



# How to count documents that has at least one annotation in FileNet Image Services?

## Question & Answer

### Question

Are there any ways to count the distinctive number of documents that has at least (or more than) one annotation stored in FileNet Image Services (IS)?

### Answer

There are several ways to count that, depending on how annotations have been added to each documents.

Before describing those procedures, it is worthwhile to learn how annotation information is connected to the target document.

### Background

Using MKF\_tool and running "desc(ribe) annotations" command, one observes there are "doc\_id" and "annotation\_id" columns. The "doc\_id" is the id number of the target document that the annotation is attached. The "annot\_id" is a number that is an incremental and sequential number assigned as created. Once saved, this number is fixed.

If one deletes one of those annotations, it will be dropped and such missing number won't be filled. For instance, if you add 3 annotations, you will have annot\_id=1, 2 and 3. And after saving this annotation, when you delete the annotation 2 and save it, you will have two annotations with annot\_id=1 and 3.

de

So, taking the above in mind, one can see there are three procedures of counting the number, depending on how annotations are added/removed to documents.

**Procedure 1)** This applies for the case where there is only one annotation per document:

One uses "count annotations \*" command in MKF\_tool. One-on-one relationship between annotations and documents ensures the result simply mean the number of annotations.

**Procedure 2)** This applies for the case where there are more than one annotations per document, but no annotations are deleted:

In this case, we can expect every document that has annotation must have annot\_id=1. So, you can count the number of "distinctive" documents like,

```
<MKF_tool>count annotations * where annot_id = 1
```

See the example 1 below for illustration of this.

**Procedure 3)** This applies for the case where there are more than one annotations per documents, ever deleted more than one annotation:

This is the most generic case. One cannot actually count the number in question within the MKF\_tool, because there's no command like "SELECT DISTINCTIVE" in SQL. So, it is advised to output the document ids that have annotations (may duplicate as mentioned), and count only the unique numbers using other tools. The below shows one example of doing this:

- a) Use "hard(copy)" command to set the output into a text file.
- b) Run "se(lect) annotations \* showing doc\_id" command and output doc\_ids into the output file.
- c) Count the unique (or distinctive) doc\_ids from the output file.

An additional remark is that since this "annot\_id" key is not a 'unique key', it may take much time if one has many annotations in his or her database.

#### **<<Example 1>>**

In the IS library, 9 documents are stored. Only 3 documents have annotations. Documents with doc\_id 100007 and 100009 have multiple annotations.

```
<MKF_tool>se docs * showing doc_id
doc_id
-----
100000
100001
100002
100003
100004
100005
100006
100007
100009
<MKF_tool>count docs *
Total of 9 records counted
<MKF_tool>se annotations * showing doc_id annot_id
doc_id annot_id
```

```
-----  
100004 1  
100007 1  
100007 2  
100009 1  
100009 2  
100009 3  
<MKF_tool>count annotations *  
Total of 6 records counted  
<MKF_tool>se annotations * where annot_id = 1 showing doc_id annot_id  
doc_id annot_id  
-----
```

```
100004 1  
100007 1  
100009 1  
<MKF_tool>count annotations * where annot_id = 1  
Total of 3 records counted
```

## **<<Example 2>>**

In this case, you deleted the annotation with annot\_id=1 in the document with doc\_id=100009.

```
<MKF_tool>se annotations * showing doc_id annot_id  
doc_id annot_id  
-----
```

```
100004 1  
100007 1  
100007 2  
100009 2  
100009 3  
100009 4  
<MKF_tool>hard annot_se.txt
```

Outputting to file 'annot\_se.txt'

```
<MKF_tool>se annotations * showing doc_id  
doc_id  
-----
```

```
100004  
100007  
100007  
100009  
100009
```

100009

<MKF\_tool>hard

And the content of annot\_se.txt is like this:

Outputting to file 'annot\_se.txt'

<MKF\_tool>se annotations \* showing doc\_id

doc\_id

-----

100004

100007

100007

100009

100009

## Was this topic helpful?

Yes 

No 

## Document Information

### More support for:

FileNet Image Services (<https://www.ibm.com/mysupport/s/topic/0TO50000000IMo2GAG>)

### Software version:

4.1.2, 4.1.1, 4.0, 4.1

### Document number:

144299

### Modified date:

17 June 2018