



Ingeniería Técnica en Informática de Gestión
TFC - Plataforma Gnu Linux

Proyecto: Scientific Ubuntu LiveDVD
Documento: Memoria

13 de Junio de 2011

Alumno: Benigno Cuevas De La Fuente bcuevas@uoc.edu
Profesor: Helena Rifá Pous hrifa@uoc.edu
Consultor: Jordi Massaguer Pla jmassaguerp@uoc.edu

Licencia.

© Benigno Cuevas De La Fuente

Esta obra está bajo una licencia Reconocimiento-No comercial-Sin obras derivadas 2.5 España de Creative Commons. Puede copiarlo, distribuirlo y transmitirlo públicamente siempre que cite al autor y la obra, no se haga un uso comercial y no se hagan copias derivadas.

La licencia completa se puede consultar en

<http://creativecommons.org/licenses/by-nc-nd/2.5/es/deed.es>

Índice de contenidos

1	Introducción.....	4
2	Referencias.....	5
2.1	Características de Scientific Linux 6 LiveDVD.....	6
	Aplicaciones de carácter científico en Scientific Linux 6.....	8
2.2	Características de Poseidon Linux.....	10
	Características de las aplicaciones de carácter científico.....	12
2.3	UbuntuScience.....	22
2.4	Ubuntu[sci] team (scubuntu).....	23
3	Objetivos.....	25
4	Metodología y entorno de desarrollo.....	26
4.1	Script de modificación de la imagen iso y recompilado.....	28
5	Desarrollo de Scientific Ubuntu.....	32
5.1	Aplicaciones.....	32
6	Conclusiones.....	46
7	Recursos.....	47
7.1	Genéricas.....	47
7.2	Referencias de otras distribuciones.....	47
7.3	Customization.....	47
7.4	Herramientas.....	47
8	Índice de imágenes.....	48
9	Anexos.....	49
9.1	Planificación.....	49
9.2	Requerimientos de hardware y software.....	49
9.3	Instrucciones de instalación.....	50
9.4	Paquetes incluidos.....	51
9.5	Glosario de términos.....	52
9.6	Agradecimientos.....	53

1 Introducción.

En el ámbito científico se deben manejar una serie de herramientas de software para el desarrollo de investigaciones y análisis de datos. Aunque hay disponibles herramientas de software cerrado, el elevado coste de las licencias de uso y la no generalidad, limitan su uso a situaciones muy concretas.

En muchos casos, las necesidades básicas del investigador se limitan a una herramienta para el análisis estadístico de datos, una hoja de cálculo, un entorno de desarrollo para un lenguaje de programación y, en ocasiones, un software de cálculo numérico.

Generalmente las distribuciones Linux no incluyen herramientas apropiadas instaladas por defecto, aunque si las hay disponibles tanto en distribuciones basadas en RedHat como en Debian/Ubuntu. En algunas áreas, como la astrofísica y los sistemas de información geográfica (GIS), incluso se podría desarrollar un LiveCD específico con las herramientas disponibles de código libre.

El usuario objetivo para el que queremos construir un LiveDVD, es el investigador/estudiante que necesita de unas herramientas concretas para la resolución de problemas de carácter matemático y estadístico, que, unido a generadores de gráficos, son comunes a la mayoría de áreas científicas y de investigación. Otras áreas de interés son construir aplicaciones en un entorno de desarrollo avanzado (IDE) e incluso un gestor de referencias bibliográficas. Añadiremos también un editor de texto basado en LaTeX muy útil para la generación de documentos de manera profesional y una herramienta de planificación de proyectos.

Para facilitar la tarea de añadir software de cualquiera de estas áreas, se incluirán por defecto los repositorios necesarios.

2 Referencias

Las distribuciones de referencia son [Poseidon Linux](#) basada en Ubuntu, y [Scientific Linux](#) basada en RedHat Enterprise Linux. Mencionaremos también los proyectos [UbuntuScience](#) y [Ubuntu\[sci\]](#) (conocido también como scubuntu) .

Tabla resumen de características:

	Poseidon Linux	Scientific Linux 6.0
Distribución base	Ubuntu 9.10	RHEL 6
Kernel	2.6.31	2.6.32
Sistema de paquetes	deb (apt)	RPM (yum)
Gestor de paquetes	Synaptic	Packagekit
Entorno de escritorio	GNOME 2.28.1	GNOME 2.28.6
Repositorios disponibles (LiveDVD)	<p>http://ftp5.gwdg.de/pub/misc/cran/bin/linux/ubuntu</p> <p>http://ppa.launchpad.net/ubuntuugis/unstable/ubuntu</p> <p>(Main, Universe, Multiverse, Restricted, Medibuntu)</p>	<p>Http://ftp.scientificlinux.org/linux/scientific/\$releasever/\$basearch/os/</p> <p>Disponibles y no activos los repositorios Testing y Source.</p> <p>Disponibles desde packagemanager (no instalados):</p> <p>http://rpmrepo.org/RPMforge-RPMforge</p> <p>http://elrepo.org/tiki/tiki-index.php - ElRepo</p> <p>http://fedoraproject.org/wiki/EPEL - ExtraPackages (EPEL).</p>
Aplicaciones de carácter científico incluidas	GIS, Cálculo numérico, Estadística, Modelado 2D/3D, Gráficos, BioInformática, Química, Astronomía, Publicación(LaTeX).	ATLAS, GNUPlot, Numpy, Units,Eclipse+OpenJDK
Soporte web e idiomas disponibles	http://www.es.poseidonlinux.org/ Brazilian, Spanish, English.	http://www.scientificlinux.org English.
Plataformas disponibles	Versión 3.2 solo 32 bits.	32 y 64 bits, ficheros fuente también disponibles.
Tamaño imagen iso	Poseidon_Linux_3.2_32bits.iso 2.4Gig	SL-60-x86_64-2011-03-07-LiveDVD.iso 2.2 Gig

2.1 Características de Scientific Linux 6 LiveDVD.

Analizaremos las aplicaciones disponibles en una sesión LiveCD, aunque para realizar una prueba más exhaustiva se ha realizado una instalación de esta distribución en una máquina virtual e indicaremos de forma expresa cuando se corresponda a ella. Hacemos esta distinción ya que las aplicaciones incluídas en el LiveDVD están en muchos casos duplicadas, al contener los entornos de escritorio GNOME y KDE (aunque usa por defecto GNOME).

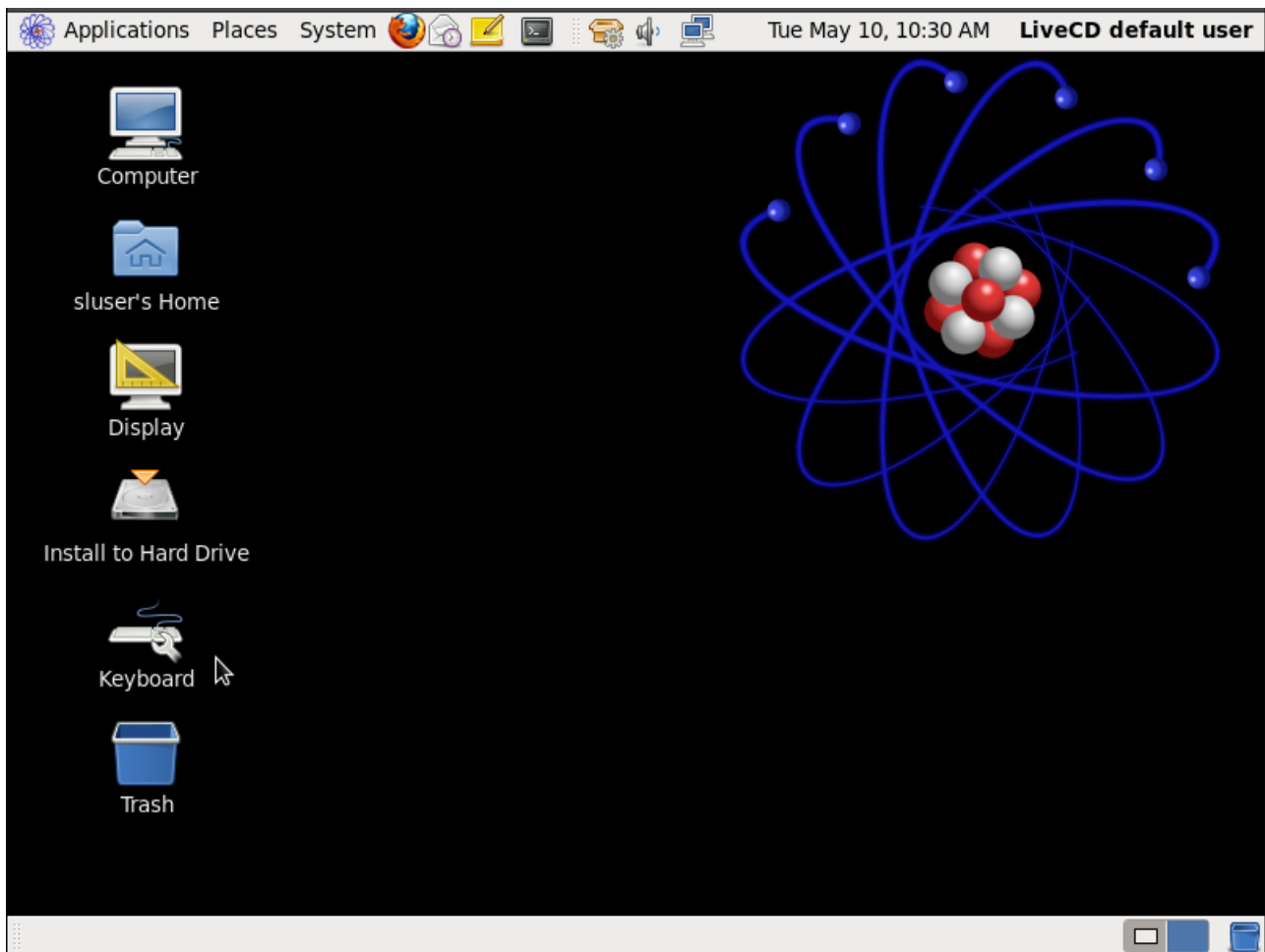


Imagen 1: Scientific Linux Desktop.

Scientific Linux 6 está basada en Red Hat Enterprise Linux (RHEL), y a pesar de su nombre apenas incluye aplicaciones de carácter científico en el LiveDVD. Está desarrollada por el [Fermilab](#) y el [CERN](#) y su objetivo primordial es evitar la gran variedad de sistemas instalados en los diferentes laboratorios. Permite crear diferentes 'spins' o variantes apropiadas a los usos de los laboratorios. Dispone de soporte para sistemas de ficheros Infiniband (conectividad GRID) y OpenAFS aunque no vienen instalados por defecto. Activando los repositorios adicionales (ver imagen 2) disponemos de las aplicaciones consideradas básicas para este proyecto: octave (pero no su front-end qtoctave), R y Lyx.

Repositorios adicionales:

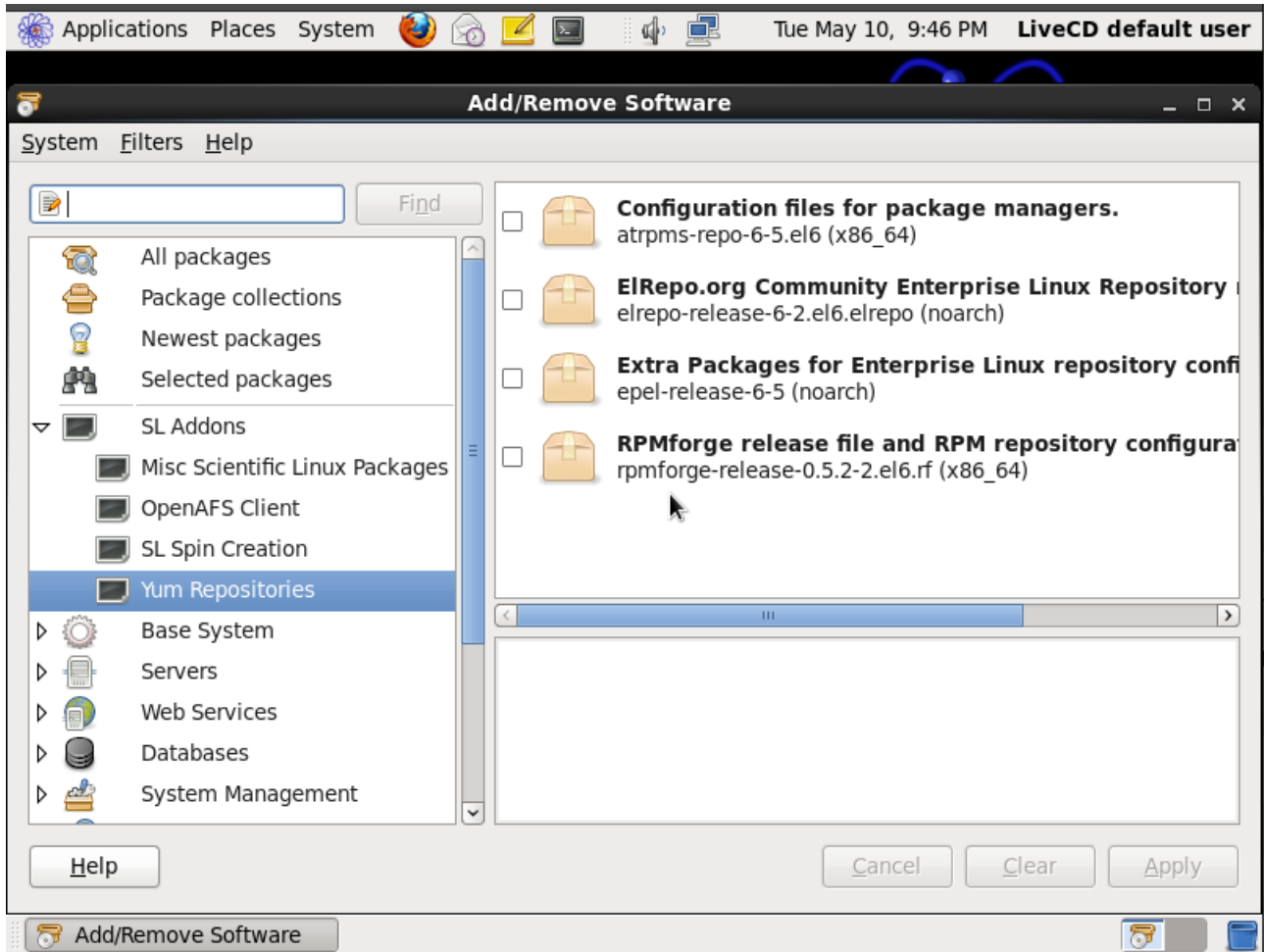


Imagen 2: Scientific Linux, repositorios adicionales.

Resumen de aplicaciones presentes en Scientific Linux.

Grupo	Aplicación
Programación	<ul style="list-style-type: none"> • OpenJDK • Eclipse
Edición de textos	<ul style="list-style-type: none"> • OpenOffice • LaTeX (texlive, no incluye editor Lyx)
Bases de Datos	<ul style="list-style-type: none"> • MySQL (cliente y servidor) • Postgres (únicamente cliente)
Ciencia	<ul style="list-style-type: none"> • ATLAS (Automatically Tuned Linear Algebra Software). • Gnuplot (generador de gráficos en consola) • numpy (librería Python de manejo de arrays multidimensionales) • units (convertor de unidades)

Aplicaciones de carácter científico en Scientific Linux 6.

[ATLAS](#)

ATLAS (Automatically Tuned Linear Algebra Software) es un conjunto de interfaces en C y Fortran77 para (todas) las librerías [BLAS](#) y unas pocas rutinas de [LAPACK](#).

ATLAS se usa en:

- [MAPLE](#)
- [MATLAB](#)
- [Octave](#)

BLAS y Lapack son dos librerías para realizar operaciones básicas de algebra lineal. Son librerías de cálculo científico de libre acceso que cuentan con rutinas especializadas y que son llamadas de acuerdo a su disponibilidad y a la naturaleza de cada operación (secuencial o susceptible de paralelización).

BLAS. Tenemos tres niveles de complejidad:

Nivel 1: Operaciones vector-vector.

Nivel 2: Operaciones vector-matriz.

Nivel 3: Operaciones matriz-matriz.

Lapack está diseñado desde el principio para aprovechar el nivel 3 de BLAS resolviendo los siguientes tipos de problemas: sistemas de ecuaciones lineales , problemas de mínimos cuadrados, problemas de valores propios y valores singulares y factorización de matrices principalmente.

Licencia BSD.

[NumPy](#)

Numpy se emplea para realizar cálculos científicos con el lenguaje Python. Fundamentalmente para realizar cálculos con matrices N-dimensionales, álgebra lineal, la Transformada de Fourier y números aleatorios.

Licencia BSD.

[Units](#)

Utilidad de consola para realizar conversiones de valores numéricos entre distintas escalas y conversiones de temperaturas (Fahrenheit to Celsius).

Ejemplo:

```
~$ units -t '500GB' 'GiB'
```

```
465.66129
```

Licencia GNU GPL.

