

A study of Search Partner Query Strings

Conducted by PyBiz, Inc to evaluate viability
Of search aggregation concepts
as requested by Daniel at PointCom.com

Partners

- ❑ [Summary](#)
- ❑ [Partner1.Kanoodle.com](#)
- ❑ [FindWhat.com](#)
- ❑ [Meta.7search.com](#)
- ❑ [Searchpartners.about.com](#)
- ❑ [Ah-Ha](#)

Summary Notes:

When we first started working with PointCom one of the fundamental rules where that the partners where adhering to a common XML DTD and query formulation for interaction.

As it turns out from reviewing the data coming back from the various partners this rule is not true and as a result it is going to make the query and aggregation process more difficult than originally anticipated.

This document summarizes what I have found as a result of preparing to implement the what we talked about today. I actually modified the queries enough to be confident I understood how each of the different systems worked and then took the results straight out of IE5 from executing these queries.

What this boils down to is that a small adapter must be written for each partner that deviates from the standard. This adapter will convert their inbound data to something that matches the expected norm.

This is needed because while XDisect handles the variability fine the XSL used to render the screens for the end users would have a more difficult time dealing with the variability and it would also make future expansion more complex.

General query result structure

```
<result>
  <record>
    <title> CDATA STRING</title>
    <url> CDATA STRING </url>
    <description> CDATA STRING</description>
    <clickurl> CDATA STRING</clickurl>
    <bidprice>CDATA STRING</bidprice>
  </record>
</result>
```

- ❑ There may be one to N records declared

Issues

- ❑ Each query string varies by partner. It has to be constructed dynamically from the form query parameters for each partner.
- ❑ Some partners are using the term <bid> rather than <bidprice>
- ❑ Some partners are using <site> rather than <record> tags for each item returned.
- ❑ Kanoodle is missing the outer <result> tag.

Kanoodle.com

<http://partner1.kanoodle.com/cgi-bin/partner.cgi?query=modems&format=xml1&id=10255842>

- format = xml1 = tells kanoodle to return as a xml
- id = pointcom account id
- query=the string of the query to be sent

Issues & Notes

- Kanoodle does not return valid XML. They are missing a valid out enclosing TAG. This will break in our parser.

Sample

```
<?xml version="1.0" encoding="windows-1252"?>
<record>
  <title><![CDATA[modems]]></title>
  <url><![CDATA[http://www.modems.buy-pricescan.com/goto/]]></url>
  <description>
    <![CDATA[Find the best selection of products, prices, and reviews with PriceSCAN.com, your
    unbiased shopping guide. Don't buy it without a priceSCAN!]]>
  </description>
  <bidprice>
    <![CDATA[<a
    ref=http://www.kanoodle.com/client_services/listings/index.cool?word=modems>(Cost to
    advertiser: $0.06) </a> ]]]>
  </bidprice>
  <clickurl>
    <![CDATA[http://www.kanoodle.com/clickthrough.cool?position=7001&tid=bpnobinobnnpbnlbg
    &bid=0.06&unq=772863943&eid=1&id=8308231&query=modems&clickid=10255842]]>
  </clickurl>
</record>
```

findwhat.com

http://findwhat.com/bin/findwhat.dll?getresults&fl=1&base=0&mt=modem&dc=3&aff_id=42394&ip_addr=64.81.35.161

http://findwhat.com/bin/findwhat.dll?getresults&fl=1&base=0&mt=modem&dc=300&aff_id=42394&ip_addr=64.81.35.161

- mt="the term we are searching for"
- dc="the maximum number of query results"
- fl= do not know
- base= do not know
- aff_id= identifier of the pointcom account

Sample

```
<?xml version="1.0" encoding="windows-1252"?>
<result status="OK" records="207" first="1" last="3"
  searchRequest="modem"
  responseTime="15ms">
  <record>
    <title><![CDATA[TDK Semiconductor]]></title>
    <url>http://www.tdkmodems.com</url>
    <description>
      <![CDATA[TDK Semiconductor Corporation is focused on the communications IC
        market. Specific data transmission market segments served include high-speed
        networking, remote data transaction terminals and satellite TV set-top boxes.]]>
    </description>
    <bidprice>0.15</bidprice>
    <clickurl>
      http://www.findwhat.com/bin/findwhat.dll?clickthrough&amp;c=11027&amp;w=74349&amp;t=6D6F64656D&amp;aff\_id=42394&amp;srch\_id=6D657F607070212B48DD47
    </clickurl>
  </record>
</result>
```

meta.7search.com

<http://meta.7search.com/scripts/meta.asp?qu=modem&language=1&affiliate=36680&r=3&pn=1>

