

CMS に基づく医学/生物学分野向け文献管理データベース

池野 英利, 西岡 拓人, 大澤 五住¹, 白井 支朗²
社会システム環境学大講座, 大阪大学大学院生命機能研究科¹,
理化学研究所脳科学総合研究センター²

CMS based bibliography database system for medical and biological fields

Hidetoshi IKENO, Takuto NISHIOKA, Izumi OHZAWA¹, Shiro USUI²

Biological Information System Laboratory
School of Human Science and Environment,
University of Hyogo/Himeji Institute of Technology,
1-1-12 Shinzaike-honcho, Himeji, Hyogo, 670-0092 Japan
Graduate School of Frontier Biosciences, Osaka University¹
1-3 Yamadaoka, Suita, Osaka, 565-0871 Japan
Brain Science Institute, RIKEN²
2-1 Hirosawa, Wako, Saitama, 351-0198 Japan

Abstract

PubMed (<http://www.ncbi.nlm.nih.gov/entrez/>) is the standard public database system for publications in medical and biological fields. Every registered paper in this database is assigned a unique ID number, PMID. Detailed information for a given paper can be obtained from the database using PMID as a key. Recently, applications of Content Management System (CMS) have become widespread for constructing WWW portal sites. Taking advantage of these resources for neuroinformatics, we have applied the CMS technologies for providing a data sharing environment in the laboratory. In this paper, we present a bibliographic database module, PubMed-PDF, which can be used on a PHP-based CMS, XOOPS. The module is capable of managing PDF (Portable Document Format) reprint files in addition to their bibliographical information provided by PubMed. Registered literature information is automatically indexed based on the title, authors, year of publication, journal and key words in abstracts. The system may be highly useful for literature management not only for personal or laboratory use, but also for neuroinformatics portal sites.

Keywords: PubMed, database, content management system, bibliography management, neuroinformatics

1 はじめに

医学・生理学、生物学分野における文献データベースとしては、NIHで構築されたMEDLINEが広く利用されており、ここに登録されている情報はPubMedサイトからインターネットを介して無料で利用することができる。PubMedに登録されている文献には固有の番号(PMID)が付けられており、このPMIDによって該当する文献に関するタイトル、著者名、発行年、掲載誌などに関する情報を参照することができる。

このようなPubMedの機能を個人の文献情報管理に活用する方法が、平成11~15年の文部科学省振興調整費「視覚系のニューロインフォマティクスに関する研究」プロジェクトにおいて提唱され、これをサポートするツール類も開発、公開されている[1, 2]。このように統一的な文献管理の方法が確立されることによって、研究室や研究テーマの近いグループ間などで書誌情報を共有することも容易になり、コンピュータ、ネットワークの効果的な活用に結び付くことが期待できる。

さて、インターネットによる情報共有の方法としては World Wide Web(WWW) が広く活用され、個人から国家レベルに至る様々な規模のサイトが日々誕生している。このような WWW システムの構築においては、最近ではデータベースをバックに持ち、コンテンツの管理およびサイト運用の業務など、ほとんどの作業を WWW ブラウザを通じて実施するシステム環境：コンテンツマネジメントシステム (CMS) が広く用いられるようになってきている。CMS の導入はサイトの構築、管理を飛躍的に簡単化し、費用と時間の節約につながることから、標準的な Web 技術として定着しつつある。

本研究では、PHP をベースとした CMS : XOOPS を基本ツールとして利用し [3]、この環境下において PubMed と連携する形で、書誌情報の集約、共有ができる機能を提供するモジュール : PubMedPDF を開発した。このモジュールを用いることによって、PMID が付加された文献については、自動的に文献の詳細情報を取得し、著者、発行年、掲載雑誌別のインデックスを作成することができる。さらに、論文リプリントが PDF ファイルとして手元にある場合には、このファイルをデータベースにアップロードしておくことによって、文献一覧から簡単に参照することができるようになる。なお、PMID のない文献については、登録ページにおいて情報をユーザが登録することによって、同じシステム上で管理することができる。

