

生成AI時代の「気づき」と「直感」

—俯瞰解析の結果を生成 AI で料理する—

"Awareness" and "Intuition" in the Age of Generative AI
Upgrading the result from Panoramic View Analysis with generative AI



VALUENEX 株式会社 代表取締役社長 CEO

中村 達生

早稲田大学大学院理工学研究科を修了後、三菱総合研究所入社、途中、東京大学工学部助手に就任。2005年に工学博士取得。2006年に株式会社創知（現 VALUENEX）を設立、代表取締役社長 CEO に就任（現任）。2014年2月米国シリコンバレーに当社現地法人を設立、President に就任（現任）。2018年、当社東証マザーズに IPO。現在、早大大学院理工学術院非常勤講師も兼務。2018年度特許情報普及活動功労者受章。2019年スタンフォード大学より Japan-US Innovation Award 企業に選定。

✉ nakamura@valuenex.com ☎ 03-6902-9833

要約

生成 AI が出現し、私たちの生活や業務おける活動の中から、検索や整理するという行為の機械依存が進み、さらには思考するという人類を特徴づける行為までも放棄してしまう危惧が見えてきている。しかし、現在のノイマン型コンピューターが、大量の情報の整理や要約はできても、創造を行うには、次の世代のコンピューターが実用化されるまでの数十年の時間が必要とされる。それまでの期間を私たちは、生成 AI から得られた結果に対して、気づきと直感で判断するハイブリッド利用で賢く生き抜くことが求められる。

ない。情報の陳腐化、検索サービスの罅、生成 AI の限界、そして、まだ人類に優位性のある「気づき」と「直感」を発揮することの意義について詳述することが本稿の目的である。

2 身の回りにおける情報爆発は加速している

ゼタバイト級の情報の総量はさらに増加を続け加速しているため、情報を収集してデータベース化し、そのデータプールから必要な情報を検索し、人間が読むという利用形態、すなわち、従来のデータベース事業における単純な検索サービスが、新しいスタイルのものに変わりつつある。ChatGPT に代表される生成系 AI は、巨大な言語空間を抱えるデータベースが肝であり、従来のデータベースの進化系であるが、人間が読み、整理する作業を、推論エンジンを組み込み自動化した対話型のツールである。まだ学習過程の段階であり、専門的な話題に対しては必ずしも精度は高くないが、臨界点を通過すると、不足する知識を推論エンジンが補い、データベースに蓄積されていない情報についても、回答することができるようになるだろう。ChatGPT がときおり言い訳とも受け取れる回答をするのは、背後に蓄積している知識が不足しているために、必要十分な推論が機能しないという結論に至るからだろう。

そもそも、情報は新規性、偏在性、集積性の掛け算で価値が決まる。新規性とは情報の新しさ、偏在性とは、どこにでも普遍的に存在するのではなく、アクセスしにくい特定の分野や特定の仕組みの中に存在することを示

1 はじめに

情報爆発は 1960 年代に提唱されたムーアの法則に従って着実に進行している。そして、現在インターネット上に溢れる情報量は 50 ゼタバイトを超えたと言われている¹。その半分はデバイス間の通信であり、人間の介在しないコミュニケーションが急増している。生活や仕事空間の様々なデバイスを活用する場面において、何かしらのアルゴリズムが機能し、無意識に私たちの把握していない方向に誘導される世界が待っている。AI の進化による利便性の向上、それにともなう、サービスの高度化と加速が、私たちをますます機械依存の深みにはめてゆく。しかし、私たちは思考することを止めてはいけ

1 2020年の時点で、インターネット上にある情報の総量は 50 ゼタバイトであると推測されている。1 ゼタバイト (ZB) = 10 億テラバイト (TB) = 1 兆ギガバイト (GB)

