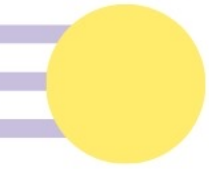


Borse di studio GARR  
Orio Carlini

# **CREDO: a friendly Customizable, REproducible, DOcker file generator for bioinformatics applications**

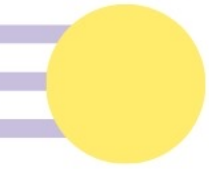
**Eliseo Martelli**

Università degli Studi di Torino – Centro Interdipartimentale di  
Biotecnologie Molecolari "Guido Tarone"



IL PROBLEMA IDENTIFICATO

Raggiungere la **riproducibilità**  
nella **bioinformatica** è difficile



**IL PROBLEMA IDENTIFICATO**

# Ma cosa significa difficile?

# Difficoltà nel tracciare le dipendenze

Mancanza di strumenti standardizzati

Impossibile ricostruire l'ambiente iniziale

Difficoltà nel tracciare le dipendenze

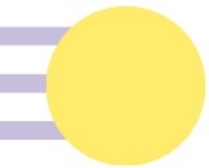
# **Mancanza di strumenti standardizzati**

Impossibile ricostruire l'ambiente iniziale

Difficoltà nel tracciare le dipendenze

Mancanza di strumenti standardizzati

**Impossibile ricostruire l'ambiente iniziale**



PROPOSTA

