

## 医療用人工知能を知る

[医療用人工知能とは \(研究者向け\)](#)

[医療用人工知能とは \(医師向け\)](#)

[医療用人工知能の応用事例](#)

医学生と民医連の情報誌「Medi-Wing」72・73号 特集

# 「AIは医療をどう変えるか ～AI研究者と語る医療の未来～」前・後編

医師向け

この記事は20分で読めます

### 講師（敬称略）

- ・奥村 貴史（おくむら たかし）  
北見工業大学 工学部・工学研究科 教授 保健管理センター長  
ピッツバーグ大学 大学院 計算機科学科  
旭川医科大学 医学部 医学科 卒

### インタビュアー（敬称略）

- ・井村 春樹（いむら はるき）  
尼崎医療生協病院 内科/感染症内科 旭川医科大学 医学部 医学科 卒  
京都大学大学院医学研究科  
社会健康医学系専攻 健康情報学分野 専門職学位課程修了  
現在、博士後期課程在籍中
- ・山田 康平（やまだ こうへい）  
横浜市立医科大学 医学部 医学科2年

[医学生と民医連の情報誌「Medi-Wing」](#)

近年、AI（人工知能）の技術を利用した製品やサービスが身近なものとなり、私たちの生活に様々な変化をもたらしています。医療の分野においてもAIの導入は始まっており、診断の精度やスピードの向上などに期待が寄せられています。AIが医療の進歩に貢献する一方で、医療のあり方にはどのような変化が起きるのでしょうか。

AIを利用した診断支援研究の分野で活躍されている奥村貴史先生を招いての対談の様子を前後編でお届けします。



(mediwingより)

## 臨床研修を通じて感じた疑問

**井村：** 尼崎医療生協病院で内科をしています、井村と申します。診療支援や意思決定支援に興味がありまして、今回参加させていただきました。よろしくお願ひいたします。

**山田：** 横浜市立大学医学部医学科2年の山田康平と申します。知識的に至らない点もあるかと思いますが、ワクワクした気持ちで参加しました。よろしくお願ひいたします。

**奥村：** 北見工業大学の奥村です。公衆衛生系の研究所にいて、公衆衛生や保健医療行政の情報化に携わっていました。そこで医療用のAI研究に出会い、10年ほど関わっています。

**井村：** 先生が、医療用AI研究の方針を決めるきっかけはあったのですか。

**奥村：** それは明確で、私が北海道の富良野で臨床研修をしていた時のことです。場所は北海道の富良野で、そこはいわゆる医療崩壊地域だったのですが、ちょうど私が研修をしていたときに、内科医が院長だけになってしまったんです。

**井村：** それは非常に厳しい状況ですね。

**奥村：** はい。見て見ぬふりをするわけにもいかず、院長と2人で内科を守りましょうということになりました。ですが、状況があまりに酷かったのです。電子カルテなど、本来仕事をサポートしてくれるはずの機械が、逆に仕事の効率を落としたり、待ち時間を増やしたりということがあまりに多く、「患者が死ぬか医師が死ぬか」というぐらいの苦境に追い込まれていました。

当時、コンピューターの基礎研究の学位を持った医師として、それはあまりにおかしいと考え、臨床研修2年が終わった後に厚労省の研究所へ行きました。こうした経験から「本来医師や看護師、パラメディカルの仕事を支えるべきコンピューターが、何故ここまで非効率なのか。それをどうにかしたい」という思いが当初からのモチベーションの1つでしたので、それに類似する依頼が来たときに「これはチャンスだ」と思い、自分のストーリーに近いところを見つけて取り組んだという経緯になります。

## 医療用AIの抱える課題

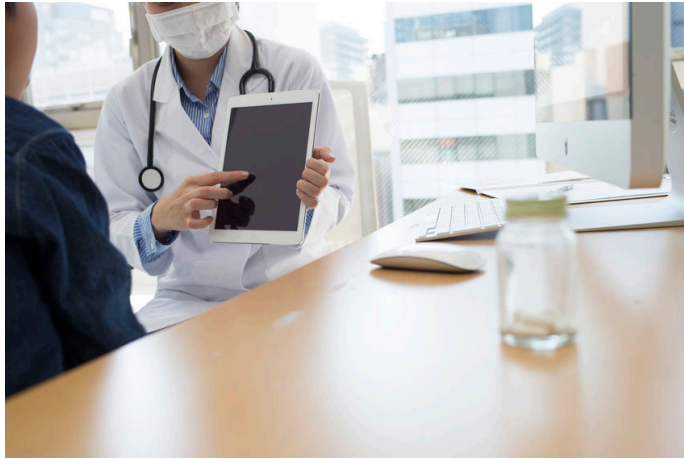
**奥村：** 医学生の山田君は、医療用AIについてどういうイメージをお持ちですか。

**山田：** 医学生としては、医療用AIがいま医療従事者がしている仕事のうち何を助けていく方針になっているのか、気になっているところです。

**奥村：** ありがとうございます。井村先生はいかがでしょう。

**井村：** 医療用AIについて初めて聞くような方、特に学生さんと一般の医療従事者だと、AIって何かすごい魔法の道具で、場合によっては「人類を脅かしかねないのではないか」「仕事を奪われるのではないか」という懸念を抱いている方は多いと思います。僕はそうじゃない、むしろ助けてくれる技術だと思っています。基本的には補助のツールで、分かりやすい例で言うと電子カルテが挙げられます。また、奥村先生が開発を担当されている診療支援システムも医療用AIに含まれますね。

**奥村：** そもそも医療に限らず、社会生活一般でコンピューターと関わりがないものは、なかなか無いですよ。医療研究を助けるAIは当然のようにあり、僕らが患者さんと向き合ったときに、診断を助けるAIもあります。治療法の選択や予後の予測、次の外来あるいは何年後かに振り返る際にも関わります。言ってみれば、医療の入口から出口まで、それぞれの過程に何らかの形でAI的な技術というのが関わり得るのです。ちょっぴりイメージが湧きますか？



**山田：** はい。今までは、医師が診断する道筋を示してくれたり、在宅の患者さんから情報を引き出して病院につなげていく、臨床の場でのAIを強くイメージしていましたが、研究利用であったり、他の分野にも広く散らばっている印象を受けました。

**奥村：** そうですね。これは医師側から見た説明ですが、患者さん側にも直接関わる技術ではありませんね。

**井村：** 患者さん自身が困っていることを自分でうまく表現できないとき、AIがどのように情報を引き出せるのか、というところは大きな課題だと思います。

例えば、患者さんが「頭が痛い」と言っているとき、本当に頭が痛いのか、痛みのステージはどうか、そういう情報をうまく引き出せるのかは、AIではなかなか代替、補助ができない部分ではないかと思っています。こういった部分は人でもバイアスが掛かるので、それを補助するものとしてAIを使った場合に、患者さん自身の誤解がシステム自体に影響を与える可能性も課題ではないかと思っています。

加えて、そもそもの課題としてAIが診断支援をするときに誰がその診断に責任を持つのか。確率的にまれな事象が起こった場合に、どう対応するのか、今後は問われてくると思っています。

**奥村：** ありがとうございます。大きく2つの課題を挙げていただきました、1つ目はAIと人間との接点、ユーザーインターフェースの問題です。相互に情報をやりとりしていく対話型のAIでないと、医師のもつバイアスに引き付けられて誤った解釈する可能性がある。それを排除するためには、人間の認知のバイアスなども含めたユーザーインターフェース研究を進める必要がありますが、なかなかそういう研究の蓄積がありません。私の行っている研究の1つがまさに、医療用AIのユーザーインターフェース研究になります。

2つ目は責任問題です。これは薬機法の問題になります。そのルール上、薬や医療機器は国が定めた品質管理基準を満たす必要があります。例えば「埋め込み型のペースメーカーで変な製品を体に入れてしまったらどうするのか」という話です。

そのため、リスクの大きなものはがちに縛る必要があります。しかし、実は今問題になっているのはソフトウェアの方です。医療用ソフトウェア自体は長らく薬事的な枠組みの外にありましたが、最近になって医療用ソフトウェアも薬事の網に掛けるという動きが出てきました。しかし、技術革新と品質管理を両立させていくための方法論が確立しているとは言い難い状況にあります。

**井村：** そうですね。

**奥村：** ただ、シンプルに考えると、人の体に入れるような命に直結するものは、安かろう悪かろうでは困りますよね。けれど一方で、診断支援において「95%はこれ、5%はこれ、0.5%でこの可能性もある」と可能性を提示して、医師に判断を委ねるシステムが直ちに患者の命を奪うでしょうか。

極論ですが、診断に困った医師を支援するシステムの出力はランダムでもいい。なぜかという、診断困難時の原因の多くは、コモンな疾患における稀な病態(atypical presentation)だからです。後から冷静になって振り返る、あるいは別の医師が関わると簡単に分かる。けれども、患者がその病院に来たときの経緯や、医師に伝えるストーリー、それまでの診療経験

など、医師は色々なバイアスに囚われがちです。その結果、本来であれば辿り着ける診断に辿り着けないというケースは、おそらく多い。

そうすると、診断支援システムも「答えはコレです」というより、「この可能性を見落としていませんか」と教えてくれる方が、実は誤診などの改善にもつながる。僕らはこのことを **debiasing** (デバイアシング) と呼び、医療用AIに求められる大きな機能の1つだと考えています。それは気付きを与えるものであれば何でも良くて、例えばチェックリストなど「この可能性を忘れていた」と気付かせてくれるものであれば、臨床的に有益です。

しかし、「AIの出力はこの品質を満たさなければならない」という品質管理基準を設けると、確率的に高い選択肢が上位に来ないとおかしくなる。それでは人間が持つバイアスを補強する結果になりかねない。逆に、ランダム出力を見せたほうが **debiasing** に資するとしたら、そもそも品質管理基準とは何なのかということになります。

**井村：** なるほど。

**奥村：** ですから、医療用のソフトウェアの品質は、生命予後に直結するような病態の判断に関わるソフトウェアか、そうでないものかで考え方を分けるべきです。前者については、僕は法的規制派です。そうでないと技術の信用を落としてしまいます。

しかし、排除できないリスクが残るものについては、品質管理のためにレギュレーションを強めるのではなく、(自動車保険のように)何か起きた時の補償で守るべきです。そうすることで、研究者も医師も技術に対するハードルが下がり、いろいろなトライアルが可能になります。

なので、僕らはいっそのこと、間に医師が介在する相対的にリスクが少ないものは、レギュレーションの枠から外すべきだと主張しています。それが井村先生が問題提起された医療用AIの研究における責任問題に対する見解となります。



## 社会的な理解はこれから

**井村：** 先生は、他にどんな課題があるとお考えでしょうか。

**奥村：** 現状の課題は、大きく分けて3つあります。1つ目は先ほど話題になった、法的なレギュレーション。2つ目はデータの問題です。日本では、研究利用できるデータが諸外国と比べて明らかに不利な状況にあります。

**井村：** それは僕も日々実感しています。元々利用できるデータが非常に少なく、守秘義務や個人情報保護の問題もあり、オープンにされていないナショナルデータベースもかなりあります。

**奥村：** 3つ目は理解の問題です。それは医療従事者や医療機関トップ、あるいは各医学に関わる研究団体の執行部の理解、究極的には社会の理解の問題でもあります。AIはうまく使うべき技術ではありますが、社会的あるいは医療従事者側の理解はまだ相互不理解というか、意思疎通ができていないところがあります。

同じように、これから社会的な理解が進んでいく分野の例では、自動運転技術が挙げられま

す。AIには人間の持つバイアスがないので、おそらく死亡事故も減る。将来、自動運転がそのレベルになったときに、初めて社会的な合意が進んでいく、そういう時代が来るはずで。しかし、そこに至るまでには犠牲も出て、それに対する社会的な合意を形成しながら、少しずつ進んでいくと思います。

そこで振り返ってみて、医療用AIはどうか。僕らが現役の間はともかく、100年後、200年後と考えたときに、場合によっては医師よりもAIの方がミスが少ない分野がいくつも出てくるでしょう。人間を補う形で、技術の支援を得ようという時代は来る。そこに向けて、今の時代の研究者としてどう社会と向き合っていくのかは、考えるべき課題です。実際に僕らのところにもそういう疑問がよく寄せられています。先の話でも「医師が要らなくなるのではないか」などが挙げられましたね。

**井村：** そうです。「どの科がいらなくなる」あるいは「新しい診療科の形ができるのでは」という話もあります。私たちの予想では、内科系や精神科系の負担は増えるけれど、外科系や画像診断、病理などの科は、かなり自動化が進んでいくのではないかと考えています。

**山田：** 学部の中では、精神科に関しては「人が診ないと精神は分からない」という意見が大多数で、「AI導入後も今までと変わらない」という見方の人が多いです。一方でそれ以外の科は「自動化して負担は減っていく」と考えている人が多いと思います。

**奥村：** 技術進歩によって、必要なくなる治療や手技、検査はあってしかるべきで、僕らはちょうど歴史の転換期に立ち会っていると捉えた方が自然でしょう。100年後や200年後に「診断の入口自体がデジタル情報で、それに対しデジタルな反応をするタイプの診療行為は、機械の方が適している」という時代が来ても驚きはしません。ただ、恐らく僕らが現役のうちはそのまでは到達しないと思います。

**井村：** そうですね。いずれは洗練された外科医の技術が、オートメーション化されていく時代も来るでしょう。例えば、CTなどの画像検査においても、自分たちが見つけられていないものを教えてくれる技術は、どんどん進んでいくと思います。

それとは逆に、内科系や精神科系など、患者さんから情報を引き出す部分に関しては、もう少し時間がかかる気もします。ただ、精神科領域の診断に関しては、「対人間だとバイアスが入ってしまうが、AIだと冷静に分析できるのでは」という印象もあります。この辺りは技術の進歩とともに、自分たちが想像していないような形で変貌していくのでしょうかね。

**奥村：** そういう意味では、懸念を持っておられる方々と、実際に研究であるとか、今の日本の医療が抱えている問題や技術が解決できるものをうまくすり合わせないと、本来は技術で突破すべきところにブレーキがかかってしまいますね。

**井村：** 「医療従事者の仕事が奪われる」という方向に向くと、そうなってしまいますね。本当はAIで業務がもっと楽になり、労働時間を短くするなどの目的が伝わればと思います。

**奥村：** 僕が言いたいのは、「医師の睡眠時間を長くする」、まさにこれです。それが意味ゴールの1つで、このような言い方をした方が社会的な理解が進むと思います。

僕らも医師という仕事に進んだ以上、自分の人生を捧げているわけですから、夜中に起こされることそのものに不満はない。けれども、コンピューターの待ち時間や、「1ヶ月前の心電図と最新の心電図とを比較したいが、古い方は隣の病院にあり電話して取り寄せる必要があるがなかなか手元に来ない」という事態には不満が出てきます。

**井村：** 画像診断でCDが送られてくるけれど読み込みに時間がかかることや、電子カルテを開く時間が長いなどもありますね。

**奥村：** そういった点は、AIを活用し業務改善につながればと考えています。

～後編～

## 社会との認識のギャップを埋める

**奥村：** AI技術はこれから発展が見込まれる、間違いなく人類に貢献する技術です。しかし、その研



導されました。まず自分の目で診断した上で、その答え合わせのつもりで見なさいと。そうでないと、機械の診断結果というバイアスに囚われるからです。しかし、そういう習慣をつけた上で読んでみると、それなりにいいことが書かれています。

**井村：**「そういえば、この可能性を見落としていた」というときがあります。

**奥村：**はい。それが医療の質を上げているか下げているかという、確実に上げていると思います。

**山田：**今の話で、医療用AIが誤診する・しないのではなく、人のミス減らすための1つの指針というか、「これも参考にしたほうがいい」という情報を提示してくれるものという印象が強まりました。

**奥村：**そうですね。そもそも医療現場では、重大な決断を単一の情報源だけでは決めません。仮に血液検査で1つの項目が異常値だったとしても、その他の項目や全身状態、これまでの経過などを総合して判断します。  
心電図の例でいうと、病棟では、各部屋、各ベッドの患者さんの心電図を遠隔でモニターへ飛ばして、様子がある程度見えるようにしています。その時、アラートに誤診があるとすると、どう思われますか。

