

# Integration of libre software applications to create a collaborative work platform for researchers at GET

Olivier Berger

GET/INT

olivier.berger@int-evry.fr

Christian Bac

GET/INT

christian.bac@int-evry.fr

Benoît Hamet

benoit.hamet@laposte.net

## Abstract

*Libre software provides powerful applications ready to be integrated for the build-up of platforms for internal use in organizations. We describe the architecture of the collaborative work platform which we have integrated, designed for researchers at GET. We present the elements we have learned during this project in particular with respect to contribution to external libre projects, in order to better ensure the maintainability of the internal applications, and to phpGroupware as a framework for specific applications development.*

## Keywords

Free software, Libre software, Open Source software, in-house applications, collaborative work environment, groupware, phpGroupware, wiki, TWiki, Sympa, WebDAV, OpenLDAP, contribution, ProGET, PicoLibre

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

# 1 Introduction

ProGET is a collaborative work platform build out of a set of specialized libre software applications integrated together. ProGET is designed for the whole of teachers/researchers at GET. It provides every GET research projects with the best features found in each application (Wiki, mailing-lists management, shared WebDAV folders, Web portal...).

We start with a description of the libre components that have been integrated and of the features that have been selected, as well as elements of architecture of the developed platform.

We will then introduce the strategy for collaboration that we have devised for our contribution to phpGroupware.

We finish with a first evaluation at the end of the initial development phase.

## 2 Research at GET

The *Groupe des Écoles des Télécommunications*<sup>1</sup> (GET) is composed of several engineering and business schools together with research centers in Paris (ENST), Brest (ENST Bretagne) and Évry (INT), in France. The research teams are made up of more than 600 full-time research equivalents.

The range of the researchers' expertise, from technologies to social sciences, enables the integrated approach so characteristic of GET research and fosters its adaptability to new application sectors and new usages in response to current challenges in the field of Information and Communication.

To give a clearer view of research at GET, the Research Office started to catalogue the activities from the different locations so that the research may be described in terms of research *projects*

---

<sup>1</sup><http://www.get-telecom.fr/>

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

and *programmes*. A project is made-up of a group of people working together on closely related subjects. For example the authors belong to the “Collaborative Platforms for Research” (PFTCR) project. A programme associates different projects loosely related; for example our project is related to the “Web and Information society” programme. Due to the fact that GET teams are located in different areas, the research office also decided to propose a Web platform to help researchers collaborate through groupware tools and animate their research work.

### **3 ProGET integrated platform for collaborative work**

*ProGet* has been launched in July 2003 with the following goals:

1. provide all research teams in GET and their external partners (more than 1250 users) with cutting-edge technologies in terms of Web based groupware tools;
2. allow the Research Office to manage the *administrative* records describing research projects;
3. generate a public *Web portal* based on information extracted from both previous components<sup>2</sup>.

The home page of the Web portal is shown Figure 1<sup>3</sup>.

---

<sup>2</sup>the GET research portal can be accessed at <http://proget.int-evry.fr/>

<sup>3</sup>The screenshots of the application provided in figures of this paper display pages in French. Not all interface elements have been provided with a translation to English yet in the ProGET application.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

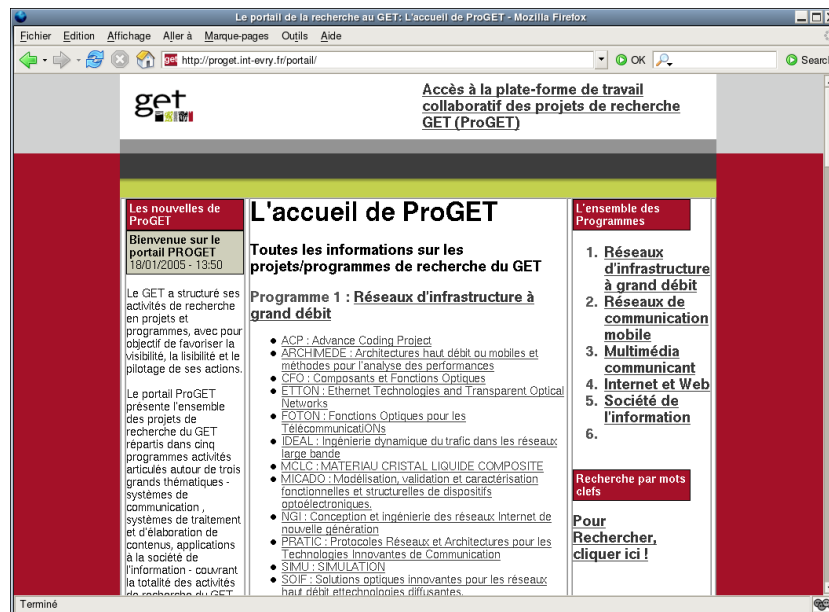


Figure 1 Homepage of the GET public research portal

### 3.1 Groupware tools for researchers/teachers

The following features have been selected as corresponding to the basic needs for collaborative work in the context of research activities at GET:

- **Document sharing:** people in the same project must have a way to share the documents they produce, whatever the type of document;
- **Asynchronous communication:** each project manages predefined *mailing-lists* and may also create mailing-lists at its needs;
- **Online editing:** people in the same project group are allowed to write easily and in collaboration simple hypertext pages using their Web browsers, to create a collaborative Web, in a Wiki-type tool [wiki];
- **Publishing short announcements:** the project manager can write short news items about the project, and have these project news published on-line on the public Web portal very easily.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

The platform must provide secure access, from any point on the Internet, to a set of tools. These tools will have to be accessible through a Web interface, to allow people to use them without the need to install any specific application on the computer from which they are connected.

