



データキュレーターと オープンサイエンス

館長 米山博志

4月以来、図書館に関係したいろいろな資料や文章を読む機会が多くなり、データキュレーターという言葉を知ったのは、10年以上前にベストセラーとなった「ダ・ヴィンチ・コード」を読んだときだ。この小説の冒頭は、ルーヴル美術館のダ・ヴィンチの絵の前でキュレーターが死体で発見されるという設定である。キュレーターとは、辞書には学芸員とあり、博物館などの施設において専門知識をもって資料の分類、管理などを行うとある。その後、この言葉に出くわすことが多くなり、そして、データキュレーターである。ここでは、オープンサイエンスというものと関わって使われているものである。

このオープンサイエンスについて、2016年6月16日から2日間、仙台で開かれた国立大学図書館協会総会で議論が行われた。今年は、「国立大学図書館協会ビジョン2020」を採択し、その理念を1. 知の共有、2. 知の創出、3. 新しい人材、としてまとめ、オープンサイエンスなどの動きに対応することが決まった。この3つの概念は目新しいものではなく、大学の図書館としては、古くから掲げられてきた、そしてワールドワイドな概念であろう。それでは、なぜ改めてこの3つの概念をビジョンの中心に据えることになったかということ、それはその質が大きく変わりつつあるからである。その最も大きな要因は、大量の情報が、瞬時に、しかも双方向的にやりとりされるコミュニケーション様式の変化であり、そのことが、教育・研究のあり方もそして大学図書館のあり方も変えようとしているようだ。

オープンサイエンスが現実的に動き出したのが、2013年、ロンドンでのG8科学大臣会合の共同声明だそうで、その後急速に広がり、日本では内閣府の検討結果が「我が国におけるオープンサイエンス推進のあり方について～サイエンスの新たな飛躍の時代の幕開け～」(2015年3月、国際的動向を踏まえたオープンサイエンスに関する検討会)にまとめられている。それによると、オープンサイエンスは、公的研究資金を用いた研究結果について、広く容易なアクセスを可能にし、効果的に研究を推進することを目的とする。これはオープンアクセスとオープンデータからなり、前者はインターネット上で、査読済み論文の自由な閲覧、利用が可能になることであり、後者は研究によって生まれたデータを他者が自由に使えるようにするものである。

そもそも科学の歴史は、成果のオープン化という歴史を辿ってきている。「オープンサイエンスと科学データの可能性」(宮入暢子「情報管理」2014 Vol57, 80)によると、この歴史の中で培われてきた(1)先駆性の確保、(2)科学の集約化、(3)第三者による正当性の担保、(4)著者による説明責任の確立、にまとめられる特徴が、現在のプレプリントサーバー、オープンデータリポジトリなどへの動きにつながっているという。

オープンサイエンスにおいては、研究成果の検証のためだけではなく、新たな価値を生み出すためにデータそのものが重要な役割を果たすことになる。G8科学大臣会合で合意されたdiscoverable(探し易い)、accessible(アクセスし易い)、understandable(理解し易い)、manageable(管理し易い)なオープンデータを保証するための、データの保存、管理の仕組みが構築されることが必須で、これらのデータに価値付けをし、そして利用し易くする役割が非常に重要になってくる。これが、データキュレーションであり、それを担う人材がデータキュレーターと呼ばれ、将来の科学のあり方に重要な役割をすることになる(村山泰啓、林和弘、オープンサイエンスをめぐる新しい潮流(その1)科学技術動向、2014, 146, 12.同(その2)科学技術動向、2014, 147, 16.同(その3)科学技術動向、2015, 148, 4.)。

と、ここまで付け焼き刃で見聞きしたオープンサイエンスについて書いて来たが、正直な所、腑に落ちるまではもう少し時間がかかりそうだ。しかし、大学図書館はこの急速な動きに対して、将来新たな役割を担うべく変化を求められていることは確かなようで、前述の図書館協会総会で行われた、深貝保則氏(横浜国立大学)の大変興味深い講演によれば、オープン化という容易でない課題にたいして「離れ業」を果たしながら、大学図書館そのものが広い意味でのキュレーターとしての役割を果たして行かなくてはならないという。