GATE Teamware User Guide

GATE Team

January 12, 2010

Contents

Chapter 1

Introduction

GATE Teamware is a software suite and a methodology for the implementation and support of *annotation factories*: it is intended to provide a framework for commercial annotation services, supplied either as in-house units or as outsourced specialist activities. This framework is a novel development in several aspects:

- it structures intervention by different actors (human and machine) into clearly-defined roles and provides the means to manage them in a unified fashion,
- it complements GATE's developer-oriented facilities with User Interfaces (UIs) oriented on other necessary staff roles,
- it is methodological instead of purely technological.

1.1 Roles

There are a number of staff roles required for an effective annotation factory including:

- **annotators**, who are largely unskilled, may be geographically distributed, and whose work is quality controlled by curators via metrics-related mechanisms,
- **curators**, who may be skilled annotators, capable of making decisions and valuing annotator's work,
- **language engineers**, these staff work on resources producing automatic annotations. These resources are combined to produce a service.
- **managers**, who are supposed to create and manage annotation projects
- **administrators**, who administrate the workflow system and its users.
- **superadministrator**, single super user who can modify the critical parts of the application along with other administrator duties

GATE Teamware, together with GATE Developer, provides tools for all these roles, and the workflow by which they may be combined with automatic Information Extraction (IE) systems to provide cost-effective annotation services.

1.2 Prerequisites

GATE Teamware is a Java web application deployable under Apache Tomcat server (5.5.x and 6.x) and fully compatible with both JDK 1.5 and 1.6. Apart from a web browser (the currently supported versions are

Internet Explorer (versions: 6, 7 and 8), Firefox 1.5, 2.x, 3.x) and Opera (versions: 8, 9 and 10)) you need to have Java Web Start (JWS) version 1.6.0 installed on client machines. JWS can be downloaded as the part of Java Runtime Environment (JRE) version 6 from the following address: `http://java.sun.com/javase/downloads/index.jsp`

1.3 Login

GATE Teamware authentication and authorisation is accomplished twofolds:

- through Database
- through Active Directory (AD), so that all users who have AD account can log in to the GATE Teamware with the same credentials. The scenario supported at the moment is:

By default the authentication and authorisation are performed against database. New user gets registered:

- either by following *Signup* link at the login page (see Figure 1.1)
- or by administrator, who also assigns role(s) to the new user.

With AD situation is a bit more complicated. The following scenario applies:

- when the AD user signs in for the first time, s/he will be added to Teamware database (DB) with the appropriate role; currently, these roles are *manager*, *curator* and *annotator*; this can be changed in administration section later.
- the user profile in the GATE Teamware DB is automatically populated by matching data from AD account. All other data have the default values, which can be changed later.

When you type teamware url like `http://[host.name]:[host.port]/[instance.name]/executive` in your browser, you will be prompt with the login page, see Figure 1.1.

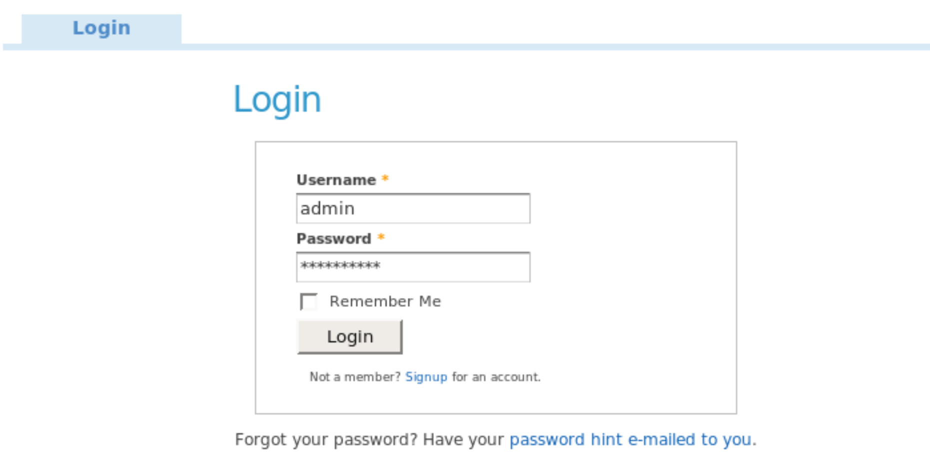


Figure 1.1: Login Form

Login with your username and password.

1.4 Home Page

After you have logged in successfully, you will be redirected to your *home page*. The menu items differ based on the type of the user role. If logged in as a *manager*, you should see the menu and shortcut links as shown in Figure 4.1. For users with annotator roles, the number of available options will be reduced (e.g., the menus *Resources* and *Projects* will not be visible).

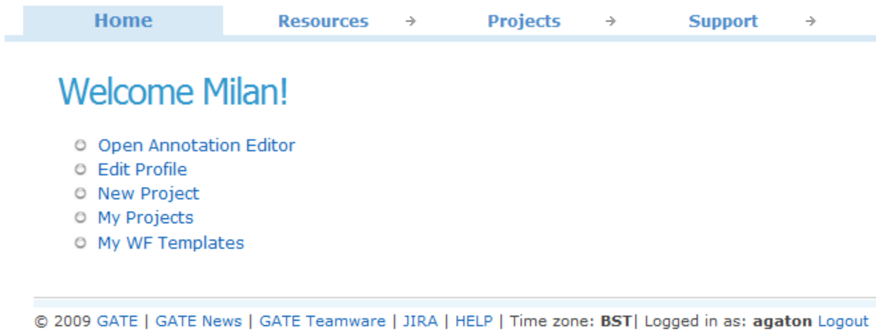


Figure 1.2: Home Page - Example

1.5 Edit Profile

You can access/change your profile by clicking on *Edit Profile* link from the home page (see Figure 1.2). The *Edit Profile* form as shown in Figure 1.3.

Note that, in case of AD authentication, only those fields which exist in your AD profile will be copied to your GATE Teamware profile (see Table 1.1). Other field values will be provisional and you should change them after the first login.

Table 1.1: The overview of matched fields in AD and Teamware DB

Active Directory attribute name	**Teamware DB Field Name**
givenName	*first_name*
sn	*last_name*
sAMAccountName	*username*
msSFU30Password	*password*
mail	*email*
telephoneNumber	*phone_number*

1.6 Logout

When you finish your session, click on *logout* link in the bottom-right corner.

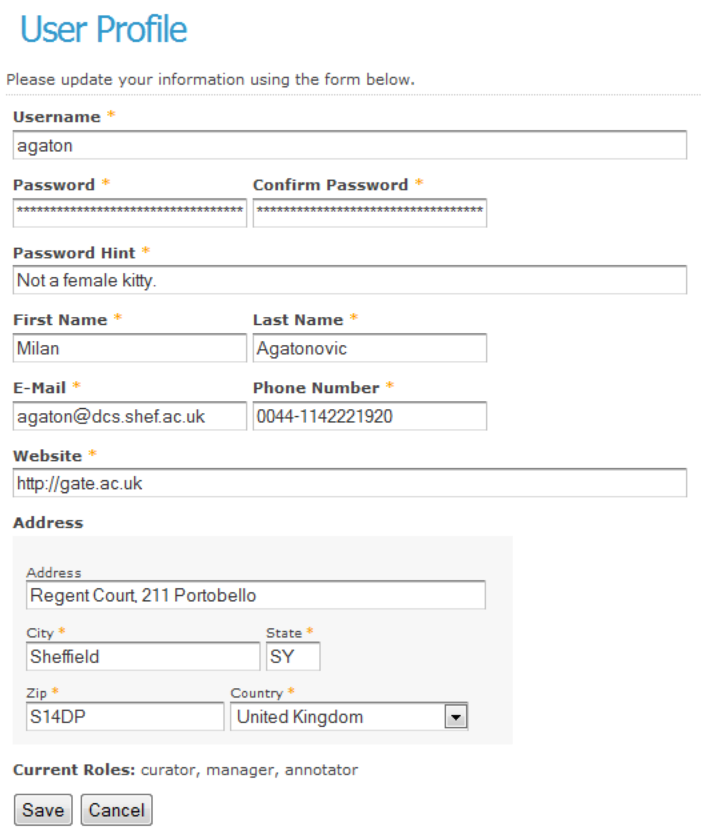


Figure 1.3: Edit Profile

Chapter 2

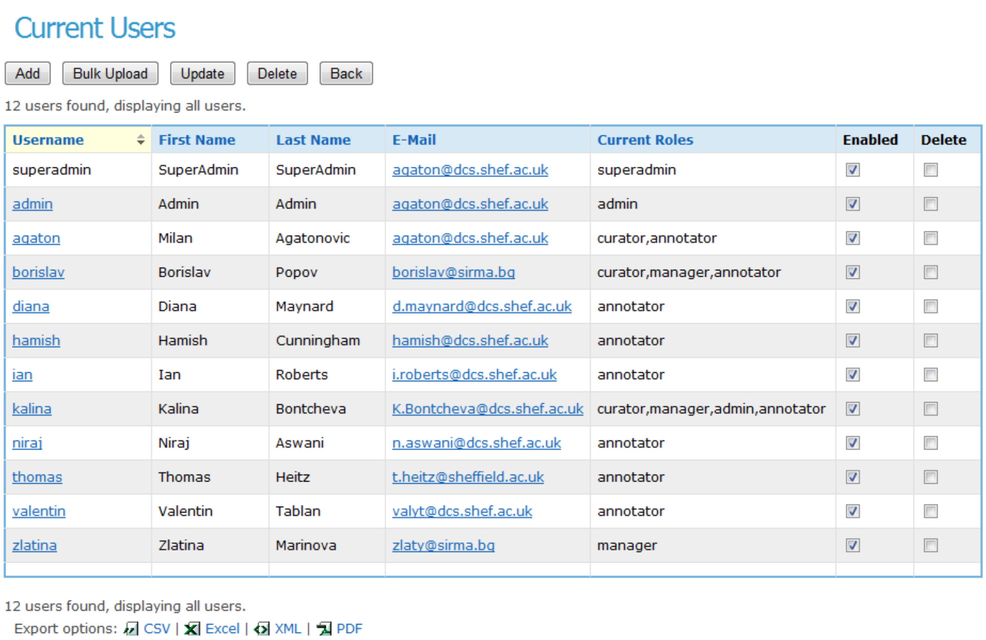
User Management

2.1 Manage Users

This chapter provides details about user and role management which you can do if you have *administrator* role. *Manage Users* option from *Admin* menu will take you to the user list page (see Figure 2.1) from where you can: add new user(s) (one by one or bulk), view, edit, enable/disable and delete users.

2.1.1 Add Users

To add a new user click on the *Add* button[1].



Current Users

Add | Bulk Upload | Update | Delete | Back

12 users found, displaying all users.

Username	First Name	Last Name	E-Mail	Current Roles	Enabled	Delete
superadmin	SuperAdmin	SuperAdmin	agaton@dcs.shef.ac.uk	superadmin	☑	☐
admin	Admin	Admin	agaton@dcs.shef.ac.uk	admin	☑	☐
agaton	Milan	Agatonovic	agaton@dcs.shef.ac.uk	curator,annotator	☑	☐
borislav	Borislav	Popov	borislav@sirma.bg	curator,manager,annotator	☑	☐
diana	Diana	Maynard	d.maynard@dcs.shef.ac.uk	annotator	☑	☐
hamish	Hamish	Cunningham	hamish@dcs.shef.ac.uk	annotator	☑	☐
ian	Ian	Roberts	i.roberts@dcs.shef.ac.uk	annotator	☑	☐
kalina	Kalina	Bontcheva	K.Bontcheva@dcs.shef.ac.uk	curator,manager,admin,annotator	☑	☐
niraj	Niraj	Aswani	n.aswani@dcs.shef.ac.uk	annotator	☑	☐
thomas	Thomas	Heitz	t.heitz@sheffield.ac.uk	annotator	☑	☐
valentin	Valentin	Tablan	valyt@dcs.shef.ac.uk	annotator	☑	☐
zlatina	Zlatina	Marinova	zlaty@sirma.bg	manager	☑	☐

12 users found, displaying all users.
Export options: CSV | Excel | XML | PDF

Figure 2.1: User List

[1] Please note that newly added users need not to have AD account before hand

2.1.2 Bulk Users Upload

To upload multiple users click on the *Bulk Upload* button. You can download the excel template file, by

