https://doi.org/10.20965/jaciii.2010.p0825

Paper:

A Method for Using Discounted Utterances in Spontaneous Conversation

Hiroki Yamaguchi, Yukio Ohsawa, and Yoko Nishihara

Dept. of Systems Innovation, School of Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
E-mail: gutsu@panda.sys.t.u-tokyo.ac.jp, {ohsawa, nishihara}@sys.t.u-tokyo.ac.jp
[Received April 1, 2010; accepted June 22, 2010]

Using Discounted Utterances (DUs) in spontaneous conversation by applying text mining technology, extraction, and evaluation, we focused on DUs where values were buried in previous conversations. We discovered DU potentials by reconsidering them through human-computer interaction. Onlinechat experiments clarified DU features and demonstrated our system's importance. We found DUs involving (1) experiences shared by the subjects, (2) subjects' unique experiences, concerns, or beliefs, and (3) apparent unimportance or unrecognized potential. Results of the experiments showed our evaluation method to be appropriate for calculating DU importance when DUs involving (3) were valued significantly lower than (1) and (2). Experiments also suggested that most DUs extracted by the system were not indeed completely ignored but included subjects' unique stories involving main contexts. Such stories were based on subjects' unique experiences and may be useful for helping subjects' metacognition. The system may also enable non-subjects to infer subjects and their thinking.

Keywords: discounted utterance, discourse analysis, spontaneous conversation

1. Introduction

Communication is an important, ongoing human concern, but not everyone is good at communicating.

We sought to clarify communication from the viewpoint of engineering by focusing on Discounted Utterances (DUs) in conversation.

The system we developed for using DUs, consists of extraction and evaluation applying text mining techniques. Using onlinechat experiments for evaluation, we had subjects fill out questionnaires for extraction validity and DU importance evaluation.

Our extraction algorithm simply applies cosine similarity [1] and is good through human fuzziness enough to consider from many viewpoints, expanding the possibilities of creating new values.

Results of experiments showed that DUs were not totally ignored but were actually partly admitted by others because these DUs may include subjects' stories. Such a "side-story" was based on subjects' unique experiences and, although related to the main context, were often not understood by all subjects. Our system mined stories by extracting and reconsidering DUs. Content as such was not important, using DUs and focusing on such stories as triggers for reconsideration was itself important.

In onlinechat experiments, we discussed DUs to evaluate stories. In conversation analysis, we think it important to focus less on reproducibility than on whether subjects can enjoy conversation and build trusting relationships, which is why we discuss conversation dynamics. Experiments showed that extracted DUs were ignored not for the participants but for computers, meaning that a DU is what computers can recognize but human cannot, and vice versa, what human can recognize but computers cannot. This is essential in the viewpoint of human-computer interaction.

Experiments also clarified features, i.e., (1) the system helps subjects reconsider their concerns (metacognition [2]), (2) non-subjects can understand subjects' personalities better, and (3) the system activates human communication.

Background

Since our system applied text mining, we used a vector space model [3] simple enough to be applied to many situations in dealing with daily conversation. Daily conversation generally seems difficult to analyze because its structure is too floating and spontaneous.

Important contents are often distributed in daily conversations, so we should not use complicated methods. We thus simply applied cosine similarity to extract utterances.

Compared to other studies on summarization [4, 5], document clustering [6, 7], etc., our proposal focuses on DUs apparently not focused on before, meaning that DUs could not bring to bear their potentials.

Among the many studies developing communication support, Nishimoto et al. [8] proposed a topic-development agent that joins a daily conversation with human participants as an equal participant by replacing restrictions to keep the conversation lively. This processed a conversation based on surface information of each utterance, so only typical words (nouns and unknown-part-of-speech words) were taken into consideration in their system, causing oversights of implied contents, which we want to mine.

Matsumura et al. [9] proposed automatic indexing, focusing on term activity in text-based communication, which is useful for understanding context or dynamism in a conversation, but, contents that could not become a topic in the conversation were not taken into account. Despite much work focusing on utterances or keywords already bringing some value in conversation, little is known about potential DUs.

Hearst [10] proposed textTilling based on unities, in turn based on repetition of the same word. TextTilling uses word distribution bias. Because many methods exist in natural language processing, it is unclear whether we can create new values with these technologies. Among related research, a concept called Chance Discovery [11] makes it necessary to consider interaction between computing and human processes. Our proposal includes focusing on the interactions.

