



## The “Barebones” Tutorial

In this tutorial, you will learn how to use CATMA's core functionality:

1. Getting started	🐼 Login as a guest	3–4
2. Loading a text	🐼 Your point of departure: <i>Manage Resources</i>	5–8
3. <i>Annotate</i> a text	🐼 This is it: markup a text using tags!	9–18
4. <i>Analyze</i> your text	🐼 Query your text and annotations	19–28
5. <i>Visualize</i> results	🐼 Take a look at the big picture	29–33
6. Answers		34–41

This tutorial includes a sequence of sample tasks through which you can explore CATMA's core functionalities in hands-on mode. Answers and results to the tasks can be found at the end of the tutorial—but don't look them up too quickly, give it a try first!

A broader range of CATMA's functions is described in our separate “Expert” Tutorial.

## 1. Getting started

For the purpose of this tutorial we will use a guest login.

The advantage: Several preparatory steps which you would usually have to deal with yourself have already been carried out for you.

CATMA will open with Edgar Allan Poe's *The Tell-Tale Heart*. A Tag Library as well as an Annotation Collection have been created for you as well. What is all that? You'll find out throughout the next sections.

The sample tasks in this tutorial require [this preloaded material](#).

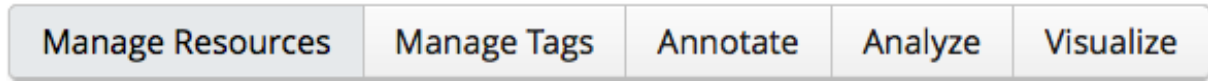
Please note: Your work in a guest account will not be saved—you need a personal CATMA account for that. (We'll show you how to create one in the “Expert” Tutorial.)

→ To login as a guest user, now click on the CATMA icon, answer the question that pops up, and you are ready to go!

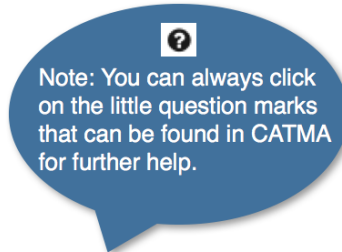


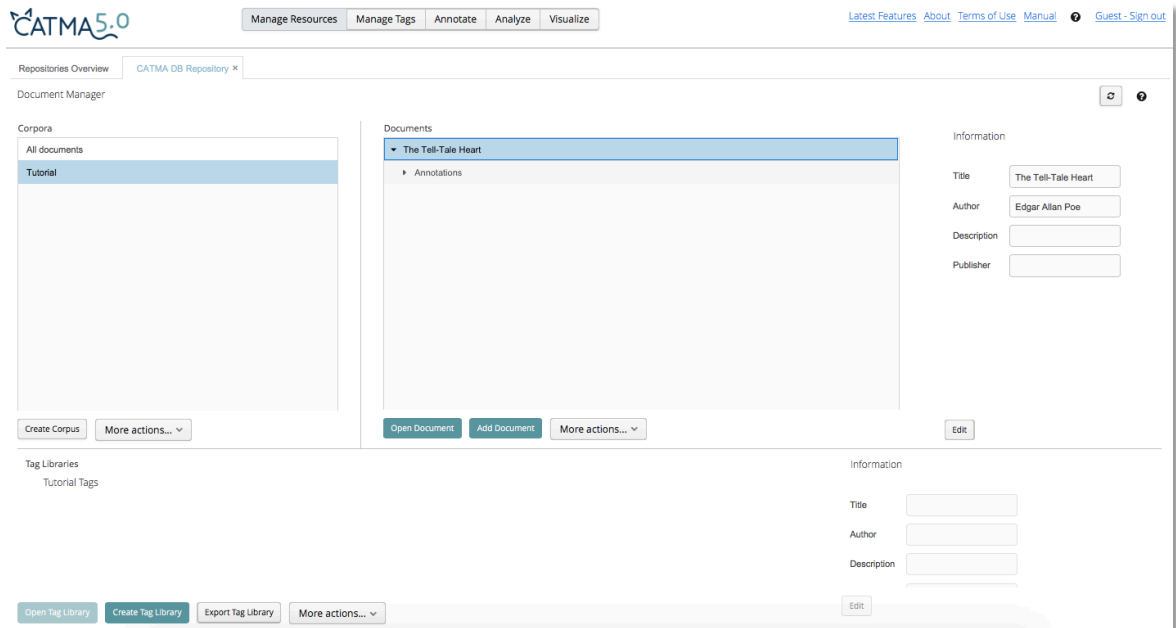
## 2. Loading a text

In CATMA, there are five modules, which you can use by clicking on the respective tab. You can open several tabs in each module, like you do in your web browser. Our point of departure is the ***Manage Resources*** tab on the left:



CATMA's Modules





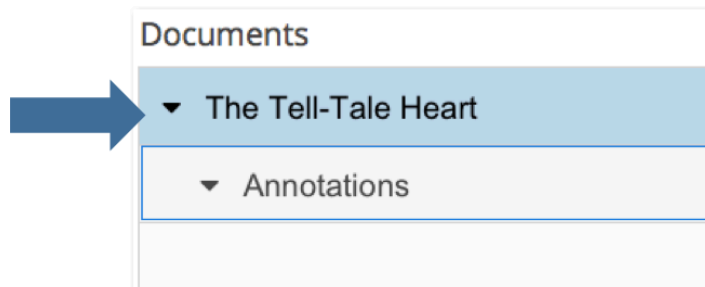
## Manage Resources window in CATMA

CATMA “Barebones” Tutorial—page 6

The ***Manage Resources*** window has three side-by-side panels on top plus a bottom panel underneath.

- ❖ In the middle panel titled *Documents* you will see the preloaded text file *The Tell-Tale Heart*. We'll focus on this panel for now and only mention the others briefly—their functionality is described in detail in the “Expert” Tutorial.
- ❖ The left panel titled *Corpora* is where you can group a number of documents together into a text corpus.
- ❖ The right panel titled *Information* displays details about what's currently selected, for example the specifics of a text document. You select items by highlighting them.
- ❖ The bottom panel *Tag Libraries* is used to organize your annotation terms, the so-called tags. We'll deal with that in the “Expert” Tutorial.

In the *Documents* panel, single-click on the document “The Tell-Tale Heart”—it will now be highlighted.



To the left of the file name there’s a little black triangle: if you click on it, a drop-down list will open. It shows you the *Annotation* collection associated with this text document. Usually you have to create these collections yourself; however, for the purpose of this tutorial we preloaded one for you.



### 3. *Annotate a text*

But wait a moment—what exactly is contained in the *Annotations* file? Let's clarify some basic terminology before we proceed.

#### **What's in an *Annotations* file?**

When you annotate a text in CATMA you highlight a piece of text and then use a tag to add a descriptive label to that text segment. Tags add up to **Markup**: information that enriches the original text by way of annotation. The *Annotations* file is where this text specific information is stored.

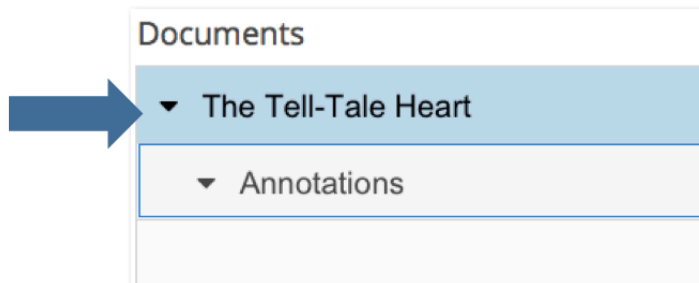
CATMA enables you to perform almost any type and granularity of annotation, from a single character to an entire text document. Some annotators need only three tags to markup an entire text; others will develop a complex system of descriptive labels that might contain a couple of hundred tags. In addition, it obviously makes sense to reuse and share tags rather than to start all over again with every new text.

This is why CATMA invites you to systematize, export, import, and share tags independently from the primary text documents. This is done as follows:

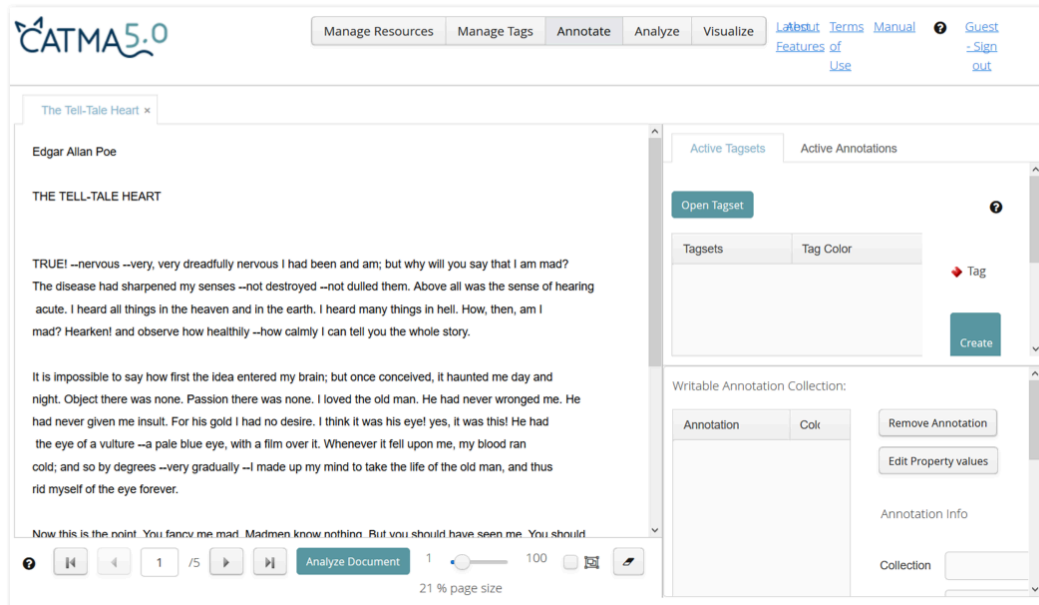
- 🔖 To begin with, you can group related tags in **Tagsets**
- 🔖 Every Tagset then automatically forms part of a **Tag Library** which may contain as many tags and Tagsets as you need.

