

COMPUTEROME 2.0 USERS WORKSHOP

CENTER FOR HEALTH DATA SCIENCE (HEADS)
FACULTY OF HEALTH AND MEDICAL SCIENCES,
UNIVERSITY OF COPENHAGEN, FEBRUARY, 2021



CENTER FOR HEALTH DATA SCIENCE

Provide **SUND** researchers with health data science **services** and provide **training**.

Establish health data science as a topic in its own right and conduct its own **research**.

Started March 2020.
Located at **Panum Institute**.



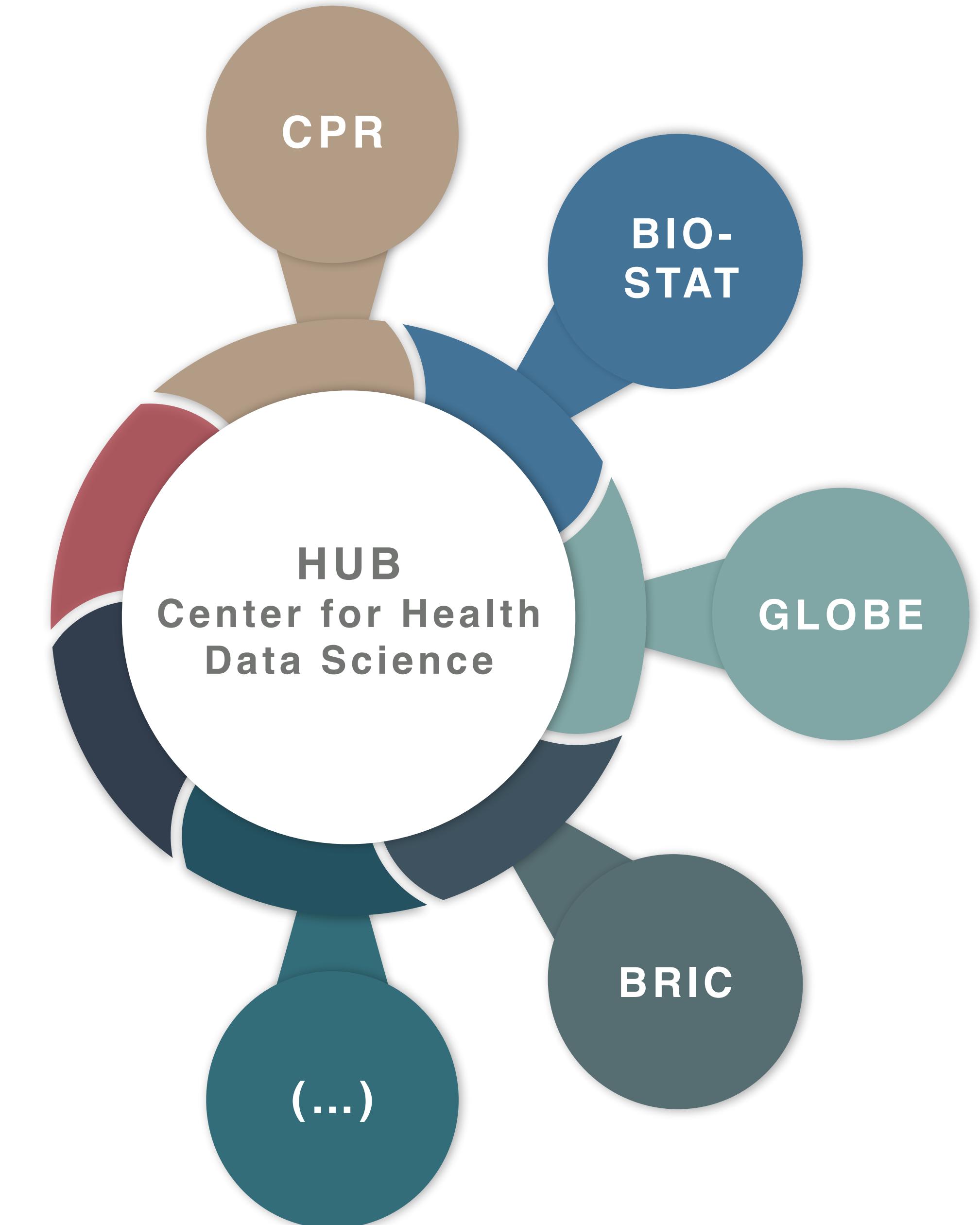
CENTER FOR HEALTH DATA SCIENCE

RESEARCH

HeaDS is a **Hub** for health data science research.

The **Hub** will connect groups and sections at **SUND** (**spokes**) within the field of data science research:

- Coordination
- Collaboration
- Workshops, Seminars, ...
- Visibility



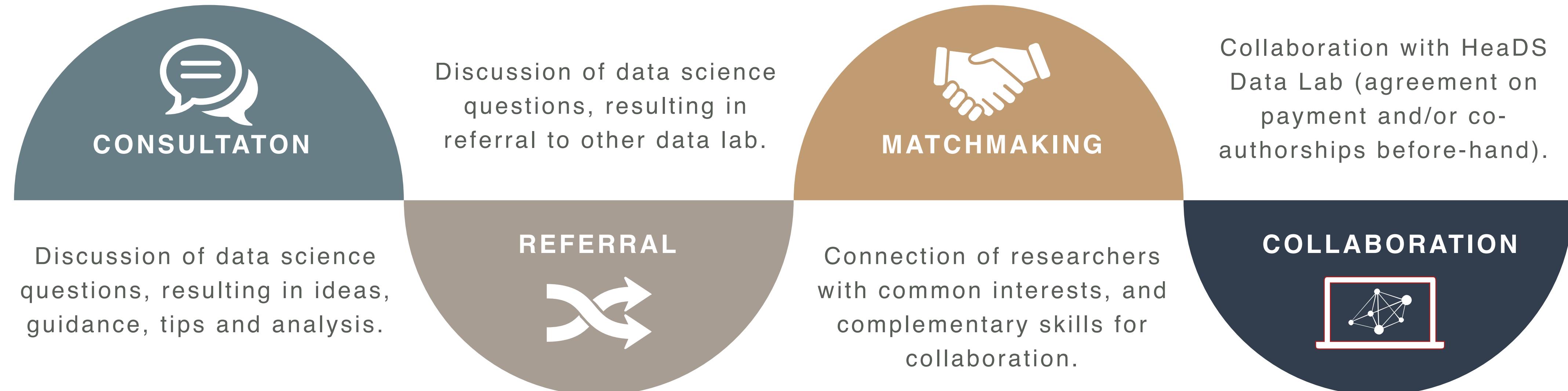
The Center will seek to obtain **external funding** to fuel more research in health data science - **groups within the Center**.

CENTER FOR HEALTH DATA SCIENCE

DATA SCIENCE LAB – SERVICES

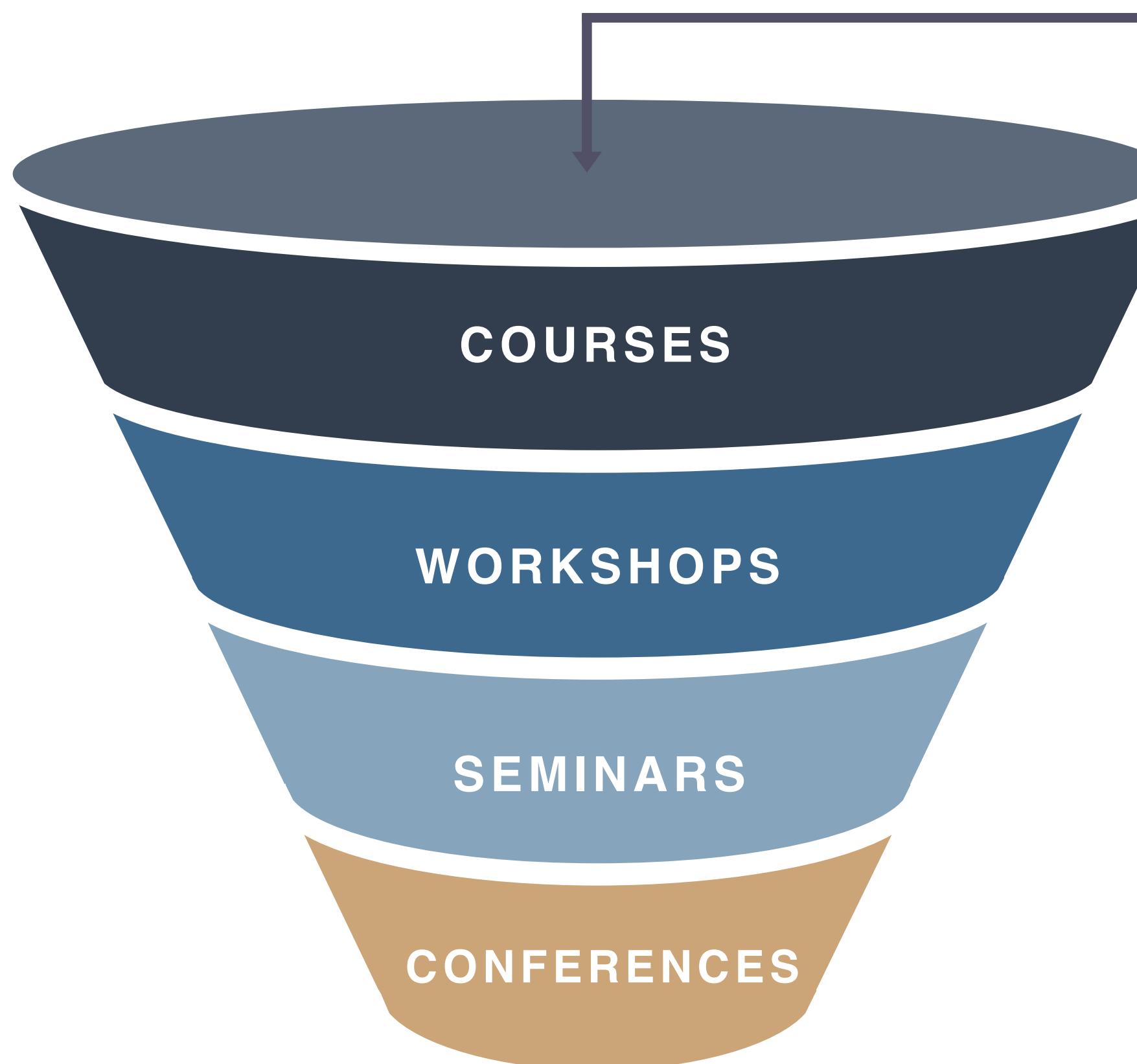
An essential part of the Centers activities is **Support and Service** for **SUND researchers**.

