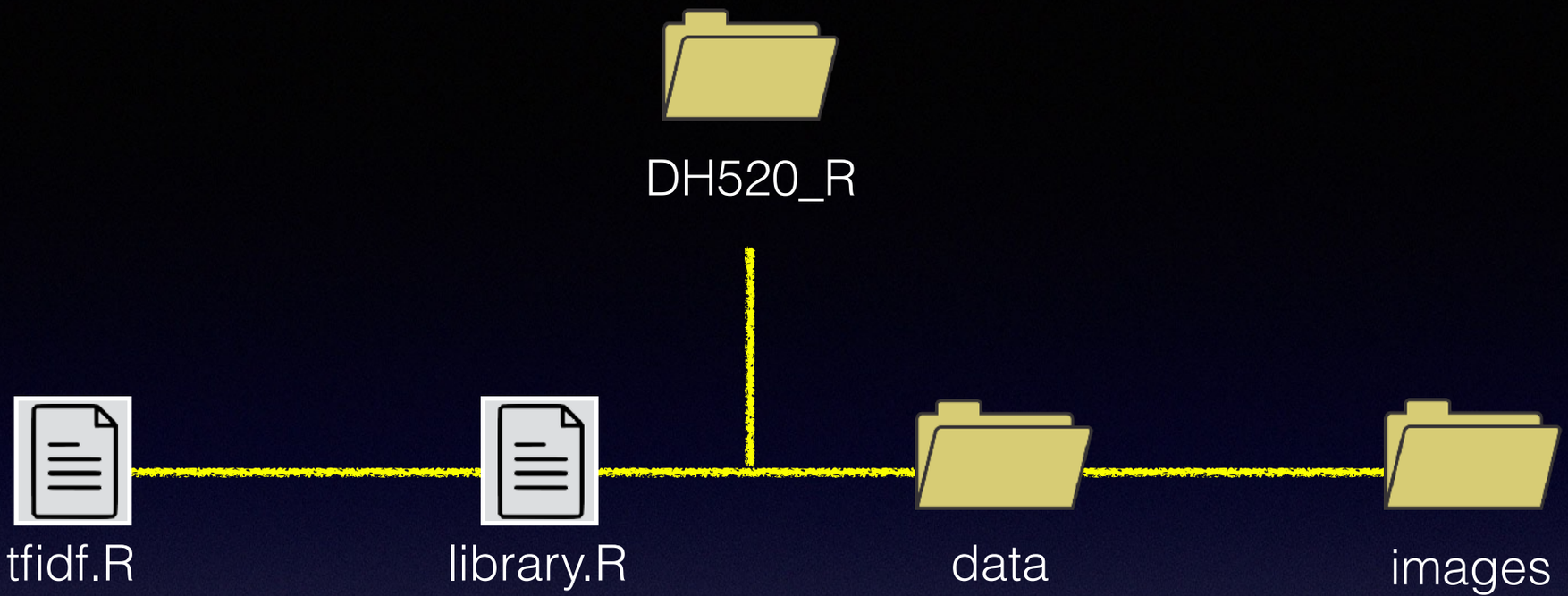
