# GitHub Enterprise as a service for academic

## Recommendation for further service development

**Short-term**: To progress with GitHub Enterprise as a service for non-for profit academic use. We expect that the service will be used for small internal research projects and teaching assignments. The service will be offered on the case-by-case request model. We will require collect (via SharePoint) user agreements to the service non-for-profit use terms. We will require develop dedicated process assisting account creation of student-groups in the service. It is expected that the service will not require substantial help-desk support apart of initial on boarding activities. We will require update our current project-naming schema, although not ideal, we could introduce department affiliation in the project names.

**Mid-term**: We propose to run exploratory project with the aim to propose solutions for some issues highlighted on this stage.

The biggest drawback of the proposed short-term approach is lack of a clear path how to open the service for external non-Durham contributors. We will require clarify weather GH Enterprise non-for-profit academic licence allows to use system by external academics. We will require propose and implement an authorisation process (possibly based on federated identity technologies) to assist enrolment of external users to the service.

It was pointed to our attention that current GH Enterprise service stores some sensitive configuration files. The project should analyse any potential risk associate with unauthorised access to those information by non-university members.

Given above we think that project scope might review available systems other than GH Enterprise (e.g. GitLab, Bitbucket). The review should include pros and cons of adopting particular solution based on various factors including technology/cost/support.

## Background

Currently, Durham University subscribes to GithHub Enterprise a private instance of GitHub (<http://github.com>) a web-based distributed source code management and revision control system based on Open Source Git solution. Our local repository (<https://source.durham.ac.uk>) is only available to use by some individuals in few professional departments responsible for University’s web applications development. Recently, we have explored GithHub non-for-profit academic use licence and we have found that this could be a very cost effective way to extend (beyond current 50 licence sits) our current service to use by unlimited members from academic and student community.

In order to gather relevant academics views on proposed new service as well as support requirements an email including intendant plans and two questions with four indicative answers had been send to Hamilton HPC cluster email list early this year:

* Q1 – asking for interest in the new service;
* Q2 – asking for support required for the new service.

## Survey findings

7% of the list users (twenty-nine individuals) responded to the original email. Following is breakdown of collected answers:

Questions 1. I would be interested to use GitHub Enterprise service for:

* 1. 7 x yes for - Research projects
  2. 8 x yes for - Research group projects
  3. 4 x yes for - Teaching assignments
  4. 4 x yes for - All previous

Question 2. About your support requirements:

* 1. 10 x yes - I know how to use Git/GitHub and might have occasional questions
  2. 5 x yes - I would prefer drop-in session to demonstrate use of the service
  3. 1 x yes - I would prefer 2h introductory training session
  4. 3 x yes - I have specific GitHub use requirements.

Although, not intended in the original questionnaire, we have collected a lot of additional comments regarding some condition of use of the new system as well as current practices.

Majority of respondent should be classified as advanced users using various, other, available solutions, to mention a few:

|  |
| --- |
| * ICC COSMA users (potential 200) run GitLab server- not clear which version Community (no LDAP groups) or Enterprise (with separate Authorisation and Access process) |
| * ECS - BittBucket +10 users server or cloud hosted unlimited academic accounts <https://blog.bitbucket.org/2012/08/20/bitbucket-academic/> |
| * Maths - BitBucket for non Durham users, SVN WebDAV for non Durham access plus some CI (Jenkins, Mira servers) |
| * Centre for Advanced Instrumentation use GitHub public and bare git. Some individuals use private Personal GitHub <https://github.com/blog/2164-introducing-unlimited-private-repositories> |
| * Mercurial repo hosted in Cambridge, 100 users, some only read some other write |

Substantial number of responses indicated that new service would be only useful if it allows access for the external collaborators.

ECS department might use extensively new service for teaching purpose. This will require dedicated assistance with students account management.

A few comments were made about needs for service allowing versioning and external access to research datasets.

Following sections contains extracts from user comments on various topics.

