Chapter 25
Geospatial Data Management with Terrafly

**Naphtali Rishe, Borko Furht, Malek Adjouadi, Armando Barreto,
Evgenia Cheremisina, Debra Davis, Ouri Wolfson, Nabil Adam,
Yelena Yesha, and Yaacov Yesha**

1 Introduction

> *March 11, 2011, 2:46 pm – A magnitude 9.0 earthquake strikes just off the northeast coast of Japan. Within an hour, a tsunami estimated at over 30 feet high hits the coast of Japan, sweeping away entire villages and killing tens of thousands of people* [1].

N. Rishe (✉) • M. Adjouadi • A. Barreto • D. Davis
NSF Industry-University Cooperative Research Center for Advanced Knowledge, Enablement (CAKE.fiu.edu) at Florida International, Florida Atlantic and Dubna Universities, Miami, FL, USA
e-mail: rishe@fiu.edu; adjouadi@fiu.edu; Armando.Barreto@fiu.edu; dledavis@cs.fiu.edu

B. Furht
NSF Industry-University Cooperative Research Center for Advanced Knowledge, Enablement (CAKE.fiu.edu) at Florida International, Florida Atlantic and Dubna Universities, Boca Raton, FL, USA
e-mail: borko@cse.fau.edu

E. Cheremisina
NSF Industry-University Cooperative Research Center for Advanced Knowledge, Enablement (CAKE.fiu.edu) at Florida International, Florida Atlantic and Dubna Universities, Moscow, Russia
e-mail: e.cheremisina@geosys.ru

O. Wolfson
Computational Transportation Science Program (CTS.cs.uic.edu), University of Illinois at Chicago, Miami, FL, USA
e-mail: wolfson@cs.uic.edu

N. Adam
Science and Technology Directorate, U.S. Department of Homeland Security (DHS.gov), Washington, DC, USA
e-mail: Nabil.Adam@dhs.gov

Y. Yesha • Y. Yesha
NSF Industry-University Cooperative Research Center for Multicore Productivity, Research (CHMPR.umbc.edu) at the University of Maryland Baltimore County, Baltimore, MD, USA
e-mail: yeyesha@umbc.edu; yayesha@umbc.edu

B. Furht and A. Escalante (eds.), *Handbook of Data Intensive Computing*,
DOI 10.1007/978-1-4614-1415-5_25, © Springer Science+Business Media, LLC 2011

> *April 20, 2010 – A massive explosion on the BP Deepwater Horizon drilling platform in the Gulf of Mexico caused the largest oil spill in US history, killing 11 workers and spilling an estimated 4.9 billion barrels of oil* [2].

Disasters, whether environmental or manmade, have catastrophic impacts that require both quick action and long-term interventions. Mitigating the effects of those disasters requires knowledge about similar events and advanced disaster planning. Major challenges in disaster planning and intervention include a lack of up-to-date information on situational and environmental conditions, major communication gaps, and a lack of effective coordination in planning and recovery operations. With a worldwide average of 387 natural disasters a year alone [3], it is imperative that solutions are found to combat and eliminate these problems. The implementation of cutting-edge technology, in particular, is key to advancement of the science and solutions for disaster mitigation [4, 5].

1.1 The Challenge

Implementing the types of cutting-edge technology needed for the diverse needs of disaster-related solutions is complex and challenging. Massive amounts of data are required, and this data is often heterogeneous, from divergent sources, and consists of both structured and unstructured data. Geospatial and remotely-sensed data, such as geo-referenced satellite imagery and aerial photography, provides particularly critical information that either is not available in other forms, or is not otherwise easily conveyed. This type of data, however, is inherently very large, thus significantly increasing the complexity of possible technological solutions [6–10].

There are also numerous challenges with the use of existing GIS tools for processing and analyzing geospatial data [8]. The primary challenges involved include:

1. The use of multiple, disparate tools are often necessary, some of which are expensive, and require specialized skills and training to use;
2. The data must often be imported into these separate tools, each of which may require the data to be in different formats; and
3. The ability to combine heterogeneous types of data is not always possible using the GIS tools currently available.

Another difficulty is that the typical end-users' technological backgrounds are very diverse, ranging from scientists to disaster mitigation and recovery planners to on-the-ground disaster responders. This requires the availability of robust systems and tools that are very flexible and easy to use, so that end-users can focus on the work that they need to get done without having to worry about the technology behind the system they are using [11, 12]. Because of the nature of disasters, new and updated information needs to be made available in a timely manner, and in a form that can quickly and efficiently be consolidated and conveyed to multiple, diverse users at the same time.

Interestingly, the challenges faced by those in the disaster mitigation field are not unique. There are numerous fields that rely on the use of geospatial and remotely sensed data. The same types of technological solutions designed for disaster mitigation can also be implemented in and have major impacts on other fields. The most apparent are fields in scientific discovery such as ecological and environmental research, archeology, oceanography, and meteorology, among many others. Other fields that have more ordinary, day-to-day impact on people's lives would also greatly benefit from similar solutions. This includes areas ranging from government operations such as public works and urban planning to business fields such as real estate and tourism.

For example, in real estate, there are many variables that affect the value and desirability of a particular property: neighborhood, local businesses, crime rates, roads and transportation, school quality, level of urbanization, etc. As with disaster preparedness and recovery, the data necessary to provide needed information often comes from diverse, heterogeneous data sources that may be structured or unstructured. Geospatial data, once again, can provide information that is otherwise not available or in a format that is difficult to interpret. All this data must be collected, consolidated, analyzed and visualized in an easy to understand format to be effectively utilized. This challenge is further compounded in that the technical backgrounds of potential users are very diverse, ranging from engineers and surveyors, to developers, real estate agents and potential buyers.

1.2 A Solution: TerraFly

Identifying a solution to these intense and complex issues may seem overwhelming. Through the use of innovative research and cutting-edge technology, TerraFly has been created to provide a flexible, robust and forward-thinking solution to these multifaceted challenges. Specifically, TerraFly is designed to efficiently and effectively deal with the challenges involved in handling and analysis of massive amounts of heterogeneous geospatial and related data, as well as with the challenges that users encounter when attempting to use traditional GIS tools.

This article discusses geospatial data analytics using TerraFly as a case study. An overview of TerraFly will first be presented, followed by discussions of TerraFly's capabilities in data handling, information visualization, flexibility and customization, and specific domains in which TerraFly is currently used, such as disaster planning and recovery, science and research, real estate and travel.