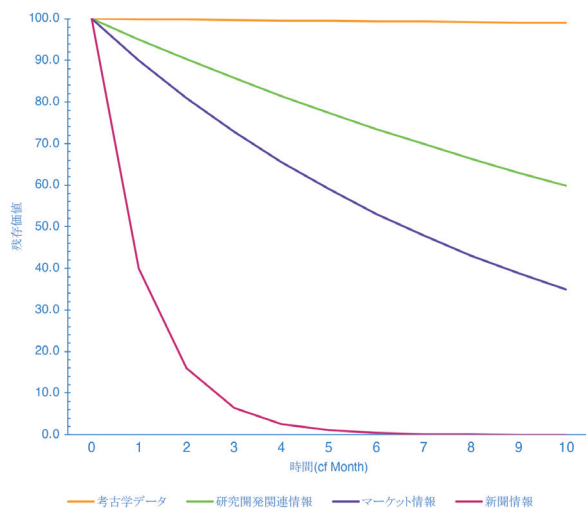
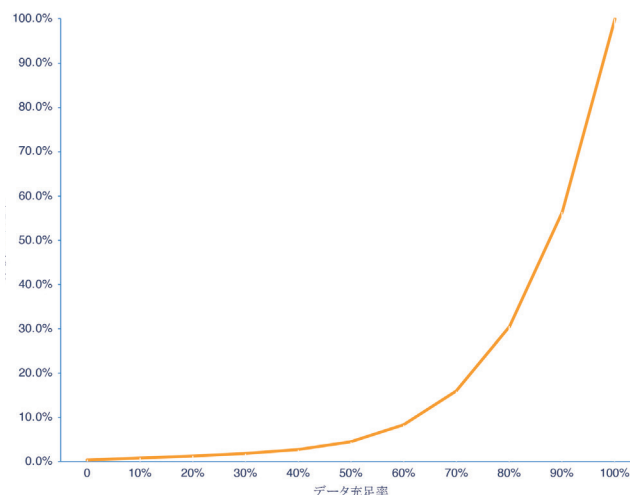


図1 情報の種類と陳腐化率 (イメージ)

図2 情報充足率とコスト (イメージ)²

し、集積性とは、その情報が異常値としてたまたま存在するのではなく、確度高く存在していることを指す。例えば、耐熱性と靱性を兼ね備えるナノオーダーの炭素繊維をとある機関の1研究チームが発明し、繰り返しそれを再現できて、その実験データをどのネットワークにもつながっていないスタンドアローンのPCの中に保存していることを想定する。おそらく世界初という新規性と、そのチームだけがアクセスできるPC上にある偏在性と、繰り返し再現できたというデータの集積性が成立し、情報としての価値は最大化している。

しかし、情報は物理的な財とは異なり、ひとたび、世の中と繋がり、拡散されると、情報価値そのものが勢いよく陳腐化する。多くの人を知っている状態になると新鮮さが失われ、誰でも簡単に詳細を知ることができることができってしまうからである。さらに、類似の実験結果や、それを超える結果や反証結果が出てくると、情報の確度そのものが瓦解する。情報の陳腐化は、時間と共に進行するが、情報の種類により拡散量が異なるため、陳腐化する速度、すなわち陳腐化率が異なる(図1)。考古学のように情報ソースが限られている分野では、情報そのものが貴重であり、偏在性が高く、情報価値の時間変化は小さい。ただし、代替される新説が現れると、価値に揺らぎが生じて、集積性が失われ、元の情報の価値が下がることになる。企業の中で行われる研究・開発に用いられる情報は、概ね数年から10年程度の残存価値がある。商品・サービスの市場における有用期間に適合した開発を行うためであり、創薬分野や社会貢献を目した基礎研究を行っている企業はこの限りではない。□□

ミ情報など、マーケット情報は、現在のトレンドを表しており、数ヶ月から年単位で、陳腐化してゆく。なお、季節に依存する商品・サービスは、一方向の陳腐化とはならず、周期性のある変化となる。陳腐化率がもっとも大きい情報ソースのひとつは新聞情報である。数日から週単位の情報残存価値と考えられ、数日のうちに、最新の情報に置き換わる。データベースを用いたビジネスでは、レアな情報をいかに多く集めるかが収益の源泉であるかのように語られることがあるが、情報を完全に収集し、データベースの価値を高める効用に対して、情報充足率を100%に近づけようとするほど、コストが急増する(図2)。価値ある情報は、新鮮であり、キーワードなどのきっかけが少なく、偏在していてアクセスがしづらく、同質の情報を十分に集めないと確度を保つことができないからである。その上、情報取得にかかる時間に対して、情報が発生する時間の方が短い場合は、永遠に充足率は高まらない。従来型のデータベースの従量課金型の多くのデータベース事業者が、データの管理コストが逼迫して利益がでない構造の理由がここにある。データベース事業者は、蓄積されたデータを利用した高付加価値サービスの提供が望ましく、ユーザーは一つのデータソースだけではなく、異なる他のデータと