[Octave](#)

GNU Octave es un lenguaje de alto nivel para cálculo numérico, compatible con Matlab, desde la consola (es necesario un front-end para tener una GUI). Octave se emplea para resolver problemas de álgebra lineal, integración de funciones y ecuaciones diferenciales, operaciones con polinomios, cálculo de raíces de ecuaciones no lineales. Permite usar módulos escritos en otros lenguajes como Fortran, C y C++.

Licencia GNU GPL.

GNUPLOT

GNUplot es una utilidad de consola para generar gráficos a partir de funciones matemáticas o conjuntos de datos. También es usada como motor de gráficos para Octave. Soporta gráficos en 2D y 3D y varios formatos de salida (jpeg, png, pdf, LaTeX, etc).

Otras licencias (Free software únicamente el código fuente).

Eclipse

Eclipse es una plataforma de desarrollo Open Source, muy conocida por desarrollar un IDE Java aunque también participa en muchos otros proyectos con el propósito de gestionar todo el ciclo de vida del software.

Existe una amplia comunidad en torno a esta plataforma y numerosos recursos (ciertamente para Java hay, virtualmente, infinidad de recursos), tantos que en su web lo califican de auténtico ecosistema.

Licencia [Eclipse Public License](#).

¿Por qué es interesante Scientific Linux 6?

Viendo la tabla anterior, no parece muy 'científico' pero con unos pocos clics del mouse ya tenemos activados los repositorios necesarios para acceder a una importante colección de aplicaciones e instalarlas desde el Package Manager no ofrece mayor dificultad. Destacar también el soporte a otros sistemas de archivos en red -infiniband, OpenAFS- orientados a computación GRID. La idea que persiguen sus desarrolladores no es tanto proveer una gran colección de aplicaciones como servir de base para que los distintos laboratorios y cumple este propósito con creces.

Scientific Linux no es una distribución 'beta', está basada en RHEL siendo más segura y estable que si fuera una variante de Fedora (la versión libre de RedHat).

El instalador es sencillo (Anaconda), en poco tiempo tenemos Scientific Linux ejecutándose en nuestro equipo y al usar el escritorio GNOME no hay demasiadas diferencias con otros 'flavours' de Linux.

2.2 Características de Poseidon Linux.

Esta distribución es desarrollada y mantenida con el apoyo de la Universidade Federal do Rio Grande ([FURG](#)) y la Universidad Federal de Pelotas ([UFPEL](#)), en el estado de Rio Grande do Sul, Brasil. Está muy influenciada por los desarrolladores originales, centrados en investigación oceanográfica y bioquímica. La versión que vamos a evaluar es la 3.2 basada en Ubuntu 9.04 para máquinas de 32 bits. Actualmente está en fase de desarrollo la versión 4 (basada en Ubuntu 10.04 LTS).

La primera impresión es que contiene un conjunto muy amplio de aplicaciones, cubriendo desde la bioinformática a la astronomía. Todas las aplicaciones están disponibles en una sesión 'live'.

Grupo	Aplicación
Programación	<ul style="list-style-type: none">• Bluefish editor (editor web)• Geany (IDE)• Qt3/Qt4 Assistant• Qt4 Designer• Qt4Linguist
Edición de textos	<ul style="list-style-type: none">• Lyx (LaTeX)• OpenOffice
Bases de Datos	<ul style="list-style-type: none">• Librerías para satisfacer dependencias de Mysql
Bioinfo	<ul style="list-style-type: none">• ClustalX• Cn3D NCBI• DDV Sequence Alignment Viewer• Entrez NCBI• fastDNAMl• OneD• PerlPrimer• PyMOL• SequinDNA• Tree-Puzzle
Química	<ul style="list-style-type: none">• Chemtool• Gperiodic• xDrawchem
GIS/CAD	<ul style="list-style-type: none">• GPS drive• Grass GIS• Imagelinker• Qcad• QuantumGIS• SagCAD

Grupo	Aplicación
Gráficos	<ul style="list-style-type: none"> • G3Data • LabPlot • QtiPlot • Blender • Gwyddion • Hagin • Inkscape • Scribus
Matemáticas y estadística	<ul style="list-style-type: none"> • fityk • GNU Octave • R Commander • Rkward • wxMaxima
Astronomía	<ul style="list-style-type: none"> • Stellarium • Open Universe Space Simulator
Otros	<ul style="list-style-type: none"> • BOINC • imagej • OpenDX Data Explorer • XTide

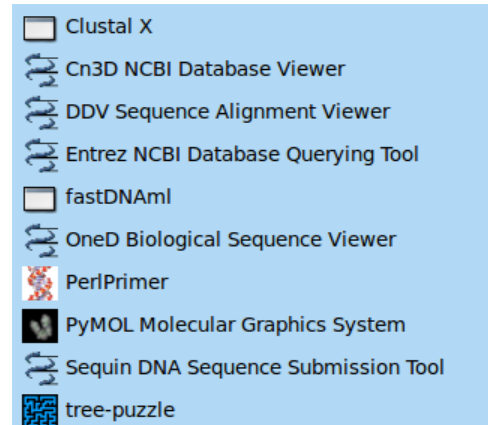
En su desarrollo se han cuidado todos los aspectos visuales incluso con menús personalizados, fondos de escritorio y temas, como podemos apreciar en la imagen.



Imagen 3: Poseidon Linux Desktop.

Características de las aplicaciones de carácter científico.

BIOINFO



ClustalX

Alineación múltiple de secuencias de ácidos nucleicos y proteínas. Clustal X es un interfaz para el programa de alineación múltiple de secuencias ClustalW. Esto proporciona un entorno integrado para realizar secuencias múltiples y alineaciones de perfil y analizar los resultados. Se puede realizar un análisis de calidad de alineación y podemos destacar segmentos de notación baja o residuos excepcionales.

Licencia GNU GPL.

Cn3D, DDV Sequence Alignment Viewer y OneD-Viewer Sequence Display.

Cn3D es un plugin para navegador web que le permite ver estructuras 3-D obtenidas de Entrez del NCBI. Cn3D muestra la estructura, la secuencia y la alineación, y permite la anotación y la alineación editando los rasgos.

Una de sus ventajas es su capacidad de correlacionar la información de secuencia y la estructura: por ejemplo, un científico rápidamente puede encontrar los residuos en una estructura de cristal que corresponden a mutaciones de enfermedad conocidas, o residuos conservados activos de una familia de secuencia homóloga.

One-D es un visor incluido con Cn3D.

Sequin DNA

Sequin es un instrumento de software independiente desarrollado por el NCBI para evaluar y poner al día entradas de la base de datos de secuencias GenBank. Maneja secuencias cortas mRNA, y secuencias largas, múltiples anotaciones, secuencias gapped, o filogenéticas y estudios de poblaciones. Por ejemplo, Sequin obtiene el código genético propio del nombre del organismo y automáticamente determina intervalos de región de codificación sobre la secuencia de nucleótido por la traducción de la secuencia de proteína.

Entrez, CN3D, SequinDNA, DDV Sequence Display y One-D Viewer Sequence Display pertenecen al mismo grupo de aplicaciones y recursos del [NCBI](#) (National Center for Biotechnology Information).

Licencia: Libre copia y distribución, pero bajo control de leyes U.S.A.

[fastDNAmI](#)

fastDNAmI es un programa para estimar los árboles filogenéticos de máxima similitud para secuencias de nucleótidos.

'fastDNAmI' optimiza la determinación de la longitud de las ramas, de acuerdo a una modificación del método de Newton.

El repositorio de Ubuntu informa de la no disponibilidad de la homepage de este software, parece un proyecto muerto. No hay información de la licencia ni está disponible el paquete para su descarga desde dicha web.

[PerlPrimer](#)

PerlPrimer es una aplicación escrita en Perl para diseñar cebadores para Reacciones en Cadena de la Polimerasa (PCR), bisulfito PCR, PCR en tiempo real (QPCR) y secuenciación. Permite automatizar y simplificar el proceso de diseño del cebador.

Licencia Free ,Open Source (sin indicar tipo).

[PyMOL](#)

PyMOL es una herramienta de visualización para biología estructural y es extensible mediante el lenguaje de programación Python, permite realizar análisis complejos de estructuras moleculares utilizando bibliotecas disponibles para Python como NumPy o pylab.

Licencia Python License (CNRI Python License))Compatible con GNU GPL v2.1.1

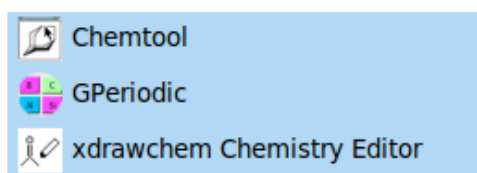
[Tree Puzzle](#)

TREE-PUZZLE es un programa para reconstruir árboles filogenéticos (relaciones evolutivas entre especies) de datos de secuencias moleculares por probabilidad.

El propósito es trazar un mapa de probabilidad, un método de investigar la relación de una rama interna sin calcular un árbol total y visualizar el contenido filogenético de una secuencia.

Licencia: no hay información, solo muestra el © de los autores.

CHEMISTRY



[ChemTool](#)

Chemtool permite dibujar moléculas orgánicas mediante una GUI e interactuando con el mouse.

Licencia GNU GPL.

[Gperiodic](#)

Es la tabla periódica de elementos, muestra información detallada de 118 elementos químicos

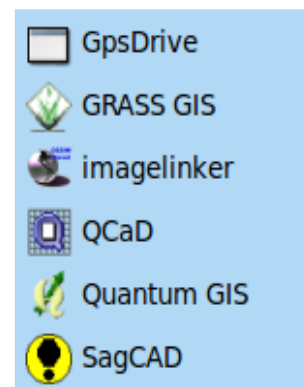
Licencia GNU GPL.

XdrawChem

Es una alternativa a [Chemdraw](#), usado en laboratorios de investigación para construir diagramas moleculares y es compatible con los archivos creados por éste. También puede leer MDL Mollfiles, un tipo de archivo que almacena información sobre átomos, relaciones y conectividad de una molécula.

Licencia GNU GPL.

GIS/CAD



GPS Drive

Es un software de navegación GPS, emplea mapas OpenStreetMap que tienen licencia Creative Commons y se construyen de modo colaborativo entre todos los usuarios de los mismos.

Licencia GNU GPL.

Grass GIS

Geographic Resources Analysis Support System (GRASS) es un sistema de información geográfica libre que permite el análisis de imágenes vectoriales o raster, visualización y creación de mapas en 2D y 3D que además permite scripting y programación de módulos a medida. Soporta varios gestores de bases de datos (hay varios drivers) entre ellos PostgreSQL y Mysql. Tiene herramientas de visualización en 3D, animación de imágenes, etc.

Es un proyecto de [Open Source Geospatial Foundation](#).

ImageLinker

ImageLinker es una aplicación de procesamiento y visualización de datos GIS u obtenidos por otros medios (radares, satélites, etc). Permite construir capas a partir de las imágenes según las características que queramos destacar.

Es también un proyecto de [Open Source Geospatial Foundation](#).

QuantumGIS

Es otro software GIS, con capacidad de trabajar con diferentes formatos de archivos (vectoriales y raster). Soporta formatos de archivo como [PostGIS](#) (que usa tablas de PostgreSQL), ESRI (formato de [ArcGIS](#)), SDTS ([Spatial Data Transfer Standard](#)), [Mapinfo](#) y [GML](#). Soporta también fotos aéreas y modelos digitales topográficos, datos online según el formato [WMS](#), etc.

Permite crear, editar y publicar mapas en Internet. Crear, modificar y exportar la información asociada a los mapas así como realizar análisis de terrenos, hidrológicos, etc. QuantumGIS requiere de GRASS, básicamente GRASS ofrece más posibilidades de análisis y usamos QGIS como una GUI para GRASS. Permite ampliar funcionalidades mediante Plugins.

Licencia GNU GPL.

Nota acerca del software GIS en Ubuntu: pertenecen a un repositorio PPA de launchpad, ubuntugis-unstable.

qCAD

Este programa soporta diseño de planos en 2D, es simple y además posee una amplia biblioteca de piezas CAD. Soporta capas, bloques y tiene las herramientas típicas de otros productos para selección y dibujo.

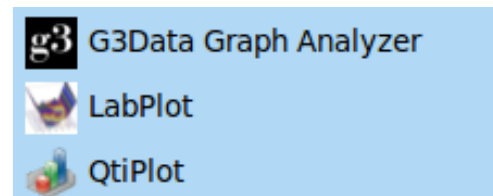
El código fuente para Linux y Mac OS X tiene licencia GNU GPL.

SagCAD

Programa CAD para diseño en 2D que permite importar y exportar archivos de AutoCAD. La última versión es del año 2009 y está etiquetada como 'beta'.

Licencia GNU GPL.

GRAPHS



G3Data

Esta herramienta permite, mediante la indicación de unos pocos parámetros, extraer los datos que originan un determinado gráfico. Su utilidad reside en que podemos extraer los datos de un gráfico desde una imagen en una publicación indicando parámetros como el valor mínimo y máximo en los ejes, las coordenadas de un punto en la imagen, etc.

Licencia GNU GPL.

LabPlot

Este software permite generar gráficos tanto de funciones matemáticas como a a partir de tablas de datos tanto en 2D como 3D. Además podemos hacer análisis de datos con cálculo de máximos y mínimos, medias, etc. Permite realizar regresión e interpolación, uso de funciones del paquete estadístico R y, generar histogramas.

Licencia GNU GPL.

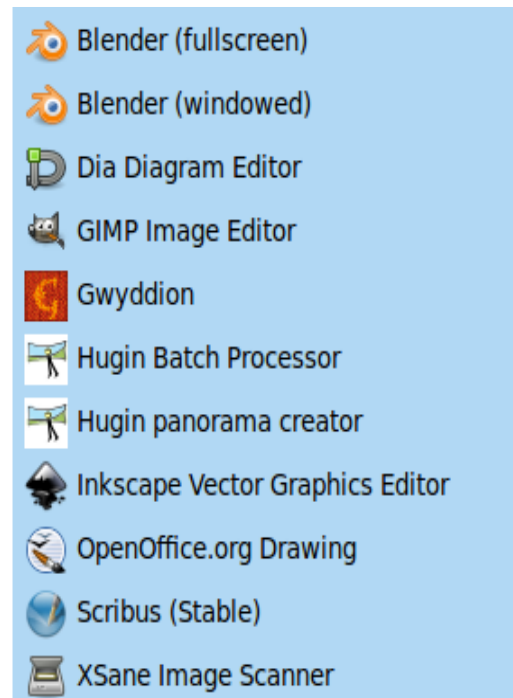
QtiPlot

Software de análisis y visualización en 2D y 3D, también soporta datos tabulados y funciones matemáticas. Los gráficos se generan dinámicamente, si los datos de origen se modifican el gráfico se actualiza automáticamente. Los gráficos obtenidos pueden ser exportados a múltiples formatos (incluyendo jpeg y pdf). Entre otras posibilidades permite realizar análisis de interpolación, FFT (Transformada rápida de Fourier), etc.

Licencia GNU GPL.

QtiPlot y Labplot son alternativas libres al software propietario Origin de [OriginLab](#) y ambos permiten importar archivos generados con ese software propietario. Origin tiene un coste que varía desde los 50\$ / año para una licencia de estudiante a los 1700\$ /año para una licencia profesional de la versión OriginPro.

GRAPHICS



Blender

Herramienta para la creación de contenidos 2D y 3D. Funcionalidades incluidas: modelado, mapeado uv, texturizado, rigging, skinning, animación, simulación de partículas, etc. Es una herramienta apropiada para la creación de contenido multimedia y modelado 3D.

Licencia GNU GPL.

Gwyddion

Este programa está enfocado a SPM (scanning probe microscopy) que usa imágenes tridimensionales (sólo superficies) incluso a escala atómica obtenidas por microscopio para su posterior análisis. Permite visualizar en 3D los materiales y realizar análisis aritméticos y estadísticos, transformaciones según la FFT (Transformada rápida de Fourier). Sus funcionalidades pueden ampliarse mediante plugins.

(Visitar la web del proyecto para listado completo de [características](#)).

Licencia GNU GPL.

Hugin

Permite componer imágenes panorámicas a partir de varias imágenes planas.

Licencia GNU GPL.

Inkscape

Este editor de gráficos vectoriales usa el formato SVG y es una alternativa de código abierto a Illustrator o Freehand, permitiendo importar desde formatos como PostScript, jpeg o png y exportar a formato png. Tiene herramientas de creación de superficies (rectángulos, elipses, etc) textos y las operaciones necesarias para manipularlos (transformaciones, clonar, rellenar, etc).

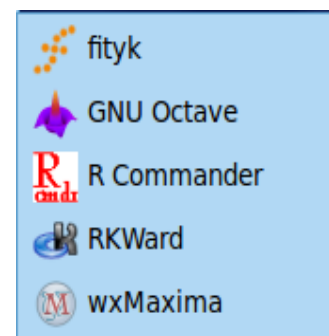
Licencia GNU GPL.