TerraFly is a technology and tools for visualization and querying of geospatial data. The visualization component of the system provides users with the experience of virtual "flight" over maps comprised of aerial and satellite imagery overlaid with geo-referenced data. The data drilling and querying component of the system allows the users to easily explore geospatial data, to create geospatial queries, and get instant answers supported by high-performance multidimensional search mechanisms. TerraFly's server farm ingests, geo-locates, cleanses, mosaics, and cross-references 40 TB of basemap data and user-specific data streams. TerraFly's Application Programming Interface allows rapid deployment of interactive Web applications; it has been used to produce systems for disaster mitigation, ecology, real estate, tourism, and municipalities. TerraFly's Web-based client interface is accessible from anywhere via any standard Web browser, with no client software to install.

2 TerraFly's Advanced Geospatial Data Processing

2.1 Overview

The cornerstones of TerraFly is in its portability, fly-over technology, ability to integrate multiple sets of geospatial and related data into customizable, multi-layered products, and inclusion of powerful but easy to use visualization, querying and analysis tools. As can be seen on TerraFly's landing page (see Fig. 25.1) [13], users are able to easily select the geographic area they are interested in exploring. By streaming incremental imagery tiles, TerraFly enables users to engage in virtual flights (see Fig. 25.2) where they maintain full control over flight speed, direction and altitude (spatial resolutions) via an intuitive, flash-based navigation system.

Navigate the map, doubleclick it to get local point-data, easily geo-query vast databases, and fly with TerraFly® Web tools, from autopilots to pro viewers. View Imagery and Data anywhere in the United States and in much of the World.

A public service of Florida International University sponsored by NSF and by NASA, USGS, and IBM.

FIU, Miami, FL 33199

ECS: Engineering and Computer Science (ECS) houses the School of Computing and Information

Examples: USA: Albuquerque Anchorage Atlanta Augusta Austin Baltimore Baton Rouge Bismarck Charlotte Chicago Cleveland Colorado Springs Corpus Christi Dayton Denver Detroit El Paso Fort Wayne Fresno Ft. Lauderdale Greensboro Honolulu Houston Jackson Jefferson City Knoxville Lexington Lincoln Louisville Lubbock McAllen Miami Minneapolis Mobile Modesto Montgomery Nashville Oklahoma City Orlando Palm Bay Philadelphia Pierre Pittsburgh Portland Riverside Sacramento Saint Louis San Diego San Francisco San Juan Seattle Tampa Washington World: Venice Berlin Havana Varadero Caracas

Applications: ***Real Estate***: Miami ***AutoPilots***: Miami Port NSF Washington Rio Grande Detroit Seattle Star Island FIU Oleta Park Spiral Authoring ***Historical Imagery***: Miami through Time Washington Dubai

01/30/2009 - TerraFly Real Estate Engine receives EAMS award of Innovative Product of the Year.

08/01/2008 - National Science Foundation increases TerraFly reserach funding by $2.5M to $7.5M.

07/06/2007 - TerraFly covered by FOX News and NPR.

02/01/2007 - TerraFly powers searches for neighborhood demographics, hotels, real estate, hospitals, clinics, dentists.

Transferring data from 131.94.133.179...

Fig. 25.1 TerraFly landing page

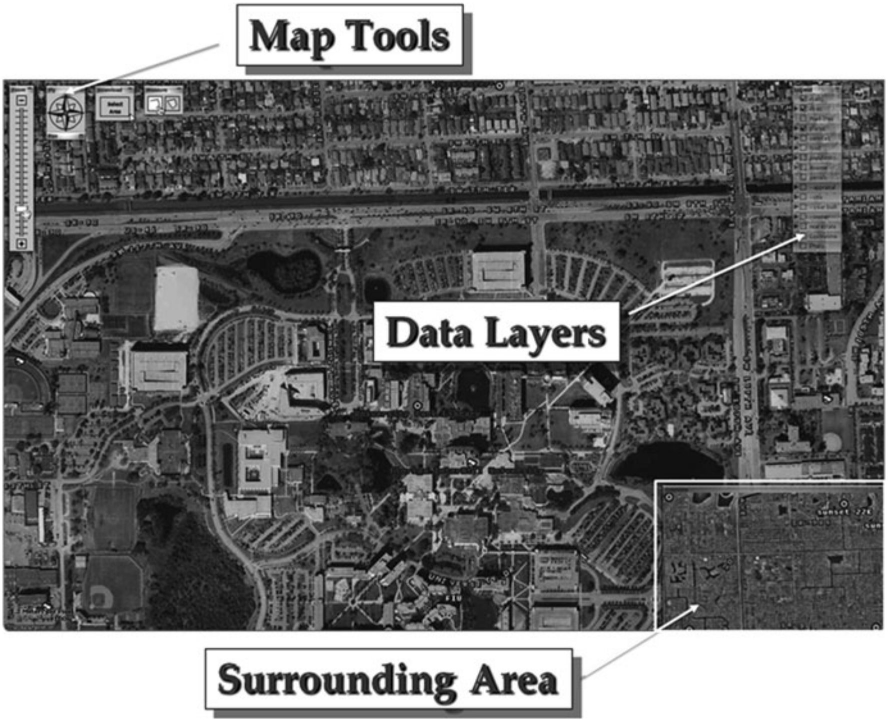


Fig. 25.2 TerraFly flight and data layers control layout

TerraFly's data-mining tools are capable of delivering an extensive amount of data related to user-specified geographical locations [14–17]. Unlike most GIS applications [18], TerraFly eliminates the need for the end-user to deal with any technical aspects of the system. Users are able to easily query for data of interest, and have that data automatically visualized in the form of non-obstructing geo-referenced overlays, or data layers, combined with spatial imagery [19–22]. The most popular types of overlaid data include NAVTEQ NAVSTREETS street vectors, World OpenStreetMaps, property parcels, Yellow pages, White pages, demographics, Geographic places (Worldwide from NGA and other sources, USGS Geographic Names Information System), services, hotels, and real estate listings. This is just a sampling of datasets.

In addition to data overlays, TerraFly provide users with a drill-down detailed information page on a point or area (see Fig. 25.3). For example, users can use TerraFly's address locator capability to "fly" to a specified address, and then request more specific information about that particular location. To do this, the user clicks on the particular point of interest on the spatial image. A preview page will pop up in the flight window that contains a summary of information about that particular location, along with links to more detailed location information. To view the more