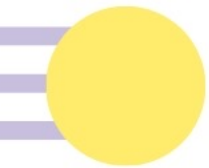
# CREDO: Il Package Manager Manager



FONDAMENTA

# La versione legacy di CREDO

10.1186/s12859-024-05695-9



OGGI

# La nuova versione di CREDO

---



OGGI

**La nuova versione di CREDO**



Consortium  
GARR

**Quali sono le principali differenze?**

OGGI

# La nuova versione di CREDO



**Architettura Modulare**

Modules

APT

PIP

CONDA

CRAN

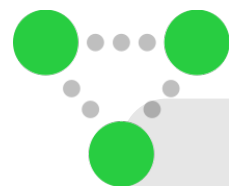
BIOCONDUCTOR

GIT

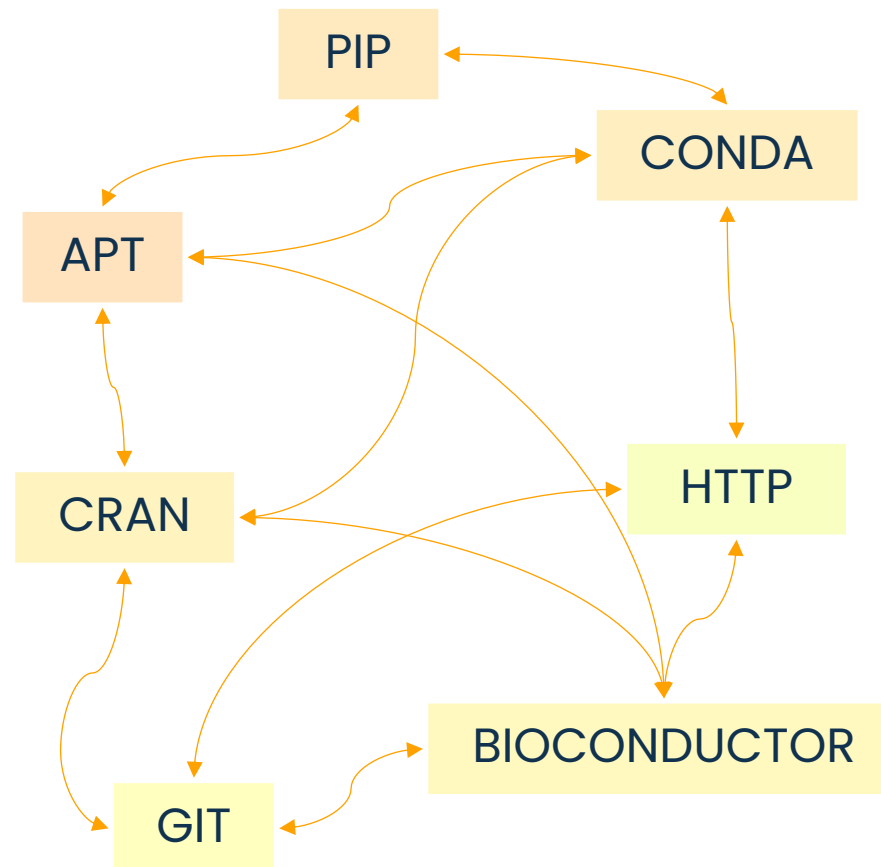
HTTP

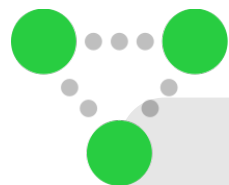
OGGI

# La nuova versione di CREDO



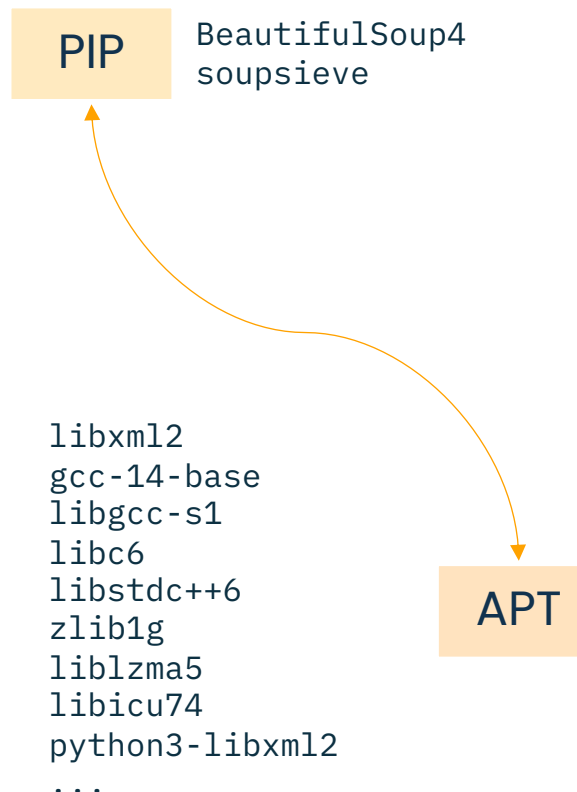
Grafo delle dipendenze





## Grafo delle dipendenze

### Esempio: BeautifulSoup 4 – Python



OGGI

## La nuova versione di CREDO

Consortium  
GARR



Multithreading

**Analizzo le dipendenze tra  
di loro indipendenti**

OGGI

## La nuova versione di CREDO



Multithreading

Analizzo le dipendenze tra  
di loro indipendenti

**Assegno l'installazione di  
una dipendenza ad ogni  
