

第5節 R&D 部門における生成 AI・AI エージェント活用に向けたデータ共有のポイント

(株) キャトルアイ・サイエンス 上島 豊

はじめに

現在の R&D 領域では、データ分析や管理は、極めて属人的な扱いである。客観的なデータ生成、分析が要求される理学、工学領域で、この属人性は大きな問題を孕んでいる。研究というものは創造的な活動であり、個人の才能、発想に起因する「なぜ、そう考えたか？」の部分に属人性が必要なことは当然だ。しかし、どのようにデータを生成し、どのように分析し、結論を導いたかは、属人的では問題で、客観的かつトレサブルであるべきである。実際、センサーや計算機的能力向上により、データの生産性が向上し、扱うべきデータが膨大になり、詳細記録の欠如、偶発的データ取り違い、主観的データ操作が発生する余地が増大し、データ生成、分析プロセスの信頼性が大きく揺らいでいる。また、機械学習・AI・生成 AI・AI エージェントと研究者の付き合い方に関しても今は属人的となっており、本来どのように研究に組み入れていくべきかが大きな課題になってきている。

実際、私は弊社を設立する前は研究機関にてコンピュータ、ネットワークの最先端技術を駆使し、自然科学、工学研究を約 10 年間行っていた。その中で R&D データが属人的に処理され、その管理状態がデータの信頼性及び有効活用性を大きく阻害し、共有化及びインフォマティクス分析、AI 化が進まないことを経験した。本節では、私自身の 10 年の R&D 経験と弊社の 20 年の R&D 支援実績から得た「R&D 部門における生成 AI・AI エージェント活用に向けたデータ共有のポイント」に関して、簡単に解説する。

1. R&D 部門において機械学習・AI に期待して良いこと、悪いこと

最近テレビやネットのニュースで機械学習・AI・生成 AI という言葉を聞かない日がないような状況になってきている。実際、R&D 部門で、「AI でこういう画期的成果が出た」、「AI を使うことで従来と比べ〇〇倍開発が早くなった」という話もよく見かける。その一方、「上層部から機械学習・AI・生成 AI などを使って、開発を高速化しろ！」と聞かれているのですが・・・という相談も私のところに頻繁に舞い込む状況もある。プレスリリースで見かけた会社からも・・・である。本項では、「機械学習・AI とは何か？」に切り込み、「機械学習・AI・生成 AI に期待して良いこと、悪いこと」を明らかにすることで、上記状況がなぜ起こっているのかの説明を行う。

生成 AI は、利用するシーンが少し違うので、後項で説明することにし、まず、本項では機械学習、AI に関して説明する。本節を読まれている方は、機械学習、AI を材料設計などで使うことを想定していると思うので、それを前提として話を進める。機械学習、AI 等は、たくさんの実験設定情報（以下、実験パラメータと呼ぶ）と実験結果情報（以下、試作物特性値と呼ぶ）を使って、予め機械学習、AI プログラム内部の設定値を決定したプログラムで、任意の実験パラメータを入力するとその実験パラメータで試作した試作物特性値を予測してくれるというものである。もちろん、プログラムの組み方により、その逆（望む試作物特性値を入力するとそれを実現する実験パラメータを予測する）も可能である。実際には、後者の方が実験者としてはありがたいのではないだろうか？ここまでで、気づいている人もいるかもしれないが、統計解析（重回帰分析など）だって、同じことを行うはずである。実際、このレベル（統計解析でも望む試作物特性値を入力するとそれを実現する実験パラメータを予測する）のことは、現場の実験者は十分理解していると思うが、経営層の上層部の人もそのことを理解して「機械学習・AI・生成 AI などを使って、開発を高速化しろ！」と言っているのかというと怪しい可能性が高い。そのような上層部の人には、雑誌のこの節部分に付箋をつけたものを机の上において、この節を読んでもらうのがまずは第一歩である。

前段落で機械学習、AI と統計解析、重回帰分析は同じことができると言った。当たり前だが全く同じことというわけではない。何が違うかというと、「予めプログラム内部の設定値を決定したプログラムを準備する」という部分の詳細レベルが少し違うのである。統計解析、重回帰分析はモデル式（ $Y = ax + b$ や $Y = ax_1 + bx_2 + c$ や $Y = ax_1^2 + bx_1 + cx_2^2 + dx_2 + e$ など）を仮定し、そのモデル式の設定値（ a, b, c, d, e など）を決定する。それに対し、

機械学習, AI はモデル式を想定せず, 機械学習, AI の内部にある膨大な設定パラメータを決定していくのである。「機械学習, AI はモデル式を想定せず」と書いたが, 実際には, 機械学習, AI には様々なモデルがあり, そのモデルの差により強い予測力の AI, 弱い予測力の AI などの差が生まれるのである (そういうものがなければ世の中の機械学習, AI は扱えるデータが同じであれば全て同じになってしまう!)。ただ, そのモデルが重回帰分析の ($Y = ax + b$ や $Y = ax_1 + bx_2 + c$ や $Y = ax_1^2 + bx_1 + cx_2^2 + dx_2 + e$) ように実験者が認識している変数 (x, x_1, x_2 など) に直結した形の意味が分かりやすい式にはなっていない。つまり, 人間が各設定値の意味を解釈できるというものをモデルと呼ぶというのであれば, 機械学習, AI の設定値はモデルとは呼べないようなものになっている。それが, 機械学習, AI はモデルなしで予測できるプログラムであると言われる所以である。

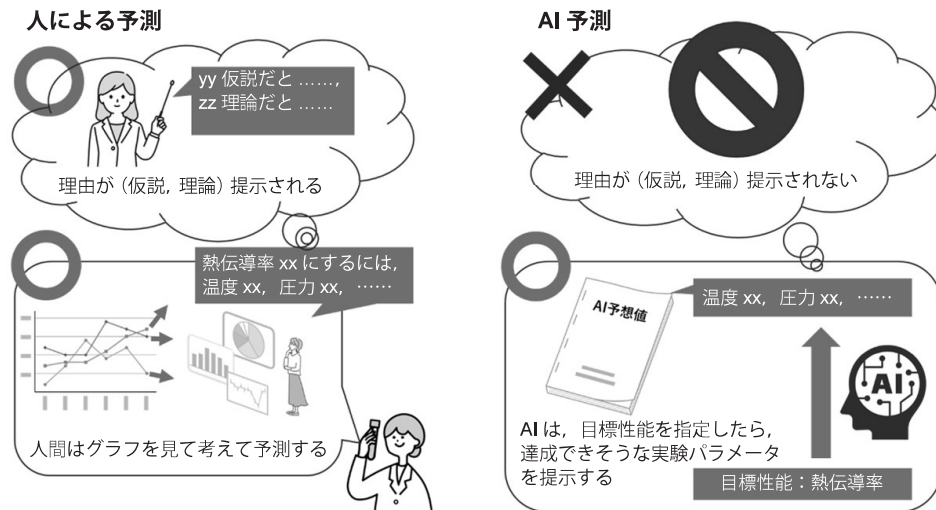


図1 機械学習, AI は予測値の理由を提示するのが苦手

統計解析, 重回帰分析はモデルありきで予測を行い, 機械学習, AI はモデルなしで予測を行う。統計解析, 重回帰分析はその本格的な解析の前に何らかの手段で適切であろうモデルを見つけ出さないといけないわけである。どちらかというと「適切であろうモデルを見つけ出さないといけない」というよりも, 「モデルを仮定して, 分析を行い, モデルの妥当性を確認する」ことが統計解析, 重回帰分析に含まれると考えたほうが良い。ただ, 全く仮説モデルが思いつかないという場合は, 統計解析, 重回帰分析はある意味お手上げな訳である。そう聞くと「モデルなしで予測できる機械学習, AIの方が優秀で, 統計解析, 重回帰分析は不要ではないか?」と思われるかもしれない。実際, これが会社の上層部の人がよく取り憑かれている機械学習, AI 万能神話に繋がってしまっているのである。当然のことながら, 世の中, そんなに甘い話はない。機械学習, AI はモデルなしで予測できるが, 同じ予測力を得ようとすると統計解析, 重回帰分析に比べ, より多くの, 実際には相当多くのデータを必要とする。つまり, 統計解析, 重回帰分析と機械学習, AI のどちらが有効かは, 適切な仮説モデル提起のし易さと大量データの取得のし易さのトレードオフで決まることになるのである。実際, マーケティングなどの社会科学的なテーマでは, 物理や化学と違い, 確実な基礎法則が明確でないため適切な仮説モデル提起は非常に難しい。一方, 年齢, 性別, 出身などの分類情報において偏りのない多くのデータを集めること自体は比較的容易という側面が強い。そのため, 社会科学的なテーマに対しては, 従来, 統計解析, 重回帰分析で悪戦苦闘していたものが, 機械学習, AI で一気に改善される例がたくさん出てきている。これも会社の上層部の人の機械学習, AI 万能神話の源泉かもしれない。しかし, 材料開発の R&D は実験パラメータ空間が広大で, 実験パラメータ空間に偏りのないデータを集める (= 実験をする) ことは, 膨大な実験数を行うことを意味し, 現実的ではなく, 社会科学的なテーマとは様相が大きく異なるのである。

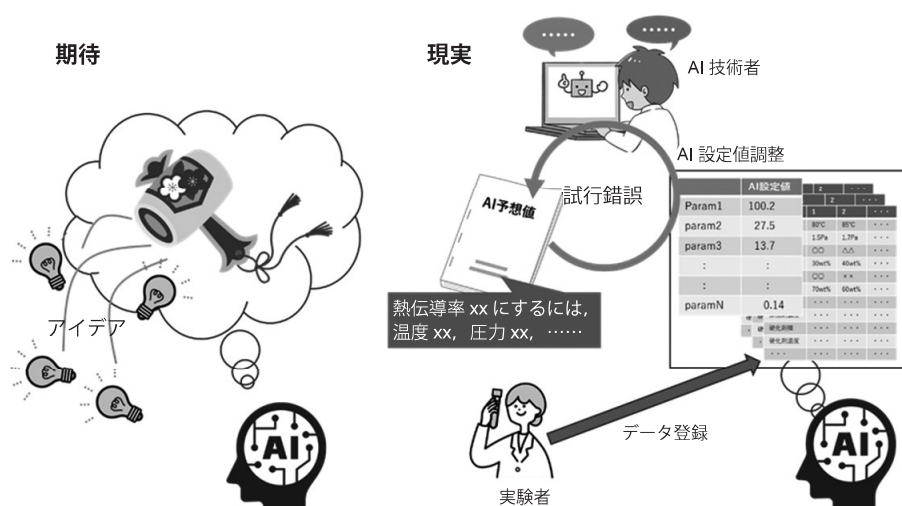


図2 機械学習，AIは打ち出の小槌ではない

一般的に実際の現象が複雑になればなるほどモデルに必要な自由度（係数の数）が膨大になり、適切な仮説モデル提起は難しくなる。物理化学的現象を意識して、係数が100個を超えるモデル式を書けるかを考えてみれば、適切な仮説モデル提起が如何に現実的でないか分かるはずである。そういう領域で、それなりの高い予測精度が必要であれば、モデルフリーの機械学習，AIに頼らざるを得なくなる。この最後の言明が独り歩きして、機械学習，AI万能神話が生まれたと考えられる。しかしながら、上記言明は適切な仮説モデル提起が難しいと言っただけで、機械学習，AIであれば高い予測精度が実現できるとは一切言っていない。頼らざるを得ないだけで、頼った結果、高い予測精度が実現できない可能性もあるのである。「高い予測精度」という言葉を何度か使っているが、どの程度が必要だろうか？当然、予測精度が高い方が良いに決まっているが、そうすると必要データ数も増えてくる（もちろん、データ数が増えれば無際限に精度が上がるというわけではないが・・・）。実は、最低限必要な予測精度というのが決まっている。それは、研究者が今まで行っていた研究開発における予測精度である。これを下回るような予測精度だと、機械学習，AIを使った研究開発は高速化されるのではなく、遅延されることにしかない。

もう少し比較しやすくするため予測精度というのを数値にしてみる。例えば、予測精度0.1という値は、10回予測すればほぼ当たるような精度だと定義しておく。つまり、必要実験数＝1／予測精度である。既に蓄積されているデータだけで、機械学習，AIの予測精度が従来型の研究者の経験，ノウハウによる予測精度を超えることができる場合は、何の気兼ねもなく機械学習，AIの活用を進めていくことができる。一方、既に蓄積されているデータだけで、機械学習，AIの予測精度が従来型の研究者の経験，ノウハウによる予測精度を超えることができない場合は、機械学習，AIのために蓄積されたデータを増やさなければならない。つまり、機械学習，AIのために追加で実験をしなければならないことになる。機械学習，AIの予測精度を上げるためというだけの実験は、はっきり言って、研究の遅延行為である。もちろん、それでも多少今の研究が遅延してでも従来型の研究者の経験，ノウハウによる予測精度を大きく超える機械学習，AIが実現できるのであれば、その遅延は取り返せるだろう。しかし、定性的な話ではいつまで追加実験（＝研究の遅延行為）をするのが決められなく、ずるずると遅延行為が続いてしまう。追加実験は、定量的な議論をして、撤退ラインを予め決めて行うべきものである。それこそ、経営者の計数感覚をもって、判断をすべきことということである。

また、「従来型の研究者の経験，ノウハウによる予測精度を大きく超える機械学習，AIが実現できた」という報告があっても諸手を挙げて、良いというわけではない。機械学習，AIの予測精度を上げる責任を負った人は、何としても予測精度を上げようとする。予測精度を上げるのに手っ取り早い方法がある。それは、適用範囲を狭くすることである。使う材料、その材料の濃度範囲、プロセス条件の変動可能なパラメータとその範囲などを絞ると、少ないデータでも予測精度は上がってくる。

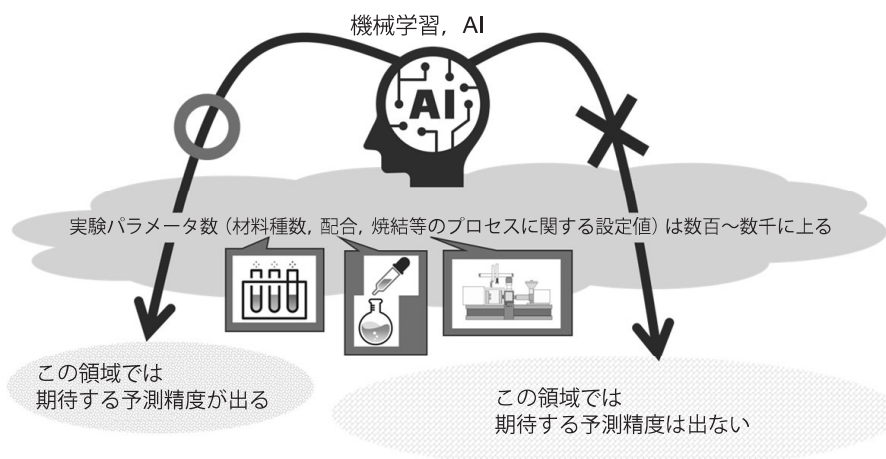


図3 機械学習, AI は期待する予測精度の出る領域の明確化が難しい

しかし、そういう風に作られた機械学習, AI は、すぐに使い物にならなくなる。その適用範囲内で目的の特性を満たす解がなければ、予測精度もある意味、役立たずだからだ。もちろん、「その範囲では望む特性の材料はない」ということがわかることは、無駄実験を減らせるので、良いことではあるが、狭いパラメータ範囲で大量の実験をしていれば、たぶん、それは従来型の研究者の経験、ノウハウでも気づくはずである。本当は、あまり実験をしていない領域に望む特性の材料があるのか？あるのであれば、どのような実験パラメータで望む特性が達成されるのかを知りたいはずである。しかし、そういう（データが少ない）領域では、機械学習, AI の予測精度も低く、ともすると従来型の研究者の経験、ノウハウの予測精度を下回ることも多いのである。実は、プレスリリースをした会社でも現場で困った状態が続いているのは、こういうからくりなのである。

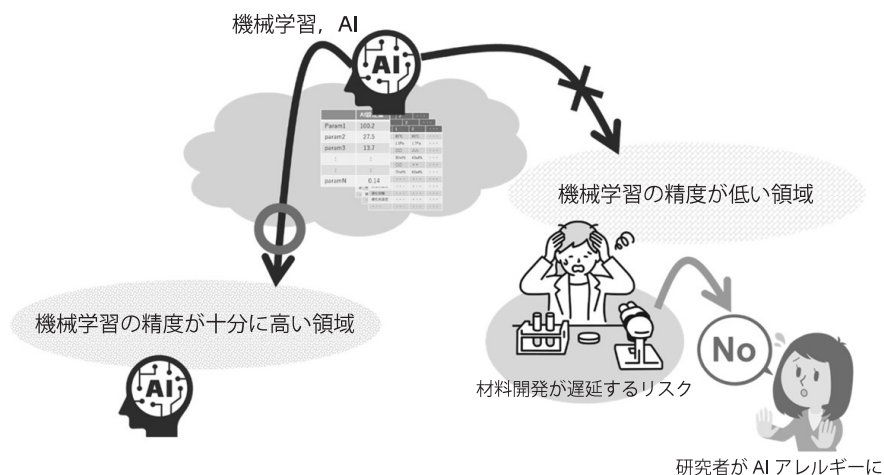


図4 機械学習, AI の予測精度の低い領域を明確化しておかないと研究者が AI アレルギーに

ここで、「R&D 部門において機械学習・AI に期待して良いこと、悪いこと」として、少しまとめておく。適切な仮説モデル提起が難しい複雑な挙動をする実験では機械学習・AI は、現研究を高速化できる可能性があり、期待はしてもよい。ただし、研究者の経験、ノウハウをベースとした従来型の研究の予測精度を超えることができるかどうかは、実際に試行してみないとわからない。また、試行するときでも適切な適用範囲での予測精度を確保しないと、すぐに使い物にならなく（その範囲では望む特性の材料はない）なる。使い物にならなくなると追加実験をして、適用範囲の拡張が必要になり、結局、研究の遅延が発生する。この研究の遅延が機械学習・AI を取り組まない時より、大きくなるようなことは許容できないはずである。したがって、機械学習・AI を進めるためには、以下のような作業が事前に必要となる。