**山田：**それを監督するのはあくまで医師なので、それを使っている医師の責任がかなり大きいのかなと思っています。

**奥村：**法的な責任の所在はおっしゃる通りです。極端な話、僕らは、僕らの責任で相当なことをするわけですね。けれども、その機械が誤っていた責任についてはどう思いますか。

**井村：**モニターが間違っていたときですね。例えば波形を見て判断して、怪しかったらもう一回心電図を取ったりするので、機械の診断のみで何かをするというのは極めて少ない。

**奥村：**そうですね。詰所から見えない病棟のベッドで寝ている患者さんの状況を、僕らは24時間見守れません。100パーセントでなくても、99.99パーセント信頼できる機器を使うことで、医師や看護師の手と目だけでやるよりも、チームや病棟、全体での医療安全的なパフォーマンスは明らかに上がります。

**井村：**夜中に伺ったら冷たくなっていた、ということは格段に減ると思います。

**奥村：**そうですね。技術の本質、あるいは今実用性があるものがどういうもので、どういうことに使うべきか、使うべきでないのかという議論こそが求められています。具体的な議論の一步前で、医療用AIが危険かどうかという話が出るとすると、研究者あるいは医療制度を持っている側の広報が、まだまだ足りないのだと思います。

**井村：**実は臨床現場で浸透しているはずなのに、なぜかディープラーニングだとか世間のAI理論とごちゃ混ぜになっているんですね。僕自身も、心電図がAIだと結び付いていませんでしたが、実際に自動判定に助けられたことは何回かあります。そういう意味で、AIが自分の診療にも深く染み込んでいるなど、今改めて認識しました。  
お話を伺って、医療人が「これが医療用AIだ」とまず認識するところから、理解が始まるような気がします。

**奥村：**認識のギャップをどう埋めるかというのも、僕ら研究者サイドの悩みではありますね。

**井村：**そうですね。先生が挙げられたポイントの中で「社会の理解」、すなわち医療用AIに対して「何か怖いもの」みたいな誤解が、この技術革新を妨げている一番大きい要因のような気がします。「実はそうじゃないよ」というところをうまく広報していく必要があると思います。医療用AIの発展を続ける中で、どうすれば正しい情報をキャッチできるのか、何か具体的アドバイスはありますか。

**奥村：**そういう意味では、僕ら研究サイドも医療従事者、とりわけ医師を対象に広報の努力を始めたところなんです。どの医師でも使える診断支援のツールをネットに設置し、体感してもらって、医療用AIがどういうものかを分かっていたらこうとしています。

また、発展途上の技術なので、ぜひいろんな方に研究に参加してほしい。医学部を目指している高校生や、これから自分の専門を決めていく医学生、あるいは大学に戻って研究すると考えておられる医師など、さまざまな立場での参加があり得えます。ひょっとすると、この技術の研究に参加してもらうことが、一番の広報なのかなという気もしています。

**井村：**「医療従事者を含む多くの方が、医療用AIの研究に携わる機会を作っていく」ということでしょうか。

**奥村：**はい。実際に今そのための準備をしているところです。

## AIで変わる医療と、変わらないもの

**山田：**どの人にも、AIが導入された医療の形をイメージしきれない部分があると思います。「アルファ碁」のように自分で考え、情報を引き出してどんどん診断を進めていく医療用AI。それが誤診をしたときには、自動運転車の死亡事故のようにセンセーショナルな記事になる。そのようなイメージを持っている人も多いと思います。AIが導入された診断や、医療の形を視覚的にイメージできるようなものがありましたら、教えていただければと思います。

**奥村：**技術予測はなかなか難しく、ある技術革新が「来ない」と言って来ることもあれば、その逆もあり得ます。ただ、少なくとも僕や山田さんが現役の時代に、AIが医師を代替するような事態にはならない。それは僕ら研究者が求めているゴールではなく、今はそこまで先の話をしていないんです。

医師が健康と家族を犠牲にして、どうにか成り立っている今の医療水準を、どう現実的なコストで持続可能にするのか。それこそが、まず僕らが取り組むべき課題です。それが達成できて初めて、「診療行為のうち、この辺は機械が置き換えていく」という話が現実味を持てきます。

ただ世界では日本より医療水準や医師数が厳しい国があって、「AIでいいから診療所を支えてほしい」という状況もあります。そういった国では導入の議論も異なるでしょうが、日本では先ほど話した通り、医師の代替ではなく支援が喫緊の課題です。

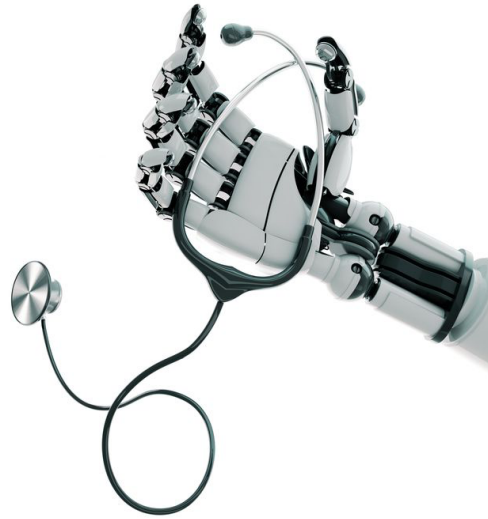
**井村：**「医師の業務負担をできるだけ軽減する方向で、医師自身が仕事を長く続けられるように支援していきたい」ということですか。

**奥村：**それが結局のところ、日本の患者の利益になると考えています。

とりわけ地方だと、そこに医師がいないと、子どもが産めません。子どもが産めないと、人口を維持できません。それをどう解決するのが、僕らの社会が抱えている大きな問題であって、医師を代替するかどうかというのは当面問題にならないし、なるべきではない。

例えば映画で医療用ロボットが手術をしたり、出産を助けるシーンがあります。ゆくゆくはそういう時代が来るのかもしれませんが、突き詰めると医療と同胞意識って切り離せないんですよ。結局「人間がどう生きてどう死ぬか」という話だからです。機械に「あなたの寿命は78歳です。しかし今の習慣をやめれば、〇パーセントの確率で80歳まで寿命が延びます」と言われて、それが三度の飯より好きなものの場合、やめられますか。





**山田：** 僕なら、それで「読書やめろ」と言われたらやめないとします。

**奥村：** そうです。結局価値観の問題で、自分がどう生きてどう死ぬのかという判断に対して、客観が入り込む余地はあまりないのです。なので、医師がなくなるかと言われれば、なくならない。じゃあ機械に負けるのかというと、負けたくはないですね。

僕は機械に負けないように学ばなければいけない。ただ、人間にもバイアスという大きな落とし穴があるので、機械と共に働くことで、どう自分たちのパフォーマンスを上げるかというのが、僕らが考えて取り組むべき課題です。

**井村：** そうですね。AIは医師の代わりにはならないだろうし、なるべきでもない。対立すべき存在ではなく、医師と医療用AIが共働して、人の支援に当たるといふか。

**奥村：** より良い医療を目指すということかと思います。

**井村：** そうですね。医師の側だけでなく患者の側も含めて、そういう未来が訪れるのを信じて、誤解が解けていけばいいな、と感じました。僕自身、目からたくさん鱗が落ちました。非常に良い機会をいただきました。どうもありがとうございました。

**奥村・山田：** ありがとうございました。



## 医療用人工知能を知る

[医療用人工知能とは \(研究者向け\)](#)

[医療用人工知能とは \(医師向け\)](#)

[医療用人工知能の応用事例](#)

# プライマリ・ケアと人工知能

医師向け

この記事は11分で読めます

奥村 貴史(北見工業大学)

## 日常となった画像認識技術

先日、大学同期による恒例の忘年会にて、集合写真を撮りました。そして、出席できなかった同期のためにと、撮影者がFacebookにアップロード。これで良いか？と確認のために見せられた画面では、各顔写真にメンバーの名前が入っていました。Facebookでは、もうお馴染みの機能です(図1)。



図1. 顔認識技術の様子 (著作権フリー素材より作成)

このような人物画像を対象とした顔認識(Facial recognition)技術は、2000年代にはまだ珍しく、防犯等限られた分野でのみ利用されていました。しかし、認識精度の向上やインターネットの普及に伴い、とても身近な技術となりました。Facebookで友人の顔写真を自動的に認識するのはその一例です。デジタルカメラを購入すると、自動的に被写体の顔を認識し、フォーカスを合わせる機能が付いてきます。近頃では、病院の受付に顔認識を利用するケースもあるようです。

コンピュータは、従来、文章や画像の作成や表示が得意とされ、1990年代以降のインターネットの普及も静的なウェブページにより支えられてきました。2000年代に入り、コンピュータの性能が向上したことで、ネットにおける音楽や動画の利用も拡大して行きました。それでも、コンピュータの役割は情報の保存や閲覧に留まっていた。こうした状況が大きく変化した背景のひとつが、画像認識技術の発展です。コンピュータが静止画や動画を認識する技術が向上したことで、情報技術の活用範囲が一気に広まりました。これを、技術革新によってコンピュータが「眼」を持つに至ったと、進化になぞらえて説明する論者もいます1)。

「眼」を持ったコンピュータは、今まで実現しえなかった様々な技術を現実にしてきました。自動運転車は、コンピュータが人間の代わりに自動車を運転する技術です。公道での無人運転はまだ許されていませんが、特殊な用途ですすでに実用水準に入っており、輸送コストの削減や渋滞緩和等、文明社会の在り方そのものを革新する可能性を秘めています。農業においては、従来自動化が困難であった収穫業務へのロボットの活用が試みられており2)、警備業の各社は、既に巡回警備用のロボットを実用化しています3) 4)。こうした技術は介護分野にも応用されており、夜間の見回りにロボットのPepperを活用するような試みも進められているようです5)。

## 画像認識技術の医療応用と人工知能

画像認識技術がこれだけ発達すると、医療へと応用したいと考えることも自然な成り行きです。実際、画像認識技術を胸部単純X線写真やマンモグラフィーに対して適用し、読影を支援する研究が長年試みられてきました。これは、検診によって発生する多量の画像を限られた医師で効率的に診断していくうえで、有意義な試みと考えられます。同様に、病院において撮影されるCTやMRI等の膨大な医療画像を対象として、肺癌や脳動脈瘤を指摘するような試みもなされてきました。このように、医療における人工知能というと、まずは画像診断が思い浮かぶかも知れません。実際、現在の人工知能技術の飛躍的な発展は、画像認識技術により牽引されてきました。そうした人工知能の応用先として、画像診断への期待は高いといえます。

実は、人工知能技術の診断への応用は、これら画像診断が一般化する遥か以前から行われてきました。Logoscope<sup>6)</sup>は、1950年代にイギリスのFirmin Nashが開発した診断支援ツールです(図1)。このツールには、主要な症状に対応した細い短冊が収められています。各短冊には、その症状を呈しうる疾患の位置に黒線が刻まれています。診断に際しては、患者の呈する症状に対応する短冊を選びだし、土台に並べます。すると、ところどころに黒線が並ぶので、土台に付属する移動式の拡大鏡を動かして黒線が揃った箇所を探し当てると、その隣に記された診断名が目に入るという寸法です。



図1. NashのLogoscope (筆者所蔵)

この仕組みは、Diagnostic slide rule(診断尺)とも呼ばれ、「症状」を入力し「診断名」が出力されるという点において、医療用人工知能の先駆であると考えられます。そして、20年を経て、実用的なコンピュータが一般化し、症状から診断名を推論するプログラムが開発されました。その代表的なシステムが、スタンフォード大学において開発されたMycinと呼ばれるシステムです<sup>7)</sup>。これは、医師が有する感染症に関する知識を、「Aであれば、Bが疑われる」といったルールの集合として表現し、この情報に基づいて感染症の診断推論を行うシステムでした。この試みは、その後、専門家の知識をコンピュータに代替させる「エキスパートシステム」と呼ばれるシステムの先駆となり、80年代におけるAIブームへと繋がりました。

そこからさらに30年を経た現在、医療用人工知能研究の主戦場は、画像認識技術の応用としての医療画像診断に加えて、ゲノムを対象とした生命情報科学にシフトしています。医療画像診断技術の発展は、いわゆる「ディープラーニング」技術によって画像認識の精度が飛躍的に向上したうえに、画像情報の標準化が進んでいたために応用研究が容易であったという背景がありそうです。後者もまた、技術革新に支えられゲノム情報の利用が低コストとなったことに加えて、画像認識技術のような「膨大な情報から一定のパターンを抽出する要素技術」の精度向上に支えられていると考えられます。つまるところ、認識の対象が画像かゲノムかという違いがあるだけで、これらはどちらも、技術革新により実現した「コンピュータの眼」を、医療の核である診断に応用した結果といえます。

## プライマリ・ケアへの応用

このように、人工知能の医療応用は、診断支援という文脈で進展しつつあります。これは、プライマリ・ケアにおいても有用な技術に違いありません。しかし、プライマリ・ケアにおいて医師の接する患者の多くは、診断自体には困らないcommon diseaseとなります。したがって、プライマリ・ケアへの人工知能の応用では、このcommon diseaseへの対応に代表される「ルーチ的な業務」に対する支援に期待が持たれる

こととなります。また、プライマリ・ケアの発展に向けては、医師を対象とした診断支援だけでなく、患者やコメディカルを対象とした活用も考えられるでしょう。そこで以下では、表1に示す分類を用いて、人工知能技術のプライマリ・ケアへの応用の可能性について、概観してみましょう。

表1. プライマリケアにおける人工知能の応用例

	医師を対象	医師以外を対象
ルーチン的な業務	外来における検査・処方支援 退院サマリの自動生成	自動問診システム お薬手帳管理
非ルーチン的な業務	診断困難時の診療支援 (狭義の診断支援システム)	タクシー簡易救急車化

まず、一番に期待されるのが、医師を対象としたルーチン業務の支援です。たとえば、私達が携帯電話にてメッセージを送る場合、今までの入力履歴に基づいて文字入力を予測し、少ない操作で本文を記すことが一般的となっています。こうした「予測入力」技術をオーダーリングシステムに応用すれば、患者の所見に基づいて、検査や処方を提案してくれるような仕組みが実現します。こうした仕組みは、現在、薬事法(薬機法)の観点からの難しさがありますが、人工知能による医師の外来業務支援技術として期待されます。病棟業務の支援技術としては、入院カルテの情報を元にした「退院サマリの自動生成技術」などに、大きなニーズがあるでしょう。

次に重要となるのが、数の上で医師を上回る看護師などのコメディカルや、患者を直接支援する仕組みです。たとえば、最近では、問診票の代わりに、タブレット端末を用いた「自動問診システム」が数多く開発されています。こうしたシステムの多くは、紙の問診票を電子化し、入力結果を電子カルテに効率的に流し込む仕組みを備えています。今後、スマートフォン等で急速に発達している「対話システム」に類する人工知能を活用することにより、単純な選択肢による問診票以上に効果的な問診が実現できる可能性があります。「電子お薬手帳」と組み合わせることにより、多剤処方を患者側でチェックし薬剤師に質問できるようにするような仕組みも、プライマリ・ケアにおける人工知能の有益な応用と考えられます。

医師を対象とした非ルーチン業務としては、診断困難時を対象とした診断支援が想起されます。臨床においては、大学病院であれ地方の診療所であれ、病態生理をうまく説明できない診断困難症例に遭遇し得るでしょう。そうした際、医師に対して可能な鑑別疾患を示したり、有力な診断仮説を提示したり、次に行うべき検査を提案したりする技術があれば、医師の負担を大いに軽減するものと考えられます。これは狭義の診断支援システムにあたり、多くの研究がなされてきました。

最後に、医師以外を対象とした非ルーチン業務の支援策です。たとえば、現在、軽症患者による救急車の利用により、救急外来の負担が増し、特に台数が少ない地方において救急車が占有されてしまう問題があります。そこで、タクシー乗務員を対象とした診断支援ツールを提供してはどうでしょうか。乗客に「病院に行ってほしい」と頼まれた際、症状を入力すると、どのような疾患が疑われ、どの病院に送れば良いかわかるスマホアプリを開発し、配布するのです。スマホにはGPSが備わっているために、救急隊が利用している救急搬送先の病院データと組み合わせることで、最適な病院へのナビゲーションが実現します。メディカルコントロール医との連携機能を持たせることで、さらに有用性は増すでしょう。こうしたアプリがタクシー向けに提供されていれば、患者も安心してタクシーを利用することが可能となり、救急車の適正利用に繋がるでしょう。

