

# 電子図書館と情報セキュリティ

木下宏揚

KINOSHITA Hirotsugu

## 1 まえがき

大学や研究機関などでは文字情報や音声、静止画像、動画像など様々な情報を研究者同士で共有したり、一般向に公開する必要がある。インターネットの普及により、これらの情報はデジタル化されたデータをネットワーク上で取り扱うことが一般的になってきている。この際、主に WWW (World Wide Web) が用いられることが多い。本稿では、WWW で情報を発信する際の技術的な問題点と著作権上の問題点について考察する。また、WWW をベースとした電子図書館の現状と課題について考察する。

## 2 情報公開とセキュリティ

### 2.1 WWW と情報発信

WWW (World Wide Web) は CERN (欧州合同素粒子原子核研究機構) の研究者が情報を共有するために開発された。当初は文字情報のみであったが、徐々に仕様が拡張され様々なメディアを扱えるようになった。WWW が研究者の道具から一般の人々の情報伝達手段として用いられるようになり、セキュリティに関する様々な問題が発生している。ひとつは、情報の保護や通信相手の確認など技術的なセキュリティの問題、もう一つは、発信する情報の著作権など法的あるいは倫理的な問題すなわちコンテンツの問題である。以下これらのセキュリティの問題について解説する。

### 2.2 サーバのセキュリティ

WWW でサービスを受ける場合、クライアントであるブラウザが WWW サーバにサービスを依頼する形をとる。不特定多数に WWW コンテンツを公開するということは、WWW サーバが不特定多数からの攻撃にさらされることを意味する。クラッカー (不正侵入者、悪意を持った攻撃者) はオペレーティングシステムやサーバのソフトウェアのセキュリティ上の欠陥 (セキュリティホール) を利用して侵入し、コンテンツの書き替え、消去、非公開情報の入手などを行ったり、他のサーバへの攻撃の足掛かり (踏み台攻撃) にしたりする。とくにクライアントからサーバへデータを転送するときなどに利用される CGI など動的に生成されるコンテンツはセキュリティホールになりやすいため十分注意が必要である。

防止策としては、CERT (<http://www.cert.org>) などのセキュリティ情報に気を配り、ソフトウ

ウェアを安全な状態に保つことが重要である。また、セキュリティ対策に必要なコストと人員を確保できない場合には、信用のできる組織にサーバの管理を委託する（ホスティングサービス）ことも有効である。

### 2.3 情報の保護と相手の確認

現在のインターネットは、特別に対策をとらない限りは暗号化されていない生の情報が流れている。しかも、インターネットの盗聴は電話など他の通信手段と比較して容易であるため、個人情報やクレジットカード番号など重要な情報を取り扱う場合、暗号化などの対策が必要となる。インターネットによる通信の場合、通信相手が本物かどうか確認することも重要である。クライアントとサーバがお互い相手を確認することを認証という。

現在用いられている認証には、証明書、パスワード、クッキー、クライアントのアドレス（またはドメイン名）がある。証明書はサーバの提示するサーバ証明書とクライアントが提示するクライアント証明書がある。ただし、現在のところほとんどの場合、サーバ証明書のみが利用されている。証明書は公開鍵暗号を用いたデジタル署名に基づいている。証明書は相手の公開鍵と認証局のデジタル署名などから構成されている。公開鍵は証明書のデジタル署名を確認するとともにセッション鍵（通信用の慣用鍵暗号の鍵）の暗号化に用いられる。次に、提示された公開鍵と相手の結び付きを保証するために認証局のデジタル署名が確認される。認証局は、サーバ（またはクライアント）の公開鍵の正当性を保証するために証明書にデジタル署名を行う。主要な認証局（日本ではセコムなど）の公開鍵は予めブラウザに組み込まれており認証局の署名を確認出来るようになっている。

図1に示すような、サーバ証明書の提示と確認と通信用の暗号化鍵の交換が終了すると、サーバが本物であることと通信路の安全性が保証される。また、この場合一般的なブラウザでは、ウィンドウの右下の南京錠のマークが施錠された状態になる。このとき、URLはhttpからhttpsに変化する。

パスワードは、サーバがクライアントを認証するときに用いられる。サーバが証明書を提示して、暗号化が行われていない状態ではパスワードが生の状態でネットワークを流れるため注意が必要となる。

