Digital Libraries

Access Management

W. Arms, Cornell
The Access Management Problem

The manager of a collection of information provides access subject to policies

- Library – donor restrictions, privacy, copyright
- Medical records – need to know
- Government – secrecy and classification
- Vendor – payment
Copyright

United States copyright law:

• Applies to literary works
  e.g., text, photographs, computer programs, musical scores, videos, audio tapes

• Initially, the creator of a work or the employer of the creator owns the copyright
  Exception: materials created by government employees

• Intellectual property -> can be bought and sold like any other property
Copyright gives the owner the exclusive right to:

- reproduce
- distribute
- perform
- display
- license to others

_Nominally for a fixed period, but the period has been steadily lengthened_

_Derivative work_: new work uses any part of another work:

- New parts are owned by new author
- Conditions that apply to old work apply to derived work
Copyright

Rights of users

• First sale
  e.g., can sell used books

• Fair use
  e.g., can quote short sections in scholarly articles or reviews

International differences -- moral rights

• In Canada: author has rights to
  attribution of authorship
  integrity

• Moral rights cannot be transferred
Fair use

Factors to consider

• the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes

• the nature of the copyrighted work

• the amount and substantiality of the portion used in relation to the copyrighted work as a whole

• the effect of the use upon the potential market for or value of the copyrighted work
Contracts, licenses and derivative works

Contracts allow intellectual property to be sold or licensed

- Almost any terms and conditions can be agreed
  - Permanent or temporary, whole or part
  - Exclusive or non-exclusive
  - Restrictive license or broad

- Enforceable by courts
Users

Authentication

Roles

Information Managers

Laws and agreements

Policies

Authorization

Attributes

Identification & authenticity

Digital objects

Access

Permitted Operations
Digital material

Authorization

Attributes

Identification & authenticity

Digital objects
Digital objects

Digital objects contain information that users may wish to access subject to policies. Properties of the digital objects that are important for access are encoded as attributes.
Attributes

Administrative metadata describes properties of the digital object, e.g.

- Registered for copyright on 1/1/1996.
- French government publication.
- Letter from donor, dated 1/1/1893, states "I donate my collected papers to the nation."
Different attributes may be associated with different elements of a digital object.

Gaudeamus igitur Juvenes dum sumus

Permitted Operations

Authorization

Digital object

Attributes

Attributes
Users and roles

Users

Authentication

Roles

Authorization
Users and roles

User

A user is a computer system, or a person using a computer system, that wishes to access digital objects. Characteristics of users are encoded as roles.
Examples of users and roles

Roles

Verifiable facts about a user, used in access management, e.g.,

- The user is a subscriber to all ACM publications.
- The user is a minister of religion.
- The user is a high school student.
- The user is physically located within the Library of Congress.
A user can change role by authentication or payment.
Examples: authentication and payment

Authentication and payment

User's roles can be modified by authentication and payment, e.g.,

- The user provided the login and password associated with William Y. Arms.
- The user has paid a fee of $10 to Visa.
- The user is verified to be located within a high school.
Permitted operations

Authorization

Permitted Operations

Access
Permitted operations

Formally defined actions that a user may take to access digital objects, e.g.,

• Replicate from one computer to another.
• Render an image on a screen.
• Extract 2 minutes from a video program.
• Create a derivative work.
• Perform in public for profit.
• Export to Australia.
Encryption and other security measures may by used to enforce the permitted operations.
Enforcement

Methods to ensure that the permitted operations are the only actions carried out on digital objects.

Enforcement may be:
- technical (e.g., encryption)
- legal (e.g., damages for violation)
- contractual (e.g., revocation of license)
- social (e.g., isolation from peers)
Subsequent use

Access management policies frequently restrict the subsequent use that a user may make of digital objects, e.g.,

- No redistribution without attribution.
- Display on screen, but not print.
- Use on a specified computer only.

Enforcement of subsequent use policies by technical methods is rarely possible without great inconvenience.
Policy

A rule that associates attributes of digital objects with user roles to permit operations, e.g.,

- Access to subscribers only.
- May be used for any non-commercial purposes.
- Prints may be made at $1 per print.
- For use only within the Cornell Library.

if (attribute) and (role) then (operation)
A Terms and Conditions digital object is a standard set of policies that are applied to many digital objects.

Example:

- T&C object, CUL1, represents the standard policies for digital materials licensed by Cornell University.

- Material received by Cornell has the attribute CU1.