General services (Computerome, data-mangement etc.) or more **specialized** (e.g., NGS analysis).



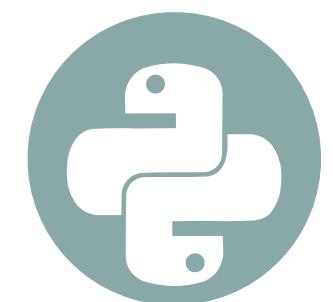
CENTER FOR HEALTH DATA SCIENCE

DATA SCIENCE LAB - COURSES



Course Coordination:

- Python Tsunami, Spring 2021 (two times).



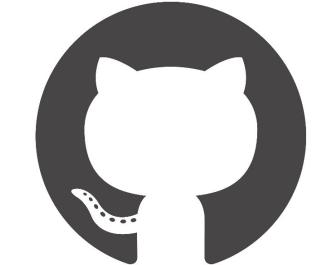
R courses:

- From Excel to R (module 1) - Right now!
- More R modules in collaboration biostatistics.



Git/Github Workshop (Reproducible Science):

- April, 2021.



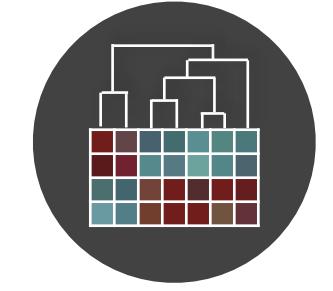
Computerome 2.0 Users Workshop:

- February, 2021 - Collaboration Globe & Danstem



Analysis of Next Generation Sequencing Data

- Summer, 2021.





COMPUTEROME 2.0 USERS WORKSHOP

CENTER FOR HEALTH DATA SCIENCE (HEADS)
FACULTY OF HEALTH AND MEDICAL SCIENCES,
UNIVERSITY OF COPENHAGEN, FEBRUARY, 2021



WHO ARE WE

CENTER FOR HEALTH DATA SCIENCE:

- Anders Krogh, Professor, Head of Center
- Thilde Terkelsen, Data Scientist



ANDERS KROGH



THILDE TERKELSEN



MAGALI MICHAUT

DANSTEM:

- Magali Michaut, Research Consultant

GLOBE INSTITUTE:

- Bent Petersen, Associate Professor
- Shyam Gopalakrishnan, Associate Professor



BENT PETERSEN



SHYAM GOPALAKRISHNAN



ALLAN HAVE

KU IT:

- Allan Have Sørensen, Head of Section

C2 WORKSHOP PROGRAM

DATE: 18th of February, 2021

TIME & PLACE: 09.00-16.00 on Zoom

09.00 - Welcome	13.00 - A Technical Intermezzo
09.10 - Introduction to Computerome 2.0	13.15 - C2 Modules and Queue System
09.50 - The Basics	13.45 - Exercise 3.0
10.10 - Exercise 1.0	14.30 - Coffee Break // Troubleshooting
10.40 - Coffee Break // Troubleshooting	15.00 - Optimization of Time and Cost
11.00 - Working on Computerome 2.0	15.20 - Exercise 4.0
11.20 - Exercise 2.0	15.45 - Q&A and Outro
11.50 - Q&A	
12.00 - Lunch	

PART1

INTRODUCTION

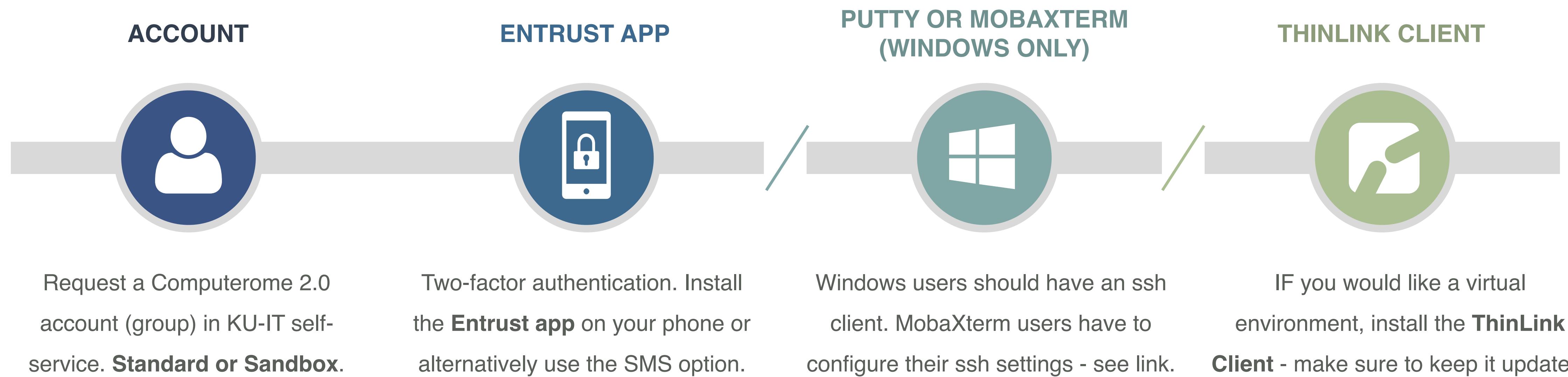
INTRODUCTION TO C2 - BENT

PART 2

THE BASIS

BEFORE LOG IN TO COMPUTEROME 2.0

GETTING STARTED: <https://www.computerome.dk/display/C2W/Getting+Started+-+new+users#GettingStarted-newusers-SSHlogintoComputerome2.0>