WWWはクリックするごとに接続開始、データ転送、接続終了を行っているため、1回目のクリックと2回目のクリックが同じクライアントのものであることを確認できない。クッキーは、クライアントの身元を保証するものではないが、サーバがクライアントの同一性を確認するものである。例えば、オンラインショッピングで、商品の選択、支払い方法の指定、配送先の指定など一連のセッションの間、あるいは後日再びセッションを開始したと

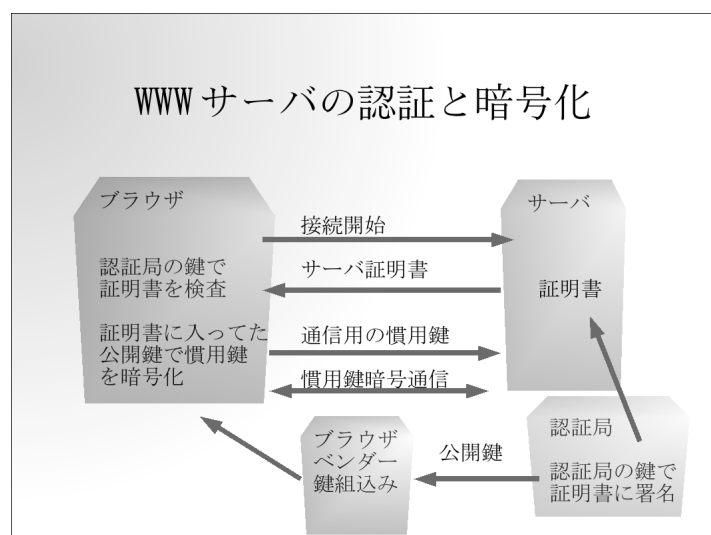


図1 WWWの暗号化と認証

きに同じクライアントであることを識別する。クッキーは最初の接続時にサーバからクライアントに送付され、次の接続時にはクライアントからサーバへ提出を行う。サーバにセキュリティホールがある場合クッキーが第三者に掠め取られる危険性があるため注意が必要である。

WWW サーバは接続を許可するクライアントを IP アドレスやドメイン名 (hertz.ee.kanagawa-u.ac.jp など) で指定することができる。

### 3 WWW のコンテンツ問題

#### 3.1 情報公開と著作権

著作権など WWW は世界規模で不特定多数の者に情報を提供可能なメディアであり、個人的なコンテンツであったとしても著作権は出版物や放送などと同等に扱われる。コンテンツで問題となるのは、他人が著作権を保有している情報を意図的あるいは無意識にコンテンツに含めてしまうことである。著作権法では、これは複製に該当するが、個人的あるいは家庭など小さな規模のグループ内での使用に限り複製は認められる。したがって、パスワード認証や IP アドレスでの認証を適切に行えば、この範囲内になる場合もあるが、個別の事例により微妙な判断が要求されるので、十分注意が必要である。

また自分で撮影したものであっても被写体によっては肖像権（主に一般人）やパブリシティ権（有名人）の侵害になる可能性がある。自分で撮影した建築物の写真は一般には著作権の対象にはならないが、テーマパーク内の建物など建物自体に高度の芸術性、創造性、商品価値がある場合には注意が必要である。国により異なるが、日本では著作権は権利者の死後 50 年で消滅し、これ以降許可なく情報を利用できるが、対象物の著作権が消滅していても書籍など出版物から転載する場合には著作権隣接権（演奏者やメディアの制作者の権利）が存在している場合がある。また、個々のデータが著作権をクリアしていたとしても、これらを体系立てて収集し、データベース化した場合、データベースに著作権が発生する。

#### 3.2 リンク

WWW の特徴としてリンクが挙げられる。リンクにより世界中の情報が有機的に結合され、紙の情報では得られない効率で情報収集することが可能となっている。リンクの問題としては、他サイトへのリンクを張る場合の許諾、リンク先のコンテンツの責任、リンクの方法に分類できる。他サイトへリンクは、著作権法上は引用に該当すると考えられるので、法的には相手に許可を求める必要はない。しかし、倫理的問題は別であり、次のように考えればトラブルとなる可能性は低いと考えられる。リンク先のページにリンクに対する指示がある場合は原則としてそれに従う。大企業や大学全体のサイトなど規模が大きく公共性が高いと考えられるサイトに対しては管理者の手間も考慮すると許可は求める必要はない。個人や研究室単位など比較的小規模のサイトの場合は許可を求めた方が無難である。

