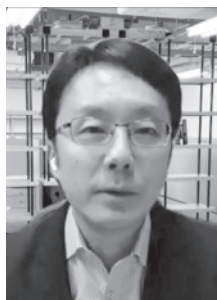




Title	人間知・脳・AI 研究教育センターにおける学際的教育研究の取り組み
Author(s)	田口, 茂
Relation	第6回 人文・社会科学系研究推進フォーラム報告書 Ⅰ 講演の記録 第2部 事例紹介「人社系が参画するさまざまな融合・連携のかたち」
Issue Date	2021-03-29
DOI	https://doi.org/10.14943/JF6.77
Doc URL	https://hdl.handle.net/2115/83473
Type	conference paper
File Information	JF6_hokudai_1-9_Taguchi.pdf





第2部 事例紹介

人事系が参画するさまざまな融合・連携のかたち

人間知・脳・AI 研究教育センターにおける学際的教育研究の取り組み

北海道大学 大学院文学研究院 教授／ 人間知・脳・AI 研究教育センター長 田口 茂

今日は、私がセンター長をやっております人間知・脳・AI 研究教育センターの事例紹介をさせていただきたいと思います。

これまでのお話と大分毛色の変った話になるかと思います。これまでは、あまり哲学といったような、一般に非常に抽象的とみなされるような学問はほとんど話題に上がりませんでしたけれども、われわれのセンターは、哲学のような人文系の学問がむしろ主導的な役割を果たすような事例になっています。そして、同時に学問同士が距離をとりながら緩くつながるというよりは、むしろ完全に融合して一つのプロジェクトを形成するといった事例になっております。その一つの具体的な事例紹介として、お話ししていきたいと思います。

1. 人間知・脳・AI 研究教育センターとは何か

まず、人間知・脳・AI 研究教育センターとは何かということを簡単にご説明します。最初に、センター設立の経緯を簡単にお話します。

このセンターは、2019年の7月1日に学内共同施設として設置されました。そもそもの発端は、2017年の3月に卓越大学院プログラムへの応募を目指して、文学研究科、これは今、文学研究院に改組されているのですが、当時の文学研究科で議論を開始したことです。その中で、脳科学やAI分野の社会実装に向けた文理融合・学際的研究および教育の実施ということで何かできるのではないかと構想が始まりました。

2018年の初めに、卓越大学院プログラムの応募は断念しまして、概算要求事業での教育プログラムの立ち上げという方向に転換し、同時にセンターを設



図1 人間知・脳・AI研究教育センター



図2 3つの分野が交差する地点に新しい「人間知」が成立

立する構想を進めていきました。2019年に、無事、概算要求プロジェクト事業として採択されて、センターを設立することになりました。2020年度からは概算要求組織整備事業にも採択され、その資金も利用して、現在プロジェクトを拡大しているところです。

センターの名称は、人間知・脳・AI研究教育センターという、ちょっとてんこ盛りの名称なのですが、英語名としては Center for Human Nature, Artificial Intelligence, and Neuroscience ということで、日本語とはちょっと語順を変えまして、頭文字を取っていくと CHAIN になるようになっていきます。この CHAIN というのが略称になります。図1はロゴです。

センターの目的は、人文社会科学、神経科学（これは脳科学とも言われます）、それから AI（人工知能）、この三つの分野が交差する地点で、学際的・文理融合的な教育・研究を行っていくことです（図2）。この交差点において新しい人間知というものが成立するだろうと考えています。この人間知とい

うのは何かについて少しお話しします。

われわれとしては、「人間知」とは、人間とは何かという問いをめぐる新しい知というように位置づけています。古来のいわゆる「人文知」にはおさまらない、もっと広いものだと考えています。それは、神経科学や人工知能（AI）の研究が、従来、人文社会科学のテーマであったような領域へと、だんだん踏み込みつつあるという現状を背景としています。ですので、従来の人文知よりももっと広がってくるわけですが、そこでは、その三者、神経科学、人工知能、人文社会科学という、この三者の相互作用による新しい知が成立してくるのではないかと、そしてこれを「新しい人間知」と呼びたいと考えています。

2. CHAIN 設立の背景と研究・教育の方向性

次に、この CHAIN 設立と研究・教育の方向性においてはどのような考え方が背景になっているかをお話ししていきたいと思います。

これはちょっと哲学的な話になってくるのですが、まず、古来人間について考えてきた学問は、哲学や宗教や文学、歴史などの、いわゆる人文学だったと思います。数千年にわたるその伝統を受け継いで、いろいろ学問が分化していった、現在、人文社会科学と呼ばれるような一つの学問群をなしています。しかし近年、人文社会科学は、そういった長い伝統にあぐらをかいているというわけにはいなくなってきたと思います。それは、神経科学や AI の研究がどんどんこの人間の領域に近づいてきたからです。

