pycsw Workshop

version 1.4

Jeff McKenna, Tom Kralidis, Angelos Tzotsos

June 06, 2013

Contents

pycsw Workshop **1**

- Background 1
  - History 1
  - Acknowledgements 1
  - Assumptions 1
  - License 1
- Workshop Requirements 1
  - 1. Install VirtualBox 2
  - 2. Download OSGeo-Live 2
    - Download 2
    - Extract 2
  - 3. Create Virtual Machine 2
  - 4. Run the Virtual Machine 6
- Metadata Background 7
  - Types 8
  - History 8
- OGC CSW Specification 8
  - Purpose 8
  - Example Requests 8
- Introduction to pycsw 8
  - Goals 8
  - Features 8
  - Use Cases 8
  - Installing 8
  - Configuring 8
  - Exercises 8
- Advanced pycsw 8
  - Harvesting 8
  - Exporting 8
  - Tips and Tricks 8
  - Community 8
  - Exercises 8
- pycsw and Open Data 8
  - GeoNode 9
  - Open Data Catalog 9
- pycsw Future Development 9

pycsw Workshop

Welcome to the pycsw workshop. This workshop is a hands-on workshop that will give you an introduction to the popular pycsw metadata publishing software.

Note

The workshop instructions are also available as a single PDF document 

Background

History

(high-level history of how/why pycsw was created...use IRC discussions)

Acknowledgements

The initial pycsw workshop materials were created through funding provided by the Oregon Coastal Management Program, through an FGDC CAP grant, in 2013.

Assumptions

License

Workshop Requirements

The pycsw workshop requires the following installed locally:

- 7-Zip (ability to extract *.7z* files)
- VirtualBox (virtual machine software, ability to load virtual disk *.vmdk files)
- OSGeo-Live Virtual Machine (which contains pycsw)

Note

We recommend using the OSGeo-Live Virtual Machine method, although OSGeo-Live is available also through a bootable DVD or USB drive.

1. Install VirtualBox

- download the VirtualBox platform package for your local machine
- run the installer, and select the default setup options (approve any device security questions)

2. Download OSGeo-Live

Caution!

You'll need a minimum of 10GB of free hard disk space, as well as a machine with 2GB of RAM.

Download

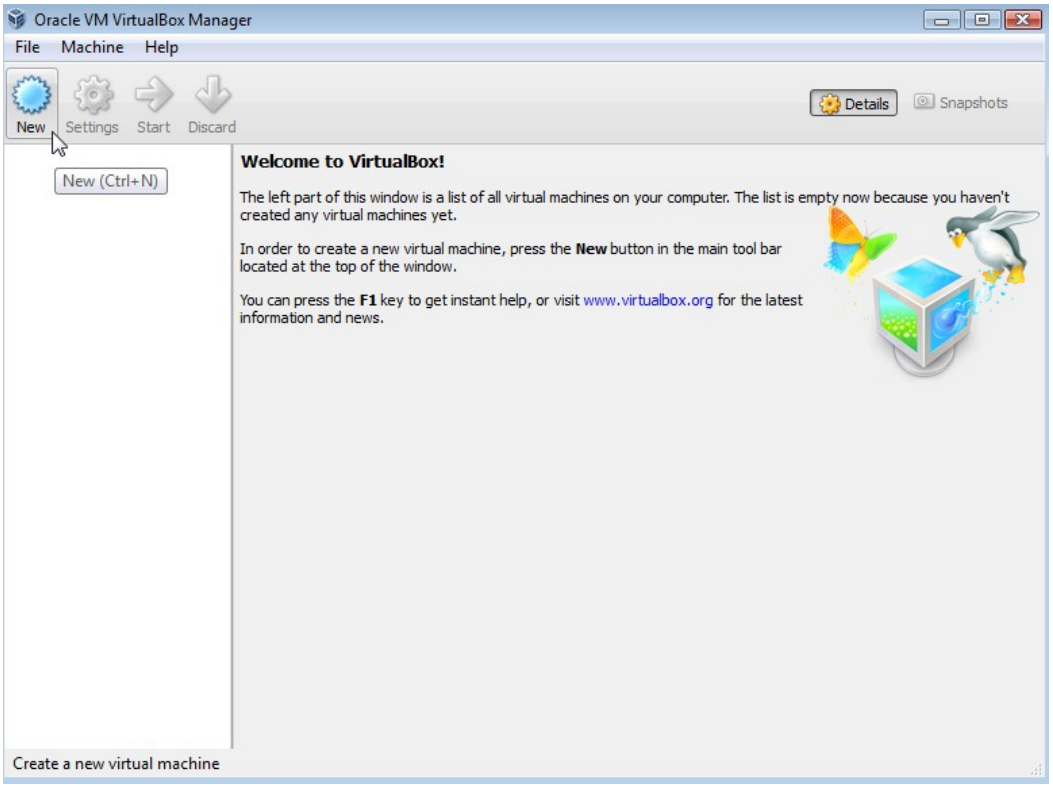
- download the OSGeo-Live Virtual Machine (*.7z) file. It will likely take you ~1 hour to download the 3GB file. There are several sites you can download this from:
  - official site
  - UC Davis mirror
  - National Technical University of Athens mirror