Some features must also be accessible through non-Web client tools, in the Microsoft Windows™, or GNU/Linux environments. For instance, as explained in section 5.4, access to the DAV repositories through DAV-compatible file managers.

Information generated in this collaborative part of the platform may be made accessible to the public portal directly by the research teams, without further interaction with a Webmaster.

### **3.2 Tools for the Research Office**

The Research Office is supervising the setup and evolution in time of the research projects and programs. In ProGET, it uses, for this purpose, a dedicated tool, which helps managing the life-cycle of the *project description forms*.

This module does not provide a replacement for the research information system under development at the present time at GET. It only allows the management of informations on the projects which are necessary for the other parts of the platform. It handles the initial list of persons participating in projects teams, people who are project managers, the projects' descriptions, and their yearly goals.

The project description forms are initially filled by the researchers who will be responsible for the projects. They are then validated by the Research offices. This validation allows the creation of the collaboration projects in the platform. Those forms are also used to feed the contents of the research public Web portal.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

### 3.3 Public Web portal

The platform publishes a public Web portal which reflects the state of research activity at GET. This portal describes essential informations relating to the research programmes and projects at GET. Figure 1 shows the home page of the portal.

This public Web site corresponds to a dynamic extraction of the contents of the project description forms stored in the database. That's why it will be updated immediately whenever the GET Research Office validates new projects creation, or updates the contents of existing forms.

Since the data extraction is dynamic, this portal website can be duplicated in several flavors according to the origin of the visit. So when the website is browsed from inside the GET intranet, it will provide more detailed informations, including some facts about the research projects which are considered to be published to a restricted audience only (teams composition, detailed yearly goals of the projects...).

## 4 Previous work with libre software collaboration tools

As described in [Cousin et al, 2002], we and other GET researchers and students, contributed in the development of a collaborative platform called *PicoLibre* [picolibre]. This platform is targeted at collaborative software development, to help students in Computer Science curricula and researchers to develop and host their software projects. It provides the necessary collaboration tools (mailing-lists, CVS repository, issue trackers, etc.) in a similar way to SourceForge [sourceforge]. Another goal of the PicoLibre project was to foster the publication of projects as libre software, since it would be a way to introduce the users to the practice of the common tools used everyday by libre software developers.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

This platform was created as a free software tool (published under the GNU GPL license<sup>4</sup>), using free software components and adapting them for PicoLibre. Several PicoLibre instances are in operation since fall 2001 in GET sites, and outside.

In some aspects, PicoLibre is a very successful project, since it allows GET to host a large number of software projects with a minimum burden for its administrators. It is also sometimes used for projects not specifically targeted at software development, for teams that need a collaboration space and associated *groupware* tools. The PicoLibre project has helped us get a better knowledge of the specifics of the design and development, but also the administration, of a collaborative Web-based platform in the context of a higher education institution.

But PicoLibre also failed to certain extents, especially with respect to its maintainability, and in its capacity to be integrated into a libre software distribution. The project was done in so a short time that developers only concentrated on the innovative parts and spent too few efforts in keeping the software they developed in the main stream of the libre projects it was based on (in particular with respect to phpGroupware [phpgw]).

Although having been used intensively by teams of researchers or students investigating software development, PicoLibre wasn't well suited for generic needs of teams of non software developers. For example, the CVS<sup>5</sup> revision management repository is a central tool in PicoLibre. It is very useful for software development, but doesn't fit well for casual researchers' use. For example, although it is possible with CVS to manage the revisions of a document written with an office suite, it is not very convenient, since it requires installation of a specific CVS client program. Most non software developers (for instance researchers in the field of finance or business administration) are thus not comfortable in using such a tool.

---

<sup>4</sup>GNU General Public licence : <http://www.gnu.org/copyleft/gpl.html>

<sup>5</sup>Concurrent Versions System : <http://www.nongnu.org/cvs/>

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

To build on top of our previous developments, and considering the aforementioned perspective, we proposed to start and build the ProGet platform using some of the free software modules that had been used to create PicoLibre, and to combine these modules with other existing libre software projects to fulfill new capabilities. We intended also to do our developments and the integration of the numerous applications necessary in a much more maintainable way.

## 5 Structure of the platform

The ProGET platform is composed of several specialized software applications, installed together on a dedicated machine running the GNU/Linux operating system (Debian).

This paper does not describe all of the components of the system, but only the modules which provide the highest level features, and the way they have been integrated into a single platform.

The development team has been participating for a long time in communities of research and practice in libre software : development of PicoLibre (cf. 4), organization of the *Autour du Libre* conference, participation in the *CALIBRE* FP6 European project<sup>6</sup>... We have then naturally preferred to use exclusively libre software for the development of the platform, in order to ensure the complete control of all technical aspects, and the conformance to open standards.

### 5.1 Integration principles

To deliver the required features, we have integrated different existing libre software applications that we will list in section 5.2. Most of these applications rely on the Apache web server, in particular for the PHP execution engine.

Each one of the products, taken apart, was often not covering all required featured. So instead of trying to add to one of these applications the missing features, we have preferred to integrate several applications rather specialized. We are therefore able to take advantage of the best

<sup>6</sup>FP6 IST Project 00433 "Coordination Action for Libre Software" : <http://www.calibre.ie/>

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

aspects of each, even if there's a risk of partial redundancy. For instance, phpGroupware provides a Wiki module, but rather limited. We have preferred to plug phpGroupware and TWiki, an advanced wiki engine, instead of engaging in a tougher effort of enhancement of phpGroupware's.

In each integrated libre software application, we have been careful to allow the possibility of a logical *partition* of the data stored in the application, to ensure that the research projects can be kept autonomous and enhance privacy (cf. 5.8).