リンク先のコンテンツには一般的には責任は生じないと考えられるが、リンクの説明内容によっては問題が発生する場合がある。同一組織内へのリンクの場合、リンクの有無に関わらず、その組織に

管理責任が生じる。ただし、基本的にはコンテンツの責任はその作者が負うべきであり、所属組織は作者に対して利用規定やガイドラインの意味を理解させる努力を行う義務が生じる。

リンクの方法は、情報がサーバ上にあるのではなく、リンクであることが読者に明確にわかる形で行う必要がある。読者が錯誤するような形のリンクでは著作権法上の引用に該当しない可能性が発生する。

### 3.3 電子透かし

現在の WWW のシステム（http の範囲内）では、情報転送後の複製を制御することは困難である。そこで画像や音声などの著作権者が誰であるか明確に示すために、電子透かしが用いられる。

図2のように、電子透かしは人間の視覚、聴覚では知覚できないような形で、著作権を示すような視覚的パターンもしくは文字データなどを保護したい対象の情報に埋め込むのである。電子透かしは透かしの情報を故意に除去しようとする、保護対象の品質が著しく劣化するような形式で埋め込まれており、一定の処理を行うと透かし情報を取り出して、著作権情報を確認することが可能である。

WWWに関連して発生しうるセキュリティの問題点について解説した。デジタル化された情報の処理の容易さ、インターネットの広域性、即時性により、問題が発生した際の被害は大きなものになる可能性がある。また、インターネット上のセキュリティ問題の解決には、技術的な対策だけではなく、法的あるいは倫理的な要因が含まれているため、多方面の専門家の協力が必要不可欠である。

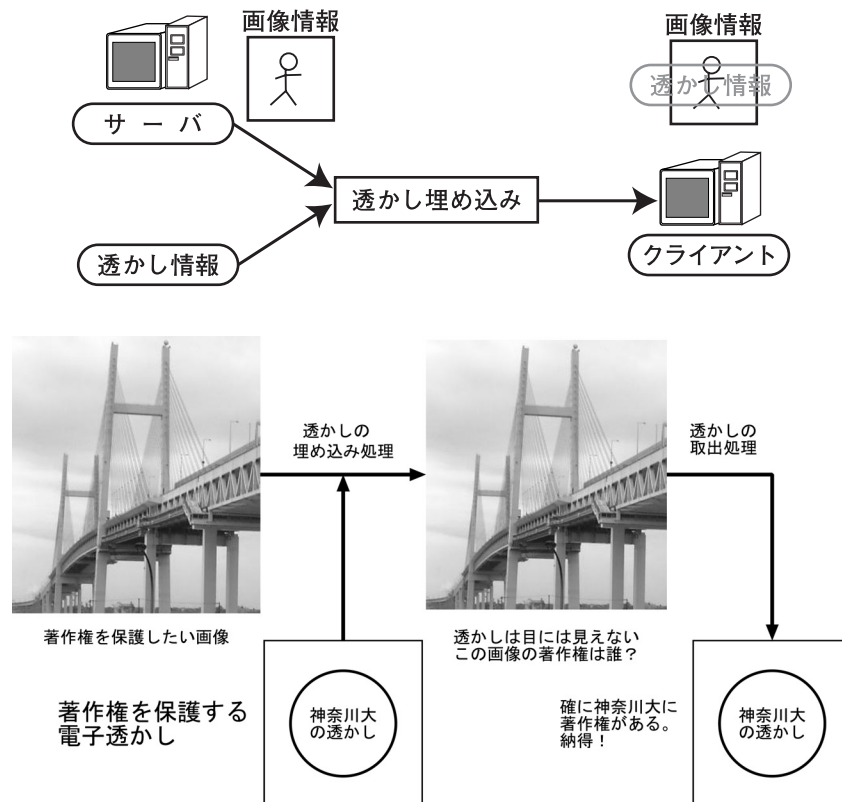


図2 電子透かし

## 4 電子図書館の動向

### 4.1 電子図書館

近年、コンピュータおよびインターネットの普及により電子図書館が注目されている。新しい図書館の流れとしては、インターネットによる情報検索、ポータルサイトとしての役割、オンラインデータベース、電子ジャーナル、CD-ROM の利用、DVD、VIDEO、CD、Microfilm などの視聴覚資料の利用などが挙げられる。