2 情報をもれなく拡充しようとする(充足率を高める)、拡充のコストが急増する。陳腐化する速度に対して、充足率を100%に近づけて高めようすると、コストが急増する。データベース事業を営むには、データの陳腐化率を勘案し、適度な充足率を受容し、極端なコスト増を抑制する必要がある。その分、異なる情報ソースとの組み合わせにより、結論を導出するデータフュージョンの考えが大切である。

の組み合わせにより、解を推論する手法を身につける必要がある。データの付加価値化の一つの手法が俯瞰的な可視化でありユーザーに気づきを、複数のデータソースの俯瞰を組み合わせるデータフュージョンは、さらなる直感をユーザーにもたらすこととなる。

3 情報技術 (IT) の進化と一般化により仕事の価値が変化している

情報爆発の発端は情報技術の進化であるが、私たちの生活や仕事の情報依存度がかつてよりも高まっていることが情報量の増加に拍車をかけている。1980年代のMS-DOS³の時代は1.44Mbの5.25inchフロッピーディスクにOSが入りaドライブに挿入し、bドライブにはデータをいれるための別のフロッピーディスクを挿入して利用していた。すなわち、ひとつのバッチ処理は、1.44Mbの範囲で完結させ、それを超える場合は、フロッピーディスクを抜き差しして、さらに少ないメモリー空間に保持させて利用する必要があった。必然的に、データサイズをいかに少なくし、最大のアウトプットを出すかが課題となり、アルゴリズム研究がもっとも進化した時代でもあった。まもなくして、1990年になる直前にハードディスクの価格が急速に下がり、PCを保有する人たちにもハードディスクが普及しはじめ、個人でもデータ蓄積が身近になり、データベースの一般化が起きた。1990年代は主に電子化の時代であった。あ

らゆる紙で所蔵されていた情報がデジタル化されて、複製が容易にできるようになった(図3)。これにより、貴重な年鑑や名簿などの偏在性が損なわれ、これらの貴重な書籍の所有をもとにビジネスをしていた人たちが存在意義を失っていった。2000年代に入るとITバブルの勃興とともに、Yahoo、Infoseek、Googleといった検索エンジンが一般に知れ渡ることとなり、検索を生業とする人たちの職域が急速に狭まった。2010年代はPCのみならずスマートフォンをはじめとするモバイル端末が急速に普及したことによって、あらゆる情報がICTでつながる時代へと突入した。これにともない地球上から偏在がほぼなくなり、地球の裏側の特定地域の明日の天気など、ちょっとした解析結果はすぐに手元に表示できるようになった。90年代から2010年代までの情報技術の進化と一般化により、それぞれの時代を経るごとに、私たちが情報システムから得られる価値が、データ、情報、知識へと昇華⁴してきた。そして、いま統計手法と確率論をベースにした推論エンジンを携えたAIが一般に出現し世間を賑わしている。2020年代は機械が私たちに「気づき」を与える時代になると考えられる。多種多様な情報ソースから確率的に推論を働かせて、さまざまな見えない壁の向こう側を気づかせてくれるサービスが出現しつつある。ただし、現在の生成型AIはノイマン型コンピューターを用いて、大規模・大容量計算をしているに過ぎず、かつてよりも計算速度が

