

人工知能が浸透する社会を考える ワークショップ第三弾 報告

はじめに: 2014年度を振り返って

始まりは、2013年度末に世に出た一枚の絵だった。その絵が世間を賑わしたのは1か月にも満たなかつた。しかし、それをきっかけとして、このワークショップや研究グループが形成されたのも事実である。

「人工知能が浸透する社会を考える」というテーマで、人工知能学会9月号に小特集を組むことで、本研究グループのコアメンバーが集まった。これで終わりにせずに、とりあえず問題を理解するためのワークショップを2014年度中に3回開催しようと関係者に声をかけ始めたのは2014年の7月半ばだった。手探りで始めた活動だったが、これは今まさにやらなければならないプロジェクトとの確信は皆に共有されていたように思う。

海外でも、2014年にスタンフォード大学がAIと未来社会に関する研究プロジェクト(AI100)を立ち上げたほか、ボストンを中心に活動をするネットワーク団体 The Future of Lifeも研究助成プログラムを立ち上げた。その意味で2014年度は、人工知能の浸透する社会についての再考が始まった年であったともいえよう。

立ち上げとなった第一弾(9月)、科学技術社会論学会での議論を行った第二弾(11月)に続き、区切りとなる今回のワークショップ第三弾では、本研究グループが来年度以降、どのような問題意識を持って研究会を重ねていくのかの議論が行われた。

第3回 WS の概要

2015年2月24日(火)の13時から21時近くまで、京都大学白眉センター2階にてワークショップを行った。

第一部は、最初の3時間を使って3名の研究者の方に話題提供をいただいた。まず、音声認識が専門の駒谷和範氏より、「最近の音声対話システムの進展と社会との関わり」というタイトルでPepperやSiriなど最近の技術を例に、技術と社会の関係について紹介いただいた。次に、大家慎也氏より、「人工知能をより善く設計し、より善い社会をつくる—媒介理論からのアプローチ」というタイトルで主にVerbeekの議論に依拠しながら、人工知能の設計論についてお話をいただいた。最後に、大谷卓史氏より「人工知能が浸透した社会の情報倫理学」というタイトルで情報倫理研究の最前線について、最近の技術やSF作品を参考しながらお話をいただいた。

コーヒーブレイクをはさみ、第二部では来年度以降の本研究グループ Acceptable Intelligence with Responsibility (AIR)の活動計画についての議論が行われた。

約8時間にわたる議論のうち、本報告では個別の話題提供ではなく、主に質疑応答や議論から浮かび上がってきた議題を簡単に紹介する。

項目

- ◊ 本研究会の目的
- ◊ 人工知能と責任問題
- ◊ 誰のための技術なのか
- ◊ 学会の社会的役割と責任
- ◊ 評価に要する時間間隔の相違
- ◊ これまでとこれから



『トランセンドンス』

(原題 Transcendence.日本での公開は2014年6月)

人工知能と化した科学者の姿を通して、行き過ぎたテクノロジーがもたらす危機を描いた

『her/世界でひとつの彼女』

(原題 Her.日本での公開は2014年6月)

人工知能に恋をする男を描いた恋愛映画

Elon Musk

"We need to be super careful with AI. Potentially more dangerous than nukes" (2014.8.3)

Stephen Hawking

"The development of full artificial intelligence could spell the end of the human race" (2014.12.2)

Bill Gates

"I don't understand why some people are not concerned" (2015.1.27)

アシモフのロボット三原則

SF作家アイザック・アシモフの小説において、ロボットが従うべきとして示された原則

第一条 ロボットは人間に危害を加えてはならない。また、その危険を看過することによって、人間に危害を及ぼしてはならない。

第二条 ロボットは人間にあたえられた命令に服従しなければならない。ただし、あたえられた命令が、第一条に反する場合は、この限りでない。

第三条 ロボットは、前掲第一条および第二条に反するおそれのないかぎり、自己をまもらなければならぬ。

Johnson, Deborah G. and Noorman, Merel (2014) "Recommendations for Future Development of Artificial Agents," IEEE Technology and Society Magazine, Vol.33, No.4(Winter 2014), 22-28.

本研究会の目的

2014年から2015年にかけて、人工知能の社会的影響や危険性を考えさせるような映画が上映されたほか、企業家でありエンジニアでもあるイーロン・マスク、マイクロソフトのビル・ゲイツ、理論物理学のスティーブン・ホーキング博士などの著名人が、人工知能の未来へのビジョンや不安などをtwitterやテレビなどのメディアを介して一般に発信する機会を多く目にした。

またアメリカでは、2014年3月にボストン(MITやハーバード大学)を中心としてボランティアベースの The Future of Life Institute (FLI) が、同年12月にはスタンフォード大学で One Hundred Year Study on Artificial Intelligence (AI100) が設立されるなど、人工知能の発展と人類の未来社会について考えるプロジェクトが推進された。特に英米においては、人工知能が暴走してしまったら人類の未来はどうなるのかというシンギュラリティ的観点からの議論が強いように見受けられる。

