

名前はなぜデータベースの問題となるのか

伊藤 希 (筑波大学 生命環境科学研究科・遺伝子実験センター)

我々は日々何らかの形で名前によって情報交換を行っており、名前をとりたてて問題のある何ものかとは考えない。名前と聞いてまず思い浮かべるのは人名であろう。言うまでもないことだが、人名すなわち個人名とは個人の名前である。個人名が指し示す対象ははっきりしており、間違っただけでその対象を呼べば「私の名前は中野です」などと名指しの対象の側で訂正してくれる。もちろん同姓同名はあるが、会話においては共通の知人という文脈からどの個人のことであるかわかる場合もあり、またわからなければ「どの鈴木一郎?」と問うことで指示対象を特定できる。同姓同名ということからわかるように、個人名というのは個人につけられた文字列でしかない。単なる文字列がどうしてデータベースの問題になり得るのだろうか。たとえば年金関連で問題となった名寄せの不完全さは、個人同定の問題であって名前そのものの問題ではない。この場合重要なのは個人個人が正しく識別され同定されることであって、名前は識別された個人の属性に過ぎない。

同じことが生物の名前、学名や和名といった生物分類群の名前にも成り立つだろうか。結論から言えば、生物分類群には個人名とは異なる問題が存在し、それは集合名についての問題とも異なるものである。我々が名前に期待する機能は、違うものは違う名前と呼ばれ、同じものは同じ名前と呼ばれることである。個人名の場合、個人個人が異なる個体である以上同じものは一つとして存在せず、何が同じで何が異なるかという問題は存在しない。これは個物であれば生物個体でも標本でも同じことである。一方、学名や和名が対象とする生物種やその上位の分類群について、何が同じで何が違うのかということは個人の場合のように明白ではない。こうした名前をデータベースで扱う際には、個物に従属する文字列を扱うだけではなく、文字列とその名指しの対象との関係を含めた扱いが必要となる。

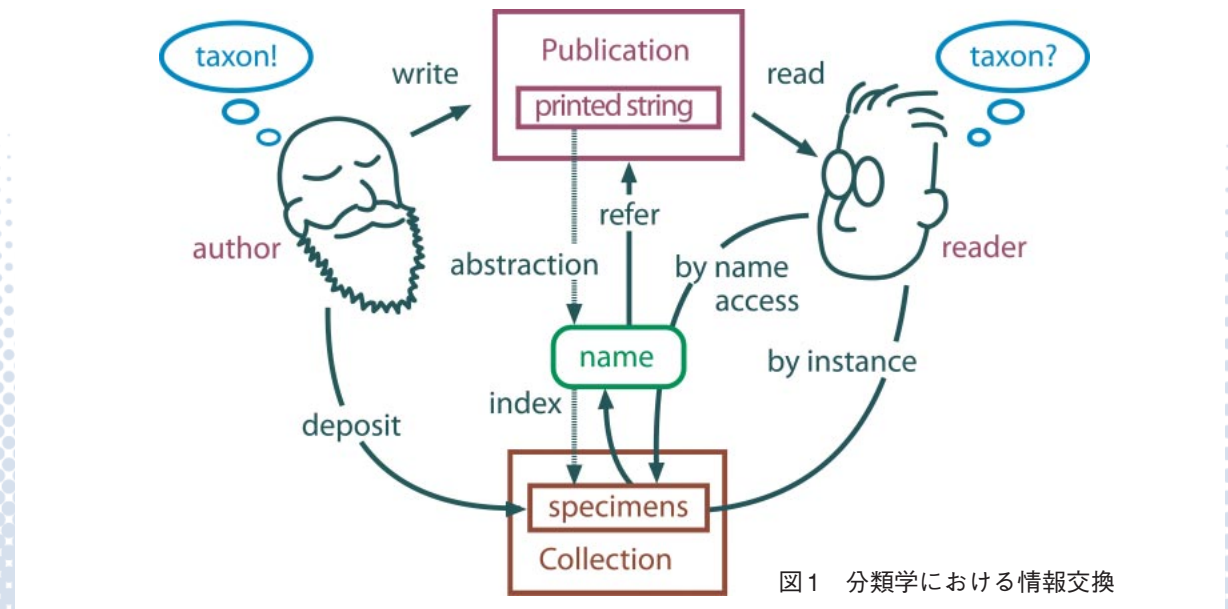


図1 分類学における情報交換

それでも、たとえば図書館の蔵書のように、指示対象が単なる個物の集まりでしかないならば、個人名同様の扱いが可能である。では、生物分類群と図書館の蔵書ではどこが異なるのだろうか。ある分類群に属する個体が増減するのと同様に、図書館の蔵書も増減する。したがって指示対象となる集合の外延が示す時間発展が問題なのではない。問題なのは、たとえば新種の発見によって、その種が含まれる属の内包が影響を受けるのかどうか、そのことが名前に影響するのかどうか、という点にある。こうした点について人間は柔軟に扱うことができるが、データベースで同様の扱いをするには相応の工夫が必要とされる。そうした工夫を含め、生物名をめぐるデータベースの現状を紹介する。

