

```
1 <?php
2 /*
3  * Visit each URL for product page in Ali feed (from step 1)
4
5  * Fake clicking the next page button to advance through all pages of
6  feedback (programmatically trigger the JS onclick() function).
7
8  * For each page of feedback, look at HTML source code and truncate the
9  code to toss out all the code before and after the feedback section.
10
11 * Then when working with just the feedback section, place each
12 "<tr>...</tr>" set of open & close table row tags as an element in an
13 array of reviews since each corresponds to an individual user's review
14
15 * Now for each review, just do a string search to look for an "<img
16 src=...>". If an image exists in the review, then go ahead and grab th
17 image's URL and store this image URL in an array with all the others
18 corresponding to that product.
19
20 */
21
22 // Visit each URL for product page in Ali feed
23 //
24 // Fake clicking next page button to loop over each feedback page
25 $start_time = time();
26
27 function trimTakeText($start_flag, $end_flag, $haystack)
28 {
29 //
30 //}
31 //}
32 // Toss out all code before & after feedback section
33 function getAllFeedback($url)
34 {
35     $data = file_get_contents($url);
36     $start_flag = 'thesrc="';
37     $end_flag = '"></iframe>';
38
39     $first_substr = substr($data, strpos($data, $start_flag) +
40 strlen($start_flag));
41
```

```
42     $second_substr = strstr($first_substr, $end_flag , TRUE);
43
44     $feedback_txt = '';
45     $feedback_count = 0;
46
47     for ($i=1; $i < 166; $i++) {
48         //$x = 1;
49         //
50         $iframe_data = file_get_contents($second_substr."&page=$i");
51         //echo $second_substr; return;
52         $start_flag2 = '<tbody>';
53         $first_substr2 = substr($iframe_data, strpos($iframe_data,
54 $start_flag2) + strlen($start_flag2));
55
56
57         $start_flag3 = '<tbody>';
58         $end_flag3 = '</tbody>';
59
60         $first_substr3 = substr($first_substr2, strpos($first_substr2,
61 $start_flag3) + strlen($start_flag3));
62
63         $feedback_txt .= strstr($first_substr3, $end_flag3 , TRUE);
64         //$x = $x + 1;
65         //echo 'x is now'.$x.'<br/>';
66         // }
67         return $feedback_txt;
68     }
69 }
70 function load_reviews_into_array($feedback_txt)
71 {
72     $delimiter = '<tr>';
73     $reviews_array = explode($delimiter, $feedback_txt);
74     return $reviews_array;
75 }
76 }
77 function countImg($review)
78 {
79     // <div class="pic-mark"> is only present when there's a
80     // user-uploaded pic
81     $needle = '<div class="pic-mark">';
82     $count = substr_count($review, $needle);

```

```

83     return $count ;
84 }
85 }
86 function getAllImg($reviews_array)
87 {
88     $product_pics = array();
89     foreach ($reviews_array as $key => $review) {
90         $count = countImg($review);
91         if($count > 0)
92         {
93             $start_flag = '<div class="pic-mark">';
94             $end_flag = '<div class="biPic"';
95             $first_substr = substr($review, strpos($review, $start_flg
+ strlen($start_flag));
96             $second_substr = strstr($first_substr, $end_flag , TRUE);
97
98             $second_substr_arr = explode(' ', $second_substr);
99             foreach($second_substr_arr as $k => $sec)
100             {
101                 $pos = strpos($sec, 'http://');
102                 if($pos !== false)
103                     $product_pics[] = $sec;
104             }
105         }
106     }
107 }
108 return $product_pics ;
109 }
110 }
111 $url =
112 'http://www.aliexpress.com/item/2014-Women-Elegant-Belted-Tartan-Long-
113 Sleeve-Patchwork-Tunic-Work-Business-Casual-Party-Bodycon-Pencil-Sheat
114 32219604594.html';
115 // $url =
116 'http://www.aliexpress.com/item/New-2015-Womens-Celebrity-Elegant-
117 //Vintage-Pinup-Bow-Ruch-Tunic-Business-Casual-Cocktail-Party-Business
118 //Bodycon/32298823720.html';
119 // $url =
120 'http://www.aliexpress.com/item/Hot-Selling-Brand-Valentine-Women-
121 //Calfskin-Pointed-Toe-Rivets-Flat-Shoes-Genuine-Leather-Flats-Sandals
122 //Big/32350592150.html?s=p';
123 // $i=31; 16sec $url =

```

```

124 // 'http://www.aliexpress.com/item/2015-Couture-Ball-Gown-Elegant-Weddi
125 //-Dress-Lace-Tulle-Plus-Size-Bridal-Gowns-Custom-Made-abiti/322884755
126 //.html';
127 // $url =
128 'http://www.aliexpress.com/item/2014-fashion-vintage-American-Appare
129 //jeans-woman-pencil-casual-denim-stretch-skinny-high-waist-jeans-pant
130 //32244547396.html';
131 // $url =
132 'http://www.aliexpress.com/item/Hot-Sale-Women-Pencil-Pants-Skinny-
133 //Zipper-Hollow-Out-Black-White-2015-New-Fashion-Casual-Slim/323223774
134 //.html?spm=2114.30010108.4.30.9a5eXh#feedback';
135 // $url =
136 'http://www.aliexpress.com/item/2013-Hot-sale-OP-20-Elegant-Mermaid-
137 //Sweetheart-Beaded-Organza-Chapel-Wedding-Dress-Party-dress-Custom/
138 //543863621.html';
139 // $url =
140 'http://www.aliexpress.com/item/Free-Shipping-Wholesale-Mix-Length-
141 //4pcs-lot-virgin-mongolian-afro-kinky-curly-hair-extensions-5A-
142 //unprocessed/1322839499.html';
143 // $url =
144 'http://www.aliexpress.com/item/Remy-Hair-Clip-In-Human-Hair-
145 //Extensions-Full-Sead-Set-27-Colors-available/1350432632.html';
146 //
147 $feedback_txt = getAllFeedback($url);
148 $reviews_array = load_reviews_into_array($feedback_txt);
149 $product_pics = getAllImg($reviews_array);
150 var_dump($product_pics);
151 foreach($product_pics as $k=>$pic)
152 {
153     echo '<br/>';
154 }
155 $end_time = time();
156 $time_elapsed = $end_time - $start_time ;
157 echo 'Time elapsed : '.$time_elapsed;
158 /*
159 function getHttpResponse($path, $data){
160     $options = array(
161         'http' => array(
162             'header' => "Content-type:
163 application/x-www-form-urlencoded\r\n",
164             'method' => 'POST',

```

```
165         'content' => http_build_query($data),
166     ),
167 );
168 $context = stream_context_create($options);
169 $result = file_get_contents($path, false, $context);
170 return $result;
171 }
172 }
173 function
174 $path = 'http://feedback.aliexpress.com/display/productEvaluation.htm?
175 $data = array(
176     '_csrf_token_' => '1b6npfcy3jl_', //why does it work
177 despite me not updating csrf token for each page?
178     "ownerMemberId" => "220754190",
179     "memberType"    => "seller",
180     "productId"     => "32322377471",
181     "companyId"     => "", //why does it work with companyId
182 being blank?
183     "evaSortValue"  => "Sort by comment",
184     "page"          => "1",
185     "i18n"          => "false",
186
187 );
188
189
190 $result = getHttpResponse($path, $data);
191 echo $result;
192 */
193 ?>
194
```