UFCEKG-20-2
Data, Schemas & Applications

Lecture 6
The Web of Data, API’s & PHP
Twitter to RSS

Searching the Twitter stream

Twitter advanced search
Bristol tweets as Atom
Tweet Steam via ATOM and XQuery
Last week:

- Tree structures
- Navigating trees with SimpleXML in PHP5
- Object construction and navigation in PHP5
- Iteration in PHP - foreach ()
- PHP for transformation form one XML vocabulary to another
- String building with .= assignment
Human versus machine-readable sites

- Most sites are built for humans to browse. HTML pages designed for humans are not good for machines.
- Page is a complex document including CSS, Javascript and HTML layout.
- Many pages are not well-formed XHTML.
- Hard to identify the real content of the page.
- Page 'scraping' - getting the content out of a page - is possible but hard.
- The structure of the HTML changes from time to time even if the content is the same.

There are applications designed to help page scrape - e.g. Dapper.
Flickr site

- Flickr site
- Lewis Hamilton in text
- Lewis Hamilton in tags
- Advanced Search
Application Programmer Interface (API)

Interface designed for a programmer to use in a program

- Provides a machine-readable content
- Can be used client-side (e.g. JSON) or server-side (e.g. XML, Atom, RSS)
- Separates the content from the presentation
- Often requires personal access key (API key) to monitor and control access
- Provides a rich language to query data
- With authentication, allows update to user's data
Flickr API

Flickr API Documentation
Supports multiple protocols - we will use REST (HTTP GET)
method flickr.photos-search
API explorer lets you experiment with queries
Response from API call
URL construction
Flickr API Discussion

Photos can be searched on a wide range of criteria not available via the usual web search. For example, geo-coded photos can be searched by location latitude and longitude and radius of circle.
PHP Flickr search (1)

This script provides a secondary interface to Flickr using the API. The output is suitable for including in a page using iframe

Peros Bridge
Peros Bridge first 2
flickrsearch.php source
PHP Flickr search (2)

A generic function to call any Flickr function with an api key.

Accepts a method (e.g. flickr.photos.search) and an array of parameters (name=value), calls Flickr and returns the simpleXML object.

```php
function call_flickr($method, $params)
{
           "&api_key=" . FLICKRKEY . "&" . 
           join($params, "&");
    $xml = get_file($uri);
    return new SimpleXMLElement($xml);
}
```

The FLICKRKEY is defined in a separate file. A key is freely available on registration but there are limits on the rate and number of requests made, and keys are bound to a specific host. (i.e. they will only work with scripts running on that host.)
PHP Flickr search (3)

- A function to construct a list of images from the response to the search.
- Accepts an XML object returned from the search and constructs a string of $number img tags

```php
function get_flickr_photos($flickr,$number) {
    $h = "";
    for ($i =0; $i < $number; $i++) {
        $photo = $flickr->photos->photo[$i];
        $pa = $photo->attributes();
        $imguri = "http://farm{$pa->farm}.static.flickr.com/{$pa->server}/{$pa->id}_{$pa->secret}_m.jpg";
        $photouri = "http://www.flickr.com/photos/{$pa->owner}/{$pa->id}";
        $h.= "<a href='{$photouri}'> <img src='{$imguri}'/> </a>";
    }
    return $h;
}
```
PHP Flickr search (4)

A function to get a URI parameter, defaulting to a value if none provided.

```php
function get_uri_parameter($name, $default) {
    if (isset($_REQUEST[$name]))
        return $_REQUEST[$name];
    else
        return $default;
}
```
PHP Flickr search (5)

The main program

The script accepts two parameters - the text to search for and the number of images required. It returns a div containing the images.

```php
<?php
require ("commonlib.php")
require ("flickrlib.php");

function get_uri_parameter (...) {
...

$text = get_uri_parameter("text","");
$number = get_uri_parameter("number",10);
$flickrXML = call_flickr(
    "flickr.photos.search",
    array("text=", rawurlencode($text),"sort=relevance")
);
echo "<div>" . get_flickr_photos($flickrXML,$number) . "</div>";
?>
```
Development Hints

- Essential: Use the isa server for debugging
- Change preferences in Explorer so that known filetypes are not hidden
- Use Firefox for debugging and Error Console for Javascript
- Insert echo statements at suitable points e.g. to show a constructed URI
- Test an example of any constructed URI directly in the browser
- Many errors are caused by the wrong case, missing semicolons and quotes.
- Use Google search to access documentation and advice on problems
Other APIs

- YouTube - search does not need an API key and results are in Atom format
- Google Maps API - needs an API key
  - Javascript API
  - REST API to geocode addresses e.g. [Prince Street Bristol](https://www.google.com/maps/place/Prince+Street+Bristol+)
  - [Static Google Maps](https://developers.google.com/maps/documentation/js/)
- see Dapper and Programmable Web for other APIs
Concepts

- Application Programmer Interface
- API documentation
- API Key
- Debugging tricks