

WorldCat Mashathon

- Google AppEngine, Python, JSON

Etienne Posthumus, DPO, TU Delft Library

24-11-2009

Aan u de keuze: NL/EN

- From Twitter:

“Listening to Dutch speakers in English can be cringeworthy, but this Englishman speaking Dutch is hard too, esp since it sounds like German”

Identity withheld to protect the innocent...

Worldcat Mashathon at IISG

IISG = International Institute of Social History www.iisg.nl

Photo: Alice Sneary

- May 2009 in Amsterdam
- Big success, fun, educational.



WorldCat Search API

<http://www.worldcat.org/devnet/wiki/SearchAPIDetails>

- You need an Access key – OCLC members
- Use OpenSearch or SRU
- Get back XML (MARC)



WorldCat Search API

<http://www.worldcat.org/devnet/wiki/SearchAPIDetails>

- You need an Access key – OCLC members
- Use OpenSearch or SRU
- Get back XML (MARC)

```
<entry>
  <author>
    <name>Sarala. Ray, Lisa, 1972- Abraham, John, 1972- Biswas, Seema. Manorma. Hamilton, David (David
      C.) Mehta, Deepa, 1949- </name>
    </author>
    <title>Water</title>
    <link href="http://worldcat.org/oclc/70243935"/>
    <id>http://worldcat.org/oclc/70243935</id>
    <updated>2009-11-19T16:02:25Z</updated>
    <summary>A Hindu widow in 1938 could throw herself on her husband's funeral pyre, marry her husband
      's younger brother or live in poverty and self-denial from then on. When eight-year-old Chuyia is w
      idowed, she is sent to a home in the Holy City of Varanasi where Hindu widows live in penitence. He
      r feisty presence affects the other residents, forcing each to confront their faith and society's p
      rejudices, especially a young woman who has fallen in love with a follower of Gandhi.</summary>
    <oclcterms:recordIdentifier>70243935</oclcterms:recordIdentifier>
  </entry>
```

An example OpenSearch query Atom results:

```
<entry>
  <author>
    <name>Fredericks, Carrie.</name>
    </author>
    <title>Water</title>
    <link href="http://worldcat.org/oclc/71350506"/>
    <id>http://worldcat.org/oclc/71350506</id>
    <updated>2009-10-29T09:54:57Z</updated>
    <summary>Discusses using water as a source of energy.</summary>
    <dc:identifier>urn:ISBN:9780737735932</dc:identifier>
    <dc:identifier>urn:ISBN:0737735937</dc:identifier>
    <oclcterms:recordIdentifier>71350506</oclcterms:recordIdentifier>
  </entry>
```

WorldCat Search API

Results from a SRU search for 'water' with MARC output

```
<recordData>
<record xmlns="http://www.loc.gov/MARC21/slim">
  <leader>00000 a2200000 4500</leader>
  <controlfield tag="001">70243935</controlfield>
  <controlfield tag="008">060705s2006 cau hin d</controlfield>
  <datafield ind1="1" ind2=" " tag="041">
    <subfield code="a">hin</subfield>
    <subfield code="b">eng</subfield>
    <subfield code="b">spa</subfield>
    <subfield code="h">hin</subfield>
  </datafield>
  <datafield ind1="0" ind2=" " tag="130">
    <subfield code="a">Water (Motion picture : 2005)</subfield>
  </datafield>
  <datafield ind1="1" ind2="0" tag="245">
    <subfield code="a">Water</subfield>
    <subfield code="h">[videorecording] /</subfield>
    <subfield code="c">Fox Searchlight Pictures presents ; Mongrel Media presents in association
with Telefilm Canada, Noble Nomad Pictures, Echo Lake Productions ; a David Hamilton production ; a
Deepa Mehta film ; produced by David Hamilton ; written and directed by Deepa Mehta.</subfield>
  </datafield>
  <datafield ind1=" " ind2=" " tag="260">
    <subfield code="a">Beverly Hills, CA :</subfield>
    <subfield code="b">20th Century Fox Home Entertainment,</subfield>
    <subfield code="c">[2006]</subfield>
  </datafield>
  <datafield ind1=" " ind2=" " tag="300">
    <subfield code="a">1 videodisc (ca. 117 min.) :</subfield>
    <subfield code="b">sd., col. ;</subfield>
    <subfield code="c">4 3/4 in.</subfield>
  </datafield>
  <datafield ind1=" " ind2=" " tag="500">
    <subfield code="a">Originally released in 2005 as a motion picture.</subfield>
  </datafield>
  ...
```

MARC for the brave among us

WorldCat Search API

JSON anyone?