- 1) 従来型の研究の予測精度を定量的に明確化する。
- 2) すぐに使い物にならないような適用範囲がどの程度かを過去の実績から明確化する。
- 3) 目標予測精度を決め、追加実験などを許す数を制限し、撤退ラインを決める。

ここまでの議論では、追加実験、適用範囲、予測精度の観点でしか話をしていないが、現実問題では、機械学習・AI に使うためのデータの収集、整理（データクレンジング）やモデル選定、設定値調整、精度評価など他にも多くの労力がかかってくる。それら全てにかかる労力、時間を考慮した上でも従来型の研究より効率的なものを目指さないと意味がないということは、覚えておいてほしい。また、「従来型の研究か？機械学習・AI か？のような二者択一ではなく、相補的に助け合えるような・・・」ということが良く提起されることがある。その考え方自体は間違っていないが、上記 1), 2), 3) とその他の労力、時間に関して考えることを止めて、なんとなく定性的に進めてしまっている状況を良く見かけるが、それは間違っている。どう相補的に助け合えるような形になるのかを決め、その上で難易度はさらに上がることになるが上記 1), 2), 3) と同じレベルの定量化に落とし込む必要があることを最後に付け加えておく。

2. R&D 部門において生成 AI に期待して良いこと、悪いこと

前項では、「R&D 部門において機械学習・AI に期待して良いこと、悪いこと」に関して、説明を行った。本項では、「R&D 部門において生成 AI に期待して良いこと、悪いこと」と題し、R&D 部門において生成 AI の活用可能性に関して論じる。

「生成 AI」という単語で一括りにしているが、「生成 AI」は、質問応答や画像生成、グラフ作成、プログラム作成など様々なことができる。例えば、最近のニュースでは、プログラム作成を生成 AI を使ってどの程度改善がみられるかを調査したところ、「実際には 20%遅くなったが、プログラムしている人は 20%ぐらい早くなったように感じている」という報告があった。また、別のニュースでは「50%以上遅くなったが、使い続けていると改善していくという傾向も見えてきている」という報告もあった。どんな便利なものでも今までと違ったものを使うと多かれ少なかれ、同じような傾向になるだろう。ただ、前項で述べたように定量的に確認できる仕組みを作っておかないと、本当のところ早くなっているのか？かえって遅くなっているのか？は判断できないということは、このニュースでも重要な教訓として取り上げられていることに注目すべきである。R&D 部門では、実際の生産をしている部門に比べそういう定量化が少ない部署であるがゆえに気を付ける必要がある。画像生成、グラフ作成、プログラム作成などは、本書籍の読者の主たる興味ではないと思うため（というか、もう少し紙面のページ数を割いて、技術よりの文章を書かないとあまり役に立たないため）、本項では、生成 AI の質問応答に限って話をしていく。

生成 AI の質問応答は私もよく使っているが、従来の Web 検索に比べると非常に強力で、知りたいことに対して、わかりやすい回答を返してくれる。たぶん、前項に書いていることも質問の仕方を工夫すれば、ほぼ同じような内容が回答として得られるはずである。「質問の仕方を工夫すれば」といったが、極端な話、質問の仕方によっては、全く逆と思われるような回答が返されることもあるのである。良く良く読んでみてやっと「これ、おかしいんじゃないの？」と気づくレベルのしっかりした回答を返すことも多いのである。生成 AI の基本ロジックは、「単語の繋がりにおいて、確率の高いものを繋げていく」というものなので、文章としては、確率の高い（＝よく見る）慣れ親しんだ文章になってしまう。つまり、表面を読んでいるだけだと何の疑問も湧かない良い文章なので、その中身とロジックをしっかり確認しないとイケないということである。また、単語の繋がりやロジックというのもそれなりに模倣されてしまうので、結構複雑なロジックの記述があったからといって安心をしてはいけないのである。

次に厄介なのが、どうも生成 AI は、質問者に媚びへつらう傾向があるという点である。会話のパートナーとしてということを考えると、ツンケンされているのは会話も弾まなく、次第に使われなくなるので、ソフトウェアの仕様としては、「質問者に媚びへつらう傾向」を持たせるのは理解できる。しかし、単なる会話ではなく、仕事として正確な情報が欲しい時は「質問者に媚びへつらう傾向」は、困った仕様である。中立的な質問だとなかなか思った回答が

返ってこないで、明示的に「○○でありますか？」という質問をするとあっている例を探し出して、それを基に答えを出す形になる傾向が高い。これら为了避免するためには、回答する側（生成 AI）に質問者に対して批判的に回答するような役割を先に明示しておいたり、生成 AI の答えにわざと批判的な指摘をして、再回答をさせて見るなどをする必要がある。また、その回答の基となっている元資料の提示を要求し、元資料を確認することも絶対に必要である。生成 AI は、いい奴で悪気のない嘘を平気でつく親友のようなもので、その嘘を見抜き騙されないように付き合う力を身につけないと、その友情は脆くも崩れ去るのである。

R&D 部門において質問したいことは、「どのような添加剤を加えれば特性改善が進むか？」などではないだろうか？ R&D 部門では、ネット上に転がっている当たり前の情報というより、自社の報告書や論文を基に質問の回答が欲しいという類のものであると思われる。前項に書いているような文章は、一般論で、具体的な研究の内容に立ち入っていないのでオープンな生成 AI でも作成できるが、「自社の報告書や論文を基にした回答」は、オープンな生成 AI では適切な回答にならないはずである。もし、オープンな生成 AI からの回答が適切っぽい回答になっている場合は、それこそ、その回答の基となっている元資料を確認すべきである。前段落でも述べたが「適切な回答」に見えるような回答を作っているだけの可能性が高い。結局、R&D 部門において質問応答用に生成 AI を導入するためには、自社の報告書や論文を生成 AI の学習データに含める必要がある。OpenAI だとか、Microsoft がクラウド上でサービスを展開しているので、自社の報告書や論文を学習データとしてとり込んだ生成 AI を作ることも可能になっている。当然、オープンな生成 AI として使われるわけではないが、社内の情報をクラウドに送り込むので、会社によっては許可されない場合もあると思う。新鋭気風の IT 企業などは、オープンソースとして公開されている学習モデル（BERT など）や、LLaMA のようにライセンス条件付きで配布されている大規模言語モデルを入手し、自社内サーバで自社の報告書や論文を学習データとして取り込んだ自社生成 AI サービスを立ち上げている。後者の方法では、生成 AI をゼロから開発する力はなくともプログラムベースでチューニング等ができ、サービス運用を行える技術スタッフがいることが必須となる。

ここまでは、質問応答用生成 AI を使う場合の注意点と R&D 部門で利用できるようにするために、どのようなことを行わないといけないかを話した。ここからは「生成 AI は何に期待して良いか、悪いか」に関して議論していく。まず、よくある質問に「報告書や論文を学習データとして取り込んだ生成 AI があれば、前項で議論した機械学習や AI なんていらんではないか？」というものがある。本気でそう思った人は、もう一度前項を読み返してほしい。少なくとも以下の 3 つは、生成 AI を使おうとも行わないといけななことである。

- 1) 従来型の研究の予測精度を定量的に明確化する。
- 2) すぐに使い物にならないような適用範囲を過去の実績から明確化する。
- 3) 目標予測精度を決め、追加実験などを許す数を制限し、撤退ラインを決める。

次に大きく異なるのは、報告書や論文に実験データ全てが記載されているかである。社内の報告書や論文を確認してもらえばわかると思うが、全実験数の 1 割も記載されていないはずである。また、報告書に記載されている実験であっても全ての実験設定情報（各材料の濃度、配合や攪拌、乾燥などのプロセス条件）や特性値計測情報（計測器名、計測モード、計測温度などの計測条件）が漏れなく記載されているだろうか？そんなことはないはずである。

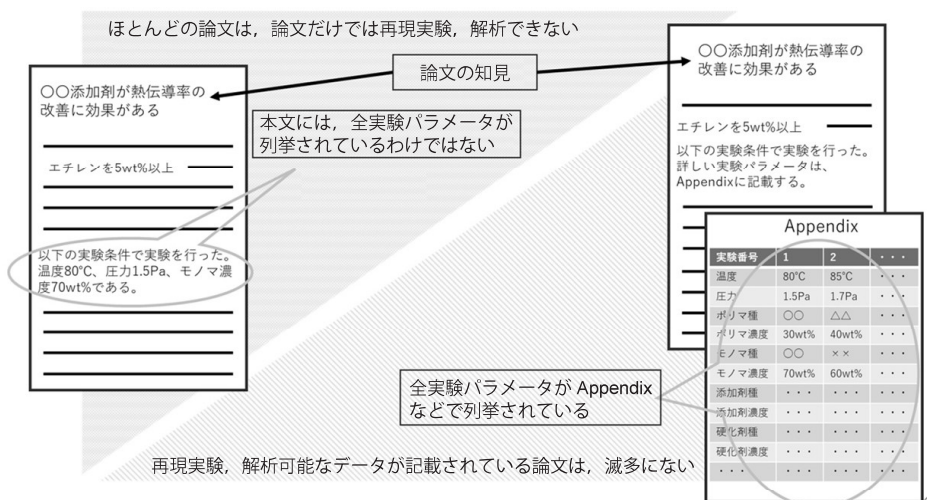


図5 報告書には実験や解析を再現するために必要な情報が書かれていない

報告書や論文というものは、いくつかの実験から何らかの法則性を導き出したという報告をすることが主であり、細かな実験情報を記録して提出することが目的ではないはずである（もちろん、本来は報告書や論文でもその報告内容の信頼性を担保するために再現実験できるレベルの情報を書いておくべきなのだが、そうなっていないのが実情である）。また、同じ材料名や同じプロセス条件の項目名を様々な報告書で違う表現で記述していたりするため、細かな実験情報や特性値計測情報の記録があったとしても報告書全体のデータを表として間違えなく纏め直すことは、生成AIであってもほぼ不可能である。実は、「間違えなく」ではなく、多少の間違いを許すと現在の生成AIではある程度できるかもしれないが、前項で書いているような予測精度を議論するような話では、「多少の間違え」は、一切許容できない。やはり、質問応答用生成AIは、機械学習、AIの代わりにはなりえないのである。

報告書及び論文＋生成AIが実験詳細データ＋機械学習、AIの代わりにはなりえないことは、ここまでの話で理解できたと思う。それであれば、報告書及び論文＋生成AIに全く意味がないのかというと、当然そんなことはない。前述した通り、論文や報告書は、複数の実験から何かしらの一般的な知見や法則性を導き出し、それを提示することが目的であり、一般的な知見や法則性という強力な実験指針を得られるので、非常に価値の高いものである。例えば、ニュートンの運動方程式やシュレディンガー方程式を使うと、世の中の多くの現象を説明、予測できることを考えれば、一般的な知見や法則性がいかに強力で価値が高いかはわかっていただけたと思う。ただ、論文や報告書を参考にする場合には、注意すべき点もある。例えば、「〇〇添加剤は、熱伝導率の改善に効果がある」と報告書に書かれていても、実際、あらゆるケースに対して、「〇〇添加剤が熱伝導率の改善に効果がある」ことを確認したわけではないはずである。この知見は、特定の実験パラメータ領域での知見のはずであり、今、自分が適用しようとしている領域でその知見、法則が確認されているとは限らない。報告書のタチが悪いのは、どの領域ならその知見が成立するかを証拠データとともに明示的に書かれていなかったり、拡大解釈（実際に確認をしていないがそう信じている）して書かれていることが多かったりする点である。これは、社内報告書だけでなく、学術論文でもその傾向がある。また、これは論文の査読レフリーでも十分に指摘しきれていない部分であり、学術論文でもその点を十分注意して、参考にする必要がある。

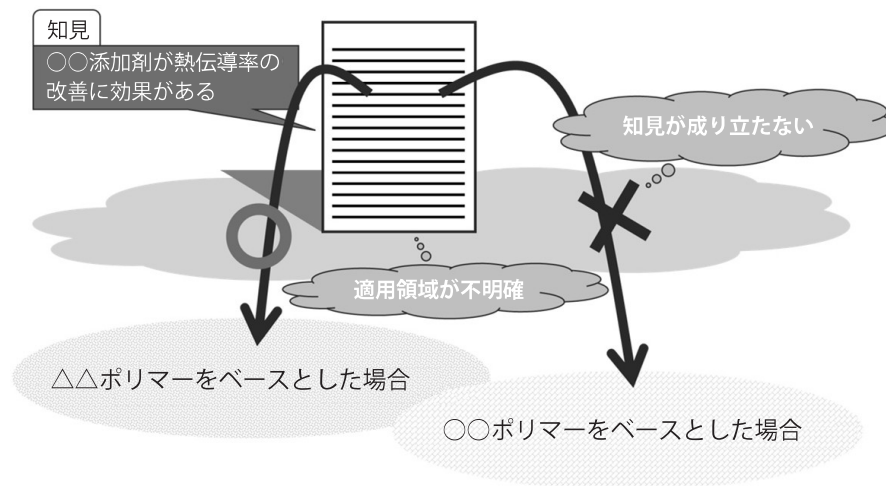


図6 報告書に書かれた知見や法則性は適用領域が不明確なものが多い

実は、ニュートンの運動方程式やシュレディンガー方程式も、どの領域ならその法則が有効かをデータとともに明示的に記して提案されたわけではなく、結構広い領域で成り立っているであろうという信念のもとに提示されている。どちらかという、それを読んだ人がその法則に懐疑的で、反例を探そうと様々な実験をし、反例が見つからないことで、その法則が成立する領域が明らかになっていったという歴史の流れになっている。そして、長い年月と持続的な反例探しに耐えることで、蓋然性の高い法則になっていっているのである。こういうスタンスで、論文や報告書を参照するのであればいいのだが、社内共有の観点では、報告書に懐疑的で反例を探そうと思ひ報告書を読む人は稀で、どちらかという書かれていることは、すでに蓋然性の高い法則、知見だと思って参照してしまう人が多い。それはかえって研究を遅延させかねない行為なので、そういう点をよく理解した上で、報告書は活用すべきであることをここで指摘しておく。

最後に少しだけ個人的認識を書いておく。一番安心な報告書の活用方法は、報告書内に書かれた知見、法則を確定的なものと考えず、研究の新たな観点、閃き、アイデアを得る源泉として利用することではないだろうか。つまり、うまくいかなかったときに「報告書にこう書いてあったから」ということを言い訳にせず、「自分が閃いたから」ということで、報告書抜きに自分だけで説明責任を引き取る覚悟をもって、報告書を活用すればいいということである。

3. R&D 部門において AI エージェントに期待して良いこと、悪いこと

前項では、「R&D 部門において生成 AI に期待して良いこと、悪いこと」と題し、R&D 部門において生成 AI の質問応答に関して、活用可能性について論じた。本項では、「R&D 部門において AI エージェントに期待して良いこと、悪いこと」と題し、R&D 部門において AI エージェントの活用可能性に関して論じる。

皆さんは十分承知しているだろうが、最初に「AI エージェント」とは、どういうものか一般的な説明を書いておく。「AI エージェントは、目標を与えると、自分で状況を見て考え、手順を決め、必要なツールやシステムを使って主体的に動く仕組み。途中で環境が変われば処理自体を自律的に変えるもの」というのが、「AI エージェント」の一般的な説明である。分かったような気もするがもやもやとするのではないだろうか？この話の前に「生成 AI」ができないことがあり、「AI エージェント」ではそれをできるように拡張したものであるということを理解しておかないと、「AI エージェント」の本質は理解できない。

「生成 AI」は、質問応答や画像生成、グラフ作成、プログラム作成など様々なことができる。そして、「目標を与えると、自分で状況を見て考え、手順を決め、主体的に動き、途中で環境が変われば処理自体を自律的に変える」こともそのようにプロンプトを書けば、できるのである。それでは、「AI エージェント」との違いは何だろうか？最初の

「AI エージェント」の説明をよく見てもらおうと、正解が分かるはずである。「生成 AI」にできなくて、「AI エージェント」にできること、それは、「必要なツールやシステムを使って」である。実は、「生成 AI」と「AI エージェント」の差は、「必要なツールやシステムを使うことができる」、この 1 点だけである。

しかし、この 1 点は、できることの幅を大きく広げることになる。例えば、直近で行った複数の実験に関して、添加剤濃度が熱伝導率にどのような影響を与えるかを確認するために、グラフを描くことを考える。生成 AI でもグラフは作れるので、この作業は生成 AI だけでもできなくはない。しかし、直近で行った実験のレシピ情報や熱伝導の計測データを自力で集めてきて、それを生成 AI に送り込み、「添加剤濃度が熱伝導率にどのような影響を与えるかを確認するために、グラフを作成してください」と指示を出す必要がある。一方、「AI エージェント」では、「直近で行った実験の ID を指定し、レシピデータは〇〇か××にあるのでそこから入手して、粘度計測データは△△か□□にあるのでそこから入手して、それらデータを使って、添加剤濃度が熱伝導率にどのような影響を与えるかを確認するために、グラフを作成し、そのグラフを私宛にメールください」と言えば、すべて自動で行ってくれるのである。ここで、必要なツールやシステムは、Windows のファイル検索機能であったり、ファイルを PC に転送する転送ツールであったり、メールを送るメーラーであり、既存のツールやシステムに関して人を介さず、「AI エージェント」から使えるのである。「生成 AI」からは、いくら指示を出しても既存のツールやシステムは一切使えない。唯一使えるのは、Internet 上での検索 (google 検索と同じと考えてもらってよい) だけである。「生成 AI」と「AI エージェント」の違いは分かってもらえただろうか？「必要なツールやシステムを使うことができる」は、結構便利さを拡大するのである……。

前文章がなぜ「……」で終わっているのは、誤植ではない。私の文章を何度か読んでいる人は、すでにこの先の展開が見えているかもしれない。実は、「AI エージェント」は便利と言いながら結構、様々な制限があるのである。AI エージェントが使えるのは、「AI エージェントがアクセスを許可した API (より正確には Built-in Tool API, Public SaaS API, Auth-Restricted Private API の 3 パターンだけで、SSH や上記以外の REST API はアクセスできない) を持つツールやシステムだけ」なのである。だから、計測器に付属している分析ソフトも社内でデータ蓄積しているデータベースも自作で作った解析プログラムもメーラー、そのままでは AI エージェントからアクセスできないのである。これは、AI エージェントを PC で動かしていたとしても、その PC 内にあるツールにもアクセスできないのである。しかし、最初から AI エージェントからアクセスできる API を持ったツールやシステムなんてないはずなので、「必要なツールやシステムを使うことができる」は机上の空論に近いものである。

