

System Requirements

PURE (PubMed article recommendation system) was implemented and tested on Fedora Core (FC) 4 and FC 5. Thus, the following instructions will vary for other distributions. PURE requires Apache, MySQL, PHP, and Perl to be installed and set up. Prior knowledge of these programs is a required before proceeding. In addition, sendmail (or compatible MTA (e.g. Postfix)) and the Perl DBI module for MySQL (which is part of a FC4 and FC5 installation) need to be installed and configured already.

Install

- Edit `/etc/php.ini` file and adjust the parameters as described below.
error_reporting = E_ALL & ~E_NOTICE
register_globals = On
max_execution_time = 600
Execution time is highly dependent on your network speed because the PHP script retrieves PubMed records over the network. Therefore, you should allow enough time when selecting a value for the `max_execution_time` parameter.
- Add a user for PURE (e.g. pure) and setup its home directory. The home directory (e.g. `/var/www/html/PURE`) is used for the WWW service and execution of the PURE scripts, which are executed by cron.
- Continue the following setup procedure as the PURE user.
- Decompress the PURE package (e.g. `tar -zxvf PURE.tar.gz`) into the home directory. A list of files in the package is shown in `file_list.txt`.
- Confirm permissions. Perl scripts should be executable, other files should be readable and the temp directory should be writable for any user.
- Create a MySQL database named "PURE".
- Import the dumped MySQL tables (e.g. `mysql -u root -p PURE < PURE.sql`).
- Edit the beginning of `config.pl` by changing the mysql host, user, and password for your system.
- Edit the "Setup section" of `main.php`. It is important to confirm the parameters related to MySQL and the proxy server. If you connect to the Internet directly, the value of `$proxy_host` parameter should be empty (VERY IMPORTANT).
- Execute the `get_ref.pl` script. The files to be generated by this execution are shown in `file_list.txt`. If you use a proxy server for Internet access, set the "HTTP_PROXY" environment variable (`HTTP_PROXY=proxy_server_name:port`) prior to execution. Since this script downloads 10,000 records from PubMed and computes a background word distribution, some time is required to complete. If execution of the script is successfully launched, a `Ref.txt` file containing 10,000 records in the MEDLINE format, is created in the temp directory.
- Set the "HTTP_PROXY" environment variable in crontab if you use a proxy server.
- Schedule the two scripts, `check_parameter_update.pl` (PS1) and `daily_check.pl` (PS2). PS1 checks if users updated the criteria for preferred articles and modifies the model parameters accordingly. Therefore, this script should be scheduled at night. PS2 downloads, analyzes and recommends the newly added records. The PS2 should be scheduled at night in Eastern Standard Time (EST) as instructed by the PubMed e-Utility site of NCBI.
- Access `index.php` with a web browser. If the login page is not displayed, check the configuration of the web service.

PURE logs its progress to the `"PURE.log"` file. When PURE runs PS1, `"#---Parameter Update..."` is written to the log file. When PURE runs PS2, `"#---Daily Check..."` is written. If you cannot find

these entries in the log file, double-check the crontab description and ensure the scripts are scheduled by the PURE user. “Count = *number*” in PS2's log file indicates the number of retrieved records. If the *number* is empty (i.e., not zero), there was a problem accessing the PubMed server from the script. In this case, check the environment settings.

How to use

PURE is easy to use and we demonstrate it through a simple tutorial.

User Registration

Click “Register new user” in the left menu frame and enter the appropriate information.

PURE (PubMed Article Recommendation System)

Menu <ul style="list-style-type: none">• Login page• Latest highly rated articles• Previous highly rated articles• Register new user• Edit user information• Delete user information• Edit preferred articles	UserID: <input type="text" value="yoneya"/>
	Name: <input type="text" value="Takashi Yoneya"/>
	E-mail: <input type="text" value="yoneya@xxx.yy.ac.jp"/>
	Password: <input type="password" value="*****"/>
	Mail Report: <input type="text" value="Title"/>
	Report Count: <input type="text" value="10"/>
	<input type="button" value="Submit"/>

You can select report type as “Title only” or “Title + Abstract”. “Report count” is the number of articles recommended everyday. After completing the form, click the submit button. If you want to edit your preferences, click “Edit user information”; if you want to delete a user, click “Delete user information”.

Registration of preferred articles

When you access PURE for the first time, the login page is displayed. Input your user id and password to log into the system. The latest list of highly rated articles is displayed after logging in, if they exist. Click “Edit preferred articles” to register preferred articles.

PURE (PubMed Article Recommendation System)

Menu

- [Login page](#)
- [Latest highly rated articles](#)
- [Previous highly rated articles](#)
- [Register new user](#)
- [Edit user information](#)
- [Delete user information](#)
- [Edit preferred articles](#)

New PubMed IDs

16798067
16798010
16798113

submit

Input PubMed IDs, one per line, and then click the submit button.