To make sure users won't have to deal with a set of authentication tokens (login + password), the authentication will be done in the same way in all the modules, although they may be operated by separate applications. This is done in a classical way, by relying on interaction with a dedicated directory implemented by OpenLDAP [OpenLDAP]. Some components also share a database stored in a dedicated MySQL server.

Integration of the applications for the realization of ProGET was done in a rather traditional way, by sharing the lowest layers (OpenLDAP, MySQL) and with adapters (code *glue*), when the applications themselves were rather monolithic. It would have been preferable to take advantage of libre applications organized in a more modular way and supporting a Web service paradigm. But very few mature versions of libre software which may have been integrated offered such interoperability mechanisms at the start of our project. Even today, this kind of interfaces is not always entirely present, even if this approach seems to be preferred now for the future versions of projects like phpGroupware or Sympa, for instance.

## 5.2 Integrated libre software high-level applications

The main libre software applications that have been integrated in the ProGET platform are:

- **the Apache Web server**

the server provides the link between the client tools (browsers) and the Web applications.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

– **A WebDAV repository**

The mod\_dav<sup>7</sup> Apache module implements the shared Web folders which implement the documents repositories for the projects (cf. 5.4).

– **The phpGroupware engine and applications**

phpGroupware provides the Web interface of the “virtual desktop” of the collaborative workspace (cf. 5.3), and brings in standard groupware components. It also serves as the basis for the software infrastructure of the public portal. In addition to standard phpGroupware modules, we have added two new modules developed for the very own needs of GET. One module which extracts informations out of the collaborative work applications to deliver the public portal (cf. 3.3) and a module which enables the administrative management for the Research Office (cf. 3.2).

– **The Sympa mailing-list manager [sympa]**

This powerful mailing-list manager provides some asynchronous communication means inside and outside of the projects.

– **The TWiki Wiki server [twiki]**

This Wiki system, among the most advanced, provides a white board-like interface to implement the “knowledge base” of the projects.

– **The Agata Reports reporting tool<sup>8</sup>**

It helps the Research Office to issue queries on the projects description forms in order to answer internal or external requests about research activity at GET.

We hereafter describe some of the characteristics of the ProGET modules and the obtained features.

---

<sup>7</sup>DAV module for Apache : [http://www.webdav.org/mod\\_dav/](http://www.webdav.org/mod_dav/)

<sup>8</sup> <http://www.agata.org.br/>

### 5.3 Researchers/teachers virtual desktop in phpGroupware

Each researcher who connects to the platform gets access to his/her own “virtual desktop”, provided through phpGroupware. Figure 2 shows this start page for a researcher who belongs to two research projects (here LICIP and PFTCR).

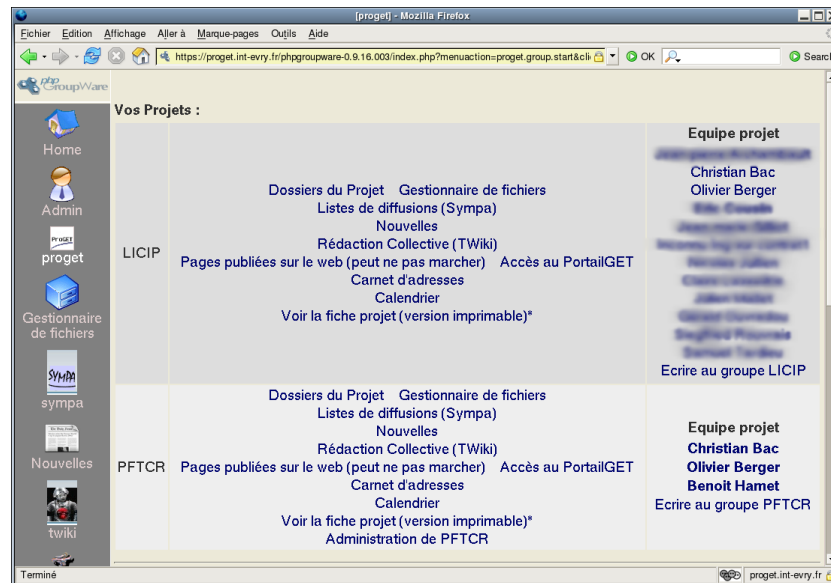


Figure 2 ProGET Collaborative workspace start page

The user will then be provided directly, for each project he/she belongs to, with a set of collaboration tools (with the corresponding links in the central column fig. 2) :

- Direct access to browsing of the contents of the documents repository of the project (cf. 5.4);
- Web-based file manager (Web interface for the management of the contents of this repository);
- Mailing-lists (access to the Web interface of Sympa for the lists of the project);
- News (management of the project's news displayed in the public portal);
- Project's Wiki (cf. 5.5);
- Project's specific Web pages (cf. 5.6);
- Access to the project's description page in the GET research portal (cf. 5.6);

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

- Shared address-book and calendar (standard phpGroupware tool);
- Printable project form (including restricted access informations).

Some of these tools are described in more details hereafter.

## 5.4 Project documents repository accessible through WebDAV

As described in [Dridi and Neuman, 1999], WebDAV suits well to create a groupware portal that helps people share their documents. In ProGet each project is provided, for its internal use, with a unique secured documents repository, in the form of a specific Web folder. It is shared among members of the project team. It can be accessed via HTTPS for browsing, or through WebDAV [WebDAV] (over HTTPS) for modification of contents of the repository. Figure 3 shows the contents of a repository through HTTPS in a Web browser (here, the public\_html subfolder in project PFTCR's folder).