「最初から AI エージェントからアクセスできる API を持ったツールやシステムなんてないはず」と言ったが、実はクラウド系のサービスは AI エージェントからアクセスできる API を持ったものが多い。例えば、Gmail, Microsoft Graph Mail, チャット・メッセージ系サービス, google doc, google sheets, google drive, One drive, Dropbox, Box などには AI エージェントからアクセスできる API を持っている。つまり、こういう最新のサービス環境だけで、業務を進めていることが AI エージェント利用の前提なのである。それでは、計測器に付属している分析ソフトや社内でデータ蓄積しているデータベースや自作で作った解析プログラムや社内の sendmail は、AI エージェント化するのは諦めないといけなくはないかということ実はそうではない。AI エージェントからアクセスできる API を自作してやり、それをそれらソフトウェアやサービスの前面にラッパーとしてかましてやれば、AI エージェント化が可能になる。例えば、exe を実行する API を Python で書くと、以下のようなものになる。

```
from flask import Flask, request
import subprocess
app = Flask (__name__)
@app.post ("/run")
def run ():
    args = request.json["args"]
    result = subprocess.run (["C:/tools/mytool.exe"] + args, capture_output = True, text = True)
```

```
return {"stdout": result.stdout, "stderr": result.stderr}

app.run (port = 5000)
```

自作のコードや計測器の分析ツールは、この方法で回避できると思うが、sendmailなどの既にサービスとして提供されていて、多くの人が使っているものはたぶん、情報システム部の管轄なので、こういうラッパーを勝手に付け加えることはできない。ラッパーをつけるためには管轄部署との折衝が必要になり、結構な高確率で、受け入れてもらえないと思うので、そのあたりは覚悟しておいた方がよい。

最後に AI エージェント化するときにはやっつけてしまいがちだが、絶対にやってはいけないことを説明しておく。実は、AI エージェント化（実は、AI は関係なくて自動化全般）に便乗して、今まで行っていないデータ探査、データ処理をこの機に AI エージェントに実装してしまいたいという要望が挙がるがよくある。しかし、これは絶対に行ってはならないことのひとつである。以下、なぜこのようなことは行ってはいけないのかを説明しておく。

あ) 第三者が手動処理で確実に実施できる手順書がない

オルゴールで例えると AI エージェントは、オルゴールの外箱であり、楽譜（＝手順書とツール）があって、それをドラムに刻み込ま（AI エージェントへの指示を書く）なければ AI エージェントは完成しない。手動処理で確実に実施できる手順書がないということは、オルゴールを作るときに楽譜がないということと同じである。楽譜なしに音楽をドラムに刻み込むのは不可能であり、また、即興で楽譜を書いたとしても天才でない限り、まともな音楽になっていることは奇跡でしかありえない。実際、すでに手動実施していることを手順書に書き下すことでさえ、悪戦苦闘するはずであり、その状況を考えると、手順書なしに AI エージェントを作ることが如何に現実的でないか理解できると思う。

い) AI エージェント化対象の作業が他の作業より優先して対応すべきか客観的に判断できない

即興で実際に実施可能な手順書が書けたとしても、「その手順通り実施した結果が価値の高いものか？」や「その実施頻度がどの程度か？」は、判断がつかないということである。つまり、そもそも AI エージェント化対象を決める序列の「作業発生頻度 × 作業時間」の正確な値が算出できず、他のものと客観的に優先度を比較できないのである。

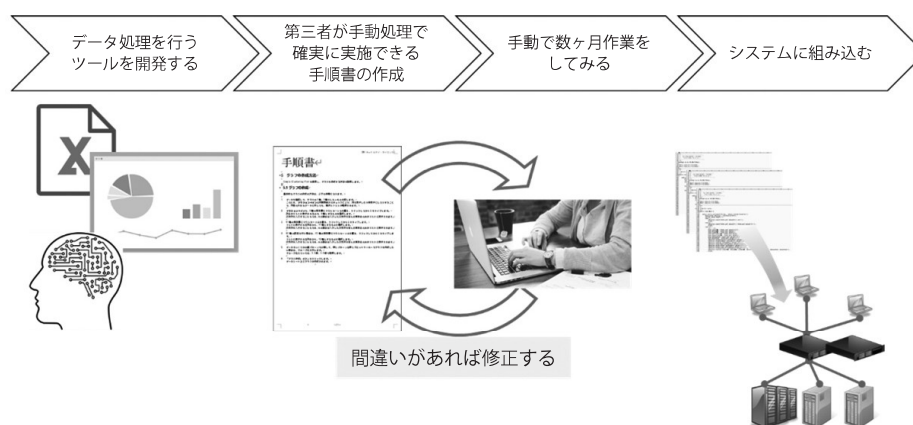


図7 手動で行えていないことを AI エージェント化するまでの手順

「今まで行っていないデータ探査、データ処理をこの機に AI エージェントに実装してしまいたいという要望」の中で、ツール自体もまだ存在していない場合もある。当然このような作業の AI エージェント化は、上記理由より行うべきではないが、「ツールが存在していない」状態では、AI エージェント化を行ってはいけないもう一つの重要な理由を以下で説明しておく。

う) ツールと AI エージェントの境界が曖昧になる

R&D 部門では項目追加, 変更だけでなく, データ処理に関しても追加, 変更が頻出する。これが業務系と異なり, AI エージェント化を難しくしている原因である。このような R&D 部門でも AI エージェント化を可能にするためには, ツールと AI エージェントの分離が重要で, AI エージェントからツールを呼び出す形でなければならない。そのようにすることで初めて, AI エージェントなしにツールを使う手順を確立することが可能になる。しかし, ツールと AI エージェントを同時に開発すると, 手順の確立が難しくなってしまう。また, 同時開発はツールと AI エージェントの境界を曖昧にしまい, その結果, ツール部のみの改良や AI エージェントのみの改良したときにその改良で全体として問題が生じないかの確認, 検証が難しくなる。つまり, ツールや AI エージェントのバージョンアップ時に AI エージェント全体の長期の停止が必要になるなど, メンテナンス性が悪くなるのである。

え) ツール部, AI エージェント部及び手順の問題の切り分けができない

ツールと AI エージェントを同時に開発すると, 何か問題があったときに, 手順に問題があったのか?, ツール仕様の問題があったのか?, AI エージェント部分に問題があったのか? の切り分けが非常に難しくなる。そもそも, AI エージェントを使わない状態での正解が示されないとその判別自体が原理的にはできないということになる。AI エージェントの問題であることを排除するためには, 手動でツールのみを使った手順の確立が必須である。もちろん, ツールの問題であることを排除するためには, ツールを使わない手順の確立が重要である。したがって, ツールを自作する場合には, 他の複数の商用ツールなどでも同じことができる手順を確立しておくべきである。また, AI エージェント開発会社にすべてを丸投げしてしまう場合は, 丸投げ体質が染みつき, もし, ツール部や AI エージェント部に問題があっても, 気づかず間違った処理をしたままになる可能性が高くなる。実際, 正確な仕様書と手順書がないと AI エージェント開発会社も何が正しいのかを把握できないことになるので, 正解がわからないところに丸投げをするという非常に危機的な状態に陥る。

「それでは, 将来展望など, 現時点でできないことを実現するためにはどうすればいいのか?」という質問がよくあるので, 説明をしておく。

3.1 データ処理を行うツールを開発する

データ処理を行うツールが存在しないのであれば, まず, 手順での実施を可能にするためのツール作成を行う。ツールの動作を確認する為のサンプルデータとその結果も整えておくこと。実際, ツールに BUG がないことを保証するのは, 相当大変だが, システムに組み込む前に十分確認しておかないと, 皆がツールには問題がないとして使いますので, 十分注意が必要である。もし, ツールをリリースした 1 年後などに問題が見つかった場合は, その連絡及び関係部署への影響範囲の確認などが発生し, 結構大きな問題になる。

3.2 第三者が手動処理で確実に実施できる手順書の作成

ツールを使って, 手動で実施可能な手順を確立し, 手順書を作成する。また, この過程で, データの項目名や項目値の単位がばらばらとか, 数値の項目値に数値以外の値が入っている (> 1 , ~ 10 , // とか, 目視では意味が分かるが, AI エージェントには意味判別ができない記述) ものは, 適切な数値に修正をしたり, 新たな項目 (下限値, 上限値, 信頼幅など) を作り, その値に変更するなどの作業をする。この作業は一般的にデータクレンジングと呼ぶ。

3.3 手順書に基づき数ヶ月間, 手動で処理を実施してみる

手順書に基づいて, 手動で処理を実施してみる。実際にどんなに完璧に作った手順書のつもりでも実際実施してみると, 問題箇所が見つかるものである。手動実施の間に手順書を修正してしまうことが重要なのである。実際, AI エージェント化した後では, 途中過程が見えず, いきなり最終結果を得てしまうことになるので, こういう問題

が露わにならないことが多いものである。手動で、ゆっくりと実施するからこそ、途中過程でおかしなことに気づくのである。また、実際手動で処理をしてみると数ヶ月間で1, 2回しか実施しなかったり、実施をしても手動処理で数分で終わるレベルの処理であることが明確化することもある。こういう（頻度が少ない、手動でも多くの時間がかからない）処理は、AI エージェント化するメリットは小さいので、対象から外すべきである。AI エージェント化が数分でできるならいいのだが、様々なパターンのテストデータを用意して、AI エージェントが意図通りに動くことを検証することを考えるとメリットより、デメリットの方が大きいのである。

3.4 手動手順書を AI エージェント化する

3.3 で AI エージェント化する意義が高いと判断されたものだけ AI エージェント化に着手する。AI エージェントの記述自体は、結構簡単かもしれないが、様々なパターンのテストデータを用意して、意図通りに動くことまで検証する必要がある。他の人に提供するツールを開発した人は、理解していると思うが、ツールのプログラムを書くより、どのようなテストケースを行えば、意図通りに動くことの検証になるのかを考え、テストデータを作り、テストを実施する方が何倍も時間がかかるものである。AI エージェントは自然言語で書けるので、ツールのプログラムを書いた経験のない人でも書けるのだが、上記、テストのことは、頭がないことが多いはずである。結局、そういうテストが不十分なものは、不具合や問題を内在することになるのだが、これはプログラムでも AI エージェントでも同じことなのである。

4. AI・機械学習などの MI の特性と注意すべき点

前項では、「R&D 部門において AI エージェントに期待して良いこと、悪いこと」と題し、R&D 部門において AI エージェントの活用可能性に関して論じた。本項では、「AI, 機械学習について、特性と注意すべき点」に関して、もう少し説明を加える。

AI, 機械学習は、万能の打ち出の小槌ではないことはすでに述べた通りで、AI, 機械学習で必要とされる予測精度を得るためには、一般的に非常に多くの過去データと AI 最適化のための設定値調整が必要である。例えば、顔認証の AI, 機械学習を現在のような使いものになるレベルにするためには、数十名規模のプロジェクトで数百万以上のデータと数年以上に及ぶ AI, 機械学習モデル及び設定値の調整が必要だったのである。また、飛行場でのバードストライクを判定する画像認識は、羽田空港で使えるレベルのチューニングができて、それをドゴール空港で使用する、誤判定が多すぎて使い物にならず、結局、各飛行場でのチューニングが必要なのが現状である。「顔認証」も「バードストライク判定」も、どのレベルの精度が必要か？という問題が実用性に大きく絡んでおり、人を雇って「判定」をしたほうが安く、精度が高いようであれば、AI, 機械学習導入は見送られる。

材料開発のような分野では、恐らく、実験パラメータ数（材料種数、配合、焼結等のプロセスに関する設定値）は数百～数千に上るので、ある領域では期待する予測精度が出るかもしれないが、顔認証のようにあらゆる対象の判定は難しいはずである。また、空港の場合は、どの空港というのは簡単に指定できるが、材料開発の場合、そもそもこの適用領域というものをしっかり指定（定義）すること自体が難しい問題なのである。

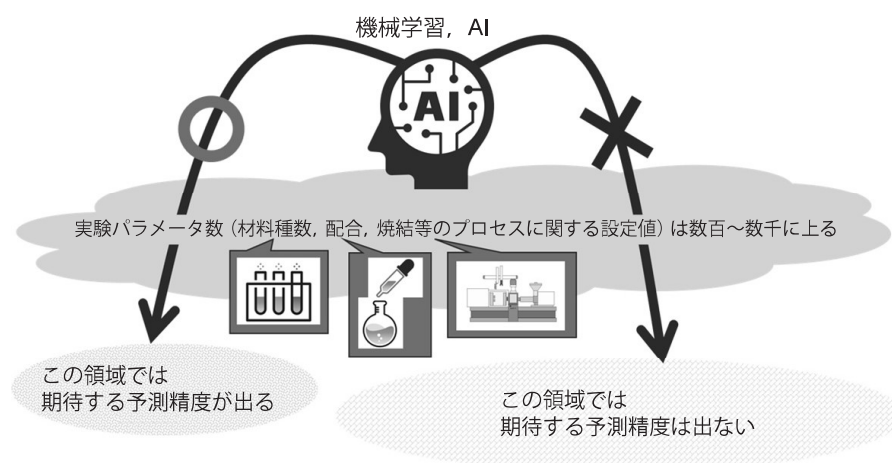


図 8 AI, 機械学習は期待する予測精度の出る領域の明確化が難しい

この領域定義がしっかりしていないと、「利用者にこういう領域はこの AI, 機械学習は高い精度を有するので使ってください」とアナウンスさえできないのである。この適用領域定義を明確にしないまま利用開放をすると, AI, 機械学習を使った方が使わない時より, 材料開発が遅延してしまいかねない。また, 利用者がアレルギー反応を起こしてしまい, AI, 機械学習を避けるようになってしまう可能性もあるので, 適用領域の明確化は非常に重要である。ただ, R&D 部門のように広大なパラメータ領域で, 常に未知の領域への挑戦を続けるような部門では, 概して, 必要とされる予測精度を確保できる適用領域が確定したころには, その領域は研究対象でなくなっている可能性も高い。逆の言い方をすると, 予測精度を確保できる適用領域を確定するためには, 当該領域で多くの実験が必要で, これは研究において, 「その狭い領域において, AI, 機械学習の性能確認のためにそんなに多くに実験を行うことが, 研究加速につながるのか?」という, 悩ましい問題にもつながるのである。

適用領域の話はここまでとし, ここからは AI, 機械学習と従来研究方法の関係に関して, 考察する。AI, 機械学習は予測結果を返すが, なぜそのような結果になるのかの理由を提示しない。一方, 研究というのは, その理由 (仮説, 理論) を考え, 検証をしていくものである。また, 前述した通り, AI, 機械学習の結果が常に正しいわけではないので, それがどれくらい正しいかを研究者の独自の方法でデータ分析をし, その理由 (仮説, 理論) 検討を行うことはやめてはいけない。それらの検討を踏まえ, 最終的に AI, 機械学習が提示する実験を行うかどうかの判断は, 研究者自身が行うべきなのである。

AI, 機械学習の精度が十分に高い領域が確立すれば, その領域の材料開発では, 仮説, 検証のためのデータ分析は不要になるかもしれないが, 研究者はデータ分析を常に行っておくべきである。そうしないと, 研究者のデータ分析能力が劣化し, AI, 機械学習の精度の低い領域の材料開発ができなくなってしまう。実際, AI, 機械学習の精度が十分に高い領域が確立されたというのであれば, そこは, データ分析ができる研究者に材料開発をさせるのではなく, 実験作業に特化した作業員が請負や派遣で実施するのが良いはずである。そして, そのような領域は将来的には AI, 機械学習と連動した自動実験に移行することも可能なはずである。



図9 AI、機械学習の予測精度の高い領域は、自動実験へ移行していくべき

ここまでで説明したように、AI、機械学習は、蓄積データがあれば大きな力を発揮するが、評価、運用には注意が必要である。実際に、AI、機械学習に関して気を付けるべきことを最後にまとめておく。

- 1) AI、機械学習は万能の打ち出の小槌ではなく、非常に多くの過去データと設定値調整が必要
- 2) AI、機械学習が適用できる領域を広げようとする現実が遠く
- 3) どのような実験パラメータ領域の材料開発でAI、機械学習が有効か明確にしないと反って開発を遅延させてしまう
- 4) AI、機械学習を利用してもAI、機械学習の結果を考察し、その理由（仮説、理論）検討を行うことはやめるべきではない
- 5) AI、機械学習の精度が十分に高い領域は、実験作業に特化した作業員が請負や派遣で、材料開発を実施するのが良い

5. AI・機械学習などのMIの研究への組み込み方法

前項では、「AI・機械学習などのMIの特性と注意すべき点」と題し、AI、機械学習などの特性を説明し、利用する場合に注意すべき点に関して、解説を行った。本項では、「AI・機械学習などのMIの研究への組み込み方法」と題し、AI、機械学習等を研究で活用するときに発生する様々な問題や研究者側のアレルギー反応などをできるだけ抑えられるような研究スタイルに関して、考察をする。

前項で、材料開発の場合、AI、機械学習等のMIが必要とされる予測精度を確保できる適用領域を確認すること自体が難しい問題であること、そして、この適用領域定義を明確にしないまま利用開放をすると、AI、機械学習を使った方が使わない時より、材料開発が遅延してしまいかねないだけでなく、利用者がアレルギー反応をしてしまい、AI、機械学習を避けるようになってしまうかもしれない点を指摘した。