PURE (PubMed Article Recommendation System)

Menu

- [Login page](#)
- [Latest highly rated articles](#)
- [Previous highly rated articles](#)
- [Register new user](#)
- [Edit user information](#)
- [Delete user information](#)
- [Edit preferred articles](#)

Are you sure to update these records?

Add as Preferred Articles

PMID	TITLE
16798067	User interpretations of future information system use: A snapshot of technological frames.
16798010	Incorporating a TEV cleavage site reduces the solubility of nine recombinant mouse proteins.
16798113	Recombination between retrotransposons as a source of chromosome rearrangements in the yeast <i>Saccharomyces cerevisiae</i> .

Update

If the titles are not displayed, confirm the PubMed IDs and network settings. If the records are correct, click the update button. When you click “Edit preferred articles” after registering the articles, PURE shows the next screen.

PURE (PubMed Article Recommendation System)

Menu

- [Login page](#)
- [Latest highly rated articles](#)
- [Previous highly rated articles](#)
- [Register new user](#)
- [Edit user information](#)
- [Delete user information](#)
- [Edit preferred articles](#)

New PubMed IDs

Delete Registered IDs

Delete	Reg. Date	PMID	Title
<input type="checkbox"/>	2006-06-28	16798067	User interpretations of future information system use: A snapshot with technological frames.
<input type="checkbox"/>	2006-06-28	16798010	Incorporating a TEV cleavage site reduces the solubility of nine recombinant mouse proteins.
<input type="checkbox"/>	2006-06-28	16798113	Recombination between retrotransposons as a source of chromosome rearrangements in the yeast <i>Saccharomyces cerevisiae</i> .

Input additional PubMed IDs in the top text box if you want to add new articles. If you want to delete some articles, check their corresponding check boxes and then click the submit button.

Browsing of the recommended articles

The latest list of highly rated articles is displayed after login.

PURE (PubMed Article Recommendation System)

Menu

- [Login page](#)
- [Latest highly rated articles](#)
- [Previous highly rated articles](#)
- [Register new user](#)
- [Edit user information](#)
- [Delete user information](#)
- [Edit preferred articles](#)

Recommended articles in 2006-06-09

	PMID	Score	Title
<input type="checkbox"/>	16760148	6.84	Co-morbidity in prostate cancer.
<input type="checkbox"/>	16760422	6.53	Following tetraploidy in an Arabidopsis ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes.
<input type="checkbox"/>	16760022	5.54	A paradigm shift: from disease to health orientation.
<input type="checkbox"/>	16760146	5.40	Prostate cancer: how much do we know and how do we know it?
<input type="checkbox"/>	16760000	5.26	Studying health outcomes in farmworker populations exposed to pesticides.
<input type="checkbox"/>	16758751	5.25	[The expectations for results of clinical trial "IDEAL"]
<input type="checkbox"/>	16758764	5.22	Alteration of gene expression in human cells treated with the agricultural chemical diazinon: possible interaction in fetal development.
<input type="checkbox"/>	16758784	5.06	Case reports of basaloid squamous cell carcinoma--an aggressive variant of squamous cell carcinoma.
<input type="checkbox"/>	16759270	4.96	Head & neck cancer incidence in Wales 1992--2001.
<input type="checkbox"/>	16761104	4.60	On using a cancer center cancer registry to identify newly affected women eligible for hereditary breast cancer syndrome testing: practical considerations.

Click the PubMed ID if you want to read the abstract. If you want to add some records as preferred articles, click the check box followed by the "Add_to_preferred_articles" button. If you want to see the previous results, click "Previous highly rated articles".

PURE (PubMed Article Recommendation System)

Menu

- [Login page](#)
- [Latest highly rated articles](#)
- [Previous highly rated articles](#)
- [Register new user](#)
- [Edit user information](#)
- [Delete user information](#)
- [Edit preferred articles](#)

2006-06-09 ▾

Submit

Select the date you want to see and click the submit button and PURE shows the list of articles for the specified date.

Results sent by e-mail.

PURE sends an e-mail alert and users can select title only or title plus abstract as the e-mail format.

```
Go To PURE
PMID: 16784829, SCORE: 6.41
Clinical, mammographic and ultrasonographic features of blunt breast
trauma.

Gatta G, Pinto A, Romano S, Ancona A, Scaglione M, Volterrani L
Department of Diagnostic Imaging, "A. Cardarelli" Hospital, 80131 Naples,
Italy.
Eur J Radiol. 2006 Jun 17;.
```

When you click the PubMed ID, the title and abstract as stored in PURE are displayed. If you click "Go To PURE", PURE shows the list as of that date. If you access PURE via e-mail, you login to PURE automatically.