Furthermore, this hierarchical structure enables you to extend and modify your tags at any time—you don't have to worry about some missing tags when you start. You can always add them later and even create them “on the fly.”

So, now let's get started for real: in the *Documents* panel, **double-click** on the document *The Tell-Tale Heart*.

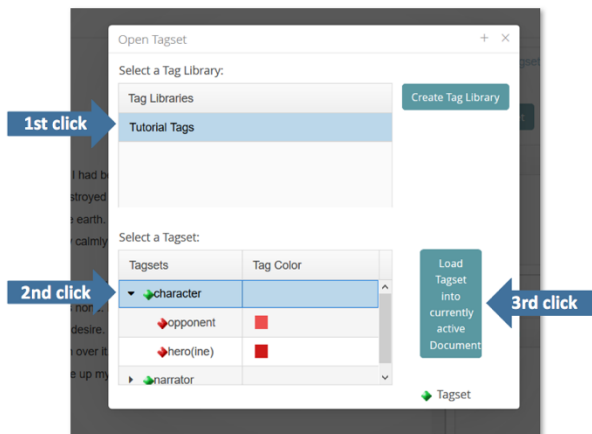


Your text will open in the *Annotate* window:



**Annotate window in CATMA displaying a source text**

- 🐱 In the **Annotate** window, click on **Open Tagset**.
- 🐱 A dialogue window will pop up. Click on **Tutorial Tags**; two preloaded Tagsets appear at the bottom: **character** and **narrator**.
- 🐱 Open the **character** Tagset by clicking the little triangle in front: you'll see the individual tags contained therein (the *opponent* and *hero(ine)* tags).

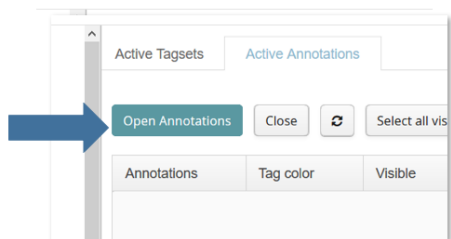


🐱 Highlight this Tagset and load it into the document. First, click on the label character, then click on the big “Load Tagset into currently active Document” button on the right.

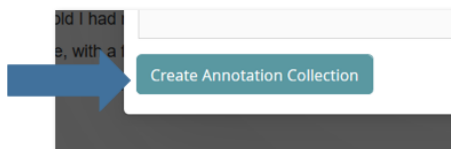
CATMA will take you back to the **Annotate** window.

Before you can begin annotating the text with the tags that have now been associated with the document, you must do one more thing: instruct CATMA where to store your annotations. This is done in three steps:

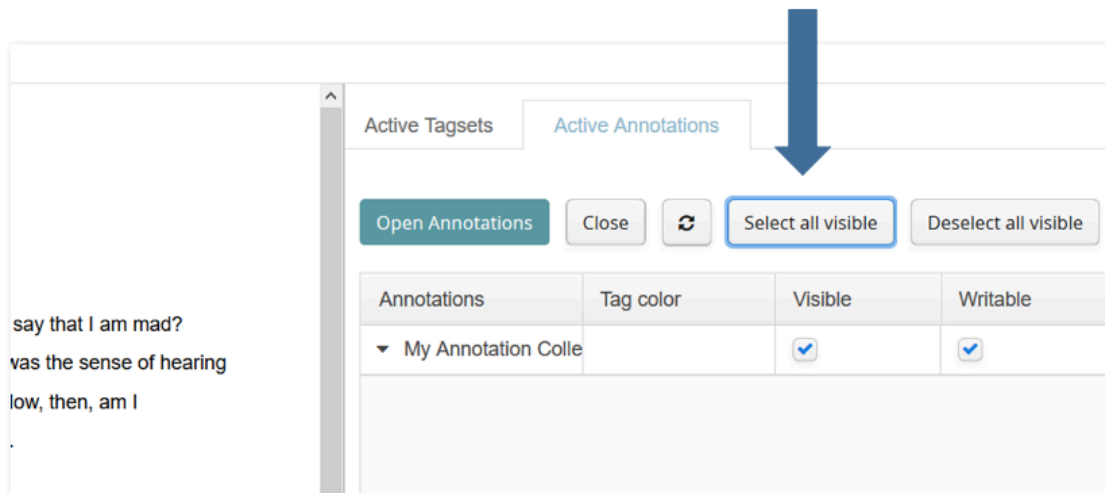
- Switch from the “Active Tags” to the “Active Annotations” tab, then click on “Open Annotations”



- A new dialogue will pop up. Click on “Create Annotation Collection.”



