

Infrastructure Reference: Institute of Computational Linguistics

Access to Servers and Systems

After choosing a username and password, the Institute of Informatics (ifi) will provide you with your personal email address (either username@cl.uzh.ch or username@ifi.uzh.ch). Moreover, your chosen login credentials are used to access all the technical infrastructure of the Institute of Computational Linguistics (ICL).

Mailing

Check your mail on the ifi webmail page <https://roundcube.ifi.uzh.ch/> or consult the ifi intranet (see below) for information on how to set up an email client.

Register for the mailing list of the CL server infrastructure by sending an empty mail to cl-linux-server-subscribe@lists.ifi.uzh.ch. You will then be informed about possible outage of or problems on the CL infrastructure.

If you need help working on the servers or you want to report a technical problem, mail to tech-cl@ifi.uzh.ch.

Useful Sites

<https://pub.cl.uzh.ch/wiki/start>: CL Tech Wiki, which provides you with information specific to the CL infrastructure.

<http://www.cl.uzh.ch>: Homepage of the Institute of Computational Linguistics.

<https://www.ifi.uzh.ch/ifi/intranet.html>: Intranet of the Institute of Informatics (ifi).

Fileserver, Storage, Quota, Resources

The dedicated fileserver `clfiles` is managed by the ifi and can be accessed from any CL machine on `/mnt/storage/clfiles`. The data is backed up daily, while snapshots are taken every hour and stored for 24 hours. **Use `clfiles` if you have static data to store. If your data changes often, you will use up all space on `clfiles` in no time as the snapshots attempt to store every change.**

We also have a local data storage on every CL machine, which can all be accessed from any other machine using `/mnt/storage/servername`. These file locations do neither have backups nor snapshots.

https://pub.cl.uzh.ch/wiki/doku/data_management gives you an overview of where to store your personal and/or project data, how backups and snapshots are organised and how to archive your old data. In short, use `/mnt/storage/servername/users/username` for personal data, `/mnt/storage/servername/projects/projectname` for project data you work on and `/mnt/storage/clfiles/projects/projectname` for static data.

When you want to work in a project you must be a member of the corresponding group, which you can check using the command `$ clcl paths`. If you have write permission in a project, its project path will be displayed. Do not work for a project in a directory outside `/mnt/storage/servername/projects/projectname`, otherwise you risk doing work that was already done for the project.

Be aware that your home directories are limited in space (check for disk space your file occupy using `$ clcl quota`), as the servers are primarily meant to provide power for your calculations. Therefore, use the fileserver and storage described above to store your data.

Under `/mnt/storage/clfiles/resources` you will find some useful resources such as aligners, taggers, tokenizers, corpora and many more, for which you can check the overview for resources on <https://pub.cl.uzh.ch/wiki/auto/resources>. Resources are organised to be used on any machine with access to `/mnt/storage/clfiles/resources`. In case you miss a useful tool or software please contact tech-cl@ifi.uzh.ch.

Connect to Servers

`$ ssh username@login.cl.uzh.ch #` You can use this to connect from your system to the login server of the CL infrastructure login.cl.uzh.ch.

When you log on to a CL server from your local machine, you need to use the full-qualified server name, such as login.cl.uzh.ch in the example above. Once you logged on to a CL server, you can jump from server to server using the server name only (e.g. using the command `$ ssh idavoll.cli`).

Caution: Do not develop or run your programs on login.cl.uzh.ch. It only serves as login server for the development infrastructure and is neither powerful enough nor equipped with the necessary programs and tools.

`$ ssh -t username@login.cl.uzh.ch ssh -t server #` You can use this SSH-Tunneling to connect to the login host and from there to another server (e.g. idavoll.cli) by using a single command.

When you log on to a server, you may be informed that the authenticity of your host cannot be established. Confirm by typing 'yes' into the console that you want to continue connecting.

`$ ssh -Y username@servername.cl.uzh.ch #` The argument `-Y` enables X11-Forwarding, which allows to use graphical interfaces such as gedit while working on the servers.

`$ exit #` Close the server connection and/or your terminal. Alternatively, you can use CTRL+D.

Check <https://pub.cl.uzh.ch/wiki/doku/howto/ssh> for further useful commands.

Basic Commands

<https://pub.cl.uzh.ch:8443/wiki/internal/basiccommands> provides you with a summary of commands that might be useful knowing.

For further commands, see for example <https://forum.ubuntuusers.de/> as a source for commands and explanations to them.

Also, you may check the internet for Linux cheatsheets, such as on <https://www.nixtutor.com/linux/all-the-best-linux-cheat-sheets/>

Commands Provided by Tech-CL

`$ clcl #` This lists a variety of predefined commands for the CL servers provided by tech-cl.

For example, you can display the disk space used on the storage by typing `$ clcl disk_usage`. Visit https://pub.cl.uzh.ch/wiki/doku/howto/clcl_command for further information to the different clcl commands.

Copy Data from and to Servers

The quick guide in <https://pub.cl.uzh.ch:8443/wiki/internal/sftp> helps when you have to copy data between servers.

Process Management

You can display the running processes of a server using the htop command. This tool also provides you with information such as memory usage and lets you end or re-prioritise processes.

`$ htop -u username #` Let htop show the processes of *username*. Alternatively, just use `$ htop` to show all processes.

For further information on server processes see <https://pub.cl.uzh.ch/wiki/doku/howto/processmanagement>. Especially, check on how to influence processes' niceness in order to save resources on the CL servers.