thread**

OGGI

## La nuova versione di CREDO



**Cache delle dipendenze**

**Nel test dell'environment**

01

OGGI

## La nuova versione di CREDO



**Cache delle dipendenze**

**Nell'installazione effettiva**

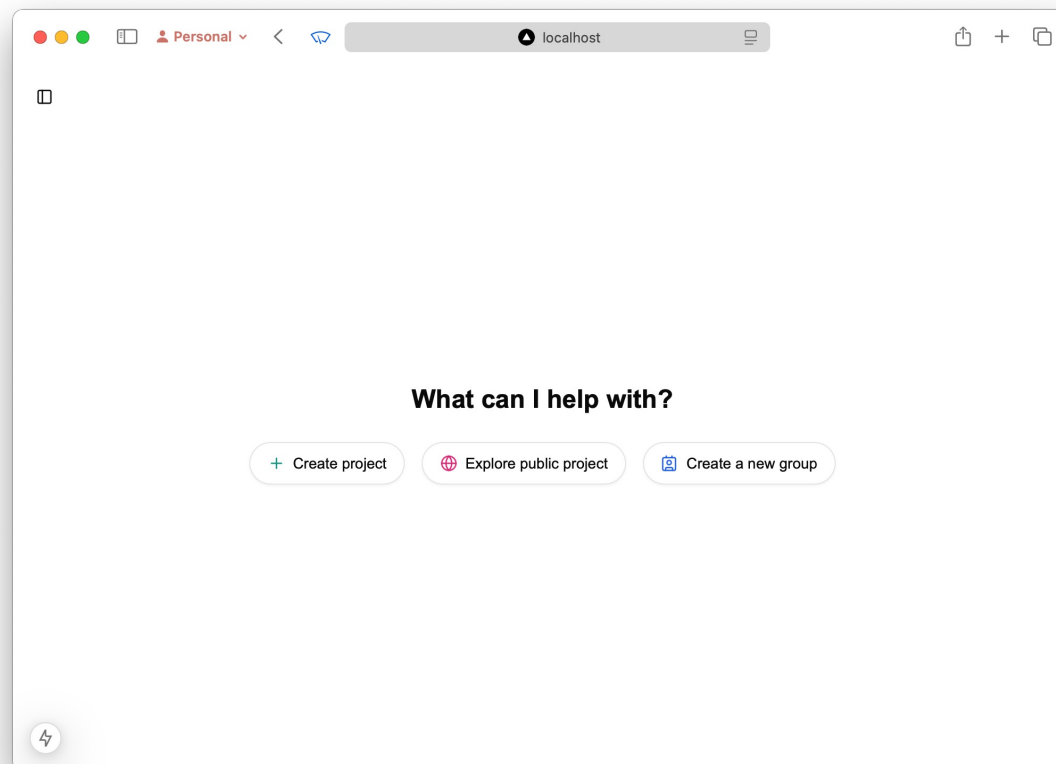
02

OGGI

# La nuova versione di CREDO



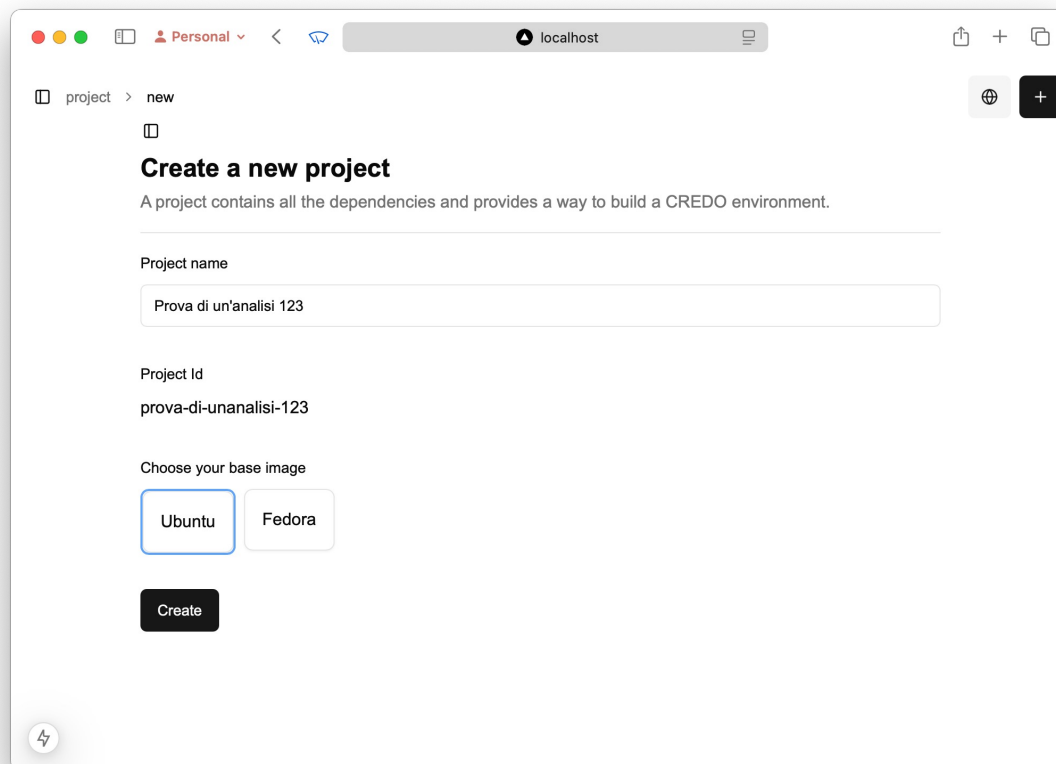
## Interfaccia Utente Migliorata



OGGI

# La nuova versione di CREDO

## Interfaccia Utente Migliorata



The screenshot shows a web browser window at localhost displaying the 'Create a new project' form. The form includes a title, a description, a 'Project name' input field with the value 'Prova di un'analisi 123', a 'Project Id' field with the value 'prova-di-unanalisi-123', and two radio buttons for 'Choose your base image' (Ubuntu and Fedora). A 'Create' button is at the bottom.

project > new

### Create a new project

A project contains all the dependencies and provides a way to build a CREDO environment.