- 🐱 Now, give your Collection a name, e.g. “My Annotation Collection.” In the dialogue window, confirm with OK. You’ll be taken back to the Annotate window and see this—make sure to click on “Select all visible.”



We haven't started annotating yet, so there is not a lot to see. Here's a preview of how things might look like once you've added a lot of annotations to the text, which we will do in a moment:

The screenshot displays the CATMA 5.0 web application interface. On the left, a document titled "The Tell-Tale Heart" is shown with various text segments highlighted in yellow and blue. A blue box labeled "the document you're working with" points to this text. On the right, a panel titled "Active Tagsets" shows a list of tags with their corresponding colors. A blue box labeled "the set of tags available to annotate the text" points to this panel. Below the tagset panel, a "Writable Annotation Collection" is shown, with a specific tag "questionable factual claims" highlighted in blue. A blue box labeled "the individual tag that you chose to annotate a specific piece of text – your markup is stored in Annotations" points to this tag. The interface also includes navigation tabs at the top: "Manage Resources", "Manage Tags", "Annotate", "Analyze", and "Visualize".

the document you're working with

the set of tags available to annotate the text

the individual tag that you chose to annotate a specific piece of text – your markup is stored in Annotations



So, let's start with some hands-on annotation exercise. This is how it's done in two easy steps:

The screenshot displays the CATMA interface. On the left, a text document is open with the instruction "Select the relevant piece of text" and a blue arrow pointing to a selected line of text: "een and am; but why will you say that I am mad?". On the right, the "Open Tagset" panel is visible, showing a list of tagsets and their corresponding color buttons. A blue arrow points from the text selection to the "style" tagset's color button.

Tagsets	Tag Color
opponent	
hero(ine)	
narrator	
attitude	
questionable morals	
questionable factual claim	
style	
address to the reader	
exclamation	
repetition	

Assign a tag by clicking on the color button

## Sample Task 1 for *The Tell-Tale Heart*

1. In **Annotate**, close-read Poe’s *The Tell-Tale Heart* and pay attention to the phenomena for which we have created the *character* and *narrator* tags. Once you identify the relevant phenomenon, select the corresponding text string and click on the color next to the tag in the “Active Tagsets” tab (you may need to unfold the respective Tagset by clicking on the arrow symbol next to it). You’ve now officially set your first annotation, congrats!

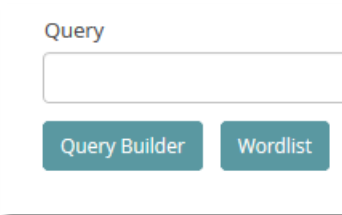
## 4. *Analyze* your text & annotations

Whenever we engage with texts in order to explore and discuss their meaning, two operations go hand in hand: analysis and interpretation. Their mutual dependency results in what is known as the *hermeneutic circle*.

CATMA supports this approach by combining text analysis and text annotation in a cyclical, iterative workflow. Once you have uploaded your digital text to CATMA, you will typically combine your work in *Annotate*, as described above, with using the merits of the *Analyze* module. Whereas you pursue a close reading in *Annotate*, *Analyze* supports a more distant reading approach. Here you can browse your texts without reading them yourself.

You reach *Analyze* through *Annotate* by clicking on the “Analyze Document” button below the displayed source text. Whenever you access *Analyze* through an active document, a new tab is created for each new document. Likewise, different analyses of a single document open new tabs (numbered serially), so you don’t lose your previous results. You can also open a new tab by clicking the “+” button in the upper right corner.

In *Analyze*, you see two buttons in the top left section:



The image shows a screenshot of a web interface. At the top, the word 'Query' is displayed in a small, grey font. Below it is a white rectangular input field with a thin grey border. Underneath the input field are two teal-colored buttons with white text. The left button is labeled 'Query Builder' and the right button is labeled 'Wordlist'.

## Create a Wordlist

CATMA automatically creates a Wordlist of the respective text if you click on the “Wordlist” button. The Wordlist can be sorted in ascending or descending frequency by clicking on the “Frequency” field once or twice.

## Sample Task 2 for *The Tell-Tale Heart*

2. What is the most common content word (= a word with “more” semantic meaning than function words such as articles, pronouns etc.)?

### **Kwic visualization (Kwic = Keyword in context)**

Select a word in the Wordlist that you are particularly interested in (e.g. “heart” in our sample). Tick the “Visible in Kwic” box for this word and look at the Kwic display on the right-hand side of the screen. If necessary, change the scope of the displayed context using the slider at the bottom (setting to 1–30 tokens possible).