- When making quick & dirty Javascript hacks
- Very useful
- The engine behind many 'Ajax' web applications
- Simple, easy(ier)



WorldCat Search API

JSON example

```
{  
  586: "Genie Awards, 2006: Best performance by an actress in a leading role (Biswas); Best achievement in cinematography; Best achievement in music - original score."  
  521: "MPAA rating: PG-13."  
  300: "1 videodisc (ca. 117 min.) : sd., col. ; 4 3/4 in."  
  260: "Beverly Hills, CA : 20th Century Fox Home Entertainment, [2006]"  
  700: "Mehta, Deepa, 1949- aus drt"  
  710: "Twentieth Century Fox Home Entertainment, Inc."  
  245: "Water [videorecording] / Fox Searchlight Pictures presents ; Mongrel Media presents in association with Telefilm Canada, Noble Nomad Pictures, Echo Lake Productions ; a David Hamilton production ; a Deepa Mehta film ; produced by David Hamilton ; written and directed by Deepa Mehta."  
  001: "70243935"  
  130: "Water (Motion picture : 2005)"  
  546: "Hindi dialogue, with optional English or Spanish subtitles; closed-captioned."  
  500: "Special features: director Deepa Mehta commentary; "Behind-the-scenes" featurette; "The story behind the making of 'Water'" featurette."  
  650: "Motion pictures, Hindi."  
  520: "A Hindu widow in 1938 could throw herself on her husband's funeral pyre, marry her husband's younger brother or live in poverty and self-denial from then on. When eight-year-old Chuyia is widowed, she is sent to a home in the Holy City of Varanasi where Hindu widows live in penitence. Her feisty presence affects the other residents, forcing each to confront their faith and society's prejudices, especially a young woman who has fallen in love with a follower of Gandhi."  
  008: "060705s2006 cau hin d"  
  655: "Melodrama. gsafd"  
  041: "hin eng spa hin"  
  538: "DVD, region 1, NTSC, widescreen (2.35:1) presentation; Dolby Digital 5.1 surround."  
}
```



WorldCat Search API

But JSON isn't one of the formats supported by the WorldCat API. ☹️



WorldCat Search API

But JSON isn't one of the formats supported by the WorldCat API. ☹️

So let's hack one in...



When you have a hammer...



When you have a hammer...
everything looks like a nail.



When you have a hammer...
everything looks like a nail.

Python is the hammer in my toolbox



So we mash up:

- The WorldCat Search API
- Google AppEngine (the Python version)

using

- Python WorldCat Module (*Mark Matienzo NYPL, WorldCat Hackathon 2008*)
- PyMarc module (*Ed Summers LOC*)



Demo

Murphy says never to demo live code in a presentation...

fingers crossed!

Javascript code

```
$.ajax({  
  
    dataType: 'jsonp',  
  
    data: 'query=srw.' + $('#field').val() + ' = "' + $('#value').val() + '"',  
  
    jsonp: 'jsonp_callback',  
  
    url: 'http://code4bib.appspot.com/worldcat/',  
  
    success: function (output) {  
  
        $('#output').html('<h1>Results</h1>');  
  
        $.each(output, function() {  
  
            $('#output').append('<p>  
  
<a href="http://www.worldcat.org/oclc/' + this['001'] + '" target="_zarg">' +  
  
this[MARCSHOW] + '</a></p>');  
  
        });  
  
    });  
  
});
```

Python code

```
from worldcat.request.search import SRURequest

from worldcat.util.extract import extract_elements, pymarc_extract

from django.utils import simplejson

class MainHandler(webapp.RequestHandler):

    def get(self):

        query = self.request.get('query', 'srw.ti = "IISG"')

        jsonp_callback = self.request.get('jsonp_callback')

        maximumRecords = self.request.get('maximumRecords', 30)

        TUD_key='xxxxx'

        s = SRURequest(wskey=TUD_key, query=query, maximumRecords=maximumRecords)

        o = s.get_response()
```

Python code

```
self.response.headers['Content-Type'] = 'application/json'

buf = []

for r in pymarc_extract(o.data):

    rr = {}

    for f in r.get_fields():

        rr[f.tag] = f.format_field()

    buf.append(rr)

if jsonp_callback: self.response.out.write('%s(' % jsonp_callback)

self.response.out.write(simplejson.dumps(buf))

if jsonp_callback: self.response.out.write('')
```

Questions?

e.posthumus@tudelft.nl