Extract

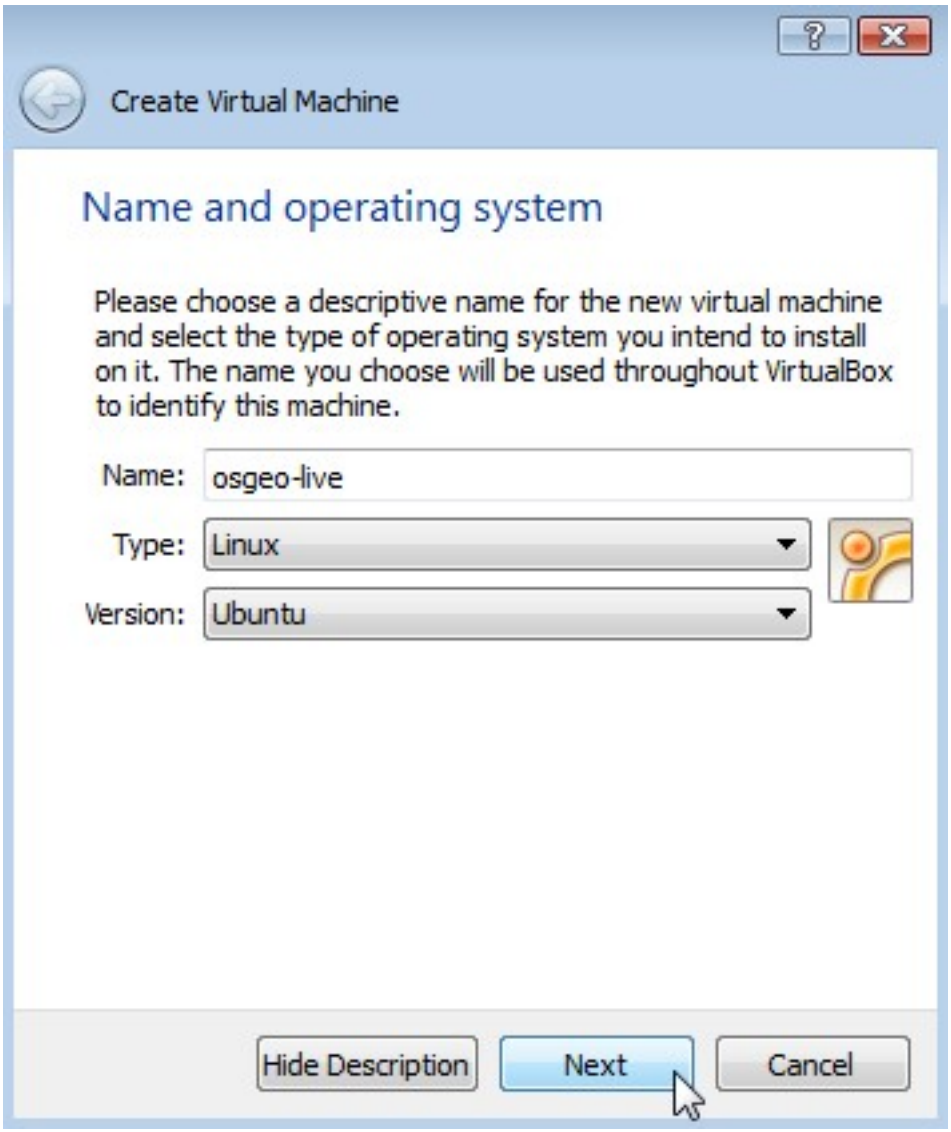
- using 7-Zip, open the *.7z* archive and extract the *.vmdk* file onto your hard disk (the extracted file is ~10GB in size)

