)

I wonder what P0 thinks. Can you tell us your opinion?

• Facilitate active participation

Any more ideas?

• Facilitate active participation

Should we vote now?

• Encourage voting to make consensus

Press enter to send message

## Subjectivity of the grading system

Problem  
Represent  
-ation

Previous: What evidence do you have to justify the cause of the problem?

**Solution  
Generation &  
Evaluation**

What can be the solution? Next >

☒ Making the whole grading process public (1/1)

☐ Make an absolute grading scale (0/1)

Press enter to give your answer

What are the advantages of the solution?



What are the disadvantages of the solution?

What makes this better compared to the other ones?


## Agenda Panel


- Provides a discussion structure [G1]
- Highlights the current stage [G1]
- Displays key opinions [G2]
- Displays vote status [G3]







Participants   You've discussed for total 12 mins


What can be the solution?


 Now let's talk about possible solutions

 Okay

 Since transparency is the key issue here, I suggest make public the process of grading to as many people as possible.


  


↳  Can you elaborate?


 I agree

**Moderator message recommendation (Inline MR)**

**Moderator message recommendation (Block MR)**

 I wonder what P0 thinks. Can you tell us your opinion?  
● Facilitate active participation

 Any more ideas?  
● Facilitate active participation

 Should we vote now?  
● Encourage voting to make consensus

...

Press enter to send message

## Message Recommendation (MR)

- Moderator focused
- Managerial and pedagogical message focused [G3-4]
- Recommended message intent related messages
- Recommends state related messages

# Message Recommendation

MR	Type	Condition	Recommended Message Example	References
Inline	Social	NLU (opinion)	"I see.", "Thank you for your opinion."	[42, 31]
	Pedagogical	NLU (opinion)	"Can you provide some evidence?" "Can you elaborate?"	[42, 31]
Block	Managerial	Discussion stage change	"In this stage, we will set the goal of the discussion."	[30]
		Every three minutes	<div> <div>P1</div> <div>Since transparency is the key issue here, I suggest make public the process of grading to as many people as possible.</div> <div> <div>I see</div> <div>Thanks for your idea</div> <div>Any evidence for this?</div> </div> <div> <div>Inline MR</div> </div> </div>	
	Pedagogical	A FO takes 3 min. after FO added	<div> <div>↳</div> <div> <div>M me</div> <div>Can you elaborate?</div> </div> </div>	
		FO count > 3	<div> <div>"Can we try to find evidence for the featured opinions?"</div> <div>"Are there any other opinions?"</div> <div>"Let's check if our featured opinions are biased."</div> </div>	[42] [31]
	Interaction	Every three minutes	"I wonder what XX thinks. Can you tell us your opinion?"	[31]

**Table 2.** The recommendation messages of Inline MR and Block MR. Inline MR recommends social and pedagogical messages for discussant's opinion messages and Block MR recommends primarily managerial and pedagogical messages based on the information of AP.

# Message Recommendation

MR	Type	Condition	Recommendation	
Inline	Social	NLU (opinion)	"I see.",	<div> <div>I wonder what P0 thinks. Can you tell us your opinion? ● Facilitate active participation</div> <div>Any more ideas? ● Facilitate active participation</div> <div>Should we vote now? ● Encourage voting to make consensus</div> </div>
	Pedagogical	NLU (opinion)	"Can you elaborate?"	
Block	Managerial	Discussion stage change	"In this stage, we will set the goal of the discussion."	[30]
		Every three minutes	"We have X minutes left for our discussion."	[25]
			"Let's talk about this for X more minutes."	
			"Can we move faster since we are running out of time?"	
		A FO takes a majority vote	"Shall we proceed to the next stage?"	[31]
		3 min. after the last FO entry	"Shall we vote now?"	[31]
	Pedagogical	3 min. after the last FO entry FO added	"Are there any other opinions?"	[31]
			"Can we try to find evidence for the featured opinions?"	[42]
		FO count > 3	"Are there any other opinions?" "Let's check if our featured opinions are biased."	[31]
Interaction	Every three minutes		"I wonder what XX thinks. Can you tell us your opinion?"	[31]

Block  
MR

**Table 2.** The recommendation messages of Inline MR and Block MR. Inline MR recommends social and pedagogical messages for discussant's opinion messages and Block MR recommends primarily managerial and pedagogical messages based on the information of AP.

