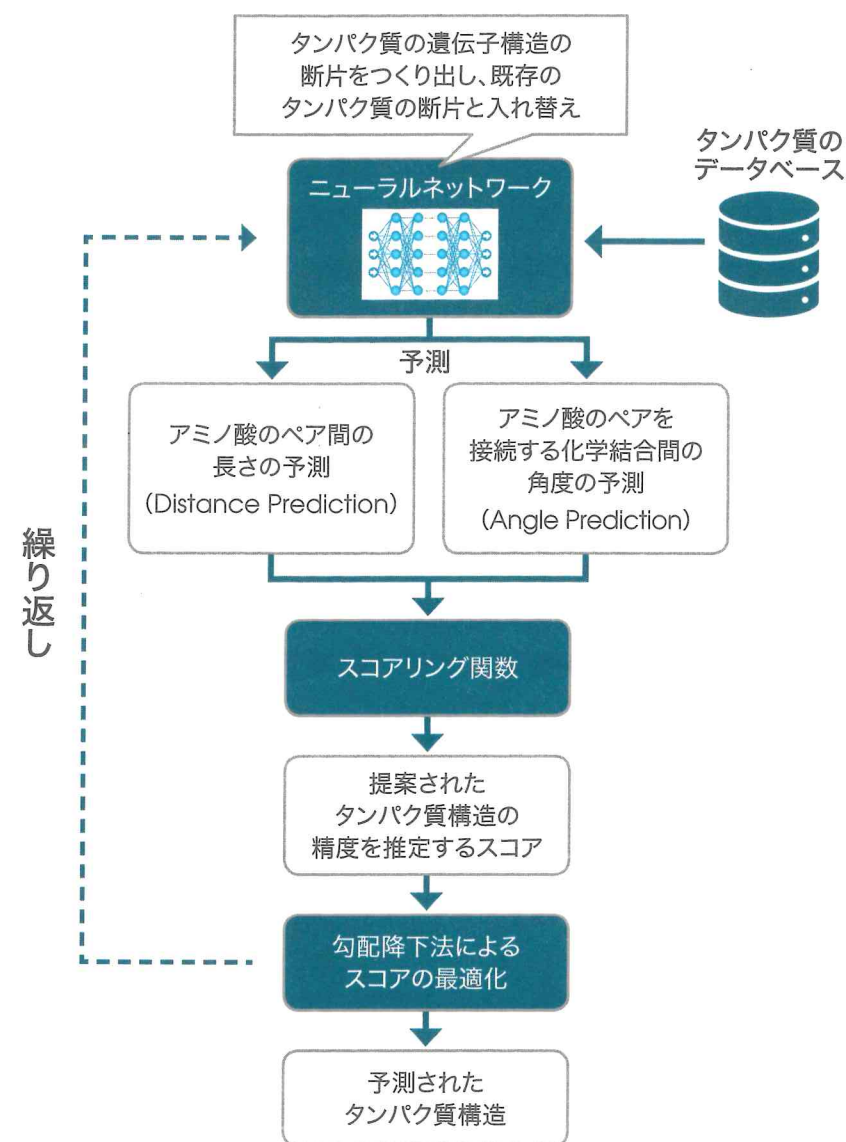


■図2-2-16 AlphaFoldによる3次元構造の予測

出典: DeepMind: AlphaFold: Using AI for scientific discovery^{※35}

参考文献

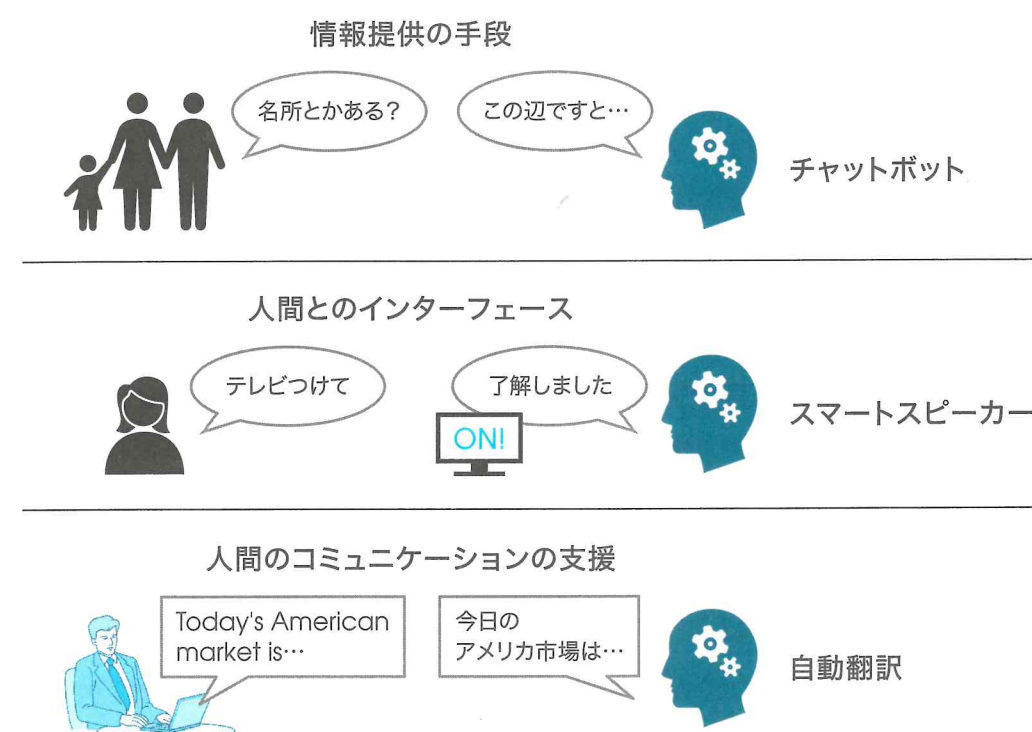
- [1] Graham, Paul (August 2002), A Plan for Spam<<http://www.paulgraham.com/spam.html>>
- [2] Gori, Marco, Gabriele Monfardini, and Franco Scarselli. "A new model for learning in graph domains." Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. Vol. 2. IEEE, 2005.
- [3] 国立研究開発法人日本医療研究開発機構. 2種類の深層学習手法の組み合わせで薬剤とタンパク質の相互作用を予測—高速で高精度な予測と相互作用部位の特定・可視化による検証を実現— <https://www.amed.go.jp/news/release_20180829-02.html>

※35 <https://deepmind.com/blog/article/alphafold>

2.2.6 言語

