講　演

デジタル記録の信用性

――インターパレス・プロジェクトの成果――

ルチアナ・デュランチ†
翻訳　古賀　崇‡

インターパレス第１・２期（1998－2006年）の目標

インターパレス（InterPARES）は1998年に開始されました。その目的は、芸術的、科学的、また電子政府上の活動に伴い生成される、データベースや業務用のシステム上のデジタル記録、および動的で感覚的経験を生じさせる、双方向的なシステム上のデジタル記録が、正確で信頼性のあるかたちで作成され、また技術の陳腐化やメディアの脆弱性にかかわらず、長期的にも短期的にも、作成者や社会全体にとって真正なかたちで維持・保存されることを保証するような、理論と方法を実体化することにあります。

言い換えると、インターパレス研究は、記録の保存はその作成時点から始まるという認識に基づいて、デジタル保存のための新しい理論および方法論を開発することを意図していました。

インターパレス第３期（2007－2012年）の目標

一方、インターパレス第３期は、プロジェクトの第１期・第２期で開発された理論と方法論を適用し、特にリソースの限られた機関や組織において存在する問題を解決していくことを目的としました。

インターパレス第１・２期の主な成果

インターパレス第１・２期の主な成果は、試験として参加する組織において試用され、必要に応じて改変されました。これらすべての主な成果およびその他の成果は、インターパレスのウェブサイトで入手可能ですが、著作権で保護されていません。つまり、誰でもダウンロード、複製、翻訳、配布、再発行でき、また特定のニーズに応じて改変できるということです。

１番目の主な成果は「方針枠組み」です。これは、記録の作成および保存を行う組織が方針を設定する上での指針となる、原則の枠組みです。「方針枠組み」は、記録作成者のための13の原則と、記録保存担当者のための13の原則から構成されています。

２番目の主な成果は、「作成者向けガイドライン」です。この中には、個人、専門家、小規模事業者、例えば医院や法律事務所といった、レコード・マネジャーを置くほど規模が大きくない組織向けに、デジタル資料を作成および維持するための提案が含まれています。これらのガイドラインは、デジタル記録を作成するいかなる個人、オフィス、組織も適用することができます。また、デジタル記録を正しく維持し、

†ブリティッシュ・コロンビア大学図書館学・アーカイブズ学・情報学研究科教授
‡天理大学人間学部准教授

その正確性、信頼性、真正性が確実に保たれるように管理、利用され、時間が経過してもそのアクセス可能性を維持できるようにするための指針を示しています。

これらのガイドラインは既にいくつかの言語に翻訳されています。ウェブサイトでは英語、フランス語、スペイン語、カタルーニャ語、ポルトガル語、中国語のものが掲載されています。日本語にはまだ翻訳されていませんが、もし有用であれば、日本の専門家コミュニティの方々がこれを翻訳していただくことを大いに歓迎いたします。

3番目の主な成果は、「保存担当者向けガイドライン」です。これは、アーカイブズ関係の機関、プログラム、ユニット、組織のためのデジタル保存に関する提案です。基本的には、「作成者向けガイドライン」をベースに作られているので、それを補完するものになっています。これらは、アーキビストや、記録の保存に関わるあらゆる人々のためのものです。作成者から受け取ったデジタル記録を保存するために必要な指針のすべてを網羅しています。これらのガイドラインは「作成者向けガイドライン」と同じ言語に翻訳されています。

4番目の主な成果は「真正性のための基準および基本要件」です。真正性のための基準は、記録を作成および維持する人が、記録の現用段階のいかなる時点においても、それが真正であると確実に証明できるための要件です。基本要件は、アーキビストまたは保存担当者が、長期間にわたって真正性を維持し、それを証明することができるための要件です。

5番目の主な成果は「ファイル・フォーマット選択ガイドライン」です。これは、保存に最も適したファイル・フォーマット、ラッパー（訳註1)、符号化のスキームを選択するための原則および基準を述べたものです。

6番目の主な成果は「専門用語データベース」で、3つの部分、つまり用語集、辞書、および3つのオントロジーで構成されています。用語集はインターパレスの文書に含まれる用語で構成されており、インターパレス研究者が使用する定義がされています。辞書には用語集と同じ用語が含まれますが、他のいくつかの定義も示しており、他の分野のものも含めた既存の用語集および辞書を出典として使用しています。3つのオントロジーは、諸々の用語と概念の関係を示す図式です。

