

Connecting to the Amazon instance

**Find your Amazon instance
IP address**

EVOLUTION AND GENOMICS

Intensive and immersive training opportunities



2020 WORKSHOP ON POPULATION AND SPECIATION GENOMICS, CESKY KRUMLOV



The 2020 Workshop on Population and Speciation Genomics will be held in [Český Krumlov](#) from 19 January – 1 February. It brings together an international collection of faculty members and Workshop participants to study and discuss current ideas and techniques for the analysis of genomic data on the level of populations and closely related species. The Workshop consists of a series of lectures, demonstrations and computer laboratories that cover theory and practice of population and speciation genomics analyses.

The Workshop registration fee is \$1,800. This amount does not cover travel, lodging or boarding. Many important



The 2020 Workshop on Population and Speciation Genomics will be held in [Český Krumlov](#) from 19 January – 1 February. It brings together an international collection of faculty members and Workshop participants to study and discuss current ideas and techniques for the analysis of genomic data on the level of populations and closely related species. The Workshop consists of a series of lectures, demonstrations and computer laboratories that cover theory and practice of population and speciation genomics analyses.

The Workshop registration fee is \$1,800. This amount does not cover travel, lodging or boarding. Many important details about the Workshop series can be read on our [FAQ](#) page.

The Workshop on Population and Speciation Genomics is organized and co-directed by:

[Walter Salzburger](#), University of Basel, Switzerland

[Michael Matschiner](#), University of Zurich, Switzerland

[Jan Stefka](#), University of South Bohemia České Budějovice, Czech Republic

[Scott Handley](#), Washington University, US

[Emiliano Trucchi](#), Università Politecnica delle Marche, Ancona, Italy

[Julia M.I. Barth](#), University of Basel, Switzerland

[Britta S. Meyer](#), Max Planck Institute for Evolutionary Biology, Plön, Germany

[Milan Malinsky](#), University of Basel, Switzerland

ANNOUNCEMENTS

We are looking forward to meet everybody at the reception dinner on Sunday at 6 p.m. at the hotel Zlatý Anděl!

LINKS

- [Lecture slides and learning activities](#)
- [Map of important locations in Český Krumlov](#)
- [Amazon instances IP addresses](#)

docs.google.com

2020 Workshop on Population and Speciation Genomics, Cesky Krumlov - Evolution a... WPSG2020 AMI IP addresses - Google Sheets

WPSG2020 AMI IP addresses

File Edit View Insert Format Data Tools Add-ons Help

100% \$ % .0 .00 123 arial 10 B I S A

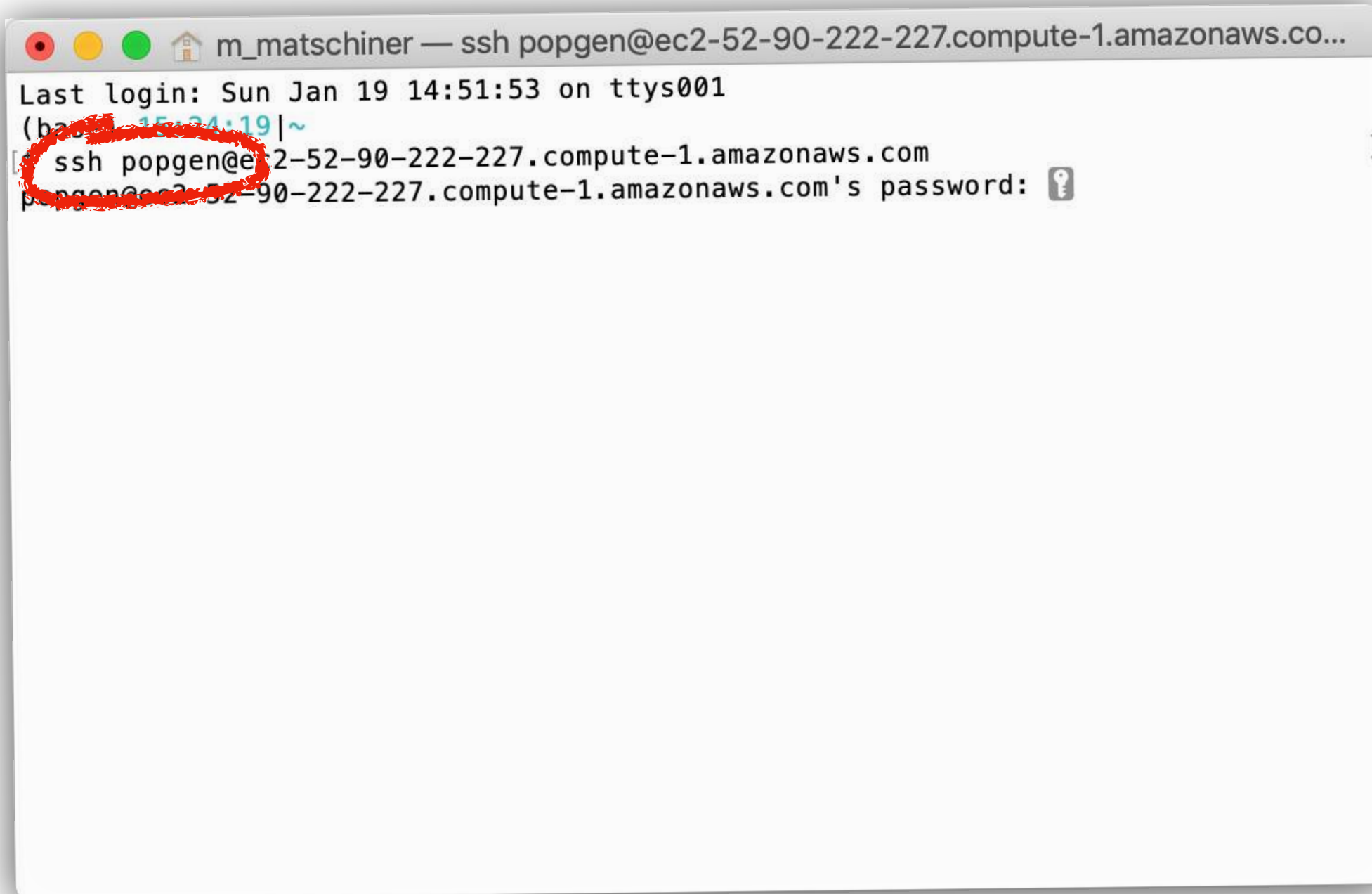
	Name	Instance Type	Launch Time	Status Checks	Alarm Status	Public DNS
1	Name	Instance Type	Launch Time	Status Checks	Alarm Status	Public DNS
29	b5 Angela	t3a.medium	running	2/2 checks passed	None	ec2-18-209-213-36.compute-1.amazonaws.com
30	b5 Angelica	t3a.medium	running	2/2 checks passed	None	ec2-3-234-228-107.compute-1.amazonaws.com
31	b5 Emiliano	t3a.medium	running	2/2 checks passed	None	ec2-54-236-59-103.compute-1.amazonaws.com
32	b5 Jan	t3a.medium	running	2/2 checks passed	None	ec2-34-236-255-182.compute-1.amazonaws.com
33	b5 Jule	t3a.medium	running	2/2 checks passed	None	ec2-3-234-228-107.compute-1.amazonaws.com
34	b5 Micha	t3a.medium	running	2/2 checks passed	None	ec2-18-215-188-156.compute-1.amazonaws.com
35	b5 Walter	t3a.medium	running	2/2 checks passed	None	ec2-35-172-169-64.compute-1.amazonaws.com
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						