Project name

Project Id

prova-di-unanalisi-123

Choose your base image

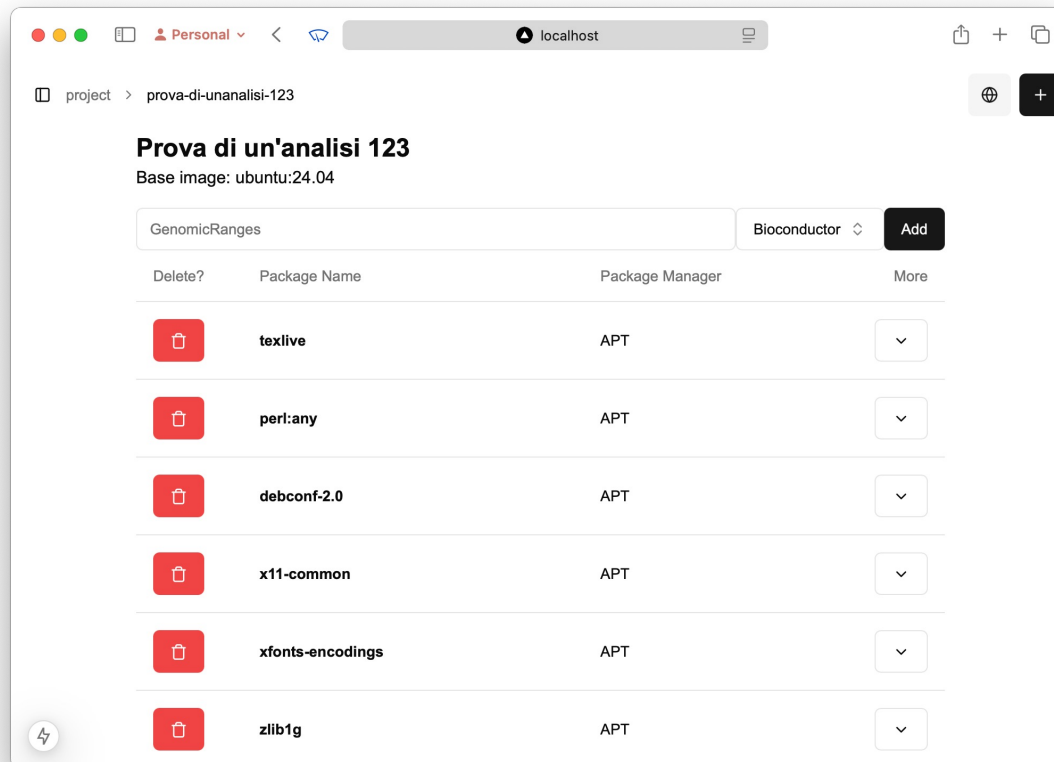
Ubuntu  Fedora

Create

OGGI

# La nuova versione di CREDO

## Interfaccia Utente Migliorata



The screenshot shows a web browser window displaying the CREDO interface. The browser address bar shows 'localhost'. The page title is 'Prova di un'analisi 123' with a subtitle 'Base image: ubuntu:24.04'. Below the title, there is a search bar containing 'GenomicRanges' and a dropdown menu set to 'Bioconductor', with an 'Add' button. A table lists installed packages with columns for 'Delete?', 'Package Name', 'Package Manager', and 'More'.

Delete?	Package Name	Package Manager	More
	texlive	APT	
	perl:any	APT	
	debconf-2.0	APT	
	x11-common	APT	
	xfonts-encodings	APT	
	zlib1g	APT	

OGGI

## La nuova versione di CREDO



**Suggerimenti di  
Dipendenze**

**Pensata per le librerie che  
non dichiarano le proprie  
dipendenze**

OGGI

## La nuova versione di CREDO



**Suggerimenti di  
Dipendenze**

Pensata per le librerie che  
non dichiarano le proprie  
dipendenze

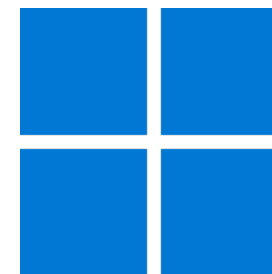
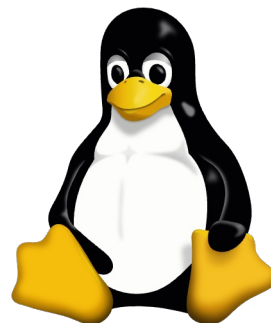
**Infrastruttura per l'analisi  
statica del codice**

