



UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO

IA369Z - REPRODUTIBILITY IN COMPUTATIONAL
RESEARCH

Best Practices for Reproducible Research

PROFA.: LETICIA RITTNER

CARLOS RONY RECALDE VITONERA, RA:229973

CAMPINAS,SP - 1S/2019

¿ POR DONDE COMENZAR ?

Saludos coordiales...

Si has llegado hasta aquí es porque igual que yo, tú también estás interesado y te encuentras aprendiendo un poco más sobre el campo de la reproducibilidad dentro de la investigación. En este diario de buenas prácticas mi objetivo es compartir contigo mi experiencia a lo largo de reproducir mi artículo, proporcionandote consejos que te puedan llegar a ser útiles para que al final consigas reproducir tu propio artículo.

Para comenzar debemos de ser selectivos a la hora de elegir un artículo para hacerlo reproducible, en esta diciplina hemos llegado a la conclusión unanime que no existe un camino correcto para poder hacer reproducibilidad, apenas hemos intentando seguir ciertas reglas que nos acerquen al objetivo final.

La ciencia busca resultados de experimentos reproducibles, las investigaciones avanzan mucho más rápido cuando es posible verificar y tener disponibles los resultados de los otros y así evitar gastar tiempo siguiendo otros caminos que al final sean poco útiles. En la práctica ni todos los artículos llegan a ser reproducibles, es por eso que desde el inicio hemos seleccionado 5 elementos importantes para trabajar en reproducibilidad.

1.Datos

Van a ver muchas ocasiones en las cuales no se pueda tener acceso a información de forma pública, es por eso que deben ser cuidadosos eligiendo el artículo que planean reproducir.

Consejos:

1. Hacer uso de repositorios online tales como Github o Bitbucket.
2. De ser posible, escoger datos que puedan ser corroborados como verdaderos.
3. Verificar que los datos se encuentren completos y no haya sido disponibilizada apenas una parte.
4. Entender los tipos de datos disponibles.
5. Clasificar los tipos de datos.

Enlaces de interés:

<https://github.com>

<https://desktop.github.com/>

2. Código

Otro punto importante sin duda es el código, es bien conocido los problemas que el usar el código de otra persona pueda traer, para eso te comparto ciertos consejos para evitarlos.

Consejos:

1. Utilizar herramientas de control de versión.
2. Usar la misma versión sobre la cual haya sido escrito el código.
3. Hacer uso de repositorios para almacenar el código.
4. Intenta escribir tu código con comentarios y asignando nombres relacionados para que sea más simple de entender.
5. Es bueno crear un archivo Readme que proporcione informaciones claves para las demás personas.

3. Documentación

En el semestre revisamos diferentes herramientas que pueden ser usadas en reproducibilidad, como sugerencia todos interactuamos con Jupyter Notebook, que es una herramienta de Literate Computing la cual permite la combinación de texto y código ejecutado en tiempo real.

Consejos para trabajar en Jupyter:

1. Es más sencillo iniciar Jupyter desde la terminal.
2. Es mejor crear varios notebooks para cada parte importante del trabajo y al final juntarlos.
3. Poner atención a los nombres asignados para cada notebook creado para llevar el orden.
4. Use bibliotecas que puedan ser útiles a su proyecto y simplifiquen el trabajo.
5. Sea organizado.

4. Ambiente

Uno de los principales problemas al intentar reproducir un trabajo ha sido el cambio de versiones, las nuevas versiones que van saliendo no siempre son compatibles con antiguas versiones, y se da el caso que por ejemplo al intentar reproducir un trabajo realizado hace dos años con una nueva versión de los programas utilizados, ya no se consigan llegar a los resultados esperados, o en el peor de los casos, no haya un soporte para la versión en que se trabajó el trabajo. Para lo cuál te ofrezco los siguiente consejos.

Consejos

1. Agrupar todas las dependencias necesarias para facilitar el trabajo de reproducción.
2. De lo posible, trabajar en más de un ambiente (windows,linux,mac os).
3. Disponibilizar el trabajo en un contenedor o una máquina virtual.
4. Tener un archivo README indicando informaciones importantes.
5. Crear un mini tutorial con todos los pasos sin obviar ninguno.

Links <https://www.docker.com/>

<https://www.virtualbox.org/>

5. Workflow

Este último elemento puede como no ser considerado, sin embargo es una gran ayuda para entender el funcionamiento del trabajo de otra persona. Aquí de forma gráfica se explica el proceso que se sigue a través de la reproducción de un trabajo. La idea es poder ofrecer un mapa de todo el trabajo para que alguien más pueda guiarse y que consiga llegar al resultado esperado.

Consejos

1. Tener claras las ideas sobre el funcionamiento de tu trabajo.
2. Ser coherentes en los pasos.
3. Usar herramientas de diseño.
4. Ser claros y creativos.
5. Incluya apenas los puntos importantes, no código.

Links <https://www.docker.com/>

<https://www.virtualbox.org/>