Scribus

Herramienta de creación de documentos al estilo de PageMaker o InDesign, menos potente pero con una interfaz sencilla, más orientada al diseño y composición de documentos. Permite exportar y editar en formato pdf.

Licencia GNU GPL.

MATH/STATISTICS



fityk

Software de análisis de datos y ajuste de curvas no lineales (interpolación no lineal). Sus ámbitos de uso principales son la cromatografía, difracción de polvo, fotoluminiscencia y espectroscopía de rayos y fotoelectrones. Su uso principal es para ajustar datos a una curva 2D. Posee muchas funciones integradas y la posibilidad de definir funciones personalizadas y añadido de plugins.

Licencia GNU GPL.

GNU Octave

GNU Octave es un lenguaje de programación de alto nivel para cálculo numérico. Mediante una interfaz de consola permite resolver problemas lineales y no lineales numéricamente, aunque hay disponible un front-end (qtoctave) que provee un entorno GUI más amigable. Octave resuelve problemas de álgebra lineal, cálculo de raíces, tratamiento de polinomios, integración. Permite cargar módulos escritos en otros lenguajes de programación como Fortran o C.

Licencia GNU GPL.

R Commander

Es una interfaz básica para el paquete estadístico R. Es una GUI simple dividida en tres partes: Script Window , Output Window y MessagesWindow (los gráficos aparecen en una ventana separada Graphics Device Window). Destacamos los menús:

- Data. Para manipular e importar datos.
- Statistics. Análisis estadísticos básicos (tablas de contingencia, medidas de centralización y dispersión, correlaciones, etc).
- Graphs. Creación de gráficos simples.
- Models. Análisis de la varianza , contraste de hipótesis, intervalos de confianza, etc.
- Distributions. Probabilidades, cuartiles, gráficos de estadísticas comunes.
- Fit Models. Modelos de regresión.

Licencia GNU GPL.

Rkward

Es otro front-end para R, diseñado para el entorno de escritorio KDE. Es algo más que una GUI, permite guardar workspaces, crear scripts, configurar los plugins de R instalados y usar las funciones estadísticas, además de generar gráficos a partir de los datos.

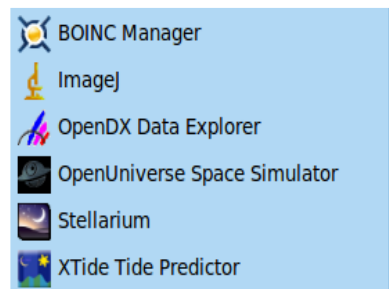
Licencia GNU GPL.

WxMaxima (maxima + GUI).

WxMaxima es la GUI para la aplicación Maxima. Maxima es indicado para el cálculo simbólico: límites, diferenciación, integración, series de Taylor, ecuaciones diferenciales y lineales, polinomios, matrices, vectores, tensores y conjuntos. Maxima también permite cálculo numérico y generación de gráficos en 2D y 3D.

Licencia GNU GPL.

Sin clasificar.



BOINC

Berkeley Open Infrastructure for Network Computing (BOINC) es un software de computación distribuída (grid) inicialmente vinculado al proyecto [SETI@home](#) pero que permite seleccionar entre varios proyectos para instituciones como el CERN, Universidades y centros de investigación. Su funcionamiento se basa en aprovechar los tiempos de inactividad del ordenador para realizar cálculos matemáticos en aras de resolver una tarea asignada por el proyecto y devolver los resultados obtenidos.

Licencia GNU Lesser General Public License (LGPL).

IMAGEJ

Es una herramienta de análisis de imágenes. Entre sus funciones encontramos: medir el área, media, desviación estándar, mínimo y máximo de selección o imagen. Medidas de longitudes y ángulos, generar histogramas y perfiles de gráficos. Soporta plugins y macros, maneja archivos TIFF, BMP, JPEG, PNG entre otros. Tiene herramientas de selección (rectangular, elipse) de mejora de la imagen, escalado, recorte y volteado. También puede realizar análisis comparativos entre dos imágenes y operaciones como análisis de Fourier o filtrado para reducir el ruido (ver [median filter](#)). Plugins disponibles en: <http://rsb.info.nih.gov/ij/plugins/index.html>.

Licencia [libre, Public domain](#).

OpenDX Data Explorer

Con este programa importamos y caracterizamos datos, los procesamos y generamos una imagen explicativa de los mismos. Ofrece herramientas para manipular, transformar y renderizar los datos, permitiendo varios métodos de análisis (basados en puntos, líneas, áreas, volúmenes o primitivas geométricas, es decir, representaciones gráficas de funciones matemáticas).

Licencia [IBM Public License](#).

Open Universe Space Simulator

Simulador del sistema solar en tiempo real, muestra imágenes en 3D permitiéndonos navegar por los planetas, orbitarlos, etc.

Licencia GNU GPL.

Stellarium

Programa de muy sencillo manejo que nos permite ver el cielo como si empleáramos un pequeño telescopio, podemos ver el cielo en tiempo real fijando nuestra posición de observador.

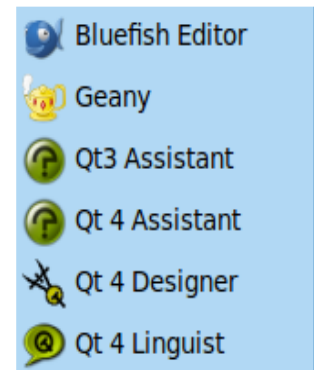
Licencia GNU GPL.

Xtide Tide Predictor

Este software nos permite conocer el estado actual de las mareas y su predicción en gran variedad de formatos (texto, gráficos, calendarios). Emplea el algoritmo del [National Ocean Service](#) de U.S.A

Licencia GNU GPL (excepto algunos ficheros fuente que son de dominio público).

PROGRAMACIÓN



[Bluefish editor](#)

Editor rápido y ligero (según describen en su web y he comprobado) que soporta múltiples lenguajes. Incluye una barra de herramientas HTML, soporta múltiples codificaciones de caracteres (por defecto UTF8). Permite trabajar con archivos remotos, crear proyectos como en IDE's avanzados. Está orientado a diseño web pero podría usarse en cualquier ámbito de programación.

Licencia GNU GPL.

[Geany](#)

Editor ligero, pensado como un IDE (con resaltado de texto según el lenguaje) que soporta multitud de tipos de archivo: C, C#, Java, Pascal, LaTeX, PHP, HTML, SQL, Haskell, por citar algunos (son más de 40). Es ligero e independiente (sólo requiere librerías GTK2 para funcionar). Permite la instalación de Plugins (hay un proyecto propio de [plugins disponibles](#)).

Licencia GNU GPL.

[Qt3/Qt4 Assistant, Qt4 Designer y Qt4Linguist.](#)

Conjunto de herramientas de desarrollo para QT que pueden integrarse en IDE's como Eclipse, sus funciones principales son:

- Qt Designer para crear GUI de manera rápida y sencilla (similar a como lo haríamos en un IDE como NetBeans para Java).
- Qt Linguist herramienta de internacionalización .
- Qt Assistant es una utilidad para poder acceder offline a toda la documentación de QT. Es similar a usar un navegador web, con una barra de marcadores lateral y usa hiperenlaces para navegar por la documentación (similar a la documentación online de Java cuando usamos frames).

Licencia GNU Lesser General Public License (LGPL).

Conclusiones acerca de Poseidon Linux.

Es llamativo el número de aplicaciones disponibles (más de 40), organizadas en sus correspondientes menús. Algunas de las aplicaciones incluidas, como Stellarium o Gperiodic son más apropiadas de entornos educativos, por contra, Blender es muy potente para crear imágenes 3D pero está muy enfocada a la creación multimedia y no tanto al ámbito científico. El editor Geany no puede considerarse un IDE integrado ya que no posee las herramientas avanzadas de éstos (gestión de proyectos, importación de librerías, desarrollo de GUI, etc).

Existe cierta duplicidad o solapamiento de aplicaciones con funciones muy similares, destacamos:

- Octave y maxima, ambas permiten cálculo numérico (maxima permite también cálculo simbólico).
- QCAD y SagCAD son herramientas que básicamente tienen las mismas funcionalidades.
- Rcommander y Rkward son el front-end del software estadístico R, aunque Rkward tiene algunas funcionalidades propias de un editor avanzado, son casi equivalentes.
- Hay dos aplicaciones que son inclasificables: primero GPS Drive, que está orientada a sustituir herramientas cerradas de navegación GPS propias de los típicos navegadores para vehículos, y Xtide Predictor para realizar calendarios de mareas.

Por otro lado, tenemos un amplio catálogo de aplicaciones, cubriendo muchos aspectos del ámbito científico: biología, química, matemáticas y estadística, además de las aplicaciones GIS y las utilidades relacionadas con la generación de gráficos y manipulación de imágenes llegando incluso a aplicaciones de análisis de imágenes obtenidas con microscopio (esta distribución en su origen estaba vinculada a la investigación oceanográfica).

Personalmente creo excesivo el número de aplicaciones propuesto, con solapamiento de funciones entre ellas (qCAD/SagCAD o R Commander y Rkward por citar algunas), muy heterogéneo con algunas propias de entornos educativos (Stellarium, GPeriodic) y otras demasiado específicas con curvas de aprendizaje muy pronunciadas (Blender). Carece también de un IDE potente y de un soporte completo para bases de datos (MySQL o PostgreSQL). Desde mi punto de vista sería una opción más interesante crear metapaquetes o scripts que permitieran instalar las aplicaciones para cada ámbito según las necesidades de cada usuario.

2.3 *UbuntuScience.*

En la documentación de Ubuntu, encontramos una referencia interesante a software de carácter científico en [Ubuntu Science](#). No es una distribución, es un listado del software disponible para Ubuntu y en que repositorio se encuentra (una forma indirecta de indicarnos el tipo de licencia).

Encontramos listados los paquetes por secciones y con enlaces a repositorios y referencias a distribuciones de carácter científico.

Las secciones presentes son:

1. [Astronomy](#)
2. [Biology](#)
3. [Biophysics](#)
4. [Chemistry](#)
5. [Earth Sciences \(Geology/Geography\)](#)
6. [Ecology and Environmental Sciences](#)
7. [Neuroscience](#)
8. [Physics](#)

Bastantes paquetes aquí referenciados han sido empleados en Poseidon Linux, en el apartado de química aparecen muchas más aplicaciones disponibles, incluso un listado de aplicaciones de todo tipo y licencia en [Linux for chemistry](#). En Ubuntu Science podemos encontrar incluso la aplicación Google Earth, presente en dos secciones: Astronomy, por la posibilidad de ver el cielo con esta aplicación, y en Earth Sciences, mas apropiada desde mi punto de vista y que podríamos asimilar con las aplicaciones GIS en Poseidon Linux .

Podemos considerar esta web como un punto de partida útil para encontrar aplicaciones en alguna de las áreas de interés, pero en mi opinión su mayor defecto es no incluir ninguna sección dedicada a software matemático o estadístico. Podemos considerar de forma genérica, que, con una hoja de cálculo ya podemos realizar estadísticas simples e incluso mostrar resúmenes de datos en diversos diagramas, pero en casi cualquier área de investigación son necesarios estudios estadísticos. Un software de cálculo numérico el estilo de Octave o Maxima también resulta de utilidad sobre todo en resolución de problemas que impliquen cálculo integral o diferencial que podemos encontrar en cualquier ámbito científico.

2.4 Ubuntu[sci] team (scubuntu).

Este proyecto, también conocido como [Ubuntu Science remix](#) y tiene 4 variantes que parten de una base común denominada como [Ubuntusci-base-desktop](#) que según reza en la web es el sistema de partida para científicos e investigadores, está basado en Ubuntu 9.10 y dispone de imágenes para arquitectura i386 y amd64 . Las cuatro variantes son:

- Ubuntusci GeoSciences/GIS
- Ubuntu^{sci} BioSciences
- Ubuntu^{sci} Mathematics
- Ubuntu^{sci} Computer Science

Esta distinción es una gran ventaja desde mi punto de vista, evita el gran número de aplicaciones existentes en Poseidon Linux.

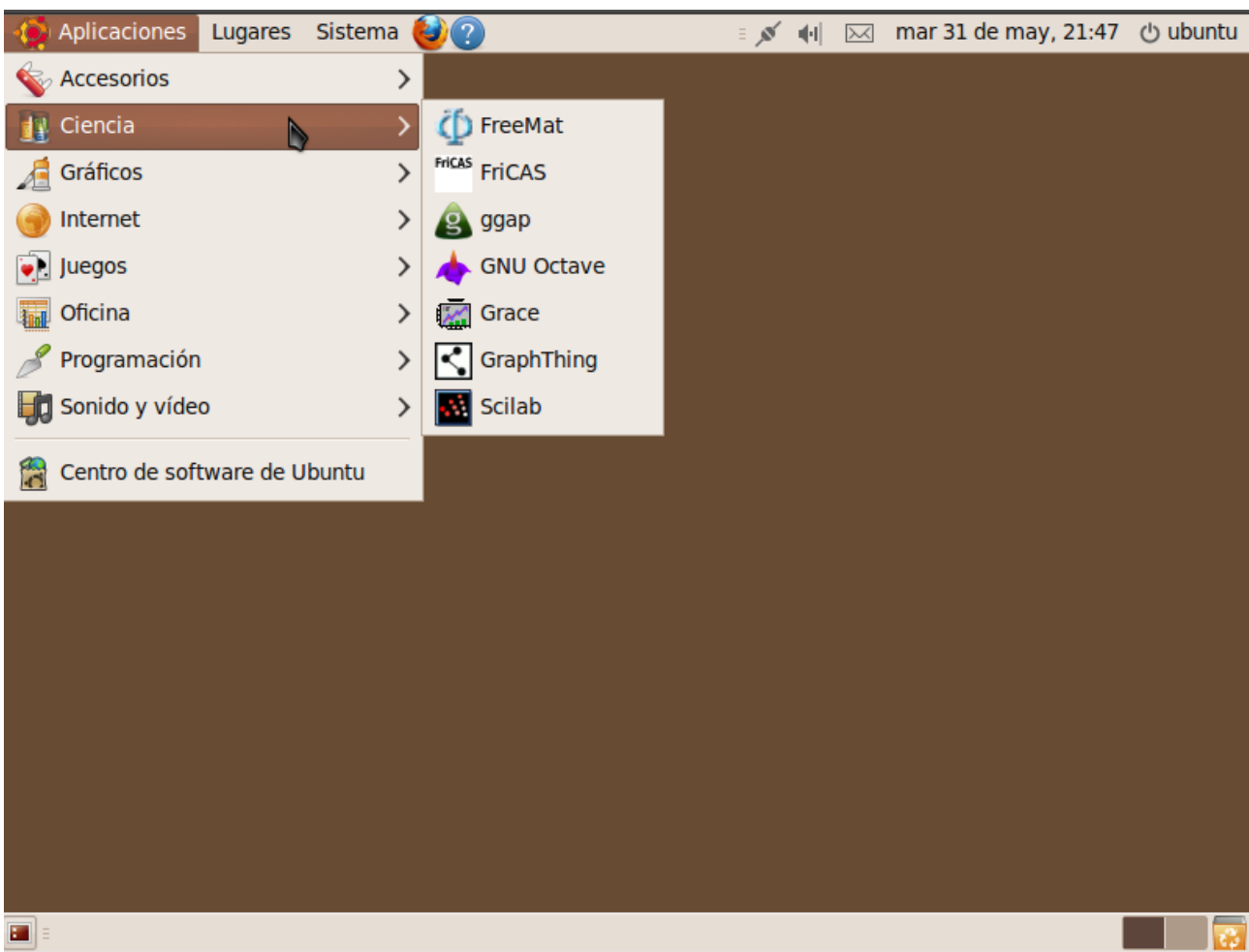


Imagen 4: Escritorio de Ubuntu sci-Math

En versiones anteriores de este proyecto solo estaba disponible una variante basada en Ubuntu 8.10 para la arquitectura i386 que comprendía aplicaciones de BioSciences y GIS.

El sistema base de Ubuntu^{sci} tiene unos objetivos muy similares a los de este Trabajo Fin de Carrera. Provee de aplicaciones para el tratamiento estadístico de datos (R) un gestor de referencias bibliográficas (realmente incluye dos: jabref y bibus), aplicaciones de cálculo numérico y algebraico (octave y maxima respectivamente), generación de gráficos (GNUplot, octave-plot) y edición de textos con LaTeX.

Es un proyecto interesante por distinguir ámbitos de uso, pero no parece muy vivo: la última modificación según la wiki del proyecto es de 22 de Abril de 2010. La última versión publicada es de hace más de un año y no hay anuncio de nuevas versiones a pesar del calendario de releases de Ubuntu y de la publicación de Ubuntu 10.04 LTS.

3 Objetivos

Se pretende construir un LiveDVD a partir de una imagen estándar de la distribución Ubuntu, ubuntu-10.04.2-desktop-i386.iso que es una versión Long Term Support (LTS), dirigida a un usuario objetivo del entorno científico tanto académico como en el laboratorio. Se intentará, en la medida de lo posible, incluir software de código abierto para cubrir las necesidades específicas de esta distribución de ámbito científico.