コミュニケーションにおいては「言語」を切り離すことはできない。コンピューターが「言語」を扱えるようになれば、人間とコンピューターとのコミュニケーションが可能になる。すでに音声でスマートフォンやスマートスピーカーに指示を出すことも日常風景になっている(図2-2-17)。また、機械翻訳や顧客からの問い合わせに対応するチャットボットなどの応用も進んでいる。本節では、自然言語処理に焦点を当て、伝統的なルールベース翻訳から、統計機械翻訳、ディープラーニングによる機械翻訳への進歩を紹介し、BERTやGPTなどの最新技術動向を紹介する。

■図2-2-17 AIによる「言語」の活用例



(1) 概要

コンピューターで自然言語を扱うことは、コンピューターの創世記から行われていた。また、コンピューター登場以前に言語学などで行われていたテキストの統計処理、テキストや発話を大規模に集積したデータベースであるコーパスの作成なども本来はコンピューター向きの作業であった。その中で機械翻訳は「自然言語処理最大のアプリケーション」[1]といわれるように、コンピューターでの自然言語処理研究の中心にあった。しかし自然言語処理は、例えば特定言語の単語数だけでも膨大なものがあり、コンピューター創世記のハードウェアには過大な処理でもあった。また自然言語は、言語学としてコンピューター登場以前から研究されてきたが、文法定義や品詞、意味分類は必ずしも一つに定まらず、コンピューター処理を想定したものでもなかった。

1960年代から盛んになったワードプロセッサなどの発展による文章の電子化に続き、「機械翻訳」は、多大な投資と、自然言語処理技術や計算機パワーの向上などにより、技術的な発展を遂げたが、最初に開発された、主に人手で整備するルールに基づいた機械翻訳(RBMT)は、その性能がすぐに頭打ちになった。1980年代後半にはIBMの研究グループが統計的機械翻訳(SMT)の研究を開始した。これは単語の翻訳確率や並べ替えの確率などの翻訳に必要な知識