

Máster Universitario en Bioinformática

Generación y mantenimiento de datos ómicos

Curso académico 2024-25



Universidad
Internacional
de Valencia

Dr. Jordi Tronchoni
jordi.tronchoni@professor.universidadviu.com

22/04/2024

De:
 Planeta Formación y Universidades

Tema 3

Gestión de entornos y paquetes

22/04/2024

Tema 1. Introducción a la bioinformática

- 1.1 Historia de la bioinformática
- 1.2 Bioética aplicada al análisis de datos

Tema 2. Principales flujos de trabajo en bioinformática

- 2.1 Genómica
- 2.2 Metagenómica y metataxonómica
- 2.3 Transcriptómica
- 2.4 Proteómica

**Tema 3. Gestión de entornos y paquetes**

- 3.1 Conda

Tema 4. Bases de datos y herramientas bioinformáticas

- 4.1 Principales bases de datos
- 4.2 Otros recursos online

Tema 5. Alineamiento de secuencias

- 5.1 Introducción al alineamiento de secuencias
- 5.2 Alineamientos Pairwise
- 5.3 Alineamientos Múltiples

Tema 6. Métodos de secuenciación

- 6.1 Primera generación de secuenciadores
- 6.2 Segunda generación de secuenciadores
- 6.3 Tercera generación de secuenciadores
- 6.4 Comparación de plataformas de secuenciación

Tema 7. Pre-procesado y calidad de secuencias

- 7.1 Calidad de secuencias
- 7.2 Pre-procesado de secuencias

Herramientas



22/04/2024

Las actividades prácticas y workshops están testados en WS de AWS y será el ecosistema donde debes realizarlas.

Herramientas



Gestor de paquetes

Instalar
Ejecutar
Actualizar
(tanto paquetes como
sus dependencias)

Gestor de entornos

Crear
Guardar
Cargar
Moverse entre entornos

Originalmente creado para resolver el problema de la dependencia de distintas
herramientas de distinta versiones de python

22/04/2024

Conda

Conda como **gestor de paquetes** te ayuda a encontrar e instalar paquetes.

Si necesitas un paquete que requiere una versión diferente de Python, no necesitas cambiar a un gestor de entorno diferente, porque Conda también es un **gestor de entorno**.

Con sólo unos pocos comandos, puedes configurar un entorno totalmente separado para ejecutar esa versión diferente de Python, mientras sigues ejecutando tu versión habitual de Python en tu entorno normal.

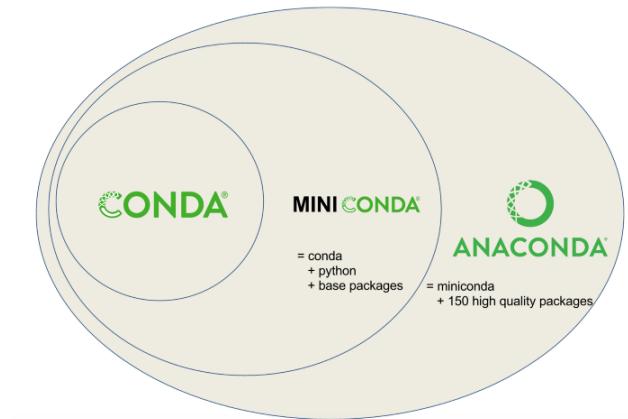
22/04/2024

¿Qué es CONDA?

Conda es un gestor de paquetes y entornos. Es la forma más fácil de manejar la instalación de la mayoría de las herramientas que queremos usar en bioinformática. Que algo pueda instalarse en Conda, requiere que alguien (el desarrollador u otros) se haya tomado la molestia de hacerlo de esa manera, por lo que no todo está disponible (cuanto más actual más difícil será encontrarlo), pero casi todo lo que probablemente queramos usar está. El otro valor de Conda es que facilita la instalación de estas herramientas y maneja muy bien los diferentes entornos.

A veces el Programa A dependerá de una versión específica del Programa B. Pero entonces, el Programa C dependerá de una versión diferente del Programa B, y esto causa problemas. Conda nos permite crear y gestionar fácilmente entornos separados para evitar este tipo de conflictos de versiones, y comprueba automáticamente por nosotros cuando tratamos de instalar algo nuevo. Además, mejora y facilita la reproducibilidad de proyectos.

