

Pipes and Y! Query Language (YQL)

Jonathan Trevor (jtrevor@yahoo-inc.com)





How do you find an apartment *near* a park?



- Apartment listings
 - For each apartment:
 - Click on map link or enter an address into a browser
 - Check distance to a park on the map
- Tedious



- Data is available
 - Apartment <u>RSS</u> feeds, craigslist, backpage.com
 - Yahoo! Local API to find "things" like parks
- Can do it in about 50 lines of Perl code

```
#!/usr/bin/perl -w
use strict;
use LWP::Simple;
use XML::Simple;
...
```



- Basically combine feeds + web services
- Yet another custom mashup
 - HousingMaps, ChicagoCrime, ...

Would be nice if there was an easier way...



Pipes

```
grep -iv yahoo.com squid.log | sort |
uniq -c | sort -n > top_sources.txt
```

- Unix Pipes for the Web
- Build useful applications from simple primitives



Pipes

- A free service that lets you remix and create data mashups using a visual editor
- No need to host, we do it for you

Craigslist

Yahoo!



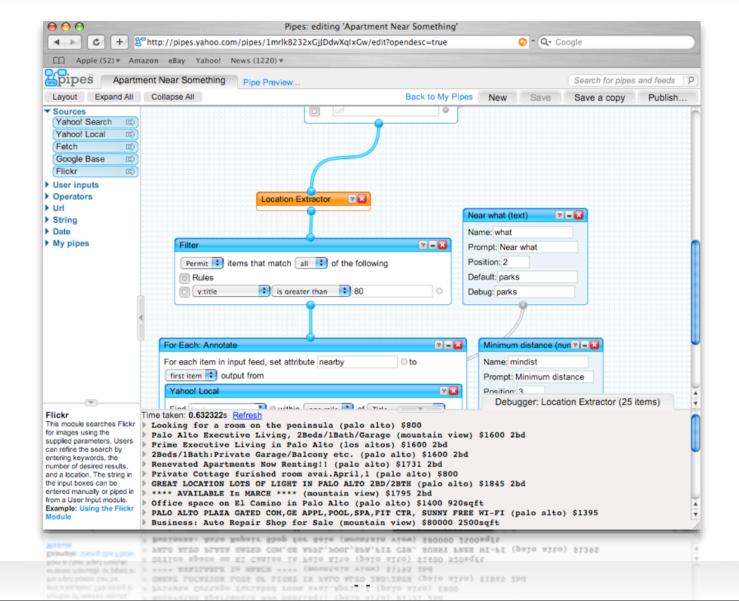
Pipes

- A free service that lets you remix and create data mashups using a visual editor
- No need to host, we do it for you





Demo





Piecing things together in the cloud



Any Input

Craigslist

Yahoo!

Google

Ebay

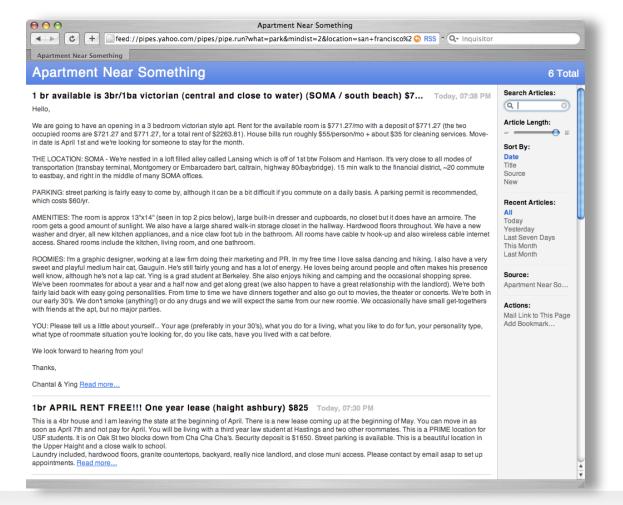
Your data here!







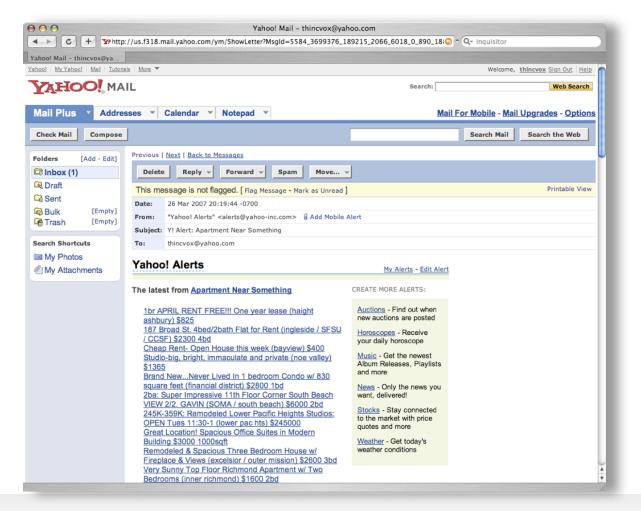






```
\Theta \Theta \Theta
                                                pipe.run
     {"count":6,"value":{"title":"Apartment Near Something","description":"Ever wanted to find an
     apartment close to a park, a school or Whole Foods? Now you
     can...","link":"http:\/\pipes.yahoo.com/pipes\/pipe.info?_id=1mrlkB232xGjJDdwXqIxGw","pubDate":"M
     on, 26 Mar 2007 19:58:58
     -0700", "generator": "http:///pipes.yahoo.com//pipes/", "callback": "", "items": [{"link": "http:///sfb
     ay.craigslist.org\/sfc\/roo\/301061840.html","nearby":{"link":"http:\/\local.yahoo.com\/details?id=
     21370452&stx=park&csz=San+Francisco+CA&ed=SGNhwa160SyE3VbtIOu6sVNhreovG.xE4iq9Wwpss_u1ys
     Nsnh6A8aFZFoTJ_.Fxoro7edIk","Categories":{"Category":[{"content":"Tourist
     Attractions", "id": "96926026"}, {"content": "Outdoor
     Parks","id":"96928354"}]},"BusinessClickUrl":"http:\/\www.sfgov.org\/","Rating":{"LastReviewIntro"
      :"Golden Gate Park is a great place to go check out and visit anytime! See the Chinese Garden, the
     Japanese Garden, the great American Bison, the Conservatory of Flowers, the Dutch windmill
     w√flowers and or many of the other things in the park...and then go just down the road a bit to
     check out Ocean Beach when you are all
     done.","TotalRatings":"7","LastReviewDate":"1172972754","AverageRating":"4","TotalReviews":"3"},"y:
     location":{"country":null,"city":"San Francisco","lat":"37.772735","street":"501 Stanyan
     St", "postal":null, "lon": "-122.454438", "quality": "100", "state": "CA"}, "Distance": "0.12", "Phone": "(415
     831-2700","ClickUrl":"http:\/\local.yahoo.com\/details?id=21370452&stx=park&csz=San+Franci
     sco+CA&amp:ed=SGNhwa160SvE3VbtIOu6sVNhreovG.xE4ia9Wwpss_u1vsNsnh6A8aFZFoTJ_.Fxoro7edIk","MapUrl":"ht
     tp:\//maps.yahoo.com/maps_result?name=Golden+Gate+Park+-+San+Francisco&desc=4158312700&csz
     =San+Francisco+CA& qty=9& cs=9& ed=SGNhwa160SyE3VbtIOu6sVNhreovG.xE4iq9Wwpss_u1ysNsnh6A8aFZ
     FoTJ_.Fxoro7edIk&gid1=21370452","BusinessUrl":"http:\/\www.sfgov.org\/","id":"21370452","title"
     :"Golden Gate Park - San Francisco"},"dc:title":"1br APRIL RENT FREE!!! One year lease (haight
     $825", "dc:type": "text", "dc:date": "2007-03-26T19:30:15-07:00", "dc:language": "en-us", "description": "T
     his is a 4br house and I am leaving the state at the beginning of April. There is a new lease coming
     up at the beginning of May. You can move in as soon as April 7th and not pay for April. You will be
     living with a third year law student at Hastings and two other roommates. This is a PRIME location
     for USF students. It is on Oak St two blocks down from Cha Cha's. Security deposit is
     $1650. Street parking is available. This is a beautiful location in the Upper Haight and a close
     walk to school. <br>\nLaundry included, hardwood floors, granite countertops, backyard, really nice
     landlord, and close muni access. Please contact by email asap to set up
     appointments.","y:location":{"country":"United States","street":"2019 Oak
     St","lat":"37.771179","postal":"94117","state":"CA","city":"San
     Francisco","lon":"-122.452583","quality":"87"},"dc:rights":"Copyright \u00a9 2007 craigslist,
     inc.","rdf:about":"http:\/\sfbay.craigslist.org\\sfc\/roo\\301061840.html","y:title":"1br APRIL
     RENT FREE!!! One year lease (haight ashbury)
     $825","dc:source":"http:\/\sfbay.craigslist.org\/sfc\/roo\/301061840.html","title":"1br APRIL RENT
     FREE!!! One year lease (haight ashbury)
     $825","pubDate":"2007-03-26T19:30:15-07:00","y:published":{"hour":"2","timezone":"UTC","second":"15
      ","month":"3","minute":"30","utime":"1174962615","day":"27","day_of_week":"2","year":"2007"},"geo:la
     t":37.771179, "geo:long":-122.452583}, {"link": "http:\/\sfbay.craigslist.org\/sfc\/apa\/301061385.htm
     l","nearby";{"link";"http:\/\local,yahoo,com\/details?id=23134331&stx=park&csz=+CA&ed=C
Line: 1 Column: 1 Plain Text
```





