



Consiglio Nazionale delle Ricerche

Automatic management of temporary web pages

P. Lazzareschi

IAT B4-22/2001

Technical Report

Dicembre 2001



Istituto per le Applicazioni Telematiche

Consiglio Nazionale delle Ricerche

Automatic management of temporary web pages

IAT-B4 2001 022

Pasquale Lazzareschi
pasquale.lazzareschi@iat.cnr.it

10 december 2001

Istituto per le Applicazioni Telematiche
Via Moruzzi 1, 56126 Pisa

Abstract

Many web sites host temporary information about short-term events and news (conferences, meetings, competition announcements, etc.).

Typically these announcements are created by manually inserting a short description into the home page with a link to a page describing the event. Of course the announcement should be inserted a certain amount of time before the event and removed afterward. Often the need for manual modification is the reason for many out-of-date web sites.

This work describes a web application to automate this process. Administrators can insert data about new events by using a web form, specifying: the title of the event, the publication date, the expiration date and the HTML page describing the event. The application saves all the information in a database and modifies the home page (inserting the event title among the links to the describing page) on the publishing date, and removing the link on the expiration date. The application interface also permits managing the information stored in the database: events can be viewed, modified or deleted. In order to maintain a history of events all information is kept in the database for one year after the expiration date. Security administrator access is provided using SSL and user identity can be checked using a digital certificate or password.

Administrative work is greatly simplified by use of this application. The home page is not modified by hand (thus reducing the possibility of errors). New announcements needn't be inserted on the publication date; the administrators are free to insert the announcement whenever he/she likes before the publication date and old announcements are automatically removed. Login to the web server machine is not necessary since all operations may be carried out from every machine on the internal network, and the use of SSL and digital certificates ensure the operation's security.

At present this application is used to manage the Italian National Research Council web site news.

Introduction

The aim of this work is to provide a tool for automatic management of event announcements on web sites. Many web site home pages contain information about short-term events or news that must be frequently updated in order to keep the site up-to-date. Typically these announcements are done by manually inserting a short description in the home page with a link to a page describing the event. Of course the announcement should be inserted a certain amount of time before the event and removed afterward. Often the need for manual modification is the reason for many out-of-date web sites.

This application provides a tool for the automatic insert, update and deletion of announcements on one web page (typically the home page). Administrators can insert data for a new event using a web form, specifying: the title of the event, the publication date, the expiration date and the HTML page describing the event. The application saves all the information in a database and puts the announcement online on the publication interval time. All the event's information is kept in the database for 1 year (a different period can be configured) after the expiration date, in order to maintain a history of events.

The original idea for this work came from the need for a tool to maintain the news announcements on the Italian National Research Council web site.

How the system works

In this section we describe how the application works. First the software used is described, then the security model, the web interfaces (how the web administrator view the application) and last, how the application works internally.

Software used

All the software used is "open source". They are free and available on various platforms. The use of "open source" software is justified in order to avoid licensing problems and to be platform-independent. The following kind of software are used:

- PHP (www.php.net) as a script language. PHP is an embedded script language, powerful and easy to use. All the web interfaces are written in PHP.
- Open SSL (www.openssl.org) for security management. All communications of the administration web interface are encrypted using SSL. The identity of administrators can be checked by SSL using digital certificates.
- MySQL (www.mysql.com) as database management system. All information about events is stored in a mysql database.
- Perl (www.perl.com). The application that publishes the events on the publication interval time is written in perl.
- Apache HTTP server (www.apache.org) as web server

The security model

The security model must ensure that definition or modification of events can only be carried out by authorized people. The following description concerns only security aspects related to this application. The security of the web server is a more complex matter and is not within the scope

of this work.

All administration web interface communications are encrypted using SSL (HTTPS protocol) to ensure that unauthorized people cannot read or modify data transmitted on the net.

A table in the database contains the names of the authorized users; the same table can also contain the password associated with each user. The password is used only for HTTP authentication.

This application can be configured to check the identity of an administrator in two ways:

1. HTTP authentication. After connecting, the application, responds with "401 Unauthorized", the browser displays a dialog box asking for user-name and password, the application checks the supplied user-name and password with the valid user-name/password contained in the database.
2. Digital certificate. Each administrator must obtain a digital certificate from a certification authority known to the web server. At connection time SSL checks the certificate validity and the application checks if the user is authorized (the certificate "common name" must be in the database table of authorized users).

The web interfaces

Each administrator accesses this application via a web browser connected to the application web site. The application URL depends on the installation and configuration chosen.

