

情報整理のルールとヒント

メタデータ	言語: jpn 出版者: 公開日: 2017-10-03 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	http://hdl.handle.net/2297/34375

第3章 情報整理のルールとヒント

森 雅秀

1. 情報整理の5つのルール

情報整理の心得を以下の5つの項目に絞って、それぞれについて解説しよう。

- ①情報は一元的に管理する
- ②情報はできるだけ規格を統一する
- ③一件のデータには一件の情報のみ
- ④整理することに時間と労力を費やさない
- ⑤自分の記憶にたよりすぎない

1.1 情報は一元的に管理する

「一元的」ということばがむずかしければ「ひとつの原理」や「ひとつの法則」と言い換えてもよい。逆の表現を用いれば「情報は分類して整理してはならない」となる。これは②の「情報はできるだけ規格を統一する」や、⑤の「自分の記憶にたよりすぎない」とも関連する。

もちろん「分類してはならない」と言っているのではない。「分類して整理してはならない」と言っているのである。「分類」とはすでにその作業において、何らかの思考の枠組みを有していて、はじめてできる行為である。「整理する」ことすでに知的活動ははじまっている。むしろ、分類することにこそ、オリジナリティを見いださなければならない。情報の整理はそのための地ならしのようなものである。分類するための枠組みを作っては壊し、作っては壊し、つねに別の可能性を探る努力を続けることで、新しい視点や方法が見つかる。それは、他者からのさまざまな批判に耐えるだけの強度を、自分自身の思考に与える作業なのである。

「分類せずに整理する」方法の典型は、野口の『「超」整理法』の中の「押し出しファイリングシステム」や「ポケット一つ原則」である。この方法の基本は「時系列で並べる」「整理する場所はひとつに決めておく」である。この場合、ふたつの点で「ひとつであること」が強調されている。すなわち「原理がひとつ」であることと、「場

所がひとつ」であることである。

はじめの「原理がひとつ」というのは、実行してみるとなかなか困難である。時系列に並べるのは「原理がひとつ」の典型であるが、野口の言うほど、ファイルとファイルの時間的な順序は、ファイルを作った本人にとっても明確ではない。とくに、時間がたてばたつほど、当然のことながら、記憶は曖昧になる。また、ひんばんに使う最近のファイルについても、出し入れがはげしいほど、時間的な順序はいいかげんになる。大まかな順序でことが足りると考えた方がよいであろう。

ファイルの数が増えるにしたがって、全体の整理をひんばんに行う必要もある。作ってから一度も見えていないようなファイルも、その中にはたくさん含まれるようになる。作ったという記憶そのものがなくなったり、同じ名前のファイルを二つも三つも作ってしまうこともめずらしくない。野口は「押出しファイリングシステム」を「使えば使うほど整理されていくシステム」と言っているが、それはやはり誇張である。どんな整理法でも、ある程度のメンテナンスはやはり必要である。

これにくらべて、もうひとつの法則である「ひとつの場所」の方が、より重要である。自分の探しているファイルが「必ずここにある」と確信を持てることは、作業をする上で心強い。少し時間がかかっても、必ずめざすファイルにたどり着くことができる。その場合、少しくらい時間的な順序が守られていなくても、ほとんど問題ない。野口の「押出しファイリングシステム」では、それぞれの見出しに作成した年月日を記入することをすすめているが、これは必ずしも重要ではない。日付を見ながら探すことはほとんどないし、何度も使うファイルには、日付そのものが無意味である。

仕事をする場所が一カ所の場合、整理する場所がひとつであることはそれほどむずかしくないが、われわれ大学教員のように、大学と自宅のどちらでも作業をするような人間には、やっかいである。私自身は押出しファイリングシステムは、大学の研究室に作ってあり、自宅にはない。一カ所にまとめることをあくまでも守っているのである。必要な書類があるときは、ファイルを持ち運び、終われば元にもどすようにしている。自宅に放置しておくことは避けている。

最近では、持ち運ぶときには袋ファイルではなく、透明なクリアフォルダーを用いることが多い。作業がすむまではそのまま、終わった段階で袋ファイルに入れて整理する。中身が見えるという状態は重要である。たまっている仕事がいくつもあるときには、かなりの数のフォルダーを持ちこぶこともあるが、いちいち取り出さなくても、全体の中身を短時間で確認できる。袋ファイルの見出しのみの情報よりも便利である。また、袋ファイル自体も多くなるとかなりかさばって、持ち運びに不便である。

押出しファイリングシステムは、パソコンの中のファイルの整理と連動させるとよい。

パソコンの中のファイルを世の人々はどのように整理しているのであろう。デスク

トップ上に、さまざまなファイルやフォルダー、あるいはソフトウェアのアイコンをならべている人を時々見るが、いったい、この人はきちんと仕事をしているのだろうかと思う。デスクトップにおさまるほどのファイルしか使っていないとすると、ずいぶん仕事の少ない人だと思ひ、どこかのドライブとデスクトップに分けているとすれば、なぜわざわざ分けなければならないのだろうかといぶかしく思う。

たいていの人は **Document** のようなフォルダーに、仕事の内容に応じていくつかのフォルダーを作り、さらにサブフォルダーを作って整理しているであろう。押し出しファイリングシステムのように、すべてのファイルをひとつのフォルダーに突っ込んで、時間順にならべている人もいるかもしれないが、パソコンでのファイルの増え方を考えると、あまりすぐれた方法ではない。ある程度の分類は必要であろう。

押し出しファイリングシステムと異なり、パソコンの場合、フォルダーを作ることは手間も材料もいらない。ならべるための書架も不要である。フォルダーやファイルのならべる順番も、時間順はもちろん、名称順でもサイズ順でも、あるいはファイルの場合、ファイルの形式でも可能である。強力な検索機能もパソコンにはそなわっているので、ひとつのキーワードからでもめざすファイルやフォルダーを探し出すことができる。

現在、仕事でも研究でも、ほとんどのファイルはワードやエクセル、PDFなどの電子ファイルになっている。実物のファイルは、むしろその予備的な役割を果たしているにすぎないことも多い。見方を変えれば、ひとつの情報は、主としてパソコン上で管理され、補助的に実物のファイルがあるという状況である。この場合、両者の名称をそろえておくことを心がけるとよい。パソコン上でひんぱんに目にするファイルやフォルダー名は記憶に残りやすく、それと同じタイトルの袋ファイルを探しやすい。

パソコンのファイルやフォルダーと、現実の袋ファイルが対応していると言っても、両者を同じように整理する必要はない。袋ファイルは時系列の押し出しファイリングシステムがやはり有効であるし、パソコン上ではある程度分類した方が探しやすい。むしろ、一件ごとの内容を両者のあいだでできるだけ統一し、一対一で対応するように心がけるべきである。「ひとつの原理」という法則は、むしろそのレベルで重要である。

1.2 情報はできるだけ規格を統一する

「できるだけ」であって「必ず」ではない。もともと情報というのは多様であり、それを無理に統一的な規格に押し込めても無意味である。

梅棹の京大式カードは、規格の統一というルールの典型である。自分の思考をカードという統一規格におさめ、それをシャッフルすることで新たな発想を生み出す。次のルールである③の「一件のデータには一件の情報のみ」とも関係し、ひとつのカー

ドにむやみに情報を詰め込まないことも、カード作成の原則である。

カードが現在でも有効かどうかについては、意見が分かれるであろう。カードの役割はパソコンでも可能と考える人は多い。タブレット型のパソコンや、スマートフォンでも同様であるし、文字にしなくても、録音することも可能である。すでにカードの時代は終わったと考える人も多いであろう。大学生協での京大式カードの売り上げは、おそらく微々たるものであろう。

しかし、それでもなお、紙に手で文字を書くという作業には、存在意義はある。それは、ノスタルジックな思い出というよりも、デジタル情報のもろさへの危惧がつねにあるからである。

パソコンに入力したデータは、あたりまえのことであるが、パソコンがなければ（あるいは立ち上がらなければ）存在しないことに等しい。そのために、バックアップのデータをせせせと作るが、それは何らかの記憶媒体（外付けハードディスクや USB メモリなど）であり、その再生（読み取り）機能がなければ同じことである。フロッピーディスクはすでに過去の遺物になりつつあるし、1990 年代の前半まで主流であった SCSI 接続の機能は、USB 接続の出現によって、あっという間に姿を消した。新しいメディアの出現とともに、過去のデータをすべて保存し直すという作業がくりかえされることになるし、それを怠れば、その情報は永遠に失われてしまう。