7番目の主な成果は、2つの記録管理モデルからなります。これらは、システムを設計したい人、および考えられるギャップを特定しつつ手持ちの要素の分析をしたい人にとって、非常に重要です。「保存の連鎖モデル」(Chain of Preservation Model、COPモデル）は、作成から保存までの記録のライフサイクル概念に沿ったものです。「業務主導型レコードキーピングモデル」(Business-driven Recordkeeping Model、BRM)は、COPモデルとは対照的に、オーストラリアのコンティニュアム概念に沿ったものです。この概念に信を置く人達は、この概念を有効にするシステムを設計するためにこのモデルを使用することができます。今の話の流れからお察しいただけるかと思いますが、私はBRMを支持してはおりません。支持しているのはCOPモデルのほうです。しかし、インターパレスは世界中の記録専門家コミュニティのために貢献したいと願っていますので、あらゆる視点で有益なモデルを開発したのです。

8番目と9番目の成果は、オンラインで入手可能な2冊の本です。1冊目はインターパレス第1期のプロジェクトに、2冊目は第2期のプロジェクトに基づくものです（原註1)。

インターパレス第3期の国際的連合組織

インターパレス第3期の国際的連合組織は、各国または各地域を代表する以下のチームで構成されています。カナダ（合衆国のパートナーを含む）、ブラジル、カタロニア（スペイン）、中国、コロンビア、イタリア、韓国、マレーシア、メキシコ、ノルウェー（2009年まで）、シンガポール（2009年まで）、トルコ。このプロジェクトは、カナダ社会科学・人文科学振興会とカナダ・バンクーバーのブリティッシュ・コロンビア大学（UBC）から資金を得ています。プロジェクトの統括事務局はUBCの図書館学・アーカイブズ学・情報学研究科にあり、私がディレクターを務めました。

インターパレス第3期の知見

インターパレス第3期で得られた知見は3つのタイプに分けられます。概念的知見（これらの内、ここでは信用性の概念についてお話しします）、方法論的知見（これらの内、ここでは保存方法についてお話しします）、戦略的知見（これらの内、ここではアーキビストの役割についてお話しします）です。

デジタル記録 対 従来型の記録

デジタル環境において、記録の内容、構造または形式、およびメディアは、もはや密接につながってはいません。それらは個別に存在しています。

システムに保存された「記録」という存在は、コンピュータのモニター上に見えるものとは別個のものです。システムには、コンテンツ・データ、形式データ、構成データがありますが、これらのデータの表出としてスクリーン上に見えるのは、別の物です。すなわち、それは記録のデジタル上の要素について、ドキュメントの形式として見えているものです。

記録を保存するとき、私たちはそれをデジタル要素に分割します。記録を呼び出すときは、それらの要素を再度組み合わせて、複製を作成します。保存した記録にアクセスするということは、複製を作ることでもあるのです。つまり、私たちはデジタル記録を保存することはできません。できるのは、それを再生産する能力を保存することだけです。

これは、従来型の記録とデジタル記録の決定的な違いです。私たちは決して、デジタル記録の真正性をその記録そのものに依拠して検証することはできません。なぜなら、デジタル記録は多くの部分―システム内部にある部分とスクリーン上に見える部分―で構成されており、ある特定の時点で呼び出したものは常に、私たちが保存したものの新しい複製であるからです。従来型の記録では、その真正性を検証するために、私たちは記録が元々存在していた時点の紙、インク、封緘を分析し、また、その記録の伝達については、原本か、草案か、あるいは複製の伝達であったかを判断することができました（原註2）。その記録が本物であるかどうかを物理的に検証することができたわけです。デジタル記録は常に新しく、その真正性の実質的な証拠を得るためには、それが存在している環境と、保存の連鎖の記録から、その信用性を推測しなければなりません。

言い換えると、デジタル記録の信用性の評価は、その記録を作成し、長期にわたって保有している個人およびシステムをどれだけ信用できるか、にすべて基づいています。記録の専門家として私たちは、コンテクストが示すことを信じると申し上げています。そのような言説を検証する時が来ているのです。

信用とその法則

信用とは何でしょうか。信用は、私たちが行動するために必要な特定の知識を持たずに行動することです。つまり、私たちが持っていない情報を他の情報と置き換えることです。例えば、私は日本語がわかりません。だから、この講演の通訳者が皆さんに何を伝えているか、わかりません。しかし、私が知っているのは、通訳者が私の話を正確に通訳できると安藤正人教授は考えている、ということです。私は安藤教授を知っており、彼の意見は正しいと信じています。このように、私は安藤教授についての知識を用いて、私が持っていない通訳者についての知識を補っているのです。このようにして、私は信用を用いているというわけです。

信用は法則に基づいており、その法則は信用を与える人と信用を受け取る人に関連しています。信用する人（信用者）と信用される人（被信用者）との間の関係は、被信用者の4つの性質に基づいています。

