

2018 年度第 6 回物学研究会レポート

「GHELIA が目指す人と AI の共生環境とは？」

清水 亮 氏

(ギリア株式会社 代表取締役社長)

2018 年 9 月 25 日



デザインにとって、AIとの共生はIoTと並んで大きな課題です。2015年の物学研究会で「プログラミングとデザインについて」のご講演をいただいた清水亮さんは現在、ギリア株式会社を共同で設立し、日々進化を遂げるAIと人間が共存できる環境の構築を目指して、さまざまなプロジェクトを立ち上げ、実践されています。

今回は、清水さんから新会社設立の背景、現在取り組んでいるプロジェクトを通して、GHELIAが目指す「人とAIの共生環境」とはどのようなものなのか、その実現のために何が必要なのかなど、ご講演いただきます。

以下、サマリーです。

「GHELIAが目指す人とAIの共生環境とは？」

清水 亮 氏

(ギリア株式会社 代表取締役社長)



01：清水 亮 氏

■コンピュータの歴史を振り返る

清水：こんばんは。GHELIA（ギリア）の清水です。「GHELIA」という社名は、物学研究会の会員でもあるデザイナーの後藤禎祐さんが考えてくれたもので、G＝グローバル、HE＝ヒューマン・エンハンスメント、LIA＝場所を表しています。

僕らが目指すミッションは、人とAIの共存関係を実現することです。現在のコンピュータは、その前身になるものができてから約1世紀が経っていますが、その原点は、人間を拡張してよりよいものにしていくことにあります。そして全地球的な動きがコンピュータリゼーションに向かい、OAやITと呼ばれる流れが生まれました。そして、その流れの延長線上にある道の途上にあるのがAIです。社名には、こうした背景が含まれています。

僕らの事業についてお話しする前に、コンピュータの歴史を簡単に振り返りたいと思います。もともとコンピュータは、1945年頃までは計算をする職業（計算手）の「人間」を表す言葉でした。それが、1955年頃になると計算をする「機械」を表す言葉に置き換わってい

ます。その当時のコンピュータは、1台で何十億円、何百億円。しかも、設置するには1部屋分、あるいはビル1棟分もの巨大なスペースが必要でしたので、大国家、大軍隊など超巨大な組織だけがもつ特別なもので、普通の人には必要ないと思われていました。まさに今のAIと同じような状況です。

そして1960年代に入ると、アイヴァン・サザランドが世界初のインタラクティブ・コンピュータ「SketchPad」を発明します。これは、操作をすると画面がリアルタイムに反応するというという画期的なものでした。つまり、初めてコンピュータに画面をつくったのがサザランドで、逆に言うとそれまでの20年間、コンピュータには画面がなかったわけです。

この発明によって、サザランドの元には優秀な弟子が集まるようになります。そのうちの一人がアラン・ケイで、彼はサザランドのアイデアをより推し進め「Dynabook」というパーソナルコンピューティングの概念を提唱します。この時点ではまだ、個人が1台のコンピュータを所有するなど正気の沙汰ではありませんでした。そして僕らは今、同じことをAIでやろうとしているわけです。

ほかにもサザランドの弟子には、Adobeをつくったジョン・ワーノックや、PIXARをつくったエドウィン・キャットマル、シリコングラフィックスをつくったジム・クラークなどがあります。クラークは、ジオメトリエンジンやグラフィックス・プロセッシング・ユニット（GPU）、レンダリングパイプラインを発明した人です。また、スティーブ・ジョブズとビル・ゲイツは、アラン・ケイのDynabookの影響を受けて、MacintoshとWindowsを生み出したと言われています。

このように、今われわれが知っているコンピューティングのすべては、サザランドの研究室から生まれています。また、サザランドは1968年の時点でヘッドマウント・ディスプレイの提案もしていて、実際にプロトタイプをつくってデモンストレーションをしているんですね。つまり、バーチャル・リアリティ（VR）やオーグメンテッド・リアリティ（AR）なども含め、今のコンピュータ業界というのは、サザランドがやってきたことをなぞっているにすぎないのです。

■生まれては消えてきたAIブーム

人工知能という言葉が初めて使われたのは1956年に行われた「ダートマス会議」です。招集されたのは、プログラミング言語を開発している人や、チェスゲームをつくっている人、コンパイラやプログラミング言語の研究をしている人、検索エンジンを研究している人たちでした。彼らに共通するのはコンピュータを計算以外の目的で使っていること。つまり、コンピュータが人間の知識を構築したり、理解して提案したりする、そういうことに使えるような研究をしている人たちを集めて、それを人工知能の研究と名付けたのです。これを機に、第一次AIブームが生まれます。