紙という情報媒体は偉大である。人類が千数百年にわたって使い続け、今なおその価値を失っていない。学生のころに作ったカードやノートが、そのままの形でカードボックスやファイルにおさまられ、いつでも取り出せる状態の人も多いであろう。そのスペースがもったいないと思う人もいるかもしれないが、古いパソコンやハードディスクなどをとっておくよりもはるかに扱いやすいし、コストもわずかである。もっとも、それを必要とする人物（基本的には作った人だけだろうが）がいなくなれば、むしろ、電子データの方が処分は簡単であろうが。

データを記入する形式はカードがベストであるとはかぎらない。すでに確立しているデータや知識であれば、ある程度の大きさをそなえたノートのような用紙の方が使いやすい。綴じられていない方がいいが、データ相互の順序に変更がない場合や、順序そのものが重要である場合は、綴じてあるノート形式でよいし、廉価である。私自身は、パソコンで入力し、プリントアウトしたものの他に、手書きでルーズリーフ形式の用紙を使って、講義ノートや、論文や本のためのデータを作成している。いずれの場合も、紙の表のみを使い、裏は使わない。コピーをとるときに片面と両面が混在した状態では不便だからである。手書きの場合、片方だけ罫線が印刷してあるレポート用紙や罫紙があるではないかと言われるかもしれないが、紙に厚みがないのが致命的である。用紙のサイズは A4 で統一している。日付や番号の記入欄も必要に応じて利用している。

梅棹はフィールドワークでこそ、カードを活用せよと述べているが、私の知る限り、

京大式カードをフィールドワークに使っている人はほとんどいない。共同研究を行うときに、データをカードに記入してストックするという方法をとることがあるようだが、共同研究が終わったときには、カードはどうなるのであろう。活用されないカードは、ただのゴミの山である。そもそも、人文系の研究者は、人に利用されるためにカードを作ることなど考えていないであろう。

フィールドワークをおこなう人は、おおむね小型のフィールドノート（野帖）を用いている（たとえばコクヨの緑色の SKETCH BOOK）。調査の期間や内容にもよるが、一回の調査に1冊使う。使っていないページが残っていても、継続して別の調査で使うのは避けた方がよい。タイトルと調査期間を表につけておけば、増えていっても整理に便利である。厚みはないが背の部分もあるので、背文字も入れることができる。野外での調査では、この手の小型のフィールドノートを用いることが多いが、図書館などの室内で調査をする場合は、A5版からB5版程度の大きさがあつた方が扱いやすい。ボールペンや万年筆を使えないことが多いので、鉛筆で書きやすい紙質のノートを選ぶことも重要である。あとでコピーがとりやすいように、ある程度、濃い色の芯の鉛筆を用いる。

「規格を統一する」というルールと矛盾するようであるが、「オリジナルの形態をできるだけ維持する」という心がけも重要である。これは、規格の統一のために必要以上の労力や時間、お金を使うべきではないという消極的な理由ばかりではなく、オリジナルの形態そのものが重要な情報であるという積極的な理由でもある。

たとえば、人文科学の場合、先行研究のほとんどは論文や本の形で残されている。研究を開始する時点では、どのような先行研究がその分野にはあるかを調べ、それを可能な限り蒐集する。本であれば現物を購入したり、論文ならば掲載誌そのものを入手することも可能であろうが、大半はコピーをとって集める。研究者間でやりとりする論文の抜刷もある。そのときに、拡大や縮小の機能を使ってサイズをそろえる必要はない。

日本のコピー機は、A4やB5などのA系統とB系統のサイズの紙がともに流通しているため、拡大・縮小の機能が充実している。ズーム機能でパーセントを細かく指定できることも多い。現在の公文書の規格でもっとも一般的なのがA4なので、すべてA4に統一することも可能であるが、できるだけ元のままの大きさをコピーすることをすすめる。紙のサイズと著しく異なる場合は、サイズ指定、枠消し、センター移動などのコピー機の機能を使えば、余白が黒くならず、きれいにコピーはとれる。

本や雑誌は、基本的にそのサイズにあわせて、文字の大きさや余白などが設定されている。拡大や縮小しても、情報そのものはかわりがないと思うかもしれないが、このような付帯的な条件も重要な情報である。人間は文字だけを読んでいるのではなく、印刷物全体をひとつのイメージとしてとらえながら読んでいる。人間は思考するときも、しばしば記憶を紡ぎながらおこなうが、そこで立ち現れる記憶は文字が示してい

る内容だけではなく、その文字を中心とした本や論文の全体的なイメージであることが多い。別の見方をすれば、人間は文字の情報を概念だけではなく、イメージとしても記憶に定着させているのである。そのときに、オリジナルのイメージと、コピーで実際に保存しているイメージが一致しないのは、むだにふたつのイメージを混在させることになる。

本や論文に図や写真がある場合は、言うまでもないであろう。最近のコピー機は精度が高く、白黒の写真などは、オリジナルと比べてもそれほど遜色なくきれいにとれるが、それをわざわざ小さくするのは、情報量をそれだけ捨ててしまうことになる。パソコンで画像ファイルを扱うときに、精度はそのままサイズを小さくすると、ほんの少し小さくしただけでも、ファイルの大きさがぐっと小さくなることを経験した人も多いであろう。縮小の比率は縦や横の長さの比であって、面積になるとその二乗の割合で小さくなる。

さて、オリジナルの規格を維持したままの資料を保存するときに用いるのが、オープンフォルダーや袋ファイルである。私自身は、論文や本のコピーの整理に、オープンフォルダーを使っている。オープンフォルダーではコストがかかると思う人は、角形2号の封筒でよいであろう。オープンフォルダーはスチール製のキャビネットに入れて保存するのが正しい方法であるし、それがベストと思うが、これにもコストがかかる。その点、再利用の封筒であれば、経費はほとんどかからないし、そのまま書架にたてられる。オープンフォルダーも書架にたてられるが、サイズとしては縦向きになるため、はさんだ書類が下から落ちてしまう。梅棹はこの方法を推奨するが、実際にやってみると不便である。フォルダーの厚みがばらばらなのも、うまく書架にならべられない原因となる。

論文や本のコピーを著者ごとのフォルダーに整理するのは、情報の一元的管理の典型である。これについては、本章の終わりで、あらためてくわしく説明する。

1.3 一件のデータには一件の情報のみ

このルールについては、すでに前項のはじめに言及した。梅棹は、情報を詰め込みすぎたカードのたとえとして、カメラの二重写しを用いている。たしかにうまい比喻であるが、カメラにフィルムを用いることがほとんどなくなった現在では、「二重写し」という言葉自体がすでに死語かもしれない。

このルールは、エクセルでデータベースを作るときに、一件のデータは一行におさめ、その中のフィールドの数や形式は、データが異なってもつねに一定にしておくという考え方に近いであろう。一件のデータが複数行にわたったり、データごとにフィールドの増減があるようなデータベースは役に立たない（もちろん、フィールドが一定であれば、一部のセルの内部が空でもかまわない）。もっとも、梅棹のカードはデ

データベースではなく、発想をつかまえるためのひらめきのカードであり、性格はかなり異なる。

梅棹が言うように、一枚のカード(一件のデータ)に情報を詰め込みすぎないのは、たしかに重要なことである。しかし、これは実際にやってみるとかなりむずかしい。情報は小分けにしすぎること、かえって全体をとらえにくくしていることもある。それよりも、人間が頭の中でおこなう「情報のシャッフル作業」(立花隆)の方が、はるかに優れているという考えもある。あるいは、野口は『「超」整理法』の中で、カード派と非カード派の違いは、知的作業において意識と無意識のどちらを優位に置くかというとらえ方でまとめている。

要するに、カードを使うか使わないかについて絶対的な尺度はなく、優劣というよりも、扱う人間のタイプによるにすぎない。ジャーナリストと研究者という立場の違いもあるし、論理的な緻密な文章を書くか、読み手を引っ張っていく勢いのある文章を書くか、あるいは、データ蓄積型の文章か、感情発露型の文章かなどの違いも重要である。基本的には、研究者の書く文章は、こなれていなくても論理的でしっかりしたデータにもとづく文章であることが優先される。それが、読み手に感動を与えたり、余韻を残すことができれば、もちろん申し分ないのであるが。