専門家ではない立場からすると、そのような発言が飛び交って

いると、「技術的に今とんでもないところに来ているのではないか」と考えてしまうのだが、実際はどうなのだろうかという質問があった。実際は、人工知能が暴走して人間がコントロールできなくなることは、すぐに実現するわけではない。SFなどの影響で、世間の反応が過剰になるほど、技術者自身はそれを冷めた目で見ているのかもしれない。しかし「シンギュラリティ」のようなことは、まだ当分起きなくても、すでに様々な技術の萌芽はある。「技術が未来を変える」という技術先行的・技術決定論的な観点だけではなく、実際今どのような技術ができるつあり、またそれを取り巻く社会・経済・文化的環境との相互作用はどのようなものなのか、足元を見据えた議論をしていくことが重要である。そのため、様々な分野の研究者と協同しながら、未来が「どうなるのか」ではなく、我々自身は未来を「どうしたいのか」を、技術を過小評価することなく考えていくことが大事であり、それを行なっていくための土台を作っていくという本研究会の目的が再確認された。

人工知能と責任問題

人工知能の課題を考える上で問題が起きた時の責任は、誰がどのようにとるべきなのかという問題は大きい。アシモフのロボット三原則のように、人工知能における原則の必要性が言及されているが、ロボットのように「ハード」を持たず物理的に目に見える損害を起こすものではない人工知能の課題をわかりやすく原則化したものは少ない。

たとえば、人工的なエージェントに関しては、応用倫理学者である Johnson & Noorman が次のような原則と勧告を提案している。

原則: 人工的エージェントの行動の責任は人間が負うべき

勧告1: 人工的エージェントは人工物と社会的実践の構築物であって、それらの相互作用を通じて特定の課題を達成すべく編成されていると理解すべき
勧告2: 責任問題は人工的エージェント技術の開発初期段階で扱うべき

勧告3: 人工的エージェントが自律的であるという主張は、法的、明示的に特定されるべき
勧告4: 人工的エージェントの責任問題は責任の「実践」という観点から扱うのが一番良い

誰のための技術なのか

今回のワークショップでは、音声認識という技術開発、テクノロジーの価値や道徳性、そして情報倫理学の現状と課題についての話題提供が行われた。それぞれの話題と質疑応答は興味深いものであったが、そこで議論されている技術とは、誰のための技術なのかという視点が重要であったように思う。たとえば、音声認識などでは、会話の規範(一時にしゃべるのは一人であるべきだ)などをもとにモデル化が行われる。研究者は「だいたい皆こういう規範に従っているから、会話が滑らかに行われる」という観点を抽出しプログラム作成する。そのため、人間を対象とした研究というのは、マジョリティ

の規範やルールを(ある種無自覚に)研究し、それをモデル化しているといえよう。

しかし、それが実際に世の中に普及すると、マイノリティの人たちにとっては想定されなかった問題や、あるいは逆に秩序などが生み出されることもある。

人工知能のように社会に浸透してインフラとして使われる可能性の高い技術は、技術の開発段階からの社会的影響を考える上でも、その技術は誰のための技術であるのかという観点からも、多様なアプローチが必要ではないかとの指摘が行われた。

学会の社会的な役割と責任

学会には学会員が順守すべき倫理綱領や行動指針が定められていることが多い。しかしその指針はどれだけ会員に周知されているだろうか。自分の所属している学会の綱領を聞かれて、答えられるだろうか。たとえばプライバシー的観点から問題があるとされるような技術開発、あるいは業務を発見した時、その技術や業務は学会の倫理綱領に抵触するという名目で問題を指摘できるなど実質的なものとして浸透しているだろうか。また、倫理綱領に抵触した技術や研究者に対して、学会が処分や制裁を下すことはできるのだろうか。もしくは、倫理綱領にさえ従つていれば、学会が研究を擁護してくれるなどのセーフネットとして機能することは可能なのだろうか。少し背景は異なるが、Winny(ウィニー)事件など裁判になることで、技術研究そのものが遅れをとってしまうような事態を避けるために、学会が果たせる役割とはどのようなものだろうか。

学会の運営は研究者のボランティアによるところが大きいため、負担を増やすことは難しく、また特定の研究活動を禁じるよう

実行力を持たせるよりは態度の表明がせいぜいであるとする意見もある。一方で、新技術の経過措置としてもセーフネットを用意することを考えることも可能ではないかとの議論が行われた。