ところが、当時のコンピュータは人工知能を実現するにはメモリが少なすぎたこと、さらにブームのさなかにサザランドがコンピュータ・グラフィックスを発明したことも重なり、み

ながグラフィックのほうにお金をかけるようになってしまいます。こうして第一次 AI ブームは終わりました。

80年代になると第二次AIブームが起こりますが、同様にハードが貧弱であったことから、うまくいきませんでした。そうしたなかで、食い詰めた人工知能の研究者たちは、知識を体系的に集めるシステム、もしくは知識の体系を相互作用させるシステムとして、まったく別のものをつくろうということになりました。こうしてできたのが **Hyper Text Markup Language (HTML)** です。つまり現在の **WEB** は、もともと人工知能の研究から生まれたものなのです。

さらに時代は進んでも、人間と人工知能の間には大きな壁がありました。なかでも AI にとって極めて難しいのは、写真や絵を見てそれが何かを当てることです。たとえば、背景を取り除いて人物だけを正確に切り抜く作業は自動的にはできません。それは、機械にはどこまでが髪の毛で、どこからが背景なのか見分けられないからです。

たとえば、猫と犬の違いがわからない人はいませんが、犬と猫を知らない人に、その見た目の違いを言葉だけで説明して相手に理解させることはできませんよね。こうした問題を突破することがなかなかできなかったのです。

■第三次 AI ブームの到来とその理由

人工知能の世界には、説明によってつくる人工知能と機械学習という二つの潮流があります。現在の機械学習の主流はニューラルネットワークを使ったディープラーニングです。人間や動物の神経細胞は電気信号を刺激として受け取っているいろいろな判断をしています。この構造を使えば AI も同じように学習できるのではないかと、という仮説から生まれたもので、たとえば猫の写真と犬の写真的違いをたくさん見せることで、二つの違いを理解させるという学習方法です。

ただ、当時のコンピュータのメモリはよくて 1MB くらいだったので、昆虫並みの脳すらつくれませんでした。こうした状況にブレイクスルーを果たしたのがゲーム機です。3D ゲームには、超高速のメモリと超高速なプロセッサが無数に必要だったからです。特に、世界で最初に GPU を搭載したゲーム機であるプレステが成功したことによって、GPU を使ったら儲かる、面白いことができる、ということが世界に浸透していきました。

そうしたなか、GPU を使うと安く人工知能がつかれるということが偶然わかりました。これによって、メモリ不足から実現できずにいた問題が解決し、長い間ずっと溜め込んでいたフラストレーションが一気に爆発したことで、現在の第三次人工知能ブームが起きたのです。つまり、過去に試した手法にディープラーニングを組み合わせると爆発的に成果が出た。これが現在の人工知能ブームの正体です。

もう一つの理由は、サザランドの **SketchPad** 以来ずっと続いてきたコンピュータ・グラフィックスの進化が、2010年の iPad で止まってしまったことです。この8年間の **Apple** の新製

品は、時計になろうがメガネになろうが、サザランドのアイデアからなんの進歩もない。つまり、8年間停滞しているんです。コンピュータの歴史にとって8年停滞するというのは、かつて経験したことがないくらい長い停滞なんですね。それをごまかすために人工知能の技術を一部取り入れ始めた。それが Siri や Google アシスタントです。

ところが、これが極めて変なものになっている。スマホはグラフィカル・インターフェイスの機械なのに、Siri 呼び出すと急にテレタイプのような画面に変わってしまいますよね。なぜそんなことになっているかという、画面がある世界とない世界では、そもそも成り立ちが違うからです。それを今、強引に結びつけようとしたために、みんな方向性がわからなくなってしまう。こうした状況が背景にあると思ってください。

■GHELIA の「みんなの AI 事業」

このような世界において、僕らは人と AI との共生環境の実現を目指しています。これまでコンピュータの世界が当たり前のように進歩できたように、楽に生み出せるビジョンはもうありません。大げさに言えば新たにサザランド的なビジョンをつくりだし、それをオーバーテイクして普及させる。GHELIA は、それくらい大きな野心をもってつくられた会社です。設立は昨年6月。僕が個人的に所有する UEI という会社と、ソニーコンピュータサイエンス研究所、WiL の3社の出資によってスタートし、さらに先月からは、東京海上、みずほキャピタル、シグマクシスの3社の出資が加わりました。