# Message Recommendation

MR	Type	Condition	Recommended Message Example	References
Inline	Social	NLU (opinion)	"I see.", "Thank you for your opinion."	
	Pedagogical	NLU (opinion)	"Can you provide some evidence?" "Can you elaborate?"	
Block	Managerial	Discussion stage change	"In this stage, we will set the goal of the discussion."	
		Every three minutes	"We have X minutes left for our discussion." "Let's talk about this for X more minutes." "Can we move faster since we are running out of time?"	
		A FO takes a majority vote 3 min. after the last FO entry	"Shall we proceed to the next stage?" "Shall we vote now?"	
	Pedagogical	3 min. after the last FO entry	"Are there any other opinions?"	[31]
		FO added	"Can we try to find evidence for the featured opinions?"	[42]
		FO count > 3	"Are there any other opinions?" "Let's check if our featured opinions are biased."	[31]
	Interaction	Every three minutes	"I wonder what XX thinks. Can you tell us your opinion?"	[31]

What can be the solution?

Next >



Making the whole grading process public

(1/1)



Make an absolute grading scale

(0/1)

Press enter to give your answer

Featured  
Opinions

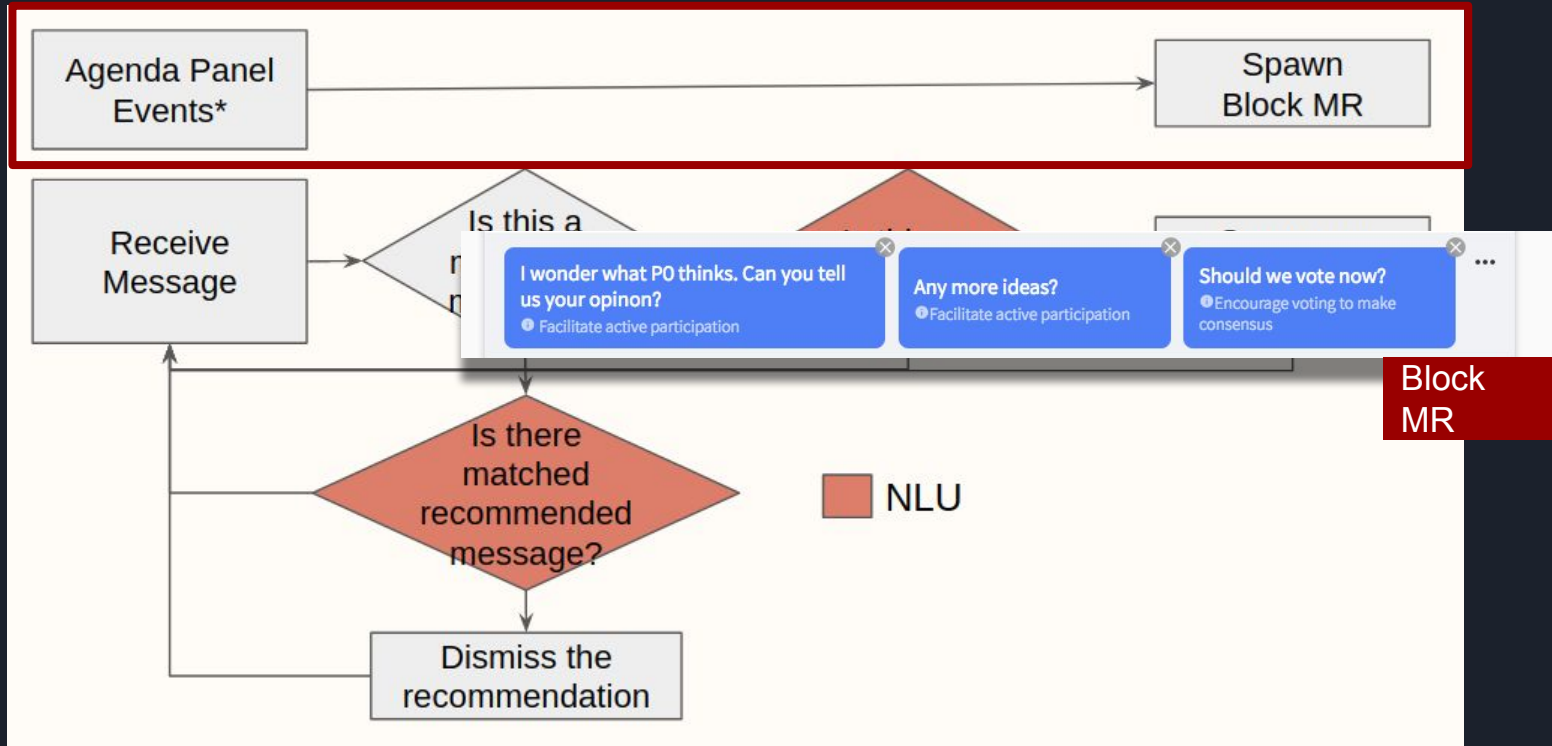
**Table 2.** The recommendation messages of Inline MR and Block MR. Inline MR recommends social and pedagogical messages for discussant's opinion messages and Block MR recommends primarily managerial and pedagogical messages based on the information of AP.