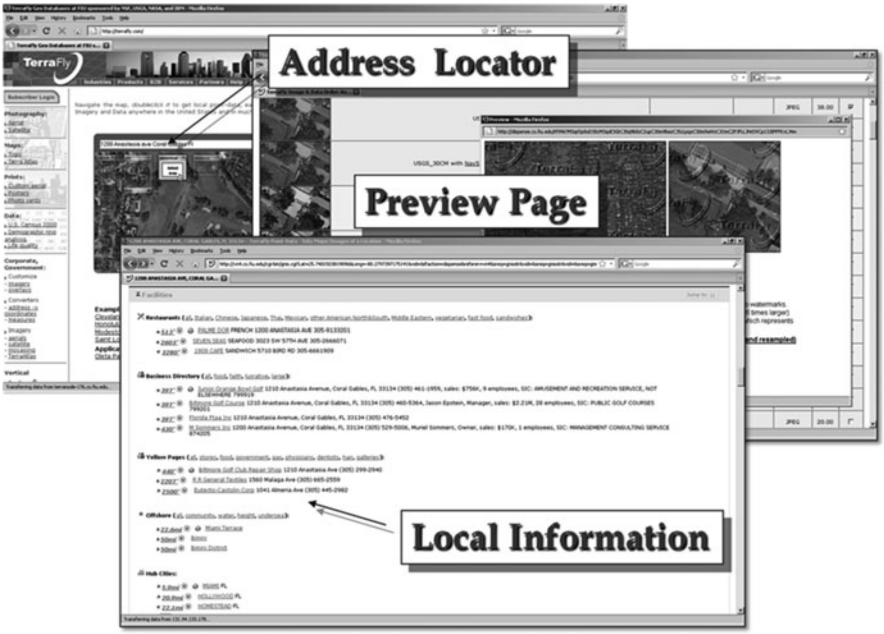
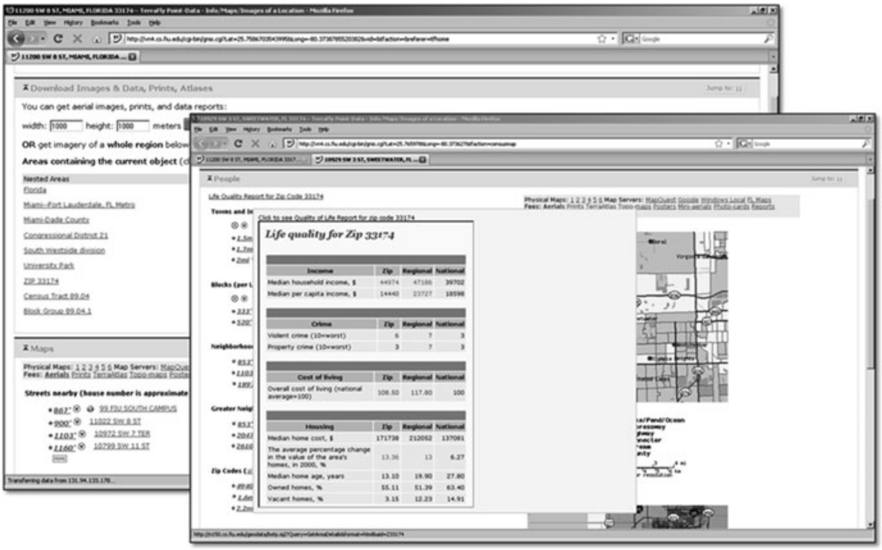


Fig. 25.3 Sample of TerraFly's data drill down

detailed information, the user clicks on the associated link in the preview page, and the user will be taken to a new page where more detailed local information, such as demographic data, local restaurants and businesses, etc. is displayed below the flight window (see Fig. 25.4).

The TerraFly system has querying and analysis capabilities that are the result of intensive, cutting-edge and innovative computing research. The tools available in TerraFly include user-friendly geospatial querying, data drill-down, interfaces with real-time data suppliers, demographic analysis, annotation, route dissemination via autopilots, customizable applications, production of aerial atlases, and an application programming interface (API) for web sites. Many of TerraFly's capabilities and the technology behind them, such as TerraFly's underlying data storage mechanism, client-server interaction, user interface, ability to overlay additional information layers, and ergonomics of use and maintenance, have been described in numerous professional publications [23–30]. More recent advancements in TerraFly's data handling capabilities, namely improved user-centric data integration and mapping tools, data repository, and advanced data indexing and preprocessing, are described in the remainder of this section. Advances in TerraFly's data visualization capabilities, such as times series visualization, the customizable autopilot and the data dispenser systems, are described in Sect. 3.



Life quality for Zip 33174

Income	Zip	Regional	National
Median household income, $	44974	47186	39702
Median per capita income, $	14440	23727	18598

Crime	Zip	Regional	National
Violent crime (10=worst)	6	7	3
Property crime (10=worst)	3	7	3

Cost of living	Zip	Regional	National
Overall cost of living (national average=100)	108.50	117.80	100

Housing	Zip	Regional	National
Median home cost, $	171738	212052	137081
The average percentage change in the value of the area's homes, in 2000, %	13.36	13	6.27
Median home age, years	13.10	19.90	27.80
Owned homes, %	55.11	51.39	63.40
Vacant homes, %	3.15	12.23	14.91

Fig. 25.4 Demographics and quality of life overview page

2.2 *Data Repository*

A critical component of TerraFly is its data repository, and a major strength lies in its integration of heterogeneous data sources including relational and semantic databases and web sources with spatial data. TerraFly's data repository was one of the first GIS databases that were able to store heterogeneous data in one database [23–26, 31, 32]. As with other GIS tools, there are two main types of geo-referenced spatial data that TerraFly must handle: raster (satellite and aerial photography) and vector data (points, lines and polygons) [6, 33]. TerraFly's data repository currently stores the following:

Raster Data. The entire collection of Digital Orthophoto Quarter Quadrangle (DOQQs) produced by the USGS (1 m-resolution orthorectified aerial photography of the entire USA, including once-over 13 TB coverage and 5 TB of multi-temporal updates); the entire collection of USGS Urban Area High Resolution Orthoimagery (15–30 cm imagery covering 135 metropolitan areas – 15 TB), and Landsat imagery [34] covering USA and parts of the world, imagery from local sources (7 cm/pixel and up), and a vast collection of satellite imagery, particularly from GeoEye and Ikonos satellites.

Vector Data. The TerraFly vector collection includes 1.5 billion geolocated objects, 50 billion data fields, 1 billion polylines, 120 million polygons, including: all World roads, Worldwide geographic places and features, 24 billion demographic data items (3,000 fields × 8 million objects), 1 billion economic data items including the US Census demographic and socioeconomic datasets [35], 110 million USA

cadastre polygons and detailed data on each parcel, DEM Elevation data, 15 million records of businesses (with company stats, management roles, contacts and radius demographics for each business), 2 million physicians with expertise detail, various public place databases (including the USGS GNIS [36] and NGA GNS), Wikipedia, extensive global environmental data (including daily feeds from NASA and NOAA satellites and the USGS water gauges), and hundreds of other datasets.
New Data is constantly being updated and added to TerraFly's data repository. For a current listing of available data, please see [37].