LOG IN TO COMPUTEROME 2.0

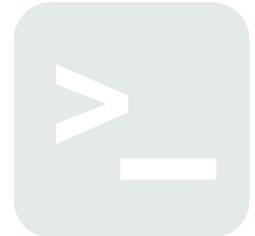
2.0

LOG IN WITH THINLINC

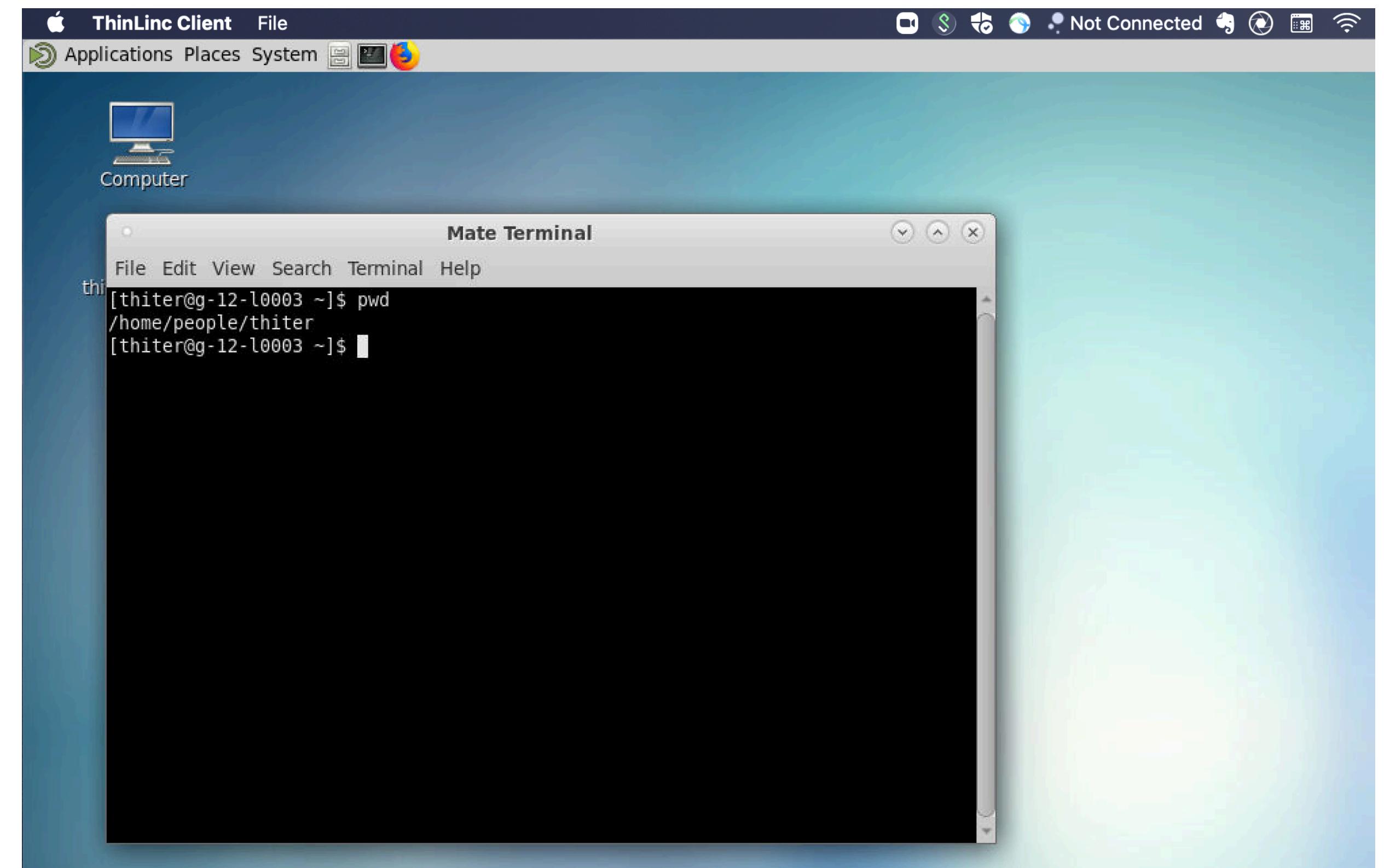


- Download and install the **ThinLinc Client**
- Open and enter your **Username + Password**
- It make take a couple of minutes to connect
- Enter the token from the Entrust App (or Confirm)
- Open a terminal or click your way around
- Exit a full screen - press the F8 key

LOG IN WITH SSH



- Open a terminal
- ssh [username@ssh.computerome.dk](ssh://username@ssh.computerome.dk)
- Enter your password
- Enter the token from the Entrust App (or Confirm)



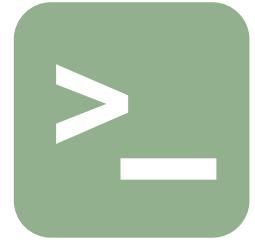
LOG IN TO COMPUTEROME 2.0

2.0



LOG IN WITH THINLINC

- Download and install the **ThinLinc Client**
- Open and enter your **Username + Password**
- It make take a couple of minutes to connect
- Enter the token from the Entrust App (or Confirm)
- Open a terminal or click your way around
- Exit a full screen - press the F8 key

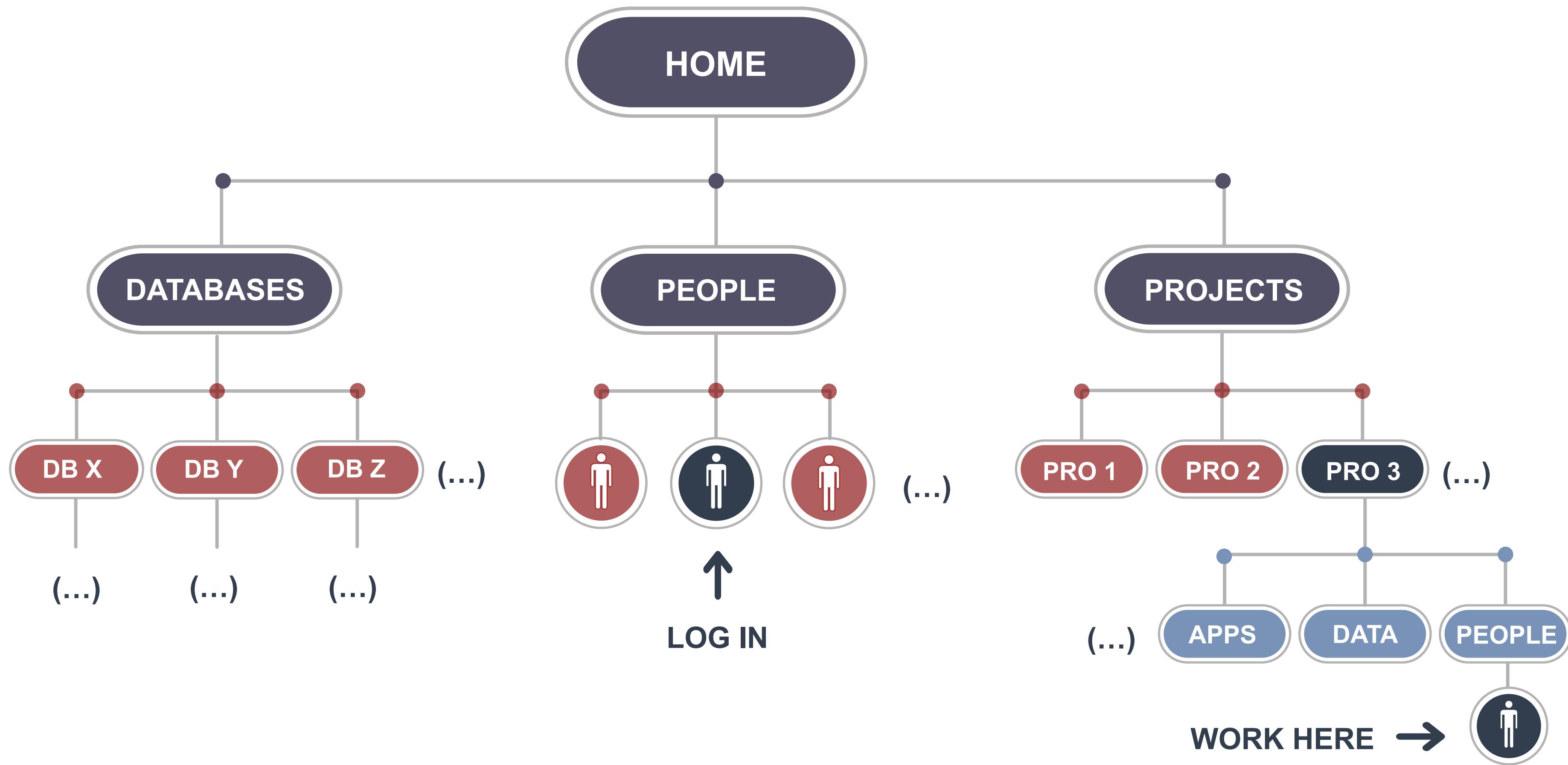


LOG IN WITH SSH

- Open a terminal
- ssh [username@ssh.computerome.dk](ssh://username@ssh.computerome.dk)
- Enter your password
- Enter the token from the Entrust App (or Confirm)

```
Last login: Mon Jan  4 15:31:24 on ttys000
[kgx936@SUN1007442 ~ % ssh thiter@ssh.computerome.dk
#####
#                                     #
#             Welcome to Computerome 2.0      #
#                                     #
#               NOTICE !!!                      #
#               This system requires 2-factor authentication   #
#                                     #
#               To login, type your password and press Enter.   #
#               This sends either an SMS passcode to your mobile   #
#               or a push-notification to Entrust IdentityGuard app.   #
#               If the push-notification times out, a new prompt line   #
#               asks for the 8 digit security code shown in the app.   #
#                                     #
#               BY LOGGING INTO THIS SYSTEM YOU ACCEPT          #
#               COMPUTEROME'S TERMS AND CONDITIONS OF USE.       #
#                                     #
#               Send support requests to computerome@dtu.dk      #
#                                     #
#####
[Password:
Last login: Mon Jan  4 13:47:55 2021 from 2.111.76.128
[[thiter@g-12-10002 ~]$ pwd
/home/people/thiter
[thiter@g-12-10002 ~]$ ]
```