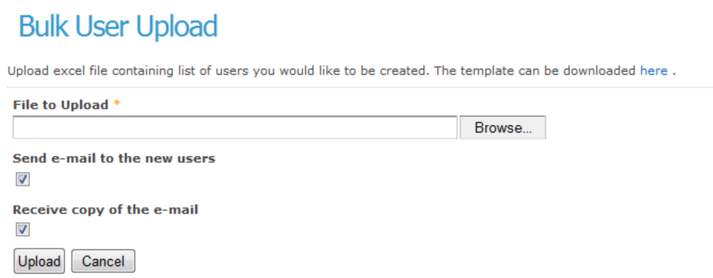
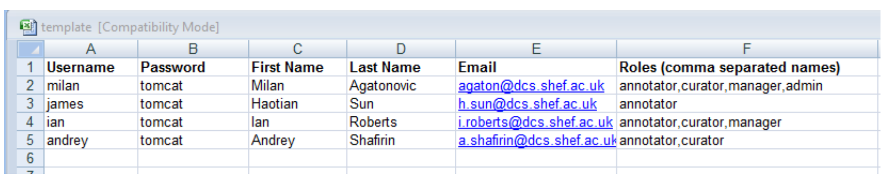


Figure 2.2: Bulk Users Upload

clicking on provided link and fill it with the user data like shown in Figure 2.3 If you want users to be notified upon their registration, check *Send e-mail to the new users*, and if you want yourself to have copy of this email check *Receive copy of the e-mail*.



template [Compatibility Mode]

	A	B	C	D	E	F
1	Username	Password	First Name	Last Name	Email	Roles (comma separated names)
2	milan	tomcat	Milan	Agatonovic	agaton@dcs.shef.ac.uk	annotator,curator,manager,admin
3	james	tomcat	Haotian	Sun	h.sun@dcs.shef.ac.uk	annotator
4	ian	tomcat	Ian	Roberts	i.roberts@dcs.shef.ac.uk	annotator,curator,manager
5	andrey	tomcat	Andrey	Shafirin	a.shafirin@dcs.shef.ac.uk	annotator,curator
6						

Figure 2.3: Excel Template File

Please note that roles assigned to the new user have to comma separated

2.1.3 View/Edit user profiles and roles

If you want to update the information of the user *agaton*, he/she can simply click on the username link of *agaton* shown in Figure 2.1.

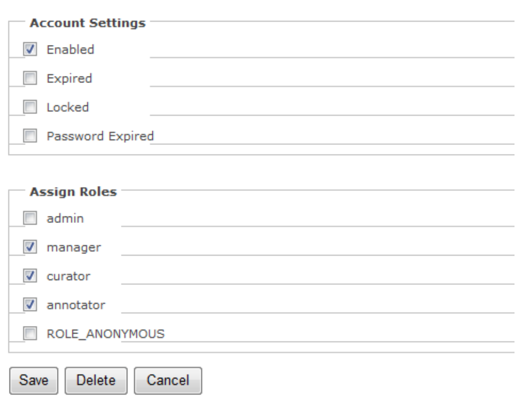


Figure 2.4: Edit User Profile

The user information form will be displayed. In addition to the information shown in Figure 1.3 previously, this form will be extended to include details of user roles and account settings as shown in Figure 2.4. Here administrator can assign multiple roles to each user.

2.1.4 Enable/Disable Users

You can quickly enable/disable multiple users by ticking/unticking the checkboxes in the user list *Enabled*, and clicking on the *Update* button.

2.1.5 Delete Users

You can delete multiple users by ticking the checkboxes in the user list *Delete* column, and clicking on the *Delete* button.

2.2 Active users

Who among users is online can be seen from the *Active Users* page as shown in Figure 2.5.

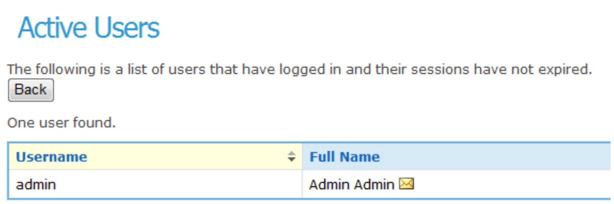


Figure 2.5: Active Users

2.3 Clickstreams

Administrator can view user request history via *Clickstreams* page, which is the last item in the Administration menu, as shown in Figure 2.6.

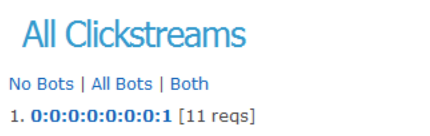


Figure 2.6: Clickstreams

Chapter 3

Resource Management

In this chapter, we explain how you can manage resources like documents (corpora), annotation schemas and annotation services if you have a *manager* role.

3.1 Corpora

3.1.1 Corpora List

To view existing corpora, choose *Resources >> Documents* from the top menu. New page will be shown with the complete list of available corpora as shown in Figure 3.1. From this page, you can choose several

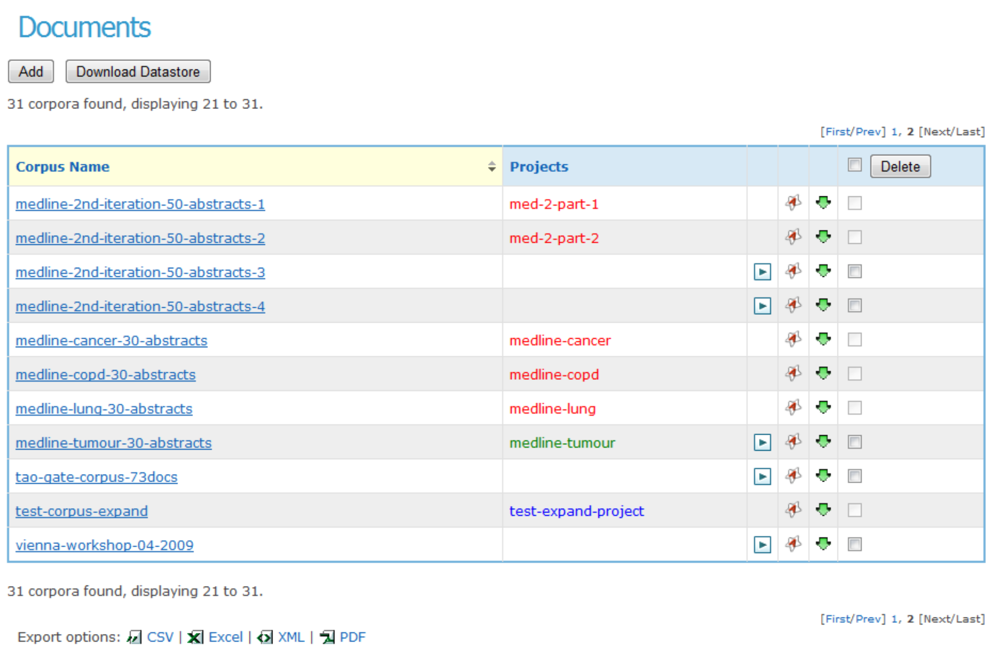


Figure 3.1: Copora List

options:

Modify Corpus To view details about already existing corpus, click on the link showing the corpus name, or the first icon in the row. If you want to add some new documents in corpus, you can do this by clicking on *Upload Documents* button.

Start Annotation Project Directly from the corpora list, you can start annotation project for each corpus, by clicking on the *start* icon. The annotation projects will be explained in details in the next chapter. For now, have this option in mind as a convenient way to start annotation from this page.

View Annotation Projects You can see if some annotation project is running, suspended or finished over each corpus. The links to the annotation project statistics can be of different colours, depending on project status: *red*: suspended; *blue*: running; *green*: completed. More about project statistics you can find in the next chapter.

Download Corpus If you click on *Download* link (second icon from the right) you will download the corpus with all documents in it, archived in zip file.

Download Datastore If you click on *Download Datastore* button you will download whole datastore with all corpora, archived as zip file.

Delete Corpus If you tick the checkbox in *Delete* column (last column on the right hand side) and press *Delete* button you will delete selected corpora with all documents in them.

3.1.2 Create Corpus

To create a new corpus, click on the *Add* button on the corpus list page. New window will open (see Figure 3.2). Specify the corpus name and encoding. By default the encoding is UTF-8 which should be fine for most of the documents. However if you need some particular encoding, you should specify it here. Click *Browse* button to provide the full path of your zip archive from your file system.

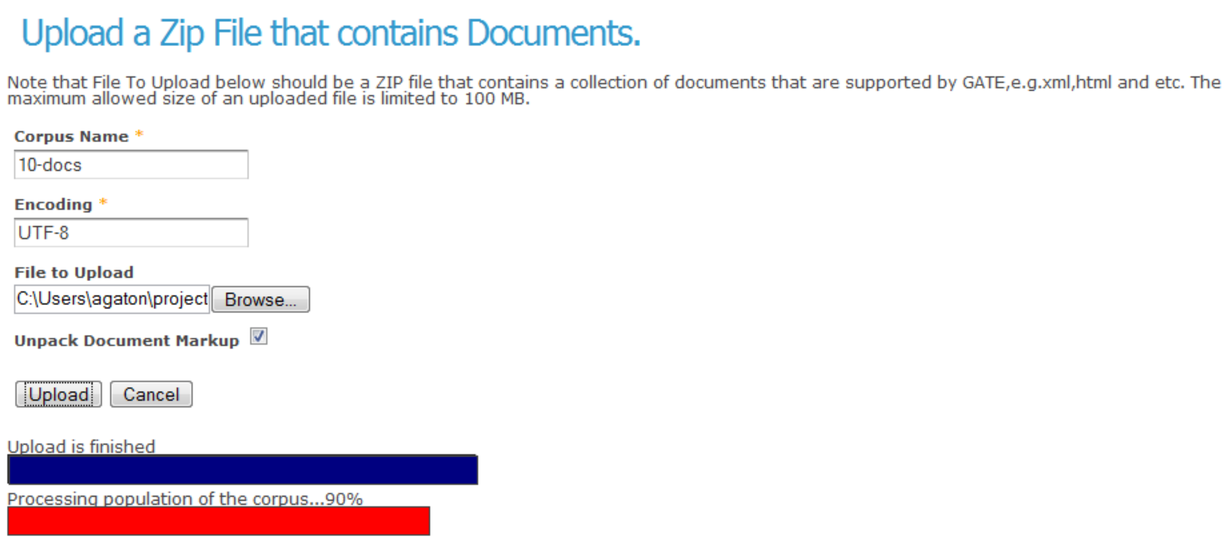


Figure 3.2: Add Corpus

GATE Teamware will try to identify the type of the document, then strip and convert any markup into GATE's annotation format. To disable this process, uncheck the option *Unpack Markup Aware*. Finally click on the *Upload* button to populate the corpus.

Please have in mind the following conditions:

- documents need to be packaged in zip archive

If the corpus is populated successfully, you should see a screen similar to the one shown in Figure 3.3.

Figure 3.3: Corpus created

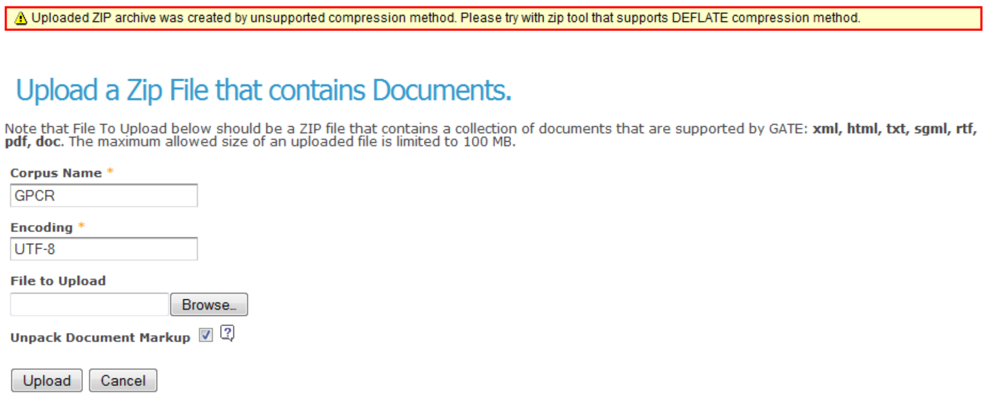


Figure 3.4: Unsupported Compression Method

- documents have to be in a format supported by GATE: Plain Text, HTML, SGML, XML, RTF, PDF, DOC or XLS

Please make sure that the zip archive you are uploading is created by standard zip archiver tool and that compression method used is DEFLATE[1]. There should be no problem with standard zip implementation under Windows and Unix. In case you get the error message like in Figure 3.4, try to zip your corpus with the different tool.

3.1.3 ANNIC Search

To use ANNIC Search facility, click on the ANNIC icon for the corpus you want to carry out the search. A Java Web Start application - ANNIC GUI will be activated.

The Figure 3.5 shows ANNIC in action. Since ANNIC allows indexing documents with annotations and features, the syntax of ANNIC query contains LHS part of the JAPE pattern/action rule.