それは具体的にどういうことでしょうか。まず、神経科学は、人間の知覚や認知を明らかにすると、感情や情動の気づかれにくい実態を解明するといったことを行っています。特に 1980 年代～90 年代ぐらいから、脳イメージング技術、特に fMRI というのが有名ですが、こういったものが発達してきたことによって、生きた脳の活動が外から見えるようになってきました。頭蓋骨をあけて脳を見たり、そこに電極を差したりしなくても（つまり非侵襲的に）、外から生きた脳の活動が画像化できるようになってきたということです（写真 3）。このような手法によって、人間の主観性や内面性にアプローチすることができるようになってきたことが背景にあります。



写真3 fMRIを使った
撮影の様子

こういう新しい技術も背景にしながら、意識とは何か、自己とは何か、社会性の本質とは何なのか、あるいは、何が本当の意味で合理的なのかといった問いを究明しようとする研究が神経科学的に進められています。かつてこれらの問いは、哲学などの人文社会科学が扱ってきたものだと思うのですが、こういう問いに神経科学者も参入しています。哲学者も含めて、神経科学者や認知科学者やいろいろな分野の研究者を巻き込んで、学際的な議論が繰り広げられています。特に、意識という分野で言うと、ASSCとかTSCといった大規模な国際会議が行われています。このように、脳を研究することで、人間らしさへの様々な手がかりが得られているという現状があります。2019年に、われわれのセンターでも国際シンポジウムを開催しましたが、ここでもいろいろな分野の研究者を国内外から招いて学際的な議論を行いました（写真4）。

他方、AI（人工知能）分野について見てみますと、AIに関しては20世紀に何度かAIブームというのが起こっています。大きなものでは2回ぐらいでしょうか。2000年代に入って、これは有名になりましたけれども、ディープラーニングと呼ばれる多層ニューラルネットワークによる機械学習手法などが登場してきたことによって、再びAIが大きな脚光を浴びています。画像認識技術などで大きな成果を上げていますが、最近は文字情報などの記号操作でも成果が上がっています。囲碁の世界で、韓国・中国のトッププロにAlphaGo



写真4 国際シンポジウムの様子

が勝利したというのが、少し前ですが、大きな話題になりました。ごく最近だと GPT-3 という AI が、人間が書いたのとはほとんど見分けがつかないような文章を生成できるという事例が話題になっています。このような AI の急速な進歩を見てみますと、AI というのは人間をもう超えていくのではないか、あるいは、人間らしい能力というのは、逆に何なのだろうか、このような問いが生まれてきて、人間への問いと AI への問いが融合するような、新しい問いが生まれてきていると思います。

次に、AI と脳科学の交差について言うと、新しい AI 技術というのは、ニューラルネットワークと呼ばれるような、人間の脳の構造や機能にヒントを得た数理モデルというのを使っています。そのため、このところブームになってきたような AI 技術というのは、神経科学からのインパクト、脳科学からのインパクトによってできたという面があるのですが、逆に計算論的 (computational) な手法で脳の機能を明らかにする計算論的神経科学などの研究也非常に盛んになっています。そういうわけで、AI と脳科学というのはお互いに深く交差し合って、影響し合って進んでいるというのが現状です。

ここにさらに人文社会系、特に哲学的な考察を結びつけてみたいと思います。AI に内発的動機づけ — intrinsic motivation と言いますが — あるいは好奇心 (curiosity) を実装するのは、今や、AI の性能向上のためにすでに確立された手法の一つになっています。最近では、退屈できるエージェン

トをつくろうといった研究も出てきています。おもしろいことに、この退屈できる AI エージェントというのをつくってみると、全然退屈しないで働き続ける AI エージェントよりも優秀だったりするということが結果として見えてきています (Yu, Chang & Kanai 2019)。