C2 SANDBOX STRUCTURE

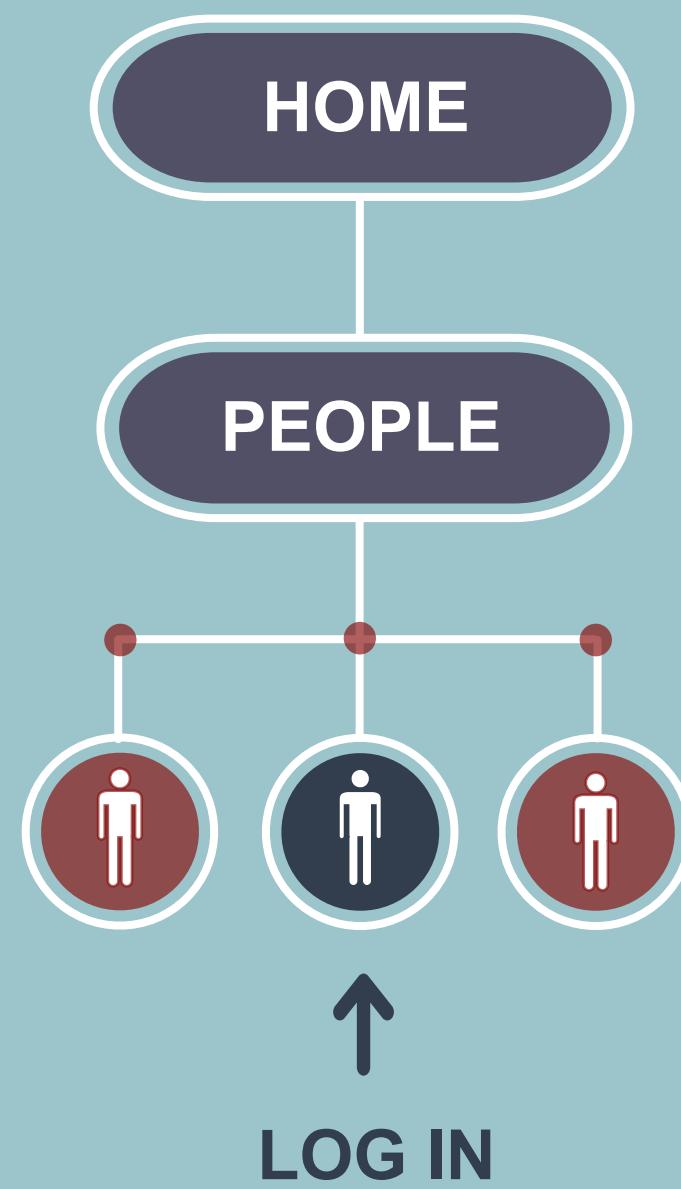


C2 SANDBOX STRUCTURE

1

/home/people/username

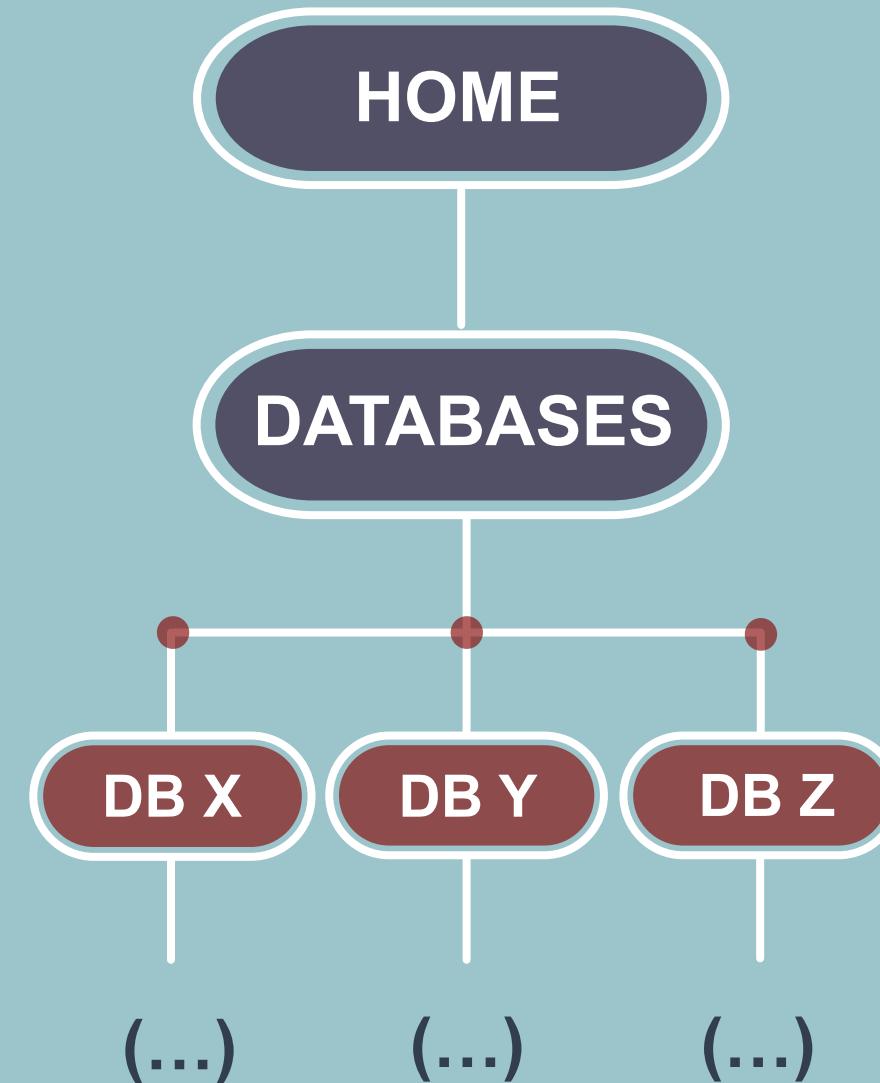
- Your private directory.
- This space is meant for configuration files.
- Do NOT store data here (max 10G)
- Do NOT run jobs from here.



2

/home/databases

- Shared database resource
- Databases include:
 - Human genomes
 - Blast
 - TCGA / GTEx
 - dbSNP
 - ClinVar / COSMIC
 - ...

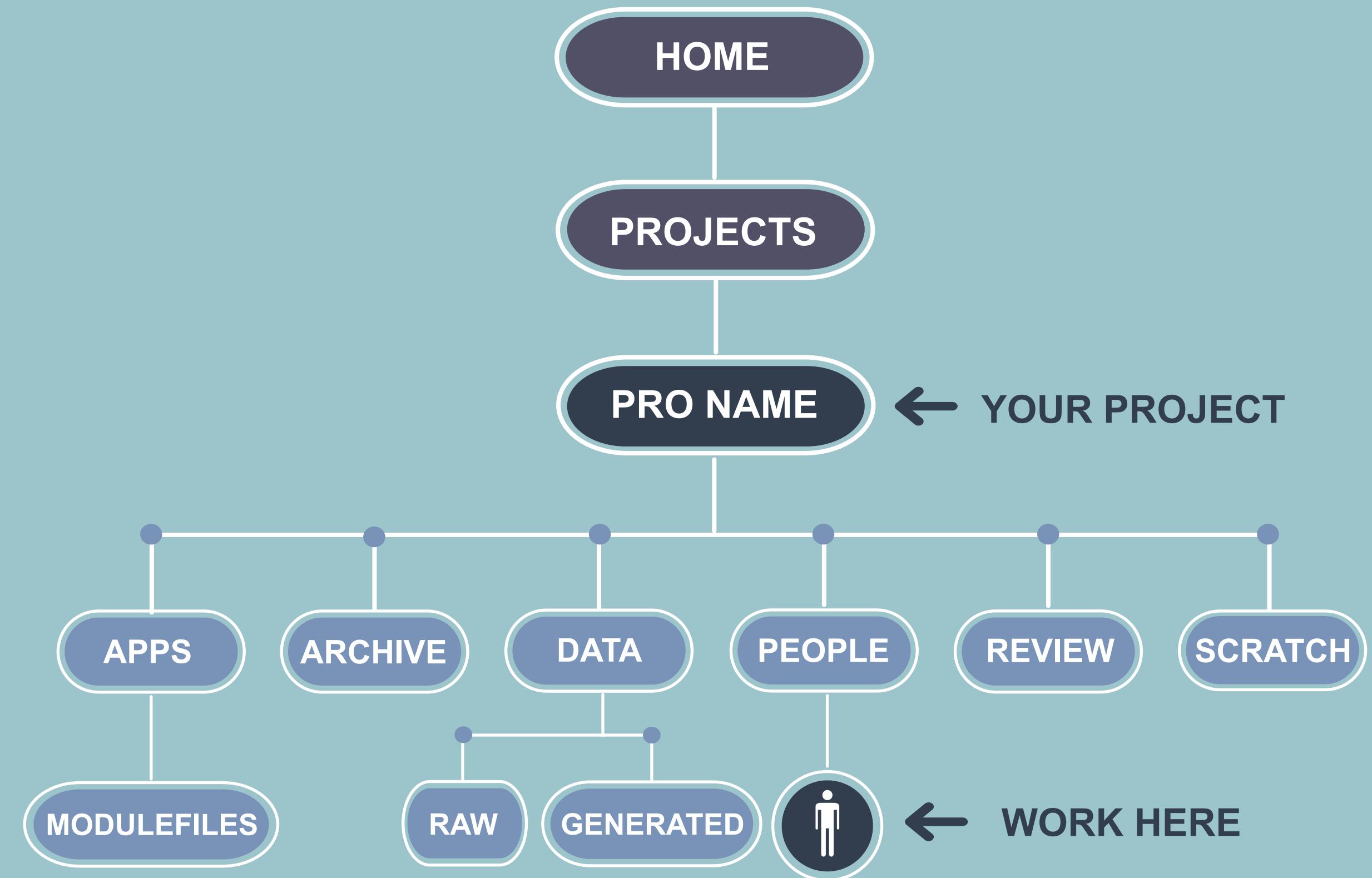


C2 SANDBOX STRUCTURE

3

/home/projects/project

- Your project directory.
- ./apps : project specific applications & modulefiles
- ./archive: dataproject specific data
- ./scratch: temporary & re-creatable data N.B will not be backed up!
- ./review used for peer-review.