2 システム構成と特徴

本研究において基本システムとして使用した XOOPS は、フリーソフトウェアとして利用できる CMS の一つであり、MySQL をデータベース、Apache を WWW サーバとして用い、スクリプト言語 PHP によって記述されている [4, 5, 6]。開発したモジュールは XOOPS が利用できる環境 (データベース: MySQL, WWW サーバ: Apache, サーバスクリプト言語: PHP4) であれば直ちに利用することができる。最近では、XSAS, XAMPP など、XOOPS に必要なソフトウェアをまとめたサーバスタックパッケージも公開されている。これらを使用することによって、Windows, Linux, Mac OS X など様々な OS において、パッケージをコピーするだけで運用を始めることができるため、コンピュー

タ、ネットワークに関する専属のスタッフがいない個人、あるいは、小規模の研究室においても極めて容易に導入、利用することができる [7]。

なお、XOOPS については導入が容易でありながら、大規模なサイトの運用にも耐えうることが示されている。また、モジュール単位で様々な機能が追加できることから拡張性も優れている。本研究においても、ユーザ管理、ネットワークアクセスなどに関しては XOOPS 自体が持つ機能を活用することによって、文献情報の登録、利用、検索などに焦点を絞った機能の実装を進めることができ、実用的な文献データベースシステムを短期間、低コストで実現できることが実証された。

開発したデータベースシステム : PubMedPDF には、以下のような特徴がある。

- 情報の入力、閲覧などのあらゆる操作が Web ブラウザから行える。
- PMID を入力することによって PubMed サイトへの自動アクセスが行われ、該当する文献情報の入手と著者別、発行年、掲載誌毎の分類が実施される。
- PMID で名前を付けられた論文リプリントファイル (ファイル名の本体は PMID, 拡張子は pdf とする) をアップロードすることにより、著者別、発行年、掲載誌毎に自動的に分類、保存される。
- データベースに登録された文献について、参照者のコメント、関連資料などを付加していくことができる。
- 著者、キーワードなどによる PubMed サイト及びデータベース内の文献検索機能を備えている。
- PMID のない場合にも、文献情報を入力することによって同じ環境下で文献の管理、共有が可能である。
- テーマ毎に文献情報をグループ化することができる。
- 論文誌の書式に応じたテキスト、BIBTeX 形式での文献リストが出力できる。
- 新たに登録された文献情報を XOOPS の標準モジュールである News モジュールの記事として自動的に投稿、掲載させることができる。

2.1 モジュールの構造

PubMedPDF は XOOPS のモジュールとして実現されており、図 1 に示すように XOOPS モジュールの構造に基づき、admin, images, language, sql ディレクトリにモジュール管理者機能、アイコン画像、英語・日本語メッセージ及びテーブル構造に関するファイル類をまとめている [8]。また、include ディレクトリには共通的に利用されるプログラムを保存している。

xoops_version.php をはじめとする basic functions は、モジュールの基本情報を設定しているプログラムであり、ここから呼び出された author.php, year.php, journal.php において文献情報の一覧表示をおこなう機能を提供している。文献情報の登録や変更、検索については、図 1 においてグループ分けしたプログラム群 (register & edit functions, search & bibliography list functions) によって実装されている。登録してある文献に関する PMID リストの作成や PMID のない文献に関する XML 情報ファイルの書き出し、読み込みなどの機能については maintenance functions において提供されており、これらの機能によりデータベースのバックアップや移行が簡単に行えるようになっている。なお、このモジュールにおいては各機能が独立したプログラムで実装されており、機能拡張やデバッグが容易な構成となっている。