2.3 TerraFly's Advanced Data Geocoding Capabilities

A commonly used feature of GIS and mapping systems is the ability to geocode street addresses. As with most GIS and mapping applications, precise and accurate geocoding of available data in TerraFly is critical. If implemented with sufficient precision, this ability can satisfy the needs of many businesses' day-to-day functions (e.g., realtors, attorneys, engineers, etc.), as well as the more complex needs of government and research (e.g., information retrieval from spatial databases) [38]. For many years, most mapping systems have used standard interpolation geocoding that estimates where on a street a particular address is located. There are a number of assumptions associated with these standard methodologies, such as consistent standardization of street numbers in all areas and substantial accuracy of the underlying data. Because it is an estimate, and not a coordinate point associated with a specific building, this type of geocoding has a certain level of inaccuracy and often results in near misses. For example, when using one of these standard methods, the resulting point often lies outside of the parcel property lines of the building of interest and sometimes a few buildings away.

To address this issue and to provide the significantly higher level of precision required for many TerraFly applications, *roof-top geocoding* has been implemented in TerraFly via the inclusion of First American Parcel Point Nationwide Cadastre data [39] along with efficient data management algorithms. This data set contains attributes that are intended to support data integration related to land parcels across jurisdictional boundaries, and includes parcel boundaries, parcel centroid, addresses, Assessor's Parcel Numbers (APN) and ownership information [39, 40] (ownership information is not available in all counties).

After initial data cleanup to remove data with missing components and incorrectly formatted records, the data was cross-referenced with other datasets and a precise geocoding component was created for TerraFly. Specifically, TerraFly's rooftop geocoder was created using spatial indexes and data structures already used in TerraFly, with street address interpolation and string matching algorithms. In short, the process is as follows:

1. When a query for a particular address is made, TerraFly uses a standard interpolation methodology to generate approximate coordinates for the requested location.

2. These coordinates are then used to perform a nearest neighbor query to retrieve any nearby parcels from the data set.
3. A local search for the best matching parcel is performed on the results.
4. If a match is found for that parcel in the database, the coordinates for that record are returned to the user.

TerraFly also has *location-sensitive geocoding*. If the user is focused in a particular location and provides a partial address or a partial description of a geographic or social place, the system will look for the best match, weighing the factors of place importance and its proximity to the user's current location.

2.4 Image Mosaics, Raster Data Analysis and Amelioration

As is seen in Sect. 2.2, TerraFly has nearly 40 TB of aerial and satellite imagery in its data repository. As with vector data, TerraFly requires high quality raster data to perform many of its applications. However, not all of the imagery is of a high enough quality for appropriate visualization, and sometimes multiple images for the same locations, are shot at different times or by different instruments. Nevertheless, some users desire to see original unaltered imagery. Most users, however, desire to see a pixel-by-pixel mosaic of the best imagery available. "Best" is a user group specific criterion, and TerraFly accepts various definitions of what "best" means for a particular user group. The default "best" is the freshest, the sharpest, and the most natural-color imagery. Thus, in preparation for mosaicing, TerraFly performs image analysis on its raster data.

Two different types of imagery analysis are conducted to improve TerraFly's image quality: (1) detection of black regions in individual tiles and (2) histogram analysis on an entire data set. When imagery data is acquired or post-processed, individual data tiles may contain areas of black interspersed or surrounded by good quality imagery. As a result, detection of black regions is performed tile by tile, and results of the analysis are stored as meta-data inside each tile. When tiles are retrieved, this meta-data is accessed and algorithms are used to determine how to best mosaic that particular tile with better quality imagery for the best quality output to the user.

Entire data sets are sometimes affected by color distortion that is not easily detected when imagery is analyzed on a tile by tile basis. Therefore, histogram analysis is conducted on the entire data set. To provide end users with the most accurate and flexible data product, results of the histogram analysis are stored separately from the original data set. This provides the user with the option of seeing either the original or corrected imagery.

Both of these types of image analysis involve data-intensive computing, particularly on the large data sets inherent to spatial data. TerraFly's raster analysis applications have been ported to MapReduce [41, 42], a highly efficient framework that automates the use of parallel processing through the use of mappers and reducers (see [43] for more details on MapReduce). For detection of black regions, each mapper is assigned a certain number of tiles to analyze, depending on the size of the data set and the number of mappers. No reducer is needed for further processing. For histogram analysis, each mapper is responsible for analyzing a portion of the data set, and the partial result is sent to the reducer. The reducer then combines all of the results and computes the final output. TerraFly's use of MapReduce has resulted in a dramatic improvement in computing time, and has shown close to linear scalability [41].

2.5 Spatial Keyword Indexing (SKI)

Spatial data is inherently very large, complex and often heterogeneous in nature. This makes meticulous and efficient data management a major challenge, particularly when dealing with extremely large databases such as TerraFly's data repository. However, appropriate data indexing can be used to make querying of data more efficient. To address this and improve performance, TerraFly includes an innovative, hybrid method to efficiently process top-k spatial queries with conjunctive Boolean constraints on textual content [16].

Specifically, this method combines an R-tree structure and text indexing into a location-aware inverted index by using spatial references in posting lists. R-Trees are often used as an indexing mechanism for spatial search query processing [18, 44, 45]. In the TerraFly-SKI hybrid method, an R-tree index is modified in the upper level nodes with the addition of a list of leaf nodes that have the same parent. An inverted file is altered to contain a list of pointers to some of the R-tree's nodes, creating a spatial inverted file. To process a query, the R-tree is traversed in a top-down method using a best-first traversal algorithm. If at least one object exists that satisfies the Boolean condition in the subtree, then a node entry is made into the priority queue. Otherwise, the unnecessary subtree traversal is eliminated. The result is a disk-resident, dual-index data structure that is used to effectively and proactively prune the search space [16].

Although this method produces improved performance, the indexing process can take a substantial amount of time. How much time is needed depends on the size of the database, as well as the size of the lexicon found in the spatial inverted file. For a database that contains N objects, where those objects are used to construct the SKI's modified R-tree, the number of insert operations is $O(N)$. When constructing the spatial inverted file, the most expensive operation involves sorting the lexicon, with a construction time of $O(N + V\, log(V))$, where V is the size of the lexicon.

MapReduce has been used to improve image processing computing time for data in TerraFly's data repository [41]. The use of parallel computing to improve the efficient construction of inverted indices has been studied by other researchers [43]. TerraFly's work with MapReduce has been leveraged to improve processing time in SKI construction. In this process, the R-tree structure is built with two MapReduce pairs as in [2]. The output includes references to R-tree nodes as intermediate data for use in the following job. Considering each object as a document, a MapReduce job also builds the spatial inverted file on the database lexicon. The MapReduce compound uses the intermediate data generated in the first iteration.

