

# Open Science@HU; Introducing the tools

Marc A.T. Teunis

2020-06-01 14:35:32

## Contents

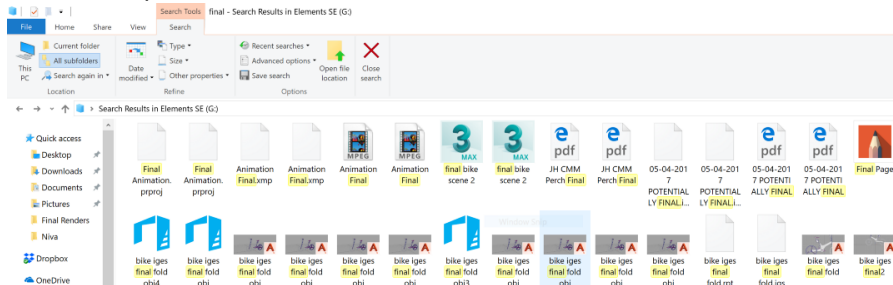
- ▶ Part 1; Introducing Open Science
- ▶ Part 2; Managing your project with 'Guerilla Analytics'
- ▶ Part 3; Open Science @HU

# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

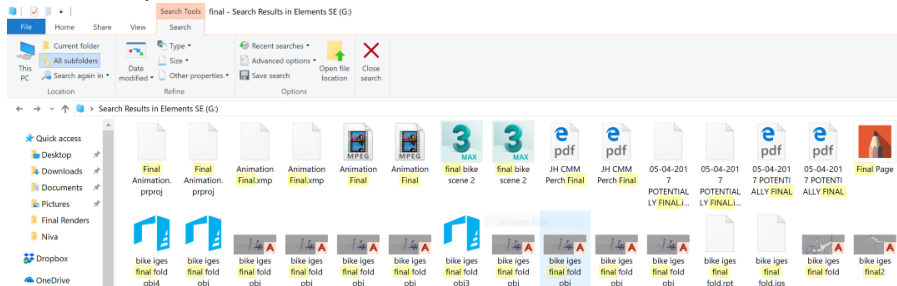


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

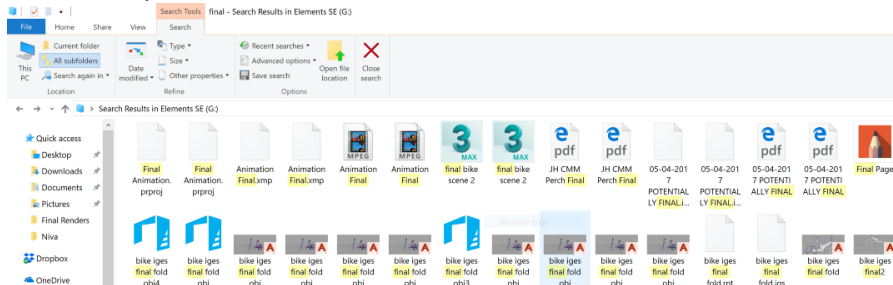


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

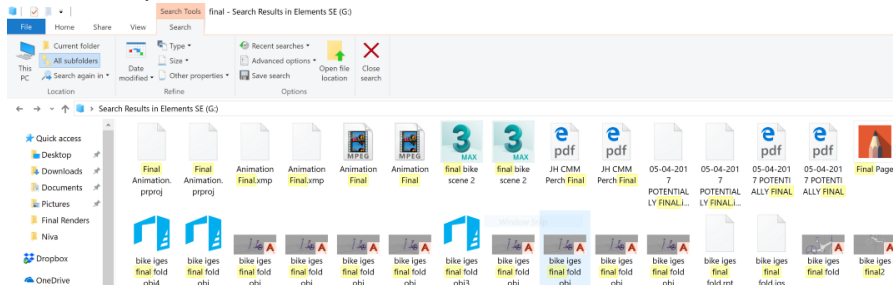


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

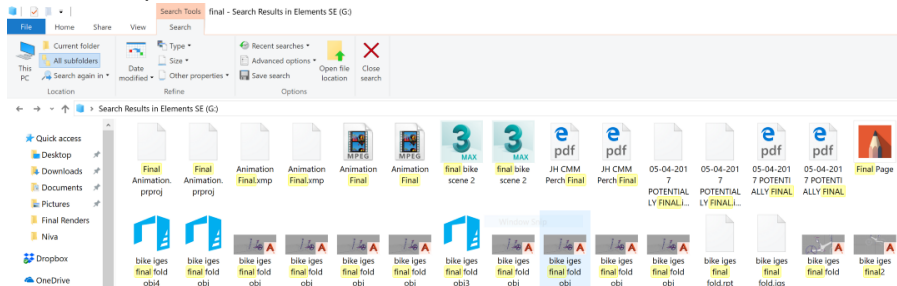


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

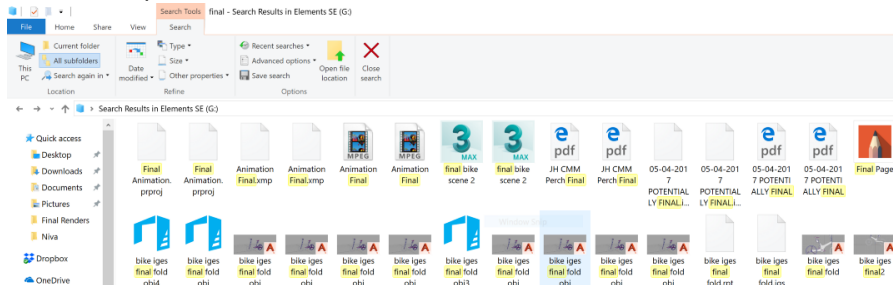


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>



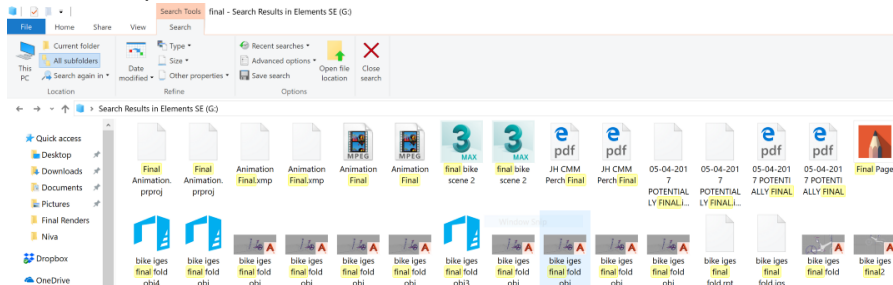


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

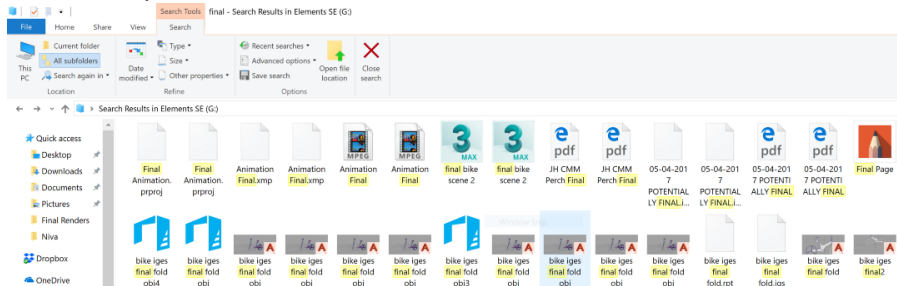


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

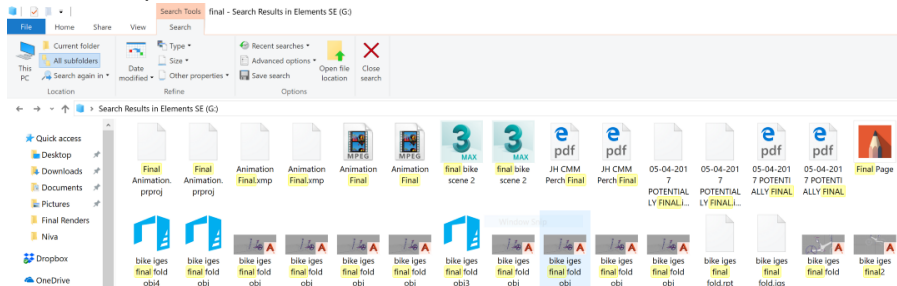


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>

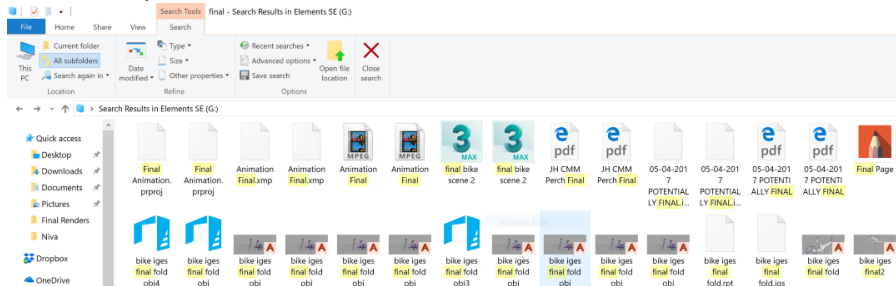


# Part 1; Introducing Open Science

1. When things go wrong
2. Why Open Science?
3. The need for learning programming
4. Open Science tools

**We all know this!**

from: <https://medium.com/@jameshoareid/final-pdf-finalfinal-pdf-actualfinal-pdf-cae61ab1d94c>



## Part2; Managing your project with 'Guerilla Analytics'

1. Files and folders / Project structure