If you double-click on one of the keywords in the Kwic visualization, CATMA will jump to the selected keyword’s full-text location back in *Annotate*.

## Query Builder

You are free to enter your query directly into the Query box, but many queries can also be generated with the Query Builder, which is much more convenient, especially if you aren't familiar with the query language. (By the way: After executing a query with the Query Builder, you can see the respective query phrase in the query mask and thus learn it as you go along.)

You use the Query Builder by clicking on the corresponding button. In the opening window, you decide what you want to search for:

- ☒ by word or phrase
- ☐ by grade of similarity
- ☐ by Tag
- ☐ by collocation
- ☐ by frequency

Choose an instance and click “Next.” If you want to search for the word “heart,” for example, use the “by word or phrase” query option. Note that queries are case sensitive.

Click “Show in preview” to test whether your query generates a result. If it does, click “Finish” at the bottom right. Your query is now being executed.

You can also use the Query Builder to create complex queries (i.e. to refine, combine, and exclude search results). Just tick “continue to build a complex query” in the query mask to do so.

### Sample Tasks 3–6 for *The Tell-Tale Heart*

3. How many words occur more than five times? Use the query function “by frequency” for this.
4. How many words occur between five and fifteen times? Which physical senses seem to be predominant in the story?

5. How many words have a 70% similarity to “heart?” Use the “by grade of similarity” query option. Open a new *Analyze* tab (by clicking the “+ New Query” button on the upper right) and increase the similarity to 80%. Do it again with 85%.
6. How often does the word “mad” appear near the word “me” (within a span of ten words). Use the “collocation” function.

## Analyze your annotations

**Analyze** also offers you to analyze your own annotations that you set in **Annotate** before. Use the Query Builder to search for a specific tag (“by Tag”), or directly type the query: tag = “tag name”. The analysis of your annotated text(s) depends on your research interest and your annotations.

After searching specific annotations, you can examine the “Result by Phrase” (as you did in the other searches) or as well the “Result by Tag” (tabs in the wordlist).







## Sample Task 7 for *The Tell-Tale Heart*

7. Search for the different “narrator” annotations you set in task 1. Can you tell something interesting about their frequency or co-occurrences? What can you infer from your annotations concerning the reliability of the narrator?

## Annotate your text(s) in *Analyze*

Since the hermeneutic workflow in CATMA is circular, **Analyze** can also help you annotate your text. Repetitive manual annotation of the same word across an entire text, for example, can be avoided by running a query first and then annotating all search results at once. Execute the following steps for this:

-  Run a query as described above.
-  Activate the Kwic display and highlight the relevant Kwic lines.

-  Click on “Annotate selected Results” in the bottom-right corner: A box displaying your Tagset and its tags will open.
-  Select and draw the relevant tag onto the Kwic lines (drag & drop).

All the selected search results have now been annotated with the same tag.

### Sample Tasks 8 & 9 for *The Tell-Tale Heart*

8. Select all occurrences of “I,” “me,” “my,” “My,” “myself,” and “mine” to be visible in Kwic and let *Analyze* annotate them with the “hero(ine)” tag. How many annotations does it set?
9. Do the equivalent thing with the “opponent” tag. Which words or phrases qualify for being annotated as “opponent”?

## Some General considerations before you start annotating and analyzing your texts

CATMA's digital infrastructure forces you to clarify your research problems. You really need to make sure what you want to find out about the source text(s) you are investigating, and thus reflect on your own work, your research questions, and your methods.

Take some time to identify the questions and problems that you would like to investigate upfront. For example, you might be interested in analyzing the presence of the characters in the text, their behavior, their character traits or the like. If you are more interested in geographical features, you could analyze types of geographic entities—countries, cities, waters, islands, but also special places, etc. If you are interested in topics, identify relevant topics, and so on.

If you look at your list of questions, they usually fall into three different categories:

- a. Questions that can be answered directly by investigating the actual words and phrases that constitute the text. For example: Which character in *The Tell-Tale Heart* is mentioned *verbatim* more often—"I" or "the man?" A wordlist analysis with *Analyze* already answers this one: "I" is mentioned more often than "the man."

- b. Questions that require you to pre-interpret textual elements in order to find out what they mean. For example: Characters in *The Tell-Tale Heart* are not identified by names, but by pronouns like “he” or function words like “officers.” But does every “he” point to the old man or maybe to one of the officers? You can make those instances of “he” that refer to the old man explicit by annotating them with CATMA tags.
- c. Questions that require you to combine both strategies—direct textual analysis and pre-analytical text interpretation. So, if you really want to find out every mentioning of the narrator in the story, you can use *Analyze* to do a combined search for literal occurrences of “I” plus a search for all occurrences of “me” etc. that you annotated as an indirect reference to the narrator.