被信用者の性質

被信用者に期待される性質とは、どのようなものでしょうか。

第一に、評判です。人は被信用者の過去の行動・行為を評価し、もしそれが良いものであれば、被信用者は信用できることになります。

第二に、実行力です。人は被信用者の現在の行動を受け入れた上で、懸案の責任を満たすために必要な要件と比較します。

第三に、確信です。つまり、責任が実行されるという期待に確信が持てることです。

第四に、これは被信用者の最も重要な性質ですが、能力です。つまり、被信用者は所定の水準で仕事を遂行するための知識、技術、適性、資質を持っているということです。

信用と記録

デジタル記録の信用性を、その記録自体で立証することができないのであれば、その記録を管理する人、すなわち保管者を信用しなければなりません。私たちは、記録の真正性を立証する代わりに、記録を管理する人（レコード・マネジャーもしくはアーキビスト）を信用しなければなりません。

もちろん、その場合、私たちは、記録を管理する人がデジタル記録を保存する仕事を実行するための知識や能力を持っていることを、その人の評判に基づいて、確信することになります。

信用の度合いは、対象となる資料の繊細さや、時間の経過と共に資料の真正性が失われることがもたらす結果に比例します。もし私たちが、ある特定の記録の真正性を証明できないとします。もし、その記録が友人への手紙であるなら、問題ありません。しかし、もしそれが家の契約書であれば、大きな問題となります。したがって、私たちが必要とする信用の度合いは、記録・情報の種類にかかわりますし、記録の真正性を証明することがどれくらい重要かにもかかわります。

さらに、私たちが信用する人、すなわちレコード・マネジャーまたはアーキビストは、誰かに対する説明責任を有し、自らの行動に責任を持ち、方針・手段・技術の枠組みのもとで作業する人でなければなりません。

デジタル記録の種類

デジタル記録の信用性の保持について議論を進めるために、私たちはまず、どんな種類のデジタル記録が存在するかを見ていく必要があります。基本的にデジタル記録は3つに分類されます。1番目の分類は、コンピュータに記憶された記録です。これらは従来型の記録と非常によく似ていますが、コンピュータの中にあるということが異なります。電子メール、レポート、あらゆる種類のテキスト文書、写真、描画などです。この種類のデジタル記録は、実際にそこに書いてあることの証拠、つまり内容の証拠として使われます。

2番目の分類は、コンピュータで生成された記録です。これらの記録は、コンピュータ・プログラムのアウトプットです。生成過程に人間の介入はありません。例えば、ある人がATMへ行ってお金を引き出します。システムが銀行に問い合わせて、その人がお金を持っているかチェックし、その人が口座からいくらのお金を引き出したかを記録するなどしていく中で、一連の記録がシステム内で生成されます。これらの記録はその内容ゆえに使用されるのではなく、ある人の取った行動、すなわちその人が銀行のATMを使って何らかの行為をし、いくらのお金が引き出されたかの証拠として使用されます。

3番目の分類は、上記2つの組み合わせです。例えば、ある人がスプレッドシートを作成すると、その人はそのスプレッドシートに自分のデータ、すなわち人間の言葉を入力します。しかし、スプレッドシートのプログラムはそのデータを加工し、その結果できた記録はコンピュータに記憶されるとともに生成もされます。スプレッドシートのような記録を持っている場合は、文書形式のみを保存することはできません。その機能、すなわちシステム内でそれが作動する方法も保存しなければなりません。

したがって、コンピュータに記憶された記録を保存する時には、スクリーン上に見えているものを維持すれば十分です。コンピュータで生成された記録を保存する時には、その記録が相互に作用する方法を保存しなければなりません。しかし、スプレッドシートを保存する時には、その両方を保存する必要があります。

信用性の種類

どのように保存するかに進む前に、従来の信用性の意味を考えておきましょう。記録の信用性には3つの属性、すなわち信頼性、正確性、真正性が含まれます。

信頼性（reliability）は、事実としての記録の信用性です。例えば、私の市民権証書は、私の市民権の証拠です。ですから、私はそれを私の市民権として信用します。私たちが記録に書いてあるとおりのことを信用するとき、私たちは額面通り、疑うことなく、それを信用してきました。私たちは誰が記録を書いたかを見て、その人物が信頼できる人であれば、それを信用します。記録が完全であり、すべての部分がそこに存在していれば、その記録を信用します。記録の作成過程が管理されていれば、その記録を信用します。私たちが記録の内容を信用できるのは、それを書いた人、形式、過程を推測によって信頼するからです。例えば、診断書に医師の署名があれば、それを信用しますが、看護師の署名だったとしたら、信用しないでしょう。

正確性（accuracy）は、記録の中のデータがどれだけ正確か、精密かに関することです。これを評価する要素は、信頼性を評価する際の要素と同じものに基づきますが、加えて、内容を記録および伝送する方法のコントロールについても評価します。例えば、行と列のある表を伝送するとき、データの場所が変わっ

てしまうことがあります。伝送は、記録の保存の連鎖において最も弱い箇所になります。

真正性（authenticity）は、記録それ自体の信用性です。つまり、記録が作成された時点から、誰によっても手を加えられていない、あるいは記録が損なわれていない、変えられていないということです。真正性は、記録がその同一性と完全性を保持していることを意味します。

