Duplicate Node Remover Explanation

The Duplicate Node Remover is to be used when a customer has had some of their data duplicated, resulting in multiple branches in the tree view that contain mostly the same information. Whether for RPC or production, this tool can be run against their data to resolve that issue. The tool is given a starting point and will recursively crawl through the customer's structure, identifying duplicates at every level and pruning them off of the tree.

Consider a tree with these two folders inside of the same drawer:

```
Folder 1
  |__Folder 2
  |     |__File 1 (10 kb)
  |     |__File 2 (10 kb)
  |     |__File 3 (10 kb)
  |__Folder 3
  |     |__File 4 (10 kb)
```

```
Folder 1
  |__Folder 2
  |     |__File 1 (10 kb)
  |     |__File 2 (15 kb)
  |     |__File 5 (10 kb)
  |__Folder 4
  |     |__File 6 (10 kb)
```

If the tool was run against this tree the result would be as follows:

```
Folder 1
  |__Folder 2
  |     |__File 1 (10 kb)
  |     |__File 2 (10 kb)
  |     |__File 2 (1) (15 kb)
  |     |__File 3 (10 kb)
  |     |__File 5 (10 kb)
  |__Folder 3
  |     |__File 4 (10 kb)
  |__Folder 4
  |     |__File 6 (10 kb)
```

What the tool has done is selected the first Folder 1 as the original and then merged the other Folder 1 into it.

- First it removed the duplicate File 1, because it was exactly the same as the original.
- Next, there was a duplicate File 2, but both copies had different sizes, so it renamed the second copy and moved it into the same location as the original.
- It then moved File 5 into the original Folder 2 because it did was a totally unique file.
- Finally, it moved Folder 4 (and all of its contents) into the original Folder 1, because it was a totally unique folder.

This tool will prune the data whether it has been doubled, tripled, or any other number of duplications. It will also traverse any level of folders and sub-folders within the tree.

Revision #3
Created 15 August 2023 17:09:02 by Abe Austin
Updated 29 September 2023 22:17:12 by Abe Austin