われわれはこの事業を「みんなの AI 事業」と呼んでいて、予算規模とプロ向けか一般・子ども向けなのかによってカテゴリを分けています。一つは予算規模1億円~100億円の B to B 向けのヒューマン・エンハンスメント事業 (HE 事業部) で、銀行のシステムや保険業務のような社会実装を扱っています。いずれは業界ごとに横展開していくことを目標にしています。

もう一つは1万円~1億円未満のコンシューマーサービス事業 (CS 事業部) です。こちらは、ディーラーニングしやすいパソコンの販売や、AI のコンテンツを提供しながら、AI 市場の裾野を広げることを目的としています。その一つが、BS フジで放送している「ちんぷいぷいプログラミング」という子ども向け番組の制作です。これは、プログラミングと人工知能を3分で教えるという番組なのですが、YouTube でも公開しているので、一時停止したり、繰り返して見たりできるので、子どもの教材としても成り立っています。

また、人工知能の精度は、地道なデータ作成作業をどれだけできるかにかかっているので、新潟県にデータ作成専門の AIUEO という会社をつくっています。

■動きをリアルに再現する「DeepMimic」

今回はあまり数学的な話をしないようにとのことなので、最近1ヵ月かけて地球を一周して、人工知能の専門家を訪ねたときの話をいくつかご紹介します。

まずは、バンクーバーで開催されるコンピュータグラフィックの SIGGRAPH（シーグラフ）という学会の展覧会です。映画やゲームなど、なんでも扱う大規模な学会なのですが、今年はディープラーニングを使って生成した動きやかたちに関する発表が 3 分の 1 を占めていました。

なかでもおもしろかったのは「DeepMimic」です。たとえば、格闘技ゲームの場合、従来のモーションキャプチャでは動きが固定されてしまうのですが、リファレンスモーションだけを学ばせると、キックするターゲットの位置が変わっても起動を描き直すことができるようになる。しかも動きをつくった人の個性をなぞるようにつくっていて、それが人間型だけでなく、ライオンやティラノザウルスでもできるんです。

■今もっとも熱いスーパーレゾリューション

もうひとつは、NVIDIA が発表した「RTX」というリアルタイムレイトレーシングシステム。システムを ON にするとカラーブリーディングもきちんと起きて、複雑な場面でもリアルタイムにレイトレできるというものです。じつは実際にレイトレしているのは全ピクセルの 5 分の 1 だけで、残りは AI が埋め合わせをしています。要するに、低解像度のものを元の高解像度の映像に戻す、ということ AI に学習させているのです。これがスーパーレゾリューション（超解像技術）という、ディープラーニングのなかで今一番熱いジャンルになります。

スーパーレゾリューションを使った映像は、おそらく 8K で撮った映像と比べても、その違いを人が見抜くことはできないだろうと僕は思っています。すると、8K のように機器を全部買い直す必要もなくなるし、もっと言えば 8K のテレビで再生するときに 8K にオーバーサンプリングして映し出すことや、圧縮手法としても使うことができる。こうしたことが、スーパーレゾリューションという非常にシンプルなアルゴリズムでできるのです。

もう一つ、スーパーレゾリューションが面白いのは、どんなジャンルのものを学習したかによって、正確さが変わるんですね。たとえば自動車をたくさん見たことがなければディテールがわからないので、画素がつぶれたときに補正できない。また、イラストを生成する学習をした人工知能に写真を見せても、イラストのように修正してしまう。逆にそれによって、別の表現手段を得ることもできるわけです。

■最強のロボットサッカーチーム

次は、ロボットサッカーの強豪チームをもつ、テキサス大学オースティン校を訪ねました。これは全部ロボットの動きをディープラーニングや機械学習を使って一番いい蹴り方やチームワークを学ぶためのシステムです。まずは 3D シミュレーションでキックやパス回しを練習して、そこで学んだ動きを現実のロボットに転送してプレイする。そこで、実際の動きとの違いをもう一度シミュレーション側にフィードバックすることによって、よりリアルな動きをつくり出すという学習方法です。

■タンジブルからラディカル・アトムズへ

MITの石井裕先生に誘われてオーストリアのリンツに「アルス・エレクトロニカ」というメディアアートのイベントを見に行きました。僕が驚いたのは、科学の文脈で語られることの多い石井先生の研究が、ここではアートになっていることでした。