Connecting to the Amazon instance

- 1. SSH / SCP**
- 2. Guacamole**
- 3. RStudio**
- 4. Jupyter Notebooks**

1. SSH / SCP

1. SSH / SCP

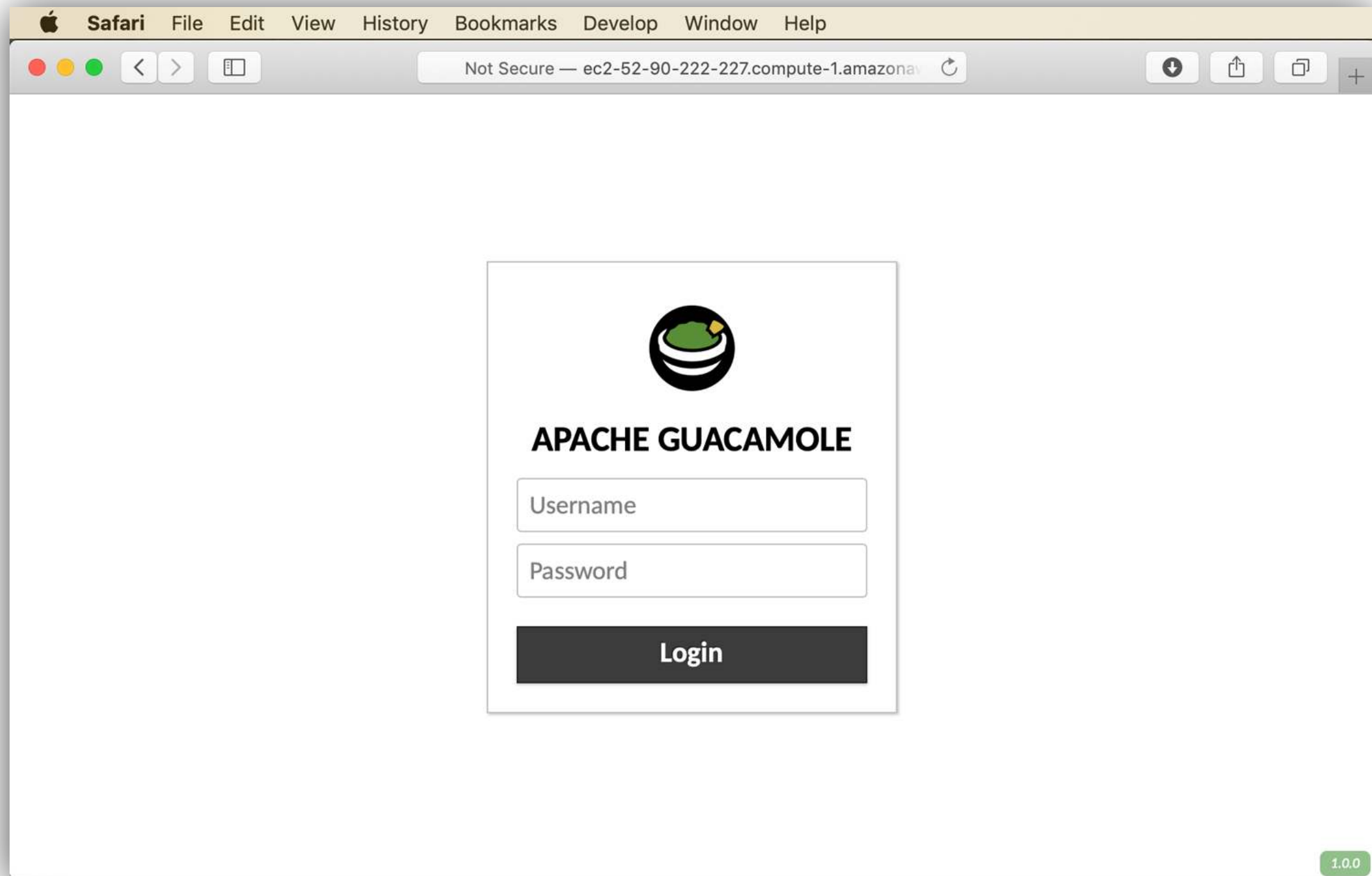


A terminal window titled "m_matschiner — ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.co...". The window shows the output of an SSH command. The first line is "Last login: Sun Jan 19 14:51:53 on ttys001". The second line is "(base) 15:24:19 | ~". The third line is "ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com", where the IP address "ec2-52-90-222-227" is circled in red. The fourth line is "popgen@ec2-52-90-222-227.compute-1.amazonaws.com's password: " followed by a key icon.

```
m_matschiner — ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.co...  
Last login: Sun Jan 19 14:51:53 on ttys001  
(base) 15:24:19 | ~  
ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com  
popgen@ec2-52-90-222-227.compute-1.amazonaws.com's password: 
```