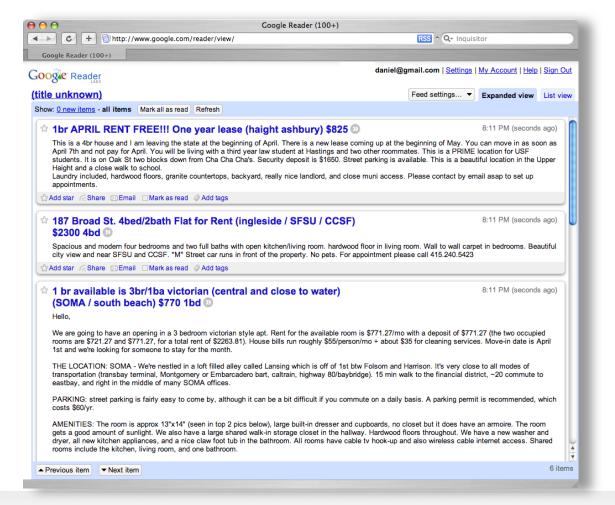




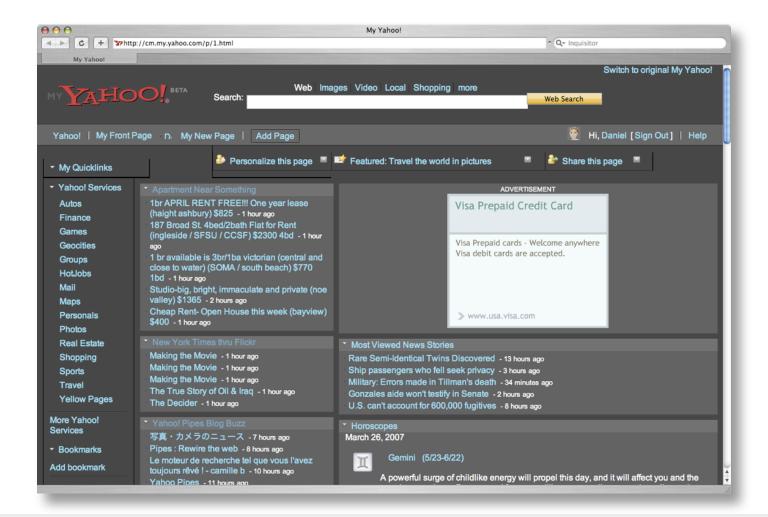




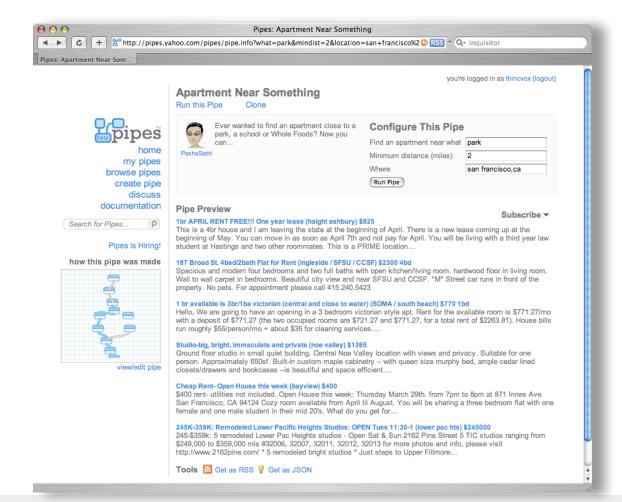




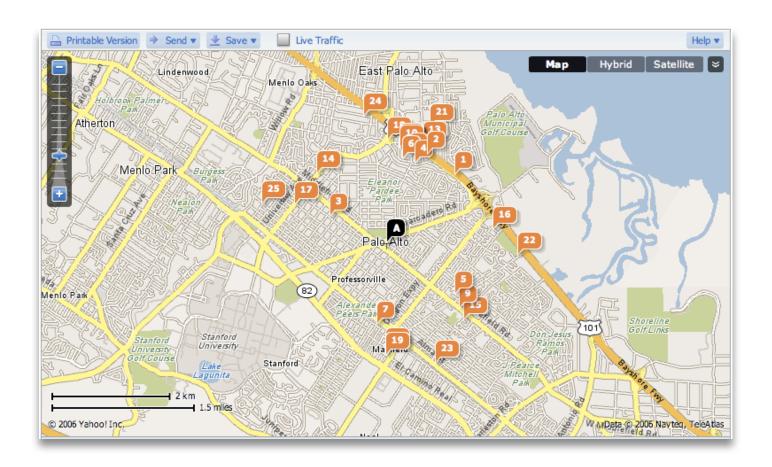


















RSS

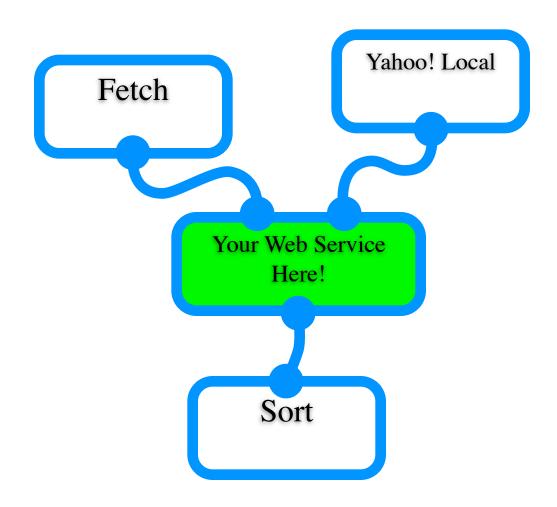
Badges

HTML

Your app here!



Any Process





Openness

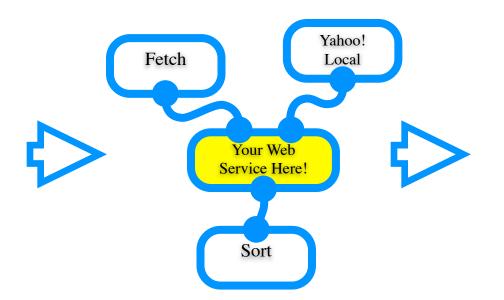
Craigslist

Yahoo!

Google

Ebay

Your data here!



RSS

Badges

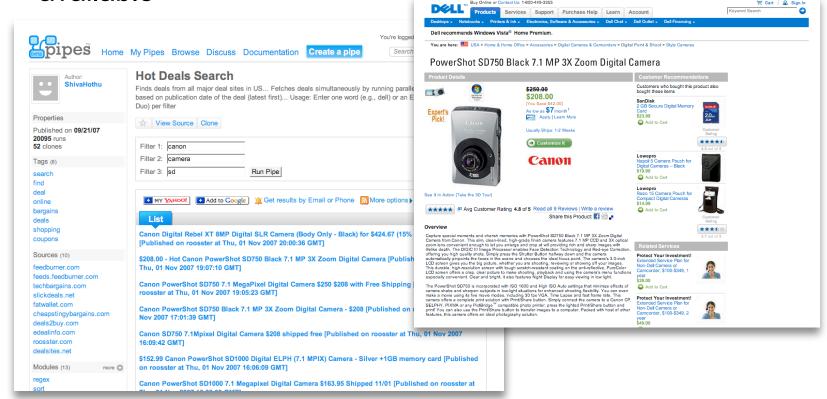
HTML

Your app here!