# Message Recommendation

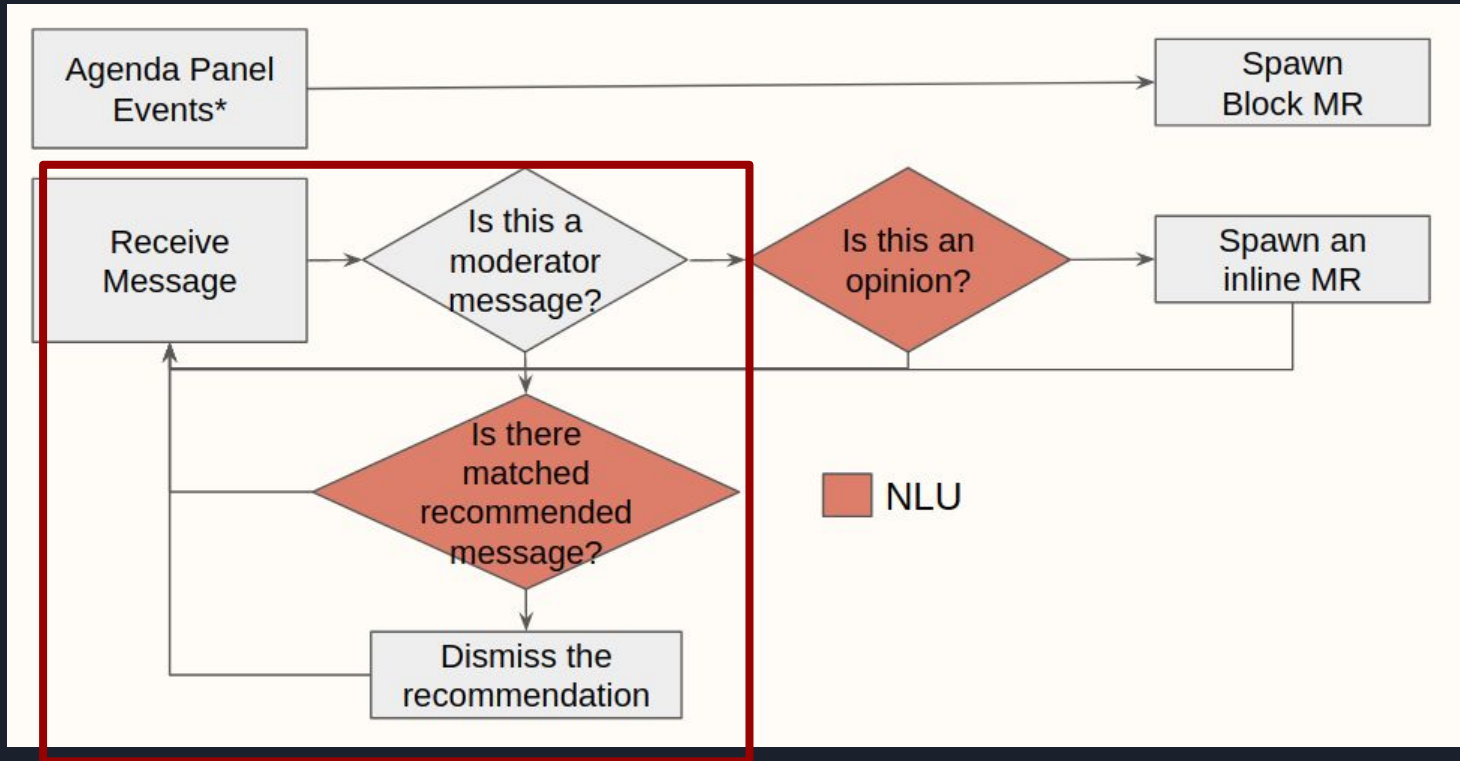
<i>MR</i>	<i>Type</i>	<i>Condition</i>	<i>Recommended Message Example</i>	<i>References</i>
Inline	Social	NLU (opinion)	"I see.", "Thank you for your opinion."	[42, 31]
	Pedagogical	NLU (opinion)	"Can you provide some evidence?" "Can you elaborate?"	[42, 31]
Block	Managerial	Discussion stage change	"In this stage, we will set the goal of the discussion."	[30]
		Every three minutes	"We have X minutes left for our discussion." "Let's talk about this for X more minutes." "Can we move faster since we are running out of time?"	[25]
		A FO takes a majority vote	"Shall we proceed to the next stage?"	[31]
		3 min. after the last FO entry	"Shall we vote now?"	[31]
	Pedagogical	3 min. after the last FO entry	"Are there any other opinions?"	[31]
		FO added	"Can we try to find evidence for the featured opinions?"	[42]
		FO count > 3	"Are there any other opinions?" "Let's check if our featured opinions are biased."	[31]
	Interaction	Every three minutes	"I wonder what XX thinks. Can you tell us your opinion?"	[31]