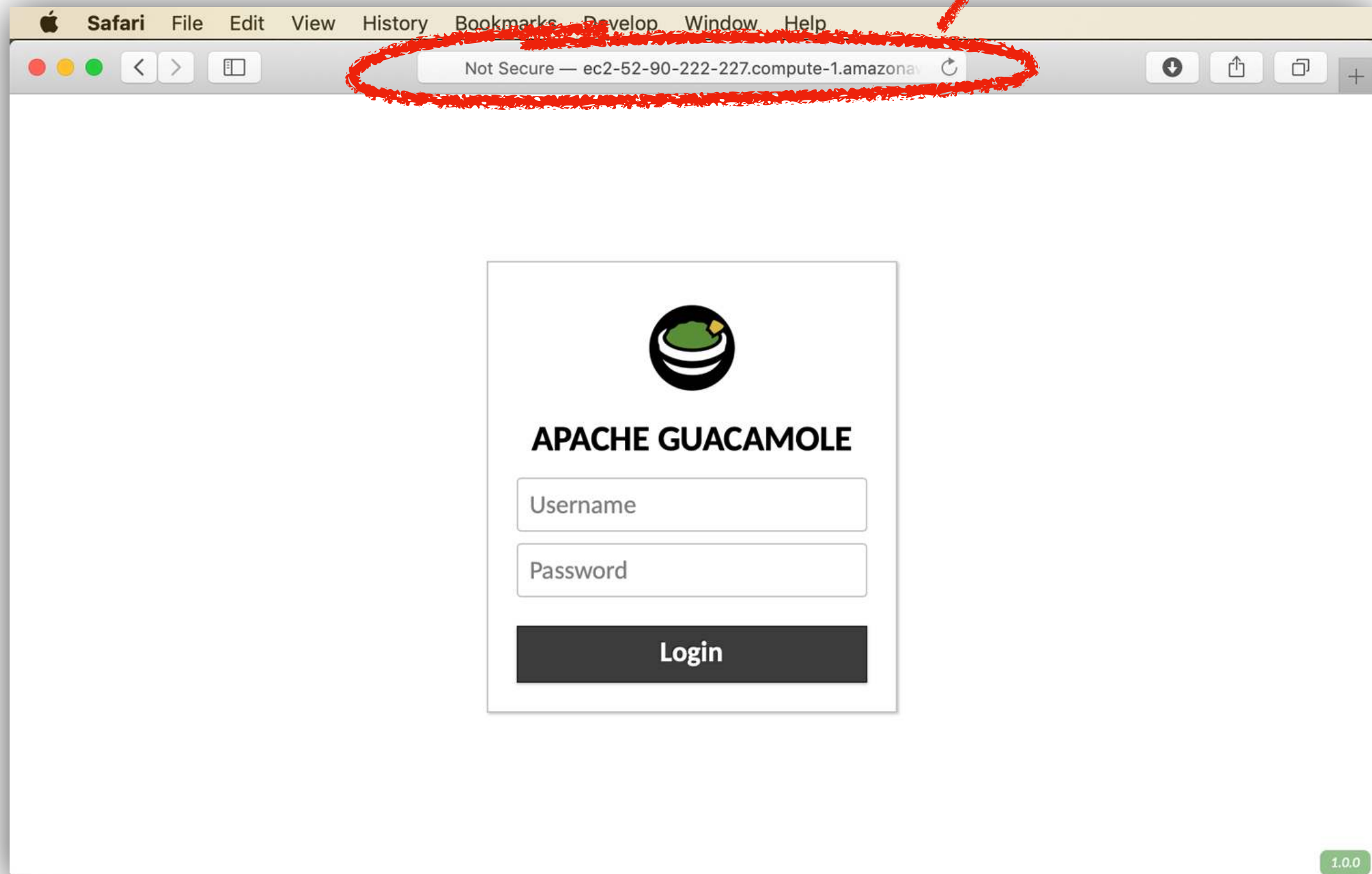
2. Quacamole

2. Guacamole

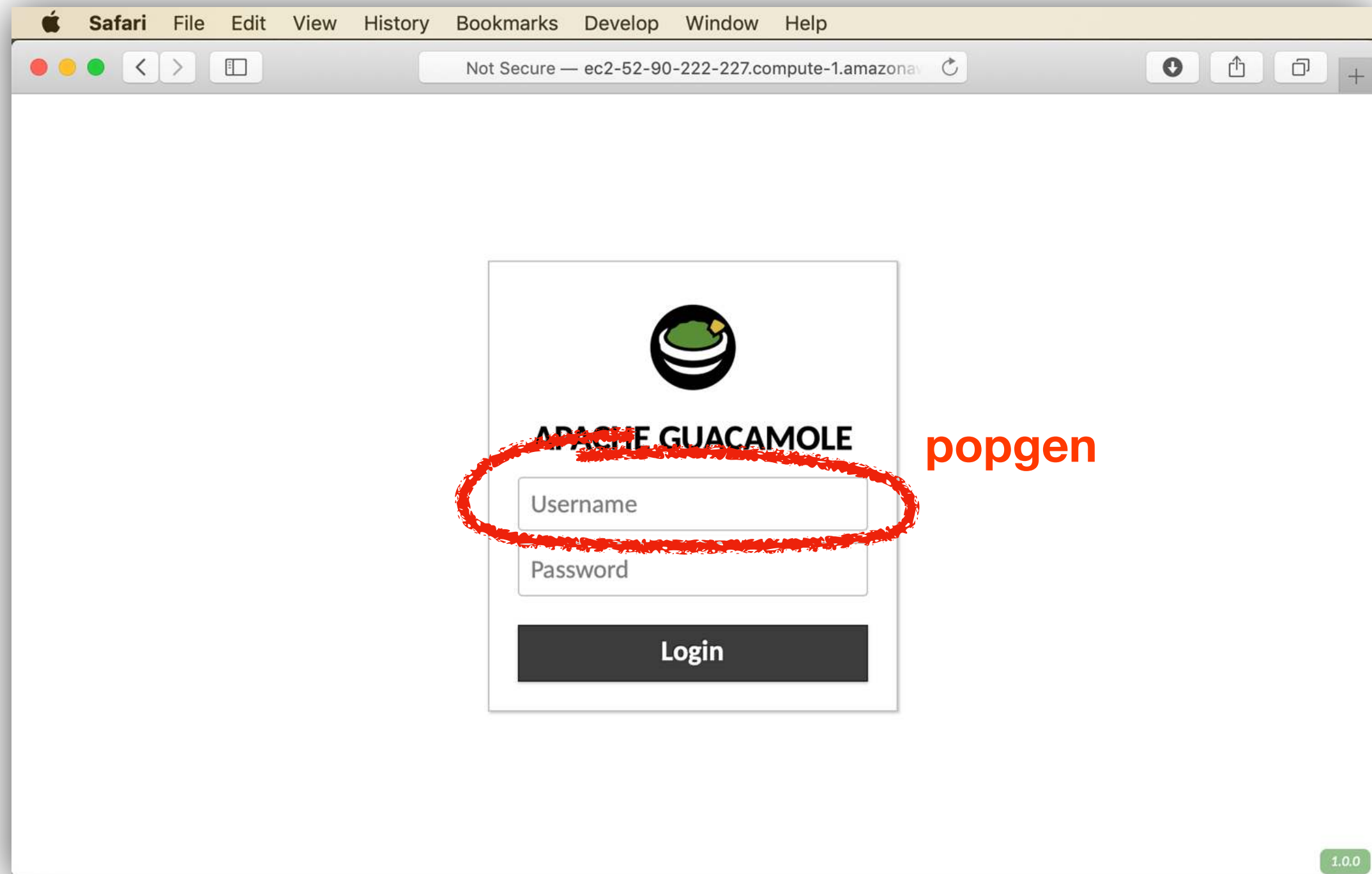


2. Guacamole

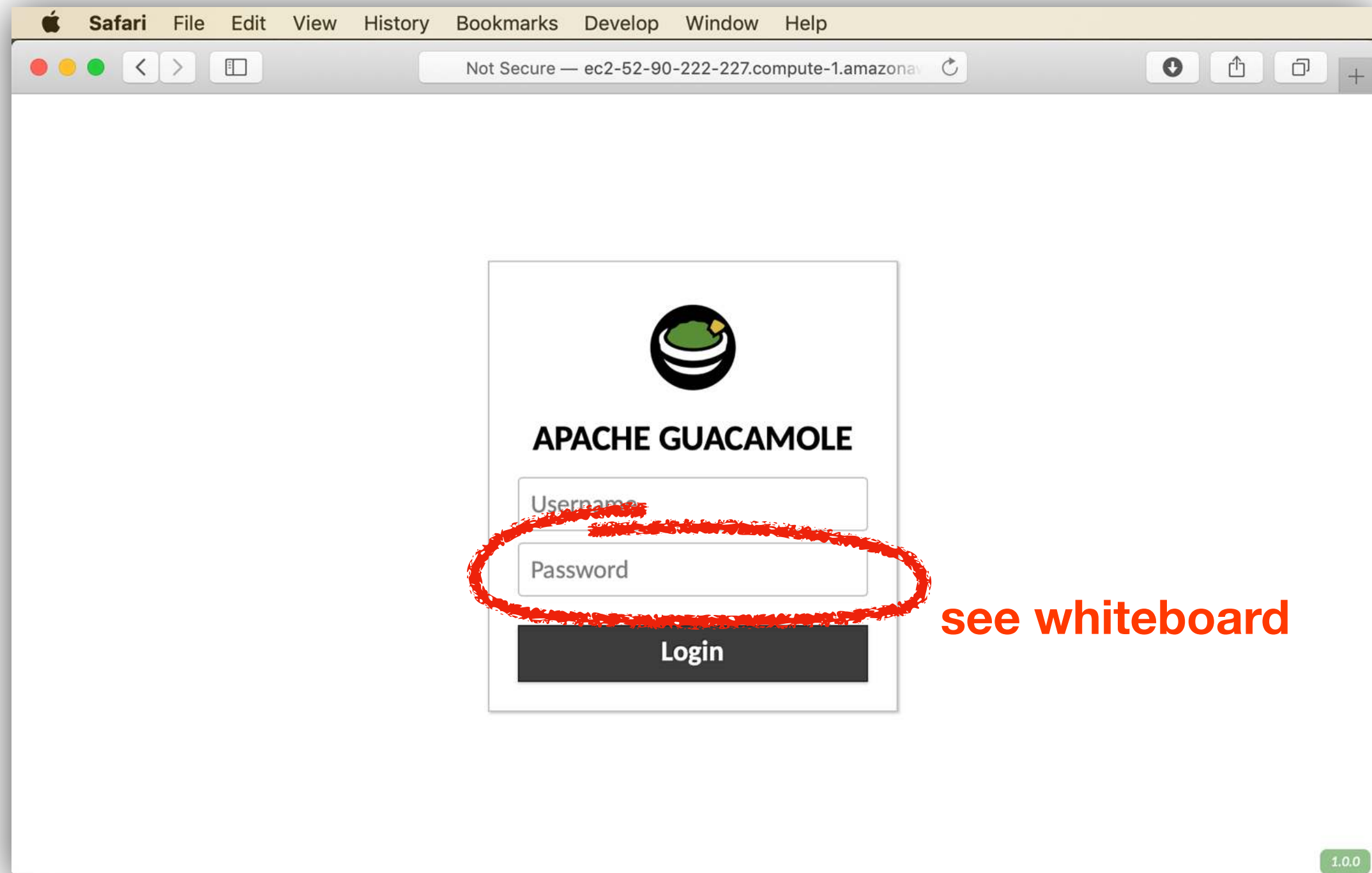
...amazonaws.com:8080/guacamole



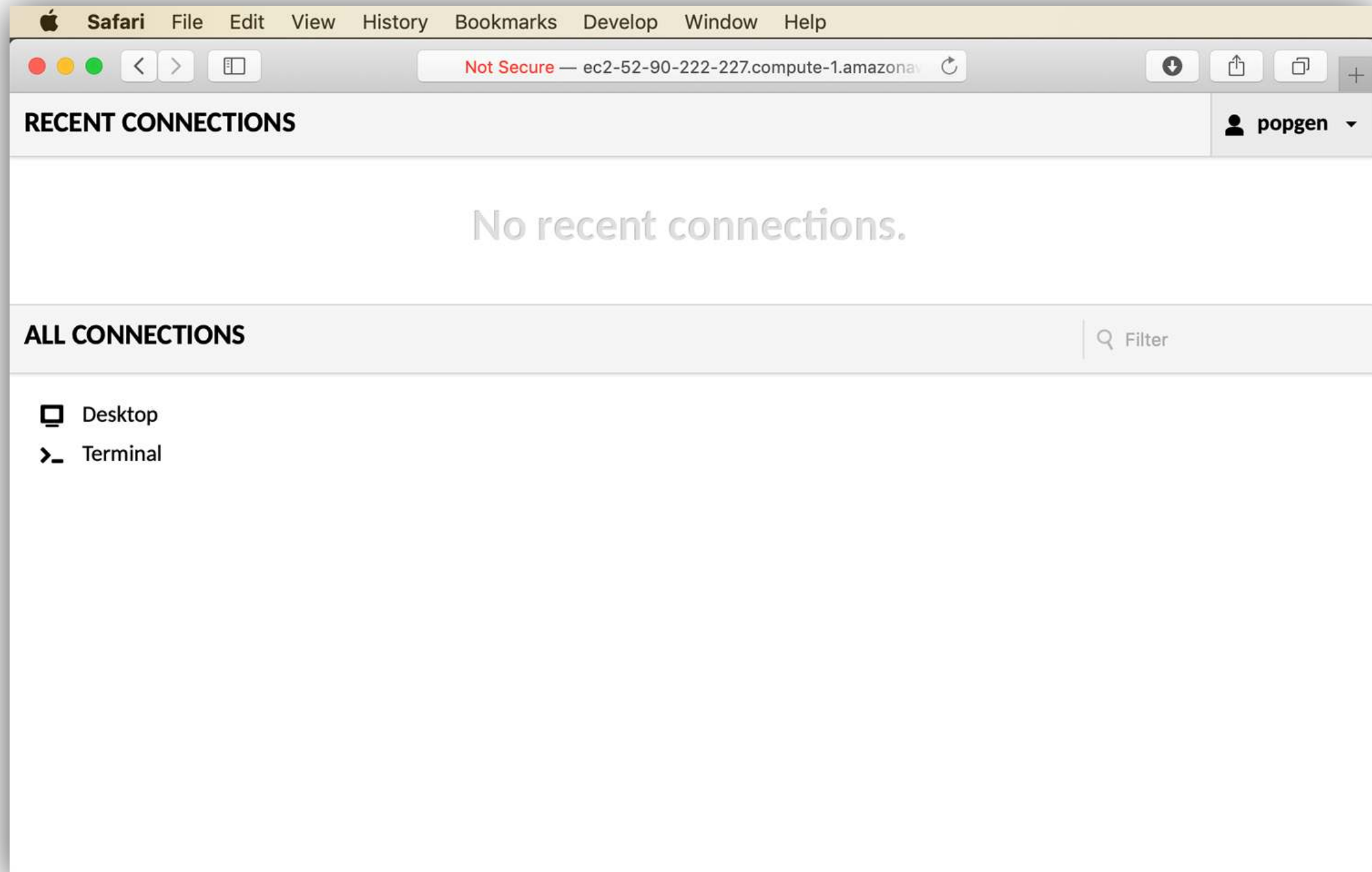
2. Guacamole