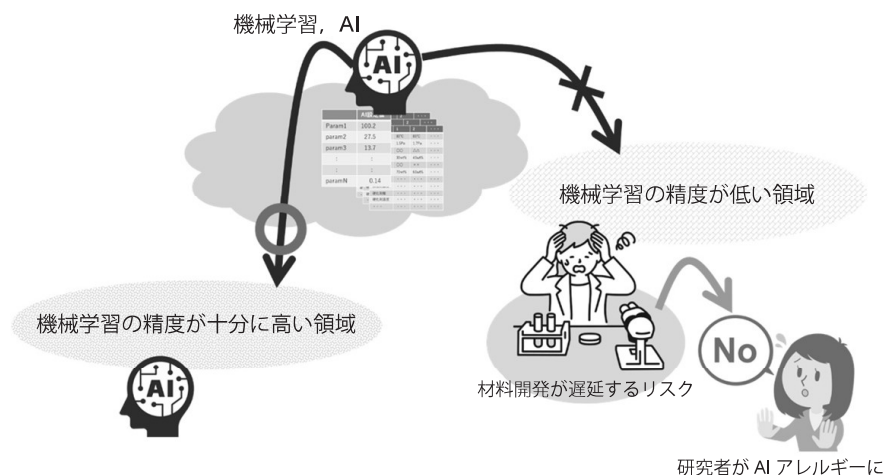


図 10 AI, 機械学習の予測精度の低い領域を明確化しておかないと研究者が AI アレルギーに

しかし、難しい、難しいと言っただけでは、いつまでたっても研究に AI, 機械学習や MI を導入できない、もしくは、研究を阻害する形での導入になってしまうことになる。

AI, 機械学習などの MI を導入しようとしている部門では、開発材料の特性を改善させるための実験パラメータを提示させる形で AI, 機械学習などの MI を使おうとしていることが多い。しかし、この形での導入では、必要とされる予測精度を確保できる適用領域が明確でない限り、前述で指摘したように様々な問題が生じる。また、AI, 機械学習などの MI から提示された実験パラメータに単に従って実験してはダメで、提示された実験パラメータがどれくらい適切かを研究者の独自の方法でデータ分析をし、その理由（仮説、理論）検討を行う必要があるが、実はそれも結構難しい。それでは、実際の研究現場には、どのようにして AI, 機械学習などの MI を導入するべきだろうか？

実際の研究現場に AI, 機械学習などの MI を導入するためには、大きく二つのハードルがある。まず、一つ目のハードルは「必要とされる予測精度を確保できる適用領域が明確でない」ことを前提にしなければならない点である。そもそも「必要とされる予測精度」＝「従来の方法で研究者が開発材料の特性改善に要した試行錯誤回数」を明確化すること自体、難しいはずである。実際、特性改善のケースごとのばらつきが多く、平均値に意味があるわけでないで、特性改善難易度毎の分類が必要だが、その分類自体相当困難であり、「必要とされる予測精度」を明確化することは諦めたほうが良い。もう一つのハードルは、「AI, 機械学習を使うべきでない領域に使ってしまうと、材料開発が遅延して、利用者がアレルギー反応をしてしまう」ことである。これは、「適用領域を明確化すること」と等価だが、そもそもの前提の「必要とされる予測精度」を明確化すること自体が困難で、もし、それができたとしても、「適用領域」というものを厳密に定義することも同じぐらいに困難で、それを克服しないと利用者に悪影響をもたらすという問題である。したがって、AI, 機械学習を使った方が使わない時より、材料開発が早くなるようにできればいいのだが、AI, 機械学習の開発者、メーカーとしては、それは保証できないはずである。「必要とされる予測精度」とその「適用領域」を明確にしないまま AI, 機械学習などの MI を利用しても、材料開発が現状よりも遅延しなければ、アレルギー反応は起らないはずである。例えば、以下のようにすればアレルギー反応は抑えられる。

a) 最初は、従来の方法で、つまり、過去データと属人的経験則も生かしながら自力で試行錯誤する。

この方法の試行錯誤が煮詰まってきたと感じたら

b) 次に、実験計画法などの統計的手法に則った実験及び解析を行なう。

この方法でも要求される特性改善に到達しなければ、

c) 最後に、AI, 機械学習などの MI に頼ってみる。

機械学習，AI を効果的に使った新しい研究の進め方

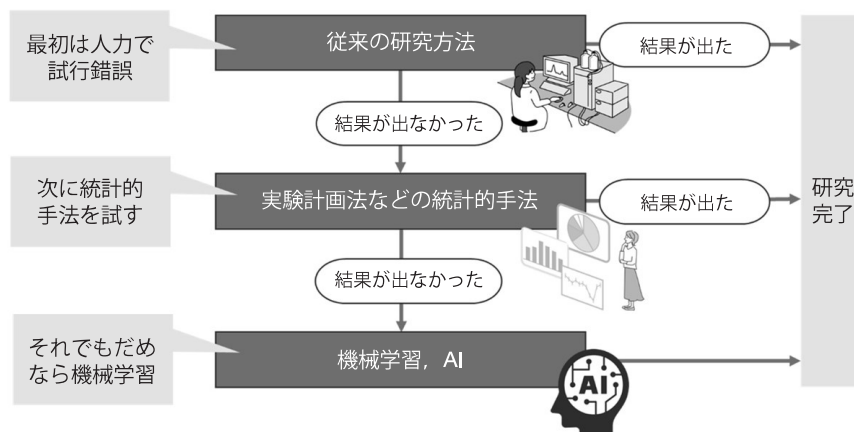


図 11 AI アレルギーを起こさない AI，機械学習を使った研究の進め方

この手順であれば，すでに自力での試行錯誤を十分行った上で，解決できなかったわけなので，AI，機械学習などの MI で解決できなくともアレルギー反応は起きないはずである。しかし，この方法だと，今までの 3 倍も試行錯誤が増えてしまう可能性がある。それを抑制するために試行錯誤を開始する前に，開発期限などを参考にして，試行錯誤に掛けられる最大試行回数を最初に見積もっておき，自力，実験計画法，AI，機械学習などの MI に 1/3 ずつ割り当てるとするとよい。ある意味，今までは自力で何とかするしかなかったので，試行錯誤が煮詰まってきたと感じても，無駄と感じながら試行錯誤をやめられなかったのだが，この方法ではあっさりギブアップが許されるのである。そして，後半は，ある意味自分の責任は軽くなり，実験計画法や AI，機械学習などの MI の責任は重くなるので，少しは気持ちは楽になれるはずであり，実験計画法や AI，機械学習などの MI に感謝の念を抱けるかもしれない。

実は，AI，機械学習などの MI は，もう少し，活用範囲を広げることができる。a)，b) を実施するとき，実験前に AI，機械学習などの MI にどんな結果になりそうか聞いてみるのである。

a) 最初は，従来の方法で，つまり，過去データと属人的経験則も生かしながら自力で試行錯誤する。

b) 次に，実験計画法などの統計的手法に則った実験及び解析を行なう。

a) ではある程度，良い結果が得られそうな算段がある実験パラメータの実験を計画しているはずである。その自分の判断と AI，機械学習などの MI の判断を突き合わせるのである。自分の判断と AI，機械学習などの MI の判断が近ければ，自分の判断にも自信がもてるだろうし，AI，機械学習などの MI に対しても，「なかなかやるな」という感情を持てるのではないだろうか？ 実際の実験の結果が予想通りであれば，自分と MI で称えあい，予想とは全く違う結果でも「お互いまだまだだな」と慰めあえる。自分の判断と AI，機械学習などの MI の判断が全く違うようであっても，a) の段階では MI の判断に従う必要はない。実際の実験の結果が自分の予想通りであれば，MI に対して，「やっぱりお前はまだ半人前だ」と言い放てばよく，MI 予想の方が近ければ，真摯に自分の負けを認め，MI の優秀さを認めることができると思う。どうだろうこの方法であれば，MI が愚かでも，賢くても MI アレルギーは起こらないし，それどころか同じ研究に対峙する同士として，親近感さえ覚えられようになるのではないだろうか？

b) の時では，自分が計画した実験パラメータではないので，どんな結果になりそうかは，a) に比べると予想できないことが多いはずである。ただ，自分なりの考えで，どのような結果になるか推測することは，推測力を鍛えるためにもよいことなので，是非，推測を試みてほしい。そして，MI にもどんな結果になりそうかは聞いてみて，a) の時と同じように一喜一憂してみればよい。こうやって，自分の推測と MI の推測を常日頃突き合わせることで，お互いの強い部分と弱い部分がわかるようになってきて，お互い頼れる相棒になっていけるのである。実際，相手がいることで，自分の弱い部分が意識できるようになり，単に弱いということがわかるだけでなく，改善もされていく。MI も自分の推定が当たったり，外れたりした結果をどんどん取り込んでいくので，改善されていく。

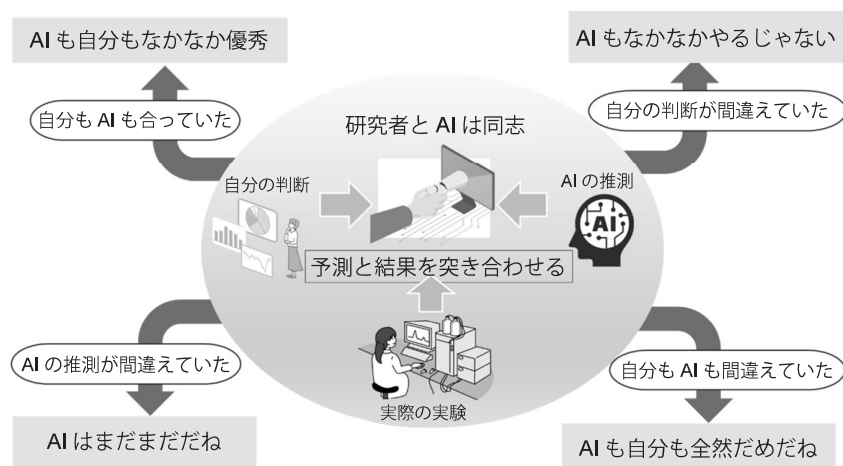


図 12 AI、機械学習との理想的な付き合い方

この話は、単なる物語として、書いているのではない。将棋の世界では、まさにこれと同じようなことをして、AI と対話をしながら、自分の弱いところ、AI の弱いところを意識し、改善していくことで、双方、強くなっているのである。研究の世界では、将棋のようにまだ AI が圧倒する状況ではないが、AI との付き合い方には学ぶところが多いはずである。

6. R&D 部門におけるデータ蓄積，共有，利活用の実情

前項では、「AI・機械学習などの MI の研究への組み込み方法」と題し、AI、機械学習等を研究で活用するときに発生する様々な問題や研究者側のアレルギー反応などをできるだけ抑えられるような研究スタイルに関して、説明を行った。本項では、「R&D 部門におけるデータ蓄積，共有，利活用の実情」と題し、AI、機械学習等を研究で活用するための大前提となるデータ蓄積，共有，利活用の実情に関して、紹介する。

R&D 部門に関わらず、データが共有、利活用されるためには「データが管理された状態」になっている必要がある。データ共有，利活用の実情の話の前に、「データが管理された状態」とは、何を意味しているのかを説明する。「データが管理された状態」とは、データを生み出した実験及び解析を第 3 者が再現するために必要な情報が記録され、それを必要な時に迅速かつ確実に探し出せる状態に保っていることを意味する。ここで言っている「再現するために必要な情報」というのは、実際上の話ではなく原理的という意味であり、「実験結果にほとんど影響しないので、無視してもよい情報」というものは省略してはいけない。それは、「実験結果にほとんど影響しないので、無視してもよい」というのは結果論であり、実験する前にそのことが確定するものではないからである。つまり、「再現するために必要な情報」とは、実験者が指定した条件のすべてと、意図的な指定はしていないが実験に影響を及ぼす条件（例えば、室温、湿度、気圧など）の全体を意味している。意図的な指定はしていないが実験に影響を及ぼす条件は、明確にすること自体に難しい点があり最初から全て列挙できるものではないので記録自体が難しいが、実験者が指定した条件の全てを記録することには、そういう意味での難しさは一切ないはずである。しかし、「実験者が指定した条件の全てを記録」することさえ、ほとんどの R&D 部門では達成されていない。



図 13 実験及び解析を第3者が再現するために必要な情報を記録

そういう観点において、ほとんどの R&D 部門におけるデータは、管理された状態というレベルには達しておらず、単なる蓄積と呼ぶのが相応しいというのが実情である。公的、民間の様々な R&D 部門を見てきた結果、データがどのように蓄積されているのかを、以下で紹介する。

一人で完結できるような実験や解析では、ほとんどの情報は研究者の頭の中のみであり、注目しているパラメータのみを研究ノートにメモ書きをされているだけの場合がある。実験や解析結果の比較評価がある程度難しい課題に対しては、実験や解析の情報が Excel に書き写され、比較しやすいように纏められ、個人 PC 内に保存されていることもある。複数の人が関わった実験や解析の場合は、他の人への実験や解析の引き渡し（依頼）に必要な情報のみは、フォーマットが揃えられた用紙もしくは Excel が準備されていることが多いが、それ以外の部分は、上記状況と変わらない。これらの状況は、「データを生み出した実験及び解析を第3者が再現するために必要な情報を記録し、それを必要な時に迅速かつ確実に参照できる状態に保っておく」という観点に立つと、「データ管理」ができていない状態ということになる。これらの状況では、研究者本人は何らかの形で頭の中では整理ができていると思っており、以下ではこの状態を「属人的なデータ管理」と呼ぶことにする。

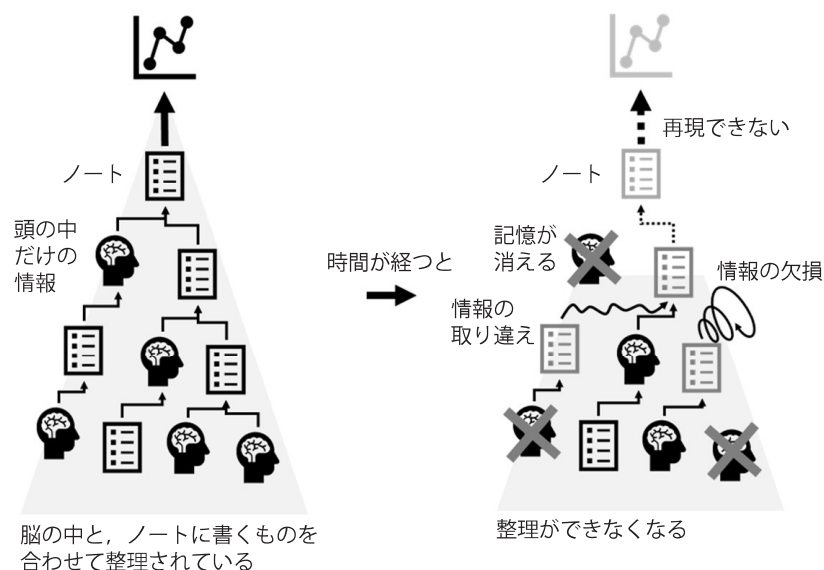


図 14 属人的なデータ管理とその特性

「属人的なデータ管理」状況では、実験及び解析の詳細なことは、実際の実施者しかわからない状況になる。当然のことながら実施者以外の人が実験条件や結果内容を知ろうとした場合、実施者にそれらを聞くしかない。実施者から実験及び解析を第3者が再現するのに十分な情報を提供してもらえれば問題はないが、実験データをそういうことができるような状態で蓄積している研究者はほとんどいない。そもそも、どれだけ整理好きの研究者でも、「〇〇の実験・解析情報、データ」が欲しいと言われたところで、該当するデータを探し出すことさえ、相当困難なことが多いはずである。該当するデータが漏れなく、間違いなく提供されることは、ほぼ不可能と考えてもいいかもしれない。このような状況の中、データの授受を行うと、間違った情報を基に実験や解析を進めることが発生し、間違った結論が導かれたり、検討を進めたのちに始まりに立ち返って、再実験や再解析を行わざるを得なくなることもある。共有した結果に間違いがあったり、再実験や再解析を行う羽目になった結果、データ提供者の信頼が失われ、データ共有の意欲がなくなっていき、データの共有、利活用は次第に廃れていってしまうのである。

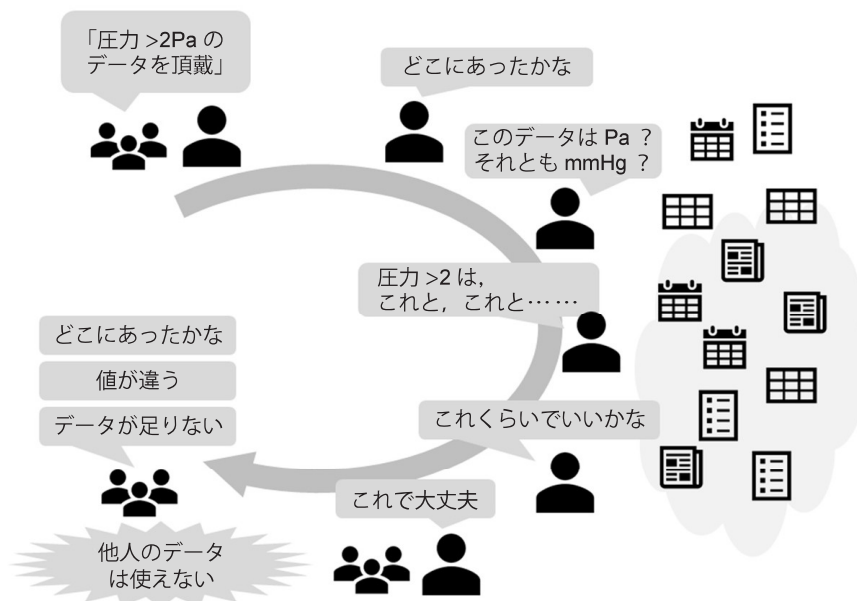


図 15 属人的に管理されたデータが引き起こす問題

このような蓄積状況のデータを使っているようでは、当然、AI、機械学習が研究の役に立たないのは理解できると思う。ただ、「AI、機械学習が研究の役に立たない」という事実には、少し注意すべき点がある。従来型の解析や分析では、X-Y プロットのように実データに立ち戻り、解析、分析結果を解釈することを行う。したがって、このような加減な蓄積状況のデータを使うと、データ自身に問題があることに気づくことが多い（もちろん、必ず気づくとは言わないが・・・）。しかし、AI、機械学習では、データを AI、機械学習プログラムに入力してしまうだけで、個々の実データに立ち戻り、結果を解釈することは行わない。したがって、データ自身に問題があることは気づきようがない。AI、機械学習の予測精度が悪いのが AI、機械学習のモデルのせいなのか？誤ったデータを入力してしまったからなのか？は、分離できなく、わからないのである。つまり、AI、機械学習を活用しようとするれば、従来型の解析や分析以上に、データの信頼性の確保が必要ということである。