## 医療用人工知能の研究開発に向けて

このように、医療用人工知能には、医療を支えていくためのさまざまな可能性があります。しかし、実際にはあまり研究が進んでいるとはいえない状況にあります。政府の投資は、研究として価値が高い、すなわち、論文になりやすい医療用画像認識やがん治療といった分野に偏っています。医療用人工知能分野は、民間の投資意欲も旺盛な分野ですが、民間投資はビジネスとして成り立つ分野に限定されます。結果として、公益性が高くても論文になりにくかったり、収益が望めなかったりする分野は、過少投資となりがちです。救急車支援技術などは、まさにこの状況が当てはまりそうです。

また、医療用人工知能の研究開発には、医療・医学と情報技術の双方に通じた人材が不可欠ですが、我が国にはそうした人材に限られており、研究開発における障害となっています。人材という点では、研究をリードする人材だけでなく、研究用の膨大なデータを準備して下さる研究支援者の不足も深刻です。現在の人工知能研究では、コンピュータは人間が作成した膨大な「お手本データ」から知識を獲得します。研究の価値を理解し、そのデータをこつこつと作成する医療従事者なしには、医療用人工知能研究は成り立ちません。こうして研究開発した医療用人工知能を試用し、また、評価を得ていく上で、各医療機関の経営者など意思決定者への啓発も重要です。

そこで我々は、今年度より、「保健医療用人工知能の技術革新と国際競争力向上に資する人材育成に関する研究」(厚生労働科学研究費補助金研究)を開始しました。この研究により、我が国における医療用人工知能研究を進めていくうえでの課題となっている人材育成に、研究当事者として貢献することを目指しています。こうした試みにおいて、プライマリ・ケアに関わる医療従事者の関与は、技術の健全な発展に向けて欠かすことが出来ません。そこで、昨年のプライマリケア・連合学会学術大会でも、シンポジウム「プライマリ・ケアにおける人工知能の可能性」を開催させて頂きました。

医療用人工知能は、プライマリ・ケアを含め、医療・医学の発展に直結する技術です。しかし、その研究開発においては、人材不足の問題に加えて、薬事法制上の問題など、さまざまな課題が存在します。今後、本稿のような機会を通じて学会員の皆様と問題意識を共有すると共に、学会での討議を通じて、議論が深まることを期待しています。こうしてプライマリ・ケア分野ならではの成果を示していくことにより、医療全体における人工知能の健全な活用に繋がることを願っています。

### 参考文献

1. 松尾豊. 人工知能は人間を超えるか. 株式会社KADOKAWA. 2015年3月.
2. 宇賀神 幸司. 熟れた実だけ採取する「農業収穫ロボ」の実力：日経ビジネスオンライン.  
<http://business.nikkeibp.co.jp/atcl/report/15/226265/093000057/> (2017年12月28日アクセス).
3. セコム株式会社. HYPERLINK "<https://www.secom.co.jp/campaign/robotx.html>" (2017年12月26日アクセス)
4. 総合警備保障株式会社. HYPERLINK "<http://www.alsok.co.jp/corporate/robot/>"  
<http://www.alsok.co.jp/corporate/robot/> (2017年12月26日アクセス)
5. エヌ・デーソフトウェア株式会社. HYPERLINK "<http://www.ndsoft.jp/bizapp/>"  
<http://www.ndsoft.jp/bizapp/> (2017年12月26日アクセス)
6. Nash, F.A. Diagnostic reasoning and the Logoscope. The Lancet 276.7166(1960):1442-1446.
7. Shortliffe, Edward H. A rule-based computer program for advising physicians regarding antimicrobial therapy selection. Proceedings of the 1974 annual ACM conference-Volume 2. ACM, 1974.

出典 奥村貴史, "プライマリ・ケアと人工知能", プライマリ・ケア, 日本プライマリ・ケア連合学会, プライマリ・ケア, Vol.3, No.1, 2018, pp.72-75.

## 医療用人工知能を知る

[医療用人工知能とは \(研究者向け\)](#)

[医療用人工知能とは \(医師向け\)](#)

[医療用人工知能の応用事例](#)

# 国立保健医療科学院における診断支援研究

医師向け

## 診断支援AI「PGY-01」の歴史・現状・展望

2017年8月10日 奥村貴史 (国立保健医療科学院 研究情報支援研究センター 特命上席主任研究官、国立情報学研究所客員教授)

## 診断支援システムの実際

本年5月13日、第8回 日本プライマリ・ケア連合学会学術大会において、「プライマリ・ケアにおける人工知能の可能性」と題したシンポジウムが開催されました。人工知能(AI)技術の医療への応用が進むことで、各学会においてもAIに関連した研究発表やシンポジウムが増えています。そうした中でも、医療用AIがプライマリ・ケアという第一線の医療現場における活用についても討議されるまでになったことは、この分野の試みが基礎研究から応用研究にまで幅広く広がってきたことを示しています。

「私たちは、このシンポジウムで、「診断困難症例の診断支援」と題した発表を行いました。臨床では、大きな大学病院から小さな診療所に至るまで、どのような環境であっても、時折、診断に困る症例に遭遇するでしょう。そうした際、全身状態が悪ければ、必要に応じて高次施設に紹介するかもしれません。しかし、他施設に紹介するにしても、適切な科に送りたいという心理は働きます。症状が軽く明らかに生命予後に関わらなければ、診断が付かない状況でも経過観察を選択するかもしれません。それでも、自分を頼って受診をしてくださった患者のためにまずは診断を付けたいという気持ちも、医師誰もが感じることでしょう。

そうした際、診断の糸口をつかむために、先生方は何をどのように調べられるでしょうか。病名が明らかとなっている際には、成書を確認するというのが正しいあり方かもしれません。しかし、そもそも診断が付いていない場合には、症状を頼りにウェブで関係しそうな疾患を検索するという方法が少なからず行われているようです。この方法の問題は、良質な情報にたどり着くためには、検索結果に現れる膨大なページや論文を一つ一つ確認する等の手間と経験が必要である点です。昔であれば図書館に行かなければならなかった文献検索が、昨今ではネットで簡単に調べられるようになりました。一方で、手に入れられる情報も膨大となり、新たな非効率を生んでいます。

私たちは、この問題に取り組み、診断困難症例への対応を支援するシステムを2009年から研究してきました。図1は、その入力画面です。画面上部にある「基本情報欄」には、患者の基本情報を入力します。例えば、年齢と性別を入力することで、診断の計算に好発年齢や性別を加味することが可能となります。同様に、急性か慢性か、あるいは亜急性かといった発症様式を問うことにより、疾患の自然歴を加味した計算が可能となります。主訴現症欄は、入力情報の中心となります。これら臨床所見を入力する毎に、画面右にある「選択症状」に追加され、必要な情報を入力した後に「病名診断」ボタンを押すと、結果画面へと切り替わります。

図2は、結果の表示画面です。「神経系の疾患」という表示の中に、中枢神経系の炎症性疾患や変性疾患といった疾患グループが並び、さらにその中にあるアルツハイマー病やレビー小体病が並んでいます。このように疾患を階層的に表示することにより、何百という鑑別疾患を効率的に閲覧することが可能となります。システムには、現在、1500ほどの疾患情報が登録されており、今年度中に5465件ほどに拡張される予定です。

さらに、図3に示すように、それぞれの疾患に関する説明文や詳細情報を効率的に確認することが可能となっています。画面例では、疾患説明文の下に、疾患情報、所見情報、関連疾患、参考文献の詳細情報タブが示されています。参考文献はとりわけ有用で、それぞれの疾患に関するオンライン上の各種文献への

リンクが整備されています。そのなかでも、メルクマニュアルや難病情報センター等、品質の高いサイトはアイコン化されており、直接移動することができます。診断支援システムがウェブや論文の検索システムと比して大幅に効率的である点を、ご理解いただけることと思います。

図1 症状入力画面 (提供：奥村氏)

図2 結果表示画面 (提供：奥村氏)



図3 疾患情報表示 (提供：奥村氏)

## 国立保健医療科学院における診断支援研究

ここで、一つ疑問に思われるかもしれませんが。ご紹介したシステムは「それなりに有用に見える」のに、なぜ今まで知られていなかったのでしょうか。この疑問に答えるためには、私たちの研究の背景についてお伝えする必要があります。

従来、日本の医療では、公費負担の対象として56種類の「難病」が知られていました。その後、2015年に制度が変更され、現在では330疾患が認められています。これは、難病対策の大きな転換であったことから、その制度変更にも先駆けてさまざまな検討がなされました。しかし、たとえ対象疾患を広げても、疾患概念として確立していない難病の患者を救うことはできません。疾患概念として確立していなければ、そもそも患者を診断することが困難です。診断基準の策定に向けた研究を進めようにも、明確な症例定義がなければ、症例を集めることも困難でしょう。

そこで2009年、「未分類疾患の情報集約に関する研究班」(研究代表：国立保健医療科学院の林謙治院長)が設けられました。医学は今までさまざまな疾患に対して多くの科学的知見をもたらしてきました。しかし、そもそも名前がなく症例定義も定まらない疾患に対する研究方法論は知られていませんでした。そこで私たちが着目したのが、オンラインの診断支援システムでした。臨床においては、大学病院や診療所といった施設区分を問わず、多かれ少なかれ診断に困る症例に遭遇します。そうした際のために医師の診断を効果的に支援するシステムをネット公開すれば、多くの医師が利用することになるでしょう。そして、もし、「未知の疾患」が既知の疾患と異なる症状を呈するのであれば、このシステムの検索履歴中に、既知疾患とは異なるパターンの検索が現れるはずで。そうした患者に地域集積性が認められれば、未知の公害病の端緒をつかめるかもしれません。もし職業集積性が認められれば、職業性胆管癌のような未知の公害が示唆されるかもしれません。国内で知られていない感染症が疑われれば、新興感染症やバイオテロの検知にもつながるでしょう。

このような体制を実現するためには、まず、臨床的に有用な診断支援システムを研究開発しなければなりません。そこで、私たちは、実際に動作する診断支援システムを試作した上で、医師を対象としたセミナーを開催して症例検討を行い、出席した先生方から頂戴したフィードバックを基にシステムを改良し、それをまた次回のセミナーで供覧する、という試みを8年間ほど続けてきました。臨床現場での実用に耐えるシステムの研究開発には、大変地道な作業を要します。開発途中の不完全なシステムに対してフィードバックを得ていくためには、システムの将来的な価値を理解して下さる限られたテストユーザーから、根気強くフィードバックを集めていかなければなりません。その際、良質なフィードバックをいただくためには、頂戴したご意見を真摯に受け止めシステムを改良していかなければならず、多くの時間がかかることになります。

また、診断支援システムを大々的に広報する上では、診断性能を担保するために臨床的な評価を行う必要があります。しかし、臨床研究には多額の予算がかかります。そこで私たちは、まずは基礎研究として要素技術の完成度を高める道を選び、主に情報系の国際学会を中心に研究成果の発表を行ってきました。私たちに



としては、臨床的な性能評価を行ったり、広報を行い渉外に要する負担を増すよりも、現場の臨床医の先生方との対話を通じて技術としての完成度を高め、研究成果を一つ一つ積み重ねていくことが重要だったのです。

## 診断支援技術とその本質

さて、私たちの診断支援システムには既に8年の研究実績があることから、さぞ高度な診断が可能かと思われるかもしれません。しかし、私たちのシステムの診断性能そのものは、決して高いものではありません。それでも、私たちのシステムは、前述のシンポジウムや今まで積み重ねてきたセミナーにおいて、臨床の先生方より高い評価をいただけてきました。次に、その理由をお伝えすることで、そもそも診断支援とは何かという技術の本質を明らかにしたいと思います。

### 医師の有する知識の補完

人類が今までに発見した疾患がどれくらいあるか、ご存知でしょうか。数え方にもよりますが、WHO（世界保健機関）は、2万～3万という数字を挙げています。分子生物学の発展によってさまざまな疾患に多くの責任遺伝子が見出されてきたことで、区別できる疾患の数が増えたことが背景にあるようです。歴史的には、18世紀頃には既に2400種程度の疾患が知られていたようですので、医学の進歩が疾患概念をおよそ10倍に増やしたと言えるでしょう。

問題は、この中で一人の医師が知っている疾患数はどれくらいかというものです。実は、この問題には正確な調査がなく、私たちの調査により概算で2000疾患弱であると推計されています。つまり、普通の医師は既知疾患の一部を知っているにすぎず、実は大多数の希少疾患については知らないことになるのです。この希少疾患に関する知識量では、生身の医師はコンピュータに決してかないません。医師が知らない希少疾患についての知識をコンピュータが支援することができれば、人間と機械が得手不得手を補い合うことができることになります。

### 医師の有する各種バイアスとデバイアシング

以上のように記すと、「医師は希少疾患を知らないかもしれないが、医学知識に基づいた診断推論が可能であり、簡単に機械には負けないのではないか」とお思いになられるかもしれません。これはその通りで、AIは網羅的なデータベースを用いることで、可能性のある疾患を列挙することは得意ですけれども、今のところ医師が有するような診断推論能力はありません。一般的な症例において医師の診断能力がAIを凌駕するのも、人間の診断推論能力が高度であることが理由の一つです。

しかし、医師の推論能力にも、大きな欠陥があります。それが、人間だれもが有する「バイアス」です。診断に際して、医師はさまざまな情報に影響を受けます。患者からの訴えや説明、患者が発する雰囲気、その日に診察した他の患者の容態や自分が今までに診た患者、とりわけ、困難な診断にたどり着いた成功体験などは、とりわけ医師の思考を制約します。CT上に見つけた重要所見に意識を奪われ、隣のスライスにあったわずかな変化を見逃すような経験は、誰もが通る道でしょう。AIの長所の一つは、人間と異なりこれらのバイアスに影響を受けない点です。バイアスがないことで、機械は時折、医師が想像もしていなかった可能性を提示します。診断支援システムは、たとえ正解を示すことができないとしても、この「デバイアシング(debiasing)」能力によって医師の有するバイアスをキャンセルすることが臨床に高い価値を有するものと考えられます。

### 関連情報の効率的な提示

進行に時間がかかる疾患においては、診察時に全ての症状が発現しているとは限りません。その場合、限られた所見から正確な診断を下すことは原理的に困難です。例えば、腹部の違和感という病歴のみから、胃軸捻転症を確定診断することはできません。したがって、診断困難症例の診察を支援する診断支援システムは、正確な診断結果を示すことよりも、「診断に役立つ情報を提供すること」に価値があると考えられます。とりわけ、診断困難な症例に接した際に、質の良い診断仮説を提示したり、その質の良い診断仮説に基づいて次に行うべき検査を示唆したりということ機能は、臨床的に極めて有益です。

そのように考えると、診断支援システムにおいては、答えが分かっている症例に対する正診率が99%あるかないかという単体性能で評価することが困難であることをご理解いただけるでしょう。診断支援システムは、医師と相互補完的に働くものであり、その総体として評価しなければ、診断支援システムの利用による臨床的なアウトカムを正しく把握することは困難なのです。したがって、診断に関連した情報を的確に、また、効率的に提示することが、診断支援システムにおいて本質的な要素となります。

## 医師と診断AIの協働に向けて

今までの研究を通じて、医療用AIが発達することによって「医師が怠惰となる」懸念を少なからずお伺いしてきました。このように、機械の発達によって人間の能力が落ちるのではないかという危惧は、AIの本格的な登場以前より指摘されてきました。とりわけ深刻であったのが、飛行機の自動操縦が発達することによってパイロットの技量が衰え、異常事態への対処能力というパイロットにとっての根源的な能力が衰えるのではないかという危惧でした。こうした事態をautomation complacencyといい、医療においても当てはまるのかもしれませんが。確かに、CTやMRIの普及によって頭部の画像診断が容易となったことで、医師の神経学的診察能力は低下してきたのかもしれませんが。この後、診断までもを機械に依存することで医師の診断能力が衰えていくことは、決して望ましい未来ではありません。

しかし、多くの医師が臨床において眼鏡という「機械」に頼っていますが、医師の臨床能力は減退したでしょうか。靴を履くことによって、人間の活動領域は明らかに広がりました。これらの事例は、機械と人間とがうまく補い合うことができれば、人間の能力が減退ではなく発展していくことも可能であることを示しています。とりわけ大切なことは、診断支援システムや医療用AIは医師や医師の診断を置き換えるものではなく、医師と相互補完的に働くものだという認識です。実際、診断学分野で有名なUCSF（米カリフォルニア大学サンフランシスコ校）のDr. Dhaliwalも、自身で診断を行った後に、診断支援システムを用いて自身自身の診断推論に見落としがないか確認されるそうです。教育への応用も重要でしょう。今後は、眼鏡が人間の能力を拡張したように、また、飛行機が人間を豊かにしたように、診断支援技術が人間の将来に貢献するような方向に研究開発を導いていくことが重要と言えます。