2. Theories and Experiments

2.1. Definition and Evaluation

We define the DUs as utterances ignored by others in conversations despite their potentials. DUs have low similarities. We propose two similarities with Eq. (1); one is between focused-on utterance and prior context (SIM_{before}), the other is between focused-on utterance and subsequent context (SIM_{after}):

$$\begin{aligned} SIM_{before} &= U \cdot C_{before} \\ SIM_{after} &= U \cdot C_{after} \end{aligned} \qquad (1)$$

U(Eq. (1)) is an utterance vector. U_k is a k-th utterance vector, $num(t_i)$ is the number of word t_i in the utterance (Eq. (2)). C_{before} and C_{after} are context vectors (Eq. (3)). Both similarities are calculated for each utterance:

$$U_k = (num(term1), num(term2), \ldots) \qquad (2)$$

$$\begin{aligned} C_{before} &= \sum_{k-w<i\leq k-1} U_i \\ C_{after} &= \sum_{k+1\leq i<k+w} U_i \end{aligned} \qquad (3)$$

$$SIM_{before} < \alpha \cap SIM_{after} < \beta \qquad (4)$$

The system evaluates similarities and extracts DUs using Eq. (4). w is used to set calculation-range. Users should set large w values for slow topic-shift conversations, and small values for fast topic-shift conversations. α and β are thresholds. Users can set these three values interactively. Although one set of parameters is for one analysis, users can and should try some patterns of the set for one analysis.

DUs are extracted by the above conditional expression, but raising a problem: short utterances consisting of a few words are extracted despite their intensity of feelings. Their similarity tends low due to the lack of word (information).

In the sections that follow, we evaluate DUs to solve the above problem using the Average Amount of Selected Information (AASI), in **Fig. 1** shown.

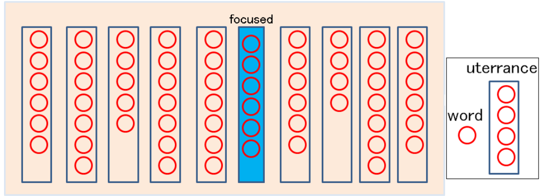

Fig. 1. AASI; information amount of a subclass in its parent class.

AASI is the amount of subset information in the parent class, defined as follows:

$$AASI(A, U_k) = \sum_{t_i \subset U_k} P_{U_k t_i} I_{At_i} \qquad (5)$$

P_{At_i}(Eq. (6)) is the probability of word t_i in word vector A, I_{At_i} (Eq. (7)) is the information amount of word t_i in word vector A, $num(A)$ is the total number of words in word vector A, and $num(t_i)$ is the number of word t_i in word vector A.

$$P_{At_i} = \frac{num(t_i)}{num(A)} \qquad (6)$$

$$I_{At_i} = -log P_{At_i} \qquad (7)$$

With AASI, the system calculates impact from an utterance on its back and forth contexts. The system infers the importance of each DU by considering impact differences between back and forth.

The system applies AASI for evaluating DU. We define two AASI values in the same way as similarities in Eq. (8), where, C_{before} and C_{after} are context vectors in Eq. (3) as follows:

$$\begin{aligned} E_{before} &= AASI(C_{before}, U_k) \\ E_{after} &= AASI(C_{after}, U_k) \end{aligned} \qquad (8)$$

E_{before} is an impact on the prior context, and E_{after} is an impact on the subsequent context. Evaluation Value (EV) is calculated as follows:

$$EV = |E_{before} - E_{after}| \qquad (9)$$

DUs with high EV are two types of DUs; (1) DUs impacting on the subsequent context but not relating to previous context or (2) those related to previous context but not to subsequent context. EV is the context shift between back and forth contexts.

Utterances with few words and utterances having no impacts on conversation have low EV and are treated as trivial by evaluation, as summarized in following.

DUs having high EV have high impact despite their low similarity. In contrast, DUs having low EV are ignored as trivial. Through EV, the system enables us to focus on and discover important DUs.

2.2. System Output and the Use

Based on EV, the system ranks DUs in descending order of rank, utterance (Utr) No., subject name, discounted utterances, and score (EV), as shown in **Table 1**.

Table 1. System output.

Rank	Utr No.	Subject	Discounted Utterance	Score
1	55	A	.....	0.38
2	34	B	.....	0.34
3	43	C	.....	0.30
4	70	D	.....	0.28
5	82	A	.....	0.25
....				