Las herramientas que considero imprescindibles son:

- Software de Tratamiento estadístico, incluiremos R en este apartado.
- Para el cálculo numérico emplearemos Octave (con un front-end, QtOctave).
- El desarrollo de gráficos recaerá en QtiPlot y GNUPlot.
- Edición de documentos científicos con Lyx.
- La gestión de referencias bibliográficas la realizaremos con Bibus y JabRef.
- Incluiremos una potente base de datos relacional y OO, PostgreSQL.
- Para el desarrollo de aplicaciones dispondremos de NetBeans IDE.

4 Metodología y entorno de desarrollo

Los pasos a seguir serán:

1. Instalar el software necesario en la máquina anfitrión: Virtual Box 4, squashfs-tools, genisoimage y Ubuntu Customization Kit.
2. Obtener la imagen iso del LiveCD de Ubuntu (ubuntu-10.04.2-desktop-i386.iso).
3. Eliminar language-packs y ubiquity con UCK, generando una nueva imagen iso.
4. Descomprimir la imagen y montar el sistema de ficheros squashFS.
5. Crear un entorno chroot para modificar la distribución.
6. Modificar la distribución (paquetes, apariencia, navegador).
7. Comprimir la distribución personalizada.
8. Compilar la imagen iso de la nueva distribución.
9. Testing del LiveDVD.
10. Publicar en [Dropbox](#).

El testing lo realizaremos en dos fases, inicialmente emplearemos una máquina virtual creada con Oracle VirtualBox y, una vez superadas las pruebas, emplearemos otra máquina para ejecutar el LiveDVD.

Entorno hardware del sistema anfitrión.

BIOS Information

Vendor: Phoenix Technologies, LTD

Version: ASUS M2NPV-VM ACPI BIOS Revision 1401

Base Board Information

Manufacturer: ASUSTek Computer INC.

Product Name: M2NPV-VM

Processor Information

Socket Designation: Socket AM2

Type: Central Processor

Family: Athlon 64

Version: AMD Athlon(tm) 64 X2 Dual Core Processor 6000+

Voltage: 1.3 V

External Clock: 200 MHz

Current Speed: 3000 MHz

Memory

MemTotal: 2831076 kB

Disk Devices

Disk /dev/sda: 500.1 GB, 500107862016 bytes

Disk /dev/sdb: 160.0 GB, 160041885696 bytes

VGA

VGA compatible controller: nVidia Corporation C51PV [GeForce 6150]

Información obtenida con:

dmidecode, fdisk -l, meminfo, cpuinfo, lspci | grep VGA.

Entorno de la máquina virtual VBOX:

Memoria de vídeo: 128 MB
Memoria base: 768 MB
Procesadores: 2 (AMD Athlon(tm) 64X2 Dual Core Processor
6000+)
Orden de inicio: CD/DVD-ROM, Disco rígido
Almacenamiento:
IDE Controller
IDE Primary Master (CD/DVD)
IDE Secondary Master (CD/DVD)
SATA Controller SATA Port 0: testing.vdi (Normal, 8,00 GB)
Acceleración: VT-x/AMD-V, Paginación anidada
RED Adaptador 1: Intel PRO/1000 MT Desktop (NAT)

El sistema anfitrión será Ubuntu 10.04 Lucid Lynx en su versión de 64 bits (Linux kernel 2.6.32 para x86/x86_64 y GNOME 2.30.2), emplearemos máquinas virtuales de tamaño fijo de 8 GB.

El entorno de desarrollo que emplearemos comprende las siguientes herramientas:

1. Ubuntu Customization Kit (UCK)
2. Oracle VirtualBox
3. SquashFS

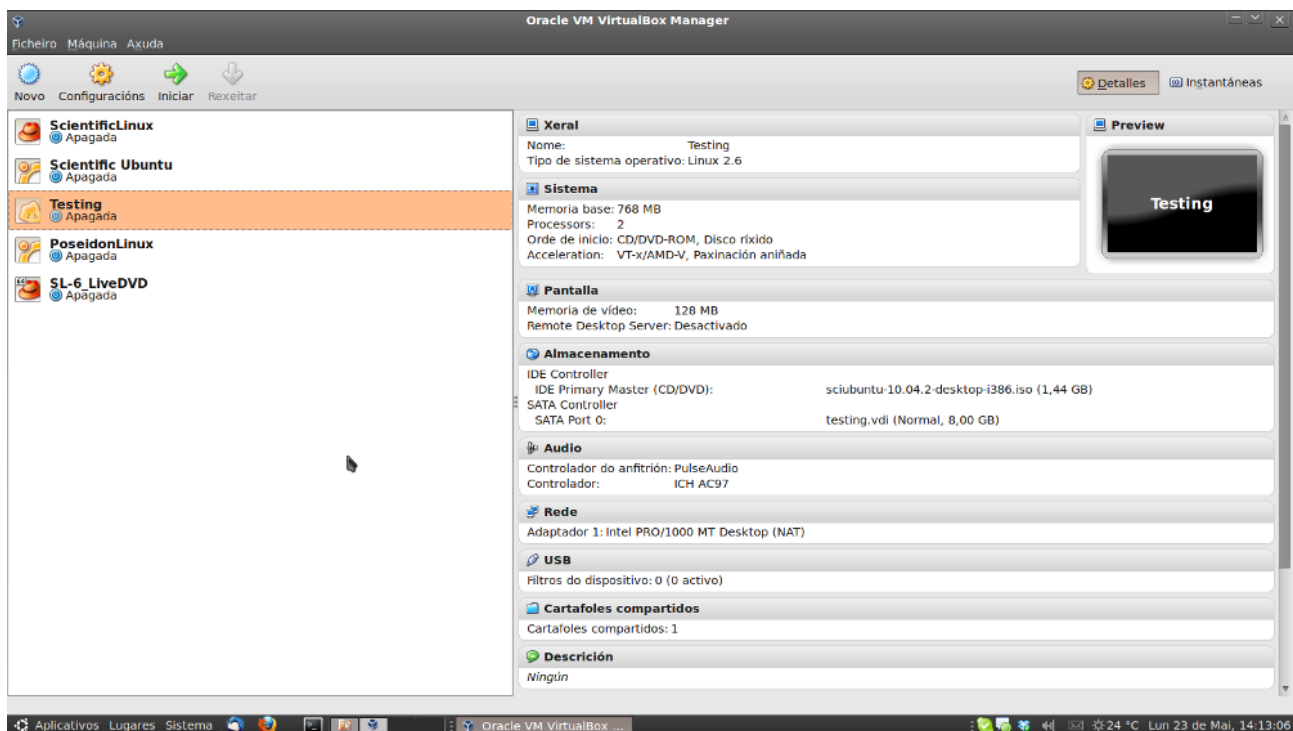


Imagen 5: Oracle Virtual Box, Pantalla principal.

4.1 Script de modificación de la imagen iso y recompilado.

```
# instalamos los paquetes necesarios para el proyecto
sudo aptitude install squashfs-tools genisoimage

# obtener el sistema base y moverlo al directorio de trabajo
mkdir livecdtmp
mv ubuntu-10.04.2-desktop-i386.iso livecdtmp
cd livecdtmp

# montar la imagen iso
mkdir mnt
sudo mount -o loop ubuntu-10.04.2-desktop-i386.iso mnt

# extraer el contenido de la imagen iso
mkdir extract-cd
rsync --exclude=/casper/filesystem.squashfs -a mnt/ extract-cd

# extraer el sistema
sudo unsquashfs mnt/casper/filesystem.squashfs
sudo mv squashfs-root edit

# pasos para preparar la jaula chroot
#
# preparar la conexion de red
sudo cp /etc/resolv.conf edit/etc/
sudo cp /etc/hosts edit/etc/

# montar directorios necesarios
sudo mount --bind /dev/ edit/dev
sudo chroot edit
mount -t proc none /proc
mount -t sysfs none /sys
mount -t devpts none /dev/pts

# para evitar problemas de localizacion e importar las GPG keys
export HOME=/root
export LC_ALL=C
```

```
# Modificaciones del sistema base
# pasos previos para modificar paquetes con apt
dbus-uuidgen > /var/lib/dbus/machine-id
dpkg-divert --local --rename --add /sbin/initctl
ln -s /bin/true /sbin/initctl

#
# primero se modifican los repositorios añadiendo los necesarios
# para el software a instalar
visudo -f /etc/apt/sources.list

# recargamos los souce.list
apt-get update

# eliminamos algunos paquetes para aligerar la distribucion
# sudo apt-get autoremove --purge nom_paquete
# instalamos los paquetes requeridos para nuestro nuevo flavour
# sudo apt-get install nom_paquete

# finalización
# hacemos limpieza
sudo apt-get clean
sudo apt-get autoclean
aptitude clean

rm -rf /tmp/* ~/.bash_history
rm /etc/hosts
rm /var/lib/dbus/machine-id
rm /sbin/initctl
dpkg-divert --rename --remove /sbin/initctl
```

```

#desmontamos los sistemas de ficheros
umount /proc || umount -lf /proc
umount /sys
umount /dev/pts

# salimos de la jaula chroot
exit
sudo umount edit/dev || umount -lf /edit/dev

# Rehacemos el manifest
chmod +w extract-cd/casper/filesystem.manifest
sudo chroot edit dpkg-query -W --showformat='${Package} ${Version}\n' > extract-cd/casper/filesystem.manifest
sudo cp extract-cd/casper/filesystem.manifest extract-cd/casper/filesystem.manifest-desktop
sudo sed -i '/ubiquity/d' extract-cd/casper/filesystem.manifest-desktop
sudo sed -i '/casper/d' extract-cd/casper/filesystem.manifest-desktop

# Volvemos a comprimir el sistema de ficheros
sudo rm extract-cd/casper/filesystem.squashfs
#sudo mksquashfs edit extract-cd/casper/filesystem.squashfs
-nolzma

# usamos un mayor ratio de compresion
sudo mksquashfs edit extract-cd/casper/filesystem.squashfs -b
1048576

# actualizamos el tamaño de sistema de ficheros
printf $(sudo du -sx --block-size=1 edit | cut -f1) > extract-cd/casper/filesystem.size

# Renombramos la imagen
sudo vim extract-cd/README.diskdefines

#Borramos la suma md5 y la calculamos para la nueva iso
cd extract-cd
sudo rm md5sum.txt
find -type f -print0 | sudo xargs -0 md5sum | grep -v
isolinux/boot.cat | sudo tee md5sum.txt

```

```
#compilamos la imagen iso
IMAGE_NAME=ScientificUbuntu
sudo mkisofs -D -r -V "$IMAGE_NAME" -cache-inodes -J -l -b
isolinux/isolinux.bin -c isolinux/boot.cat -no-emul-boot -boot-
load-size 4 -boot-info-table -o ../sciubuntu-10.04.2-desktop-
i386.iso .
```

5 Desarrollo de Scientific Ubuntu

Aplicaciones que se incluyen en el LiveDVD.

Área	Aplicación	Licencia
Edición de documentos científicos	<ul style="list-style-type: none">• Lyx	GNU GPL
Programación	<ul style="list-style-type: none">• NetBeans• gFortran	CDDL GNU GPL
Planificación de Proyectos	<ul style="list-style-type: none">• planner	GNU GPL
Gráficos	<ul style="list-style-type: none">• Octave-plot• gnuplot• qtiplot	GNU GPL
Referencias bibliográficas	<ul style="list-style-type: none">• Jabref• bibus	GNU GPL
Bases de datos	<ul style="list-style-type: none">• PostgreSQL	Licencia BSD
Estadísticas	<ul style="list-style-type: none">• OpenOffice.org Calc• R	GNU Lesser General Public License GNU GPL

5.1 Aplicaciones.

Octave

GNU Octave es un lenguaje de alto nivel para cálculo numérico, compatible con Matlab, desde la consola (es necesario un front-end para tener una GUI). Octave se emplea para resolver problemas de álgebra lineal, integración de funciones y ecuaciones diferenciales, operaciones con polinomios, cálculo de raíces de ecuaciones no lineales. Permite usar módulos escritos en otros lenguajes como Fortran C y C++.

El siguiente script resuelve la Ecuación Diferencial Lineal de orden dos $y''+y'+2y=0$ con condiciones iniciales $y(0)=0$ $y'(0)=1; t \in [0,1]$

```
function xdot=fcn(x,t)
xdot(1)=x(2)
xdot(2)=-3*x(2)-2*x(1);
endfunction
t=linspace(0,1,101); x0=[0,1];
[x,info,msg]=lsode('fcn',x0,t);
plot(t,x)
```

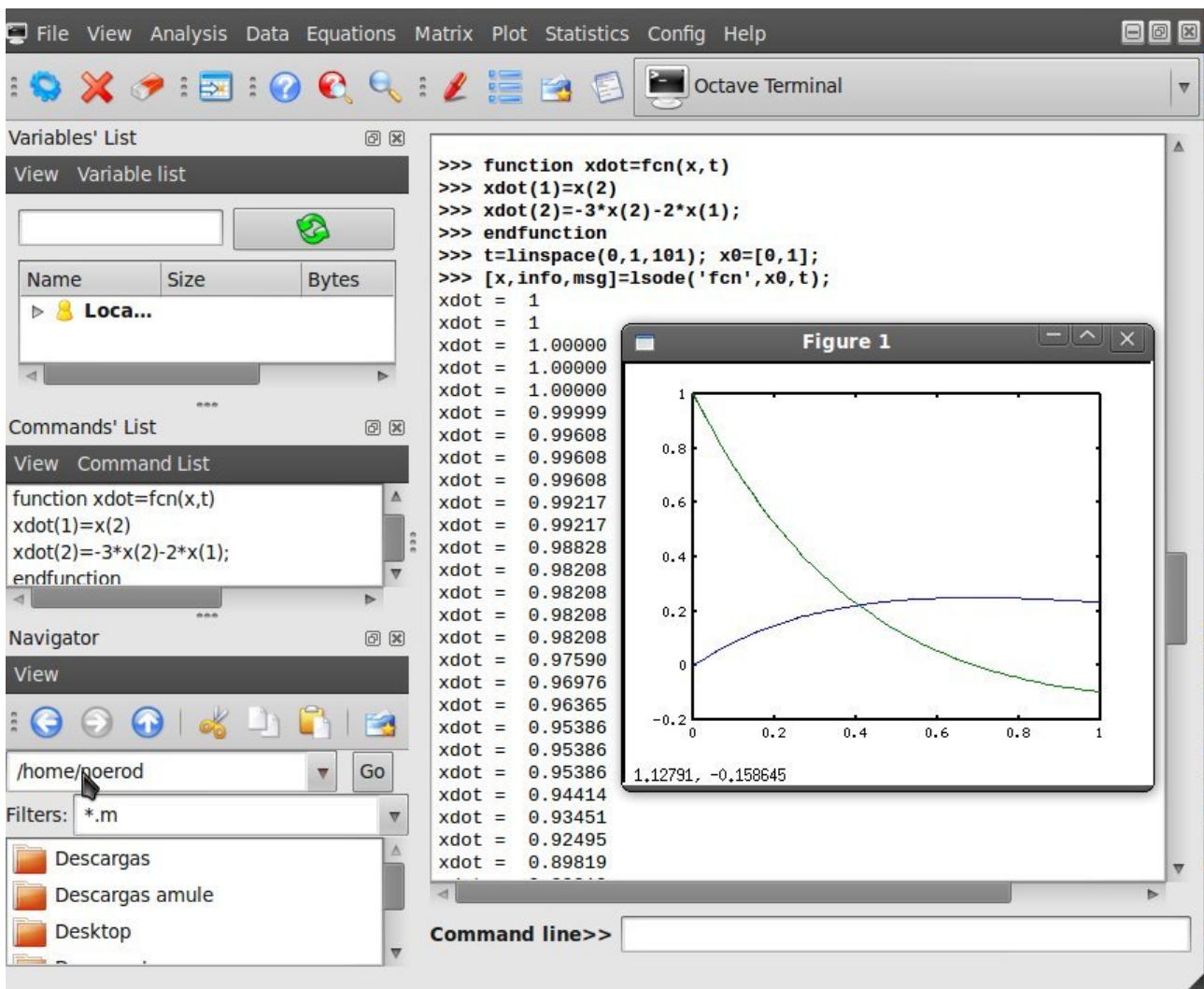


Imagen 6: QtOctave, ventana de edición y gráfica.

Licencia GNU GPL.

LyX

LyX es un procesador de textos para LaTeX que permite centrarse en la estructura del documento y no solo en la apariencia, esto conlleva un esfuerzo mayor a los principiantes. Está basado en LaTeX, no es un procesador de textos 'al uso' y cuesta adaptarse - doy fe de ello - pero es muy útil para la creación de documentos matemáticos ya que posee un editor de funciones muy potente y fácil de usar.

Permite integrar referencias bibliográficas (BibTeX con JabRef) y crear documentos con un acabado profesional.

Podemos exportar los documentos a diversos formatos como ps, pdf o dvi.

