



CARARE Training Workshops

Stein Runar Bergheim
Asplan Viak Internet as



CARARE is funded by the **European Commission's ICT Policy Support Programme**





Introduction to Repox

An OAI-PMH tool developed within the TEL/Europeana family by the Technical University of Lisbon



CARARE is funded by the **European Commission's ICT Policy Support Programme**



Outline

Common OAI-PMH implementation scenarios in data providers

The REPOX system

- For data providers
- For service providers and aggregators

REPOX current status and future releases

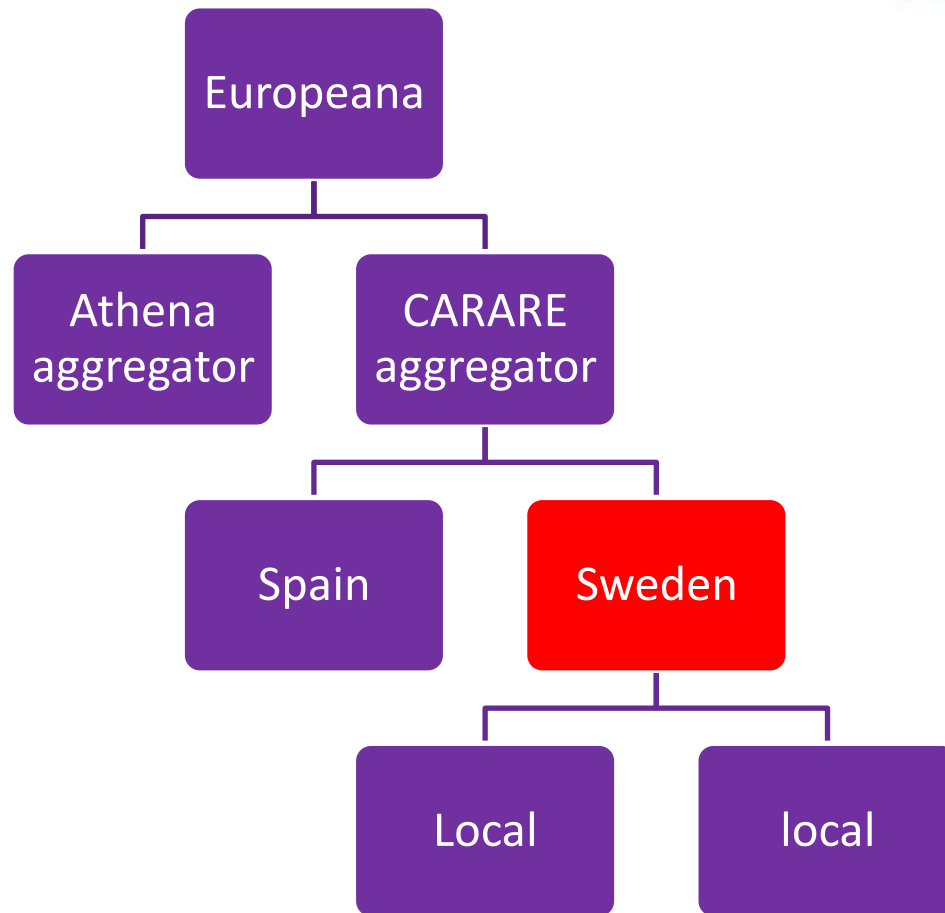


Harvesting Hierarchy

"Pure" Service Provider
Harvester

Data & Service Providers
Repository / Harvester

"Pure" Data Providers
Repository



Three OAI-PMH Scenarios

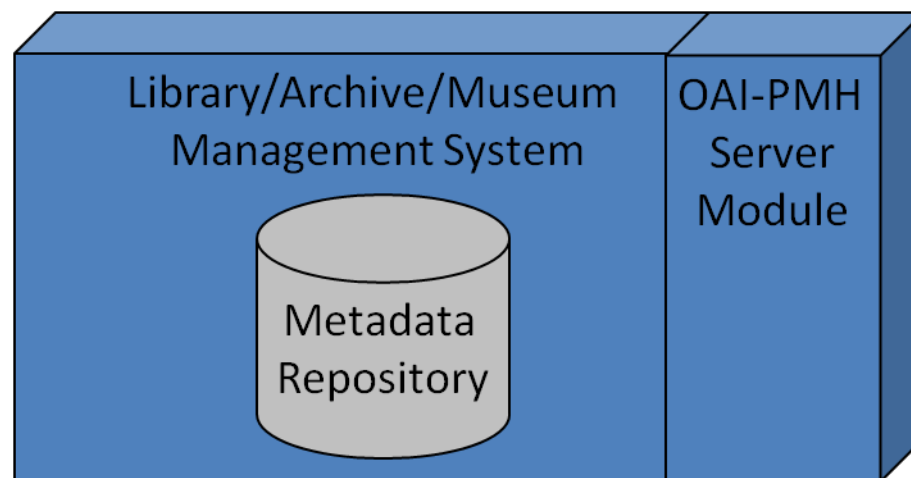
Library/Archive/Museum
Management System OAI-module
provided by the vendor

In-house OAI-PMH server
development

Standalone OAI server



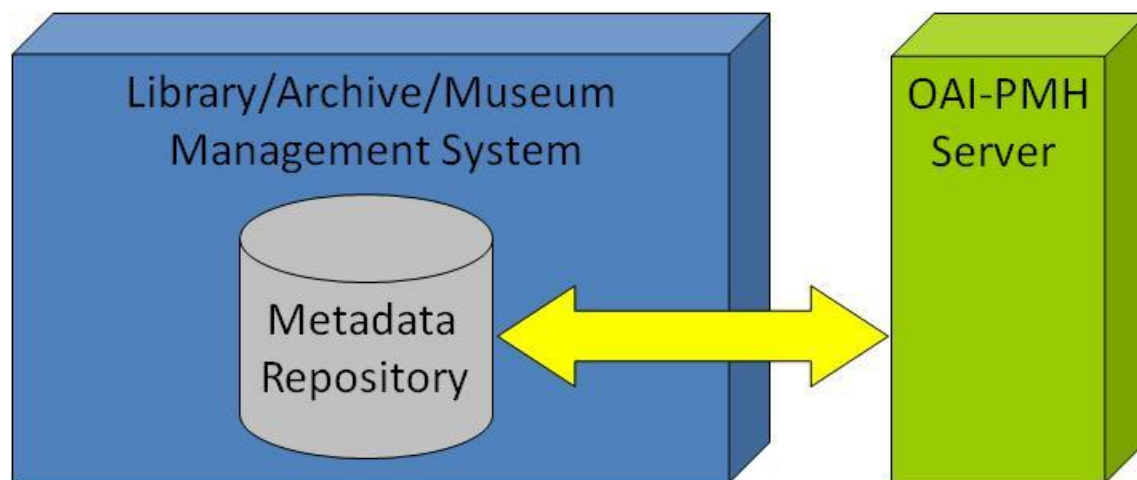
Vendor System



OAI- module provided by the vendor



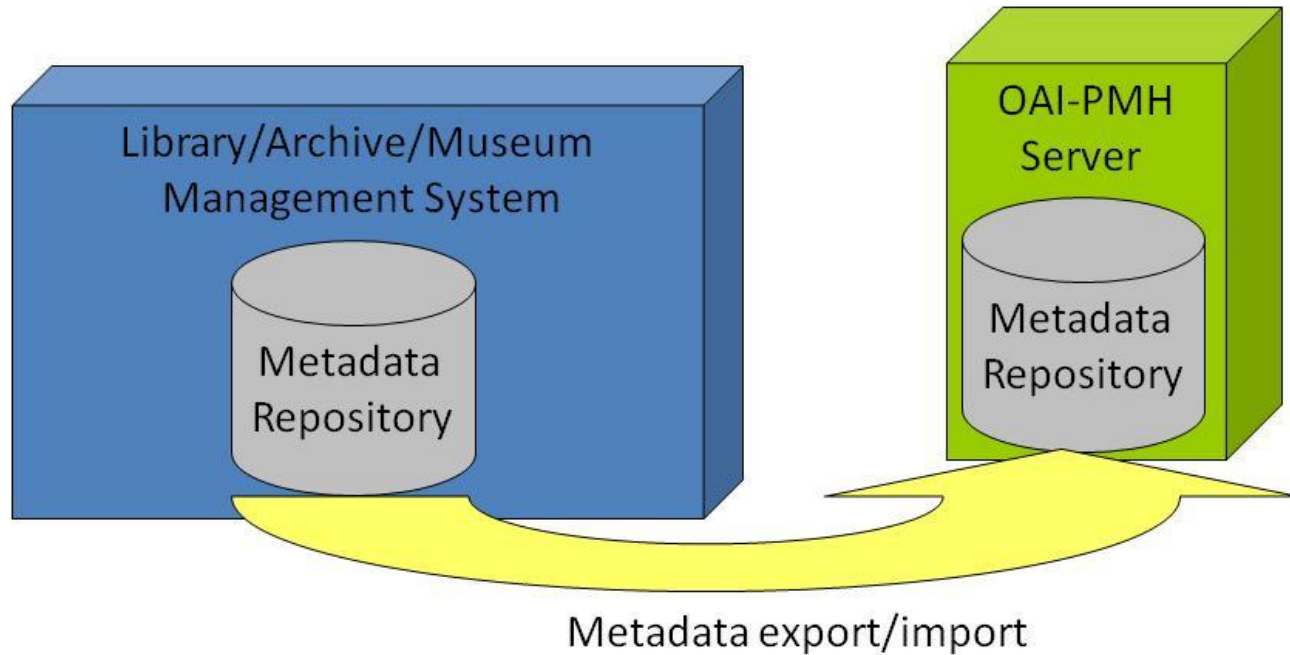
Homemade



In-house OAI-PMH server development



Standalone



Standalone OAI-PMH server



europeana

carare project



Standalone OAI-PMH server

Implementing an procedure in the library to transfer the metadata records into the OAI-PMH server

- Exporting the metadata records
 - In libraries, typically in ISO2709 or MarcXchange
- Execute the metadata conversion
 - Generating xml records in TEL, ESE, DC, CARARE
- Import the records into the OAI-Server

This process should be automated as much as possible

- Ideally fully automatic



Standalone OAI-PMH server

Advantages

- Very easy to deploy
 - May take just a few hours
- Requires only free open source software
- No software development needed
 - In some cases, the metadata crosswalks may require software development

Disadvantages

- High maintenance in cases where metadata transfer can't be fully automated



The REPOX system

Currently being developed by the
Technical University of Lisbon

May be deployed as a standalone
OAI-PMH server

- For data providers

May be deployed as an OAI-PMH
aggregator

- For service providers



The REPOX system

An open-source solution (all developed in JAVA) that can be installed locally and provide an OAI-PMH Server to the exterior

Able to manage multiple internal data sources and export multiple collection sets

REPOX reuses other open-source software...



REPOX for data providers

Provide a standalone OAI-PMH server which may be deployed without software development

– Data import:

- Supports any XML metadata format
- Supports several variations of the ISO2709 standard

– Metadata crosswalks

- User interface to define the metadata crosswalks



System Requirements

For training workshop

- Oracle Virtual Box (w/Windows XP)
- Java SE (JDK) 6.24
- Repox 1.6.3

Can also run on

- Linux and Unix servers
- Any Java SE 1.6+ enabled platform



Alternative repository tools

Many tools of varying scope, complexity and platforms

- Dspace
- Greenstone
- Fedora
- OAI-CAT
- Eprints.org
- PHP OAI Data Provider
- XMLFile v2.2 (flat files)



PMH is a standard – not a software!





OAI-PMH & Repox Exercises



europeana

carare project

CARARE is funded by the **European Commission's ICT Policy Support Programme** 



VirtualBox

Installing Oracle Virtual Box

Creating a safe, portable playground for our OAI-PMH installation and configuration



CARARE is funded by the **European Commission's ICT Policy Support Programme**



Downloading Virtual Box



The screenshot shows a Mozilla Firefox browser window displaying the VirtualBox website. The address bar shows the URL <http://www.virtualbox.org/wiki/Downloads>. The page features the VirtualBox logo on the left and a navigation menu with links for [Login](#), [Settings](#), [Register](#), and [Help/Guide](#). The main heading is "Download VirtualBox", followed by a sub-heading "VirtualBox binaries". The content explains that downloading implies agreement to the license terms and lists three categories of binaries: platform packages, the Oracle VM Extension Pack, and the Software Developer Kit (SDK). Each category includes links to specific download pages for different operating systems and architectures.

VirtualBox

[Login](#) [Settings](#) [Register](#) [Help/Guide](#)

Download VirtualBox

Here, you will find links to VirtualBox binaries and its source code.

VirtualBox binaries

By downloading, you agree to the terms and conditions of the respective license.

- **VirtualBox platform packages.** The binaries are released under the terms of the GPL version 2.
 - [VirtualBox 4.0.4 for Windows hosts](#) ⇨ [x86/amd64](#)
 - [VirtualBox 4.0.4 for OS X hosts](#) ⇨ [Intel Macs](#)
 - [VirtualBox 4.0.4 for Linux hosts](#)
 - [VirtualBox 4.0.4 for Solaris hosts](#) ⇨ [x86/amd64](#)
- **VirtualBox 4.0.4 Oracle VM VirtualBox Extension Pack** ⇨ [All platforms](#)
Support for USB 2.0 devices, VirtualBox RDP and PXE boot for Intel cards. See [this chapter from the User Manual](#) for an introduction to this Extension Pack. The Extension Pack binaries are released under the [VirtualBox Personal Use and Evaluation License \(PUEL\)](#).
- **VirtualBox 4.0.4 Software Developer Kit (SDK)** ⇨ [All platforms](#)

See the [changelog](#) for what has changed.
You might want to compare the

- ⇨ [SHA256](#) checksums or the
- ⇨ [MD5](#) checksums

to verify the integrity of downloaded packages.

In this workshop we will use Windows as a joint platform – but binaries exists for both Windows, OS X, Linux and Solaris

Step 1



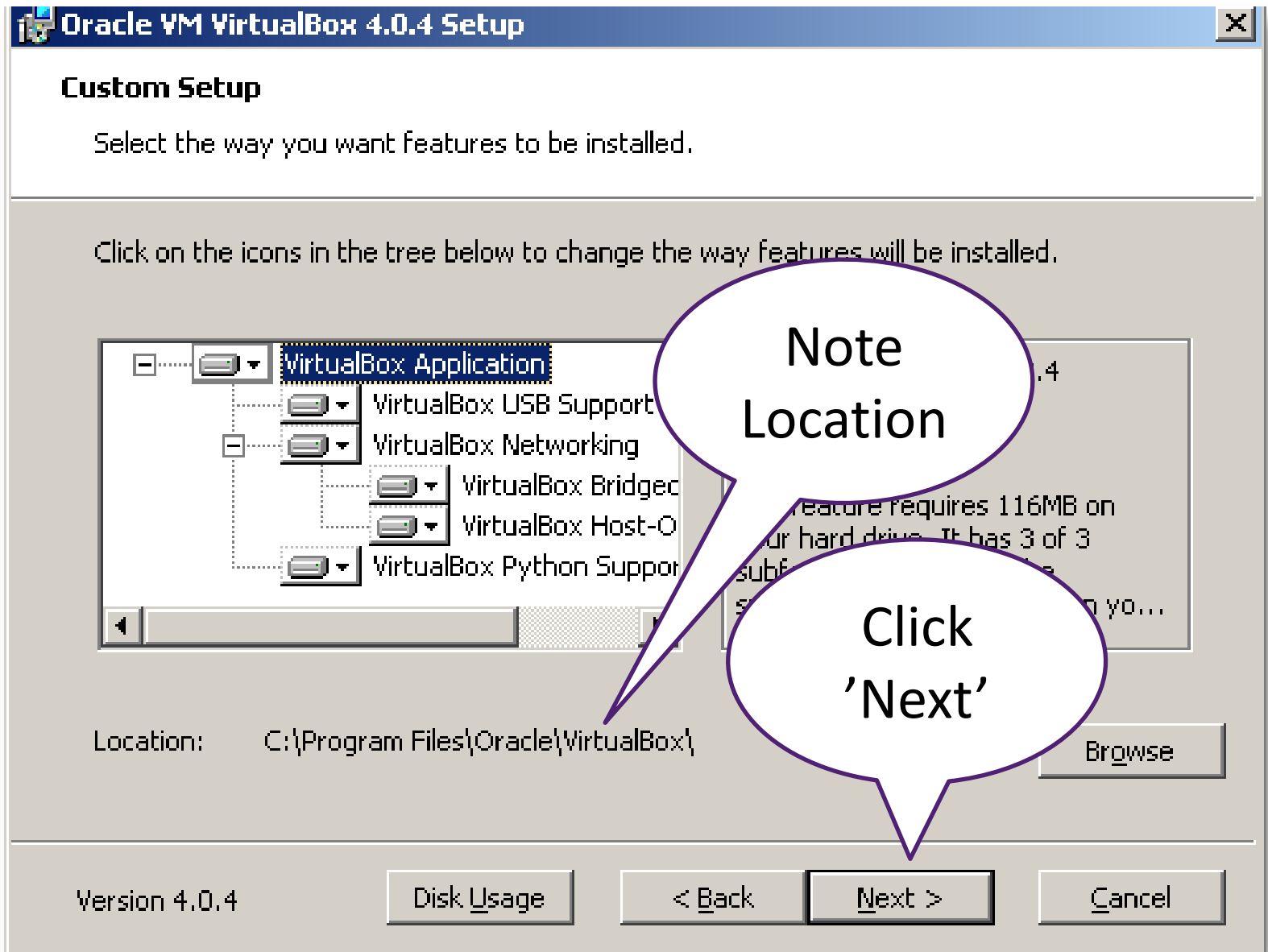
After double-clicking the 'VirtualBox' executable, the following screen may appear on Windows Vista and Windows 7.

