

平成 21 年 6 月 1 日現在

研究種目：基盤研究 (C)

研究期間：2007 年度～2008 年度

課題番号：19500104

研究課題名 (和文) エージェントプレゼンテーション多言語間翻訳技術の研究

研究課題名 (英文)

Translating agent presentations into multiple languages

研究代表者

中野 有紀子 (NAKANO YUKIKO)

成蹊大学・理工学部情報科学科・准教授

研究者番号：40422505

研究成果の概要：本研究では、(1) 日本語とドイツ語の 2 カ国間での比較対象可能な、音声、ビデオ、書き起こし、非言語行動の注釈つき対話データを収集・構築し、(2) 各国の社会・文化的特徴、および非言語行動の特徴をそれぞれパラメータ値として表現し、(3) 社会・文化的特徴から非言語行動を予測するためのモデルをベイジアンネットワークにより構築した。(4) さらに、このモデルをエージェント動作決定機構に組み込むことにより、文化に応じた非言語コミュニケーション行動を生成する多言語対応エージェント生成機構を構築した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
19 年度	1,200,000	360,000	1,560,000
20 年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	2,400,000	720,000	3,120,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：知的エージェント，異文化コミュニケーション，非言語情報

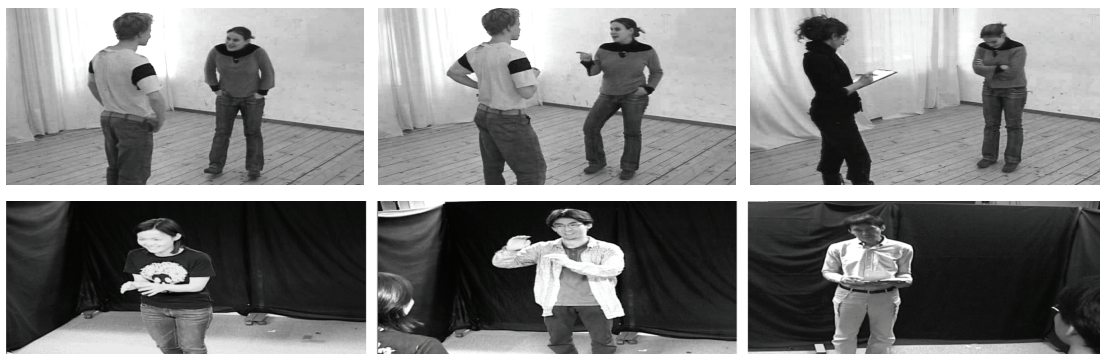
1. 研究開始当初の背景

多様な言語による文書がインターネットを通じて閲覧可能となり、それを支援する技術として、google 翻訳など、ネット上での翻訳サービスが急速に成長している。このようなサービスを利用することにより、母国語で作成した広告を自動的に翻訳し、海外向けに広告を掲示することも可能となった。

一方、初心者や一般ユーザ向けの情報提供には、アニメーションキャラクターを用いた会話エージェントが説明役となる映像コンテンツの方がテキストのみによる提示より

もわかりやすく、親しみやすいといわれている。しかし、テキストが自動的に翻訳できても、異なる言語で適切なジェスチャや表情をまじえながら説明できる会話エージェントは存在しない。これは、ジェスチャや表情が言語や文化により異なる場合があり、このような非言語コミュニケーションの違いを自動的に生成する技術が確立されていないためである。

また、社会・文化的要因を考慮して、エージェントの振る舞いを適切に決定することは、コミュニケーション教育を目的としたシ



(a) First time meeting

(b) Negotiation

(c) Higher Status

図 1: 日独共同での対話データの収集

システム開発において重要であり、数多くの研究がおこなわれている。

イギリス、ドイツ、ポルトガル等が中心となる、eCIRCUS[1]プロジェクトでは、主に学校教育の場面で、情報技術を使って、社会的、情緒的な発達を支援することを目的とし、ヴァーチャルキャラクタとのロールプレイによって、体験的なコミュニケーション教育を行うシステムが研究・開発されている。このプロジェクトで開発された FearNot! ははじめへの対処法を学ばせることを目的としており、異なる文化・社会的背景をもつ子どもたちが集まるヨーロッパの国々で、実践的な研究が進められている。