3. Create Virtual Machine

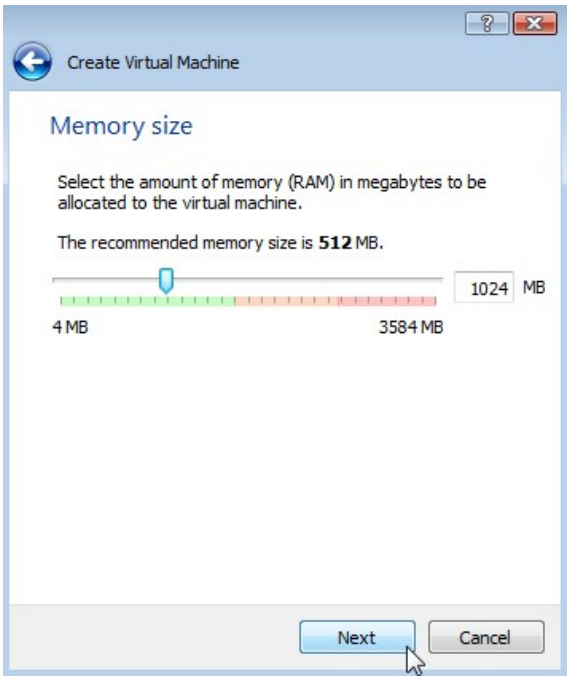
- start VirtualBox ("Oracle VM VirtualBox)
- click on the *New* button to create a VM



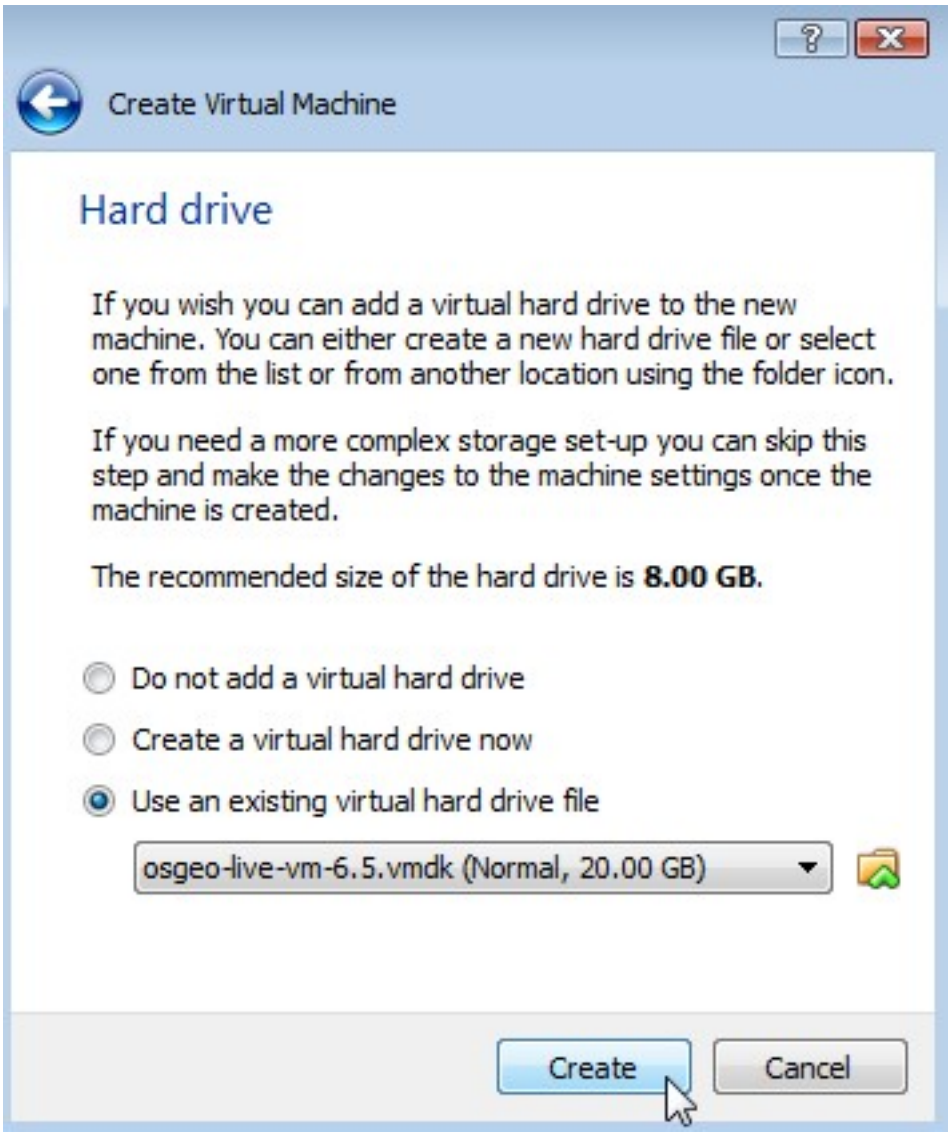
- enter "osgeo-live" for the name, and select *Type: Linux* and *Version: Ubuntu*