「属人的なデータ管理」状況が生み出される直接的原因は、そもそも第3者が再現するのに十分な情報が記録されていないことと、記録されているものに関してそれが何を示す値なのか、つまり、その値の項目名に関して、人によって、また、同じ人でも時期によって、その項目名の同一性が保たれていないことになる。これらの解決方法は、研究開発リーダーのバックナンバー及び書籍「研究開発部門への DX 導入による R&D の効率化、実験の短縮化」の私の執筆部分に詳しく記載しているので、そちらを参照してほしい。

7. データ蓄積，共有で研究の何が改善できるのか？

前項では、「R&D 部門におけるデータ蓄積，共有，利活用の実情」と題し，AI，機械学習等を研究で活用するための大前提となるデータ蓄積，共有，利活用の実情に関して，紹介した。本項では、「データ蓄積，共有で研究の何が改善できるのか？」と題し，そもそもデータ蓄積，共有で，「どのようなことが改善できるのか？」，逆に言うところ「どのような改善は期待してはいけないのか？」を論じる。

「データ蓄積，共有で研究の何が改善できるのか？」を論じる前に，まず，「研究」というものの自身がどういう要素から成っているかを考えてみる。研究は，閃きや発想という独創的な力と，データを分析し，傾向を把握する力という2つの力から成っている。閃きや発想は，属人的，主観的なもので，それが研究者の個性的能力ともいえる。一方，データを分析し，傾向を把握する力は，一見，閃きや発想と同じように属人的，主観的なもののように感じるが，よく考えてみると数学や統計などの客観的方法で確立されるべきものである。実際，同じデータを使って，分析し，傾向を把握したとすると分析結果の解釈以外は，同じ結果になるべきものである。また，分析結果の解釈も単なる閃きや発想と比較すると属人的，主観的というよりも結果解釈時に何に重きを置くのか？何を無視するのか？という取捨選択のパターンという客観的なものとして考えることも可能である。つまり，データを分析し，傾向を把握する力は，主観的でなく，客観的なものだということである。

前述のことを念頭に置きながら，「研究とは何か」を少し考えてみる。大きく想像を膨らませてほしいのだが，「大航海時代に冒険家が世界の海を渡り，次々に新発見をしていった過程」と，「研究者が研究で新しい発見をしていく過程」は近いものがあるのではないだろうか？そう考え，研究を航海に例えてみると，「航海士の経験，勘は，研究者の閃きや発想」に対応し，「航海での海図と羅針盤とその利用技法が研究でのデータとデータ抽出，分析手法」に対応するのではなかろうか？どれだけ優秀な航海士もでたらめの海図と羅針盤とその利用技法では，まともな航海はできないように，研究もデータとデータ抽出，分析手法がでたらめでは，まともな研究はできない。実際，海図と羅針盤とその利用技法が航海士で共通のものであるべきであるように，データとデータ抽出，分析手法も研究者で共通であるべきで，それによって研究の客観性を担保するのである。優秀な航海士は，その客観的情報（海図と羅針盤とその利用技法）に，自身の主観的な経験，勘を組み合わせることで，他の航海士より，危険な海をより安全に航海できる。当然研究者も客観的情報（データとデータ抽出，分析手法）に，自身の主観的な閃きや発想を組み合わせることで，他の研究者ではできない発見が可能なのである。「研究と航海」，非常によく似た対応関係になっているように見えないだろうか。

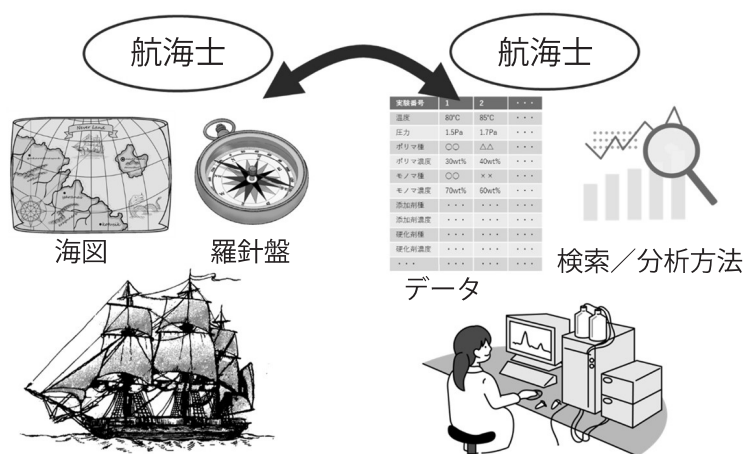


図 16 研究者のデータと検索／分析手法は，航海士の海図と羅針盤と同じ

本題に戻って、「データ蓄積，共有で研究の何が改善できるのか？」について，考察を進めよう。研究は，閃きや発想という独創的な力とデータを分析し，傾向を把握する力という客観的な力の2つの力から構成されているので，「データ蓄積，共有で研究の何が改善できるのか？」は，「データ蓄積，共有で閃きや発想が改善できるのか？」と「データ蓄積，共有でデータ分析や傾向把握が改善できるのか？」に言い換えることができる。「閃きや発想」は，属人的，主観的なものなので，改善できるとしても研究者ごとにその改善内容や改善度合いは異なってしまう。また，「閃きや発想」というものは，定量的に評価しにくいものであり，改善されたかどうかを判定すること自体が非常に難しい。一方，「データ分析，傾向把握」は，根本的には数学的，統計的データ処理やグラフ化などの客観的方法で実施されるべきもので，主観ではなく客観的な対象データ自体によって決定づけられるものである。つまり，数学的，統計的データ処理やグラフ化の結果解釈以外，すなわち，数学的，統計的データ処理やグラフ化自体は主観的でなく，客観的なものだということである。

原理的には，データ蓄積，共有で「閃きや発想」も「データ分析や傾向把握」も改善は可能ではあるが，主観が絡み合った部分は，改善指針や改善評価自体を研究者個人個人に合わせこむ必要があるので，負担が大きく，取り組みにかかる負担が改善による効用を上回ってしまう可能性が高く，最初に取り組むべき課題ではない。したがって，改善指針や改善評価が人によらない数学的，統計的データ処理やグラフ化部分をデータ蓄積，共有での最初の改善対象と考え，その改善に道筋ができたあとで，数学的，統計的データ処理やグラフ化の結果解釈部分をデータ蓄積，共有での改善対象として，取り組むべきである。

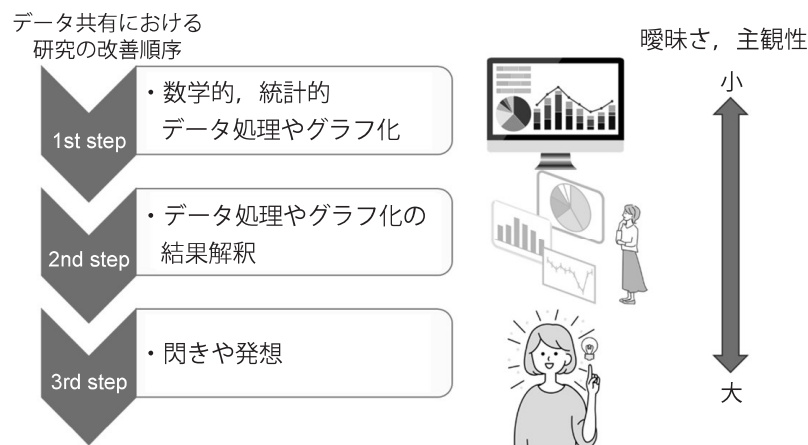


図 17 データ蓄積，共有における研究の改善順序

「閃きや発想」は，直接的な改善を期待するのではなく，上記の改善で「閃きや発想」にかけられる時間的余裕ができ，新しい「閃きや発想」が生まれる機会が増えると考えておくのが現実的である。もちろん，直接的な改善は絶対ないということではない。もしそういうことがあれば，蓄積，共有したデータをどのように扱えば「閃きや発想」が改善されるのかを客観的，統計的に検証し，他の人でも「閃きや発想」の改善が確実にできるような方法を確立していけばよい。実際，こういう客観的，統計的検証を経るまでは，「閃きや発想が改善している」ということ自体が主観的な思い込みの可能性が高いので，信用してはいけない。

ここで，少し逆説的な話をしておく。ここまでの，データ蓄積，共有での最初の改善対象は，データ処理やグラフ化部分と考えるべきと述べた。それは本来客観的に扱える部分を主観的に扱っていたものを客観的に扱えるようにすることで，研究の質的改善を促そうというものであった。航海との比較でもこの質的改善の意義は理解して頂けたと思う。しかし，実はデータ解析やグラフ化，AI，機械学習を使って研究を高速化（目標の特性改善が図れるまでの時間を短縮）することを目的として，データ蓄積，共有は行ってはいけないのである。航海の話に例えると，高速に航海をするということを目的として，海図データを増やそうとしてはいけないということである。海図データを蓄積，共有するのは，航海を客観的に行えるようにし，航海の安全性，安定性を確保するという観点で航海を質的に改善す

るものである。実験データの蓄積、共有も実験自身に客観性を付与し、実験の再現性や客観的知見を共有することで、研究を質的に改善するものなのである。

この言明は、「蓄積されたデータを使って、データ解析やグラフ化、AI、機械学習を行ってはいけない」ということを言っているのではない。この言明で何を伝えたいかというと、「蓄積されたデータを使って、データ解析やグラフ化、AI、機械学習を行ったところで、研究が高速化される保証はない」ので、「データ解析やグラフ化、AI、機械学習による研究の高速化を目的として、データ蓄積、共有を行おうと考えてはいけない」ということである。研究の質的な改善が高速化というような量的な改善にはつながらないということを言っているのでもない。データ蓄積でデータ解析やグラフ化、AI、機械学習も改善することは間違いない。しかし、データ解析やグラフ化、AI、機械学習による研究の改善が研究自身の高速化に影響を及ぼす程度の量的改善になる保証は何もないということなのである。研究だけに限らず、こういう因果関係の話を単純に考えすぎて、間違った判断をしてしまうことは多いので、これを機に心にどめておいてほしい。「研究を高速化しろ」と言われて取り組みを行っているのに、そんなことを言われても……。そもそも、「データ解析やグラフ化、AI、機械学習を使わずに研究を高速化することなんて、できないのでは？」と今思ってしまった人は、思考が固定パターンに嵌り過ぎている状況で、研究を進める上では危険な状態である。

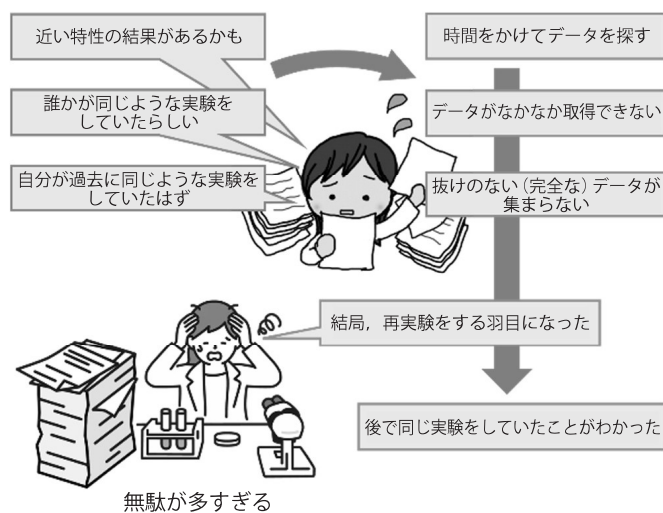


図 18 データ蓄積、共有が不完全だと研究に様々な無駄が発生

私が様々な組織でデータ蓄積、共有を推進してほしいと頼まれるときに、最初に行うことがある。それは、現在の研究、実験において、困っていること、時間がかかっていることをブレインストーミング的に洗い出すことである。もちろん、「目標の特性改善を達成するまでに時間がかかり過ぎる」、「安定した計測ができる装置の使い方が分からない」というようなものが出てくると思う。ただ、その他にも「過去にこんな実験をしたと思うのだけど、Windowsのフォルダを探し回ってもなかなか見つからない」とか、「レシピを記録したと思われるファイルを見つけたのだけど、材料名が曖昧に書かれていて、どんな材料かわからなかった」とかもあるはずである。このような単純な困りごとは、そういうことが年に何回ぐらいあって、1回あたりどれくらいかかっているかを積算してみると、結構な時間がかかっていることに気づくはずである。

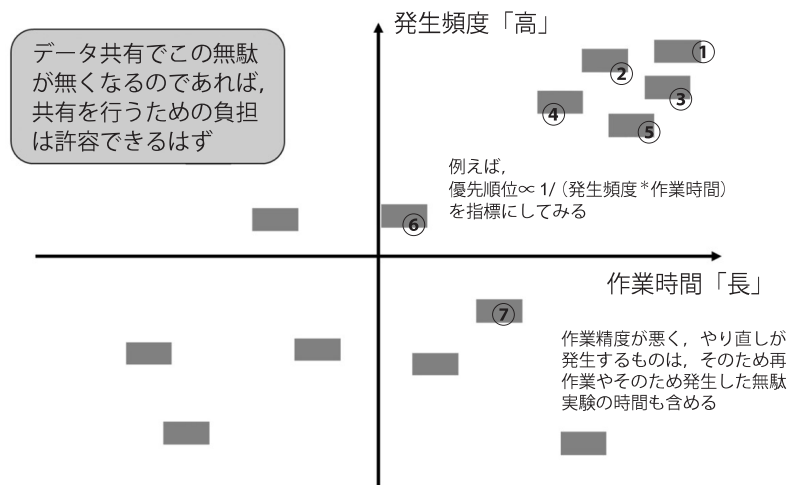


図 19 困っていること、時間がかかっていることを定量化

データ蓄積、共有をしっかりと、データベースなどにデータを格納しておき、いつでも素早く探し出せるようにしておく、これらの時間が大幅に短縮される。これは、確実に「目標の特性改善を達成するまでの時間が短縮される」ことになるのである。コロブスの卵のようなことではあるが、特性改善をするのに解析や AI、機械学習などの高度な技法を用いなくとも、確実に改善できる方法は山のようにあるわけである。「特性改善をする高度な技法」は、そういうコロブスの卵で節約できた時間で挑戦すればいいわけである。宝くじは買わなければ当たらないのと同じで、「特性改善をする高度な技法」もその開発に挑戦しなければ手に入らない。しかし、「特性改善をする高度な技法」は、必ずそういう技法が見つかる保証はない。だから、生活に支障のない程度で購入する宝くじのようにコロブスの卵で節約できた時間を超えない範囲で、挑戦し続けなければならないわけである。

最後にもう一言。「目標の特性改善を達成するまでに時間がかかり過ぎる」という課題はどうしようもないのだろうか？「特性改善をする高度な技法」を見つけようとする前にやるべきことがある。「目標の特性改善を達成するまでに時間がかかり過ぎる」のは、何故なのか？をしっかりと問うべきなのがある。いわゆる、「なぜなぜ分析」の要領で、要因を分解していくのである。そうすると、その要因のいくつかは、「過去にこんな実験をしたと思うのだけど、Windows のフォルダを探し回ってもなかなか見つからない」とか、「レシピを記録したと思われるファイルを見つけたのだけど、材料名が曖昧に書かれていて、どんな材料かわからなかった」とかのように、確実に改善策が見つかる課題に分解される可能性がある。そういう確実なものを確実に改善していき、確実とは言えないけれど、改善できれば大きな改善になりうる挑戦すべきものを明確化、先鋭化させていくのである。明確化、先鋭化されないと挑戦しようにも何をどう挑戦すればよいか分からないはずで、右往左往するだけになってしまう。したがって、挑戦すべきものは他の要因を削ぎ落とすことで挑戦方針を明確化、先鋭化してから、他の改善で節約できた時間で挑戦すべきなのである。

8. そもそもデータ蓄積、共有は何のためにやるのか？

前項では、「データ蓄積、共有で研究の何が改善できるのか？」と題し、R&D 部門におけるデータを蓄積、共有することで、データ探しなどの無駄が大幅に削減できることを説明した。そして、R&D 部門では広大なパラメータ空間の割には実験数が少ないので、蓄積データを使った分析や MI は研究を大幅に加速することもあるが、そうでないことも多々あり、期待しすぎるべきでないことも論じた。本項では、「そもそもデータ蓄積、共有は何のためにやるのか？」と題し、データ蓄積、共有の動機に関して、考察を行う。

まず、「なぜ、データを蓄積し、共有しようと思ったのか？」に戻って動機を見つめなおしてみよう。それは、恐らく、「自分の記憶の実験結果だけでは得られない有用な情報が、他者の実験や記憶から消えてしまっている自分の過去実験にあるのではないか？」、「そういう実験結果を簡単に確実に閲覧し、分析できれば、新しい傾向が見つかり、研究

が加速するのではないか？」といったものではないだろうか？双方、漠然と参照できるデータが増えれば、「研究にポジティブインパクトがあるのではないか？」とか、「機械学習などのMIが使えるようになり、それらから良いアイデアが提示されるのではないか？」と思っているのではないだろうか？それは、大きな認識間違いである。参照できるデータが増えれば、自分の考えや推測に合う都合な情報に見えるデータを選び出すことがより容易になり、それはある意味データを間違っただけで解釈してしまう可能性を増やすことにつながる。これは、認知心理学や社会心理学では、よく知られた現象「仮説や信念を検証する際にそれを支持する情報ばかりを集め、反証する情報を無視または集めようとしない傾向」で、確認バイアスと呼ばれており、人間の心理構造に深く根差した傾向なので、よほど意識的に注意（完全に主観を外して評価する）などをしない限り逃れられない傾向である。しかし、完全に主観を外した評価として機械学習などのMIに頼ってしまうと、4項で話をしたように必要とされる予測精度を確保できる適用領域を明確化すること自体が難しく、「良いアイデアが提示されるのではないか？」と思って、利用すると期待を裏切られ、利用者は強いMIアレルギーになってしまいかねないのである。

「有用な情報」とか、「良いアイデア」というのは、結果論であり、「他者の実験や記憶から消えてしまっている自分の過去実験」自身や「機械学習などのMIの予測値」自身は、常に「有用な情報」とか、「良いアイデア」であるわけではない。つまり、「他者の実験や記憶から消えてしまっている自分の過去実験」や「機械学習などのMIの予測値」が「有用な情報」とか、「良いアイデア」かを判定する装置のようなものがあれば、もちろん、データ共有は研究進展を確実に改善できるのであるが、そんな判定装置がない以上、そういうことに大きな期待を寄せてはいけないのである。それであれば、データ共有では研究進展を確実に改善できることはないのだろうか？