Hot Deals Search

 Searches across many different deal hunting sites on the internet looking for the best prices. You can search for particular items or just let the pipe find the best of what's available





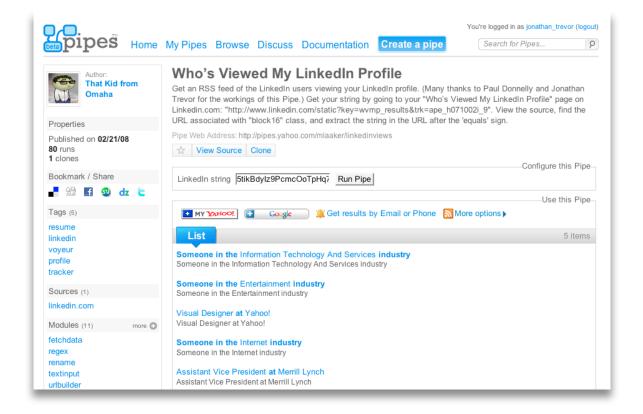
Geoannotated Reuters News

 Takes an RSS feed from the Reuters news service, and "geocodes" each item - making it possible to show where that news item is happening on a map of the world.





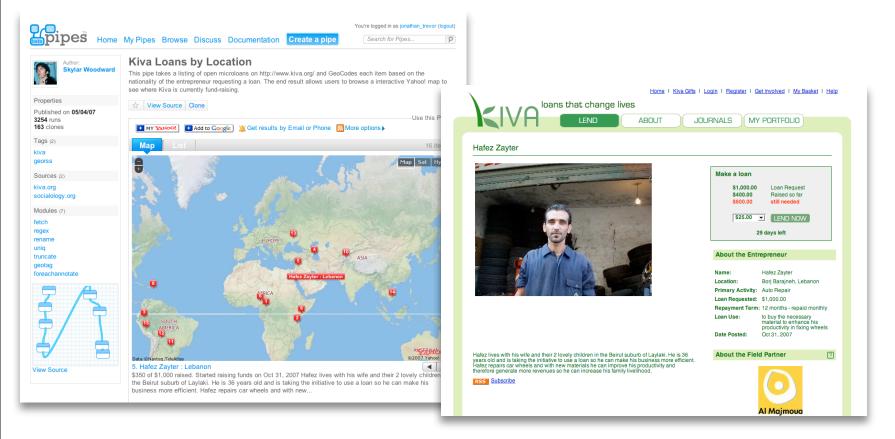
Who's Viewed My LinkedIn Profile





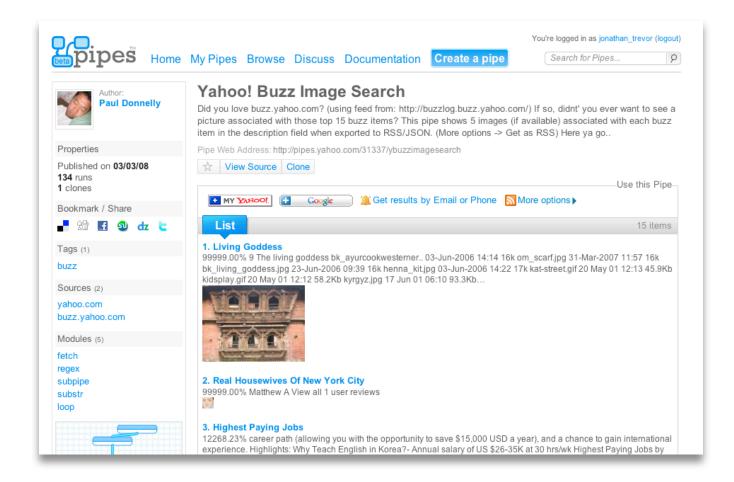
Kiva Loans by Location

 Gets a list of the micro-loans people have been making through the Kiva site, and shows the amazing variety of people and places that these loan are helping out.





Yahoo! Buzz Image Search





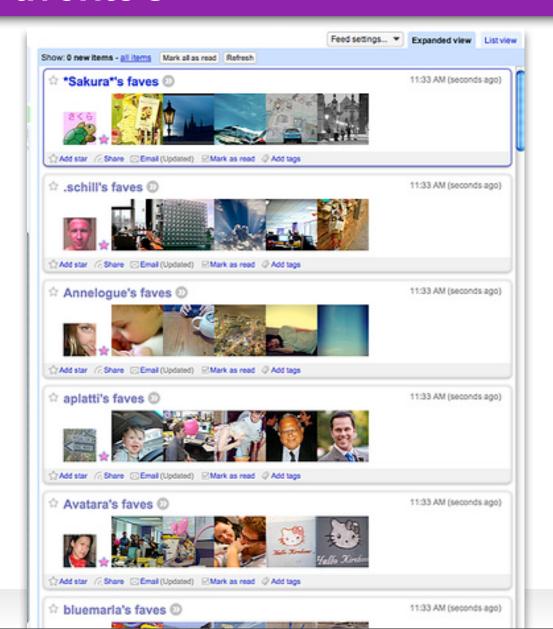
_ _

Yahoo Finance Stock Quote Watch List Feed w/ Chart





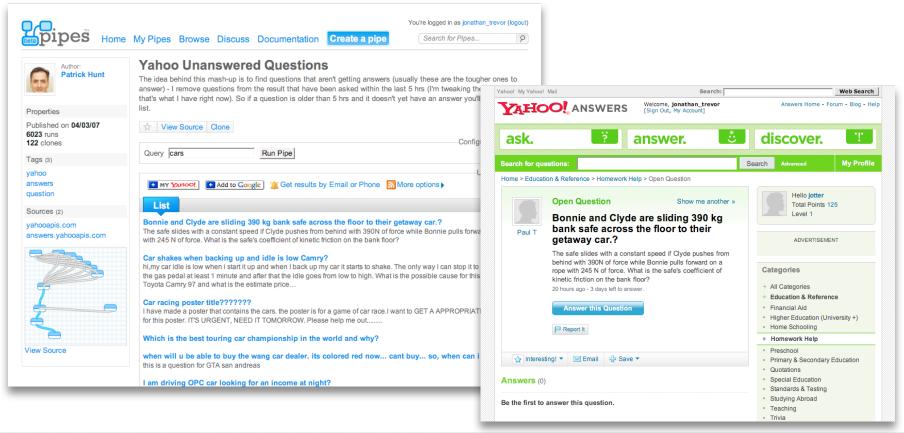
Contact's Favorite's





Yahoo Unanswered Questions

 Finds those questions in the Y! answers site that don't currently have an answer - so you can show how smart you are and answer those tricky questions.





Babbler by Max Case

Translates IM messages in Second Life





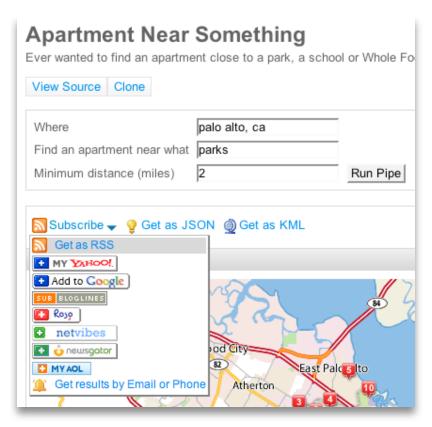
Advantages to developers

- Why use an online service to do this?
 - Leveraging large infrastructure
 - Faster access to network resources
 - Faster access to network services
 - System-wide knowledge
 - Leverage inter-organizational agreements
 - Easy to "string" together with other services
 - Easy to use (REST-style URLs)