そのためには、診断支援システムの要素技術がバランス良く健全に発展していく必要があるでしょう。具体的には、医学知識をコンピュータが利用可能な形に整備した「知識ベース」、医師とのやり取りを担う「ユーザーインターフェース」と、入力された所見等から実際の診断を行う「診断エンジン」、それぞれが高度化していく必要があります。そこで、私たちは、「疾患知識ベース」から研究に取り組み始め、現在、医師とコンピュータが協働していく要となる診断AIのユーザーインターフェース研究へと辿り着きました。

私たちのシステムは、1年目研修医の状況に似ていると考えています。つまり、「国家試験に通ったばかりであることから覚えている疾患の単純な数はベテランよりも多い」ものの、「学んだ知識を組み合わせる診断を行ったり適切な治療を考えたりという能力は一般的な医師には到底及ばない」という状態です。そこで、研修医が先達に育てられ経験を積むことで臨床能力を高めていくように、私たちのシステムも先生方からの指導を通じてより有用で効果的なシステムへと育つことを願って、1年目研修医を意味するPGY-01(※)と名付けています。今後、医療用AIのユーザーインターフェース研究を中心に基礎研究を進めつつ、臨床の先生方との共同研究を少しずつ増やすことにより、より有益なシステムへと育つこと、そして、医学の発展に資する研究成果につながることを願っています。

※ PGYは、本来Post-Graduate Yearですが、ここではProgrammed General phYsician。

※M3.com『[医療維新](#)』より許可を得て転載

## 医療用人工知能を作る

[イベント報告・論文紹介](#)

[医療用人工知能と政策](#)

[医療用人工知能の作り方 / 研究開発の中心となるデータ作成の実例](#)

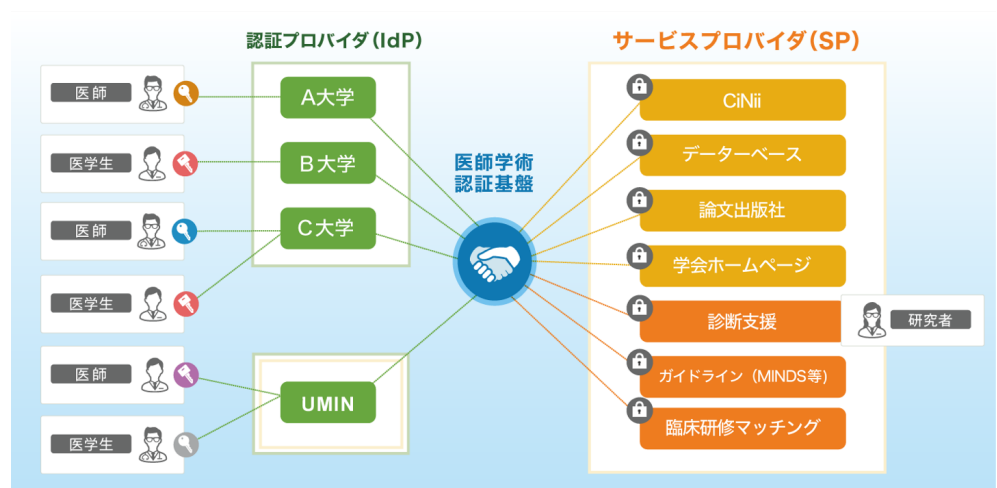
[医療用人工知能研究に求められるデータ作成作業の海外委託](#)

行政官向け

## 医師学術認証基盤構想

この記事は9分で読めます

日本の医療用人工知能は、米国、中国に大きく差をつけられつつあります。その背景には、いくつかの理由があり、とりわけ、研究開発と技術の実用化に要する様々なコストが高い点への対応は急務です。「医師学術認証基盤構想」は、この問題への有力な解決策の一つとして、我々が提案してきた施策です。この研究開発基盤により、医療用人工知能研究への対応を切り口に、「医師認証基盤の普及」、「医療用人工知能の発展」、「医師キャリアの効率的な追跡」という、日本の医療を支える多彩な波及効果が期待されます。



[資料ダウンロード \(PDF: 1.7MB\)](#)

### 医師学術認証基盤とは

医師学術認証基盤とは、医師を対象としたオンラインの認証サービスを提供する仕組みです。医師が普段利用するアカウントを利用し、中医療関連分野の様々なサービスを利用することを実現します。政府は今まで、保険医療分野の公的な認証基盤(HPKI)の運用を支援してきました。しかし、この基盤は法的な効力を有する電子署名システムとして構築されており、医師にとっては紹介状の電子署名以外の使い道がないことから、利活用も進んでいませんでした。医師学術認証基盤は、普段から利用されているアカウントを活用することで、この厳密な認証を必要としない幅広いアプリケーションを低コストに実現し、医療の情報化にさまざまなメリットと波及効果をもたらすことが期待されます。

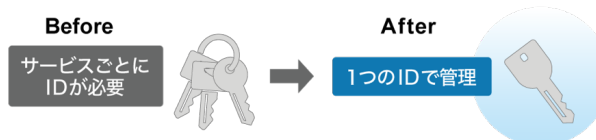
### 医師学術認証基盤が実現するメリット

医師学術認証基盤には、[パンフレット]が示すように、12項目にも及ぶメリットが期待されています。ここでは、医療用人工知能と保健医療福祉行政にとっての代表的なメリットを抜粋します。

[メリット1-医師向けサービスの統合認証化]

医師は、日常的に各種学会や医療サービスをオンライン上で利用していますが、それぞれのサービスでIDを作成し、管理する必要がありました。医師学術認証を用いることで、所属大学やUMIN等、日常的に利用しているアカウントを用いて、連携する各種学会・医療情報サービスにアクセスできるようになります。これにより、医師は様々な医師向けオンラインシステムのために、いくつものアカウント情報を管理する手間から開放されます。

こうした認証システムは、コンシューマ向けのサービスでは既に実現されています。たとえば、FacebookやTwitter、Google等のアカウントを使って様々なウェブサービスを利用できることをご存知ではないでしょうか。このように、単一のアカウントにより様々なサービスを利用できるサービスを、「シングルサインオン」と称します。医師学術認証基盤により、医師向けな多彩なサービスのシングルサインオン化が実現します。



### 〔メリット 2－医療用AIの研究開発促進〕

医療用人工知能の研究開発には、人工知能に正解を伝えたり、開発した人工知能の精度管理を行なうために、多くのデータが必要となります。こうしたデータの生成に関しては、医師でなければ品質を確保することが困難であることが少なくありません。しかし、膨大な数のデータを生成する単純作業に従事して下さる臨床医は決して多くはありません。また、その単純作業の意義を啓発するためにも、時間と手間を要します。

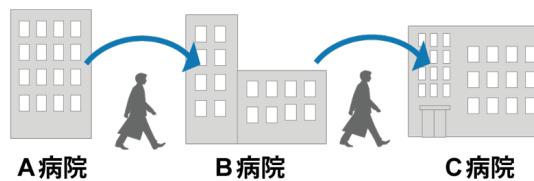
医師学術認証が実現することにより、たとえば、研究用データの生成タスクをインターネット上で公開し、医師のみを限定して受け入れるような体制が簡単に実現できるようになります。医師学術認証は、このように、研究協力をして下さる医師の参加促進を通じて、医療用人工知能研究のコストを大幅に低減します。これは、医師学術認証基盤を医療用人工知能の開発に必要なデータを作成する「クラウドソーシング」(別記事：「[医療用人工知能研究に求められるデータ生成・監査タスクを海外委託する](#)」参照)に活用する試みです。実際にクラウドサービス上で医師にデータ生成を依頼する試みも行われていますが、国内ではクラウドサービスを通じてタスクを請け負う医師は多くありません。医師学術認証基盤は、医師の資格確認を大幅に軽減すると共に、アカウント管理の負担をなくすことにより、医療用人工知能研究のコスト削減を通じて研究開発の促進に貢献します。



### 〔メリット 3－医師キャリアの効果的な追跡〕

医学部においては、卒業生の進路や勤務先の把握によって大学教育や入学試験制度の改善に繋げることができます。そこで、同窓会組織等を通じた卒業生の組織化が行われていますが、自主的な報告に基づく動向把握は完全ではなく、また、医師側、同窓会組織側の双方への負担が大きいものでした。適切な同意のもとに在籍医学部がアカウント発行を行うことは、そのアカウント情報を通じた卒業生のキャリアの効率的な把握を実現します。

政府は、医師需給に関する政策の検討に向けた基礎資料として、医師の卒後動向の調査システムの確立を望んできました。しかし、こうしたシステムの確立のためには大きな予算を要することに加えて、たとえ予算化が実現したとしても、医師側に報告するインセンティブを設けることは容易ではありません。出身大学を含む各組織が協力する分散型の医師認証基盤を実現することにより、医師キャリアの効率的な追跡を低コストに実現します。



## 低コスト化の工夫

このように医師学術認証基盤には様々なメリットが期待されています。しかし、そのような施策の実施には、従来、多大なコストがハードルとなってきました。そこで次に、医師学術認証基盤が、なぜ、どのようにして低コストであるのかを紹介します。

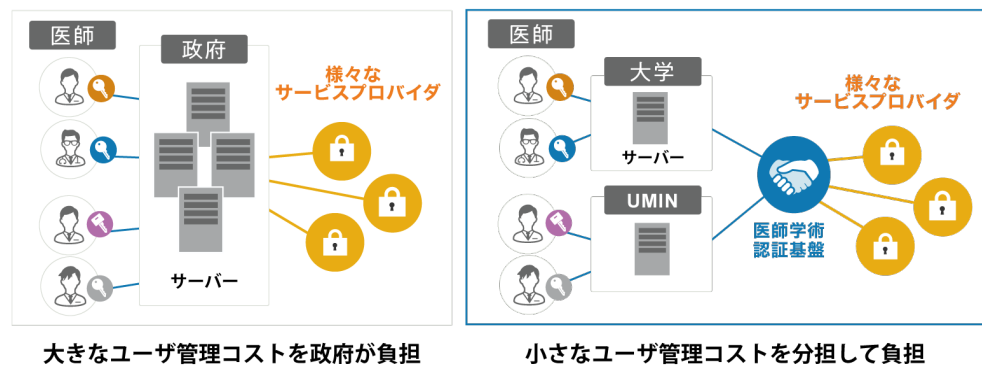
### 〔工夫1－フェデレーション型認証の利用〕

医師学術認証基盤では、新たな政府認証局のような公的基盤を構築「しません」。代わりに、各大学やUMIN(大学病院医療情報ネットワーク)等にて医師が利用しているアカウントを、相互接続します。これにより、各医師は、自分がメインで利用しているアカウントを維持するだけで、インターネット上の様々な医師向けサービスを利用することが可能となることが期待されます。FacebookやTwitterのアカウントにより、他のさまざまなオンラインサービスが利用できるようなイメージです。このような認証システムは、フェデレーション型の認証と呼び、世界の学術認証において利用されています。我が国においても、国立情報学研究所が、主として大学等の高等教育機関を対象として運用しているフェデレーション型認証基盤「学認」があり、その仕組みを再利用することで、低コストに医師向けの学術認証基盤を実現します。



### 〔工夫2－ユーザー管理コストの軽減〕

認証基盤の維持には、インフラの維持コストだけでなく、ユーザーの登録・維持といった管理コストが掛かります。提案手法では、ユーザーの管理は個々の大学医学部やUMINという既存組織に委ねられていることから、新たに政府は大きな管理コストを予算化する必要がありません。

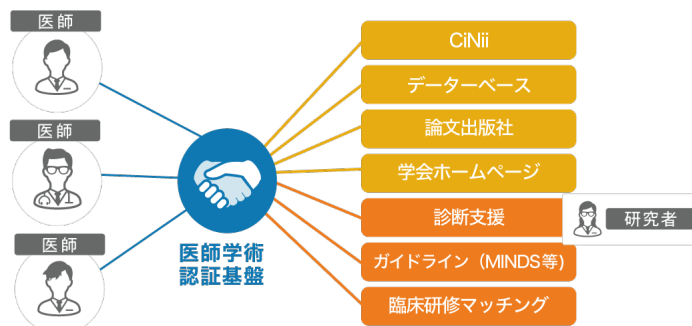


また、今まで政府は、「臨床研修マッチングプログラム」のために大きなコストを負担し全国の医学生にマッチングプログラム参加のためのアカウントを配布して来ましたが、このアカウントはマッチング終了と同時に利用が終わり、無駄が生じていました。医師学術認証基盤を利用することにより、医学生は、在学時に利用していたアカウントにより「臨床研修マッチングプログラム」への参加が実現します。これは、臨床研修予算を、医師学術認証基盤の実現に振り向けうることを意味しています。

### 〔工夫3－サービス提供コストの軽減〕

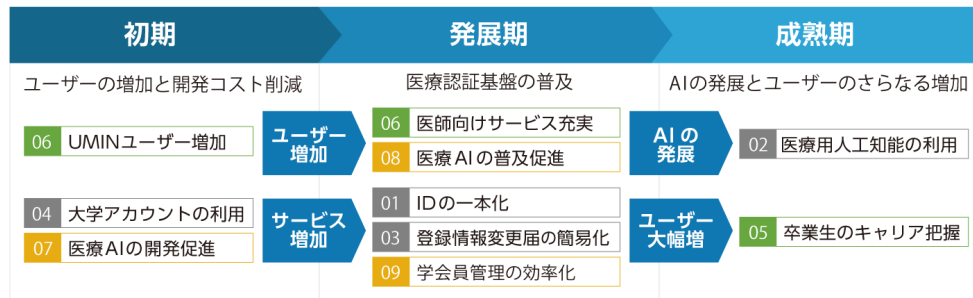
従来、医師向けサービスを提供する学会や事業者、公的機関は、ユーザーの獲得と維持のための事務コスト、システム運用コストと大きなコストを負担する必要がありました。このコストがハードルであることから、医師向けサービスを立ち上げたり拡大することは容易ではありません。結果として、医師を対象とした様々な情報システムの開発と普及が進まないという問題がありました。

医師学術認証基盤によって、サービス提供側は認証基盤への接続に要するコストを負担するだけで、多くの医師ユーザーを獲得することが可能となります。フェデレーションの運用のために要するコストは、商用サービスの接続に対して課金を行う等することにより容易に確保することが可能となると考えられます。



## 医師学術認証基盤の実現へ向けて

以上のように医師学術認証基盤は、サービスを利用する医師、サービスを提供する研究者や各種事業者、アカウントを提供する大学等組織のそれぞれに大きなメリットがあります。そして、このように参加各位に様々なメリットが存在することで、事業に参加する各種の主体が増し、サービス種別の拡大が情報基盤としての価値を高めます。こうして利用医師が増えることで、サービス提供者が拡大し、さらに利便が高まるサイクルが成立します。



こうした仕組みの整備は、インターネットが爆発的に普及した背景にある「イノベーションのためのシステム」の確立に似ています。政府は1980年代、日本社会の情報化に向けてキャブテンシステムやISDNといった公的情報技術の普及を目指し、多くの投資を行いながらも挫折を続けてきました。そうしたなか、1990年代に大きく普及したインターネットは、利用者が技術革新を主導することで絶え間ない技術革新を実現し、これがネットワークの価値を高めてさらなるユーザーを獲得するイノベーションシステムを実現し、爆発的な普及を実現しました。

我が国の医療用情報技術は、コストが高いうえに品質が低い特徴を有しています。そうした状況において技術革新を実現していくためには、コストを低廉化し、品質を上げる施策が不可欠です。そのためには、その公的な性質が故に目的が制限された政府認証基盤(HPKI)に加えて、ユーザー数とサービス種別の拡大を実現する医師学術認証基盤の導入が合理的です。今後、啓発活動を通じて関係者による理解を得ることで、施策化に向けた検討が進められることを願っています。





## 医療用人工知能を作る

[イベント報告・論文紹介](#)

[医療用人工知能と政策](#)

[医療用人工知能の作り方 / 研究開発の中心となるデータ作成の実例](#)

[医療用人工知能研究に求められるデータ作成作業の海外委託](#)

