



News

Dr: Ehab Rushdy

WEB MAP NEWS



PROJECT NAME

WEB MAP NEWS

IT DEPARTMENT

DR:EHAB RUSHDI

DR : OSAMA EL KOMY

STUDENTS

- 1. AHMED MAHMOUD DAOUD**
 - 2. EMAN ALI EL SHAHAT**
 - 3. RANA ESSAM ABOZAHIA**
 - 4. MOHAMED AHMED GOMAA**
 - 5. MOHAMED THARWET MOHAMED**
 - 6. NAHLA HOSSAM EL SHEHRY**
-

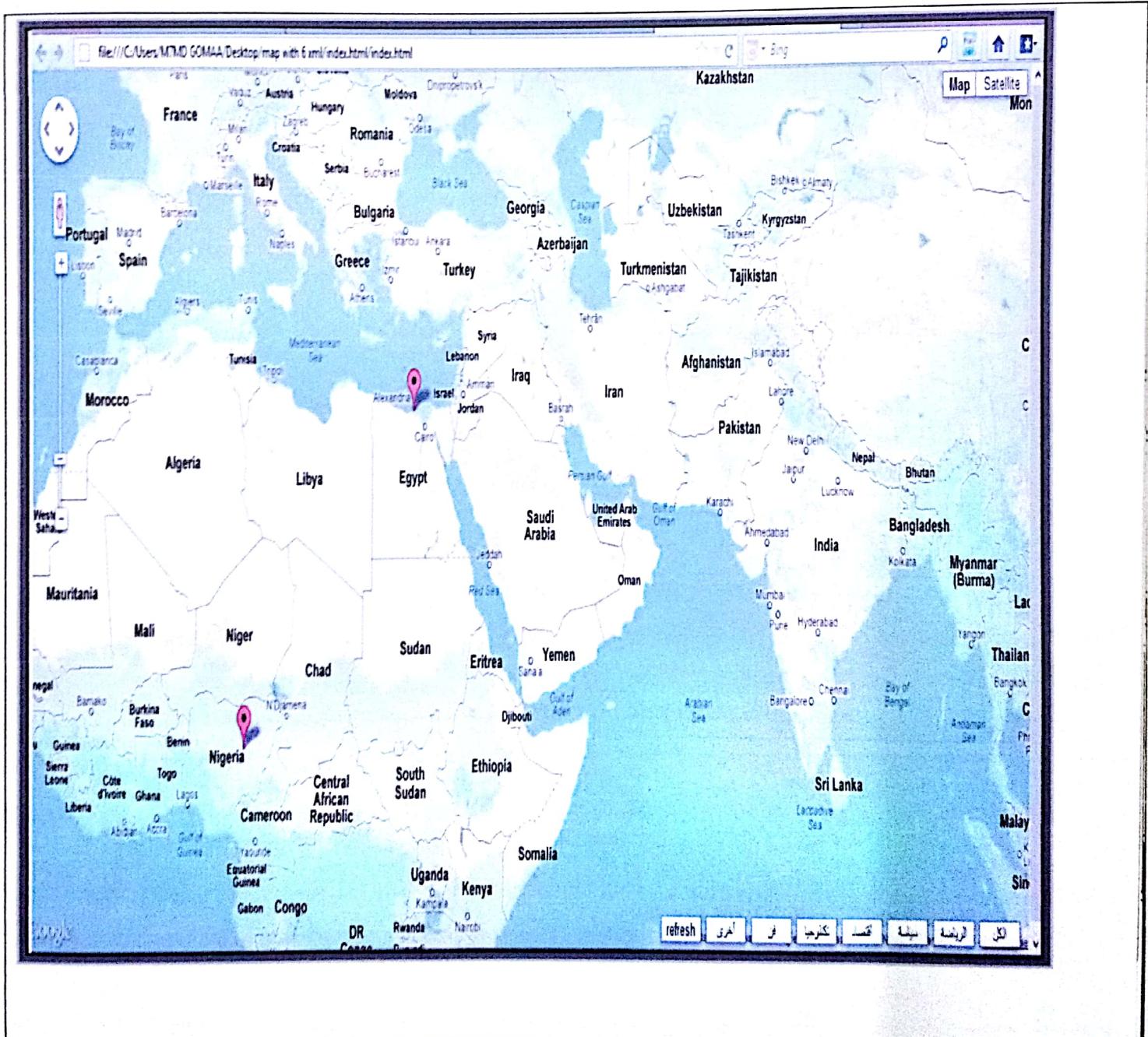
Chapter 1

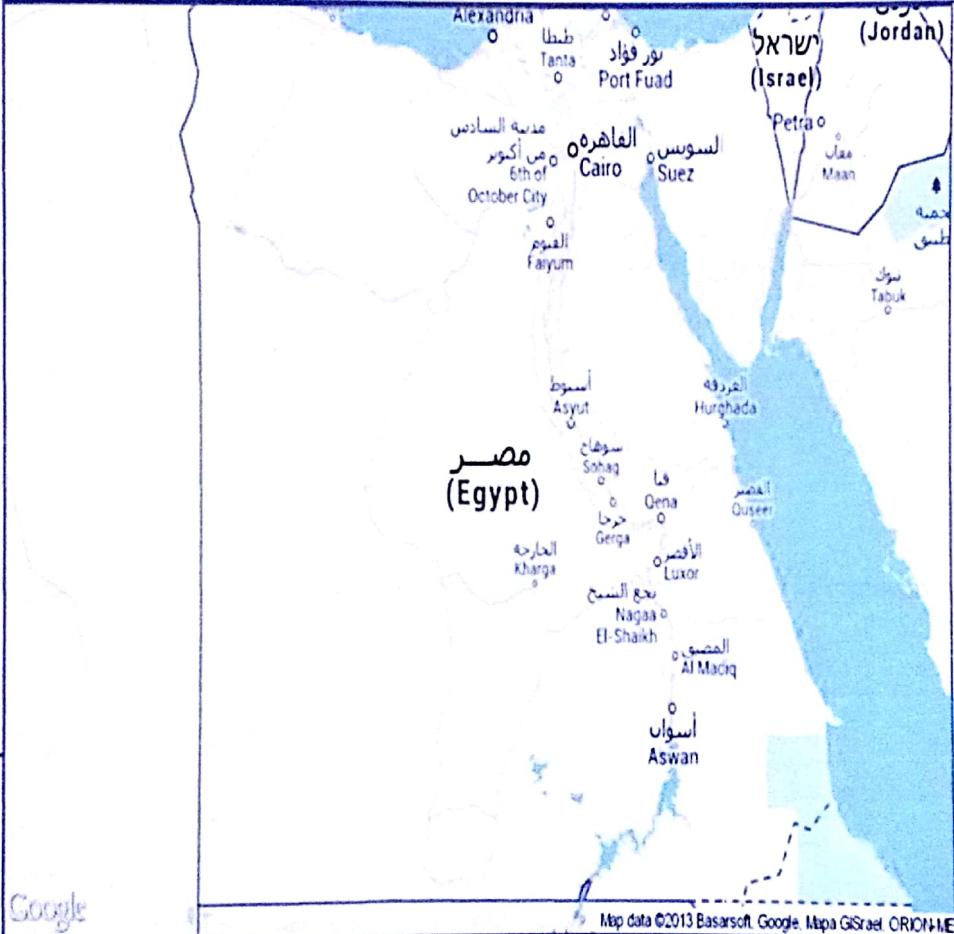
Introduction:-

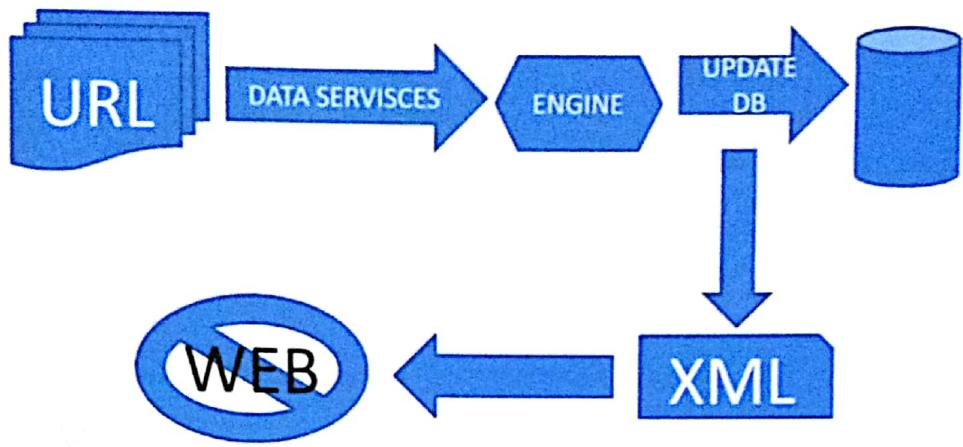
Idea of the project

The idea of the project is the work of Map news for all the governorates of Egypt

Where it appears on the map all the terms of villages and centers within each province of the political news and economic and technical

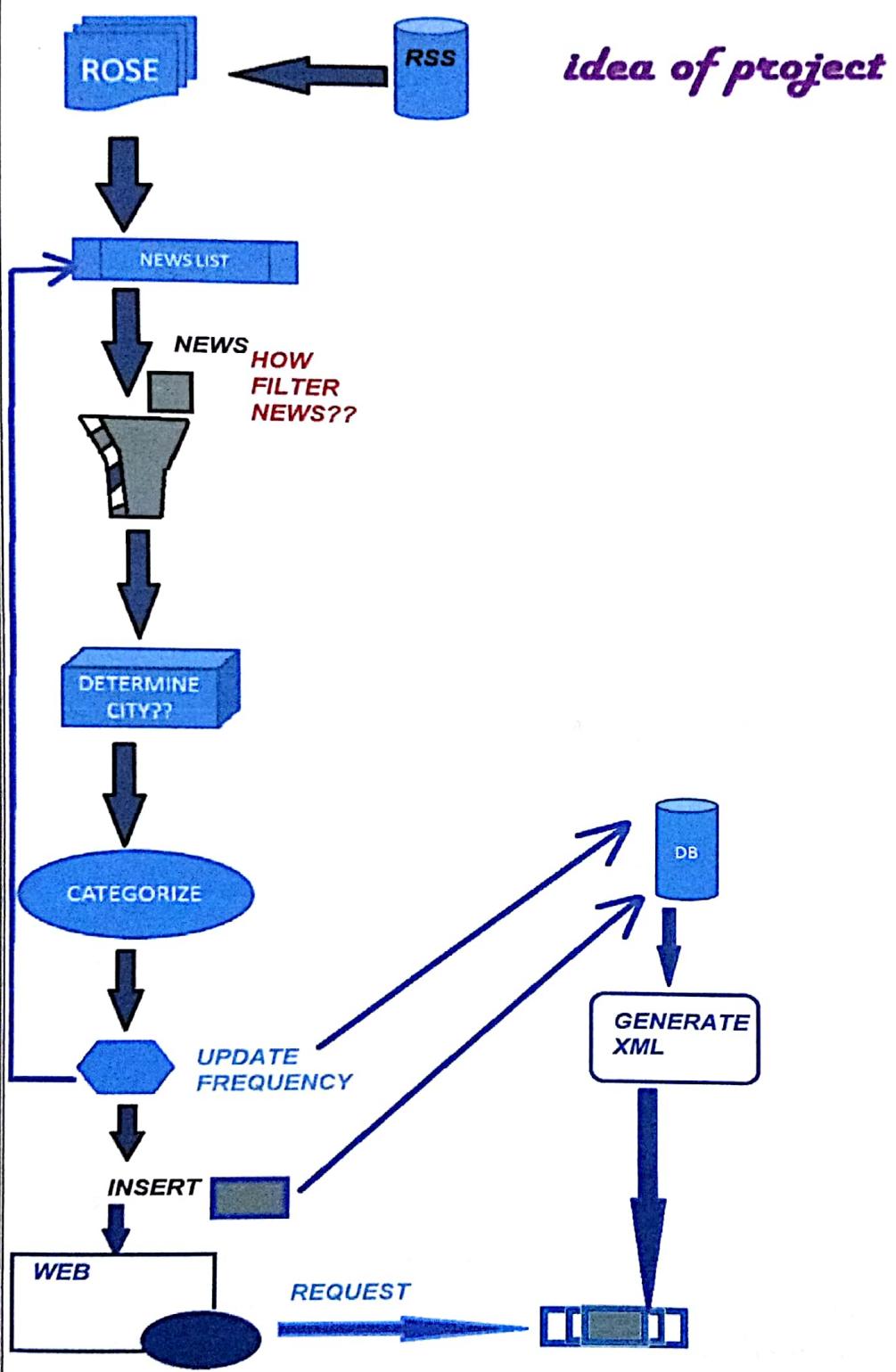






THIS IS DIAGRAM FOR:-

- 1- WHEN TAKE THE URL OF THE NEWS
- 2- DO DATA SERVICES WHICH UPDATE NEWS
- 3- FINALLY PUT THE XML TO THE WEB



THIS DRAWING FOR :-

- 1-REED RSS LINKS
 - 2-REED NEWS LIST
 - 3-FILTER NEWS
 - 4 -DETERMINE WHICH CITY THE NEWS ABOUT ??
 - 5-DETERMINE HOW CLASSIFY NEWS TO CATEGORIES??
 - 6-HOW UPDATE FREQUENCY??
 - 7-HOW INSERT DATA BASE ??
-

when collect sites that read RSS and read the RSS news

What is RSS?

A basic introduction to RSS FEEDS and aggregators for non-technical people from Software Garden , Inc.

What is RSS?

RSS stands for "Really Simple Syndication". It is a way to easily distribute a list of headlines, update notices, and sometimes content to a wide number of people. It is used by computer programs that organize those headlines and notices for easy reading.

What problem does RSS solve?