OGGI

## La nuova versione di CREDO



Supporto per  
più piattaforme



OGGI

## La nuova versione di CREDO



Suggerimenti



Grafo dipendenze



Architettura Modulare



API



Multithreading



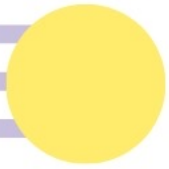
Caching



Interfaccia utente migliorata



Supporto per più piattaforme

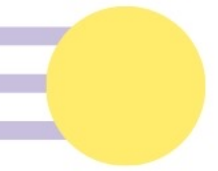


# La nuova versione di CREDO

# 01

VANTAGGI

## Miglioramenti nella riproducibilità



# 02

VANTAGGI

## Accessibilità degli ambienti



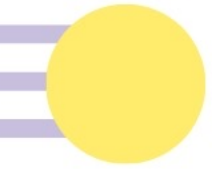
OGGI

La nuova versione di CREDO



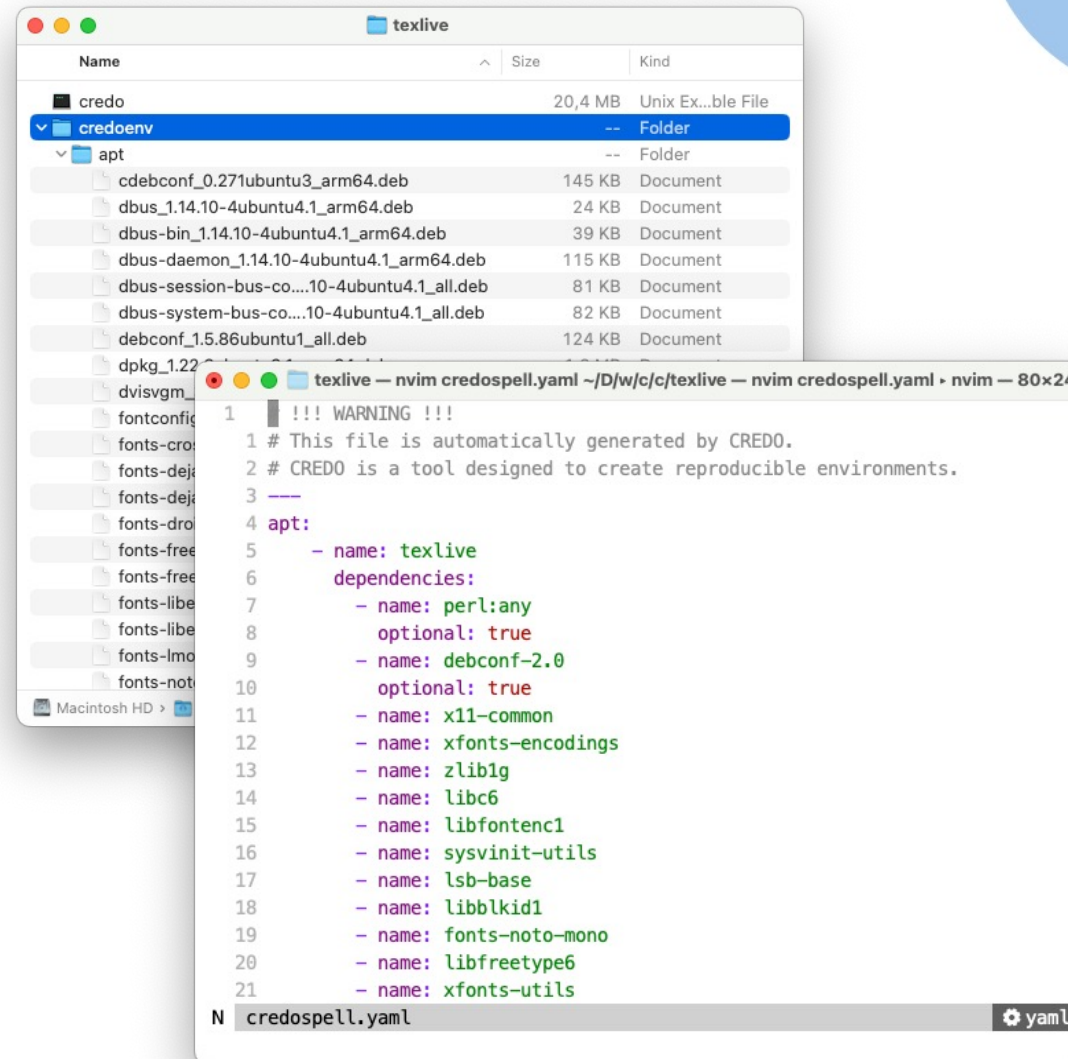
Consortium  
GARR

Qual è il risultato di CREDO?