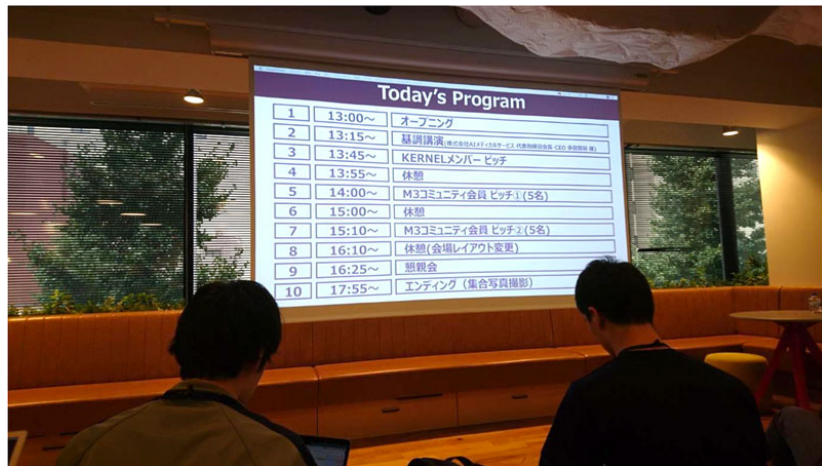
# 医療用AI開発者イベント参加記

## -DEEPCORE+M3交流ピッチイベントに参加して-

この記事は7分で読めます

旭川医科大学 医学部医学科 小笠原亜美

(監修: 北見工業大学 奥村 貴史)



## 医師とITエンジニアの交流からみえた現代医療が抱える課題

2018年9月29日午後、東京文京区本郷にて、エムスリー社「AIラボ」とAI特化型インキュベーターであるディープコア社による、医療AIについて医師とエンジニアの交流を目的としたイベントが開催された。講演や懇親会での交流を通じて、医師側とエンジニア側の双方が抱える疑問や課題について意見交換をすること、そこから新たな課題を発見することが今回のイベントの目的である。

主催者による開会の挨拶の後、多田智裕さん(医師AIメディカルサービス代表取締役予定)が基調講演を行った後、高橋秀徳さん(医師自治医科大学)をトップバッターに、医師、エンジニアを含む計10名による「ピッチ」と称するショートプレゼンテーションが行われた。

5時間に渡ったイベントにおいて医師側は、AIを医療の現場でどのように活用しているのか、また現場ではたらく医師がどのようなことをAIに期待しているのかといった話題を提供した。それに対してエンジニア側は、具体的にどのようにエンジニア側が現場や事業に介入できるのか、またビジネスとしての市場価値はどこにあるのかといった疑問を医師側に直接質問したり、意見を述べるなどした。

医療用AIに興味を持つ医学生として今回イベントに参加し、医師とITエンジニア双方から現状や今の状況に対する意見を聞いた結果、医療の現場に医療AIの導入やITエンジニアの参入を行うために解決した方がよい課題がみえてきた。イベント参加記として、以下に紹介したい。

## 医師の抱える困りごとから始まる医療用AI開発

開会式の後、最初に医師であり起業家でもある多田智裕さんによる基調講演が行われた。世界に挑戦する内視鏡AIというタイトルで、内視鏡画像から癌を識別するAIを作成した経緯や今後の展望について語った。「はじめは現場の困りごとだった」と多田さんは述べた。さいたま市の市民検診で撮影された内視鏡画像のダブルチェックは終業後に行われる。検診を受けた市民の数は4~5万人、1人につき4、50枚の画像が存在



するので医師一人あたり毎週2,800枚もの内視鏡画像をチェックしなければならない。技術の進歩によって大量に生産されるようになった医療画像は専門医の処理能力を超えていた。このような日常業務の困りごとから、トップ医療機関と共同研究により、50万枚の内視鏡画像から、AIを用いたダブルチェックができるソフトが製作された。現在はリアルタイムで内視鏡を使用している間に癌を診断できるシステムを開発中である。AI内視鏡画像診断での癌の見逃しゼロを目指すとして述べながら、「治療方針や手技のサポートなどといった用途があり、AIは医師の道具である」と強調していた。

次に、眼科医師である高橋秀徳さんがプレゼンテーションを行った。タイトルは「たたみ込みニューラルネットワークで解決する画像診断」で、機械学習を用いた糖尿病網膜症病期分類のための画像診断AIについて発表した。現在AIは一般画像識別において画像識別に特化した人の認識率を超えたという。加えて人工知能に慣れていなくても、ディープラーニングが登場したことにより、人工知能の学習方法が洗練されたことで広いユーザーが使えるようになったと高橋さんは述べた。高橋さんが作成したAIは、診断、病期分類、数値推測これら全てがたたみ込みニューラルネットワークで画像から自動的にいき、開発したAIの病期的中率81%であったという。一方人工知能の開発で苦労した点の1つとして学習パラメータやハイパーパラメータの調整を挙げている。他には人工知能に学習させるためのデータを獲得するために様々な課題をクリアしなければならないと述べていた。精度の高い人工知能を作るためには大量の学習用データが必要とする。大学病院と周辺クリニックなどからデータをもらうためにデータの契約形態と協力形態の方法、個人情報のマスク方法を倫理委員会に通したり、大学病院の外の有識者委員会から許可を貰ったりと人工知能を開発する前の準備段階からの苦労がみえた。フロアからの質問で、今から作るには最小何例必要かという問いに対して、「特化型AIでは、1万枚は必要」と述べていた。

続いて現役医師10名による「ピッチ」が行われた。匿名、非公開を前提とした発表も多いことから氏名と内容については伏せるが、今の医療の課題を解決する手段としてAIをどのように利用しているか、AIに現場のどのような困りごとを解決してもらいたいと期待しているかなど、一人一人様々な視点で発表していた。

## 交流を通じて明らかとなる様々な困りごと

イベントの最後には懇親会が行われ、AIに造詣が深いITエンジニアと、AIを利用していないし興味のある医療関係者との、交流の場となった。本イベントを開催した方々、講演を行った医師やイベントに興味を持って参加した大学生など様々な立場の方と話す機会があった。

懇親会中、ディープコア社の田中さんは「いま医療用AIに興味を持ち、今の医療現場の状況に危機感をもって行動している有識者を日本中から集めてこの規模は本当に異常なこと」との懸念を述べていた。エンジニアと医師の交流イベントを設けても都内のビルのワンフロアで済んでしまった点について、今の医療の現場に対して危機感をもって行動に移そうとしている医師やエンジニアが少ないとの認識に立っての発言だ。

また同じくディープコア社の兩宮さんらに話をうかがったところ、医療という現場が特殊すぎることから、医学生や医師にとって当たり前の情報や状況を非医療従事者側が承知していないことが判明した。これは、興味がないと言うより、医療の現場の情報にITエンジニアらがアクセスする手段がなく、そういった情報を知る機会や手段がそもそも与えられていない、分からないということが大きいようだ。交流イベントや求人もほとんどなく、医療というフィールドは、当事者以外にはかなり閉鎖的であるということが理解できた。

続いて講演者の一人である高橋先生と話した。電子カルテや医師の問題意識の低さが、医療AIの導入の足かせになっているという。高橋先生は、「電子カルテのデータが先生毎に好き勝手にかかれていて規格化されていない、データ同士が紐付けされていない、簡単にアクセスできない、集約されていないデータが多すぎる。だから患者のデータを利用するときに無駄な作業が多くなってしまったり、データを抽出するのに時間がかかったり、一度取り出したデータが使い切りになったりしてしまいとても不便だ」と述べる。また電子カルテは導入も更新も高額であり、電子カルテを維持するのに数千万円がとんでしまい、ユーザーである医師が使いやすく改良したいという要望にまで手が回らないことも分かった。一方で、高橋先生の専門である眼科では、診察、検査、診断、手術といった一連の流れが1つの科でまかなえるため、AIやIT関連の導入が比較的しやすいとも述べていた。

## コミュニケーションを目的としたイベントへの需要と効果

今回の参加者の割合、講演やピッチにおける質疑応答、懇親会を通じて、医療の現場にはIT側が介入できる課題(日常の困りごと)がとてたくさんあることが分かった。しかし、ビジネスの領域として、医療現場と医師はとて扱いづらい。医師側は「とにかく困っていて、患者さんを救いたい」という気持ちが強い。それ自体は悪いことではないが、必要な情報提供がうまくできない例も存在するのではないかとこの考えが浮かぶ。「どのくらいの症例があって、どのくらいの患者がいて、それをどのようにしたら問題が解決するのか、ビジネスたり得る利益はあるのか、そういった具体的な話や背景が分からなければエンジニア側も仕事を受けられない」、ピッチの質疑応答でのエンジニアのコメントである。

また、解決が求められる困りごとだけでなく、医療の現場には、大量にデータがある。それは、外来カルテ、入院カルテのみならず、各検査や画像、さらに健康診断のデータなどにおよび、AIの研究開発に利用できるデータの原石はたくさんあることが今回のイベントで再確認できた。しかし、そのデータの管理が画一化されていないため、いわゆる「汚いデータ」であるのが難点となる点も認識できた。こうした問題を解決していくためには、電子カルテのデータベース化、ユーザーインターフェースの簡便化と画一化が求められる。

こうした問題を解決していくためには、ITエンジニア側が、常に医療者側のオーダーを聞きながらシステムを改良し続けることができる、そして医師もエンジニアを信頼して定期的にプロジェクトの会議に参加したり要望を具体的に話したりといった基盤が必要だと感じた。医師側のAIへの興味・期待、ITエンジニア側の医療領域への興味は共に高い。しかし、AIやプログラミング言語の勉強会、医師の困りごとを定期的にフォローするシステムなど人工知能を臨床現場に導入する前に、準備しなければ成らない課題がまだ数多くある。

AIに興味のある医療者でもAIを万能機だと思っている方が多く、難しいことや分からないことは人間にできないことはAIに全部任せれば解決するのではないかと、この考えの方もいた。医療用AIの発展に向けた基盤を整備し、また、医療従事者とITエンジニアの相互理解を深めるうえで、双方のコミュニケーションを目的としたイベントには潜在的な需要が大きく、また、効果的な情報交換が実現すると実感できた。



## 医療用人工知能を作る

[イベント報告・論文紹介](#)

[医療用人工知能と政策](#)

[医療用人工知能の作り方 / 研究開発の中心となるデータ作成の実際](#)

[医療用人工知能研究に求められるデータ作成作業の海外委託](#)

# 医療用人工知能の作り方

～研究開発の中心となるデータ作成の実際～

この記事は15分で読めます

## 概要

機械学習を用いた人工知能に開発において、学習データは必要不可欠なものとなってきました。この学習データの生成には人間による作業が必要不可欠あり、学習データの種類によってはその作業に高度な専門的な知識を要求します。そんな中、医療用データを用いて学習を行なう医療用人工知能は学習データの生成に医師の協力が必要となることがあります。

ところが、このデータ生成には膨大な単純作業が含まれ、医師の多くがそういった作業を好みません。こうした状況は医療用人工知能の開発を遅らせる大きな原因となり得ます。こうした作業に関して医師の理解を得られれば、医療用人工知能の実用化に向けて大きな一歩となります。

## 1. 人工知能と機械学習

近年、人工知能が急激に進歩したことで、関連技術があらゆる領域で話題になっています。顔認識技術がスマートフォンなどのセキュリティ分野に応用され、記者の代わりに記事を書き、症状から疾患を提示して医師の手助けをするようになってきました。人間の知能を再現するという人工知能の試みは長い年月を経てようやく形になってきました。こうした人工知能の成長を支えているのがディープラーニングを始めとする機械学習です。

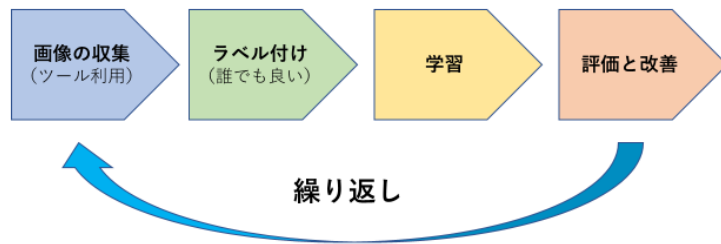
今までは人工知能の認識・判断・行動に関わるあらゆるルールを人間が自ら教えていました。しかし、人間が直接教えられることには限界があり、特に人間が無意識に行っている感覚的なタスクである画像認識や音声認識などの領域では長らく壁にぶつかっていました。それがディープラーニングによって壊れたのです。

大きく進歩した人工知能による認識能力によって人工知能は改めて注目を集めるようになり、様々な応用に期待がかかるようになってきました。その一方で、人工知能がどのように学習し、成長しているかという点についてはよく理解されていません。そこで、画像認識システムの学習プロセスを例に挙げ、機械学習の特性について簡単に解説していきます。

## 2. 画像認識を行なう人工知能の開発

人工知能の開発と一口に言っても様々なアプローチがあります。人工知能の目的と目的達成の手法、使われる技術によって開発方法は全く異なっており、全ての人工知能が同じプロセスで作られているわけではありません。今回、人工知能の応用技術の中でも社会に大きなインパクトを与えた「画像認識システム」における画像の学習プロセスについて紹介します。

機械学習には大きく分けて人間が学習データに正しい情報を付与する「教師あり学習」と正しい情報を付けない「教師なし学習」があります。ここでは学習効率に優れた教師あり学習について解説します。この場合、学習プロセスは「データの収集」「データへのラベル付け」「ラベル付きデータでの学習」「性能評価と改善」の4つです。性能評価で満足が行く結果が出ればこのプロセスは終了しますが、出なければアルゴリズムや学習方法に改善を加え、必要に応じて新たにデータ収集・ラベル付け・学習・評価を繰り返します。次に、各プロセスで何が行われているのかを細かく見ていきましょう。



### <データの収集>

まずは学習させたい対象の画像を収集します。人間の顔であれば人間の顔を、犬の画像であれば犬の画像を集めることになります。ここで収集される画像の枚数は時に数千万に及び、手軽に集められるものではありません。さらに、同じ人間の画像であっても角度・背景・光によって見え方が異なるように、全く同じ物体の画像であっても、バリエーションに富んでいることが望ましいのです。

当然ながら「どこからそんなに沢山の画像を集めてくるか」という点が問題になります。一昔前であれば、この段階で画像が集められず人工知能の性能は頭打ちになっていたところですが、今ならインターネットがあります。キーワードで検索すれば大量の画像がヒットする上に、動画を使えば1つの動画から同じ対象の角度の違う画像を複数枚入手できるため、「人間や動物の画像」であればさほど学習用データの収集には困らなくなりました。

### <ラベル付け>

ところが、こうして得られたデータはそのまま学習に使うことはできません。データを集める際に人間が一枚一枚選んでいた場合はともかく、通常はより多くのデータを一度に収集できるようにするためにデータ収集用のプログラムを作り、そのプログラムにデータを集めさせます。そのプログラムはあくまで「特定のキーワードで得られた画像を全て収集する」もしくは「特定のキーワードで検索した動画の中からランダムに画像を抽出する」などのタスクを行なうだけのものであることが多く、全く関連性の無い画像も多数混ざっています。

そのため、収集した画像は人間が目で見分けて分類していかなければなりません。人間の顔を認識させたいのであれば、人間の顔が含まれている画像に「人間の顔」というラベルを付け、犬を識別したいのであれば犬が含まれている画像に「犬」というラベルを付けていきます。この作業は人間の顔や犬を識別出来る人間なら誰でも出来る極めて単純な作業ですが、それを1人で何千何万と行なうため、ここでも効率化のためにプログラムを組むことが多いです。

例えば、犬を識別するために、画像と一緒に「犬である」「犬ではない」と書かれたボタンが表示され、クリックすると次が表示されるようなプログラムを作れば、人間が行う作業は画像を見てマウスやキーボードを押すだけになります。単純作業でも、最低限の作業で済むプロセスを作ることによって時間と労力は最小限に抑えられます。

### <データでの学習>

退屈なラベル付けの作業が終わると今度はそれを使って人工知能に画像を識別するための学習を施します。この際、ラベルの付け方によって成果は大きく変わってきます。例えば、「犬である」「犬ではない」というシンプルなラベルしか貼っていないのであれば、人工知能が学習するのは犬か犬ではないかだけで、犬の種類までは識別しません。もし、「犬であるかどうか」に加えて「犬の種類」も加えれば犬の種類も学習できますし、犬ではないものに「人」や「猫」を加えれば、そのデータを使って人や猫の識別について学習させることもできるでしょう。

そして、人工知能の学習時間は画像の枚数・アルゴリズム・コンピューターの性能にもよりますが、学習プロセス全体から見てそれほど多くの時間はかかりません。場合によっては一瞬で終わります。人工知能の開発プロセスにおいて「学習」は重要なプロセスであることに変わりありませんが、手間と時間がかかるのはそれ以外のプロセスなのです。

