◆新連載◆

《隔月連載 全 5 回》　第 1 回

# R & D 部門における機械学習・AI・生成 AI 活用へのデータ共有の重要性

上島 豊　（株）キャトルアイ・サイエンス　代表取締役

《PROFILE》

略歴：

| | |
|---|---|
| 1992 年 3 月 | 大阪大学工学部　原子力工学科　卒業 |
| 1997 年 3 月 | 大阪大学大学院工学研究科　電磁エネルギー工学専攻　博士課程修了 |
| 1997 年 4 月 | 日本原子力研究所　博士研究員 |
| 2000 年 4 月 | 日本原子力研究所　研究職員 |
| 2006 年 3 月 | 日本原子力研究開発機構（旧日本原子力研究所）　退職 |
| 2006 年 4 月 | キャトルアイ・サイエンス設立　代表取締役　就任 |

主な参加国家プロジェクト：

文部科学省 e-Japan プロジェクト「ITBL プロジェクト」,「バイオグリッドプロジェクト」
総務省 JGN プロジェクト「JGN を使った遠隔分散環境構築」
文部科学省リーディングプロジェクト「生体細胞機能シミュレーション」

主な受賞歴：

| | |
|---|---|
| 1999 年 6 月 | 日本原子力研究所　有功賞<br>「高並列計算機を用いたギガ粒子シミュレーションコードの開発」 |
| 2003 年 4 月 | 第 7 回サイエンス展示・実験ショーアイデアコンテスト文部科学大臣賞「光速の世界へご招待」 |
| 2004 年 12 月 | 第 1 回理研ベンチマークコンテスト　無差別部門　優勝 |

主な著作：

培風館「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（基礎編）」ISBN：456301558X
培風館「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）」ISBN：4563015598
培風館『ペタフロップス コンピューティング』ISBN978-4-563-01571-8
臨川書店『視覚とマンガ表現』ISBN978-4-653-04012-5

## 1　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，客観的かつトレーサブルであるべきである。実際，センサーや計算機の能力向上により，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え，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 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。本記事では，私自身の 10 年の R & D 経験と弊社の 19 年の R & D 支援実績から得た「R & D 部門における機械学習・AI・生成 AI 活用へのデータ共有の重要性」に関して，簡単に解説する。

## 2　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 と言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○○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！と言われているのですが・・・」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。本章では，「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，後章で説明することにし，まず，本章では機械学習，AI に関して説明する。本連載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（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。しかし，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，雑誌のこの記事部分に付箋をつけたものを机の上において，この記事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 $Y = ax1 + bx2 + c$ や $Y = ax1^2 + bx1 + cx2^2 + dx2 + e$ 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 など差が生まれるのである。（そういうものがなければ扱えるデータが同じなら世の中の機械学習，AI は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 $Y = ax1 + bx2 + c$ や $Y = ax1^2 + bx1 + cx2^2 + dx2 + e$）ように実験者が認識している変数（x，x1，x2 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しまっている。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。

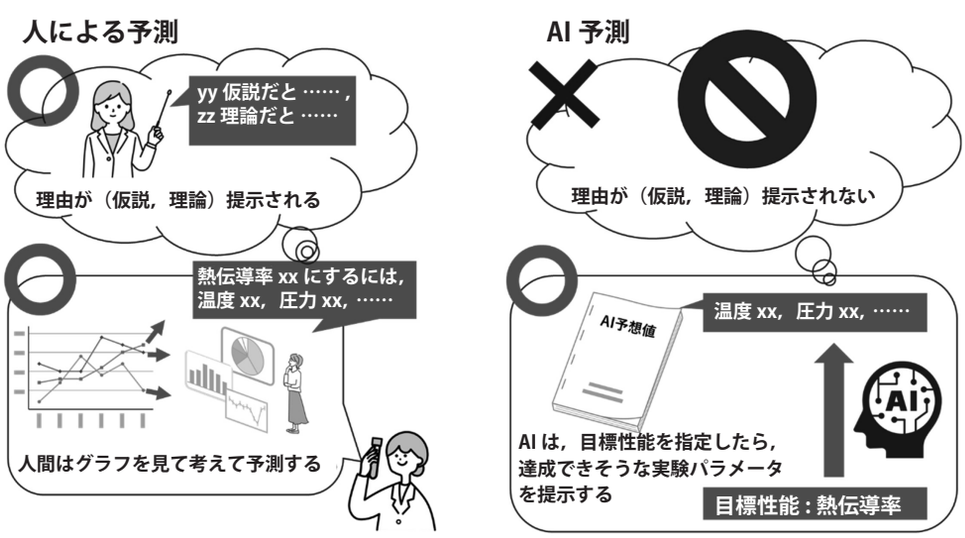



図 1　機械学習，AI は予測値の理由を提示するのが苦手

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わけである。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良く取りつ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。

