

社会変革に向けたNTT研究所の取組み —IOWNの研究開発と万博での適用事例—



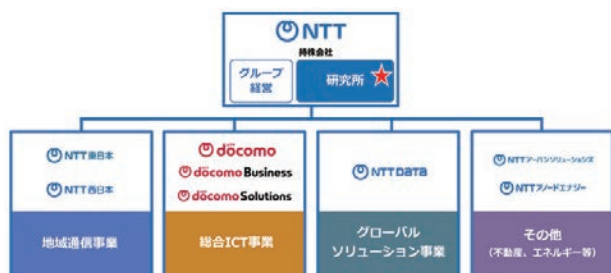
NTT株式会社 執行役員研究企画部門長 **きのした しんご**
木下 真吾

1. はじめに

本日は、NTT研究所の概要に加え、現在注力している3つの研究テーマ、「IOWN (Innovative Optical and Wireless Network)」「GenAI (生成AI)」「Quantum (量子)」についてご紹介する。このIOWNの文脈の中で、万博での活動についても触れていきたい。

2. NTT研究所の組織とミッション

まず、研究所の概要について述べたい。NTTの研究所は、グループ内の各事業会社に直接属しているわけではなく、持株会社の中の組織として、各事業会社から研究費を受け取って基礎的な研究を行い、その成果を事業に還元するというミッションを担っている。



■図1. NTTグループ

現在、研究員の数は約2,300名であり、設立は1948年に遡る。組織体系としては、国内の4つの総合研究所の下に14の研究所を配している。研究領域は非常に幅広く、物理や数学といった基礎的な分野から、ネットワーク、暗号、AIといった応用的な分野までを網羅しているのが特徴だ。また海外にもカリフォルニアのサン・ノゼに3つの拠点を構えており、こちらでも量子、暗号、メディカルといった極めて基礎的な研究を行っている。

我々は、年間を通じて非常に多くの特許出願や論文執筆を行っている。研究所内には実用化を担う部隊も多く事業に対して具体的な成果を提供しているが、その土台となるアカデミアへの貢献は非常に重要であると考えている。私自身もこれを重要視しており、KPI (重要業績評価指標) として、論文数や難関国際会議への投稿数を定めて管理している。具体的な指標を挙げると、IT系企業の中では、

論文数は現在世界で第9位、日本国内では第1位となっている。さらに分野を絞れば、光通信、量子、暗号といった領域では、世界でも1、2位を争う組織となっている。

最近では、研究所の技術を活用したスピノフベンチャーの立ち上げも活発に行っている。まず、音響技術を活用したスピーカー専用の会社 (NTT sonority) がある。次に「SPACE COMPASS」は、スカパー JSATとの協業により、宇宙通信や宇宙データセンターの構築を目指すものである。3番目の「NTTグリーン&フード」は、NTTとしてはかなり特殊な取組みだ。ここでは京都大学発のベンチャーであるリージョナルフィッシュと連携している。彼らはゲノム編集によってタイなどの成長を早める技術を持っている。一方で、NTT側の出身者は、中性子線加工を用いて藻のCO₂吸収量や栄養価を高める技術を持っている。これらを組み合わせ、栄養価の高い藻をタイやエビなどの餌とすることで、成長を促進させる。さらに、魚の背骨などにCO₂を吸着・固定化させることで、CO₂削減までも視野に入れている。現在、この陸上養殖はタイだけでなく、ヒラメやエビにまで拡大しており、またエビの陸上養殖業としては日本で一番の規模となっている。最後が「NTT AI-CIX」である。こちらは物流や流通における最適化を担う企業だ。最近では小売りのトライアル社などと協力し、サプライチェーンの最適化に携わっている。



■図2. 研究所スピノフ

本日は、こうした広範な研究体制を背景に、IOWN、生成AI、そして量子の3テーマについて順を追って解説していくが、まずはこれらのテーマを取り上げた背景を説明する。

3. AI巨大化が突き付ける課題

昨今のAI、特に生成AIは急激な大規模化を遂げており、それに伴い開発コストが深刻なまでに高騰している。スタ

ンフォード大学のレポートによれば、ChatGPTのベースとなった「GPT-3」は、1750億パラメータ数の規模を持ち、1回の学習に約5億円を要したとされる。しかし、続く「GPT-4」等の次世代モデルでは、パラメータ数は数兆規模に達し、開発コストは150億～200億円規模にまで跳ね上がっている。我々も生成AIを開発している立場から言えば、最新モデルのリリースには更に数倍から数十倍のコストが投じられているのが実態と思われる。

