

Chapter 14: Annotation Tasks

The manual Annotation Task is specifically designed for quality control and other special projects. In this task, a registrar codes data items while reading text-based reports and highlights supporting text for each data item. The task itself is flexible and can be adapted to many different projects. The data items collected and the modes by which the data are collected are part of the project definition.

These collection modes can be used in an Annotation Task: Single Annotation, Concurrent Annotation, and Sequential Annotation.

- Single Annotation – A single set of values is collected by a registrar. One person reviews the path report and codes values for the data items required by the project. The coder is asked to highlight supporting text for each data item.

Two values are collected for each data item in Concurrent & Sequential Annotation modes. The first is typically calculated by an automated SEER*DMS process. The second is manually entered by the coder completing the Annotation Task.

- Concurrent Annotation – The coder sees the calculated value as they code.
- Sequential Annotation – The coder does not see the calculated values until they review the path text, code and submit their values.

Example of an Annotation Project

As an example, the Annotation Task is being used to evaluate the Natural Language Processing algorithms developed as part of the NCI-DOE collaboration. The API developed by the DOE applies NLP algorithms to the pathology text. The API returns values for site, histology, behavior, laterality, and grade. A person codes the same fields manually via one of the modes listed above. If the manually coded values differ from the API then the registrar is asked to review and confirm the entered values. The registrar highlights supporting text which is used by the NCI, DOE, and IMS staff to improve the NLP algorithms. If the original coder's values differ with the API values, then the path report moves to an Annotation Review task. A second registrar adjudicates the differences.

Creating Annotation Tasks

Currently, IMS staff must configure each Annotation Project per NCI and registry specifications. IMS staff will also create the Annotation Tasks using the cohort defined in the project's protocol.

Completing an Annotation Task

Annotation Tasks will be assigned to each project participant. Click Annotation Tasks in your list of tasks on the SEER*DMS home page.

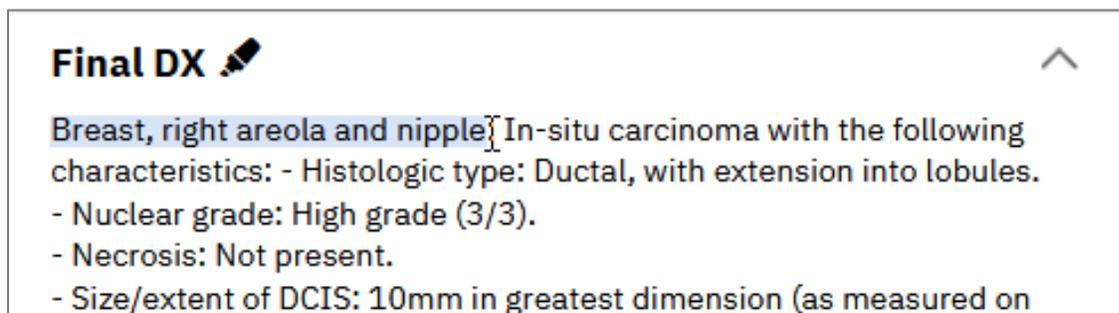
As a task is completed, the next task assigned to that person will automatically open (similar to a Path Screening Task). The interface is also similar to that of Path Screening, except the fields are entered in the left column, and text highlighting is supported.

The registrar should follow these steps when completing an Annotation Task:

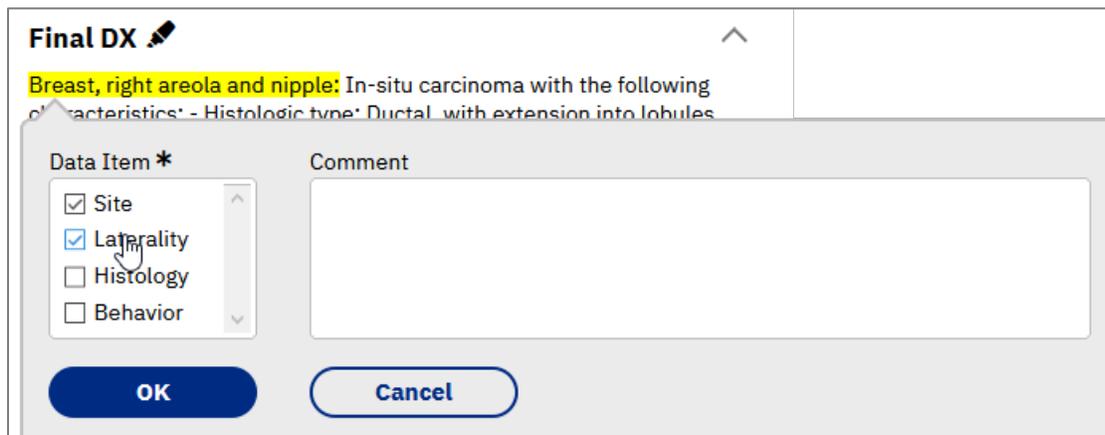
1. Read the Annotation Task instructions in the bottom left panel.
2. Review the pathology text and code the data fields.
3. **Single Annotation** (Registrar Only)
 - a. Highlight supporting text for each field for which there is supporting text.
 - b. Enter a comment if there is no text to support the value of a data item; this may be necessary if you need to enter unknown for a data item.
 - c. Click Complete
4. **Concurrent Annotation:**
 - a. Highlight supporting text for any field where your entry differs from the API value. You have the option of highlighting text for every field, but highlighting or a comment is required if your value differs from the API value.
 - b. Click Complete
5. **Sequential Annotation:**
 - a. Click Save. The API Values will be displayed and the Instructions Pane will list the fields that had different values. Review the pathology text and decide whether to:
 - i. Modify the codes that were originally assigned OR
 - ii. Keep the codes as originally assigned and highlight supporting text for each data item (required).
 - b. Click Submit.