Most people are interested in many websites whose content changes on an unpredictable schedule. Examples of such websites are news sites, community and religious organization information pages, product information pages, medical websites, and weblogs. Repeatedly checking each website to see if there is any new content can be very tedious.

Email notification of changes was an early solution to this problem. Unfortunately, when you receive email notifications from multiple websites they are usually disorganized and can get overwhelming, and are often mistaken for spam.

RSS is a better way to be notified of new and changed content. Notifications of changes to multiple websites are handled easily, and the results are presented to you well organized and distinct from email.

How does RSS work?

RSS works by having the website author maintain a list of notifications on their website in a standard way. This list of notifications is called an "**RSS Feed**". People who are interested in finding out the latest headlines or changes can check this list. Special computer programs called "**RSS aggregators**" have been developed that automatically access the RSS feeds of websites you care about on your behalf and organize the results for you. (RSS feeds and aggregators are also sometimes called "**RSS Channels**" and "**RSS Readers**".)

Producing an RSS feed is very simple and hundreds of thousands of websites now provide this feature, including major news organizations like the New York Times, the BBC, and Reuters, as well as many weblogs.

What information does RSS provide?

RSS provides very basic information to do its notification. It is made up of a list of items presented in order from newest to oldest. Each item usually consists of a simple title describing the item along with a more complete description and a link to a web page with the actual information being described. Sometimes this description is the full information you want to read (such as the content of a weblog post) and sometimes it is just a summary.

For example, the RSS information for headlines on a local news website could contain the following information:

Item 1:

Title: Sidewalk contract awarded

Description: The city awarded the sidewalk contract to Smith Associates. This hotly contested deal is worth \$1.2 million.

Item 2:

Title: Governor to visit

Description: The governor is scheduled to visit the city on July 1st. This is the first visit since the election two years ago. The mayor is planning a big reception.

The RSS information is placed into a single file on a website in a manner similar to normal web pages. However, the information is coded in the XML computer language for use by a program (the RSS aggregator) and not by a person like a normal web page.

RSS aggregator programs

Think of an RSS aggregator as just a web browser for RSS content. RSS aggregators automatically check a series of RSS feeds for new items on an ongoing basis, making it is possible to keep track of changes to multiple websites without needing to tediously read and re-read each of the websites yourself. They detect the additions and present them all together to you in a compact and useful manner. If the title and description of an item are of interest, the link can be used to quickly bring the related web page up for reading.

Here is a screen shot of an RSS aggregator in action. On the left is a list of the RSS feeds being monitored, along with an indication of the number of unread items in each feed in parenthesis. On the right are the details of the most recent items in a selected RSS feed (in this case, the New York Times).

The screenshot shows a software interface for an RSS aggregator. On the left side, there is a list of 13 subscriptions with their names and the number of unread items in parentheses: Boston Globe -- City / Region News (10), CNET News.com, New York Times: NYT HomePage, Reuters: Top News (10), washingtonpost.com - Politics (8), and Wired News. There is also a link to 'Mark All Read'. On the right side, the New York Times homepage is displayed with a blue header bar containing the title 'New York Times: NYT HomePage', a subtitle 'Headlines, abstracts and links for the latest Ne', and the text 'Updated: Sun, 27 Jun 2004 02:03 PM'. Below this, a news article titled 'Houston Finds Some Pain In Car-R' by Shaila K. Dewan is shown, with a snippet of the article text: 'Last year, when Houston finally got a rail line,'.

There are many RSS aggregators available. Some are accessed through a browser, some are integrated into email programs, and some run as a standalone application on your personal computer.

How do I find out if a website has an RSS feed?

It is getting more and more common for websites to have RSS feeds. They usually indicate the existence of the feed on the home page or main news page with a link to "RSS", or sometimes by displaying an orange button with the letters "XML" or "RSS". RSS feeds are also often found via a "Syndicate This" link. Text "RSS" links sometimes (there are lots of variations) point to a web page explaining the nature of the RSS feeds provided and how to find them. The buttons are often linked directly to the RSS feed file itself.

Once you know the URL of an RSS feed, you can provide that address to an RSS aggregator program and have the aggregator monitor the feed for you. Many RSS aggregators come preconfigured with a list to choose from of RSS feed URLs for popular news websites.

HOW IS THE RSS feed file produced?

Unless you are maintaining a website or want to create your own RSS feed for some other purpose, how the RSS feed is produced should not be of concern and you may skip this section.

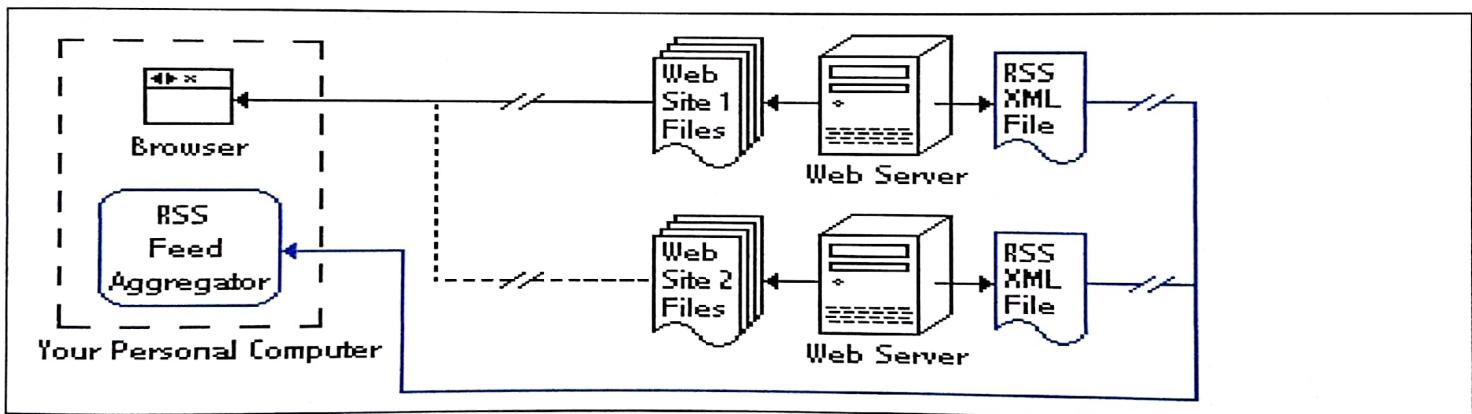
The special XML-format file that makes up an RSS feed is usually created in one of a variety of ways.

Most large news websites and most weblogs are maintained using special "content management" programs. Authors add their stories and postings to the website by interacting with those programs and then use the program's "publish" facility to create the HTML files that make up the website. Those programs often also can update the RSS feed XML file at the same time, adding an item referring to the new story or post, and removing less recent items. Blog creation tools like Blogger, Live Journal, Movable Type, and Radio automatically create feeds.

Websites that are produced in a more custom manner, such as with Macromedia Dreamweaver or a simple text editor, usually do not automatically create RSS feeds. Authors of such websites either maintain the XML files by hand, just as they do the website itself, or use a tool such as Software Garden, Inc.'s List Garden program to maintain it. There are also services that periodically read requested websites themselves and try to automatically determine changes (this is most reliable for websites with a somewhat regular news-like format), or that let you create RSS feed XML files that are hosted by that service provider.