Step 2

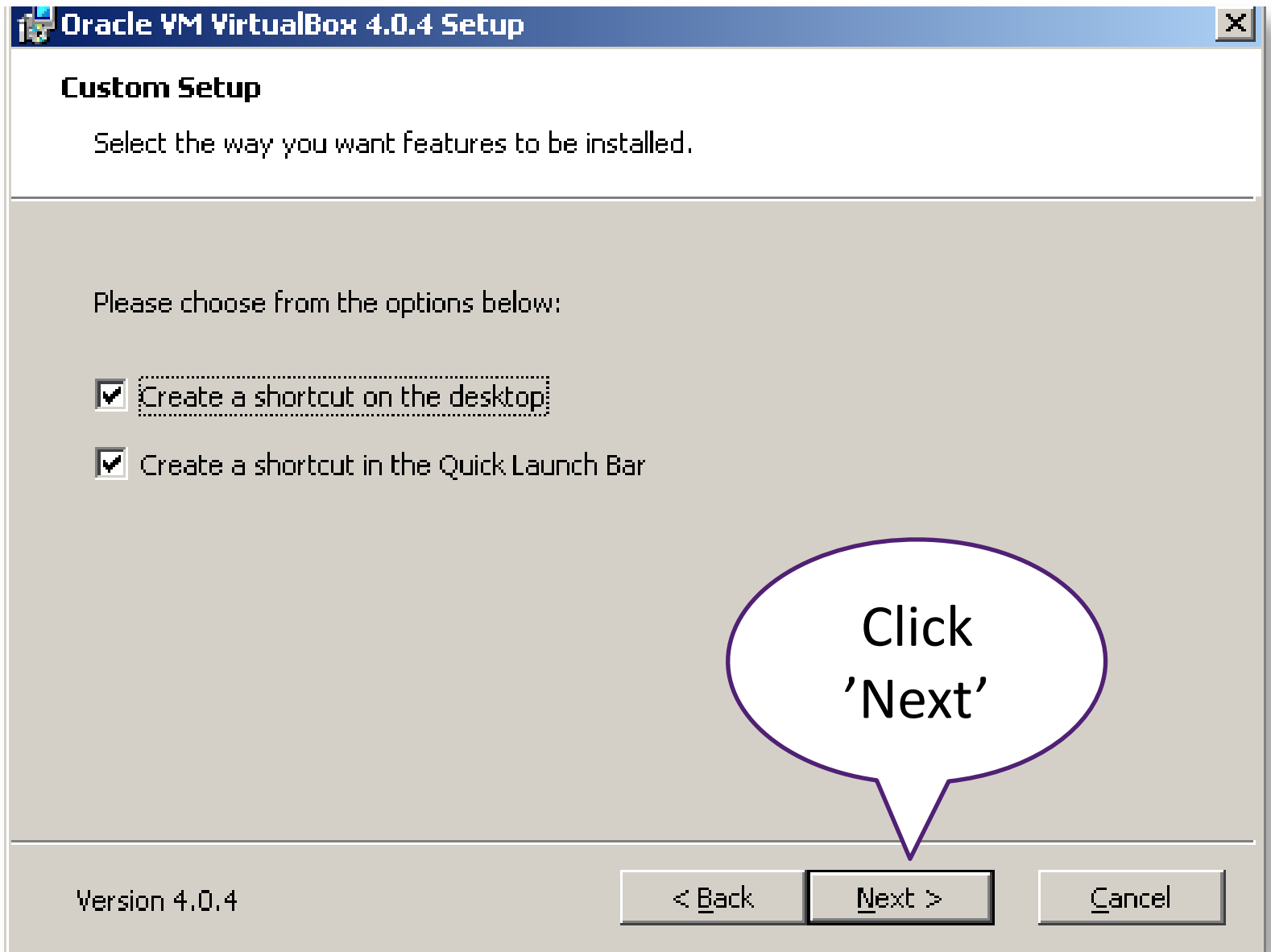


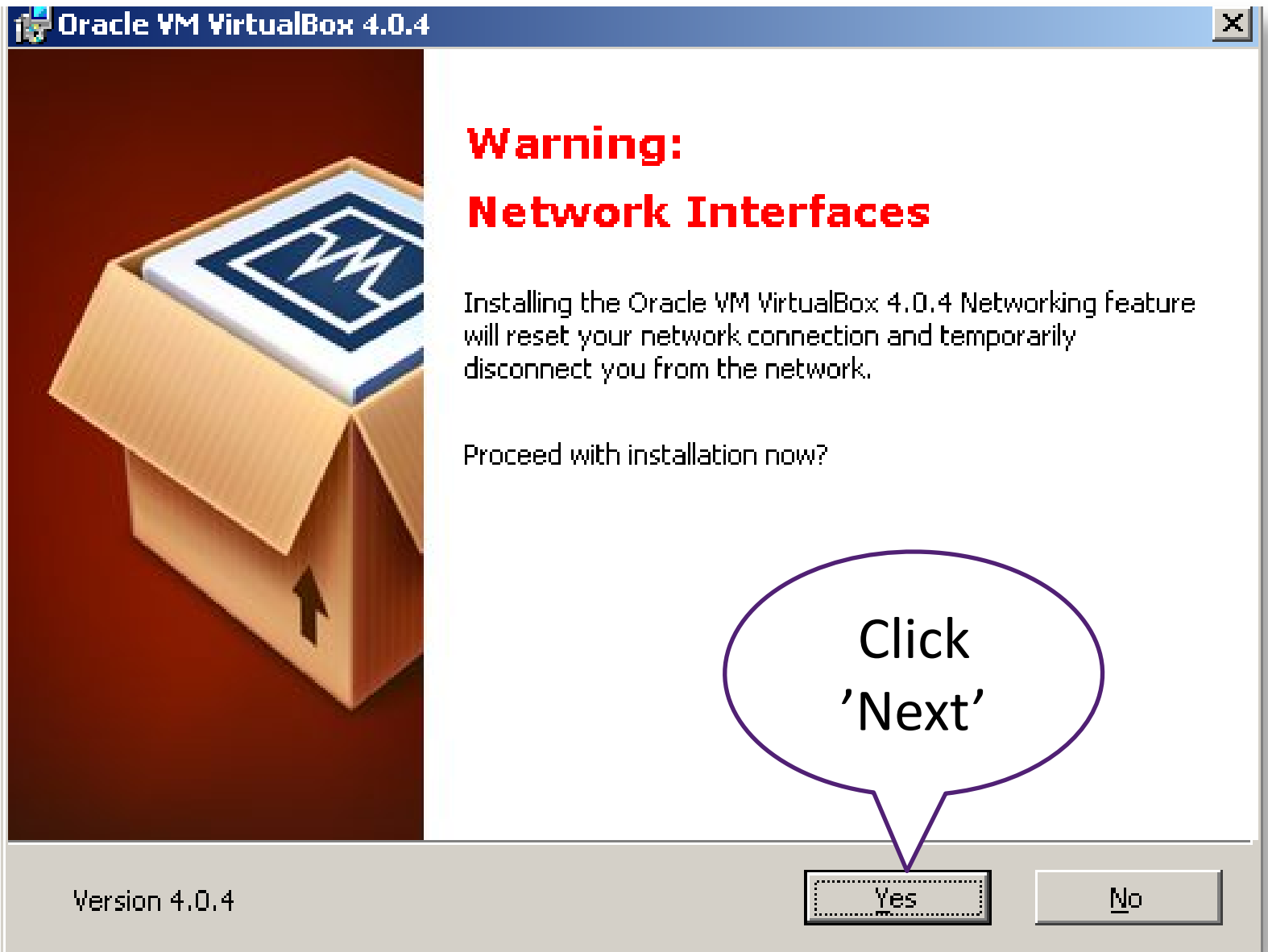
Click 'Next' to start the installation.

Step 3

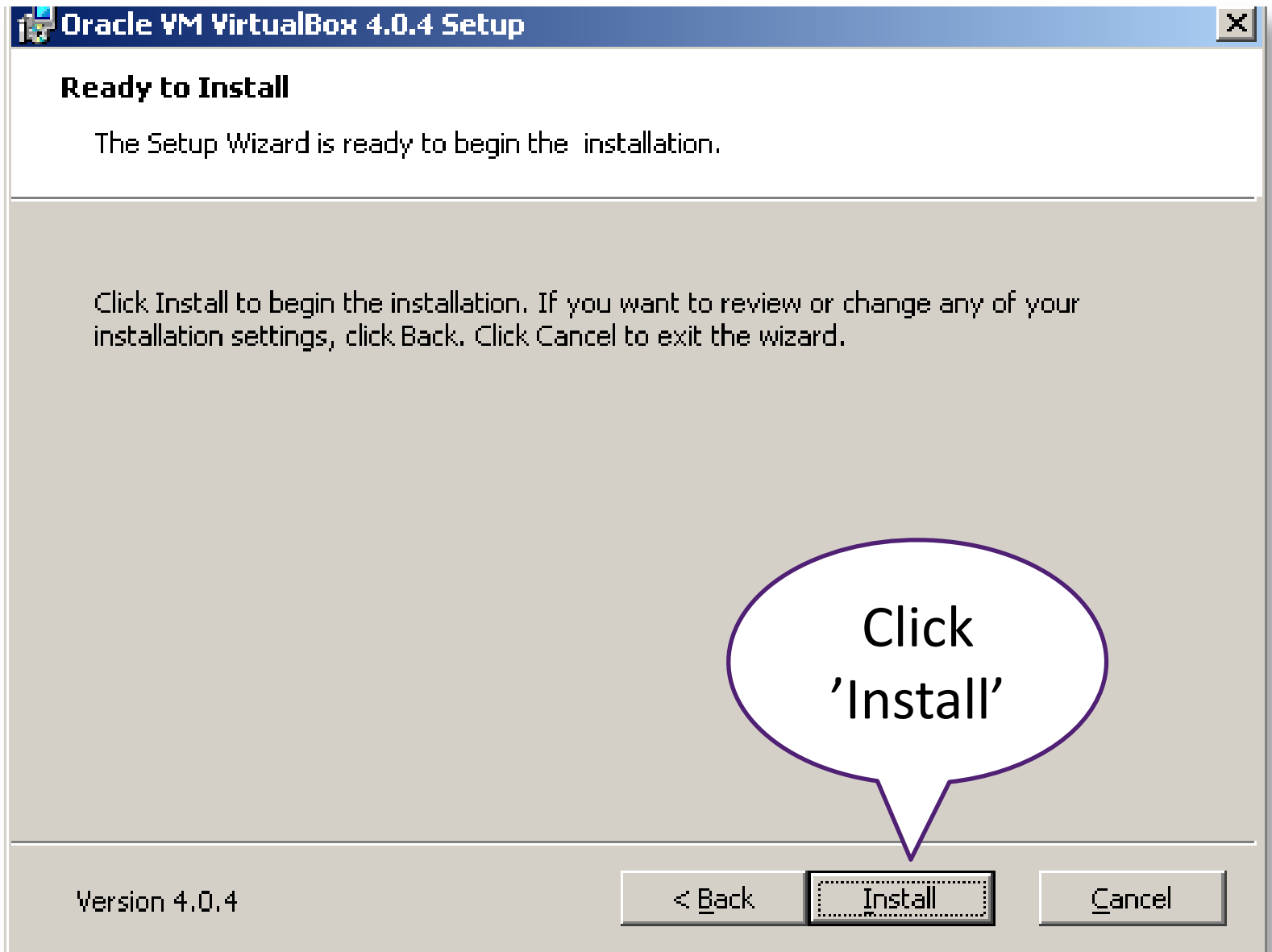


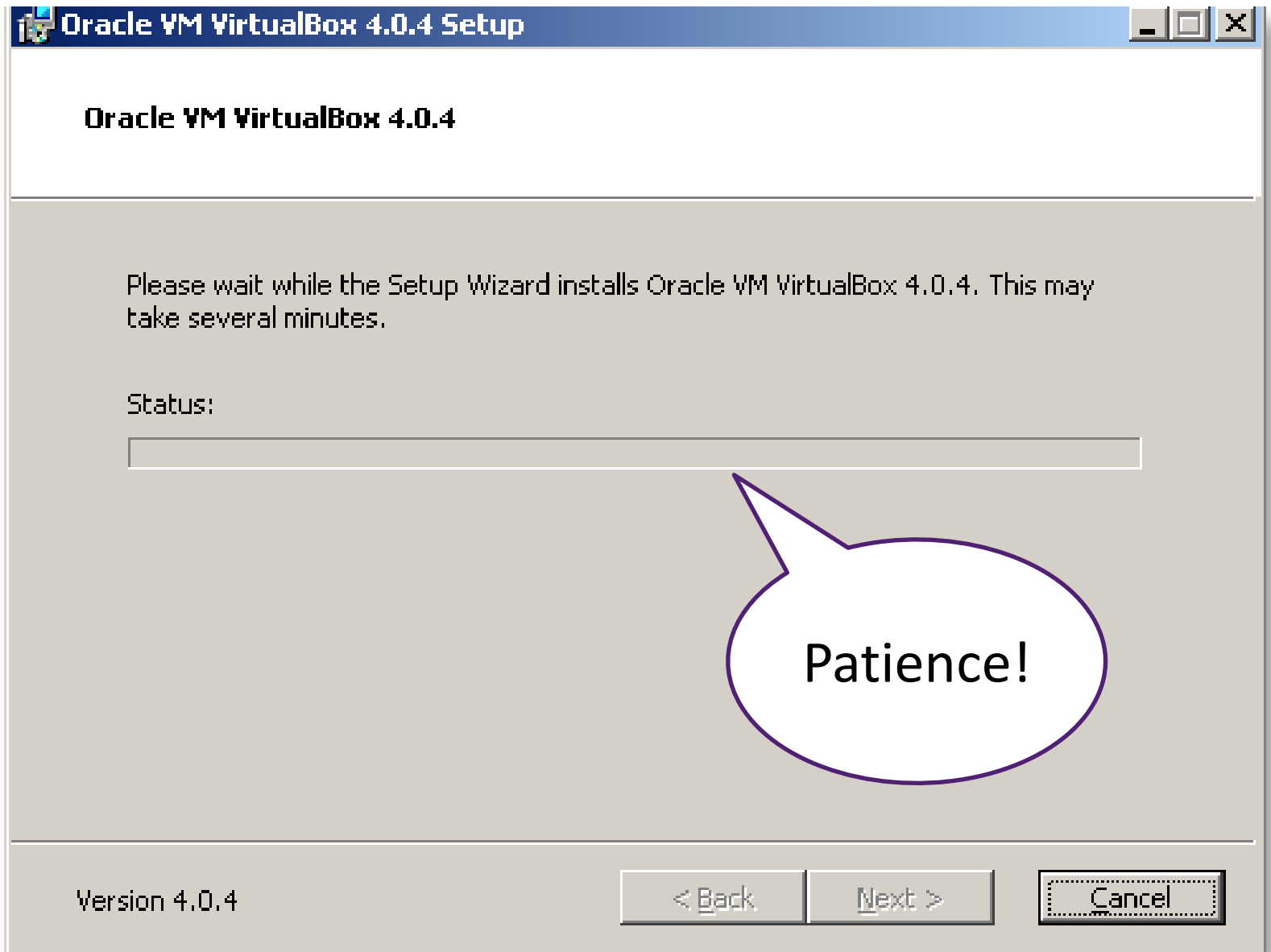
Leave the default selection, note the installation location and click 'Next' to proceed.



