PAN 2013
User Guide for the Virtual Machines

1   Welcome

Welcome to the PAN 2013 evaluation lab. A novelty this year is that PAN switches from the submission of runs (say, the output of a piece of software) to the submission of the software itself. From a scientific viewpoint, software submissions are attractive since they allow anyone to re-evaluate and compare a task's solution on any new or rearranged datasets in the future (given your consent). Last year, we have experimented with software submissions on a subtask of plagiarism detection which turned out to be a success. Hence, this year we decided to scale up software submissions. In order to make them convenient and transparent, we provide to each participant a personal, fully accessible virtual machine on our servers. We ask you to deploy your software to this virtual machine. The purpose of this user guide is to describe how to access these virtual machine and how to make a submission to PAN.

2   Getting a Virtual Machine

Once you have registered for PAN using the registration form found on the website (http://pan.webis.de), you will receive a welcome e-mail in which we ask for your preferred operating system (Ubuntu 12.04 or Windows 7). After we have set up your virtual machine, you will receive an e-mail with all the necessary information to access it (`XX` refers to a two-digit number):

Host address	webisXX.medien.uni-weimar.de
User name	`<name>`
User password	`<password>`
SSH port	444`XX`
RDP port	555`XX`

3 Accessing your Virtual Machine

You can access your virtual machine via SSH and via remote desktop.

From Ubuntu to a Ubuntu Virtual Machine

Use the following command to access your virtual machine via SSH:
```
ssh <name>@<host-address> -p <ssh-port>
```