For example, the query: $Reference.type == Claim(Token) + 10Measurement$ will find first an annotation of type *Reference* that has a feature named type with the value *Claim*, then to allow from 1 to 10 annotations of type *Token*, i.e. a word or punctuation, and eventually to find an annotation of type *Measurement*. This query can be interpreted as a search for finding a reference to a claim that involve measurement.

To execute a query, enter it in the query text field, and click on the magnifying glass icon. You can choose annotation sets, then the annotation type, and features to add to the middle part of the GUI with the bottom arrow icon. If there is more than one page of results you need to use the right arrow icon to see the next page of results.

[1] `http://en.wikipedia.org/wiki/DEFLATE`

You can see the results in the bottom part of the GUI that lists in a table the matching patterns, their contexts, the name of the document and annotation set. In the middle part, you can see the annotations and features with their values for the line selected in the bottom part.

Help window is provided in the GUI if you click on the information sign icon in the top tool bar. For more information about Annic and JAPE, please check Gate User Guide.

3.2 Documents

3.2.1 Document list

To see the list of documents inside the corpus click on *View* link next to the corpus name. New page will open as shown in Figure 3.6.

To delete a document from the corpus, simply tick the checkbox next to document and click on the *Delete button*. You can choose to delete all listed documents by ticking the box from the top of the table.

3.2.2 View Document with Annotation Editor

To open/edit a specific document, click on the *View Document* icon (the first icon from the left hand side shown in Figure 3.6). Annotation Editor will be activated and start loading the document. As shown in Figure 3.7, users can now carry out the annotation tasks within the Annotator GUI.

3.2.3 Compare Annotations with Annotation Differ

To use the Annotation Differ facility on the document, click on the icon *Annotation Differ* on the documents list page (the second icon from the left hand side shown in Figure 3.6). Then, you should see the Annotation

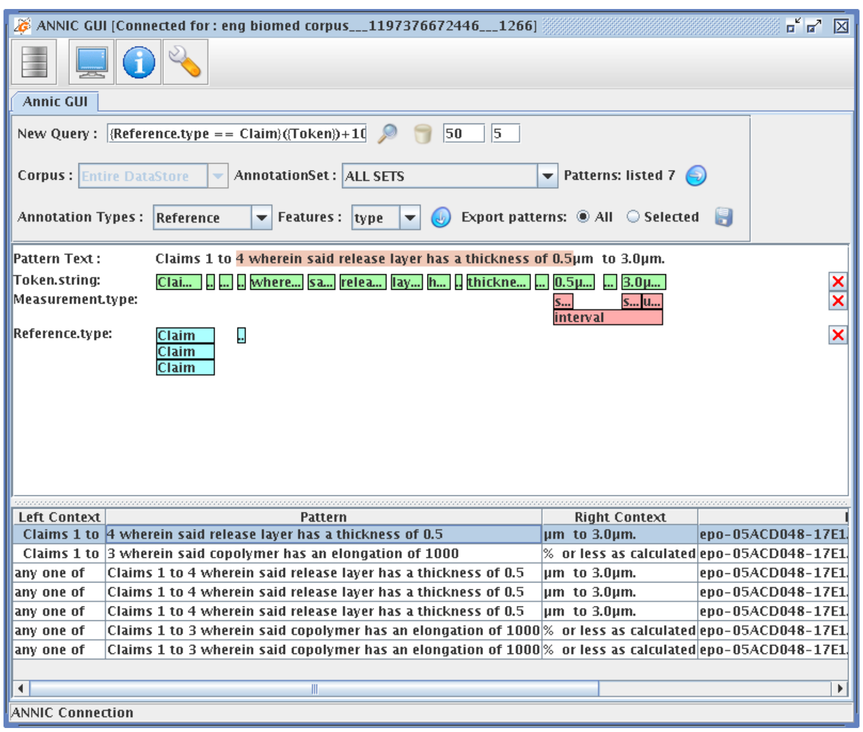


Figure 3.5: Annic Search Result

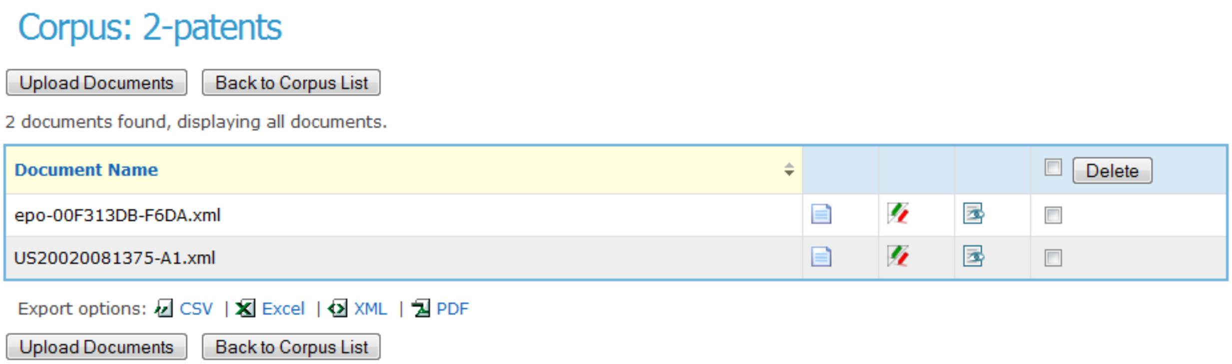


Figure 3.6: Document List

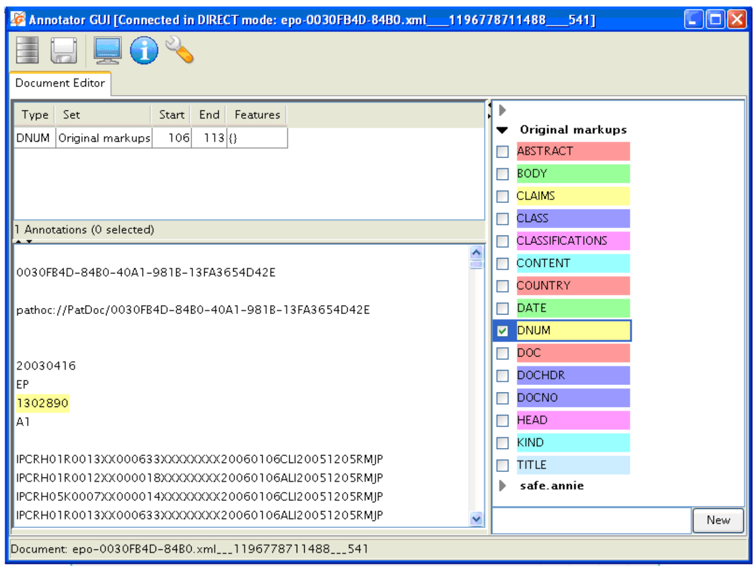


Figure 3.7: Annotation Editor

Diff Editor, where you need to specify which two annotation sets and the common annotation type you want to compare.

Figure 3.8 depicts a comparison between two annotators who have their annotations contained respectively in *annotator1* and *annotator2* annotation sets. The first is the key, i.e. the gold standard, and the second is the response, i.e. the tested data. There is a common annotation type *SectionTitle* in these annotation sets. It is thereby possible to do a comparison on this annotation type.

The option *Features* has been set on *All* – this means that all values of the features are taken into account during comparison. For example, the first line in the table shows for the same text, *Field of invention*, two different feature values, *Background* and *Field* for the same feature named *type*.

You can see different colors according to the type of difference between the annotations from each annotation set that is also shown in the central column of the table.

Help window is provided in the GUI if you click on the information sign icon in the top tool bar. For more information, please refer to the *Chapter 13*, Performance Evaluation of Language Analysers, in Gate User Guide (`http://www.gate.ac.uk/sale/tao/index.html`).

3.2.4 IAA Calculation

From the document list page, you can do IAA calculation over annotation sets on each document. Click on the IAA icon (third icon from the left hand side shown in Figure 3.6), and the new page should open as shown in Figure 3.9.

You need to select at least two annotation sets from the left available annotation sets box and move them to the right box to do the comparison over a shared annotation type. Also there are four algorithms available for you to choose. The screenshot shown in Figure 3.10 is an example of the result based on the chosen annotation sets, type and algorithm. Note: four algorithms offer different result formats because of their difference in various values.

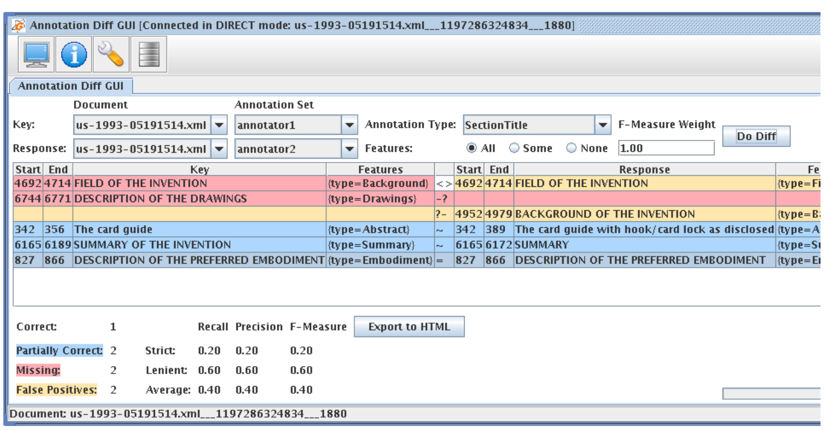


Figure 3.8: AnnDiff Editor

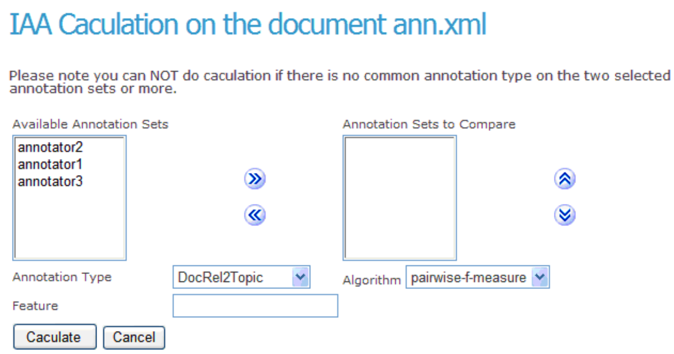


Figure 3.9: IAA Calculation

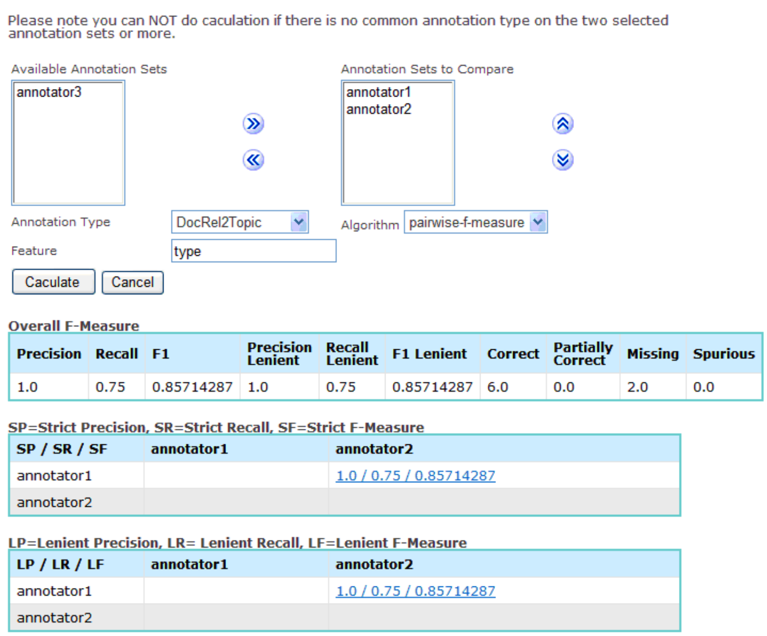


Figure 3.10: IAA Calculation Result

3.3 Annotation Schemas

Users can upload their annotation schemas in XML format to the server so that they can be used in the process of manual annotation. Click on the *Resources >> AnnotationSchemas* option from the top menu and you will reach the annotation schema list page, as shown in Figure 3.11.