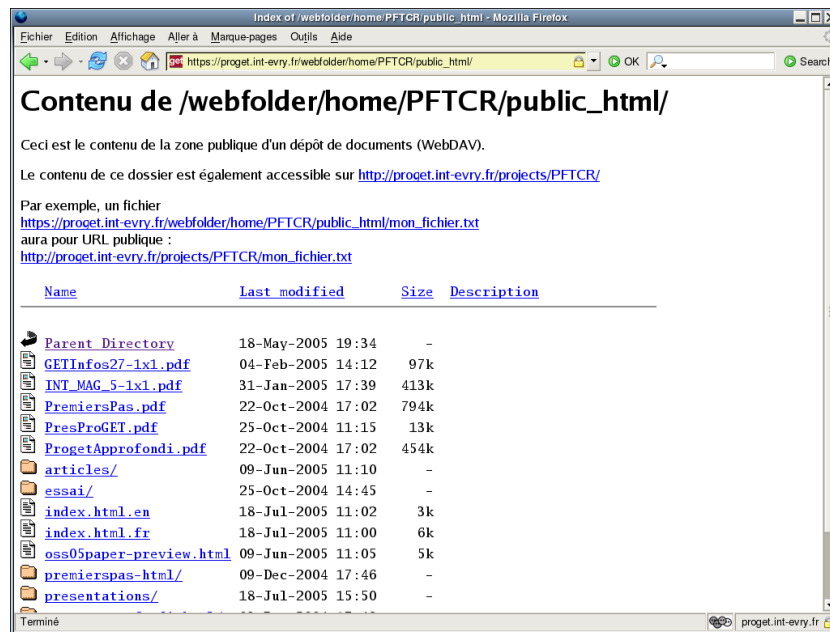


Figure 3 Contents of a project's DAV folder (webfolder)

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

The use of the WebDAV protocol enables transparent access in browsing or modification in the office applications, allowing for instance drag-and-drop in file managers for uploading. The uploading is thus extremely simplified<sup>9</sup> compared to the use of protocols like FTP or CVS. DAV compatible clients exist on all platforms, for instance under KDE (cf. Figure 4) or on Windows XP.

A simple file manager is also available as a Web application in the phpGroupware collaborative work environment of the project, for cases where the nomad users won't have the possibility to access the repository with DAV clients.

A specific sub-section of each project's documents repository, corresponding to the **public\_html** subdirectory, may contain the public Web site specific to the project (cf. 5.6). This sub-tree can then host a set of public pages very easily maintained, in addition to the stereotyped informations displayed in the project's form on the portal. Figure 3 shows the contents via DAV of this sub-directory, which is displayed, in this case, as the contents of a folder, a list of documents, instead of as a set of Web pages (pages would be displayed as in Figure 6).

---

<sup>9</sup>the counterpart of this simplification, compared to CVS in particular, is that the repository doesn't support documents versioning on its own. This could be enhanced in a future version of ProGET by using *Subversion* (<http://subversion.tigris.org/>) instead of the `mod_dav` module used today, which would bring the platform closer to what exists in the recent Trac software development platform (<http://projects.edgewall.com/trac/>), which combines a Wiki and Subversion.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

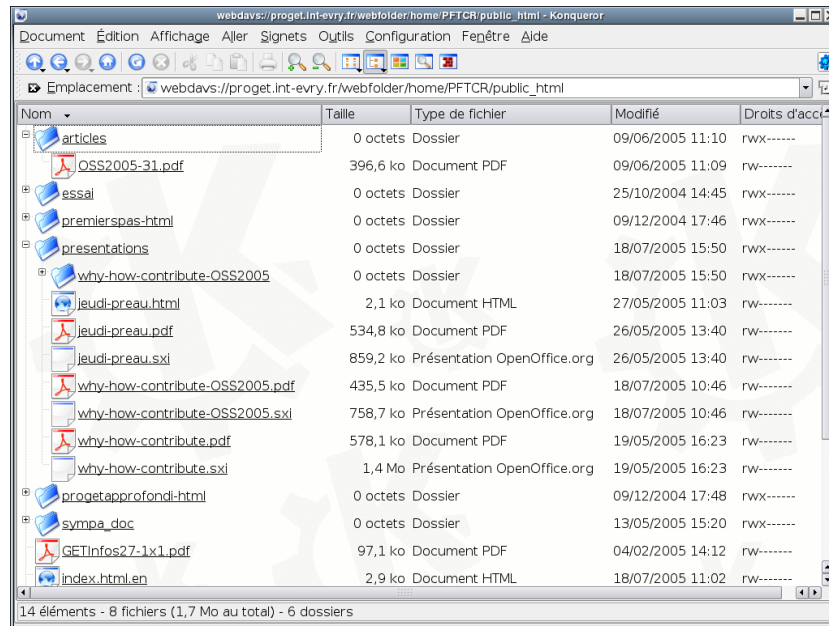


Figure 4 Contents of a project's DAV folder in KDE

## 5.5 Project's Wiki

Each project is provided, internally, with a *Wiki*<sup>10</sup> which can be used to share, in an authenticated way, in the form of Web pages, a set of simple hypertext documents, not necessarily very structured up-front.

This Wiki can serve many uses. Each project may for example use it as a hypertext repository to setup a knowledge base, which will hold all elements of the project's life, its productions, ongoing works. The wiki enables the progressive elaboration of informations, in a way less structured and more opened than by using the project's documents repository.

Of course, the TWiki wiki engine provides all classical features of these tools (most recent modifications list, notifications, search tool, back-links...) allowing users to better apprehend the

---

<sup>10</sup>the authors assume that the reader is already familiar with this kind of tool which have become very popular in the latest years. For more details, refer to [wiki] and to the Wikipedia project which hosts it, for a good example of application of Wikis.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

mass of hypertextual informations such a tool can host. Figure 5 shows an example of a list of most recent changes in a project's wiki.