信頼性

ここまでは、従来型の記録をめぐる環境において私たちが信頼性、正確性、真正性を評価してきた方法を説明しました。デジタル環境においては、信頼性の観点から見ると、記録の出所はやはり重要であると考えられます。私たちは記録の出所が信用できれば、その記録には信頼性があると考えます。しかし、デジタル記録の場合、出所は信頼できる人物もしくは信頼できる手段だけではなく、プロセスもしくはソフトウェアも含まれます。つまり、その記録を生成するソフトウェアを信用できれば、その記録は信用できるものであるということです。

このことから、ソフトウェアはオープン・ソースでなければならないと言えます。なぜなら、もし私たちが記録作成プロセスの信頼性に基づいてその記録の信頼性を評価しようとするなら、そのプロセスがどんなものであるかを知る必要があるからです。つまり、ソフトウェアが独占的なもの（proprietary）であるとすると、私たちはそれを知ることができないのです。したがって、私たちはプロセスもしくはある結果を生み出すシステムを説明できなければならないし、そのプロセスまたはシステムを実際に示して、同一の結果を生み出すということを証明する必要があります。

正確性

従来型の記録の正確性を評価するためには、その記録が原本であることを証明すれば十分でした。なぜなら、記録の中のデータのみを、気づかれないままに変えることは不可能だったからです。しかし、デジタル記録においては、データを気づかれないままに変えることが可能なので、記録が正確であることを証明できるのは、同じ作成プロセスを繰り返した時、同じ結果が得られる場合だけです。したがって、証拠を同定し、それがねつ造されていないことを立証するために開発されたデジタル・フォレンジック（訳註2）の基本的な教訓の一つは、反復性です。正確性の検証は、その記録に関して実行されたすべての行為に関する文書化作業（documentation）によって支えられる必要があります。そのためには、記録に対してなされたすべての事項を記録できなければなりません。

もちろん、プロセスを再現するためには、そのプロセスを知るために、オープン・ソース・ソフトウェアを持っていることが必要です。このことは、デジタル記録の変換やマイグレーション（移行）を行うアーキビストにとって特に重要です。ここでの「変換やマイグレーション」とは、記録を異なるメディアに移行させることや、そのメディアが旧式になった場合には異なるOSに移行させることなどが挙げられます。アーキビストは、誰がプロセスを実行しても、またいかなる条件下でプロセスが実行されても、同じプロセスが同じ結果を出すということを証明できなければなりません。

真正性

真正性については、デジタル記録のコンテクストに依存しなければなりません。すなわち、手続き的コ

ンテクスト、文書的コンテクスト、技術的コンテクストです。また、記録の同一性と完全性にも依存することになります。

記録の同一性は、日付、筆者、署名、封緘、分類コード、登録番号等によって規定されていました。現在では、記録の同一性はメタデータの中にあります。したがって、必要なメタデータが保存されることが非常に重要になります。特に、その記録と他の記録との関係、すなわち記録の文書的コンテクスト(documentary context)を示すメタデータが重要です。

完全性とは、記録が伝えようと意図したメッセージが実質的に変えられていないということを意味します。しかし、それはどういうことでしょうか。

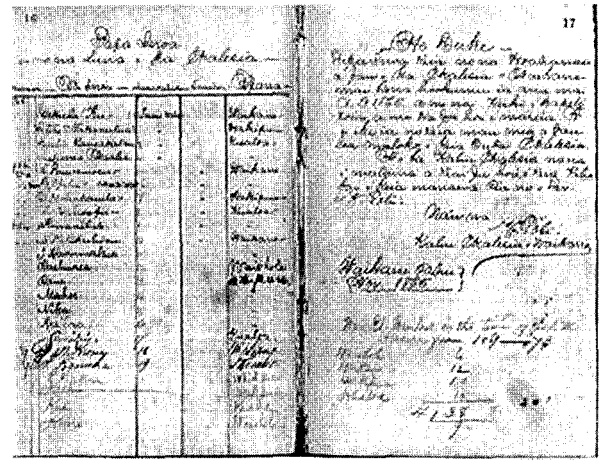

図1

完全性

従来型の記録において、ある記録がもはや完全性を持たないというのはどのような時でしょうか。穴が1つ開いてしまった時でしょうか。穴が2つになった時でしょうか。あるいは黄ばんだ時でしょうか。インクがにじみ出してしまった時？切れてしまった時？しわが寄った時？このような判断をするのに、私たちは常識に頼ってきました。