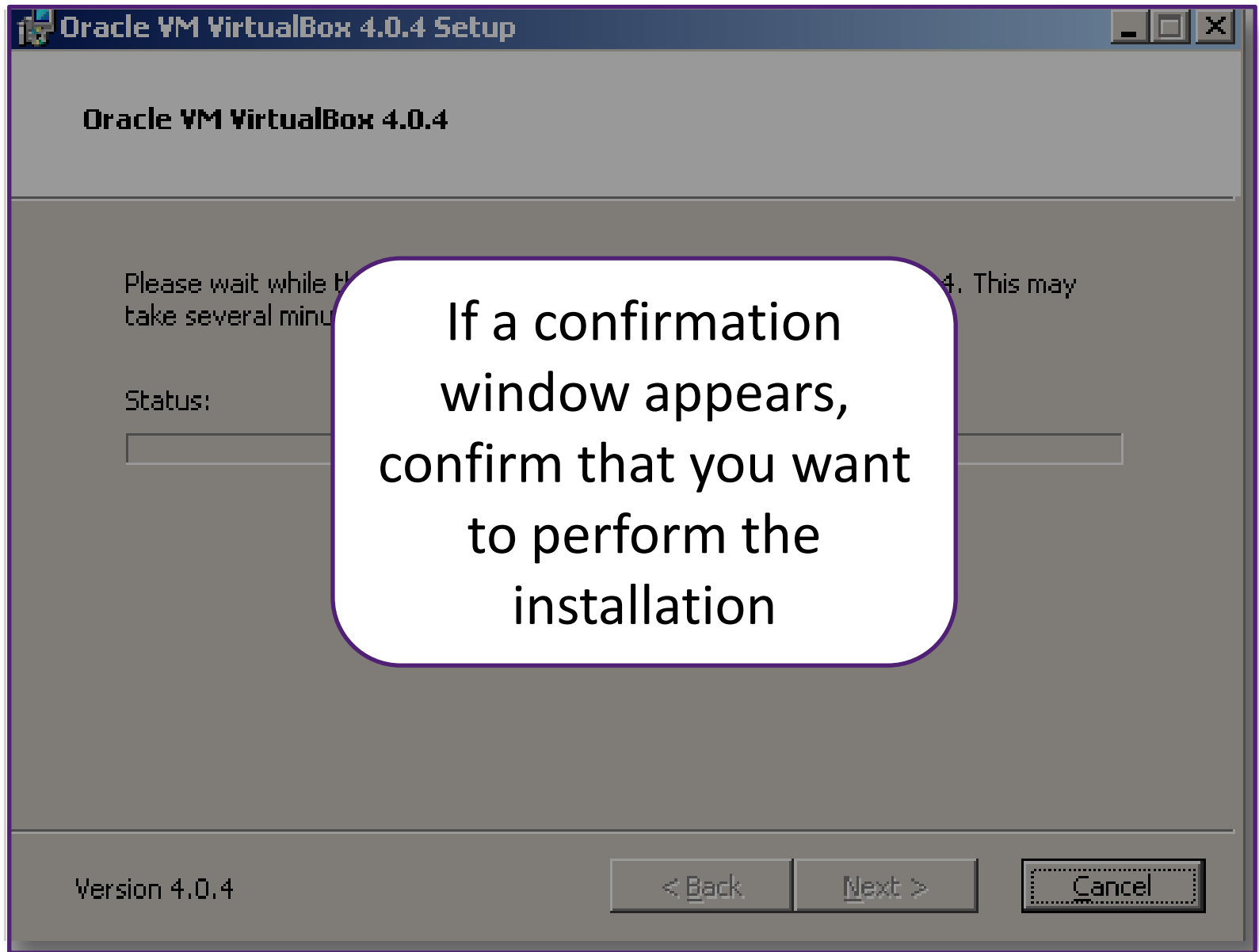


Step 6

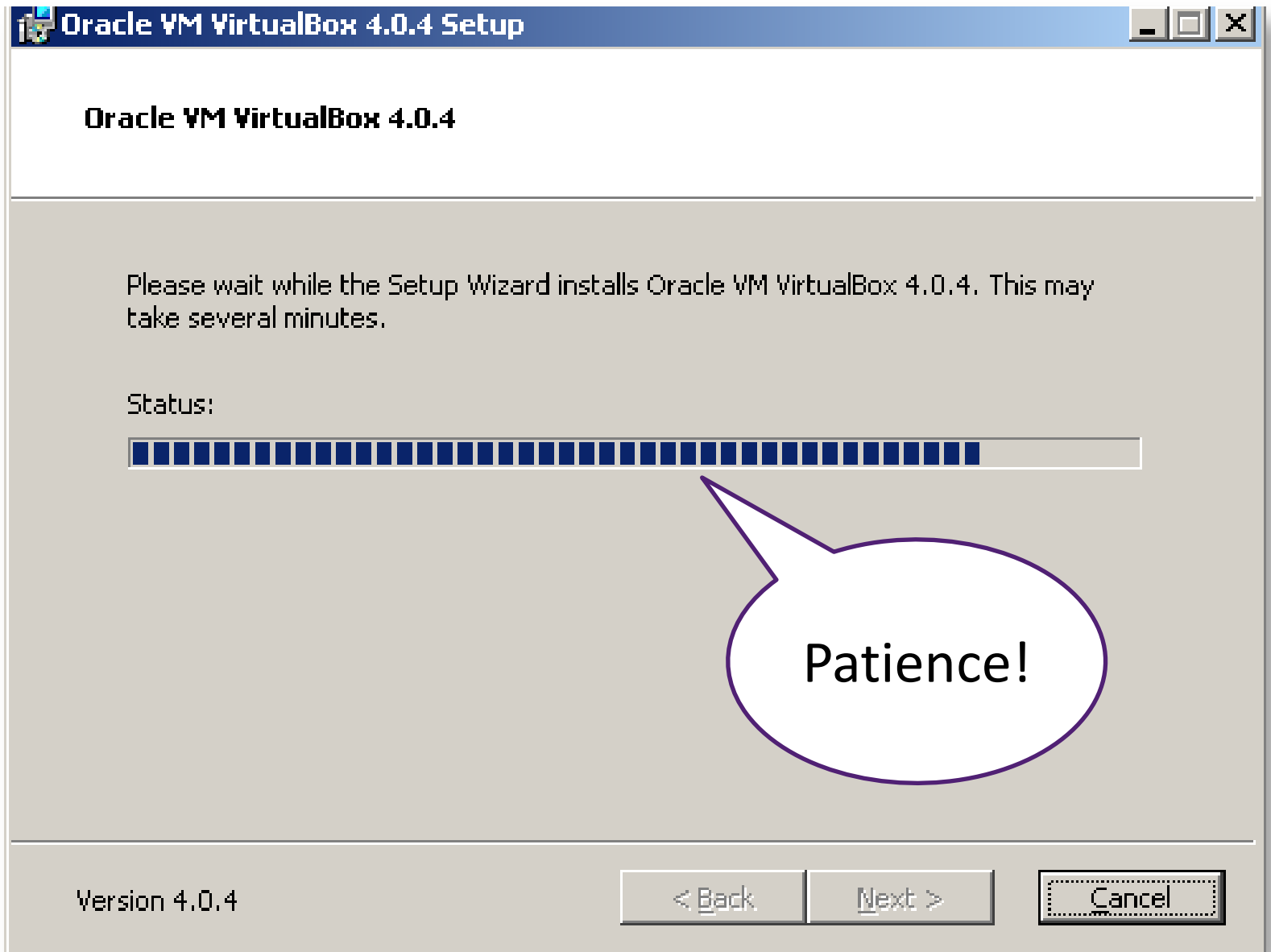




Step 8



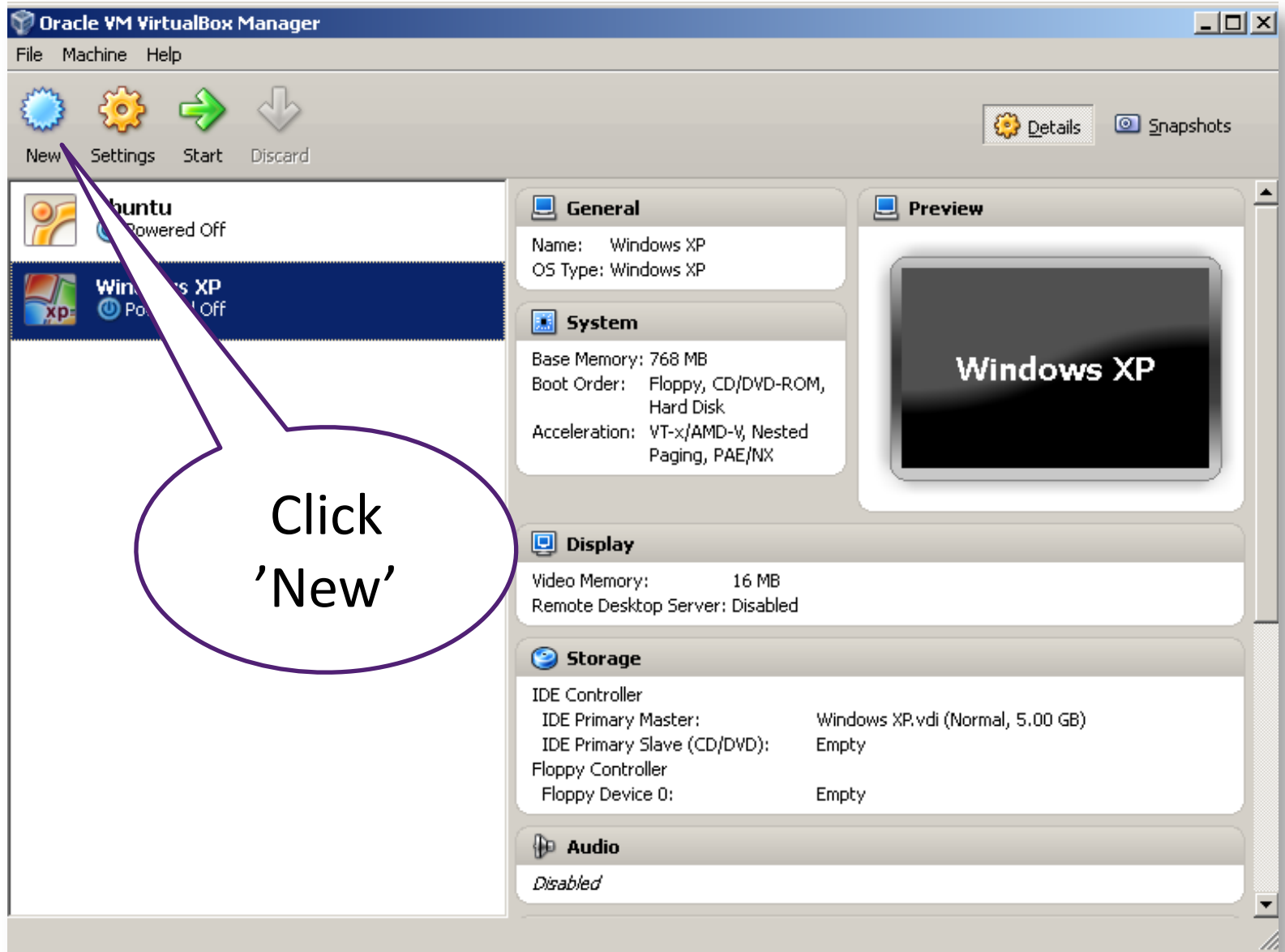
Step 9



Step 10



Step 11



New Virtual Machine

Welcome to the New Virtual Machine Wizard!

This wizard will guide you through the steps that are necessary to create a new virtual machine for VirtualBox.

Use the **Next** button to go to the next page of the wizard and the **Back** button to go to the previous page. You can also press **Cancel** if you want to cancel the execution of the wizard.



Click
'Next'

< Back

Next >

New Virtual Machine

VM Name and OS Type

Enter a name for the new virtual machine and select the type of the guest operating system you plan to install onto the virtual machine.

The name of the virtual machine will be used by all VirtualBox components and hardware configuration files for your virtual machine.

Name:

OS Type:

Operating System:

Version:

< Back Next >

Any name

Always
XP!

New Virtual Machine

Memory

Select the amount of base memory to be allocated to the virtual machine.
The recommended base memory is 512 MB.

Base Memory Size

4 MB 3072 MB

< Back Next >

New Virtual Machine

Virtual Hard Disk

Select a virtual hard disk to be used as the boot hard disk of the virtual machine. You can either create a new hard disk or select an existing one from the drop-down list or by clicking the corresponding button (to invoke file-open window).

If you need a more complicated hard disk configuration, you can create multiple virtual hard disks later using the VM Settings dialog box. You can also skip this step and attach a hard disk later.

The recommended size of the boot hard disk is 5.00 GB.

- Boot Hard Disk
- Create new hard disk
- Use existing hard disk

Windows XP.vdi (Normal, 5.00 GB)

Choose
'Use
existing'

< Back

Next >

New Virtual Machine

Summary

You are going to create a new virtual machine with the following parameters:

Name: OAI-PMH XP
OS Type: Windows XP
Base Memory: 512 MB
Boot Hard Disk: Windows XP.vdi (Normal, 5.00 GB)

If the above is correct press the **Finish** button. Once you have created a new virtual machine you can start it.

Note that you can alter these and all other settings using the **Settings** dialog accessible through the **Settings** button.

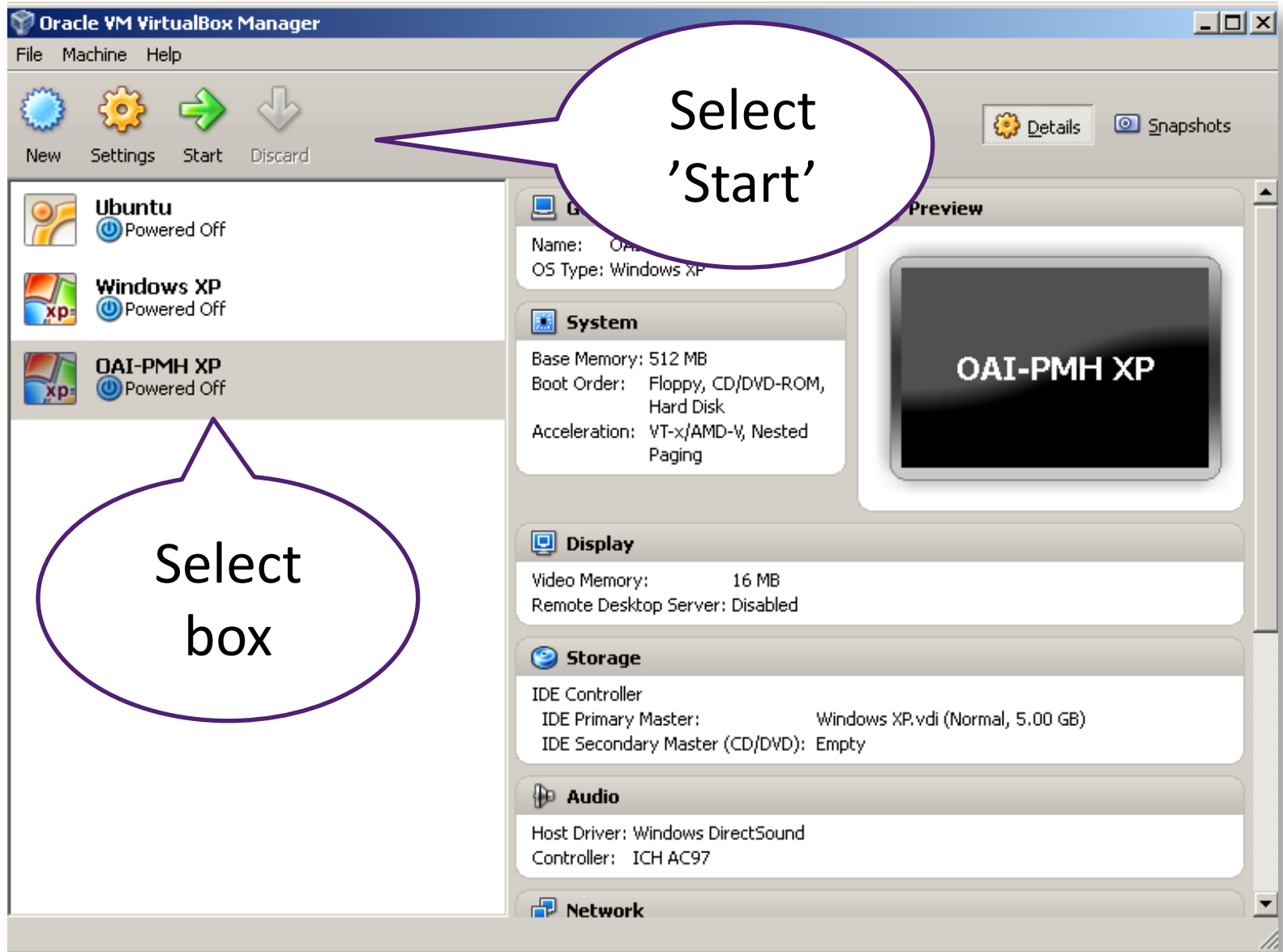


Click
'Finish'

< Back

Finish

Step 17



Installing Java SE



CARARE is funded by the **European Commission's ICT Policy Support Programme**

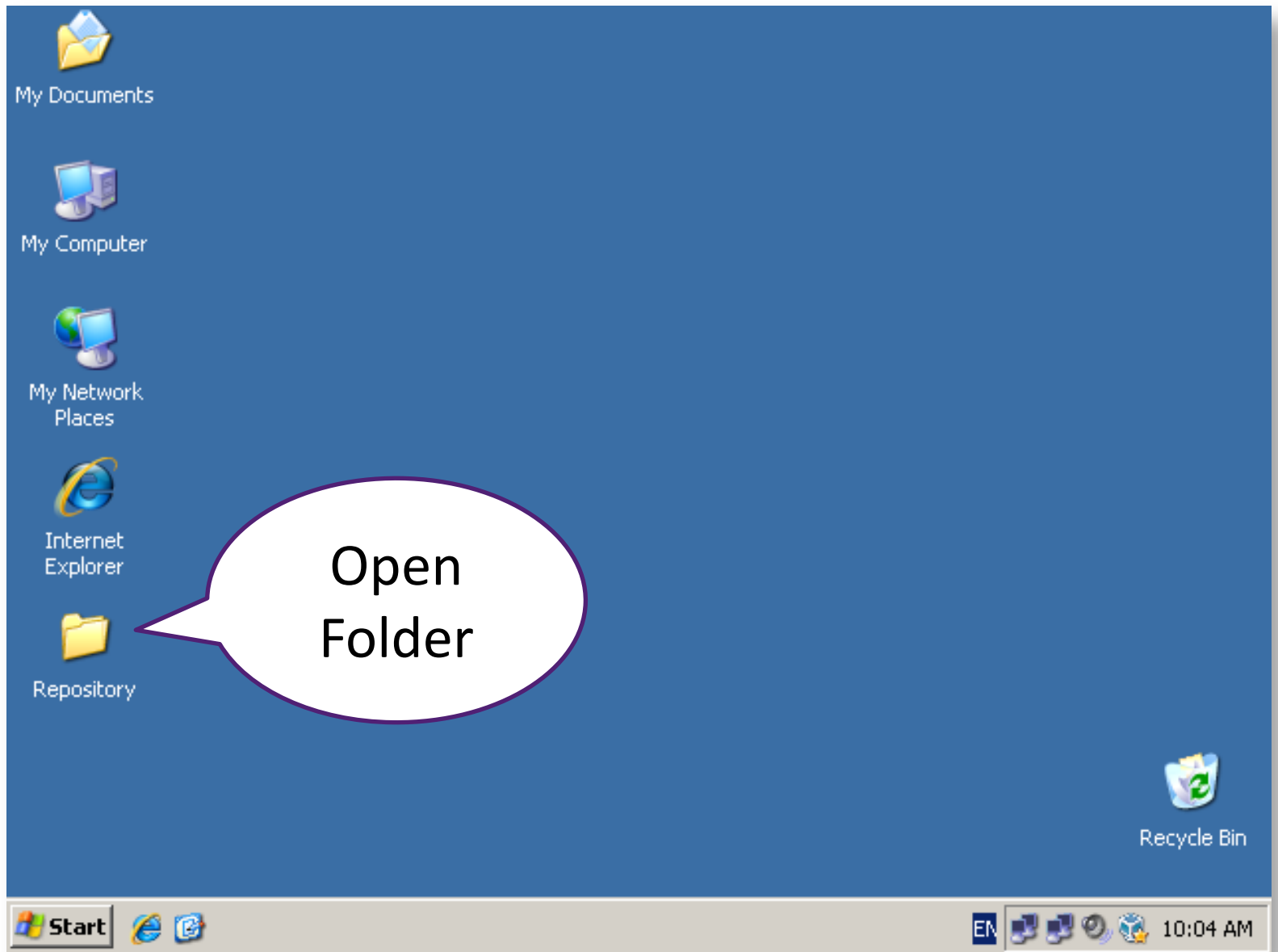


Downloading Java SE JDK

The screenshot shows the Oracle Java SE Downloads page in a Mozilla Firefox browser window. The browser's address bar displays the URL: <http://www.oracle.com/technetwork/java/javase/downloads/index.html>. The page features the Oracle logo at the top left, a navigation menu with options like 'Products and Services', 'Downloads', 'Store', 'Support', 'Education', 'Partners', and 'About', and a search bar. The main content area is titled 'Java SE Downloads' and includes a sidebar with a list of Java products (Java SE, Java for Business, Java Embedded, Java EE, Java ME, Java FX, Java DB, Web Tier, Java Card, Java TV, Community). The main content area has tabs for 'Overview', 'Downloads', 'Documentation', 'Community', 'Technologies', and 'Training'. Under the 'Downloads' tab, there are four download buttons: 'Latest Release', 'Next Release (Early Access)', 'Embedded Use', 'Real-Time', and 'Previous Releases'. Below these are four download buttons for 'Java Platform (JDK)', 'JDK + JavaFX Bundle', 'JDK + NetBeans Bundle', and 'JDK + Java EE Bundle'. The 'Java Platform (JDK)' button is highlighted, and the text below it reads: 'Here are the Java SE downloads in detail.' and 'Java Platform, Standard Edition Java SE 6 Update 24'. The status bar at the bottom of the browser window shows 'Done' and 'Tor Disabled'.

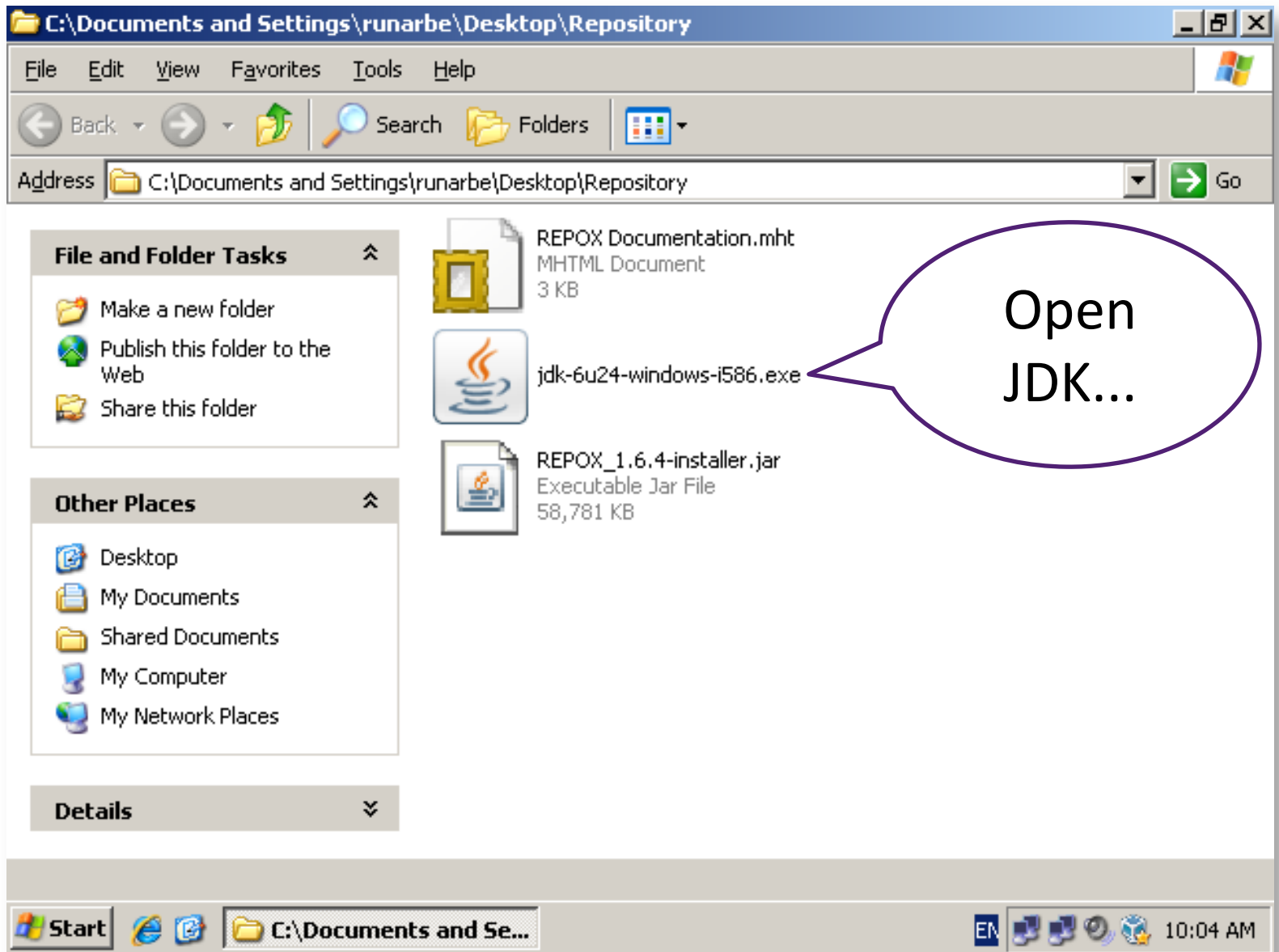
Similarly as for Virtual Box, binaries exists for all platforms including both Windows, OS X, Linux and Solaris