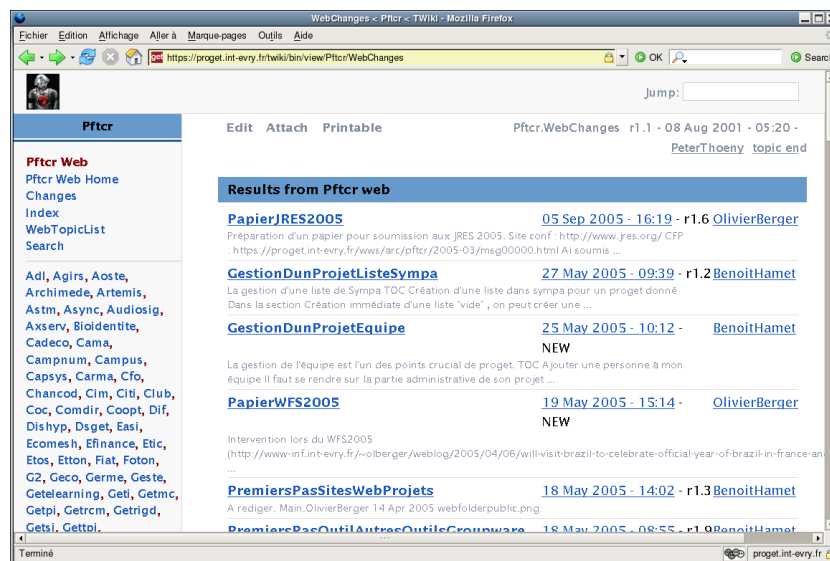


Figure 5 Most recent changes in a project's wiki

A portion of the Wiki can be used to put online hypertextual contents in the project's public Web site, in an automated way (cf. 5.6). To this end, the Wiki pages must have been identified individually as *public*. This feature may be used for instance to pre-publish on the Web some of the informations being elaborated inside the project, which have not yet achieved a sufficiently finalized or structured form required for publication of the documents available on the project's the web site. This is a way to offer some means of transparency on the research works, by allowing the public to directly view the “lab notepads” of the projects.

## 5.6 Project dedicated Web sites

On the same platform, each research project benefits of several dedicated publication spaces, for its external communication on the Web.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

### ***Stereotyped page in the public portal***

This page describes briefly the research activity and the team. It represents a stereotyped description of a project which has been validated by the Research Office. In addition, the page contains a dedicated news zone for the project which is updated directly by each project manager. This zone helps spreading news about the life of the project. To this end, an RSS<sup>11</sup> feed comes alongside each project's news pages.

Project stereotyped pages are created in the research portal through a phpGroupware module named **sitemgr**. This module is configured specifically and relies on informations extracted from the administrative database of project forms.

### ***Specific Web space***

This Web site is made of static HTML pages. The contents of this space is managed directly by each project, at will. It may be produced with the HTML editing software of its choice. The upload is made directly by members of the project via publications in the `public_html`<sup>12</sup> sub-directory of the project's WebDAV folder (cf. 5.4).

No specific competence is required in the teams to be able to contribute to the project's web site. Figure 6 shows an example of such a Web page. The URL to access this Web space is stereotyped from the project's acronym : `http://proget.int-evry.fr/projects/PROJECT_NAME/`

---

<sup>11</sup>Really Simple Syndication : <http://blogs.law.harvard.edu/tech/rssVersionHistory>

<sup>12</sup>the name "public\_html" is a convention adopted after web server Apache's naming policy.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

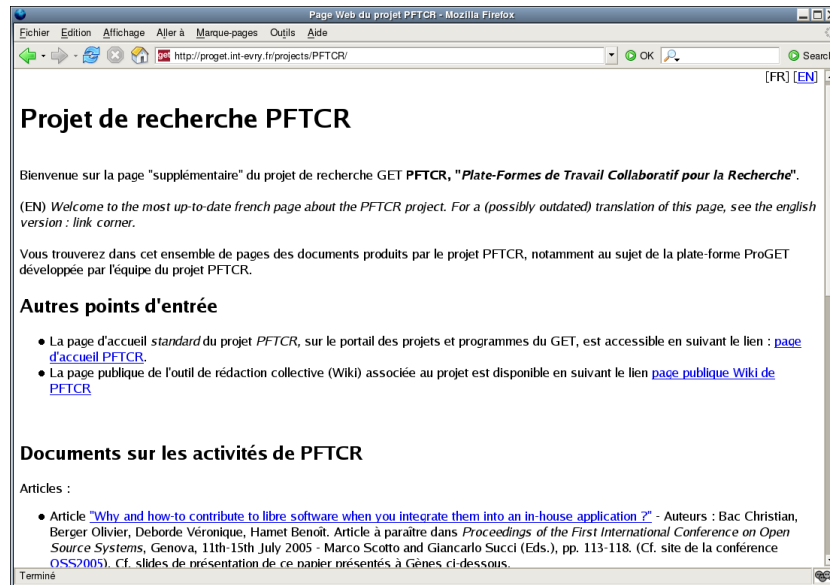


Figure 6 Dedicated project's website for project PFTCR

### Wiki excerpts publication space

This space complements the previous spaces by allowing the direct publication of excerpts (pages declared *public*) of the resources elaborated by the project inside its internal Wiki (cf. 5.5). This may help with pre-publishing, issuing short reference papers, providing a knowledge base, etc.

## 5.7 GET research projects administrative management module

Whereas other modules of the platform existed previously as generic libre software tools fulfilling the classical needs of every organization, this project administrative management module has been developed specifically for the needs of the GET Research Office.