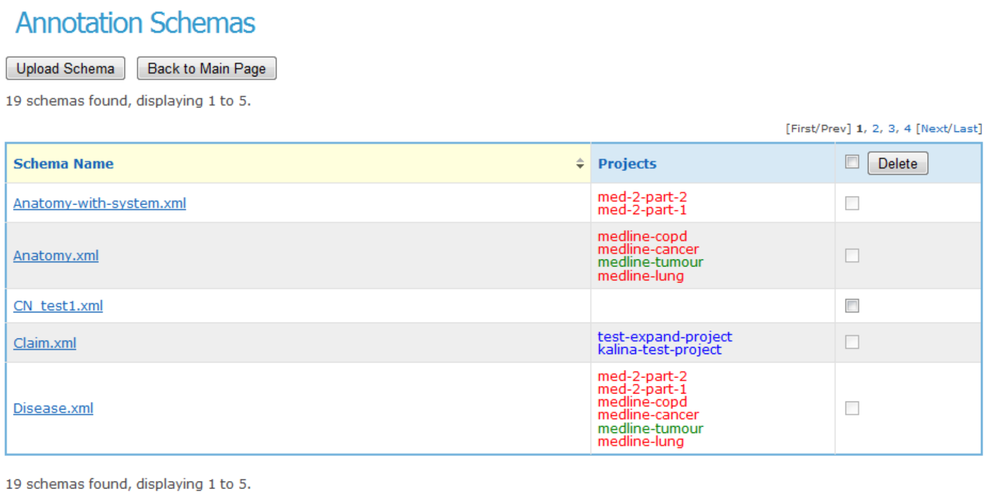


Figure 3.11: Annotation Schema List

You can see if some annotation project, using particular schema is running, suspended or finished. The links to the annotation project statistics can be of different colours, depending on project status: *red*: suspended; *blue*: running; *green*: completed. More about project statistics you can find in the next chapter.

To upload a new annotation schema to the server, click on the *Upload Schema* button and you should reach the page where you can select the schema file on your file system (see Figure 3.12).

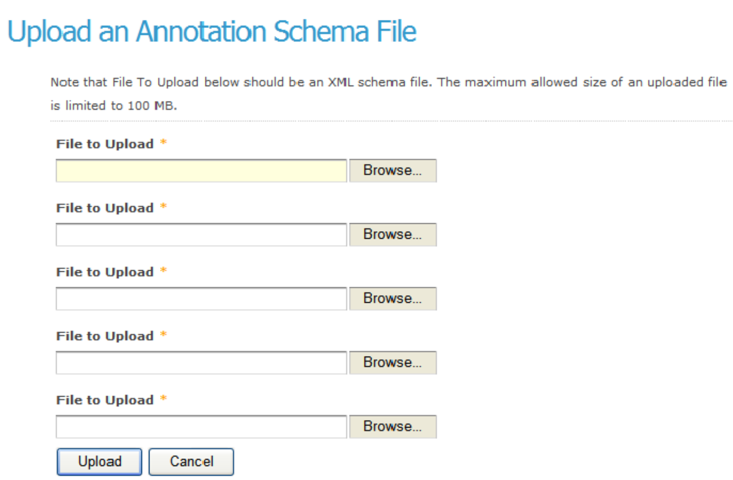


Figure 3.12: Upload Annotation Schema

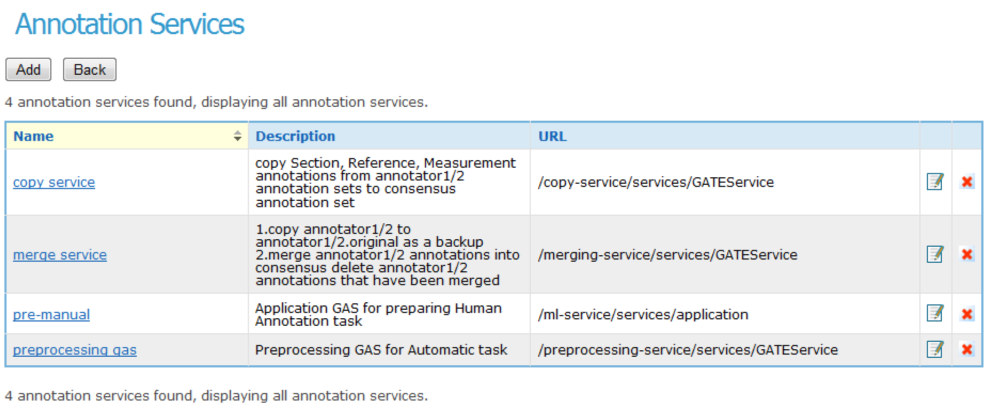
After clicking on the *Save* button from , you will see the confirmation message and the annotation schema you have added.

To delete one or more schemas, check the tickbox in the last column of the table (see Figure 3.11) and click *Delete*.

If you want to view annotation schema click on *View* icon.

3.4 Annotation Services

Users with *Administrator* role can view, add, update, and delete annotation services. To view available annotation services, select *Services* option from *Resources* menu (see Figure 3.13 below).



Name	Description	URL		
copy service	copy Section, Reference, Measurement annotations from annotator1/2 annotation sets to consensus annotation set	/copy-service/services/GATEService		
merge service	1.copy annotator1/2 to annotator1/2.original as a backup 2.merge annotator1/2 annotations into consensus delete annotator1/2 annotations that have been merged	/merging-service/services/GATEService		
pre-manual	Application GAS for preparing Human Annotation task	/ml-service/services/application		
preprocessing gas	Preprocessing GAS for Automatic task	/preprocessing-service/services/GATEService		

Figure 3.13: Annotation Services

Here you can delete or update existing annotation services or create a new one. Note that currently, GATE Teamware supports only GATE services (GAS).

To create a new GAS service, click on *Add* button and fill in the form as shown in Figure 3.14.

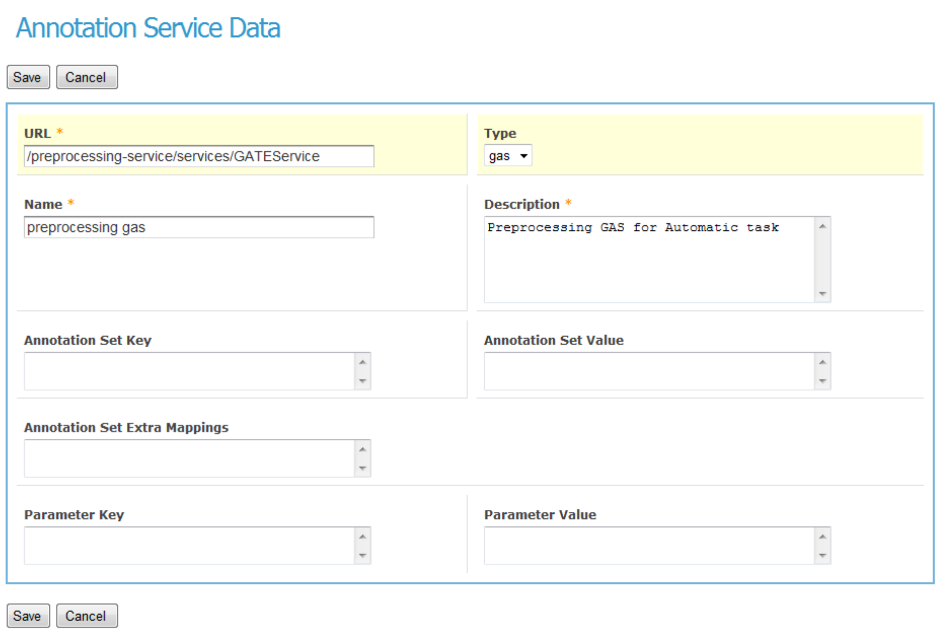


Figure 3.14: New Annotation Service

This form is used to inform Teamware about a published service – the service must already be deployed at a URL accessible to the Teamware server. The GATE user guide[2] explains how to save a GATE application in a suitable package for publishing as a Teamware service.

Mandatory fields are:

- *url*: the URL of the GAS service you have published;
- *name*: the friendly title, you will refer to, when you configure your *Workflow Template*;
- *description*: a few words about what GAS is doing.

It is important to mention that if you ommit *http* in the URL, it will be assumed that GAS is published under the same context as the web application. For example on Figure 3.14, the preprocessing-gas has URL */preprocessing-service/services/GATEService*. That means if GATE Teamware is hosted on *http://somehost/teamware* the GAS URL will be resolved as: *http://somehost/teamware/preprocessing-service/services/GATEService.*

[2] `http://gate.ac.uk/userguide/sec:howto:export`

Chapter 4

Project Management

4.1 Introduction

Project Management functionality is accessible to the *manager* and *administrator* roles only. Managers can access and modify content created by themselves only, while administrators can see and edit all content available in the system.

From the home page you can either use available shortcuts, or you can access even more options from the *Projects* menu on the top. As a manager you will see options as shown in Figure 4.1, while as an administrator you will see more options (see Figure 4.2).

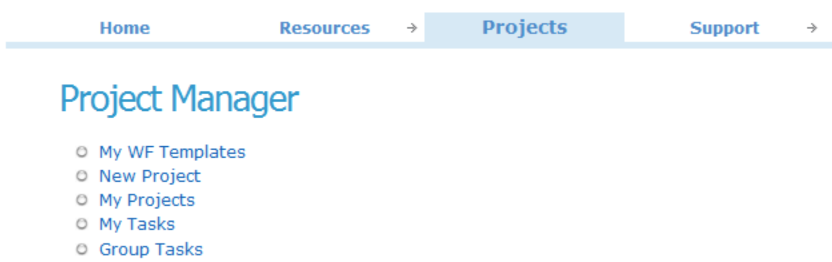


Figure 4.1: Project menu for *managers*

Annotation projects are meant to support the following annotation modules:

- *automatic*
- *manual*
- *post-manual (consensus)*
- *review*
- *post-processing (final consensus)*

You can use any particular module or all of them in the predefined order. If you want use *manual* or *review* module, you will be able to manage annotators and curators and specify annotation rules.

In case of *automatic*, *post-manual* or *post-processing* mode you can choose from existing GATE services (GAS) which can be executed in sequence. Please see the *Annotation Services* chapter for more details.

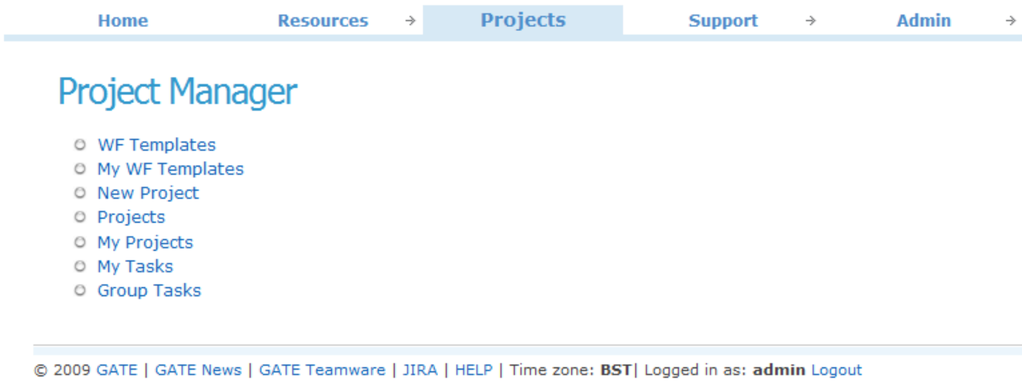


Figure 4.2: Project menu for *administrators*

The motivation for combining all modules is to use pre-processing to bootstrap the manual annotation process. The result of manual annotation result is consolidated by using post-manual module, in other words the annotation sets created by human annotators, e.g. *annotator1/2* are inspected and merged. The identical annotations are moved to *consensus* annotation set and the different ones are left in *annotator1/2*. Review process should validate the annotations in *annotator1/2* and delete false ones. This will leave only reviewed annotations in *annotator1/2*.

Finally, the remaining - valid annotations in *annotator1/2* are copied to *consensus* set by post-processing module.

4.2 Create new WF template

To create new project, it is mandatory to create Workflow (WF) template first. Templates are later accessible for creating similar projects.

To create new template, you need to choose *Projects >> My WF Templates >> Create New WF Template with Wizard* option from the top menu. The wizard will then guide you through the process of creating template as explained next in Section 4.2.1.

Another way to create new template is during the creation of New project (explained later in Section 4.4) by selecting option *Blank template* at the list of available templates. Click *Select* button, and you will be redirected to the *Workflow Template* page, from where you are entering *Workflow Template Wizard*.

4.2.1 Workflow Template Wizard

You need to provide the *name* and *description* of the template. These two fields will be later used to show overview of the templates and therefore should contain useful information to help you distinguish between various templates. Click *Go* button as shown in Figure 4.3).

The wizard comprises several steps, the number of which is not fixed as each of them depends on the preselected options in the previous step.

During the wizard, the *progress diagram* on the right hand side will show your progress by highlighting your current step.

The first step is the *Configuration: "select annotation modules"* step, where you need to choose one or more annotation modules: *automatic*, *manual*, *post-manual*, *review* and/or *post-processing*, as shown in Figure 4.4).

Click *Next*. Depending on the chosen modules, the steps to follow will vary. Here is the detailed explanation of each possible step in case when all available modules are selected.