- If the standard policies change, only CU1 is changed.
Techniques of Access Management

• Roles and permitted operations
• Policies
• Encryption
• Authentication
• Subsequent use
Users

Authorization

Authentication

Roles

Information Managers

Information Managers

Laws and agreements

Policies

Access

Permitted Operations

Attributes

Digital objects

Identification & authenticity
A publishing example

**Collection** consists of:
- current journals, back list, promotional materials

**Subscribers** have access:
- current and back list - general, no redistribution

**Other users** have access:
- current - list price, no redistribution
- back list - 50% of list price, no redistribution

**Promotional** materials - unlimited access
## Attributes of digital objects

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Current</th>
<th>Back list</th>
<th>Promotional</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>back</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Roles of users

<table>
<thead>
<tr>
<th>Roles</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>subscriber</td>
<td>User is a subscriber</td>
</tr>
<tr>
<td>other</td>
<td>Other user</td>
</tr>
<tr>
<td>list</td>
<td>Has paid list price</td>
</tr>
<tr>
<td>discount</td>
<td>Has paid 50% of list price</td>
</tr>
</tbody>
</table>
## Permitted operations

<table>
<thead>
<tr>
<th>Operations</th>
<th>General access</th>
</tr>
</thead>
<tbody>
<tr>
<td>general</td>
<td></td>
</tr>
<tr>
<td>dist</td>
<td>Redistribution</td>
</tr>
</tbody>
</table>
## Policies

Each row of the table represents a policy.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Role</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>current or back</td>
<td>subscriber</td>
<td>general, not dist</td>
</tr>
<tr>
<td>current</td>
<td>other and list</td>
<td>general, not dist</td>
</tr>
<tr>
<td>back</td>
<td>other and discount</td>
<td>general, not dist</td>
</tr>
<tr>
<td>promo</td>
<td>any</td>
<td>general, dist</td>
</tr>
</tbody>
</table>
The publisher changes its policies.

Current and back list will be treated the same, with a 20% discount on all journals.
**Example: Revised Role**

Define a new **role**: 

| standard | Has paid 80% of list price |
Revised policies

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Role</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>current or back</td>
<td>subscriber</td>
<td>general, not dist</td>
</tr>
<tr>
<td>current or back</td>
<td>other and standard</td>
<td>general, not dist</td>
</tr>
<tr>
<td>promo</td>
<td>any</td>
<td>general, dist</td>
</tr>
</tbody>
</table>
The basic decisions

Providing access is harder than blocking access

Intrusive technology drives people away
People value their privacy

It must be clear what the technology is trying to achieve

Technology serves economic or organizational goals

*Every technical question has an organizational context*
Technical strategies

Technology can support alternative market strategies:

**Strong enforcement:**

- Emphasis is on strict control by technical means.
- Subsequent use is barred by technology.

**Weak enforcement:**

- Emphasis is on customer satisfaction and market growth.
- Technology augmented by economic and social forces.
Trade-offs in enforcing access management

Convenience to users  ➔ Strength of enforcement

What is the cost of failure of authorization systems?
• Loss of revenue
• Harmful effects of security failure
• Loss of privacy
• Local compromise of security
• Global compromise of security

In digital libraries, the harm from security failures may be small
The loss from unhappy customers may be great
Encryption

Decryption
Dual key encryption

Each individual is given a key pair:

- **public** key -- known to the public
- **private** key -- kept private
You wish to send me an encrypted message

1. I tell you my **public** key (public information)

2. You encrypt the message using my **public** key and send it to me

3. I decrypt the message using my **private** key
Encryption in practice

Key management is difficult

- Single key encryption needs shared private keys.
- Dual key encryption needs public key infrastructure.

- One-time keys are good for secure transmission.

Government policies are misguided
Authentication of users

The issue: Cornell University has a site license to ACM journals. Is this user a member of Cornell University?

Approaches:

• IP address of user
• IP address of proxy
• login ID and password
  -> separate for each application or system
  -> campus authentication (e.g., Kerberos)
Authentication of users

Approaches to authentication

- What you know -- password
- What you have -- smart card, IP address
- Who you are -- finger print

Trade-off

*Simple, but insecure*

- Address of computer
- ID and password

*Expensive and intrusive*
Authenticity of digital objects

The issue:

• Content can easily be changed by error or maliciously.

• Authentication systems based on digital signatures fail if one bit changes.

• Authentication of content should be invariant over changes of font, format, encoding, and layout.

Examples:

• Copyright registration.

• International document delivery.
Hashing as test of identity

If $a = b$ then A is identical to B. Chance of error is tiny.
I wish to prove a message came from me

1. I calculate a hash of the message.

2. I encrypt the hash using my private key.

3. I send you:
   the message
   the encrypted hash

4. You decrypt the hash using my public key.

5. You calculate the hash on the received message.
If $b = b'$ then:
(a) Message is unaltered, $A = B$.
(b) Encryption used correct private key.
Subsequent use

Access management policies frequently restrict the subsequent use that a user may make of digital objects, e.g.,

• No redistribution without attribution.
• Display on screen, but not print.
• Use on a specified computer only.

Enforcement of subsequent use policies by technical methods is rarely possible without great inconvenience.
## Secure container (Cryptolope)

<table>
<thead>
<tr>
<th>Bill of Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear Text</td>
</tr>
<tr>
<td>Encrypted fingerprinting and watermarking instructions</td>
</tr>
<tr>
<td>Encrypted document part</td>
</tr>
<tr>
<td>Encrypted document part</td>
</tr>
<tr>
<td>Encrypted document part</td>
</tr>
<tr>
<td>Encrypted document part</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Terms and Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrity protection and signatures</td>
</tr>
</tbody>
</table>
Trusted systems

If all computers in a system can trust each other, powerful and flexible access management is possible.

General purpose personal computers are unlikely to be trusted.

Special purpose computers may be trusted, e.g., smart cards, printers.