



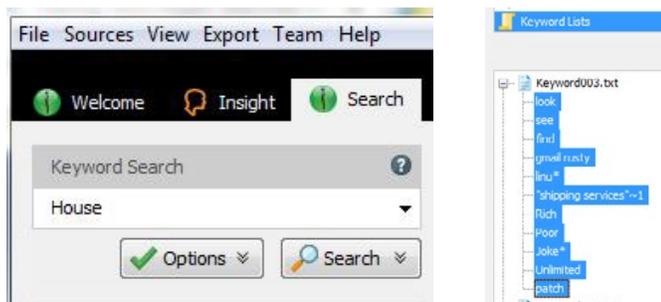
Regex Searches

A Better Understanding of Searching in Intella

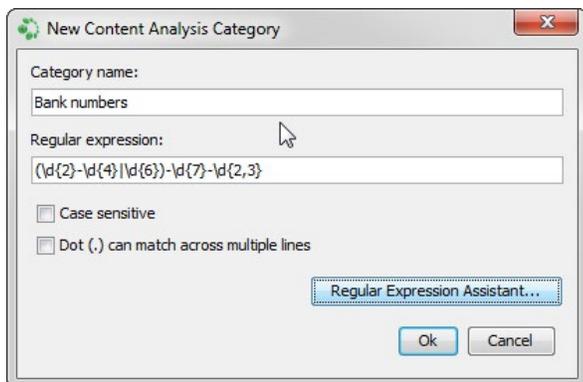
Searching in Intella and Intella Connect

Two methods of searching in Intella and Connect:

- Using the normal keyword search box and keyword lists.



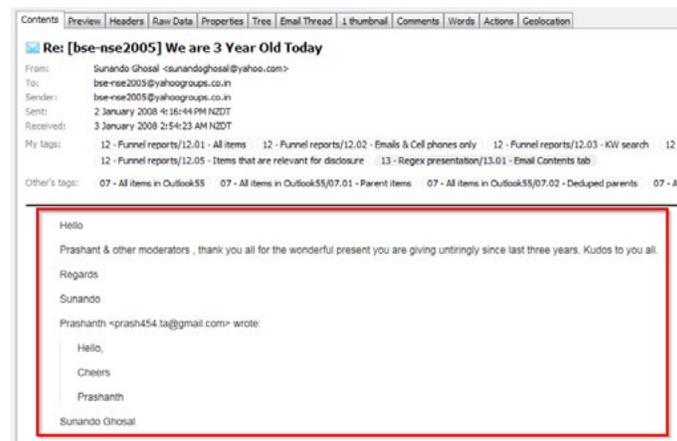
- Using the more advanced Regular Expressions feature in Content analysis



- In regards to special characters, when indexing, most of these characters are typically filtered out and will never make it into the index. The rules for handling specific characters depend on the context in which they occur. For instance, punctuation characters like dots (.) or dashes (-) are significant within numbers, email addresses or host names and are shown. But, they will be ignored (i.e. interpreted as whitespaces) when they are between regular words. In the latter case, escaping them in the query will not make them searchable.

Searching over extracted text with Regular Expressions

During processing Intella also stores a copy of all of the text contained in documents and items. This is the information that you see in the Contents tab when reviewing an item, and this text can also be searched.



The Index in Intella and Intella Connect

Disadvantages/Limitations:

- When text or words are added to the index, they are tokenised. This means that the search must match the word for it to be responsive. For example, if a dataset contained emails to, and from support@vound-software.com, searching for this email address will show hits. But, if you searched for support@vound-software, you would not get any hits as the .com part of the email address was omitted.

To work around those issues, you can use wild cards with your searches. In this case searching for support@vound-software* would return the items containing the full address.

In addition, the special characters seen in the Contents tab remain in this text store. But, this text can not be searched with the normal KW searching features. It needs to be searched using the Regular Expression feature under Contents Analysis.

Limitations:

Although the regular expression search shows respondent items, a limitation is that in the current version is that no hit highlighting is shown in the preview window. This will be improved in a future version.

