# Scholastic Site Based Authentication

## Use Model

Robert Lacatena

Scholastic Site Based Authentication Use Model

written by Robert N. Lacatena
for Scholastic, Inc.
February 3, 2011

as part of the SBAM System implementation effort

This document was prepared as an introduction to the concepts, restrictions and techniques that underlie the business model in use at Scholastic with the intent of delivering online web services through the Internet to a selected customer base without individual (personal) user authentication.

# Scholastic Site Based Authentication Use Model

## Table of Contents

INTRODUCTION 1

THE BUSINESS MODEL 1
CUSTOMERS AND PRODUCTS 1
AUTHENTICATION 1

AUTHENTICATION 2

IP ADDRESSES 2
USER NAMES AND PASSWORDS 2
REFERRER URLS 3
PROXY NETWORKS 3
REMOTE SETUP URLS 3

SUBSCRIPTION TERMS 5

CUSTOMER PROFILES 6

CUSTOMERS AND SITES 6
DISTRICTS 6
CONSORTIA 6
LOCATIONS AND PREFERENCES 7
CONTACT INFORMATION 7

AGREEMENTS 8

AGREEMENT STRUCTURE 8
MIXED AGREEMENTS 8
UNRESOLVED AUTHENTICATION 8

USAGE STATISTICS 9

REPORTING 10

GENERAL 10
COMMISSIONS 10
RENEWALS 10
DEFERRALS 10

PUTTING IT ALL TOGETHER 12

THE SITE BASED AUTHENTICATION DATA MANAGEMENT SYSTEM 12
AUTHENTICATION EXPORT 12
EXISTING DATA CONVERSION 12

GLOSSARY 13

# Scholastic Site Based Authentication Use Model

## Introduction

### The Business Model

The Scholastic Site Based Authentication System is designed to support a “site based authentication” business model. In this context, a “site” defines a customer location, rather than individual persons at that location. Authentication refers to the task of determining which sites are or are not entitled to access to which products (Internet services).

In this business model, the products delivered by the business come from a menu of Internet based services which will be restricted to paying customers through an authentication method. That method will identify the customer, and so determine the products for which a visitor is eligible (i.e. has paid for or should otherwise receive) based on the site from which the products are accessed.

### Customers and Products

The customers in this business model are not individuals, but rather institutions with some sort of enrolled patronage (a student body for a school, or patrons for a library). The service is paid for by that institution or some umbrella institution. The individual customers receive the service (with a few notable exceptions) through computers belonging to the customer institution, and which are identifiable as such, i.e. “authenticated.”

### Authentication

Authentication is the act of determining whether or not site access is allowed for a visitor, and what products (database services) the authenticated customer can access. It is also necessary to track usage statistics for a customer, so that each product and page access can be tabulated for and reported to the customer.

A variety of authentication methods are used to identify visitors. The simplest is by IP address, but a visitor may also be identified by user names and passwords or a referrer URL, and these latter methods may also be further restricted by IP ranges (usually in the event of a proxy).

# Authentication

## IP Addresses

The most direct method of authenticating a user request would be through the user's IP address. In theory, the IP addresses used by a particular customer (an institution, or a particular building owned by an institution) will all fall within a known range, and no other customers will ever use those addresses. Under this model, a customer can be authenticated simply by associating a range of IP addresses with the customer, and then identifying visitors (and associating them with a customer and a customer's agreement) by their IP address.

This simple scenario fails under various circumstances. In some cases, the network in use by a customer (for instance, a district wide network for an entire school system) may randomly use IP addresses in a range for the network, so that individual schools or buildings within the network cannot be distinguished. Some customers may also use proxy services, which are shared with many other customers, and the IP address which is seen by the destination service will be any IP address from the range assigned to the proxy network, regardless of what visitor originated the request.

## User Names and Passwords

The easiest way to overcome IP address resolution issues is to assign user names and passwords. It should be noted that these usernames and passwords are assigned only to customers, not to individual users at the customer site.

The advantage to this method is that a user name can be assigned for use from any computer, regardless of its network or physical location. The primary disadvantage of this method is that the individual user may need to know the user name and password. A secondary disadvantage is that it is less secure, and would allow persons who are not actual enrolled patrons of the customer institution to use the services.

To overcome this second problem, whenever possible "cookie" user names/passwords are used. In this scenario, an administrator (or technical contact) is given a user name and password. At each computer workstation, the administrator can visit a cookie page, where they can enter the valid username and password, and the page can install a permanent authentication cookie. In this way, future visits to the product Internet sites from that workstation will be authenticated from that cookie, without the actual user ever entering or even knowing the user name and password.