石井先生は、アラン・ケイがMITのメディアラボにスカウトした人物で、タンジブル・ユーザーインターフェース（TUI）の提唱者です。TUIは、直接手でデジタル情報に触って操作できるインターフェースのことで、石井先生の言葉で言うと、デジタルのビットはとてもシャイなので外に出てこない。だから、マウスやキーボードを使ってリモートコントロールして動かす世界なんだと。そんなシャイなビットを界面の外に出して、デジタルのものを直接手で触れるようにしてあげよう、という概念です。

石井先生がこのことを言い始めた90年代の初めの頃は、まったく理解されませんでした。現在当たり前のように使われているタッチスクリーン、ICチップ内蔵のマンションのキー、Suicaなども、もともとは触れたら動くというタンジブルの概念から始まった技術なんです。

タンジブル・インタフェースに続くものとして石井先生が研究しているのは、今お見せした映像にある「ラディカル・アトムズ」というビジョンです。今までは直接画面を触らなければいけなかったのが、画面の方から現実世界に出てきて、現実世界にある玉を動かす。フィジカルをラディカルに変えるという世界です。最終的には、素手だけで動かせることができ初めてラディカル・アトムズになると石井先生は考えているので、今はその前段階にあります。

今のところはまだ製品化できる見込みは立っていませんが、サザランドのSketchPadから始まってiPadで停滞していたコンピュータの歴史の先には、もしかしたらラディカル・アトムズが来るのかもしれません。

■テクノロジーとアートの融合

石井先生はよく「エステティックを大事にしておかない」「アートという文脈があることになぜ気づかないのか」ということをよくおっしゃっています。それで僕も、アートという言葉について考えるようになりました。

AIはArtificial Intelligenceのことですが、Artificialに似た言葉でArtifactという言葉があります。コンピュータ・グラフィックスの世界では、Artificialが人工的なものであるのに対して、Artifactはマスマスプロダクトではなく自然発生的にできたものという意味になります。

この二つの言葉に共通する「アート」についてオックスフォード英語辞典で調べると、人間のクリエイティブスキルとイマジネーションによってつくられた表現という意味のほか、「紳士」という意味もあることがわかりました。これは、紳士とは「人間らしさ」「理性をも

っている」という意味に解釈できるのではないかと僕は勝手に思っています。

先日、僕の尊敬する親友、猪子寿之君がやっているチームラボプラネッツに行って、衝撃を受けました。テクノロジーとは、アーティファクトもしくはアーティフィシャルなものなのだ実感したからです。これまでテクノロジーだけを求めていた僕は、じつは矮小で、その一つの側面だけで物事を見ていたら、足をすくわれるかもしれない。テクノロジーとアートは同じ視点を共有できるのではないかと思います。

もうひとり、僕の大好きな河口洋一郎先生の作品についてお話します。彼の作品は、改めて見るとわけがわかりません。でも、たくさんの作品を見ていると、彼がやりたいことを感じ取ることができるんです。たとえば、数学的に美しい精巧なりサーージュ曲線や、フラクタルを応用したもの、アルゴリズムミックなものであることがわかってくる。すると、まったく別の価値が出てくるのです。

こうした、わけのわからないものの中にうっすらと感じる哲学というものが、じつは人工知能とすごく似ているんです。たとえば以前、アメリカで育ったAIに「つけ麺」の写真を見せたら「カルボナーラ」と答えたんですよ。ぜんぜん違うけれども、なんだかわかる。そんな感覚です。同じように、河口先生の作品もなんだかよくわからないけど、同じ作者だとわかる。けれども、それを先ほどの犬と猫の違いの話と同じように言葉だけで説明することは、かなり難しいということです。

■未来を切り開くための「清水三力」

最後に、石井先生がおっしゃる「石井三力（出杭力、道程力、造山力）」にちなんで、「清水三力」というものをつくってみましたので、ご紹介します。

まずは「拡張力」。一つは、未来を予測するには過去の文脈を理解することが必要だという意味での拡張。もう一つは想像力を広げるという意味での拡張。サザランドの Sketchpad の想像力のすごさを考えると、逆にいかにこの後の人たちの想像力に独創性がなかったかということがわかります。

二つ目は「愚行力」。世の中で活躍している人がみんな言うのは、最初は誰にも理解されなかったということです。それは、人がやらないことをやるからであって、これが正しいんだと思って確信犯的にやる愚行には、必ず意味があると僕は思っています。そして、それを誰かに言われても、勝手に言っていればいい、と思えることが大事かなと。