2. Data structures
3. Wide data-formats
4. Long data-formats
5. Coding variables
6. Data integrity (md5sums)
7. Factor levels (labels)

### The Guerilla Analytics Principles

- ▶ Principle 1: Space is cheap, confusion is expensive
- ▶ Principle 2: Prefer simple, visual project structures and conventions
- ▶ Principle 3: Prefer automation with program code
- ▶ Principle 4: Maintain a link between data on the file system,

# File names and file formats

# Data structure -> A context free solution

# Wide data-formats



# Long data-formats

## Part 2

1. Data integrity (md5sums)
2. Factor levels (labels)
3. FAIR principles
4. HU ResearchDrive - using Rsync to transfer and check files

## Meta data

<http://rd-alliance.github.io/metadata-directory/standards/>

## Part 3; Open Science @HU

### Access Research Drive via Rclone

#### HU-RD Wiki

In this part you will find documentation about rclone. Rclone is the rsync for cloud storage. Information on how to install rclone and other things may be found at: <https://rclone.org>.

Apart from being an rsync-type tool for cloud storage, it has the following features:

MD5/SHA1 hashes checked at all times for file integrity

Timestamps preserved on files

Partial syncs supported on a whole file basis

Copy mode to just copy new/changed files

Sync (one way) mode to make a directory identical

Check mode to check for file hash equality

Can sync to and from network, eg two different cloud accounts

Optional encryption (Crypt)