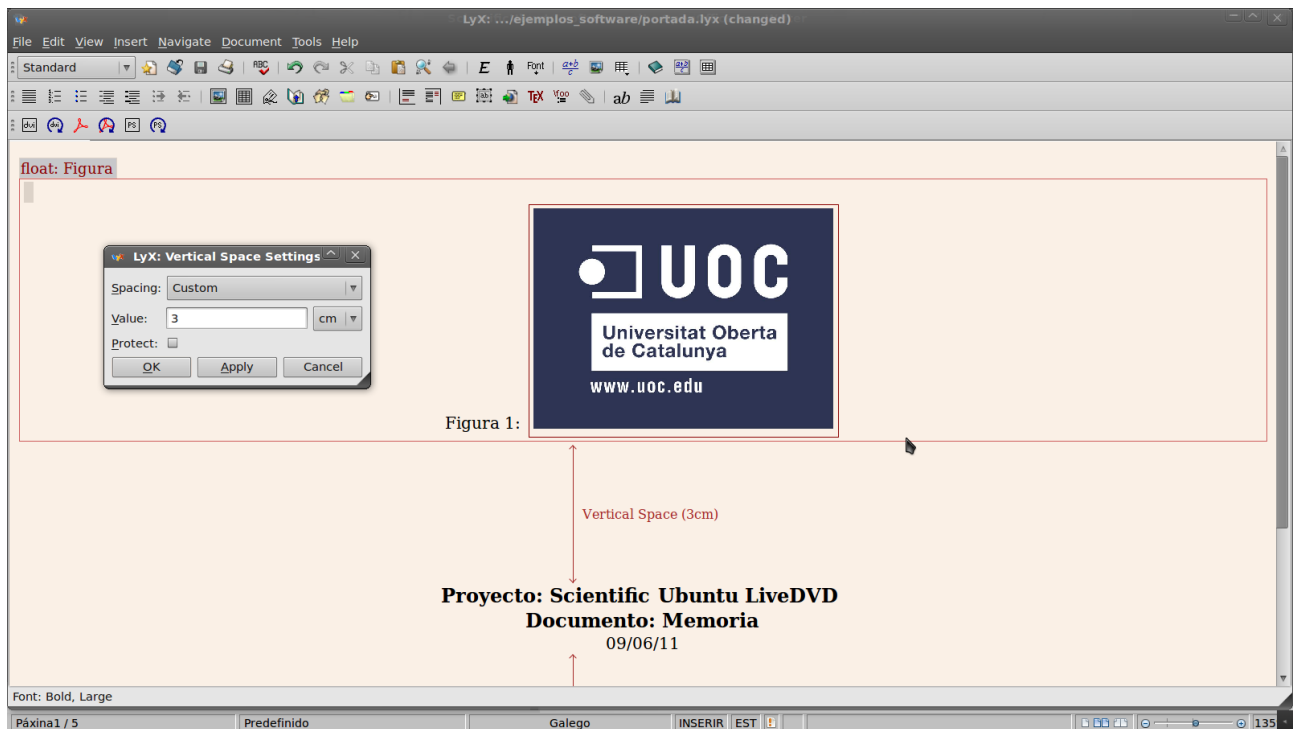


Imagen 7: LyX, espaciado vertical.

En esta imagen observamos como se establecen los espaciados verticales en LyX, que si editamos el código fuente con un simple editor de texto como gedit veremos:

```
\begin_layout Standard
\align center
\series bold
\size large
\begin_inset VSpace 3cm
\end_inset
\end_layout
```

Para hipereenlaces como los del correo electrónico, nos encontramos con el bloque:

```
\begin_layout Standard
\align_center
Alumno: Benigno Cuevas De La Fuente
\begin_inset CommandInset href
LatexCommand href
target "bcuevas@uoc.edu"
type "mailto:"
\end_inset
\end_layout
```

Y el resultado final que obtenemos es el que muestra la siguiente imagen.

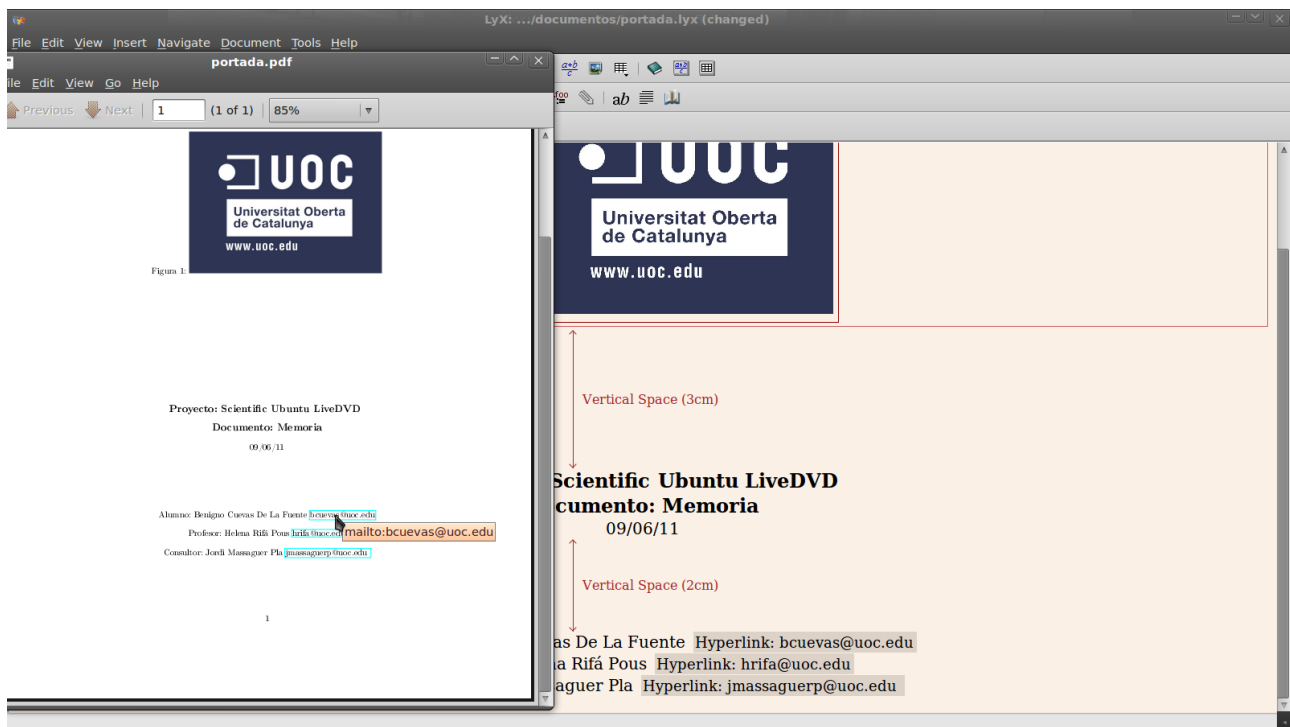


Imagen 8: LyX, documento de salida en formato pdf

El editor LyX nos evita explicitar la apariencia del documento por código, pudiendo ver en todo momento el resultado en alguno de los formatos de salida, en este caso pdf.

He de reconocer que aunque es potente y la salida que obtenemos es muy profesional, no resulta fácil adaptarse al modo de trabajar de LyX: mi primer documento me supuso casi 20 minutos de 'enfrentamiento' con la aplicación. Como muestra, para poder insertar una imagen que podamos referenciar automáticamente en un índice lo haremos en dos pasos, primero `Insert`⇒`Float`⇒`Figure` y posteriormente insertamos la imagen propiamente `Insert`⇒`Graphics`. Pero nos va a permitir modificar fácilmente el ángulo, escala e incluso la escala de visualización en el editor.

Podemos ver el código fuente del documento en un panel, pero no editarlo!

Licencia GNU GPL.

PostgreSQL

Gestor de bases de datos objeto-relacional, muy potente, con una comunidad de usuarios y desarrolladores muy activa. PostgreSQL se basa en la arquitectura Cliente/Servidor aunque lo ejecutemos en una única máquina, necesitaremos un servidor PostgreSQL y nos conectaremos como un cliente (local o remoto) a ese servidor. Existen herramientas como pgAdmin (se incluye en el LiveDVD) que facilitan la gestión de las BBDD con una interfaz gráfica. Su configuración es sencilla, el servidor está escuchando en un puerto, y el cliente debe conectarse a la BD deseada a través de ese puerto. El proceso servidor arranca al finalizar la instalación. Desde pgAdmin creamos una nueva conexión (habrá que indicarle un usuario, contraseña y nombre) para conectarnos a una BD activa o para crear una nueva.

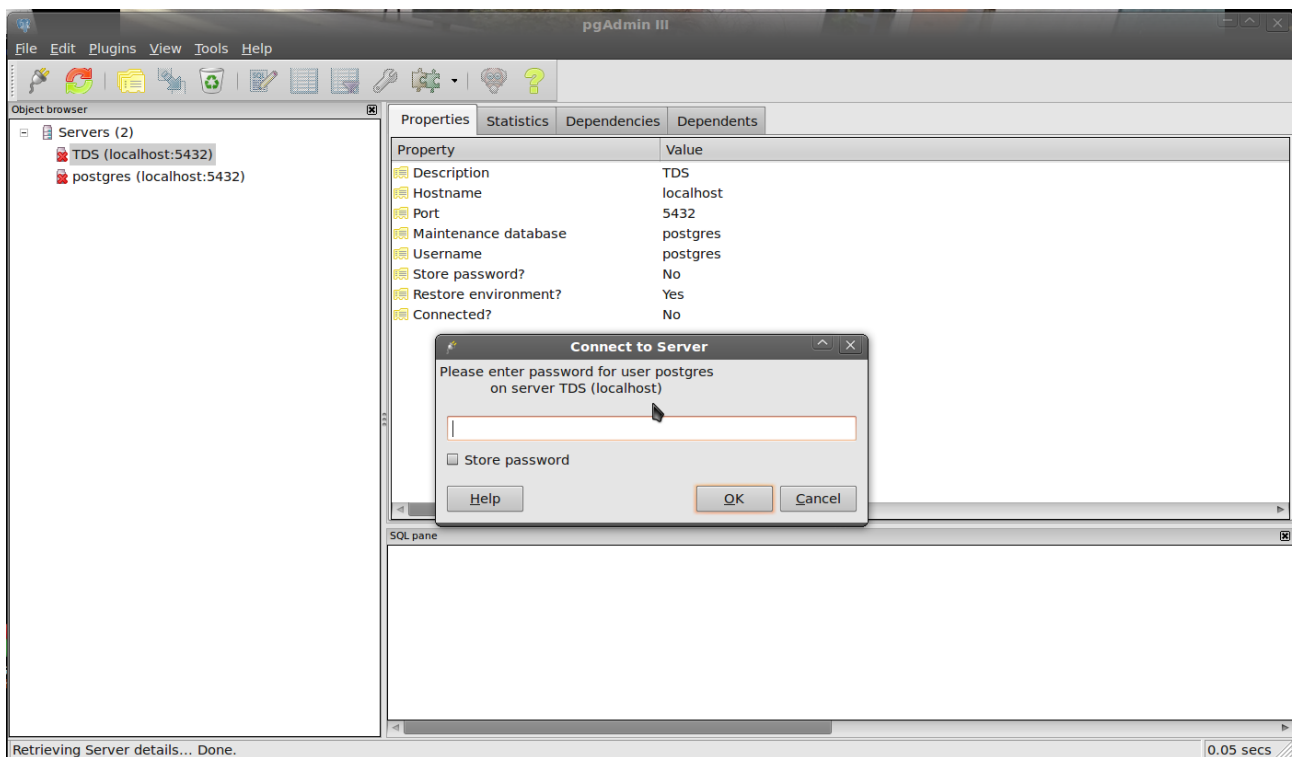


Imagen 9: pgAdmin, conexión al servidor.

Una vez nos hemos logeado, podemos realizar todas las acciones desde este entorno: crear BD, Tablas, realizar estadísticas, navegar por los objetos de la BD y ver sus propiedades, etc.

Trabajar a 'golpe de ratón' es bastante lento y tedioso, pero pgAdmin nos permite cargar o editar scripts SQL, que resultan más productivos y nos permiten un mayor control sobre los objetos creados y sus restricciones.

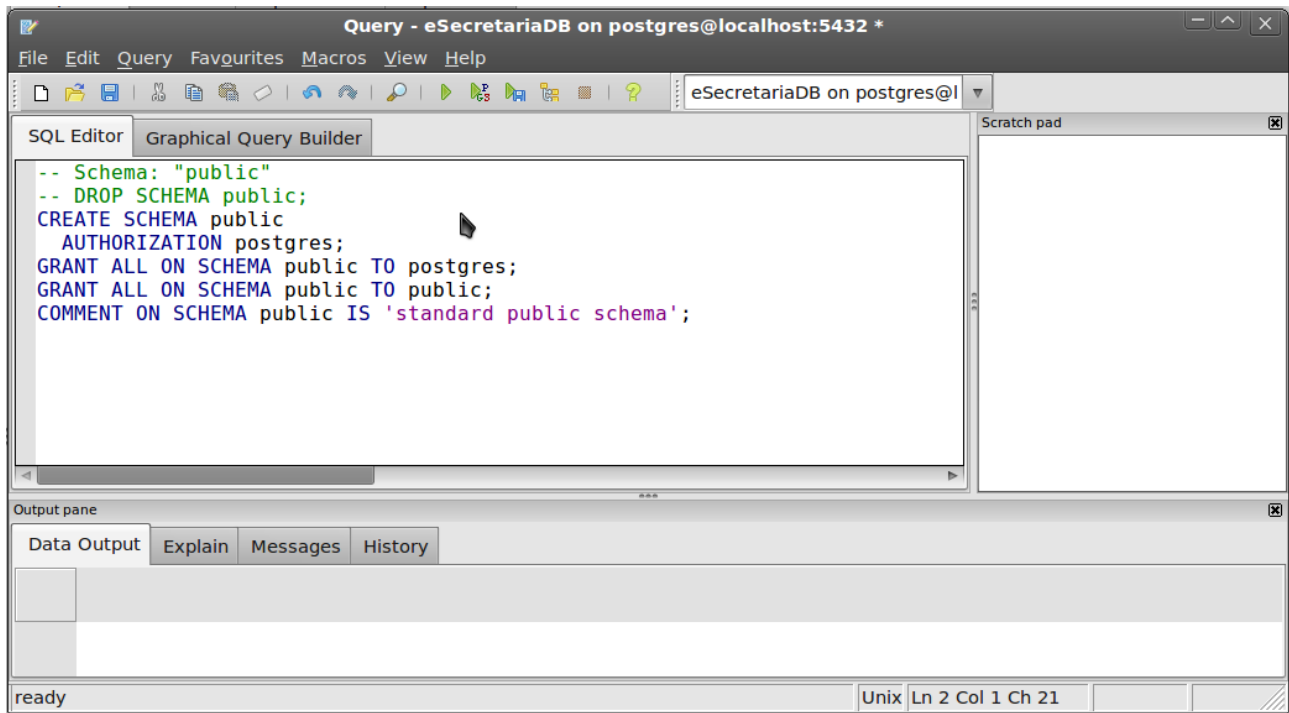


Imagen 10: pgAdmin, SQL Editor.

Una utilidad interesante es poder realizar consultas simples a las tablas (ver datos) realizar un recuento de registros (estadísticas) o navegar por los diferentes objetos de la BD para comprobar sus propiedades, ver relaciones con otros objetos y realizar incluso operaciones de borrado (Truncate), todo desde menús contextuales.

Licencia BSD.

JabRef

JabRef es un software de gestión bibliográfica que utiliza [BibTeX](#) como formato nativo. Proporciona una interfaz fácil de usar para la edición de archivos de tipo BibTeX, para la importación de datos de bases de datos científicos en línea, y para la gestión y la búsqueda de archivos BibTeX.

Veamos con un ejemplo de <http://www.citeulike.org/> como podemos importar una referencia bibliográfica.

Primero seleccionamos la referencia y la exportamos en formato BibTeX (ver las imágenes), con esto tendremos un fichero con extensión .bib que podemos abrir con JabRef e incluirlo en la base de datos de referencias.