電力消費も極めて大きな問題だ。GPT-4クラスのモデルを1回学習させるのに必要な電力は約4万MWhに達し、これは原子力発電所約40基分の1時間の発電量に相当するとも言われている。また、学習時のみならず「推論（ユーザーの問いかけへの回答生成）」時の消費電力も膨大である。Google検索と比較すると、ChatGPTのクエリ1回当たりの電力消費は約10倍の2.9Whを要するとされ、同様の質問でもAIを利用すれば10倍のコストがかかる計算になる。

1回の利用にかかる消費電力



■ 図3. AIの大規模化：膨大な消費電力（利用時）

さらに、投資規模の格差も顕著である。米国の民間企業によるAI投資額は年間約10兆円に達し、中国がそれに続く。対する日本は約1000億円規模にとどまっており、米国とは約100倍もの開きがある。こうしたコスト、電力、投資という多面的な課題が、現在のAI開発に突き付けられている。

AI開発におけるリソースの格差は、資金面だけでなくエンジニアの数にも現れている。正確な調査は難しいものの、一例として生成AIを開発する各社が発表した論文の共著者数を見ると、2023年末のGoogle「Gemini」では約1,400名が名を連ねている。通常の論文であれば多くても10名程度であることを考えると、1つのモデル開発に、まるで映画産業のような規模の人員が投入されているのが実態である。現在のAI開発は、純粋なアルゴリズムの競争というより、莫大な資金と人員を投じる「パワーゲーム」の様相を呈している。

4. 将来のAIの進化に向けたNTTのアプローチ

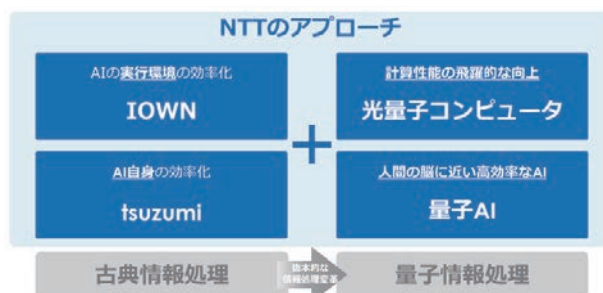
このような背景を考えると、現在私たちは、将来のAIの進化に向けて単純に現在の計算リソースを増大させ続けるべきかという問いに直面している。投資額の差を考慮すれば、日本が同様のリソースを確保し続けられるのか、また、この競争を続けることが本当に人と地球のために役立つのかという観点から、多角的に検討しなければならない。

この課題に対し、NTTは主に2つのアプローチをとっている。まず、「IOWN」、これはAIの開発や推論の実行環境において、同じ処理を行うにしても消費電力効率に優れた基盤を実現することを目指している。そして2つめにAI自体を非常に軽量化・効率化することを目指し、「tsuzumi」の研究開発を進めている。

この2つは現在の情報処理、いわゆる「古典情報処理」の世界でいかに努力するかという観点での取組みである。しかし、昨今ではこの手法も限界に達しつつあると言われており、この限界を打破するためには、「量子情報処理」の世界へもいち早く踏み込む必要があると考えている。

具体的には、計算性能を飛躍的に向上させる「光量子コンピュータ」への挑戦、そして、人間の脳に近い僅かな電力で極めて効率的な学習を行う、高効率な「量子AI」の研究などに取り組んでいる。

このように古典と量子の両面から、IOWN、生成AI（tsuzumi）、そして光量子コンピュータの3領域をバランスよく推進している。これが、本日のアジェンダに3つのテーマを据えた背景である。



■ 図4. NTTのアプローチ

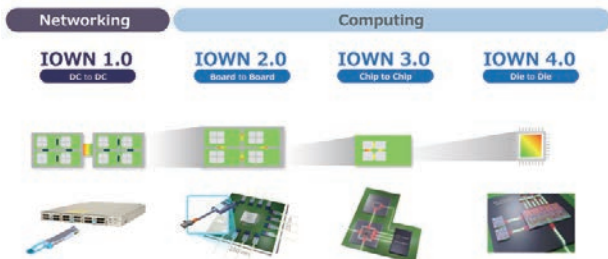
5. IOWNのロードマップと狙い

IOWNは、情報通信基盤を電気から光へと段階的に転換していくための構想であり、4つのステップで整理されている。第一段階である1.0では、ネットワークの世界に光をとということで、データセンター間の通信をエンドツーエンドで完全に光化することを目指している。現在も光ファイバー