The Figure 1 shows the first page (customized for CNR web site) appearing after the connection. The page contains a link "define new news" to allow the definition of new events. The rest of the page contains a list of existing events ordered by publication date. Currently published events are show on a white background. The title of each event can be clicked on in order to modify its data.

Netscape: Gestione news CNR

File Edit View Go Communicator Help

NEWS management for CNR web site

[Define a new news](#)

News:

Title	Start publ.	End publ.	File or URL	Note	Modified on	Defined on
Nomina Direttore Generale del CNR	25/07/2001	23/10/2002	(ADD) 1 NominaDG.htm		25/07/2001 18:08:01	25/07/2001 18:08:01
Esito finanziamenti Agenzia 2000	06/07/2001	31/12/2002	(ADD) 1 NewsAgenzia2001-3.htm		18/07/2001 18:20:13	18/07/2001 18:20:13
Proroga finanziamenti Agenzia 2001	27/06/2001	31/12/2002	(ADD) 1 NewsAgenzia2001-2.htm		17/08/2001 19:31:45	17/08/2001 19:31:45
Nuove Commissioni selezione Direttori Istituti CNR	24/06/2001	15/12/2002	(ADD) 1 Dast1.htm		17/07/2001 19:24:53	17/07/2001 19:24:53
Selezione Direttori di n. 21 Istituti CNR	19/06/2001	31/12/2002	(ADD) 1 NewsConc21DirettoriCNR.htm 2 bandiera_eng.gif		17/08/2001 19:34:37	17/08/2001 19:34:37
Finanziamenti Agenzia 2001	04/06/2001	31/12/2002	(ADD) 1 NewsAgenzia2001.htm		17/08/2001 19:48:57	17/08/2001 19:48:57
Nomine al CNR	02/10/2000	22/08/2001	(ADD) 1 Nomine.htm		12/07/2001 16:21:10	12/07/2001 16:21:10

NOTE: Active news are showed with a white background

100%

Figure 1. Initial page with list of existing events

Figure 2 shows the page (customized for CNR web site) used to define a new event. The administrator specifies the title, the start publication date, the end publication date and a pointer to an event description. The event description can be specified with a file name of a previously written HTML page or a URL pointing to an existing page. The field "note" is never published, it can be used by the administrator as he/she likes.

Events, in the publication interval time, are automatically put online by an off-line procedure. The procedure is started automatically by a "cron" job at regular intervals (a new event in the publication interval time will be put online at the next run of the publication procedure).

