

『AIの遺電子』に学ぶ未来構想術：2 貴方の考える未来社会像

10 アーカイブ衰亡史

応
般

宮本隆史 編 | 東京大学

〈ネットワーク〉に接続するシステムが排出する履歴情報を指す「アーカイブ」という用語を目にする機会は、今日では技術者による専門的な議論のほかではまずないだろう。しかし、近代の幕開けの時期である2000年前後の人々は、現在とは異なる意味でこの言葉をさかんに使っていた。当時の人々にとって「アーカイブ」とは、さまざまな情報の「集積」を意味した。ここでいう「アーカイブ」とは、書籍の画像データやテキスト・データ、映画作品の動画データ、音楽などの音声データ、建築物の立体データなど、非常に多様な種類のデータが集められ保存される場所のことであった。デジタル・データだけでなく、紙などさまざまな素材を媒体とした資料も、そこに収蔵されるべき対象とされていた。「アーカイブ」という語が意味した内容の広さには目を見張るものがある。

さらに興味深いのは、当時の人々が情報について、「集積」するような性質のものとして捉えていたということである。わたしたちは、百数十年前の人類が現在の我々とはまったく異質な認識枠組みの中に生きていたことを理解しなければならない。認知歴史学者デイヴィッド・サルカールによれば、21世紀の人々は情報を水や空気のような流体として捉えていたという（デイヴィッド・サルカール、『情報流体説と21世紀社会』、2123年）。このような世界観においては、情報は流れたり溜まったりする性質を持つものとしてイメージされ、その「フロー」や「ストック」について語られていたのである。しかし、情報が「集積」するという近代初期の捉え方について、現在のわたしたちはどのように理解できるだろうか。

そこで、「アーカイブ」の概念と制度の歴史を振り返ってみよう。20世紀までのアナログ・メディアの時代、「アーカイブ」という語が意味したのは、主に紙などのメデ

ィアに記録された資料の集積やそのための制度、とりわけ「文書館」のことであった。この文書館という制度は、21世紀の自然環境とエネルギー市場の急激な変化のために絶滅した数多くの公共的制度の1つであった（同じように失われた類似の制度として、図書館や博物館などがある）。意外なことに、当時の人々は紙という媒体が非常に堅牢なものだと考えていた。適切に取り扱えば、紙に記録された情報は世紀を超えて保存できるとさえ信じられていたのである。もちろん、この信念は無根拠なものではなく、実際に21世紀前半までは1000年以上前の紙資料が数多く残存していたのであった。

ところが、21世紀後半の急激な気候変動とエネルギー環境の変化は、それまでの2千年間に人類が蓄積してきた資料の大半が失われる原因となった。世紀前半のうちは、地球の温暖化は当時危惧されたよりも緩慢にしか進まなかった。温暖化に対する不安はたびたび表明され、時として激しい政治的論議の焦点にもなったが、結局は雇用や経済成長といった政治経済的課題のほうが優先されつづけたのである。しかし、事態は世紀後半に急展開を迎えることになった。地球の気温は50年間のうちに5.8度も上昇し、それまでの人類が経験したことのない規模の災害が起こった。中でも紙媒体資料の生存にとって決定的だったのは、アジアと北米の沿海地域における台風や豪雨などの水害であった。当時の世界経済の中心であった超大国インド・中国・アメリカ合衆国およびその周辺諸国の沿海部では、6～11月に巨大な台風や激しい豪雨が頻発した。この気候変動は非常に急速に起こったため、被災国の住民の生活基盤は大打撃を受けた。政府予算の多くは住民の生活維持に振り向けられ、「文化」事業は後回し

とされた。その結果、図書館、博物館、文書館などの資料保存施設には保守管理費用が回されなくなり、台風や豪雨による破壊を生き延びた資料も、多くがカビによる被害を回避することはできなかった。20世紀後半から21世紀前半に建設された資料保存施設は、比較的大規模な施設内に資料を集中的に保管するものであったため、風雨が吹き荒れる環境で空調設備が停止すると、カビへの耐性は非常に低かった。自動化された集中管理方式があだとなったのである。システムが停止すると数千年を生き延びると信じられた紙資料でさえも、短期間のうちに朽ち果ててしまった。もちろん、21世紀前半の人類の怠慢がこの惨事を招いたのか、それとも気候変動を人類が左右することなどそもそも不可能だったのか、環境学者たちの見解はいまでも割れている。

これら失われた紙資料には、歴史的な公文書も多く含まれていたが、幸運なことにそのほとんどはすでにデジタル化されていた。公文書以外にも、刊行されていた多くの文書や書籍がデジタル化されており、実際に紙媒体の資料が参照されることはすでに非常にまれになっていた。また、21世紀半ばまでには、ほとんどの文書は最初からデジタル形式で作成され処理されるようになっていた。紙媒体の文書は当時でもほとんどが歴史的な記録だったのである。そのため、紙資料が失われたこと自体は、国家運営や私企業のビジネスにとって大きな打撃とは考えられなかった。被災諸国の経験は、暴風雨の影響を受けない赤道近くや内陸部の国家にとっても先例として受け止められ、デジタル代替物がある資料については、実物資料を保存するための予算が各国で切り下げられていった。その結果、実物資料のほとんどが廃棄されるか放置されることになった。アジアの西方の砂漠には今日でも、朽ち果てた本や文書の廃墟が残っており、動物や浮浪者のすみかとなっているという。