は用いられているが、途中で電気変換が介在しており、完全な光通信にはなっていない。この電気変換をなくし、光をエンドツーエンドで、波長レベルで専有することで、高性能かつ高効率な通信を実現しようというのがIOWN 1.0の考え方である。

そして、IOWN2.0からは、光がコンピュータの内部に入り始めるステップであり、IOWNの真骨頂となってくる。まずはデータセンター内において、サーバー間やボード間の接続を光化する。続くIOWN3.0では、ボード内のチップ間接続を光に置き換え、さらに、IOWN4.0では、チップ内部、すなわちダイ間の接続まで光化するというロードマップを描いている。



■図5. IOWNロードマップ

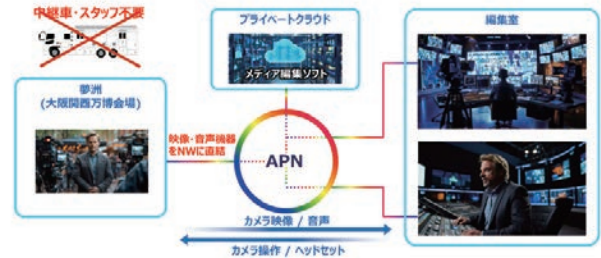
まずはIOWN1.0の状況について説明する。ここでユースケースとして紹介したいのが、弘前大学と行っている遠隔手術の実証実験である。遠隔手術に関しては、遅延時間の長さだけでなく、人間が医療機器を操作する上ではその幅が一定で揺らぎがないことが非常に重要であると現場の医師から言われており、その点IOWNは光の波長をエンドツーエンドで占有することで揺らぎをほぼなくせる点が高く評価されている。

また、最近では建設現場でも労働力の不足が深刻な問題となっている。現在、安藤ハザマと進めている実験では、重機の遠隔監視や遠隔操縦を可能にすることで、移動時間を削減し、労働効率を大幅に上げることができると考えている。

また放送分野においても、従来はイベントなどがあると現地に中継車や多人数のスタッフを派遣して編集室までの回線を整備するなど非常に手間とコストがかかっていたが、APNで現場と編集室をつなぎ、クラウドベースの編集ソフトを使うことで現地に大規模な設備を必要としない制作形態が実現しつつあり、2025年度の世界陸上など実際の大型イベントで活用が進んでいる。

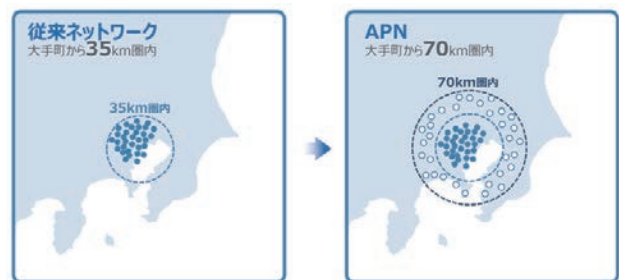
以上は1対1の光通信を活用したユースケースだが、IOWN

TBS、NHK、朝日放送、関西テレビ放送、読売テレビと連携して万博中継等で実証
編集室と中継先をAPNで接続することで遠隔番組制作を実現



■図6. ユースケース：放送リモートプロダクション

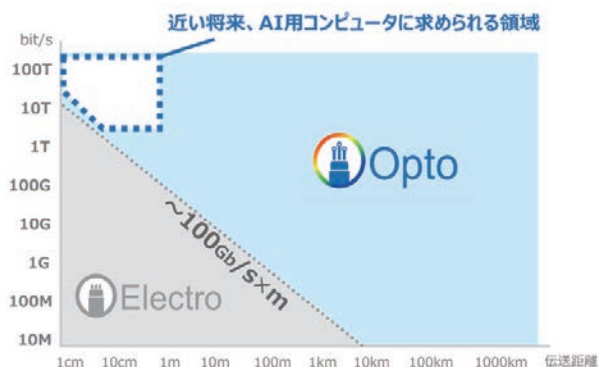
の最大のターゲットはデータセンターでの活用である。現在日本国内ではデータセンターが都市部に集中しており、用地確保や電力供給が制約となっている。しかしIOWNのAPNを活用することによって通信遅延を圧縮できれば、データセンターを都市から離れた郊外に分散配置しつつ、仮想的には一体のデータセンターとして運用することが可能になる。例えば、従来であれば1つのアベイラビリティゾーンの目安は半径35km程度であったが、これを70km程度まで伸ばすことができるのではないかと考えており、用地の選定が容易になる。更にワット・ビット連携との相性も非常に高い。通常、送電コストに比べ通信コストはおよそ1000:1程度と桁違いに低いため、都心から離れた再生可能エネルギー拠点の近傍にデータセンターを設置し、そこから高速通信で都市部につなぐことで、全体として性能・費用のバランスが極めて高い合理的なソリューションが提供できる。