### <性能評価と改善>

学習が終わった後に行なうのは性能評価です。ここでは、学習用のデータとは別に作っておいたテスト用のデータを使って検証します。人工知能は学習を経て「犬だと思った画像」「犬ではないと思った画像」を出力します。こうして出力されたデータを人間がチェックし、その精度などを評価します。

ここで良い結果が出ても、悪い結果が出ても、評価と改善のプロセスは重要です。以前と比べて良くなった原因、悪くなった原因を推察し、アルゴリズムにさらなる改良を加えていきます。もし機械学習のアルゴリズムではなく学習データに問題があった場合は、次に収集する画像データの集め方やラベルの付け方についても検討していかなければなりません。場合によっては、画像をゼロから集め直すこともありますし、ラベルを最初から付け直すこともあります。

ここまでで画像認識システムの学習プロセスはとりあえず終了です。しかし前述の通り、必要に応じて何度も何度も繰り返されます。作り始めたばかりのシステムであれば、繰り返す回数は1回や2回ではなく、場合によっては何十回と繰り返すこととなります。そのたびに学習データを集め直し、ラベルを付けるわけですので、これは途方もない労力を伴うプロセスであることが分かります。

### 3. 医療用人工知能で同じことができるのか

前述の画像認識システムのプロセスは、画像・文字・音声などの認識タスクを行なう人工知能の多くで共通しています。この学習プロセスをきちんと回すことができれば、一定水準の認識タスクを実行できる人工知能が完成するでしょう。

しかし、一見単純そうに見える学習プロセスにも大きな問題があります。学習に必要なデータが膨大なのです。データはインターネットから適当に拾ってくるだけでは不十分で、ラベル付けなどの作業によって学習に使えるデータに変えていかなければなりません。前述の画像認識システムのように「犬」を識別するだけの作業であればそれほど難しくはないでしょう。犬が何なんなのか認識してラベルを付けるだけであれば小中学生だって構いませんし、日本以外にもいくらでも依頼できる相手が存在します。

その一方で、認識すべき情報の「答え」が分かる人間が少ない場合には非常に厄介です。例えば、肺のレントゲン写真を見てその「答え」となる疾患を当てることができる人間は医師など医療の専門家だけに限られます。必然的にラベル付けの作業は経験豊富な医師に依頼せざるを得ませんが、こうした単純作業のために医師を確保し、医師に見合った高額な報酬を支払うことは容易ではありません。

問題は人手や費用が集めにくいだけに留まらず、医師であっても疾患があるかないかの判別は犬がいるかいないかという判別に比べて遥かに難解で、画像によってははっきりと断定できない「疑いがある」レベルに留まるようなケースも多いのです。開発者と作業者である医師間での基準に関する議論も欠かせないため、ラベルの判定基準ですら簡単には決められません。

医療用人工知能に限らず、人工知能の学習は学習に用いる情報の認識・理解に高度な専門知識が必要なケースになるほど難しくなります。しかも、医療用人工知能の場合は学習に使うデータが個人情報に関わるケースも少なくないため、インターネットで何でも集まる時代であるにもかかわらず、データを集めることにも苦労します。そして、苦労して収集したデータにラベル付けをしようにも頼める医師がいなければ雇うお金もないのです。

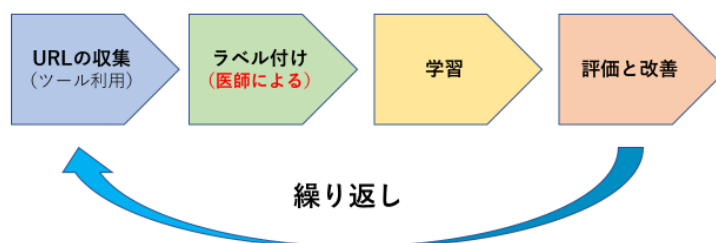
医療用人工知能の開発が画像認識の人工知能のように行かない理由はわかりました。では、実際に医療用人工知能のケースでどのような作業が行われるのかについて説明していきましょう。

### 4. 文献のリファレンスを表示する医療用人工知能の開発

医療用人工知能にも様々な種類が存在し、中でも注目を集めているのが診断支援を行なう人工知能です。今までは医師にしか識別することができなかった様々な疾患を、画像を含む各種検査データや症状から自動で判別してくれるようになりました。これだけでも十分優れているのですが、もしこうした診断支援システムが疾患に関連する文献のリファレンスを上げてくれたらどれほど医師の仕事が楽になるのでしょうか。

そこで、診断支援システムに文献のリファレンスを表示する機能を追加するとします。そのためには疾患別のリファレンスリストが必要となりますが、次々に増えていく膨大なリストを人力で整理するのは難しいでしょう。そこで文献を確認し、自動で整理してくれる分類用の人工知能を作ることになります。

この場合、膨大な医学文献を適切に分類するためのルールを人間の手で入力していくことは現実的ではないため、機械学習を用いて人工知能に自ら学ばせるのが現実的です。その際の学習プロセスは前述の画像認識システムの学習プロセスと殆ど変わりません。ただ、医療用人工知能であるため、各プロセスで必要となる作業が微妙に異なってきます。



※学習プロセス自体は画像の学習と変わらない

### <データの収集>

まず、文献データの収集を行います。画像認識システムの学習と同様にこれも自動で行いません。ただし、検索時のキーワードは疾患名であったり医学用語であったり、より専門性の高いものとなるため、ヒットする件数が少ないこともあれば無関係なものも少なくありません。また、今回はインターネットを使った文献の収集であるため、実際に収集するのはウェブページのURLとなります。URLにはWikipediaから公的機関の医療サイトまで含まれており、情報の形態は多岐に渡ります。画像認識システムの学習データのように一つの動画から複数の画像ファイルを入手することもできなければ、ヒットした画像ファイルそのまま集めれば良いというわけでもないので収集用のツールの開発も簡単ではありません。

### <ラベル付け>

そうしてデータを収集した後は、URLのリンク先のウェブサイトを確認し、医療用の文献としてリファレンスリストに掲載する価値があるかをチェックすることになります。この時点では、キーワードでヒットしたURLが集められているだけであり、疾患の名前が入っているだけで無意味なURLもあれば、疾患とは全く無関係のURLも存在します。

分類用の人工知能には、まずこうした無価値な文献を可能な限り取り除いて貰う必要があります。価値のある文献かどうかの確認は専門知識を持った人間にしかできないため、この作業は医師に依頼します。ただ、医師に依頼するからといって実質的な作業内容が複雑になるわけではありません。タスクとしては極めて単純で、URLリストに含まれている文献を1つ1つ確認し、価値があるか無いかのラベルをひたすら付けていくだけです。

### <データでの学習>

ラベル付きデータが完成したら、そのデータを使って分類用の人工知能が価値のあるデータだけを出力するように学習させます。この際、価値のあるデータだけではなく「不要」というラベルが付いたデータも同時に入力することが重要です。インターネット上に存在する文献には無数の種類があり、価値ある文献に似た無価値な文献も存在します。こうしたものも可能な限り正しく分類し、不要なデータを正確に「不要である」と分類できるようにならなければなりません。

### <性能評価と改善>

学習完了後、改めて文献を分類させ評価します。価値あるデータが出てくれば良いのですが、そう簡単にはいきません。価値のあるデータとして主力されたものの中に含まれる無価値なデータがなぜ含まれたのかを分析し、アルゴリズムの改善やデータの収集・ラベル付けの改善に生かして行かなければなりません。この際、ラベル付けには「価値がある」「価値がない」だけではなく、「なぜ価値があるのか」「なぜ価値がないのか」についてのラベルもあることが望ましいでしょう。

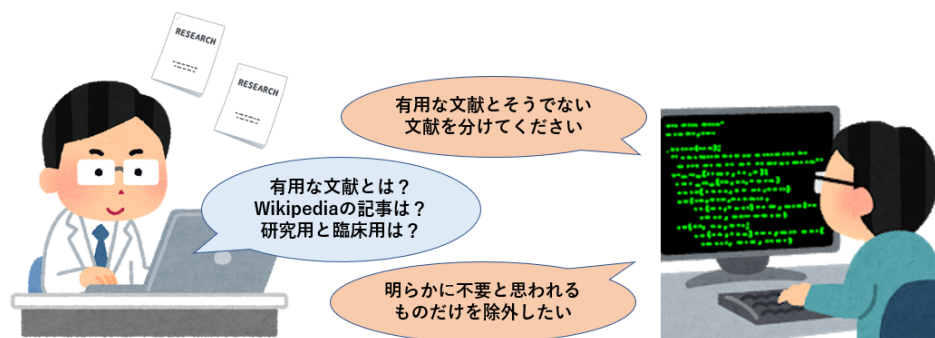
分類プログラム自体は「価値がある」「価値がない」という分類はしませんが、人間が評価・改善を行うためにより詳細なラベルが付いていることが望ましいからです。もちろん、この評価・改善プロセスで新たにラベルが追加されることもあればラベルが修正されることもあり、収集・ラベル付け・学習・評価を繰り返すことはより大きな意味を持ってきます。

## 5. 医師による医学文献のチェック作業 – ガイドラインの策定 –

ここまでの解説で、内容に微妙な違いはあれども、大まかなプロセスは画像認識システムの学習と医療用人工知能でさほど変わらないことは分かって頂けたと思います。ただ、小さな違いが大きな問題を生むことがあります。それは「ラベル付け」の工程です。分類用人工知能に出力してもらいたいのは「医師がリファレンスとして読む文献として価値のあるもの」であり、その答えを知っているのは医師だけです。医学的な知識がない人間には頼りません。そのため、医師に文献が使えるか使えないかのラベル付け作業をしてもらうこととなります。

もう少し具体的に掘り下げてみましょう。まず、医師は大量のURLリストを渡され「有用な文献とそうでない文献を分けてください」と依頼されます。ところがすぐに、このケースで言う「有用な文献って何だ」という話になります。つまり、判定基準に関する情報が不足しているのです。すぐにラベル付けに関するガイドラインの策定が必要になります。画像認識システムの学習のように「犬である」「犬ではない」という話であれば簡単ですが、医療に関わるデータの分類はそう簡単な話ではありません。

例えば、「Wikipediaなどのサイトでは、医学的に有用な情報はあるものの信頼できるとは言えない文献がある。これは価値があるといえるのか?」「研究領域では価値があるものの、臨床現場では価値のない論文はどうなるのか?」などの疑問が生まれます。こうした疑問に対し、人工知能の開発者は「明らかに不要と思われるものだけを除外したい。そのため、用途によっては価値がある文献は価値があると判定して欲しい」という回答を返します。



最初にURLを数件チェックしただけでもこれぐらいの質疑応答が行われます。こうしたやり取りを経てようやく暫定的なガイドラインが作られるのです。当然ながら、素人に「研究領域で価値のあるもの」と「臨床現場で価値のあるもの」の区別はつきません。今回のケースでは両方共価値があるとされたましたが、臨床現場だけで利用するものであれば区別する必要があったかもしれません。

## 6. 医師による医学文献のチェック作業 – ラベル付け作業 –

暫定的なガイドラインが完成してようやく実作業に入ります。ただ、膨大なURLを1つ1つ開いてチェックするのではどれだけ時間がかかるかわかりません。そのため、画像認識システムの学習時と同様に、チェック作業を効率化するツールが作られます。今回のケースでは、URLをいちいち開くのは手間なので、リンク先のキャプチャー画像と一緒にラベル選択のボタンが表示されるツールが作成されました。これによって医師はいちいちURLを開く必要がなくなり、少なくともキャプチャー画像のみで判断できるレベルのデータであれば簡単に分類できるようになりました。



## 単純作業が続く

しかし、ツールを用いたとしても単純作業がひたすら続くことには変わりありません。医療関係者向けの公的機関のサイトで、内容的にも明らかに価値があるだろうと見ただけで判断できるケースもあれば、リンク先に飛び、文献をしっかり読まなければならないケースもあるため、集中力の必要な作業が続きます。さらに、その過程で先に作成した暫定的なガイドラインでは対応できない問題にも遭遇します。

例えば、「疾患をまとめた便利なリストが含まれているケース」「医学的に価値があるが、そもそも動物に関する疾患であるケース」「医学的な文献だけど情報不足なケース」「英語で書かれていない文献」などです。これらの文献は状況によっては価値を持つ文献であり、一概に無価値と分類するのは避けたいとこです。そこで、人工知能の開発者はラベル付けのガイドラインを改定し、「リスト」「ミスマッチ」「情報不足」「非英語」「その他」などのラベルを追加しました。

それに合わせてツールを修正し、合わせてツールの使い勝手についても医師のフィードバックを受け、ツールについてもより効率的に作業できるように改善していきます。こうしたやり取りを繰り返しながらガイドラインを確定することで、ようやく安定して作業が進むようになります。

## 7. 医療用人工知能の完成に一步近づいた

分類用の人工知能はこうして分けられた文献を使って「価値がある文献」が何なのか学習していきます。医学生なら、分類された文献を何件か読めば「価値がある」とされる基準を理解できるかもしれませんが、人工知能ではそうはいきません。何千何万もの文献を読みながら、少しずつ学習していきます。また、何千何万もの価値があると判定された文献の中に大量に実は「価値のない文献」が混ざっていると学習が進まないため、ラベル付けの質は重要です。単純作業だからと十分な知識のない人間に頼むと、せっかく集めたデータが無駄になってしまいます。人工知能の性能向上において、このラベル付けの作業は非常に重要な作業と言えるでしょう。

ラベル付けを終え、データを使って学習させ、評価と改善が終われば医療用人工知能が完成するかというとそうは行きません。評価の内容次第では新たに文献を集めて同じ工程を繰り返す必要があります。そうして何とか分類用の人工知能が価値ある文献を見つけ出してくれるようになったとしても、これは診断支援システムに搭載する追加機能であるため、実際にどのように表示するのか、表示する順番はどう選定するのか、関連のある文献をどう絞り込むのかなど、問題は山積しています。

診断支援システムが鑑別と合わせて文献を表示し、それが医師の業務に役立つようなレベルに達するまでの道のりはまだまだ遠いです。しかし、医療用人工知能が現場で使えるようになり、患者に治療に役立つようになるためには、この長い道のりを少しずつ進んでいかなければなりません。

## 8. まとめ



機械学習を用いた人工知能の開発において、学習用データは必要不可欠です。しかし、学習用データの作成過程で人間の手が必要になり、それが膨大な単純作業であることはあまり認識されていません。また、学習に用いるデータによってデータの加工作業を行う人間に求められるスキルが違い、それによってこの単純作業の担い手が限られ、人工知能の開発に大きな影響を与えることもあまり知られていません。

残念ながら、医療用人工知能の学習では医師によるデータ加工が必要であるにも関わらずこの作業の重要性が正しく理解されておらず、協力者を得ることが難しくなっています。データの加工作業自体は非常に単調ではあるものの、こうした地道な作業が医療用人工知能の実用化には必要不可欠です。本記事を通じて、医療用人工知能の開発には医師の協力が非常に重要である点をご理解頂ければ幸いです。



[お問い合わせ](#)

[プライバシーポリシー](#)

Copyright © hosted by Medical Crisis WG, WIDE Project, Japan.