Pre-processing: To configure *pre-processing* you need to select one or more pre-processing services. When you are done with selection press *Next*. Alternatively you can add as many services as you want to the pipeline by clicking on the button *Save and add another*. This step is shown in Figure 4.5).

Manual annotation: Manual annotation needs a few parameters to be specified. This is done in the *Configuration step: "select annotation rules"* as shown inFigure **??**. Here you will need to:

- select number of annotators per document. For example if you select 2, each document should be annotated by 2 different annotators;
- leave *cancel task allowed* checked; this will allow annotators to cancel tasks if they feel unsure about how to annotate the document correctly;
- leave *anonymous annotation* checked; this will create generic annotation set names for annotators (e.g., *annotator1*, annotator2, etc.) instead of using their usernames. This option is useful when you later search annotation sets within documents in generalised fashion (for example, using regular expression 'annotator*');
- select some of the existing annotation schemas which will be loaded inside the Annotation Editor during the process of manual annotation (you also have the option to add new schema, if you want).
- choose the service which will prepare manual annotation task (copy preprocessed annotations into annotator's annotation set)

Post-Manual: and Post-Processing In *Configuration steps: "select post-manual service"*, you can select one or more post-manual/post-processing services as shown in Figure 4.6. These steps are identical to *pre-processing*. The main point is that you can chain services in the order you like and make them executed in a pipeline.

From the *Configuration: "template overview"* page (Figure 4.7) the template can be saved by clicking on the *Save template* button. You will be redirected to the *My WF Templates* page. Note that if you are not happy with the settings, you can repeat the setup procedure by pressing *Start Again* button.

At any time you can get more information on the current step or instructions on how to proceed by hovering over the *Help* button. This is especially useful in cases when you are not sure on how to complete the step. Quitting the wizard is always possible by pressing the *Quit* button. After the warning, and choosing the option *Yes*, the wizard will exit, but it can be resumed any time by selecting *Projects >> My WF Templates*, and then by clicking *Resume* button for the selected template.

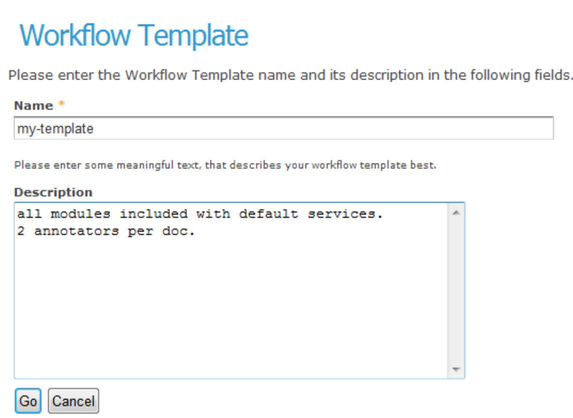


Figure 4.3: Create New Workflow template

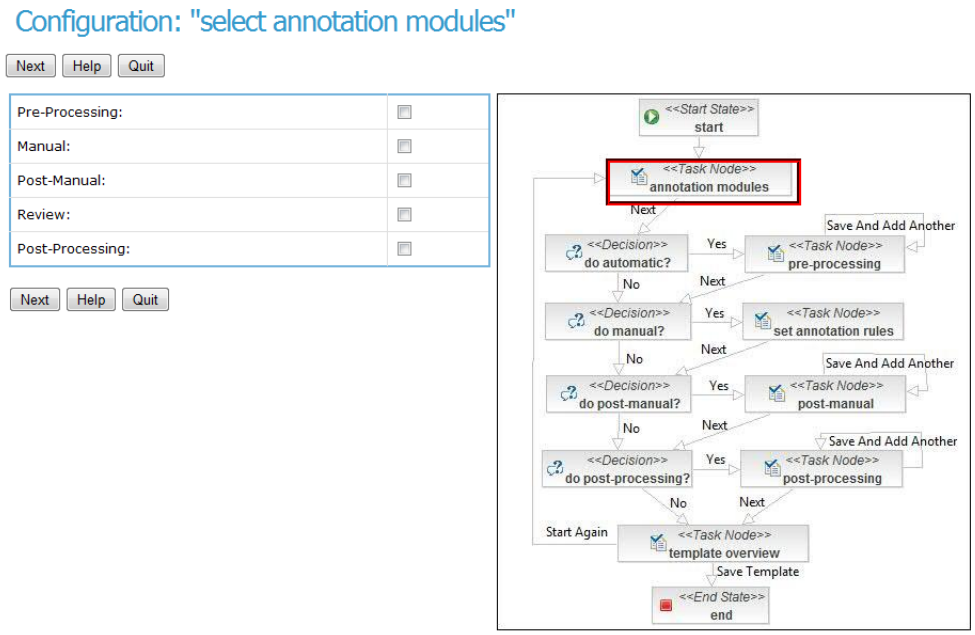


Figure 4.4: Configuration: Select Annotation Modules

4.3 My WF templates

Choose *Projects Menu >>My Workflow Templates* option from the top menu or *My Workflow Templates* link at your home page. You will enter *My WF Templates* page as shown in Figure 4.8. As we have

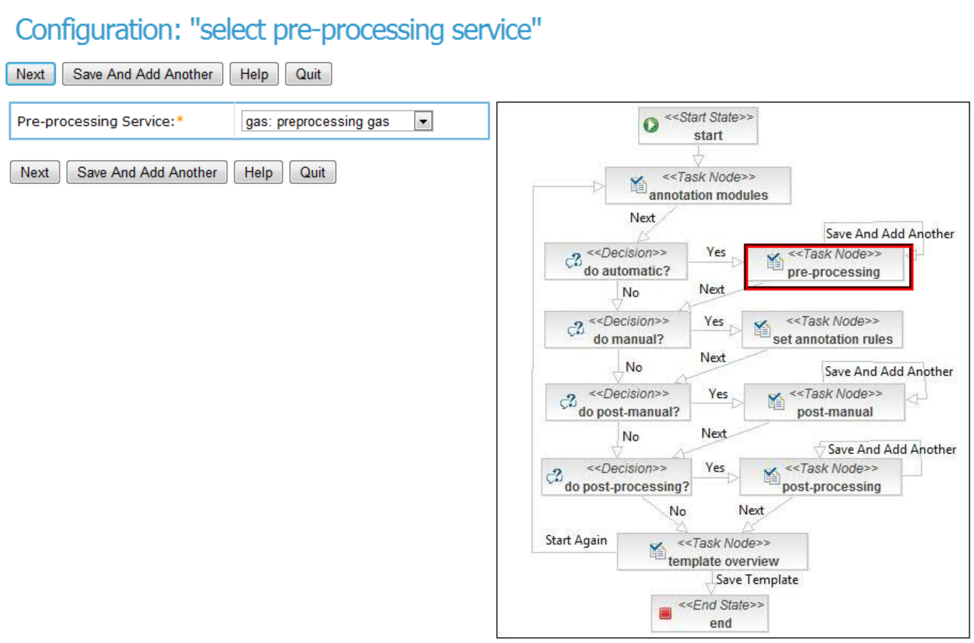


Figure 4.5: Configuration: GAS Pre-processing

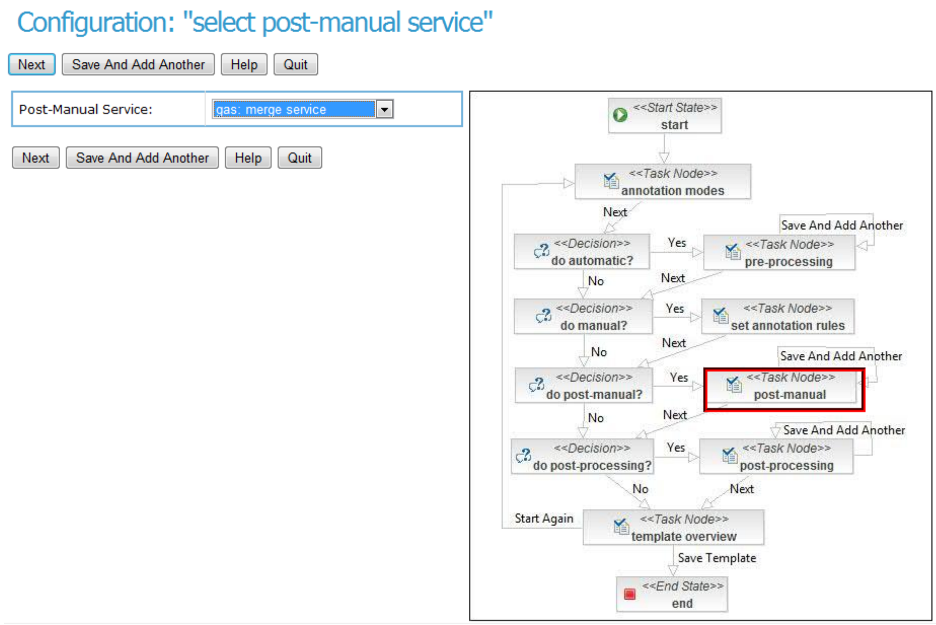


Figure 4.6: Configuration: Select Post-Manual

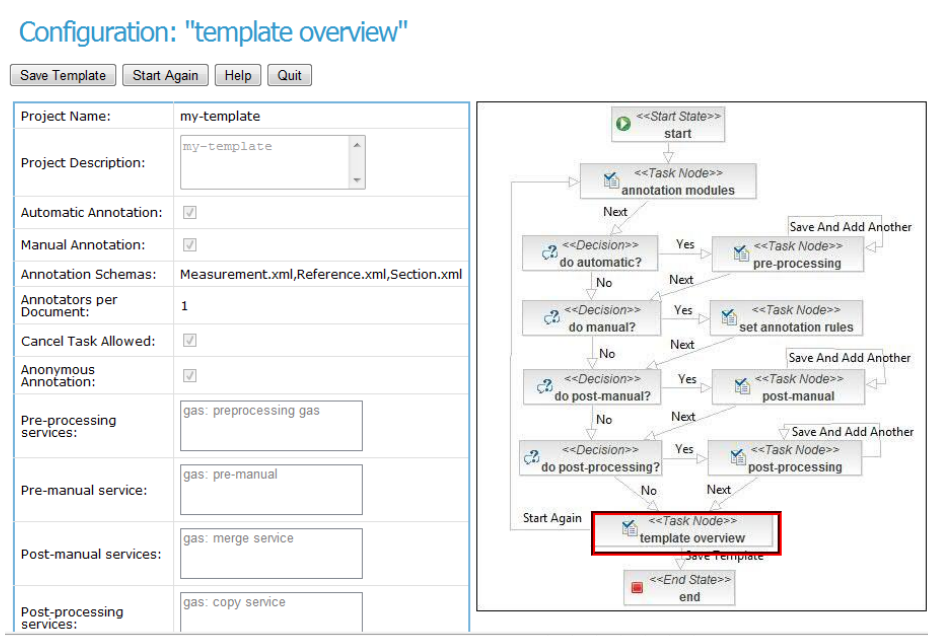


Figure 4.7: Configuration: Template Overview page

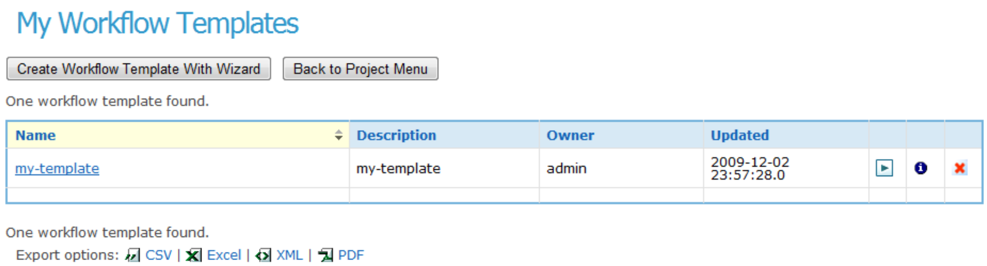


Figure 4.8: My Workflow Templates

mentioned previously in Section 4.2, you will land on this page after finishing wizard successfully. In this case, the status of workflow template will be *Ready*, meaning that you can create project based on this template (the *Create Project* icon will be visible only if the template is ready). If during the wizard you quit the process before finishing template setup, the status of the template will be *Not Configured*. By clicking on *Resume* icon, you will be able to resume the setup and complete unfinished steps.

If you want to delete a *Workflow Template*, you can do so by clicking on the *Delete* icon. Note that all projects based on selected template will be deleted.

Of course, you will often need to check what configuration is saved in template. This can be done by hovering *I* icon.

4.4 Create new project

For your convenience, there are several ways to create project. Two have been already mentioned:

- from the corpora list page, you can create project for the specified corpus by clicking on *Start Project* icon (see Section 3.1.1).
- from the My Workflow Templates page, you can create project for the specified template by clicking on *New project* icon (see Section 4.3).