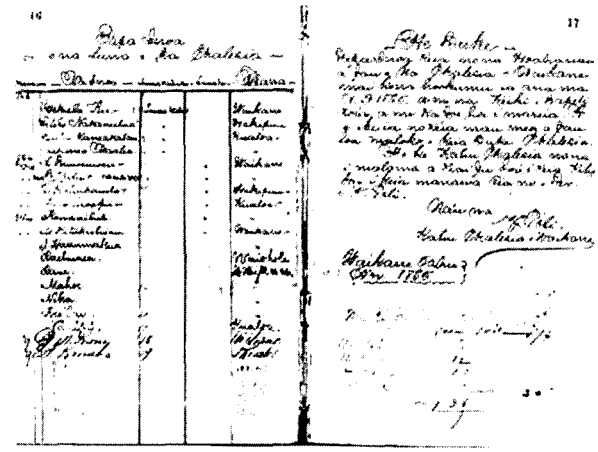

図2

データの完全性

デジタル環境において、完全性に関して最も重要なことはビット毎の完全性です。すなわち、意図的か、偶発的かにかかわらず、適正な権限なくデータが改変されていないということです。

完全性の喪失：アナログ対デジタル

この従来型の記録を見てみましょう。この記録の完全性を評価しなければなりません。

[図1] この記録に完全性があるということは、どうしたら分かるでしょうか。

[図2] この部分は少し色褪せていますが、まだ読むことはできます。

[図3] この部分は少しぼやけていますが、それでも完全性があります。もし拡大鏡を使えば、私たちはこれを読むことができます。したがって、私たちは、これが真正な記録であると言うことができます。な

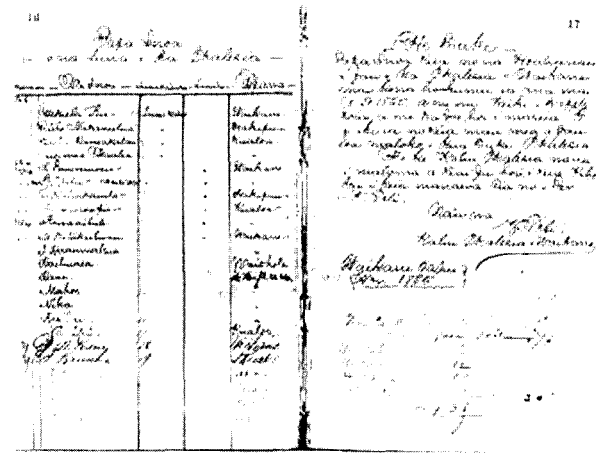

図3

ぜなら、全部がそこにあるからです。

オリジナルのビットが［二進法で］101であるデジタル記録を考えてみましょう。つまり［十進法で表記すると］5です。

しかし、順序を変えて110にしてしまうと、それは6になります。

011に変えると、3になります。

同じビットでも異なる意味になります。ビットの順序を変えると、同じビットでも異なる意味になってしまいます。

したがって完全性は、劣化の度合いとはもはや関係が無いのです。

データ改変から記録を守る

デジタル環境において完全性が失われることとは、主にデータの改変にかかわります。改変は、許可およびアクセス・コントロールによって防ぐことができます。パスワード、ファイアウォール等、またチェックサム（訳註3）やハッシュ・アルゴリズム（訳註4）といった強力な手段もあります。私たちは、記録を基本的に一連の数字に変換するソフトウェアを使いますが、同じ記録に同じソフトウェアを使用する度に、同じ数字の列を得る必要があります。もしそうならない場合は、その記録の中の何かが変えられたということになります。

偶発的な改変を避けることは、意図的な改変よりも困難です。なぜなら、記録を立証し続けるための特殊なハードウェアとソフトウェアが必要だからです。

デジタル環境では、コンピュータをオン・オフする度に、システム内の記録に変化をもたらす可能性があります。また、システムにアクセスする度に、記録が存在する環境を変えてしまいます。犯罪の記録を押収する警察の専門家は、まずコンピュータの電源を切り、それからハード・ドライブのイメージを作成します。彼らはコンピュータ内部の何物をも探そうとはしません。なぜなら、そのことにより偶発的にビットを変える可能性があるからです。そうではなく、ハード・ドライブのイメージを捜索するのです。

［変更を］防ぐことは非常に重要ですが、変更がいつ発生したかを見つける能力も同じくらい重要です。デジタル記録の保存が非常に困難な仕事である理由はここにあります。つまり、変化が起こっていないかを常にチェックしなければならないのです。

どうやってチェックするのでしょうか。それには、ログを使います。ログは、システムにインタラクトした人が行なったすべての行為をたどれるようにする為に、システムが自動的に作成するファイルです。

ログについては後ほどお話しますが、その前に複製の完全性についてお話しなければなりません。

複製の完全性

従来の環境では、私たちが記録の複製を作る時、その正確性について問題視しましたが、オリジナルがまだ残っているので、複製の完全性に関してはそれほど懸念していませんでした。

デジタル環境においては、私たちが複製を作る度に、それは私たちの記録になってしまいます。もはやオリジナルは無いのです。複製の完全性は、もし私たちが記録またはデータセットを持っている場合、複製を作るプロセスがその記録またはデータを変えることがなく、その複製はオリジナルの記録またはデータの正確なデジタル複製であるという事実と定義されます。