In some rare cases a "permanent" user name and password can be used instead, although the term "permanent" is a misleading label. This is an actual user name and password which must be entered each time the site is accessed from a workstation. It is "permanent" in that it must be retained to be functional, while the "cookie" username and password previously described is not needed, once it is applied, for workstations to continue to have access.

## Referrer URLs

A superior alternative to user names and passwords, when applicable, is the use of a “referrer URL.” Whenever a user clicks a link to visit a web page, that request comes with a “referrer URL,” i.e. the URL of the page on which they clicked the link. If the customer has a secure page delivery service, one which can only be visited by users from within the institution, then this referrer URL can be set up as an authentication method. That is, if the product site can match the referrer URL for a web page to one assigned to the customer, then visits which are directed from that URL can be assumed to be for that customer. There would be no way (without spoofing an Internet request) for a person to link to a product page, from a customer page, as long as that page is guaranteed to reside within and only be accessible from within the customer’s network.

## Proxy Networks

A common cause for the use of usernames and passwords or referrer URLs instead of IP addresses is the use of a proxy network for the school or district. A proxy service, as it applies to the problem of authentication, acts as an Internet portal for the visitor, effectively masking the visitor’s actual IP address (any authentication system will only see an IP address for the proxy, not the visitor who originated the request).

When this is the case, while the range of IP addresses for the proxy network as a whole cannot be used to distinguish a particular customer, they can be used as an additional constraint on another authentication method. This provides some level of added insurance that, for example, a password is not abused and used from outside of the customer site (or, rather, from outside of the proxy service, and therefore outside of the customer site).

## Remote Setup URLS

Some institutions pay extra for the privilege of allowing their patrons to use the services off-site. This requires the assignment of individual user names for each patron, which would be an onerous task for the business to undertake.

Instead, the referrer URL mechanism is used as a method of allowing customers to select their own usernames and to receive passwords. A “remote setup URL” is established, distinct and separate from referrer URLs. The customer is given a page to link to from a secure page within their own network (i.e. a page that can only be accessed from a workstation inside of their institution).

When a customer visits that page, the referring URL will be validated to determine the customer. If the customer authenticates, then a remote setup username and password can be established within the authentication system for that individual, associated with that customer and the services allotted to that customer. It should be noted that such remotely set up user names and passwords are not visible to or needed by the business. The business only defines the method of authentication for getting to the remote setup page, but the business never sees the remote user names and passwords which are set up on that page.

## Subscription Terms

Site based authentication products are sold on subscription terms. That is, the customer pays for site access to selected products for a predetermined period, usually a full year defined by start and end dates. Customers may amend the product list, either by adding additional products in the same period (or overlapping periods), or by terminating one, some or all products early.

Subscriptions may be (and often are) renewed from one year to the next. A renewal will consist of an additional date range, with any required product changes. In the event of large agreements with multiple sites, the recipient sites may also change (adding or dropping schools).

## Customer Profiles

Customers for site based authentication products are usually large institutions, such as schools, school districts, libraries or library networks, or colleges and universities.

### Customers and Sites

There are typically three sorts of customer entities as defined by this business model.

One entity is the customer paying for the service. This may be an institution (like a school), a parent institution (like a district to which a school belongs), or a somewhat artificial "institution" (such as a consortium, i.e. a collection of schools or school districts).

The second entity is the customer receiving the service. This is termed a "site" or a "recipient." This is usually a school. It could be an entire district (in a case where it is not possible to, or no one has taken the effort to, resolve the individual schools) or other collection of actual, physical sites.

The third entity is the "official" parent of the recipient, if there is one. For a school, this may be a district. For a branch library, this may be the main library. The primary reason for this distinction comes in gathering and reporting usage statistics, which would be relevant to (and should be made available to) the customer paying for the service, the recipient, and the parent institution of the recipient, if one exists.

### Districts

In many cases, services may be purchased district wide. In these cases, the district is both the parent institution and the customer paying for the service.

A main library is the equivalent of a district. A library network could also, in theory, have a parent like a school/district, but often there is no such actual parent entity, or the entity may not be carried in the database as a distinct customer (if there is no one in any capacity to review usage statistics for the library network as a whole, or to need to view the stats for libraries other than their own within the agreement.

### Consortia

In some cases, the paying customer is a consortium. A consortium is a collection of schools, sometimes related in some way (for instance all in the same county, although in some cases, the schools have crossed state boundaries), and sometimes not, which are engaged together in receiving a service. A consortium can consist of hundreds or even a thousand schools. In the case of a consortium, the consortium pays for the service, and the member schools are all recipients..

A consortium must be set up as a customer in the database in order to be used as the paying customer for an agreement.