ここでもう一度、「一件のデータには一件の情報を」というルールに立ち返ってみると、これは、情報の収集、管理の場面だけではなく、情報の公開の段階でも重要な指摘であることに気づく。すなわち、論文や本などの発表の段階である。

学生のレポートを読んでいて、しばしば感じられるのは、言いたいことが何かわからないことである。これは、言いたいことがないからというよりも、言いたいことがありすぎて、まとめ切れていないことが多い。

そのときに心がけるのは、文章や段落のレベルでの「一件の情報をしぼる」ことである。文章は長いよりも短い方がわかりやすい。長い文章で、かつわかりやすい文章を書くためには、ある種のテクニックや慣れが必要であるが、短ければたいてい必要としない。短い文章では、その中で伝えられることはかぎられているし、曖昧な文章やねじれた文章になりにくい。

長い文章を書いたら、思い切って短く切って、複数の文章に分けてみるといい。意味のない「～であるが」というような接続の言葉や、冗長な「～であり、～である」というような表現が姿を消す。短い文章にしても意味がわかりにくければ、もともと論理的な展開や思考方法に問題があったからであろう。長い文章の方が知的な文章であるとか、レベルが高いなどということはまったくない。むしろ、短い文章をむだな接続詞(「また」「さらに」「そして」など)をできるだけ用いずに続けて書く方が、はるかに優れているし、技術が必要である。そのような文章には、しっかりした論理的な思考が必ず背景にある。

同じことは段落についてもあてはまる。むやみに長い段落を見ると、それだけで作

者の水準がわかるような気がする。ひとつの段落で言いたいことは、ひとつに絞るのが原則である。

段落ごとに主張する内容が明らかになれば、段落と段落のつながりがわかりやすく、文書全体で、その段落が果たしている役割がはっきりしてくる。それがさらに拡大されれば、節、章、そして論文や本などの文章全体で、言いたいことが明確になる。当然、これは逆算のようなもので、主題や結論がまず確定し、それを前提にして、文章のすべての要素、すなわち、章、節、段落、文、そして単語までもが、それぞれの役割を果たすことになる。そのときに、重要な意味を持つのが、「一件のデータには一件の情報のみ」というルールなのである。曖昧さ、わかりにくさの排除と言い換えることもできる。

1.4 整理することに時間と労力を費やさない

これはその人の自由である。時間とお金、労力がいくらでもあれば、好きなだけ情報の整理につき込めばよい。立花は情報の整理に疲れ果てた青年の話を紹介しているが、それを喜々として行っている人も世の中にはたくさんいるし、整理法とセットで廃棄法を示すことで、人々の共感を得ている人たちもいる（「捨てる技術」とか「断捨離」のたぐいである）。

むしろ、われわれ研究者にとっては、情報の収集と整理から、情報の発信にどこで切り替えるかが重要である。特定のテーマで何かを調べていると、どこまでも道が続いていくように思える。新しいことがどんどん出てくるし、読むべき本や論文がイモヅル式に現れる。何かを読めば、また何か別のものを読まなければならない事態が果てしなく続いていく。卒論や修論のテーマを決めて、データを集めても、なかなか実際に執筆に取りかかれない例を、毎年のように見ている。他人事ではなく、自分自身においても、「もう少し早く書き始めていれば」という思いは、どんな文章を書いても必ずつきまとう（この文章でもそうである）。

ここで重要なのは、一種のあきらめである。もう少し肯定的な表現をすれば、「きびしい自己観察」である。どこまで資料を集め、どこまでそれを咀嚼し、自分の考えが形成できているか、その見極めが求められる。どこまで行っても完全ということはありえない。すべての論文は「見切り発車」から出発している。本の執筆に至っては、ひとつの章を書きながら、次の章の資料収集を同時に進めているのが、むしろ一般的であろう（少なくとも私はそうである）。

このような場面においては、ある種の尺度を自分の中に持っていることがのぞましい。この程度まで準備できれば、なんとか最後まで書き通せるという判断の尺度である。それは絶対的なものではないし、他人が決めるものでもない。直感的な感覚であり、試行錯誤の中から身につけるものである。経験を積むことによって、自然に醸成

されるものであろう。

書く文章の長さや読み手によっても変化する。それはしばしば、一般に考えるのとは逆になることにも注意すべきであろう。たとえば、学術的な論文の方がハードルが高いというのではなく、一般向けの短い文章の方が、実際に書き出すための準備に、より時間がかかることもめずらしくない。

1.5 自分の記憶にたよりすぎない

梅棹の『知的生産の技術』の名言のひとつは、「カードは覚えるためではなく、忘れるために作る」である。今ではあたりまえのように感じるかもしれないが、発表された当時は、おそらく読者に「コペルニクスの転回」をもたらしたであろう。パソコンの普及とともに、その流れは加速度的にはよくなっている。紙に書いて資料を作るような場面はまれになり、ワード、エクセル、パワーポイントに代表されるソフトウェアと、インターネットの情報を操るのが、現代の「知的生産の技術」のようになっている（「情報リテラシー」ということばが、それに相当する）。そこでは記憶することとははじめから放棄されているかのようだ。

むしろ、そうして作り上げられたデータや文章が、自分のものと思いつくことの危うさに注意しなければならない。「過去の自分」も「未来の自分」も、「現在の自分」から見ればすべて「他人」であることを、つねに意識する必要がある。それは梅棹が想定していたのとは異なる局面ではあるが、結論は同じである。

2. デジタル時代の知的生産の技術

パソコンが職場でも、学校でも、家庭でも、その中心を占めるようになってかなりの時間がたった。研究生活どころか、日常生活ですら、パソコンなしでは考えられない時代となった。

パソコンによる「知的生産の技術」の本も、世の中にはたくさん出ている¹。しかし、その多くはテクニックを解説することに終始し、思考法や発想法を根底から覆すような本にはお目にかかったことがない。知的興奮を覚えることもほとんどない。「裏ワザ」のような情報を満載にしている本もあるが、あくまでも「裏」のレベルにとどまり、知的な活動の本筋に関わる内容は期待できない。

ここでは、そのようなテクニックについては極力取り上げない。デジタルデータを扱うための基本的な心構えのような内容にとどめる。ハードウェアにせよ、ソフトウェアにせよ、付属のマニュアルは必ず準備されているし、重要なソフトウェアには解

1 たとえば、最近のものとしてはトリプルウイン（2008）、長谷川（2011a, 2011b）など。

説書も市販されている。インターネット上でも、さまざまなアドバイスを見つけることができる。そのそも、パソコン関連商品のサイクルの短さを考えれば、あらゆる情報は瞬く間に時代遅れになってしまう。少しでも「日持ち」しそうな情報に限定して紹介する。そのために、私自身の経験からはじめたい。

2.1 累々たる屍

情報処理関係の授業の最初に、大学生（その多くは1年生や2年生）に、「耐久性を考えた場合、情報は何に保存すればよいと思うか」という質問をする。情報メディアについての質問で、USB メモリやハードディスク（HD）、DVD、ブルーレイディスクなどの答えが出てくるが、私があらかじめ準備しているベストアンサーはいずれでもない。それは「木簡と墨」である。あるいは、木簡よりすぐれたものとして「竹簡」もあるが、竹簡そのものを大学生は知らないこともあるので、考古学の発掘などを通して、比較的知られている木簡にしている。「平城京跡から出土した8世紀の木簡に、租税に関する記述があった」というような報道の形で知っていることが多い。

さすがに、木簡に墨でノートをとれとは言わないが、それに続いて、3種類のフロッピーディスク（FD）を見せる（図1）。大中小の3種類で、一番小さいのが3.5インチ、これは大学生でもたいていは知っている。しかし、使ったことのない学生も、少しずつ増えてきているようだ。次は中で、5インチ（正確には5.25インチであるが、一般には5インチと呼ばれていた）。3.5インチに比べると、ペラペラの黒いシートのように、はなはだたよりない。最後は一番大きい8インチである。外見は5インチのFDをそのまま拡大したような形である。5インチも8インチも、ほとんどの大学生にとっては生まれてはじめて見るものである。そして、いずれもそこにある情報は、今では読むことができないし、3.5インチもそうなりつつあると言うと、だんだん、こちらの意図がわかってくるようだ。