実際、私たちが複製を作る度に、私たちが生み出す物はそれまで持っていた物とほんのわずかに異なります。複製を作った時、その複製にタイム・スタンプを付けておくと良い理由はここにあります。つまり、異なる時点で複数の複製が作られた場合、それらは異なった時点で作成されたために違いがありますが、捏造されたものではないということが分かるのです。

複製とイメージとの違いを理解することは重要です。なぜなら、フォレンジックの専門家にハード・ドライブの複製が欲しいと言った時、フォレンジックの専門家はそれを「ハード・ドライブのイメージ」と捉えますが、この２つは全く異なるのです。

ディスク・イメージとは、ハード・ドライブのビット毎の再生です。ハード・ドライブ上のデータの完全なディスク・コピーは、すべての空きスペース及びすべての削除済みファイルも含んでいます。

ある大学の文書館は、学部長たちのすべてのハード・ドライブのイメージを作成している、と聞いています。なぜなら、こうすれば、コンピュータが故障したり、資料が偶発的に失われたとしても、何が生成されたかについての完全で正確な記録が残るからです。しかし、それらは法的に非常に不安定な基盤に立っています。というのは、イメージは学部長たちが削除したすべてのファイルをも保存するからです。したがって、この手続きにはプライバシーの問題と同時に、倫理的な問題もあります。

もっと良いやり方は、リアルな複製をつくること、すなわちファイルを選択的に複製することです。ハード・ドライブ上に刻まれたものすべてではなく、見える物だけを複製します。なぜなら、特定のファイルにアクセスする為には許可が必要になるからです。したがって、デジタル・デバイスに関する不完全な形相を有することになりますが、私たちの保存責任において、私たちが完全な形相を持つことは期待できないのです。

一方で、文書館が記録を入手し、すべての資料が入ったその館自身のハード・ドライブを持っていれば、複製を保存する最良の方法は、そのハード・ドライブのイメージを作ることです。

コンピュータとシステムの完全性

ビット毎の完全性と複製の完全性に加えて、コンピュータの完全性について考えることが重要です。コンピュータの完全性とは、コンピュータのプロセスが正しく使用され、操作された時、正確な結果を生み出すこと、また、記録が生成された時にも同様に作動したことを意味します。

システムの完全性も必要です。つまり、ビットから記録、コンピュータ、システムへと、すべてのレベルの完全性が必要なのです。システムに完全性があるとは、意図的であれ、偶発的であれ、権限のない操作をされることなく、システムが意図した機能を正常に遂行し、記録が生成され、使用された時点においても同様であったということです。

コンピュータとシステムの完全性は、ユーザ許諾、パスワード、ファイアウォールなど通常の物によって保護されますが、最も重要なのは、前にも述べた通り、ログです。ログはシステムによって自動的に生成されます。

ここではすべてのログ（スライドに列記しています）について説明しませんが、ここで指摘したいのは、すべてのログの内、大部分のログは立証後に破壊されても構わないのですが、維持しておくべき重要なログは、監査ログ（auditing log）です。監査ログは飛行機のブラック・ボックスのようなもので、誰が、何を、いつ、何処で行なったかを教えてくれるものです。

プロセスの完全性

プロセスの完全性は、アーキビストにとって最も重要な完全性です。デジタル環境において、私たちは、記録を受け取った瞬間からその後ずっと永遠に、私たちがその記録に対して何を行なったかを一つ一つ順を追って立証できなければなりません。また、私たちが立証しなければいけないことは、その記録に干渉しなかったかどうかです。すなわち、私たちが記録を収集、取り込み、使用、管理、保存する為に使用した方法によって記録が変わらなかったかどうか、変えられたとしたらその変更を文書化したかどうかを立証しなければなりません。

認証

ここで、認証について一言整理します。多くの国における法律の条文の多くは、特にヨーロッパでは、真正性と認証とを混同しています。立法者は、認証の手順を規定すれば、記録の真正性を保証したと考えていますが、それは違います。

認証というのは単に、ある記録が真正であることを宣言する手段の一つにすぎません。しかし、それはある特定の時点で、つまり宣言した時点で記録が真正であると宣言できるだけです。認証は、記録の真正性を維持するわけではありません。

例えば、デジタル署名は、署名の機能というよりは、封緘の機能を持っています。というのは、署名は記録の内容に責任を賦課し、記録の必要な構成要素であるのに対し、デジタル署名は完全な記録に添付されるものだからです。

デジタル署名が持つ問題は、記録と共に保存することができない点です。したがって、記録の伝達には役立ちますが、その記録を受け取った人は、デジタル署名を記録と共に保存することができません。なぜなら、デジタル署名は記録より先に古くなってしまい、記録とともにマイグレーションを行うことができないからです。