Tying it all together

Here is a diagram showing how the websites, the RSS feed XML files, and your personal computer are connected:



The diagram shows a web browser being used to read first Web Site 1 over the Internet and then Web Site 2. It also shows the RSS feed XML files for both websites being monitored simultaneously by an RSS Feed Aggregator.

Other uses

In addition to notifying you about news headlines and changes to websites, RSS can be used for many other purposes. There does not even have to be a web page associated with the items listed -- sometimes all the information you need may be in the titles and descriptions themselves.

Some commonly mentioned uses are:

Notification of the arrival of new products in a store

Listing and notifying you of newsletter issues, including email newsletters

Weather and other alerts of changing conditions

Notification of additions of new items to a database, or new members to a group

One RSS aggregator is all that you need to read all of the RSS feeds, be they headlines, alerts, changes, or other notifications. RSS is shaping up to be a very popular and useful means for communicating.

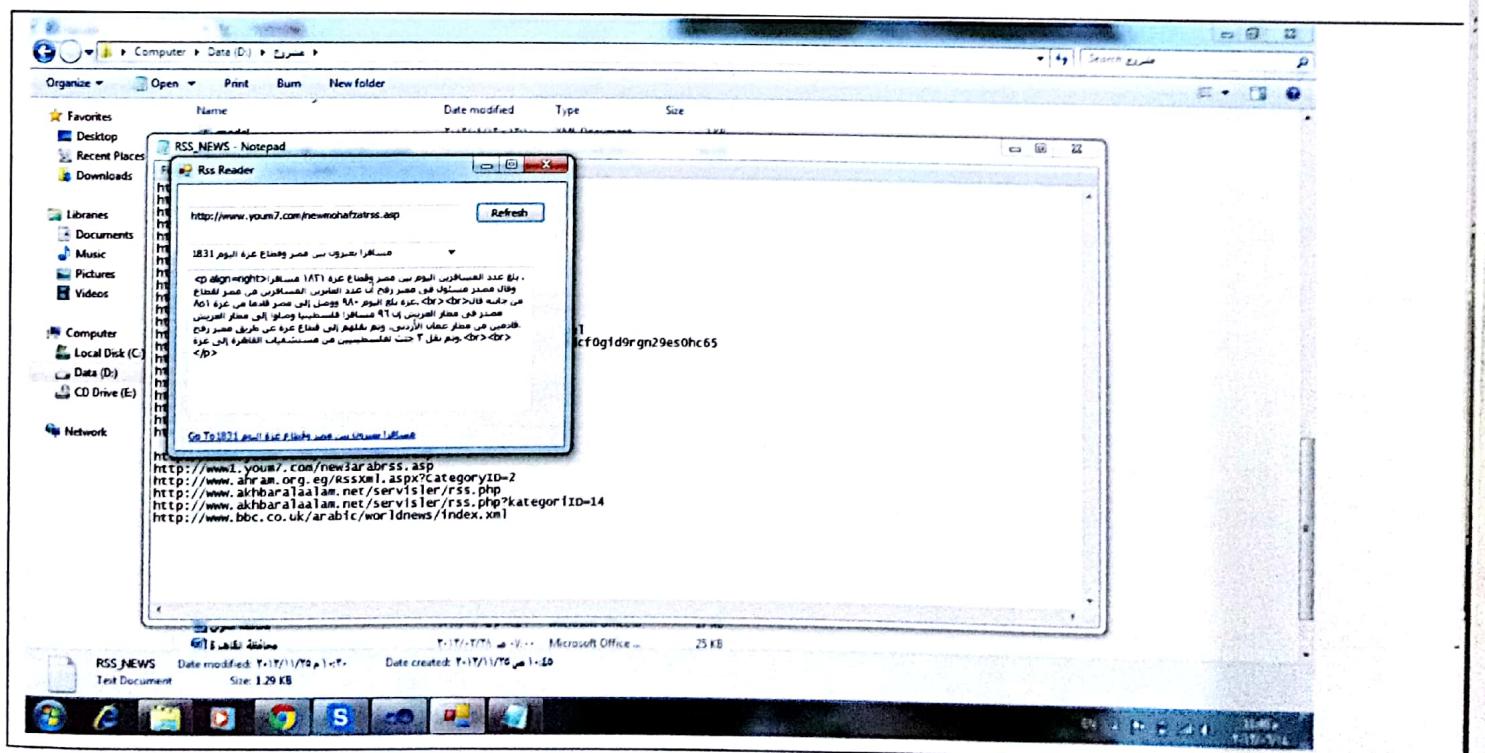
Feedster: A search engine for public RSS feeds that can also provide the results of a search as an ongoing RSS feed.

ListGarden: An open source program from Software Garden to create RSS feeds. For Windows, Mac OS X, and Linux.

Last updated: 6-Jul-2004 1:10 PM

Copyright (c) 2004 Software Garden, Inc.
All Rights Reserved.

Software Garden and Garden are registered trademarks of Software Garden, Inc.
List Garden is a trademark of Software Garden, Inc.



THEN AFTER READING RSS

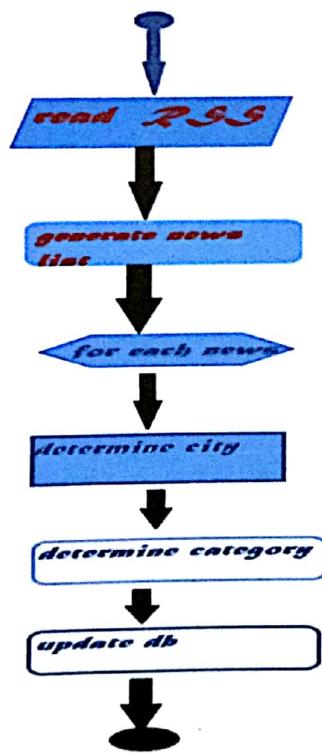
WE NEED TO CLASSIFY THE

NEWS .AND KNOW THE CITY WHICH THE

NEWS ABOUT AND RETURN ITS LATITUDE AND LAGITUDE .TO DO THAT WE NEED DATA BASE ABOUT ALL CITES AND VILLAGES .ALSO
NEED DATA BASE ABOUT CLASSIFICATION
THE NEWS

HOW TO FILTER NEWS?