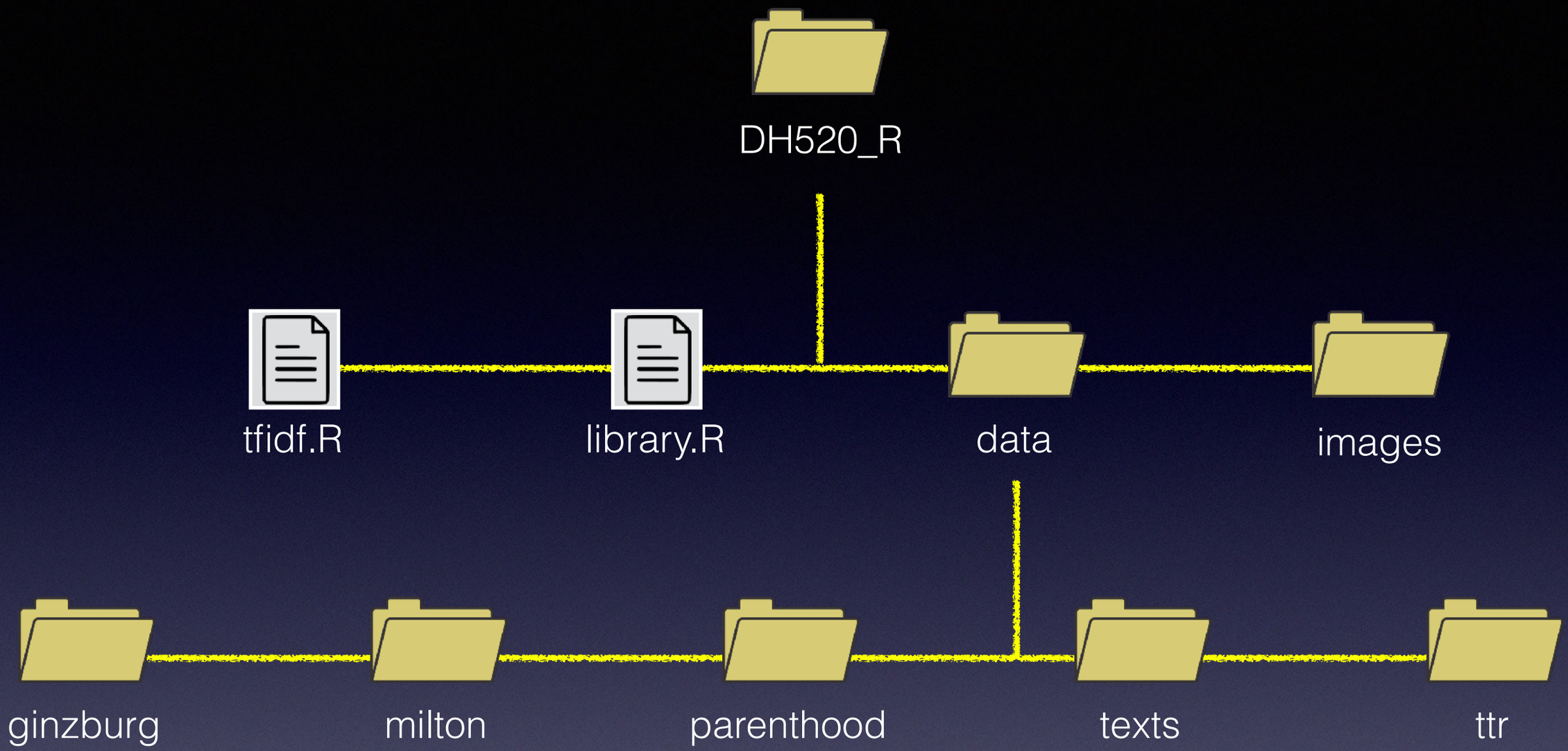
File Paths and R