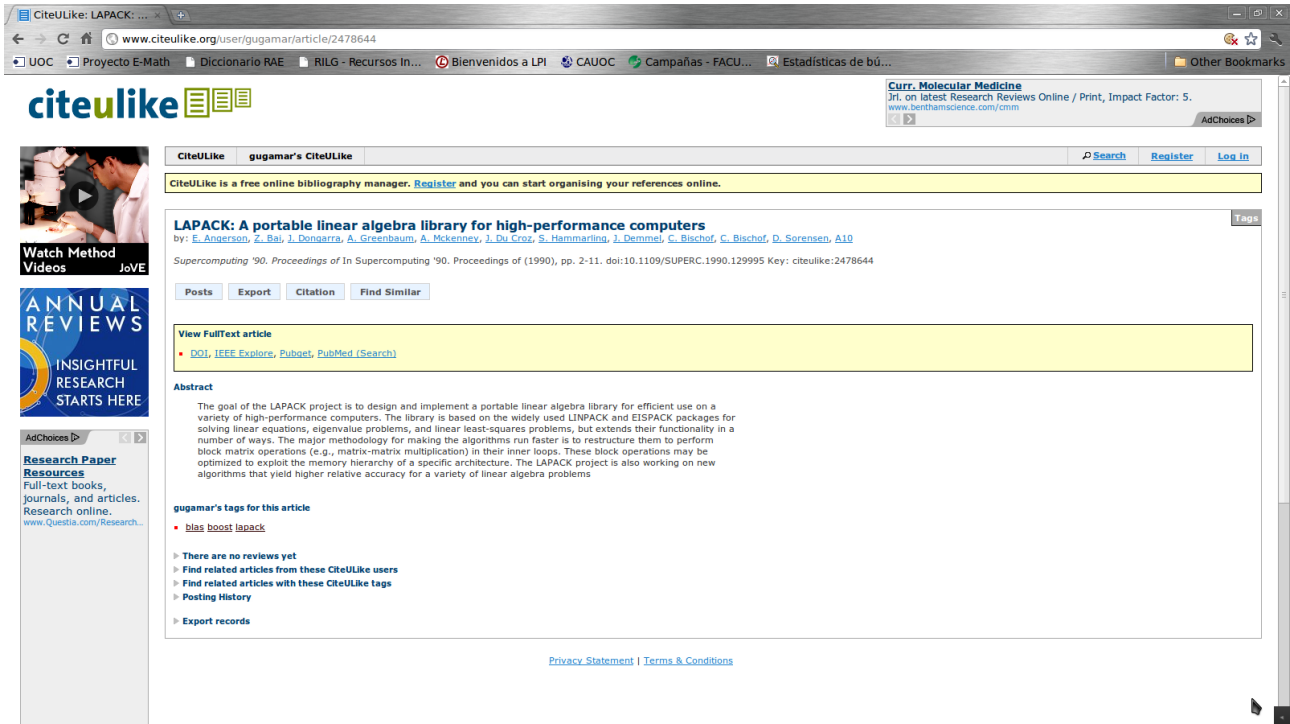


Imagen 11: Ejemplo de referencia bibliográfica en citeulike.

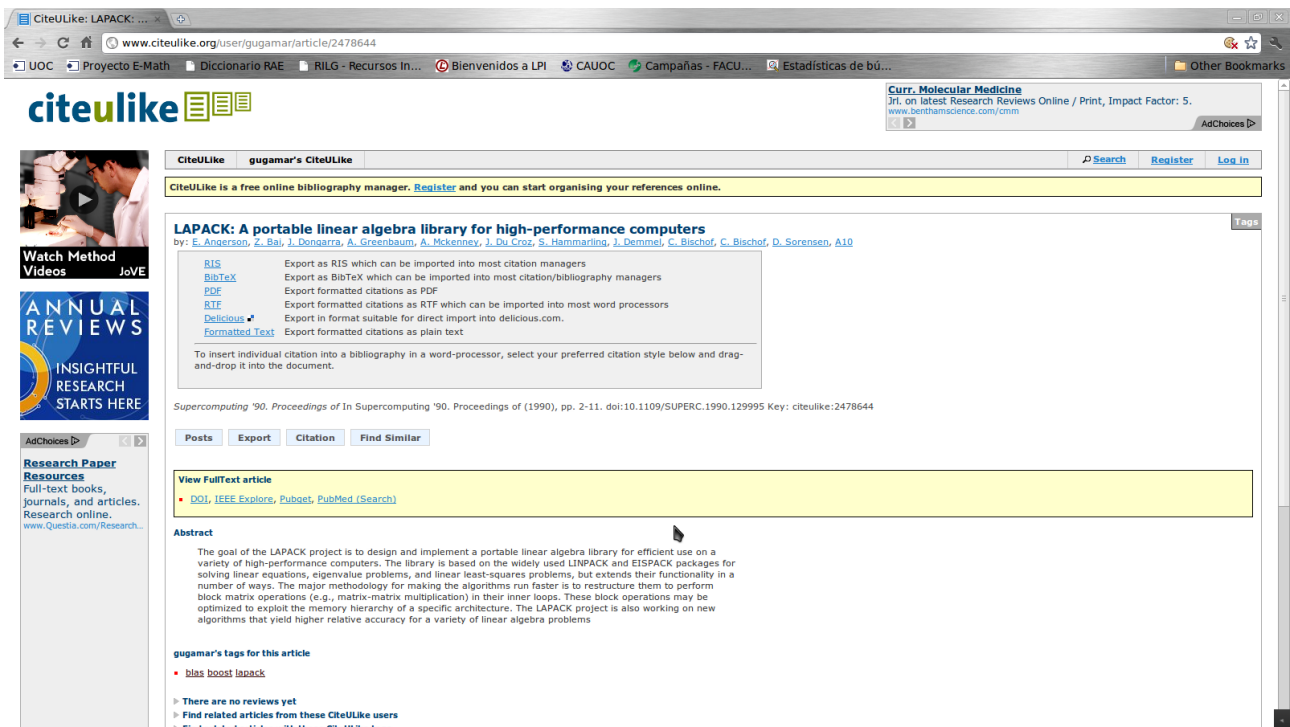


Imagen 12: Exportando referencias desde citeulike a formato BibTeX

Una vez obtenida la referencia, la podemos consultar en JabRef.

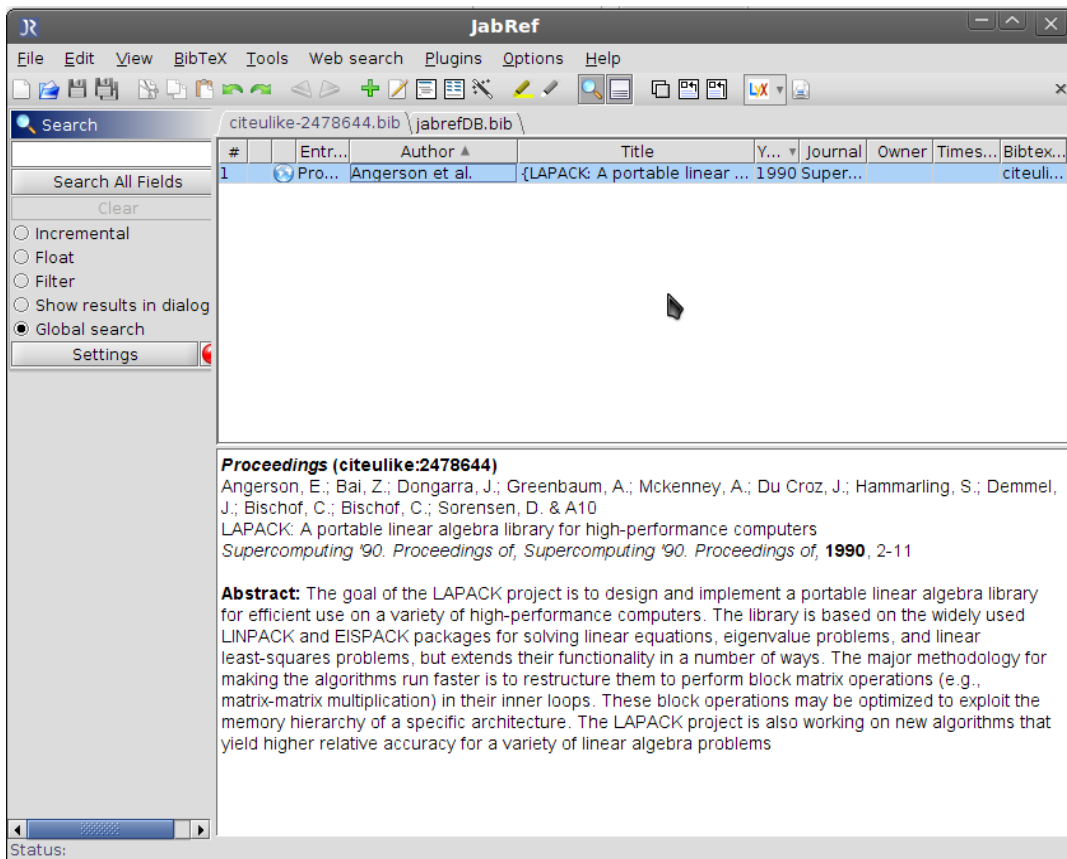


Imagen 13: Consulta de referencias en JabRef

Y ahora la podemos exportar para poder usarla con Bibus.

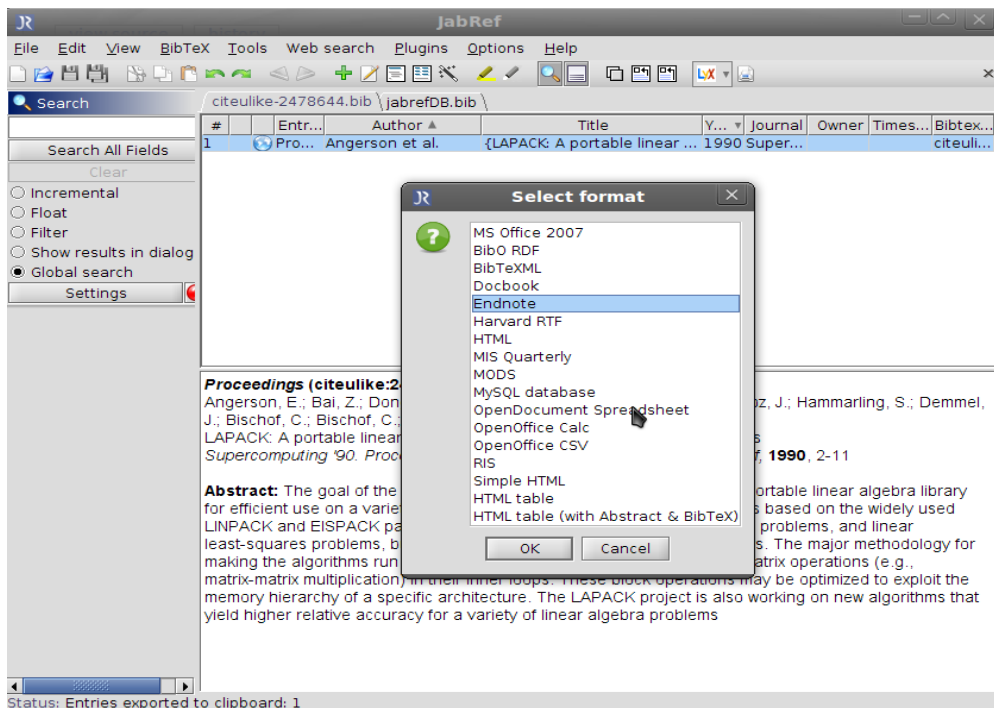


Imagen 14: Exportar referencias desde JabRef

Licencia GNU GPL.

BIBUS

Bibus es un programa para manejar referencias bibliográficas. Como otras herramientas de este tipo, Bibus nos permite buscar, editar y ordenar los datos bibliográficos. Puede integrarse con LibreOffice.org, OpenOffice.org y MS Word y permite trabajar con PubMed Entrez y eTBLAST.

La configuración es muy sencilla, en tres pasos nos preguntarán por el motor de base de datos que deseamos usar, la ubicación de nuestra BD de referencias bibliográficas y el modo de conectarnos a openoffice.org.

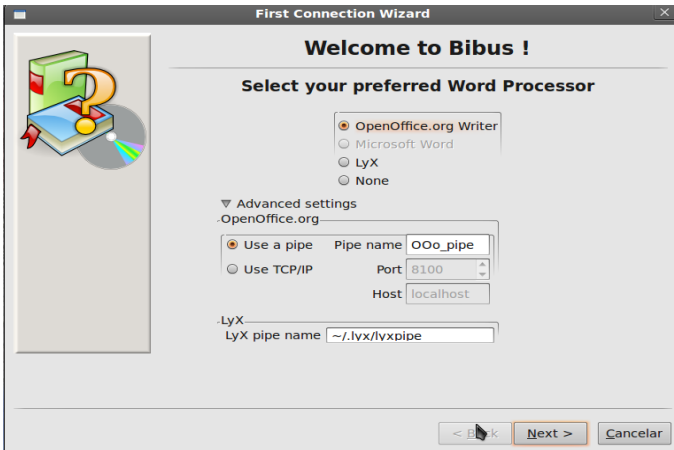


Imagen 15: Bibus, configuración del procesador de textos.

La utilidad de Bibus es que podemos literalmente pegar las referencias que tenemos en JabRef, como podemos ver en la imagen.

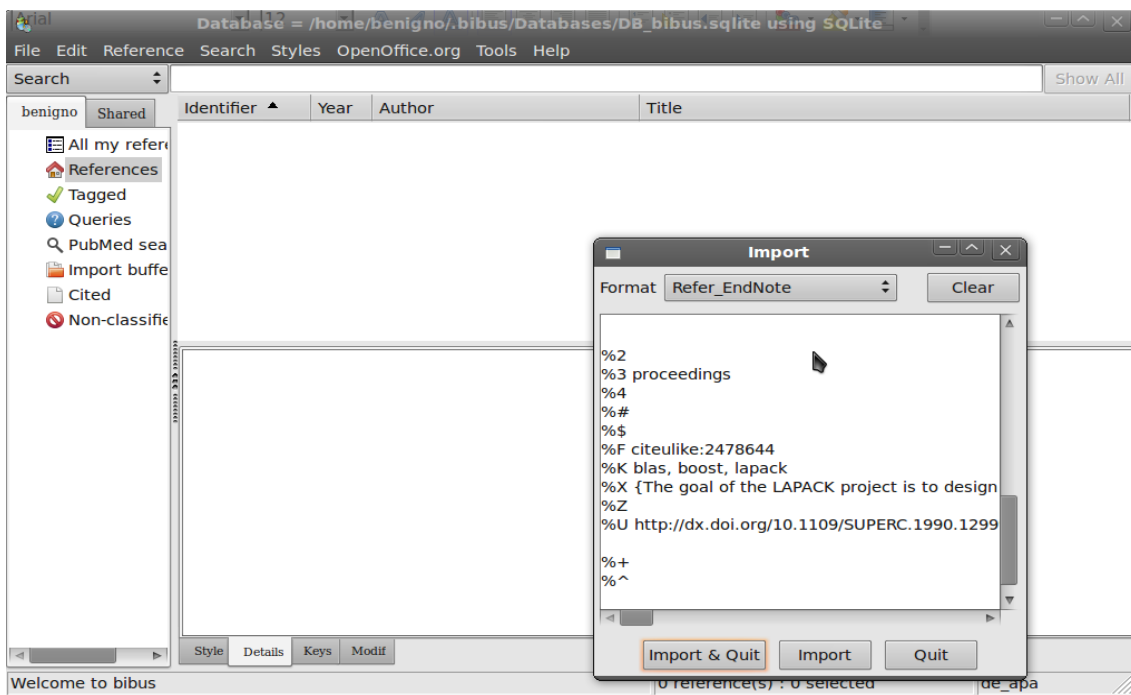


Imagen 16: Bibus, importando referencias.

Desde JabRef exportamos la referencia y con Bibus la importamos a nuestra base de datos y ya la podemos usar en openoffice writer.

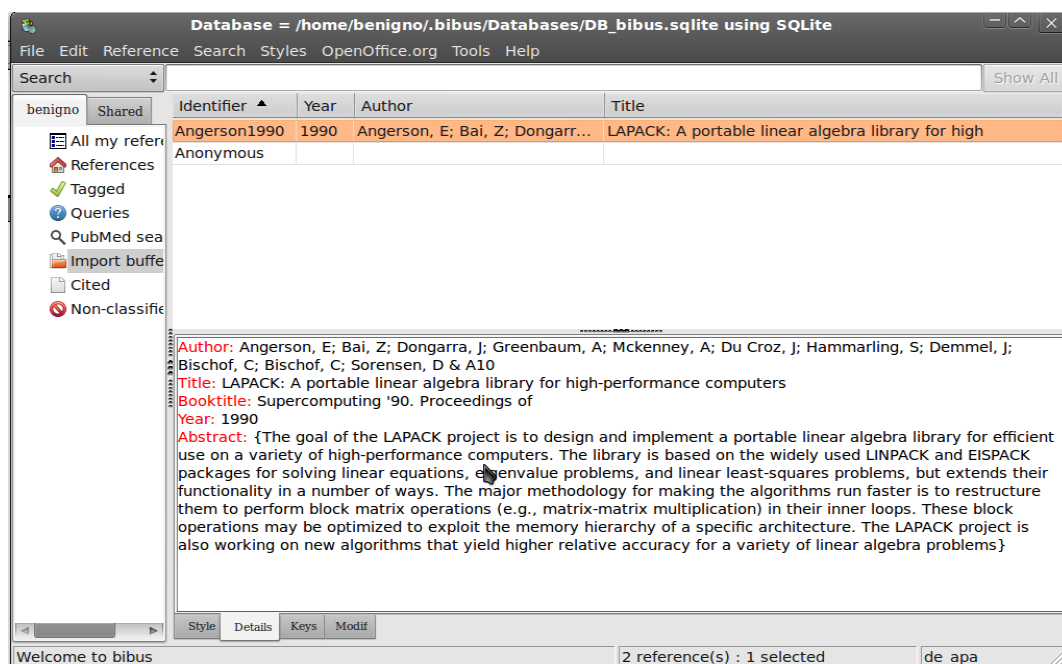


Imagen 17: Consulta de referencias en Bibus.