Run / Get the data

Each Pipe gets its own "hosted" page

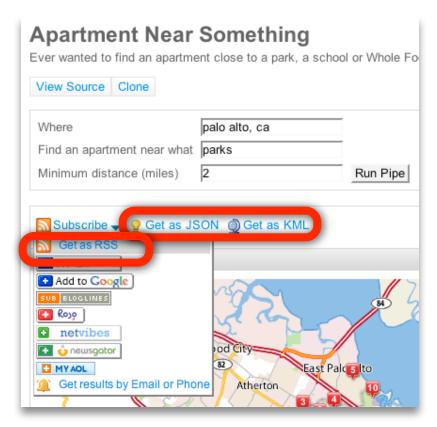


Use the REST-style URLs to get the data



Run / Get the data

Each Pipe gets its own "hosted" page



Use the REST-style URLs to get the data



```
http://pipes.yahoo.com/pipes/pipe.run?
_id=1mrlkB232xGjJDdwXqIxGw
&_render=json
&location=palo+alto%2C+ca
&mindist=2
&what=parks
&_callback=foofunction

The ID of the Pipe
```



```
http://pipes.yahoo.com/pipes/pipe.run?
   id=1mrlkB232xGjJDdwXqIxGw
&_render=json
&location=palo+alto%2C+ca
&mindist=2
&what=parks
&_callback=foofunction

The format of the output (rss, json, kml, ical, csv)
```



```
http://pipes.yahoo.com/pipes/pipe.run?
_id=1mrlkB232xGjJDdwXqIxGw
&_render=json
&location=palo+alto%2C+ca
&mindist=2
&what=parks
&_callback=foofunction
```

The per Pipe user customizable parameters



```
http://pipes.yahoo.com/pipes/pipe.run?
 id=1mrlkB232xGjJDdwXqIxGw
& render=json
&location=palo+alto%2C+ca
 &mindist=2
 &what=parks
 & callback=foofunction
Optional JSONP callback function
```



A year and a half in the wild: a few observations and lessons

- 20+ releases, 600k+ Pipes later
- Unexpected breadth
 - Experts who want to exploit the service
 - Non-programers with much simpler needs



Web addressable data...

- is very malformed
- can be slow
- needs considerate access
- can be untrustworthy
- can be inaccessible from "here" (behind firewall etc)



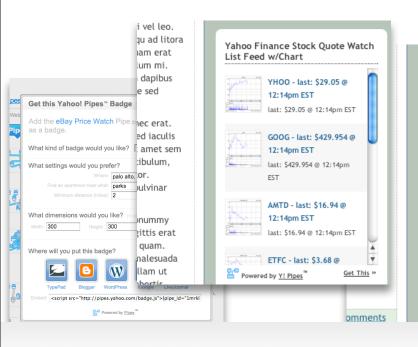
Data in the Engine...

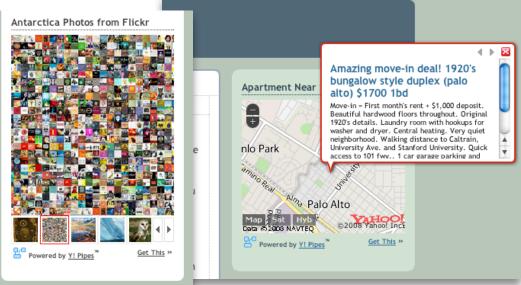
- is "cleaned" (and repaired) into UTF-8
- is cached for
 - performance
 - playing well with others
 - several HTTP proxy layers
 - · serve stale and force caching
- is "sanitized"



Making it easy to consume

- Its easy to make useful data in the cloud
 - Its not easy enough (for many) to use it after
 - Visualization beyond lists in RSS readers
- Badges are frequently requested
- Three variants for common types of data in Pipes







Typical Pipes/mashups

- Four types of mashup
 - Feed aggregation with filtering
 - Two-source mashups
 - Data transformation and geocoding
 - Complex mashups using REST APIs
- Geocoding remains a "mashup" favorite



Reasons for adoption

- Lower barrier to use
 - Graphical editor made it quick to write Pipes, attracted nondevelopers
 - "View Source" and "Clone" for learning/tweaking
- Wide array of data input formats and data output formats enabled Pipes to become a useful "component" in a larger ecology
- Web 2.0 responsiveness to community



Inaccessible data

- Lots of requests for more rich and personal data
 - Text documents, word documents, mail, Excel spreadsheets
 - Also organizational data
- Workarounds (to some) emerged
 - Online spreadsheets, calendars (gcal) with private RSS feeds and so on



Power...

- We started by focusing on RSS
 - high-level building blocks and operations
 - good for common tasks and novice users
- We listened to our user's desires



...vs Complexity

- Added sources for parsing JSON, XML, CSV, ICAL ...
- Added modules that could do more and be combined in many ways
- At the cost of simplicity
 - Harder to explain, use, compose
 - Stretching the capabilities of many users and a visual development environment



Yahoo! Query Language (YQL)



YQL

- Part of the recent Y!OS release
 - Social APIs, Universal profile, Application platform...
- Mediator service that enables developers to query, filter and combine Y! data and beyond
 - Yahoo! web services and any URL-addressable structured data sources
- Exposes a SQL-like SELECT syntax that is both familiar to developers and expressive enough for getting the right data
 - YQL operates on hierarchical documents, not relational tuples
- Like Pipes but with a simple textual language



The language and service

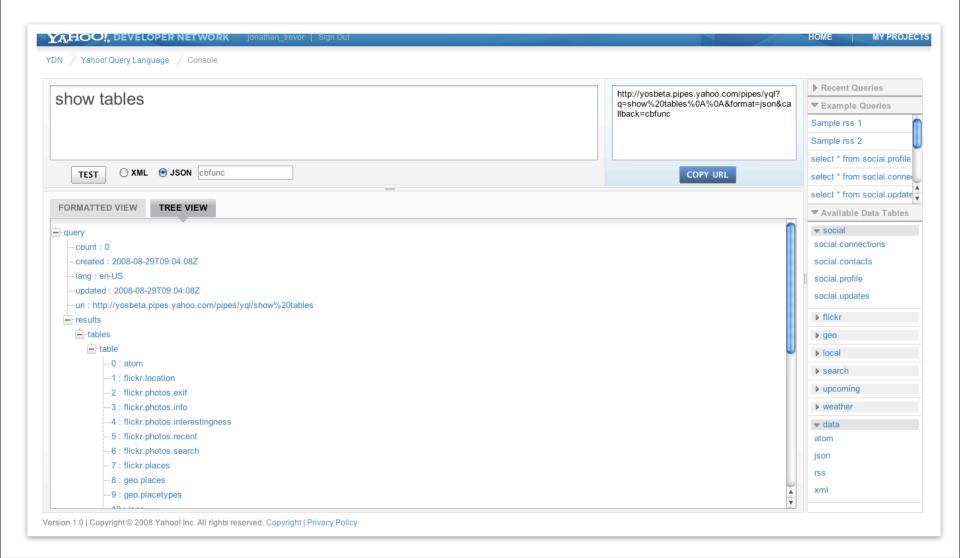
- Provides three SQL-like statements: SELECT, SHOW, DESC
- Single URL endpoint for executing everything
 - Mix and match external data and Yahoo! APIs

http://query.yahooapis.com/v1/yql?q=show%20tables

- Uses Oauth for authentication
 - Open standard that enables users to grant applications access to (selected) private data



Testing your queries: interactive console





- -



http://query.yahooapis.com /v1/yql?q=select%20*%20from%20social.profil %20%3D%20me&format=xml \rightarrow COPY URL

4 b

TEST





FORMATTED VIEW

TREE VIEW

```
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                    yahoo:count="1" yahoo:created="2008-09-30T02:46:57Z"
                                                                                                     yahoo
   <diagnostics>
      <url execution-time="162">http://social.yahooapis.com/v1/user/UNOY4FVPZVTT7C7KGLGZY4SIRM/profile</url>
      <user-time>353</user-time>
      <service-time>162</service-time>
      <build-version>2008.09.25.09:48</build-version>
   </diagnostics>
   <results>
      <guid>UNOY4FVPZVTT7C7KGLGZY4SIRM</guid>
          <created>2008-08-12T17:17:18Z</created>
          <familyName>Trevor</familyName>
          <gender>M</gender>
          <givenName>Jonathan</givenName>
          <image>
             <height>192</height>
             <imageUrl>http://l.yimg.com/us.yimg.com/i/identity/nopic 192.gif</imageUrl>
             <size>192x192</size>
          </image>
          <location/>
          <nickname>JonathanT</nickname>
          <isConnected>true</isConnected>
      </profile>
   </results>
</guery>
```