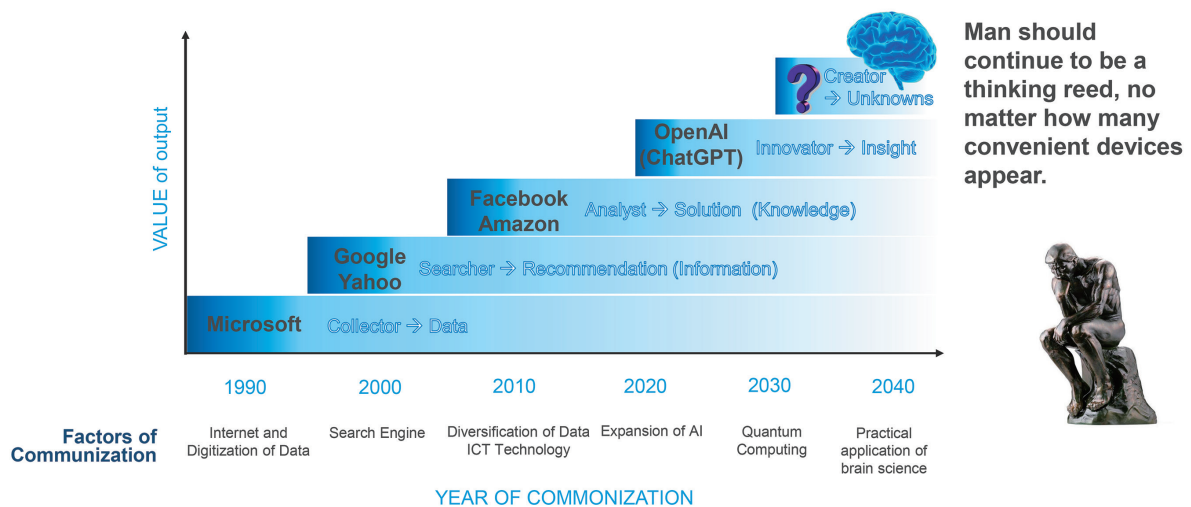


図3 情報技術の進化と一般化に伴う仕事の価値の変化

3 Microsoft Disk Operating System の略

4 アルビン・トフラーは「富の未来」講談社(2006)の中で、データ、情報、知識への昇華がなされることを予言している。

高速になったとは言っても逐次処理しかできないため、計算リソースと有限な時間内のサービスを考慮すると、適用範囲には限界があることがわかっている。かつてのフロッピーディスク時代のアルゴリズム開発のように、制約条件の中でいかに近似解を見つけるかがサービスの質と成功を決すると考えられる。一方、2020年代の後半になると、量子コンピューターの最初の実用化サービスが次々と市場に投入され、身の回りにあるさまざまな問題の最適解を機械が教えてくれる時代になるだろう。すなわち、少なくない分野において、気づきを機械が出してくれることになる。しかし、その気づかせてもらった内容に基づいて、何をどのようにするべきかの判断は、人間の直感が決めることになる。気づきというもの、何かしらの制約や思い込みで、アクセスできなかった範囲にある情報であるのに対して、直感はその人の生活や職場での経験が無意識のうちに表出する結論だとすれば、直感的判断を機械が行うには、さらに情報技術の進化が必要であり、2040年頃の脳科学とコンピューターサイエンスの融合を待つ必要がある。しかし、それが実現すると、ゼロ・トゥ・ワンの創造を機械ができるようになる。発明や創作活動も機械が行うようになると、人間は何をするべきだろうか。筆者はあくまでも人間は考え続ける輩であるべきだと考える。利他の心をもった人がいないと、世の中は前に進まないからだ。

4 Google 検索は広告表示に最適化されている

ホモ・サピエンスがアフリカ大陸に現れたという20万年前から、人類の脳の容量はほとんど変化していない。藪の中に潜む猛獣から自分自身と仲間を守るために、コミュニケーション手段と言語が発達してきた。少ない情報から推論を働かせ、正しい判断を行うことで生き延びてきた。

時が経ち西暦2000年前後の検索エンジンの一般化に始まり、直近のたった20年間における情報技術の進化は、私たちの身の回りを、信号、テキスト、画像、動画などの情報で埋め尽くしてしまい、スマートフォンやタブレットを使って、食事中も移動中も受動的に情報にさらされ続けている。情報量が多くなればなるほど、思考や推論を行わなくなり、機械が出す情報の真偽検証が行われず、鵜呑みにする姿が当たり前になりつつある。