Khaled[2]らは文化による説得ストラテジの違い分析し、それを反映したゲーム環境を実装している。また、Johnson [3]らは、ジェスチャの文化的な違いを考慮した言語教育システムを提案している。このシステムでは、ユーザはアバタのジェスチャを適切に選択する、あるいはアバタによるジェスチャの解釈をするといった訓練を行う。Miller [4]らは、ユーザとコンピュータとのインタラクションにおける礼儀正しさ(etiquette)について、文化によるあらわれ方の違いを調査し、定量的なモデルを提案している。

一方、社会学や心理学における異文化コミュニケーションの研究では、南欧の人々は、北部に住む人々と比べてジェスチャが多い[5]等、全体的な傾向についての報告がなされているが、コンピュータシステムに実装可能な定量的、決定論的なモデル化は行われていない。従って、会話エージェントの振る舞いを自動的に決定するには、社会・文化的な要因を定量的に表現する手法の開発と、さらにこれらのパラメータ値と実際に観察される非言語行動との相関関係の分析が必要である。

(参考文献)

[1] <http://www.e-circus.org/>

[2] Khaled, R., Barr, P., Fischer, R., Noble,

J., and Biddle, R. Factoring Culture into the Design of a Persuasive Game. In OZCHI'06, pages 213–220, 2006.

[3] Johnson, W. L. et al. Tactical Language Training System: Supporting the Rapid Acquisition of Foreign Language and Cultural Skills. In Proc. of InSTIL/ICALL — NLP and Speech Technologies in Advanced Language Learning Systems, 2004.

[4] Miller, C. A.. Etiquette and Politeness in Human-Human and Human-Machine Interactions: A Summary of Work at SIFT. In Proceedings of the IUI workshop on Enculturating Conversational Interfaces, 2008.

[5] Ting-Toomey, S. Communicating Across Cultures. The Guilford Press, New York, 1999.

2. 研究の目的

本研究では、(1) 複数の言語、特に日本語、とドイツ語において、言語・文化に普遍/固有の非言語コミュニケーションスタイルを各種ジェスチャ、視線、表情等の行動の出現頻度をパラメータ化した定量的なモデルにより定式化し、(2) これをエージェント動作決定機構に組み込むことにより、言語・文化に応じた非言語コミュニケーション行動を生成する多言語対応エージェント生成機構を構築する。(3) また、これらの研究の基盤となる、2カ国間での比較対象可能な、音声、ビデオ、書き起こし、非言語行動の注釈つき対話データを収集・構築する。

3. 研究の方法

(1) 対話コーパスの収集

まず、文化による非言語行動がどのように異なっているかを調べるために、2カ国間での比較対象可能なコーパスを収集するために、ドイツのアウグスブルグ大学と協力し、まっ

たく同じ設定でドイツと日本で会話収集実験を実施した。

(2) 非言語行動の分析

収集したコーパスを両機関で共有し、アウグスブルグ大学では、ジェスチャの形態について、我々日本側では、姿勢について分析を行った。

(3) 社会・文化的要因と非言語行動の関連付け

社会・文化的要因とこれらの非言語行動とを関係づける定量的なモデルを構築するために、Hofstede 理論を採用し、日本とドイツの社会的特徴を Hofstede の 5 つのパラメータの値によって表現した。さらに、これらのパラメータ値と各国での特徴的な非言語行動との関係を分析した。

(4) 非言語行動予測モデルの確立

上記(3)の分析結果に基づき、文化特徴のパラメータ値から特徴的な非言語行動を予測する統計的なモデルを確立した。

(5) 文化に適合したエージェントアニメーション決定機構の構築

最後に、(4) のモデルを実装し、国名を指定すると、それに応じた適切な姿勢アニメーションを生成するシステムを構築した。

4. 研究成果

(1) 対話コーパスの収集

・文化によって非言語行動がどのように異なっているのか、文化を特徴付ける非言語行動は何であるのかを明らかにするために、ドイツと日本で比較対照可能なマルチモーダルコーパスを作成した。状況設定として、社会的関係の異なる 3 種類を用意し、各ペアから、1 対話目：初対面の会話（約 5 分間）、2 対話目：ある目的に対しての議論（約 10 分間）、3 対話目：目上の人との会話（約 5 分間）の 3 対話を収録した。