こんなを見ると、非常に面白いと思います。退屈するとか、飽きてしまう、あなたは飽きっぽいですねなんていうのは、人間のネガティブな性質と思われるかもしれませんが、実はここに非常に重要な役割があるのではないかと、ということが AI の研究から逆に見えてきたのではないかと。退屈できるエージェントが優秀だということは、もしかしたら、人間も退屈できるからこそここまで高度な発達、発展、進化を遂げられたのではないかと、そのような推測も成り立ってくるかと思うのです。こんなふうに、AI を通して人間が見えてくると同時に、人間は退屈もするよねと、だったら退屈できる AI をつくってみたらどうだろうというような形で、人間の知を通して AI が進歩するという側面もあるのではないかと。この両方がぐるぐると循環して相互解明が加速していくという面があるのではないかと思います。

AI と人間の相互解明ということについてもうちょっと言っておきますと、AI はどんどん人間らしくなっていますけれども、まだ総合的にみれば、人間の能力にはほど遠いところがあると思います。特に、一つのことだけなら人間よりうまくできるというケースが多く出てきていますが、なかなかあれもこれもできません。これについてもいろいろ取っかかりは見えていますが、まだ普遍的にあれもこれもできる万能の AI は実現されていないという状況なのだろうと思います。ここから、逆に、「なぜ人間はあれもこれもできるのだろうか」という人間らしさへの問いが新たな形をとっていくということがあがあるかと思っています。こんなふうに、AI を鏡として人間の自己理解が進んでいく、という面が一方であります。他方で、また逆に、こうした問いを哲学的、心理学的、脳科学的に考えていくことで、新たな AI が生み出せるかもしれないということが考えられます。このように、人文社会科学、脳科学、AI の間で、それらが一つに交わるような、極めて刺激的な知の領域が生まれているとわれわれは考えています。

これはほんの一例でありまして、個人の心だけではなくて、もっと集団や社会というものを対象としたような研究もわれわれとしては扱っています。このような交差領域で新しい研究・教育を行っていくというのがわれわれのセンターのプロジェクトです。

3. 運営体制

次に、センターの運営体制を紹介しておきますと、現在3名の専任教員が専任教員として本センターに所属しています。これから2名が新しく着任してきます。これらの教員が、センター長、ならびに中心的な役割を果たすコアメンバーと呼ばれる兼務教員とともに、センターの運営に当たっています。中心メンバーだけをとってみても、神経科学、機械学習、哲学、認知心理学、社会心理学、論理学といった多様な専門をもった研究者が集まっています。兼務教員としては、北大の様々な研究院から計25名の教員が参画しています。その専門分野は非常に多彩、多様です。博士研究員も現在2名所属しています。

4. 大学院教育

次に、大学院教育について紹介したいと思います。

人間知・脳・AI 教育プログラムを今年度から開始しています（今年は新型コロナウイルス感染症拡大のせいで少し遅れて、7月からプログラムを開始しています）。北大の全大学院生から1学年20名が参加する特別プログラムになっています。希望者多数の場合は、選抜を行います。

教育プログラムには三つ柱があります（図5）。履修内容でみますと、コースワーク、サマースクール・ウィンタースクール、インターンシップの三つです。コースワークでは、それぞれの学生に合わせてテラーメイドで履修科目を決めています。真ん中の柱のサマースクール・ウィンタースクールは、全員参加の議論を通じてネットワーキングを行い、第一線で活躍する研究者を招聘して、最先端の議論に学生が触れられるようにしていくというものです。学年が上がっていくとインターンシップを行って、学内外の他の研究室に出ていくとか、海外の研究機関に留学するとか、連携企業にインターンシップに出ていく



図5 「人間知・脳・AI 教育プログラムの概要」

といったようなことをやります。修了者には最終的にディプロマが授与されます。

三つの柱のコンセプトを説明します。

第一の柱は「プログラム・ベースド・ラーニング」で、ここではコースワークを通して各分野の基本となる知識やスキルをしっかりと身につけていきます。必修科目と選択科目があります。たとえば人文社会系の学生だったら、ベイジアン・モデリングやディープラーニングを勉強するといったことをやります。あるいは、情報科学系、もともとAIなどを研究している学生さんだったら、脳科学や哲学を勉強していく、あるいは、脳科学をやっている学生さんだったら、AIや哲学に関連するものをとっていく、そのような形でそれぞれに不足する基礎的な知識を身につけていきます。

二つ目の柱、「プラットフォーム・ラーニング」では、サマースクール、ウィンタースクールを毎年開催していきます。第一線の研究者を複数招聘して、学際的で文理融合的なディスカッションを行っていきます。こうやって履修生が国際的な最先端の議論に触れられるような環境をつくっていきます。学生による議論とプレゼンテーションも行います。ここで人脈をつくって、国内外のインターンシップに出ていくということもできます。