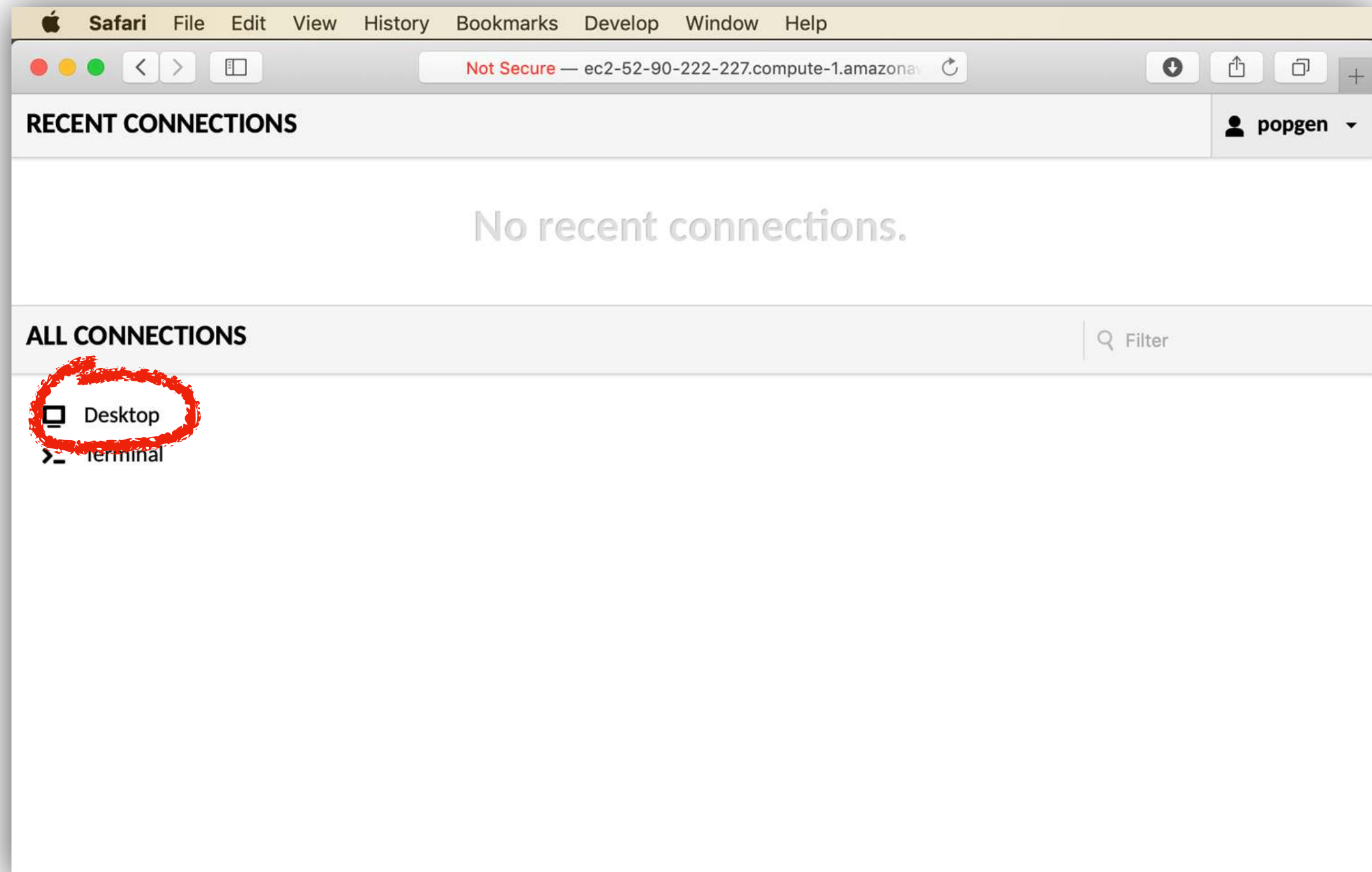
2. Guacamole



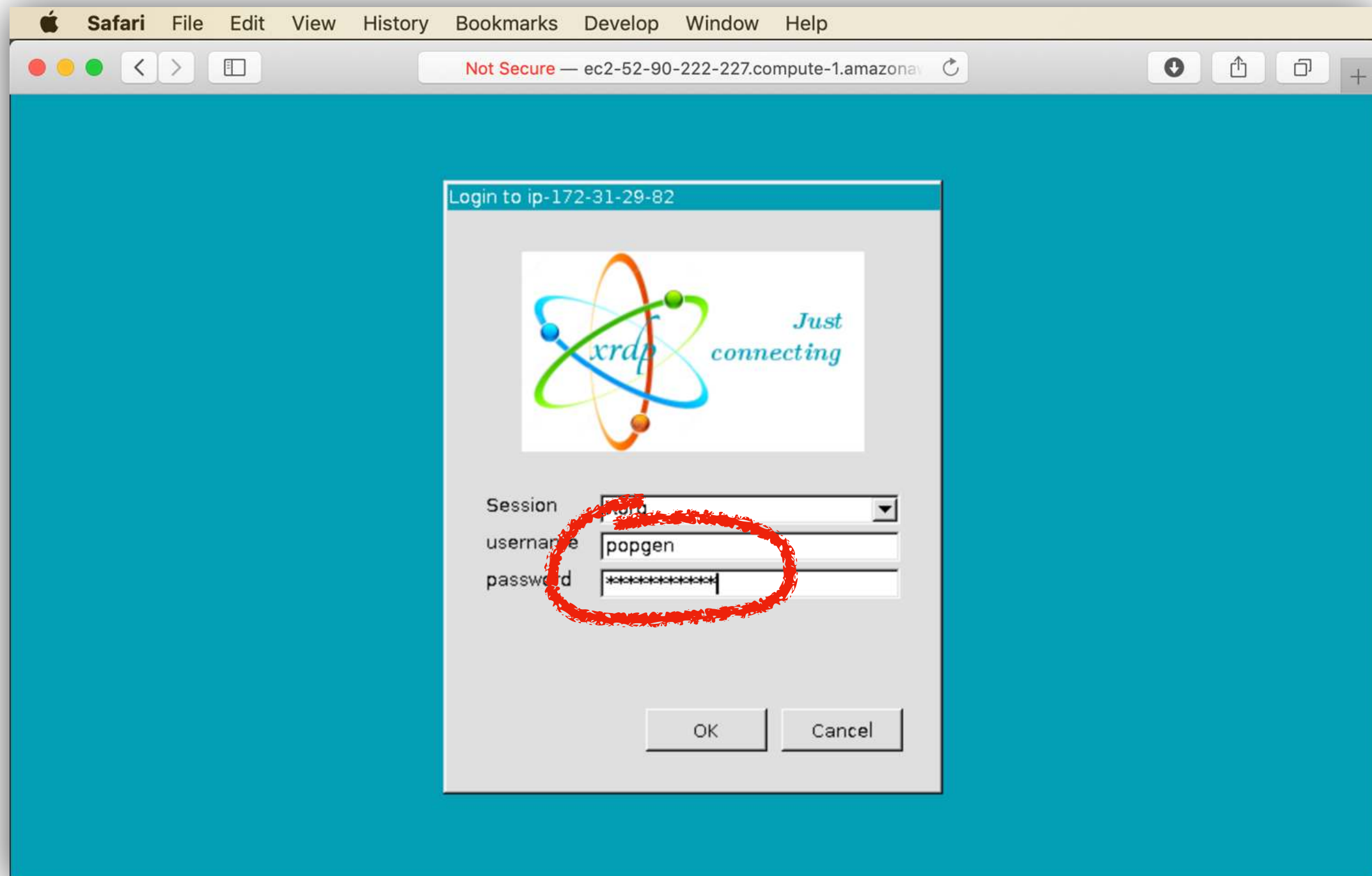
2. Guacamole



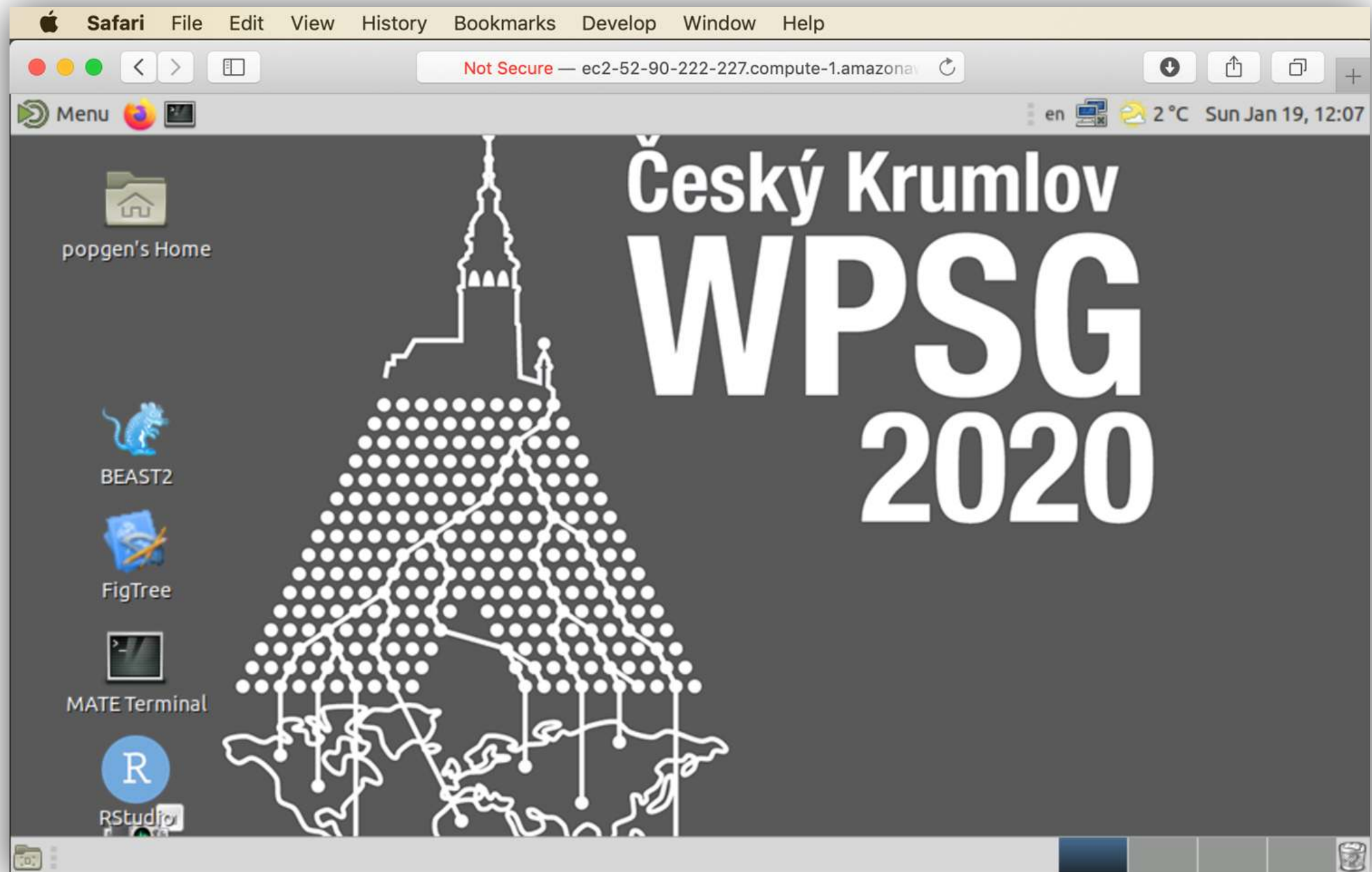
2. Guacamole



2. Guacamole



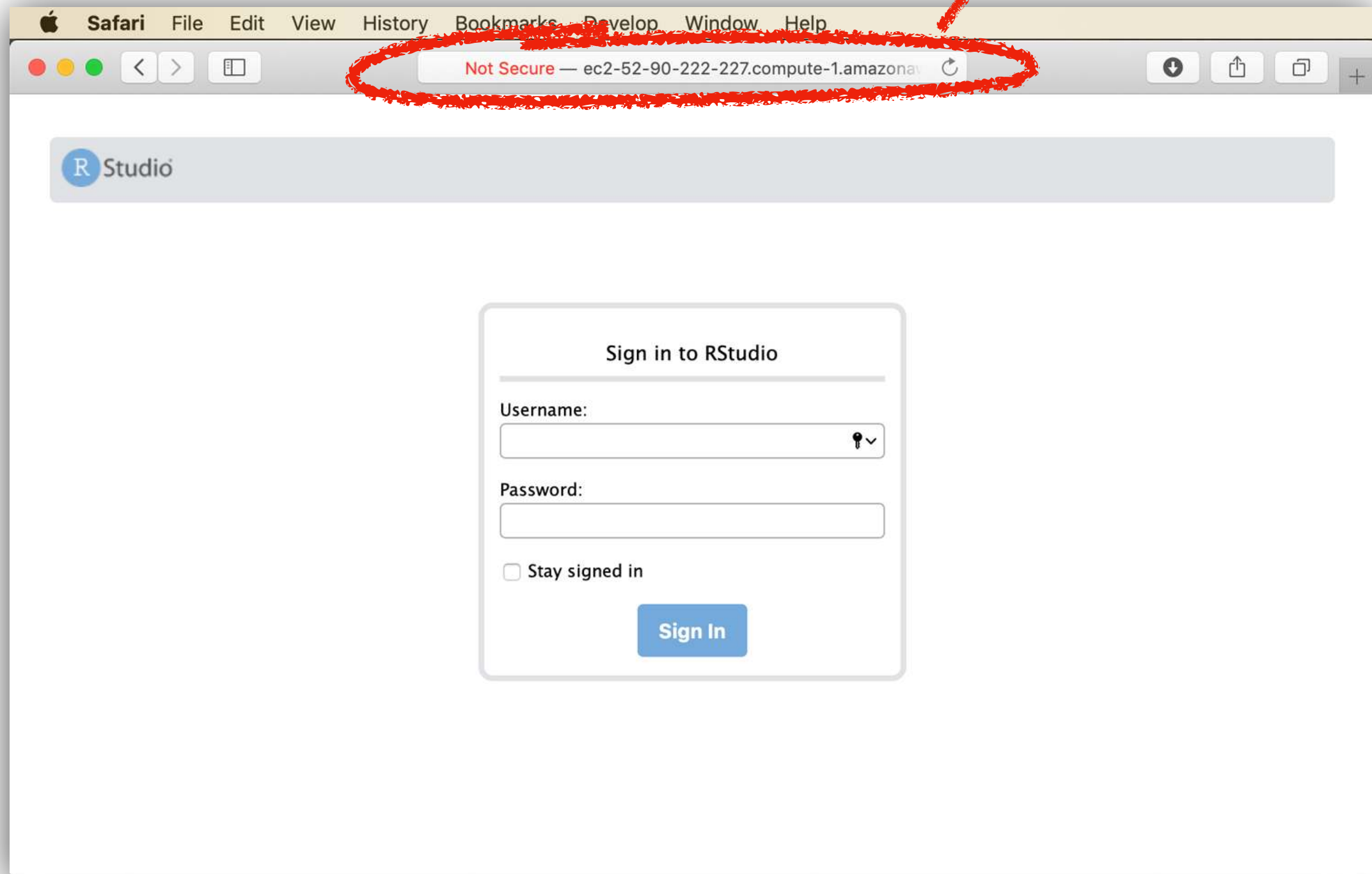
2. Guacamole



3. RStudio