図1 3種類のフロッピーディスク

私をはじめでパソコンに触れたのは、大学院生の1986年頃であった。文系なので、理系よりも少し遅い。その少し前にはワープロ専用機の時期もあった。この頃の主流だった記憶メディアが5インチのFDであった。8インチの時代は私自身はほとんど知らない。使用していたパソコ

ンは NEC の PC9801 シリーズで、OS は MS-DOS であった。MS-DOS の FD をドライブに差し込み、レバーを下げてロックすると、おもむろにカチカチという音がした。FD には悲しいほどわずかの量しか情報を保存できず、少し長い文書になると、いくつかの FD に分けて保存していた。読み取り中にうっかり取り出そうとして、FD をだめにしてしまったことも何回かあった。

ちなみに、このとき使っていたワープロソフトは一太郎 (Ver. 3) で、ひとつの文書が 3 つのファイルで構成されていた。ソフト本体のディスクに加えて辞書ディスクが必要で、自分の作成したファイルを保存するためには、辞書ディスクと入れ替えなければならなかった。しばらくすると、パソコン本体に仮想ディスクを作り、立ち上げる時にそこに辞書をコピーするという方法を知って、作業が少し楽になったことを覚えている。

その後、80 年代後半には FD は 5 インチから 3.5 インチに変わり、その時代がかなり続いたが、すでに述べているように、現在ではほぼ終わりを迎えている。私自身は、NEC の PC9801 シリーズから Mac に乗り換え、ワープロソフトも WORD を使い始めた。これは、留学先の大学で、学生が使えるパソコンが Apple の Macintosh Plus (あるいは Macintosh SE) だったことによる。記憶メディアは 3.5 インチの FD で、しばらくすると、HD などを SCSI で接続する時代があった。

SCSI 接続は 90 年代の半ば頃に USB1.0 接続に変わった。SCSI 接続の機種が姿を消したのもあつという間であった。現時点では、USB2.0 が中心で、USB3.0 が少しずつ増えてきているが、大きく見ると USB 接続の時代がまだ続いている。これは、パソコンの世界ではかなり寿命が長い方であろう。

記憶メディアについては、手軽に使うのであれば USB メモリ、容量が大きい場合は外付けの HD というのが、現在もっとも一般的であろう。光学式ディスク (CD、DVD など) もある。

私自身は、FD から USB メモリあるいは外付け HD のあいだに、Zip というメディアの時代があった (図 2)。FD よりも一回り大きいだけであるが、はじめは 100MB、あとになると 250MB の容量があった (750MB もあったそうだが、使ったことはない)。今から考えれば、ほんのささやかな容量であるが、FD に比べるとはるかに大きく、またアクセスも速かった。同じ頃には 640MB の MO が日本では流行していたが、私の扱うワードやエクセルのファイル程度であれば、どれだけ保存しても、いっぱいになることはなく、Zip で十分だった。職場と自宅の間でのデータの持ち運びに重宝したし、SCSI



図 2 Zip

接続のあとには USB 接続のドライブもあり、さらに Mac 本体のベイにドライブを増設して使っていたこともある。

さて、Zip も含め、私の手もとには、この 30 年近くの間に使った記憶メディアが山のようにある。FD、HD、Zip、USB メモリ、そしてさまざまな光学式ディスクなどである。3.5 インチの FD と Zip は、外付けの専用ドライブ（USB 接続）が保存してあるので、今でも何とか読むことはできるが、5 インチはドライブがないので無理である。HD はクラッシュしたものも多く、その場合、ゴミでしかないが、いつかよみがえるのではというはかない望みから、捨てずにとってあるものもある。よみがえったところでゾンビのようなものであろうが……。もちろん、パソコンも何回も買い換えているので、それらをすべて残していたら、部屋は中古パソコンの倉庫（あるいは廃棄場）のようになっているであろう。

要するに、デジタルデータを扱うのは、このようにやっかいなデータとの戦いなのである。パソコンであれ、記憶メディアであれ、何らかの機材を通してしか、その内容を見ることができない。しかも、その形式は容赦なく変更され、前の時代への配慮はほとんどない。すでに死んだしまった（つまり読み取れない）データや、死ぬかもしれないデータを生きながらえさせるための延命を施したり（新しいメディアへの移行やデータ形式の変更）、いつ死んでもいいようにクローンを作っておいたり（バックアップのこと）、まことに世話が焼けるのである。それでも、何とか延命できているデータはまだよい。完全に死んでしまったデータは、どんなに嘆き悲しんでも、よみがえることはない。この世にまったく存在しないことと同じである（蓮如の有名な「白骨の御文」の「あわれというもなかなかおろかなり」という一節を思い出す）。パソコンの道を歩き続ける限り、その後ろには累々たる屍が連なっているのである。

最近の経験では、外付けの HD が異常を示すようになり、「修復できないが、読み取ることはできる」というメッセージが現れるようになった。これはさしずめ、脳死状態か、植物人間といったところであろうか。重要なファイルだけは、何とか別の HD にコピーすることができた。臓器移植の感がある。

2.2 バックアップについて

マーフィーの法則

大学でもっともよく耳にする「マーフィーの法則」は「卒論（あるいは修論）の提出日の朝に、パソコンのデータは必ず壊れる」である。「パソコンのデータ」のところに、パソコンそのものやプリンターを入れたり、あるいはプリンターのインクカートリッジがなくなったという別バージョンもあるが、いずれも基本は同じである。最も壊れてほしくないときに、パソコンやそのデータは壊れるようにできているのである。大げさでも何でもなくて、私自身の経験からしても、このような話を聞かなかつ

た年の方がめずらしい。卒論や修論を提出しようとしている持ち主の執念のようなものが、パソコンにはプレッシャーになるのかもしれない。「転ばぬ先のバックアップファイル」なのである。機材そのもののバックアップについては、研究室や友人のもので、いざとなったら借りられそうなものをあらかじめ確保しておくとして（友人が同級生の場合はむずかしいが）、ここではデータのバックアップについて、いくつか重要な点をまとめておこう。

オリジナルとは別の場所に

常識であると思うが、バックアップはオリジナルとは別のハードウェアに作るのが鉄則である。同じパソコンの中に複数のドライブがあれば、異なるドライブにコピーを作成することができる。ファイル名を変えれば、同じドライブでもコピーはできる。いったん別のドライブやハードウェアにコピーを作れば、もとのドライブでも、オリジナルとは別のフォルダーにコピーを作ることができる。いずれの場合も、本体のパソコンが壊れれば、バックアップを作る意味がない。どんな記憶メディアでもよいので、本体から物理的に離れた場所に、バックアップは作らなければならない。

バックアップは作りすぎない

コピーをいくつでも作ることができるのは、パソコンのよい点でも悪い点でもある。ワードの文書のファイル程度であれば、瞬時にコピーを作ることができる。現在のパソコンやHDなどの大容量化を考えれば、たいていのデータは二つでも三つでもコピーを作っておくことができる。とくに、外付けHDの容量は、ギガを超えてテラになってきたが、文字だけのファイルでテラのレベルに達しようとしたら、天文学的な数の文字が必要である。

バックアップのコピーを複数作るとは、いけないことではないが、重要なのは、どれが最新のデータであるかを把握できるようにしておくことである。最新の状態にあるのはオリジナルのファイルであるのが一般的であるが、最新の状態ではないバックアップの方のデータを更新してしまうと、不完全なファイルがふたつになってしまう。見かけだけ新しいバックアップの方のデータで、オリジナルのデータを上書きしてしまうと、オリジナルにのみ含まれていた新しい情報が、消えてなくなってしまう。コピーが複数あれば、その危険性はさらに増すことになる。

同じファイル名の古いデータで新しいデータを上書きするときにも、その逆の場合も、パソコンは許可を求めてくる。「新しいファイルで古いファイルを上書きする」ことのみ許可する習慣を身につけておくことも重要である。「古いファイルで新しいファイルを上書きしてもよいか」と聞かれたときは、それが本当に正しい操作であることを確信した上でおこなわなければならない。