■図7. ワット・ビット連携

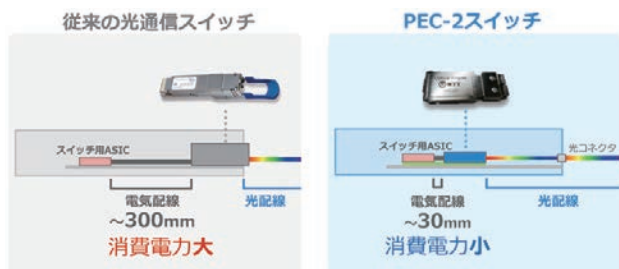
次にIOWN2.0以降に焦点を当て、コンピューティングの領域で起こっていることを説明する。現在、AIデータセンターが数多く構築されているが、このタイプのデータセンターでは大量に設置されるGPU間の通信が大きな課題となっている。GPUの進化に伴い、チップ間で求められる通信帯域は飛躍的に増大しており、従来の電気配線では消費電力と発熱のため限界に近付きつつある。このため、NVIDIA自身も光電融合を使って光化していく計画を発表しているが、

NTTもブロードコムやアクトン・テクノロジーと組んで、光電融合デバイスを使った光のスイッチを展開していこうと考えている。背景として、光と電気の適用領域を伝送距離と通信容量の関係で考えると、従来長距離は光、近距離は電気が有利というのが常識であった。しかし通信容量が急激に増大する中で、数cmの距離でも光化が必要な領域に移ってきていると言える。



■ 図8. 電気による通信の限界と、光による解決

具体的な装置のイメージで言うと、我々が開発を進めている「PEC-2スイッチ」では、従来の光通信スイッチよりも電気配線部分を大幅に短くすることで、通信容量の拡大と消費電力削減の両立を図っている。2026年末までに通信容量が102.4Tbps程度のスイッチを提供し、GPU間通信の効率化を実現する予定である。

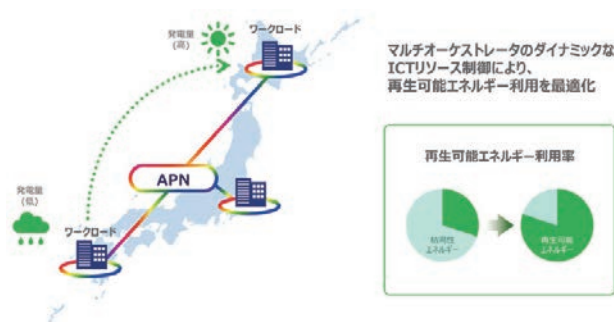


■ 図9. PEC-2 スイッチにおける消費電力削減

もう1つ重要なのがソフトウェアによる制御である。光通信スイッチなどの活用で、今後は数万台規模のGPUをコントロールすることが必要になる。現在私たちが運用しているGPU1000台規模のデータセンターでもGPUの故障率はそれなりに高い一方、AIの学習では一度開始すると1か月程度マシンを連続稼働させることが必要になる。こうした条件を考えると、利用状況や装置の状態、消費電力などを考慮しながらタスクを最適に配置する仕組みを実現する必要がある。例えば、複数拠点のGPUデータセンター環境にお

いて、新しい推論タスクに対してはレスポンス優先で地域間に分散して処理を行い、新しい学習タスクに対してはGPU間通信優先で、1か所で集中的に処理するなど、推論と学習では求められる配置条件が異なるため状況に応じたスケジューリングが不可欠である。