Here we will cover other ways.

From the top menu select "Projects" >> "New Project", or access it directly from the home page. You will be redirected to *Create New Project* page where you need to choose *WF Template*. Here you have two options (see Figure 4.9):

- Choose from the list of available templates and click *Select*.
- Create new template, by selecting option *Blank template*. If you choose this option, you will need to follow the steps explained in Section 4.2. After you have finished creating the template, you will be redirected to the list of WF templates.

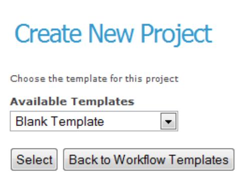


Figure 4.9: Create New Project

You will be redirected to the *Project Data* page where you need to provide details of the project. Requested details vary depending on the annotation modes defined in the WF template. In case you excluded manual annotation mode, the three options will need to be selected (see Figure 4.10): project name, corpus, and manager (you will be selected by default); as a manager of the project you will be able to see this project in *My projects* list and manage it (e.g., change actors, monitor project).

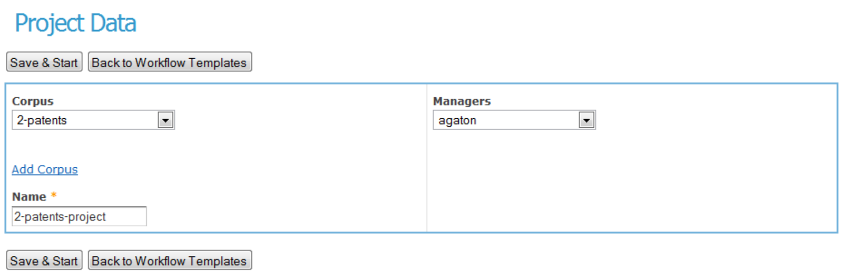


Figure 4.10: Configure Project without Manual Annotation

When *manual annotation* module is defined, there will be two options in addition to the previously selected ones: curators and annotators (see Figure 4.11). For example, you can leave yourself as *manager*

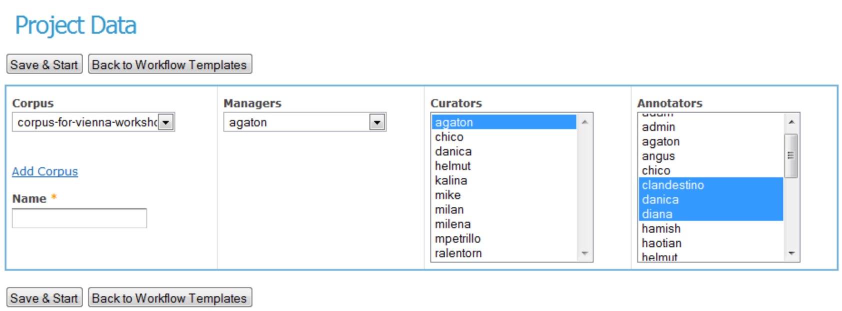


Figure 4.11: Configure Project with Manual Annotation

and *curator*, and you can select as many annotators as you wish in the list (hold Ctrl while selecting annotators for multiple selection of actors).

On the other hand, if you chose *review* module and ommit *manual annotation* module, there will be an option to select only curators and manager, since annotators do not take part in review process.

Click *Save & Start* button and the annotation project will be started.

4.5 Annotation Tasks

If used template contains manual annotation module all annotators should be able to get annotation tasks after logging in and clicking on *Open Annotation Editor* from the home page. We now discuss how annotators (users with *annotator* role execute their tasks and annotate documents.

As a general rule, users with *Annotator* role assigned, participate in the project via annotation editor. The annotator's home page contains restricted operations as shown in Figure 4.12. The annotator can either open Annotator editor, or edit his/her profile.

After clicking on *Open annotation editor* link, the annotator has to ask for the new task. To do this, s/he needs to click on the *Get New Task* button (the first one from the left side). Annotation Editor will start to search for available tasks and the annotator will see the dialogue similar to one shown in Figure 4.13.

Figure 4.12: Annotator Home Page

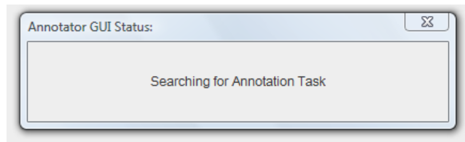


Figure 4.13: Searching for a new task in Annotation Editor

If there is an available task, it will be loaded in annotation editor. The following actions are available (see Figure 4.14):

- Finish task (click on the second button from the left) when you finished with annotations.
- Save task (click on the third button from the left) when you want to save your annotations, but not finish the task.
- Reject Task (click on the fourth button from the left) when you feel unsure how to annotate document and want this to be assigned to other annotator.

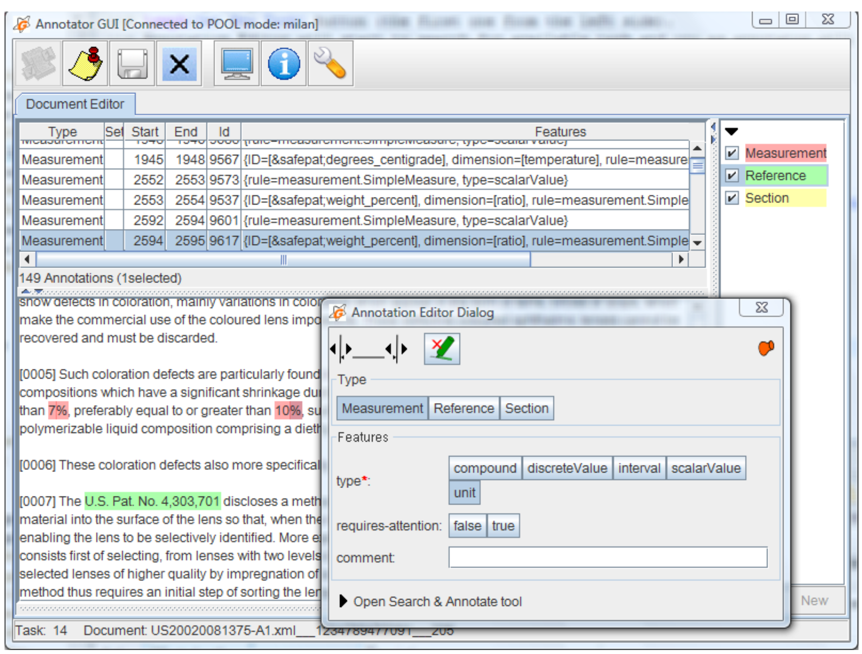


Figure 4.14: Get New Task in Annotation Editor

4.6 Curation Tasks

If used template contains *review* module all chosen curators (remember, you had to select them during project startup) should be able to get curation tasks after logging in and clicking on *Group Tasks* from the *Projects Menu* likd shown in Figure 4.15.

Figure 4.15: Project menu for curators

Now we will explain how curator claims and executes his/her tasks and review documents.

In general, users with *Curator* role assigned, participate in the project through task form. Since, more than one curator can be assigned to a particular project, it is necessary for curator to claim his/ger task first, by clicking on *Accept* icon (see Figure 4.16).

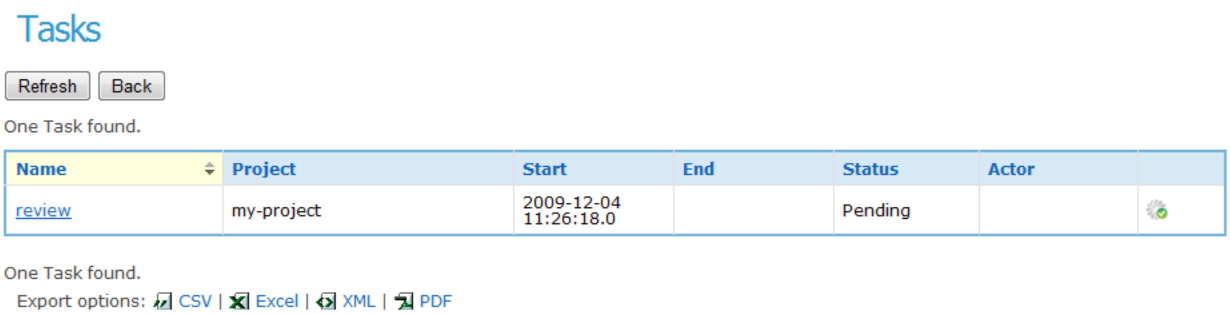


Figure 4.16: Group Tasks

That task will be moved then from *Group Task List* to his/her *Personal Task list*, as shown in Figure 4.17

Curator then executes the task by clicking on *Start* icon. He will be shown the task form with the instructions which he/she needs to follow - Figure 4.18

.

The review procedure briefly comprises the following steps:

- Download corpus you need to review as a zip archive, by clicking on *Corpus*
- Unpack the zip archive and copy documents somewhere on your hard drive
- Launch GATE Developer, create new corpus and populate it with the unzipped documents.
- Run Corpus QA tools (IAA, Annotation Diff, etc) to find out to which extent how annotators agree. Note that Corpus QA Tools are added in GATE 5.1.
- Delete annotations that you consider false from human annotation sets, e.g. *annotator1/2*.

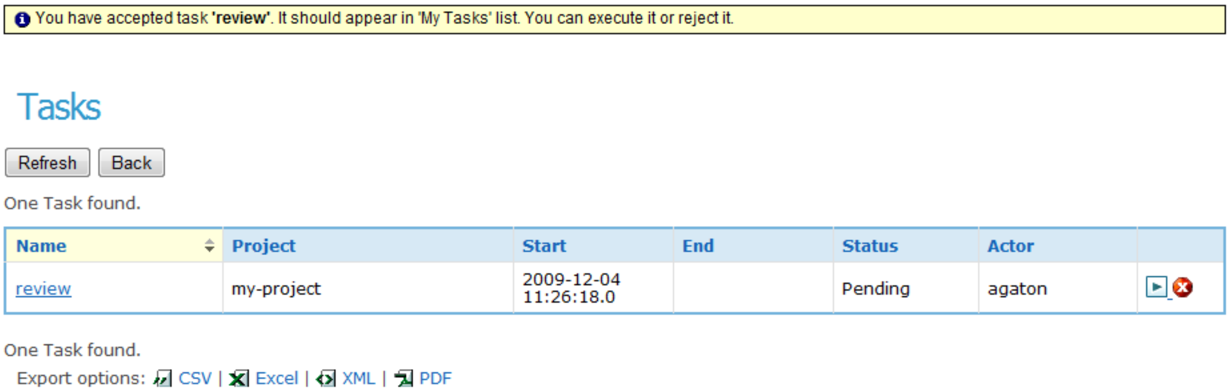


Figure 4.17: My Tasks

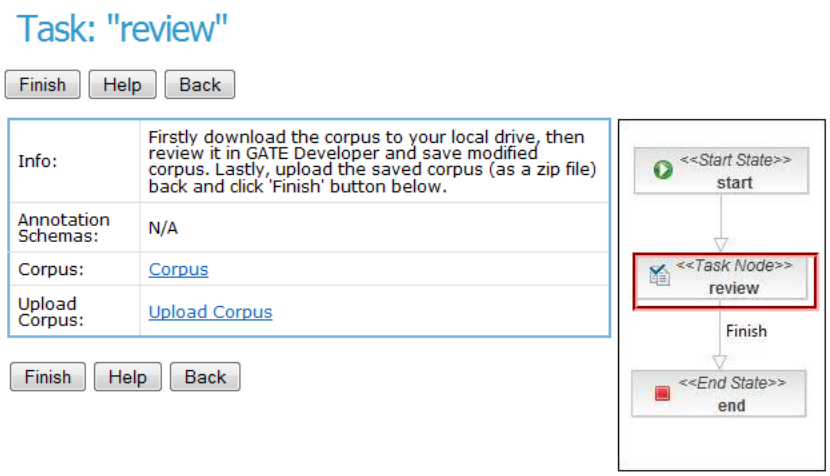


Figure 4.18: Review Task

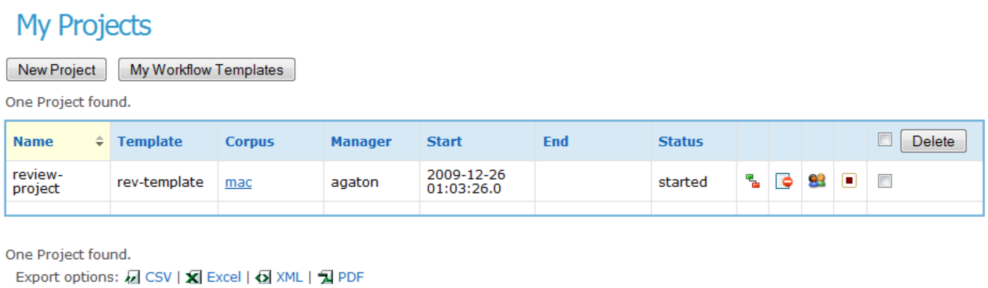
- Go back to Teamware, find your task among *My Tasks* and update the documents in corpus by clicking on *Upload Corpus* link.
- Press *Finish* button in your task form.