・収録した対話数は、ドイツ人同士 20 組から計 60 対話、日本人同士 22 組から計 66 対話である。図 1 にデータ収集の様子を示す。

表 1: 話し手の姿勢変化

区間	ドイツ						日本					
	割合	頻度1位	頻度2位	割合	頻度1位	頻度2位	割合	頻度1位	頻度2位	割合	頻度1位	頻度2位
足	F	39%	LRL 25%	WRL 25%	28%	WRL 43%	SLSF 21%					
	M	22%	WRL 50%	LRL 29%	26%	WRL 42%	LRL 23%					
	E	34%	LRL 33%	WRL 28%	31%	WRL 32%	LRL 24%					
頭	F	36%	SHd 37%	THdAP 30%	31%	SHd 43%	THdAP 33%					
	M	29%	SHd 33%	THdAP 30%	23%	SHd 49%	THdAP 20%					
	E	28%	SHd 36%	DsHd 19%	25%	THdAP 49%	SHd 33%					
腕	F	33%	PHIPt 61%	FAs 11%	27%	PHFe 30%	PHB 25%					
	M	30%	PHIPt 52%	FAs 15%	23%	JHs 26%	PHFe 23%					
	E	32%	PHIPt 40%	FAs 19%	32%	JHs 40%	PHW 20%					

(2) 対話コーパスの分析

・収録した対話データの中から、ドイツ人同士 8 組、日本人同士 10 組の対話について、発話の書き起こし、ジェスチャ発生箇所、ジェスチャ表現特徴、姿勢（頭部、腕、上半身）

について、ビデオアノテーションツールを利用して、詳細なコーパスデータを作成した。特に、日本側では、姿勢に着目して分析を進めた。

・特に、腕の姿勢において両国間で大きな違いが見られた。表 1 に示されるように、話し手の姿勢変化における腕の姿勢では、ドイツ人は腕全体を用いる姿勢、例えば腕を組む姿勢 (FAs) や肘をもつ姿勢 (PHEw) を頻繁に行っているのに対し、日本人は、手を顔に付ける姿勢 (PHFe) や両手を組む姿勢 (JHs)、手首を持つ姿勢 (PHWr) など、主に手を使った姿勢が用いられているため、ドイツ人より、動作が小さいという印象を受ける。また、ドイツ人ではポケットに手を入れる姿勢 (PHIPt) が最も多いのに対して、日本人では、この姿勢は全く観察されなかった。同様に聞き手の姿勢変化についても分析したが、同様の結果が得られた。

・さらに、姿勢変化の頻度と継続長について分析した結果、ドイツのデータでは、腕の姿勢変化頻度は 1 会話あたり平均 40.38 回、日本のデータでは 22.8 回であった。また、継続長に関しては、ドイツでは各姿勢の平均継続長は 7.79 秒、日本では、平均 14.08 秒であった。これらの結果から、ドイツのデータでは継続長が短く頻度が多いという特徴があるのに対し、日本は継続長が長く頻度が少ないという特徴があることがわかる。

(3) 社会・文化的特徴と非言語行動との関係のモデル化

・社会・文化的特徴と非言語行動との関係をモデル化し、文化的に適切な会話エージェントの非言語行動、特に姿勢を決定する方式の開発に取り組んだ。まず、社会・文化的特徴を記述する方法として、以下に示す 5 次元の特徴値により各国の社会・文化的特徴を記述する Hofstede 理論を採用した。

1. 権力格差の大小 (power distance) : より権力をもつものによる意思決定の方法、あるいは権力的に下位の者が上位の者に望む意思決定の方法に関する次元。権力格差大の社会では、より強制的な力が用いられ、権力格差小の社会では、階層はよりフラットであり、部下が意思決定に参加することがより好まれる。

2. 個人主義－集団主義 (individualism vs. collectivism) : 個人がどの程度集団帰属的であるかに関する次元。個人主義の強い社会では、個人間のつながりが弱く、自分のことは自分で考えることが期待されている。集団主義の強い社会では、人々はつながりの強い社会の中に取り込まれている。

3. 男性らしさ－女性らしさ (masculinity vs. femininity) : 性別による役割の違いに

関する次元. 女性らしさの高い社会では, 男性らしさの高い社会に比べて, 性別による役割の違いが小さい. 一方, 男性らしさの高い社会では, 競争が受け入れられ, 社会的地位の象徴が重視される.