セーフネットの1つの在り方として、査読システムへの倫理規範的観点の導入の可能性について議論が行われた。現在、論文の審査基準に、技術の倫理的な配慮という項目がない場合は、積極的に倫理的な配慮をしようとするモチベーションが働く、現場の研究者の自覚を促せない。しかし、最終的に倫理的に問題のある技術について書かれた論文が査読を通り出版され、万が一社会から問題提起が行われた場合、その査読を通じた学会本体にも責任が問われる可能性がある。表紙問題などはその1つの例であろう。そうならないためにも、トップダウンではあるが、査読の評価軸の一つとして倫理的な配慮について指摘できるような項目を取り入れていくことはできるのではないかとの議論が行われた。

Winny(ウィニー)事件

ファイル共有ソフト

Winnyの製作者である

金子勇氏が2004年5

月に著作権法違反帮助の疑いにより逮捕、起訴された事件。この事件を受けて、情報処理学会と情報ネットワーク法学会は「Winny事件を契機に情報処理技術の発展と社会的利益について考えるワークショップ」を開催した。

学会としてWinny事件に声明を出すよう求める声もあったが「当時の学会内にはさまざまな意見があったうえに、全員が全容をつかんでいるわけではなかった」ために見送ったという。

IT media ニュースより:

<http://www.itmedia.co.jp/news/articles/0406/28/news023.html>



評価に要する時間間隔の相違

Alessandra Maria Sabelli, Takayuki Kanda, Norihiro Hagita (2011), A conversational robot in an elderly care center: an ethnographic study HRI '11: Proceedings of the 6th international conference on Human-robot interaction

2014年の表紙問題のまとめについては、大澤博隆氏の「2015年表紙更新にあたってー前年の「表紙問題」のまとめとこれからー」(人工知能30巻1号, p.2-6, 2015)に詳しい

実際に、技術の社会的影響調査について開発段階からの問題洗いだしを行おうとして、うまくいかない場合、どのような理由があるかの議論も行われた。その中で研究者間での評価に要する時間間隔の違いが指摘された。

ある新技術の影響評価を外部機関に委託するとしよう。その影響評価の審査が1年で終わると言われた場合、あなたはそれを短いと感じるだろうか長いと感じるだろうか。委託された側は、その技術を世に出すときの責任を負うため、倫理的、法的、社会的、経済的などあらゆる観点からの評価をするため1年は妥当であると考えるかもしれない。しかし、開発者側からすると、学会や研究費の報告書など成果発表に出すことを考えると、審査が1年待ちと言われたら、長すぎると言わざるを得ないかもしれません。

共同研究がうまくいった1つの例として、ロボットの社会実験調査を3.5か月行った研究がある。ロボットの影響評価でこれだけ長期的に行われたものは珍しいという。今後人工知能の影響評価を行う時も、何を目的とした影響評価でどのくらいの時間がかかるのかなどに関して、相互の意識のすり合わせが必要となってくるだろう。

これまでとこれから

冒頭の「はじめに」でも述べたように、2014年は人工知能と社会についての話題が噴出した年であった。2014年1月号表紙に関する議論を受け、人工知能学会は3月号で表紙問題に関する小特集を組み、6月には倫理委員会が発足した。

他学会においても、議論が展開された。表象文化論学会においては、7月に「知性、そこは最新のフロンティアー人工知能とジェンダーの表象」というテーマで、11月の研究集会では「人工知能とジェンダー」ワークショップ開催された。また、日本SF大会で行われたニコニコ学会βの特別セッションでは「人工知能キャラクタ

ーとの付き合い方をデザインする～人工知能の作る未来社会」というセッションが組まれた。

我々の2014年度の活動は「はじめに」に記載した通りである。2015年度も更なる議論を継続していく予定であり、4月に韓国ソウルで開催されるCHI2015、東北大大学で開催される応用哲学会、5月に函館で開催される人工知能学会などすでにいくつかの学会での共同発表やワークショップが企画されている。また、国立情報学研究所の研究企画会合公募に採択されたこともあり、今後も積極的に共同研究と情報発信を行っていく予定である。

(文責: 江間)

AIR ワークショップ参加者一覧

秋谷 直矩(エスノメソドロジー, 会話分析, 科学技術コミュニケーション)

市瀬 龍太郎(機械学習, 知識処理)

江間 有沙(科学技術社会論, 監視研究)

大家 慎也(テクノロジーの政治・倫理)

大澤 博隆(HAI, 社会的知能)

大谷 卓史(情報倫理学)

神崎 宣次(倫理学, 環境倫理学)

久木田 水生(哲学・倫理学)

駒谷 和範(音声認識)

西條 玲奈(分析哲学, ロボット倫理, ジェンダー論)

服部 宏充(マルチエージェントシステム・社会シミュレーション)

本田 康二郎(現象学, 科学技術倫理, 技術哲学)

宮野 公樹(学際についての学問論)

八代 嘉美(幹細胞生物学・科学技術社会論, SF)



AIR Web サイト

URL:

<http://web4ais.wpblog.jp/>

AIR NEWS LETTER vol.1 (3)

編集: ARISA EMA