好ましい認証手段

デジタル記録を認証する手段は他にもあり、それらは手続き的方法であることが多いです。正当な保管の連鎖は、真正性を示し、記録を認証するための最善の方法です。なぜなら、その記録が生成された瞬間から、責任のある、信頼された保管者の下で保管されたということが証明できるからです。

デジタルでの保管の連鎖とは、記録とその変更に関する情報を記録することであり、特定のデータが特定の時間・日付において特定の状態にあったということを示すものです。したがって、これも適正な認証方法です。

記録の保持および保存システムの信用性に関して、専門家によって行われた宣言は、いかなるデジタル署名よりも重要です。

保存

ここまでお話してきたことは、保存の概念の意味にとって、またアーキビストの機能にとって非常に重要な意義を持っています。

デジタル環境における保存の概念は、記録の作成時点からずっと永遠にわたって、記録の伝達に必要な、

変換や移行を含むすべてのプロセスを含まなければなりません。保存というのは、我々が持っているものを維持するだけではなく、記録を保存できる方法で確実に記録を作成することです。

保存の連鎖が途切れないでいるためには、記録の作成から始め、記録作成システムから記録保持システム、そして記録の保存システムまでを継続しなければなりません。私たちが保存プロセスを記述する場合、記録が作成される予定のシステムが設計された時点から始めなければなりません。

信頼される保管者としてのアーキビスト

重要なことは、アーキビストは記録についての信頼される保管者として、自らを提示しなければならないということです。

信頼される保管者とは、中立的な第三者として行動し、記録の内容について利害関係が無く、保管下にある記録を変更する理由を持たず、偶発的・意図的にかかわらず、いかなる者にも記録の変更を許可しないことを証明する人物です。

信頼される保管者は、責任を全うするために必要な知識と技術を持ち、信頼の基礎としての能力を持ち、正規の教育によってそれらを取得しなければなりません。

信頼される保管者は、作成者の記録の正確で真正な複製が確実に取得・保存されることを保証できる、信頼性のある保存システムを確立しなければなりません。

警察署の場合を考えてみましょう。従来、警察は、証拠を証拠室に保存していました。しかし、裁判が5年、10年と長引く可能性があり、その間、デジタル記録は古くなってしまいます。

つまり、それらの記録をアクセス可能にしておき、裁判で使えるようにするために、新しいシステムおよび新しいフォーマットに移行させる必要があります。しかし、こうした操作により記録のアクセス可能性を維持しようとすると、それらの記録を変えてしまう可能性があります。したがって、双方の利害関係者の内の1つである警察に記録を移行させることはできません。たとえ彼らが記録を変えなかったとしても、私たちはそれを信じることはできないし、それを証明する方法も無いからです。最も良い方法は、裁判所のアーキビストにデジタル記録を委託することです。なぜなら、アーキビストは中立的な第三者だからです。彼らは、記録の改ざんを疑われることなしに、記録の移行を行うことができます。

アーキビストの新しい役割

アーキビストは新しい役割を持つようになりました。私たちは今や、以前とは全く異なる職責を担う者なのです。

アーキビストは今や、信頼される保管者として「指定された」役割を負う者として、記録のライフサイクルの開始点に居合わせていなければなりません。また、レコード・マネジャーと協力することが必要ですし、記録作成者に対し、信頼できる記録保管者として指定された者だと認められなければなりません。

アーキビストは、長期にわたって記録の真正性を評価し、記録が存在する期間中、その真正性を監視しなければなりません。なぜなら、作成者によって記録が1つのシステムから別のシステムへ移動される度に、その真正性、完全性、あるいはメタデータが失われるかもしれないからです。

アーキビストは、記録が作成された瞬間に、保存されるべき記録を同定しなければなりません。たとえこれが保守的な評価選別であったとしても、何が保護されるべきかを即座に同定する必要があります。そ

うすることにより、アーキビストはこれらの記録が、時を経て、システムの変化により、どのように変換されるかを監視するのです。

アーキビストは、アーカイブズの技術的能力に基づき、記録の保存の可能性を判断しなければなりません。これは非常に重要なことです。なぜなら、今までであれば、私たちはアーカイブズに移管されてきた物は何でも受け入れてきました。もし私たちが、保存不可能なデジタル記録を受け入れてしまったら、それはその記録を破壊することと同じことになります。私たちは、記録を受け入れる前に、自らの資料保存能力を評価しなければなりません。

アーキビストは、取得した記録を移行させるためのコントロールされたプロセスに基づいて、保存戦略を決定しなければなりません。また、忘れてはいけないことは、私たちは取得したすべての記録を移行させなければならないけれども、原初のフォーマットも取っておかなければならないということです。なぜなら、移行後にすべての記録をチェックすることは不可能であり、もし記録の改変を研究者が発見したとしても、原初のフォーマットを保存していなければ、取り返しのつかないことになるからです。