Conda



22/04/2024

Obtener e instalar Conda

Conda se puede obtener mayoritariamente de dos formas: como Anaconda y como Miniconda. Anaconda es grande y con muchos programas dentro, Miniconda es más ligero, y a partir de ahí, instalas lo necesario.

Miniconda, combina Conda con Python 3 (y un pequeño número de paquetes de sistemas básicos) en lugar de la distribución completa de Anaconda. Instalar sólo Miniconda anima a crear entornos separados para cada proyecto (y a instalar sólo aquellos paquetes realmente necesarios para cada proyecto) mejorando la portabilidad y la reproducibilidad de la investigación y flujos de trabajo.

Miniconda: ventajas

Conda proporciona paquetes o binarios pre-compilados (lo que generalmente evita la necesidad de tener que compilar paquetes desde el código fuente).

Conda es multiplataforma, con soporte para Windows, MacOS, GNU/Linux, y soporte para múltiples plataformas de hardware, como x86 y Power 8 y 9.

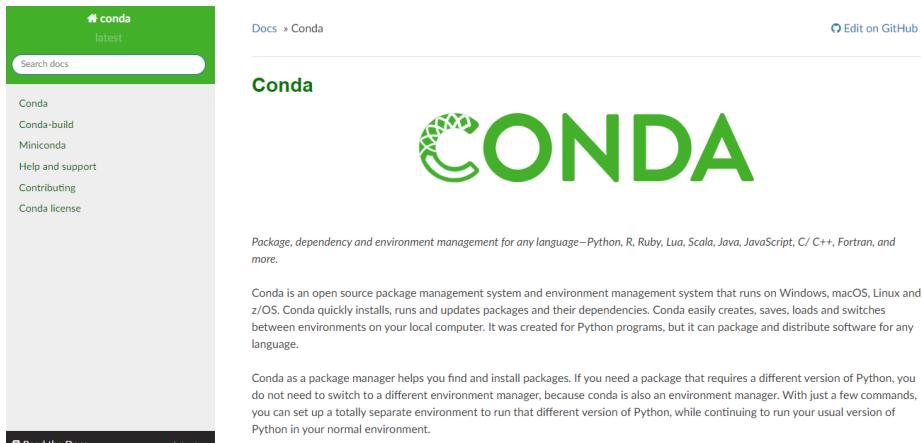
Conda está pensado para proyectos en ciencia de datos (*data scientists*), con los que compartimos diseño y ejecución de proyectos (*workflows*).

Instalación

La página de descarga de Miniconda es:

<https://conda.io/en/latest/miniconda.html>.

(Deberemos elegir el adecuado para nuestro sistema operativo)