とくに日常的に利用している検索サービスが広告ビジネスモデルであることの影響を考慮して、出力された検索結果のバイアスを取り除いてから受け入れるプロセスを実践している人は、そう多くはないのではないだろうか。Googleの検索は広告ビジネスで成り立っているため、当然のことながらSEO⁵対策が施された広告主の情報が検索結果の上位に入ってくる。検索時の多くの場合は、上位の情報のみを目を通し、リストの下位やページを何枚もめくっての情報収集は稀であろう。また、入

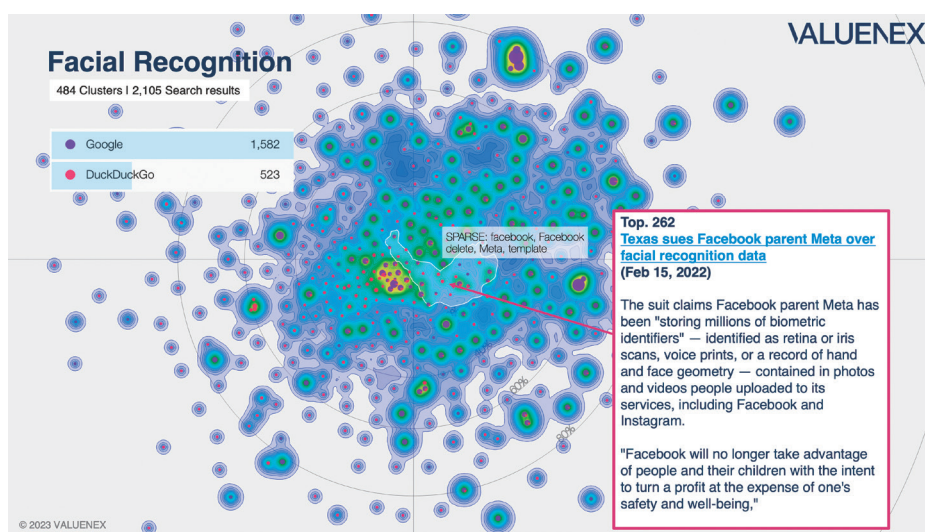


図4 GoogleとDuckDuckGoの検索結果の違い

5 SEOとは、検索エンジン最適化(Search Engine Optimization)のことであり、検索リストの上位に現れるように対策をすること。



力した質問キーワードに対して、類似するページがリスト化される仕組みでは、掲載漏れしている情報は見つけられない。検索エンジンのボットがアクセスしていない特殊な分野の新しいサイトの情報は、クローリングされていないかもしれない。しかし、最新であり、偏在であり、集積している情報にこそ価値があり、広告ビジネスに適用されている検索エンジンが出力する検索結果は、広告効果を最大化するため、プロフィールが適合する多くの人に関連する情報が上位に上がる仕組みとなっている。すなわち、情報価値としては高くない一般的な内容のものが多くなる。また、Google が意図的に表示していない情報も存在している。公序良俗に反する情報や、風説の流布に該当する場合などは、システム管理者として削除することが規約に明記されている。ただ、民間私企業が運営するサービスであるため、自社にとって都合がよろしくない情報も隠されているのではないかと

いう疑念も湧いてくる。そこで、ユーザーの情報をプロフィールしないことを標榜している検索エンジン DuckDuckGo の検索結果と、Google の検索結果を混ぜて検証を行った。

両方の検索エンジン（米国にて実施）に、“Facial Recognition” を検索文として入力し、出力されたスニペット（要約）付きのリストの全 2,105 件を用いて俯瞰図を作成すると（図 4）、中央部分に巨大な疎な領域が出現する。ここには、Google の出力結果がなく、DuckDuckGo の検索結果のみが現れている。俯瞰図の特徴として、中心付近に位置する情報は残りの情報も含めて当該母集団と多く共通する内容が配置される。“Facial Recognition” のように一般的な用語を入力しているので、中心に空白がくることは珍しい。プロフィールしていないという DuckDuckGo から出力された当該中心領域の情報を参照すると、その中には