アーキビストは、アーカイブズの技術的環境が更新された場合に、記録に起こるすべての変更を文書化しなければなりません。

アーキビストは、アーカイブズの中で誰がその記録を閲覧する権利を持つか、記録に影響を及ぼしかねない行為を誰が実行するか、誰が記録を使用することができるか、誰が記録を複製することができるかを指定しなければなりません。これは、アーキビストが自らの機能を果たすにあたり、自身が記録の完全性を改変してしまうことを防ぐためです。

アーキビストは、記録の損失または破損を防ぎ、発見し、修正するための手続きを確立しなければなりません。それとともに、メディアの劣化および技術的変化の際にも、記録の完全性が継続されることを保証しなければなりません。

アーキビストは、認証責任を規定するルールにしたがって個々の記録を認証しなければなりません。すなわち、変換もしくは移行が行われる度に記録の正確性を制御し、知的財産権およびプライバシーの問題を扱うための手続きを開発する必要があります。

アーキビストは、個々のアーカイブズ資料群をひとまとめに認証する手段として、また、その資料群の中の記録同士の関係性を認証する手段として、アーカイブズ記述を使用しなければなりません。なぜなら、従来型の記録において、記録同士の関係は、記録が保存されている場所によって明示されていたけれども、デジタル記録においてはそうではないからです。アーカイブズ記述のみがその関係性を明らかにすることができます。したがって、従来型の記録目録は、紙環境においてよりもデジタル環境においてずっと重要になっており、記録の文書的コンテクストを認証する機能として働きます。

最後に、アーキビストは常に、企業と同様に、研究・開発プロジェクトに携わっていなければなりません。なぜなら、私たちは、学者いわんや立法者や官僚が教えてくれるのを待ってはいられないからです。状況は非常に速く変化しますから、アーキビストは常に研究の最先端にいなければならず、世界中で実行されているすべての研究プロジェクトによる成果を試してみなければなりません。そうしなければ、追いつくことは決してできないからです。

インターパレス第3期の成果

まずは、インターパレス第3期の成果・リストをご覧下さい。私はそれらについて説明しませんけれども、インターパレスのウェブサイトの「Products」という項目の中の、「InterPARES 3」で詳しく見ることができます。ケーススタディ、総合研究、各チームのレポート、キーワード、またはテーマによって、各成果にアクセスすることができます。また、ご希望であれば、すべてのインターパレス3の成果は、自由に使うことができます。

ご清聴ありがとうございました。

[原註]

（1）ここに列記したすべての主な成果は、インターパレスポータル www.interpares.org で簡単にアクセスできる。

（2）記録の伝達状態は、その完成度の程度を意味する。つまり、草案は不完全で、修正を意図したものである。原本はその記録が意図した目的に到達できている最初の完全な記録である。複製は草案の複製、原本の複製、あるいは複製の複製である。紙の世界では、各記録はそれがファイルに綴じられた時の伝達状態のままで存在し続けた。デジタルの世界では、原本は受領した後、保存されるまでの間しか存在しない。それ以後は複製しか存在しない。草案は、作業している間だけ草案として存在する。保存された後は、その複製を再生できるのみである。

[訳註]

（1）ある機能やデータを別の形態で提供するためのプログラム。

（2）コンピュータや関連機器を対象とし、デジタルデータから証拠となるものを探し出す、鑑識ないし科学捜査の活動。

（3）デジタルデータの送受信の際に、そのデータの完全性を検証する手法のひとつ。送信するデータをブロックに区切り、それぞれ符号化して加算した値を算出し、それをデータと一緒に送信する。受信した側でも同様の方法で値を算出し、双方の値が同一であればそのデータは完全である、と見なされる。

（4）デジタルデータから、固定長の乱数（ハッシュ値）を生成する演算アルゴリズム。同一データからは常に同一のハッシュ値が出力されるしくみであり、ハッシュ値がわずかにでも異なるとそのデータは改変されたと見なされる。「チェックサム」と同様、データの送受信の際などにその完全性を検証する手法として用いられる。

[編集委員会註記]

本講演は、学習院大学大学院アーカイブズ学専攻および京都大学大学文書館の共催による国際セミナー「デジタル記録とアーカイブズ」（2012年6月23日、於・芝蘭会館別館）において行われたものである。同セミナーの概要については、『京都大学大学文書館だより』第23号（2012年10月）を参照されたい。講演は英語で行われ、天理大学の古賀崇准教授が通訳を担当した。ここに掲載したのは、講演者のルチアナ・デュランチ教授による加筆修正を経た講演録の日本語訳である。あわせて、講演録の原文（英文）とセミナー当日の講演スライドを次ページ以降に収録した。なお、本文中の［　］内は翻訳者の古賀氏または編集委員会による註記である。