デジタル資料という知識基盤に全面的に依拠して国を統治する新たな体制においては、すべての情報がデジタル化され、ネットワーク上で常時呼び出すことが可能であった。「インターネット」と呼ばれた当時の

支配的ネットワーク上では、常時稼働するサーバ・コンピュータが情報を提供しており、利用者はインターネットに接続していればいつでも情報を得ることが可能であった。この時代には、歴史家たちはほとんど絶滅危惧種と考えられた。誰でもいつでもインターネットを通じて歴史的な資料を読むことができ、分析プログラムにかけさえすれば、前世代の歴史家たちが書いていたようなモノグラフを自分のために自動生成することができたのである。2050年代には、この知識基盤の総体が、慣習的に「アーカイブ」と呼ばれるようになっていた。

60年代には記憶装置の高騰という問題が浮上する。それまでの通念では、技術革新により半導体の価格は継続的に下がり続けるとされていた。実際に50年代までは半導体価格は下がり続け、記憶装置も際限なく大容量化できるかのように思われた。しかし、60年代に入ると、気候変動に伴う電力料金の高騰と、デジタルデータの増加による半導体の需要の高まりによって、記憶装置の価格は上昇に転じる。人類は、集中と圧縮という方法で費用を節約することで、この問題を当面はやりすごすことができた。高騰した記憶装置を購入できずデータを自力で保持しきれなくなった組織や個人のために、各国政府が公共サービスとしてデータを保存・管理する政策が広く実施された。これに加えて、プラットフォームを持つ企業が、比較的低価格でデータを預かるサービスを提供した。かくして、ほとんどのデジタルデータは、個人の端末上ではなく、政府や大企業の記憶装置に集中的に保存されるようになった。

データの保存を担うことになった政府と企業は、圧縮技術を用いて効率化を図ることになった。鍵となったのは、メタデータ関連の技術であった。21世紀初頭には、インターネット上のすべてのデータに標準化されたメタデータを付与し、その関係を機械可読なたちで記述することで、コンピュータにその意味を与えるという、セマンティック Web の構想が提案されていた。しかし、こうした情報量の豊かなデータを作成することは、当時は費用に対する効果が見合わないと判断され、実験的な試みに終わっていた。ところが2060年代には、

この技術がデータの圧縮に有用であるとして、再び注目されるようになった。人工知能の開発が進んでいたことも費用を下げることに寄与し追い風となった。

この方式について、ごく単純化してその仕組みを素描してみよう。たとえば、次のような文があるとする。

1857年にインド大反乱が勃発すると、翌58年にイギリス東インド会社は解体され、英領インドはイギリス王の植民地となった。1877年には、ヴィクトリア女王がインド皇帝として戴冠し、英領インドの統治は帝国主義的色彩が濃くなっていく。

このとき、「イギリス東インド会社」や「ヴィクトリア女王」などの名詞（下線を引いた部分）について、識別子を付与して辞書に登録しておくのである。実際には、より詳細で複雑な辞書が作られたのだが、ここでは理解を助けるために単純化したものを示しておく（表-1）。こうして標準的な方法で語に識別子を与えてネットワーク上に辞書を公開しておけば、メタデータを通じて参照することができる。つまり、つぎのように文のデータを書くことができるのである。

1857年に<n id="01"/>が勃発すると、翌58年に<n id="02"/>は解体され、<n id="03"/>は<n id="04"/>の<n id="05"/>となった。1877年には、<n id="06"/>が<n id="07"/>として戴冠し、<n id="03"/>の統治は<n id="08"/>的色彩が濃くなっていく。

id	name
01	インド大反乱
02	イギリス東インド会社
03	英領インド
04	イギリス王
05	植民地
06	ヴィクトリア女王
07	インド皇帝
08	帝国主義

■表-1 簡略な辞書のイメージ

もちろん、ここに示したのは単純化した例であるが、仕組みは理解していただけだろう。コード化された文を利用者がダウンロードすれば、端末上の再生プログラムがネットワーク経由で辞書を参照し、完全な文を構成することができたのである。動画や音声についても、基本的に同様の考え方で圧縮が行われた。これによって効率的な情報交換が可能となった。こうした「アーカイブ」のシステムは、総体として2080年代までは非常にうまく機能しているように思われた。