Step 1



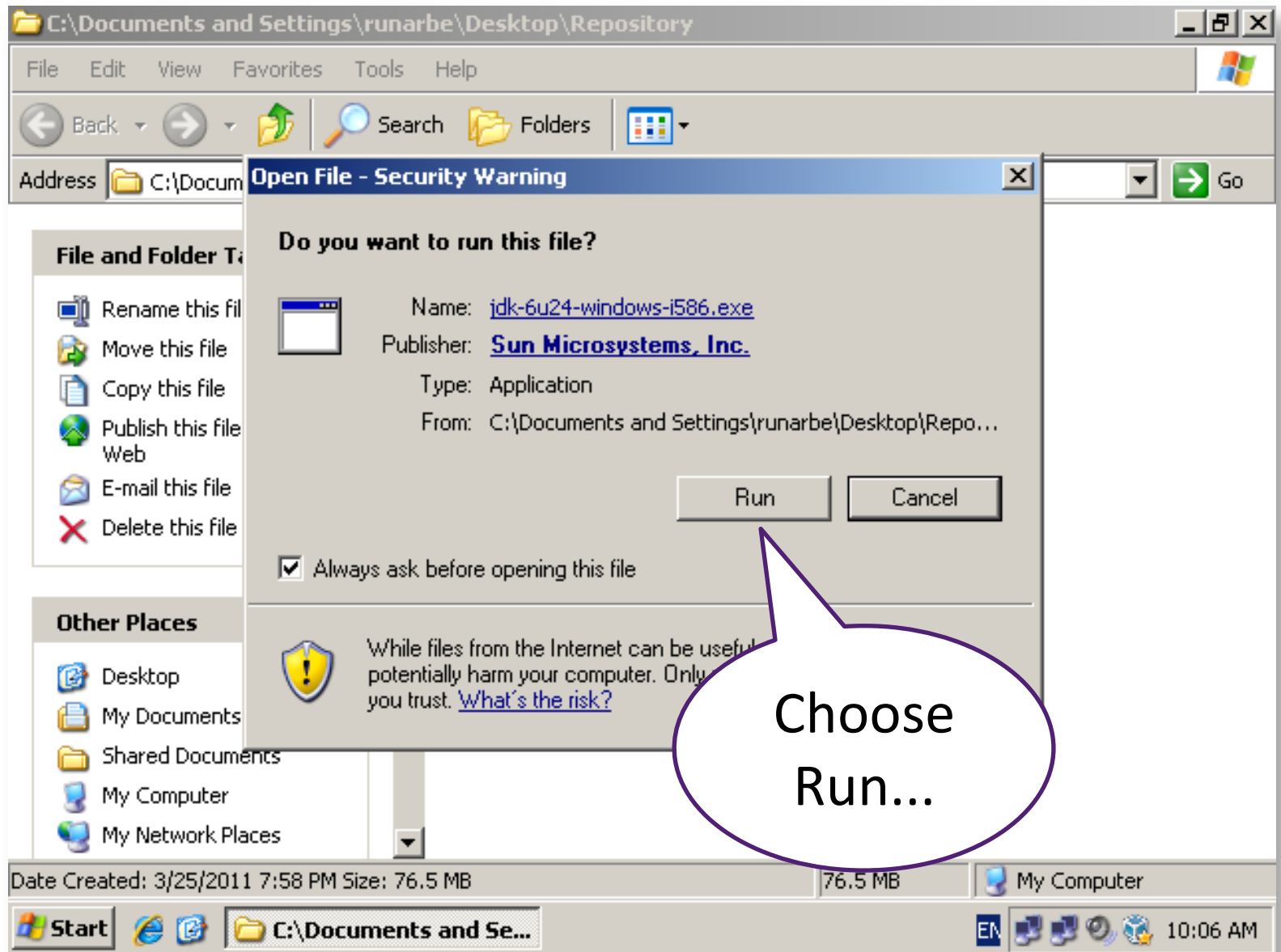
All installation files have been placed in a folder on your desktop

Step 2



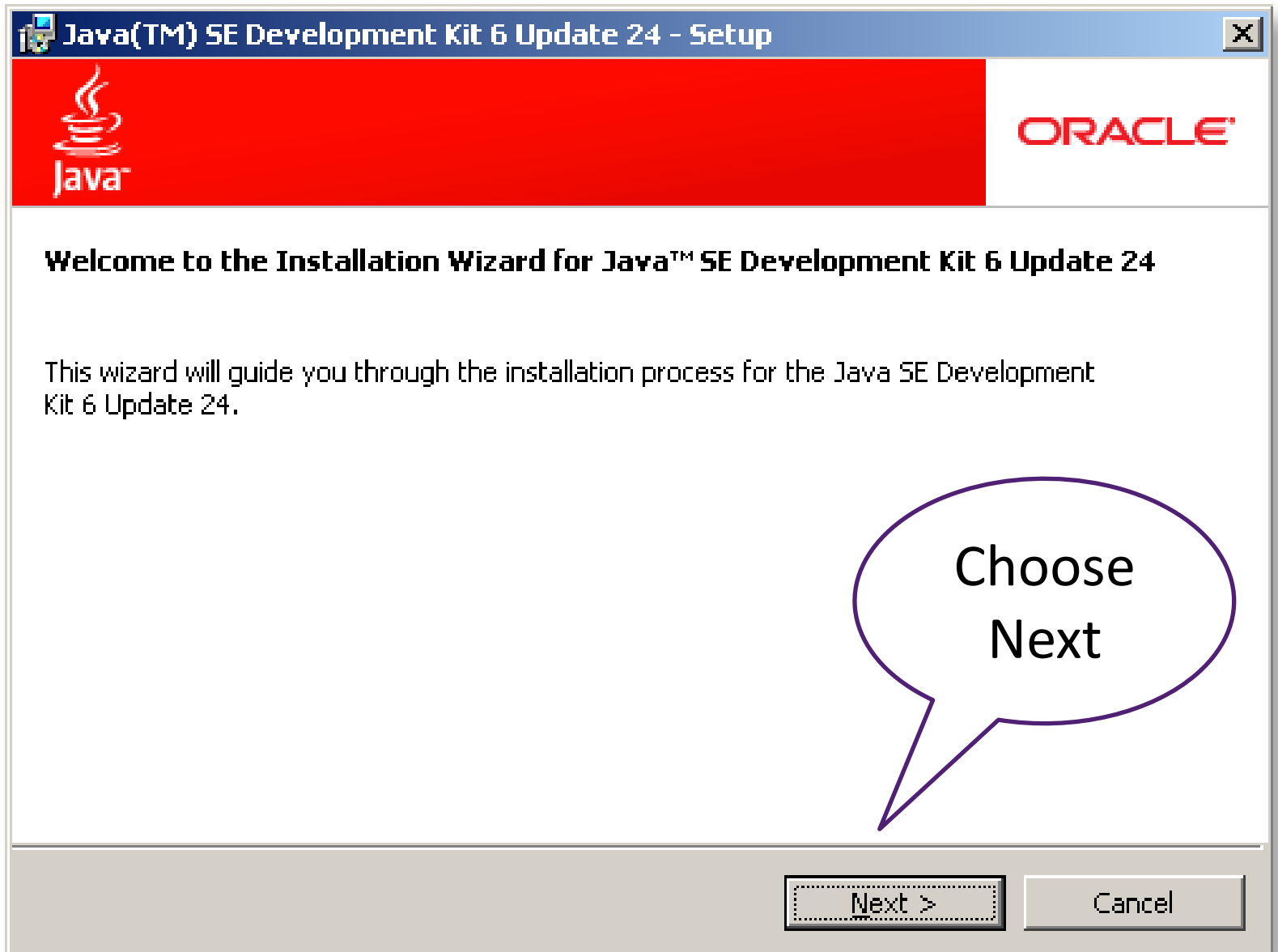
Wait until the installation starts – it may take some time – especially on slower/older machines

Step 3



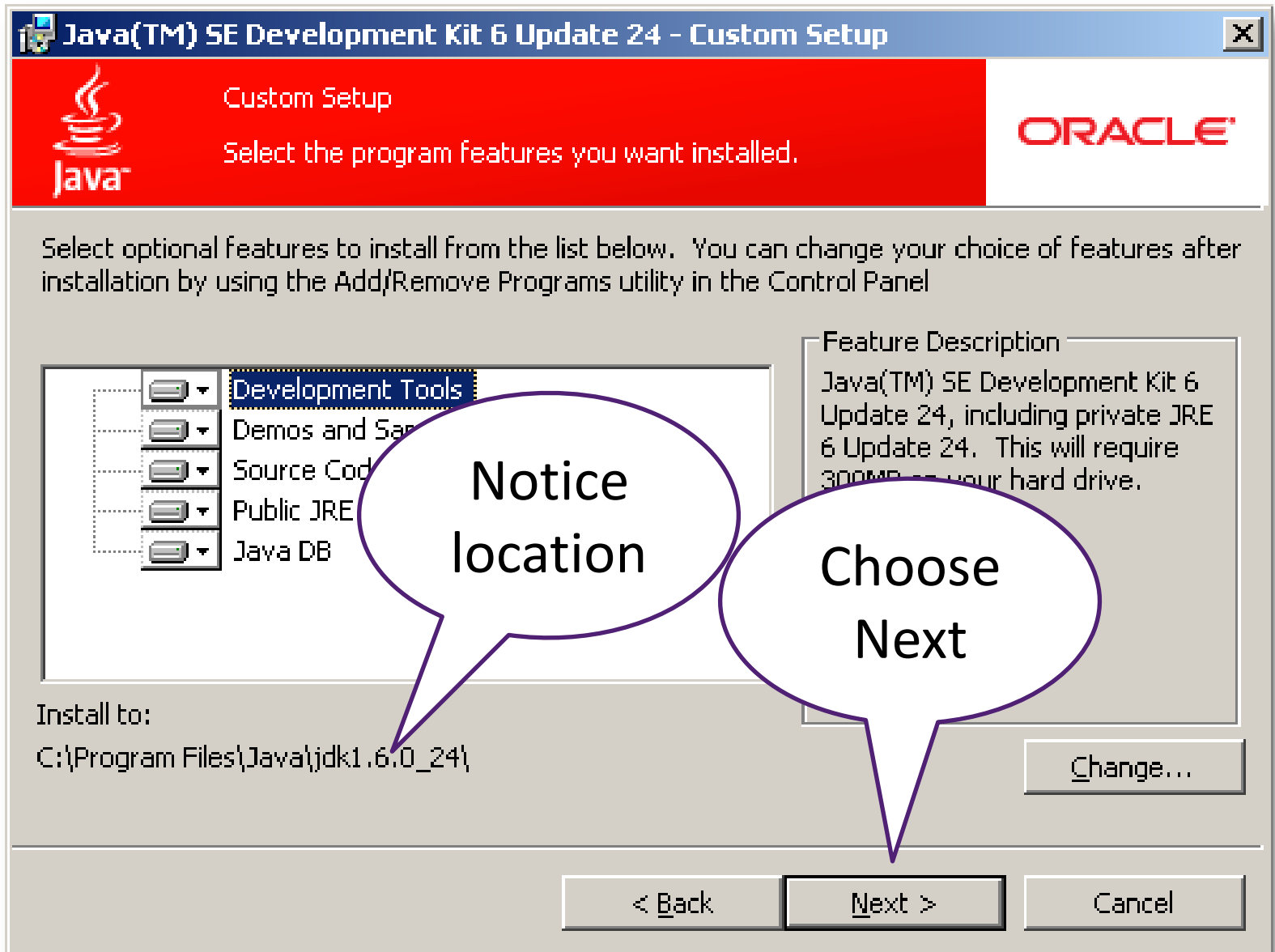
Follow on-screen instructions...

Step 4



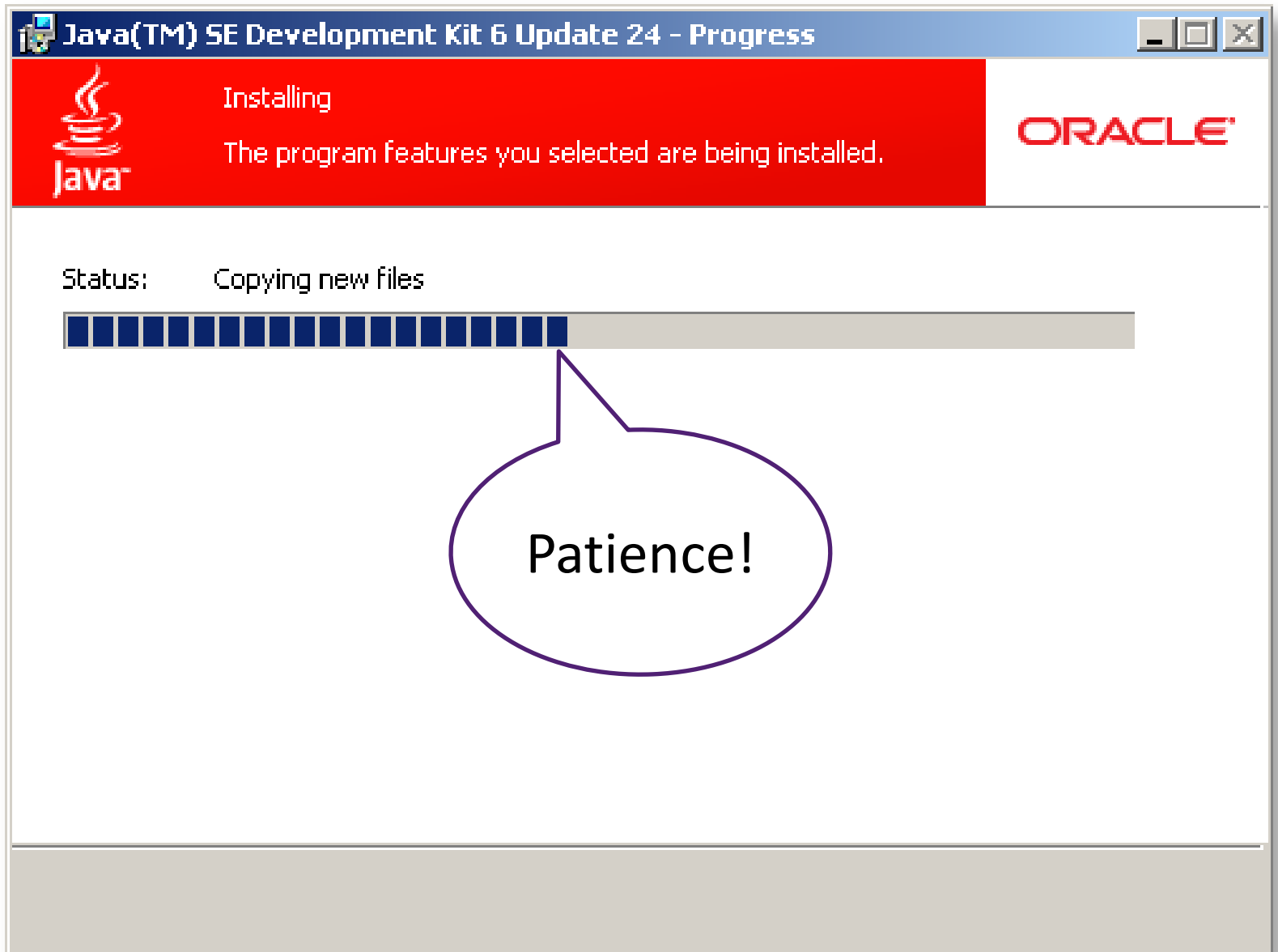
Follow on-screen instructions...

Step 5



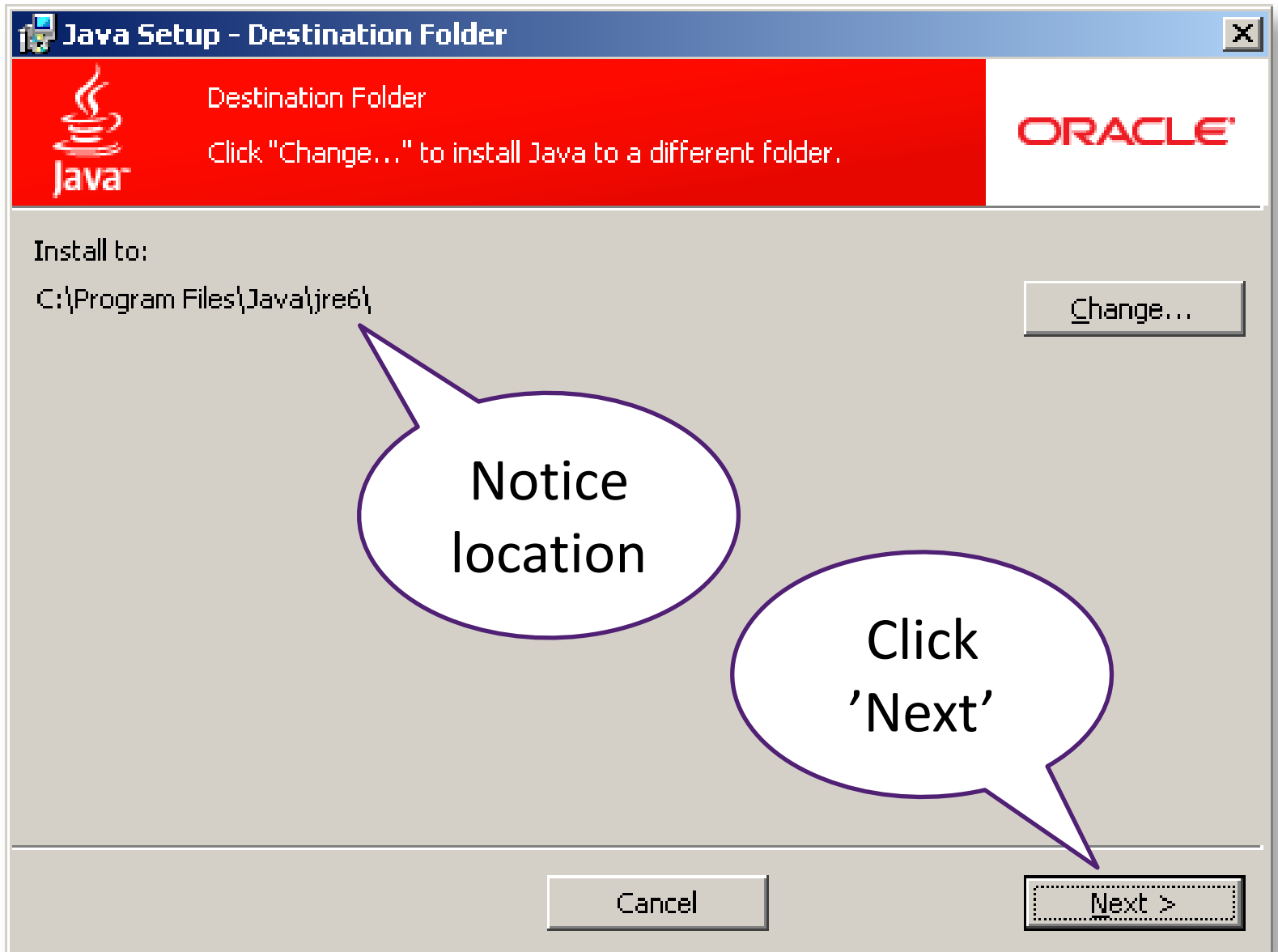
Make no changes in this dialog window but notice where you install the software and choose next...

Step 6



Follow on-screen instructions...