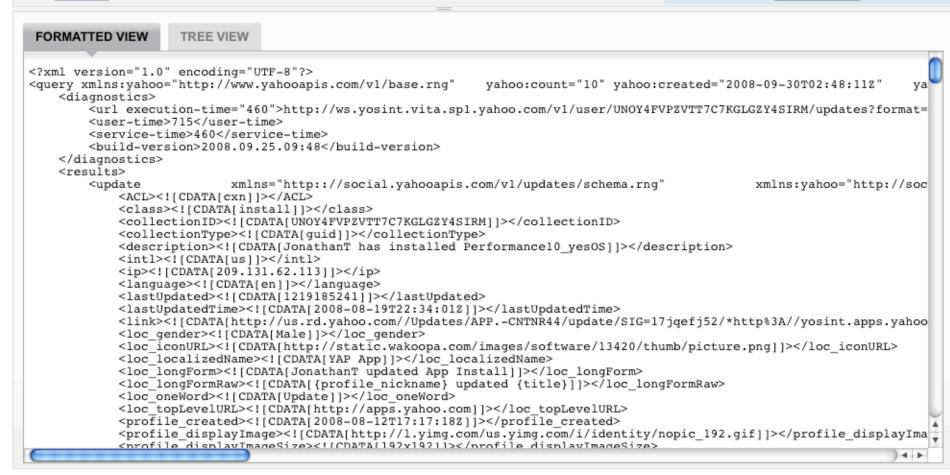
http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20social.upda format=xml

) 4 b

TEST



COPY URL





http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20social.conn format=xml

TEST

XML JSON



COPY URL

4 1

```
FORMATTED VIEW
                     TREE VIEW
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                               yahoo:count="5" yahoo:created="2008-09-30T02:48:54Z"
    <diagnostics>
        <url execution-time="185">http://ws.yosint.socdir.spl.yahoo.com:4080/v1/user/UNOY4FVPZVTT7C7KGLGZY4SIRM/connecti
        <user-time>221</user-time>
        <service-time>185</service-time>
        <build-version>2008.09.25.09:48</puild-version>
    </diagnostics>
    <results>
        <connection xmlns="http://social.yahooapis.com/v1/schema.rng"</pre>
                                                                                    yahoo:created="2008-08-28T20:52:06Z"
            <quid>54DY7TNQ6ZMUWA5C6EWEGFDAWY</quid>
            <contactId>0</contactId>
        </connection>
        <connection xmlns="http://social.yahooapis.com/vl/schema.rng"</pre>
                                                                                    vahoo:created="2008-08-25T19:52:07Z"
            <quid>7C7SIF3MMHEOH3MJ5HYPYOOVOI</quid>
            <contactId>1</contactId>
        </connection>
        <connection xmlns="http://social.yahooapis.com/vl/schema.rng"</pre>
                                                                                    yahoo:created="2008-08-28T20:52:06Z"
            <guid>GAUV42G4YMWOYCRCG7DP5PJJ6U</guid>
            <contactId>0</contactId>
        </connection>
        <connection xmlns="http://social.yahooapis.com/v1/schema.rng"</pre>
                                                                                    vahoo:created="2008-08-28T20:52:06Z"
            <quid>PLQBS2ATTNUELZMA2QNTF2I6RI</quid>
            <contactId>0</contactId>
        </connection>
        <connection xmlns="http://social.yahooapis.com/v1/schema.rng"</pre>
                                                                                    yahoo:created="2008-08-28T20:52:06Z"
            <guid>ZJP7P4HM2PGEZLGNXAMSO4Z5TM</guid>
            <contactId>0</contactId>
        </connection>
    </results>
</mierv>
```

select * from social.profile where guid in (select guid from social.connections where owner_guid=me)

http://query.yahooapis.com /v1/vgl?q=select%20*%20from%20social.profil format=xml

TEST



XML JSON

COPY URL



TREE VIEW

```
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                          yahoo:count="5" yahoo:created="2008-09-30T02:49:28Z"
                                                                                                                 vah
   <diagnostics>
       <url execution-time="170">http://ws.yosint.socdir.spl.yahoo.com:4080/v1/user/UNOY4FVPZVTT7C7KGLGZY4SIRM/connecti
       <url execution-time="162">http://social.yahooapis.com/v1/user/PLOBS2ATTNUELZMA2ONTF2I6RI/profile</url>
       <url execution-time="156">http://social.yahooapis.com/v1/user/GAUV42G4YMW0YCRCG7DP5PJJ6U/profile</url>
       <url execution-time="160">http://social.yahooapis.com/v1/user/7C7SIF3MMHEOH3MJ5HYPYOOVOI/profile</url>
       <url execution-time="165">http://social.yahooapis.com/v1/user/54DY7TNQ6ZMUWA5C6EWEGFDAWY/profile</url>
       <url execution-time="164">http://social.yahooapis.com/v1/user/ZJP7P4HM2PGEZLGNXAMSO4Z5TM/profile</url>
       <user-time>557</user-time>
       <service-time>977</service-time>
       <build-version>2008.09.25.09:48/build-version>
   </diagnostics>
   <results>
       <guid>PLQBS2ATTNUELZMA2QNTF2I6RI</guid>
           <created>2008-08-21T00:11:49Z</created>
           <familyName>Doe</familyName>
           <gender>M</gender>
           <givenName>John</givenName>
           <image>
               <height>192</height>
               <imageUrl>http://l.yimg.com/us.yimg.com/i/identity/nopic 192.gif</imageUrl>
               <size>192x192</size>
           </image>
           <location>New York</location>
           <lookingFor>NETWORKING</lookingFor>
           <nickname>Agent Smith</nickname>
           <prefileUrl>http://yosint.profiles.yahoo.com/u/PLQBS2ATTNUELZMA2QNTF2I6RI
           <relationshipStatus>COMPLICATED</relationshipStatus>
           <e+a+11e>
```

XML JSON





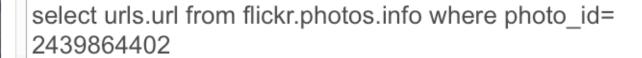
http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20flickr.photos %3D%202439864402&format=xml

COPY URL

TEST

FORMATTED VIEW TREE VIEW <?xml version="1.0" encoding="UTF-8"?> <query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre> yahoo:count="1" yahoo:created="2008-09-30T02:51:03Z" yah <diagnostics> <url execution-time="177">http://api.flickr.com/services/rest/?method=flickr.photos.getInfo&photo id=2439864 <user-time>184</user-time> <service-time>177</service-time> <build-version>2008.09.26.09:09/build-version> </diagnostics> <results> <photo dateuploaded="1209077470" farm="4" id="2439864402"</pre> isfavorite="0" license="0" media="photo" <owner location="South San Francisco, USA"</pre> nsid="8614440@N03" realname="Alex Ho" username="ge <title>IMG 7106</title> <description>Justine Ezarik and Sam Pullara</description> <visibility isfamily="0" isfriend="0" ispublic="1"/> <dates lastupdate="1209671816" posted="1209077470"</pre> taken="2008-04-23 03:12:34" takengranulari <editability canaddmeta="0" cancomment="0"/> <usage canblog="0" candownload="1" canprint="0"/> <comments>1</comments> <notes/> <tags> machine tag="0" raw="web">web< <tag author="8614440@N03" id="8591386-2439864402-5913"</p> <tag author="8614440@N03" id="8591386-2439864402-26493"</p> machine tag="0" raw="2.0">20< <tag author="8614440@N03" id="8591386-2439864402-4251"</p> machine tag="0" raw="expo">exp <tag author="8614440@N03" id="8591386-2439864402-11227"</pre> machine tag="0" raw="web20">w id="8591386-2439864402-10061458" machine tag="0" raw="web20 <tag author="8614440@N03" <tag author="8614440@N03" id="8591386-2439864402-34067"</p> machine tag="0" raw="moscone" <tag author="8614440@N03" id="8591386-2439864402-46"</p> machine tag="0" raw="sanfrancisc <tag author="8614440@N03" id="8591386-2439864402-37813"</pre> machine tag="0" raw="2008">20 <tag author="8614440@N03" id="8591386-2439864402-42024"</pre> machine tag="0" raw="crawl">c <tag author="8614440@N03" id="8591386-2439864402-17079"</p> machine tag="0" raw="bayarea"





http://query.yahooapis.com /v1/ygl?q=select%20urls.url%20from%20flickr.j %3D%202439864402&format=xml **COPY URL**