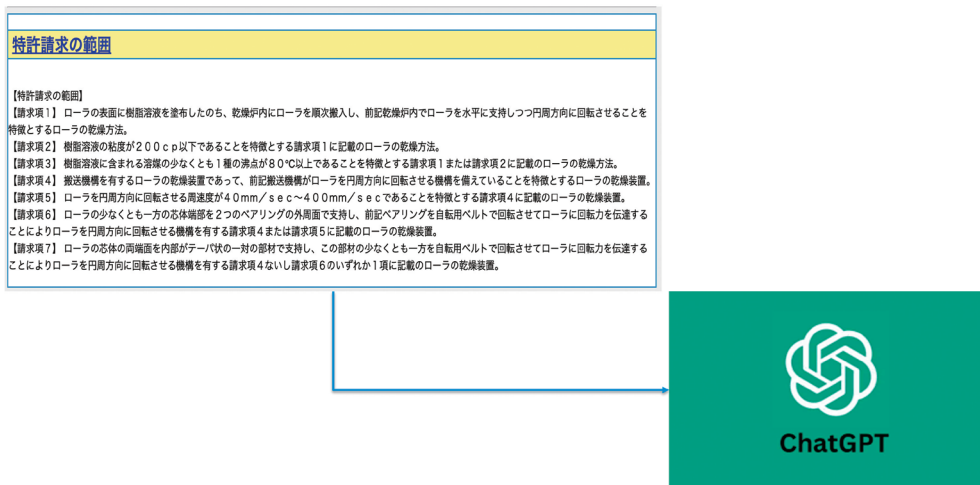


図 5 特許公開公報を ChatGPT に投入

Q : 重要な数値と単位を抽出してください。



重要な数値と単位を抽出します。

1. 樹脂溶液の粘度: 200 cp 以下 (200 cp or less)
2. 溶媒の沸点: 80°C 以上 (80°C or higher)
3. 周速度: 40 mm/sec ~ 400 mm/sec (40 mm/sec to 400 mm/sec)

Q : 当該分野の専門用語をTFIDF法に従って重要順に抜き出してください。



TF-IDF法に基づいて、当該分野の専門用語を重要順に抜き出します。

- | | |
|-------------------------------|------------------------------------|
| 1. 乾燥炉内 (Drying furnace) | 8. 芯体端部 (Core end) |
| 2. 樹脂溶液 (Resin solution) | 9. 回転させる機構 (Rotational mechanism) |
| 3. 融媒 (Solvent) | 10. 乾燥装置 (Drying apparatus) |
| 4. ベアリング (Bearing) | 11. 回転させる周速度 (Rotational speed) |
| 5. 回転力 (Rotational force) | 12. テーパー状 (Tapered) |
| 6. ロール (Roller) | 13. 支持する (Support) |
| 7. 搬送機構 (Conveying mechanism) | 14. 乾燥方法 (Drying method) |
| | 15. 特許請求の範囲 (Claims of the patent) |

これらの用語は、与えられた文章中で重要性が高いと判断されました。

図 6 ChatGPT を使って重要な数値と用語を抽出

Q:次の言葉のいくつかを使って、物語を作成してください。

Key Phrases:  

バンド | 時計ケース | バンド抜き耐久試験用装置 | 装着バンド | 通 | 原時計 | ピボットピン | 時計本体 | アクセサリー時計 | 連結部
| バンド本体 | バンド連結具 | 裝飾部材 | バンド片 | 連結足 | バネ棒 | 中留 | バンド離別防止枠 | 巻着装置 | 帯状装身具 | 尾錠 | バンド部材
| バンド部品 | バンド連結部材 | バンド部 | ロックピン | バンド連結部 | 枠 | ベゼル | 中留具 | 貫通孔 | 連結部 | 駒 | 裏面バンド部材
| 裏蓋 | バンド取付部 | 紐状部材 | 連結ピン | 操作部材 | フッ素樹脂繊維