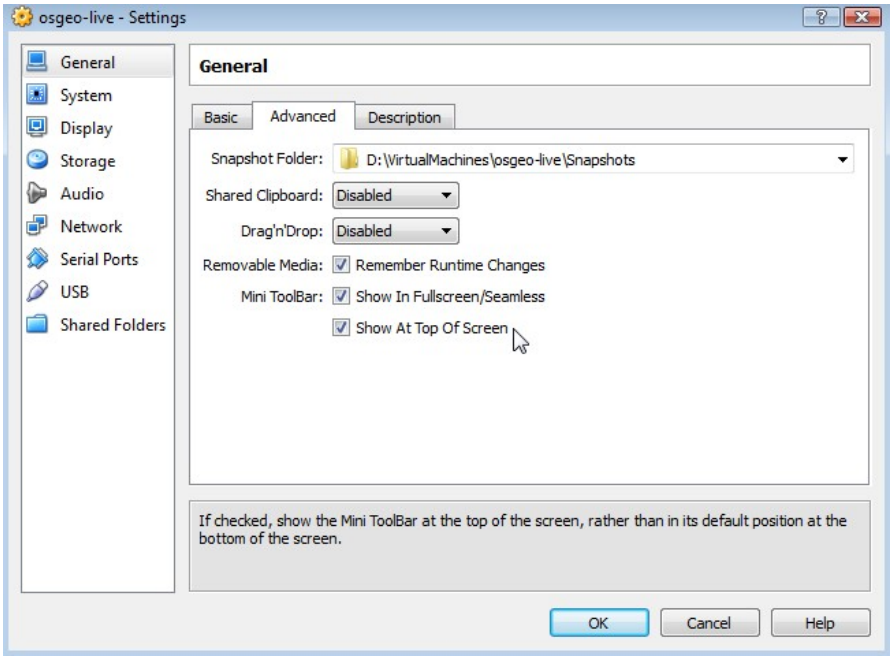
- In the next screen set the memory to 1024 MB (or more if your host computer has more than 4GB).



