RealDecoy

EndecaCommunity.com

Everything you want to know about Endeca   Search   or   Ask a Question

Endeca Community Solutions

Endeca Search Features

Dear Endeca Community user,

On July 31, 2015 we closed EndecaCommunity.com, and asked users to instead take advantage of the collaboration tools and communities now provided by Oracle.

Why we did this? Because EndecaCommunity.com has fulfilled its purpose.

We created EndecaCommunity.com to fill a temporary void, to support teams working with Endeca technology in both Commerce and Data Insight capacities. The intention was to provide continuity through the transition period that followed Oracle's acquisition of Endeca, until Oracle had ramped up its own Endeca support community.

Now that Oracle's Endeca communities and formal user groups are well established, it's better for you as an Endeca user to find all the support you need from OAUG Special Interest Group and Oracle sources, rather than split your time having to cross-post to multiple disparate sites to find the answers you need.

We recommend joining the [Endeca Community on LinkedIn](#) or the [OAUG Endeca Special Interest Group](#). More support for Endeca can be found on the Oracle community website at [Technical Questions](#) and [Endeca Information Discovery](#).

On the following pages you will find the most frequently asked questions about Endeca Search Features on EndecaCommunity.com and the best solutions as determined by members of the Community.

Thank you for taking advantage of EndecaCommunity.com and for trusting RealDecoy as your Endeca authority.

Sincerely,

Richard Isaac
President & CEO, RealDecoy

Table of Contents

Endeca Search Features

Endeca Phrases not working

Q: I am currently working on Endeca 6.2.2. In the case of automatic phrasing, I have added an entry in the AppConfig.xml to take the phrases from Workbench.

      `<property name="webStudioMaintainedFile4" value="phrases.xml"/>`

Upon adding phrases in the workbench, the phrases xml is not being updated in the pipeline even after performing the baseline update. I tried using the other approach to import phrases using the developer studio. This is updating the phrases xml in the pipeline. But the search is not working as expected after baseline update. Suppose if red white wine is added as a phrase entry, the search for red white wine is giving results whereas the search for "red white wine" is not giving any results. Spelling correction and 'Did you mean' are working as expected.

Please let me know in case I missed any configuration settings.

Also please let me know if phrases testing can be verified using the Endeca JSP Reference Implementation while performing search under any match mode.

A: I was able to solve this issue. Just to summarize with my findings:

Case 1: Added Phrase does not exist in any of the Endeca records. Initially I was giving some phrases which do not exist in any of the records. Since these phrases do not exist, Endeca search tries to break these words and get results instead of displaying "no results found".

So in this case even adding the parameters Ntpc=1 & Ntpr=1 does not matter. The results vary depending on the match modes used.

Case 2: Added phrase exists in the Endeca records. In this case I have added the phrase which does exist in some of the Endeca records and have seen that Endeca does consider the phrase and retrieve the results. For Endeca to consider the phrase we need to add the parameters Ntpc=1 & Ntpr=1. The result is same even with different match modes.

Wildcard search in Russian, or OLT and wildcard search?

Q: We have Russian customers. Per the manuals, Endeca uses Oracle Language Technology (OLT) analysis as the default for Russian, but it doesn't support wildcard search. Also, type ahead is working through wildcard. How do I resolve it? How can I switch into Latin-1 (I think I will find static dictionaries in Russian)? Has anyone had such a problem with another language?

A: As per my investigation, an issue existed with the wild-carding for Dimension search with OLT. For implementing wildcard search you should create a separate index for it. MDEX version - 6.4.0

How to get thesaurus list in classic Endeca with PageBuilder

Q: In OEC 3.x, I am able to query thesaurus entries and view all items in the list using the following soapUI request:

    <soap:Envelope xmlns:soap=http://schemas.xmlsoap.org/soap/envelope/>
    <soap:Header/>
    <soap:Body>
      <config-service:configTransaction xmlns:config-
    service=http://www.endeca.com/MDEX/config/services/config/2011>
        <config-service:getConfigDocuments>
        <mdex:configDocumentName
    xmlns:mdex="http://www.endeca.com/XQuery/mdex/2008">thesaurus</mdex:
    configDocumentName>
        <config-service:getConfigDocuments>
        <config-service:configTransaction>
    </soap:Body>
     </soap:Envelope>

Is there a way to perform this request or something similar to retrieve the thesaurus entries in classic Endeca with PageBuilder?