やっかいなのは、複数のファイルを含むフォルダーの場合である。フォルダーでフ

フォルダーを上書きすることは、ファイルでファイルを上書きするのと同様、可能である。そのときに「古い」か「新しい」かについても、同じように確認が求められる。注意しなければならないのは、ここで新旧を比較してるのは、フォルダーそのもので、中のファイルではないことである。たとえば、同じ複数のファイルを含むオリジナルのフォルダーとバックアップのフォルダーがあるとき、それぞれ別のファイルを加工したとすると、あとで加工した方のフォルダーが、当然、新しいフォルダーになる。しかし、もうひとつのフォルダーの中の、その前に加工したファイルは、新しいはずのフォルダー内の同一名のファイルよりも新しい。フォルダーごと古い方で新しい方を上書きすると、新しいはずの一部のデータが、古いファイルに置き換わってしまうのである。

このような事態を避けるためには、上書きによるバックアップファイルの更新は、できるだけファイル単位でおこなうようにすべきである。そして、フォルダーごとバックアップを作るときは、古いフォルダーを残して、別途、新しくバックアップファイルを作ることとする。その場合、バックアップ同士の前後関係をはっきりさせておくことが必要である。これについては、次項でも述べる。

バックアップのフォルダーには日付を入れる

ワードやエクセル程度のファイルであれば、一回にあつかう数は知れているし、ファイルごとにバックアップをとることもたいした手間ではないが、これがデジタルカメラの写真のように、一度に大量のデータをあつかうときにはやっかいである。ファイルごとにコピーしていたのでは手間がかかりすぎる。フォルダーごとにまとめてバックアップを作る人が多いであろう。

そのときに、フォルダーに日付を入れておくとよい。そして、あらたにバックアップを作るときには、新しい日付にして、別のフォルダーとして、フォルダー単位で複数のバックアップを作っておく。現在の大容量の HD であれば、そのくらいの保存は十分できるであろう。つまり、上書きせず、フォルダー単位でいくつかのファイルのセットを作っておくのである。もちろん、10 も 20 も作る必要はない。バックアップのセットが 5、6 個以上になったら、古いものからゴミ箱に捨てていけばよいのである。ただし、デジタルカメラのデータの場合、オリジナルは加工せず、コピーを一式作っておいて、それを加工するのが原則である。オリジナルはもちろん、オリジナルとまったく同じバックアップのセット（初代のバックアップ）も捨ててはならない。

なぜ、画像データを例にして、このようなやっかいなことを説明しているかという点、大量のファイルをひとつのフォルダーで管理していることが多いのに加えて、画像データは「不可逆的」なデータ形式が多いためである。

画像データについては、あとであらためて説明するので、ここでは簡単に述べるにとどめるが、最も一般的な画像データ形式である JPG が、代表的な「不可逆的なファ

イル形式」に相当する。つまり、いったん変更してしまうと、もとはもどせないものである。ワードなどの文字のデータであれば、記憶を頼りに、もとの形に戻すことも不可能ではないし、プリントアウトしてあれば、それを手がかりに復元できる。しかし、画像データを加工して、JPG形式で上書き保存すれば、もとのデータは二度とよみがえらない。印刷してあっても同様である。何代か前のファイルがバックアップの形で保存されていれば、失われてしまったデータを取り戻すことが可能である。オリジナルの画像データはそのまま保存し、加工はコピーにするのもこのためである。

機械に依存する記憶媒体を信用しない

すでに本章のはじめに長々と述べたように、読めないデータは存在しないし、それを保存するメディアはただのゴミである。パソコンそのものも、壊れれば扱いに困る粗大ゴミである。

これまでに出現した記憶メディアの中で、今でもひんぱんに使っているのは、USBメモリと外付けHDくらいである。FDはMacのパソコン本体からドライブが消えた90年代半ばには、ほぼ主役の座を降りた。私自身は、その後も出版社へのデータの送付などで少しは使用したが、メールの普及とともにその役目も終えた。CDにバックアップを作っていた時期もあったが、焼くのに時間がかかりすぎて、実用的ではなかった。保管しておくのもやっかいである。画像データなど、比較的、大きなデータを郵便で送るときにも使っていたが、これもインターネット上でファイルのやりとりができるようになってからは、無用となった。

これらは、現役の記憶メディアであるが、いつまでもそれが続くか保証はない。利用しなくなれば、あつというまにそのためのデバイスが消えるのは、ZipやMOで体験しているし、SCSI接続のような接続方式も同様である。

現在、主流であるUSBの接続は、しばらくは安泰と思われるが、いつ新しい方式にとってかわられるかは、われわれ素人には予想できない。一時期、MacではFireWire（IEEE1394）方式の接続方式にかなり力を入れていたが、今では装備されていない。最近Thunderboltという方式があるが、これもどれだけ普及するかは不透明である。

機種や接続方式が永遠ではないことに加え、ケーブルや対応ソフトの問題もある。USBではUSB1.0とUSB2.0では同じケーブルが使えたが、USB3.0では差し込み口（コネクタ）の形態が異なる。記憶メディアが継承できても、ケーブルがないため使えなくなったという事態もこの先ありうるであろう。読み取るために特別なソフトが必要なことや、パソコンのOSが対応しないこともある。メディア側とパソコン側の双方で、バージョンアップによる旧バージョンの切り捨てが起こりうるのである。このような問題は、パソコンを使うときには避けて通れないが、記憶メディアの場合、そこに入っているのが、個人で作成した「世界にひとつだけのデータ」であることが多いだけに、失われれば永遠によみがえることはない。

これに対抗するのは、結局は物理的な手段でしかない。ほおって置けば消えてしまう（実際は読めなくなるだけであるが）データを、何とか新しい状況でも使えるように、こまめに引き継いでいくことである。新しいメディアや接続方式が出現し、それを導入したときには、旧の環境がまだ当面は維持されるとわかっている、新しい状況に移し替えておくことが必要である。その時は、重要なものだけ選んで保存し直しておくというようなまどろっこしいことはしないで、一切合財、すべてひとつのフォルダに入れておくぐらいの大胆さが必要である。選んだ時点では気がつかなくて、後から必要になったファイルが、すでに手遅れで読めなくなることはいくらでもある。自分自身の判断力など信用しないことである。

最近では、いわゆるクラウドが普及して、作業用のデータをインターネット上に保存できるようになった。便利な機能であるし、この先、ますます広がっていくと思われるが、データの脆弱性については、自分のところにある場合と何ら変わりはない。ある日突然、データが読めなくなったとしても、誰も助けてはくれない。バックアップのデータをとっておくことが重要なのは、どこにあっても同じである。

とりあえずプリントアウトする

アナログ的な発想と思われるかもしれないが、すべてのデータは、いったんプリントアウトしておくといよい。

パソコンの導入は、紙の消費の爆発的増大をもたらした。会議資料などをペーパーレスにした組織などを別にして、個人のレベルでは、簡単にしかも美しくプリントアウトできる状況で、あえてそれを使わないという選択はまずありえない。文書や表を作成すれば、まずプリントアウトするのが普通であろう。少し手直しすれば、また新しくプリントアウトして、古い方はいらなくなる。かくして A4 の反故が、職場でも家庭でも山と積まれていく。

これはやむをえないことである。むしろ、時代に逆行するようであるが、とりあえずプリントアウトすることは、記録を残すということで重要と考えられる。もとのデータが失われたとしても、プリントアウトしたハードコピーがあれば、もう一度入力すればすむことである。新しく考えて作り直す手間に比べれば、格段に楽であるし、すでに失われたデータは、どんなにがんばってもその百パーセント復元することはできない。そのときの後悔の念や不愉快な思いは、永遠に消えない。「逃した魚は大きい」のである。

プリントアウトする紙であるが、つぎつぎと増える反故、すなわち紙の裏を使っている人も多いが、これはやめた方がよい。ひとつは、はじめに印刷されている情報が、予想もしないところで見られるおそれがあるからだ。会議の資料などでは、必ずしも「マル秘」や「取扱注意」のような指示がない文書でも、場合によっては機密漏洩のような内容が含まれるし、組織全体の情報管理を揺るがすような可能性もある。

もうひとつの理由はまったく物理的なことで、反故は紙詰まりを起こしやすいことである。プリンターにとって、最もやっかいなトラブルのひとつが紙詰まりである。いったん印刷された紙は、たとえばレーザープリンターの場合、「そり（反り）」ができていて、それをもう一度使うと、しばしば紙詰まりが起こる。詰まった紙を無理に引っ張ったことで、プリンターに致命的なダメージを与えることがある。マニュアルを見ながら、ていねいに処理すれば、たいていの紙詰まりは解消するのであるが、いちいちマニュアルを確認しない人も世の中には多い。そういう人ほど、マニュアルをしまった場所を覚えていないし、詰まった紙は力づくで引っ張れば何とかかなると思っているようだ。

