

From: [Shulman, Stu](#)
To: [Willey, Seth](#)
Cc: [Kasdin, Alexandra](#); [Sarah Backsen](#)
Subject: Re: Question from FWS about DiscoverText
Date: Tuesday, June 14, 2016 3:24:35 PM

On it...

On Tuesday, June 14, 2016, Willey, Seth <seth_willey@fws.gov> wrote:

Hey Stu,

I just wanted to ask if you could put this towards the top of you list. We have a group of folks standing by and are hoping to sort this out sooner than later.

Thanks,
Seth

Seth L. Willey, *Acting* Chief of Staff
Office of the Regional Director
Mountain-Prairie Region, USFWS
Seth_Willey@fws.gov
303-236-4257

ON DETAIL - May 31st through June 24th, I'll be on a detail as acting Chief of Staff for the Regional Director's Office. Sarah Backsen will be acting for me during this period (sarah_backsen@fws.gov, 303-236-4388).

On Tue, Jun 14, 2016 at 10:20 AM, Kasdin, Alexandra <alexandra_kasdin@fws.gov> wrote:

Hi Stu,

My name is Alex Kasdin and I work with Seth Willey at the U.S. Fish and Wildlife Service. I am currently managing a team of coders analyzing the public comments we received on our proposal to delist the grizzly bear. We are using DiscoverText and it has been enormously helpful. I have a technical question about coding vs. annotating in DiscoverText.

We have already sorted through most of our comments and coded them as substantive and non-substantive. I then split the dataset to create a new dataset with only the substantive comments. So now we are at the stage where we want to apply topic codes to these substantive comments. Our preferred method of doing this is *annotating* the comments so we can have a spreadsheet at the end which pulls out only the substantive lines of each comment and sorts these lines by topic. For example, if we have a three sentence substantive comment and one sentence is non-substantive, one sentence relates to habitat degradation, and one sentence relates to post-delisting management, I'd want this comment to be split into two rows in the final Excel spreadsheet we export from DiscoverText. One row would have the sentence related to habitat degradation and a note that this sentence received the habitat degradation code; the second row would have the sentence related to post-delisting management and a note that this sentence received the post-delisting management code. The non-substantive sentence wouldn't be in the spreadsheet. It seems like we can do this with the annotation function in DiscoverText. However, when we have a lot of people coding all at once, it seems like you need to give the entire comment a code to mark it as complete (i.e.

that's the only way DiscoverText understands that the comment has been analyzed by someone and ensures two people don't analyze the same comment). Do we need to apply codes for the entire comment in addition to our piece-meal annotations? Because it seems that if we just apply annotations to particular pieces of the comment, DiscoverText does not consider that comment "analyzed/coded." I don't want to "double-dip" the analysis and create issues for the Excel file we download at the end so I want to make sure I am not confusing things by giving comments codes *and* annotations (I want to ensure the Excel file we download at the end will have a separate row for each line of the comment we gave a different annotation).

We are hoping to start the annotation process this afternoon so hope you can help us as soon as possible. I apologize for the short notice. It is easier to reach me on my cell today: (609) 610-2060.

Thanks!

Alex Kasdin
Princeton University SINSI Fellow
Ecological Services, Mountain-Prairie Regional Office
134 Union Boulevard
Lakewood, CO 80228
Phone: (303) 236-4217

--

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