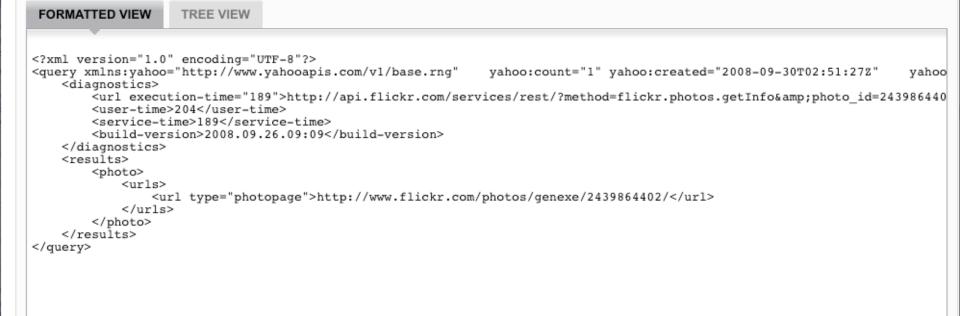
 \rightarrow













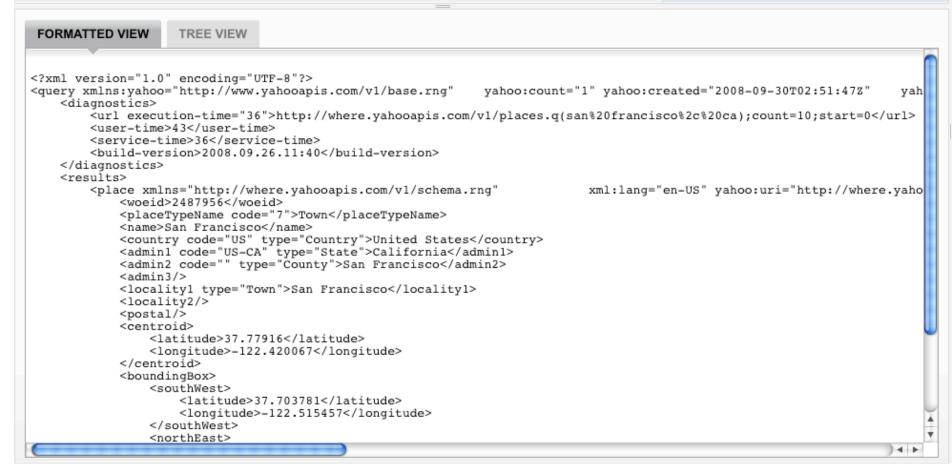
http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20geo.places1 %3D%22san%20francisco%2C%20ca%22& format=xml

TEST



COPY URL

4 1



XML JSON



TEST

select * from upcoming.events where woeld in (select woeld from geo.places where text="North Beach")

http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20upcoming.e %3D%22North%20Beach%22)&format=xml 4 1 COPY URL

description="She & Him Doors at 7 p.m. S

description="Multi-instrumentalist, Meric Lo

description="Brooklyn?s Gang Gang Dance crea"

```
FORMATTED VIEW
                    TREE VIEW
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                              yahoo:count="12" yahoo:created="2008-09-30T02:54:00Z"
                                                                                                                         ya
   <diagnostics>
        <url execution-time="63">http://where.yahooapis.com/v1/places.q(North%20Beach);count=10;start=0</url>
        <url execution-time="67">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=1097818&a
        <url execution-time="73">http://upcoming.yahooapis.com/services/rest/?per_page=50&amp;page=1&amp;woeid=23511188&
        <url execution-time="79">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=2460639&a
        <url execution-time="89">http://upcoming.yahooapis.com/services/rest/?per_page=50&amp;page=1&amp;woeid=28746208&
        <url execution-time="99">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=2460640&a
        <url execution-time="62">http://upcoming.yahooapis.com/services/rest/?per_page=50&amp;page=1&amp;woeid=23499175&
        <url execution-time="73">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=2460636&a
        <url execution-time="64">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=2460633&a
        <url execution-time="75">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=23417381&
        <url execution-time="70">http://upcoming.yahooapis.com/services/rest/?per page=50&amp;page=1&amp;woeid=2460634&a
        <user-time>253</user-time>
        <service-time>814</service-time>
        <build-version>2008.09.26.09:09</puild-version>
   </diagnostics>
    <results>
        <event category id="2" date posted="2008-09-05 18:35:59"</pre>
                                                                              description="America's ONLY weekend-long con
        <event category id="2" date posted="2008-09-06 21:06:22"</pre>
                                                                              description="America's ONLY weekend-long con
        <event category id="1" date posted="2008-07-28 13:22:32"</pre>
                                                                              description="Iron and Wine (solo) with Sholi
        <event category id="1" date posted="2008-09-08 22:41:07"</pre>
                                                                              description="" end date="" end time="-00001"
        <event category id="1" date posted="2008-07-15 09:47:59"</pre>
                                                                              description="Tickets are $19.50. 21+ Jolie H
        <event category id="1" date posted="2008-07-11 12:55:12"</pre>
                                                                              description="Dengue Fever?s psychedelic take
                                                                              description="doors @ 7 p.m. 18+" end date=""
        <event category id="1"
                               date posted="2008-08-06 09:47:35"
                                                                              description="" end date="" end time="-00001"
```

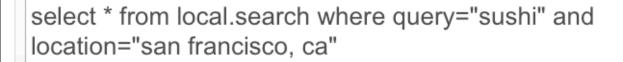
<event category id="1" date posted="2008-06-26 15:46:01"</pre> <event category id="1" date posted="2008-09-02 11:39:43"</pre>

<event category id="1" date posted="2008-09-22 12:46:30"</pre>

<event category id="1" date posted="2008-09-08 16:57:35"</pre>



YDN / Yahoo! Query Language /



http://query.yahooapis.com /v1/vgl?q=select%20*%20from%20local.search %3D%22sushi%22%20and%20location %3D%22san%20francisco%2C%20ca%22& format=xml

TEST

XML JSON

COPY URL

FORMATTED VIEW

TREE VIEW

```
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                             yahoo:count="10" yahoo:created="2008-09-30T02:52:57Z"
   <diagnostics>
       <url execution-time="1379">http://local.yahooapis.com/LocalSearchService/V3/localSearch?results=10&amp;start=1&a
       <user-time>1393</user-time>
       <service-time>1379</service-time>
       <build-version>2008.09.26.09:09</puild-version>
   </diagnostics>
   <results>
       <Result xmlns="urn:yahoo:lcl" id="21373494">
            <Title>Sushi Bistro</Title>
            <Address>455 Balboa St</Address>
            <City>San Francisco</City>
           <State>CA</State>
            <Phone>(415) 933-7100</Phone>
            <Latitude>37.777099</Latitude>
            <Longitude>-122.463517</Longitude>
            <Rating>
                <AverageRating>4.5</AverageRating>
                <TotalRatings>43</TotalRatings>
                <TotalReviews>36</TotalReviews>
                <LastReviewDate>1189988903</LastReviewDate>
                <LastReviewIntro>On 09.15.07 Saturday evening we walked in and were told it will be 45 minutes. After 40
            </Rating>
            <Distance>2.55</Distance>
           <Url>http://local.yahoo.com/info-21373494-sushi-bistro-san-francisco</Url>
            <ClickUrl>http://local.yahoo.com/info-21373494-sushi-bistro-san-francisco</ClickUrl>
            <MapUrl>http://maps.yahoo.com/maps result?q1=455+Balboa+St+San+Francisco+CA&amp;gid1=21373494</MapUrl>
            <BusinessUrl>http://www.sushibistrosf.com/</BusinessUrl>
            <BusinessClickUrl>http://www.sushibistrosf.com/</BusinessClickUrl>
```

YDN / Yahoo! Query Language / Console

select * from local.search where query="sushi" and location="san francisco, ca" and Rating.AverageRating>4