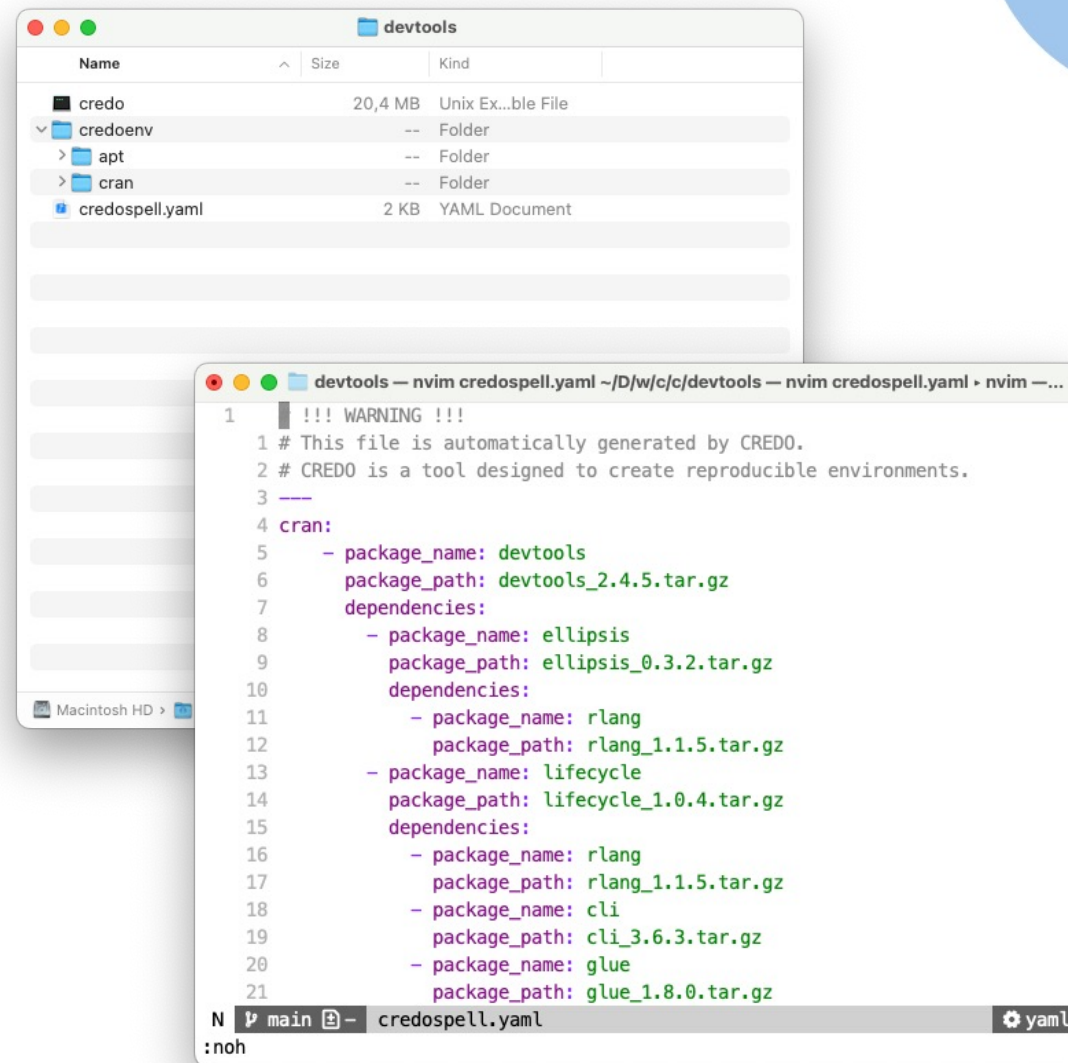


# Utilizziamo come esempio tre librerie

# Texlive



# Devtools



The image shows a macOS file explorer window titled 'devtools' with the following contents:

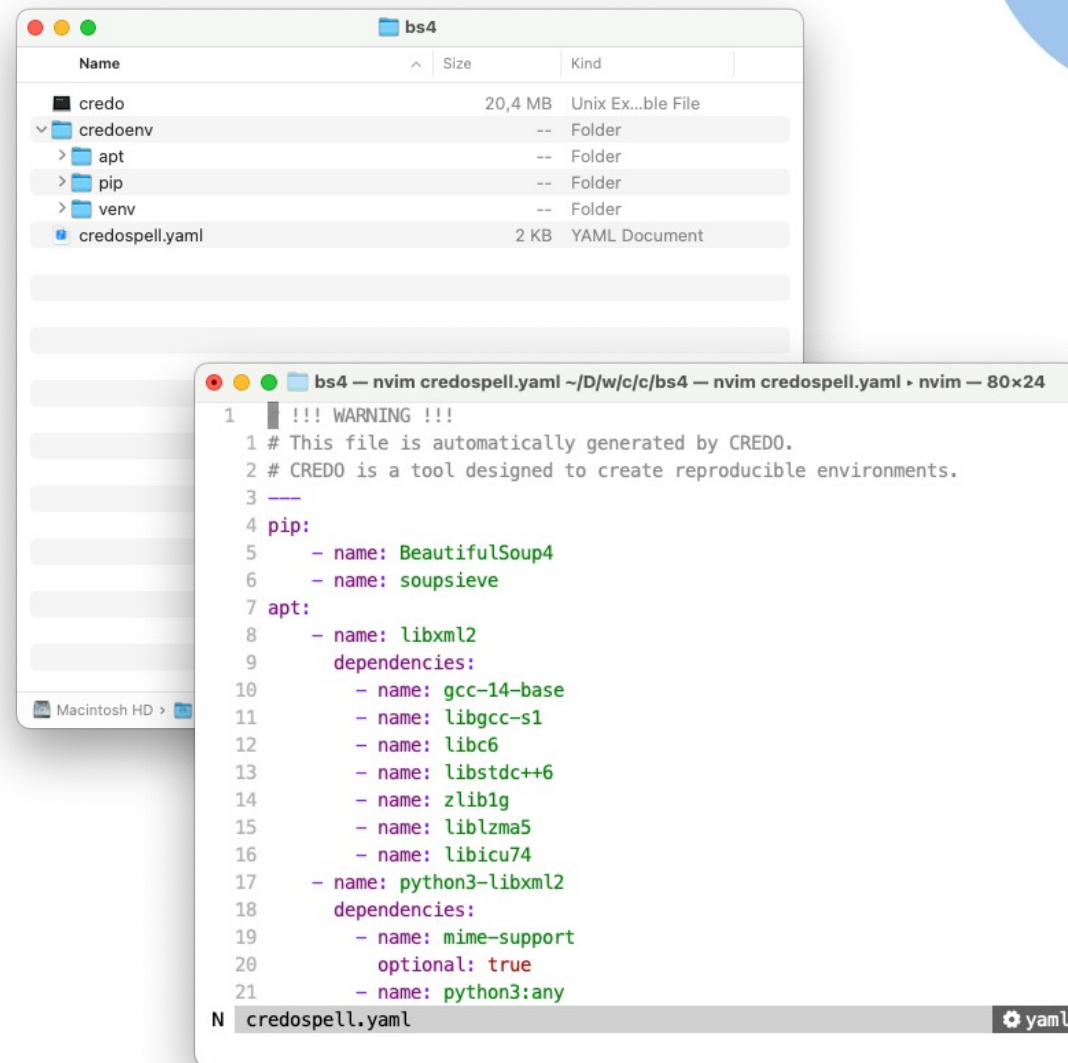
Name	Size	Kind
credo	20,4 MB	Unix Ex...ble File
credoenv	--	Folder
apt	--	Folder
cran	--	Folder
credospell.yaml	2 KB	YAML Document