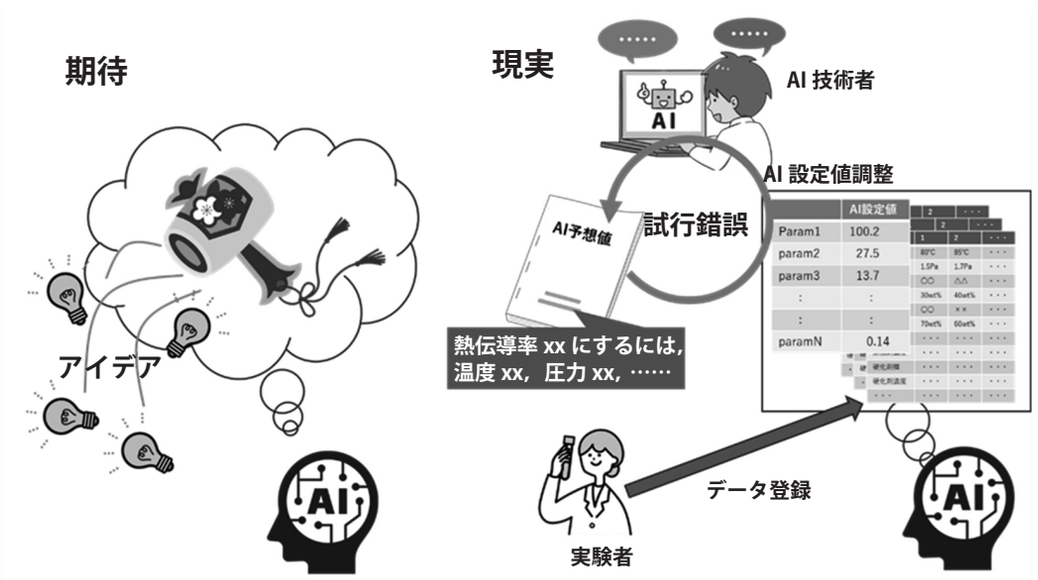



図 2　機械学習，AI は打ち出の小槌ではない

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らない。

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での実験は，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）をするのかが決められなく，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のである。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。

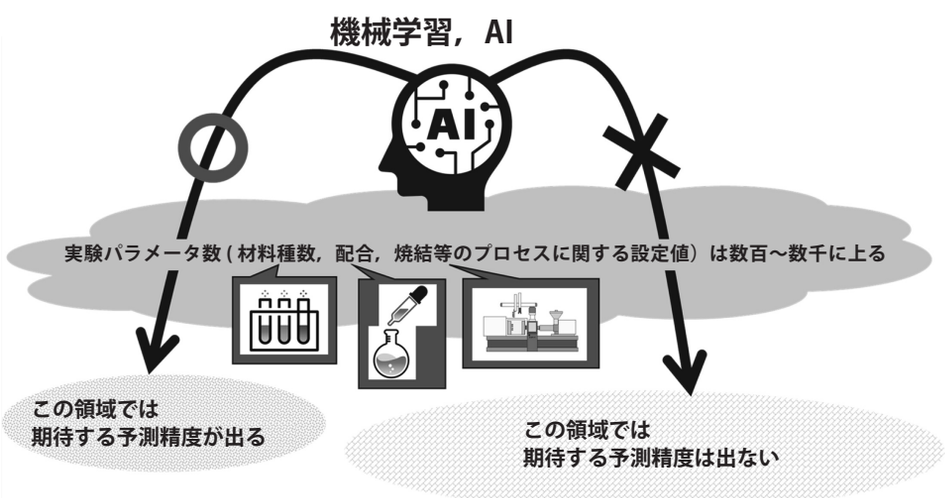



図 3　機械学習，AI は期待する精度の出る領域の明確化が難しい

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，「その範囲では求める特性の材料はない」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。

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（その範囲では求める特性の材料はない）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 2 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はいないが，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。