3. RStudio

...amazonaws.com:8787



3. RStudio

Safari File Edit View History Bookmarks Develop Window Help

Not Secure — ec2-52-90-222-227.compute-1.amazonaws.com

R Studio

Sign in to RStudio

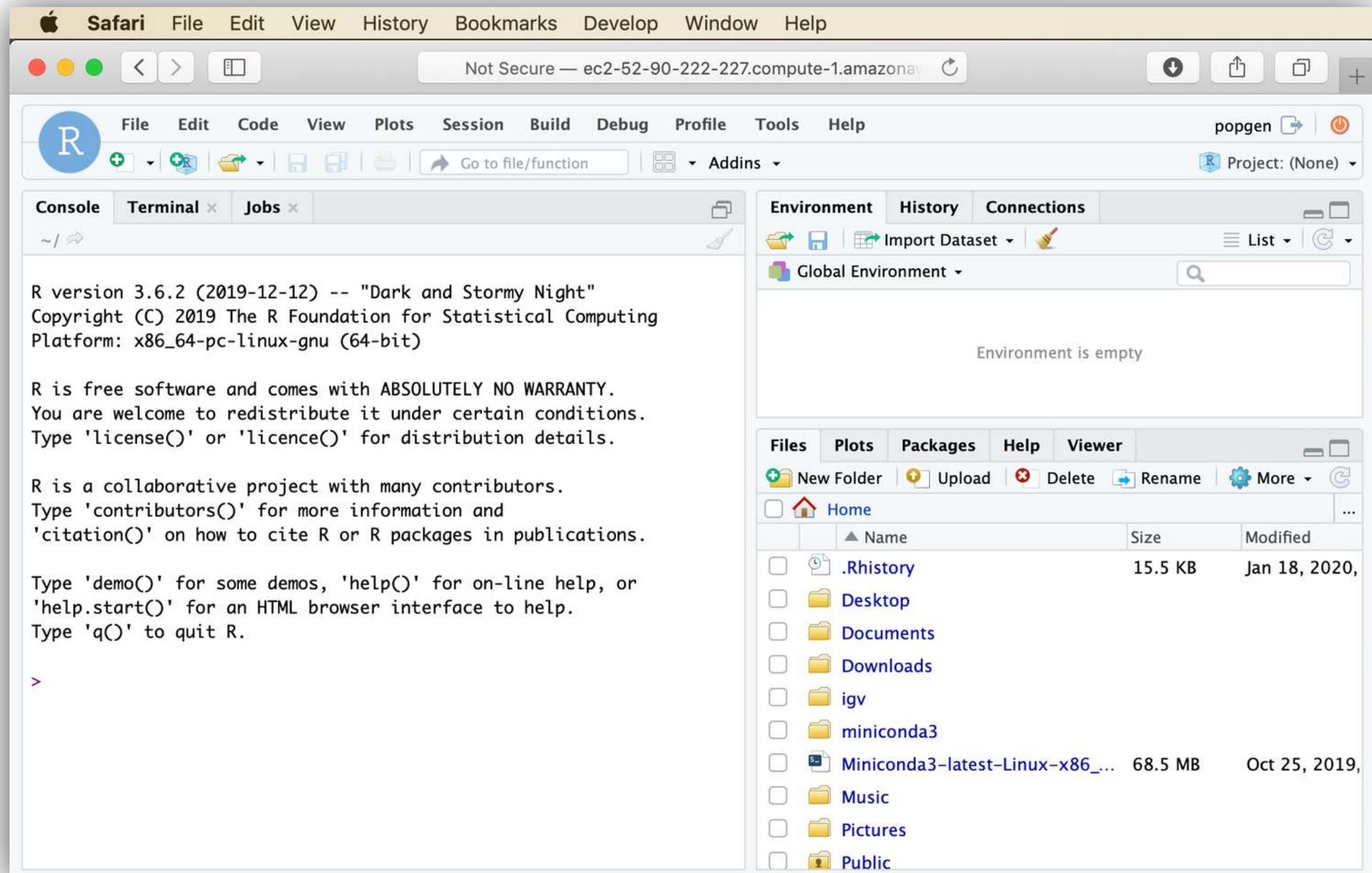
Username:

Password:

☐ Stay signed in

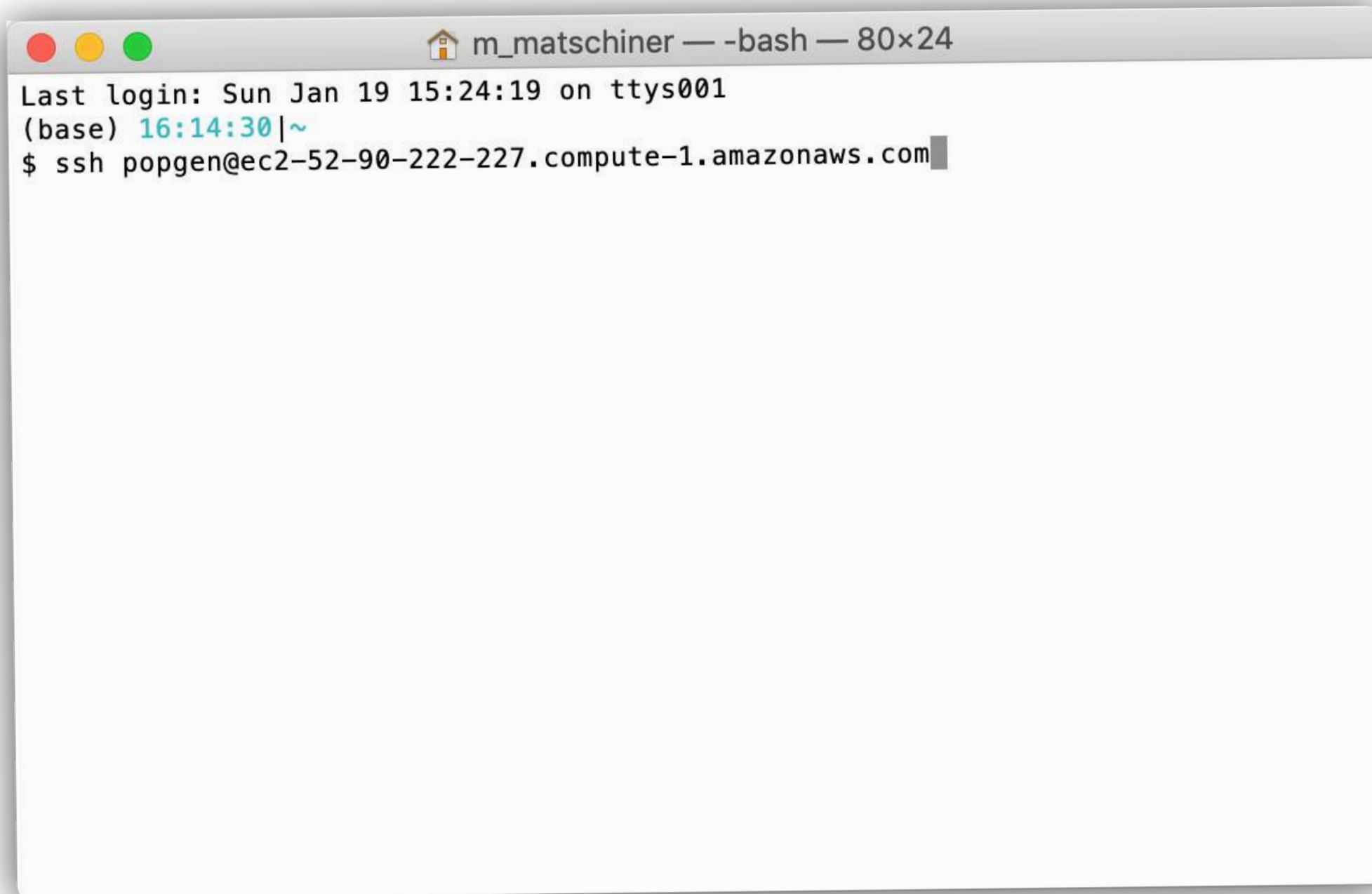
Sign In

3. RStudio



4. Jupyter Notebooks

4. Jupyter Notebooks



```
m_matschiner — -bash — 80x24
Last login: Sun Jan 19 15:24:19 on ttys001
(base) 16:14:30|~
$ ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com
```

4. Jupyter Notebooks

```
m_matschiner — ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.co...
Last login: Sun Jan 19 15:24:19 on ttys001
(base) 16:14:30 | ~
[$ ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com
[popgen@ec2-52-90-222-227.compute-1.amazonaws.com's password:

#####
##          Workshop on Popgen 2020          ##
##              Cesky Krumlov              ##
##              @evomics #wpsg2020          ##
#####

Welcome to Ubuntu 19.10 (5.3.0-1009-aws).

Last login: Sun Jan 19 16:04:15 2020 from 90.177.198.83
(base) 16:14:52 :~ $
```

4. Jupyter Notebooks

```
m_matschiner — ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.co...
Last login: Sun Jan 19 15:24:19 on ttys001
(base) 16:14:30|~
[$ ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com
[popgen@ec2-52-90-222-227.compute-1.amazonaws.com's password:

#####
##          Workshop on Popgen 2020          ##
##          Cesky Krumlov                    ##
##          @evomics #wpsg2020                ##
#####

Welcome to Ubuntu 19.10 (5.3.0-1009-aws).

Last login: Sun Jan 19 16:14:32 2020 from 93.177.198.83
(base) 16:14:32 :~ $ conda activate conda
```

4. Jupyter Notebooks

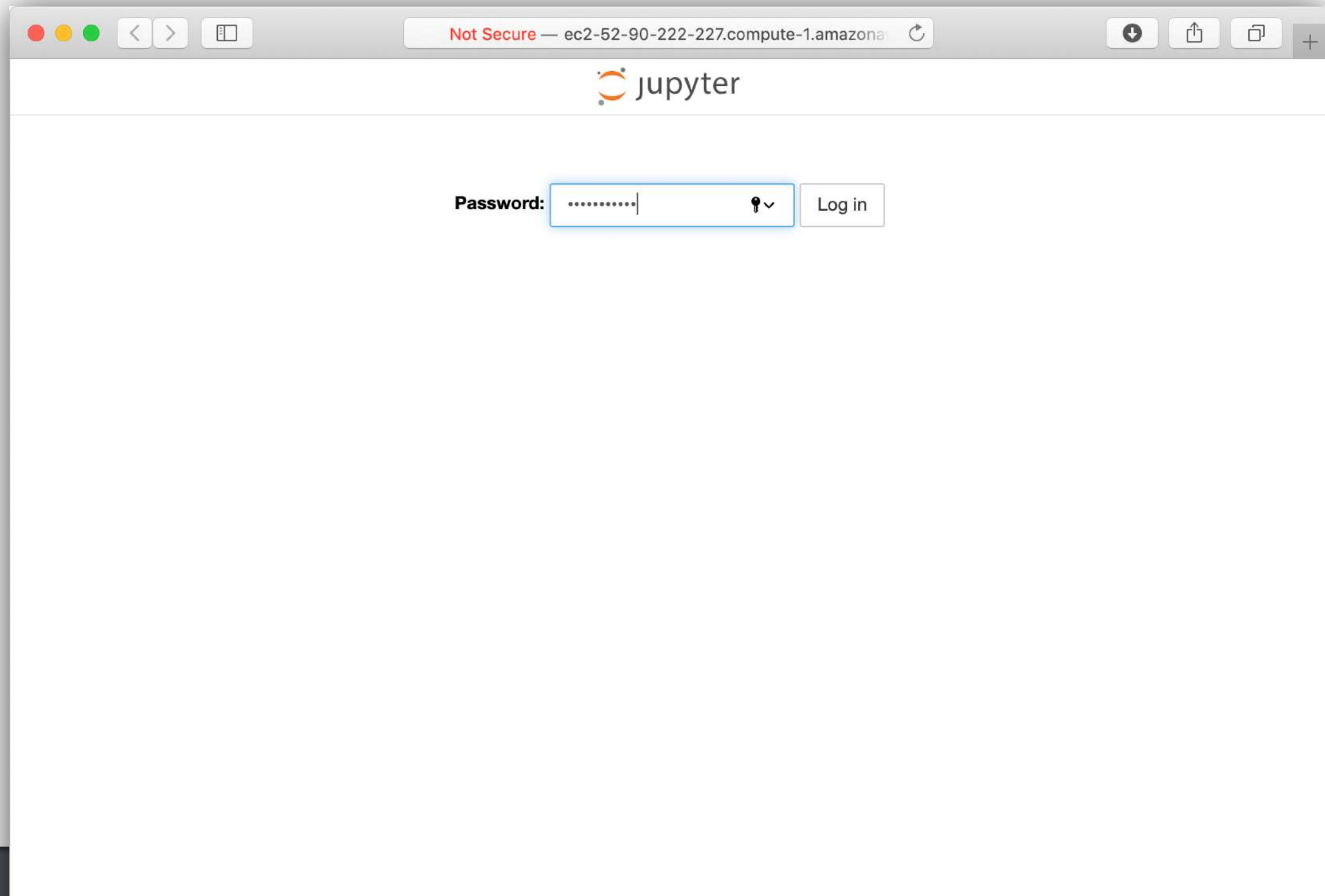
```
m_matschiner — ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.co...
Last login: Sun Jan 19 15:24:19 on ttys001
(base) 16:14:30 | ~
[$ ssh popgen@ec2-52-90-222-227.compute-1.amazonaws.com
[popgen@ec2-52-90-222-227.compute-1.amazonaws.com's password:

#####
##          Workshop on Popgen 2020          ##
##              Cesky Krumlov              ##
##          @evomics #wpsg2020              ##
#####

Welcome to Ubuntu 19.10 (5.3.0-1009-aws).

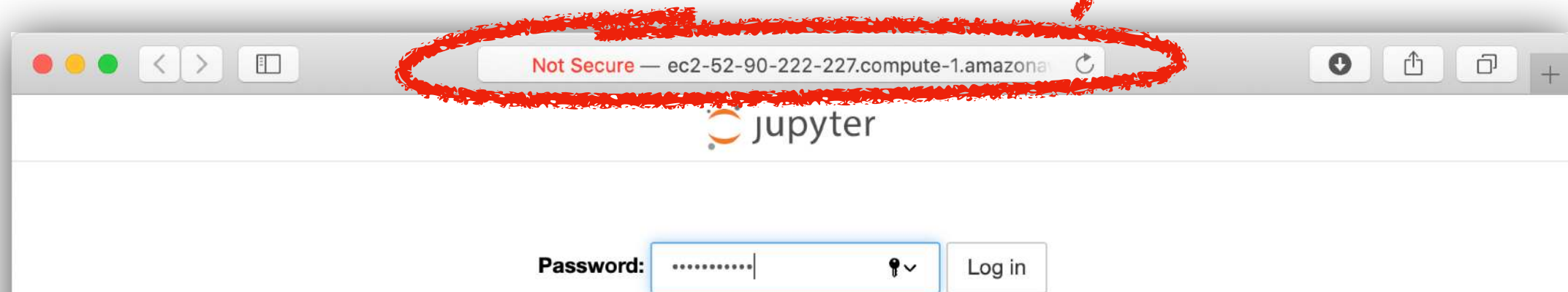
Last login: Sun Jan 19 16:04:15 2020 from 90.177.198.83
(base) 16:14:52 : ~ $ conda activate conda
(conda) 16:15:40 : ~ $ jupyter notebook
```