The resulting SKI's data structures are stored remotely in the Cloud, and are downloaded to a local host to serve interactive queries. In this process, the SKI data structures are partitioned because the number of "smaller" SKI structures is equal to the number of Reducers used in the MapReduce job. As a result, queries are processed with a modified version of the search algorithm proposed in [16, 27].

3 Advanced Data Visualization Capabilities

New and innovative technologies and functionality are continually being developed and added to TerraFly's already extensive capabilities. Section 2 presented key advances in TerraFly's backend and data handling capabilities. Those technological advances are functions that end-users may not be entirely aware of, but that affect them and the quality of their work substantially. In this section, we discuss some of TerraFly's most recent advancements in data visualization and user-centric capabilities. Although equally as important as the advances made on the back end, end users ability to interact with the system is more directly affected by advances to TerraFly's visualization engine. Specifically, this section discusses TerraFly's time series visualization capabilities, auto pilot, and data dispenser.

3.1 Time Series Visualization

A powerful capability is the TerraFly TimeSeries application. This application has a unique ability to provide efficient and ergonomic dissemination of imagery with spatio-temporal data overlays. In other words, the TerraFly TimeSeries application can retrieve geospatial and remotely sensed imagery of the same geographic location that was acquired during different time periods. The system is then able to create an animated sequence over time to clearly show historical changes.

To accomplish this, the TimeSeries application uses the coordinates of the current map center that it has received from the TerraFly API to send a request for information on available images to the TerraFly SO service (imagery source service). The SO service provides the TimeSeries application with an XML formatted string that contains information about available imagery sources, resolutions and acquisition dates. Upon receipt of this information, the TimeSeries application parses the XML [46] and orders the sources by acquisition date, starting with earliest date to the latest date. It then creates a time-line panel that accurately reflects the proportion of the number of days between acquisition dates, and then searches for the closest resolution images of the requested location. The TimeSeries application then generates URLs needed to request the corresponding imagery. Once the imagery servers provide the requested images, the TerraFly API loads the imagery and creates an overlay in the viewing area. The TimeSeries application

Fig. 25.5 Ishinomaki, Japan, before the earthquakes and tsunami of March 11, 2011

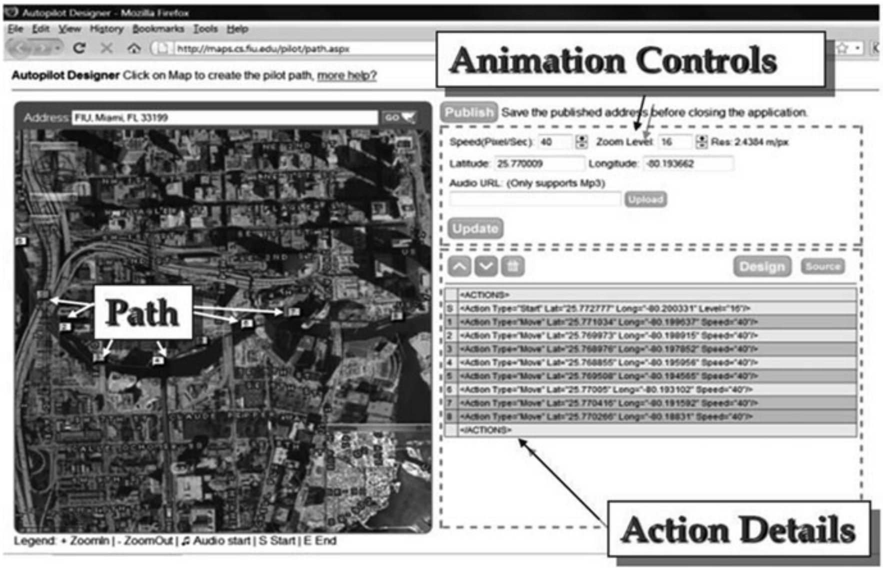
creates the animated time sequence by fading-in and -out the corresponding images in the timeline. Specifically, this fading-in and -out effect is achieved by changing the transparency parameter of these images from low to high (fade-in) or high to low (fade-out) every 40 ms at a rate of 25 frames/s.

The resulting time series can be quite dramatic and useful, particularly in disaster mitigation applications. When an extreme event occurs, certain features or characteristics of a particular area might be radically changed. With this unique use and visualization of historic data, the TerraFly TimeSeries application enables taking the important information and creating a new synthetic view of an emergent reality. This can not only aid in response to the disaster, but it also provides researchers with a rich source of data that can be studied to find better ways to plan and respond to similar types of disasters.

The recent earthquakes and tsunami that hit Japan in March of 2011 provides an example of the power of this technology. As can be seen in Figs. 25.5–25.7, the before, during and after geospatial images, respectively, of the tsunami are quite dramatic. In the image from April 4, 2010, entire neighborhoods can clearly been seen as having been established and intact. The image from March 12, 2011, the day after the tsunami hit, shows that water has rushed well inland, inundating those neighborhoods. The final image, from March 19, 2011, shows the aftermath once the waters have finally receded. Those neighborhoods were completely decimated. Although images from the ground would show the devastation, the remotely sensed images allow researchers to better study the overall impact and patterns of this catastrophic event.

Fig. 25.6 Ishinomaki, Japan, the day after the earthquakes and tsunami of March 11, 2011

Fig. 25.7 Ishinomaki, Japan, 8 days after the earthquakes and tsunami of March 11, 2011



Address: FIU, Miami, FL 33199

Save the published address before closing the application.

Speed(Pixel/Sec): 40 Zoom Level: 16 Res: 2.4384 m/px

Latitude: 25.770009 Longitude: -80.193662

Audio URL: (Only supports Mp3)

```
  <ACTIONS>
S <Action Type="Start" Lat="25.772777" Long="-80.200331" Level="16"/>
1 <Action Type="Move" Lat="25.771034" Long="-80.199637" Speed="40"/>
2 <Action Type="Move" Lat="25.769973" Long="-80.198915" Speed="40"/>
3 <Action Type="Move" Lat="25.768976" Long="-80.197852" Speed="40"/>
4 <Action Type="Move" Lat="25.768855" Long="-80.195956" Speed="40"/>
5 <Action Type="Move" Lat="25.769508" Long="-80.194565" Speed="40"/>
6 <Action Type="Move" Lat="25.77005" Long="-80.193102" Speed="40"/>
7 <Action Type="Move" Lat="25.770416" Long="-80.191592" Speed="40"/>
8 <Action Type="Move" Lat="25.770266" Long="-80.18831" Speed="40"/>
  </ACTIONS>
```

Fig. 25.8 TerraFly's autopilot flight sequence creation tool

3.2 *Auto Pilot*

TerraFly includes the autopilot technology that allows users to preplan and map out a customized flight path of interest. With this tool, end users can quickly and easily select specific destinations over spatial and remotely sensed images, and the system will automatically create that flight path as a series of point destinations at the speed and altitude (resolution) desired by the user. The speed and altitude need not remain static during the automated flight. At any point in the predefined flight sequence, users may include changes in speed, and zoom in (i.e., view higher resolution data) or zoom out (i.e., view lower resolution data) at will. The user is also able to determine which additional features will be displayed while traveling in the flight path. In essence, any feature or information that is possible to view when the user is in manual mode can be added as a component of the requested flight path. The Autopilot flight sequence creation tool can be seen in Fig. 25.8.