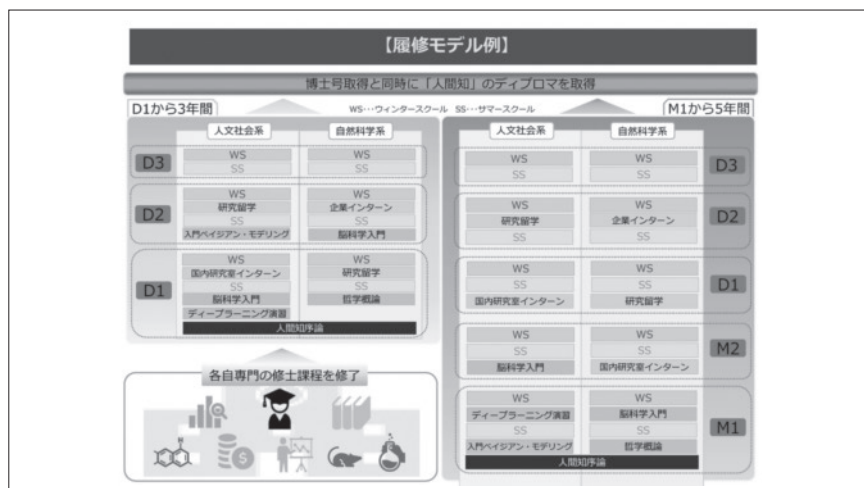


図6 履修モデル例

三つ目の柱が「プラクティカル・ラーニング」で、これは国内の研究機関や海外の研究機関への留学を行ったり、国内外の企業でのインターンシップを行うというものです。これまで学んだ知識やスキルを用いて、現場でその知識やスキルを生かし、体験的に学んでいきます。

5年間で、マスター1年からドクター3年までのプログラムを終えられるようにしてあります。ちょっと忙しいですが、博士の1年から3年間でも終われるような設計にはしてあります（図6）。

5. CHAINの研究

最後に、CHAINの研究についてです。研究上の目的は、北大内の文理融合的研究を集約・加速していくということです。外部機関との連携も積極的に行ないます。センターの専任教員とともに文理融合的研究の国際的な拠点をつくっていくということも目的となっています。主要なテーマとして、意識・自己・社会性・合理性というテーマを設定しています。これは非常に抽象的で普遍性の高いテーマに見えると思うのですが、こういう哲学的とも見えるようなテーマを設定することによって、逆に、非常に緊密な文理融合が実現されてい

るという点が特徴かと思います。また、国内外の多様な研究機関との連携や産学連携も進めています。

CHAINの研究環境については、人文社会科学の非常に豊かなリソースがあるということ、そして、北大内の連携機関である社会科学実験研究センターが実験専用のMRIを持っているということも恵まれた環境と言えるかと思います。それから、これは世界唯一の施設らしいのですが、Dual MEGという装置があったりします。それから、動物実験施設もあり、情報科学研究院・工学研究院ではロボット・AI研究も行われています。そして、これらすべてが徒歩圏内にあるということが特徴です。こういう点で、非常に恵まれた研究環境が実現できているのかなと思います。

6. まとめ

最後にまとめです。本センターは、人文社会科学、脳科学、AIのハイブリッドな研究・教育に特化した——これに特化しているという点では、おそらくは国内唯一かと思うのですが——そういう研究教育機関になっています。いろいろな連携を通して、このような融合研究を推進していくハブとなることを目指しています。そして、教育面では文理融合、双方の知識やスキルを持った先端的人材の育成、研究面では、理論的なブレイクスルーの探究を行っています。今後、規模・人員・教育プログラム・研究体制の拡充を計画しています。

最後に強調したい点をいくつか挙げます。本センターは、そもそも人社系主導のプロジェクトです。文学研究科（当時）の教員の発案で始まったプロジェクトです。文学研究院の山本文彦前研究院長が主導して、文学研究院の教員が中心となってプロジェクトを推進してきました。社会全体を見渡して、時代の新たな方向性を見据えて、文理をまたぐ学際的プロジェクトを動かしていくには、人文社会系の理念や知見、それから構想力が不可欠になってくるのではないかと私としては考えています。無理に文理を融合させようとするのではなくて、同じ目標、たとえば意識や自己、社会性、合理性とは何かといった同じ研究目標を実現するために諸学問が集うなら、そこには必然的に融合が生まれてくるのではないのでしょうか。本センターの活動によって、その生きた実例をつくることができればと考えています。