4. Jupyter Notebooks

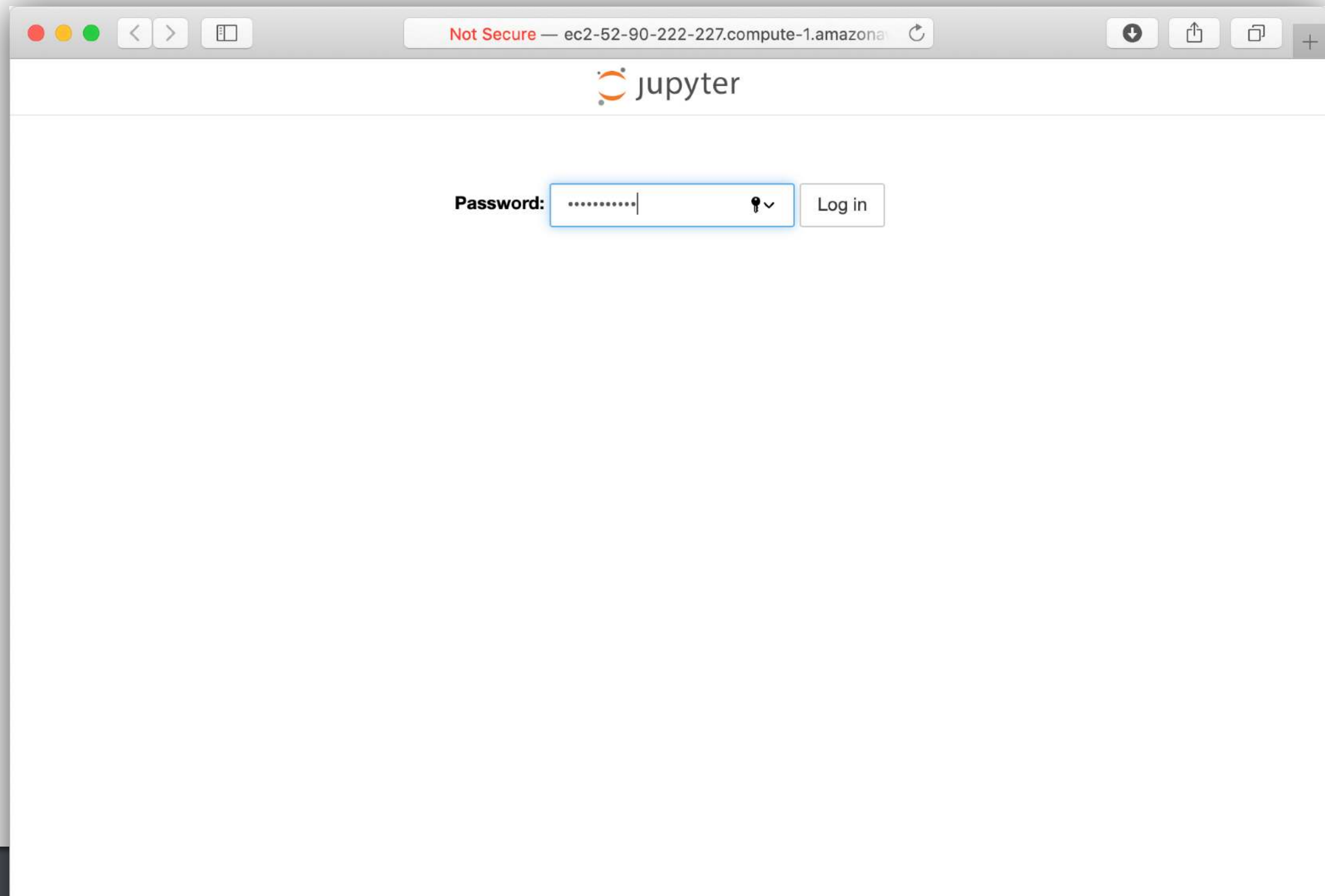


4. Jupyter Notebooks

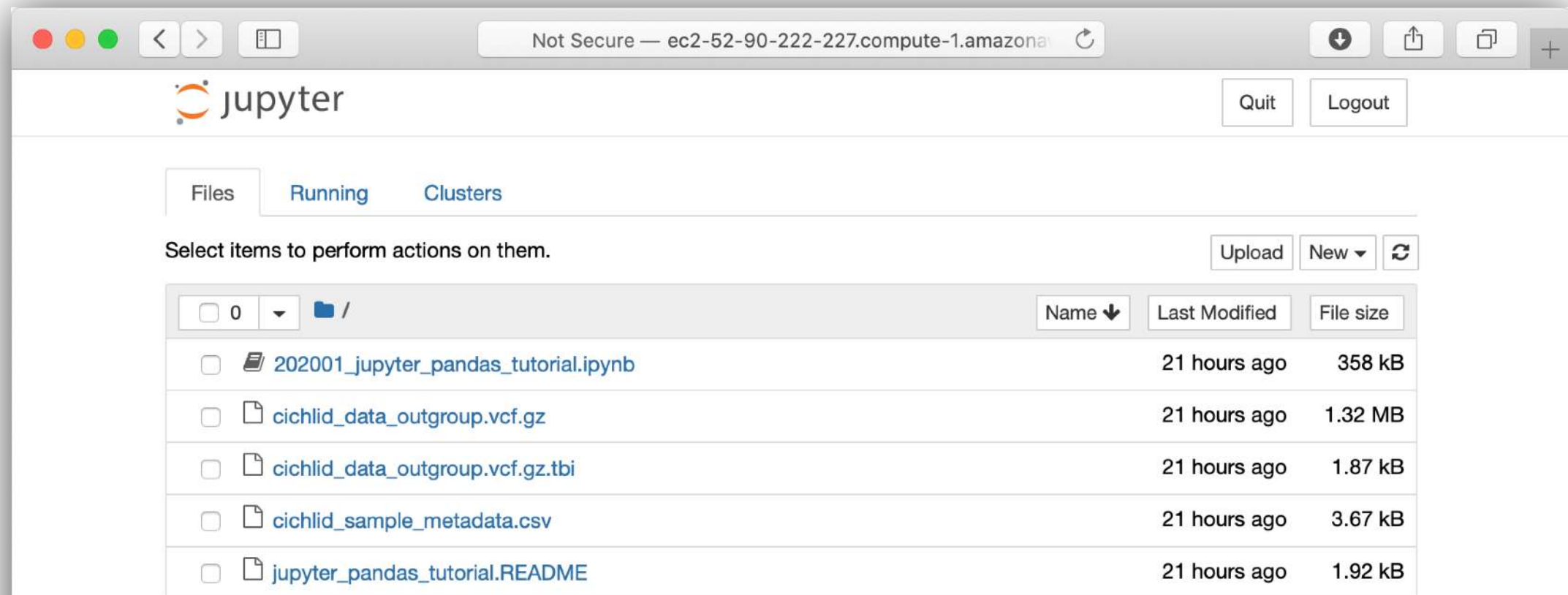
...amazonaws.com:8888



4. Jupyter Notebooks



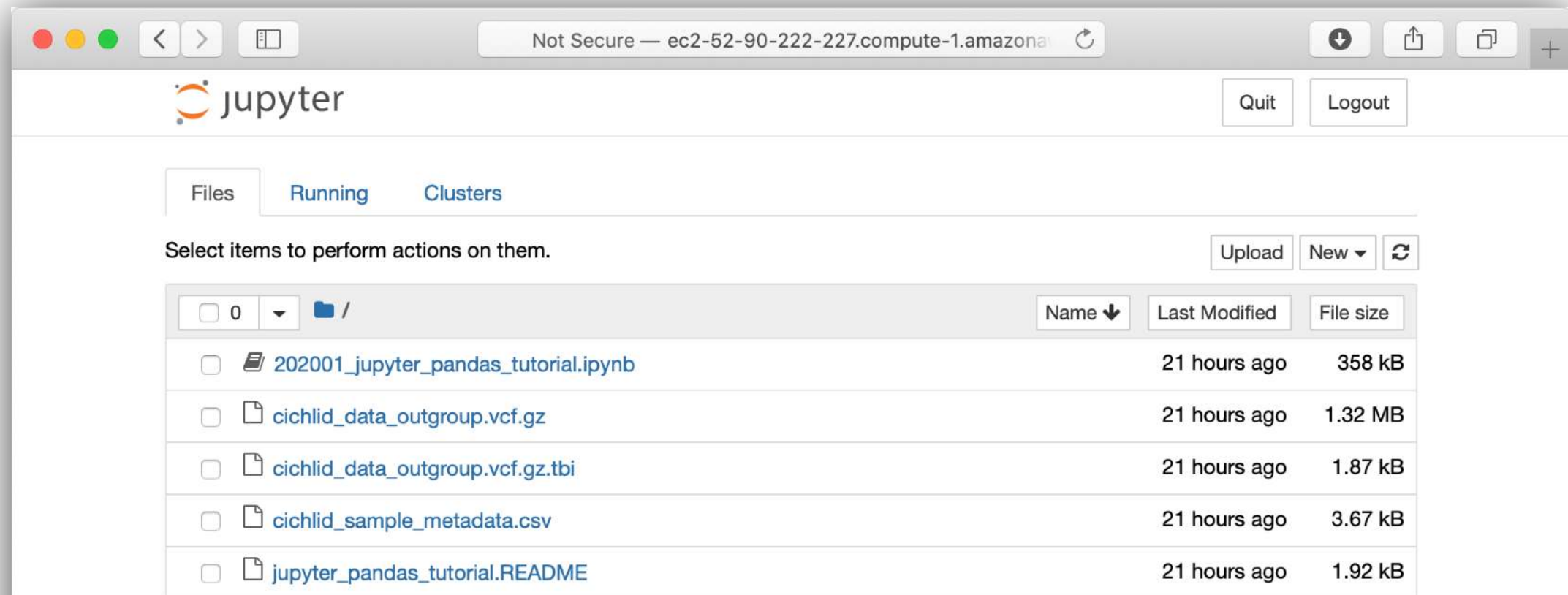
4. Jupyter Notebooks



The screenshot shows the Jupyter web interface in a browser window. The address bar indicates a "Not Secure" connection to an Amazon EC2 instance. The Jupyter logo and "Quit" and "Logout" buttons are at the top. Below the logo, there are tabs for "Files", "Running", and "Clusters". The "Files" tab is active, showing a file browser interface. At the top of the file browser, there is a prompt "Select items to perform actions on them." and buttons for "Upload", "New", and a refresh icon. Below this is a table of files in the current directory. The table has columns for "Name", "Last Modified", and "File size". The files listed are:

	Name	Last Modified	File size
☐	0		
☐	202001_jupyter_pandas_tutorial.ipynb	21 hours ago	358 kB
☐	cichlid_data_outgroup.vcf.gz	21 hours ago	1.32 MB
☐	cichlid_data_outgroup.vcf.gz.tbi	21 hours ago	1.87 kB
☐	cichlid_sample_metadata.csv	21 hours ago	3.67 kB
☐	jupyter_pandas_tutorial.README	21 hours ago	1.92 kB

4. Jupyter Notebooks



The screenshot shows the Jupyter web interface in a browser window. The address bar indicates a "Not Secure" connection to an Amazon EC2 instance. The Jupyter logo and "Quit" and "Logout" buttons are at the top. Below the logo, there are tabs for "Files", "Running", and "Clusters". The "Files" tab is active, showing a file browser interface. At the top of the file browser, there is a prompt "Select items to perform actions on them." and buttons for "Upload", "New", and a refresh icon. Below this is a table of files in the current directory. The table has columns for "Name", "Last Modified", and "File size". The files listed are:

	Name	Last Modified	File size
☐	0		
☐	202001_jupyter_pandas_tutorial.ipynb	21 hours ago	358 kB
☐	cichlid_data_outgroup.vcf.gz	21 hours ago	1.32 MB
☐	cichlid_data_outgroup.vcf.gz.tbi	21 hours ago	1.87 kB
☐	cichlid_sample_metadata.csv	21 hours ago	3.67 kB
☐	jupyter_pandas_tutorial.README	21 hours ago	1.92 kB

4. Jupyter Notebooks

Home Page - Select or create a notebook

202001_jupyter_pandas_tutorial - Jupyter Notebook

jupyter 202001_jupyter_pandas_tutorial (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3

Run

Markdown

Genome data analysis in Python

A brief tutorial on the use of *jupyter notebooks* and the python data analysis library *pandas* for genomic data analysis.

Workshop on Population and Speciation Genomics, Český Krumlov, January 2020.
By Hannes Svardal (hannes.svardal@uantwerpen.be)

This is a jupyter notebook running a Python 3 kernel. The Jupyter Notebook App (formerly IPython Notebook) is an application running inside the browser. Jupyter notebooks can run different kernels: Python 2/3, R, Julia, bash, ...

Further resources about jupyter notebooks can be found here:

- <https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/>
- <https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook>

Jupyter notebooks can run locally or on a server. You access them in your browser.

Activity

EVOLUTION AND GENOMICS

Intensive and immersive training opportunities



2020 WORKSHOP ON POPULATION AND SPECIATION GENOMICS, CESKY KRUMLOV



The 2020 Workshop on Population and Speciation Genomics will be held in [Český Krumlov](#) from 19 January – 1 February. It brings together an international collection of faculty members and Workshop participants to study and discuss current ideas and techniques for the analysis of genomic data on the level of populations and closely related species. The Workshop consists of a series of lectures, demonstrations and computer laboratories that cover theory and practice of population and speciation genomics analyses.

The Workshop registration fee is \$1,800. This amount does not cover travel, lodging or boarding. Many important