A: Using the **emgr_ update** script you are able to retrieve all web studio configurations, including thesaurus entries. You need to pass **get_ ws_ settings** as argument for the parameter **--action**

The sample query would look like:

```
<pre><code>emgr_ update --action get_ ws_ settings --host &lt;hostname&gt; --prefix
&lt;prefix&gt; --dir &lt;directory&gt; --app_name &lt;application name&gt;

</code></pre>
```

In the directory specified, an xml file called **< prefix >.thesaurus.xml** will be created with the extracted entries.

Programmatically disabling autocorrect

Q: I am using Mdex 6.2.0.
Dgraph received these parameters: <arg>--spl</arg> <arg>--dym</arg>
I need to perform a request without any word correction. I tried to set
NavigationCommand.AlternativePhrasingMode to Disabled but it was not successful.
Is there a way to do this?

A: Auto-correct is configured using a Dgraph flag and thus cannot be disabled
programmatically. However, you can tell when auto-correction is done and handle the
scenario accordingly. Here is a snippet for displaying an auto-corrected search term:

```
// Get the Dictionary of ESearchReport objects
IDictionary recSrchRpts = nav.ESearchReports;
// Get the user's search key
String searchKey = Request.QueryString["Ntk"];
if (searchKey != null) {
   if (recSrchRpts.Contains(searchKey)) {
      // Get the first Search Report object
      IDictionaryEnumerator ide = recSrchRpts.GetEnumerator();
      ide.MoveNext();
      ESearchReport searchReport = (ESearchReport)ide.Value;
      // Get the List of auto-correct objects
      IList autoCorrectList = searchReport.AutoSuggestions;
      // If the list contains Auto Suggestion objects,
      // print the value of the first corrected term
      if (autoCorrectList.Count > 0) {
         // Get the Auto Suggestion object
         ESearchAutoSuggestion autoSug = (ESearchAutoSuggestion)autoCorrectList[0];
        // Display autocorrect message
         %&gt;Corrected to &lt;%= autoSug.Terms %&gt;
      }
   }
}
```
The "**searchReport.AutoSuggestions**" property is what you'd be interested in.

RealDecoy

Beyond synonyms

Q: Assume my index has only two records, and each of the record has one field called 'title'. The values for title are 'cougar town' and 'cougartown' respectively. In my search term, I want to use either 'cougar town' or 'cougartown' and I expect both records to be returned. For this particular example, synonyms will solve it. But what if there are thousands of terms like this (eg, 'cougar town' and 'cougartown', the only difference is that one with space and the other does not), I don't want to have to create thousands of synonyms. Is there a more elegant approach to deal with this problem?

A: You don't want to do this at query time, you want to fix this in your etl. It actually sounds like you need to do some data cleansing.

My first recommendation is clean the data in the source system. My next suggestion is create a new property or dimension "corrected/suggested title". Then pull the entire list of titles into a tool like "open refine" (formerly google refine) or some tool that has a feature for word distance http://en.m.wikipedia.org/wiki/Levenshtein_distance (cloverETL has one). This will enable you to map dirty titles to clean titles. After you have created a cleaned list (that has mappings to dirty titles), join the clean titles to each record. You will get some false positives but certainly if you get it right 80% of the time you are doing well.

I do not in any way suggest you create a synonym file with that many entries. Your mdex will not perform as well as when it has a smaller synonym file.
My main points: don't use synonyms, clean the data in the source system, don't do it at query time, if you have to, do it in etl, don't write it yourself, make a new field leave the dirty field intact.

Precedence Rules in Endeca

Q: Please let me know the advantages of using precedence rules in Endeca with the help of some business use cases. While using Endeca Experience Manager, I feel that precedence functionality can be controlled by cartridges.

A: Yes, precedence rules can be 'done' by cartridges. Precedence rules pre-date guided nav cartridges and is really nice in setting up a global set of rules (when to show, what dimensions.) You could create the same thing in Experience Manager but it would be a lot more work. They do the same thing. I'd say Experience Manager is more flexible, at the cost of being more labor intensive (you must configure a template for each navigation you wish to hide/show). Precedence rules are less flexible at the cost of being less labor intensive (make a central change that effects how all dimensions are displayed).

Endeca Auto Suggestion

Q: I want to reconfigure my auto suggest feature such that for each user the most frequently searched item would appear on the top of the list rather than just according to the relevance ranking algorithm. In other words, 3-5 frequently searched items should appear on the list first and then according to the ranking algorithm.

A: I suggest you to limit 'Frequent words' dimension values, because if this is going to be increased in the future, you will face performance issues in dimension search.