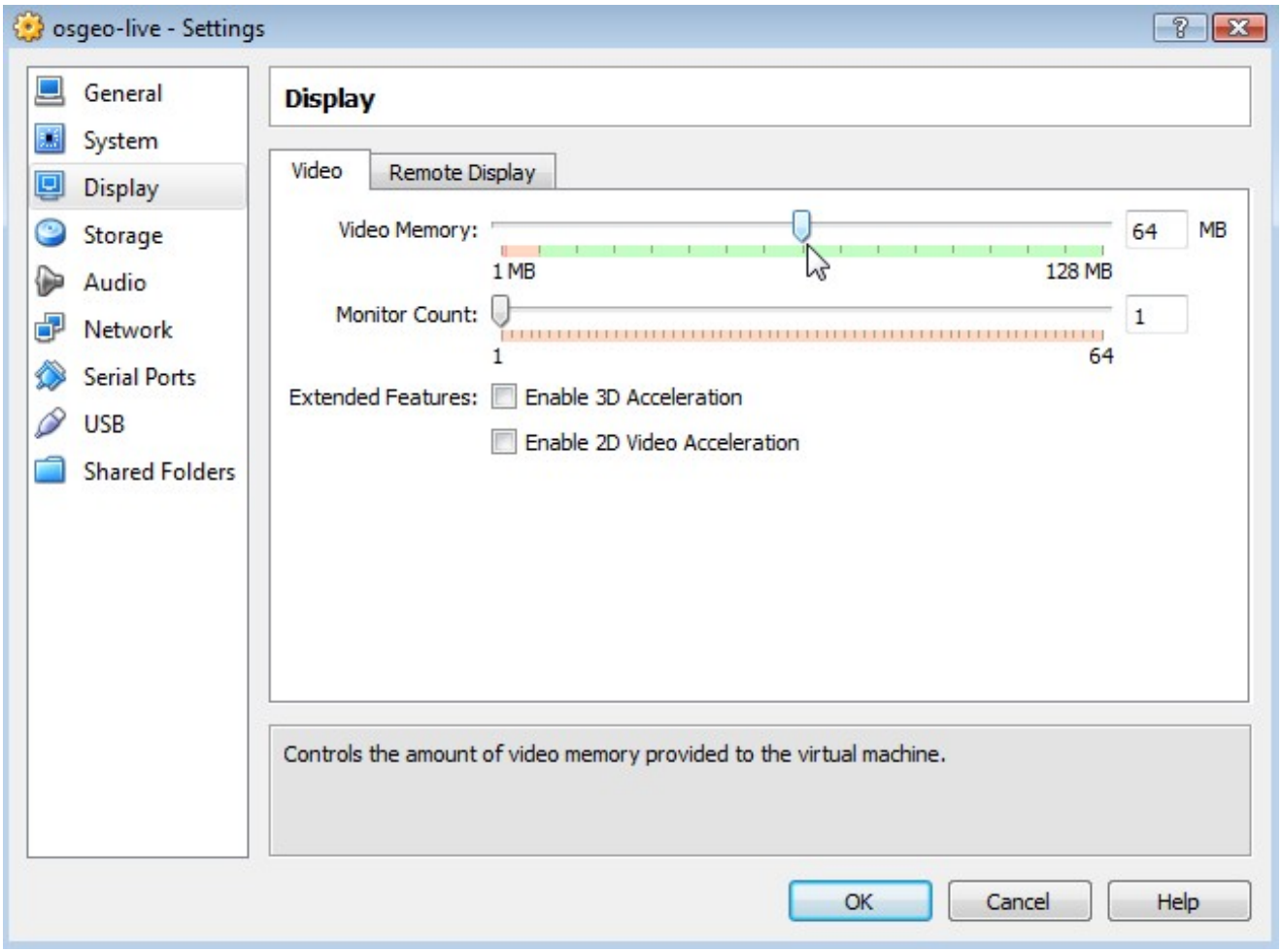
- Continue to the next screen and choose "Use existing hard disk". Then click on the button (a folder icon) to browse to where you saved the *.vmdk file. Select this file, press *Next* and *Create*.



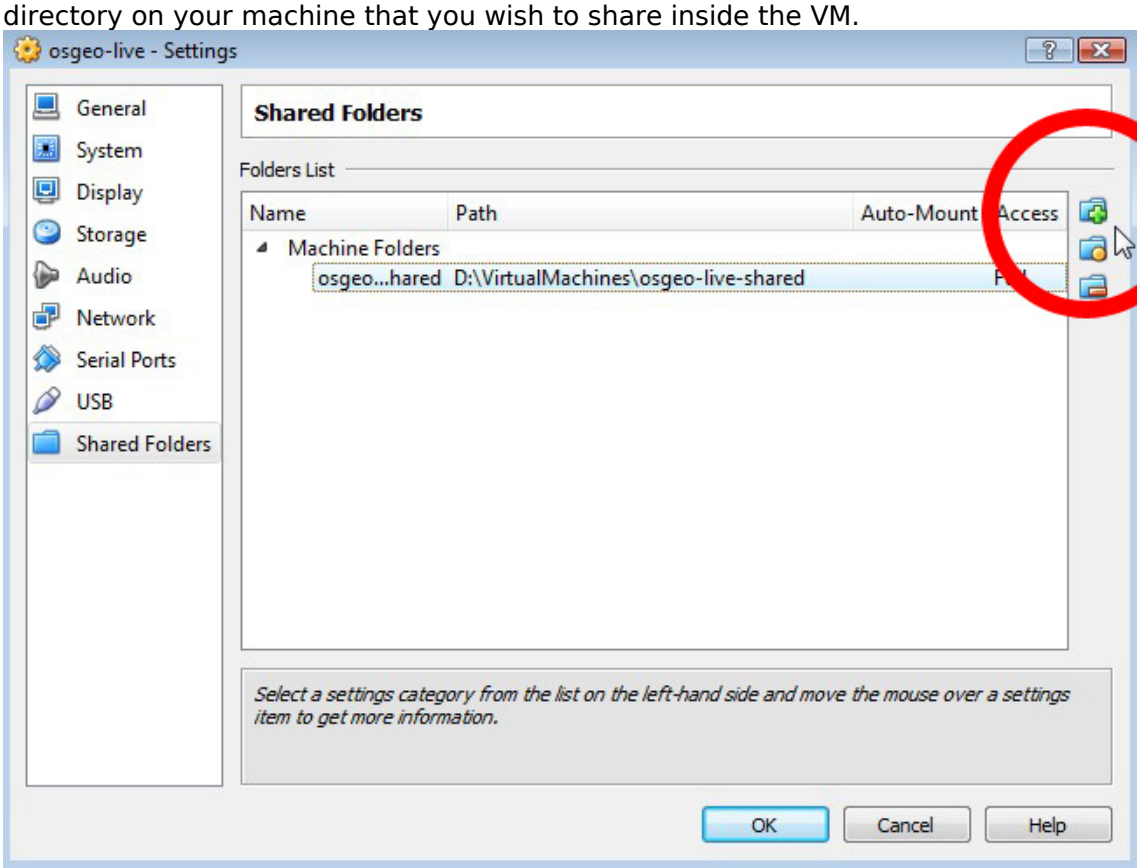
- Once the VM is created, click on the Settings button. In the "General" section, go to the Advanced tab, and click to select "Show at top of screen" for the Mini toolbar.