結論から言うと、データ共有で研究進展を確実に改善できることはある。それはデータ探索、データ処理に網羅性、均質性、再現性を付与することである。そして、まさしく、データベースの価値はここに見出すべきである。逆に言うと、研究過程において網羅性、均質性、再現性のないデータ探索、データ処理は、研究の品質を著しく損ねるので、行ってはいけないはずである。素粒子実験や天体観測などの大規模実験では、「網羅性、均質性、再現性のないデータ探索、データ処理」は厳しく禁止され、監査されている。それは「網羅性、均質性、再現性のない」データ探索、データ処理の結果から導き出された仮説から次の実験を行うと無駄実験になる可能性が高いからであり、大規模実験では無駄実験は金銭的にも相当大きい無駄になるからである。しかしながら、これは大規模実験にのみ適用されるべきことというわけではない、例えば、再実験が安価なものであったとしても本当に間違っただけでデータ分析から導かれた無駄実験をしてもいいのだろうか？もちろん、データ分析時点でのデータが少なく、結果的に無駄実験になってしまうのは致し方なしであるが、実際にはデータが十分にあったり、全く同じ実験をしているのにデータ抽出、分析方法を間違っただけで、そのことに気が付かず、無駄実験をすることはあってはならないはずである。小規模実験では、実際実験をする過程で、一番高価なものは研究者の時間のはずである。無駄実験で研究者の時間は相当無駄に消費されることになり、それがそもそも研究開発の進展速度を遅らせている大きな原因の一つのはずである。データベースにより網羅性、均質性、再現性を担保されたデータ探索、データ処理を使い、研究を実施することで、このような無駄実験はなくすることができるはずなのである。

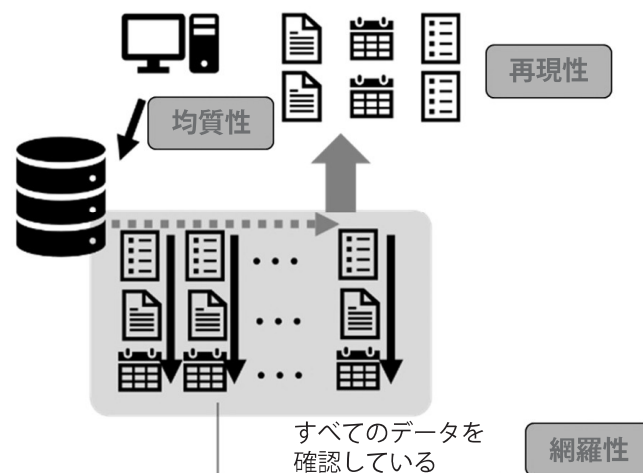


図 20 データベースは、網羅的、均質的、再現性を保証する

データベース化は、これ以外にも作業者負担軽減、作業時間短縮というメリットもある。手動処理でも網羅性、均質性、再現性を簡単に達成できるような業務ではこれが主たるメリットであるが、R&D 部門に限って言うと、手動処理（ばらばらに保存されているデータが記録されたファイルをフォルダを駆け回りながら探し、ファイルを開いて・・・）で網羅性、均質性、再現性を達成できることはまずないため、作業者負担軽減、作業時間短縮はあくまで副次的なメリットであり、データベース化の目的は、データ探査、データ処理に網羅性、均質性、再現性をもたらすことと考えるべきである。以下で、網羅性、均質性、再現性に関して、具体的にどのようなものか？研究にどのように影響してくるか？の解説を加えておく。

網羅性：手動で全データを対象に探査、処理を行うことは現実的ではないが、データベースだと全データを網羅した探査、処理が可能である。

例）熱伝導度が〇〇以上の材料を探したい場合に、記憶からいつぐらの実験にあったはずと目星をつけて、いくつかのデータを確認する。しかし、目星が間違っていることもあるし、そもそも「熱伝導度が〇〇以上の全ての材料」に目星が付けられるわけではない。このような網羅的でない抽出データは傾向が偏っている可能性もあり、それを分析した結論は、信頼性を大きく棄損している。

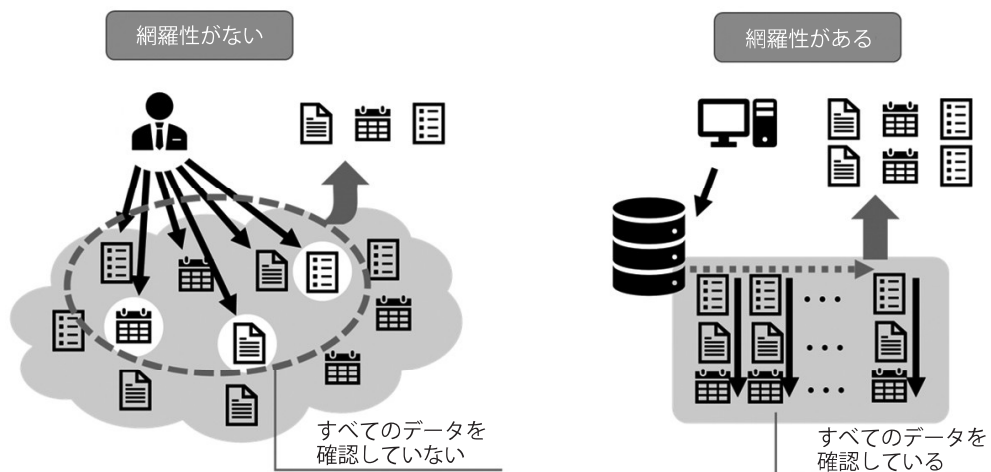


図 21 データベースの利点の網羅性とは

均質性：手動で探査、処理すると意識バイアスがかかったり、初期と終期での探査、処理の同一性が維持できないが、データベースだと完全に均質な探査、処理が可能である。

例）耐電圧が 500 V 以上の材料を探している場合に、いくつかのデータを見つけた後、あまりにも該当データが少なかったため、480 V 以上のものも含めるようにした。ただ、最初の方から探査をやり直していないので、いくつかのデータは取りこぼしている可能性がある。また、明らかに他の特性が合致しないデータは、500 V 以上でも抽出しないこともあるだろう。しかし、それは明示的なルールで一貫しているかといわれると自信をもって Yes と言い切れないのではないだろうか。このような均質でない抽出データを分析した結論は、信頼性を大きく棄損している。

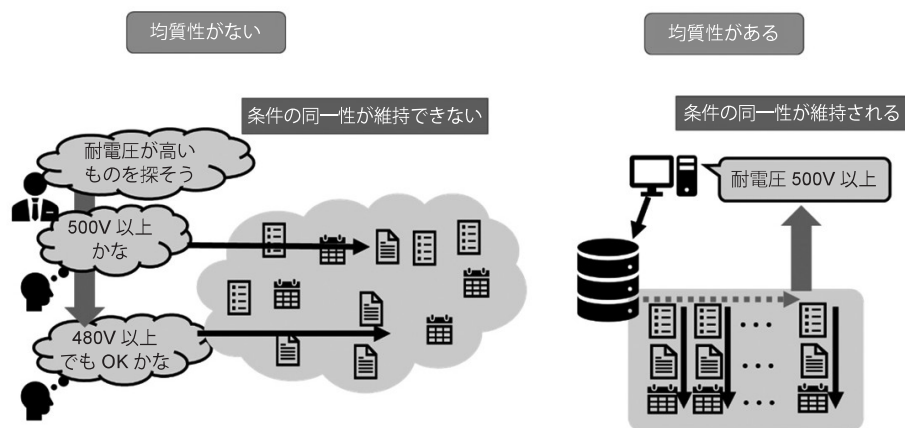


図 22 データベースの利点の均質性とは

再現性：半年前のデータ探査や処理は、手動ではまず再現できないが、データベースであれば、データ探査や処理の再現が可能である。

例) 手動処理では、上記で述べたように網羅性も均質性もほとんどの場合で担保されていないので、当然、再現性も期待できない。こういう再現性のない抽出データを分析した結論は、信頼性を大きく棄損している。

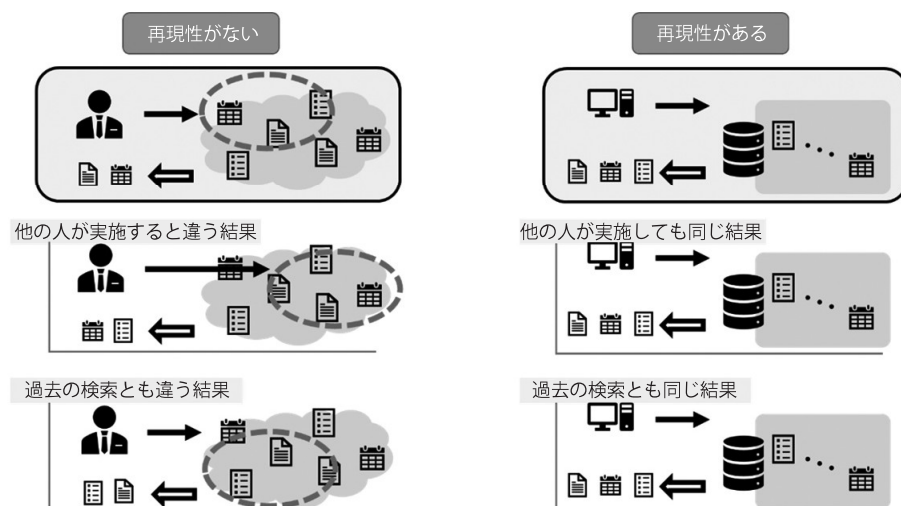


図 23 データベースの利点の再現性とは

9. 教科書などであまり触れられていない多変量データ分析の重要な注意点

前項では、「そもそもデータ蓄積、共有は何のために行うのか?」と題し、データ蓄積、共有は AI や機械学習などのために行うのではなく、データ探査、データ処理に網羅性、均質性、再現性を付与することで、研究の品質、客観性を担保するために必要であるということを説明した。本項では、「教科書などであまり触れられていない多変量データ分析の重要な注意点」と題し、多変量データ分析において、非常に重要にもかかわらず、データ分析の教科書にはあまり触れられていない注意点に関して、説明する。

世の中では、MI、AI が叫ばれているが、MI、AI はアイデアを提供してくれるという点では心強いが、「それを良いと思った理由」は提示してくれないので、データ分析、理解という観点では無力である。実際、MI、AI を活用するにしても、MI、AI が提示した案について、実データを分析し、その背景理由を探ることを辞めてしまっはいけない。そもそも、それを辞めるなら研究者はいらないということになる。背景理由を探るための最も初歩的なデータ分析方

法が義務教育時代から親しんだ X-Y プロットである。たかが、X-Y プロットであるが、されど X-Y プロットである。結局、様々な学術論文でも賑やかなしの 3D 可視化や画像などがあつたとしても、論文で一番重要な部分は X-Y プロットになっているのは、偶然ではない。

X-Y プロットは、「Y 軸項目値の X 軸項目値依存性」を表すもので、データ分析、理解というものの基本的な部分は、まさしくこれを使う必要がある。一般的に X 軸項目に実験パラメータを割り付け、Y 軸項目に実験結果項目を割り当てる。実は、注目している実験パラメータを X 軸項目に指定し、依存性を確認したい実験結果項目を Y 軸項目に指定しただけでは駄目なのである。前項でも少し触れたが、実際のデータ分析では、いくつかの注意が必要である。まず、どんな実験でも実験パラメータが一つということとはあり得ないはずで、注目している実験パラメータ以外にも実験パラメータがたくさんあるはずである。その実験パラメータを全く無視して、X-Y プロットを描いてはいけない。実際、そのようなことをすると「Y 軸項目値の X 軸項目値依存性」を見ているようで、表には出てきていない X 軸項目以外の実験パラメータ依存性を X 軸項目値依存性と誤認してしまう恐れがある。

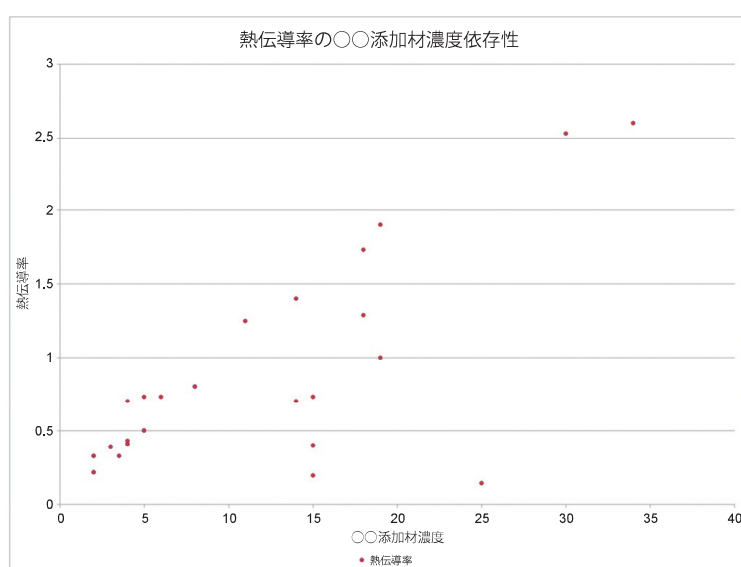


図 24 X 軸以外の実験パラメータが一定であるデータ毎にデータ分類されていない X-Y プロット

上記 X-Y プロットは、様々な実験結果に対して、注目している〇〇添加剤濃度を X 軸項目に指定し、依存性を確認したい熱伝導率を Y 軸項目に指定したもので、添加剤濃度を上げていけば、添加剤濃度 1%あたり、熱伝導率は 0.07 ~ 0.08 ずつ比例して上がっていく傾向にあることが読み取れてしまう。

実は、図 24 の X-Y プロットは問題なのである。「注目している実験パラメータを X 軸項目に指定し、依存性を確認したい実験結果項目を Y 軸項目に指定」する前に X 軸項目以外の実験パラメータが一定であるデータ毎にデータを分類しておく必要があるのである。多工程の実験では、実験パラメータは数百から千程度になるので、「X 軸項目以外の実験パラメータが一定であるデータ毎にデータを分類しておく」作業は、結構大変な作業になる。しかし、この作業を省略して、「Y 軸項目値の X 軸項目値依存性」をグラフ化してはいけない。検討したい X 軸項目以外の実験パラメータが確定している場合は、「X 軸項目以外の実験パラメータが一定であるデータだけにデータを絞り込む」ことは、ひたすら実験パラメータの数だけ Excel の filter をかけていけばいいだけなので、大変な作業にはなるができてしまう。しかし、そもそもその一定にしたい値がよくわかっていない場合は、何らかの方法で「X 軸項目以外の実験パラメータが一定であるデータ」毎に分類しておき、分析をしたい実験パラメータが一定であるデータ塊を選ぶ必要がある。

Excel で数百から数千の filter をかけることも大変だが、「X 軸項目以外の実験パラメータが一定であるデータ」毎に分類するのは、手動処理では永遠に終わらないほどの作業量になってしまう。例えば、Excel で 100 個の実験パラ

メータ項目を「実験パラメータが一定であるデータ毎に分類」しようとする、1 個目のパラメータ項目の値を filter で 1 個指定し、次のパラメータ項目も値を filter で 1 個指定しという作業を 100 項目すべてに行い、やっと実験パラメータが一定であるデータが 1 塊だけ分類できる。その後、最初のパラメータ項目の値を別の項目値に filter を変え、同じことを全項目の全項目値のパターンを抜けなく filter をかけていく必要がある。もし、項目値が 2 個ずつだとして、filter を 1 秒で行なえたとしても、全パターンの 2^{100} (2 の 100 乗) パターンを確認するには、 2^{100} 秒 = 2 の後に 0 が 30 個ぐらい付いた数 (秒) = 宇宙の寿命よりはるかに長くなってしまい、現実的ではない。実際には、何列か filter をかけると結果が 0 件 (0 行) になることがほとんどなので、ここまで時間はかからないが、たぶんやりきる人はいないはずである。したがって、蓄積データ全体でごく少数の実験パラメータしか変動させていない R&D 部門 (そんなところがあるとは思えないが…) 以外では、こういうことが簡単に実施できるツールを整備しておくことが必須ということになる。

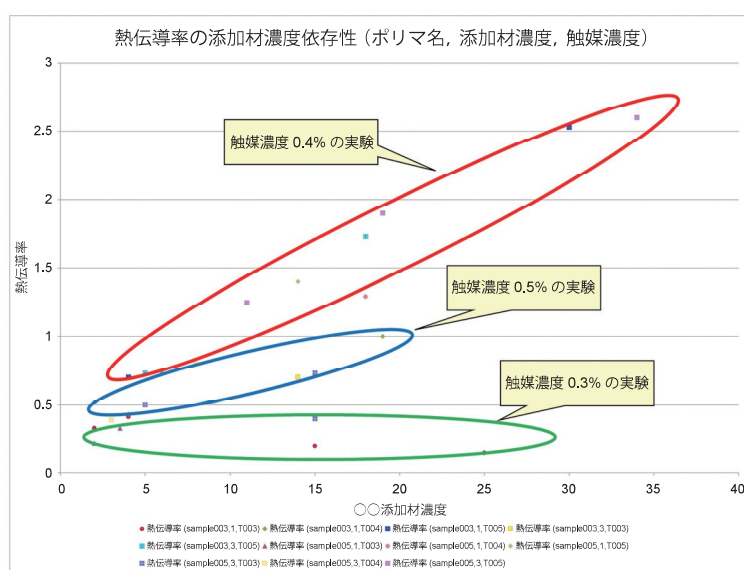


図 25 X 軸以外の実験パラメータが一定であるデータ毎にデータ分類された X-Y プロット

先ほどのデータは、X 軸項目以外の実験パラメータ値で分類をしないで、X-Y プロットを描いたが、図 25 では追加しているポリマー種類や触媒濃度などの X 軸以外の実験パラメータの値が同じデータ毎にデータを分類 (マーカー種を分け) し、データをプロットしたものである。そうすると「添加剤濃度 1%あたり、熱伝導率は 0.07 ~ 0.08 ずつ比例して上がっていく傾向がある」のは、触媒濃度を 0.4%にした時だけで、0.3%の場合は、そもそも添加剤濃度依存性がないことがわかる。X 軸以外の実験パラメータの値が同じデータ毎にデータを分類しないと、結論をミスリードしてしまいかねないのである。