Licencia GNU GPL.

Con JabRef crearemos las bases de datos de referencias bibliográficas y con Bibus podremos importarlas para usarlas con el procesador de textos deseado.

Nota: las siguientes webs son habituales en la búsqueda de referencias bibliográficas: [Medline](#), [Citeseer](#), [IEEEExplore](#), [arXiv](#), [PubMed](#), [eTBLAST](#), <http://scholar.google.es/> y <http://www.citeulike.org/>.

Planner

Esta herramienta permite gestionar proyectos, creando tareas, recursos y asignado el tiempo de desarrollo. También gestiona calendarios (jornada laboral, festivos, etc). Podemos guardar los archivos en formato XML o exportarlos como pdf o HTML, también permite importarlos en formato XML desde MS Project pero este apartado no siempre ofrece los resultados esperados, al intentarlo con la planificación de este TFC no se consiguió completamente al no recuperar correctamente el encadenamiento y el orden temporal de las tareas.

La siguiente imagen muestra los primeros pasos en la planificación de este TFC, con las posibilidades de relación entre tareas.

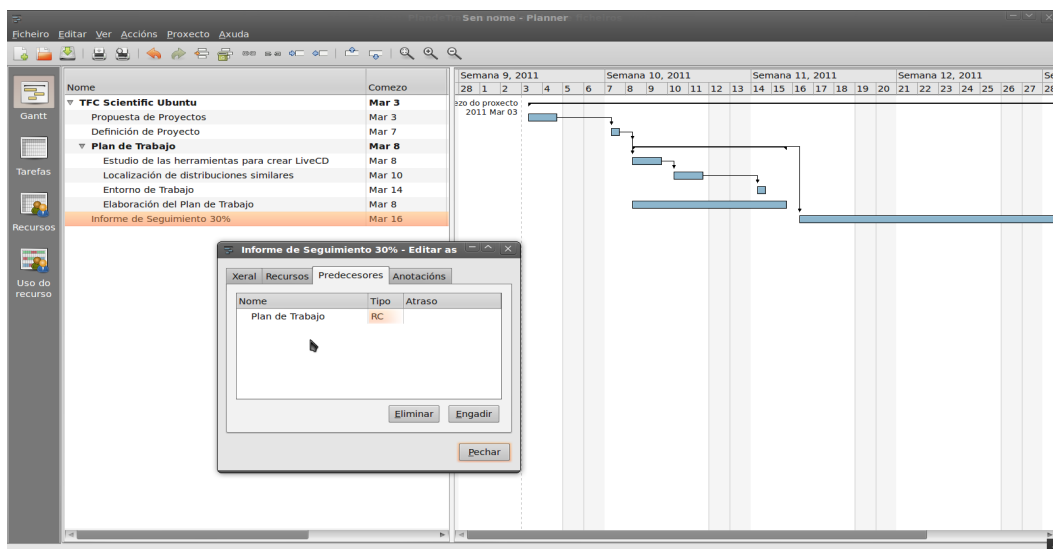


Imagen 18: Planner, ejemplo de proyecto.

Es una alternativa a herramientas como MS Project aunque con funcionalidades más limitadas.

Licencia GNU GPL.

QtiPlot

Software de análisis y visualización en 2D y 3D, también soporta datos tabulados y funciones matemáticas. Los gráficos se generan dinámicamente, si los datos de origen se modifican el gráfico se actualiza automáticamente.

Permite la generación de superficies en 3D a partir de funciones matemáticas dadas de forma explícita y de forma paramétrica.

Los gráficos obtenidos pueden ser exportados a múltiples formatos (incluyendo jpeg y pdf). Entre otras posibilidades permite realizar análisis de interpolación, FFT (Transformada rápida de Fourier), etc.

El siguiente gráfico de la 'Botella de Klein' nos sirve para mostrar las potencialidades de esta herramienta.

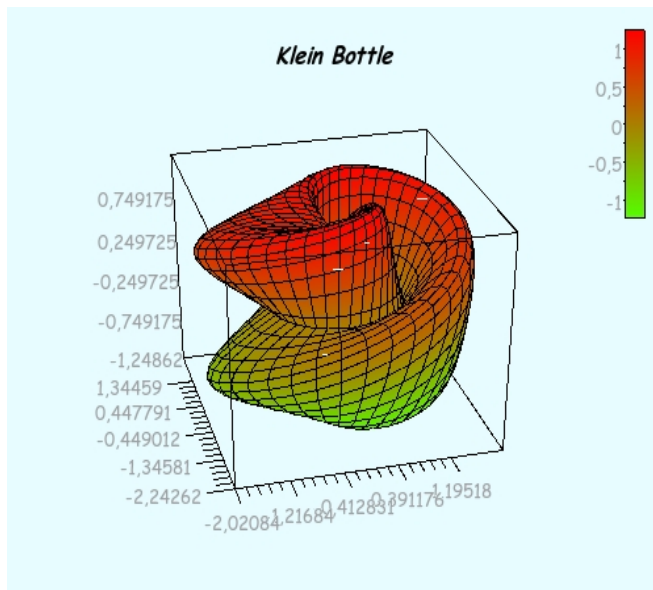


Imagen 19: QtiPlot, Botella de Klein.

$$(1 + \cos \frac{u}{2} \cdot \text{sen } v - \text{sen } \frac{u}{2} \cdot \text{sen } 2v) \cdot \cos u \quad (1 + \cos \frac{u}{2} \cdot \text{sen } v - \text{sen } \frac{u}{2} \cdot \text{sen } 2v) \cdot \text{sen } u \quad \text{sen } \frac{u}{2} \cdot \cos v + \cos \frac{u}{2} \cdot \text{sen } 2v$$

$$u \in [0, 2\pi] \quad v \in [0, 2\pi]$$

Licencia GNU GPL.

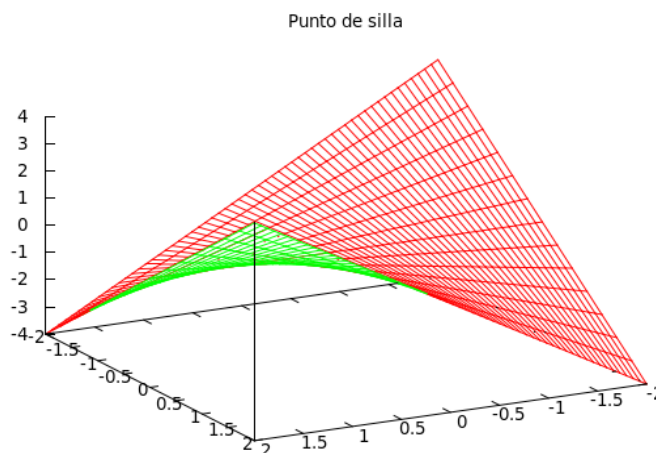
GNUPLOT

GNUplot es una utilidad de consola para generar gráficos a partir de funciones matemáticas o conjuntos de datos. También es usada como motor de gráficos para Octave. Soporta gráficos en 2D y 3D y varios formatos de salida (jpeg, png, pdf, LaTeX, etc).

Como ejemplo, el siguiente script genera la representación gráfica de un 'punto de silla':

```
set nokey
set parametric
set hidden3d
set title 'Punto de
silla'
set view 80,50

set isosamples
60,15
set xrange [-2:2]
set yrange [-2:2]
set zrange [-2:2]
set ticslevel 0
splot [-2:2][-2:2]
u, v, u*v
```



Otras licencias (Free software el código fuente).

NetBeans

Este IDE, inicialmente desarrollado por Sun Microsystems y actualmente perteneciente a Oracle, permite el desarrollo en diversos lenguajes: Java, PHP, C, C++, etc. Aunque algo pesado, es muy potente, tiene un editor visual de GUI con generación automática de código que permite un rápido prototipado, herramientas avanzadas de edición, posibilidad de autocompletar código y ejecución del proyecto dentro del propio IDE. La inclusión de librerías es sencilla, y además se muestran en el árbol de ficheros del proyecto. Permite la eliminación segura de clases (búsqueda dentro del proyecto de todos los usos), y edición avanzada: sustituir o reemplazar las apariciones de un objeto, renombrar de forma segura un atributo, etc. La inclusión y organización de paquetes e interfaces es sencilla y el acceso a las funciones que proveen es asistido por el editor para evitar errores tipográficos.

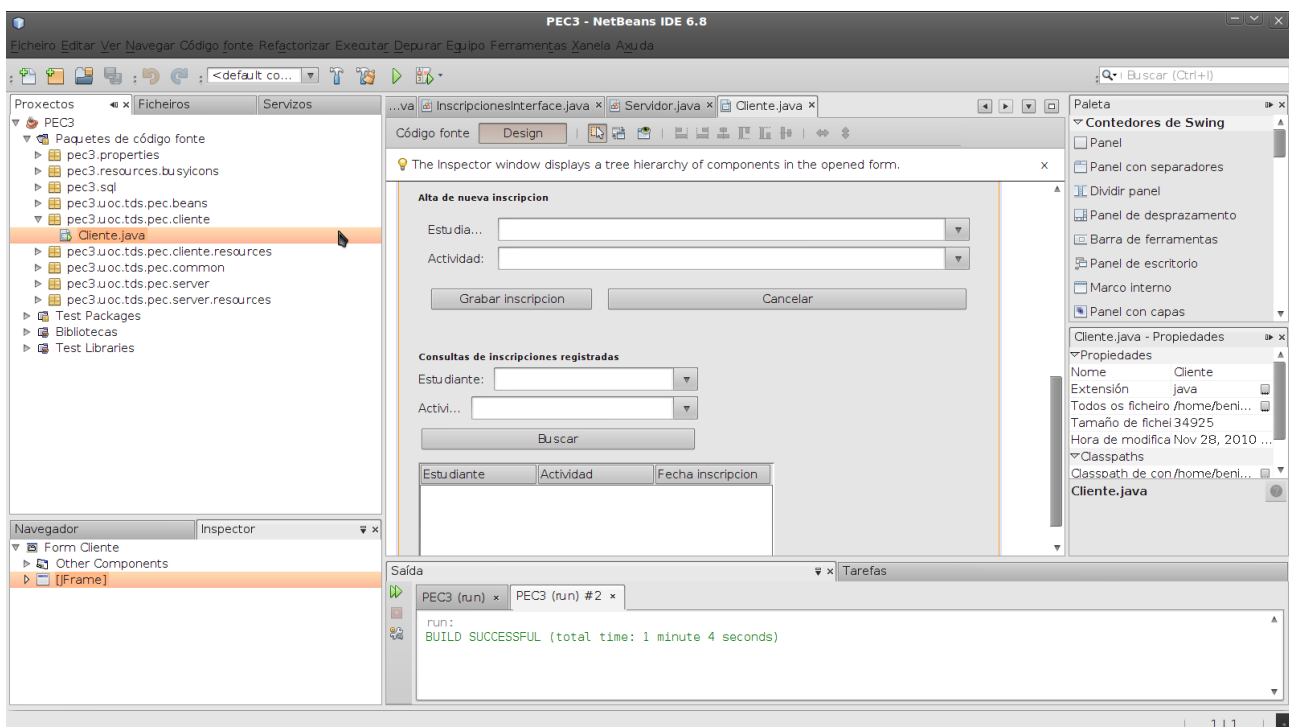


Imagen 20: NetBeans, ejemplo de proyecto (tomado de la asignatura TDS).

Licencia [CDDL](#).

R y R Commander

R es un software de cálculo estadístico (creado por Ross Ihaka y Robert Gentleman) que permite realizar múltiples análisis de los datos, pero también es un lenguaje de programación orientado a objetos que nos permite implementar nuestros propios procedimientos para análisis de datos y generación de gráficos. Además de los cálculos estadísticos básicos (medidas de dispersión, medidas de centralización, histogramas y curvas) también podemos realizar análisis de regresión (modelos lineales y no lineales) y el análisis de varianza (ANOVA).

Algunos ejemplos del paquete gráfico 'lattice':

```
> library(lattice)
> demo(lattice)
```

```
Terminal
Eicheiro Editar Ver Terminal Ayuda
> densityplot(-x)
Hit <Return> to see next plot:
> histogram(x, type = "density",
+           panel = function(x, ...) {
+             panel.histogram(x, ...)
+             panel.densityplot(x, col = "brown", plot.points = FALSE)
+           })
Hit <Return> to see next plot:
> ## Using a custom panel function to superpose a fitted normal density
> ## on a Kernel Density Estimate
> densityplot(~ height | voice.part, data = singer, layout = c(2, 4),
+           xlab = "Height (inches)",
+           ylab = "Kernel Density\n with Normal Fit",
+           main = list("Estimated Density", cex = 1.4, col = "DarkOliveGreen"),
+           panel = function(x, ...) {
+             panel.densityplot(x, ...)
+             panel.mathdensity(dmath = dnorm,
+                               args = list(mean=mean(x),sd=sd(x)))
+           })
Hit <Return> to see next plot: █
```

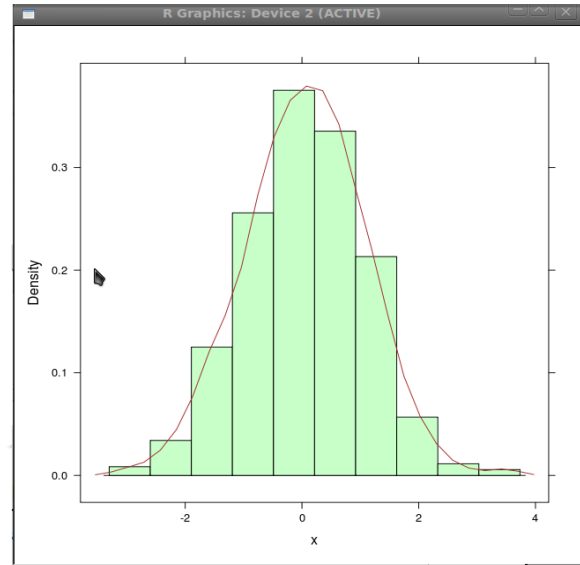


Imagen 21: R, histograma y polígono de frecuencias del paquete lattice.

Con el front-end R Commander, podemos emplear Data Sets y realizar análisis de las variables, histogramas, cálculo de medidas de centralización y dispersión, etc. desde una interfaz mas amigable. También podemos hacer análisis de probabilidad basándonos en distribuciones conocidas o generar gráficos de las mismas.

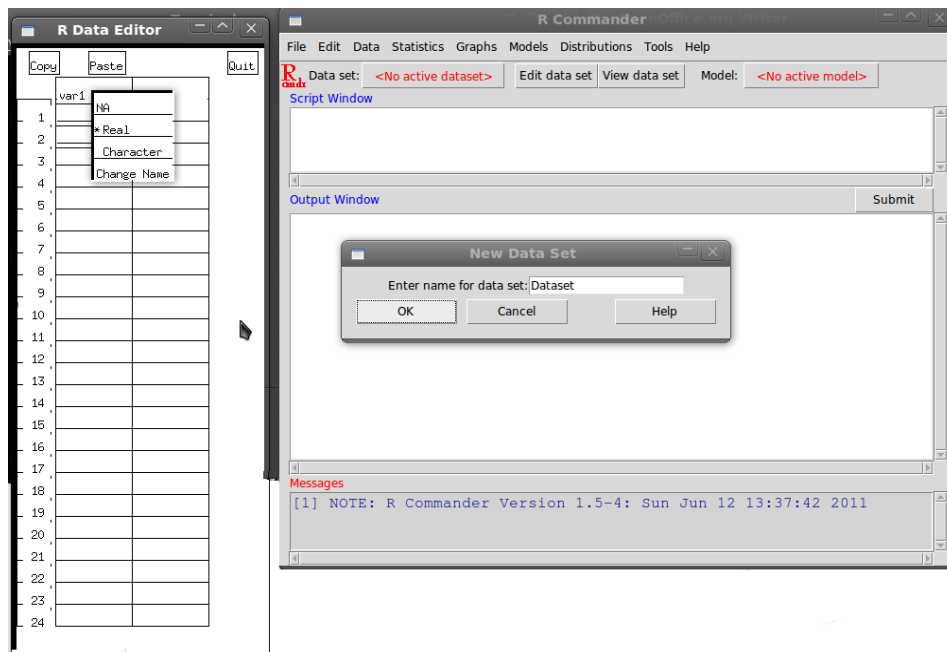


Imagen 22: R Commander, creación de Data sets.

Es muy flexible y existen multitud de [librerías disponibles](#) para ampliar sus funcionalidades (según la web, 3049).

Licencia GNU GPL.

6 Conclusiones

El objetivo de este Trabajo Fin de Carrera es obtener un producto software con herramientas básicas para el trabajo en ámbitos científicos y académicos, el LiveDVD entregado aporta una posible aproximación usando herramientas de software libre, de las cuales he aportado algunos ejemplos para justificar su inclusión. He intentado llegar a un consenso de mínimos en cuanto a las aplicaciones disponibles, evitando abarcar demasiado y teniendo en mente las necesidades comunes de un estudiante/investigador. Para esta cuestión nada trivial he consultado con personas que por su formación y/o experiencia en el ámbito científico me han ayudado a delimitar el alcance de este trabajo.