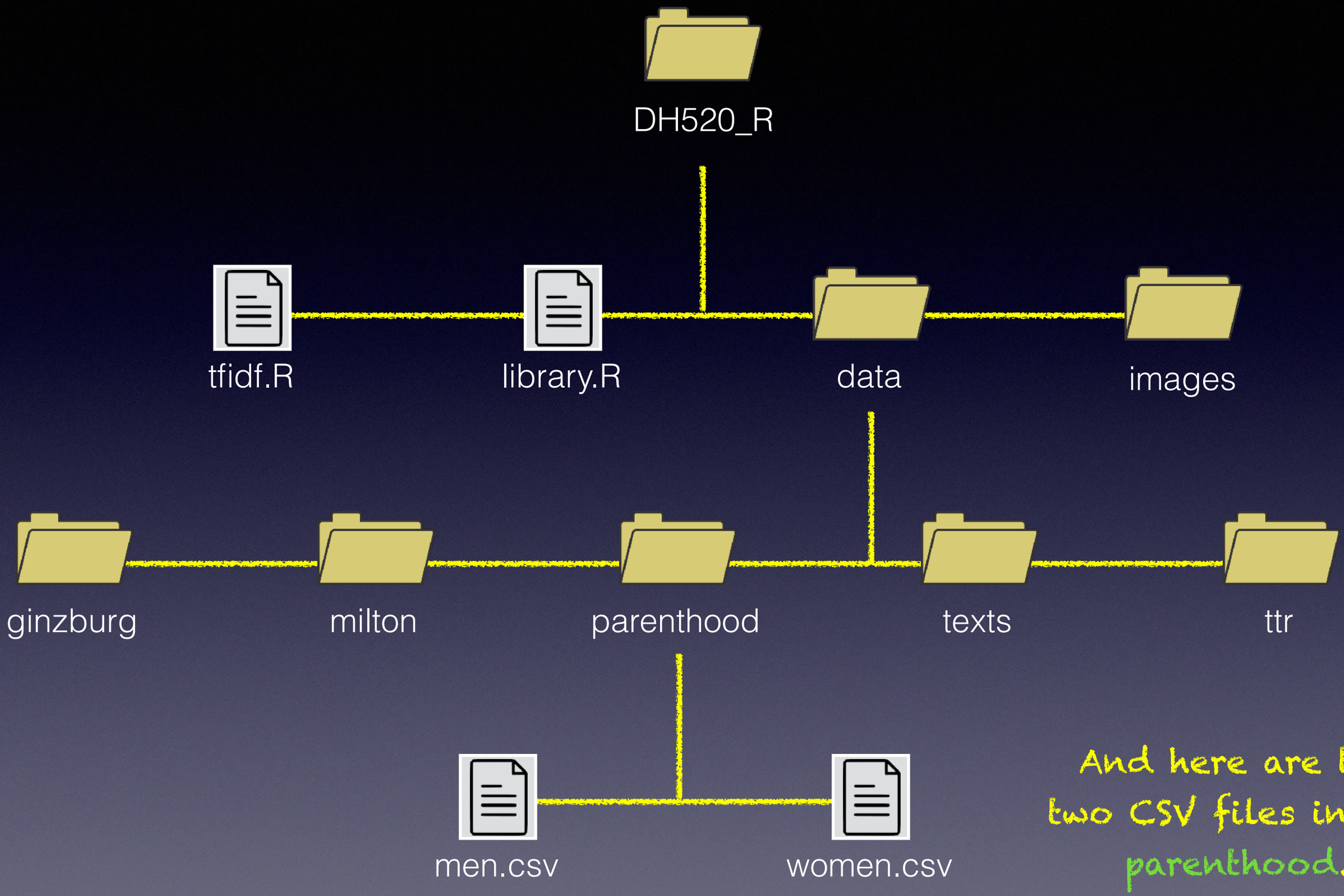
or,
Why Can’t R Find My Data?