- In the "Display" section and increase video memory to 32 or 64 MB.



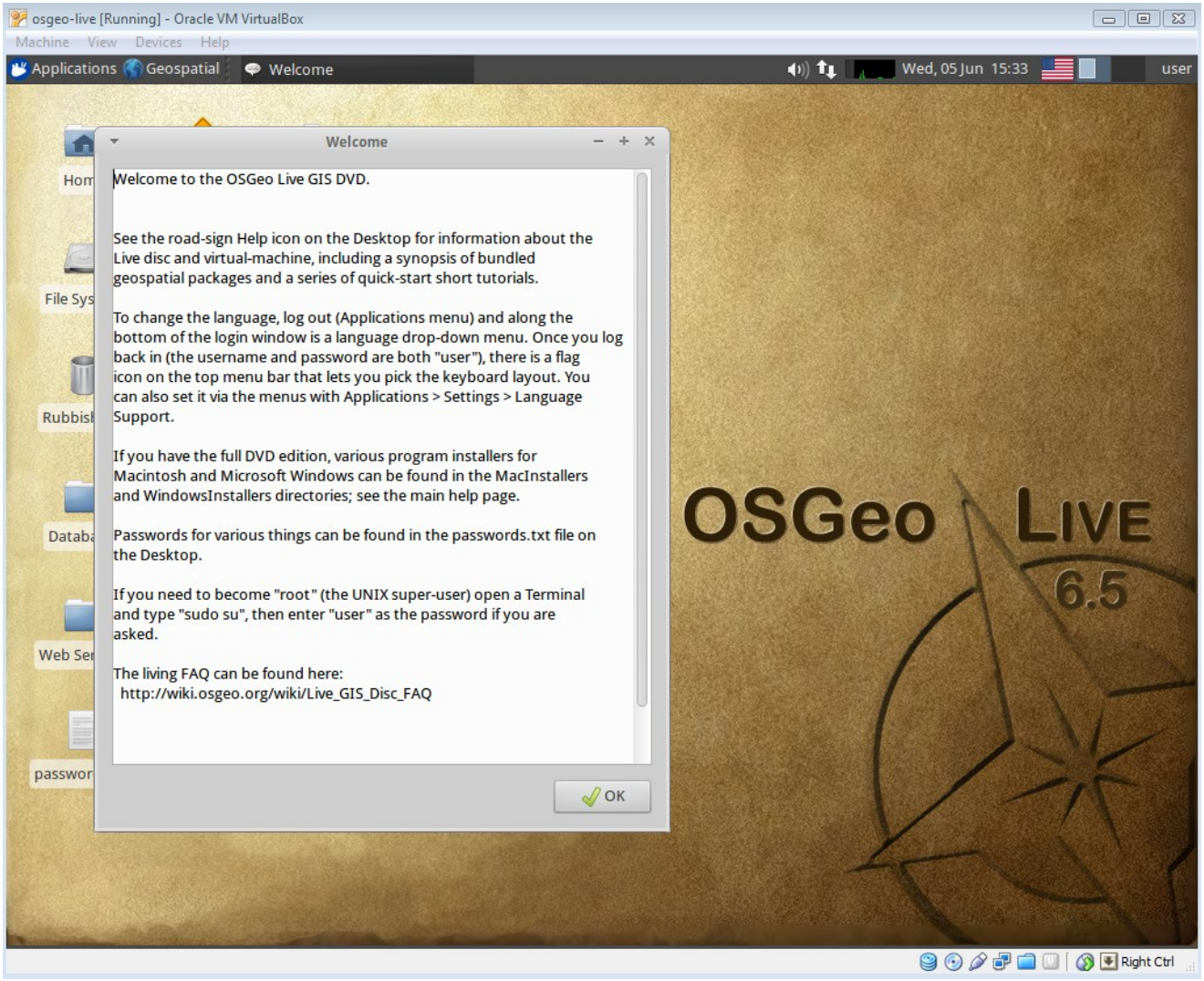
- In the "Shared Folders" section, click the "Add folder" (green + icon on the right) to find a directory on your machine that you wish to share inside the VM.



