

微化石ビッグデータの蓄積と利用

鈴木紀毅*・西 弘嗣**

*東北大学大学院・理学研究科地学専攻・**東北大学学術資源公開センター・東北大学総合学術博物館

Input and output of “Big Data” for micropaleontology

Noritoshi Suzuki* and Hiroshi Nishi**

*Department of Earth Sciences, Graduate School of Science, Tohoku University, 6-3 Aramaki Aoba, Aoba-ku, Sendai City, Miyagi, 980-8578 Japan (*norinori@m.tohoku.ac.jp); **The Center for Academic Resources and Archives, Tohoku University Museum, Tohoku University, 6-3 Aramaki Aoba, Aoba-ku, Sendai City, Miyagi, 980-8578 Japan

Key words: Big Data, database, Naples, PaleoTax, Schewiakoff

微化石データは、膨大な標本を連続的に検討した結果からなり、「高解像度」、「高頻度」、「多様性」という総務省（2015）が述べるビッグデータの3点の特徴を備えている。これらデータの根本は「標本」にあり、その「標本」がどのように情報基盤のデータベースとして整備されているかを知ることは、標本の情報を正しく活用するために重要である。データの質を保証するものは標本で、その記述は印刷論文だけでなく、研究ノートなどにも残されている場合もある。標本と記述の記録を解読して情報基盤とすることで、新たな研究標本として活用できる（図1）。標本のデータベース化は様々な形でなされているが、いずれも登録当時のままの分類名で留めるか、スライド単位での登録など「コンテナ」で登録するのが現

実的である（図2）。こうやって蓄積されてきた情報基盤を活用するにあたり、データの精査が必要である。精査した上でアウトプットをしてもそれはまだ「データ」で（図3）、そのデータが事実を適切に反映しているかさらに精査してはじめて情報基盤を正しく活用できる。

木原 章・Travaglini, A.・Bentivegna, F., 2016. ナポリ臨海実験所標本データベースの分類データベースとしての役割. 化石, 99, 33-46.

鈴木紀毅, 2016. シノニム・データベース PaleoTax for Windows に基づく放散虫研究の現状. 化石, 99, 15-31.

総務省, 2015. 平成24年度版情報通信白書. <http://www.soumu.go.jp/johotsusintokei/whitepaper/h24.html>

(2016年1月14日受付, 2016年1月16日受理)

Schewiakoff Collection に関する資料

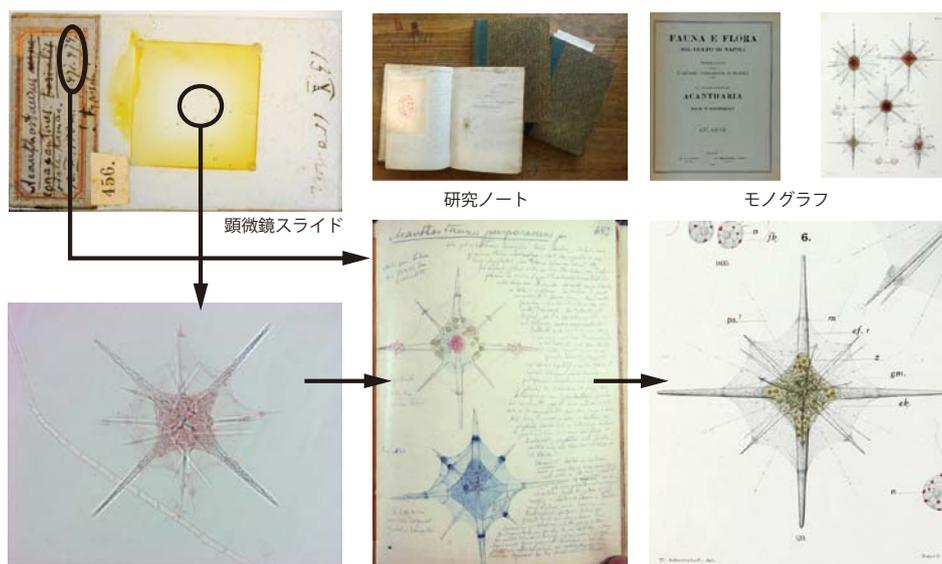


図1. Schewiakoff Collection と呼ばれる放散虫アカンタリア目の研究に使われた顕微鏡スライド、研究ノート、公表論文（モノグラフ）を照合した事例（木原ほか, 2016）。標本記述の照合が重要であり、両者の一致を確認すればデータベースとして活用することが初めて可能となる。特に、研究ノートの記述はきわめて有益となる。



図2. ナポリ臨海実験所の標本資料を事例としたデータベース。異なる分類群が一つの「瓶詰め標本」になっていたりと、学名は分類概念の違いで常時流動的であることから、データベースは、データの一つとして「コンテナ」単位でデータベース化している（木原ほか, 2016より）。

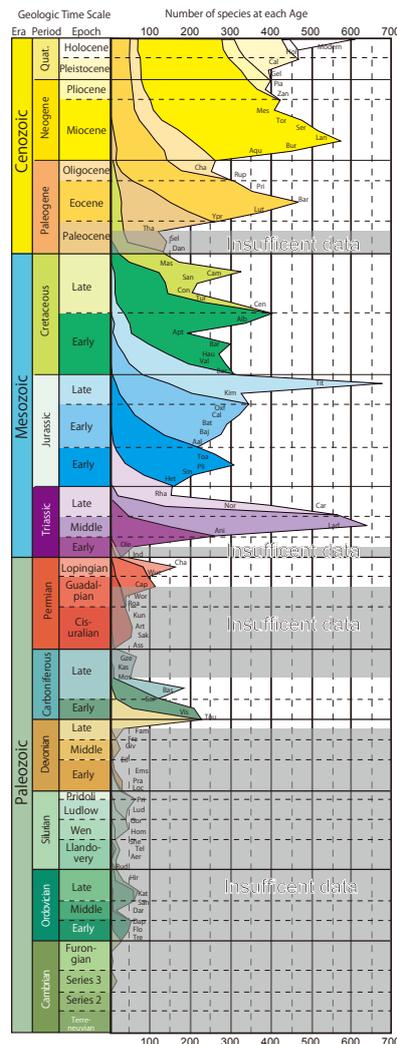


図3. 放散虫の既知種全種（15,040種）を現行の分類概念に整理した上（9,870種）で、地質時代ごとに存在種数を整理した図（2015年3月現在の入力データ）。このグラフが示す特徴には産地がそもそも少ない地質時代など、様々なバイアスがかかっているため、解釈する前にデータの精査が欠かせない（鈴木, 2016）。