- qu = query string being searched for
- language = don't know
- affiliate = account id for pointcom
- r = do not know
- pn = do not know

Basic Structure

```
<results>
  <site>
    <index>unique identifier</index>
    <url>CDATA STRING</url>
    <name>CDATA STRING</name>
    <bid> CDATA STRING</bid>
    <httplink>CDATA STRING</httplink>
  </site>
</results>
```

Issues & Notes

- there is an array of 1..N sites
- Error - use of the <site> tag rather than the <record> tag
- note the use of <httplink> rather than <clickurl>

Sample

```
<?xml version="1.0" standalone="yes"?>
<RESULTS>
  <SITE INDEX="1">
    <URL>
      <![CDATA[http://7search.com/scripts/redirect.asp?urlid=13361064&affiliateid=36680&url=http%3A%2F%2Fwww%2Erjross%2Ecom&keyword=modem&s=pls&u=http%3A%2F%2Fpointcom%2Ecom]]>
    </URL>
    <NAME>
      <![CDATA[RJ Ross Discount Computer Hardware Upgrades]]>
    </NAME>
    <DESCRIPTION>
      <![CDATA[Discount Computer Hardware Upgrades and Computer Accessories including intel processors, asus motherboards, memory chips, creative sound, ati video, maxtor hard drives, etc.]]>
    </DESCRIPTION>
    <BID>
      <![CDATA[$0.06]]>
    </BID>
    <HTTPLINK>
      <![CDATA[http://www.rjross.com]]>
    </HTTPLINK>
  </SITE>
</RESULTS>
```

searchpartners.com

<http://searchpartners.about.com/taxis/metaresults?site=pointcom&terms=search&abc=3>

Issues

The use of name rather than title

Error - use of bid rather than bidprice

Error - use of httplink rather than clickurl

Warning - It appears that searchpartners is using url the way others are using clickurl and that they are using httplink in the way others are using url.

Sample

```
<?xml version="1.0" standalone="yes"?>
<RESULTS>
  <SITE INDEX="1">
    <URL>
      <![CDATA[http://7search.com/scripts/redirect.asp?urlid=13361064&affiliateid=36680
        &url=http%3A%2F%2Fwww%2Erjross%2Ecom&keyword=modem&s=pls&u=http%3
        A%2F%2Fpointcom%2Ecom]]>
    </URL>
    <NAME>
      <![CDATA[RJ Ross Discount Computer Hardware Upgrades]]>
    </NAME>
    <DESCRIPTION>
      <![CDATA[Discount Computer Hardware Upgrades and Computer Accessories
        including intel processors, asus motherboards, memory chips, creative sound, ati
        video, maxtor hard drives, etc.]]>
    </DESCRIPTION>
    <BID><![CDATA[$0.06]]></BID>
    <HTTPLINK><![CDATA[http://www.rjross.com]]></HTTPLINK>
  </SITE>
</RESULTS>
```

Ah-ha.com

Query

<http://partner.ah-ha.com/9999.asp?Q=modem&CID=9999&IP=207.142.38.5&CK=182937465>.

- ❑ See.: http://partner.ah-ha.com/xml_dtd.asp for the official documentation.
- ❑ q = query string
- ❑ qt = query type C-complete, L-logo links only, F-fast
- ❑ ft = format type f1- xxx f2 - xxx
- ❑ s = size of response
- ❑ cid = Pointcom customer ID =
- ❑ o = Offset, begin the next search after offset items,
- ❑ ck = unique cookie to identify the users.

Notes

In General the Ah-ha seems to be one of the better done designs. And they do not use the CDATA where it is unneeded which will make parsing easier.