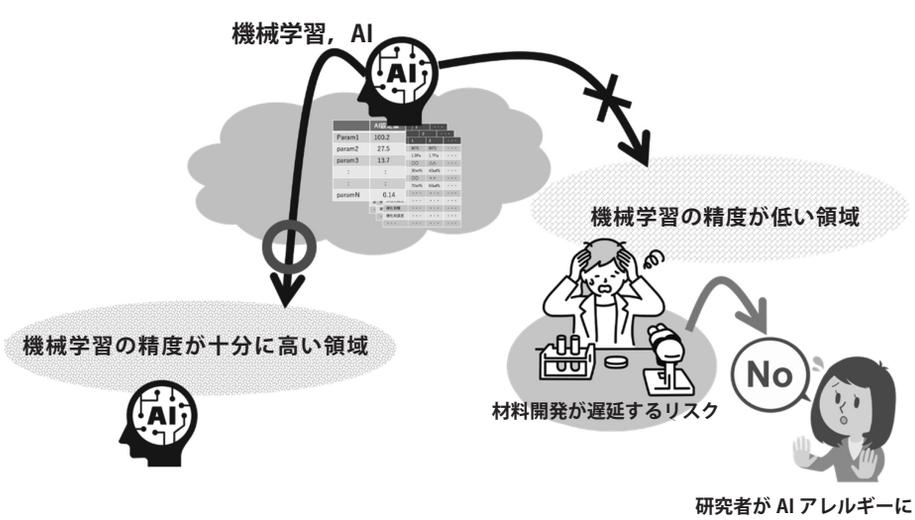



図 4　機械学習，AI の精度の低い領域を明確化しておかないと研究者が AI アレルギーに

## 3 R & D 部門において生成 AI に期待して良いこと，悪いこと

前章では，「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。本章では，「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，前章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，本記事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），本章では，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，2 章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いけないということである。また，単語の繋がりで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，ツンケンされていては会話も弾まなく，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ので，明示的に「○○であっていますか？」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。これらを避けるためには，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は絶対に必要である。

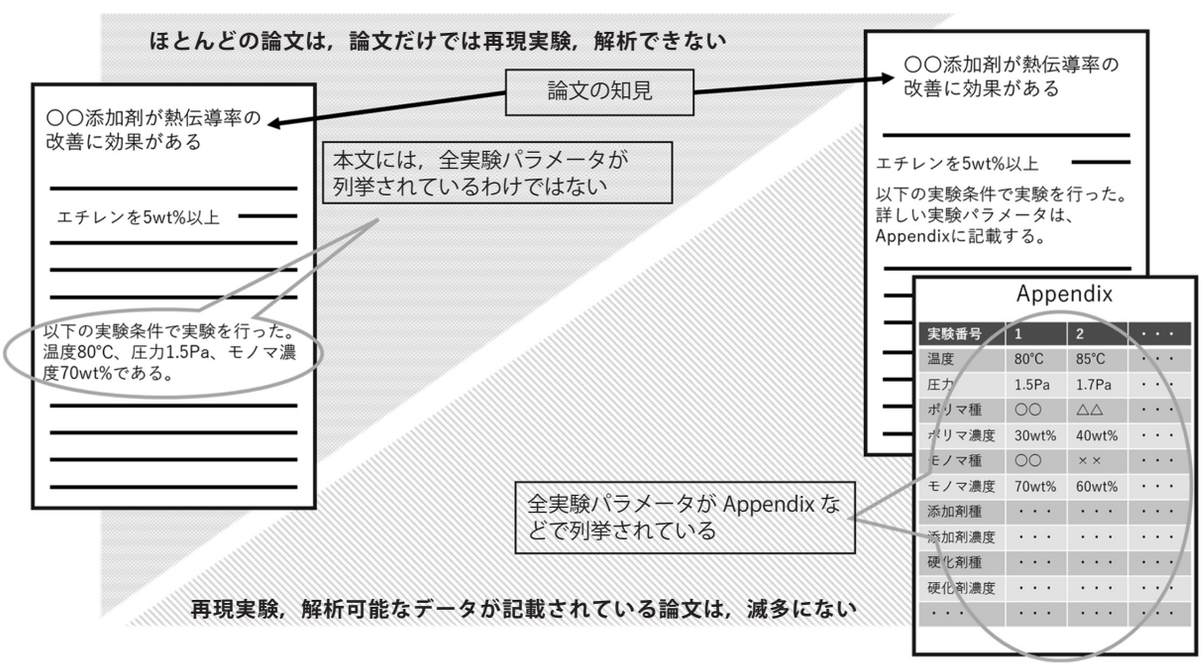
R & D 部門において質問したいことは，「どのような添加剤を加えれば特性改善が進むか？」などではないだろうか？ R & D 部門では，ネット上に転がっている当たり前の情報というより，自社の報告書や論文を基に質問の回答が欲しいという類のものであると思われる。2 章に書いているような文章は，一般論で，具体的な研究の内容に立ち入っていないのでオープンな生成 AI でも作成できるが，「自社の報告書や論文を基にした回答」は，オープンな生成 AI では適切な回答にならないはずである。もし，オープンな生成 AI からの回答が適切っぽい回答になっている場合は，それこそ，その回答の基となっている元資料を確認すべきである。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。オープン AI 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，Open Source として公開されている学習モデル（BERT,LLaMA,GPT-3.5 Turbo など）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をした。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，2 章で議論した機械学習や AI なんていらないのではないか？」というものがある。本気でそう思った人は，もう一度 2 章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いことである。