Step 7

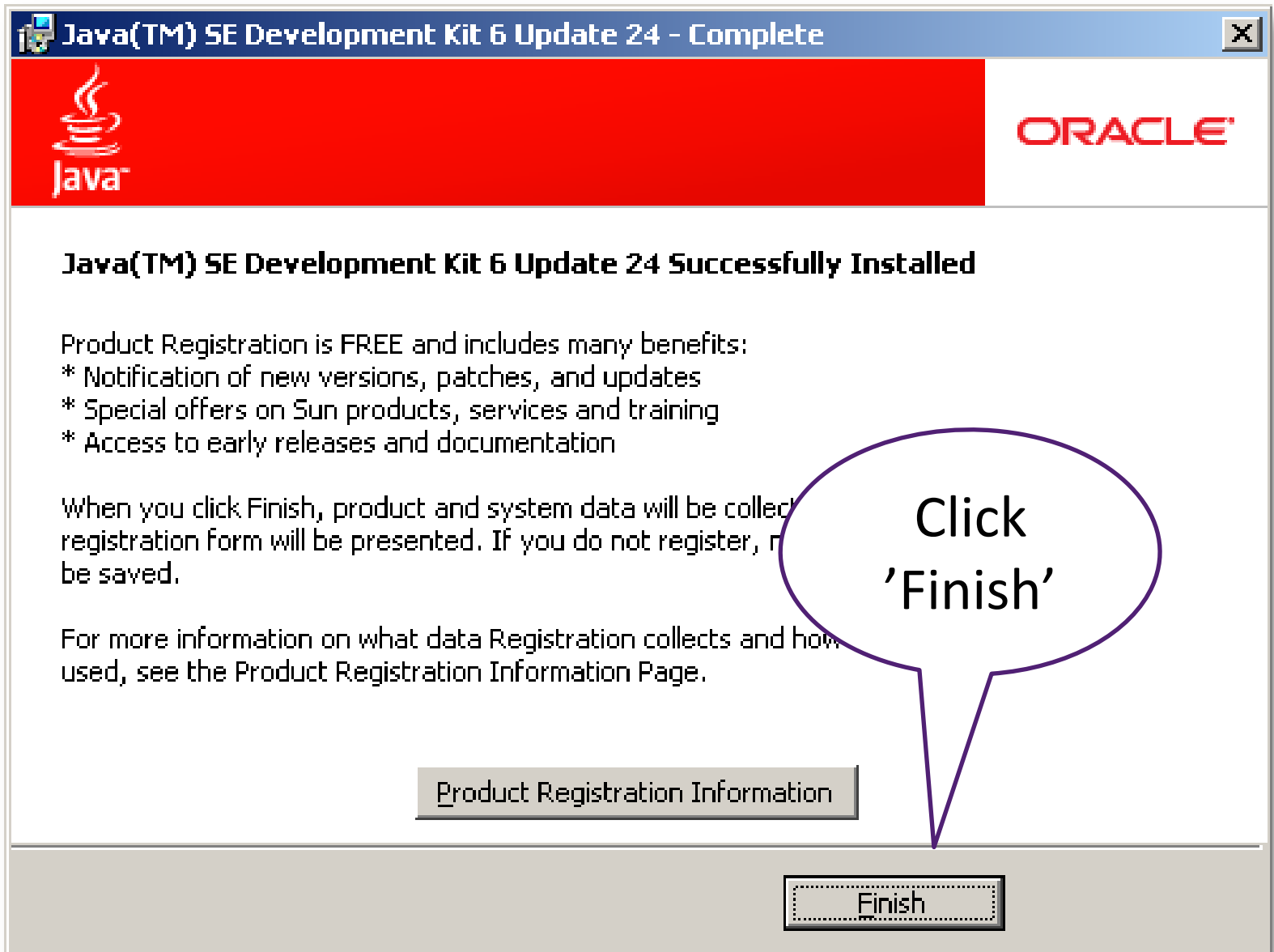


Follow on-screen instructions...

Step 8



Follow on-screen instructions...



Follow on-screen instructions...



A Metadata Space Manager

Installing Repox



CARARE is funded by the **European Commission's ICT Policy Support Programme**



Download Repox

REPOX - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://repop.ist.utl.pt/download.html

Webcam Tools BasecampHQ Carare Tool

REPOX

INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

inesc id
lisboa

Home Download

Documentation

Data formats

Download

Known Issues

Contacts

REPOX installations

To install REPOX, Java 6 or higher is required. Java 6 can be downloaded from:
<http://java.sun.com/javase/downloads/index.jsp>.
See also the [installation instructions](#). [Version History](#)

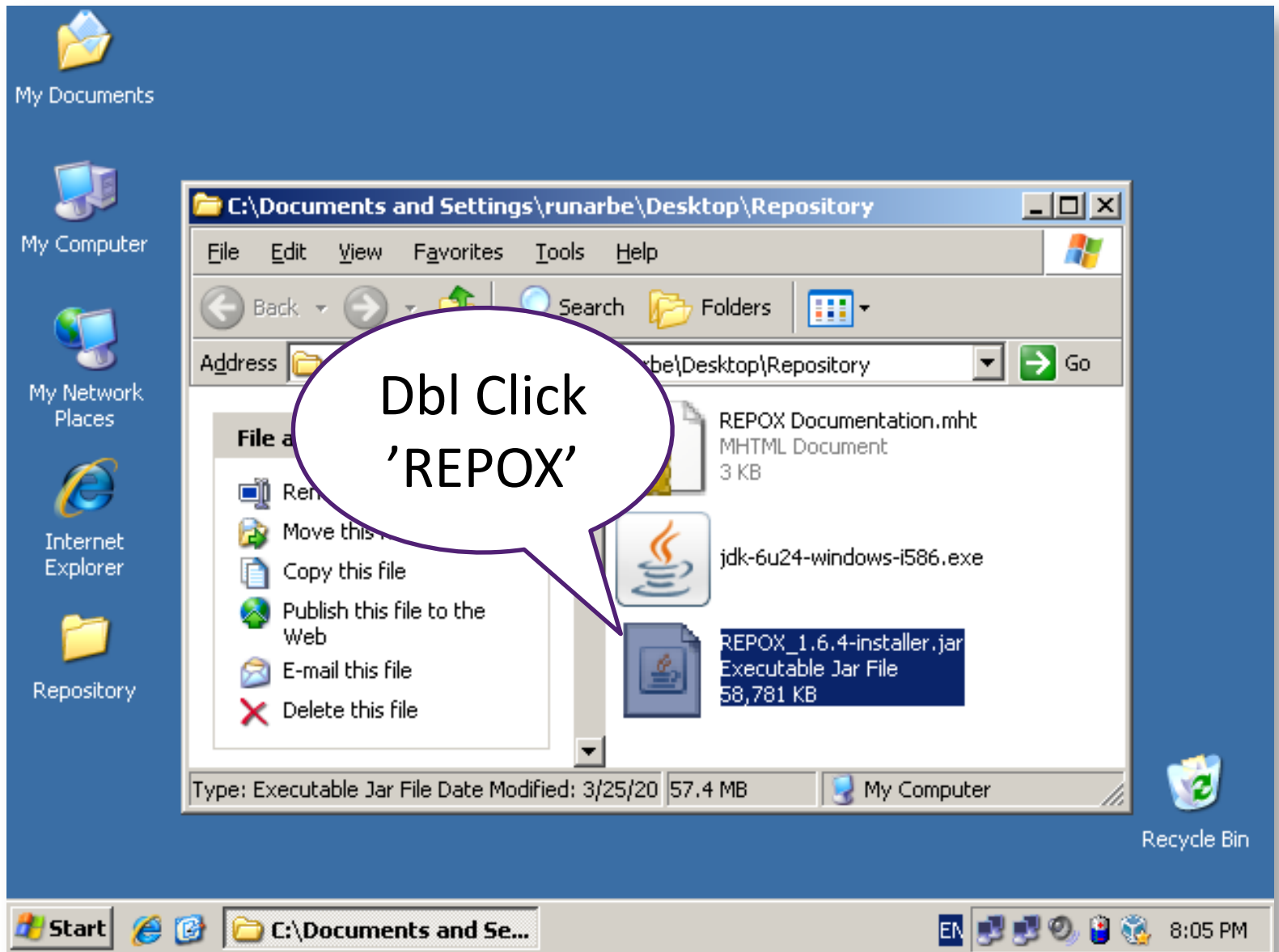
REPOX

- [REPOX v1.6.4](#) (2011-03-03)
- [REPOX v1.6.4 Linux](#) (2011-03-03)
- [REPOX v1.6.3](#) (2011-01-19)
- [REPOX v1.6.3 Linux](#) (2011-01-19)
- [REPOX v1.6.2](#) (2011-01-06)
- [REPOX v1.6.2 Linux](#) (2011-01-06)
- [REPOX v1.5.6](#) (2010-11-24)
- [REPOX v1.5.6 Linux](#) (2010-11-24)
- [REPOX v1.5.5](#) (2010-10-20)
- [REPOX v1.5.5 Linux](#) (2010-10-20)
- [REPOX v1.5.4](#) (2010-10-07)
- [REPOX v1.5.4 Linux](#) (2010-10-07)
- [REPOX v1.5.3](#) (2010-07-12)
- [REPOX v1.5.3 Linux](#) (2010-10-07)

Done Tor Disabled

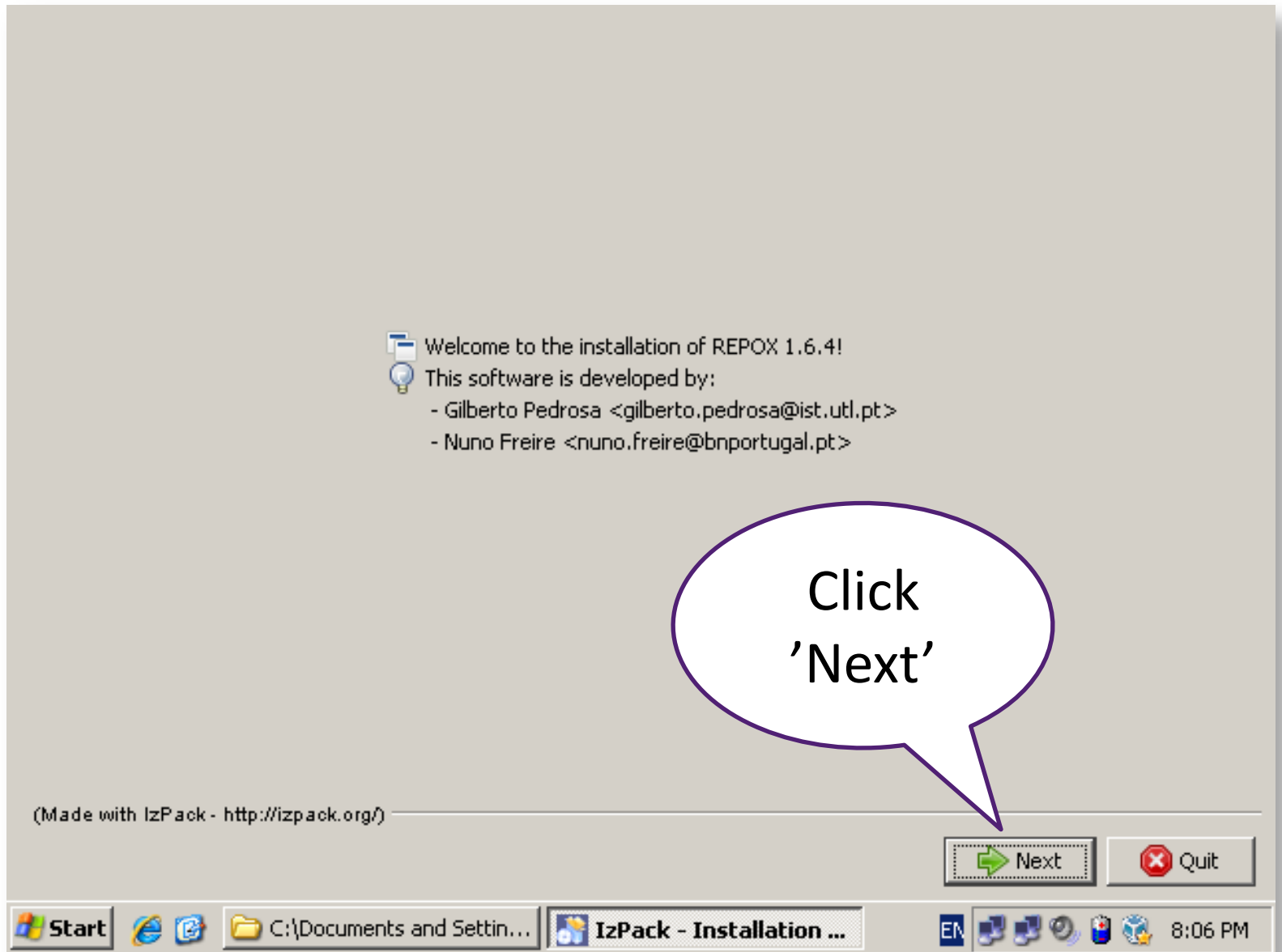
Repop is written in Java and runs independent of operating system. Most recent version is 1.6.4 as per 28th of March 2011

Step 1



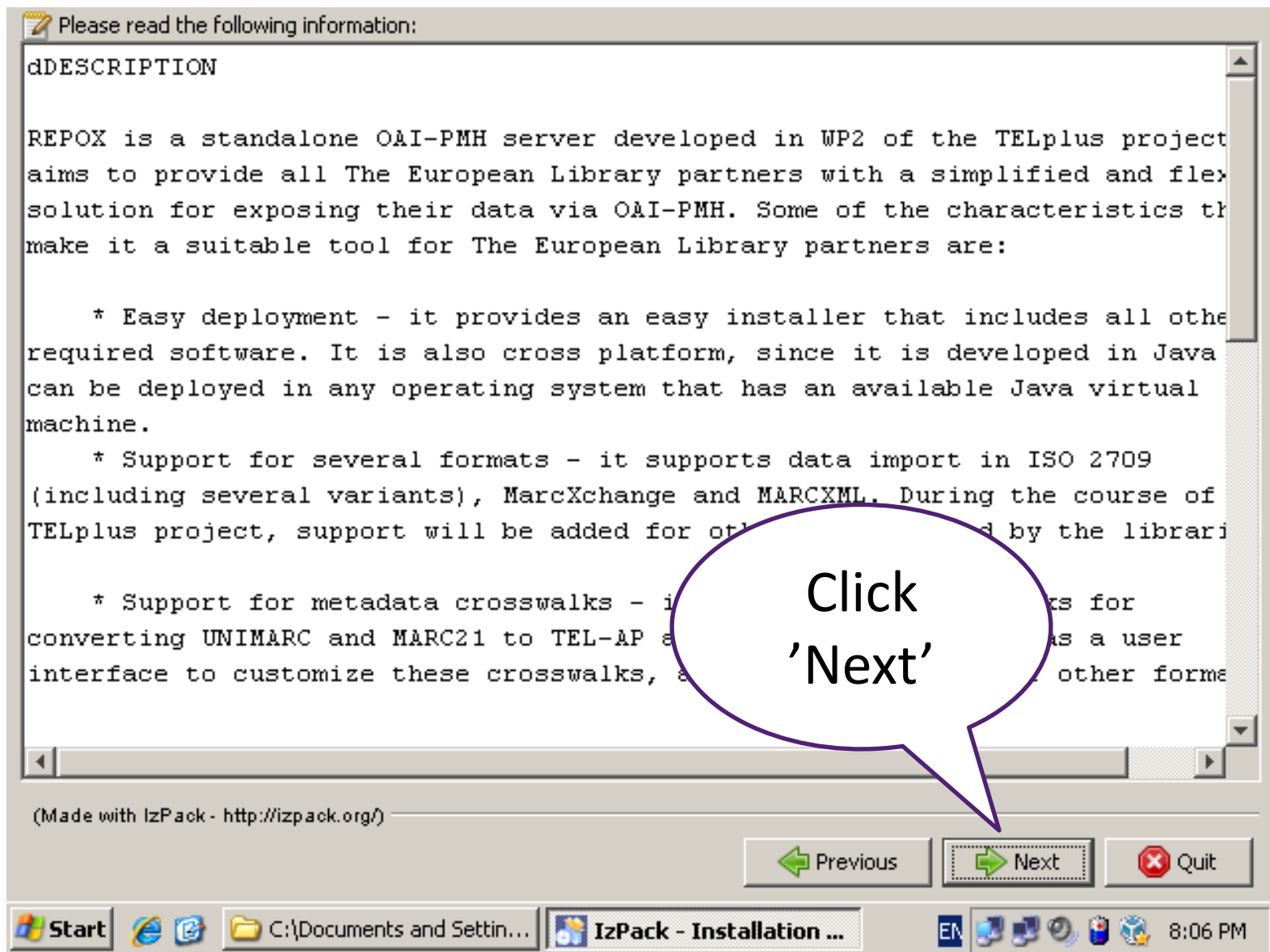
Double-click the JAR-file to start installation of Repox

Step 2



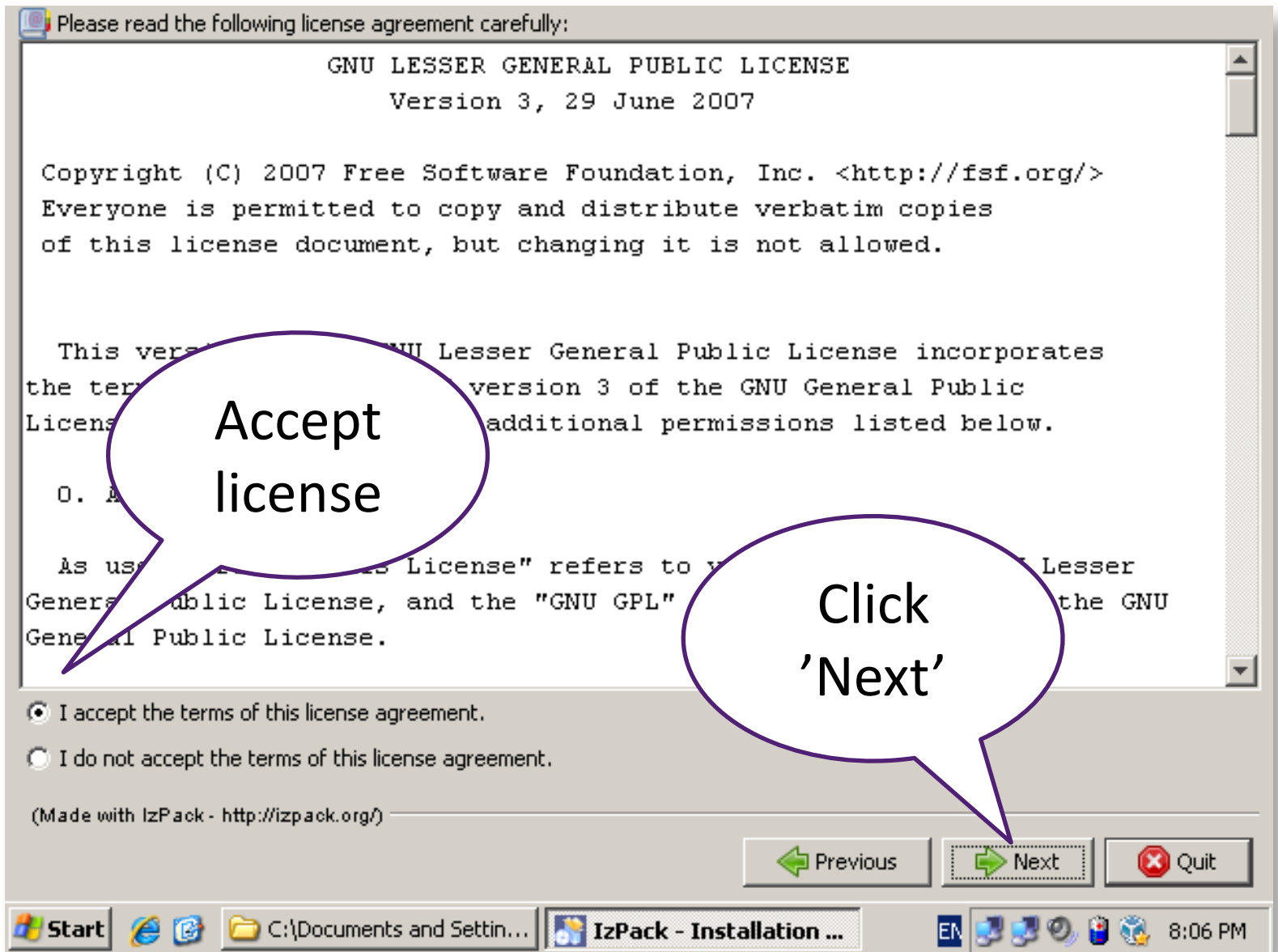
Proceed to the next screen by clicking 'Next'