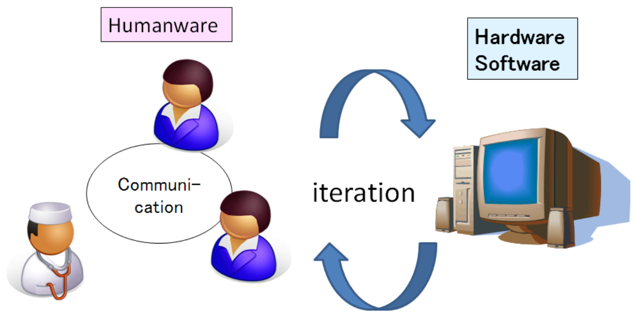

Fig. 2. System use: human-computer interaction.

Table 2. Questionnaire criteria.

value	Contribution	Intensity
1	Ignored	No intensity
2	A few contribution	Not so much
3	Chicken-and-egg	Chicken-and-egg Cannot answer
4	Adopted somewhat	Some intensity
5	Adopted	Strong intensity

Table 3. Extraction results: several DUs are picked up.

No.	Rank	Contents	Score
A2	3	The function is great!?	0.44
A2	5	I heard Prof. Y has that one.	0.036
A1	11	So it remains in the end	0.028
B1	4	Instead, please permit in Y	0.021
B2	6	They firmly pay the contribution.	0.066
B3	12	Courtyard after all	0.24
C1	1	Christmas song with Te (electronic musical instrument)	0.29
D1	2	Now, I want to play b-ball with the lab members!?	0.69
D2	4	Te(electronic musical instrument) featuring Y	0.29
D2	5	Shall we use that directivity loudspeaker?	0.24

From results, subjects can infer stories behind DUs and suggest new topics, as shown in **Fig. 2**.

The system consists of human and computing processes. (1) Subjects make face-to-face or non-face-to-face conversations, using conversation logs as input, (2) computers analyze logs and show results to subjects. (3) Subjects make new conversations based on the results. Subjects repeat processes from (1) to (3) until they are satisfied with the results. Subjects therefore can recognize their concerns which were not considered, and discuss the solutions more deeply and multilaterally.

2.3. Onlinechat Experiments

2.3.1. Conditions

We implemented onlinechat experiments to evaluate the system, in which subjects made onlinechats with themes. The themes were simple and had no obvious answers. Subjects were all Japanese and the chats were in Japanese. We need daily conversation data, because daily conversation often includes subjects' real intentions. For graduate students in the same laboratory, the experiments were conducted with 3- to 5-person groups. Each group made a few chats. The themes of the first chats were "Life in the laboratory." From the second chats themes were determined before the chat by the subjects after feeding back a previous chat's analysis.

The subjects were 14 people in 4 groups. Each experiment was about 30 minutes long, for seven experiments in all.

2.3.2. Post-Chat Questionnaires

After the chat-log was analyzed, questionnaires about the chat were generated automatically. The questionnaire consists of chat-log text (HTML files) and an answer sheet (CSV file) containing several utterances. Some were extracted DUs, and the others were randomly-selected utterances. These utterances were shuffled so that subjects could not identify which utterances were DUs. Each subject in the chat then answered for all utterances in the questionnaire, eventually, answering for both their own utterances and those of others'.

The questionnaire has two viewpoints respectively for evaluating by five-stages. One is the "Contribution." Contribution is defined as a measure of "whether the utterance was adopted or not." Strong responded utterances have high contribution. The other is the "Intensity." Intensity is defined as the strength of feelings within both the appearance and implied stories (no way to say, needless to say, etc.). Subjects answered both their own utterances and other speakers'. Against the other speakers' utterances, subjects supposed and answered the speakers' intensity. **Table 2** shows the evaluation criteria.

3. Results and Discussion

3.1. Extracted DUs

Table 3 shows a part of results. The rank in the table indicates the rank in each chat where the DU was uttered. No. indicates the experiment ID consisting of the group name and number of times. These are some DUs the authors could mine stories from. The set of parameters w,

Table 4. Experiments parameters sets.

No.	w	α	β
A1	20	0.20	0.25
A2	20	0.20	0.25
B1	20	0.30	0.35
B2	20	0.30	0.35
B3	20	0.20	0.25
C1	20	0.25	0.30
D1	20	0.30	0.35
D2	30	0.24	0.30

α, and β in each experiment is listed in **Table 4**.