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。




| 実験番号 | 1 | 2 | ・・・ |
|---|---|---|---|
| 温度 | 80℃ | 85℃ | ・・・ |
| 圧力 | 1.5Pa | 1.7Pa | ・・・ |
| ポリマ種 | 〇〇 | △△ | ・・・ |
| ポリマ濃度 | 30wt% | 40wt% | ・・・ |
| モノマ種 | 〇〇 | ×× | ・・・ |
| モノマ濃度 | 70wt% | 60wt% | ・・・ |
| 添加剤種 | ・・・ | ・・・ | ・・・ |
| 添加剤濃度 | ・・・ | ・・・ | ・・・ |
| 硬化剤種 | ・・・ | ・・・ | ・・・ |
| 硬化剤濃度 | ・・・ | ・・・ | ・・・ |
| ・・・ | ・・・ | ・・・ | ・・・ |

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のは，いくつかの実験から何らかの法則性を導き出したことを報告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（ただし，本来は報告書や論文でもその報告内容に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，２章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ると思う。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，「○○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，「○○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。

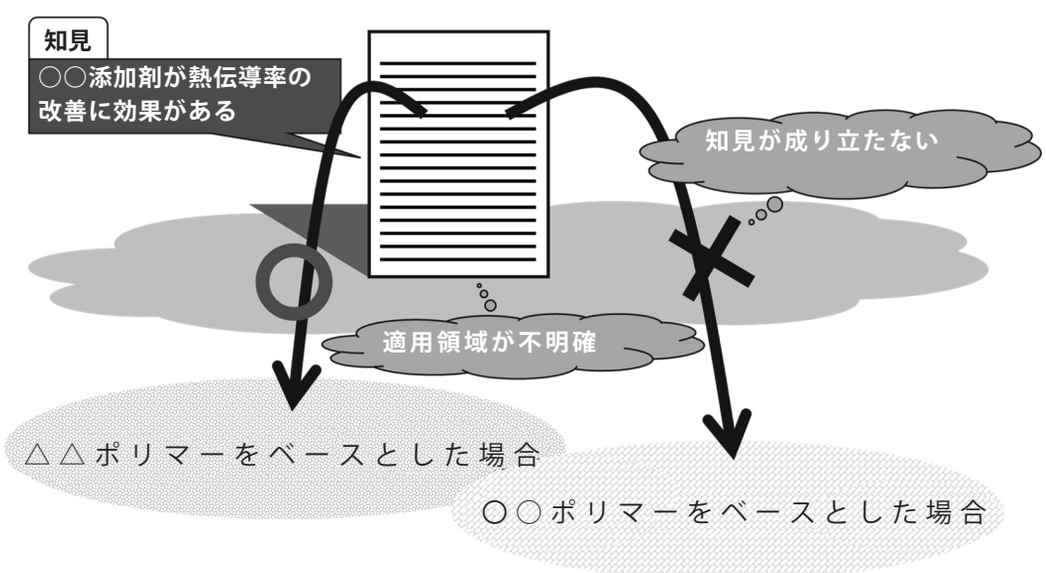



図 6　報告書に書かれた知見や法則性は適用領域が不明確なものが多い

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。どちらかというと，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い報告書を読む人は稀で，どちらかというと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」とういうことで，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。

## 参考文献


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」6 月号，技術情報協会，（2023），p63-68

5）上島豊，月刊「研究開発リーダー」7 月号，技術情報協会，（2023），p86-91

6）上島豊，月刊「研究開発リーダー」8 月号，技術情報協会，（2023），p78-82

7）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

8）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

9）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

10）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

11）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

12）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

13）上島豊，他
ケムインフォマティッ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

14）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

15）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