2.2 モジュールの導入と利用

XAMPP を使った運用の場合においては、XAMPP のパッケージを指定されたディレクトリ (Linux の場合は /opt, Windows の場合は C:) にコピーし、XAMPP を起動する。XAMPP ディレクトリ下にあるhtdocs に XOOPS のパッケージファイルをコピーし、Web ブラウザから XOOPS の環境を設定することでの利用環境が整う。PubMedPDF については、モジュールファイルを XOOPS ディレクトリ下の modules ディレクトリにコピーし、管理者メニューのモジュール管理においてインストールすればよい。インターネットアクセスに Proxy を使用する場合は、PubMedPDF モジュールの設定メニューで Proxy のアドレスとポートを設定しておく。なお、付録において XAMPP に PubMedPDF などのモジュールを加えた研究室向け

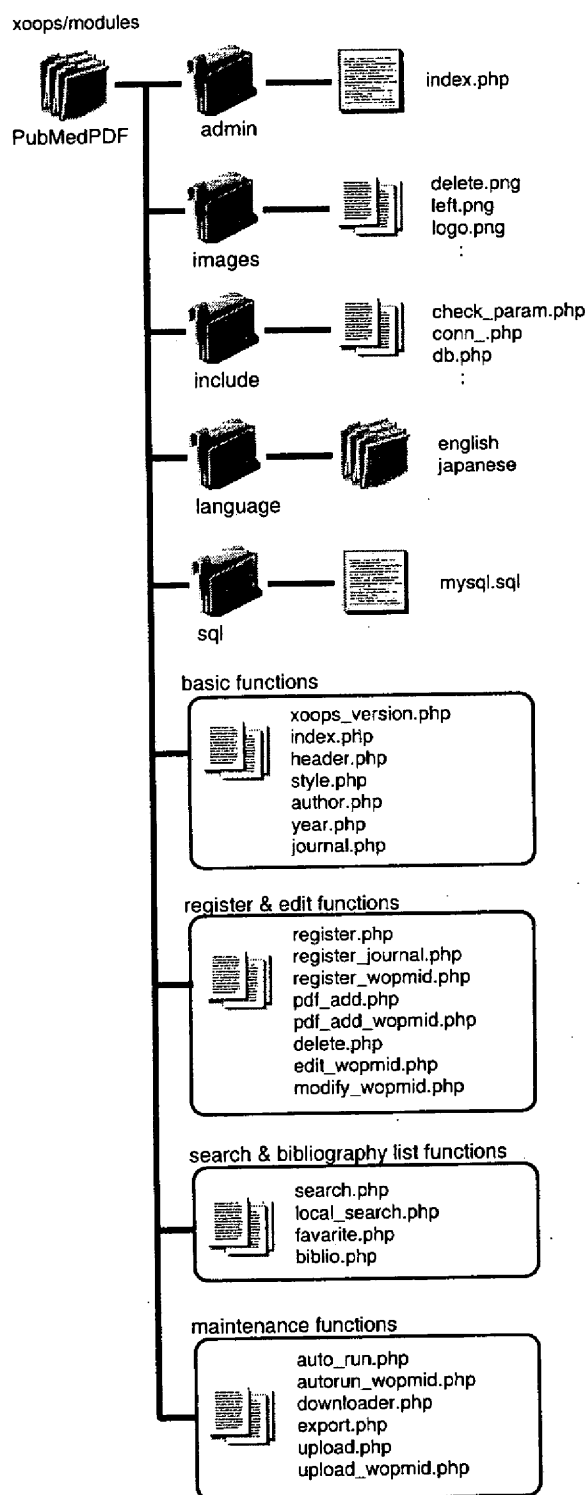


図 1: PubMedPDF モジュールのプログラム構造

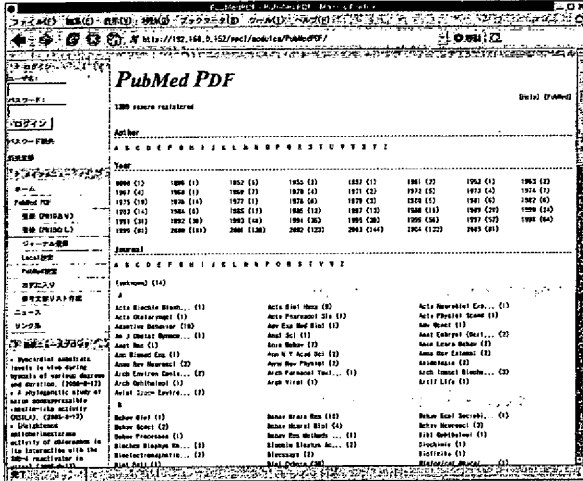


図 2: PubMedPDF のトップページ

パッケージ XLP(XOOPS Laboratory Portal) の導入方法を記載した。

本システムの起動により、ネットワークに接続しているコンピュータからデータの登録、閲覧が可能となる。サーバに対して Web ブラウザからアクセスすることによって、図 2 のように、登録してある文献情報について著者名 (Last name のイニシャル)、発行年、発行誌からのアクセスページがトップページとして表示される。