Here, we have shown only four results with short parts of the conversation logs in translated English. Underlined utterances indicate DUs under discussion. Note that this analysis phase is the most important for using DUs. The system can be useful only if the subjects conduct this phase properly.

[DU1] "So it remains in the end"

Ta > It took a long time for pizza enough to be delicious.
To > It was delicious but I didn't eat so much.
To > It can make us full easily.
Ta > That's right.
N > So it remains in the end.
To > We must eat it by tonight, or it goes bad.
Ta > Have you already finished eating, N?
N > Not yet. I'll eat later.

This DU was simply about a fact everyone was familiar with. There were few stories in the DU to be reconsidered, so we concluded the DU was not so important. The DU was ignored because it was trivial mentioning a fact merely.

[DU2] "Instead of me, please have Y stay there"

G > We need an application to use fire.
Z > We must escape if any urgency happening without application in advance.
Z > So, let's leave F?
F > Instead of me, please have Y stay there.
Z > Y > "May I remain?"
F > It's so funny.
Z > What is Y doing now?
F > He is crazy about the game.
K > Really? What kind of game?
F > It's just a joke, ha ha.
Z > Do you play games, F?

In the above context, the topic was what to do if an emergency bell rings while using fire in the laboratory without permission. Although Y did not take in the chat, his name was mentioned suddenly. In response, Z pretended Y for a joke a short time later. The topic then shifted to games. "About Y" became the next chat theme by showing the results to the subjects, meaning that this theme was so interesting for the subjects to reconsider.

[DU3] "Christmas song with Te (an electric musical instrument)"

M > Let's talk about the coming year-end party.
H > Shall we put on some short performances?
Ta > Christmas song with Te.
H > Yeah, Ta's performance decided!?
M > Do you have anything to do on Christmas?

In this chat, the topic was what to do for a Christmas party. The DU was a suggestion to use Te (an electric musical instrument) for the party, but nobody paid much attention, and few seems to think this instrument was useful, and some have tried to find good uses for it. This was a concern shared by the subjects. Through the DU, even non-participants can suggest the subjects' concern. DUs can be useful to understand some characteristics of a conversation group.

[DU4] "Shall we use that directivity speaker?"

M > Great, it's Y's stage.
S > Singing along to Te.
M > Let's prepare the stage and mike.
S > Shall we use that directivity speaker?
Y > Y's beam against Prof. K.
M > Amazing!? And Y will become the legend!?

This DU implied to use a directivity speaker. Although the utterance was not ignored in the real context, no one used the terms, "directivity" or "speaker," words understood in this group in the subsequent conversation, meaning that this utterance was extracted as a DU. This was understood as an example of an implied utterance extracted as a DU.

Note that the speaker of this DU has a unique thinking for the directivity speaker, which was found afterward in an interview with the speaker. Although the speaker was trying to discuss uses of the directivity speaker, other subjects paid little attention to the issue, and continued to talk just for joking. So this DU was based on the speaker's unique experience, and the subjects other than the speaker seemed to misunderstand the speaker's intention.

Although we demonstrated only four DUs as examples here, we found three types of DUs – (1) those involved in experiences shared by subjects (DU2 and DU3); (2) those involved in speakers' unique experiences, concerns, or beliefs (DU4), and (3) those not important even if reconsidered (DU1).

Regarding AASI, we described three types of DUs and validated the EV of each type as shown in **Table 5**.

The EV average of type (3) is lower than that of other types. Evidently, t-test proved significant differences between (1) and (3) and between (2) and (3), so we can suggest the evaluation is appropriate for calculating DU importance.

Table 5. EV averages of three DUs.

type	(1)	(2)	(3)	total
number	11	17	14	42
EV average	0.49	0.48	0.16	0.37
correlation coefficient	0.30	0.28	0.15	0.29