Table 2. The recommendation messages of inline MR and Block MR. Inline MR recommends social and pedagogical messages for discussant's opinion messages and Block MR recommends primarily managerial and pedagogical messages based on the information of AP.

# Recommendation Flow

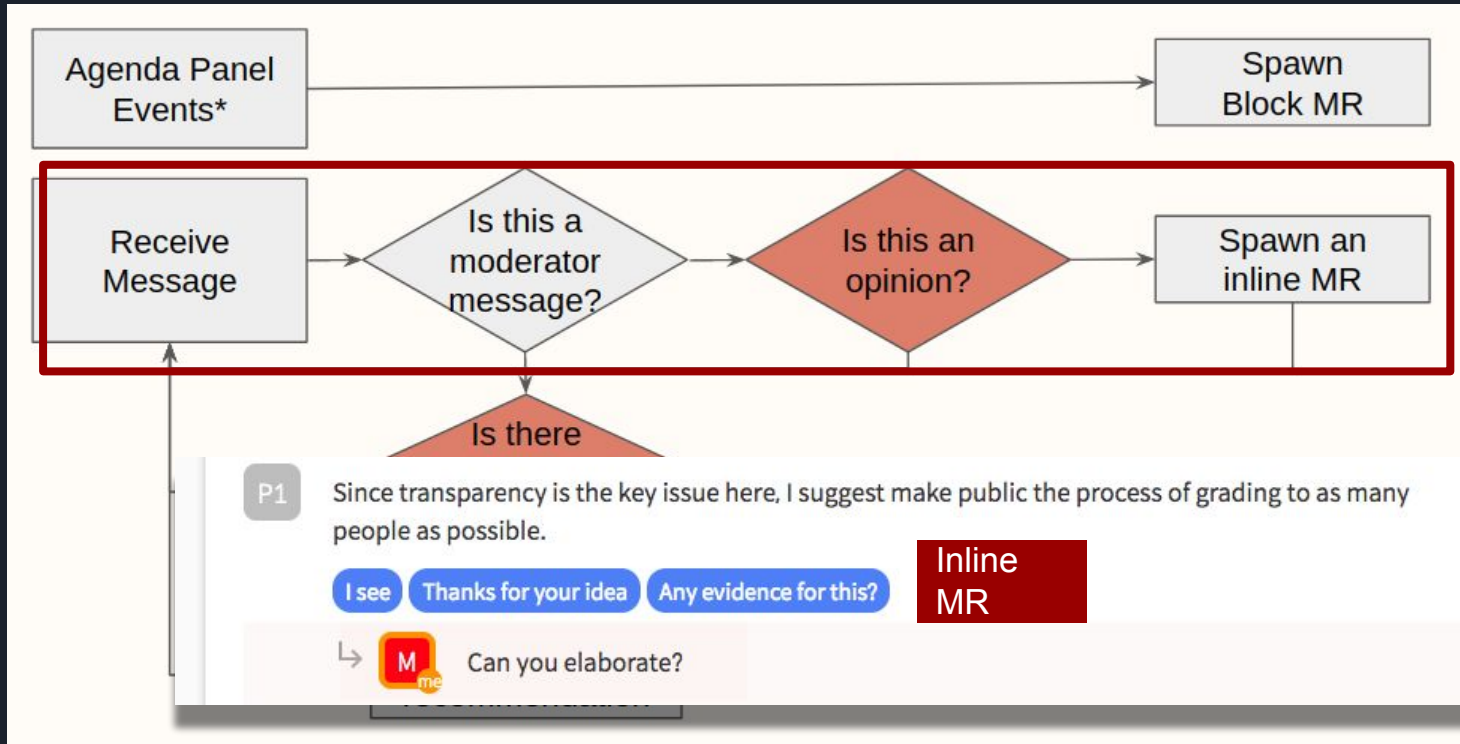


# Recommendation Flow





# Recommendation Flow







# Evaluation

Participants	55 participants from two Korean universities
Group	4-5 members, totals 12 groups
Compensation	18 USD for two hours
Configuration	Within-subjects

## Discussion topics

- Subjectivity in the academic grading system
- The inconvenience of the course registration system