Once a flight sequence is defined, the flight path is overlaid the geospatial images in the flight path window. As can be seen in Fig. 25.9, users are able to fly along the requested flight path in the main TerraFly flight window without any intervention by the user. Users will also be able to see a zoomed out overview of the overall flight sequence in a smaller window found in the lower right-hand corner of the flight area.

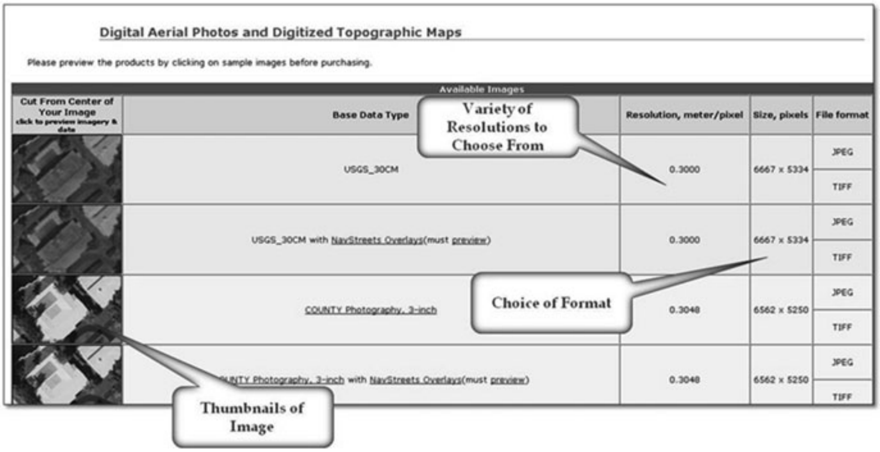
TerraFly's autopilot technology is not merely for entertainment value. There are numerous applications for this technology, including education, emergency preparedness in urban areas, and the study of crops in rural areas. To best illustrate this, imagine this scenario. A local office of emergency preparation and response

Fig. 25.9 TerraFly's autopilot flight path window

is notified of the imminent approach of a hurricane to their region. Officials and employees must make preparations to both mitigate effects of the hurricane, as well as be prepared to respond in the aftermath. With TerraFly's autopilot technology, coordinated autopilot flight paths could be created prior to the storm event and assigned to specific responders. Once the storm as passed and new data is available regarding storm impacts in the area, each responder could quickly and easily engage their assigned flight sequence to view and interpret impact results. The responders would not have to manually control movement over the area of interest and could, instead, just focus on gaining the information that they need. Further, this same flight path could easily be used over and over as updates and new information come in. The auto-piloted path can be annotated with voice clips, images, and data.

3.3 Data Dispenser

There are times when end users are in need of being able to work with spatial data outside of TerraFly. For example, an environmentalist may need to run complex, domain specific analyses that cannot be accomplished in TerraFly itself. Alternately,



Digital Aerial Photos and Digitized Topographic Maps

Please preview the products by clicking on sample images before purchasing.

Available Images

Cut From Center of Your Image click to preview imagery & data	Base Data Type	Resolution, meter/pixel	Size, pixels	File format
	USGS_30CM	0.3000	6667 x 5334	JPEG TIFF
	USGS_30CM with NavStreets Overlays(must preview)	0.3000	6667 x 5334	JPEG TIFF
	COUNTY Photography, 3-inch	0.3048	6562 x 5250	JPEG TIFF
	UNTY Photography, 3-inch with NavStreets Overlays(must preview)	0.3048	6562 x 5250	JPEG TIFF

Fig. 25.10 Customized web page created by TerraFly's data dispenser

a business may need a large, high quality poster print of a particular area but lack to equipment to produce the poster. To address these needs, TerraFly provides a flexible and user friendly data dispenser that end users can use to acquire any spatial and related data for a location of their choice [47].

TerraFly's data dispenser provides users with fast and convenient access to a map or a remotely sensed image. The data dispenser has been designed with an easy to use, intuitive interface that allows users to acquire data without any specialized training or tools. Users are able to easily choose, mark and dispense satellite images or aerial photos of any size, and in varied formats. TerraFly's dispenser can also provide the user with textual geo-referenced data associated with a dispensed image. When combined with the requested imagery, this data gives the user a unique information package associated with the geographical area of interest. Numerous products, both digital and printed, are available to the user. End users can order large poster prints, aerial photos and topomaps, atlases as auto-formatted PDF files, photo prints, GeoEye satellite products, reports, and more [48].

To the user, requesting data via TerraFly is very simple. The download button is clicked, the user selects the area of interest using the mouse to manipulate a bounding box, and then selects "download" from the menu. Once this request is sent to the system, TerraFly's data dispenser module searches TerraFly's imagery database for all image tiles that can be used to generate the imagery for the user-defined area. TerraFly's information databases are searched for all the possible data reports related to the selected area. TerraFly also has the capability to search data sources on other sites. Currently, TerraFly also searches the GeoEye Archives for the availability of GeoEye and Ikonos imagery. As can be seen in Fig. 25.10, once the searches are complete, the system generates a customized web page that presents users with all of the unique product options available for the area of interest.

4 Application Domains

TerraFly's use of innovative research and cutting-edge technology has created a flexible, robust and forward-thinking solution that has multiple domain applications. The wide range and types of data available in TerraFly makes the system useful to a much broader user base than conventional geographic information systems. In this section, an overview of the primary domains that currently use TerraFly are presented, namely, disaster mitigation and response, research and scientific inquiry, real estate, travel and tourism, and government operations and public interest.