一方、従来型の図書を中心とした図書館は今後も存続していくと考えられるが、インターネット時代の図書館はどうあるべきか、紙媒体としての図書はどうなるかについて触れ、図書館の未来について考える。ここでは、インターネットによる情報検索や電子媒体資料と図書館の関係について考察する。

### 4.2 電子図書館の目的

図書館の基本的機能は、書籍の収集および分類所蔵、利用者に対する蔵書の提供、利用者が求める情報探索の支援、文献複写などがある。

一口に電子図書館といっても、その目的や電子化の度合いは様々である。一般に電子図書館と言われているものには、従来型の図書館と電子図書館の両方の機能をもった図書館とインターネットでのみサービスを行う図書館がある。

いずれにしても電子図書館のコンセプトは

- (1) 知識情報コンテンツはデジタル書式で記録され、体系的に蓄積保管される。
- (2) 知識情報コンテンツに対して、いつでも、だれでも、どこからでも、アクセス可能となる。  
そのためにインターネットが使われる事が必須条件となる。
- (3) 知識情報コンテンツへのアクセス手段の使い勝手が良い。ユーザから見て、「デジタルデバイス」を引き起こす様なことはない透明なシステムである。
- (4) 個人個人の役割や要求に応じて、必要とされる知識情報コンテンツにアクセス可能な制御が自然に行われる。

などが考えられる。

具体的なサービスおよびコンテンツの種類としては、

- (1) 著作権の保護期間が終了した著作物をデジタル化したもの
- (2) 図書館あるいは図書館が所属する組織に著作権があるもの
- (3) 有料の電子化雑誌記事・論文、電子ジャーナル
- (4) DVD、CD などパッケージ系電子出版物
- (5) オンライン商業データベース
- (6) 他機関の WWW サーバなどネットワーク情報資源へのリンクを集めたポータルサイト



### 4.3 利用と管理

図書館の利用者と管理者の枠組みについては従来の図書館でも重要な項目であるが、電子図書館の場合、より広い利用範囲が想定されるためいっそうの注意が必要となる。

利用者については、無条件に公開する場合と登録されたユーザにのみ公開する場合がある。また公開に際して、情報を無償で提供する場合と有償で提供する場合がある。登録されたユーザを識別する方法としては、パスワード認証、ホスト認証、X.509 証明書を用いる方法がある。

管理者については、各管理者の資源へのアクセスの範囲の決定、ファイル名やディレクトリの構成などデータ構造の決定権などを決定する必要がある。

### 4.4 情報の表現と検索

電子図書館で取り扱う情報には、対象となる情報本体の1次情報とタイトル、著者、索引、あらまし、書評など1次情報に付随し、検索などの手がかりとなる2次情報がある。

従来の紙媒体やアナログの音声や映像情報などは、そのままの形で2次情報のみデジタル化する場合と、1次情報を含めてデジタル化する場合がある。

1次情報の電子化方法としては、書物の場合、人手によるもの、スキャナにより画像データに変換したもの、これをさらにOCR (Optical Character Reader) で符号化したテキスト情報に変換する方法などが考えられる。

デジタル化した情報の蓄積媒体としては、ハードディスク、CD、DVDなどの光ディスク、磁気テープなどが考えられるが、取外しが可能なメディアは、読み出し装置を継続的に入手することが困難な場合が多く、1次的なバックアップや情報交換以外の用途には不向きである。現状ではハードディスクによるディスクアレイによる記録が容量、信頼性、スピード、継続性などにおいて優れていると考えられる。

非デジタル情報をデジタル化し、あるいはもともとデジタル化されている情報は、そのままの形では扱いにくいので、計算機で処理が容易な何らかの記述言語で表現する必要がある。現在このような目的で有力視されているのはXML (Extensible Markup Language) である。XMLではHTMLと同様にタグにより情報を構造化して記述するがHTMLではタグの意味が予め規定されているのに対して、XMLではタグの意味を定義可能で柔軟な記述が可能である。