This is a representation of how we will (eventually) set up our DH520_R folder.

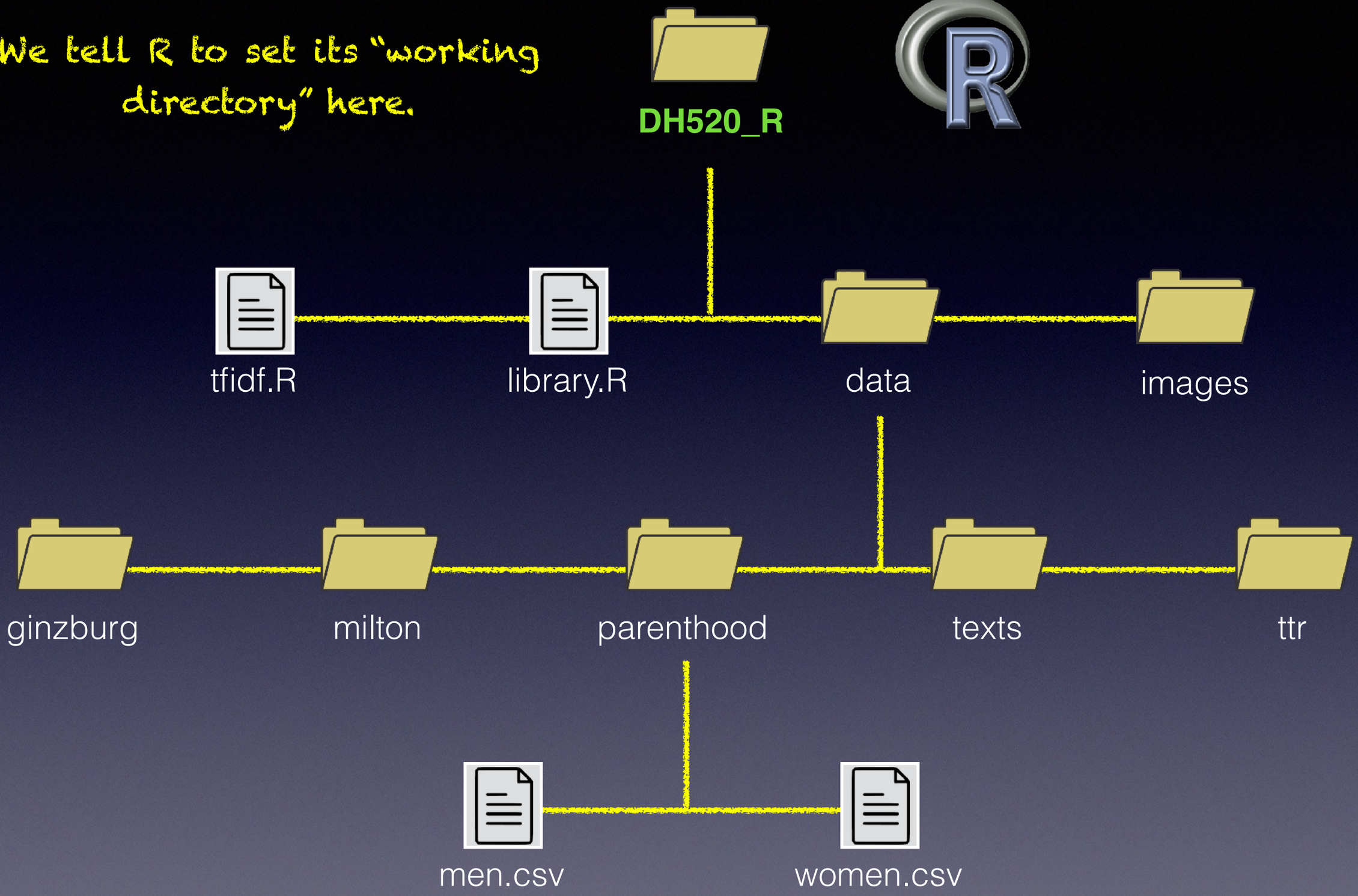


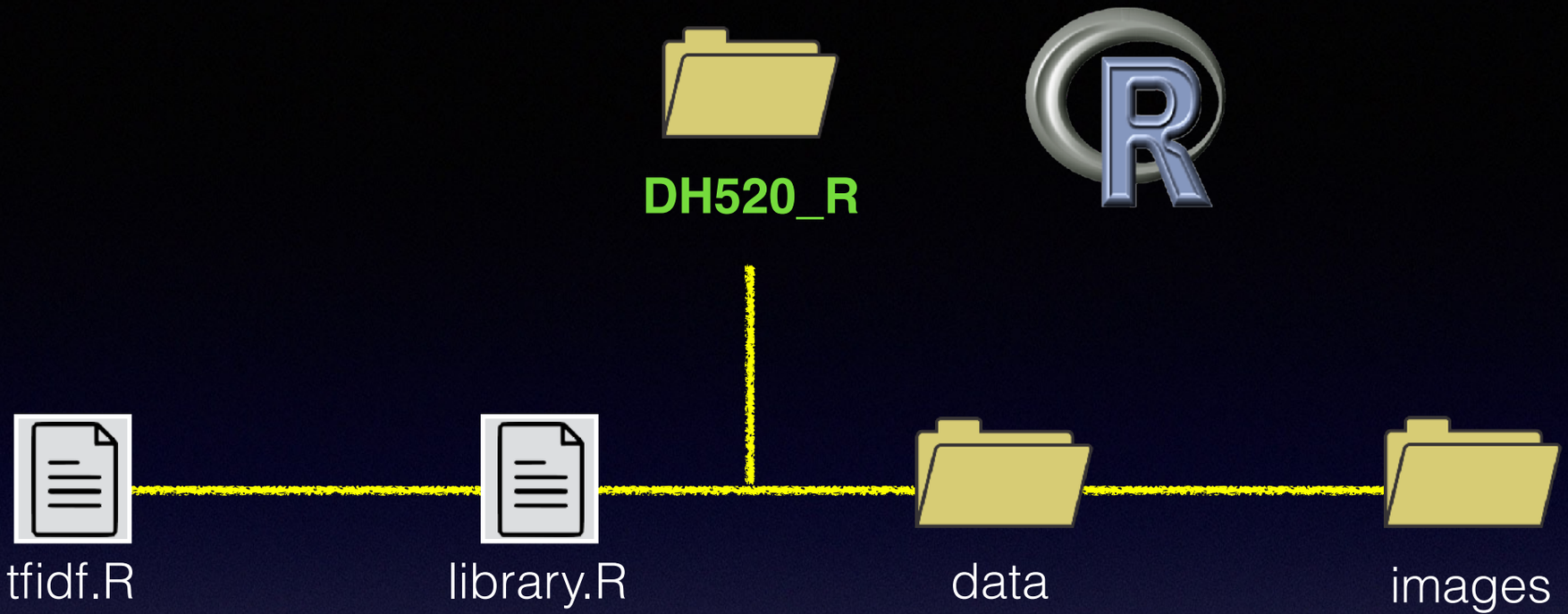
If we add the sub-folders inside data, it looks a bit like this (I missed a few....)



And here are the two CSV files inside parenthood.

We tell R to set its "working directory" here.





However, R can only see
things inside that one folder!

Unless I tell R exactly how to
find other folders or files,
this is all that R can see.

DH520_R

tfidf.R

library.R

data

images

ginzburg

milton

parenthood

texts

ttr

men.csv

women.csv

So if I want R to load women.csv down here, I need to tell R how to find it.