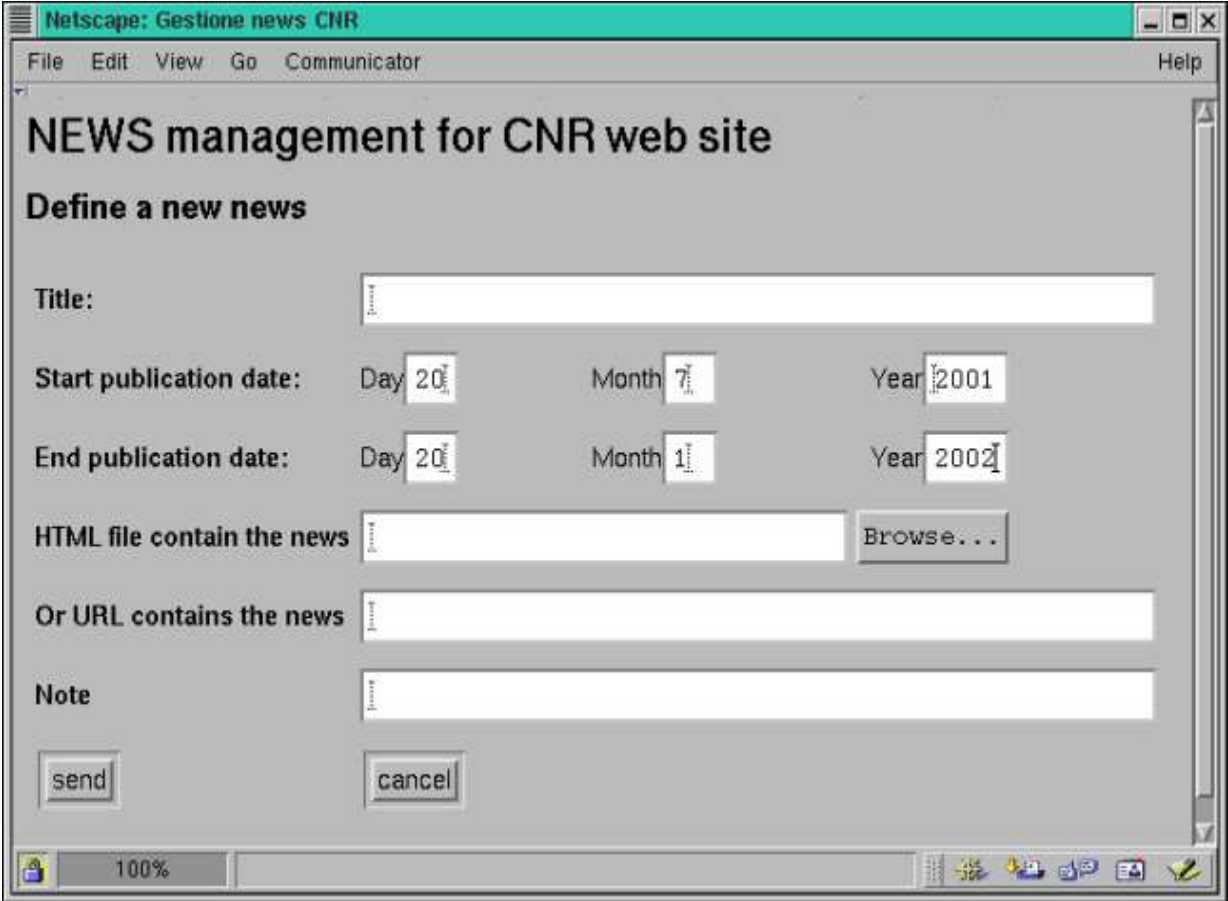


Figure 2. A form to define a new event

An internal overview

Figure 3 shows the logical relationships between application components. The application is composed of two parts: the web interface and the publication procedure.

The web interface consists of HTML pages with embedded PHP code. This interface is used by the administrators to define, update and view events. The work carried out by PHP code is:

- At each connection checks the authorization of the user trying to connect.
- Dynamically generates parts of HTML page from database data (the list of existing events, a form to modify a existing event, etc)
- Checks the validity of the data specified in the input forms
- Inserts data in the database (for new events or modification). All the data are saved in the database: title, start publication data, expiration date, HTML pages describing the event or URL and note. A unique ID is automatically assigned to each new event using an auto_increment database field.

The publication procedure consists of a perl program started at regular intervals by "cron" (a unix daemon used to execute commands at specified dates and times). This program extracts data from the database and generates its files in the directory NEWS of the web site home page.

The operations are as follows:

- Extracts from the database the data for all the events for which the current date is within the publication interval time.
- Generates a file named NEWS/news.html containing an HTML unordered list of event titles. Each list item is a link to the page describing the event (if the event description was a URL the link points to that URL; otherwise the link points to an HTML file as described later).
- If the event description is one or more HTML pages: for each event generates a directory, in the NEWS directory, named from the event ID and puts the HTML pages describing the event in this directory.
- Removes all the data of expired events (directories and files) from the news directory.

In order to use this application to automatically manage short-term events or news, the web administrator must enable the "server side include" feature in the web server configuration. See apache documentation for the module mod_include to enable server side include. The administrator must also insert an include statement like "`<!--#include virtual NEWS/news.html -->`" in the document where he/she wants the automatic generated list of events to be placed.