Step 3



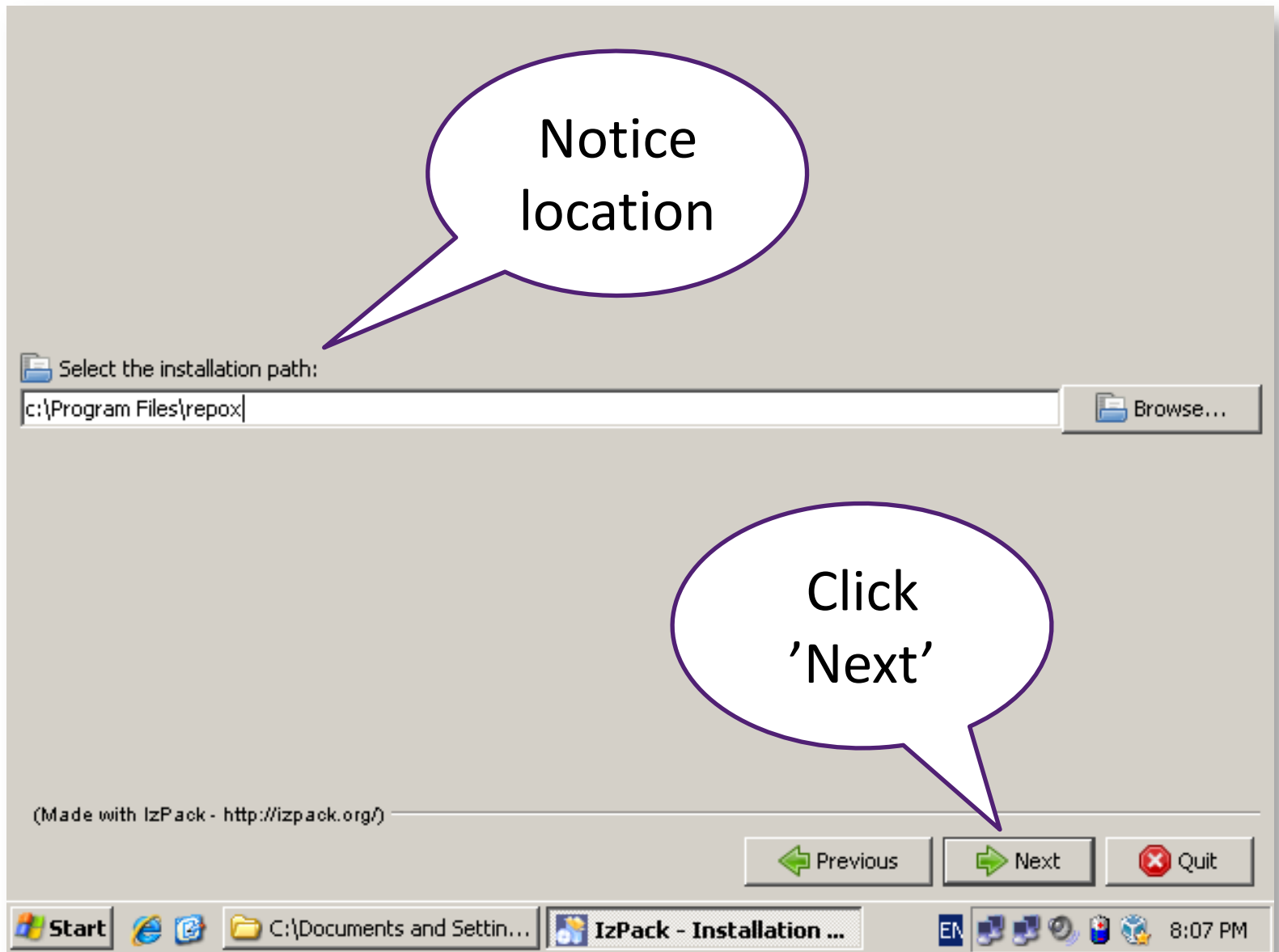
Proceed to the next screen by clicking 'Next'

Step 4



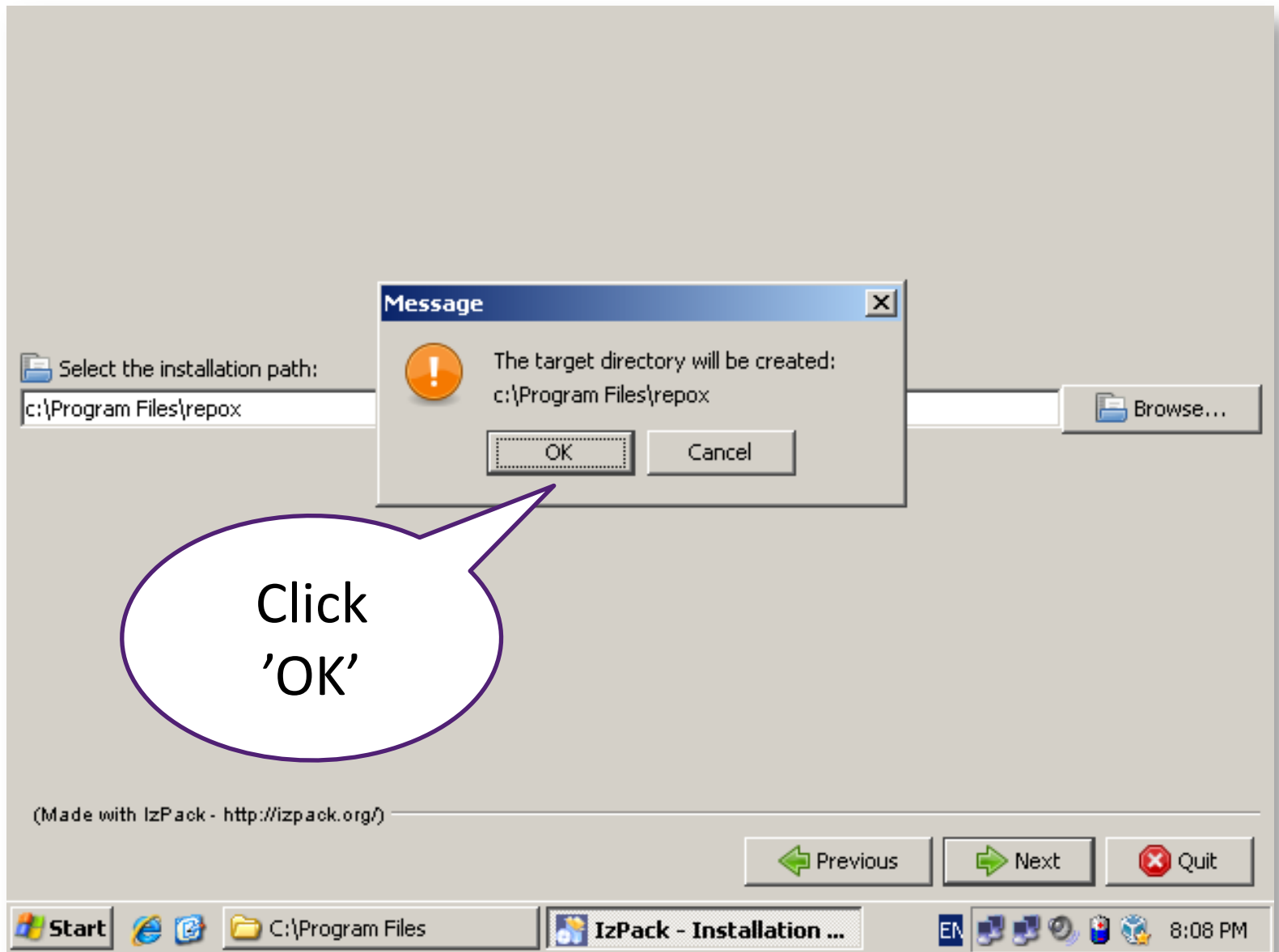
Repos is distributed under the GNU L-GPL License, check the radio-button to accept it and click 'Next'

Step 5



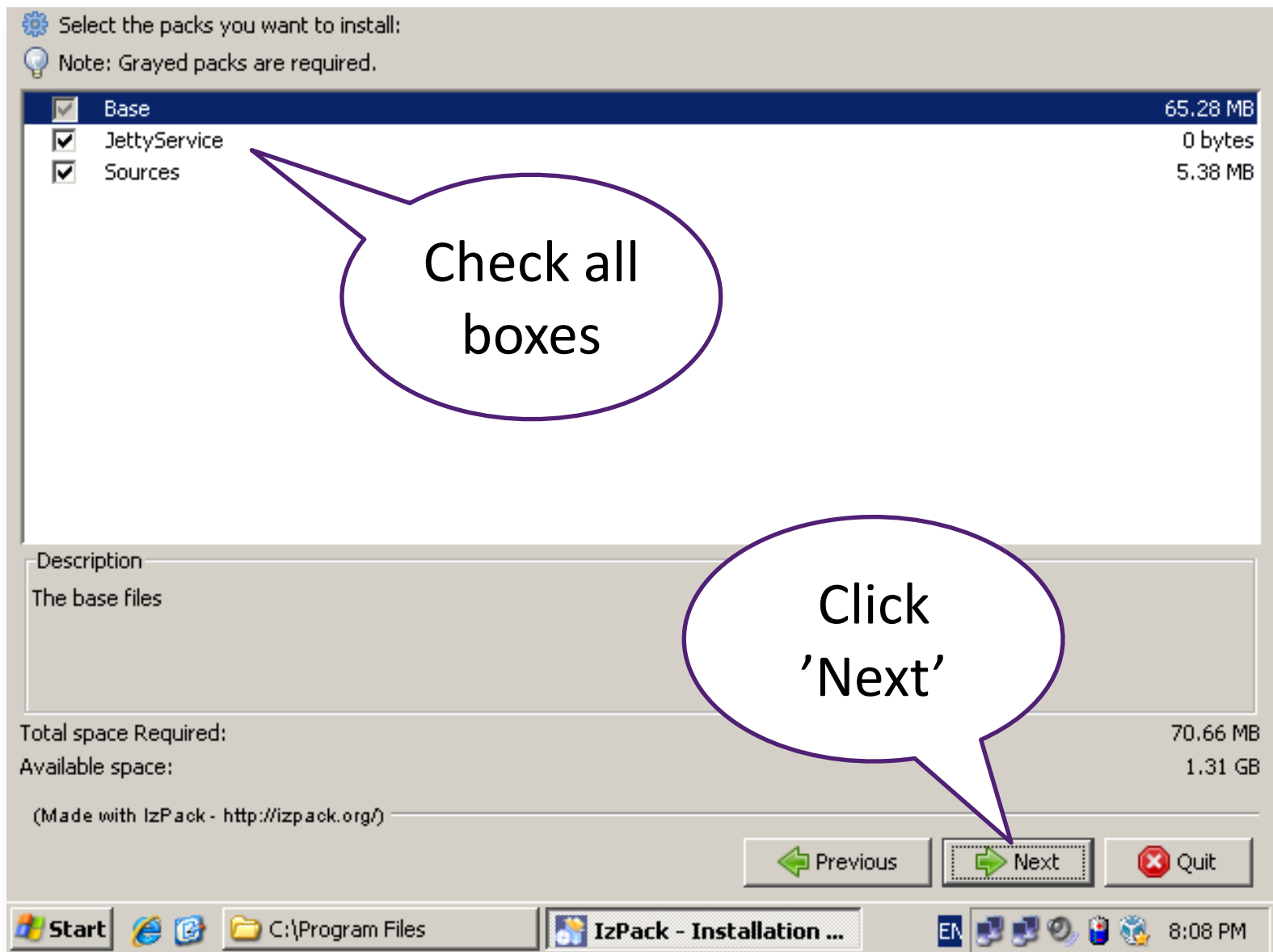
Don't change the default values here – but notice where you install RepoX for future reference

Step 6



The program directory will be created for you. Press 'OK' to accept and proceed

Step 7



Leave the default settings and click next to proceed. Sources are not strictly necessary but may be useful for advanced users.

Step 8

Advanced Configuration

REPOX URN syntax is urn:[NAME]: where [NAME] should be an identifier of this rebox like the dns of
urn:rebox.ist.utl.pt:

REPOX URN	<input type="text" value="urn:rebox.local:"/>
Jetty Server Install Directory	<input type="text" value="c:\Program Files\rebox"/>
Jetty Port	<input type="text" value="8080"/>
Admin email	<input type="text" value="admin@admin.com"/>
SMTP server (empty for no admin emails)	<input type="text"/>
User Login	<input type="text" value="admin"/>
User Password:	<input type="password" value="*****"/>
Retype Password:	<input type="password" value="*****"/>

(Made with IzPack - <http://izpack.org/>)

Previous Next Quit

!!!

Note password

Click 'Next'

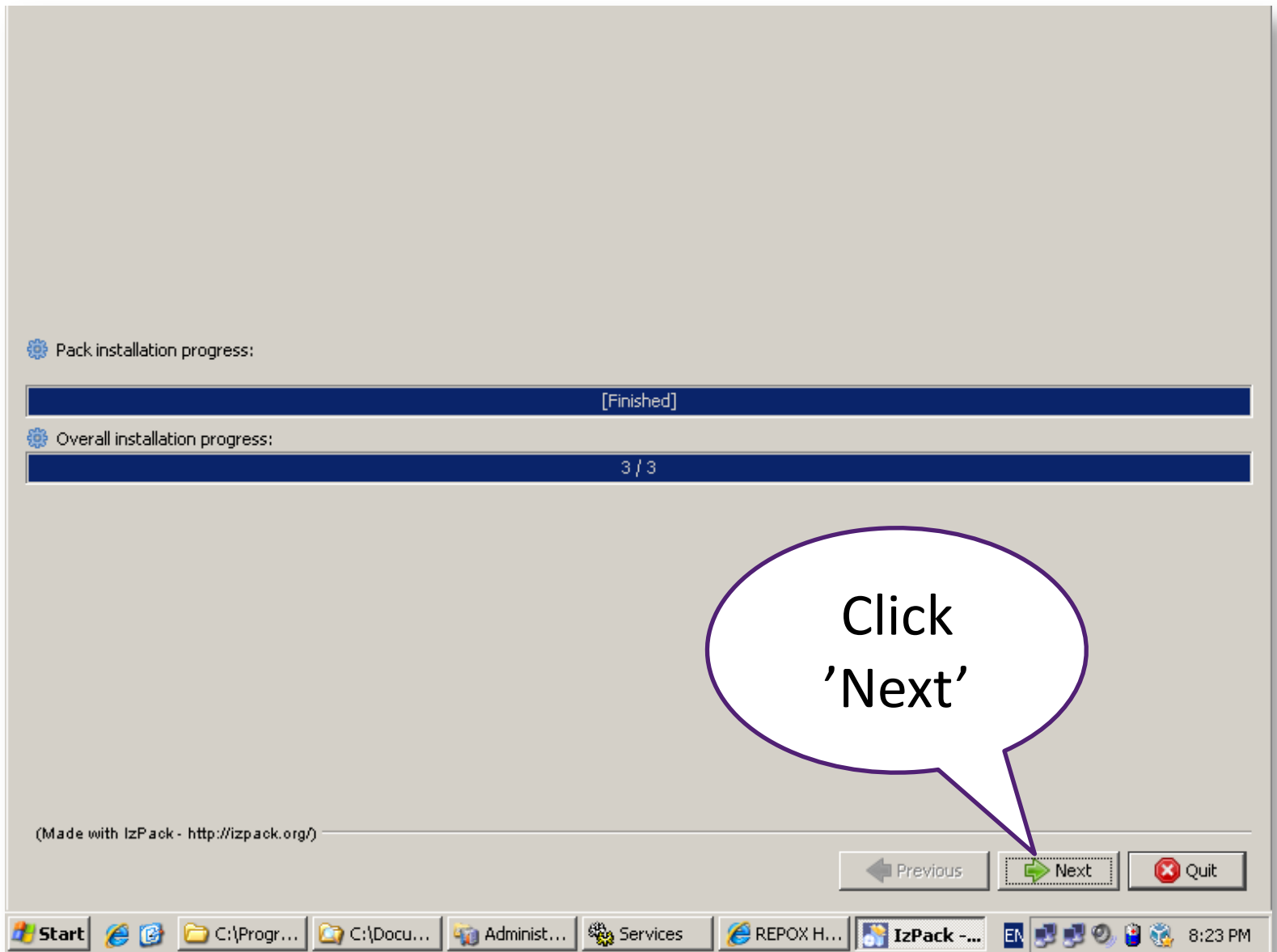
The REPOX URN should be a unique namespace, e.g.
urn:rebox.yourdomain.com: | Port number could be 80 – if free...

Step 9

The screenshot shows an IzPack installation window. At the top center, a white speech bubble with a purple outline contains the word "Patience". Below it, the window displays two progress bars. The first bar is titled "Pack installation progress:" and shows the path "c:\Program Files\rebox\jetty\webapps\rebox\jsp\WireIt\test\planarGame\planarGame.js" with a sub-label "Base" and a dark blue progress bar. The second bar is titled "Overall installation progress:" and shows "1 / 3". At the bottom of the window, there are three buttons: "Previous" (with a green left arrow), "Next" (with a grey right arrow), and "Quit" (with a red X). The Windows taskbar at the bottom shows the Start button, several open folders (C:\Progr..., C:\Docu..., Administr...), Services, REPOX H..., and IzPack -... The system tray shows the time as 8:23 PM.

Please wait while the installation proceeds

Step 10



Once the installation is completed, click 'Next' to proceed.

Processing

3/3

```
C:\Documents and Settings\runarbe\Desktop\Repository>echo "generating Jetty password"  
"generating Jetty password"
```

```
C:\Documents and Settings\runarbe\Desktop\Repository>echo off  
wrapper | Jetty6-Service installed.  
wrapper | Starting the Jetty6-Service service...  
wrapper | Waiting to start...  
wrapper | Jetty6-Service started.
```

Patience

Click
'Next'

(Made with IzPack - <http://izpack.org/>)

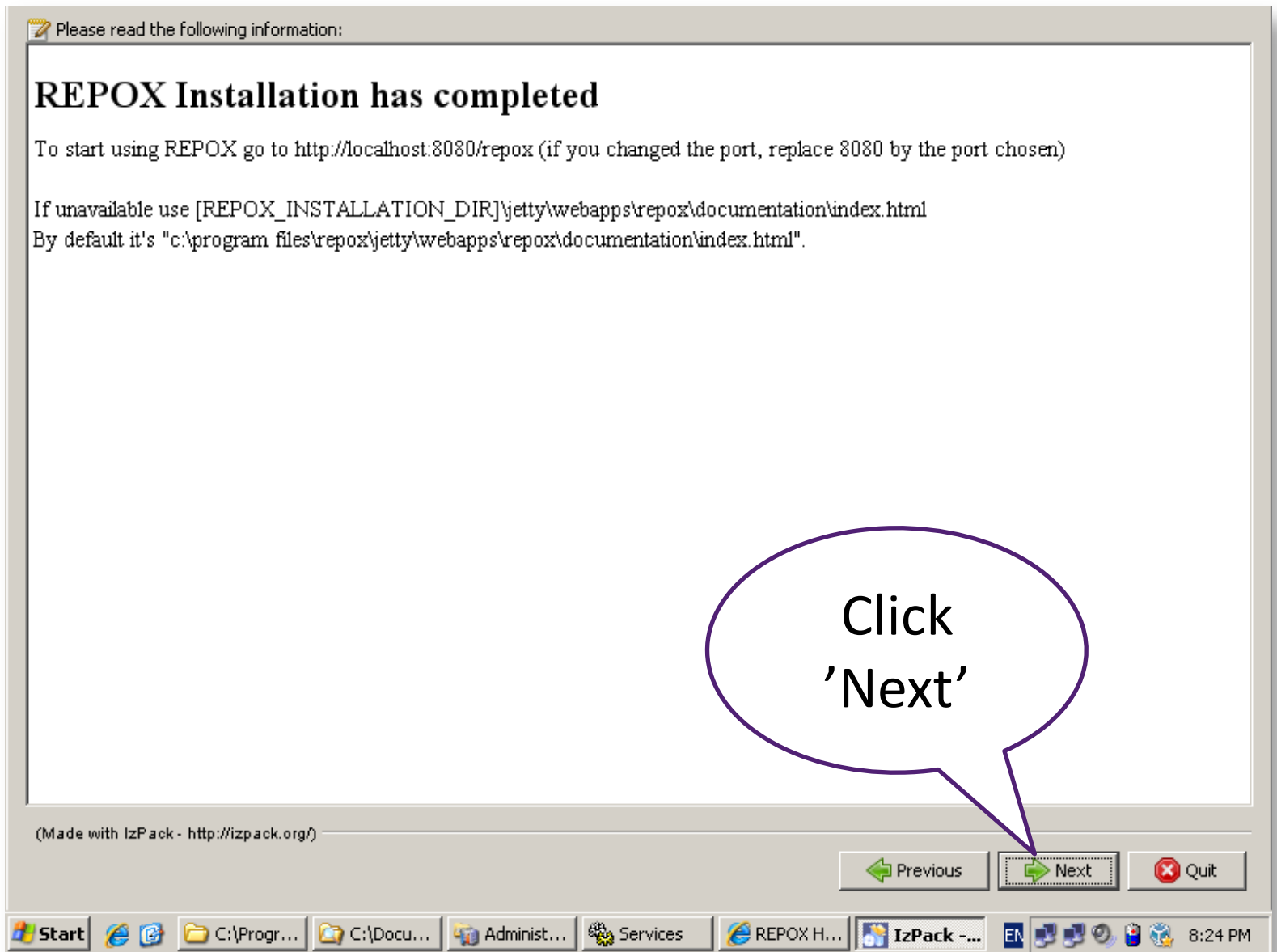
← Previous **Next** → × Quit

Start Internet Explorer C:\Progr... C:\Docu... Administr... Services REPOX H... IzPack -... EN 8:23 PM

This step may take some time for the service to be installed, verified and started. Please wait until completed.