This module implements the workflow of the contents of research projects forms, during the life-cycle of GET research projects (creation, approval, yearly updates...).

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

It was developed in PHP on top of the phpGroupware API, by using the **etemplates** component, which allows the creation of input masks and screens for the update of the data stored in the underlying MySQL database.

It also uses a set of query and report templates built with *Agata Reports*, a tool for creation of database queries without programming skills, in Web mode. An expert of the domain, who knows the database schema can then improve the set of queries initially integrated in ProGET, for instance to produce complex reports in the form of spreadsheet files or even synthetic booklets in the form of word-processor or PDF files.

## 5.8 Data partition

Inside the same collaborative work platform, each research project must be able to host, its own informations and documents, which may be confidential. Thus, a partition of the whole data handled by ProGET was setup, relative to the project groups. This applies in particular to the documents stored in the document repositories (including its dedicated website), the messages on the mailing-lists in Sympa, the informations from the projects description forms in restricted access, and some of the informations in the phpGroupware groupware tools.

In a classical way, the OpenLDAP directory is used to allow the authentication of the user's work sessions with a unique password for the different tools integrated in the same platform. The directory is also used as a reference for the description of project teams and their access authorization for the data handled by the different applications. The system relies on persons and groups defined in a single place, in the LDAP directory. These informations are managed through the users and groups management modules in phpGroupware.

A list of persons responsible for the project and a list of participants are associated to each research project. Members of the first group have the right to modify the list of members of the second group via the Web interface in phpGroupware, and to add new co-opted members in their

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

group of responsible persons. The management of the teams for the collaborative work is then done directly in the projects, in an autonomous way, for a better reactivity.

In each of the applications integrated in the platform, access to the data relating to the research projects is controlled by checking if one belongs to these groups. For instance, Sympa is configured to query OpenLDAP directly, which helps defining automatic subscription of members of a project group to a mailing-list. That way, each research project receives automatically a discussion list whenever it's created. The record of discussions on this list is only accessible to its subscribers, being the members of this research project.

For the phpGroupware groupware tools (calendar, address-book, news,...), the partition of the data is implemented on the basis of stereotyped phpGroupware *categories*. So each project group defined in the underlying OpenLDAP directory receives a corresponding phpGroupware category automatically, which is then used as an umbrella for classifying the informations.

In TWiki, this principle is implemented through standard mechanisms which are devised to host several autonomous wikis on the same site (known as *Webs* in Twiki). All the pages of a project are contained in a TWiki Web, i.e. a dedicated autonomous wiki. The creation of this space is done at the time a group of users is created, and corresponding privileges are granted to this group (allowed to modify pages, to add new members to the group, ...). Unlike the other applications integrated in ProGET, it was not possible to let TWiki (at least the standard version we used) interface directly with the LDAP directory for the definition of groups of persons. An adapter module has been developed, to synchronize the groups defined in TWiki with the contents of the research project teams.

## 5.9 Modification and customisation efforts

Most of our software development effort has been targeted on :

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

- **phpGroupware custom components** for administrative management of research projects (cf. 3.2) and Web portal (cf. 3.3). These developments were very specific to the needs of GET, and shouldn't be released.
- **phpGroupware low-level components.** We have in particular improved the interaction mechanisms between phpGroupware and OpenLDAP, and the code used to access WebDAV. These developments do not specifically fit only our own needs, and improve in a sensible way the standard phpGroupware platform.
- **Bug fixes** and modifications to the other components. Many small modifications and bug fixes were necessary in particular on the WebDAV server, and in the LDAP server interface of Apache, or on the *Twiki* and *Sympa* programs. These modifications make it easier to integrate these components on top of OpenLDAP. These developments were contributed to the community in the form of patches when they were of general interest.

## 6 Strategy for contribution

This section deals with some methodological aspects of the development project of the ProGET platform. Our project is more an inhouse project than an Open Source project and thus cases studies like [Mockus et al, 2002] does not always apply to it. Despite the fact that the project was not open to outsiders, it relies profoundly on libre software and we felt the need to set up a strategy for contributing to libre projects. The strategy was adopted in order to establish the conditions of a greater maintainability of the developed system and a clean integration of our patches and improvements to the original projects. These aspects have been described in more details by the authors in a previous paper, see [Berger et al, 2005].

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

## 6.1 Benefits of libre software for their integration in in-house applications

Libre software offer today an opportunity to lower the duration and cost of software projects. They also help to raise the quality of the applications [Bauer and Pizka, 2003].

The software components are numerous and available for integration or adaptation in order to create applications that fit the needs of organizations [Wheeler 2005].

The development of the initial version of an in-house application can be done quickly, even in a “brutal” way, by modifying components in order to combine them.

To be able to use them for its internal projects, one may not be forced to participate directly, or in an indirect way, to their development in the libre communities<sup>13</sup>.

These tasks of integration can then be done in a classical way, as if the components had been completely developed internally, without trying to contribute to the initial “upstream” projects. But it's also possible, and even advisable, to contribute, whenever it's possible.

## 6.2 Maintainability of in-house specific developments

After the initial integration phase, comes the issue of the mid-term evolution of the obtained product. This issue becomes crucial when it becomes necessary to react to recurring external evolutions of the integrated components (for instance in case of a published fix for a security problem).

---

<sup>13</sup>Note that contrary to some rumors, it is in no way mandatory to publish the modifications that you've done on a libre software program, as long as this software is not distributed to third parties, even if it was initially published under the GNU GPL.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

We profess, out of the experience gained during this project, that it is necessary to define a policy for contributions on the most important integrated components.