PMID を持つ文献については、図 2 の左側に表示されたメニューから、「登録 (PMID あり)」を選択することで表示される登録ページ (図 3) において PMID を入力する、あるいは、PMID に基づく名前の PDF ファイルをアップロードすることによって登録することができる。この操作によって、モジュールは PMID に基づき PubMed サイトへ自動的にアクセスし、該当する文献情報をデータベースに登録する。

なお、登録データの削除や、データベースに登録されている文献情報に関する PMID の一覧をファイルとしてダウンロードしたり、逆にそれを使って一括登録することもこのページから行うことができる。この機能は、データベースの継続的運用、バックアップにとって非常に有効であり、さらに、各研究室において集約された文献情報は、特定のテーマに分類された文献リストという新たな価値を生むことになる。

一方、PMID を持たない文献情報については、まず、掲載誌に関する情報 (英語名称、日本語名称、発行所

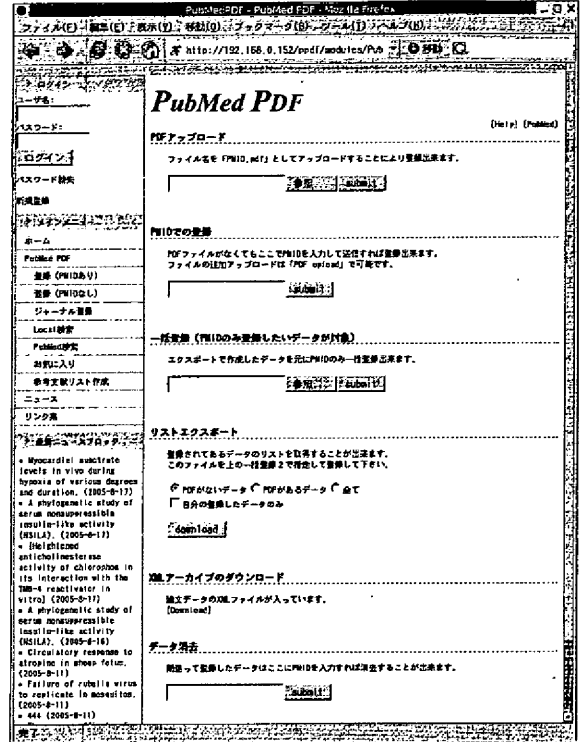


図 3: PMID による文献情報の登録

など) をジャーナル登録機能によって登録する。文献情報については、「メニューから登録 (PMID なし)」をクリックし、登録ページにおいて文献に関する情報を入力する (図 4)。この登録操作により、文献情報はデータベースに書き込まれると同時に XML ファイルが作成される。この XML ファイルを用いることにより、別のマシンへの転送、共有を行うことが簡単にできる。データベースに登録された文献情報の XML ファイルを保存しておくことによって、登録情報を再現することができることから、ユーザレベルでのデータバックアップにも利用することができる。

こうして登録された文献に対しては、トップページの著者名のアルファベット、発行年などからアクセスできる。PDF ファイルが添付されている文献については、文献名が太字で表示される。右下にある [PDF] という PDF ファイルへのリンクをクリックすることにより、図 5 に示すように該当する文献の PDF ファイルが表示される。