- Low quality of cafeteria food

# Conditions

- **Baseline :** Structure
- **AP :** Structure Agenda Panel
- **AP + MR :** Structure Agenda Panel Message Recommendation
- Within-subjects, each group experienced all three conditions with randomized order

Session A

Chat (20min)

Survey (5min)

Rest (5min)

Session B

Chat (20min)

Survey (5min)

Rest (5min)

Session C

Chat (20min)

Survey (5min)



# Hypotheses

H1. Participants hold higher awareness of the discussion structure with AP.

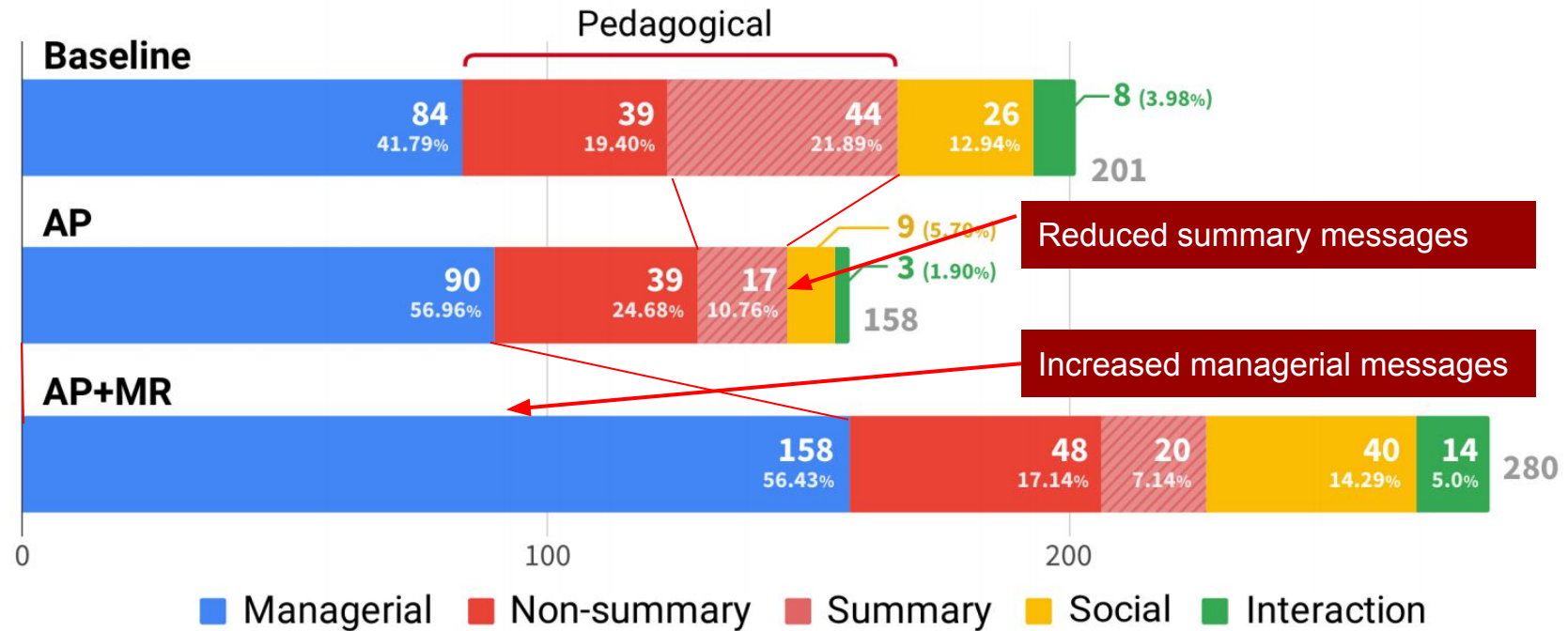
H2. Moderators provide better summarization support with fewer summarization messages with AP.