4. 不確実性の回避 (uncertainty avoidance)

曖昧性に対する耐性に関する次元. 人々が不確実な状況や未知の状況に対してどの程度, 快/不快を感じるかを示す. 不確実な状況とは, 新規で, 不明な点が多い反面, 驚きや通常とは違った点がある.

5. 長期的志向-短期的志向 (long-term orientation vs. short-term orientation) : 長期的志向は, 持続性や忍耐, 節約に関連し, 短期的志向は, 伝統の尊重, 社会的義務の遂行, 面子の維持に関連する.

以上, Hofstede の理論では, 社会・文化的特徴は, これらの次元を合成することにより表現できるため, より小さい集団の特徴を記述することもできる. つまり, 国の違いという地理的な情報に基づき, 社会・文化的特徴を記述的に表すのではなく, 抽象化したレベルを設定することにより, パラメータ値の合成により, ある地域, ある年代など, 様々な下位集団の社会・文化的特徴をプロファイリングすることができる. 本研究では, 日独の違いに着目しているが, このモデル化の方法は, これら2カ国に限定したのではなく, あらゆる粒度の集団に適用できる.

表 2: Hofstede 理論による日独の社会・文化的特徴

	ドイツ	日本
権力格差	35	54
個人主義	67	46
男性らしさ	66	95
不確実性の回避	65	92
長期的志向	31	80

・以上の Hofstede の 5 次元に関する日本とドイツの特徴値を表 2 に示す. ドイツに比べて日本の方が権力格差に対する意識が強く, ドイツは日本より階層に対する意識が薄い. ドイツは日本に比べてより個人主義的であり, 日本はより集団への帰属意識が強い. 日本は男性らしさが非常に高く, Hofstede の調査対象の国では 1 位であった. また, 日本は不確実性の回避傾向が強いのに対し, ドイツは日本に比べて弱い. 日本はより長期志向的であるが, ドイツは比較的短期的志向である.

表 3: 姿勢の特徴

パラメータ名	定義	日本	ドイツ
(a) 総体的活性度	頻度平均	58.6	71.9
(b) 空間的広がり	姿勢分類より	手中心	腕中心
(c) 時間的広がり	継続長平均(sec)	13.75	10.1
(e) 力強さ	7 段階評定	3.86	3.96

・次に, 非言語行動の特徴パラメータとして, 時間的広がり, 空間的広がり, 姿勢の硬さ, 頻度等 (総合的活性度) 5 つのパラメータを定義し, 被験者による評定の収集や, ビデオデータの分析により, ドイツと日本における非言語行動特徴パラメータ値を設定した. 設定したパラメータ値の一覧を表 3 に示す.

(4) 非言語行動予測モデルの確立

・以上の結果を統合し, 社会・文化的特徴と非言語行動の特徴をノードとしたベイジアンネットワークモデルを構築した. 本モデルでは, Hofstede の理論に基づく社会・文化的特徴を中間層とし, 国を特定すると, 非言語行動の表現形態の特徴が予測され, 逆に, 観察された非言語行動の特徴から, 国を予測することが可能になる. 姿勢についてのベイジアンネットワークモデルを図 2 に示す.

ネットワークの構成やパラメータの調整などは今後さらに検討する必要があるが, ドイツと日本の 2 カ国については, ほぼ直観にあった予測結果が得られていることに加え, 他国についても, 従来研究で報告されている特徴について, 適切に予測ができていることを確認した.

(5) 文化に適応したエージェントアニメーション決定機構の構築

(4) で作成したベイジアンネットワークモデルをキャラクターの姿勢アニメーション決定機構に組み込み, 国名を入力すると適切な姿勢候補を出力する機構を実装した. 簡略化したシステム構成図を図 3 に示す.

(6) 国際ワークショップの開催

本プロジェクトとドイツアウグスブルグ大学との共同研究が中心となり, Intelligent User Interface 2008 (IUI2008) にて, Enculturating Interfaces と題したワークショップを開催した.

その中で, ユーザインタフェース, 特に入話エージェントの設計における文化的側面についての研究発表, 及び議論を行った.