しかし、2090年代初頭に事態が変わりはじめる。ユーラシアと北米の大国が瓦解し、数多くの継承国家に分裂すると、どの公文書をどの国家が引き継ぐのかという問題が発生したのである。その過程で、何者か（その文書の利害関係者たちであろう）によって膨大な数の文書が書き換えられ、内容が異なる公文書の異本が複数の国家に引き継がれることになった。デジタル形式の文書は効率的に扱うことができる反面、統治体制が揺らぎ監視が行き届かなくなれば、改竄も非常に効率的に行うことができた。紙媒体の文書は失われて久しく、ある文書のどの異本が正確なものなのか、判断することはすでに困難となっていた。大国の瓦解に伴って生じた騒乱の犠牲者の補償問題は大いにこじれ、今日でも継続している訴訟がある。

さらに問題を難しくしたのは、2101年9月の同時多発攻撃である。現在でもその実態は明らかにされていないが、「テロリスト」グループの犯行とされる攻撃によって、各国政府だけでなくそれまで世界規模で情報技術産業を1世紀にわたって支配していた主要企業や教育研究機関のデータ・センタが物理的に破壊されたのである。それだけではなく、この物理的攻撃に先立って、インターネットには通称「ディコンストラクション」ウイルスが放たれていた。このウイルスは、デジタル・データに与えられていたメタデータをランダムな文字列に書き換えるものであった。先に見たように、21世紀のインターネット上では、ほぼすべてのデータに標準化された規格で豊富なメタデータが付与されており、これによって情報が構造化され効率的な情報交換が可能と

なっていた。しかし、感染力の強いこのウイルスによって、メタデータが無意味なものにされてしまったことで、コンピュータが文を正確に読み取れなくなってしまった。アナログ形式のコピーが失われている世界では、コンピュータが読めない文はヒトが読めない文と同義であった。同時多発攻撃は、こうした状況において発生したのである。インターネットは、物理的にもデータ構造上も断片化されてしまった。

人類はこのとき歴史を失った。寸断されたインターネットにかわり得るネットワークを構築できる体力を持つ企業や国家はすでに存在しなくなっていた。メタデータを失ったデジタル・データは、支離滅裂な代物になってしまっていた。もちろん、紙媒体の情報もほとんどが失われていた。22世紀の人類に残されたのは、個々の端末上に残された断片的な情報だけであった。これらの諸断片どうしがどのように繋がるかを明らかにする作業が、現代を生きる歴史家の役割となった。異本どうしの関係を推論して断片的なデータをつなぎ、あるいは著者や読者の証言を取ることで、もとの文書を復元するための努力が続けられている。

前世紀の歴史家たちが聞けば驚くかもしれないが、いまや職業としての歴史家の将来は約束されている。歴史家の多くが公的機関に雇われ、公文書の復元業務にあっている。また、企業も有能な歴史家たちを引き抜きあい、過去のさまざまな情報の復元にいそんでいる。政府や企業の情報独占に対抗し、普遍的な知の構築という目標を掲げる市民運動であるプロジェクト・ライブニッツの活動もさかんである。このプロジェクトの近年の成果としては、書籍ではウィリアム・シェイクスピア全集、音楽ではグレイトフル・デッド全集、動画ではモンティ・パイソン・シリーズなどの復元・無料配布が話題になった。シェイクスピアやグレイトフル・デッドの作品は、前世紀のうちに自由に無料で配布されており、数多くの異本データが残されていたことが復元の役に立ったとされている。モンティ・パイソンの作品は、ある種のプログラマたちのコンピュータの中で、ウイルスによる破壊をまぬがれて保存されていた幸

運なケースであったという。

データの復元の過程で徐々に形成されてきたのが現在のわたしたちが基盤とする〈ネットワーク〉である。同時多発攻撃後のわたしたちは、どこか1カ所に情報を集中させることについて強い恐怖を共有する文化に生きている。〈ネットワーク〉は、断片的な情報をつなぎ復元する作業の過程でたちあがってきた、コンピュータどうしが対等に情報を交換しあう自律分散型のネットワークである。現在では、ほとんどの情報はこの〈ネットワーク〉上で交換されている。22世紀の情報は、もはや流れたり集積したりはしない。すべては断片化されており、絶えず相互に交換されつづけるものとなったのである。現在では「アーカイブ」の語は、壮大な知識の「集積」を意味することはなくなり、〈ネットワーク〉上での情報交換の履歴情報を指す技術用語としてかろうじて使われているのみである。

*編者解題：この文章は、複数の端末に残されていた異本データから、編者が復元したものである。サルカールの研究書（未復元）を参照していることから、発表年は2123年から新型ディコンストラクション・ウイルスが再流行した2148年の間のことであると考えられる。著者は明らかではないが、文体からこの時期に精神的に執筆していたイスマット・チャンによるものではないかと推察される。編者の呼びかけに応じて異本を提供してくれた〈ネットワーク〉利用者の諸氏に感謝する。

(2019年10月17日受付)

■宮本隆史 mt@jasmineorange.com

1979年、京都府出身。東京大学大学院総合文化研究科博士課程単位取得退学。現在は東京大学文書館特任助教。専門は歴史学。インド近代史、刑罰の社会史、デジタル技術とアーカイブの制度史に関心を寄せる。