2.3 ファイル形式とデータの扱い方

拡張子

ファイルの形式を意識することは、パソコンを使う上できわめて重要である。ファイル形式を示すのは、ファイル名の後ろについている拡張子である。拡張子についてのポイントを、安直ではあるが Wikipedia の「拡張子」の項から抜粋してまとめるとつぎのようになる。

- ・拡張子（かくちょうし、**extension**）とは、ファイルの名前（ファイル名）の末尾につけられたファイルの種類を識別するための文字列。
- ・通常、ファイル名の本体と拡張子は "."（ピリオド、ドット）で区切られる。
- ・拡張子はアルファベットと数字の組み合わせで、歴史的には3文字以内が好まれたが、それ以上の場合もある。
- ・オペレーティングシステム（OS）においてファイルの判別のために任意につけられる。
- ・拡張子が適切でない場合は、システムまたはアプリケーションソフトウェアの動作に影響を及ぼす場合がある。
- ・逆に、これを応用して、拡張子を変更することで、バックアップとして利用したり、動作の変更を行うこともある。

おもな拡張子

txt	プレーンテキスト
doc, docx	Microsoft Word
xls,xlsx	Microsoft Excel
ppt, pptx	Microsoft Powerpoint
rtf	リッチテキスト形式
csv	コンマで区切られたテキスト
jpg	JPEG 形式で圧縮された静止画像ファイル

gif	256 色まで扱える静止・アニメ画像ファイル
zip	zip 形式で圧縮されたファイル
lzh	lzh 形式で圧縮されたファイル

パソコンを使っている学生を見ていると、拡張子を示さずに作業をしている者が多いことに気がつく。デフォルトでは「拡張子を隠す」という設定になっているらしい。ファイルを保存するときにも、拡張子を隠して名称を入力させるような初期設定になっている。逆に、何かファイルを開くときにも、拡張子があらかじめ設定された状態で、その中から選ぶようになっているが、それに気がついていない人も多い。たとえば、ワードであれば、ワードの書類のみを選んで表示するように「*.doc」（*はワイルドカード）や「*.docx」になっていて、はじめから、それ以外のファイルが表示されない²。自分で作ったファイルであれば、それでたいてい問題はないが、他の人が作ったテキストファイル（拡張子が txt）や、リッチテキストファイル（rtf）、あるいはデータベースに多い csv ファイルなどをワードで開こうとしても、見つからなくて困っている。いずれもファイル形式、とくに拡張子が何であるかを知っていれば、まったく問題にならないはずである。

ファイル名

ファイル名に何をつけるかも、その人のパソコン観のようなものを示して興味深い。かつて、MS-DOS の時代には、ファイル名はアルファベットや数字で 8 桁までで、ドットの後に拡張子がついていた。いつからか 8 文字という制約はなくなり、漢字や仮名も使えるようになった。人によっては、長々としたタイトルをつける人も多いが、短いに越したことはない。日本語のファイル名は、Windows と Macintosh のあいだでデータのやりとりをする際に、別の文字に置き換わるいわゆる文字化けが起こることを知っておいた方がよい。自分だけで使うのならばかまわないが、メールの添付ファイルにしたり、メーリングリストで送るときは、ファイル名が正しく伝わらないことがある。

ファイル名の中にピリオド（ドット）を入れると、そこから後ろは拡張子と見なされるので、避けなければならない。このほかにも、カンマやハイフン、カッコなどのいわゆる「やくもの」は、問題になることも多いので、避けておいた方が無難である。その代わりに、よく用いられる句読点のような記号はアンダーバー（_）である。半角カタカナや丸付き文字などの機種依存文字も使用禁止と思った方がよい。

文書ファイルを作るときに知っておくとよいのが、前にも述べたテキスト形式とリッチテキスト形式である。現在、日本語ワープロのシェアは圧倒的にワードが優位であるが、かつてはさまざまな入力ソフトがあったし、今でもその一部は健在である。

2 ワイルドカードについては本稿末の参考資料④を参照。

ワード以外のワープロソフトは、独自のファイル形式で保存することがあるが、これはワードでは読めないこともある（代表的なのは一太郎の jtd 形式）。あるいは、現在では読むことができても、将来的にバージョンアップされたときに、サポートしない形式（つまり読み取れない形式）になる可能性もある。ここでも「累々たる屍」が出現する状況がある。

そのときに効力を発揮するのが、文字情報のみに限定したテキスト形式と、それにフォントの種類やポイント数、行間などのワープロデータの基本的な情報も加えたりリッチテキスト形式である。とくに後者は、脚注なども含めて、ほぼオリジナルどおりのテキストとして、ワードをはじめとするさまざまなソフトで再現することができる。あまり一般的ではないソフトウェアで作成した文書ファイルは、リッチテキスト形式でバックアップファイルを別にひとつ残しておくことをすすめる。

2.4 画像データ

やっかいな画像データ

デジタルカメラの普及、低価格化にともない、多くの人々が画像ファイルと関わるようになった。携帯電話で写真を撮る人の方が多いかもしれないが、学術的なレベルで、携帯電話で撮影したデータを使うことはあまりないであろう。また、デジカメの場合、新しい商品が出るたびに、画素数は著しく増大し、データ一件あたりのファイルの大きさも、以前とは比べものにならないほど大きい。そのために記憶メディアの容量の増加や、パソコンの処理能力の向上が競われ、さらにそれがデジカメの性能向上をもたらすという「倍々ゲーム」がくりかえされている。

デジカメで写真を撮るときには、ファイル形式を何にするか意識しなければならない。最も一般的なものは JPG 形式であろう。サイズによって三段階の中から選べるものが多い。しかし、すでに述べたように、JPG は不可逆的な圧縮方式をとる。一件あたりのデータの大きさは小さく抑えられるが、圧縮をもとに戻すことはできない（デジカメの場合、はじめから JPG 形式で保存するので、オリジナルという概念そのものが成り立たない）。そのため、多くのカメラメーカーが準備しているのが RAW 形式である。これはちょうど、フィルムカメラにおけるリヴァーサル・フィルム（ポジ・フィルム、いわゆるスライド）に相当する。とくに、カラー印刷を前提としたデータの場合、RAW で撮っておくと、業者に喜ばれる。

RAW データの注意点

RAW 形式のデータの難点は、JPG に比べてデータが大きくなることである。大きいということは、メモリーカードあたりの撮影枚数が少なくなったり、パソコンでデータを開くときに時間がかかったりするということである。もちろん、バックアップ

などを作るときにも、コピーに時間がかかるし、HD上で大きなメモリを占める。

あつかいやすさを考慮してRAWとJPGの両方のデータが保存できるような設定がある。通常はJPGの方のデータで閲覧や作業を行い、イメージの加工や実際の印刷など必要に応じてRAWを用いるのである。しかし、ただでさえファイル数が膨大になりつつある現在、一件につき二種類の形式のデータを残していくのは、正直なところ、かなりやっかいである。また、JPGが不可逆的な圧縮形式のファイルであるため、RAWやJPGからはTIFFなどの可逆的な圧縮形式にして保存することを、メーカーは推奨しているようであるが、これもファイル数の増加をもたらす、さらに管理に時間や労力、コストがかかる。

RAWデータでもうひとつ注意しなければならないのは、RAWがひとつの共通した形式ではないことである。わかりやすくいえば、メーカーごとに独自のRAWがあり、互換性はほとんど期待できないということである。ニコンもキャノンもソニーも、いずれもRAW形式のファイルで撮影できるが、それぞれ別である。ちなみに、写真店のラボにRAWのデータを持って行っても、プリントは受け付けない。

たしかに、デジカメの写真データとしてRAWはもっともすぐれた状態のデータであるが、それは永遠に保証されたものではない。たとえば、あるカメラメーカーが倒産すれば、その会社のRAWデータは、しばらくすればサポートされなくなり、いつかは読み取れなくなる。もちろん、そこにデータはあるが、画像としてのイメージは再現されない。ここにも屍が生まれる。