YOUR PROJECT AND GROUP

PROJECT

<project> usually refers to the projects funding.



USERNAME

<username>
Your unique user ID. Can be connected to multiple projects and groups.

PROJECT NAME

<project_name> is the name given by the C2 team. Almost exclusively the same as <project>.

GROUP NAME

<group_name>
ID assigned by DTU.
Refers to the UNIX group, e.g. the projects/users access to resources.
Often the same as <project_name>.

Type in <id> to get this information

```
#####
#           Welcome to Computerome 2.0
#
#           NOTICE !!!
#           This system requires 2-factor authentication
#
#           To login, type your password and press Enter.
#           This sends either an SMS passcode to your mobile
#           or a push-notification to Entrust IdentityGuard app.
#           If the push-notification times out, a new prompt line
#           asks for the 8 digit security code shown in the app.
#
#           BY LOGGING INTO THIS SYSTEM YOU ACCEPT
#           COMPUTEROME'S TERMS AND CONDITIONS OF USE.
#
#           Send support requests to computerome@dtu.dk
#
#####
Password:
Last login: Wed Jan  6 13:04:54 2021 from 2.111.76.128
[[thiter@g-12-10002 ~]$]
[[thiter@g-12-10002 ~]$ id
uid=113800351(thiter) gid=113800351(thiter) groups=113800351(thiter),113800352(ku_00033)
[thiter@g-12-10002 ~]$ ]
```

GOOD TO KNOW

COMPUTEROME 2.0 WIKI - FIND EVERYTHING HERE

<https://www.computerome.dk/>

VIDEO TUTORIALS - HELPFUL TO WINDOWS USERS

<https://www.computerome.dk/display/C2W/Video+Tutorials>

BRUSH UP ON YOUR UNIX SKILLS

[https://www.computerome.dk/display/C2W/Getting+Started++new+users?
preview=/42205237/42205239/Hackinars_in_Computerome-wiki.pdf](https://www.computerome.dk/display/C2W/Getting+Started++new+users?preview=/42205237/42205239/Hackinars_in_Computerome-wiki.pdf)

CHECK YOUR USAGE ON C2

module load tools usage_script/2.0
usage -u username

GETTING HELP

Write to computerome@dtu.dk

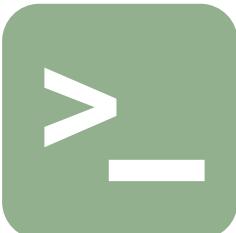
- Subject description
- System, e.g. Computerome main, sandbox, or <named> cloud system
- Login issues, ThinLinc vs SSH login, username & error message(s)
- Command line(s) used
- Describe the problem including error messages

PART3

WORKING ON COMPUTEROME 2.0

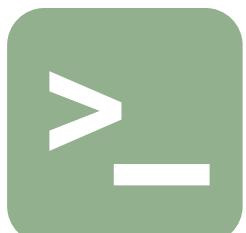


MODULES



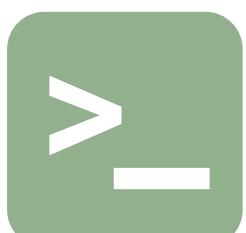
LOADING A MODULE: MODULE LOAD

- You need to load each software you want to use (in its specific version)
- module load bowtie/2.3.2



SEARCHING FOR A MODULE: MODULE AVAIL

- First load the "tools" module: module load tools
- Use the avail function with the start of your software name
- module avail bowtie



ADDING THE DEPENDENCIES

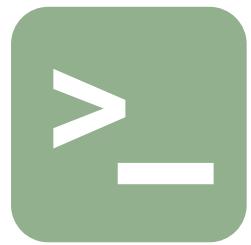
- When there is a missing dependency, follow the suggestions

```
[magmic@g-12-10002 ~]$ module avail bowtie
----- /cm/local/.modulefiles_cache/tools/modulefiles ---
bowtie/1.1.1    bowtie2/2.2.4  bowtie2/2.3.2
bowtie/1.2.1.1  bowtie2/2.2.5  bowtie2/2.3.4.1
[magmic@g-12-10002 ~]$ ]
```

```
[[magmic@g-12-10002 ~]$ module load R/4.0.3
Loading R/4.0.3
ERROR: R/4.0.3 cannot be loaded due to missing prereq.
      HINT: the following module must be loaded first: gcc
[[magmic@g-12-10002 ~]$ module load gcc R/4.0.3
Loading R/4.0.3
ERROR: R/4.0.3 cannot be loaded due to missing prereq.
      HINT: the following module must be loaded first:
            intel/perflibs
[[magmic@g-12-10002 ~]$ module load gcc intel/perflibs R/4.0.3
[magmic@g-12-10002 ~]$ ]
```

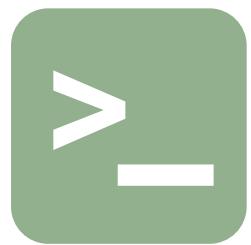
More on C2 wiki: <https://www.computerome.dk/display/C2W/Installed+Software>

REQUEST RESOURCES



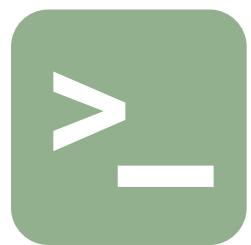
DO NOT START ANY JOB ON THE LOGIN NODE

- After you log-in, you can move around and have a look (e.g. cd, ls, cat....)
- When you know what you want to do, you need to request the resources first



INTERACTIVE MODE WITH IQSUB

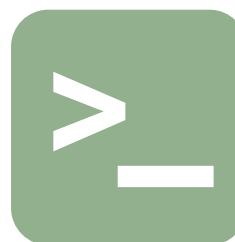
- Just type iqsub in the command line
- This will open a window for you to put the parameters for what you need



PLACE A SCRIPT IN THE QUEUE WITH QSUB

- Specify the resources needed with PBS commands (e.g. in a PBS script)
- Place the script in the queue with qsub
- Explored in mode details this afternoon

IQSUB PARAMETERS



PROJECT / GROUP

- As part of which project / C2 group are you working
- This affects who is paying

Interactive job

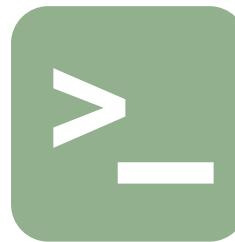
Select group

1	cu_10056
2	danstem
3	cu_10077
4	dan_bri
5	dan_ser
6	dan_gra

↓(+)

26%

< OK > <Cancel>



TIME NEEDED

- This is the wall time. After this time, the resources are close.
- Everything you are doing that is not saved is lost.
- Better to put a bit longer than what you think you need (you can close before the end)
- If you are done, you can exit with command exit ;-)

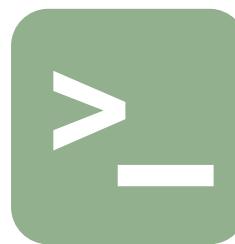
Interactive job

Select time needed (non extendable)

1	1 hour
2	2 hours
3	4 hours
4	8 Hours
5	24 hours
6	Custom

< OK > <Cancel>

IQSUB PARAMETERS

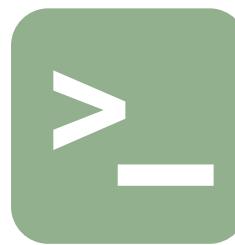


NUMBER OF CPUs

- Only useful for jobs that are parallelized and can benefit from several CPUs
- If you don't know, you can leave 1

Interactive job
Enter number of Processors needed
(1-40)

< OK > <Cancel>

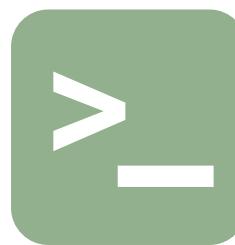


NUMBER OF GPUs

- Unless you know you need some GPU, leave it to 0
- If you want to use GPU, you put 1

Interactive job
Enter number of GPUs needed (0-1)

< OK > <Cancel>



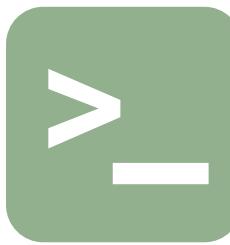
MEMORY NEEDED

- Depends on the methods / software you use
- E.g. mapping with STAR, you might want to increase to 60-80g