各文献の PMID をクリックすることで該当文献の PMID サイトに移動することができ、[favorite] によ

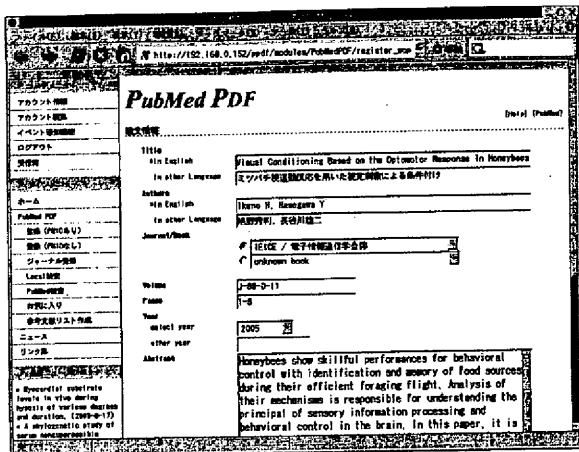


図 4: PMID を持たない文献情報の登録

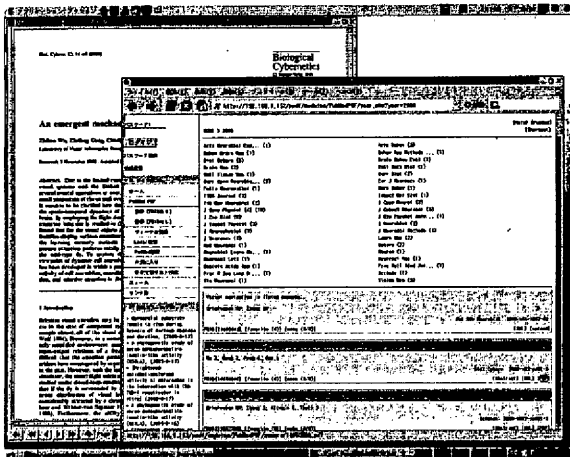
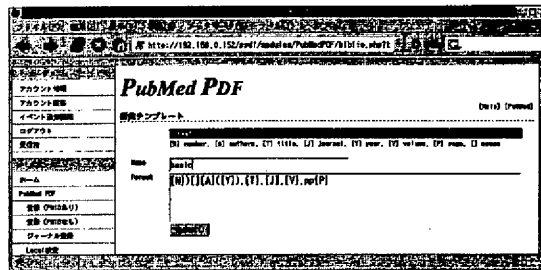


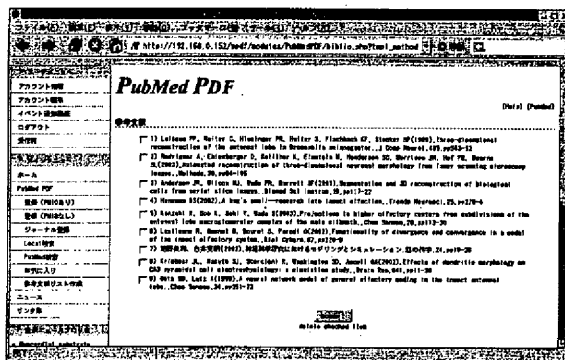
図 5: 文献情報とリプリントファイルの参照

りテーマなどに基づき文献情報のグループを作成することができる(お気に入り機能)。なお、このグループはユーザごとに作成、管理することができる。また、[note]によって閲覧者のメモやファイルを追加することが可能であることから、関連した資料などを文献とともに管理することができる。付加されたメモについても追加したユーザのみ、あるいは、全てのユーザが閲覧可能に設定できる。

グループ化された文献情報へは、メニューの「お気に入り」から参照することができる。グループは階層的に構成することができることから、テーマ別、著者別など細かい分類が可能となっている。本モジュールの開発方針として、データの登録の手間は極力排除し、