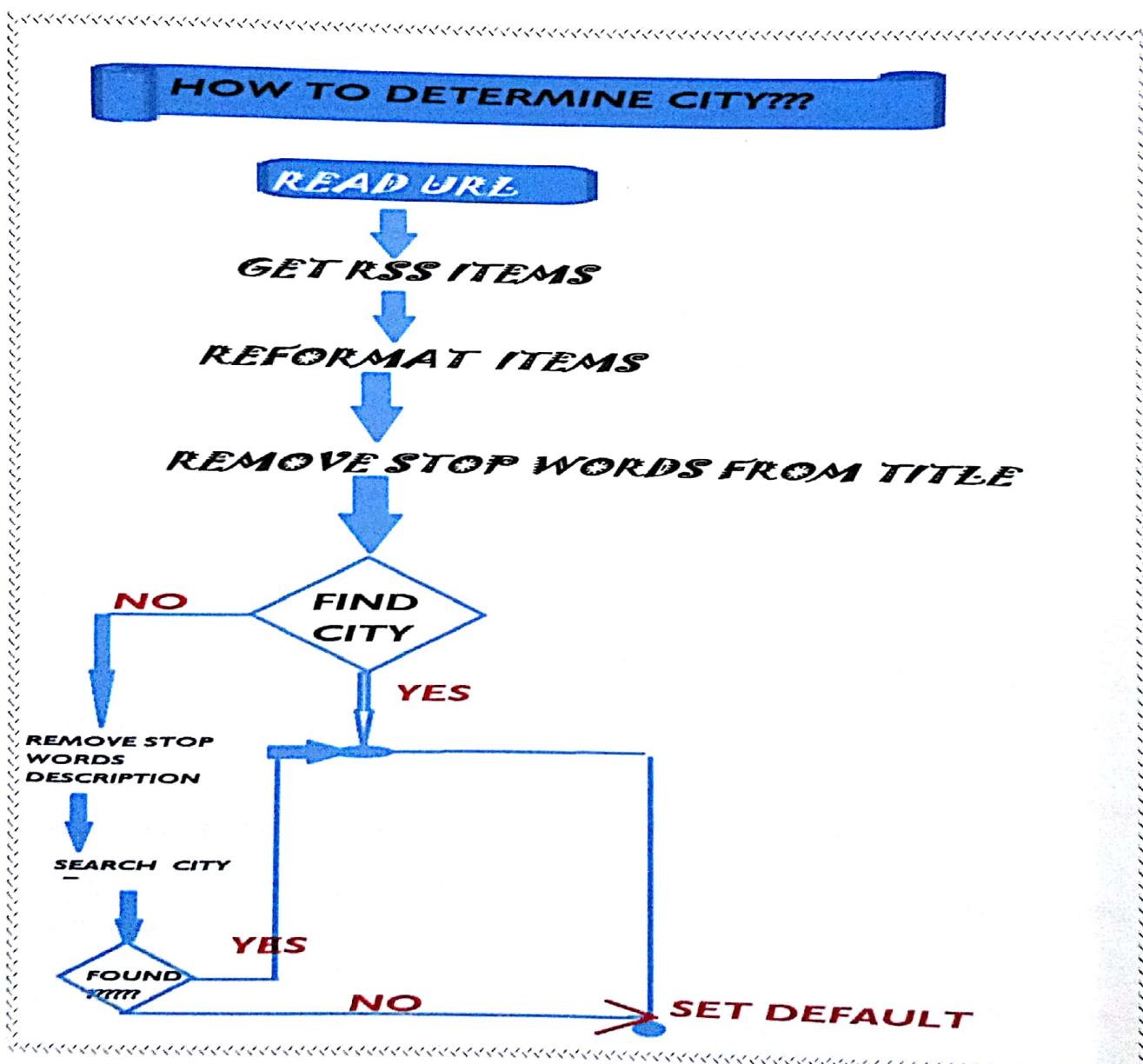
flowchart (web services)



THIS DRAWING FOR WINDOWS SERVICES :-

HOW FILTER NEWS ??

- 1- READ RSS NEWS
- 2- GENERATE NEWS LIST
- 3- FOR EACH NEWS DETERMINE THE CITY
- 4- CLASSIFY NEWS THEN UPDAATE THE DATA BASE



chapter 2

Data mining

What data mining is?

- The automatic search of patterns in huge databases, using computational techniques from statistics, machine learning and pattern recognition;
- The non-trivial extraction of implicit, previously unknown and potentially useful information from data;
- The science of extracting useful information from large datasets or databases;
- The automatic or semi-automatic exploration and analysis of large quantities of data, in order to discover meaningful patterns

• What is not data mining ?

What data mining could not be: Searching for particular information on Internet (about cooking on Google).

identify three characteristic steps of the data mining process:

1. *Exploring data*, consisting of data 'cleansing', data transformation, dimensionality reduction, feature subset selection, etc.;
 2. *Building the model and its validation*, referring to the analysis of various models and choosing the one who has the best performance of forecast –*competitive evaluation of models*;
 3. *Applying the model* to new data to produce correct forecasts/estimates for the problems investigated.
-

a classification process is characterized by:

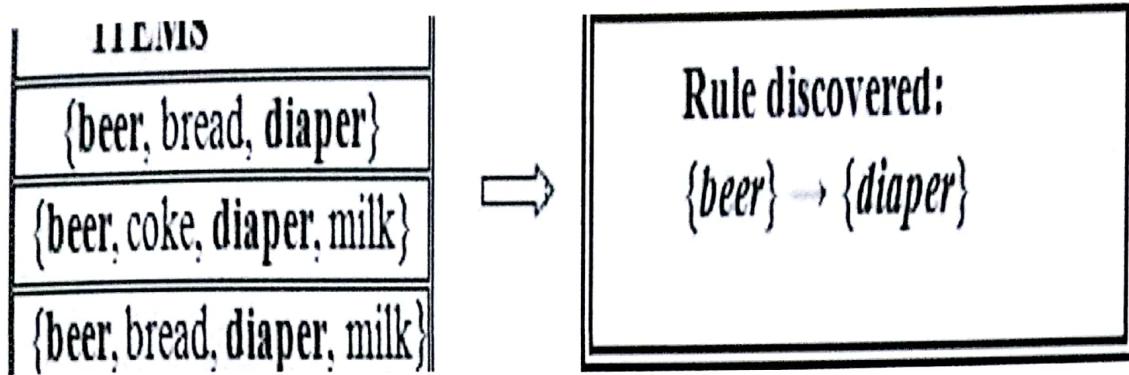
- **Input:** a training dataset containing objects with attributes, of which one is the class label ;
- **Output:** a model (classifier that assigns a specific label for each object (classify the object in one category), based on the other attributes;
- **The classifier :** is used to predict the class of new, unknown objects. A testing dataset is also used to determine the accuracy of the model.

Cluster Analysis

By *clustering* we mean the method to divide a set of data (records/tuples vectors/instances/objects/sample) into several groups (clusters), based on certain *Document clustering*, which aims to find groups of documents that are similar to each other based on the important terms appearing in them, based on their similarity, usually determined using the frequency with which certain basic terms appear in text

Association Rule Discovery

In principle, by the *association rule discovery/association rule learner* we understand the process of identifying the rules of dependence between different groups



Sequential Pattern Discovery

In Meteorology -at a general scale- discovering patterns in global climate change (see global warming, for instance), or particularly, discovering the occurrence moment of hurricanes, tsunamis, etc., based on previous sequences of events.