El desarrollo del LiveDVD me ha permitido descubrir un importante número de aplicaciones Open Source disponibles para la investigación que en muchos casos son igual de adecuadas que alternativas de código propietario más costosas. Se puede hacer ciencia sin salir del universo GNU, se pueden publicar los resultados con calidad sin necesidad de costosos programas de edición dependientes de una plataforma software concreta. Que instituciones como el CERN apuesten por plataformas Linux basadas en Open Source muestra el potencial de Linux en el ámbito científico.

En cuanto a la metodología seguida, a pesar de existir herramientas muy tentadoras como remastersys o incluso UCK que me hubiesen permitido realizar un producto de contenido casi exacto, no son comparables a la experiencia de controlar todo el proceso desde una consola y poder usar herramientas menos habituales hoy en día como el editor vi.

En este proyecto no se programa casi nada, salvo en algún script, algo casi 'contra natura' en un informático, pero se bucea en las interioridades de un sistema Linux basado en Ubuntu para su adaptación a los objetivos descritos.

7 Recursos.

Referencias usadas en el desarrollo de este TFC.

7.1 Genéricas

[InstallCDCustomization - Community Ubuntu Documentation](#)

[Apt - doc.ubuntu-es](#)

[GRUB - Guía Ubuntu](#)

Linux man pages

Linux Professional Institute Certification. Guía de estudio LPIC-1 (Ed. Anaya)

Ubuntu Powerful Hacks and Customizations, Dr. Neal Krawetz (Ed. Wiley)

Cálculo vectorial e ecuaciones diferenciais, M. Besada Moráis, F.J. García Cutrín y C. Vázquez Pampín (Servicio de publicaciones Universidade de Vigo)

7.2 Referencias de otras distribuciones.

[Scientific Linux - Welcome to Scientific Linux \(SL\)](#)

[UbuntuScience - Community Ubuntu Documentation](#)

[SIE de Investigación y Enseñanza - Software list](#)

<http://www.es.poseidonlinux.org/>

[GetUbuntuSci - ubuntu\[sci\]Wiki](#)

[PPA for scubuntu team : "Ubuntu\[sci\] team" team](#)

7.3 Customization

[LiveCDCustomization - Community Ubuntu Documentation](#)

[Personalizar distribución de Ubuntu Live CD - Wikilibros](#)

[UCK - Ubuntu Customization Kit](#)

[LiveCDCustomizationFromScratch - Community Ubuntu Documentation](#)

[LFS LiveCD Documentation](#)

[LiveCDCreator - Ubuntu Wiki](#)

[Remastersys Homepage](#)

7.4 Herramientas.

[Linux_Downloads - VirtualBox](#)

[Dropbox](#)

Repositorios de Ubuntu.

8 Índice de imágenes.

Imagen 1: Scientific Linux Desktop.....	6
Imagen 2: Scientific Linux, repositorios adicionales.....	7
Imagen 3: Poseidon Linux Desktop.....	11
Imagen 4: Escritorio de Ubuntu sci-Math.....	23
Imagen 5: Oracle Virtual Box, Pantalla principal.....	27
Imagen 6: LyX, espaciado vertical.....	33
Imagen 7: LyX, documento de salida en formato pdf.....	34
Imagen 8: pgAdmin, conexión al servidor.....	35
Imagen 9: pgAdmin, creación de una nueva tabla.....	36
Imagen 10: pgAdmin, SQL Editor.....	36
Imagen 11: Ejemplo de referencia bibliográfica en citeulike.....	37
Imagen 12: Exportando referenciasde citeulike a formato BibTeX.....	38
Imagen 13: Consulta de referencias en JabRef.....	38
Imagen 14: Exportar referencias desde JabRef.....	39
Imagen 15: Bibus, configuración del procesador de textos.....	39
Imagen 16: Bibus, importando referencias desde JabRef.....	40
Imagen 17: Consulta de referencias en Bibus.....	40
Imagen 18: Planner, ejemplo de proyecto.....	41
Imagen 19: QtiPlot, Botella de Klein.....	42
Imagen 20: NetBeans, ejemplo de proyecto (tomado de la asignatura TDS).....	43
Imagen 21: R, histograma y polígono de frecuencias del paquete lattice.....	44
Imagen 22: Boot Priority en Award BIOS.....	48
Imagen 23: Escritorio del LiveDVD Scientific Ubuntu.....	49

9 Anexos

9.1 Planificación.

Id	Nombre de tarea	Duración	Comienzo	Fin	Predecesora
1	TFC Sicientific Ubuntu LiveCD	82 días	jue 03/03/11	vie 24/06/11	
2	Propuestas de Proyectos	2 días	jue 03/03/11	lun 07/03/11	
3	Definición del Proyecto	1 día	mar 08/03/11	mié 09/03/11	2
4	Plan de Trabajo	13,13 días	jue 10/03/11	lun 28/03/11	3
5	Estudio de las herramientas para creación de LiveCD	2 días	jue 10/03/11	sáb 12/03/11	
6	Localización de distribuciones similares	2 días	lun 14/03/11	mié 16/03/11	5
7	Configuración del entorno de trabajo	1 día	jue 17/03/11	vie 18/03/11	6
8	Elaboración del Plan de Trabajo	6 días	lun 21/03/11	lun 28/03/11	
9	Informe de seguimiento 30%	13 días	mar 29/03/11	vie 15/04/11	4
10	Selección de paquetes a incluir en la distribución	5 días	mar 29/03/11	mar 05/04/11	
11	Documentación	4 días	mié 06/04/11	mar 12/04/11	10
12	Compilación versión preliminar LiveCD	1 día	mar 12/04/11	mié 13/04/11	11
13	Testing versión preliminar del LiveCD	2 días	mié 13/04/11	vie 15/04/11	12
14	Informe de seguimiento 70%	26 días	lun 18/04/11	lun 23/05/11	9
15	Solución de errores en la versión Preliminar del LiveCD	6 días	lun 18/04/11	lun 25/04/11	13
16	Personalización de la distribución -temas, navegador web,arranque-	3 días	mar 26/04/11	jue 28/04/11	15
17	Compilación del LiveCD Sicientific Ubuntu LiveCD	4 días	mar 26/04/11	vie 29/04/11	15
18	Documentación y Borrador de la memoria del proyecto	14 días	mié 04/05/11	lun 23/05/11	16
19	Memoria y LiveCD final	14,88 días	mar 24/05/11	lun 13/06/11	14
20	Presentación TFC	4 días	lun 13/06/11	vie 17/06/11	19FC-1 día
21	Tribunal de evaluación	4 días	lun 20/06/11	vie 24/06/11	20

9.2 Requerimientos de hardware y software.

Ubuntu requiere para funcionar al menos 256MB de memoria, y un procesador de al menos 1 Ghz, si deseamos instalar el sistema necesitaremos al menos 5 Gb para el sistema base y cerca de 1GB extra para las aplicaciones añadidas de carácter científico, nótese que en la máquina de prueba hemos dedicado 756 MB de memoria, 128 MB para la memoria de vídeo y 8 GB de espacio en disco duro para disponer de unas condiciones óptimas de ejecución en caso de desear instalar la distribución en nuestro Pc.

La BIOS del Pc debe estar configurada para arrancar desde la unidad de DVD, si no fuese el caso consulte el manual de su placa base para modificar la prioridad de los dispositivos en el arranque de su equipo, generalmente la tecla **Supr** permite acceder a la utilidad de configuración de la BIOS.

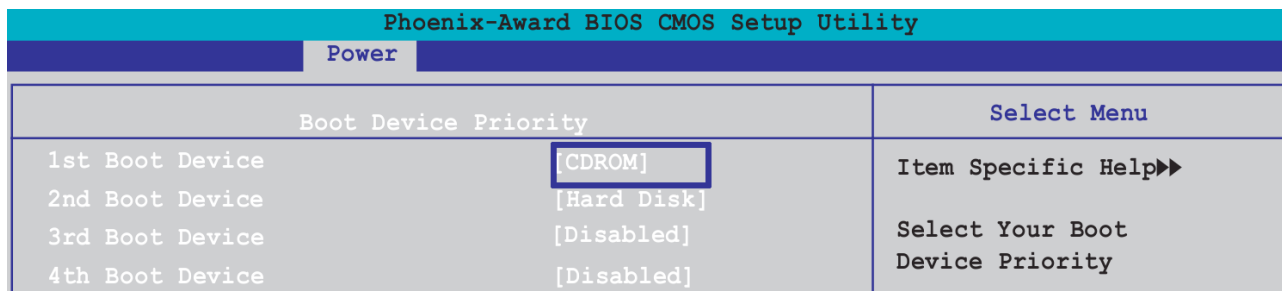


Imagen 23: Boot Priority en Award BIOS

9.3 Instrucciones de instalación.

La documentación y el LiveDVD final que se entrega como producto de software están disponibles en un espacio de almacenamiento gratuito ofrecido por [DropBox](#). Se han dispuesto dos carpetas compartidas y se cursará la invitación correspondiente para permitir su acceso al consultor de la asignatura, en este caso D. Jordi Massaguer Pla (jmassaguerp@uoc.edu).

Si así lo considera el consultor o el tribunal de evaluación, se crearán carpetas compartidas con el contenido que puedan solicitar y se cursaran las invitaciones correspondientes por correo electrónico.

Una vez descargada la imagen iso del LiveDVD, podremos usar una herramienta de grabación, por ejemplo [k3b](#), para poder disponer de un DVD y realizar un testing del producto software entregado.

Otra alternativa es crear una máquina virtual como la descrita en la sección 4 de este documento y ejecutar directamente la imagen iso (es necesario configurar la máquina virtual a tal efecto, ver instrucciones en [VirtualBox Manual](#)).

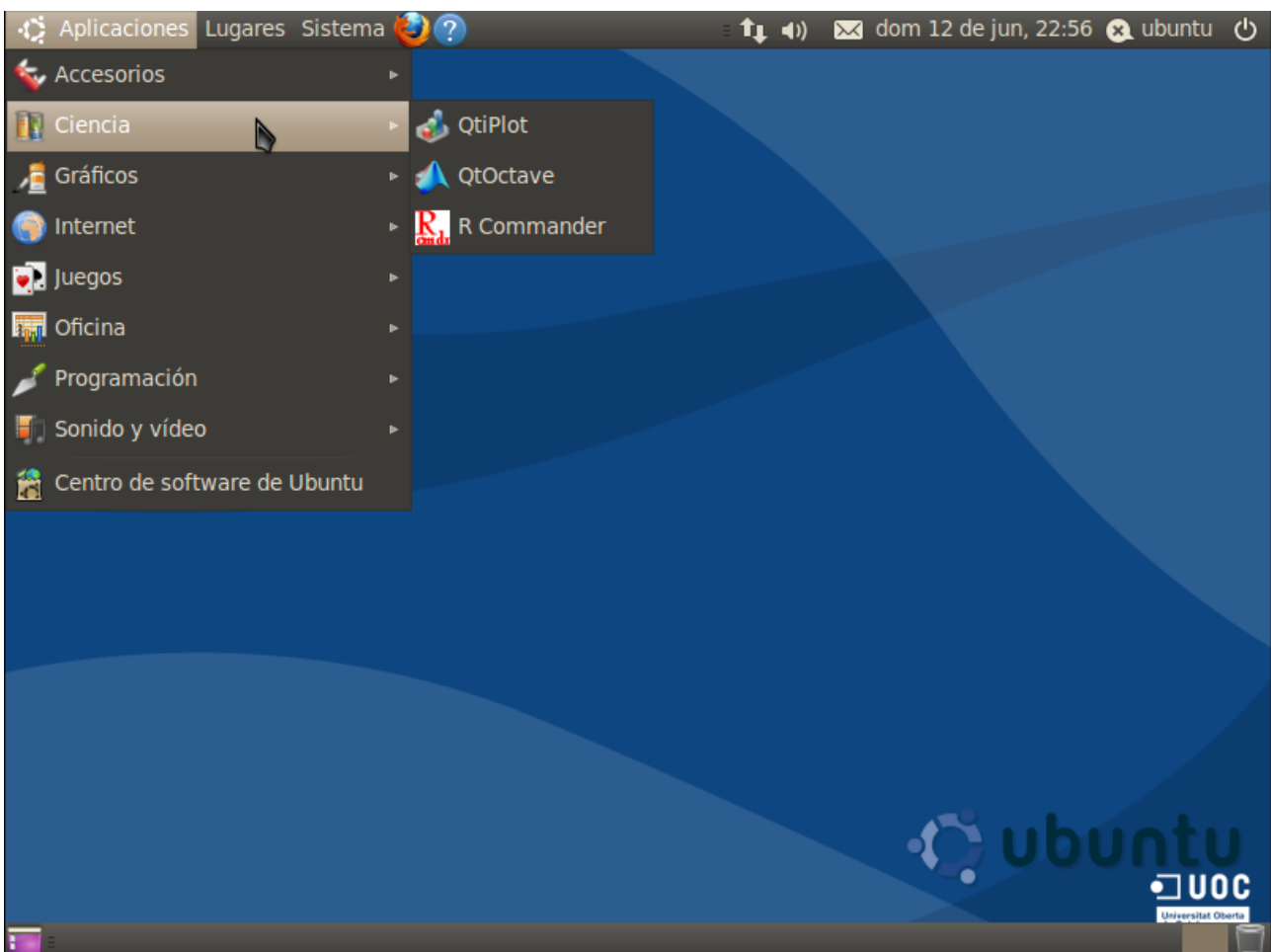


Imagen 24: Scientific Ubuntu, escritorio.

9.4 Paquetes incluidos.

octave3.2 qtoctave netbeans lyx postgresql postgresql-client
pgadmin3 netbeans gnuplot qtiplot r-base r-base-core r-base-latex
r-recommended r-doc-pdf r-doc-html r-mathlib r-cran-rcmdr jabref
bibus planner gfortran

9.5 Glosario de términos.

[UCK](#) (Ubuntu Customization Kit) es una herramienta pensada para modificar los paquetes instalados en una imagen iso de ubuntu.

[Oracle VirtualBox](#) es una herramienta multiplataforma para virtualizar en una máquina anfitrión otros sistemas operativos con la única limitación del espacio en disco y la memoria necesarias.

[Squashfs](#) es un sistema de ficheros de sólo lectura que permite comprimir archivos, directorios e inodes y se usa para construir los LiveCD.

[ATLAS](#) (Automatically Tuned Linear Algebra Software) es un conjunto de interfaces para las librerías BLAS y Lapack.

GRID. Computación distribuida o como usar equipos diversos para la resolución de una tarea.

[OpenAFS](#) es la versión abierta de AFS (Andrew File System), un sistema de archivos distribuidos en una arquitectura Cliente-Servidor que es 'transparente' en el lado del cliente.

[Infiniband](#) es una arquitectura de transmisión de información entre dispositivos que provee velocidades de hasta 120GB/s.

GIS. Software que emplea una base de datos con información geográfica (datos alfanuméricos) que se encuentra asociada por un identificador común a los objetos gráficos de un mapa digital. De esta forma, señalando un objeto se conocen sus atributos e, inversamente, preguntando por un registro de la base de datos se puede saber su localización en la cartografía.

Entrez. Sistema de búsqueda de información en bases de datos de genes, cromosomas cancerígenos, genoma de diversos organismos, publicaciones en PubMed, etc.

[CERN](#) es el centro europeo de investigación nuclear, muy conocido actualmente por el acelerador de partículas Large Hadron Collider (y también dónde trabajó Tim Berners-Lee).

IDE. Un entorno integrado de desarrollo, comprende las herramientas avanzadas de un potente editor de textos con herramientas asociadas al desarrollo de software (compiler, debugger, generador de código y diseño de GUI).

Cebador o primer. Cadena de ácido nucleico o de una molécula relacionada que sirve como punto de partida para la replicación del ADN.

LTS. Ubuntu tiene un calendario de lanzamientos predefinido, cada 6 meses lanza una versión nueva y cada 2 años una versión LTS, que tienen soporte durante más tiempo.

9.6 Agradecimientos.

Las siguientes personas han colaborado con indicaciones muy valiosas para delimitar el ámbito de este Trabajo Fin de Carrera, aportando algunos ejemplos de su experiencia académica o profesional.

Noelia Rodríguez Casal, Licenciada en Matemáticas ([Universidad de Santiago de Compostela](#)).

Julia Villanueva Alonso, Doctora en Química ([Universidad de Santiago de Compostela](#)).

Luisa Blanco Fernández (Master en Astrofísica, [Universidad de la Laguna - IAC](#)).

Mauricio Refojo, Licenciado en Matemáticas ([Universidad de Santiago de Compostela](#), actualmente colaborador en el [Trinity College](#) de Dublin).

Raquel Sanclaudio (Licenciada en Físicas, [Universidad de Santiago de Compostela](#)).

Este trabajo es un punto y seguido en la relación de este alumno con la Universitat Oberta de Catalunya, sin cuyos tutores, consultores y profesores no hubiese sido posible llegar hasta aquí.