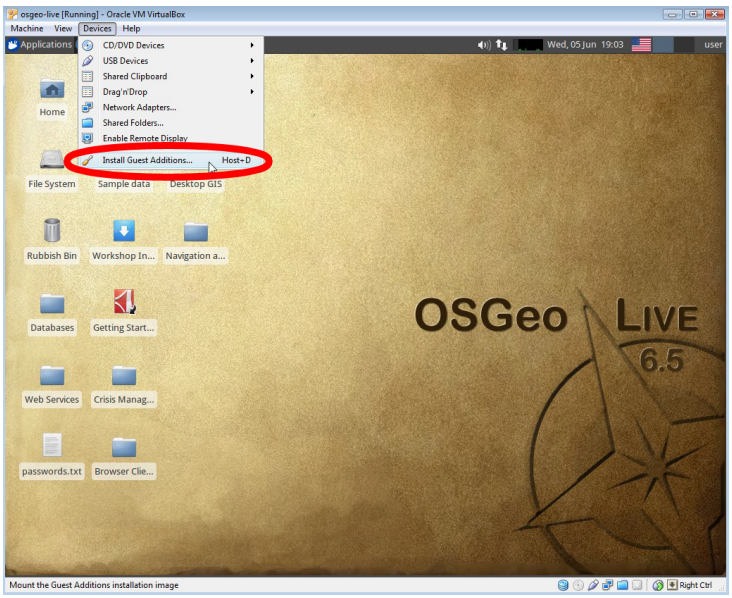
Once the "Folder path" and "Folder name" are defined, click OK, and close the settings window.

4. Run the Virtual Machine

- Now bootup the VM by clicking the Start (green arrow) button. OK any warning messages.



- To improve video performance and enable the shared folders, open the Devices menu and click "Install Guest Additions".



- Next, on the desktop you will see an icon named "VBOXADDITIONS_4.2.12_84980", click it (this mounts the drive). You can then close this window.
- Open a Terminal window (in top left click "Applications" / "Accessories" / "Terminal Emulator"
- In the Terminal, execute the following:

```
user@osgeolive:~$ sudo apt-get install linux-headers-`uname -r`

   password: user

user@osgeolive:~$ cd /media/VBOXADDITIONS_4.2.12_84980

user@osgeolive:/media/VBOXADDITIONS_4.2.12_84980$ sudo ./VBoxLinuxAdditions.run
```