H3. Moderators provide better managerial support with more managerial messages with MR.

H4. Moderators provide better pedagogical support with more pedagogical messages with MR.

# Results (Moderator message count)

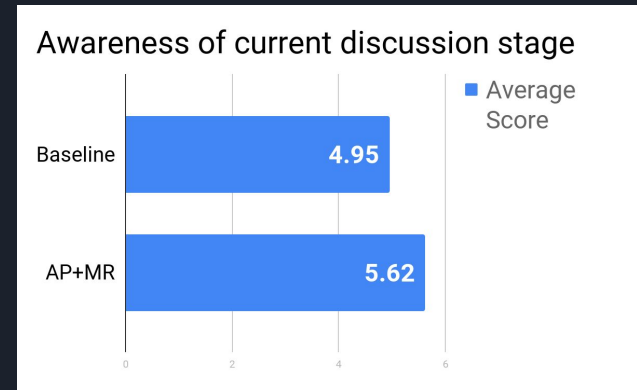
(x-axis = message count)



# Results

H1. Participants hold **higher awareness of the discussion structure with AP.**

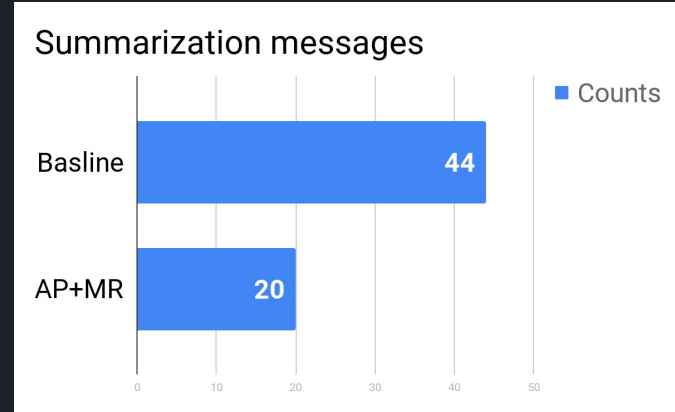
- Participants showed no significant differences on overall structure awareness between conditions
- **Participants showed a higher current stage awareness level in AP and AP+MR than the baseline (AP, AP+MR > Baseline)**
  - $p < 0.005$  for baseline - AP
  - $p < 0.05$  for baseline - AP+MR



# Results

H2. Moderators provide **better summarization support with fewer summarization messages** with AP.

- The moderators' summarization messages counts are significantly decreased in AP and AP+MR than the baseline (AP, AP+MR < Baseline)
  - $p < 0.05$  for baseline-AP,
  - $p < 0.05$  for baseline-AP+MR

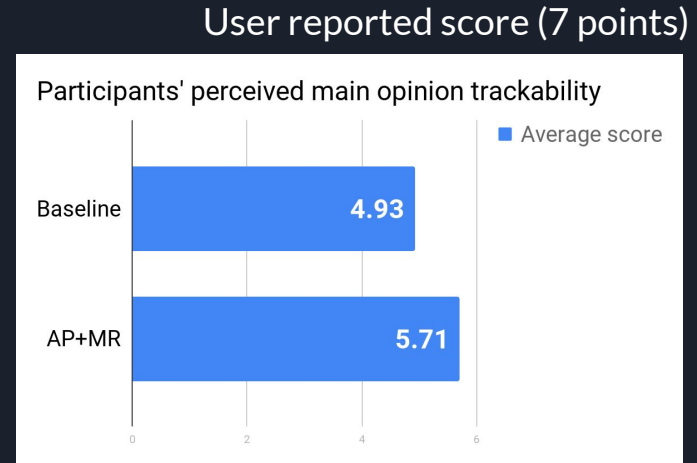


# Results

H2. Moderators provide better summarization support with fewer summarization messages with AP.