三つ目は「独尊力」で、誰にも頼らなくてもやっていけるという根拠のない自信をもつこと。根拠のある自信は、根拠が崩れるとガラガラってなってしまうけれども、根拠がなければそもそも崩れるポイントがないので、ずっと続けていける。人の話は聞くとしても、結局判断するのは自分だと。

じつは、自分がそう思うからいいんだという感覚は、人工知能の認知にわりと近いんです。誤解を恐れずに言うと、今の人工知能は直感力を持っているんですよ。

以前、仕事で知り合った方から、連絡先とキーワード以外はメモを取るな、と言われたことがあります。俺にとって大事なことがお前にとって大事なことは限らないから、自分の心に残ったものだけが本物だということだったのですが、これは人工知能の教育法とほとんど同じなんです。

ただ、自分の中でおもしろそうだなと思ったことだけを、言葉でなく感覚として残すというのは、じつはとても難しいことなんです。常に神経を研ぎ澄ませて観察していないと大事なことを取りこぼすので。ですから、一瞬、一瞬、さぼらないで何が起きているか見極めることが大事なのだらうと思います。

以上、拡張力・愚行力・独尊力という清水三力の話を最後にして、今日は終わりにしたいと思います。ありがとうございました。

関：ありがとうございました。では、質疑応答に入りたいと思います。

Q&A

Q1 今やられていることのゴールは、いつぐらいを想定されていますか。また、それに向けてやりたいこと、やろうとしていることがあれば教えてください。

清水：50年くらいかかるんじゃないですかね。サザランドのビジョンが実現するのも、ビジョンにたどり着いてスタートしてさらに20年くらいかかっているので、そんなに簡単なことではありません。今は、登る前にどのルートから攻めるかを考えているぐらいの感じだと思ってくれればいいかなと思います。

Q2 Siriの画面が矛盾しているという話がありましたが、清水さんからみて正しいと思うグラフィック・インターフェースはどのようなものだと思いますか。

清水：そもそもSiriを本当のアシスタントのように機能させるのならば、画面や文字がなくても受け答えさえできればいいわけです。たとえば、明日歯医者に予約しといて頼んだとき、わかりましたと言ってそのまま予約すればいいのに、わざわざ画面を使って近くの歯医者が10件見つかりました、などと答える。そんなアシスタントは無能ですよ。

つまり、指示型の音声入力インターフェースが不完全だから、画面で補おうというさもしい発想が裏側にあるんです。「Hey Siri」で100%自分が呼ばれたことに気付けるぐらい洗練されていれば、ボタンを長押しして呼び出す機能などいりません。精度が悪いから、こういうエスティックが失われているインターフェースにせざるを得ないわけです。

——精度を上げるのは難しいのですか。

清水：真面目につくったらできますよ。それなのに、手を抜いているから腹が立つんです。アシスタントなら、なんでこっちから呼ばないと何もしないのかとか、外出するときに持って行くのを忘れたら、人間のアシスタントに「私、忘れられているんですけど」と連絡するとか、もっと言うと自分で追いかけて来るとか、それくらい踏み込んだことが夢としても語られていないし、ロードマップにあるようにも思えない。なんかしょぼいなというのが僕の印象です。

Q3 AIは直感力までもってきているというお話がありましたが、新しいクリエイティブを生み出すには、間違えることも大切ではないかと思うのですが、AIはそういったところまでもっていけるのでしょうか。

清水：よくある誤解なのですが、AIは基本的に間違えるものなんです。間違えたときに学習をして、間違いを減らす。たとえば先ほどの話にあったように、つけ麺をカルボナーラに間違えたら、これはつけ麺というものと教えればそれがわかるようになります。また、絵を描かせて良し悪しの判定を人間がすると、どんどん絵がうまくなっていくんです。このように、間違わないと学習できないのがAIのおもしろいところで、むしろ間違いこそがAIにとっては宝物なんです。

関：清水さん、みなさん、ありがとうございました。

以上

2018 年度 第 6 回物学研究会レポート

「GHELIA が目指す人と AI の共生環境とは？」

清水 亮 氏

(ギリア株式会社 代表取締役社長)

写真・図版提供

01 ; 物学研究会

編集=物学研究会事務局

文責=関 康子

- [物学研究会レポート] に記載の全てのブランド名および商品名、会社名は、各社・各所有者の登録商標または商標です。
- [物学研究会レポート] に収録されている全てのコンテンツの無断転載を禁じます。

(C)Copyright 1998~2018 BUTSUGAKU Research Institute.