



ネガティブデータの共有と AI

●
吉田潤一 Jun-ichi YOSHIDA

京都大学大学院工学研究科 教授・有機合成化学協会 会長



研究成果は学术论文などとして発表されるが、よい結果のみを記載するのが常であり、いわゆるネガティブデータを記載することは少ない。しかし、よい結果が得られなかったからといって、その事実は、本人はもちろんほかの研究者にとっても重要な情報である。事実を明確に記録し報告することは重要である。いかに多くのきちんとしたネガティブデータを持っているかが、新たな創造や研究開発のための鍵を握っているといっても過言ではない。大学などの場合には、教員の頭の中にそういったネガティブデータが蓄積されている場合が多いのではないかと考えられるが、企業の場合には、ネガティブデータの蓄積・共有を組織的・系統的に行っているところが多いのではないだろうか。

さて、問題はそのようなネガティブデータの社会的な共有である。以前から、ネガティブデータを集めた学術誌を発行すればよいのではないかと言われてきた。しかし、新たな発見・発明のための宝の山かもしれないネガティブデータを他者（他社）に公開・共有するのは躊躇される。また、きちんとしたデータを取るには時間と労力が必要で、短期間で多くの成果を求められる現状では、いわゆるポジティブデータに集中せざるを得ず、質の高いネガティブデータの蓄積・共有はなかなか困難である。

しかし、最近、この問題に関して新たな局面が出てきた。近年、人工知能（AI）が社会のあらゆる分野で利用され始めるようになってきた。蒸気機関などの人工動力の利用が第1の産業革命と呼ばれるのに対して、AIの利用は第2の産業革命と呼ばれることもある。近年のAIは、膨大な数のデータを機械学習や深層学習で扱うことが特徴である。化学も例外ではなく、AIの利用が様々な領域で始まっている。化学分野で機械学習の研究を行っている研究者の話を国際会議などで聞くと、データの質が担保できないこととネガティブデータが蓄積・共有されていないことが問題とのことである。ネガティブデータは境界条件を明らかにするため必須であるが、論文などからはなかなか得られない。また、前述のようにネガティブデータの蓄積を人間が行うには負担が大きすぎる。ところが、最近、実験ロボットなどを活用してネガティブデータを含めた質の高いデータを蓄積し、それを共有する動きが始まっているようである。質の保証されたデータをどれだけ素早く大量に蓄積するとともに、それらを有効に共有できるかが成功の鍵を握るであろう。そして、それらのデータに基づいたAIを研究・開発にどのように活用するのかが、化学の発展にとって重要な課題となるであろう。今後のこの分野の発展に期待したい。

© 2017 The Chemical Society of Japan