Regression analysis relates in principle to:

- Determination of a quantitative relationship among multiple variables;
- Forecasting the values of a variable according to the values of other variables

Deviation/Anomaly Detection

The detection of deviations/anomalies/outliers, as its name suggests, deals with the discovery of significant deviations from 'normal behavior'. Fig. 1.12 below suggests
evely illustrates the existence of anomalies in data.

Data Mining Applications

in banking, data mining methods were intensively used (and are still successfully used) in:

- modeling and forecasting credit fraud;
- risk assessment;

- trend analysis;
- support for direct marketing campaigns.

What Are Data ?

"Raw" data, as they were directly obtained by various processes of acquisition refer to numbers figures , images, sounds These data, once collected, are then processed, thus information that is stored,

Types of Datasets

- *Records:*

- Data Matrix; – Document Data; – Transaction Data.

- *Graphs:*

- World Wide Web (WWW); – Molecular Structures

- *Ordered datasets:*

- Spatial Data; – Temporal Data;

- Sequential Data; – Genetic Sequence Data

-Record data consist of collections of records, each record (object/sample) being characterized by a set of attributes. In principle, each object-data has a fixed number of attributes

Data Quality

- Noise and outliers;
- Missing values;
- Duplicate data

Types of Attributes

Numerical data. Numerical data (quantitative data) are themselves of two kinds:

discrete data and continuous data

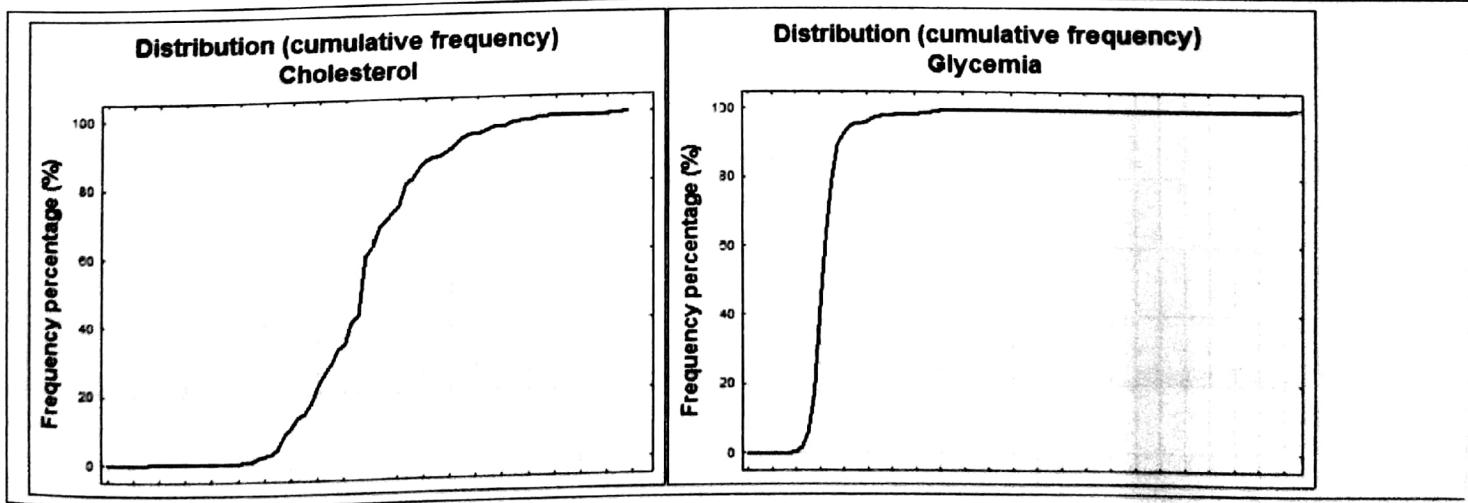
Categorical data. Unlike the numerical data

1. male/female;
2. married/single /divorced;
3. smoker/non-smoker

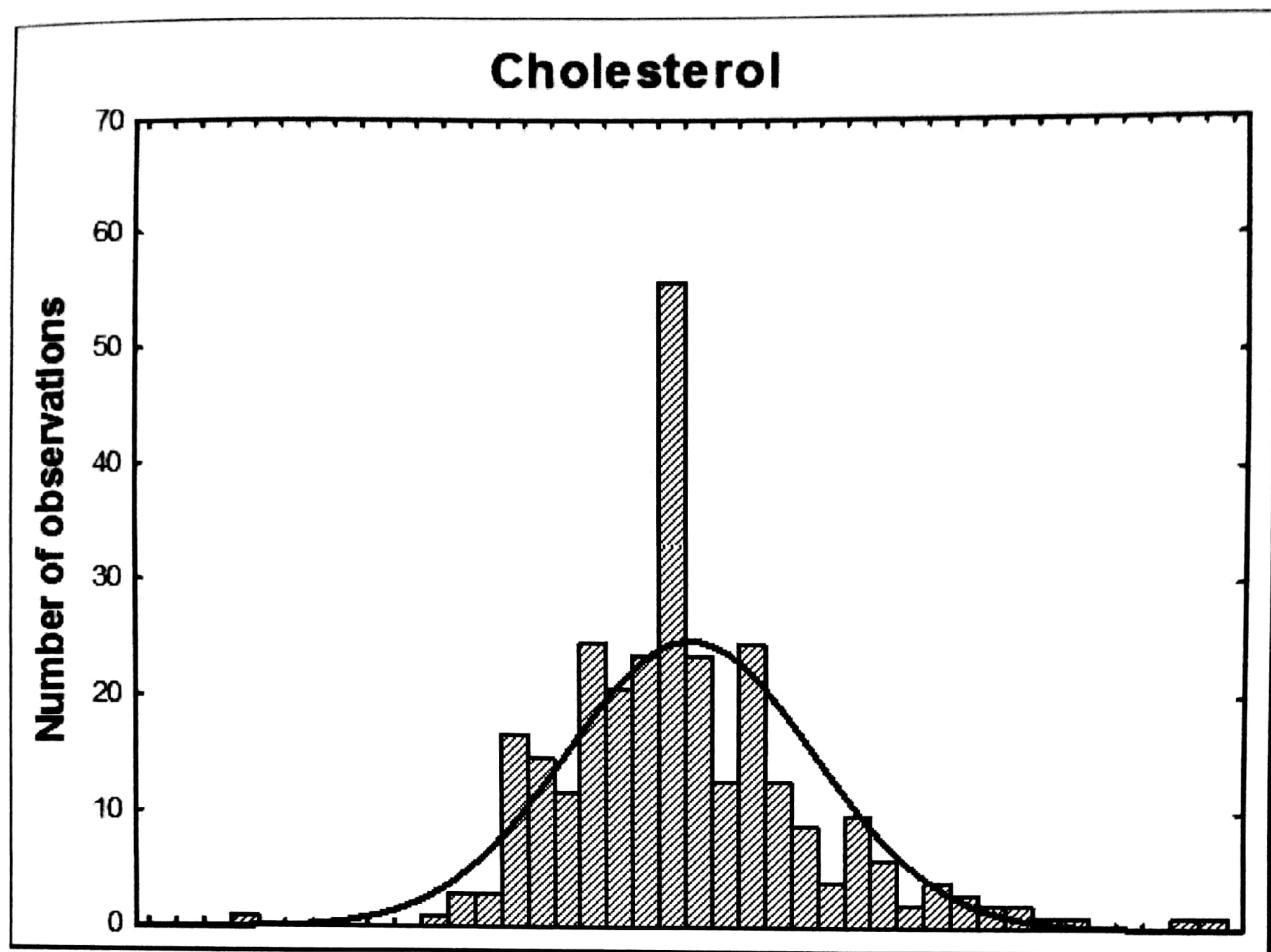
Decision tree

Descriptive Statistics Parameters

Descriptive statistics aims to summarize a large number of observations concerning a set of objects, using different specific methods, thus highlighting the main features of their attributes. There are two main ways to achieve this goal: either using a simple graphical representation, about which we shall speak later during this chapter, or by using numerical representations containing the main statistical