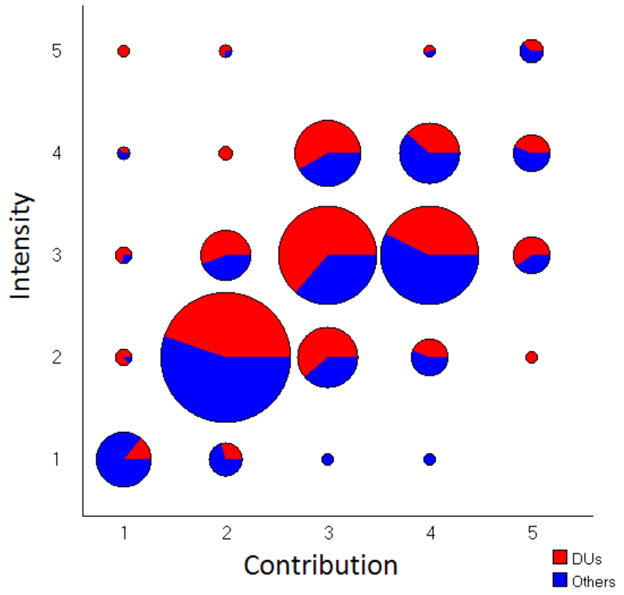

Fig. 3. Results of questionnaires. Size of circles indicates the number of answers. Red and blue areas indicate the ratios of DUs and others.

3.2. Questionnaires

Post-chat questionnaires experiments involved 20 subjects and 253 answers for 88 utterances (DU: 41, others: 47) in total. **Fig. 3** shows the questionnaire results.

Figure 3 shows that "Contribution" and "Intensity" were in the proportion, meaning that conversation in experiments went well and questionnaires were implemented appropriately.

Regarding DU features, red in circles shows the DU ratio. Note that large amount of red parts in the upper part of **Fig. 3** (average, DU: 2.98, others: 2.71), indicating that DU tends to include stronger intensity, t-test showed a significant difference between DU and others on "Intensity" ($P<0.05$). In contrast, red parts at right and left are evenly distributed (average, DU: 2.99, others: 3.01). t-test did not prove a significant difference between DUs and others on "Contribution." We should improve our extraction, but even poor extraction achieved a certain performance. The system could also evaluate DU importance to display by EV. Although we can improve extraction through, e.g., annotation, such methods require too much efforts to apply to real situations. Our method needs only conversation-log to analyse. This is the great advantage for realizability, which is most important in engineering.

4. Conclusions

The system we developed for using DUs involves extraction, evaluation, and its use is simple enough to handle spontaneous conversation. Onlinechat experiments showed extracted DUs tended to include stronger intensities than other utterances, i.e., many DUs should have been better considered essentially. Our system extracted DUs to be reconsidered, helping users discover subjects' unique stories. The stories are useful for hitting on new ideas. In experiments, DUs could trigger reconsidering new events or existing concerns.

Results of post-chat questionnaires pinpointed a problem in our system the lack of validity in extraction. Although we should improve extraction, we do not consider it a fatal problem, however, because the system includes both human and computing processing. Human can deal with these errors appropriately through flexible thinking, and errors may sometimes produce breakthrough, we concluded that extraction performs well enough. We should consider extracted DUs ignored not for the participants but for the computers. In other words, both "what computers can recognize but human cannot" and "what human can recognize but computers cannot," is a viewpoint essential to human-computer interaction.

Chat experiments showed three types of DUs; (1) those involved with experiences shared by the participant, (2) those involved with the speakers' unique experience, concerns, or beliefs, and (3) DUs mentioning a fact merely and which are not important even if reconsidered. While (3) is not important enough to reconsider, (1) and (2) are important for understanding the group and helping subjects' metacognition. Experiments showed that our evaluation was appropriate for calculating DU importance. Evaluation can sort DUs by the importance because evaluation of (3) were significantly lower than those of others.

We now plan to apply our proposal to dynamic situations such as medical situations, development meetings, public talks, etc. The method is simple enough to apply to many situations, even though some differences exist between the analysis of face-to-face conversations and non-face-to-face conversations. People in face-to-face conversations, for example, tend to ignore DUs more than non-face-to-face conversations. While we can easily review a previous conversation in onlinechat, we cannot do so in face-to-face conversations. People read the air and sometimes stop saying what they want to say, so our system works better in face-to-face conversations than in non-face-to-face conversations. We will study the difference between those two types of conversation in projected work.

Our system features; (1) helping participants reconsider their concerns (metacognition), (2) helping nonparticipants infer characteristics of participants, and (3) making communication more efficient and lively. We plan to use extracted DUs as triggers to mine stories of subjects.