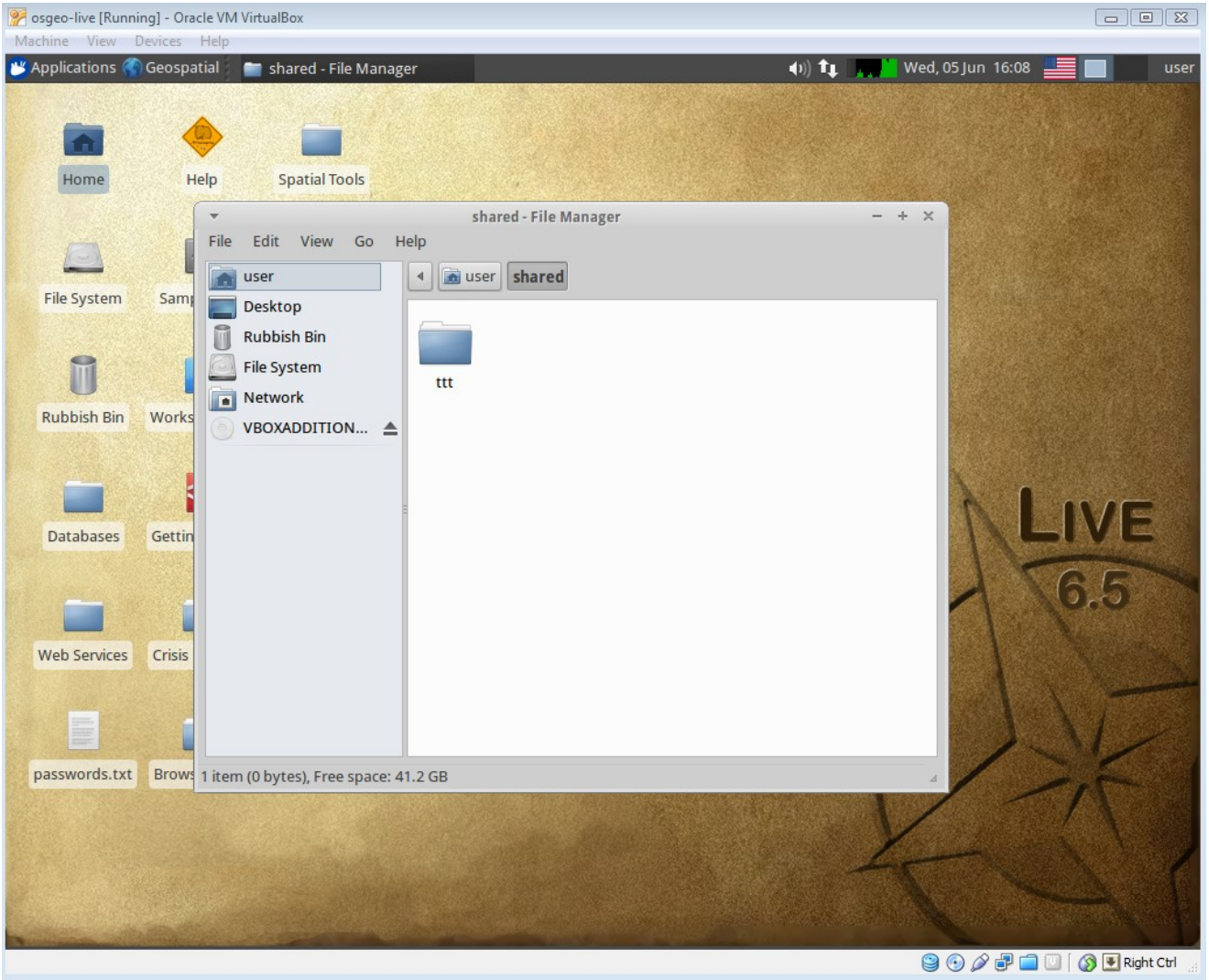
- reboot the machine (click on "user" in top-right of desktop, and select "Reboot")

- Open a Terminal window again, and execute the following (where "osgeo-live-shared" is the name you entered earlier in the Settings for the shared folder):

```
user@osgeolive:~$ mkdir shared

user@osgeolive:~$ sudo mount -t vboxsf -o uid=user,rw osgeo-live-shared /home/user/shared
```

You can now create a test folder on your local machine (in my case "ttt") and then view it within the virtual machine.



Metadata Background

Types

History

OGC CSW Specification

Purpose

Example Requests

Introduction to pycsw

Contents:

Goals

Features

Use Cases

Installing

Configuring

Exercises

Advanced pycsw

Contents:

Harvesting

Exporting

Tips and Tricks

Community

Exercises

pycsw and Open Data

- pycsw is embedded in various Open Data portal software
- metadata editing / management done with portal
- portal exposes CSW service automagically
- easy integration into existing apps/workflows

Contents:

GeoNode

- Open Source Geospatial Content Management System
  - geospatial data / metadata management
  - interactive mapping
  - collaboration
- pycsw enabled out of the box
  - embedded CSW

Open Data Catalog

- Code for America app
- open data publishing
- pycsw enabled out of the box
  - CSW embdedded

pycsw Future Development

- CKAN integration
- native spatial databases
- backends (GeoCouch)
- PostgreSQL full text search (FTS)
- enhanced harvesting / additional formats / APIs
- search engine libraries
- CSW 3.0