JPGは現在では画像データとして最も普及しており、それに対応する機種やソフトがすぐになくなるような事態はそれほど心配する必要はない。しかし、各メーカーに依存するRAWデータは、それよりもはるかに脆弱な基盤しか持っていないのである。

ここでも、アナログ的な発想であるが、「とりあえずプリントアウト」の原則を守った方がよいようである。これは、学術的な写真よりも、思い出や記念になるスナップ写真のようなものにとくにあてはまる。とりあえずプリントしておけば、それ自体は捨てない限り、必ず残るし、アルバムに整理しておけば、パソコン内部のデジタルデータよりもはるかに目にしやすい。プリントした写真から、スキャナーで画像データにすることができるし、そのようなサービスは写真店でもおこなっている。画像データのプリントアウトは、自分のカラープリンターでも可能であるが、品質を考えれば、写真店のラボに出した方がよいであろう。お金を取るだけあり、さすがに家庭用のプリンターとはレベルの違う仕上がりになる（見た目にはそれほど違わないように見えるかもしれないが）。

画像データのファイル名の付け方

ファイル名に関しても、画像データ固有の注意点がある。たいていのデジカメは、撮影した順に0001、0002のように連番がデータにふられる。その前にはメーカー固

有のアルファベットの文字列がつくことが多い。何もしないでいると、ずっと連番が続いて、一巡したら、またはじめからになる。あるいは、メモリーカードを入れ替えたり、フォーマットすると、また 0001 から始まる（メーカーによっては、前の番号を引き継ぐものもある）。その結果、同じ番号のデータがいくつも作られることになるが、これは、一件のデータにつき、ひとつのファイル名という原則からはずれ、たとえば、同じフォルダにまとめたときに、一方が他方を上書きして消してしまうおそれもある。

少し別の視点になるが、画像ファイルのやっかいなこととして、検索がむずかしいこともあげられる。パソコンの検索システムが文字列を基本にしているため、当然といえば当然なのだが、画像データのファイル名に現れるアルファベットの文字列と連番のみから、特定のデータを探し出すことは、ほとんど不可能である。撮影時期でしぼって、サムネールを表示させて、内容を確認するというのが、一般的であろう。

カメラ自体が任意のファイル名をつける機能がある場合は、それを利用するとよい。あらかじめ設定しておけば、特定のグループのデータ（日時や調査場所など）には、すべて共通したわかりやすい名称と、その内部での連番が付けられる。撮り終わったあとの場合、ひとつひとつのファイルの名称を変えることはわずらわしいが、リネーム (rename) のソフトが入手できれば、それによって一括で変えるという方法もある。若干、手間はかかるかもしれないが、検索や保存のことを考えれば、十分それに見合う作業である。

3. ビブリオグラフィーの作り方

3.1 ビブリオの基本

人文系の学問は、まずビブリオグラフィーを作ることからはじまる。ビブリオグラフィー（たいてい「ビブリオ」と略される）は参考文献のリストのことで、論文の末尾にも「文献」「引用文献」、英語論文であれば、“Bibliography”、“Works cited”などの名称であげられる。

どんな分野であろうと、研究のテーマを決めると、先行研究の確認を行わなければならない。単行本や論文の形で、すでに発表されている文献を手に入れ、それらに含まれているビブリオから、さらに別の先行研究を知る。これらの作業は、そのテーマについての知識を学ぶために行うと考えがちであるが、むしろ逆で、どのような研究がなされていないかを知ることの方が重要である。見方を変えれば、自分の言おうとしていることが、すでに誰かによって言われてしまっていないかを確認するのである。もし知らずに言ったとすれば、無知をさらすことになるし、知りながら言ったらもっと悪質で、剽窃や盗作とみなされ、研究者生命さえ失われる。本や論文は、自分の主

張がそこに含まれていないことを保証してくれる保険のような存在なのだ。先行研究の意義を十分理解し、自分の研究を正しくその中に位置づけることが重要なのである。

本や論文を書くときに、どんな先行研究でも手当たり次第に入れて、できるだけぎやかな参考文献のリストにしようとする学生や研究者もいるが、これも気をつけた方がよい。どんな先行研究を見ているかは、その研究がどの程度のレベルであるかを示すバロメーターである。私の知り合いのある有名な研究者によれば、ビブリオと注を見れば、本文は読まなくても、その論文や本のレベルがわかるそうだ。貧弱な参考文献や注はもちろんであるが、量だけはたくさんあっても、あげられている文献が二流や三流のレベルであれば、本文がそれよりもすぐれていることはありえない。これは、典拠に関する情報を多く含む注に関してもあてはまることである。

さて、このように本や論文に掲載されているビブリオは、必要かつ最小限にしなければならぬが、その背後には、取捨選択する前の総合的なビブリオがあることがのぞましい。日頃から論文や本を集めながら、その文献に関するデータをきちんと加えていくことが肝要である。論文や本を書くときに、まとめて入力すればよいというのは、順番を完全に間違えている。

3.2 ビブリオ作成の実際

文献を入手するまで

論文のコピーを入手する場合を例にとろう。

おそらく、その論文についての情報は、何か別の本や論文、あるいはインターネット上のデータベースに示されているであろう。すでにその時点で、基本的な書誌学的情報はわかっていることが多いので、ビブリオの形式に従って入力しておく（形式については、本稿末の「参考資料①」を参照）。すでにビブリオのファイルが存在すれば、それに追加すればよい。

このとき、どのようにならべるかは、決定的に重要である。テーマにしたがって分類する人も多いが、やめた方がよい。配列はひとつの原理でという整理法のルールを守るべきである。このとき、最も一般的な原理とは著者名である。日本人であれば五十音順、アルファベット表記の外国人であれば、アルファベット順である。このふたつは混在させるわけにはいかないもので、それぞれ別のファイルにした方がよい。日本人でも外国語の論文や著書であれば、当然、外国語の論文に分類されるし、その逆も同様である。分野によっては、論文のビブリオはひとつにまとめることが一般的なこともあるが、その場合は、アルファベット順になるであろう。なお、中国人や韓国人で、漢字表記の著者名となる場合は、その分野の慣習にしたがう。

発表年代にしたがって文献をならべる分野もあるようだが、これはわかりにくい。論文末尾のビブリオ程度であれば、探すのはそれほどやっかいではないかもしれない

が、単行本ともなると、数ページにわたることもあり、発行年から目的の文献にたどりつくことは困難である。

発行年代は著者名の次にくる配列の原理として用いる。すなわち、同じ著者に複数の著作がある場合、古いものから新しいものへという順序で並べる。同じ発行年に複数の著作があれば、a、bなどのアルファベット加えればよい(2013a、2013b …)。このとき、発行の月や日まで調べて、前後関係を決める必要はない。ビブリオには発行年の情報は必ずあげなければならないが、それよりも細かい時間の情報はもともと不要である。a、bなどは単に区別するための添え字である。

ビブリオの一件のデータのどこに発行年代をあげるかは、とくにきまりはないが、配列順にしたがって、著者名のあとにあげることがのぞましい。これは、本文や注で著者名と発行年だけで、特定の文献を指示する方式をとる場合はなおさらである。この方式は、どちらかという、社会科学系で一般的であったが、近年は人文学でも広まりつつある。かつては、人文学ではビブリオそのものをあげる習慣も少なく、注ですべての情報を示していた。その場合、発行年は一件のデータの末尾にあげられるのが普通であった。この方式は、今でも人文学の単行本ではしばしば見かけるが、同じ文献をくりかえし注や本文であげるときに不便である。「前掲書」のような指示で済ませることがあるが、どこが「前掲」かわからないときも多く、合理的ではない。

文献の入手と整理

さて、必要な書誌情報をビブリオのリストに加え、実際に論文を入手したとする。直接、掲載誌や掲載書が入手できる範囲にあれば、現物を借りてコピーをとればよい。しかし、近くの図書館などがないときは、所属の大学図書館を通してILL(文献複写・貸借)依頼制度を利用して、コピーを取り寄せることになる。このとき、ネット上で依頼するのが現在では一般的であるが、すでに基本的な情報が入力されているので、コピー&ペーストですむ。

コピーの現物が届いたとしよう。すでにその書誌情報は入力されているが、それが本当に正しいかを、現物を見てチェックする。巻数や号数、ページ数が間違っていることはときどきあるし、タイトルや著者名の誤記、とくに表記の誤り(漢字と仮名、送りがななど)はひんばんにある。