- ❑ One to N <link> elements may occur

Issues

- ❑ The <ahha_search> tag encloses the <results> tag. Extra layers.
- ❑ The <link> tag is used instead of the <record> tag.
- ❑ <domain> tag was added by do not know what to do with it.
- ❑ The use of <bid> rather than <bidprice>
- ❑ <logolink> add but do not know what to do with it
- ❑ <nk> added but do not know what to do with it
- ❑ <site_url> used rather than <url>
- ❑ <redirect_url> used rather than <clickurl>
- ❑ We will have no ck, cookie parameter readily available since we are not tracking users and for ah-ha this is a mandatory query parm.

Basic Structure

```
<ahha_search>
  <results>
    <link>
      <type> a code</type>
      <redirect_url>a URI</redirect_url>
      <site_url> a URI </site_url>
      <domain> unkown</domain>
      <title> a string</title>
      <description> a string</description>
      <logolink> a URI</logolink>
      <bid>a number</bid>
      <nk> unkown </nk>
    </link>
  </results>
</ahha_search>
```

Sample

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<AHHA_SEARCH>
  <COPYRIGHT>
    <META NAME = "COPYRIGHT" CONTENT="COPYRIGHT © 1999-2000 ah-ha.com
      Inc. All rights reserved."/>
  </COPYRIGHT>
  <BANNER>
  </BANNER>
  <RESULTS>
    <LINK>
      <TYPE>S</TYPE>
      <REDIRECT_URL>
        http://partner.ah-ha.com/tools/redirect\_nk.asp?cid=9999&amp;b=11790&amp;bid=0%2E2&amp;k
          eyword=beetle&amp;nk=2155769&amp;t=S&amp;r=1
      </REDIRECT_URL>
      <SITE_URL>http://www.buy-carsdirect.com/beetle?source=1030</SITE_URL>
      <DOMAIN></DOMAIN>
      <TITLE>Beetle</TITLE>
      <DESCRIPTION>
        Come to CarsDirect.com for great prices on virtually any new car. Then order
        online and even apply instantly for a lease or loan.
      </DESCRIPTION>
      <LOGOLINK>
        http://a1356.g.akamai.net/f/1356/2288/7d/images.ah-ha.com/1790%2Eqif
      </LOGOLINK>
      <BID>0</BID>
      <NK>2155769</NK>
    </LINK>
  </RESULTS>
</AHHA_SEARCH>
```

Sample Input Form

```
<table cellspacing="0" cellpadding="0" border="0" align="left">
  <tr><form METHOD="GET" NAME="q" ONSUBMIT="submitForm()">
    <td width="30"></td>
    <td align="left"><input TYPE="TEXT" NAME="q" VALUE="beetle"></td>
    <td width="5"></td>
    <td><input TYPE="SUBMIT" VALUE="Submit"></td>
  </tr>
</table><br><br>
<table cellspacing="0" cellpadding="0" border="0" valign="top" align="left">
  <tr>
    <td width="30"></td>
    <td align="left"><b>Query Type:</b><br>
      <input TYPE="radio" VALUE="c" name="qt" checked>&nbsp;C (<u>C</u>omplete)<br>
      <input TYPE="radio" VALUE="l" name="qt">&nbsp;L (<u>L</u>ogolinks only)<br>
      <input TYPE="radio" VALUE="f" name="qt">&nbsp;F (<u>F</u>ast only)<br>
    </td>
    <td width="20"></td>
    <td align="left" valign="top"><b>Format Type:</b><br>
      <input TYPE="radio" VALUE="f1" name="ft" checked>&nbsp;F1 (normal)<br>
      <input TYPE="radio" VALUE="f2" name="ft">&nbsp;F2 (separated)<br>
    </td>
    <td width="20"></td>
    <td align="left" valign="top"><b>Rev share:</b><br>
      <input TYPE="radio" VALUE="0" name="r" checked>&nbsp;0 (false)<br>
      <input TYPE="radio" VALUE="1" name="r">&nbsp;1 (true)<br>
    </td>
    <td width="20"></td>
    <td align="left" valign="top"><b>Size:</b><br>
      <input TYPE="text" VALUE="20" name="s" size="1" maxlength="2"><br>
    </td>
    <td width="20"></td>
  </tr>
```



```
<td align="left" valign="top"><b>CID:</b><br>
  <input TYPE="text" VALUE="9999" name="cid" size="2" maxlength="4"><br>
</td>
</tr>
</table></form>
```