Cholesterol



Data Visualization

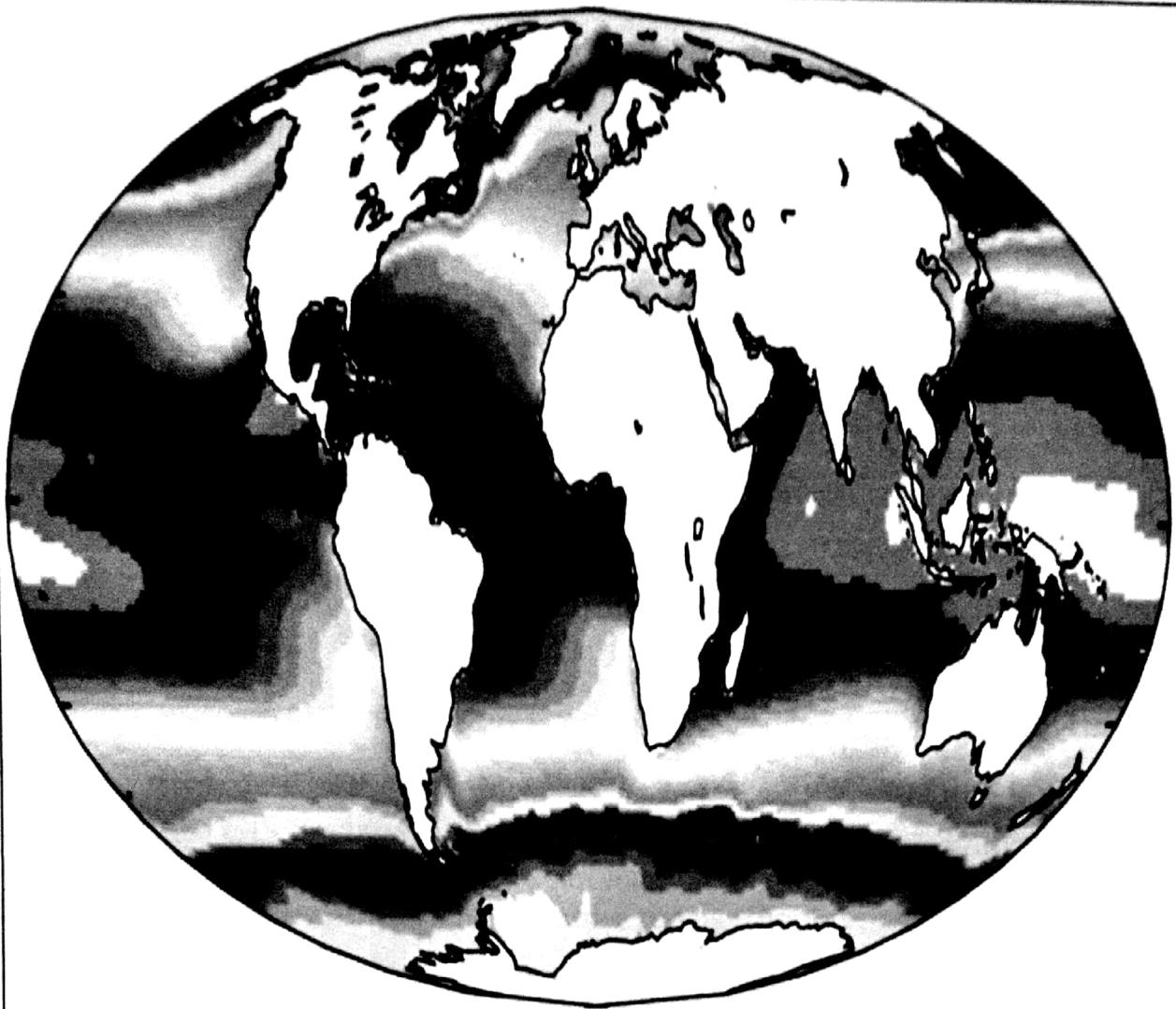
governorate_id	governorate_name	capital	distinct_id	distinct_name	governorate_id	village_id	village_name	district_id
1	الإسكندرية	الإسكندرية	1	حي الجمرك	2	1	شياحة بو شوشة	1
2	القاهرة	القاهرة	2	حي غرب الإسكندرية	2	2	شياحة البركة	1
3	الدقهلية	المنصورة	3	حي العاشرية	2	3	شياحة البلاعورية	1
4	الإسكندرية	الغردقة	4	حي المتنزه	2	4	شياحة التمزازة	1
5	الإسكندرية	المنوفية	5	حي شرق الإسكندرية	2	5	شياحة العجاري	1
6	الإسكندرية	المنوفية	6	حي وسط الإسكندرية	2	6	شياحة الحلواني	1
7	الإسكندرية	الفيوم	7	حي الصحوة	2	7	شياحة النسابة	1
8	الإسكندرية	الإسماعيلية	8	مركز منطلا	3	8	شياحة الشمرلي وزاوية الأصرح	1
9	الإسكندرية	الجيزة	9	مركز السسطة	3	9	شياحة العيادين	1
10	الإسكندرية	المنوفية	10	مركز المحلة الكبرى	3	10	شياحة العزاز	1
11	الإسكندرية	المنوفية	11	مركز بسيون	3	11	شياحة المخواش	1
12	الإسكندرية	المنوفية	12	مركز ريزقى	3	12	شياحة حارة موزدة	1
13	الإسكندرية	المنوفية	13	مركز سمنود	3	13	شياحة رأس النبى	1
14	الإسكندرية	المنوفية	14	مركز قطور	3	14	شياحة القلبانية	1
15	الإسكندرية	المنوفية	15	مركز كفر الزيات	3	15	شياحة زاوية حطاب	1
16	الإسكندرية	المنوفية	16	المنصورة	4	16	شياحة سوق السماد الغربية	1
17	الإسكندرية	المنوفية	17	أجا	4	17	شياحة صقر باشا	1
18	الإسكندرية	المنوفية	18	المنبلتون	4	18	شياحة قبو العلاح	1
19	الإسكندرية	المنوفية	19	العلالية	4	19	قسم مينا البصل	2
20	الإسكندرية	المنوفية	20	المنزلة	4	20	قسم البابان	2
21	الإسكندرية	المنوفية	21	بنقاش	4	21	قسم الجنشية	2
22	الإسكندرية	المنوفية	22	دكورة	4	22	قسم العازورية	3

	Results	Messages	
	id	category	
1	1	رياضة	
2	2	سياسية	
3	3	اقتصاد	
4	4	الحوادث والقضايا	
5	5	فنية	
6	6	تكنولوجيا	
7	7	أخرى	

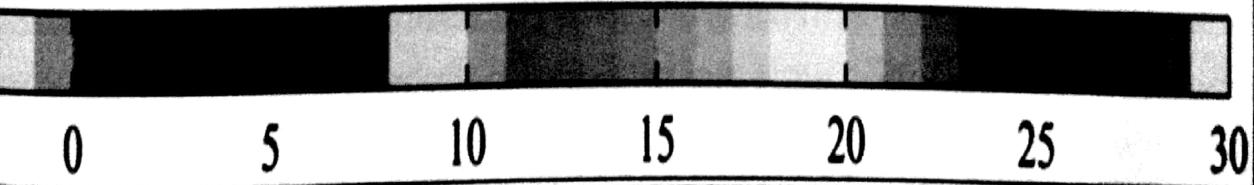
	word	categoryword	rank	rec	category
1	جي	ني	NA	NA	
2	معن	ني	NA	NA	
3	الكل	ني	NA	NA	
4	دلي	ني	NA	NA	
5	ألا	ني	NA	NA	
6	الأبله	ني	NA	NA	
7	حفل	ني	NA	NA	
8	البيان	ني	NA	NA	
9	الآباء	ني	NA	NA	
10	كأس	ني	NA	NA	
11	جزء	ني	NA	NA	
12	الأفغان	ني	NA	NA	
13	الم	ني	NA	NA	
14	دون	ني	NA	NA	
15	الله	ني	NA	NA	
16	المرأة	ني	NA	NA	
17	آباء الآباء	ني	NA	NA	

This is examples of data base

Data visualization is one of the most powerful and appealing techniques for data exploration, one of the fundamental pillars of the EDA. Unlike the elementary graphical representation that we talked about in the previous subsection, which is practically attached to the statistical description of data, the visualization techniques



Sea-surface temperature [$^{\circ}\text{C}$]



0 5 10 15 20 25 30

Chapter 3

BACKGROUND

Web mapping

is the process of designing, implementing, generating and delivering maps ON the World Wide Web and its product. While web mapping primarily deals with technological issues, web additionally studies theoretic aspects: the use of web maps, the evaluation and optimization of techniques and workflows, the usability of web maps, social aspects, and more. Web is similar to web mapping but with an emphasis on analysis, processing of project specific and exploratory aspects. Often the terms web GIS and web mapping are used synonymously, even if they don't mean exactly the same. In fact, the border between web maps and web GIS is blurry. Web maps are often a presentation media in web GIS and web maps are increasingly gaining analytical capabilities.

A special case of web maps are mobile maps, displayed on devices, such as "The use of the web as a dissemination medium for maps can be regarded as a major advancement in cartography and opens many new opportunities, such as real time maps, cheaper dissemination, more frequent and cheaper updates of data and software, personalized map content, distributed data sources and sharing of information. It also implicates many challenges due to technical restrictions (low and limited in particular with mobile computing devices, many of which are physically small, and use slow connections), and security issues, reliability issues and technical complexity. While the first web maps were primarily static, today's web maps can be dynamic and integrate multiple media.

Learn the history and applications of web mapping.

Web mapping involves the design, creation, and distribution of maps using the world wide web. The technology that enables web mapping was first introduced in 1993, with the advent of dynamically generated content and functionality allowing hyperlinked images. The use of the web as a medium for mapping is a major breakthrough in cartography. The web medium allows for more frequent, cheaper, and efficient updates of data, personalized map content and facilitates the sharing of geographic information. Before web mapping cartography as an activity was restricted to a few companies and organizations and required expensive tools and skilled labor. The widespread availability of online mapping tools and geodata have enabled a greater number of people to produce maps.

Web Maps

The primary distinction between web map types is between static and dynamic maps. While the first web maps were primarily static due to technical restrictions, today's web maps can be fully interactive and often integrate multiple forms of media. Static web maps are fixed images and non-interactive. These maps are similar to paper maps and may not be optimized for screen viewing. Dynamic maps are created on demand each time a webpage is accessed, usually from information stored in a database. They are often interactive, allowing users to change parameters such as zoom and the area of view.

Broader categories of web maps include the following:

Collaborative web maps rely upon content creation by a large number of users. Examples include OpenStreetMap and Google Map Maker.

Customizable web maps are designed to be modified or added to and used in others people's web pages. Examples include the Open Layers Framework and Google Maps.

Distributed web maps are created from multiple independent data sources.

Personalized web maps allow users to filter map data and select the content to be displayed.

Realtime maps display time-dependent information on demand. Examples are weather maps and traffic maps.

History

Interactive maps have been one of the most significant and practical uses of the internet, and the development of web mapping closely mirrors the development of the world wide web. In 1993 released the first web mapping service, the Xerox PARC Map Viewer, which enabled other people to create sites that combined their own content with others' base maps.

In the following years other mapping services began to appear which allowed for this activity, known as mashups. Among the notable early services was MapServer, which was created by the Census Bureau. In 1995 ForNet at the University of Minnesota was launched, which incorporated the first widely-used, free, open-source web mapping software.

, the first commercial online map service, was released in 1996. Terraserver, a popular web map service by the USGS, Microsoft, and HP using aerial images and USGS graphics, was released in this period as well.

The early 2000s marked the introduction of OpenStreetMap, a web-based collaborative project to create

a world map under free licence, and in 2005 Google Maps was introduced, allowing the easily integration of map services into existing websites.

Benefits of Web Maps

Web mapping has a number of advantages over traditional mapping forms. These include:

Affordability- Web server hardware and tools for producing web maps are either relatively inexpensive or free. Products can be distributed and reproduced either at no cost or at very limited expense.

Collaboration- Web maps enable greater collaboration between users GOOGLE EARTH is an example of a collaborative tool, allowing various users to share and disseminate information.

Integration- It is possible to integrate other forms of media, such as photos, into web maps.

Real time information- Web maps can easily deliver timely information and update as new information is available.

Lessons

Lesson: Differentiating between map types

Objective: Be able to find and identify different types of web maps

Group Size: individual

Materials: computer, internet

Instructions:

1. Go to www.google.com
2. Search for at least ten mapping sites.
3. Mark the name, address, and type for each site (collaborative; customizable; distributed; interactive; personalized; real time). Determine which category of site is the most popular. Note that some sites may fit into more than one category.

Cartography refers to the practice of making maps.

Dynamic Maps allow users to see multiple forms of data in a single map, enabling them to better understand the interaction between data types.

GIS, short for Geographic Information System, is an electronic system used to store, retrieve, and archive map data.

Static Maps are non-changing, non-interactive images.