## 医療用人工知能を作る

[イベント報告・論文紹介](#)

[医療用人工知能と政策](#)

[医療用人工知能の作り方 / 研究開発の中心となるデータ作成の実例](#)

[医療用人工知能研究に求められるデータ作成作業の海外委託](#)

# 医療用人工知能研究に求められるデータ生成・監査タスクを海外委託する

この記事は12分で読めます

## サマリー

医療用人工知能の学習データの作成には時間とコストがかかります。そこで、インターネットを活用して軽作業を委託する「クラウドソーシング」を通じて海外の医療従事者に作業を委託する試みを行いました。今回の試みでは、利用者の多いクラウドソーシングサイトである「UpWork」を利用し、スリランカ在住の医師に自動的に生成した文献データの医学的な監査作業を委託しました。UpWorkでは委託管理のシステムが確立されていたため、実作業において、応募・作業・決済はスムーズに進みました。また、作業履歴も遠隔で確認することができ、結果的に研究利用するうえで十分な品質のデータが完成しています。最終的に、国内医師を雇用するケースに比べて5分の1程度のコストで監査作業を行うことができました。合わせて行った作業者へヒアリングを行ったところ、作業側にとっても大きなストレスなく作業を行えることが確認できています。クラウドソーシングを利用した作業委託は応用範囲が広いと、今後は医療用人工知能の研究開発への活用が期待できそうです。

## 海外に作業を委託する理由

近年、人工知能の研究開発が活発になっています。その中でも目覚ましい成果を上げているのが機械学習によって能力を向上させていくタイプの人工知能です。この種の人工知能では多かれ少なかれデータが必要になるため、データをどのように集め、学習させ、検証するかが人工知能研究の鍵になってきます。この人工知能研究に必要なデータの作成には、人間が関与せざるをえない部分が少なからず存在し、殆どの場合は膨大な単純作業を必要とします。これは人手と時間がかかる作業で、大きな課題となっていました。

これは医療用人工知能研究でも変わりません。プログラムが出力したデータが医学的に正しいか、有用であるかはその領域に関する知識と経験の豊富な人間が判断するしかなく、信頼のおける医療従事者が不可欠となります。しかし、国内の医療従事者の人件費は高く、一般的な人工知能研究に比べてデータ生成にかかるコストが高額になってしまうのです。この状況は、日本の医療用人工知能研究の研究コストを増加させることとなり、人工知能の研究開発が進まない一因となっています。また、問題は開発コストだけではありません。医師は精神的な集中を要求される膨大な単純作業を好まないため、予算があったとしても協力者の確保は簡単なことではありません。

もし、海外の医療従事者を効率的にリクルートし、データ生成業務を委託することが出来れば、この人件費の問題と研究協力者の確保の問題が解決する可能性が出てきます。とりわけ、データ作成の多くは単純作業であることから、コミュニケーションにハードルがある海外医療従事者であっても必要事項の伝達は容易であり、データ生成業務の委託も現実的な選択肢として考えられるようになります。加えて、クラウドソーシングサービスが発達したことで、最近では海外の単純作業者を探すことも容易になっています。大規模なものになると世界数十カ国から一千万人以上のワーカーが集まり、総額一千億円以上の取引がなされる巨大な労働市場が形成されています。言語を問わなければ有望な労働者をすぐに見つけられるでしょう。具体的には、作業ジャンルを問わないクラウドサービスとして「UpWork」「Freelancer.com」が存在し、特定の作業に特化したサービスとしては「Amazon Mechanical Turk」などが知られています。

このクラウドソーシングを利用することで、委託者は海外の賃金の安い作業者を簡単に確保でき、作業者は自国の賃金よりも高い報酬で作業ができることも多いため、多くの医師やエンジニアが海外の作業を受注しています。そこで、こうしたサービスを利用することで、安価に医療用人工知能研究のデータ生成ができないか検証を試みました。

## 委託内容と作業手順

今回クラウドサービスで委託する業務として取り上げるのは「医療に関するウェブページの内容検証作業」です。ここで検証するウェブページは、医療関係のウェブサイトの中から有用と考えられるものをアルゴリズムが自動的に選んだ結果です。具体的な検証作業としては、ウェブサイトの名称やURLなどが含まれている一覧ファイルに対し、作業者は各URLの内容を確認し、それが特定の疾患に関する医学的に有用なウェブサイトかどうかを判断するというものです。

66687	Partial epilepsy with pericentral spikes	omim_607221_0_omim_607221	0	74769	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20183">http://www.ncbi.nlm.nih.gov/pubmed/20183</a>	OK	abstract
6218	Familial primary cryofibrinogenemia	omim_123540_0_omim_123540	0	34473	20141006	google	<a href="http://en.wikipedia.org/wiki/Cryofibrinogenemia">http://en.wikipedia.org/wiki/Cryofibrinogenemia</a>	OK	wikipedia
127	Achard syndrome	omim_100700_0_omim_100700	0	36468	20141006	google	<a href="http://www.righdiagnosis.com/ai/achard_sy">http://www.righdiagnosis.com/ai/achard_sy</a>	OK	textbook
7775	Dyslexia	omim_127700_2_omim_127700	2	74426	20141006	google	<a href="http://en.wikipedia.org/wiki/Reading_disability">http://en.wikipedia.org/wiki/Reading_disability</a>	OK	wikipedia
37398	Enterocolitis	omim_226150_0_omim_226150	0	26845	20160501	google	<a href="https://www.nlm.nih.gov/medlineplus/ency/0">https://www.nlm.nih.gov/medlineplus/ency/0</a>	OK	textbook
26999	Seborrheic keratosis	omim_182000_0_omim_182000	0	32977	20141006	google	<a href="http://www.mayoclinic.org/diseases-conditions">http://www.mayoclinic.org/diseases-conditions</a>	OK	textbook
9000	Spontaneous thrombosis with factor V excess	omim_134400_1_omim_134400	1	34619	20141006	google	<a href="http://www.update.com/contents/consent/">http://www.update.com/contents/consent/</a>	OK	textbook
31477	Total intestinal aganglionosis	omim_202550_0_omim_202550	0	53839	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19939">http://www.ncbi.nlm.nih.gov/pubmed/19939</a>	OK	textbook
1135	Internal anal sphincter myopathy	omim_105565_0_omim_105565	0	76291	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/89637">http://www.ncbi.nlm.nih.gov/pubmed/89637</a>	OK	abstract
17984	Multiple exostoses with spastic tetraparesis	omim_158345_0_omim_158345	0	76368	20160504	bing	<a href="http://www.ncbi.nlm.nih.gov/pubmed/15012">http://www.ncbi.nlm.nih.gov/pubmed/15012</a>	OK	abstract
75031	Megarbane-lalkh syndrome	omim_612785_1_omim_612785	1	75213	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17407">http://www.ncbi.nlm.nih.gov/pubmed/17407</a>	OK	abstract
45046	Ophthalmoplegic neuromuscular disorder	omim_258470_0_omim_258470	0	74823	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/16093">http://www.ncbi.nlm.nih.gov/pubmed/16093</a>	OK	abstract
61582	Juvenile recurrent parotitis	omim_603588_0_omim_603588	0	72178	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17407">http://www.ncbi.nlm.nih.gov/pubmed/17407</a>	OK	abstract
60242	Pseudoacromegaly with severe insulin resis	omim_602511_0_omim_602511	0	140586	20141006	google	<a href="http://en.wikipedia.org/wiki/Acromegaly">http://en.wikipedia.org/wiki/Acromegaly</a>	OK	wikipedia
42732	Progressive lymphoid system deterioration	omim_247630_0_omim_247630	0	133500	20141006	google	<a href="http://en.wikipedia.org/wiki/Central_nervous">http://en.wikipedia.org/wiki/Central_nervous</a>	OK	wikipedia
58925	Sebacous nevus syndrome and Hemimeg	omim_601359_0_omim_601359	0	74486	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/76254">http://www.ncbi.nlm.nih.gov/pubmed/76254</a>	OK	abstract
73479	Hunter-MacDonald syndrome	omim_611962_0_omim_611962	0	37333	20141006	google	<a href="http://rare Diseases info.nih.gov/gard/27514">http://rare Diseases info.nih.gov/gard/27514</a>	OK	unknown
40888	Male hypogonadism	omim_241100_0_omim_241100	0	54933	20141006	google	<a href="http://www.update.com/contents/testes/">http://www.update.com/contents/testes/</a>	OK	textbook
55054	Paine syndrome	omim_311400_0_omim_311400	0	210081	20141006	google	<a href="http://en.wikipedia.org/wiki/Payne_Stewart">http://en.wikipedia.org/wiki/Payne_Stewart</a>	OK	wikipedia
15106	Atopic IgE responsiveness	omim_147050_0_omim_147050	0	36871	20141006	google	<a href="http://www.wikibooks.org/wiki/Handbook_of_">http://www.wikibooks.org/wiki/Handbook_of_</a>	OK	wikipedia
20174	Voluntary nystagmus	omim_164170_0_omim_164170	0	23439	20141006	google	<a href="http://www.righdiagnosis.com/medical/vol">http://www.righdiagnosis.com/medical/vol</a>	OK	textbook
60680	Apraxia of eyelid opening	omim_603119_0_omim_603119	0	45651	20141006	google	<a href="http://eyewiki.aao.org/Diagnosis">http://eyewiki.aao.org/Diagnosis</a>	OK	wikipedia
15639	Kaloid formation	omim_146100_0_omim_146100	0	73421	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/94945">http://www.ncbi.nlm.nih.gov/pubmed/94945</a>	OK	abstract
65298	Paraphernal arterial occlusive disease	omim_606787_0_omim_606787	0	104095	20141006	google	<a href="http://www.aefp.org/aefp/2004/03/16/525.htm">http://www.aefp.org/aefp/2004/03/16/525.htm</a>	OK	unknown
75143	Glaucosel syndrome	omim_612917_1_omim_612917	1	46853	20141006	google	<a href="http://en.wikibooks.org/wiki/Handbook_of_">http://en.wikibooks.org/wiki/Handbook_of_</a>	OK	wikipedia
52286	Calvarial hyperostosis	omim_302030_0_omim_302030	0	25791	20141006	google	<a href="http://www.righdiagnosis.com/medical/calva">http://www.righdiagnosis.com/medical/calva</a>	OK	textbook
18454	Hartung type myoclonic epilepsy	omim_159600_0_omim_159600	0	101785	20141006	google	<a href="http://en.wikipedia.org/wiki/Chorea_grandis">http://en.wikipedia.org/wiki/Chorea_grandis</a>	OK	wikipedia
11238	Giant platelet syndrome with thrombocytopen	omim_137560_0_omim_137560	0	76351	20141006	google	<a href="http://www.ncbi.nlm.nih.gov/pubmed/14601">http://www.ncbi.nlm.nih.gov/pubmed/14601</a>	OK	abstract
49010	Tardive dyskinesia	omim_272620_0_omim_272620	0	34991	20141006	google	<a href="http://www.nytimes.com/health/guides/disea">http://www.nytimes.com/health/guides/disea</a>	OK	textbook
19724	Nevus flammeus of nape of neck	omim_163100_0_omim_163100	0	66081	20141006	google	<a href="http://en.wikipedia.org/wiki/Nevus_flammeus">http://en.wikipedia.org/wiki/Nevus_flammeus</a>	OK	wikipedia
39879	Normal-pressure hydrocephalus	omim_236690_0_omim_236690	0	59966	20141006	google	<a href="http://www.update.com/contents/normal_">http://www.update.com/contents/normal_</a>	OK	textbook
46778	Infantile polymyoclonus	omim_263550_0_omim_263550	0	78851	20141006	google	<a href="http://en.wikipedia.org/wiki/Opsoclonus_my">http://en.wikipedia.org/wiki/Opsoclonus_my</a>	OK	wikipedia
47685	Retinal telangiectasia and hypogammaglob	omim_267900_0_omim_267900	0	56125	20141006	google	<a href="http://en.wikipedia.org/wiki/Macular_telangi">http://en.wikipedia.org/wiki/Macular_telangi</a>	OK	wikipedia

(ウェブサイトの名称・URL・疾患名などが含まれる一覧ファイル)

この一覧データに含まれるURLを「明らかに病気と無関係のデータ」や「有用性の低いデータ」として正しく判定することで、アルゴリズムの性能向上に役立てることができます。リストに含まれるウェブサイトの種類は多岐に渡り、公的機関の情報提供サイトや医学論文の概要を含むサイトもあればニュースの記事や製薬会社のホームページなども含まれています。こうした理由から、作業者は専門知識を有する医師が望ましいでしょう。

判定作業自体は極めてシンプルです。病名とURLが含まれるリストを渡された作業者はURLを1つ1つ確認しながら、有用なサイトには「Useful」とラベルを付けます。そうでないサイトには「Non medical」「Mismatch」「Poor」「List」「Other」「Garbage」「Non-English」のいずれかのラベルをつけます。この場合、「Useful」以外は不適切なサイトとして扱われることになります。

## 作業委託に使ったツール

この作業には時間がかかるため、Googleスプレッドシートのスクリプトを活用し、作業効率の向上が図られました。具体的には、リストの中で判定に必要となる「病名」「URL」「ウェブサイトのサムネイル」「ラベル(選択肢)」のみが画面表示され、作業者はサムネイルだけで判断できる場合はラベルをクリックして先に進み、判断が難しい場合にURLをクリックして詳細を確認します。マウスのみでの作業も可能となっており、100個のURLを20-30分程度で判定できるGoogleスプレッドシートのスクリプトを用意しました。

このツールを実現するために「判定に利用するウェブサイトのサムネイル作成」「URLリストの分割」「Googleスプレッドシートの作成」「スプレッドシートへのスクリプトの埋め込み」等の環境整備を予め行い、作業者が判定作業を行った後に「データのダウンロード」と「集計」を行って結果を出力する体制も整えておきました。

## 246925: McLeod neuroacanthocytosis syndrome

<http://rarediseases.info.nih.gov/gard/10731/mcleod-neuroacanthocytosis-syndrome/resources/>

(ラベルとサムネイルが表示されている判定画面)

## クラウドサービス上で作業者を募集する

(今回利用したクラウドサービス「UpWork」のホームページ)

今回、作業者を探すのに使ったクラウドサービスは「UpWork」です。oDeskとして2003年にサービスをスタートさせたUpWorkは、2017年の時点で全世界180カ国1400万人がユーザー登録しており、クラウドサービスとしては世界トップレベルの規模を誇ります。雇用者側が負担する手数料は支払額の2.75%、作業者側は報酬に応じて5%-20%の手数料が取られます（2018年1月時点）。高額の報酬になるほど手数料が減るシステムとなっているのも大きな特徴の一つです。

会員登録の段階ではクレジットカード情報などは不要で、必要な情報を入力したらすぐに仕事を作成して募集をかけることができます。一般募集はせずにワーカーを検索して直接依頼する事も可能です。また、ユーザー登録せずにワーカーを探すこともでき、利用の敷居を下げるための多くの工夫がこらされています。

す。ただし、ユーザー登録をしなければコンタクトを取ることはできず、プロジェクトを作成しての募集もできません。とりあえずワーカーを探してみて、見つかったら登録するという形で利用することできるでしょう。

## Freelancer Preferences

### Do you want freelancers to find and apply to your job?

- Freelancers using Upwork.com and public search engines can find this job.
- Only Upwork users can find this job.
- Only freelancers I have invited can find this job.

### Do you have freelancers in mind that you would like to invite?

Select freelancers

### Preferred Qualifications

Specify the qualifications you're looking for in a successful application. Freelancers may still apply if they do not meet your preferences, but they will be clearly notified that they are at a disadvantage.

Hide Qualifications

#### Freelancer Type

No preference

#### Job Success Score

90% Job Success & up

#### Rising Talent

Include Rising Talent

#### Hours Billed on Upwork

Any amount

#### Location

Any location

#### English Level (self-assigned)

Basic - Only communicates through written comm...

#### Group

No preference

### Screening Questions

(UpWorkの求人票情報作成画面。ワーカーのレーティングなどを絞りこめる)

求人票を作成するために入力可能な情報は様々ですが、業務内容の説明さえ記入すればとりあえずの求人はできます。ただし、時給制かプロジェクト単位の支払いになるかは最初の段階で決めておく必要があります。業務内容を説明する項目では必要なスキルなども記載することが一般的で、通常はこちらに記載する形で募集するのが良いでしょう。その他にも、UpWork内での実績やレーティングを指定して絞り込むことも可能で、UpWorkのシステムを理解してきたらより細かな条件で作業者を探すこともできるようになっています。また、応募の際にワーカーが答える質問事項なども用意できるため、スキルの有無や実績についての質問を用意しておけば、応募者の選定が容易になるはずです。

求人票の提示段階では支払い方法の登録などは不要です。しかし、ワーカーと実際に契約する際には支払いに必要な情報を登録する必要があります。方法はクレジットカードの他にPaypalがあり、クレジットカードの場合は承認後に返金される数ドル程度事前課金（デポジット）が行われます。クレジットカードの課金がすぐに確認できない場合には承認にやや時間がかかるため、早めに対応しておくことが望ましいでしょう。

## 応募者の確認と契約まで

VIEW JOB POST > INVITE FREELANCERS > REVIEW PROPOSALS (15) > HIRE (1)

ALL PROPOSALS (15) SHORTLISTED (0) MESSAGED (0) ARCHIVED (3)

Invite a coworker to help you hire ×  
 Once on your team, they can easily help you find, interview, and evaluate freelancers

Type email address and press enter

Also allow these coworkers to hire and pay with this account.

Search for applicants    Sort: Best Match

**Best Match** Psychiatry writer    
 \$40.00 / hr \$0 earned 📍 United States  
 What past project or job have you had that is most like this one? - I have worked as a clinic ... Cover letter - Dear Employer, I am a psychiatrist from India with 6 years of experience in different medi ...