こうした制御が実現すれば、再生可能エネルギーの発電状況に応じて処理負荷を地域間で移動させるといった、ダイナミックなワット・ビット連携も可能になると期待されている。



■ 図10. ダイナミックワット・ビット連携

6. 大阪・関西万博におけるIOWN活用

大阪・関西万博では、IOWNを用いた様々な取組みが実施された。1970年の大阪万博では、NTTが世界で初めてコードレスホンを使えるようにし、会場から通話できる体験を提供したが、今回の万博も同様に、新技術を社会に提示する場として位置付けた。

まず規模の面では、NTTはパビリオンを出展し来場者数は約39万人に達した。また、バーチャル万博の運営も行い、約280万アクセスを記録している。会場内の地図表示や最適ルート案内するパーソナルエージェントアプリも提供し、約250万ダウンロードと、期間限定の施策としては大きな利用実績となった。

IOWN 1.0 (All-Photonics Network) は、万博会場内14か所、国内14か所、更に台湾との接続に利用され、国際的なネットワークを構成した。この基盤の上で会期中には10を超えるユースケースが提供された。開幕日に実施された大規模合唱イベント「1万人の第九」では、分散した演奏者を一体化するため、低遅延で揺らぎの少ない通信が活用された。加えて、日本と台湾を結ぶパフォーマンスや、遠隔地にある農機の操作体験など、距離を意識させない体験を実現した。

NTTパビリオンでは3つのゾーンが設けられ、その中心



■ 図11. 大阪・関西万博におけるNTTの取組み

となるゾーンでは、Perfumeとの連携ということで、IOWNを用いた新しい空間体験が展開された。これは、「空間を伝送する」というコンセプトで、離れた場所で行われるPerfumeのパフォーマンスを多数のセンサーやカメラで取得し、リアルタイムに立体映像として再現する試みである。床振動や照明とも連動させることで、空間全体として体験を構成した点が特徴であった。通信容量は約25Gbpsと極端に大きいものではないが、点群データを含む映像をリアルタイムに処理・再構成した点は技術的にも挑戦であった。

さらに、IOWN 2.0、すなわちコンピュータ内部への光技術導入についても、万博の場で複数の実証実験が行われ、将来の計算基盤を見据えた検討が進められた。

7. 国産生成AI「tsuzumi」の狙いと特長

tsuzumiは、NTTが自社で開発している国産の生成AIである。2024年3月に最初のバージョンとして約70億パラメータのモデルを公開し、その後2024年10月には約300億パラメータまで拡張したtsuzumi2モデルを開発した。300億というサイズは、比較的安価なGPU1台で動作可能な最大規模であり、On-premiseやプライベートクラウドにも載せやすい。その結果、低コストかつ高いセキュリティを確保したAIとなっている。



Why 30B? → Only 1 GPU → On-premise

→ 低コスト&高セキュア AI

■ 図12. tsuzumi2の特長

tsuzumi2は、日本語の各種ベンチマークにおいて同規模の海外モデルと比べて高い性能を示しており、更に一段大きなモデルと比較しても優位性が確認されている。一部の超大規模モデルには及ばない点もあるが、小型でありな

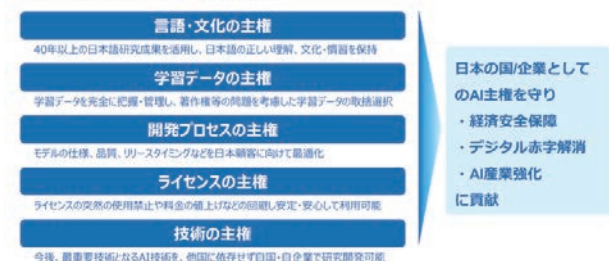
から高い性能を持つ点が特長である。

また、チューニングのしやすさも重要な特性である。分野を絞っての学習効率が高く、例えば、金融分野の「ファイナンシャル・プランニング技能士検定2級」を集中的に学習させた場合、比較的少ない学習量で専門的な試験に対応できる水準に到達した。これは同様のタスクにおいて、他モデルよりも高い学習効率を示している。