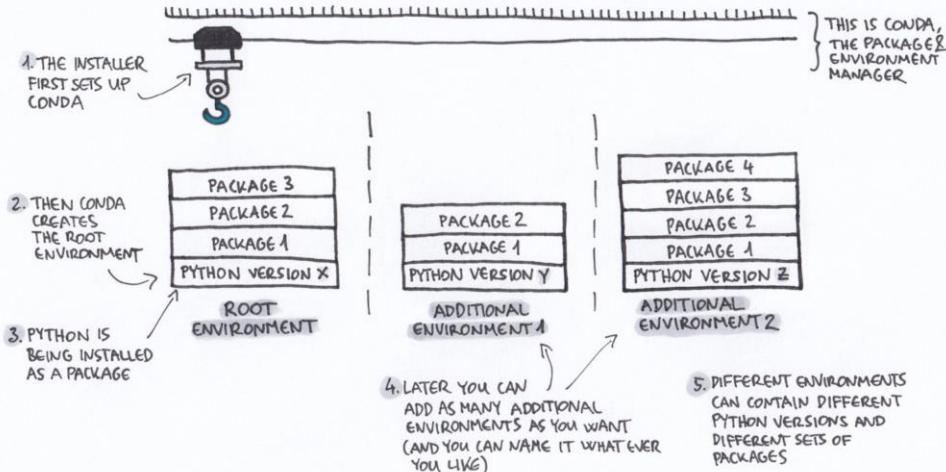


The screenshot shows the Conda documentation website. At the top, there's a navigation bar with a green header containing the Conda logo and the word "latest". Below the header is a search bar labeled "Search docs". To the right of the search bar are links for "Docs" and "Edit on GitHub". The main content area has a large green title "Conda" followed by a large green "CONDA" logo. Below the logo, there's a subtitle in small text: "Package, dependency and environment management for any language—Python, R, Ruby, Lua, Scala, Java, JavaScript, C/C++, Fortran, and more." A detailed description follows: "Conda is an open source package management system and environment management system that runs on Windows, macOS, Linux and z/OS. Conda quickly installs, runs and updates packages and their dependencies. Conda easily creates, saves, loads and switches between environments on your local computer. It was created for Python programs, but it can package and distribute software for any language." At the bottom of the page, there's a date: "22/04/2024".

<https://docs.conda.io/en/latest/>

Pero en la siguiente diapositiva veremos como instalarlo

Entornos en Conda



22/04/2024

https://angus.readthedocs.io/en/2019/conda_tutorial.html

Crear y navegar por los entornos

Aquí, los "entornos" representan la infraestructura en la que nuestro ordenador está operando actualmente. Esto incluye las variables y cómo se establecen (como nuestro PATH).

Entorno base

El entorno "base" de conda es, como suena, una especie de base de operaciones dentro de conda. No queremos instalar muchos programas aquí, ya que mientras más cosas se agreguen, es más probable que algo termine teniendo un conflicto.

Entornos en Conda



22/04/2024

Al empezar un proyecto nuevo, podemos crear un nuevo entorno e instalar lo necesario para ese proyecto en ese entorno, con un segundo proyecto haremos lo mismo. Si necesitamos volver a correr un pipeline del primer proyecto tiempo después, nos aseguramos de que todo funciona adecuadamente ya que no hemos vuelto a instalar nada en ese primer entorno y no deberíamos tener ningún conflicto entre las distintas herramientas.

Nosotros usaremos Conda de forma similar, inicialmente, nuestros proyectos serán nuestras asignaturas. Si realizáramos todas las instalaciones necesarias en este curso en el mismo lugar, tendríamos conflictos entre las distintas herramientas que veréis en las distintas asignaturas. Por lo general, usareis al menos un ambiente distinto en cada asignatura (en ocasiones más de uno).

Entornos en Conda



22/04/2024

Una vez instalado conda, debemos ser siempre conscientes del lugar en el que nos encontramos. Por defecto al abrir un terminal, ahora nos encontraremos en el entorno base. Desde aquí, podemos activar un entorno. Si salimos del entorno activado volvemos al entorno base.