**Best Match** Writer, editor, translator    
 \$30.00 / hr \$0 earned 📍 Poland  
 What past project or job have you had that is most like this one? - I am managing editor of t ... Cover letter - Hello, I am interested in taking this job. I have medical degree from Warsaw Medical Univ ...

**Best Match** Pharmacist    
 \$32.00 / hr \$8k+ earned 📍 Brazil  
 What past project or job have you had that is most like this one? - Analyzing medical informa ... Cover letter - I'm pharmacist with 09 years experience working as a consultant. I work in State Secretary ...

(応募者の一覧画面。UpWorkにおける平均時給や実績などが表示される)

今回の委託では、医療分野の専門知識を持つワーカーを募集しました。その結果、1週間で10名前後、3週間で18名(辞退3名)の応募者が来ました。新規案件はユーザーに参照されやすいため、募集直後が最も応募者が増えるようです。応募者に対してこの段階でメッセージのやり取りなどをすることが可能で、応募時に用意した質問への回答を深掘りし、実績の詳細情報や医療従事者としての証明書の提示を求めることもできるようになっています。

今回の応募者を見てみると、職種は医師・薬剤師・研究者・医学生・ライターなどと幅広く、医師だけでも5名いました。他にも実績の豊富な医療の専門家と思しきワーカーは6-7名おり、半数以上が作業に適した医療関係者だったこととなります。ただ、あくまでプロフィールや質問への回答から確認できるものだけであり、これだけでは信頼できるワーカーかどうかは分かりません。

また、募集してきたワーカーのUpWorkにおける平均時給は時給10ドルから65ドルと幅広いが、これは現在の職業の他に応募者が居住する地域にも大きな影響を受けているため、平均時給が高いからといってスキルのあるワーカーだとは限りません。むしろ、先進国の医師で時給が50ドル前後だった場合には却って注意が必要かもしれません。

## Writing and editing a medical case report

APRIL 2015

Technical Writing english-proofreading, editing, content-writing, medical-writing

Report

The client had a partly written case report on a patient with a lichenoid drug eruption. He needed help in modifying the document with regard to content editing, text formatting and proofreading. He also wanted the conclusion and the abstract of the case report written within a word limit.

I rearranged... [more](#)

[Case report - A patient with a lichenoid drug eruption.pdf \(252.54 KB\)](#)

(UpWork上で確認できるワーカーの実績に関する資料)

作業者に目星をつけたら、直接メッセージを送って実績を証明できるものを提示してもらうこともできますし、UpWork上でこれまでの仕事の実績やフィードバックを確認することも可能です。過去の仕事の内容や成果物が提示されているケースもあるため、過去に同様の仕事をこなし、高く評価されているのであれば判断材料となるでしょう。

しかし、あくまで「本物の医師」に作業を依頼したいという場合には「UpWork上での実績」だけでは不十分です。日本や米国（州ごと）など医師データベースを提供している国ならそれを使用できますが、途上国のようにそうしたシステムが無い場合やそもそも相手がデータベースに登録していない場合には医師免許や学位証明書で医師であることを確認する必要があります。また、過去の実績や所属病院のウェブサイトを用いた存在確認の方法もあります。所属病院の医師リストと病院のドメインが入ったメールで本人確認すれば良いため、こちらは手軽に行なえます。実際のスキルを確認する場合は、求人票により詳細な質問を加えるか、Skypeでのインタビューの実施、小規模のプロジェクトによる試雇などが考えられます。実施するプロジェクトに応じて、それに合わせた資格・スキルの確認を行うと良いでしょう。

## 採用した作業者と賃金について



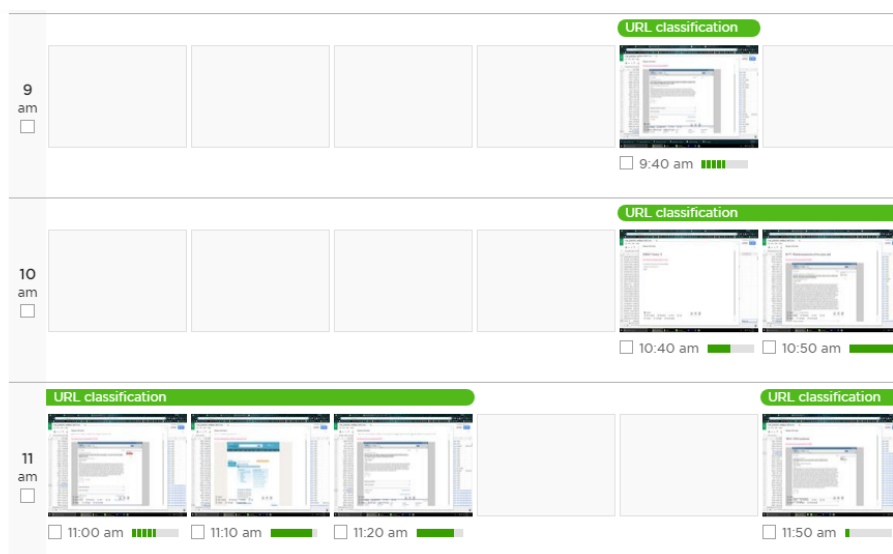
(委託したワーカーのプロフィール画面)

検討の結果、今回はスリランカのA医師に依頼することにしました。A氏はスリランカ最大の大学であるコロポ大学を卒業しており、国の研究機関での勤務経験もありました。UpWork上での実績も豊富で、論文の著者名でも名前が確認できていたため、A氏が専門的な知識を有するワーカーであることはほぼ間違いありませんでしたが、念のために学位証明書の提示を依頼すると快く証明書の写真を送ってくれました。

TIME & PAYMENTS	MESSAGES & FILES	WORK DIARY	TERMS & SETTINGS
<b>Rate and Limits</b>			
Rate	\$20.00 /hr		
Weekly Limit	30 hrs/week		
Manual Time	Manual time allowed		
Start Date	January 11, 2018		
<a href="#">View history of contract changes</a>			

(UpWorkの契約内容画面。時給や上限などが確認できる)

契約を結ぶ前に金額などの条件についてワーカーと交渉します。A氏のUpWorkにおける平均時給は20ドル。この時給額をそのまま提示したところあっさり契約が成立しました。この時に「Weekly Limit」として週あたりの最大労働時間を設定することができるため、時給制でも予算オーバーは避けられます。また、UpWorkではワーカーの活動ログを測定するアプリを提供しており、アプリを使うことで労働時間の計測もできるようになっていました。アプリによる活動ログの結果をワーカーと依頼者双方の合意のもとに修正できるようにしているため、適宜柔軟に対応することも可能です。



(UpWorkで確認できる活動ログ。スクリーンショットと活動頻度が表示される)

この活動ログアプリを用いると、作業中に定期的に撮影されたスクリーンショット画像やキーボードとマウスの使用頻度がUpWork上に保存されます。この画面だけでも記録としては十分ですが、より細かく確認したければ1分おきに記録されたキーボードとマウスの使用状況を見ることができます。

TIME (UTC)	KEYBOARD	MOUSE
2:41:44 pm	22	1
2:42:44 pm	15	11
2:43:44 pm	16	4
2:44:44 pm	17	10
2:45:44 pm	15	26
2:46:44 pm	22	8
2:47:44 pm	23	11
2:48:44 pm	13	6
2:49:44 pm	10	20
2:50:44 pm	37	21
2:51:44 pm	13	15
2:52:08 pm	6	1

(マウスとキーボードの使用ログ)

このように、かなり細かくワーカー側の労働状況を確認することができるため、少なくともPC上で作業するタスクを依頼する場合には、労働時間をサバ読んで請求されることはなさそうです。もちろん、スクリーンショットや活動ログを見て作業をしていないと判断できる場合には合意の上で依頼者側が作業時間を調整することも可能です。



<b>From:</b>		<b>INVOICE</b>	
<b>Bill to:</b>		<b>INVOICE #</b>	T164817331
		<b>DATE</b>	Jan 28, 2018
		<b>DUE DATE</b>	Jan 29, 2018
		<b>TOTAL AMOUNT</b>	\$210.00
		<b>TOTAL DUE</b>	\$210.00

DESCRIPTION / MEMO	AMOUNT
- 10:30 hrs @ \$20.00/hr - 01/22/2018 - 01/28/2018	210.00
<b>TOTAL AMOUNT:</b>	<b>\$210.00</b>

Invoice created via **Upwork**

(UpWorkより発行される請求書)

作業が実施されると、作業ログを元に週毎にUpWorkから請求が行われます。請求書は以上の様な形態となっており、この金額にUpWork手数料(2.75%)を足した金額が週毎にUpWork経由で請求されます。支払いは登録された決済方法で自動的に行われるため、活動ログのチェックはそれまでに済ませなければいけません。

さらに、契約途中での金額変更やボーナスの支払いも可能なシステムになっており、タスクの状況が変化すれば同意の上で時給を変更することができます。予想以上の成果を挙げてくれた場合には追加報酬を支払うことで次の仕事に繋げることもできるため、システム上の限られたやり取りであったにも関わらず、かなり柔軟に報酬の設定を行うことができました。

## 委託の結果、時間あたりの作業量とコスト感

TIME & PAYMENTS	MESSAGES & FILES	WORK DIARY	TERMS & SETTINGS	...					
<b>Last 24hrs</b> 3:50 hrs <small>Last worked 10 hours ago</small>	<b>This week</b> 9:10 hrs <small>Of 30hr limit</small>	<b>Last week</b> 7:10 hrs <small>\$143.33 Paid <span style="color: red;">Dispute</span></small>	<b>Since start</b> 35:30 hrs <small>\$526.67 Paid</small>						
<b>Timesheet this Week</b>									
MON 2/12	TUE 2/13	WED 2/14	THU 2/15	FRI 2/16	SAT 2/17	SUN 2/18	HOURS	RATE	AMOUNT
5:20	3:50	—					9:10	\$20.00/hr	\$183.33
<b>All Timesheets and Other Payments</b>								Last 30 days: \$526.67	
DATE	DESCRIPTION	CHARGE	INVOICE						
Feb 12	Invoice for 02/05/2018-02/11/2018 - 7:10 hrs @ \$20.00/hr	\$143.33							
Feb 5	Invoice for 01/29/2018-02/04/2018 - 4:30 hrs @ \$20.00/hr	\$90.00							
Jan 29	Invoice for 01/22/2018-01/28/2018 - 10:30 hrs @ \$20.00/hr	\$210.00							
Jan 22	Invoice for 01/15/2018-01/21/2018 - 1:50 hrs @ \$20.00/hr	\$36.67							
Jan 15	Invoice for 01/08/2018-01/14/2018 - 2:20 hrs @ \$20.00/hr	\$46.67							

(全体及び曜日毎の労働時間と週毎の支払額)

今回の委託における最終的な作業時間はおよそ35時間で、支払額は約700ドル程度となりました。完了した判定作業は約7000件。作業初期に発生した判定基準の変更による再判定や別件で行ったヒアリングなども作業時間に含まれているため、それらを除いた概算で「1時間あたり約300件」の作業効率となっています。

作業に慣れる前はもう少し時間がかかっていたものの、最終的にはこの作業速度で落ち着きました。判定作業の品質も問題ありませんでした。さらに、作業に際しては作業の効率化に資する意見なども出して頂き、環境そのものを改善しながら進めることができたため、非常に有益な業務委託となりました。日本の医師の時給はおよそ1万円前後とされているため、仮にこの作業を日本の医師に依頼した場合は5倍のコストがかかることになります。

日本人の感覚からすると、医師を時給20ドルで働かせることに不安を覚えるかもしれません。しかし、スリランカの医師の月収は600-700ドルの水準であり、これは日本の医師の20分の1です。それを踏まえれば20ドルという時給は高額な部類に入り、今回の支払いだけでも月収に匹敵する収入になります。

スリランカ以外にも同じ賃金レベルの国は数多く存在するため、最適な作業者を見つけられさえすれば、より大規模な判定作業を行なうとしても同じようなコスト感で作業ができるワーカーを見つけられる可能性は十分にあるでしょう。

## 作業者の反応

今回はコストパフォーマンスの極めて高い作業委託となり、依頼者としては満足のいく契約になったが、作業側はどう感じているのかは疑問です。こうした委託が作業者にとって大きな負担となる場合、より大規模且つ長期間に渡って作業をする場合に問題になります。そのため、A氏に今回の案件に関連して簡単なヒアリングを行ったところ、以下のような返答が得られました。

Q「今回の賃金には満足していますか？」

A「ええ、当時の平均時給でしたので十分に満足しています。」

Q「作業についてどのような印象を持ちましたか？またやりたいと思いましたか？」

A「今回の作業は私の医学的な知識をフルに使い真剣に情報を評価するもので、退屈どころか興味深いものでした。コンピューターだけで作業ができましたし、作業時間も自由です。気持ちよく作業できましたので、こういった仕事はまたやりたいですね。」

Q「クラウドサービスを通じて作業をしたい医師や医療関係者は多いと思いますか？」

A「途上国では給料も少ないので、副業をしている医師は多いです。私の回りにもUpWorkで作業をしたいと考えている医師は何人かいますが、競争が激しいので最初の仕事を見つけるに苦労していますね。」

Q「UpWork以外でこうした作業を委託する作業者をどうやって見つければ良いと思いますか？」

A「同僚や知人を辿って探すことになるのではないのでしょうか。UpWork以外にも作業者を見つける方法はたくさんあると思いますが、先進国のクライアントにとってUpWorkは安価に雇用できる途上国のワーカーを探す魅力的な場になると思います。」

Q「UpWorkで仕事を見つける時にどのような問題がありますか？」

A「興味深いタスクが中々見つからないことですね。提示額が低すぎるケースも多いです。あと、一番の問題は他の応募者との競争が激しいことですね。応募しても中々採用されません。ただ、最近は定期的に招待されるようになったので特に困っていませんが。」

## まとめ

委託者としては国内水準の5分の1という想定以上に安い費用で作業を委託できました。作業側も満足の行く賃金で作業できたようで、双方にとってメリットのある作業委託になったといえます。海外の作業者とのコミュニケーションは殆どUpWorkのシステム上で行ったものの、契約・作業・支払いまでのプロセスは非常にスムーズに進みました。実作業における作業プロセスが確立していれば、人員を増やしたとしても特に大きな問題はなさそうです。

一方、UpWorkのシステムではUpWork上で作業した実績しか参照できないため、実績のある作業者に依頼が集中する傾向がみられます。実際、委託する側としても実績のないワーカーには依頼しにくいという事情があるため、結果として最初の1件までのハードルが高くなり、潜在的なワーカーが仕事を得られていない事もヒアリングから明らかとなりました。こうしたワーカーの中から十分なスキルのある人材を見つけ出すこともコストパフォーマンスを高めるためには重要になります。

プロフィールやメッセージのやり取りを通じて実績やスキルの確認は可能です。単にUpWork上の実績を確認するだけでなく、手間を惜しまずUpWork上で実績の少ないユーザーとメッセージのやりとりをしてスキルの確認をすれば、隠れた有望なワーカーを見つけることもできそうです。

そして、一旦信頼できるワーカーを見つけられれば、同僚を紹介して貰う形でワーカーを増やすこともできます。大規模な作業の前に小規模な作業を委託し手順を確立した上で、ワーカーを紹介してもらう形で作業を増やせば、より大きなプロジェクトにも応用できるのではないのでしょうか。

今回の作業はデータの検証作業でしたが、遠隔で作業が可能なシンプルなタスクであれば医療用人工知能のデータ検証以外のタスクにも応用はできます。もちろん、医療分野の他にも各種専門家にタスクを依頼するような作業には有用でしょう。特に人工知能研究ではデータ生成関連の単純作業が多いため、応用範囲は幅広いです。賃金についても、国内と海外の賃金格差の大きな職種ほどクラウドソーシングを使った海外業務委託のメリットは大きく、相対的に賃金の高い日本の研究プロジェクトであれば十分にメリットがあるでしょう。

また、今回の委託によって研究中のアルゴリズムの精度が概算で現状87%程度あることが確認できました。今回作成したデータを使って再度学習を進めることで、分類プログラムのさらなる性能向上が期待されます。今回の海外作業委託は成功だったと結論づけても良さそうです。このように、クラウドソーシングが人工知能の研究開発に大きなインパクトを有することがご理解頂けたら幸いです。



[お問い合わせ](#)

[プライバシーポリシー](#)

Copyright © hosted by Medical Crisis WG, WIDE Project, Japan.