## Locations and Preferences

Many products include customer preferences. These preferences affect behaviors such as the style of presentation for a product, or whether or not sound or multi-media features are turned on, or perhaps specifies some customer specific value, such as the name of a logo image to be displayed on the web pages (for instance, a consortium's logo). Preferences are distinct from products, and are turned on or off (or are otherwise set) by recipient site.

It is possible, either for the gathering of usage statistics or for the application of preferences, that one recipient site will need to be separated into multiple locations. For example, a service might be implemented for a school with audio and multi-media features turned on, but those features need to be turned off in the school library. Or the site may be a library, where those features should be turned off except in their media room.

To manage this, it is sometimes necessary to subdivide a site into locations, so that each location can be associated with different authentication methods and so be given separate preferences. This division is another reason for the use of user names and passwords or referrer URLs (if, for instance, there is no way to distinguish the IP addresses for workstations in the different locations).

## Contact Information

Individual, personal contact information is sometimes associated with either a customer or an agreement. At the time that accounts are first set up a technical contact is usually recorded. This will be needed frequently to work through all of the authentication issues which may be encountered in the effort to get the product services properly delivered. Other contacts may also be recorded, such as the person responsible for the contract, a person responsible for testing access.

# Agreements

## Agreement Structure

For the business, contracts are organized into agreements for business purposes (i.e. an agreement to deliver certain product services for a specified period to specified customers).

For the purposes of actual implementation, however, with the need to deliver different product services to different sites or locations, using varying authentication methods, one “actual” business agreement may be broken into multiple “system” agreements, purely for the purposes of organization.

## Mixed Agreements

Sometimes a customer site will receive different services from multiple agreements (“business” or “system”). This could occur for “real world” reasons, such as because there are different actual business agreements. For example, a school could decide to purchase one product on its own, while the parent district purchases a set of other products for the entire district (or all schools of a certain grade range in the district), while yet other products are being purchased for the school by a consortium (for instance, by the state).

In these cases the same site (or site locations) will appear on multiple agreements, but should necessarily have the exact same authentication methods and preferences.

Because of the complexity of managing the data (the number and variety of authentication methods, and the problems that can arise in a more complex situation versus another, simpler case) this is sometimes not always properly realized.

## Unresolved Authentication

There are circumstances where authentication methods are established within an agreement (such as a range of IP addresses) with the intent of making the product services available to the customer (i.e. meeting the terms of the business agreement) without necessarily resolving the distinction between different recipient sites. This is perfectly acceptable if the customer is not interested in usage statistics by site, and there is no business reason to make a distinction (e.g. there is no site within those covered by the authentication methods which should not be receiving the services).

As a result, there may be cases within the data where authentication methods are not tied to customers or customer sites at all. They are instead tied only to an agreement, and so can be used to activate the services for those authentication methods (i.e. for anyone using those IP addresses, usernames/passwords, or referrer URLs), but they cannot be associated with any particular customer site.

## Usage Statistics

Customers are often interested in viewing usage statistics which show when and how which product services were accessed by visitors from which sites. These statistics are gathered by the authentication system, but are then compiled and reported through a separate usage statistics reporting system. That system is tasked with authenticating the visitor using a user name and password (one which is separate and distinct from any product services authentication password – it's an administrative ID) and associating the user with a particular customer and rights to view particular customer stats.

Generally, a user will be allowed to view any stats for their associated customer which is a site, for any sites for which their associated customer is the parent, and for any sites for which their associated customer is paying for the service.

While the task of collecting and reporting the statistics falls to other systems, it is still the role of the business to maintain the user names and passwords which will be used to log onto the reporting system, and to associate them with customers.

The reporting system also includes an "aggregation" feature. By declaring an aggregation code with a site, the business is able to let the customer group related sites. For example, aggregation codes in a large district or consortium might be set up to allow the user to group schools by grade range (high school, middle school, elementary school) or by geographic location (east , west , north , south).

# Reporting

## General

Within the business systems in use are a variety of reporting needs. Any system which supports site based authentication is primarily concerned with fulfilling that task, however other parts of the business will need reports from the system since in many cases it may be the only repository of the required information.

## Commissions

Commissions for site based authentication are complicated by several factors. First, the products offered often originate from several different business units, so the credited sales rep may vary by product. At the same time, these services are available to a wide variety of customers (public schools, private schools, school libraries, school districts, public libraries, other public and private institutions, and consortia). Again, often different business units are responsible for different types of customers, but an agreement may mix recipients of different types, without any distinction as to the dollar value of the service allotted to each receiving site.

As such, the business may need to accurately record a sales representative (or territory) assignment for agreements, customers, products and agreement terms. In some cases, it simply is not possible to properly distinguish sales rep assignments.

The commission structure behind this business model also usually involves different rates for renewals versus new service set ups.