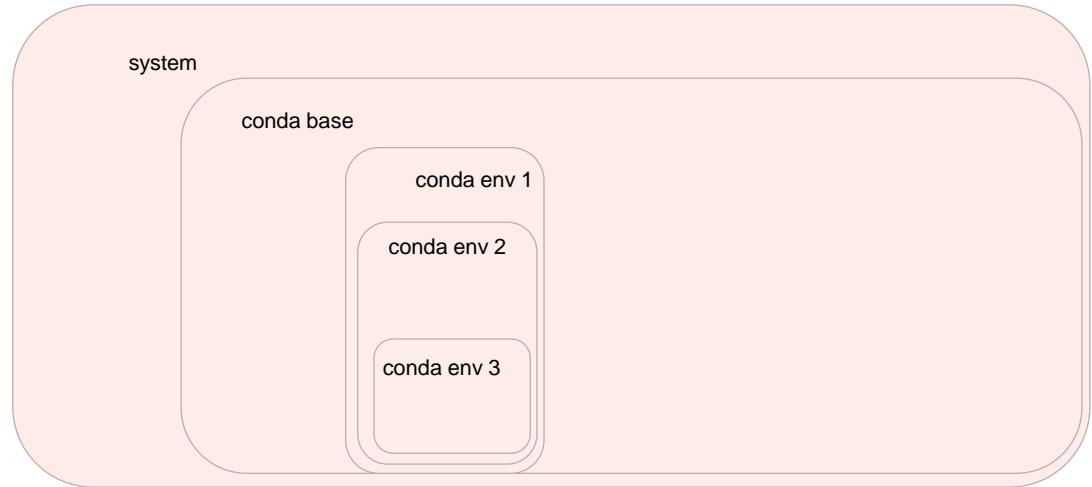
Entornos en Conda



22/04/2024

De esta forma podemos tener múltiples entornos e ir activando y desactivando para "entrar" y "salir" de ellos.

Entornos en Conda



22/04/2024

¡OJO! Es posible activar un entorno dentro de otro entorno y así sucesivamente... recordar siempre salir del entorno en el que nos encontramos a base antes de volver a entrar o pueden suceder cosas inesperadas. Como por ejemplo llamar a una versión de una herramienta pensando que al estar en otro entorno estamos llamando a otra...

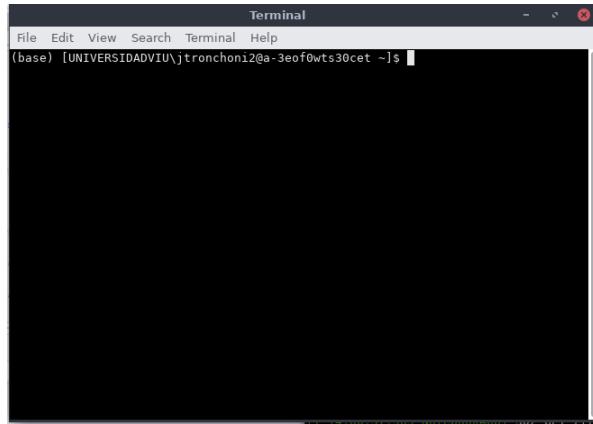
Instalación en WorkSpaces de AWS

```
#instalación de Conda (Miniconda):  
#creación de carpeta, preferiblemente en /home/USUARIO  
mkdir -p ~/miniconda3  
#descargar el instalador  
wget repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh -O ~/miniconda3/miniconda.sh  
#correr el instalador  
bash ~/miniconda3/miniconda.sh -b -u -p ~/miniconda3  
#eliminar el instalador  
rm -rf ~/miniconda3/miniconda.sh  
#ejecutar conda por defecto en el terminal  
~/miniconda3/bin/conda init bash  
#cierra y abre el terminal, deberías encontrarte en:  
#(base) [UNIVERSIDADVIU\USUARIO ~]$  
#(base): indica que nos encontramos en el entorno por defecto o base de Conda
```

22/04/2024

```
#abrimos una ventana del terminal  
#instalación de Conda (Miniconda):  
#creación de carpeta, preferiblemente en /home/USUARIO  
mkdir -p ~/miniconda3  
#descargar el instalador  
wget repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh -O  
~/miniconda3/miniconda.sh  
#correr el instalador  
bash ~/miniconda3/miniconda.sh -b -u -p ~/miniconda3  
#eliminar el instalador  
rm -rf ~/miniconda3/miniconda.sh  
#ejecutar conda por defecto en el terminal  
~/miniconda3/bin/conda init bash  
#cierra y abre el terminal, deberías encontrarte en:  
(base) [UNIVERSIDADVIU\USUARIO ~]$  
#(base): indica que nos encontramos en el entorno por defecto o base de Conda
```

Instalación en WorkSpaces de AWS



#(base) [UNIVERSIDADVIU\USUARIO ~]\$

#(base): indica que nos encontramos en el entorno por defecto o base de Conda

22/04/2024

¿Tengo Conda instalado?

Si al abrir un terminal vemos (base) al inicio del prompt, tenemos conda instalado
Comprobamos si tenemos conda instalado con

```
[UNIVERSIDADVIU\jtronchoni2@a-3eof0wts30cet ~]$ conda
```

¡Gracias!



Universidad
Internacional
de Valencia

universidadviu.com

De:
 Planeta Formación y Universidades