This active contribution brings the possibility of transferring the evolutions that were required by the internal developments into the libre modules that were used. Once they have been adopted by the external projects, they will become part of all future releases of these components, which will remove the need to keep on applying internally some specific patches each time a new update is issued for the original module. In this respect, some amount of the maintenance effort for the internal solution is “outsourced” to the libre community, which helps lower the internal costs.

### **6.3 Strategy adopted**

So, to make some adaptation on phpGroupware components, we have not limited ourselves to only modifying internally this software (as it had been done before for PicoLibre), but on the contrary, we have decided to collaborate actively with the phpGroupware project. This led us to undertaking the process for the official application of a member of our team to the phpGroupware project, and the transfer of the copyright on our developments to the FSF which holds the rights for phpGroupware.

Again, our goal, by doing so, is to have our modifications integrated in the standard code base of phpGroupware, so that they will be part of future releases of the software, and we may not have to apply again and again the same changes whenever a new version of phpGroupware is released. This way, we intend to lower the cost for the future maintenance of ProGET.

## **7 First outcomes and prospects**

Even if integration of existing applications requires extensive adaptation [Adams et al. 2005], the effort is far lower than what would have been necessary for a project entirely developed

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

internally. The vast range of features that we thought necessary for our platform was only achievable at a reasonable cost by taking advantage of already existing libre software.

This platform brings generic tools for non-programmer researchers but also features appealing to the most demanding researchers, like DAV repositories or a Wiki.

The ProGET platform is in production now and opened for its users since June 2005. The GET research portal and the projects administrative management application are already used on a day to day basis. It is however too early to proceed to an assessment of the use of collaborative work tools, which are still used regularly by only a few research projects.

The use of the collaborative work environment is at the present time proposed for volunteers. Its roll-out to the whole GET research teams is not yet envisioned, some uncertainties remaining on the general infrastructure to be setup, for instance for the necessary training phase of the users, or the financial costs to ensure the maintenance of the production platform with maximum availability.

Some findings can still be learned from the return on experience of the first users.

## **7.1 Conformance to requirements**

The tools which have been integrated offer a far range of features, which address most of the generic needs of the research teams at GET. The TWiki tool, quite versatile, provides in particular a great potential, although it requires some amount of training.

Some shortcomings have however appeared during the first months of usage of the platform.

ProGET was devised by focusing on an organization structured around the notion of research project with a clearly defined perimeter, validated by the GET Research Office. However, needs for collaborative work tools also come from contexts outside the precise boundaries of already

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

validated research projects. Thus, during upstream exploratory phases leading to the creation of projects, for instance, some collaboration is often necessary between members of teams which have already been established independently, and who often work part-time on these exploratory activities. Several users have requested that the kind of tools available in ProGET be offered also for these upstream phases, where they could greatly facilitate participation, reactivity and communication between various actors. It then seems necessary to be able to handle also collaborative work groups linked to exploratory phases, outside the rigid frame of validated research projects already approved for the GET research portal.

It was also identified, through the first steps of take-up of the tools by research teams, that the principle of partition of data that was implemented, based on a lowest granularity corresponding to the research projects, was not fitting that well. Inside a GET research project, teams have to work for different research contracts, which all have their own list of partners, and their own privacy requirements. The granularity chosen for teams which have to access the same set of documents and informations in the project is thus not well adapted in the current version of the platform. It seems necessary to be able to define, for collaborative work, some sub-projects more or less autonomous in the realm of the same research project, which will each define a specific list of participants and associated access privileges.

## **7.2 Evolution towards other collaborative activities**

The users of the platform are also participating, besides their research activities, in teaching activities, and numerous other work groups, for which Web-based collaborative work tools may not always be available, either at GET or at its partners'.

There is thus a need for the availability of one or more platforms supporting collaborative work, which would be as feature-full as the ProGET platform, but which may allow the easy and quick

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

creation of projects and their work groups, progressively, in a loose environment, as was the case with the PicoLibre platforms.

Need for such a platform could lead to the creation of a “new generation” PicoLibre system, which could integrate the whole set of generic components that have been integrated in ProGET, together with a generic module for creation of collaborative work spaces, and which could be more maintainable than the previous generation.

The ProGET platform could then be split in two parts : on the one hand a module for the administrative management of GET research projects and a GET research Web portal, and on the other hand the use of a “PicoLibre V2” kind of platform, in which the groups corresponding to the GET research projects could be hosted, among others.

### **7.3 Constraints for administration**

The ProGET platform was devised in an effort for reuse and modularization in integrating the existing libre components, in order to lower the maintenance efforts to be made internally (cf. 6).

However, the installation and maintenance tasks for such a platform are quite demanding and require specific competence (good knowledge of the internals of the various integrated software), which imply a necessary precise planing in the event of the outsourcing of its administration to the regular IT department which is in charge of the GET information systems.

In this respect, one first hypothesis for improvement would be the availability of the whole set of components in a standard packaging. It would be interesting to have all the components used, and their new adapters, available in the form of packages for the libre software distributions already known by the support teams. For instance, in the Debian distribution, the phpGroupware, Sympa and TWiki tools are already packaged in the ordinary way. It would be interesting if all modules that have been developed for ProGET were packaged too. This could be achieved by

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

taking advantage of the recent efforts made to package the modules of the PicoLibre application for Debian.

## 7.4 PhpGroupware for the development of custom applications

Several findings can be stated, looking at the use of phpGroupware as the basis for the development of custom applications.

PhpGroupware brings a large choice of components which have not all achieved a common level of quality. Some functions suffered from numerous bugs preventing them to be put into production as such, in particular regarding the conditions required for ProGET, for instance if using underlying OpenLDAP and WebDAV layers.