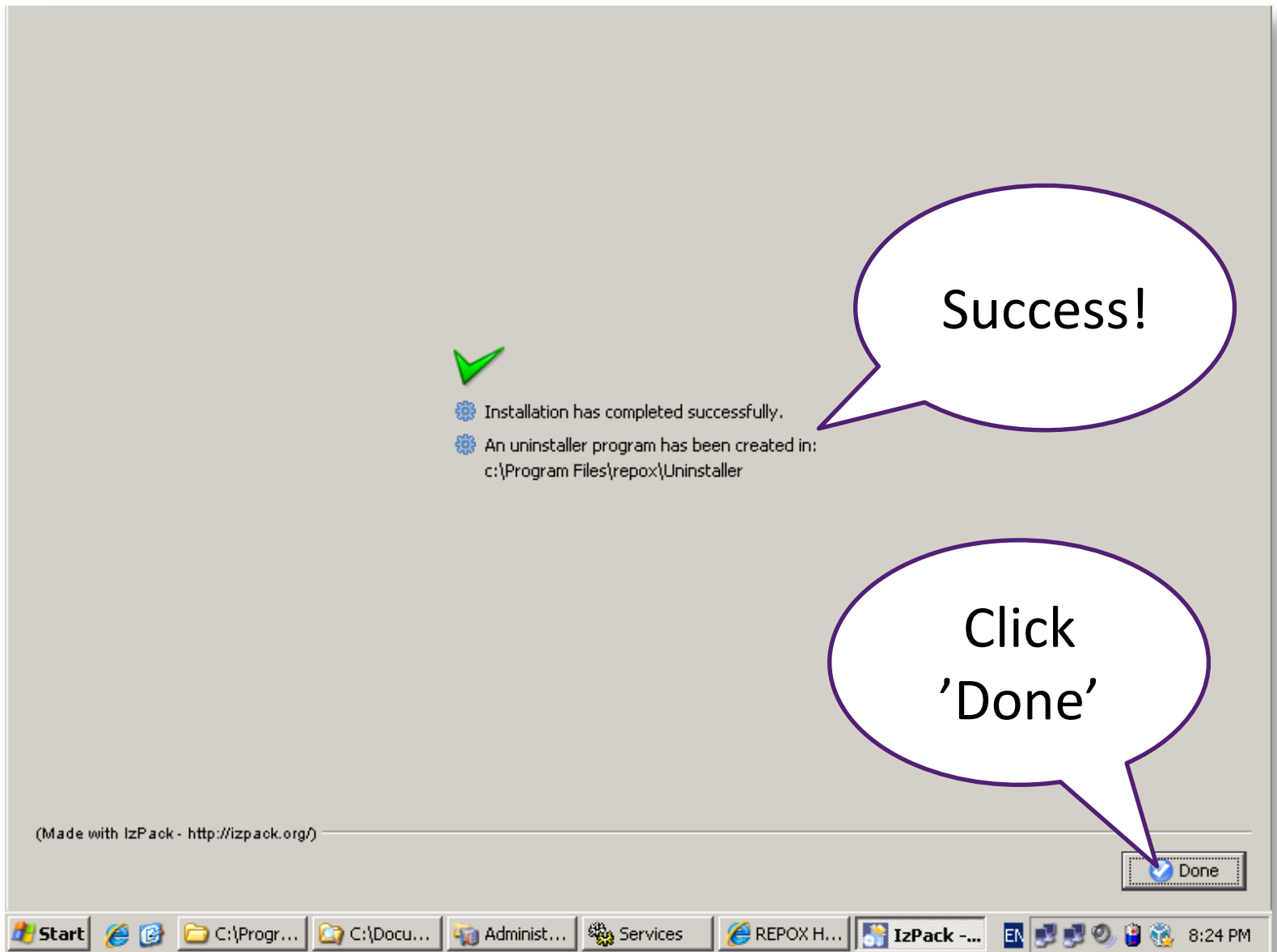
Step 11

Step 12



The final window will give you details on where you have installed your Repox and how you may access it. <http://localhost:8080/repoX>

Step 13



Upon completion of the installer, the program will confirm whether it has been installed successfully.



Testing Repox

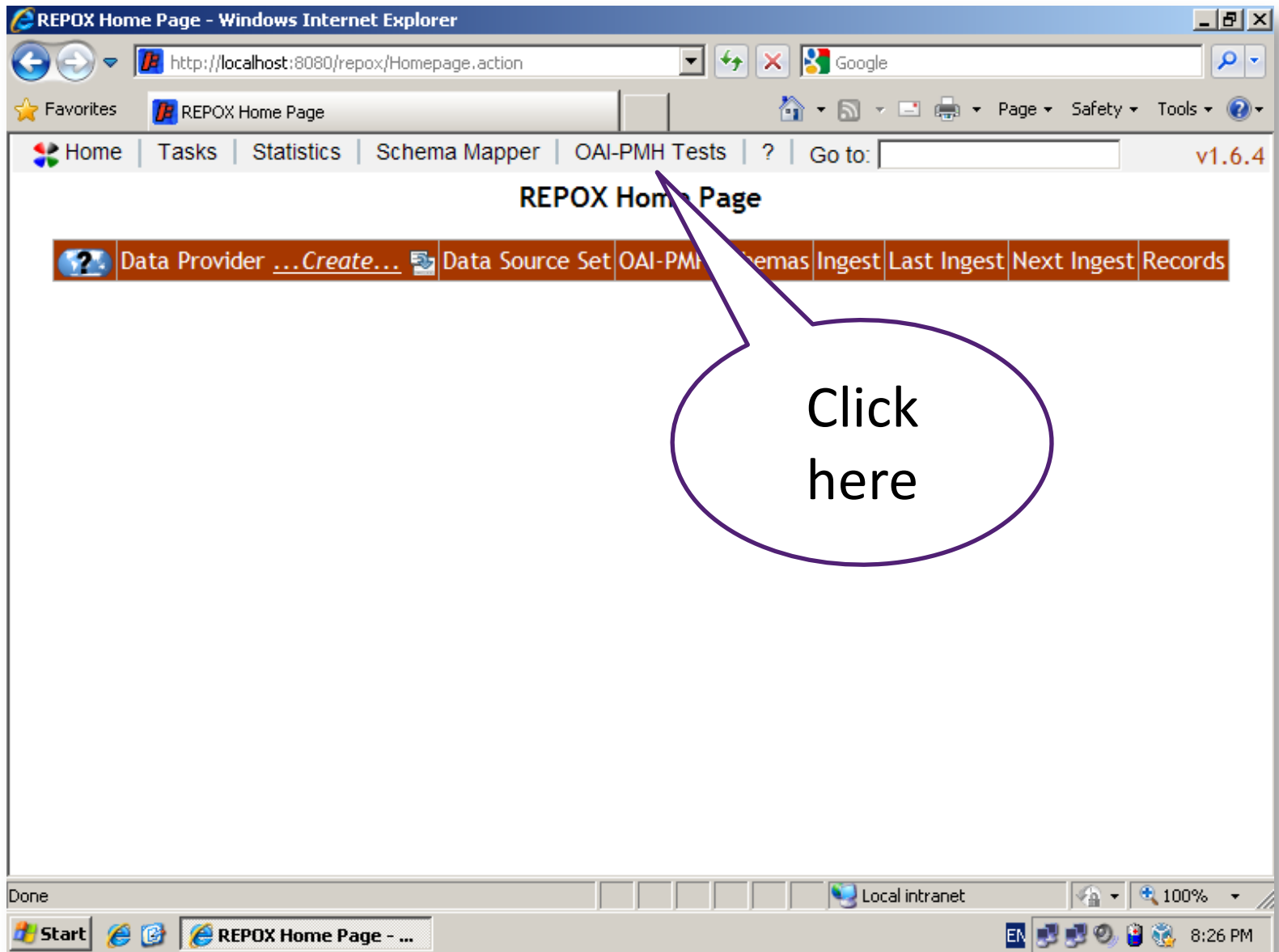
First look at the OAI-PMH web services for an empty repository



CARARE is funded by the **European Commission's ICT Policy Support Programme**



Step 1



Open <http://localhost:8080/repoX> in your web browser to verify that your Repox installation is up and running

Step 2

OAI-PMH Testing - Windows Internet Explorer

http://localhost:8080/repox/jsp/testOAI-PMH.jsp

OAI-PMH Testing

Home | Tasks | Statistics | Schema Mapper | **OAI-PMH Tests** | ? | Go to: | v1.6.4

OAI-PMH Testing

Parameters

Server URL:	http://localhost:8080/repox/OAIHa
metadataPrefix:	
from:	
until:	
set:	
identifier:	
resumptionToken:	

Operations

- ? Identify
- ? ListMetadataFormats
- ? ListSets
- ? ListIdentifiers
- ? ListRecords
- ? GetRecord

Response:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-in:
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/OAI-PMH.x:
  http://www.openarchives.org/OAI/2.0/OAI-PMH.x:
  <responseDate>2011-03-28T20:23:19Z</responseDate
  <request verb="Identify">http://localhost:8080/repo:
- <Identify>
  <repositoryName>REPOX Repository</repositoryName
  <baseURL>http://localhost:8080/repox/OAIHandle
  <protocolVersion>2.0</protocolVersion>
  <adminEmail>mailto:admin@admin.com</adminEmail:
  <earliestDatestamp>1970-01-01</earliestDatestamp>
  <deletedRecord>persistent</deletedRecord>
  <granularity>YYYY-MM-DD</granularity>
  <compression>gzip</compression>
  <compression>deflate</compression>
- <description>
  - <oai-identifier xmlns="http://www.openarchives.org/OAI/2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-in:
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/OAI-PMH.x:
  http://www.openarchives.org/OAI/2.0/OAI-PMH.x:
  </oai-identifier>
</Identify>
</OAI-PMH>
```

Done

Local intranet | 100%

Start | OAI-PMH Testing - Wi... | 11:23 PM

Click on the link 'OAI-PMH Tests' to activate the test interface which allows you to send requests to your OAI-PMH repository web service



Changing static properties



europeana

carare project

CARARE is funded by the **European Commission's ICT Policy Support Programme**



Changes to installation

Directory:

- C:\Program Files\repox\jetty\webapps\repox\WEB-INF\classes

Config files:

- oaicat.properties
- configuration.properties

Edit the files using Word Pad

Restart Jetty Server after editing

- Services > Jetty > Restart





Creating a data provider



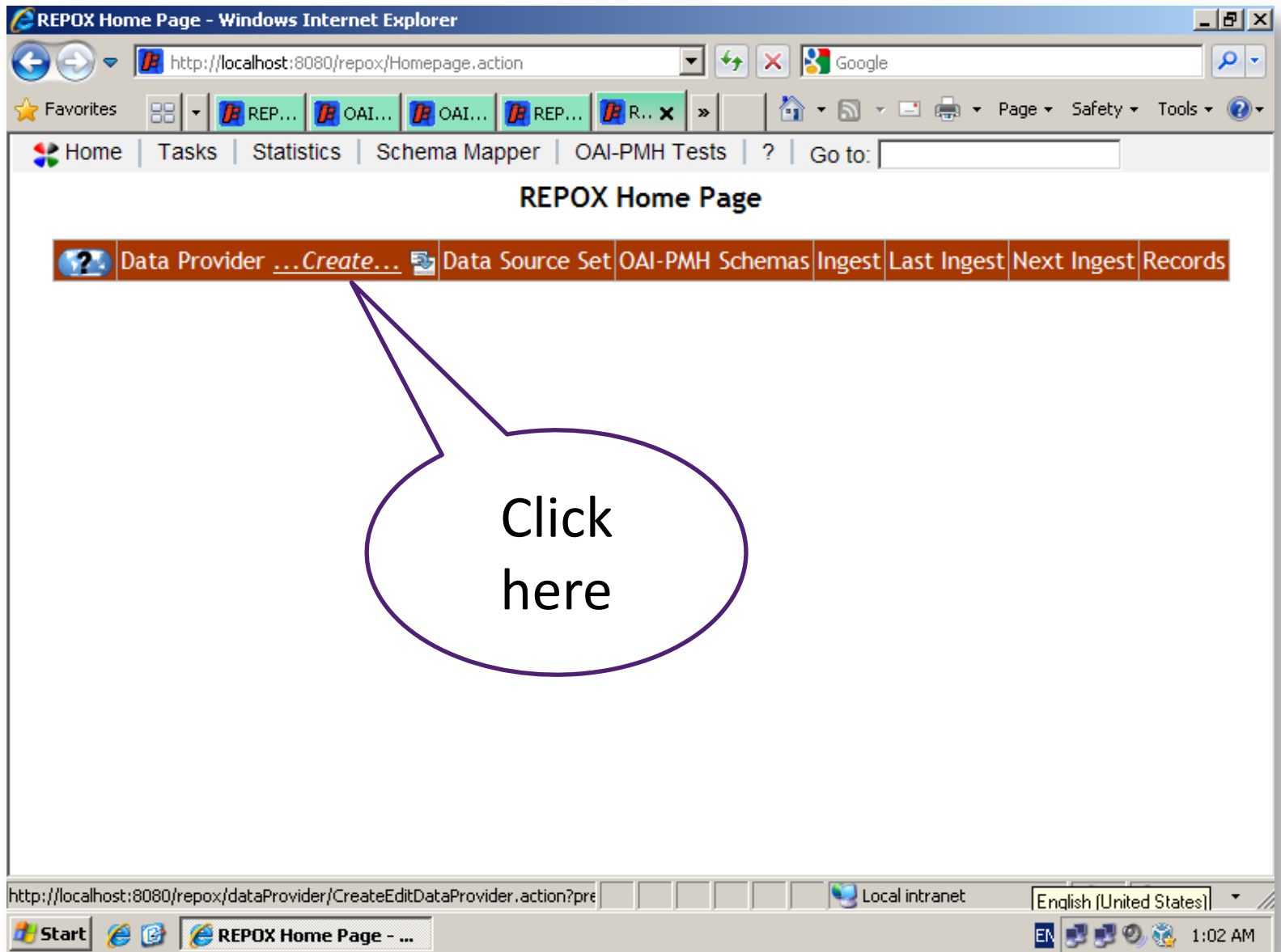
europeana

carare project

CARARE is funded by the **European Commission's ICT Policy Support Programme**



Step 1



From the Repox 'home' screen, click 'Create' next to 'Data Provider'

Step 2

REPOX Authentication - Windows Internet Explorer

http://localhost:8080/repoX/jsp/login.jsp

REPOX Authentication

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to: v1.6.4

REPOX Authentication

User:

Password:

Don't ask for my password for two weeks

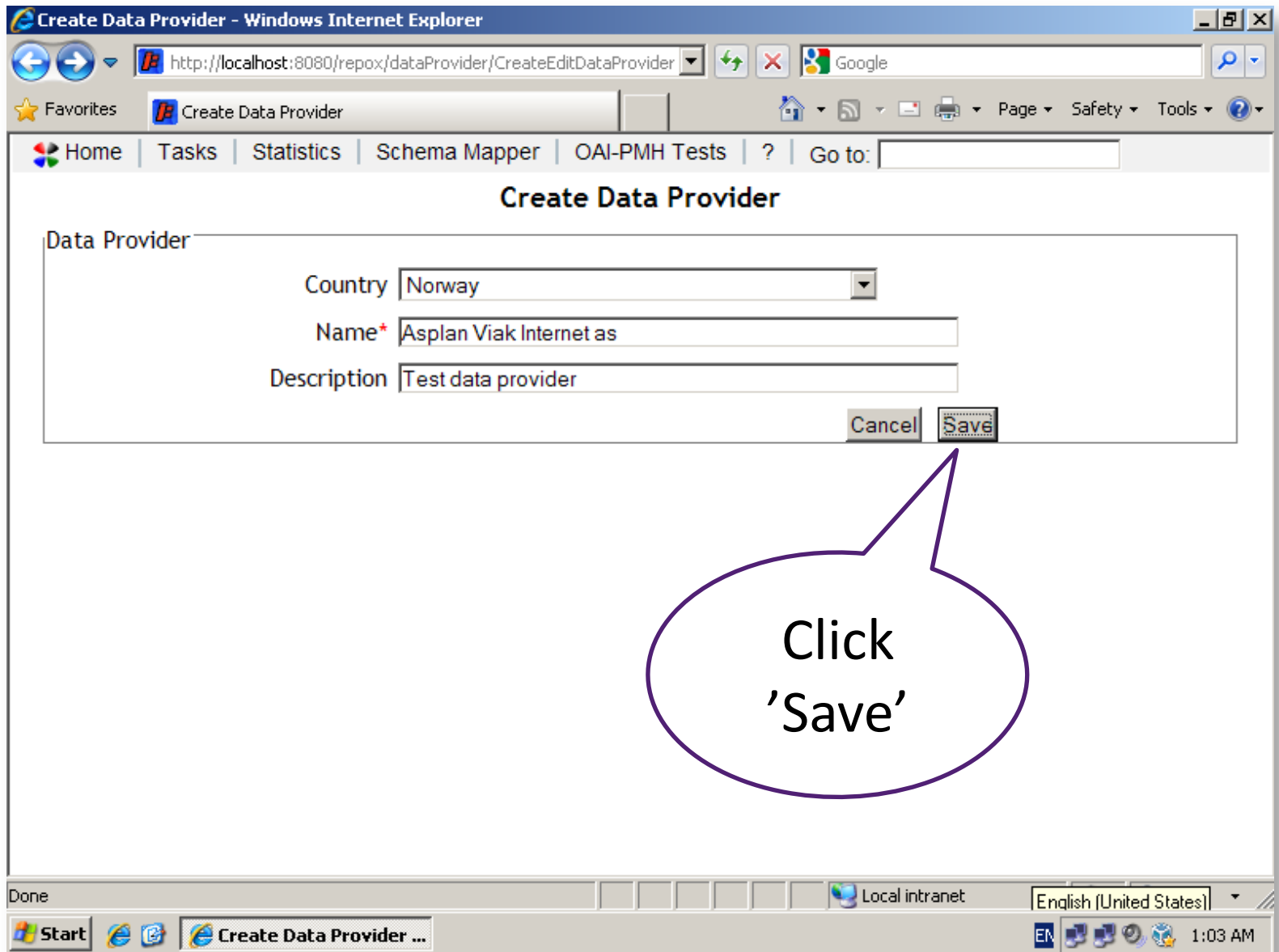
Click 'Submit'

Done Local intranet 100%

Start REPOX Authentication... 11:28 PM

Before being allowed to create a data provider, it is necessary to log in using the username/password you configured during installation

Step 3



The data provider is configured with only three simple metadata items. 'Country', 'Name' and 'Description'. None are critical to the operation.

Step 4

The screenshot shows a web browser window with the following content:

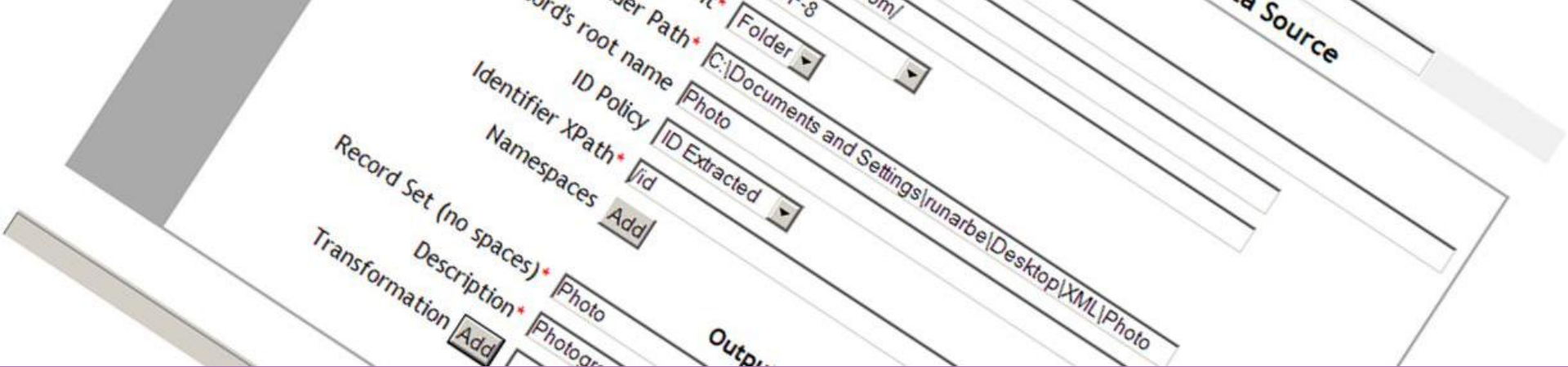
- Address bar: `http://localhost:8080/repoX/dataProvider/ViewDataProvider.action`
- Page Title: **Data Provider Asplan Viak Internet as**
- Message: Data Provider with name Asplan Viak Internet as was created successfully. To create a new Data Source for this Data Provider click [here](#).
- Table:

Asplan Viak Internet as - Test data provider (Refresh)						Edit	Delete
Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records		
...Create Data Source...							

A callout bubble labeled "Tools" points to the "Edit" button in the table header.

Windows taskbar: Start button, Data Provider Asplan ... taskbar icon, Local intranet, English (United States), 1:04 AM.

Your data provider has now been successfully configured and you may proceed to create a data source or edit the values entered in Step 3



Creating a data source



europeana

carare project

CARARE is funded by the **European Commission's ICT Policy Support Programme**



Step 1

REPOX Home Page - Windows Internet Explorer

http://localhost:8080/repoX/Homepage.action

REPOX Home Page

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to: v1.6.4

REPOX Home Page

Data Provider ...Create...	Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
Asplan Viak Internet as						

Click here

Local intranet 100%

Start REPOX Home Page - ... 11:31 PM

A 'Data Source' belongs to a 'Data Provider'. Choose the 'Data Provider' name you created in the previous section.

Step 2

Done

Local intranet 100%

Start Data Provider Asplan ... 11:31 PM

To create the new 'Data Source', click the link as shown above

Step 3

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

Data Provider: Asplan Viak Internet as - Create Data Source

Folder	Data Source
<u>OAI-PMH</u>	Metadata Format* <input type="text" value="other"/> <input type="text" value="test"/>
<u>Z39.50</u>	Schema* <input type="text" value="http://test.com/"/>
	Metadata Namespace* <input type="text" value="http://test.com/"/>
	Character Encoding* <input type="text" value="UTF-8"/>
	Retrieve Variant* <input type="text" value="Folder"/>
	Folder Path* <input type="text" value="C:\Documents and Settings\runarbe\Desktop\XML\Photo"/>
	Record's root name <input type="text" value="Photo"/>
	ID Policy <input type="text" value="ID Extracted"/>
	Identifier XPath* <input type="text" value="/id"/>
	Namespaces <input type="button" value="Add"/>
	Output
	Record Set (no spaces)* <input type="text" value="Photo"/>
	Description* <input type="text" value="Photography Database"/>
	Transformation <input type="button" value="Add"/> <input type="text" value="-"/> <input type="button" value="New"/>
	<input type="button" value="Cancel"/> <input type="button" value="Save"/>

!!!

Local intranet | 100%

Configure your 'Data Source' as shown above. A yellow box indicates that you are free to change the value.

Step 4

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

Data Provider Asplan Viak Internet as

Data Source with name Photo, from Data Provider with name AsplanViakInternetasr0, was created successfully

Asplan Viak Internet as - Test data provider (Refresh) Edit Delete

Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
Photo	test	Folder test			0
Type	Folder				
Folder Path	C:\Documents and Settings\runarbe\Desktop\XML\Photo				
Record's root name	Photo				
Namespace Prefix - URI					
Local Metadata Format	test				
ID Policy	ID Extracted				
Identifier XPath	/id				
Namespace Prefix - URI					
Record Set	Photo				
Description	Photography Database				
Schema	http://test.com/				
Metadata Namespace	http://test.com/				
Transformations					
Number of Records	0				
Ingest	Ingest Now Schedule Ingest				
Export	Full Path: <input type="text"/> Records per file: <input type="text" value="1"/> Export Now Schedule Export				

Done Local intranet 100%

Once successfully created, the key information about the 'Data Source' and harvesting/export options will be shown on the screen.

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

Data Provider Asplan Viak Internet as

Harvest will start in a few seconds.

Asplan Viak Internet as - Test data provider (Refresh) Edit Delete																																					
Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records																																
Photo	test	Folder test			0																																
<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 30%;">Type</td> <td>Folder</td> </tr> <tr> <td>Folder Path</td> <td>C:\Documents and Settings\runarbe\Desktop\XML\Photo</td> </tr> <tr> <td>Record's root name</td> <td>Photo</td> </tr> <tr> <td>Namespace Prefix - URI</td> <td></td> </tr> <tr> <td>Local Metadata Format</td> <td>test</td> </tr> <tr> <td>ID Policy</td> <td>ID Extracted</td> </tr> <tr> <td>Identifier XPath</td> <td>/id</td> </tr> <tr> <td>Namespace Prefix - URI</td> <td></td> </tr> <tr> <td>Record Set</td> <td>Photo</td> </tr> <tr> <td>Description</td> <td>Photograph</td> </tr> <tr> <td>Schema</td> <td>http://test.co</td> </tr> <tr> <td>Metadata Namespace</td> <td>http://test.co</td> </tr> <tr> <td>Transformations</td> <td></td> </tr> <tr> <td>Number of Records</td> <td>0 </td> </tr> <tr> <td>Ingest</td> <td> <input type="button" value="Ingest Now"/> <input type="button" value="Schedule Ingest"/> </td> </tr> <tr> <td>Export</td> <td> Full Path: <input type="text"/> Records per file: <input type="text" value="1"/> <input type="button" value="Export Now"/> <input type="button" value="Schedule Export"/> </td> </tr> </table>						Type	Folder	Folder Path	C:\Documents and Settings\runarbe\Desktop\XML\Photo	Record's root name	Photo	Namespace Prefix - URI		Local Metadata Format	test	ID Policy	ID Extracted	Identifier XPath	/id	Namespace Prefix - URI		Record Set	Photo	Description	Photograph	Schema	http://test.co	Metadata Namespace	http://test.co	Transformations		Number of Records	0	Ingest	<input type="button" value="Ingest Now"/> <input type="button" value="Schedule Ingest"/>	Export	Full Path: <input type="text"/> Records per file: <input type="text" value="1"/> <input type="button" value="Export Now"/> <input type="button" value="Schedule Export"/>
Type	Folder																																				
Folder Path	C:\Documents and Settings\runarbe\Desktop\XML\Photo																																				
Record's root name	Photo																																				
Namespace Prefix - URI																																					
Local Metadata Format	test																																				
ID Policy	ID Extracted																																				
Identifier XPath	/id																																				
Namespace Prefix - URI																																					
Record Set	Photo																																				
Description	Photograph																																				
Schema	http://test.co																																				
Metadata Namespace	http://test.co																																				
Transformations																																					
Number of Records	0																																				
Ingest	<input type="button" value="Ingest Now"/> <input type="button" value="Schedule Ingest"/>																																				
Export	Full Path: <input type="text"/> Records per file: <input type="text" value="1"/> <input type="button" value="Export Now"/> <input type="button" value="Schedule Export"/>																																				

Click here

/repoj/dataProvider/ViewDataProvider.action | Local intranet | 100%

Step 5

Click the button 'Ingest Now' and wait for 10-15 seconds to allow the initial harvesting to complete

Step 6

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

Data Provider Asplan Viak Internet as

Asplan Viak Internet as - Test data provider (*Refresh*)

Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
Photo	test	Folder test	2011-03-29		1027

Type: Folder
Folder Path: C:\Documents and Settings\runarbe\Desktop\XML\Photo
Record's root name: Photo
Namespace Prefix - URI:
Local Metadata Format: test
ID Policy: ID Extracted
Identifier XPath: /id
Namespace Prefix - URI:
Record Set: Photo
Description: Photograph
Schema: http://test.oai.org
Metadata Namespace: http://test.oai.org
Transformations:
Number of Records: 1027

Update

Ingest

Export Full Path: Records per file:

Scheduled Tasks

Log Files [2011-3\1_20110329_010704.log](#)

Done Local intranet 100%

Check progress by pressing update. Once the harvesting is complete, you can view the outcome of the operation in the 'Log Files' section.

Step 7