Interactive job
Enter amount of memory needed

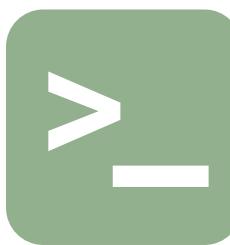
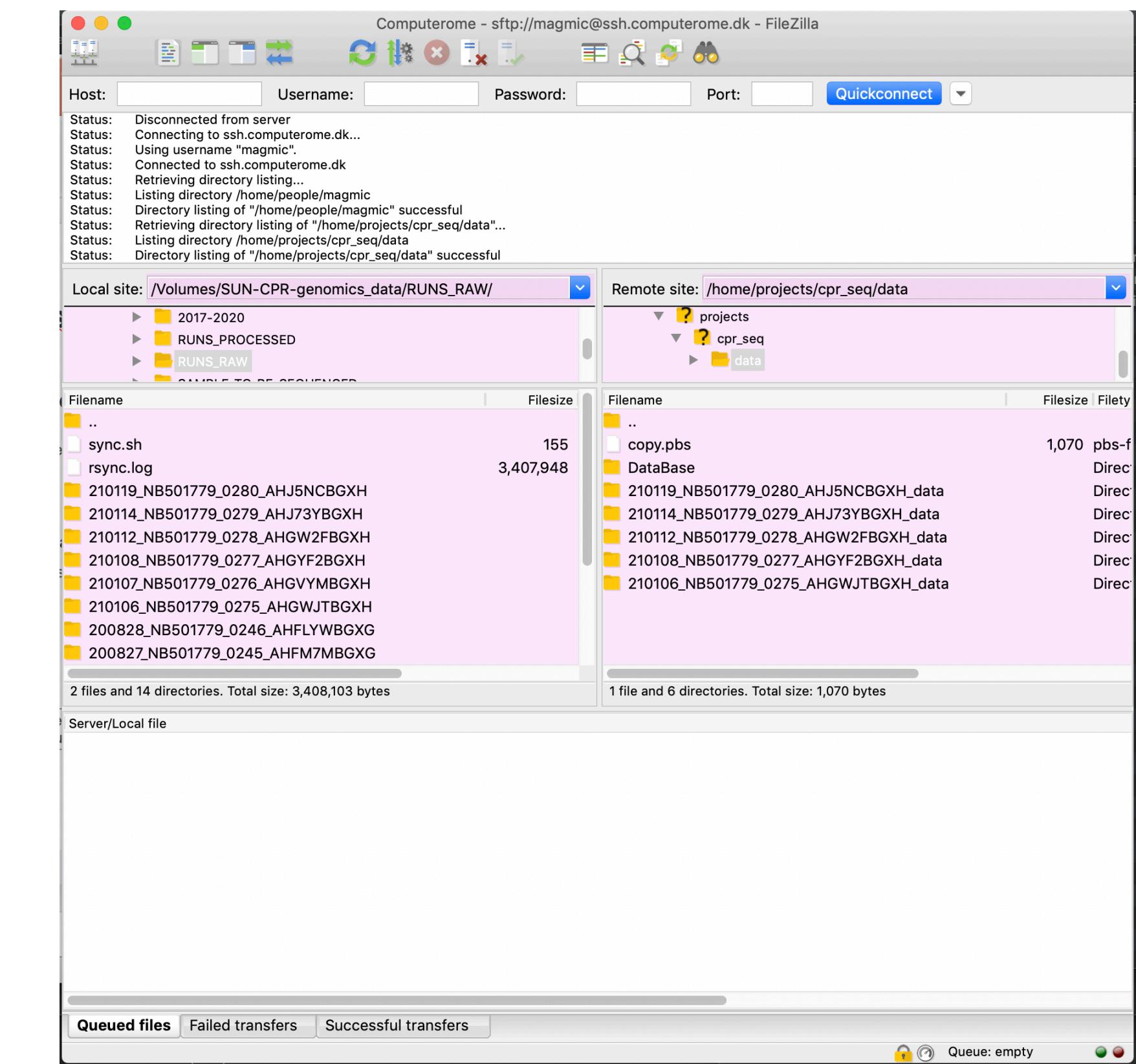
< OK > <Cancel>

FILE TRANSFER: scp & FileZilla



COPY DATA FROM/TO C2 via SCP command

- From local to computerome:
•scp /Users/path/to/file username@ssh.computerome.dk:/home/path/to/the/destination/
- From computerome to local:
•scp username@ssh.computerome.dk:/home/path/to/file/ /Users/path/to/destination/
- If it's a whole folder, use -r
- For more info, study the scp command (man scp)



Use FileZilla graphical interface

- You can also use a software with a graphical interface
- Find on the wiki how to set up the connection

Filezilla configuration

SFTP Client Configurations

In order to use most FTP/SFTP clients on a 2 factor system you will need to limit the maximum simultaneous connections to avoid multiple login prompts.

Here you can see how to setup the mostly used clients

Filezilla

⚠ Changes has been made in resent versions which means that Quickconnect no longer works with 2 factor authentication

GLOBAL

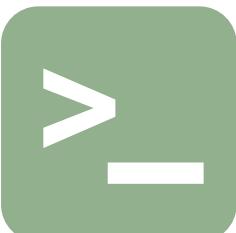
Go to Edit->Settings->Transfers
Change Maximum simultaneous transfers to 1

SITE PROFILE / BOOKMARK

Go to File->Site Manager->(your profile)->Transfer settings
Check Limit number of simultaneous connections
Change Maximum number of connections to 1
Change Login type to Interactive

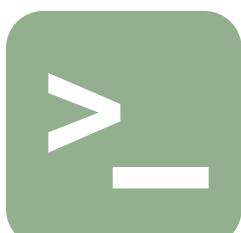
From C2 wiki: <https://www.computerome.dk/display/C2W/SSH+login+to+Computerome+2.0>

MODULES



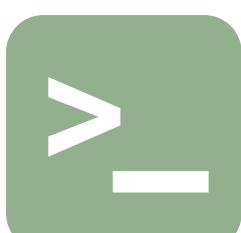
LOADING A MODULE: MODULE LOAD

- You need to load each software you want to use (in its specific version)
- module load bowtie/2.3.2



SEARCHING FOR A MODULE: MODULE AVAIL

- Use the avail function with the start of your software name
- module avail bowtie



ADDING THE DEPENDENCIES

- When there is a missing dependency, follow the suggestions

```
[magmic@g-12-10002 ~]$ module avail bowtie
----- /cm/local/.modulefiles_cache/tools/modulefiles ---
bowtie/1.1.1    bowtie2/2.2.4  bowtie2/2.3.2
bowtie/1.2.1.1  bowtie2/2.2.5  bowtie2/2.3.4.1
[magmic@g-12-10002 ~]$ ]
```

```
[magmic@g-12-10002 ~]$ module load R/4.0.3
Loading R/4.0.3
ERROR: R/4.0.3 cannot be loaded due to missing prereq.
HINT: the following module must be loaded first: gcc
[magmic@g-12-10002 ~]$ module load gcc R/4.0.3
Loading R/4.0.3
ERROR: R/4.0.3 cannot be loaded due to missing prereq.
HINT: the following module must be loaded first:
intel/perflibs
[magmic@g-12-10002 ~]$ module load gcc intel/perflibs R/4.0.3
[magmic@g-12-10002 ~]$ ]
```