実際、「X 軸以外の実験パラメータの値が同じデータ毎にデータを分類」をしてみると、分類毎にデータが 1 個ずつしかなくなるようなことがよくある。そうすると分類毎の傾向は全く見えてこない。当然、分類間の傾向の違いも見えるわけがない。だから、分類をやめて、図 24 のような X-Y プロットに戻して、傾向を把握しようとする人がいる。実は、それは本末転倒なのである。分類毎の傾向が見えず、分類間の傾向の違いも見えないのは、事実なので、それを受け入れるしかないのである。無理やり、正しくない処理をしてそこで傾向が見えたとして、それは意味があるわけではない。よく聞くのが、「そこまで厳密な傾向把握をしたいわけではないので、X 軸以外の実験パラメータの値が同じデータ毎にデータを分類するまでは不要だ」という言葉だが、「そこまで厳密な傾向把握をしたいわけではない」は認めたとしても、その後の「ので」には、論理的な正しさがない。つまり、そこまで厳密な傾向把握をしたいわけではない」を認めたとしても、だからといって「X 軸以外の実験パラメータの値が同じデータ毎にデータを分類するまでは不要」という結論は導き出せないのである。結論が導き出せないのに、今できる方法として、「X 軸以外の実験パラメータの値が同じデータ毎にデータを分類しない」グラフ化を「厳密ではないがそれなりに正しい

傾向把握」として認識してしまっているのである。実際には、「厳密ではないがそれなりに正しい傾向把握」をより具体的に定義して、それを満たすようなグラフを作成し、そのグラフから傾向を把握すべきなのである。「厳密ではないがそれなりに正しい傾向把握」をより具体的に定義することができない、難しい場合は、そこで立ち止まって、傾向把握を諦めるべきである。逆に言うと、今までのデータでは傾向が把握できないので、傾向把握をするために追加実験の計画を立てるべきなのである。「傾向把握をするための追加実験計画」は、究極的には一因子実験しかないのである。例えば、便利な手法と思われている実験計画法などは、実は以下のような非常に厳しい前提条件が課せられているが、その詳細を確認せず、使われていることが多い。一部であるが、重要な条件を列挙しておく。

- あ) 因子を束ねる（エイリアス）ことにより、実質的な変数を 10 個程度にする
- い) エリアス化する因子は、目的変数に対して依存性がほとんどない
- う) エリアス化する因子を含んだ交互作用も目的変数に対して依存性がほとんどない
- え) すべての条件で誤差分散が同じ大きさ
- お) 誤差が正規分布に従う

さて、話を少し戻す。実は、データ蓄積、共有を行うまでは、「X 軸項目以外の実験パラメータ値で分類する」作業は発生しないか、したとしても大きな負担ではない。データ蓄積、共有を行うまでは、自分が依存性を調べたいと思う実験パラメータ数種だけを変動させて、それ以外の実験パラメータは固定した実験を行うはずであり、また、実験直後にデータ分析を行うはずである。その場合、ほとんどの実験パラメータは固定されており、また、自身も何を固定したという記憶が残っているため、ここで議論している「X 軸の項目以外の実験パラメータが一定であるデータ」を分類することは簡単である。実際、X 軸の項目以外の実験パラメータを一定にして実験をするので、そもそも分類する必要がないことがほとんどである。つまり、こんなグラフを描きたいと思ってから、そのグラフが簡単に描けるような実験をしているということになる。

しかし、データ蓄積、共有をし、遥か過去の自分のデータや他研究者のデータを含めて分析をする場合、「どの実験パラメータが固定されているのか？」は、実データを確認するまで分からなく、また、殆どの場合は、実験パラメータの大多数が変動してしまっているはずである。実は、データ蓄積、共有をしない場合は、「実験前にグラフが描きやすいように実験パラメータを固定しているか」、少し前の実験であれば自身の記憶を活用することで「どの実験パラメータが固定されているのか」が分かるので、比較的楽にデータ分析ができるだけなのである。データ蓄積、共有を始めると自身の記憶を使わずにデータ分析をすることが強いられ、その分析の大きな変化を研究者が認識できず、当惑してしまうのである。データ蓄積、共有をする前に、自身の記憶を活用せず、実データをしっかり確認しながらデータ分析をする習慣をつけていくことが重要なのである。

こういう話をするとき、「折角データベースがあるのであれば、自分が確認したい実験パラメータ（それこそ、X 軸以外の殆どの実験パラメータ）を検索条件に指定し、その結果を X-Y プロットするのであれば、データ分類（検索条件で既に分類されている）もいらなくて、簡単にグラフが描けるのではないか？」という人もいる。そうすれば確かにデータ分類はいらなくなるが、本当にそれで、データ蓄積、共有をしたメリットが出るだろうか？実は検索条件でデータ分類が不要なぐらいに絞り込みをすると自分の直近のデータしか抽出されないことになるのである。X 軸以外の実験パラメータすべてが同じ値になる実験なんて、過去の自分でもしていないはずである。実は、データ蓄積、共有をする本懐は、自分が実施していない、意識していない実験パラメータの実験結果と自分が意識して実施した実験結果を並べ、そこから広いパラメータ空間での関連性を発見することなのである。

自分は添加剤濃度依存性を確認したいと思っていて実験をしていた時に、過去の経験から触媒濃度を 0.3 で十分と判断し、触媒濃度を固定して実験をしていたとする。自分が確認したい実験パラメータ（それこそ、X 軸以外の殆どの実験パラメータ：触媒濃度は当然 0.3）を検索条件に指定し、その結果を X-Y プロットすると、熱伝導率は添加剤濃度に依存しないという結論になってしまう。データ蓄積、共有をするときには、検索条件はあまり厳しく設定せず、自分が意識できていない実験パラメータでの実験結果も含めて傾向を把握しようという態度が重要なのである。これは、今までのデータ分析の傾向の見方である局所依存性とは反対の大域構造の把握であり、頭の考え方もコペルニク

ス的転回が必要ということになる。

さて、データ分析を行うときに、X 軸に実験パラメータでなく、特性値を設定することもあるはずである。例えば、試作物の粘度と熱伝導度の関係性を知りたい場合は、X 軸項目に粘度を設定し、Y 軸項目に熱伝導度を設定して、X-Y プロットを描くはずである。ただ、この時にも幾つか気を付けなければならないことがある。特性値は実験パラメータのように項目値が実験パラメータに対して一対一の関係のある項目でない為、「粘度が〇〇の時に熱伝導度が××」というプロットがあっても、あらゆる実験でそれが成り立つかと言えば違う。実験パラメータが違って「粘度が〇〇」になることは、十二分にあり得て、その時、「熱伝導度が××」にならないことも十二分にあり得るのである。だからと言って、「粘度が〇〇」になるあらゆる実験パラメータを実験で確かめ、「粘度が〇〇」になる実験パラメータのすべてで、「熱伝導度が××」になることを調べるのは、現実的でない。もちろん、こういうことは、皆は承知で、いくつかの異なる実験パラメータの結果で、この手のグラフを描いているはずだが、どんな実験パラメータでもその傾向が同じかどうかはしっかりと確かめられていないことが多いはずである。つまり、たまたま、偏ったパラメータの実験での限定的な依存性（偽依存性）を一般的な依存性と勘違いしてしまう可能性が高いのである。したがって、この手のグラフを描く時には、事前に実験パラメータと当該 X 軸、Y 軸に設定する予定の特性値の相関を事前に分析しておき、どのパラメータ領域での粘度と熱伝導度の関係性を評価しようとしているかを明確化しておく必要があるのである。

ところで、実験計画法や統計解析のツールで、全ての 2 変数の組み合わせの散布図（X-Y プロット）を描き、2 次的に並べたものを見たことが無いだろうか？これは、散布図行列と呼ばれ、変数の相関関係を視覚的に捉えることができると説明されている。例えば、変数が身長、体重、座高、握力だとすると低段：身長－握力、体重－握力、座高－握力、中段：身長－座高、体重－座高、高段：身長－体重の三角形上に散布図（X-Y プロット）が並ぶ。

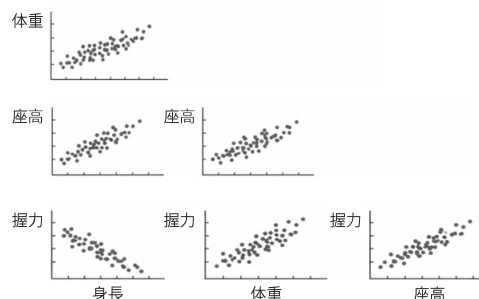


図 26 散布図行列の例

例を見てわかる通り、散布図行列は社会科学的なデータ分析によく使われる。つまり、変数がすべて結果変数なのである（社会科学的なデータは、実験室実験のように実験側が制御して実験を行うような実験条件変数はない）。そういう意味で、前段落で説明した X 軸、Y 軸に特性値（結果変数）を割り付けた X-Y プロットを全ての 2 変数の組み合わせで書いたものである。したがって、前段落で指摘した注意点は同じように確認しておく必要があるのである。しかし、散布図行列の説明部分では、こういうことを書いていないことが多い。それは、社会科学的なデータは、実験室実験のように実験側が制御して実験を行うような実験条件変数はない（もしくは本当はあっても実験側が制御できないし、記録もできていない）ので、前段落で指摘した注意点を守ろうにもそれができないためである。その代わり、社会科学的なデータは、労力がかかるがデータを増やすことは可能なので、分析結果が変わらなくなる程度まで十分膨大なデータを集めれば、「たまたま、偏ったデータでの限定的な依存性（偽依存性）」の可能性を排除できる。しかし、実験室実験では、分析結果が変わらなくなる程度まで十分膨大なデータを集めるということは、膨大なデータが必要ということになるため現実的ではないのである。つまり、散布図行列から何らかの意味を見出す場合は、書籍やツールの説明書には書いてはいないが、前段落の注意点は意識しないといけないということである。

また、R&D 部門で、X 軸、Y 軸に特性値（結果変数）だけでなく、実験パラメータ（実験条件変数）を含め散布図行列を書いている例もよく見る。X 軸が実験パラメータで、Y 軸が特性値の X-Y プロットは、それこそ本項で説

明したように「X 軸項目以外の実験パラメータが一定であるデータ」毎に分類し、マーカー種別を変えるなどの処置が絶対に必要であり、それをしないで何らかの意味を見出すことは間違った傾向把握することになってしまう。X 軸が実験パラメータで、Y 軸も実験パラメータの X-Y プロットは、実験結果が全く含まれない X-Y プロットなので、実はどう実験したかという現在までの実験計画のパターンが見えるだけであって、実験パラメータ同士の適切な関係性が見えているわけではない。つまり、この X-Y プロットからは、「こういう実験パターンは行われていないから今後は、こういう実験パターンも行っておこう」という実験パターンの粗密傾向から粗密をなくすような指針を見出すのはいいのだが、「この変数がこの値の場合は、あの変数はいつもこの値になっているので、今後もそう設定しよう」というような実験の傾向指針は導き出せないことを理解しておく必要がある。

10. 実験パラメータ項目は、どのような独立変数系にするかが重要

前項では、「教科書などであまり触れられていない多変量データ分析の重要な注意点」に関して、説明を行った。本項では、「実験パラメータ項目は、どのような独立変数系にするかが重要」と題し、実験パラメータ項目に関しては、単なる単独の項目だけを見て、項目名を決めるのではなく、全体を俯瞰し、データ分析のためにはどのような独立変数系が適切かを考える必要があるということを論じる。

具体的な実験パラメータ項目を想像しやすいように、図 27 のような仮想的な実験を考える。ここでは、液体としてモノマー a、モノマー b、触媒 c、気体としてモノマー c、モノマー d、水素 h を温度、圧力が一定の容器に密閉し、容器内でバブリングすることで気液混合をし、合成反応を行う実験を仮定している。

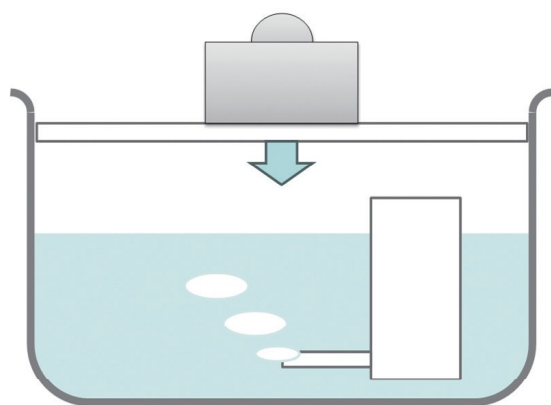


図 27 仮想的な実験環境のイメージ

この時の実験パラメータは、実験の計量、設定を念頭に置き、前項のルールを加味すると、以下のような項目名が自然である。ちなみに、この実験パラメータの独立変数の数は 8 個になる。

m_a_液体モノマー a 使用量 (g)

m_b_液体モノマー b 使用量 (g)

m_c_気体モノマー c 使用量 (g)

m_d_気体モノマー d 使用量 (g)

m_h_水素ガス使用量 (g)

m_s_液体触媒 s 使用量 (g)

P_容器内圧力 (MPa)

T_容器内温度 (°C)

ここで、実験結果として得られる合成物分子量が、容器内温度に対してどのような依存性を持つのかを蓄積データから調べたいとする。X 軸に T_容器内温度 (°C) を設定し、Y 軸に合成物分子量を設定した X-Y プロットを描けば、依存性は分かるはずである。

実は、この言明は半分は合っているが、半分は間違っている。合成物分子量の容器内温度依存性を明らかにしようとし、今まさに実験をする場合は、たぶん、容器内温度以外の実験パラメータ（各種使用量と圧力）を固定して、容器内温度だけを変えた実験をいくつか行うはずである。したがって、上記 X-Y プロットを描けば、容器内温度依存性は分かるはずである。しかし、過去に様々な実験パラメータで実施し、蓄積されているデータでは、容器内温度以外の実験パラメータもバラバラなはずであり、そのまま X-Y プロットを描いても、容器内温度依存性は分からないはずである。・・・と言うか、グラフを描いてしまうと、依存性らしきものが見えることもあり、容器内温度依存性を読み取ってしまうことの方が危険である。依存性を見て取るには、図 28 のように X 軸以外の実験パラメータの値が同じデータごとに分類し、それぞれのデータごとに X-Y プロットを描く必要があるのである。



図 28 合成物分子量の容器内温度依存性を示す X-Y プロット

さて、図 28 を見て、何かが分かるだろうか？実は、蓄積データを使ったデータ分析で最初に躓くのは、ここなのである。せっかく項目名を揃えてデータ蓄積をしても、このように意味の分からない大量のグラフが生成され、そこからどのように検討を進めればいいのか分からなくなる。そして、分からないので、機械学習や AI を勉強して・・・というのがよく聞く流れである。ここで、強調しておくが、X-Y プロットで読み取れないことが機械学習や AI で分かるようになると思うのは危険で、それであれば分析のための研究者は不要で、実験を精緻に実施できる技術者がいればいいだけということになる。X-Y プロットでの読み取り補助のために機械学習や AI を使うべきなのである。社会科学のようにそもそも対象が相当複雑で、実験パラメータを振った人為的実験ができない分野で、データが大量にある場合は、X-Y プロットでの読み取り自体も主観的にならざるを得ないので、機械学習や AI に主観排除としての大きな意味がある。しかし、実験パラメータを任意に設定し、他の環境要因を排除できる実験が可能な工学領域の R&D では、データ分析に客観的意味、解釈が十分付与できるので、X-Y プロットでの読み取りを省略することはあり得ない。話が脱線してしまったので話を少し戻す。図 28 から傾向が把握できないのは、実は X 軸以外の実験パラメータの項目がデータ分析に適していないからなのである。

図 28 の実験パラメータ項目 (= 独立変数) は、原材料 6 個に対して、全て「○○使用量 (g)」となっている。当然この実験では、何らかの化学反応が起こるはずなのであるが、そもそもその基礎となる化学反応式は、絶対量に対する式ではない。化学反応を決めているのは、絶対量ではなく比率である。したがって、実験パラメータ項目は、以下のように比率で定義した方がデータ分析は行い易いはずである。もちろん、ナノ粒子反応や界面反応を考慮すると絶対量に依存する化学反応もある。しかし、今、注目しているところが絶対量に依存する化学反応でないのであれば、比率を項目にした方が分析はし易いはずである。変数を変換しているだけなので、比率の方が分析がし易いだけであっ

て、ナノ反応や界面反応を無視した分析になってしまうわけではないことを付け加えておく。

$$\begin{aligned}
 \text{モノマー b 液体部数} &= m_b / m_a \\
 \text{触媒 s 液体部数} &= m_s / m_a \\
 \text{モノマー d 気体部数} &= m_d / m_c \\
 \text{水素ガス h 気体部数} &= m_h / m_c \\
 \text{総重量 (g)} = m_total &= m_a + m_b + m_s + m_c + m_d + m_h \\
 \text{気体液体重量比} = r_g_l &= (m_c + m_d + m_h) / (m_a + m_b + m_s) \\
 P_容器内圧力 \text{ (MPa)} & \\
 T_容器内温度 \text{ (}^\circ\text{C)} &
 \end{aligned}$$

実際、 m_total 総重量 (g) や r_g_l 気体液体重量比が十分大きなところでは、 m_total 総重量 (g) や r_g_l 気体液体重量比の合成物分子量に対する依存性はなくなるはずである。 m_total 総重量 (g) が十分大きなところでは、表面反応やナノ粒子反応は無視でき、 r_g_l 気体液体重量比が十分大きなところでは気体は液体に対して飽和濃度に達し、液体中反応は変化しないと考えられるからである。そういう物理傾向を捉えた変数系を実験パラメータ項目にしておくと、データ抽出、分類、分析が容易になる。先ほどの変数系で、X 軸に総重量を設定し、X 軸以外の実験パラメータの値が同じデータごとに分類し、それぞれのデータごとに X-Y プロットを描くと図 29 のようになり、ごく一部の実験を除いて、総重量依存性がないことが確かめられる (はずである)。