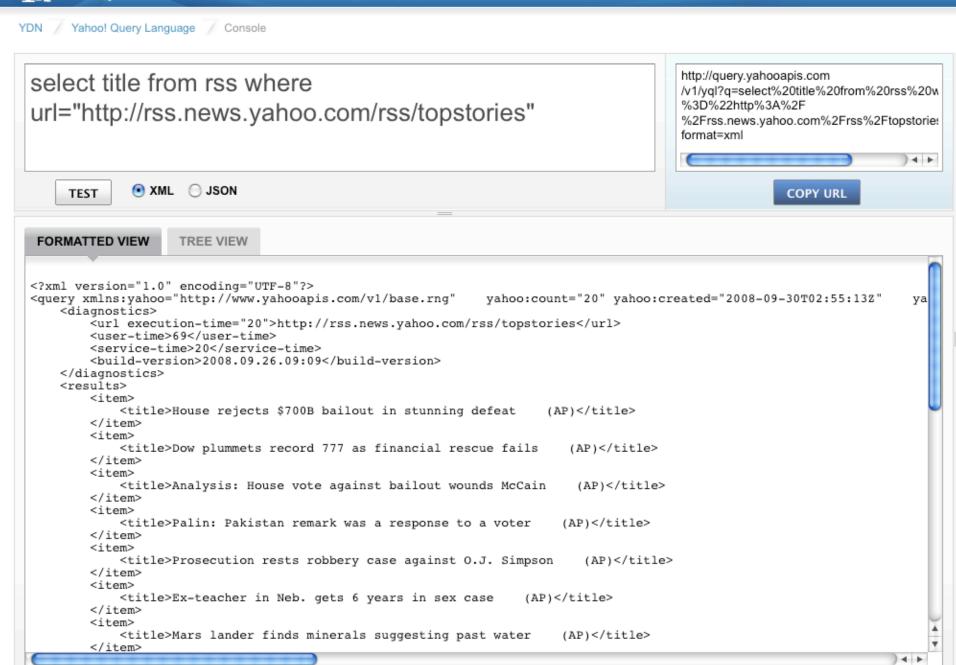
http://query.yahooapis.com /v1/ygl?g=select%20*%20from%20local.sea %3D%22sushi%22%20and%20location %3D%22san%20francisco%2C%20ca %22%20and%20Rating.AverageRating%3E format=xml **4 b**

TEST

XML JSON

COPY URL

FORMATTED VIEW TREE VIEW <?xml version="1.0" encoding="UTF-8"?> yahoo:count="9" yahoo:created="2008-09-30T02:53:39Z" <query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre> <diagnostics> <url execution-time="2650">http://local.yahooapis.com/LocalSearchService/V3/localSearch?results=10&start=1&a <user-time>2670</user-time> <service-time>2650</service-time> <build-version>2008.09.26.11:40/build-version> </diagnostics> <results> <Result xmlns="urn:yahoo:lc1" id="21373494"> <Title>Sushi Bistro</Title> <Address>455 Balboa St</Address> <City>San Francisco</City> <State>CA</State> <Phone>(415) 933-7100</Phone> <Latitude>37.777099</Latitude> <Longitude>-122.463517</Longitude> <Rating> <AverageRating>4.5</AverageRating> <TotalRatings>43</TotalRatings> <TotalReviews>36</TotalReviews> <LastReviewDate>1189988903</LastReviewDate> <LastReviewIntro>On 09.15.07 Saturday evening we walked in and were told it will be 45 minutes. After 40 </Rating> <Distance>2.55</Distance> <Url>http://local.yahoo.com/info-21373494-sushi-bistro-san-francisco</Url> <ClickUrl>http://local.yahoo.com/info-21373494-sushi-bistro-san-francisco</ClickUrl> <MapUrl>http://maps.yahoo.com/maps result?q1=455+Balboa+St+San+Francisco+CA&gid1=21373494</mapUrl> <BusinessUrl>http://www.sushibistrosf.com/</BusinessUrl> <BusinessClickUrl>http://www.sushibistrosf.com/</BusinessClickUrl>



Yahoo! Query Language

select * from html where url="http://finance.yahoo.com /q?s=yhoo" and xpath='//div[@id="yfi_headlines"]/div[2] /ul/li/a'

http://query.yahooapis.com /v1/ygl?q=select%20*%20from%20html%20 %3D%22http%3A%2F %2Ffinance.yahoo.com%2Fg%3Fs%3Dyho %22%20and%20xpath%3D'%2F%2Fdiv %5B%40id%3D%22vfi_headlines%22%5D

COPY URL

) - | -

TEST

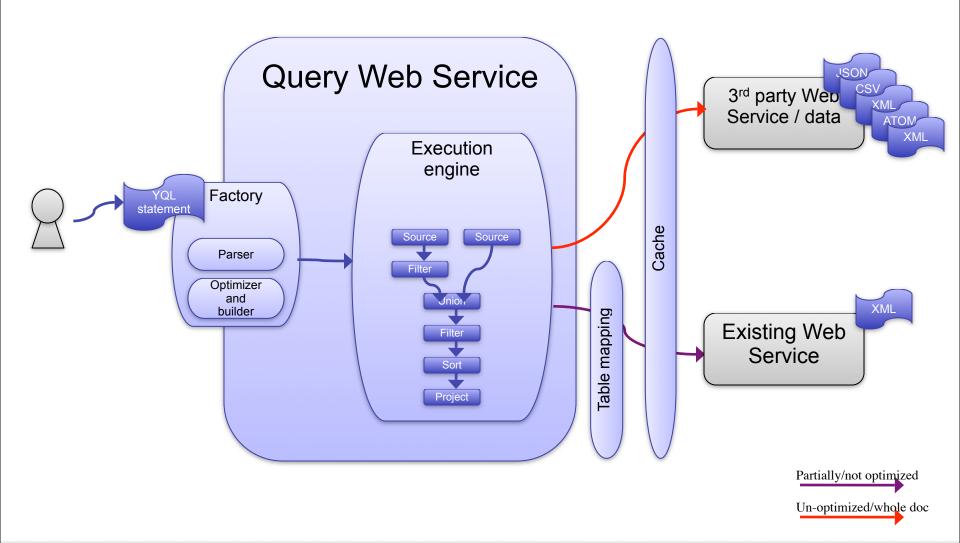
XML JSON

FORMATTED VIEW

TREE VIEW

```
<?xml version="1.0" encoding="UTF-8"?>
<query xmlns:yahoo="http://www.yahooapis.com/v1/base.rng"</pre>
                                                              yahoo:count="10" yahoo:created="2008-09-30T02:56:29Z"
                                                                                                                        yaho
   <diagnostics>
        <url execution-time="162">http://finance.yahoo.com/g?s=yhoo</url>
        <user-time>196</user-time>
        <service-time>162</service-time>
        <build-version>2008.09.26.09:09</puild-version>
   </diagnostics>
    <results>
        <a href="http://us.rd.yahoo.com/finance/external/wsj/SIG=11puff229/*http://online.wsj.com/article/SB12227249660218
        <a href="http://biz.yahoo.com/ap/080929/internet radio royalties.html?.v=4">Congress</a>
        <a href="http://us.rd.yahoo.com/finance/external/cbsm/SIG=12jf4c6ng/*http://www.marketwatch.com/news/story/new-sig
        <a href="http://biz.yahoo.com/ap/080929/internet final glance.html?.v=1">Final
        <a href="http://us.rd.yahoo.com/finance/external/forbes/SIG=13erp9gmp/*http://www.forbes.com/2008/09/29/small-busi
        <a href="http://us.rd.yahoo.com/finance/external/allthingsd/SIG=13fhf49ie/*http://kara.allthingsd.com/20080929/yus
        <a href="http://biz.yahoo.com/ap/080929/internet_stocks_sector_wrap.html?.v=1">Sector_
        <a href="http://us.rd.yahoo.com/finance/external/mfool/SIG=12ri9nkas/*http://www.fool.com/investing/high-growth/20
        <a href="http://us.rd.yahoo.com/finance/external/bizj/SIG=12i4sbj1j/*http://sanfrancisco.bizjournals.com/sanfranci
        <a href="http://us.rd.yahoo.com/finance/external/tsmfe/SIG=127j315vh/*http://www.thestreet.com/ yahoo/newsanalysis
   </results>
</query>
```

(Very) High Level Architecture





Mapping tables to data sources

- YQL wants to push as much of the query as possible to the remote data provider/service
- Typically REST query/path parameters do not map closely to result structure
 - We call these "keys" and are named differently than dot-path
 - Simple REST definition language describes how YQL executes queries on "table" providers



Remote and Local filtering, paging

- Table data can be filtered in the WHERE clause either:
 - Remotely by the table data source provider or
 - Locally by the YQL engine
- YQL tries to present "rows" of data
 - Abstracts away "paging" views of data sources
 - Presents a "subset" of paging tables by default

```
select * from local.search(500,1000) where
zip='94085' and query='pizza'
```



IN (SELECT...): Joining across data sources

- No left joins, sub-select only
- Get an international weather forecast? Join two services in different companies:

```
select * from weather.forecast where location in
  (select id from xml where
    url="http://xoap.weather.com/search/search?where=prague"
    and itemPath="search.loc")
```

- Sub-select works the same as normal select except it can only return a "leaf" element value or attribute
- Parallelizes execution



Post-query manipulation

 Simple post-SELECT processing can be performed by appending the "pipe" symbol to the end of the statement:

```
SELECT ... | sort(field=item.date)
SELECT ... | unique(field=item.title) | ...
```

 Functions only operate on the data being returned by the query, nothing to do with the tables or data sources themselves



Use it!