DH520_R

tfidf.R

library.R

data

images

ginzburg

milton

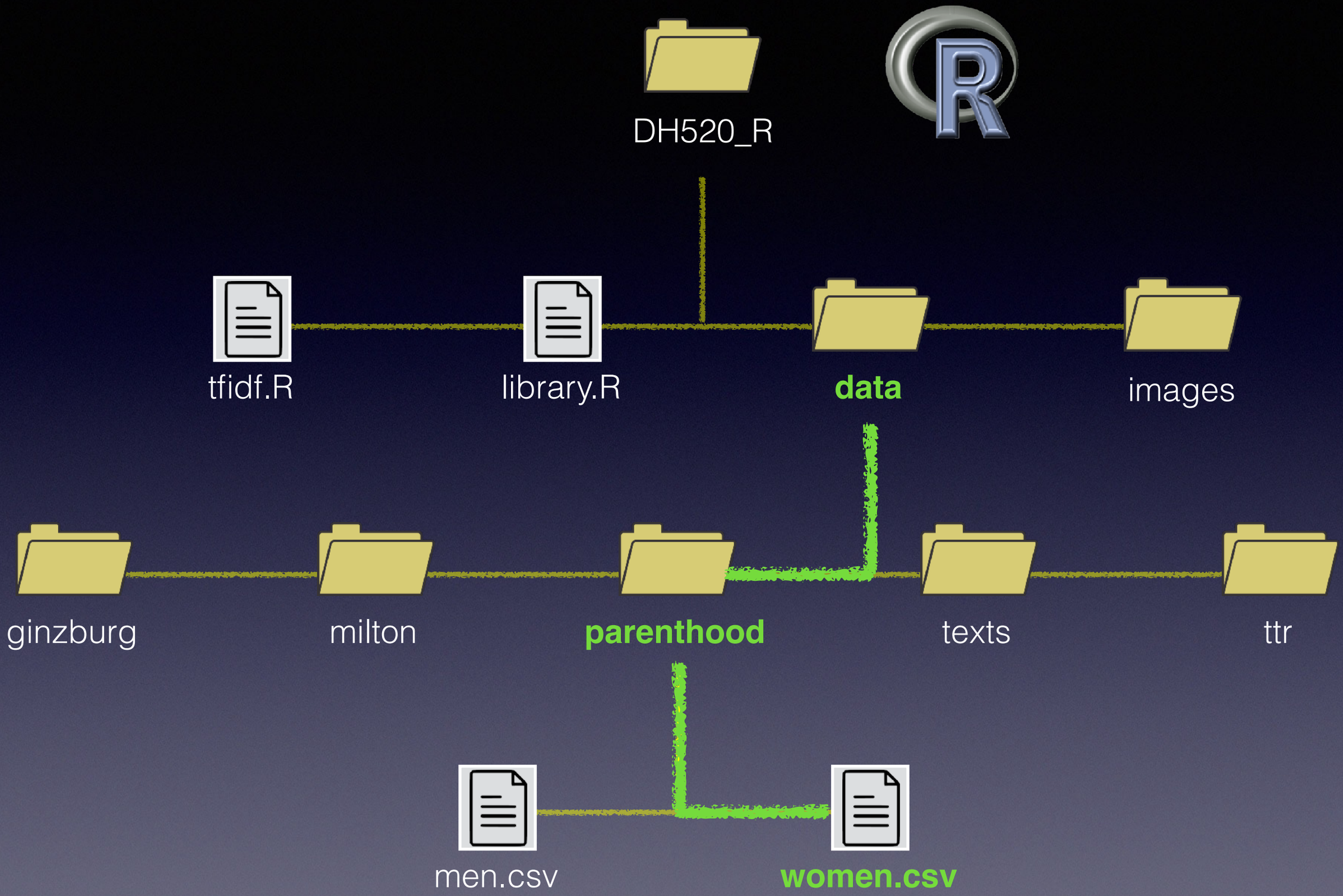
parenthood

texts

ttr

men.csv

women.csv



So the path is "data/parenthood/women.csv"

The / represents a folder, so this is a string representation of the folder hierarchy that I've created.

```
women <- read.csv("data/parenthood/women.csv")
```

That is the path "relative" to R's current working directory.

The / represents a folder, so this is a string representation of the folder hierarchy that I've created.

```
women <- read.csv("data/parenthood/women.csv")
```

That is the path "relative" to R's current working directory.

Alternately, I could specify a "full path" that lists all the folders on my hard drive. For example, on my laptop it's this:

```
/Users/hquamen/Documents/Courses/DH520_R/data/parenthood/women.csv
```

/Users/hquamen/Documents/Courses/**DH520_R**/data/parenthood/women.csv

When I tell R to set its "working directory" as DH520_R, I'm telling R to move to this location. Now everything R does starts from here.

~~/Users/hquamen/Documents/Courses/~~DH520_R/data/parenthood/women.csv

Now, R doesn't need to know any of the preliminary path information that originally got us to this location (that is, R needs no information that is "higher" in the directory tree).

~~/Users/hquamen/Documents/Courses/~~DH520_R/**data/parenthood/women.csv**

Consequently, the file paths needed to access my data all become shorter and simpler.

Interestingly, this is also how web URLs work:

- an access protocol followed by
- a server followed by
- a number of folders, followed by
- a file:

http://www.ualberta.ca/efs/courses/english150.html

File Paths and R

or,
Why Can’t R Find My Data?