(a) テンプレートの編集



(b) 生成された文献リスト

図 6: テンプレートと文献リストの生成結果

データベースに集約されたものを活用する際にユーザが目的に応じて編集していくものとした。すなわち、利用者が登録された情報を使って「お気に入り」を増やしていくことによって、テーマ、目的に応じたデータベースへと発展させることができる。

グループ化された文献情報については、ユーザが設定した形式に整形した参考文献リストを作成することができる。文献リストの作成にあたっては、図 6 (a) に示した新規テンプレート作成によって文献タイトル [T]、著者名 [A]、発行年 [Y] などの出力順、形式などを設定し、これをテンプレートとして保存する。こうして保存されているテンプレート名を選択し、出力先を参考文献リストの表示、テキストファイルでのダウンロード、BIBTEX フォーマットファイルのダウンロードの中から選択することで文献リストが作成される(図 6 (b))。

3 評価と考察

開発したモジュール PubMedPDF は, sourceforge.jp などを通じて公開しており, 2004年12月下旬以来, 2,100件以上のダウンロードを記録, 研究室における文献情報管理に留まらず研究成果の公表などにも利用されている [9]. XAMPP のようなパッケージを用いることで, XOOPS 及び PubMedPDF モジュールは, OS に依存せず簡単に利用することができ, 大学の研究室などにおいても容易に導入できるシステムとして提供ができた. なお, インターネットに公開されているサイトで利用する場合には, サーバや登録情報に関して, XAMPP にセキュリティ設定機能や Apache のアクセス制限によるセキュリティの確保が重要となる.

今やインターネットを通じて膨大な情報を入手することができることから, 本当に役立つ, 必要な情報を集約, 共有していく環境の構築が重要となってきた. 文献情報についても, 研究者レベルで EndNote など市販のソフトウェアを活用しているケースが多いが, 各研究者が蓄積, 整理した情報はそれ自身が重要な資産であり, 研究室, ソサエティなどでの共有していくことが大切であると考えられる. 我々が進めているニューロインフォマティクスに関する研究プロジェクトは, 神経科学分野における情報技術活用を推進していくことを目的としており, 書誌情報から実験データ, 数理モデルなど様々な情報の統合によって, 新たな学際的研究領域への発展を目指している.

PubMedPDF については, 集約された文献情報から重なるキーワードの自動抽出およびこれに基づく分類や PubMed に対する自動巡回機能などの導入など, さらなる改良, 発展が考えられる. また, 登録情報を公開するユーザ, グループの設定, 情報の登録, 閲覧状況を把握するためのログ機能なども, 幅広く活用していくために必要な機能であり, 現在, その導入を進めている. さらに, 現在, XOOPS をベースに実験データなどを管理するモジュールの開発を進め, 実験データから文献情報, 研究資料にいたる様々な研究リソースを統合的に管理するシステムへの展開を進めている.

4 むすび

本研究では, PubMed で公開されている文献情報を活用し, 研究室や個人レベルで目的, テーマに応じた

データベースシステムを構築するためのツールを開発した. 医学, 生物学分野は, 日々, 多くの報告, 論文が発表される研究分野の一つであり, 最新の研究動向を把握し, 新たな研究手法, テーマを進めていく上で, 効率の良い文献調査, 管理は極めて重要な技術である.

このような情報技術を活用した研究情報の管理, 共有について, 脳・神経科学分野においては, OECD の勧告によって設立され国際ニューロインフォマティクス統合機構 (the International Neuroinformatics Coordinating Facility: INCF) を中心に急速に進展している. 我国においても神経情報基盤センター (Neuroinformatics Japan Center: NIJC) が設立され, 神経科学分野を中心とするポータルサイトの設置及びそのためのソフトウェアの開発が進められている [10]. 本システムもこのプロジェクトの一貫として開発が進められており, 個人, 研究室レベルでの情報の集約, 管理に加え, こうした情報資源を社会的に共有していくシステムの構築, 運用を進めていくことが今後の課題である.