Type (b) and (c) questions are the ones where tags come into play—that is, interpretive annotation labels that you attach to a word in order to make its implicit meaning or function explicit to CATMA.

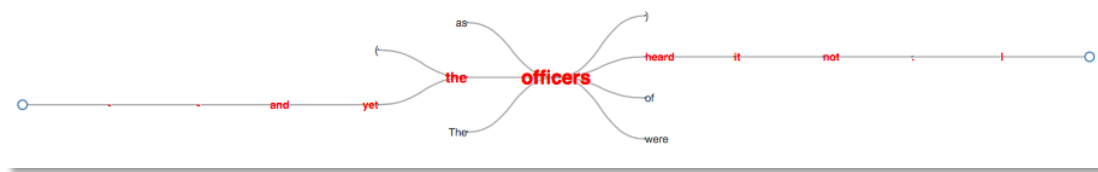
## 5. *Visualize results*

Next to annotating and analyzing your text(s), you will want to get an image of them—literally. You either want to get a quick overview of phenomena in your text or corpus or you want to visually display results of your work in *Annotate* and *Analyze*. This is where *Visualize* comes into play. Currently, CATMA offers two different forms of text visualization: a DoubleTree and a Distribution Graph. There will be more visualization functions in the future.

### DoubleTree visualization

The DoubleTree is a very compact and more interactive way to visualize your Keywords in context (Kwic). Choose a word you are interested in from the wordlist, highlight it, and click on the DoubleTree button below it. You can expand the DoubleTree to both sides.





**DoubleTree of “officers”**

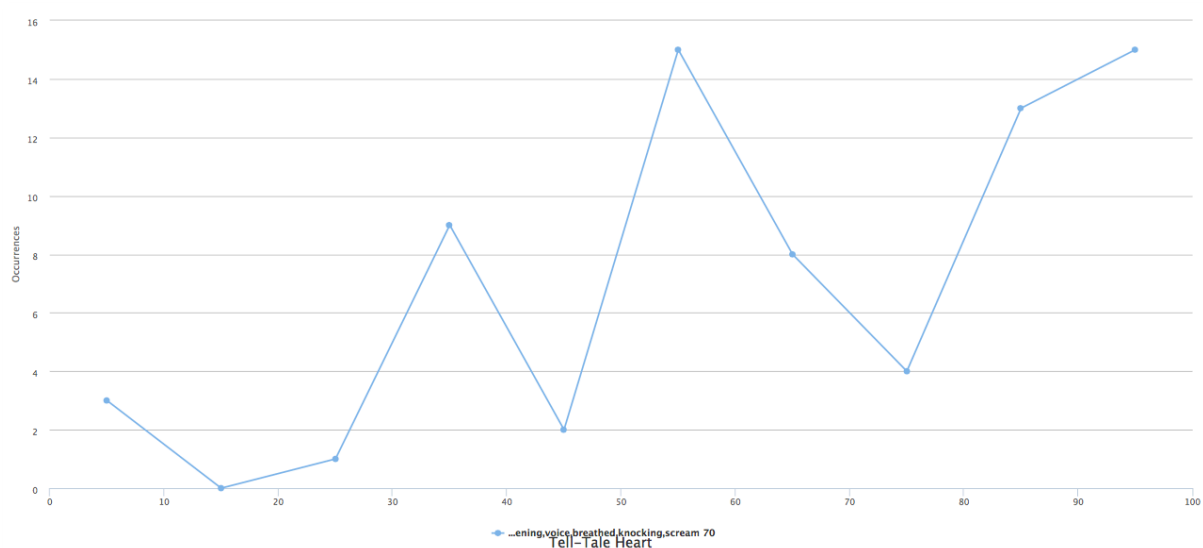
### Sample Tasks for *The Tell-Tale Heart*

10. Start from the wordlist in *Analyze*. Look for the word “man” in the list, highlight it and click on the DoubleTree button. Click on the words in the DoubleTree visualization and try to figure out how the DoubleTree works. Can you tell something about the character?
11. Try the same with “I.”

## Distribution Graph

The Distribution Graph allows you to visualize and compare the occurrence of single or multiple phenomena and their distribution and potential development in your text—or even to compare these phenomena throughout several texts in a corpus (in this case multiple diagrams will be displayed next to each other in *Visualize*). You can either combine several words or tags in one graph or show one graph for each word/tag in a diagram. To combine words or tags into one graph, select a group of words or tags that will be interesting to you in the word or tag list of *Analyze*. This is done by selecting the appropriate lines (by holding the command key). If you now click the distribution graph button, the words will be displayed in a single distribution graph. You can also click on the graphs to get back to the respective text passages in *Annotate*.