Lastly, agreements may often be split or merged due to the complexity of managing the authentication data, which is the primary task at hand. In this case, it may not always be possible to assign an exact dollar value to an individual agreement.

Commission reporting will need to account for all of these issues when generating commission reports, and in any event, some amount of manual processing will be required to sort out ultimately unavoidable conflicts and ambiguities in the reports.

## Renewals

As a general area of business, management will look at reports on renewal rates across several different demographics (product, type of school, geographic location, sales rep, etc.).

## Deferrals

Because this business model offers a subscription based service which is not received in its entirety until the full term of the agreement has elapsed, the business cannot necessarily record the sale as income, even if the accounts receivable is collected before the term of service is completed.  Reports are therefore necessarily supplied to accounting, including the terms of service and dollar values, to help in determining accruals and deferrals.

# Putting it All Together

## The Site Based Authentication Data Management System

The business requires a computer system to support the authentication data management task of site based Internet product service delivery.  This system must support all of the functionality described so far, as well as several other features.

In addition, the system should include some functionality and organizational considerations which as much as possible minimizes duplication of data and other referenced difficulties in the data management.

## Authentication Export

A main function, in fact the primary task, of the system must be the generation and delivery of the authentication data on a daily basis in a fashion which supports the actual authentication system responsible for managing the delivery of product services to the customers and recording usage statistics.  The format for these files and their content is defined by that authentication system.

## Existing Data Conversion

The Grolier Online business currently has a site base authentication data management system (labeled the Global GO system).  If that business cuts over to use a newly developed, Scholastic wide system, that data must be converted.  Such a conversion will necessarily restrict what can effectively be implemented in any new system.

## Glossary

Agreement

An agreed to list of products, recipient sites, and term of service for a predetermined price.

Agreement Term

The list of products and the period, defined by a start and end date, over which this particular phase of an agreement is applicable.

Authentication

The act of determining whether or not site access is allowed for a visitor, and what products (database services) the authenticated customer can access.

Authentication Data Management System

See “Business System.”

Authentication System

The system which is supplied by the Business/Authentication Data Management System with the authentication methods and products for any particular data, and is then responsible for the actual task of authenticating visitors, providing access to product services, and gathering usage statistics.

Business System

The Authentication Data Management application responsible for tying together the complex web of information surrounding agreements, authentication methods, usage statistics acces, products and product terms.

Consortia

A collection of schools, sometimes related in some way (for instance all in the same county), and sometimes not, which are engaged together in receiving a service.

Contract

See “Agreement.”

Cookie

A value which can be stored permanently on a particular workstation.

Cookie Username/Password

A username and password combination that are used to install a permanent cookie on a workstation, which in turn provides permanent authentication (as long as the cookie is not deleted) for that workstation.

| | |
|---|---|
| Database | See “Product Service.” |
| Deferrals | Sales dollars which cannot be “put on the books” as a sale as soon as the accounts receivable is received, and must instead be deferred and accrued in installments over time. |
| IP Address | The unique Internet address that identifies a particular computer workstation in the Internet.  This value is a 32 bit number, often expressed visually as a series of “octets” such as 123.53.63.21. |
| Permanent Username/Password | An actual username and password that is used to authenticate a session. |
| Product | See “Product Service.” |
| Product Service | An individual Internet service (also termed a product, service or a database) which provides a specific user experience (user interface and available subject information). |
| Proxy | A proxy service, as it applies to the problem of authentication, acts as an Internet portal for the visitor, effectively masking the visitor’s actual IP address (any authentication system will only see an IP address for the proxy, not the visitor who originated the request). |
| Proxy IP | An IP address associated with a proxy service (which therefore cannot be directly associated with an end customer, but may be used to add further restrictions to non-IP based authentication methods). |
| Referrer URL | An Internet address (URL) supplied in a browser request as the page from which the customer linked to the product site, and which therefore can be used as a means of authentication. |
| Remote Setup URL | An Internet address (URL) used to authenticate a visit by a user, like a referrer URL, but with the intention of assigning them permanent username and password for future *remote* access (meaning access from outside of the customer institution). |
| Service | See “Product Service.” |

| | |
|---|---|
| Site | In the context of this document, a site is a particular customer location which has access to a specific set of products, using a specific set of preferences.  This location may be an entire institution, a building, or even a room or set of rooms within a building. |
| Statistics Reporting System | The system which is tasked with compiling and reporting usage statistics to customers. |
| Usage Statistics | Statistics which show when and how which product services were accessed by visitors from which sites. |
| Usage Statistics Reporting System | See "Statistics Reporting System." |
| User ID | A value supplied by a visitor for the purposes of unique authentication. |