The 2020 Workshop on Population and Speciation Genomics will be held in [Český Krumlov](#) from 19 January – 1 February. It brings together an international collection of faculty members and Workshop participants to study and discuss current ideas and techniques for the analysis of genomic data on the level of populations and closely related species. The Workshop consists of a series of lectures, demonstrations and computer laboratories that cover theory and practice of population and speciation genomics analyses.

The Workshop registration fee is \$1,800. This amount does not cover travel, lodging or boarding. Many important details about the Workshop series can be read on our [FAQ](#) page.

The Workshop on Population and Speciation Genomics is organized and co-directed by:

[Walter Salzburger](#), University of Basel, Switzerland

[Michael Matschiner](#), University of Zurich, Switzerland

[Jan Stefka](#), University of South Bohemia České Budějovice, Czech Republic

[Scott Handley](#), Washington University, US

[Emiliano Trucchi](#), Università Politecnica delle Marche, Ancona, Italy

[Julia M.I. Barth](#), University of Basel, Switzerland

[Britta S. Meyer](#), Max Planck Institute for Evolutionary Biology, Plön, Germany

[Milan Malinsky](#), University of Basel, Switzerland

ANNOUNCEMENTS

We are looking forward to meet everybody at the reception dinner on Sunday at 6 p.m. at the hotel Zlatý Anděl!

LINKS

- [Lecture slides and learning activities](#)
- [Map of important locations in Český Krumlov](#)
- [Amazon instances IP addresses](#)

2020 Workshop on Population and Speciation Genomics, Cesky Krumlov - Evolution a...

Population and Speciation Genomics 2020 - Evolution and Genomics

Howdy, Michael Matschiner

EVOLUTION AND GENOMICS

Intensive and immersive training opportunities

POPULATION AND SPECIATION GENOMICS 2020

Here you will find lecture slides and learning activities for the [2020 Workshop on Population and Speciation Genomics](#). Slides and activities are designed by the faculty and the workshop team.

This page will be updated during the workshop.

Week I

Workshop introduction and orientation

...

Computer lab, UNIX, and R

- [Connecting to your Amazon instance](#)
- [Basic UNIX, some tips and tricks, and more!](#)
- [Introduction to R](#)
- [Genome data analysis in Python](#)

2020 Workshop on Population and Speciation Genomics, Cesky Krumlov - Evolution a...

Amazon instances - Evolution and Genomics

Howdy, Michael Matschiner

EVOLUTION AND GENOMICS

Intensive and immersive training opportunities

AMAZON INSTANCES

Connecting to your Amazon instance

Michael Matschiner, Milan Malinsky, and the Workshop Team; 20th January 2020

Amazon Web Services and Amazon instances

For this workshop we are going to use a high performance cloud platform called [Amazon Web Services](#) (AWS). This is a flexible resource to get computational work done quickly and relatively inexpensively. We have set up individual virtual machines called "Amazon instances" for each participant. These have all the software and data you will need along with sufficient CPU, RAM, and storage resources. Your instance IP address can be found on [this list](#) and the list of IP addresses is also linked from the [main workshop page \(in the box with links\)](#).

The IP addresses are going to change regularly during the workshop, as we'll stop the instances overnight and resume them on the next day to cut down the cost that we need to pay to Amazon. Whenever instances are restarted, the IP addresses will change; however, the content of the instance will remain the same. So the files that you write on any given day on your instance should usually still be present on the next. We are going to announce when we stop instances overnight, and we will accept requests to leave instances running in case that you'd like to keep working on an analysis.

In addition to stopping instances overnight, it may also occur that we need to issue completely new instances at some point. This may be the case when/if we need to install additional software or add datasets for activities. In such cases, we'll terminate old instances, meaning that the files from these instances will not be accessible