As a manager you will have more responsibilities than curators and annotators. One, especially worth mentioning is that managers will be able to manage projects and have insight into the project progress at any time. In the following sections we will provide more details on this.

4.7 My Projects

To manage project, select *Projects >> My Projects* from the top menu. You will see the list of projects that you have created/started, as shown in Figure 4.19. For each project, there are several available options (four icons next to the Status field as shown in Figure 4.19):

- `view processes` for the particular project,
- `suspend/resume` the project,



My Projects

New Project | My Workflow Templates

One Project found.

Name	Template	Corpus	Manager	Start	End	Status					Delete
review-project	rev-template	mac	agaton	2009-12-26 01:03:26.0		started					

One Project found.
Export options: CSV | Excel | XML | PDF

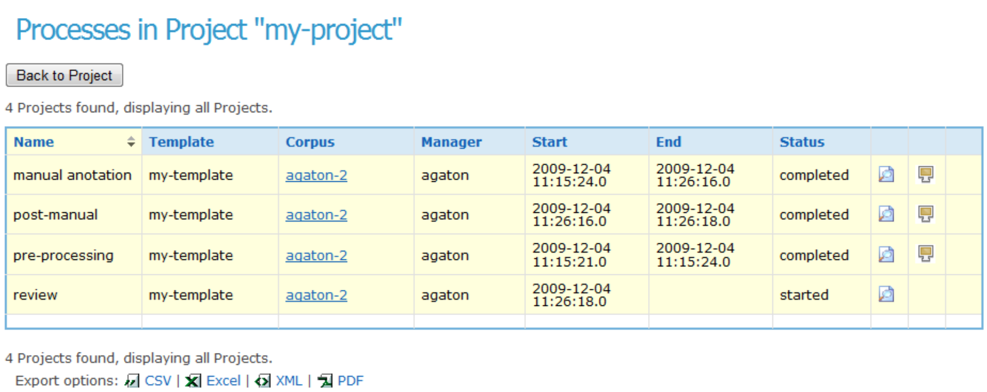
Figure 4.19: My Projects

- `change the project actors`: managers, curators, annotators,
- `end` the project,
- `delete` the project.

Besides that, after project is completed, you will be notified by email. All these options are described in detail in the following subsections.

4.7.1 Project Processes

In order to see project processes, click on *Processes* icon (red and green icon, see Figure 4.19) for the specific project listed in the *My Projects* page. The process list will be shown as in Figure 4.20: : From the *Processes*



Processes in Project "my-project"

Back to Project

4 Projects found, displaying all Projects.

Name	Template	Corpus	Manager	Start	End	Status		
manual anotation	my-template	agaton-2	agaton	2009-12-04 11:15:24.0	2009-12-04 11:26:16.0	completed		
post-manual	my-template	agaton-2	agaton	2009-12-04 11:26:16.0	2009-12-04 11:26:18.0	completed		
pre-processing	my-template	agaton-2	agaton	2009-12-04 11:15:21.0	2009-12-04 11:15:24.0	completed		
review	my-template	agaton-2	agaton	2009-12-04 11:26:18.0		started		

4 Projects found, displaying all Projects.
Export options: CSV | Excel | XML | PDF

Figure 4.20: Project Processes

page, you can:

- `see the status` of each process (started, suspended or completed),
- `monitor processes` by clicking on *Process Monitoring* icon. For details on process monitoring console see details in Section 4.8.
- `view tasks` created for each process by clicking on *View Tasks* icon (see Figure 4.21). When annotators start to annotate, you will be able to see the list of tasks and their status. For example, if there are 2 documents in the corpus and 2 annotators are assigned to annotate each document, there will be 4 tasks in total as shown in Figure 4.21. In this example, one annotator finished task, the other annotator

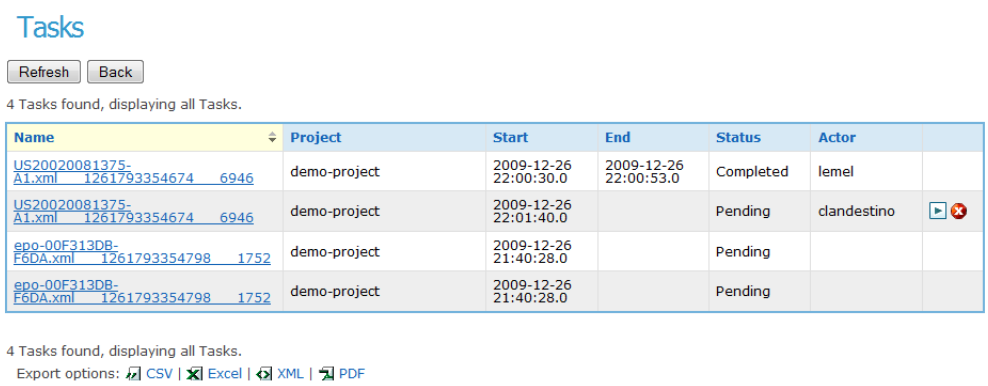


Figure 4.21: Task List

accepted the task, and two left tasks have not been taken yet. From the Task List page you can also see who is the assigned user, and start and end times for specific tasks. Should you want to finish or cancel annotator task, you can do that by clicking on relevant icons. These actions correspond to *canceling document annotation* and *returning document to the pool* respectively.

4.7.2 Suspend/Resume Project

A project can be suspended at any time, by clicking on *Suspend* icon. This means that all processes and tasks in that project will be suspended. The project can be resumed later, which analogly will result in resuming its processes and tasks. This is shown in Figure 4.22:

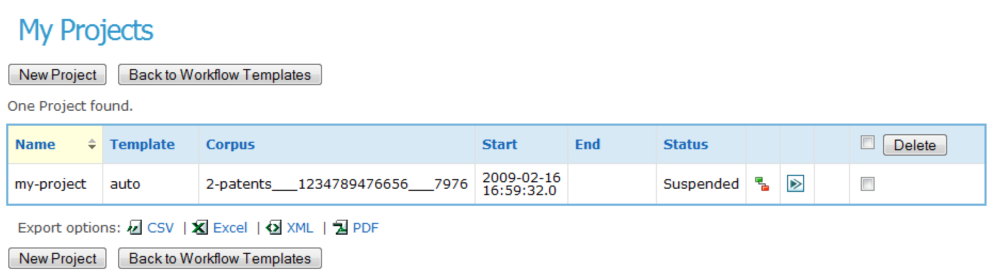


Figure 4.22: Suspend Project

4.7.3 Change Project Actors

There is often a need to change actors during project execution. For example, you might want to pass manager role to someone else or to change curators and annotators due to some changes in personnel. This option is available in GATE Teamware. From the *My Projects* page you need to click on *Change Actors* icon for the specific project, and you will be redirected to the page where you can change all actors as shown in Figure 4.23. As it is visible in Figure 4.23, the current actors will be preselected.

4.7.4 End Project

A project can be ended at any time, by clicking on *End* icon (the first on the right). This means that all processes and tasks in that project will be ended. The project cannot be resumed later. This option is useful

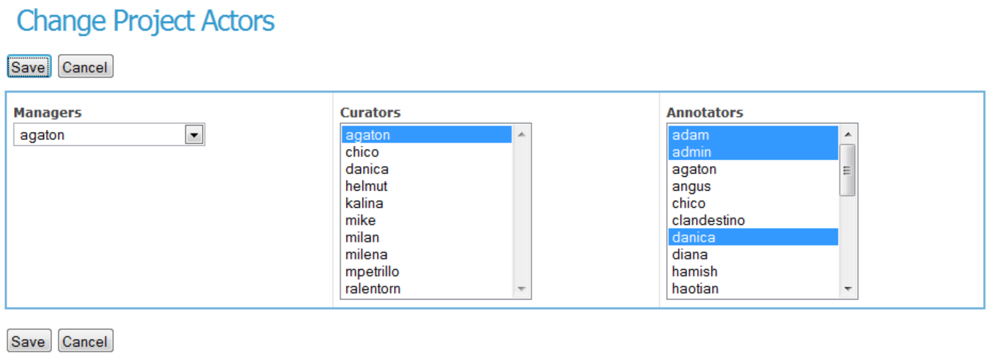


Figure 4.23: Change Project Actors

if manager wants to stop the work on the project, but leave the project history intact. This is shown in Figure 4.24:

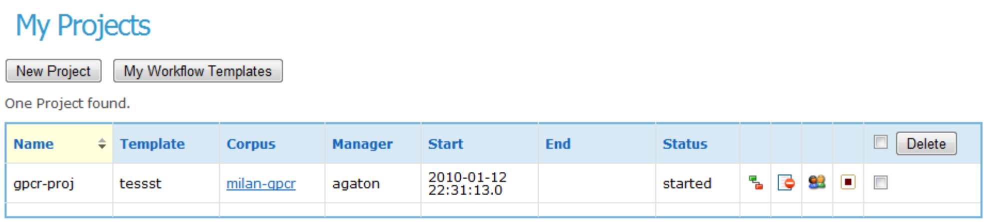


Figure 4.24: End Project

4.7.5 Delete Project

With time, the number of your projects will increase, therefore we provide the option to delete old projects. It is not restricted only to completed projects, so you can delete even running projects. Deleting a project will also delete the corresponding processes and tasks.

4.7.6 Email Notification

As a manager you will be notified by email, upon your project completion. This can be very useful, if you are not actively monitoring execution. The example of email you can expect is shown on Figure 4.25.

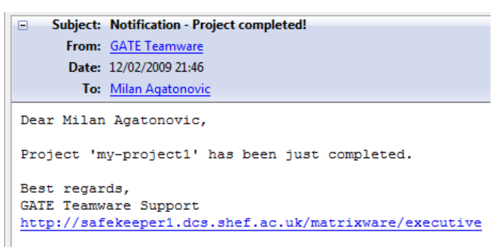


Figure 4.25: Project Completed Notification

4.8 Process Monitoring

Here we will introduce Process monitoring console for real-time monitoring and gathering information about currently executing processes. The main idea behind this functionality is to:

- isolate annotation process failures,
- improve annotation process performance,
- help making decisions, and
- improve resource management.

At the moment, GATE Teamware provides four different views in process monitoring console:

- Document Status Summary
- Annotation Status Details
- Annotator Record Summary
- Record Summary for All Annotators

Each view is explained in details in the following subsections.

4.8.1 Document Status Summary

To access this page, click on *Process monitoring* icon in the process list for particular project. You will see the progress of the annotation process both in tabular and graphical view using pie chart as shown in Figure 4.26. The progress is displayed in real time so that it is possible to monitor how many documents

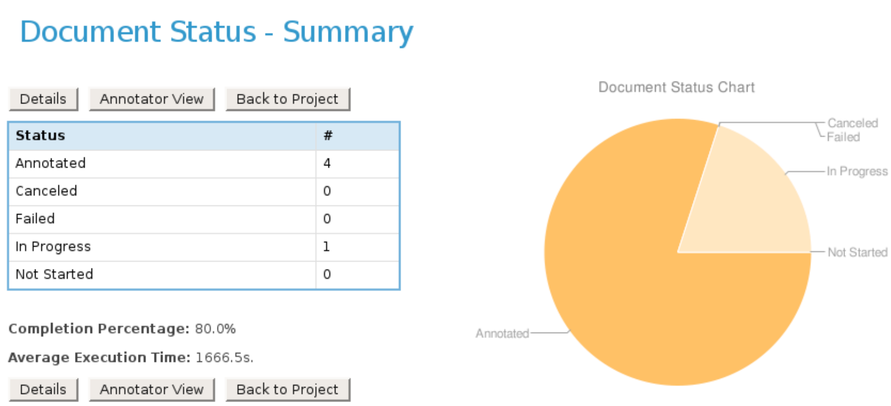
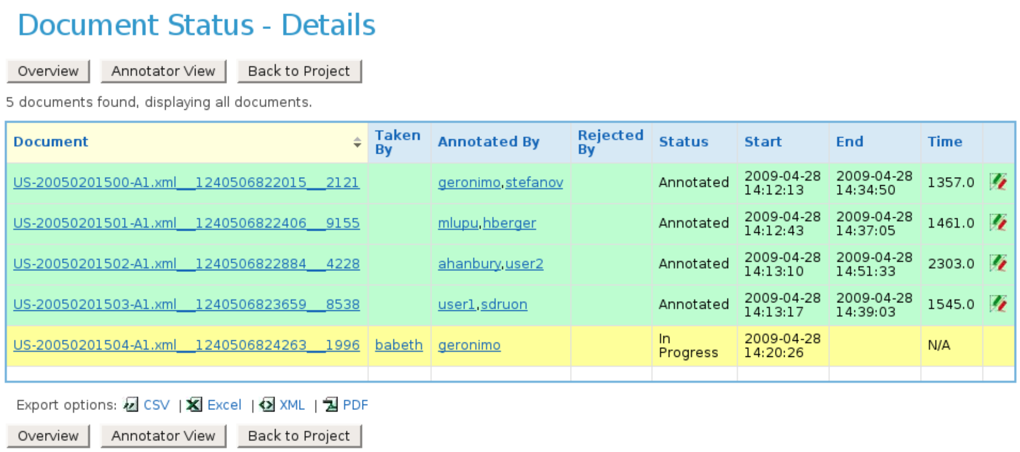


Figure 4.26: Document Status Summary

have been finished, canceled, annotated, failed, in progress, and not started. When the annotation process ends, the pie chart will disappear.

