

From: Shulman, Stu
To: [Bowman, Randal](#); [Mark Hoy](#)
Cc: [Marcia Cash](#)
Subject: Re: DiscoverText concern
Date: Thursday, August 10, 2017 9:21:26 AM
Attachments: [image.png](#)

In cluster 102 and 103 I do see a discrepancy. Mark, can you look at this?

On Thu, Aug 10, 2017 at 10:18 AM, Shulman, Stu <stu@texifter.com> wrote:
I see 10 in cluster 74 of June 13.

Inline image 1



On Thu, Aug 10, 2017 at 10:10 AM, Bowman, Randal <randal_bowman@ios.doi.gov> wrote:

In trying to get ready to count the number of near duplicate comments for June 13, June 28 and and July 20, I found some apparent mistakes in the system, in that in some cases the numbers shown for items in the cluster are lower than the number that are actually in the cluster.

It appears universally that for all clusters showing 1 item - and before I looked I had wondered how 1 would even show here - actually have 2 items.

In addition, for June 13 - cluster 74 - says 10, are 11, but for clusters 75-79, 10 = 10; clusters 102-7 say 6, have 7; cluster 112 - says 5, has 6; clusters 149-181 - say 3, have 4; cluster 301 and following says 2, have 3. And for July 20, starting at cluster 181, all that say 3 items have 4, all 2's have 3, 1's have 2. This appears to be a problem only for the low-number-of-items clusters.

I had planned to count these using the numbers shown - i.e. if clusters 200-300 were all 3s, and 93 of them opposed the review, multiply to get the number of opposed, do same with support and uncodable, then go to next 100.

But if we can't rely on those numbers I am a loss on how to get to the final count. Is there something in the system that will allow us to count the numbers of comments in the lower-end clusters? Or can it be re-set to show accurately how many items are in each cluster? Only those 3 dates are involved; there were few enough near dups in the other that I have

| already coded them, but there are too many to code for these 3.

--

Dr. Stuart W. Shulman
Founder and CEO, Texifter
Cell: 413-992-8513
LinkedIn: <http://www.linkedin.com/in/stuartwshulman>

--

Dr. Stuart W. Shulman
Founder and CEO, Texifter
Cell: 413-992-8513
LinkedIn: <http://www.linkedin.com/in/stuartwshulman>

Search and Browse Project Cluster

Through June 13 / Cluster_74 ([Project Cluster](#))

Showing 1 to 10 of 10 total



Add all to Bucket: [New](#) | [Existing](#)

Add only s



[DOI-2017-00002 2017-05-27 02-00-40_](#)

Public lands are our national treasure.

DOI-2018-05 00990