ある日、時計職人のミカエルは新しい時計の開発に取り組んでいました。彼はバンド時計のデザインに独自のアイデアを取り入れ、特許を取得しました。その特許は、バンド抜き耐久試験用装置の装着バンドに関するものでした。ミカエルは、バンドの遊腕時計ピボットピンに特別な注意を払いました。このピンは時計本体と連結され、バンド本体を支える重要な部品でした。ミカエルは、アクセサリー時計の連結枠を使い、バンド本体とバンド連結具を一体化させることで、美しい裝飾部材を実現しました。

さらに、ミカエルはバンド片の連結足に特殊なバネ棒と中留を組み込みました。これにより、バンドの離別を防止する新しい構造を生み出しました。バンド巻着装置もバンドの装着を簡単かつ確実にするために開発されました。

この新しいバンドは帯状の装身具と尾錠で固定され、バンド部材とバンド部品が組み合わさっています。バンド連結部材とバンド部ロックピンを使用することで、バンド連結脚枠がしっかりと時計のベゼル中留具に貫通し、連結部駒が確実に固定されます。

ミカエルはバンド裏側の部材にも気を配りました。裏蓋にバンド取付部の紐状部材が組み込まれ、連結ピン操作部材によってバンドの取り外しや調整が容易に行えるようになっています。

最後に、ミカエルはフッ素樹脂繊維を使用してバンドを強化しました。この繊維は耐久性に優れ、時計愛好家に高い評価を受けました。

ミカエルの新しいバンド時計は、その革新的なデザインと高品質な製品部品によって、多くの人々から注目を浴びました。彼の努力と情熱によって生み出されたこの時計は、世界中で愛されるアイテムとなりました。

図7 創作らしきもの

GAF A にとって不都合な情報が含まれていることが確認できた。これらは、Google 検索の結果には存在しないが、意図的かどうかはわからない。ただ、この実験を通じて、Google の検索結果に出てきた情報が全ての情報とは限らないことは確かめられた。

5 情報整理に有効な生成 AI

検索エンジンの検索結果がどのようなものであれ、その情報が正しいと判断すれば、そこに記載された情報のひとつふたつを使って、宿題のレポートに取り入れるためには、自分で理解し、何かしらの加工と作文をしなければならない。昨今世間を賑わせている生成 AI は、その作業を私たちがする機会すら奪ってしまう。一部の学校や組織では、生成 AI の代表格である ChatGPT の学習や業務での利用を禁止する方針を示している。成長期や習得期の人間の教育と思考訓練に障害となりうる可能性があると考えているからだろう。現時点での ChatGPT の出力を見ると一般的な答えに終始しており、多才だがまだまだ学習が必要なようである。しかし、情報爆発下の時代、要約、キー抽出に生成 AI は大いに活躍できる。人間が不得意なルーチンとなる作業は生成 AI に依頼して時間を保全し、人間は思考することに集中するハイブリッド利用が、現段階の最適な付き合い方だろう。たとえば、特許公開公報の内容を ChatGPT に投入し(図5)、文字数を指定しての要約や、重要キーワードやキーとなる数値の抜き出しをお願いすると、もの数秒で計算を終了し、対話らしく一文字ずつ表示し

てくる。出力内容は見たところ正確であるように思われる(図6)。

続いて、文献から生成されたキーワード群を ChatGPT に投入し、物語を作成してもらった。質問の仕方にさまざまな前提条件をつけることは可能だろうが、ここではあえて指定せずに、あいまいな質問形式をとった。出力された物語は、幼い頃に読み聞かせで知ったような、ヨーロッパの児童書の翻訳版の味付けがする物語であった。おそらく、ChatGPT の学習では日本よりもヨーロッパの「物語」に馴染みがあるのだろう。この生成された物語は、何かの模倣に近い創作であり、ストーリー自体には意味はないが、技術キーワードの意味を理解して、前後の文章を組み入れていることには素直に驚かされる。この能力を使えば、技術や素材の新しい組み合わせから想定される使い方などの気づきや発想支援に使える可能性がある。ただ、ここで出力されている結果はあくまでも組み合わせ最適化であり、創出ではなく、専門家のように直感的にひらめくことは期待できない。