4.8.2 Document Status Details

Annotation Status Overview gives pretty condensed view about what is going on inside process. For more detailed view, click *Details* button, and you will see the details of the annotation project progress by document, including names of annotators who started the work on it, finished or rejected the document as shown in Figure 4.27



Document	Taken By	Annotated By	Rejected By	Status	Start	End	Time	
US-20050201500-A1.xml__1240506822015__2121		geronimo,stefanov		Annotated	2009-04-28 14:12:13	2009-04-28 14:34:50	1357.0	
US-20050201501-A1.xml__1240506822406__9155		mlupu,hberger		Annotated	2009-04-28 14:12:43	2009-04-28 14:37:05	1461.0	
US-20050201502-A1.xml__1240506822884__4228		ahanbury,user2		Annotated	2009-04-28 14:13:10	2009-04-28 14:51:33	2303.0	
US-20050201503-A1.xml__1240506823659__8538		user1,sdruon		Annotated	2009-04-28 14:13:17	2009-04-28 14:39:03	1545.0	
US-20050201504-A1.xml__1240506824263__1996	babeth	geronimo		In Progress	2009-04-28 14:20:26		N/A	

Figure 4.27: Document Status Details

From this list, the status of the documents is also visible together with the execution time per document. For specific documents in the list it is possible to view details such as:

- `view (annotated) documents`, by following the document name link. This will launch *Annotation Editor*, which is described in Section 3.2.2
- `compare document annotations`; to do this follow the icon at the end of the *annotated* document row. The document with annotations will be loaded into the *Annotation Differ* application, which is described in Section 3.2.3

4.8.3 Annotator Record View

Both annotation status views explained above are focused on the document itself, while there might be the cases when the manager wants to see how particular annotator performed. Therefore, we have implemented *Annotator Record View*, which you can access by clicking on the annotator name from *Document Status Details* page. You will be able to see how many documents annotator completed, canceled, started to annotate, along with some useful metrics. A sample *Annotator Record View* page is shown in Figure 4.28:

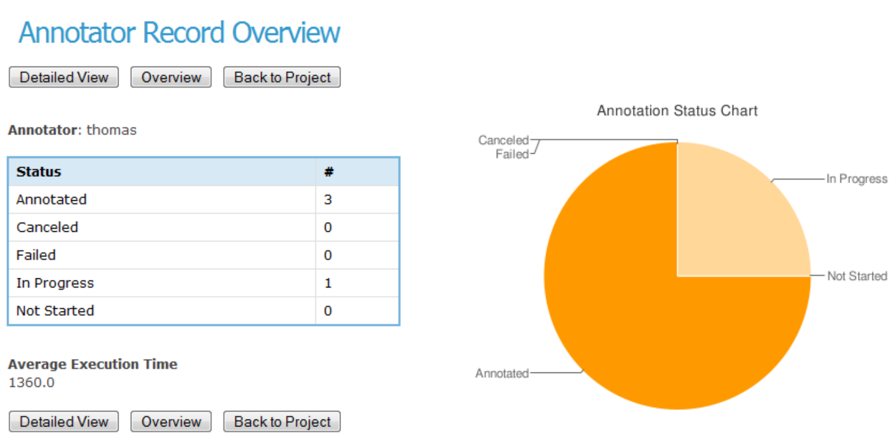
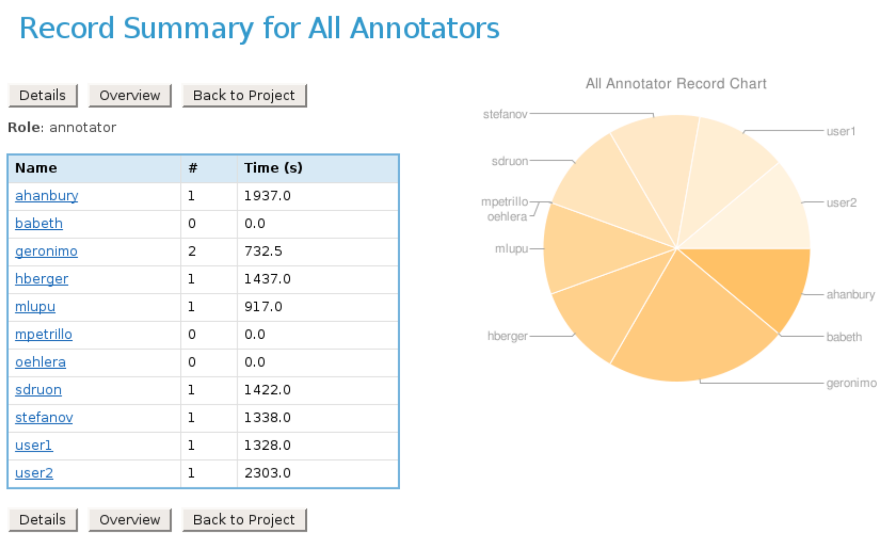


Figure 4.28: Annotator Record View

4.8.4 Record Summary for All Annotators

In case that manager is interested to see how workload is shared among annotators with an indicator about the average time spent per document, *Record Summary for All Annotators* screen can be very useful. You can see this screen by clicking on *Annotator View* button. An example of this screen is shown in Figure 4.29:



Name	#	Time (s)
ahanbury	1	1937.0
babeth	0	0.0
geronimo	2	732.5
hberger	1	1437.0
mlupu	1	917.0
mpetrillo	0	0.0
oehlera	0	0.0
sdruon	1	1422.0
stefanov	1	1338.0
user1	1	1328.0
user2	1	2303.0

Figure 4.29: Global Annotator Record View

Chapter 5

Support

This chapter details Support module which includes: forum, chat and tutorials.

5.1 Forum

GATE Teamware includes forum powered by `http://www.jforum.net`. The complete list of features is available at: `http://www.jforum.net/features.jsp`.

The forum is envisaged to serve as the primary way for interactive support, instead of mailing list. As with any advanced forum, users can post new messages and replies, or create new topics.

Figure 5.1 depicts the main forum page with list of available forums.

Feel free to use this forum to post any particular questions, report bugs and suggest any new features for the GATE Teamware.

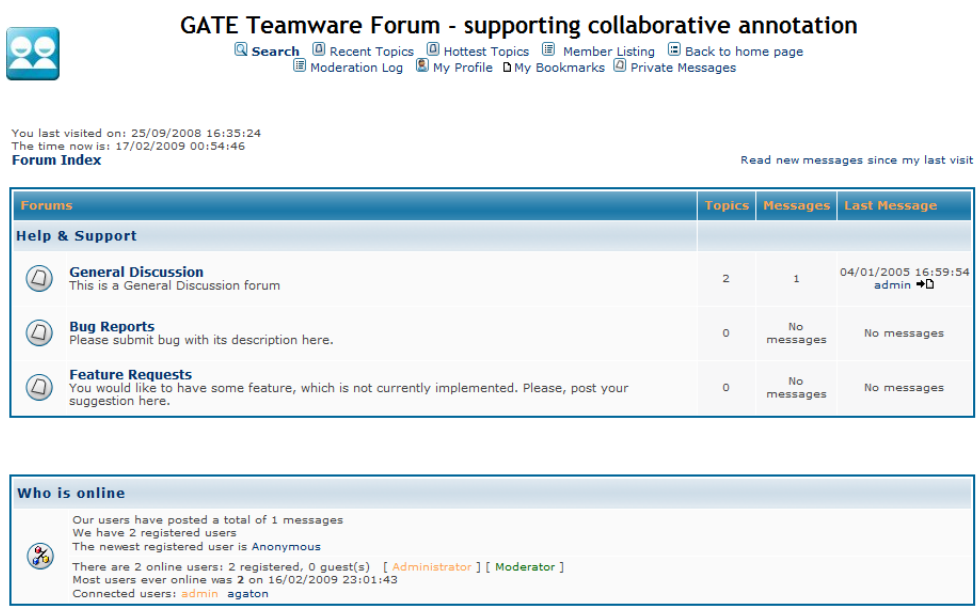
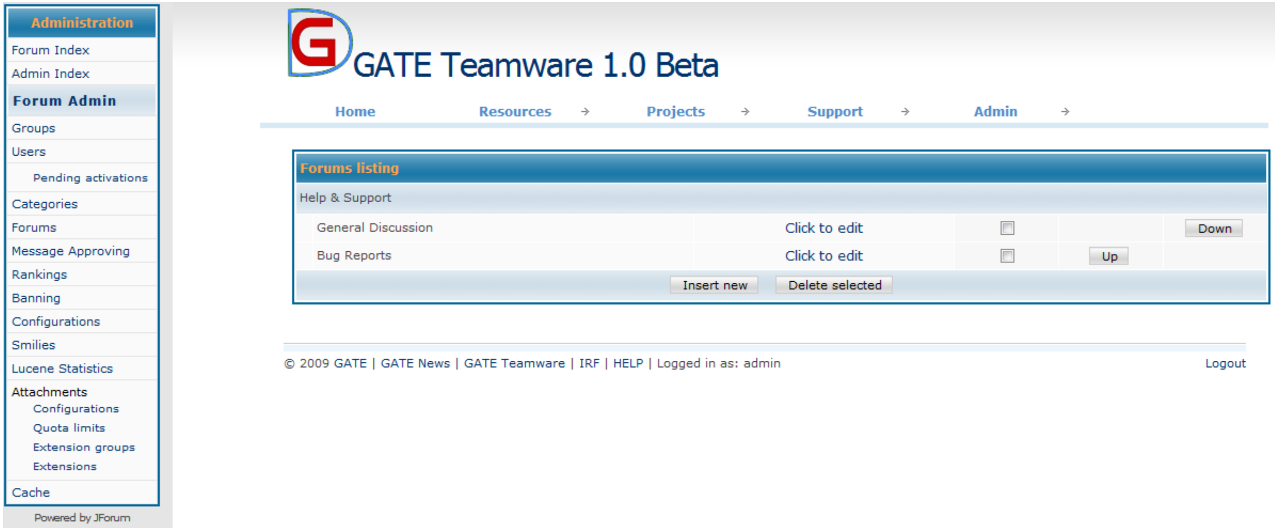


Figure 5.1: Forum home page

Users with *administrator* role can change forum settings and create new forums and categories. This is accomplished through the adiministration panel shown in Figure 5.1 below.



5.2 Chat

Chat is still in experimental phase. Currently, this service is very basic, as shown in Figure 5.2).

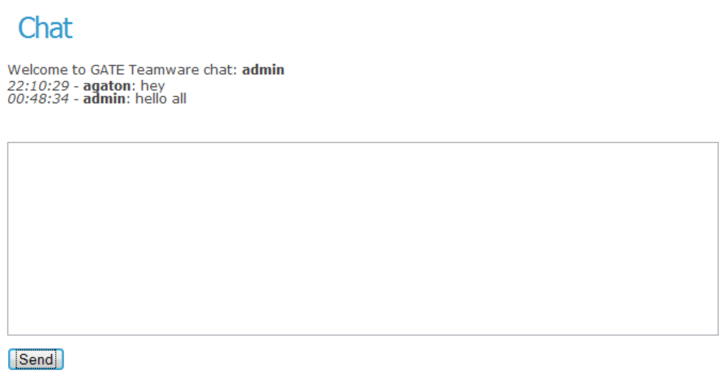


Figure 5.2: Chat Window

Message is broadcasted to all logged users. This is implemented with the idea to see if such functionality can be of interest to the users.

5.3 Help

Apart from this user guide, in this section you will find several movie tutorials which can help you to familiarise yourself with the GATE Teamware interface and functionalities.