```
-----  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vid.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
DataSource 1000 records imported  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
vider.dataSource.SimpleFileExtract$RecordIterator Extracting record from file: samplePhoto.xml  
DataSourceDirectoryImporter 27 records imported  
DataSourceDirectoryImporter Finished importing from Data Source with id Photo. Exit status: OK  
-----  
Done Local intranet 100%
```

Press
'back' in
browser

The log file shows the progress of the import. In this example, chunks of 1000 records are imported at a time – giving us 1000 + 27 records.

Step 8

OAI-PMH Testing - Windows Internet Explorer

http://localhost:8080/repox/jsp/testOAI-PMH.jsp

OAI-PMH Testing

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

OAI-PMH Testing

Parameters

Server URL:

metadataPrefix:

from:

until:

set:

identifier:

resumptionToken:

Operations

- Identify
- ListMetadataFormats
- ListSets
- ListIdentifiers
- ListRecords
- GetRecord

Response:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ OAI-PMH.xsd" responseDate="2008-08-08T10:08:00" requestVerb="ListSets">
- <ListSets>
- <set>
  <setSpec>Photography Database</setSpec>
  <setName>Photography Database</setName>
</set>
</ListSets>
</OAI-PMH>
```

Click 'ListSets'

Click 'OAI-PMH'

Local intranet 100% 1:08 AM

You can now test your 'Data Source' by choosing 'OAI-PMH Tests'. Click 'ListSets' to see that your 'Data Source' is available for harvesting.

Step 9

The screenshot shows a web browser window titled "OAI-PMH Testing - Windows Internet Explorer". The address bar shows the URL `http://localhost:8080/repox/jsp/testOAI-PMH.jsp`. The page content is divided into two main sections: "Parameters" and "Response".

Parameters:

- Server URL: `http://localhost:8080/repox/OAIHa`
- metadataPrefix:
- from:
- until:
- set:
- identifier:
- resumptionToken:

Operations:

- [Identify](#)
- [ListMetadataFormats](#)
- [ListSets](#)
- [ListIdentifiers](#)
- [ListRecords](#)
- [GetRecord](#)

Response:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd"
  responseDate>2011-03-28T22:09:04Z</responseDate>
  test
  <verb>ListMetadataFormats</verb>
  <ListMetadataFormats>
    <metadataFormat>
      <metadataPrefix>test</metadataPrefix>
      <schema />
      <metadataNamespace />
    </metadataFormat>
  </ListMetadataFormats>
</OAI-PMH>
```

A callout bubble with a purple border and white background contains the text: "Click 'ListMetadataFormats'". A purple arrow points from this bubble to the "ListMetadataFormats" link in the Operations list.

The browser's status bar at the bottom shows the JavaScript code `javascript:ListMetadataFormats();` and the system tray includes the Start button, taskbar icons, and the system clock showing 1:09 AM on 3/28/2011.

OAI-PMH can transport many metadata formats, choose 'ListMetadataFormats' to verify the name of your format.

Step 10

OAI-PMH Testing - Windows Internet Explorer

http://localhost:8080/repoX/jsp/testOAI-PMH.jsp

Google

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

OAI-PMH Testing

Parameters

Server URL:	http://localhost:8080/repoX/OAIHa
metadataPrefix:	test
from:	
until:	
set:	Photo
identifier:	
resumptionToken:	

Operations

- Identify
- ListMetadataForIdentifiers
- ListSets
- ListIdentifiers
- ListRecords**
- GetRecord

Response:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd
  http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd"
  <responseDate>2011-03-28T22:09:50Z</responseDate>
  <request verb="ListRecords" set="Photo"
    metadataPrefix="test">http://localhost:8080/repoX
- <ListRecords>
- <record>
  - <header>
    <identifier>urn:nbn:de:hbz:5:1-30001-p0111-8
    <datestamp>2011-03-28T22:09:50Z
    <setSpec>Photo
  </header>
  - <metadata>
    - <Photo>
      <type>Foto
      <motive>brøderi, bakerverer, redskap, menn, p
      ovn,</motive>
      <photograph>Fosse, Ole Pedersen</photograph>
```

Click 'ListRecords'

View results here

Local intranet | 100% | 1:09 AM

Add the values you identified in Steps 8, 9 to the web form and choose 'ListRecords'. After some seconds, your XML will appear on the right.



Harvesting another repository



CARARE is funded by the **European Commission's ICT Policy Support Programme** 

Step 1

REPOX Home Page

?	Data Provider ...Create...	Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
🇩🇰 ?	Asplan Viak Internet as	Photo	test	Folder test	2011-03-29		1027

Choose 'Data Provider' name

http://localhost:8080/repoX/dataProvider/ViewDataProvider.action?dataProv

Local intranet 100%

Start REPOX Home Page - ... 9:38 AM

Select the data provider for which you will create a new data source

Step 2

Asplan Viak Internet as - Test data provider (*Refresh*) Edit Delete

Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
Photo	test	Folder test	2011-03-29		1027

[...Create Data Source...](#)

Choose 'Create Data Source'

Choose create 'Data Source' (well known by now)

Step 3a: Murberget

Choose 'OAI-PMH' in the left sidebar

Click after entering OAI-URL

When done, click 'Save'

Data Provider: - Create Data Source

Folder

OAI-PMH

Z39.50

Data Source

OAI URL *

OAI Set

Metadata Format *

Schema *

Metadata Namespace *

Output

Record Set (no spaces) *

Description *

Transformation

This time, instead of the default 'Folder' choose 'OAI-PMH' and enter the values as shown in the form above

Step 3b: Sogn og Fjordane

Choose 'OAI-PMH' in the left sidebar

Click after entering OAI-URL

When done, click 'Save'

Data Provider: - Create Data Source

Folder	Data Source
OAI-PMH	OAI URL * <input type="text" value="http://93.94.14.7/phpoai2/oai2.php"/> <input type="button" value="check"/> <input type="button" value="Add All"/>
Z39.50	OAI Set <input type="text" value="KL"/>
	Metadata Format * <input type="text" value="other"/> <input type="text" value="abm"/>
	Schema * <input type="text" value="http://to_be_decided/"/>
	Metadata Namespace * <input type="text" value="http://to_be_decided/"/>
	Output
	Record Set (no spaces) * <input type="text" value="KL"/>
	Description * <input type="text" value="Cultural Heritage Atlas for Sogn og Fjordane"/>
	Transformation <input type="button" value="Add"/> <input type="text" value="-"/> <input type="button" value="New"/>
	<input type="button" value="Cancel"/> <input type="button" value="Save"/>

Repeat Steps 1,2

Enter the values as shown in the form above.

Step 3c: your neighbour

Choose 'OAI-PMH' in the left sidebar

After entering correct URL, click 'Check'

When done, click 'Save'

Repeat Steps 1,2

All values shown in angular brackets <value> can be entered by you

Step 4

Harvest will start in a few seconds.

Asplan Viak Internet as - Test data provider (Refresh) Edit Delete

Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
KL	abm	OAI-PMH abm			0
Ljustorp_photo	ese	OAI-PMH ese			0

Type: OAI-PMH
Local Metadata: ese
Format:
OAI URL: http://repx.murberget.se:8080/repx/OAIHandler
OAI Set: Ljustorp_photo
ID Policy: ID Generated
Record Set: Ljustorp_photo
Description: Photos from Ljustorp
Schema: http://www.europeana.eu/schemas/...
Metadata Namespace: http://www.europeana.eu/sch...
Transformations:
Number of Records: 0
Ingest: **Ingest Now** Schedule Ingest
Export: Full Path: Records per file: 1 Export Now
Schedule Export

Choose 'Ingest Now' from 'Data Source'

Once you have configured all your data sources, go to the overview screen and press 'Ingest Now'

Step 5

The screenshot shows a web browser window displaying the 'Data Provider Asplan Viak Internet as' application. The main content area features a table with the following data:

Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
KL	abm	OAI-PMH abm			0
Ljustorp_photo	ese	OAI-PMH ese	2011-03-29		1987

Below the table, the details for the 'Ljustorp_photo' set are shown:

- Type: OAI-PMH
- Local Metadata: ese
- Format:
- OAI URL: http://repoj.murberget.se:8080/repoj/OAIHandler
- OAI Set: Ljustorp_photo
- ID Policy: ID Generated
- Record Set: Ljustorp_photo
- Description: Photos from Ljustorp
- Schema: http://www.europeana.eu/schemas/ese/...
- Metadata Namespace: http://www.europeana.eu/sch...
- Transformations:
- Number of Records: 1987 (with a refresh icon)
- Ingest: Ingest Now, Schedule Ingest
- Export: Full Path: [input field], Records per file: 1, Export Now

A callout bubble points to the refresh icon next to the 'Number of Records' field, containing the text: "Click 'Update' after some time".

Have patience, and press the update icon after some time to check the progress.

Step 6

OAI-PMH Testing - Windows Internet Explorer

http://localhost:8080/repox/jsp/testOAI-PMH.jsp

Google

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

OAI-PMH Testing

Parameters

Server URL:	http://localhost:8080/repox/OAIHa
metadataPrefix:	ese
from:	
until:	
set:	Ljustorp_photo
identifier:	
resumptionToken:	

Operations

- Identify
- ListMetadataFormats
- ListSets
- ListIdentifiers
- ListRecords
- GetRecord

Response:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd
  http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd"
  <responseDate>2011-03-29T07:06:17Z</responseDate>
  <request verb="ListRecords" set="Ljustorp_photo"
  metadataPrefix="ese">http://localhost:8080/repox,
  </request>
  <response verb="ListRecords" set="Ljustorp_photo"
  metadataPrefix="ese">
    <listRecordsInfo>
      <totalRecords>1</totalRecords>
      <recordsPerPage>1</recordsPerPage>
      <startRecordNumber>1</startRecordNumber>
      <nextRecordNumber>1</nextRecordNumber>
      <datestamp>2011-03-29</datestamp>
      <setSpec>Ljustorp_photo</setSpec>
      <metadata>
        <europeana:record
          xmlns:europeana="http://www.europeana.eu/ontology/record/europeana"
          xmlns="http://www.openarchives.org/OAI/2.0"
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
```

Test your repository

Done

Local intranet

100%

Start | OAI-PMH Testing - Wi... | 10:06 AM

Once the harvesting is done, you may test your repository using the 'OAI-PMH Tests' tab and the detailed specified in Steps 3a, b, c