それでも、たとえば図書館の蔵書のように、指示対象が単なる個物の集まりでしかないならば、個人名同様の扱いが可能である。では、生物分類群と図書館の蔵書ではどこが異なるのだろうか。ある分類群に属する個体が増減するのと同様に、図書館の蔵書も増減する。したがって指示対象となる集合の外延が示す時間発展が問題なのではない。問題なのは、たとえば新種の発見によって、その種が含まれる属の内包が影響を受けるのかどうか、そのことが名前に影響するのかどうか、という点にある。こうした点について人間は柔軟に扱うことができるが、データベースで同様の扱いをするには相応の工夫が必要とされる。そうした工夫を含め、生物名をめぐるデータベースの現状を紹介する。

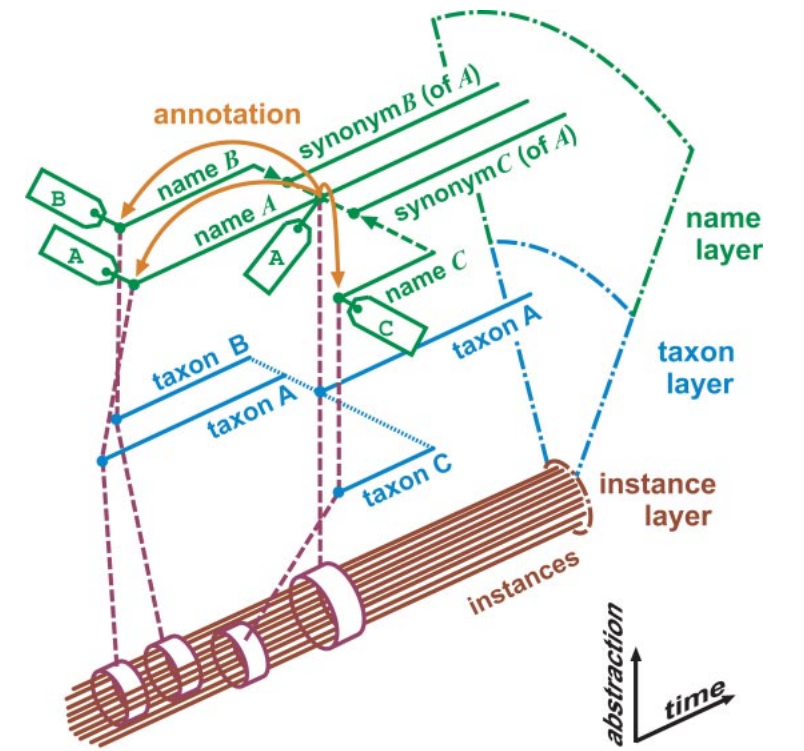


図2 分類概念と名前の変遷

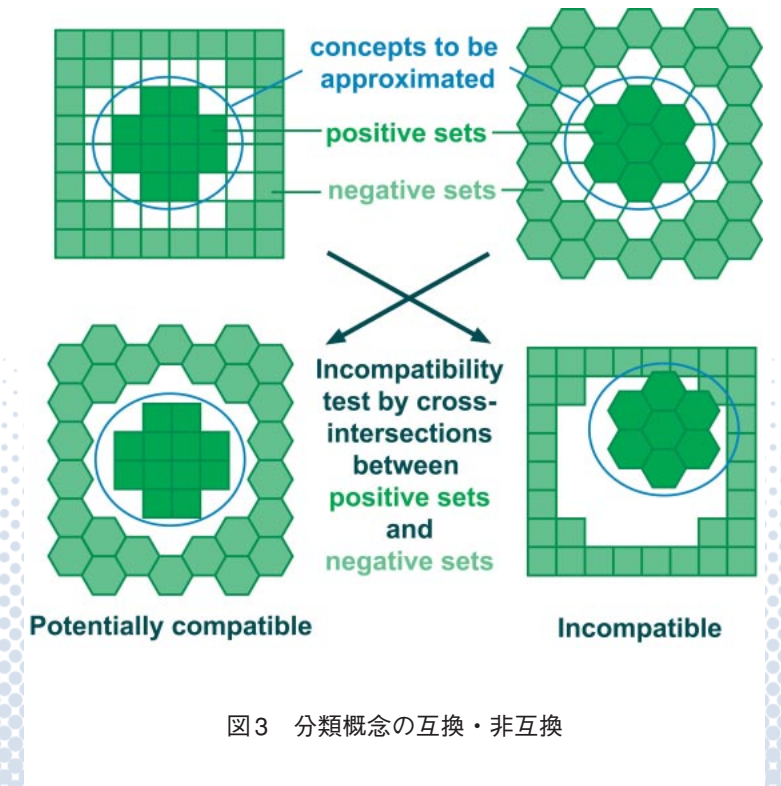


図3 分類概念の互換・非互換