**Distribution Graph for sounds and hearing in *The Tell-Tale Heart***



## Sample Tasks for *The Tell-Tale Heart*

12. Build a complex query and search for the “hero(ine)” tag as well as the “opponent” tag in *Analyze*. (This will only work if you executed tasks 8 and 9, or set the tags manually.) Highlight the “hero(ine)” tag in the Tag list and click on the distribution button below. What can you tell from the diagram?
13. Go back to *Analyze* and select the tag “opponent.” If you now click the distribution graph button again, it will be displayed in the same diagram. Is there something interesting that can be seen from the representation of the two tags?

## 6. Answers for the sample tasks

### Question 2:

The word “was” appears 33 times (or, in case you don’t consider pronouns function words, “I” appears 124 times).

### Question 3:

There are 79 words that occur more than five times.

### Question 4:

There are 73 words that occur between five and 15 times. Especially sound but also vision are the predominant physical senses.

### Question 5:

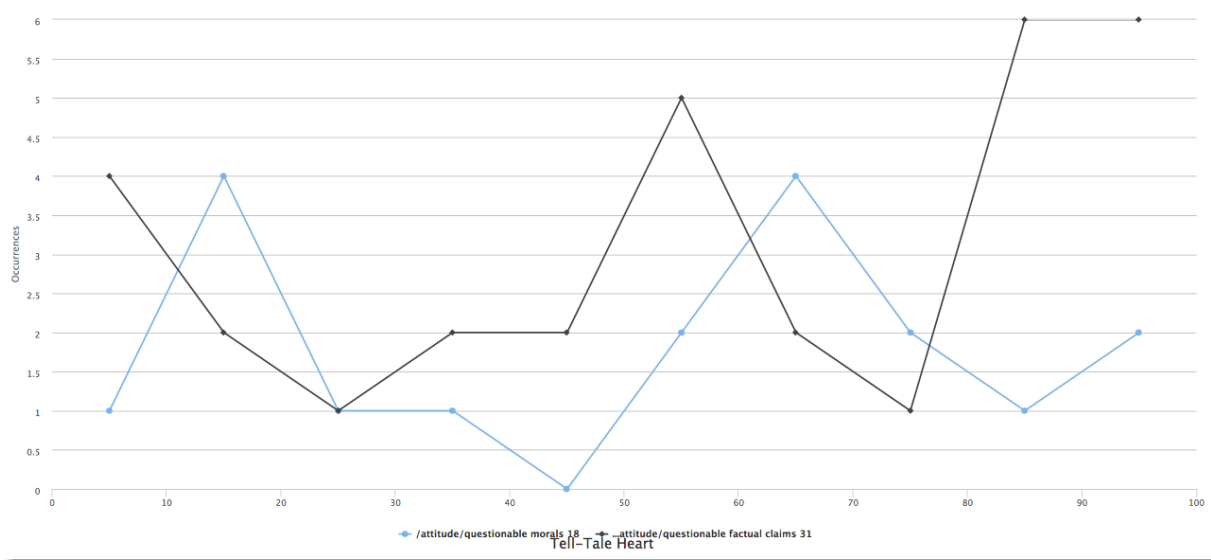
Eleven words have a 70% similarity to “heart,” eight words with 80%. With 85%, only “hear,” “HEART,” “tear,” “heartly,” “earth,” and “heart” are listed.

### Question 6:

The words “mad” and “me” co-occur three times within a span of ten words.

### Question 7:

The first distribution chart shows the frequencies of the narrator’s attitude (which is “questionable morals” and “questionable factual claims”) throughout *The Tell-Tale Heart*. The most striking thing might be the increasing amount of questionable factual claims at the end of the story, while the narrator’s morals seem more questionable at the beginning and the beginning of the last third of the story.



**Chart 1: narrator's attitude in *The Tell-Tale Heart***

Chart no. 2 displays co-occurrences of these attitude tags with annotations of the narrator's style (which is addresses to the reader, exclamations, parallelisms, and repetitions). One could observe that stylistic devices and a questionable attitude tend to peak at the same points of the story. An augmented emotional involvement of the narrator could be inferred by this, and also a trace of his unreliability on the level of form. However, you should note that especially the attitude tags tend to be set interpretively rather than in a way that generates intersubjectively reproducible results.