論文のコピーの一枚目には、おそらく著者名とタイトルが含まれると思うが、掲載誌(あるいは所収の単行本)のタイトル、巻数、号数、ページ数、刊行年を書き加えておく。自分でコピーをとるときは、全体のタイトルと奥付のページもとっておくとよい。ビブリオに入力してあるから不要と思うかもしれないが、入れたはずのデータがなかったり(間違えて消してしまうこともある)、入力ミスをしていることもありうる。また、論文を書くときに参照し、注や本文で著者名と発行年、当該ページ数を書く程度であれば、いちいちビブリオのファイルを開く手間がいらない。

文献の現物を手に入れ、その情報を正しく入力したことを確認すれば、文献を整理する段階になる。ここでも分類はしないことだ。内容にしたがってテーマごとに分類しても、すぐにどこに分類したかわからなくなる。文献の数が増えれば増えるほど、混乱の度合いは進む。ここでも「ひとつの原理」にしたがって、著者名の順序をならべるべきである。最もわかりやすいのは、著者ひとりにつき、ひとつのフォルダーを準備し、見出しの部分は著者名を記入し、五十音順、あるいはアルファベット順にならべる。全体をファイリング・キャビネットに収容すれば、出し入れも簡単で、場所もそれほどとらない。書架にならべたり、専用のボックスに入れてならべることもできるが、効率の点ではファイリング・キャビネットがもっともすぐれている。

ビブリオの配列と実際の文献（コピーや論文の抜刷）の配列が一致しているのは、あたりまえのように思われるかもしれないが、重要なことである。キャビネットの中身が実際のデータ群であるのに対し、パソコンの中のビブリオのファイルは、その仮想的な存在である。実際のデータのインデックスと言うこともできる。それを「くる」ことは、実際のキャビネットの中身を見ることに相当する。

文献の数が10や20とかであれば、それをまとめておき、必要なものをそこから探し出すのはそれほどむずかしいことではない。しかし、文献の数が数百の単位、さらに千の単位になると、それをざっと見ることはほとんど不可能である。それに対し、タイトルなどの書誌情報のみを示したビブリオであれば、数百件のデータを見るのはそれほどむずかしくないし、何よりも、ワープロの検索機能が活用できる。特定のキーワードをタイトルに含む論文を、パソコンは瞬時に示してくれる。あとは、その文献が必要なものかどうかを確認し、必要とわかれば、キャビネットのところに行って、ビブリオと同じ配列で整理されている論文の実物を取り出せばよい。

このように、論文や本を執筆する最初の段階で、すでに手もとにある資料に関しては、たちまちひととおりそろえることができるし、未入手の文献であれば、取り寄せればよい。用が済めばキャビネットに正しくもどす習慣を付け、新たに入手したものはデータ入力を済ませておく。

日頃から、関係雑誌や論文集、単行本などのチェックを怠らず、すぐに必要ではなくても、関心をひいた文献については、コピーをとり、データをビブリオに追加しておくことを心がけたい。このようにして論文や著書のアイデアや素材が蓄積されていくのである。

ビブリオにはワードを

ビブリオのデータは、データベースの一種なので、ワードではなくエクセルで入力しようと考えている人もいるかもしれない。実際、エクセルで整理している人も見かける。しかし、経験上、ビブリオにエクセルの利用はあまり適切ではない。理由のひとつは、ビブリオのデータが、雑誌論文、論文集所載の論文、単行本とで、それぞれ

形式が異なるからである。論文であればタイトルは一重カギ括弧（「 」）であるのに対し、単行本は二重カギ括弧（『 』）に入れる。雑誌論文は発行元は不要であるが、巻数（あるいは号数、もしくは両者）が必要である。これに対し、講座ものなどの論文集は、編者名や発行所の情報を加えなければならない。外国語論文についても同様である。エクセルはフィールドの数や内容が一定のものには威力を発揮するが、このようばらばらのデータには不向きである。あえて使うのであれば、一件のデータをひとつのセルに入れて処理しなければならない。

エクセルを使わない別の理由は、エクセルの場合、書式の設定がわずらわしく、とくにセルの中で特定の文字列のみの書体を変えるのがやっかいなことがある。ビブリオでは外国語の単行本はイタリック体、雑誌の巻数はボールドなどの約束事がある。また、論文や本を書き上げたあと、出版社や編集者にデータを渡すとき、本文はワード、ビブリオはエクセルというのはいやがられる。ワードの中にエクセルのデータを入れることは可能であるが、コピー&ペーストにすると、セル形式が維持されたままである。テキストファイル（あるいは csv ファイル）にすると、書式がくずれ、もう一度、書式を設定し直さなければならない。わざわざテキストファイルにするのであれば、はじめからワードにしておけば、何の問題もないのである。

3.3 電子ファイルの論文データベース

キャビネットをパソコンに

これまでは、論文はコピーや抜刷などの現物で保管してきたが、最近は電子ファイルが増えてきた。各種のデータベースやリポジトリからダウンロードしたときはもちろんであるし、研究者同士の抜刷のやりとりも、PDF ファイルで済ませることが一般的になってきた。電子ファイルの普及によって、抜刷そのものが作られなくなった雑誌もある。そもそも、電子ジャーナルのみの刊行が、近い将来、紙媒体での出版に取ってかわるであろう。

このような電子ファイルについても、整理は同じ原理でおこなうべきである。一件の文献はひとつのファイルとして管理する。PDF ファイルが最も普通であろう。紙のハードコピーから PDF ファイルにスキャンする機械も普及している。一文献につき一件のファイルを基本として、それを同じ著者のフォルダーにまとめる。このとき、ファイル名を工夫しておくといよい。たとえば、私をはじめに刊行年を示し、そのあとに論文や本のタイトルのはじめの語をあげておく。デスクトップでフォルダーの中身をカラム形式で示したとき（Mac の場合である）、ファイル名でならべるようにしておく、年代順、すなわち、ビブリオの配列にほぼ一致することになる（同一年の複数のファイルは完全には一致しないが、その程度の不一致はまったく問題ない）。タイトルの一語目も入れておくのは、年代だけでは中身がすぐにはわからないためであ

る。はじめの語だけでおおむね区別はつくので全文は必要ない。

日本語論文の著者のフォルダー名は、はじめにひらがなの読みを入れ、続いて通常の漢字にする（たとえば「もり森」「やまだ山田」）。わざわざひらがなの読みを入れるのは、五十音順に並べたいためである。漢字の場合、漢字ごとで音読みと訓読みのいずれかの読みが優先的に選ばれていて、実際の名前の読み方に一致しないことが多い。外国語論文のアルファベット表記の場合、このような問題はない。

こうしてできた著者ごとのフォルダーは、五十音順のそれぞれの文字（「あいうえお…」）ごとのフォルダーにまとめ、さらに日本語論文は「あかさたな…」の各行のフォルダーにまとめると、4つのカラムを順にたどっていけば、必要なデータに必ずたどりつくことができる（外国語論文の場合、アルファベットですむので3列）。

このように、電子ファイルによる論文データのファイル群も、基本的に、ビブリオや実際のファイリング・キャビネットに正確に対応している。キャビネット全体がパソコンのHD上のひとつのフォルダーにすっぽりおさめられ、ビブリオのファイルがそのインデックスになるのである。

今後の見通し

これからは論文のデータが電子ファイル化する勢いはますます強まるであろう。このような電子ファイルによる論文データベースが、個人のレベルでも文献整理の中心になっていくことが予想される。私自身、現時点で日本語論文約4千件、外国語論文約2千件を電子化し、データベースを整え、活用している。パソコンに入れておくだけで、いくつものキャビネットをならべるスペースが省略でき、必要な論文をパソコンの前に座ったままで、簡単に取り出せる。しかし、実際のキャビネットを廃棄することは考えていない。それは、これまでに述べてきたように、電子データに対する根本的な不信感が消えることはないからである。

むしろ、このデータベースは教育的な場面で有効性を感じる人が多い。かつては、学生に論文を紹介するときに、実際のコピーや抜刷を貸したが、返ってこないことや、貸しているあいだに必要なことがあった。電子ファイル化されていれば、データをコピーしたり、メールで送るだけですむ。