More on C2 wiki: <https://www.computerome.dk/display/C2W/Installed+Software>

PART4

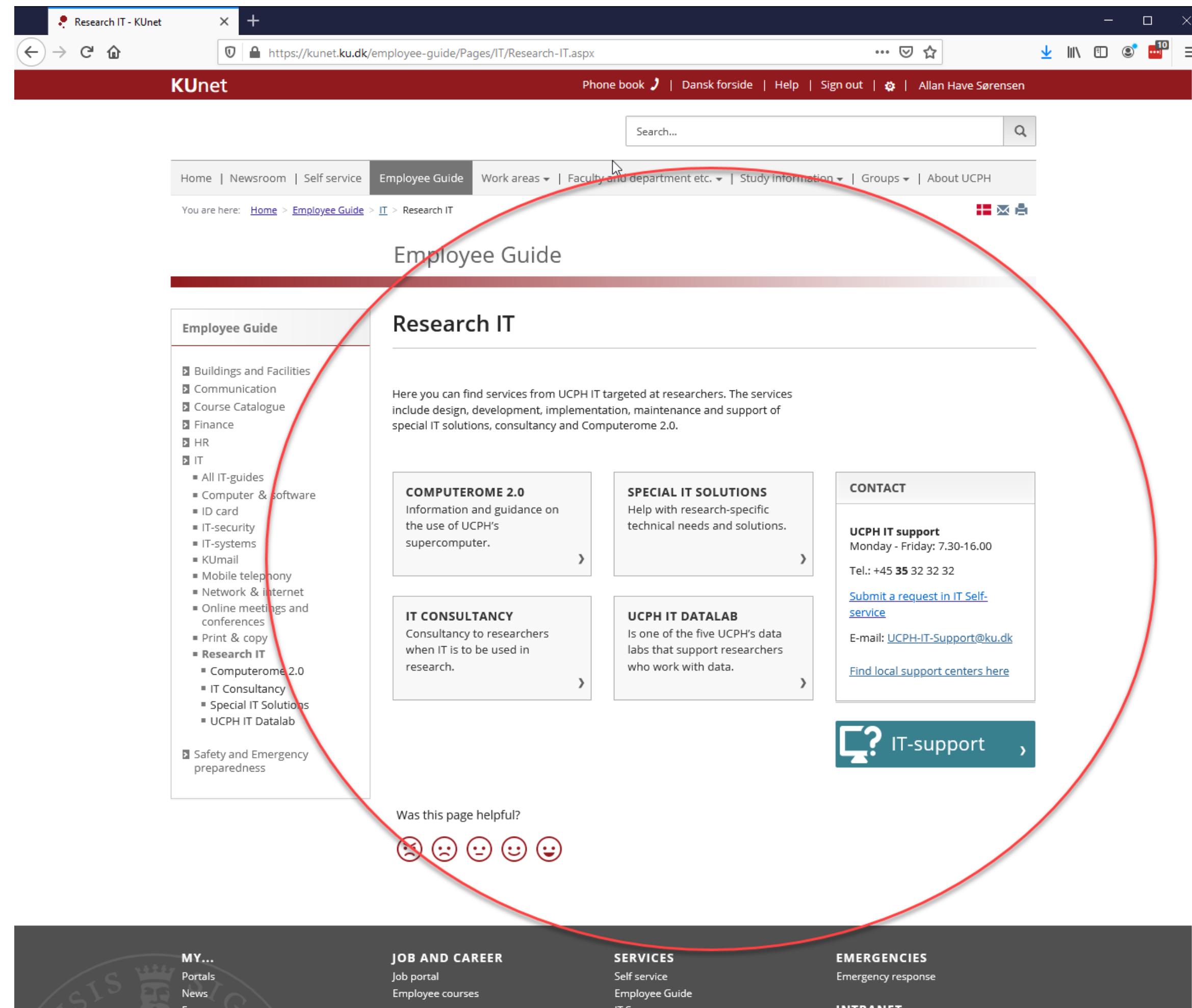
A BRIEF TECHNICAL INTERMEZZO

A BRIEF TECHINCAL INTERMEZZO

HOW DO YOU GET ACCESS OR HELP TO COMPUTEROME AT KU?

- Get access via KU-IT's selfservice at <https://sd.ku.dk>
- If you need help, KU-IT can assist with a basic C2 training
- More on C2 at KU at KUnet, Employment guide and soon Research Portal:

COMPUTEROME AT KUnet



COMPUTEROME AT KUnet

The screenshot shows a web browser window titled "Computerome 2.0 - KUnet". The URL in the address bar is <https://kunet.ku.dk/employee-guide/Pages/IT/Computerome-2.0.aspx>. The page title is "Employee Guide".

The main content area is titled "Computerome 2.0". It contains the following text:

Computerome 2.0 (in short C2) is a Danish high performance computer, a supercomputer, that is managed by the University of Copenhagen (UCPH) and the Technical University of Denmark (DTU). C2 is available to all UCPH users, regardless of their field of research. You have also as an employee or a PhD student at UCPH, the opportunity to become a free-of-charge C2 "sandbox" user, and try C2 for a limited amount of time. The UCPH IT's Research Support section is responsible for supporting the UCPH users of C2. You can contact Research Support via [IT Self-service](#).

On the left side, there is a sidebar titled "Employee Guide" with the following navigation links:

- Buildings and Facilities
- Communication
- Course Catalogue
- Finance
- HR
- IT
 - All IT-guides
 - Computer & software
 - ID card
 - IT-security
 - IT-systems
 - KUmail
 - Mobile telephony
 - Network & internet
 - Online meetings and conferences
 - Print & copy
 - Research IT
 - Computerome 2.0
 - IT Consultancy
 - Special IT Solutions
 - UCPH IT Databab- Safety and Emergency preparedness

The main content area also includes sections for "Access", "Support", "Sandbox", "Wiki", "Prices", "Technology", and "Resource allocation", each with a "+" button to expand.

On the right side, there is a "CONTACT" section with the following text:

You can contact Computerome 2.0 by submit a request to UCPH IT Research Support section via: [IT Self-service](#)

Read more under Access and Support.

There is also a "LOCATION" section with the following text:

Computerome 2.0 is physically located at DTU's Risø Campus, at Roskilde municipality, in water-cooled containers from where the generated excess heat is transferred into the municipality's heating network.

At the bottom right, there is a teal button labeled "IT-support" with a computer monitor icon.

The footer of the page contains the following text:

FOLD ALL OUT

Is C2 devoted only to Life Sciences? +

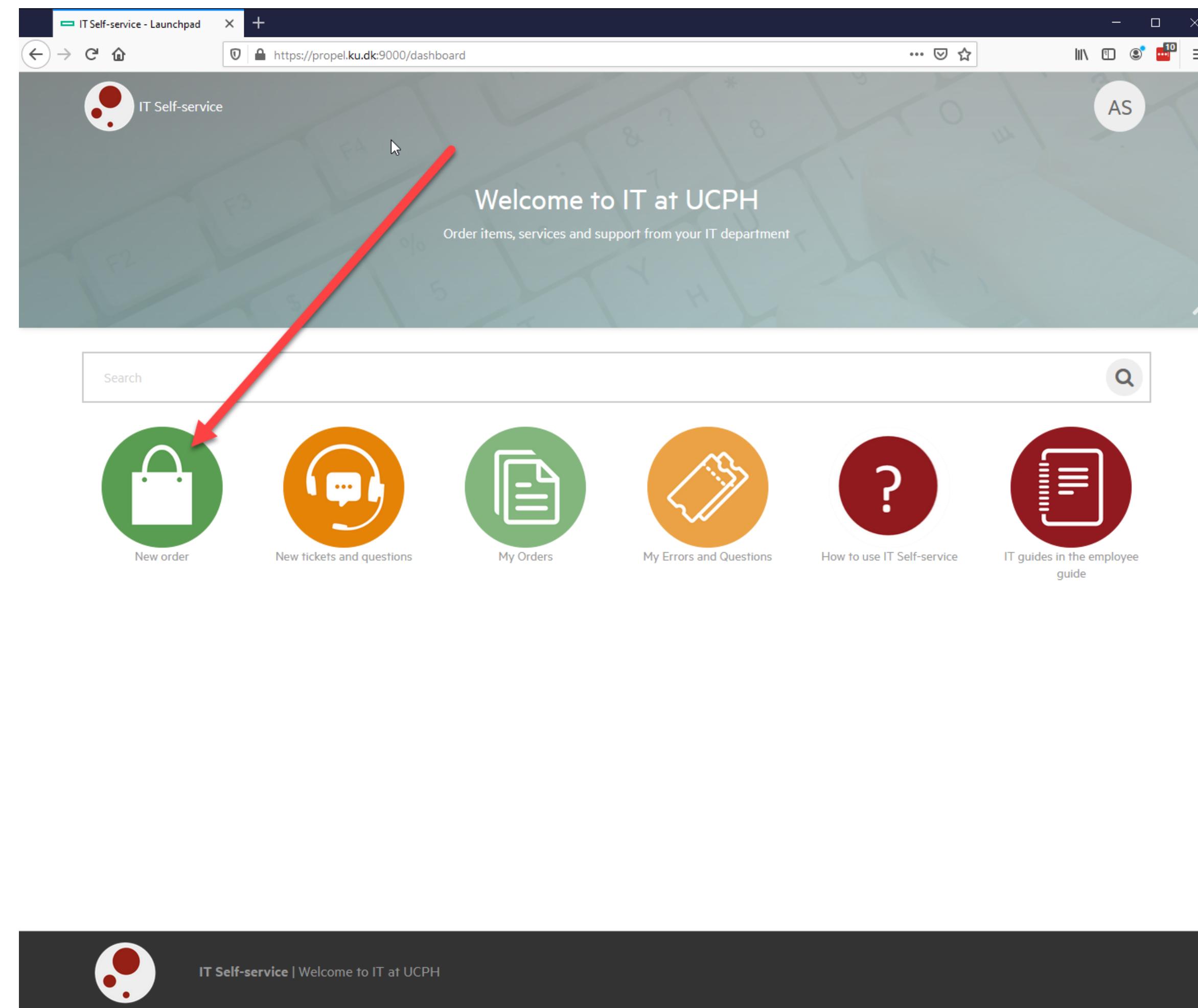
How do I communicate with C2? +

If I own a group at C2, and a new member has to be added, who should order the new membership? +

Can collaborators who are not employed at UCPH, for example collaborators who work outside Denmark, get an account on C2? +