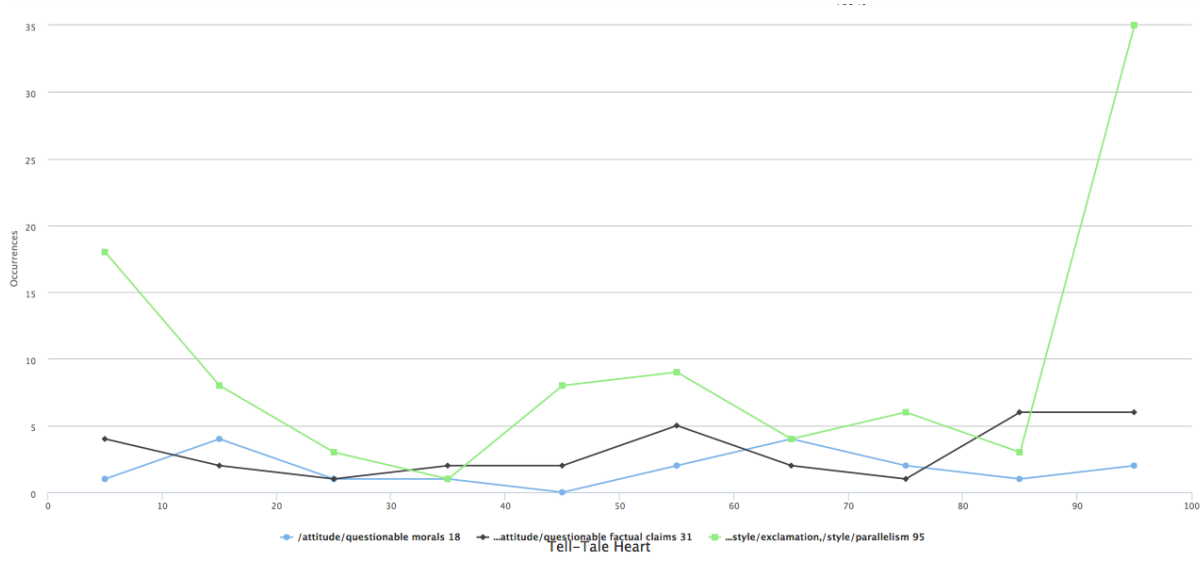


Chart 2: narrator's attitude and style in *The Tell-Tale Heart*

### Question 8:

There are 178 hero(ine) annotations set.

### Question 9:

Words/phrases like “old man,” “eye,” “heart,” “corpse,” “he,” “him,” “his,” “body,” “police,” “officers,” “they,” “them,” “their,” etc. qualify to be annotated with the “opponent” tag.

### Question 10:

The DoubleTree allows you to explore a word in its various contexts. For example, the man is obviously old, scared and he dies. This may lead to conclusions about the character of the man as well as his feelings to and/or his behavior towards the narrator.

### Question 11:

In the case of “I,” a clear picture cannot be seen at first glance—there are significantly more and clearly more variable results.

### Question 12:

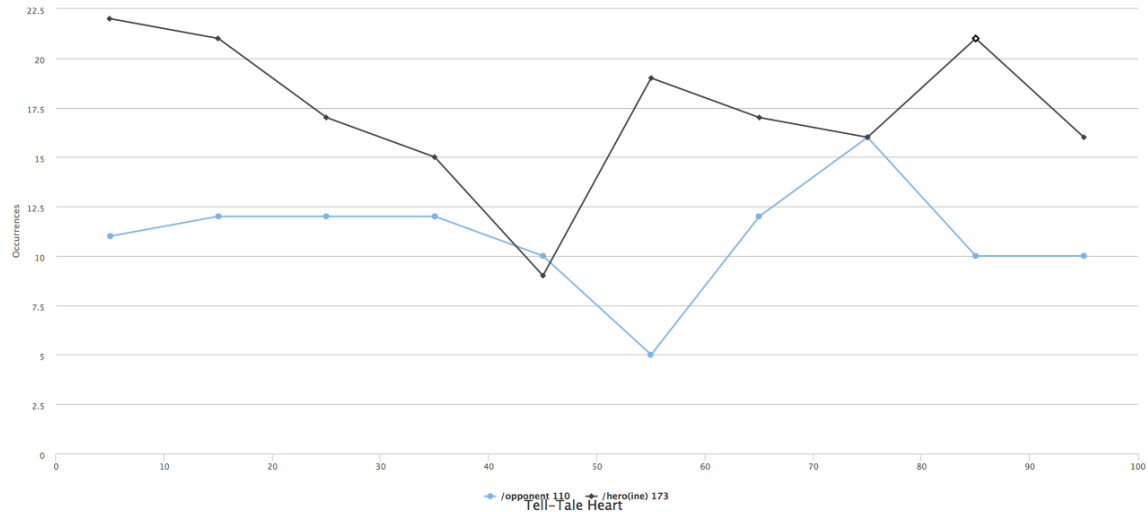
The hero is predominant throughout the story. Since the hero is the homodiegetic narrator of the story, too, this doesn’t come as a surprise. Interesting might be the low point in the middle of the text. By clicking on this point, we find out that this is exactly the moment shortly before the killing, a moment where the narrating I feels overpowered by the “evil eye” of the old man.

### Question 13:

The distribution graph makes clear where there are two important moments of decision in the story: when the narrator decides to kill the old man and when he meets the police



officers. In these moments, his “opponents” overpower or nearly overpower his presence in the text. Further hypotheses of interpretation could be built on this.



Distribution Graph for “hero(ine)” and “opponent” annotations in *The Tell-Tale Heart*