- The participants showed higher trackability on main opinions in AP and AP+MR than the baseline
  - $p < 0.005$  for baseline-AP
  - $p < 0.005$  for baseline-AP+MR

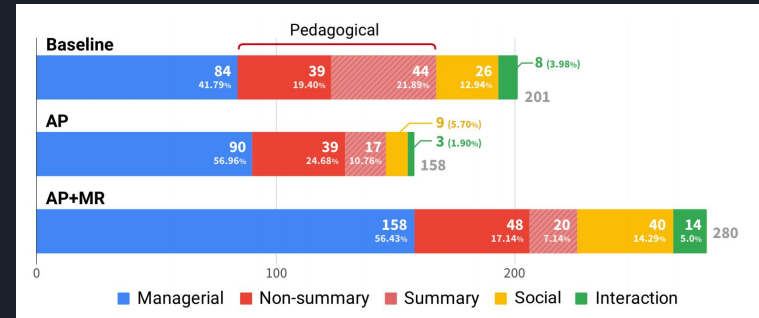
Used fewer summarization messages  
while retained user's perceived trackability



# Results

H3. Moderators provide better managerial support with more managerial messages with MR.

- The number of managerial prompts was significantly higher in AP+MR than the baseline
- Also discussants perceived that the moderators were better at
  - Stage introduction ( $p < 0.05$ )
  - Discussion management ( $p < 0.05$ )
  - Time management ( $p < 0.058$ )

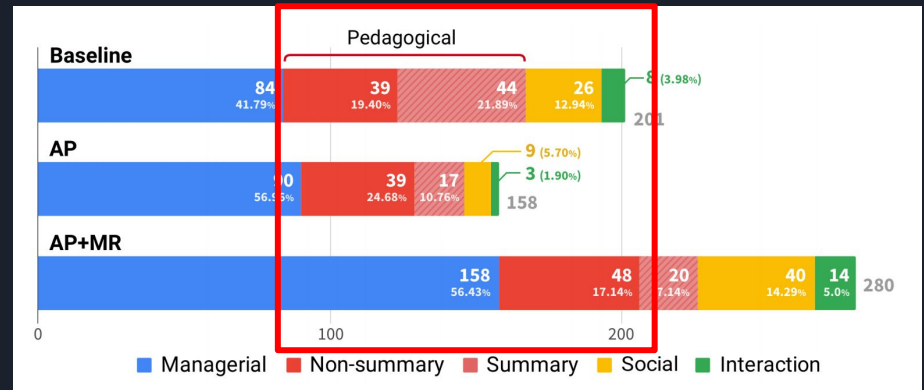




# Results

H4. Moderators provide better pedagogical support with more pedagogical messages with MR.

- No significant differences
- Some pedagogical messages (non-summary) are used to guide the direction or scope of the discussion, while discussants' need for such guide might have diminished with higher awareness and understanding of the discussion.





# Discussion

1. The importance of repetitive managerial messages
  - a. Moderators found it useful to get recommendations for repetitive messages.
2. The system should support diverse messaging styles of moderators.
  - a. ( "To go faster, " - Manually typed )  
("Shall we vote now?" - Taken MR)
3. The system should minimize the cost of inaccurately recommended messages.



## Takeaway messages

1. Reduce moderator's burden  
and promote other productive supporting  
tasks
2. Repetitive managerial supports are typical  
but demanding in real time setting

## Subjectivity of the grading system

Problem Representation

Previous: What evidence do you have to justify the cause of the problem?

**Solution Generation & Evaluation**

What can be the solution? [Next >](#)

- Making the whole grading process public (1/1)
- Make an absolute grading scale (0/1)

Press enter to give your answer

What are the advantages of the solution?

What are the disadvantages of the solution?

What makes this better compared to the other ones?

Participants **M** **P1** You've discussed for total **12 mins**

What can be the solution?

**M** Now let's talk about possible solutions

**P1** Okay

**P1** Since transparency is the key issue here, I suggest make public the process of grading to as many people as possible.

**I see** [Thanks for your idea](#) [Any evidence for this?](#)

**L** **M** Can you elaborate?

**P2** I agree

[I wonder what P0 thinks. Can you tell us your opinion?](#) [Any more ideas?](#) [Should we vote now?](#) ...

[Facilitate active participation](#) [Facilitate active participation](#) [Encourage voting to make consensus](#)

Press enter to send message

[solutionchat.kixlab.org](https://solutionchat.kixlab.org)

This work was supported by the Office of Naval Research (ONR: N00014-18-1-2834), and by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding).