

DatiExpo

“DatiExpo” è stato realizzato da Davide Mori e Magdalena Kukiz per l'esame di “laboratorio e progettazione web” del corso di laurea di Informatica Umanistica dell'università di Pisa.

Si tratta di un progetto di “web data journalism”, un metodo a cavallo tra ricerca e inchiesta giornalistica che analizza e approfondisce un evento o un fenomeno con l'ausilio di database, applicazioni web e mappe interattive.

L'argomento del nostro progetto è l'esposizione universale avvenuta a Milano nell'anno 2015 dal tema “Nutrire il pianeta, energie per la vita”. L'obiettivo della nostra analisi è quello di fare una statistica a livello internazionale e nazionale italiano per quanto riguarda le aziende che hanno partecipato all'esposizione in veste di espositori o fornitori. Poiché la partecipazione ad un evento come Expo ha un impatto economico rilevante la nostra analisi verte a esaminare quali paesi abbiano partecipato più attivamente.

L'idea iniziale prevedeva un'analisi economica in relazione agli investimenti da parte delle imprese, ai costi sostenuti e ai possibili ricavi al termine dell'evento.

Tuttavia, ancora oggi, non si hanno dati certi sul ritorno economico di expo2015 pertanto la nostra indagine si è incentrata nella disamina del numero di aziende e della loro provenienza.

L'analisi è stata supportata da articoli giornalistici (es. Fatto Quotidiano, Wired).

Come primo step abbiamo esaminato a livello internazionale le aziende espositrici e quelle fornitrici grazie a dei grafici a torta, tramite l'ausilio delle librerie fornite da highcharts.

Una volta effettuata l'analisi abbiamo avuto modo di notare che l'Italia ha avuto un ruolo predominante in entrambi gli ambiti, così abbiamo optato per suddividere l'analisi ulteriormente.

Sia per quanto riguarda gli espositori che i fornitori abbiamo realizzato, sempre tramite l'ausilio delle librerie highcharts, altri due grafici a torta escludendo le aziende italiane.

Grazie a questo si evince in maniera più chiara ed immediata il numero delle aziende europee e di quelle extracontinentali.

Infine come ultima analisi a livello internazionale abbiamo unito in un grafico le due statistiche rilevate in precedenza, andando a delineare quali paesi abbiano partecipato attivamente ad expo2015 sia come espositori che in veste di fornitori.

A livello italiano la nostra analisi è stata resa possibile grazie ad una mappa interattiva. Abbiamo esaminato il numero delle aziende suddividendole per province; la mappa, di tipo

“choropleth” permette immediatamente all'utente di determinare le zone dalle quali provengono il

Maggior numero di imprese che hanno partecipato in veste di fornitori.

Come ultima cosa, relativamente al nostro argomento "Expo Milano 2015", abbiamo voluto dare all'utente che visita il nostro sito, una visione di ciò che è stata l'esperienza dell'esposizione universale. Questo è stato reso possibile grazie all'integrazione, in ultima pagina, di un piccolo motore di ricerca by hashtag di foto instagram.

Il tutto, creato con jquery e la documentazione ufficiale dell'API.

Le foto, editate dinamicamente dal foglio di stile presente, permettono all'utente di visualizzare immediatamente il numero dei "likes ricevuti" e altresì di accedere al contenuto originale sul social network.

Api utilizzate:

- Highchart
- Highmaps
- Instagram

Composizione del database:

- Tabella Fornitori
- Tabella Espositori

DATABASE:

Il nostro Database è composto da due tabelle: espositori e fornitori.

Espositori è articolato in : azienda e paese, mentre la tabella fornitori è composta da: ragione sociale, paese e provincia.

ESPOSITORI

AZIENDA	PAESE
Technovaa	AE
Darvesh	AE
Joma Kunststofftech nik	AT
Haas Food Equipment	AT
[..]	

FORNITORI

RAGIONE SOCIALE	PAESE	PROVINCIA
A. MANZONI & C. SPA	IT	MI
1 & 1 MODELS SNC DI ERMOLLI - BADER	IT	MI
1 a 100 DI LUCA GIUSEPPE STALLA	IT	MI
A.E.M. G. RECCHIA SRL	IT	RM
[..]		

I dati sono stati presi dal sito ufficiale <http://dati.openexpo2015.it/it>

Per poter inserire i dati del nostro database nel grafico, per avere il numero di aziende per paese, abbiamo utilizzato la seguente query:

```
8  
9  $sql = 'SELECT COUNT( DISTINCT `AZIENDA` ) AS num, `PAESE`  
10 FROM `ESPOSITORI` WHERE 1 GROUP BY `PAESE`  
11 ORDER BY num DESC ' ;  
12
```

Con la stessa query abbiamo estratto i dati per quanto riguarda i fornitori e per l'analisi fatta escludendo le aziende italiane.

Abbiamo, in questi casi, preso in analisi le due colonne, paese e azienda, e abbiamo contato tramite COUNT quante aziende, in numero, ha ogni paese. Perciò, per l'estrazione del numero di aziende abbiamo considerato le singole tabelle, e le query sono state fatte all'interno di esse.

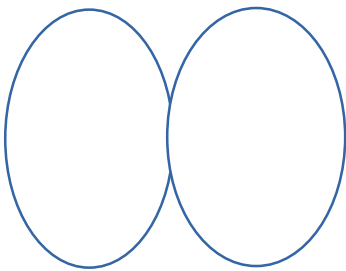
Poichè alcuni dati che abbiamo preso da OpenExpo non erano del tutto puliti, alcuni paesi non erano specificati ma indicati con un trattino (-), abbiamo dovuto escludere in alcune query le aziende determinate in questo modo, poiché non pertinenti alla nostra analisi.

Per quanto riguarda l'estrazione di informazioni per il grafico che unisce i dati degli espositori e fornitori, quindi somma il numero di aziende là dove sono dello stesso paese, e allo stesso tempo inserisce anche paesi che non abbiano avuto il doppio ruolo, dato comunque per noi rilevante, abbiamo riscontrato dei problemi nella somma del coefficiente numerico.

Così abbiamo eseguito metà della procedura tramite query, ovvero abbiamo eseguito l'unione tra le due tabelle, mentre l'altra parte, riguardante la somma, la abbiamo affrontata tramite il codice php:

UNIONE TRA DUE INSIEMI:

ESPOSITORI + FORNITORI = Un insieme contenente gli elementi di entrambi gli insiemi



```
$sql = 'SELECT COUNT( DISTINCT `AZIENDA` ) AS num, `PAESE` FROM `ESPOSITORI` WHERE 1
GROUP BY `PAESE` UNION SELECT COUNT( DISTINCT `RAGIONE SOCIALE` ) AS num, `PAESE`
FROM `FORNITORI` WHERE 1 GROUP BY `PAESE` ' ;
$unione = $db->query($sql);

$datiU = [];

while ($selemento = mysql_fetch_assoc($unione)){
    if( ! isset($datiU[$selemento["PAESE"]]) ){
        $datiU[$selemento["PAESE"]] = intval($selemento["num"]);
    }else{
        $datiU[$selemento["PAESE"]] += $selemento["num"];
    }
}
$resultato=array();
foreach ($datiU as $key => $value) {
    $resultato[] = array($key , $value);
}
```

Per quanto riguarda l'analisi a livello nazionale italiano, abbiamo preso in esame la tabella

fornitori, poiché nella tabella dei fornitori le aziende erano distribuite genericamente per paese, senza ulteriori informazioni sulla località o la provincia.

L'analisi verte ad individuare quante aziende per provincia hanno partecipato a expo2015 nel ruolo di fornitore.

La mappa interattiva è stata realizzata tramite highcharts; i dati sono stati estrapolati dalla tabella tramite la query sottostante.

```
$db = new DataBase();
$db->connetti();

$sql = 'SELECT COUNT( DISTINCT `RAGIONE SOCIALE` ) AS num, `PROVINCIA`
FROM `FORNITORI` WHERE `PROVINCIA` != "-" AND `PAESE` = "IT"
GROUP BY `PROVINCIA` ORDER BY NUM DESC ' ;

$sr = $db->query($sql);
$datiI=array();
while ($row = mysql_fetch_assoc($sr)){
    $datiI[] = array($row['PROVINCIA'], $row['num']);
}
```

Abbiamo modificato il codice in modo tale che i dati relativi al numero di aziende per provincia fossero inserite dinamicamente a partire dai dati presenti sul database.

Abbiamo creato un ciclo for che scandisce l'array contenente i risultati ottenuti con la query e unisce gli elementi dell'array associativo ([0] e [1]) alla chiave "hc-key:it-" predefinita da highcharts, e la stringa "value" che accompagna il numero di aziende:

```
var dati = [];
for(var i=0; i<data.length; i++){
    dati[i]= { "hc-key": 'it-'+data[i][0].toLowerCase(), "value": data[i][1]};
}
```

INSTAGRAM:

In conclusione al nostro progetto, abbiamo deciso di inserire una sorta di piccolo motore di ricerca di foto instagram. Sulla pagina social expo, vi è un campo di testo dove è possibile inserire l'hashtag desiderato e cliccando sul bottone cerca verranno visualizzati le venti foto più recenti.

Gli hashtag sono stati i più disparati, dal classico expo15, fino a 4expo, passando da

expomilano2015. Cercare di racchiudere tutte le foto più emblematiche dell'evento in un hashtag solo era pressochè impossibile e comunque avrebbe offerto una visualizzazione meno accattivante e limitata.

In questo modo è possibile cercare tutti gli hashtag più disparati in relazione a expo2015. E' inoltre possibile visualizzare immediatamente quanti "likes" ha ricevuto una foto ed è possibile accedere al contenuto originale per permettere all'utente di avere un'esperienza ancora più immersiva e interattiva.

Il problema maggiore riscontrato è stato quello che l'api di instagram non permette di visualizzare più di venti immagini per ogni hashtag di ricerca, nè di effettuare ricerche per hashtag multipli.

Sostanzialmente è stato questo che ci ha spinto ad optare per questo tipo di soluzione, piuttosto che

Una visualizzazione statica e predefinita di un set di immagini con un determinato hashtag.



STATO DELL'ARTE:

- socialmediahub.expo2015:



E' una pagina web dedicata a Expo Milano 2015 prettamente improntata sul lato social. Vengono visualizzati tweet inerenti all'Expo, foto e video. Tutto il materiale può essere altresì condiviso in rete tramite altri social network.

Il sito non prende in considerazione alcuna statistica, ne altro dato evidente e rilevante per ciò che riguarda il numero delle imprese e dei paesi coinvolti.

- Censimento popolazione [da www.istat.it]



E' un applicazione web realizzata dall'Istat, Istituto nazionale di statistica le cui attività comprendono: censimenti di popolazione, censimenti sull'industria, sui servizi e sull'agricoltura, indagini economiche ecc.

Questa applicazione è molto simile, nell'approccio di trattamento e di analisi dei dati, al metodo con il quale è stato realizzato il nostro progetto. Vengono utilizzate banche dati dalle quali vengono estratte le informazioni e rappresentate in forma di grafici o mappe interattive con relative informazioni.

-