4.1 Disaster Mitigation and Response

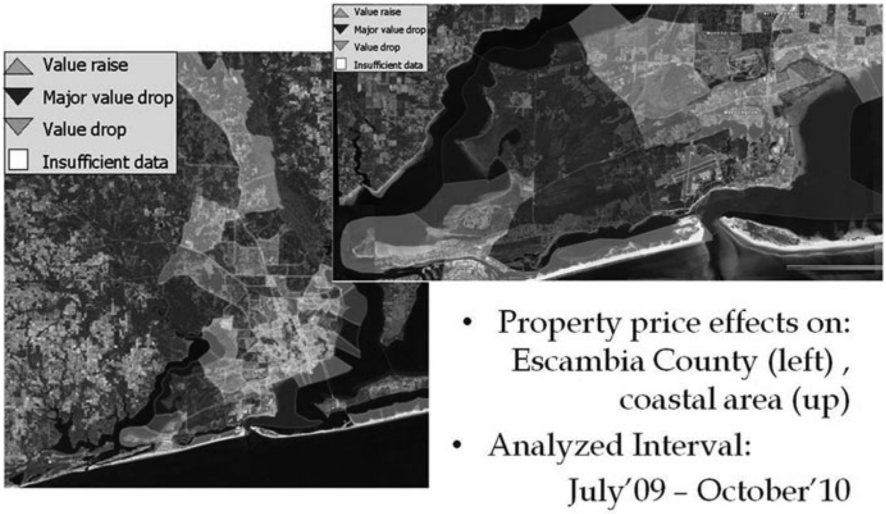
When a disaster occurs, fundamental aspects of life dramatically change. Major impacts often include changes to our physical environment that are far reaching and easy to discern. The importance of the positive impact that information technology can have on disaster mitigation and response has been discussed in several US government and international reports. They all agree that greater resources in information technology will improve our ability to plan for and respond to disasters, ultimately saving lives and property [4, 5].

Discussions throughout this paper have touched upon TerraFly's prominence as a tool for disaster mitigation and response. As has been noted, TerraFly has multiple capabilities that researchers, planners and responders can use to vastly improve disaster planning and response. For example, as was seen in Sect. 3.1 TerraFly's TimeSeries visualization capabilities can quickly and easily show users physical changes to our environment as a result of a disaster. However, other types of impacts, such as economic impacts, can also be visualized in TerraFly.

As can be seen in Fig. 25.11, TerraFly has the ability to provide visualization of economic changes, such as changes in property values, as the result of a disaster. The example presents patterns of changes in property values affected by the BP Deep Water Oil Spill. Unlike major impacts on the physical environment, this type of impact is one that is not clearly seen with the naked eye. Instead, combining, overlaying, and analyzing different types of pertinent data are required to gain this understanding.

4.2 Research and Scientific Inquiry

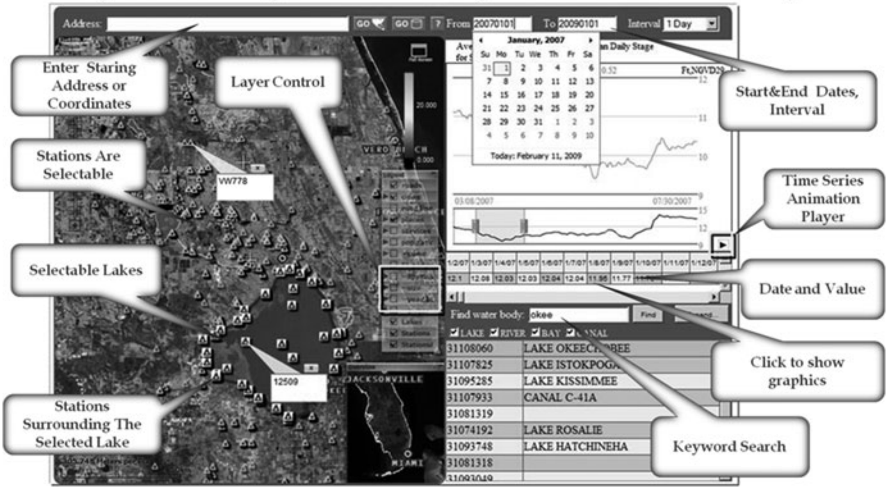
TerraFly is also extensively used as a tool for research and scientific inquiry. TerraFly is used to support and engage in many domains of computing research such as data processing, query optimization, parallel processing, storage of massive amounts of data, etc. TerraFly also provides numerous data analysis and



- Property price effects on: Escambia County (left) , coastal area (up)
- Analyzed Interval: July'09 – October'10

Fig. 25.11 The effect of the BP deepwater oil spill on property values

Average of Surrounding Stations' Mean Daily Level for Selected Water Body



Address:
From 20070101 To 20090101 Interval 1 Day

Find water body: okee

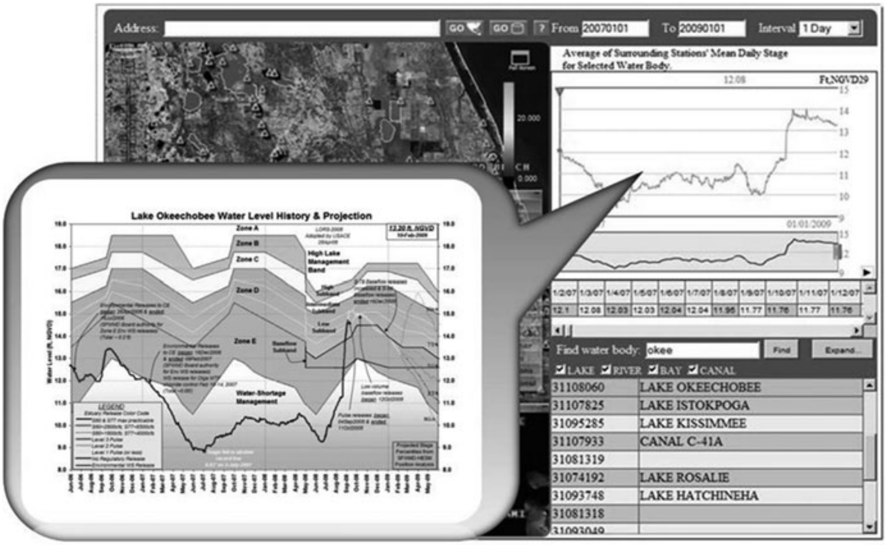
☑ LAKE ☑ RIVER ☑ BAY ☑ CANAL

31108060	LAKE OKEECH BEE
31107825	LAKE ISTOKPOG
31095285	LAKE KISSIMMEE
31107933	CANAL C-41A
31081319	
31074192	LAKE ROSALIE
31093748	LAKE HATCHINEHA
31081318	

Fig. 25.12 In situ data and graphs in TerraFly for use in hydrological analysis

visualization tools for scientists in various domains. Figure 25.12 shows Hydrology data analysis tools in TerraFly. With this tool's intuitive interface, users are able to view and analyze data, and have results of their analyses displayed as imagery, charts and tables.