加えて、ハードウェアコストも大規模LLMに比べて大幅に低く抑えられており、実用面での導入障壁が低い。

また、tsuzumi2の開発においては、国産あるいはソブリンAIとしての意義も強く意識している。米中など海外製LLMへの依存が進む中で、経済安全保障の視点が重要になっており、言語・文化の維持、学習データ、開発プロセス、更にライセンスを自らコントロールできる、「AI主権」を守ることが重要になっている。特に、生成AIは、今後半導体やOS、スマートフォンに匹敵する、あるいはそれ以上の基盤技術になる可能性が高く、この分野の技術開発を放棄すれば将来的な技術力の空洞化につながり、技術の主権を失ってしまう。こうした認識を踏まえ、NTTとしては、短期的なビジネスだけに左右されることなく、継続的に自前で生成AIを開発していくという覚悟で取り組んでいる。

NTTがゼロから開発した国産モデル、ソブリンAIに最適



■ 図13. 特長：国産AI

生成AIの応用例としては、音声対話AIがある。従来の音声対話AIが持つごちなさに対し、NTTの技術では人間同士の会話に近い自然な応答が可能になっている。ただし、この仕組みは音声認識と言語モデルを単純に組み合わせただけのものではなく、実際の通話音声音を音として大量に学習する音を使ったエンドツーエンド型の学習に基づいている。そのため、人間らしい会話は実現できる一方で、論理的な制御には課題が残り、今後は複数技術の組合せで制御面も改善していく予定である。

さらに、将来を見据えた基礎的な研究として、脳波解析によって脳から高い精度で知覚内容をテキスト化する「マインド・キャプション」技術や、LLM内部の振る舞い

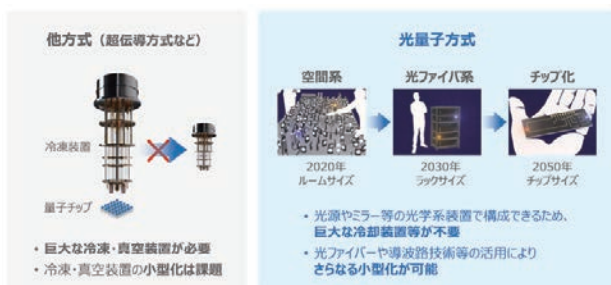
を解明する「LLMの心理的メカニズム」の研究も進めている。後者は生成AIのブラックボックス性を低減し、より高度な人間の脳に近い人工知能へと発展させるための基礎研究と位置付け、取り組んでいる。

8. 光量子コンピューティングへの挑戦

量子コンピュータは実用化までまだ時間がかかると見られており、技術的課題も多い分野である。ただし米国や中国を中心に、AIの次を見据えた大規模な投資と人材集積が進んでおり、ブレイクスルーが起こる可能性も期待されている。

NTTが取り組む光量子方式は、東京大学の古澤明教授を中心に研究が進められている。現在主流とされる超伝導方式や中性原子方式にはそれぞれ強みがあるが、光量子方式の最大の特長は、常温かつ常圧で動作する点にある。超伝導方式では量子ビットを増やすために大型の冷却装置が必要であり、物理的・電力的な制約が大きい。一方、光量子方式ではレーザー光源や光ファイバー、光学系装置といった、既存の光通信技術に近い構成要素で実現できる。

実際、光量子コンピュータに用いられる光源や観測装置



■ 図14. 特長：常温・常圧動作

は、通信分野で使われている技術と高い共通性を持っている。光通信で培われた技術をそのまま活用できる点は、開発面での大きな利点である。

また、他方式では量子ビット数を増やす際に空間的な並列化が必要となるが、光量子方式は通信を基盤としているため、周波数を高めることで処理能力を拡張できる。電気ではGHz帯が限界であるのに対し、光ではTHz帯まで視野に入り、超高速化に適した特性を持つ。現時点では量子ビット数において他方式に劣るものの、将来的なスケラビリティは高く、長期的には優位に立てると考えている。NTTはOptQCという東京大学発のベンチャー企業と連携し、2027年に1万量子ビット、2030年には100万量子ビット規模を目指して研究開発を進めている。



スケラブルで信頼性の高い光量子コンピュータの実現

2030年：世界トップレベル 100万量子ビットの実現
2027年：国内トップレベル 1万量子ビットの実現

2025年のスタンダード 国内：数百量子ビット、世界：数千量子ビット

■ 図15. 光量子コンピュータの実用化にむけて

以上、本日はIOWN、AI、量子領域についてのNTTのアプローチをご紹介します。

ご清聴ありがとうございました。

※本記事は、2026年1月26日開催「第54回ITUクラブ総会」での講演をリライトしたものです。（責任編集：日本ITU協会）