Use the following command to access your virtual machine via remote desktop (RDP):
```
rdesktop <host-address>:<rdp-port> -u <name> -p <password>
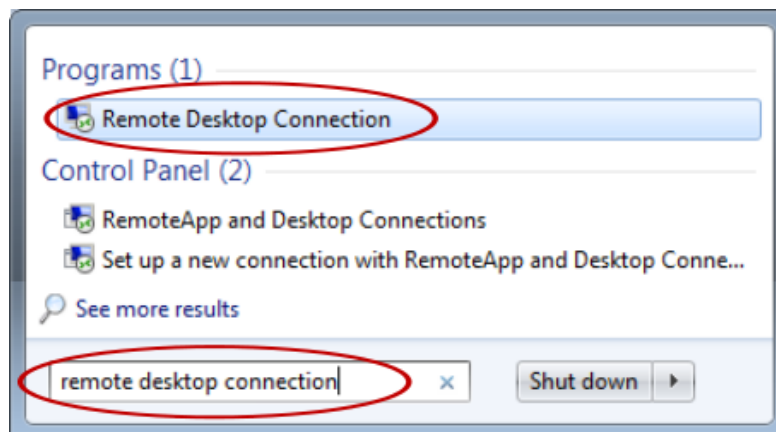```

## From Windows to a Windows Virtual machine

In order to access a Windows virtual machine via SSH, you first need to install a SSH client on your PC. We recommend to use OpenSSH which you can download from Sourceforge.[1]. After installing OpenSSH use the following command to access your virtual machine via SSH:
```
ssh <name>@<host-address> -p <ssh-port>
```

In order to access a Windows virtual machine via Remote Desktop, we recommend you use the Windows Remote Desktop Connection Dialog.
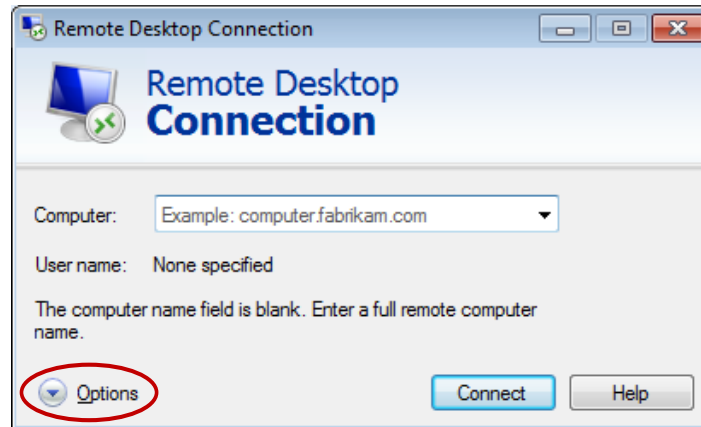
### Step 1

Press the Windows Start button. Enter "remote desktop connection" into the search field. Start Remote Desktop Connection Dialog which is listed in the category Programs.



---

[1]http://sshwindows.sourceforge.net/download/

**Step 2**

Press Options.



**Step 3**
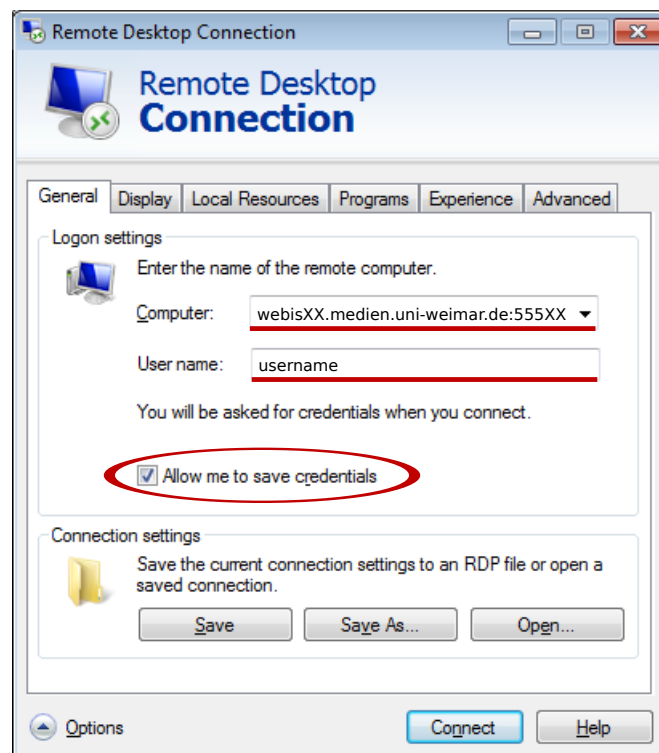
Enter your virtual machine access information:

Computer: `<host-address>:<rdp-port>`

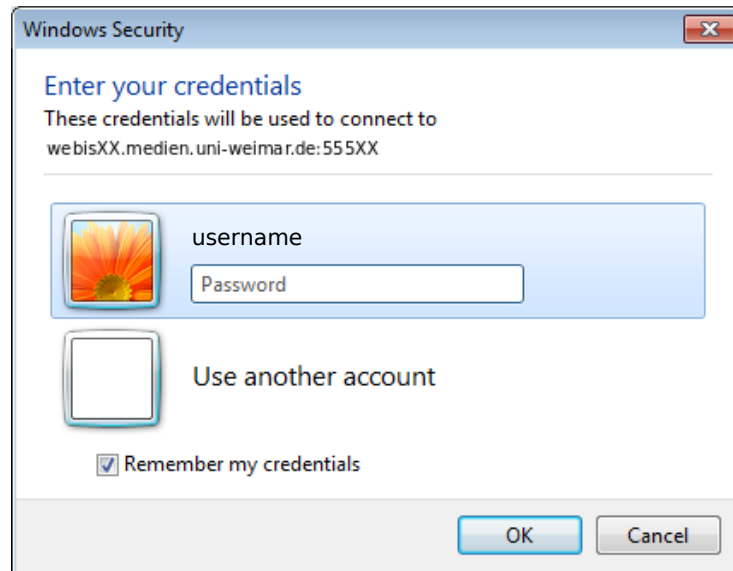User Name: `<name>`

Check "Allow me to save credentials" (which is very important!). Press Connect.

**Step 4**

Enter your `<password>` and press OK.



## 4 Permissions on the Virtual Machine

Your user is member of the sudoer group (Ubuntu) or the admin group (Windows). Hence, you can install any additional software you may need to run your software. The disc space allocated to your virtual machine should be sufficient to store your software as well as any runtime data you need. Should you still need more disc space than is available, please send an e-mail to pan@webis.de with a short explanation how much disc space you need. Please do not shut down your virtual machine; you can not bring it up again on your own. Rather restart your virtual machine using the corresponding option of your operating system.

## 5 Training Data on the Virtual Machines

Once your software is deployed on your virtual machine, you may want to try it out to see if it works. The training data of each PAN task is mounted as a shared folder into your virtual machine, so that you don't need to copy the data yourself. On Ubuntu you can find the training data in the folder `/media/pan13-training-data`. On Windows, the training data is found on drive `E:\`.

# 6   Completing a Submission to PAN

To complete a submission, deploy your software to your virtual machine and send an e-mail to pan@webis.de. Your e-mail must contain the following information:

| | |
|---|---|
| Subject: | PAN submission: `<task>` |
| Software location: | `<path to executable>` |
| Execution command: | `<execution command>` |

The `<task>` may be "Plagiarism Detection: Text Alignment," "Plagiarism Detection: Source Retrieval," "Author Identification," or "Author Profiling." Your software has to be executable from any directory on your virtual machine by issuing the `<execution command>` you provide in your submission e-mail. The parameters of the execution command must be in accordance with the task-specific information on the PAN website (for example it must be possible to specify an absolute path to a data set similar to the training data).