また, 南カリフォルニア大学 ICT の David Traum 氏を招待講演者に迎え, エージェントと文化についての最先端の研究についてご

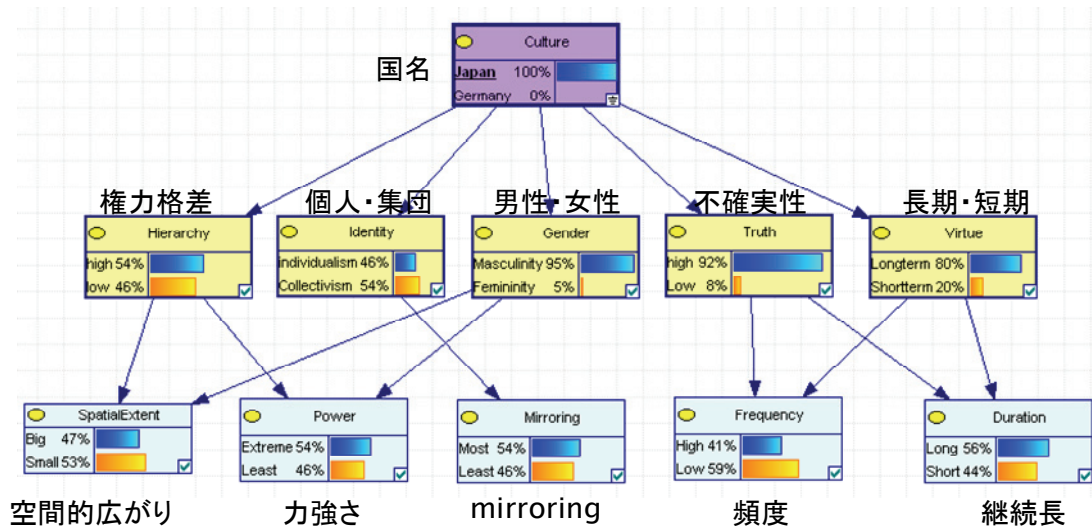


図 2: 姿勢予測のベイジアンネットワーク

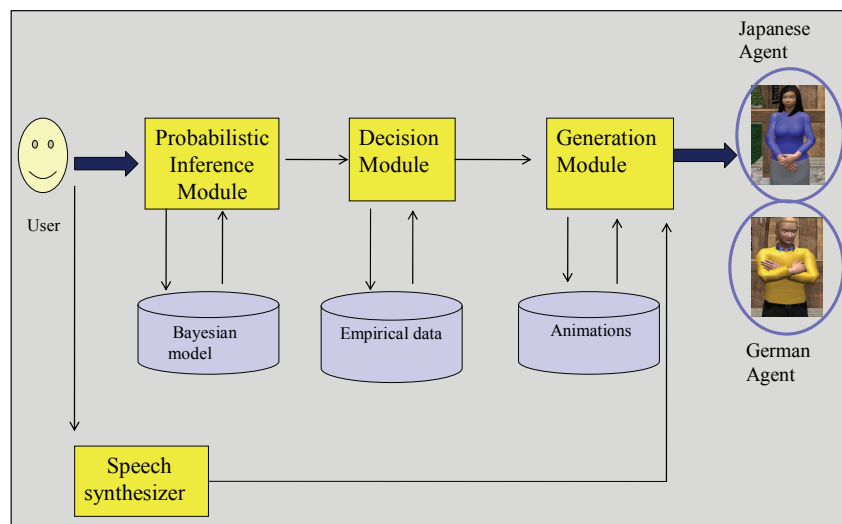


図 3: エージェント姿勢決定機構のアーキテクチャー

講演いただいた。

(7) 海外雑誌での特集号の企画

(6)のワークショップで集まった論文の中から、特に優れたものを選び、さらに新たな投稿を募集し、文化的背景を考慮したユーザインタフェースに関する特集号を国際雑誌 AI&Society にて企画した。現在、ほぼ査読を完了した状況である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

1. Huang, H., Cerekovic, A., Tarasenko, K., Levacic, V., Zoric, G., Pandzic, I., Nakano, Y., and Nishida, T.: An Agent Based

Multicultural Tour Guide System with Nonverbal User Interface, the International Journal on Multimodal Interfaces, Vol. 1 No. 1, pp 41-48, Springer Press, April 2008. 査読有
2. 榎本美香, 中野有紀子.: 人-人, 人-ヒューマンエージェントの社会的インタラクションにおける言語・非言語行為の配置規則, 日本知能情報ファジィ学会誌, Vol. 20, No. 4, pp. 540-556, 2008. 査読有

[学会発表] (計 12 件)

<国際会議>

1. Matthias Rehm, Yukiko Nakano, Elisabeth André, and Toyoaki Nishida. Culture-specific first meeting encounters between virtual agents (Finalist Best Paper Award). In Intelligent Virtual

Agents 2008, pp. 223-236, Springer, Berlin, Heidelberg, 2008.