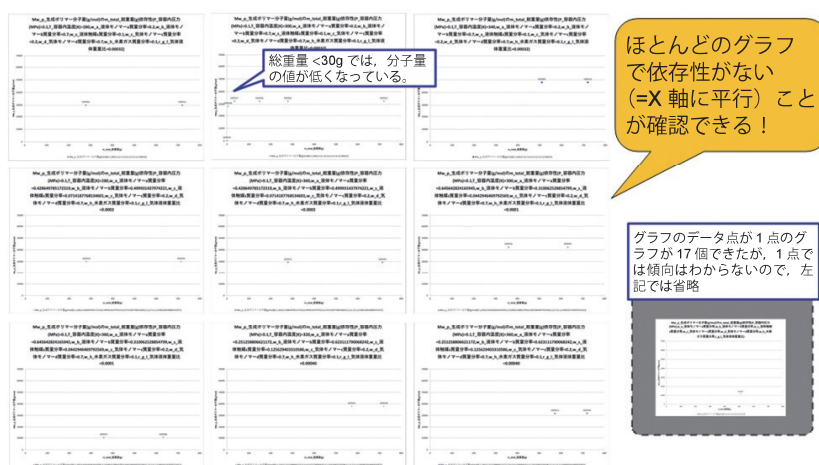


図 29 合成物分子量の総重量依存性を示す X-Y プロット

さらに、X 軸を気体液体重量比に変え、X 軸以外の実験パラメータの値が同じデータごとに分類し、それぞれのデータごとに X-Y プロットを描くと図 30 のようになり、一部の実験を除いて、気体液体重量比依存性がないことが確かめられる (はずである)。



図 30 合成物分子量の気体液体重量比率依存性を示す X-Y プロット

ここまでで確かめられた依存性がない部分のデータだけで、X 軸を温度、X 軸以外の実験パラメータ（依存性がない部分なので、総重量と気体液体重量比を除く）の値が同じデータごとに分類し、それぞれのデータごとに X-Y プロットを描くと図 31 のようになる。気体、液体の材料の比率毎の合成物分子量の容器内温度依存性がよく分かる形になっていることが分かると思う。

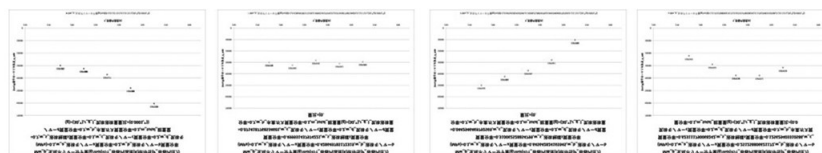


図 31 総重量、気体液体重量比依存性がないデータでの合成物分子量の容器内温度依存性を示す X-Y プロット

「X-Y プロットを描いて人が目視で傾向を確認するから変数軸をそこまで気にしないといけい訳ではないのではないか？」と思っている人もいると思う。実際、IT 系や AI 系ツールの営業担当は、そういうことを明言する人も多い。しかし、事実は違っている。m_a_液体モノマー a 使用量 (g), m_b_液体モノマー b 使用量 (g) の変数系で機械学習やベイズ推定、実験計画法などでデータ分析を行うのと、使用量比、総重量、気体液体重量比などの変数系で機械学習やベイズ推定、実験計画法などでデータ分析を行うのでは、推定精度を同じ程度に高めるのにデータ数が最低でも数十倍から数百倍必要になってくるのである。無限と思われるぐらい大量のデータ数がある場合は、変数系は意識しなくてもよいのだが、R&D 部門のように変数に比して、データ数が十分多くない（変数数の数十倍以下）の場合は、変数系により推定精度が大きく変動してしまうのである。

これは、水素原子の電子波動関数をシュレディンガー方程式で解くときを思い出してみればよい。平面波展開で解くのと球面調和関数展開で解くことを思い出してみれば、理解できると思う。平面波展開で解く場合と、球面調和関数展開で解く場合とで、解の収束性が大きく異なることは、大学の授業でも周知の通りである。実際、水素原子の場合は球面調和関数展開が良いが、結晶系であれば平面波展開の方が良い。平面波展開や球面調和関数展開というのは、まさしく変数系である。つまり、対象により適切な変数系は異なるのである。系の対称性や変数の平等性から適切な変数系は決まるのである。実験の変数に関しても同じように系の対称性や変数の平等性から適切な変数系を決める必要があるということである。

IT 系や AI 系ツールは、マーケティングや証券データなどデータが時間とともにほぼ無料で増え続ける領域が主戦場であり、そこでは無限と思われるぐらい大量のデータ数があるので、このような変数系は意識しなくてもいいのである。また、IT 系や AI 系ツールの開発者や営業担当は「R&D 部門が変数に比して、データ数が十分多くない」という事実を知らないの、注意喚起もできないのである（まあ、知っていても売れなくなったら困るので、そういうマイナス情報は言わないと思うが・・・）。

「どのような独立変数系にするかをあまり気にしないでデータ分析をしても、大きな問題は起こっていない」という感覚を持っている人はいないだろうか？実は、その感覚自体は間違っていない。だからといって、どのような独立変数系にするかに関しては、そこまで細かく考えて決める必要はないと判断を下すのは早計であり、むしろ危険である。皆がなぜ、「どのような独立変数系にするかをあまり気にしないでデータ分析をしても、大きな問題は起こっていない」と感じ、実際、問題が起こっていないのかに関して、以下で理由を説明する。

前項でも触れたが、合成物分子量の容器内温度依存性を明らかにしようとし、今まさに実験をする場合は、「たぶん、容器内温度以外の実験パラメータを固定して、容器内温度だけを変えた実験を行うので、全てのデータで X 軸以外の実験パラメータが固定されている」のである。もちろん、X 軸以外の実験パラメータが 1 パターンだけでなく、数パターンある場合もあるだろうが、それぞれで分析すればいいだけである。ここで、注目すべきなのは、「重量が一定なら重量比は必ず一定になる」が、「重量比が一定だからと言って重量は必ずしも一定ではない」という点である。自分で実験をするときは、重量を一定にして実験をすると重量比も必ず一定になるので、データ分析の観点では、重量でも重量比でもどちらでも良いのである。しかし、蓄積されたデータを使って分析をするときは、変数系が重量では問題である。蓄積されたデータ全ての重量が一定に揃ったデータではないので、データ分析をするためには、様々な重量パターンから、同じ重量比のデータを分類しなおす必要がある。全て 1 g の実験も全て 10 g の実験も比が同じなので、「総重量依存以外では分析上は区別したくない」ということである。

したがって、変数系は重量でなく、重量比である必要がある。つまり、データ蓄積、共有をして初めて変数系の問題が噴出するのである。そして、これがデータ蓄積をする前に、変数系の問題が認識されなく、実際に蓄積し出してから実験者が蓄積データを有効に利用できない原因の一つである。

最後に、蓄積データ分析及び蓄積する実験自体に関して、いくつかアドバイスをしておく。蓄積されたデータを分析する場合は、今回分析をしたように通常、依存性がないと思われる実験パラメータに関しても、依存性が認められる領域の実験もあることを前提に分析をすべきである。実は、このような分析を行うことで依存性が認められる領域が見つかるということだけでなく、外れ値から失敗実験（依存性がないと思われるところで依存性がある）も検知でき、今まで何となく外れ点として除去していた実験をしかりした理由で、除去することができる。

また、新規実験や追加実験を行う場合、総重量、気体液体重量比などの依存性がない領域がどの辺りかがある程度確かめた後は、それら依存性のない（総重量、気体液体重量比等）実験パラメータは固定して実験をすることで、効率的かつ分析精度の高い実験が可能になることも覚えておくとよい。

最後に少しだけ、付け加えておく。実は適切な変数系というのを見出すのは結構骨の折れる仕事であり、以下の変数系を思いつくためには結構深い洞察力が必要である。

$$\begin{aligned} \text{モノマー b 液体部数} &= m_b / m_a \\ \text{触媒 s 液体部数} &= m_s / m_a \\ \text{モノマー d 気体部数} &= m_d / m_c \\ \text{水素ガス h 気体部数} &= m_h / m_c \\ \text{総重量 (g) = } m_{\text{total}} &= m_a + m_b + m_s + m_c + m_d + m_h \\ \text{気体液体重量比 = } r_{g/l} &= (m_c + m_d + m_h) / (m_a + m_b + m_s) \end{aligned}$$

あまり深く考えない場合は、以下のような変数系にしてしまうのではないだろうか？

$$\begin{aligned} \text{モノマー a 液体重量\%} &= 100 \times m_a / (m_a + m_b + m_s + m_c + m_d + m_h) \\ \text{モノマー b 液体重量\%} &= 100 \times m_b / (m_a + m_b + m_s + m_c + m_d + m_h) \\ \text{触媒 s 液体重量\%} &= 100 \times m_s / (m_a + m_b + m_s + m_c + m_d + m_h) \\ \text{モノマー c 気体重量\%} &= 100 \times m_c / (m_a + m_b + m_s + m_c + m_d + m_h) \\ \text{モノマー d 気体重量\%} &= 100 \times m_d / (m_a + m_b + m_s + m_c + m_d + m_h) \\ \text{水素ガス h 気体重量\%} &= 100 \times m_h / (m_a + m_b + m_s + m_c + m_d + m_h) \end{aligned}$$

まず、この変数系は変数変換としては間違っている。変数の数は確かに変換前後で同数（6 個）だが、重量%はすべての変数の和は、100 になるので、独立変数ではなくなっている。この変数系では、独立変数として、総重量が必

要なのである。

$$\text{総重量 (g)} = m_{\text{total}} = m_{\text{a}} + m_{\text{b}} + m_{\text{s}} + m_{\text{c}} + m_{\text{d}} + m_{\text{h}}$$

まあ、この点は気づいている人の方が多いと思うが、問題はここ先である。例えば、合成物分子量のモノマー b 液体重量%依存性は、どのようになるかを調べたい場合はどうすればよいのだろうか？X 軸にモノマー b 液体重量%を設定し、Y 軸に合成物分子量を設定した X-Y プロットを描けばいいのではないかな？もちろん、9 項で説明した X 軸以外の実験パラメータの値で分類をして……。本当にそれだけで、大丈夫だろうか？X 軸以外の実験パラメータは何になるのだろうか？モノマー a 液体重量%, 触媒 s 液体重量%, モノマー c 気体重量%, モノマー d 気体重量%, 水素ガス h 気体重量%, 総重量 (g) だろうか？

モノマー a 液体重量%, 触媒 s 液体重量%, モノマー c 気体重量%, モノマー d 気体重量%, 水素ガス h 気体重量%が決まってしまうと、モノマー b 液体重量%の値も決まってしまう。つまり、それではモノマー b 液体重量%依存性なんて確認できないのである。それは、すべての重量%変数の和は 100 になるからそんなことになるだけで、何か一つ変数を削ればいだけだと思っている人も多いのではないだろうか？例えば、モノマー a 液体重量%を削ってみよう。そうすれば、モノマー b 液体重量%依存性のグラフは確かに描ける。しかし、モノマー b 液体重量%が大きくなるとモノマー a 液体重量%は小さくなる。このグラフは、純粋にモノマー b 液体重量%依存性を表しているのだろうか？

少し頭を冷やして、モノマー b 液体重量%依存性というのは、どういうものを見たいと思っているのかを考えてみる。まず、「モノマー b 液体重量%以外の実験パラメータは固定で依存性を見る必要がある」という言葉を少し深く考察してみよう。「モノマー b 液体重量%以外の実験パラメータが固定」といった時点では、「モノマー a 液体重量%, 触媒 s 液体重量%, モノマー c 気体重量%, モノマー d 気体重量%, 水素ガス h 気体重量%のそれぞれが固定」ということを明確に意識しているだろうか？これらを固定するとモノマー b 液体重量%も 1 つの値しかとらなくなってしまうので、少し考えるとそうではないことに気づくはずである。問題はここでも、「少し考えない」と気づけないというのが厄介ではある。

こういう場合は、抽象的に物事を考えるとどうあるべきかが見えにくくなってしまうので、少し具体的な例で考えてみると良い。重量%が問題をややこしくしている原因のようなので、まず、重量で考えてみよう。モノマー b 液体重量依存性を確認する一因子実験を考えてみると、モノマー a 液体重量, 触媒 s 液体重量, モノマー c 気体重量, モノマー d 気体重量, 水素ガス h 気体重量を一定にし、モノマー b 液体重量だけ増減させる実験を行うのではないだろうか？ここで少し抽象度をあげてみる「モノマー a 液体重量, 触媒 s 液体重量, モノマー c 気体重量, モノマー d 気体重量, 水素ガス h 気体重量を一定にし」は、本当にそれぞれの重量を一定にすることに実験としてそこまでの意味があるだろうか？実際、一定にしたいのは、「モノマー a 液体重量, 触媒 s 液体重量, モノマー c 気体重量, モノマー d 気体重量, 水素ガス h 気体重量の比率」ではないだろうか？実は、モノマー b 液体重量依存性を確認する一因子実験で一定にしたい、すべきものは、モノマー a 液体重量：触媒 s 液体重量：モノマー c 気体重量：モノマー d 気体重量：水素ガス h 気体重量の連比なのである。連比が一定というのは、触媒 s 液体重量／モノマー a 液体重量, モノマー c 気体重量／モノマー a 液体重量, モノマー d 気体重量／モノマー a 液体重量, 水素ガス h 気体重量／モノマー a 液体重量が一定ということと等価である。そうするとモノマー b 液体重量依存性というのも、モノマー b 液体重量／モノマー a 液体重量依存性ということになる。つまり、モノマー b 液体重量%依存性を確認したいということは、深く考えると「モノマー b 液体重量／モノマー a 液体重量依存性を確認したい。その時、モノマー b 液体重量／モノマー a 液体重量以外の変数：触媒 s 液体重量／モノマー a 液体重量, モノマー c 気体重量／モノマー a 液体重量, モノマー d 気体重量／モノマー a 液体重量, 水素ガス h 気体重量／モノマー a 液体重量及び全重量を一定にして」ということなのである。

重量であろうと重量%であろうと、比は同じ値になることや重量%というのは、なんとなく比率という感覚でとらえているのが、混乱のもう一つの原因なのである。同じ比率でもモノマー a 液体重量%, 触媒 s 液体重量%, モノマー c 気体重量%, モノマー d 気体重量%, 水素ガス h 気体重量%が一定ということと、モノマー a 液体重量%：触媒 s 液体重量%：モノマー c 気体重量%：モノマー d 気体重量%：水素ガス h 気体重量%が一定というのは、意味が違う

のだが、相当意識しないとその違いに気づかないのである。

結局何が言いたいかというと、重量%や濃度というのは日常でも良く使われる単語で、日常で使う場合には同じ単語を場面、場面で少しずつ違う定義で使っていることがあるのである。実はよく考えるとモノマー b 液体重量%依存性という言葉自体、実際にはその時頭で意図しているものではないのである。頭にあるのは、モノマー b 液体重量相対比率依存性で、モノマー a 液体重量：モノマー b 液体重量：触媒 s 液体重量：モノマー c 気体重量：モノマー d 気体重量：水素ガス h 気体重量：モノマー b 液体重量の連比で、モノマー b 液体重量の比率だけを変化させた場合を意識しているということなのである。そして、この連比を個々の比率にすると本項の最初の方でデータ分析に使った

$$\text{モノマー b 液体部数} = m_b / m_a$$

$$\text{触媒 s 液体部数} = m_s / m_a$$

$$\text{モノマー d 気体部数} = m_d / m_c$$

$$\text{水素ガス h 気体部数} = m_h / m_c$$

のようになるのである。

おわりに

本節では、R&D 部門における機械学習・AI・生成 AI 活用の可能性と限界について説明をし、それらを活用するために必要なデータ蓄積、共有に関して、実情と改善方法を示した。また、蓄積データの分析は、従来の自分が行ったデータの分析と大きく異なる点があることを実例ともに紹介し、機械学習・AI を使う場合には、そのような従来型の分析も行っていく必要があることを説いた。読んでいただいてお分かりだと思うが、R&D のデータ共有を行なうということは、何かを買って、解決するという問題ではなく、企業風土及び R&D 従事者の意識変革が必要である。R&D 従事者は、業務内容が極めて専門的で、会社内でも対話する機会が少ない人達だと思う。研究者は、とっつきにくいところはあるが、地頭は良く、非常にまじめなので、これを機に話をしてみれば、仲良くなり、本節で話したようなことを一緒に議論できるかもしれない。そういう人とのつながりを作ることが、R&D 部門におけるデータ共有の第一歩である。

文 献

- 1) 川田重夫, 田子精男, 梅谷征雄, 南多善, 上島豊, 他 PSE book— シミュレーション科学における問題解決のための環境 (応用編) 川田重夫, 田子精男, 梅谷征雄, 南多善 共編, 培風館, (2005), p69-82
- 2) 谷啓二, 奥田洋司, 福井義成, 上島豊 ペタフロップスコンピューティング, 矢川元基 監修, 培風館, (2007), p183-202
- 3) 牧野圭一, 上島豊, 視覚とマンガ表現, 臨川書店, (2007), p1-5, 221-229
- 4) 上島豊, 月刊「研究開発リーダー」7月号, 技術情報協会, (2024), p77-84
- 5) 上島豊, 月刊「研究開発リーダー」9月号, 技術情報協会, (2024), p73-81
- 6) 上島豊, 月刊「研究開発リーダー」11月号, 技術情報協会, (2024), p85-96
- 7) 上島豊, 月刊「研究開発リーダー」1月号, 技術情報協会, (2025), p74-82
- 8) 上島豊, 月刊「研究開発リーダー」3月号, 技術情報協会, (2025), p68-73
- 9) 上島豊, 月刊「研究開発リーダー」9月号, 技術情報協会, (2025), p89-97
- 10) 上島豊, 月刊「研究開発リーダー」11月号, 技術情報協会, (2025), p90-98
- 11) 上島豊, 月刊「研究開発リーダー」1月号, 技術情報協会, (2026), p88-94
- 12) 上島豊, 月刊「研究開発リーダー」3月号, 技術情報協会, (2026), p87-94
- 13) 上島豊, 他 研究開発部門への DX 導入による R&D の効率化, 実験の短縮化, 技術情報協会, (2022), p195-221
- 14) 上島豊, 他 ケムインフォマティクスにおけるデータ収集の最適化と解析手法, 技術情報協会, (2023), p39-74

- 15) 上島豊, 他 実験の自動化・自律化による R&D の効率化と運用方法, 技術情報協会, (2023), p159-199
- 16) 上島豊, 他 少ないデータによる AI・機械学習の進め方と精度向上, 説明可能な AI 開発, 技術情報協会, (2024), p112-127