How long does it take for the creation of a new account at C2? +

HOW TO CREATE A GROUP



HOW TO CREATE A GROUP

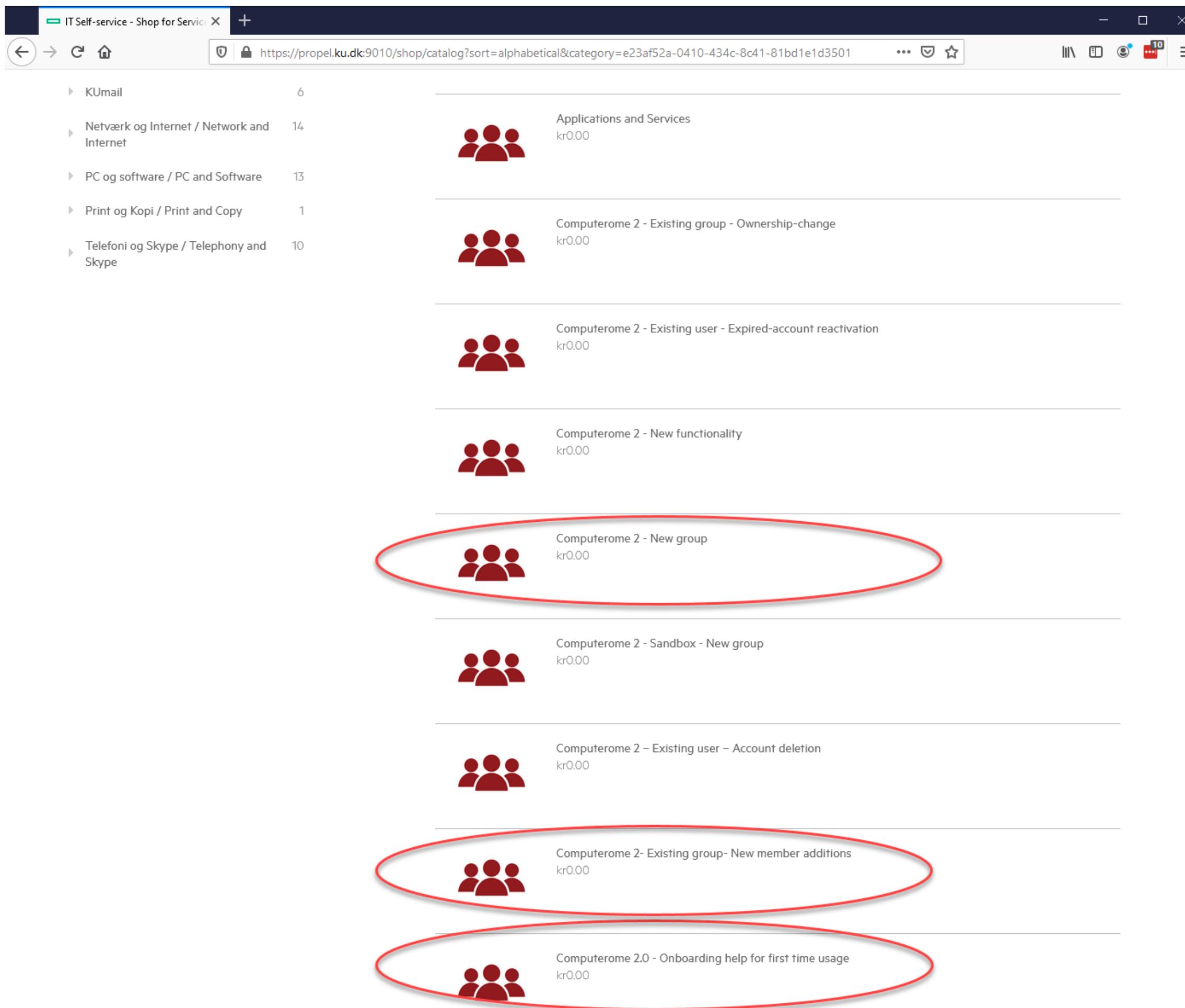
The screenshot shows a web browser window titled "IT Self-service - Shop for Services". The URL is <https://propel.ku.dk:9010/shop/catalog?sort=alphabetical>. The page displays a catalog of items, with a search bar at the top. On the left, there's a sidebar with icons for "Orders" and "AS". The main content area shows a list of categories under "Refine Results". One category, "Forsknings-IT / Research-IT", is circled in red. The list includes:

Category	Count
Andet / Other	6
Forsknings-IT / Research-IT	10
IT- og fagsystemer / IT-systems	30
KUmail	6
Netværk og Internet / Network and Internet	14
PC og software / PC and Software	13
Print og Kopi / Print and Copy	1
Telefoni og Skype / Telephony and Skype	10

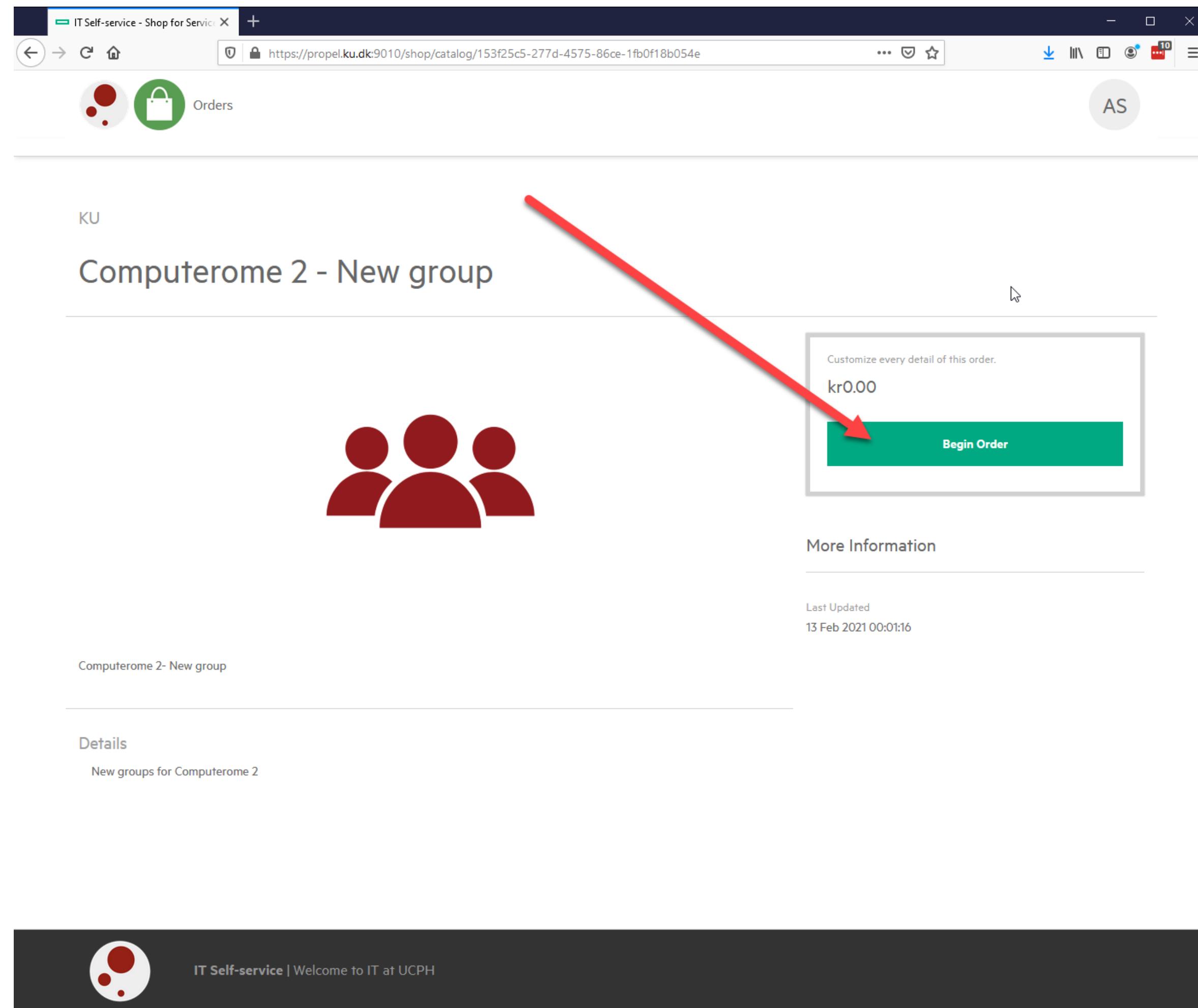
On the right, there's a list of items with icons and descriptions:

- Access to national HPC resources kr.00
- Accessories (Computer) kr.00
- Accessories (Telephony) kr.00
- Adgang til data i KU BI Tableau kr.00
- Aktivitetskontrol kr.00
- Applications and Services kr.00
- Autorisationsansvarlig kr.00

HOW TO CREATE A GROUP



HOW TO CREATE A GROUP



HOW TO CREATE A GROUP

The screenshot shows a web browser window titled "IT Self-service - Shop for Services" with the URL <https://propel.ku.dk:9010/shop/catalog/153f25c5-277d-4575-86ce-1fb0f18b054>. The page is titled "Computerome 2 - New group".

The form is divided into two main sections:

- Faculty**:
 - Name of faculty *: A dropdown menu labeled "Select a value".
 - Institute/center *: An input field.
 - Section *: An input field.
 - Paying Unit Code "stedkode" (8 digits, only numbers) *: An input field.
 - Alias (10 digits, only numbers) *: An input field.
 - Zip code and city *: An input field.
- Group owner**:
 - Name and KU username of group owner *: A dropdown menu labeled "Select a value".
 - Email address *: An input field.
 - Cell phone number *: An input field.

To the right of the form, there is a summary section:

- Subtotal**: kr0.00
- Order Now**: A green button.
- Configuration**: None

A large red arrow points from the top left towards the "Order Now" button.

PART5

C2 MODULES & QUEUING SYSTEM

MODULES, QUEUING SYSTEM & SCRIPTS - SHYAM

PART6

OPTIMIZATION OF TIME & COST

OPTIMIZATION OF TIME AND COST ON C2 - SHYAM