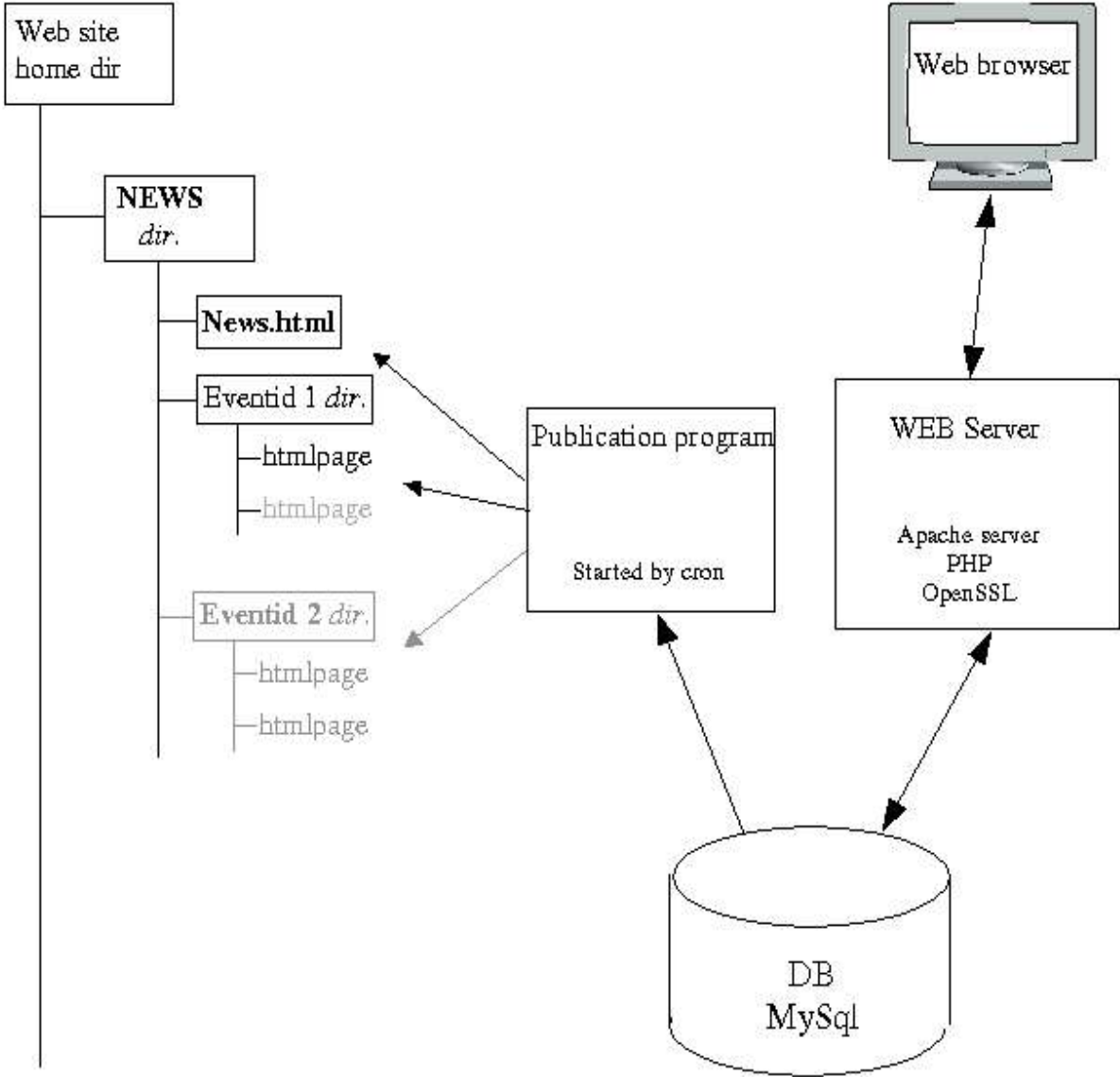


Figure 3. An internal overview

Conclusion

This application can be used to simplify the announcement management of short-term events or news on organization web sites.

- The web interface provides an easy to use tool to insert and/or update announcements.
- No particular skill is required to use the web interface.
- The possibility of error is reduced.

Data in the database permits to maintain a history of events (future, present and past) that can be viewed via the web interface.

Starting from July 2001 this system is used for the news management of the Italian National Research Council web site.

References

[1]W3C web site. www.w3c.org

[2]RFC 2616: Hypertext Transfer Protocol -- HTTP/1.1

[3]HTML 4.01 Specification W3C Recommendation 24 December 1999

[4]PHP web site. www.php.net

[5]Stig Sæther Bakken, Alexander Aulbach, Egon Schmid, Jim Winstead, Lars Torben Wilson, Rasmus Lerdorf, Zeev Suraski, Andrei Zmievski, Jouni Ahto, "PHP Manual" Edited by Stig Sæther Bakken, Egon Schmid. 19-08-2001. www.php.net/docs.php

[6]Larry Wall, Tom Christiansen & Jon Orwant. Programming Perl, 3rd Edition. O'Reilly & Associates, 2000

[7]Perl web site. www.perl.com

[8]Alligator Descartes & Tim Bunce. Programming the Perl DBI Database programming with Perl. O'Reilly & Associates, 2000

[9]Scott Guelich, Shishir Gundavaram & Gunther Birznieks. CGI Programming with Perl, 2nd Edition. O'Reilly & Associates, 2000

[10]mysql web site. www.mysql.com

[11]MySql reference manual. www.mysql.com/doc

[12]Apache HTTP server web site. www.apache.org

[13]Apache HTTP server documentation. www.apache.org/docs

[14]OpenSSL web site. www.openssl.org