Below the file explorer is a terminal window showing the contents of the 'credospell.yaml' file:

```

1  !!! WARNING !!!
2  # This file is automatically generated by CREDO.
3  # CREDO is a tool designed to create reproducible environments.
4  ---
5  cran:
6    - package_name: devtools
7      package_path: devtools_2.4.5.tar.gz
8      dependencies:
9        - package_name: ellipsis
10          package_path: ellipsis_0.3.2.tar.gz
11          dependencies:
12            - package_name: rlang
13              package_path: rlang_1.1.5.tar.gz
14            - package_name: lifecycle
15              package_path: lifecycle_1.0.4.tar.gz
16            dependencies:
17              - package_name: rlang
18                package_path: rlang_1.1.5.tar.gz
19              - package_name: cli
20                package_path: cli_3.6.3.tar.gz
21              - package_name: glue
22                package_path: glue_1.8.0.tar.gz
  
```

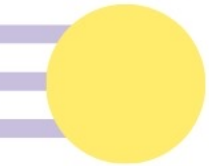
# BeautifulSoup4



The image shows a file explorer window for a directory named 'bs4'. It contains a 'credo' file (20.4 MB, Unix Executable File), a 'credoenv' folder, and subfolders for 'apt', 'pip', and 'venv'. A 'credospell.yaml' file (2 KB, YAML Document) is also present. An overlaying code editor shows the content of 'credospell.yaml' with the following configuration:

```

1  !!! WARNING !!!
2  # This file is automatically generated by CREDO.
3  # CREDO is a tool designed to create reproducible environments.
4  ---
5  pip:
6    - name: BeautifulSoup4
7    - name: soupsieve
8  apt:
9    - name: libxml2
10   dependencies:
11     - name: gcc-14-base
12     - name: libgcc-s1
13     - name: libc6
14     - name: libstdc++6
15     - name: zlib1g
16     - name: liblzma5
17     - name: libicu74
18   - name: python3-libxml2
19   dependencies:
20     - name: mime-support
21     optional: true
22     - name: python3:any
  
```



# CREDO apply

Prepara e esegue un ambiente generato con CREDO

```
bs4 -- ~/D/w/c/c/bs4 -- -fish -- 80x9
[+] bs4 git:(main) ./credo apply
Looking in links: /Users/eliseomartelli/Developer/work/credo/core/bs4/credoenv/p
ip
Requirement already satisfied: BeautifulSoup4 in ./credoenv/venv/lib/python3.13/
site-packages (4.12.3)
Requirement already satisfied: soupsieve>1.2 in ./credoenv/venv/lib/python3.13/s
ite-packages (from BeautifulSoup4) (2.6)
Looking in links: /Users/eliseomartelli/Developer/work/credo/core/bs4/credoenv/p
ip
```



OGGI

**La nuova versione di CREDO**



Consortium  
GARR

**Ma.. Ci sono ancora dei limiti..**



OGGI

La nuova versione di CREDO



Consortium  
GARR

01

**Migliorare la GUI di CREDO con i  
feedback degli utenti**



OGGI

La nuova versione di CREDO



Consortium  
GARR

FAIR

02

**Abbiamo ancora problemi con la  
Findability**



OGGI

La nuova versione di CREDO



Consortium  
GARR

02

**Dobbiamo ancora integrare DOI  
agli artefatti generati per  
facilitarne citazioni e riferimenti.**

OGGI

La nuova versione di CREDO

Consortium  
GARR

03

# Publicare la nuova versione di CREDO



(GIGA)<sup>n</sup>  
SCIENCE

BMC  
Bioinformatics



OGGI

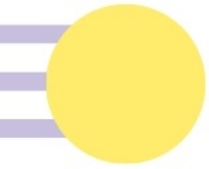
**La nuova versione di CREDO**



Consortium  
GARR

04

**Catturare parti dell'ambiente che  
non sono ancora gestite da CREDO**



**FUTURO**

# Cosa ci riserva il futuro?

01

**Vogliamo rendere CREDO una runtime sulla quale poter eseguire le proprie analisi**

# 02

## **Implementazione di un protocollo peer-to-peer per distribuire la computazione**

03

**Rendere l'onboarding e  
l'utilizzo di CREDO ancora più  
semplice**

04

**PRIVACY: per rendere CREDO  
adatto all'elaborazione di  
dati dei pazienti**

05

## **Valutazione di alcune librerie per Homomorphic Encryption**

06

**Modifica di STAR per poter  
effettuare allineamento di  
sequenze biologiche criptate**

07

**Testing sul campo di questa  
evoluzione di CREDO e  
pubblicazione di un nuovo  
articolo**

Borse di studio GARR  
Orio Carlini

**Grazie per l'attenzione**

Eliseo Martelli