- Public+private YQL tables can be accessed at: http://query.yahooapis.com/v1/yq1?q=...
- Oauth protected, URLs must be signed
 - 2-legged for public tables
 - 3-legged for social tables
- Public YQL tables (soon)
 - No signing required



Next steps, challenges

- "Open" tables
- Multiple authentication authority support
- Better YQL query optimization for endpoints
- Foreign key consistency
- Scripting/language bindings



Finally: Pipes without the GUI editor

Very popular Pipe pattern is easy to represent in YQL:

```
SELECT * FROM rss WHERE url in

(SELECT title FROM atom WHERE url="http://
   spreadsheets.google.com/feeds/list/pg_TOM/
   od6/public/basic")
   AND description LIKE "%wall street%" LIMIT 10
   | unique (field=title)
```



Conclusion: Pipes and YQL

- Provides powerful data functions to any client
- Consumes data from many services
- Common data formats means any part of the cloud can become the input
 - Dapper, AWS, Google spreadsheets
- ...or take the output
 - 1/3 Google mashups are powered by Pipes



Conclusion: Pipes and YQL

- Enable developers to easily access, combine, and filter data to fit their application requirements
 - Self-documenting model
- YQL provides developers with consistent and unified semantics for accessing data, not just Yahoo! services
- Low overhead
- Reduce roundtrip traffic by reducing the number of requests



Thank you

- Pipes
 - http://pipes.yahoo.com
- YQL
 - http://query.yahooapis.com/v1/yql
 - http://developer.yahoo.com/yql
 - http://developer.yahoo.com/yql/console
- Get in touch
 - jtrevor@yahoo-inc.com
 - yql-questions@yahoo-inc.com

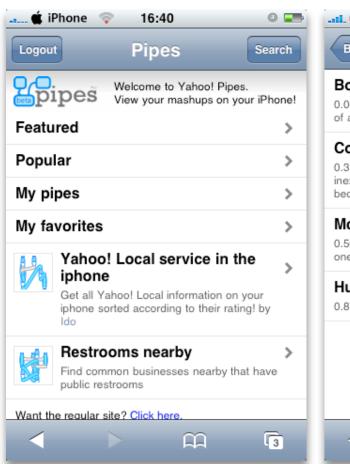


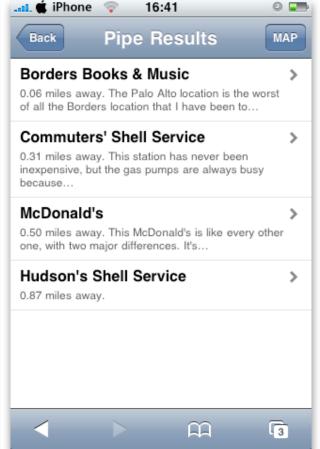
REST def

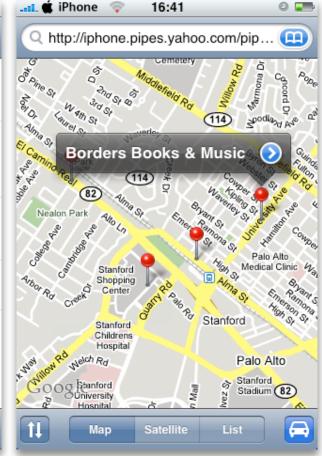
```
<?xml version="1.0" encoding="UTF-8"?>
<sampleQuery>select * from geo.places where text="sfo"</sampleQuery>
 <endpoints>
   <endpoint itemPath="places.place" format="XML">
     <urls>
       <url env="all">http://where.yahooapis.com/v1/
places=dol=and(.q(\$text\$,\$focus\$),.type(\$placetype\$))?appid=xxx</url>
     </urls>
     <paging model="offset">
       <start id="start" default="0" matrix="true" />
       <pagesize id="count" max="10" matrix="true" />
       <total default="10" />
     </paging>
     <keys>
       <key id="text" type="xs:string" />
       <key id="focus" type="xs:string" />
       <key id="placetype" type="xs:string" />
     </keys>
   </endpoint>
```



Doing the mobile mash





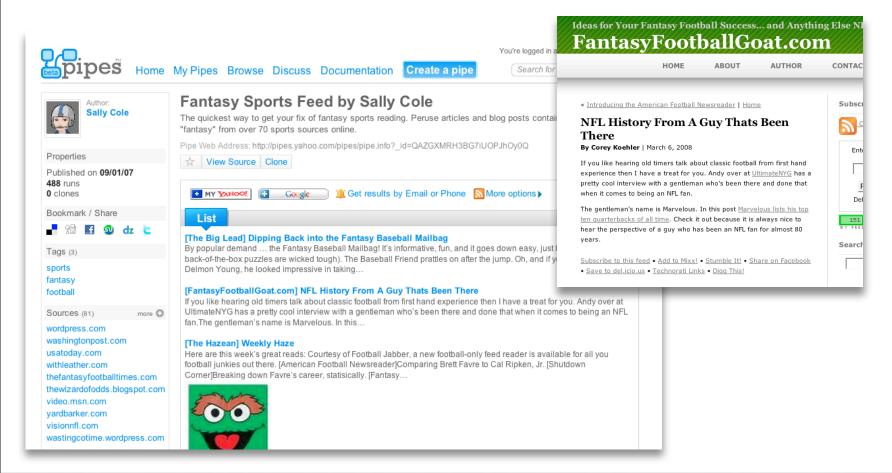




_ .

Fantasy Sports search

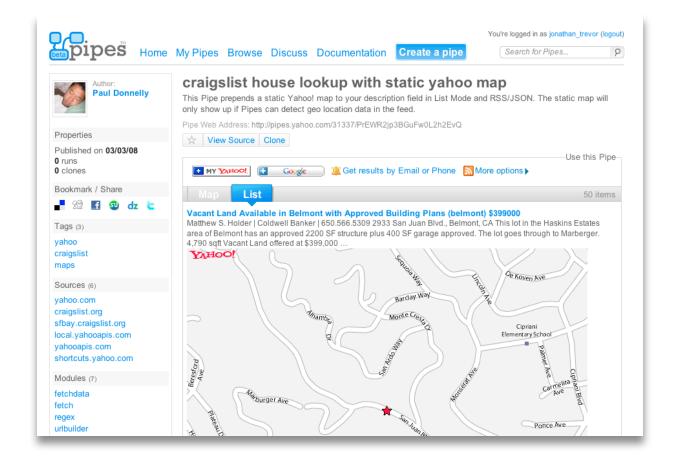
 Get the edge on your friends with a single RSS feed based on searching 70 sites for fantasy sports blog articles





- -

Craigslist house lookup with static Yahoo map

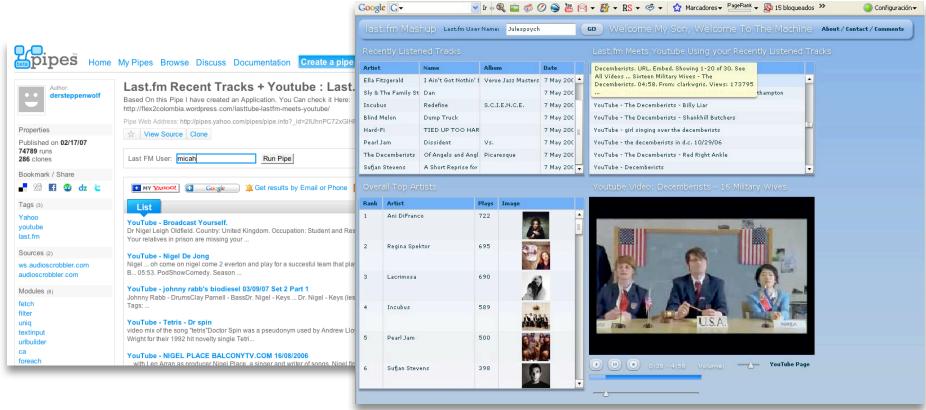




_ _

LastTube

 Uses content from Last.fm and YouTube. You can watch Youtube's content based on your Recently Listened Tracks scrobbled to Last.fm.





- -