#### Comments about teaching

*This would be very useful to me in my teaching, as I have been using github increasingly for distributing resources for programming teaching as well as enabling group working. At the moment I distribute practical exercises via github to my L1 Introduction to Programming  class (165 students) as well as encouraging its use amongst L2 Programming Paradigms (60+ students) and MSc software engineering groups.*

*In terms of support, I’m OK with basic github usage, but my needs are probably a little different to others. One of the things I would like to be able to do is to get students to submit work as pull requests, but this requires me to have a mapping between Durham IDs and github accounts which is a pain. Also, while getting hold of private repositories on github for education isn’t too bad, it takes a lot of tedious work to create repositories for individuals or groups, and to add the users to the groups. If it were possible to do this sort of thing for a whole class of students (as defined through Banner or duo) then that would be great, but I suspect it would be fairly difficult and not really something that many other people want. After that, just the means for students or groups of students to set up individual private repositories would be great.*

…

*By using the Git interface to monitor commits to projects and size of commits you can easily track those students who are working hard, those who are not (apparently) working and those that might be struggling. This allows for targeted support for students during supervision or practical classes – sometimes students who are failing don’t ask for help and this allows staff and demonstrators to approach students as needed.*

…

*There shouldn’t be any training support needs from our students, as really that’s what we’re teaching them, but I guess it would potentially add to the weight of technical support required,*

#### Comments about Collaboration

*An important part of the system should be the ability for people outside Durham university to access the repository and gain submit privileges.*

*We use a git repository (with free git software running on CIS Unix platforms) to manage versions of our MOLSCAT and BOUND programs. The only important facility that this lacks is the ability for external users to update their versions and for a very limited number of external users (one or two) to be allowed write access. We'd be interested in a solution that provided this facility, but otherwise I see no advantagein a paid-for solution.*

*It would be a great thing for Durham U to host these things in house, with the caveat that if our own services are any less flexible or more difficult to use than these public ones, the inevitable will occur and our GitHub Enterprise enterprise will gather dust.*

*Myself and collaborators currently have a private version control system for our code based out of Cambridge using Mercurial with a couple of hundred users (about 10 with write access and the rest read-only). We set it up there because of the difficulty in getting access for external users from Durham, but that was some time ago so I don’t know if it would be easier today. Mercurial, along with git has free availability and we find it works well. Unless there is demand for a paid-for version (what else do you get for that and how much does it cost?) and allows simple access for external collaborators then I suspect such a service would be underused. We all tend to use linux command-line or a free interface such as MacHg.*

*Any repository must bee accessible not only by members of the University, but also by people outside our organisation. Most of us are engaged in national or international collaboration, so a tool only accessible by members of the University is pretty much useless.*

*I think that in general having a proper source code management system in the university would be good, but it would need to be accessible to collaborators outside the university to be useful. I would use this for my Cadabra computer algebra software (which is currently on github, but that's not ideal because I cannot do private (sub) projects there without paying). I also might use this for Project IV student projects. I doubt that we have many students in lower years who can handle git (though this was initially on my list of things to teach them for Mathematical Modelling ;-)). They have difficulty storing files for different practicals in different folders...*

*Bitbucket already provides private repositories for free to academics as you probably know, but yes they might stop supporting this one day... There is no day that I do not use git so if the university would run an enterprise instance of github then I would use it as long as it allows also non-durham people to push to the repositories, and as long as the service is reliable enough.*

*In Maths, we have used subversion repositories (with DAV svn access via the department web server) for collaborations. These have been mostly collaborative research rather than code. Maintaining that setup was a pain for us but necessary because there was no CIS (or predecessor) provision.*

#### Comments about price and long term support

*Why don’t you trial a free service when the data server comes on line, with suitable training, and then see if a paid-for service is required? Of course, starting such a service would imply suitable, long term support.*

*Is GitLab cheaper and more sustainable solution?*

*… what else do you get for that and how much does it cost?*