6 気づきと直感を生かしたチームの組成

組織がデータに基づき戦略を立てるとき、社内外のさまざまな情報ソースを収集し、解析を行い、ストーリーを作成する。この一連の作業を一人でこなせるデータサイエンティストがいる場合は問題ない。多くの場合は、情報を整理する部門と、現場を理解している事業部門が、情報交換を行いながら進められる。情報を整理する部門では、課題や目的を定義化し、適切な情報母集団を

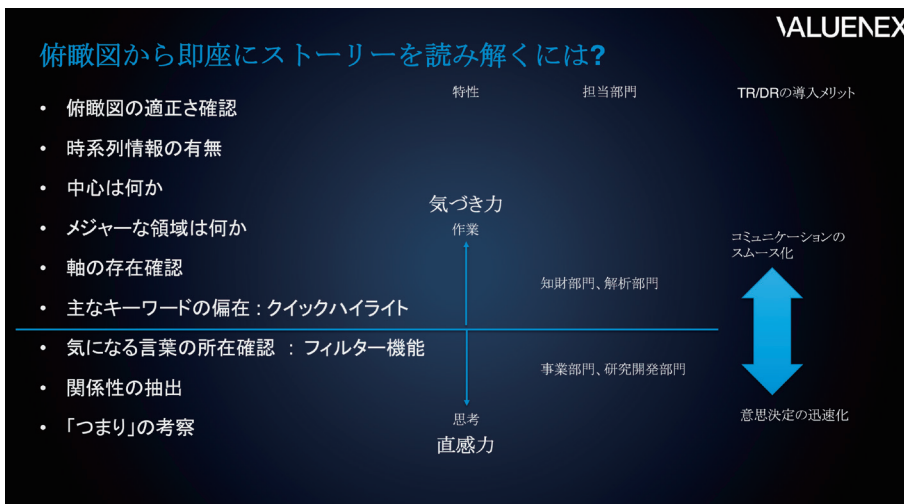


図8 気づきと直感を生かすチームプレー

収集する。集計と整理、可視化を行い、過去のパターンから気づきを得るという作業を行う。事業部門など現場感覚がある人は、可視化された情報に対して、自分の把握している事例や会社名などを手がかりに、直感力をはたらかせて、「つまり」の導出行う。二つの異なる組織を繋ぐ共通言語は可視化手法であり、スムーズなコミュニケーションを可能にする。

7 おわりに

2020年を境に情報爆発は加速し、生成AIの登場により、業務フローの中におけるホワイトカラー職の作業の機械への置き換えは、1980年代後半にレコードがCDに置き変わった速度よりも早く進むと考えられる。インフラも含めて多くの物理的なものが置き変わった転換とは異なり、生成AIを利用する意思を示した瞬間に、インターネットにてサービスが受けられるからである。情報セキュリティ、教育、トレーニング、その他さまざまな理由で利用を阻止する動きはあるが、利用方法の工夫で乗り越えられ、利便性の前には打ち勝てず、次第に少数派になるだろう。しかし、生活が人中心であり、業務が人のために仕えることであるかぎり、人の気づきと直感は、最適化に特化した機械よりも優れたアイデアを創出することは間違いない。少なくとも、2040年の脳科学とコンピューターサイエンスが融合した世界の機械が出現するまでは、人が考えるつづけることの価値は失われないだろう。

参考文献

- 1) 小林雅一「ゼロからわかる量子コンピュータ」、講談社現代新書、2022,3
- 2) ジェレミー・デシルヴァ、「直立二足歩行の人類史 人間を生き残らせた出来の悪い足」、文藝春秋社、2022
- 3) トレヴァー・コックス、「コンピューターは人のように話せるか?—話すこと・聞くことの科学」、白揚社、2020
- 4) 笹原和俊、「フェイクニュースを科学する: 拡散するデマ、陰謀論、プロパガンダのしくみ」、化学同人、2021
- 5) 中村達生、「DX時代における俯瞰解析を用いた公平・公正な未来社会の実現にむけて」法とコンピューター学会「法とコンピューター」No.40、2022,7
- 6) 中村達生、「俯瞰図を用いた技術動向分析とその活用方法」、株式会社情報機構「特許調査へのAI導入と業務効率化および特許情報の有効活用方法」発刊2022年9月