In addition, the structure of data in phpGroupware is in general organised from the perspective of one phpGroupware instance deployed in an organisation where, by default, the whole of the informations handled by the system will be available to all the users. This may fit, for instance, to SMEs or very small virtual organisations where all actors have to be informed of everything happening in the organisation. This model happened however to be poorly convenient in our case which requires that projects hosted on the same phpGroupware platform be kept separate, and that a strong level of confidentiality be preserved on the informations (cf. 5.8).

The use of phpGroupware categories partially helps to solve this kind of constraints, but by paying the price of some usage precautions which degrade the global usability of the application.

PhpGroupware provides an interesting API for the PHP developer, for Web based collaborative work applications that will be relatively simple. It happens to be somehow limited though, for advanced needs for the development of complex applications, in particular concerning the ACLs, the corresponding data model and the independence of the lowest level modules. PhpGroupware lacks also of a module providing applicative workflow facilities.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

The use of the *etemplate* “RAD tool” integrated in phpGroupware turns out a poor choice in the case of a complex application, despite its ease of use for relatively simple needs. Indeed, its associate documentation and tutorial only scarcely refer to the *MVC* model which is however generalized in the more complex modules of phpGroupware. In the latest modules developed for ProGET, we thus have preferred to use these more classic *MVC* patterns and rely more on the HTML templating system, that happen to be more maintainable. This choice was made with a mid-term vision, as the learning curve for developers is bigger in this case.

PhpGroupware is still an interesting project for a rather simple use of its existing modules. And although it provides rich and quite generic APIs, which allow the development of new applications, it cannot constitute a really generic application development *framework* like other libre environments (Zope, Apache Tomcat...).

## 8 Conclusion

Although it may be enhanced on numerous aspects, the ProGET platform delivers a range of features without competition in existing libre software collaborative work platforms, for a modest development cost.

There are strong needs in higher-grade research and teaching institutions for tools supporting collaborative work. ProGET can then constitute a reference point for organizations wishing to integrate, for their custom needs, existing libre software applications for collaborative work.

Even if a policy is adopted for the contribution to the libre software projects used, in order to lower certain maintenance costs, the generalization of the use of the present platform at GET and the future developments necessary to enhance it will only be possible through a substantial investment, which may be far more important than what was spent for the first initial developments.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article “Integration of libre software applications to create a collaborative work platform for researchers at GET”, submitted for publication in *International Journal of Information Technology and Web Engineering* (IJITWE), vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

One possibility for the reduction of these costs could be the mutualization of the development between several organizations. It could be articulated around the industrialization as well-packaged libre software programs of the most generic elements of ProGET, leading the way to a new generation PicoLibre platform.

## Bibliography

- [Adams et al 2005] Adams, Paul and Boldyreff, Cornelia and Nutter, David and Rank, Stephen (2005) *Adaptive Reuse of Libre Software Systems for Supporting On-Line Collaboration*. In: 5th Workshop on Open Source Software Engineering, 17 May 2005, ICSE '05, St. Louis, MO, USA.
- [Bauer and Pizka, 2003] Andreas Bauer, Markus Pizka, *The Contribution of Free Software to Software Evolution*, Proceedings of the International Workshop on Principles of Software Evolution (IWPSE) (2003)
- [Berger et al, 2005] C. Bac, O. Berger, V. Deborde, B. Hamet, *Why and how to contribute to libre software when you integrate them into an in-house application ?*, Proceedings of the First International Conference on Open Source Systems Genova, 11th-15th July 2005
- [Cousin et al. 2002] Cousin, E., Ouvradou, G., Pucci, P. and Tardieu, S., *PicoLibre: a free collaborative platform to improve students' skills in software engineering*, 2002 IEEE International Conference on Systems, Man and Cybernetics
- [Dridi and Neuman, 1999] Fredj Dridi, Gustaf Neumann, *How to implement Web-based Groupware Systems based on WebDAV*, in Proc. of WETICE'99, IEEE 8th Intl. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, Stanford, USA, June 1999.

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.

- [Mockus et al, 2002] Audris Mockus, Roy T. Fielding, James D. Herbsleb: Two case studies of open source software development: Apache and Mozilla. *ACM Trans. Softw. Eng. Methodol.* 11(3): 309-346 (2002)
- [OpenLDAP] Website of the OpenLDAP project: <http://www.openldap.org/>
- [phpgw] Website for the phpGroupWare project: <http://www.phpgroupware.org/>
- [picolibre] Website for project PicoLibre: <http://www.picolibre.org/>
- [sourceforge] SourceForge project hosting platform: <http://sourceforge.net/>
- [sympa] Sympa project's website: <http://www.sympa.org/>
- [twiki] TWiki project's website: <http://twiki.org/>
- [WebDAV] Y. Goland, E. Whitehead, A. Faizi, D. Jensen, *HTTP Extensions for Distributed Authoring WEBDAV (1999) – RFC 2518* - See also : <http://webdav.org/>
- [Wheeler, 2005] David A. Wheeler, *Why Open Source Software / Free Software (OSS/FS, FLOSS, or FOSS)? Look at the Numbers!* - [http://www.dwheeler.com/oss\\_fs\\_why.html](http://www.dwheeler.com/oss_fs_why.html)
- [wiki] "Wiki" article in the Wikipedia encyclopedia: <http://fr.wikipedia.org/wiki/Wiki>

This document, (© by its authors + GET/INT 2006) is a **preprint** version of the article "Integration of libre software applications to create a collaborative work platform for researchers at GET", submitted for publication in *International Journal of Information Technology and Web Engineering (IJITWE)*, vol 1(3) in summer 2006. This version published under the Creative Commons Attribution-NonCommercial-NoDerivs License.