1次情報あるいは2次情報はXMLで記述されるが、これらの情報をデータベースとして構成し、さらにネットワークで有機的に結合し、情報検索や

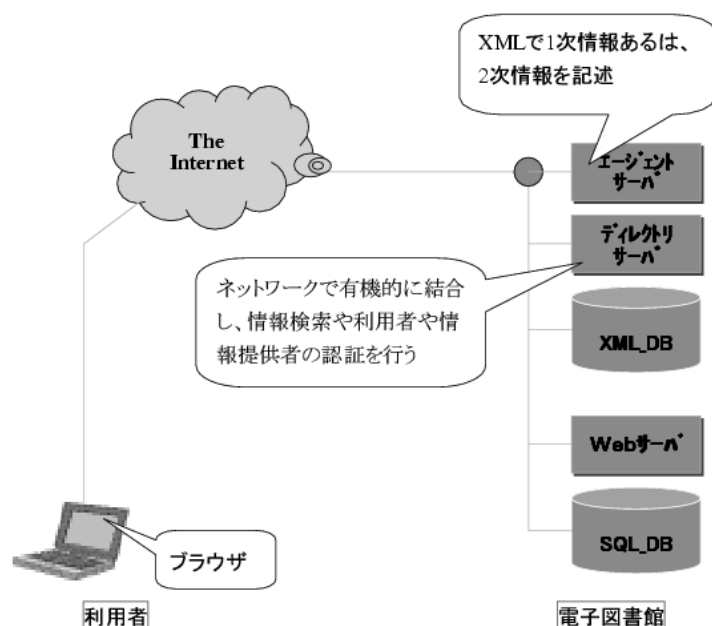


図3 XMLとディレクトリシステムの関係

利用者や情報提供者の認証を行うためには、ディレクトリシステムが用いられる。ディレクトリシステムとして現在もっとも普及しているものにLDAP (Light weight Directory Access Protocol) がある。

利用者と電子図書館など情報提供者の間のインターフェースとしては、WWW (World Wide Web) が用いられる。WWW の基本は情報提供者から利用者への情報の転送であるが、CGI (Common Gateway Interface) により双方向の通信が可能となる。XML とディレクトリシステムの関係を図3に示す。

#### 4.5 電子図書館が抱える課題

今後、電子図書館がさらに普及していくためには、以下のような課題を解決する必要がある。

(1) 電子図書館に蓄積する知識情報コンテンツの増大に対処する手法

知識情報コンテンツを効率よく分類整理して、データベースに収集するシステムが必要になる。また、膨大なデータの圧縮・分散等の技術も導入されることになる。

(2) インターネットを使用する事に関連する情報セキュリティ

知識情報コンテンツにアクセス可能なユーザが、その時必要な知識情報コンテンツに対して、安全にアクセス可能な通信システム

(3) 知識情報コンテンツを蓄積・管理するデータベースの設計

(2) を確保した上で、効率良く目的の知識情報コンテンツにアクセス可能な検索システム。特に電子図書館に於いては、従来の図書の電子化に留まらず、画像・音等、非文字のデータに対しても、効果的に蓄積・検索可能なシステムが望まれる。

(4) 電子図書館どうしのネットワーク化、知識情報コンテンツの共有化

従来の紙ベースの図書館でも蔵書リストの共有化による連携機能があるが、電子図書館ではネットワークをフルに活用し、複数の図書館による知識情報コンテンツ自体の共有化が、「リンク」という形で実現できる可能性がある。

## 5 むすび

WWW を利用した情報発信に関わるセキュリティの問題と電子図書館の動向について考察した。今後の課題としては、XML ベースで構築されたデータベースの効率的な索引付けと検索手法や著作権の管理システムの構築などが挙げられる。

### 参考文献

- (1) 山口英, “電子図書館への道”, UNIX MAGAZINE, Vol. 17, #5, pp. 49-56 (2002).
- (2) 澁川雅俊, “デジタル・ライブラリーのコンセプト”, [http://www.humi.keio.ac.jp/japan/docs\\_j/report/annual/1998/1\\_4\\_text.html](http://www.humi.keio.ac.jp/japan/docs_j/report/annual/1998/1_4_text.html)
- (3) 内藤求: “電子図書館の動向”, <http://www.ctc-g.co.jp/caua/viewpoint/voll/3-2.pdf>

(COE 共同研究員)