謝辞

システム開発に関して貴重な御意見をいただいた, 兵庫県立大学/姫路工業大学環境人間学部池野研究室のメンバー各位に感謝いたします. 特に, 鉢田卓也氏, 金井涼子氏には, テスト運用, ドキュメント作成などに関して, 多大なるご尽力をいただいた. さらに, 本システムをダウンロード, 利用いただき, コメントいただいたユーザ各位に感謝いたします.

参考文献

- [1] 白井支朗, 池野英利 (2004): 視覚系のニューロインフォマティクスサイト: Visiome Platform, 日本神経回路学会誌, Vol.11, No.4, pp.193-199
- [2] PubMedPDF Tools:
<http://www7.bpe.es.osaka-u.ac.jp/pubmedpdf/>
- [3] XOOPS 日本公式サイト: <http://jp.xoops.org/>
- [4] 高井 守 (2004): XOOPS コミュニティサイト構築ガイド, 東京: 技術評論社

- [5] 坂井 恵, 天野龍司 (2004): XOOPS 入門, 東京: 翔泳社
- [6] 堀田倫英, 石井達夫, 廣川 類 (2000): PHP4 徹底攻略, 東京: ソフトバンクパブリッシング
- [7] XAMPP:
<http://www.apachefriends.org/en/xampp.html>
- [8] 後藤峰陽, 上松秀彦 (2005): Customizing Xoops, 東京: 毎日コミュニケーションズ
- [9] PubMedPDF:
<http://sourceforge.jp/projects/ppdf>
- [10] XooNIps:
<http://sourceforge.jp/projects/xoonips>

A XAMPP/XLP の利用方法

ファイルのコピーのみで, PubMedPDF を始めとする XOOPS に基づく研究室グループウェア環境を提供するパッケージ XOOPS Laboratory Portal (XAMPP/XLP) を用いることによって, 研究室レベルでの情報共有環境を構築することができる。本付録では, その手順について簡単に述べる。なお, Mac OS X については, 大澤研究室 (<http://ohzawa-lab.bpe.es.osaka-u.ac.jp/resources/portalpackage.html>) で配布されているパッケージを利用するのが便利である。

1. パッケージのダウンロード
このパッケージは, 配布サイト
<http://bis.shse.u-hyogo.ac.jp/xampp-xlp/>
からダウンロードできる。
2. インストール
ダウンロードしたパッケージを展開後, 次のように展開したディレクトリ (フォルダ) をコピーするだけでよい。

Windows: xampp フォルダ以下を c:\ にコピーする

Linux: 管理者となって, 次のような操作で, lampp フォルダ以下を /opt/ にコピーし, 全てのファイルの所有者を nobody に変更する。

```
# cp -r lampp /opt
# chown -R nobody /opt/lampp
```

3. XAMPP を起動する。

Windows: c:\xampp フォルダの xampp_start を実行する。

Linux: 管理者権限で, 以下のよう

```
# /opt/lampp/bin/lampp start
```

を実行する。

4. XAMPP を終了する。

Windows: c:\xampp フォルダの xampp_stop を実行する。

Linux: 管理者権限で, 以下のよう

```
# /opt/lampp/bin/lampp stop
```

を実行する。

5. アンインストール

Windows: c:\xampp フォルダを削除する。

Linux: 以下のようにして /opt/lampp を削除する。

```
# rm -r /opt/lampp/
```

6. WWW ブラウザから <http://localhost/xlp> にアクセスする。ユーザ名 root (パスワードも同じ) でログインすることにより, ユーザメニューから管理者メニューに移動することができる。なお, ネットワーク接続に Proxy を使用している場合には, 管理者メニューにおいて, PubMedPDF アイコンをクリックすることで表示される configure で Proxy の IP アドレス, ポート番号を設定する。
7. ユーザ, モジュールの追加やページのレイアウトなども全てこの管理メニューにおいて行うことができる。詳細については, XOOPS 関係の書籍を参照されたい。

(平成17年10月3日受付)