Highlighting Supporting Text

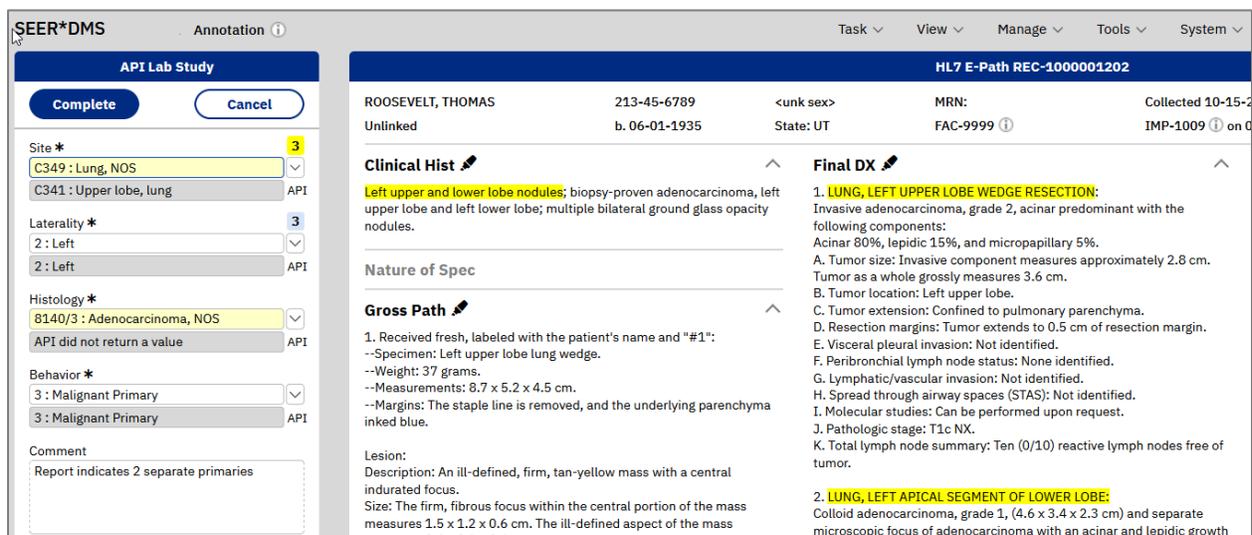
To highlight text using your mouse, position your cursor at the beginning of the text you want to highlight. Press and hold your primary mouse button (typically the left-button). While holding the mouse button, drag the cursor to the end of the text and let go of the mouse button. The screen shot below shows text being highlighted in the Final DX field.



You will be prompted to select data items when you release the mouse. You may select one or more data items for the same text. In this example, the same phrase is used as supporting text for site and laterality.



You may highlight multiple phrases for the same data item(s). This screen shot shows 3 separate phrases that were highlighted for site and laterality. The number of highlighted phrases is shown for each data item.



Annotation Review Tasks

A second manual task is created to adjudicate differences between two sets of values. Consider the annotation project to evaluate the path coding API. If the coder enters a value that differs from the API value then an Annotation Review task is created. This task is assigned to a second coder; it would typically be assigned to a senior registrar for adjudication.

The adjudicator reviews data items for fields that the API returned a value, but that value differs from the original annotator's values. In the Annotation Review task, the senior coder sets each field to a value that they believe is correct. In the Annotation Review task, the senior coder should:

- Review each field. If the coder's value is correct then leave the value as is.
- Review the supporting text for each field.
 - If the appropriate text is highlighted then leave it as is.
 - If you believe that additional text should be highlighted then highlight the phrase and associate it with the field.
 - You can also remove highlighting, if you do not think it is supportive for a data item. To remove highlighting, click the text that is highlighted and then uncheck the data item associated with that text.