Acknowledgements

H. Y. was a Research Assistant of Global COE (Global center of excellence for Mechanical Systems Innovation, GMSI). We would like to thank Professor Shigehiko Kaneko and Kaneko and Yamazaki Laboratory staffs (Department of Mechanical Engineering, School of Engineering, The University of Tokyo) for assisting us in our experiments.

References:

[1] G. Salton and M. J. McGill, "Introduction to Modern Information Retrieval," McGraw-Hill, 1983.

[2] M. Suwa, "A Cognitive Model of Acquiring Embodied Expertise Through Meta-cognitive Verbalization," Information and Medea Technologies, Vol.3, No.2, pp. 399-408, 2008.

[3] G. Salton, A. Wong, and C. S. Yang, "A Vector Space Model for Automatic Indexing," Communication of the ACM, Vol.18, No.11, pp. 613-620, 1975.

[4] R. Barzilay and M. Elhadad, "Using lexical chains for text summarization," Advances in Automatic Text Summarization, pp. 1-12, The MIT Press, London, 1999.

[5] C. Hori and S. Furui, "Summarization: An Approach through Word Extraction and a Method for Evaluation," IEICE Trans. INF. & SYST., Vol.E87-D, No.1, pp. 15-25, 2004.

[6] K. M. Hammouda and M. S. Kamel, "Efficient Phrase-Based Document Indexing for Web Document Clustering," IEEE Trans. on Knowledge and Data Engineering, Vol.16, Issue 10, pp. 1279-1296, 2004.

[7] A. Leuski, "Evaluating Document Clustering for Interactive Information Retrieval," Proc. 2001 ACM Int. Conf. on Information and Knowledge Management, pp. 33-40. Atlanta, Georgia, USA, Nov. 2001.

[8] K. Nishimoto, Y. Sumi, and K. Mase, "Enhancement of Creative Aspects of a Daily Conversation with a Topic Development Agent," Coordination of Technology for Creative Applications, Vol.1364, pp. 63-76, Springer Berlin/Heidelberg, 1998.

[9] N. Matsumura, Y. Ohsawa, and M. Ishizuka, "Automatic Indexing Based on Term Activity," J. of Japanese Society for Artificial Intelligence, Vol.17, No.4 F, pp. 398-406, 2002.

[10] M. Hearst, "TextTilling: Segmenting text into multi-paragraph subtopic passages," Computational Linguistics, Vol.23, No.1, pp. 33-64, 1997.

[11] Y. Ohsawa, "Chance Discoveries for Making Decisions in Complex Real World," New Generation Computing, Vol.20, pp. 143-163, Ohmsha, Ltd. and Springer-Verlag, 2002.

Name:
Hiroki Yamaguchi

Affiliation:
Dept. of Systems Innovation, School of Engineering, The University of Tokyo

Address:
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Brief Biographical History:

2006 Received B.E. from Keio University
2009 Received M.E. from The University of Tokyo
2009- Ph.D. Student in The University of Tokyo

Main Works:

- human-centered system, human factors, medical communication, system for supporting human communication

Membership in Academic Societies:

- The Institute of Electrical and Electronics Engineers (IEEE)
- IEEE SMC
- Japanese Society on Artificial Intelligence (JSAI)

Name:
Yukio Ohsawa

Affiliation:
Dept. of Systems Innovation, School of Engineering, The University of Tokyo

Address:
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Brief Biographical History:

1995 Received Ph.D. from The University of Tokyo
1995-1999 Research Assistant, School of Engineering Science, Osaka University
1999-2005 Assistant Professor, School of Business Sciences, The University of Tsukuba
2005- Professor of Engineering, The University of Tokyo

Main Works:

- "Chance Discovery: Discovery of events significant for decision making"

Membership in Academic Societies:

- The Institute of Electrical and Electronics Engineers (IEEE)
- Japanese Society on Artificial Intelligence (JSAI)
- Japan Association for Social and Economic Systems Studies

Name:
Yoko Nishihara

Affiliation:
Dept. of Systems Innovation, School of Engineering, The University of Tokyo

Address:
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Brief Biographical History:

2003 Received B.S. Eng. degree from Osaka University
2005 Received M.S. Eng. degree from Osaka University
2007 Received Ph.D. Eng. degrees from Osaka University
2007- Lecturer in The University of Tokyo

Main Works:

- communication design focusing on health communication and social communication
- service science

Membership in Academic Societies:

- Japanese Society on Artificial Intelligence (JSAI)