2. Akhter Lipi Afia, Yuji Yamaoka, Yukiko Nakano, and Matthias Rehm. Enculturating Conversational Agents based on a Comparative Corpus Study. In Intelligent Virtual Agents, pp. 465-467, Springer, Berlin, Heidelberg, 2008.

3. Rehm, M., Gruneberg, F., Nakano, Y., Lipi, A. A., Yamaoka, Y., Huang, H.: Creating a Standardized Corpus of Multimodal Interactions for Enculturating Conversational Interfaces, Workshop on Enculturating Conversational Interfaces by Socio-cultural Aspects of Communication, 2008 International Conference on Intelligent User Interfaces (IUI2008), Canary Islands, Spain, January 2008.

4. Endrass, B., Rehm, M., André, E., and Nakano, Y.: Talk is silver, silence is golden: A cross-cultural study on the use of pauses in speech, Workshop on Enculturating Conversational Interfaces by Socio-cultural Aspects of Communication, 2008 International Conference on Intelligent User Interfaces (IUI2008), Canary Islands, Spain, January 2008.

5. Huang, H., Cerekovic, A., Pandzic, I., Nakano, Y., and Nishida, T.: Toward a Culture Adaptive Conversational Agent with a Modularized Approach, Workshop on Enculturating Conversational Interfaces by Socio-cultural Aspects of Communication, 2008 International Conference on Intelligent User Interfaces (IUI2008), Canary Islands, Spain, January 2008.

6. Rehm, M., Andre, E., Bee, N., Endrass, B., Wissner, M., Nakano, Y., Nishida T., Huang, H.: The CUBE-G approach - Coaching culture-specific nonverbal behavior by virtual agents. Proceedings of the 38th Conference of the International Simulation and Gaming Association (ISAGA), Nijmegen, New Zealand, July 2007.

<国内会議>

7. 中野有紀子, Matthias Rehm, Afia Akhter Lipi. 日独間における非言語行動の比較と特徴記述パラメータの提案, HAI シンポジウム 2008, 1A-4, 2008.

8. 山岡 雄治, Akhter Lipi Afia, Rehm Matthias, 中野有紀子. 会話エージェントにおける非言語行動の異文化適応, 第 22 回人工知能学会全国大会, 1D2-05, 2008.

<学会発表 他 4 件>

[図書] (計 3 件)

1. Nakano, Y., Murata, K., Enomoto, M., Arimoto, Y., Asa, Y., and Sagawa, H.: Modeling Human-Agent Interaction Using Bayesian Network Technique, New Frontiers in Artificial Intelligence: JSAI 2007 Conference and Workshops, Miyazaki, Japan, June 18-22, 2007, Revised Selected Papers, Lecture Notes in Computer Science, pp. 5-12, Springer Berlin / Heidelberg, 2007.

2. 中野有紀子. エージェントによるしぐさと視線のコミュニケーション, 山田誠二(編), 人と共生するエージェント 3.2 節, 東京電機大学出版, 2007.

3. Nakano, I. Y., and Nishida, T.: Attentional Behaviors as Nonverbal Communicative Signals in Situated Interactions with Conversational Agents, In Nishida, T. (Ed.), Engineering Approaches to Conversational Informatics, John Wiley & Sons Inc, 2007.

[その他]

Workshop on Enculturating Conversational Interfaces by Socio-cultural Aspects of Communication, 2008 International Conference on Intelligent User Interfaces (IUI2008)

<http://mm-werkstatt.informatik.uni-augsburg.de/documents/ECI/>

Special Issue Enculturating Human Computer Interaction, Journal of AI & Society

<http://mm-werkstatt.informatik.uni-augsburg.de/documents/EHCI/>

6. 研究組織

(1) 研究代表者

中野 有紀子 (NAKANO YUKIKO)

成蹊大学・理工学部情報科学科・准教授

研究者番号: 40422505