Additional functionality found in TerraFly, such as time series data, key word searches, and layer control, is also available on this screen. For example, users



Address:
From 20070101
To 20090101
Interval 1 Day

1/2/07	1/3/07	1/4/07	1/5/07	1/6/07	1/7/07	1/8/07	1/9/07	1/10/07	1/11/07	1/12/07
12.1	12.08	12.03	12.03	12.04	12.04	11.96	11.77	11.76	11.77	11.76

Find water body: okee

☑ LAKE ☑ RIVER ☑ BAY ☑ CANAL

31108060	LAKE OKEECHOBEE
31107825	LAKE ISTOKPOGA
31095285	LAKE KISSIMMEE
31107933	CANAL C-41A
31081319	
31074192	LAKE ROSALIE
31093748	LAKE HATCHINEHA
31081318	
31093049	

Fig. 25.13 Geospatial-temporal plots

are able to select a date range and view time series animation of changes, as well as graphs that plot these changes over time. As can be seen in Fig. 25.13, very detailed geospatial temporal plots are available to users of this application. As with all TerraFly flight windows, locations are searchable by address, and more detailed data for specific points is available at the click of the mouse. For example, clicking on one of the stations provides information about that particular site.

4.3 Real Estate

TerraFly also provides tools for various business domains, such as Real Estate. Real estate related services are provided to real estate agents, buyers and sellers. TerraFly downloads real estate sale listings from the Multiple Listing Service (MLS) and overlays this data onto spatial imagery. As can be seen in Fig. 25.14, variables such as property types, prices and square footage are overlaid, with more detailed information and additional variables available at the click of the mouse. Users are able to quickly and easily see the number and location of homes for sale in a neighborhood, as well as nearby features that may affect the desirability of a property such as proximity to a park or school. Realtors can take potential buyers on virtual tours of a neighborhood, and map out driving and transportation routes of interest.

TerraFly's Real Estate component also provides powerful data mining capabilities. Again, as can be seen in Fig. 25.14, end users are able to search for properties

$4.8M 4 bedroom 4.5 bathroom 3017 sf CONDO at 50 S POINTE DR # 2601 built in 2008 in CONTINUUM ON SOUTH BEACH

$4.8M 4/4 PRO
3017 s/f
50 S POINTE DR # 2601

$4.88M 4/4 PRO
2901 s/f
50 S POINTE DR # TWN7

$4.9M 3/4 PRO
3760 s/f
125 OCEAN DRIVE # 503

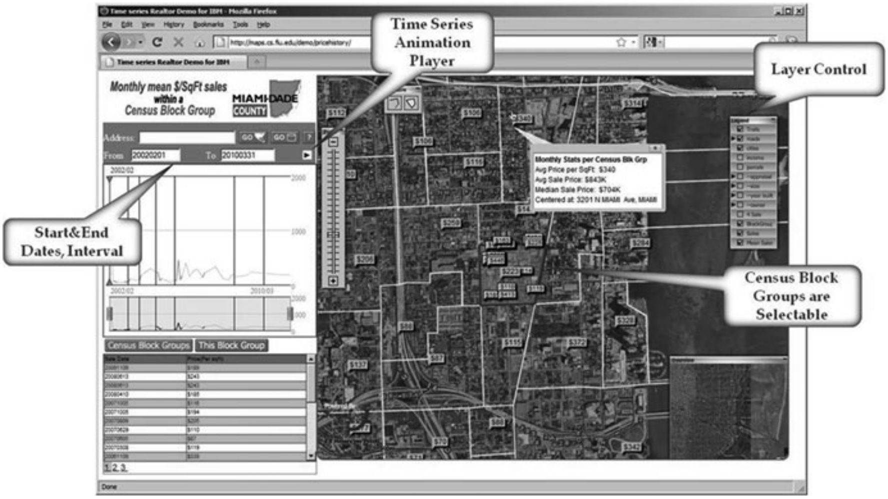
$5M 4/4 PRO
4100 s/f
300 S POINTE DR # 4006

Fig. 25.14 TerraFly's real estate consumer application interface

based on attributes such as asking price, number of bedrooms, square footage and various other keywords. A common search in South Florida, for example, is for ocean front condominiums or single family homes with a pool.

The power of this type of capability is not fully appreciated until one attempts to glean the same information without the ability to visualize it in a tool such as TerraFly. For example, imagine you are a buyer looking to purchase a home in a specific neighborhood. What you would typically be given is a list of properties for sale with their addresses and other relevant data such as square footage. Trying to visualize the locations of these properties in your mind is rather difficult, particularly in relation to desirable and undesirable features in the neighborhood. Further, visualization in TerraFly provides significantly more information about a particular property than a standard MLS listing. For instance, a home's address may indicate that it is on a small quiet street, but there is no information on what surrounds that property. It may be that the property backs up to a noisy highway or busy street. This would easily be seen in TerraFly, but would not be apparent to anyone using the typical MLS listing until the home is visited in person.

Another real estate application available in TerraFly can used to study real estate value trends over time. As can be seen in Fig. 25.15, users can view and analyze valuation trends of specific census block groups by selecting the census block of interest and entering start and end dates. Data such as average price per square foot is overlaid onto aerial imagery, and can be animated over time. The application also includes a trending graph so that users can see average changes in prices over their specified dates, as well as a list of sales prices by date.


Time series Realtor Demo for IBM - Mozilla Firefox

http://maps.cs.fiu.edu/demo/pricehistory/

Monthly Stats per Census Blk Grp
Avg Price per SqFt: $340
Avg Sale Price: $843K
Median Sale Price: $704K
Centered at: 3201 N MIAMI Ave, MIAMI

Fig. 25.15 Historical property value trends

4.4 Travel and Tourism

TerraFly's capabilities are also useful to the travel and tourism industry. As can be seen in Fig. 25.16, TerraFly's auto pilot mapping tool can be used to create virtual tours of vacation destinations of interest. Tour operators can provide potential clients with a bird's eye view of tour packages that highlight points of interest with pertinent information and high resolution geo-referenced images of particular locations. Tour operators could even use TerraFly to customize tour packages for individuals, creating the tour route with the direct input of their clients.

For example, tour operators can create a tour through Washington, DC, to show to potential clients. As they take their clients on a virtual tour, the clients can see various points of interest and let the tour operator know which destinations and points of interest they are interested in including in their tour package and which are of not interest to them. The tour operator could immediately make changes in the auto pilot to reflect the client's wishes and immediately show them the resulting new planned tour.

4.5 Government and Public Interest

The wide range and types of data available in TerraFly, along with TerraFly's robust, flexible and easy to use functionality is especially useful for government agencies and public interest. TerraFly's value as a tool for disaster planning and recovery has been illustrated throughout this paper. There are, however, numerous day-to-day functions that government agencies are responsible for that TerraFly is well

Fig. 25.16 Sample tour using TerraFly's auto plot tool

suited to support. In fact, as can be seen in Fig. 25.17, several local municipalities in Florida (The City of Coral Gables, The City of North Miami Beach, and The City of Miami Gardens, among others) have adopted TerraFly on their web sites to provide citizens with up to date information in their area, as well as tools to help make it easier for residents to work with their local government agencies and offices.

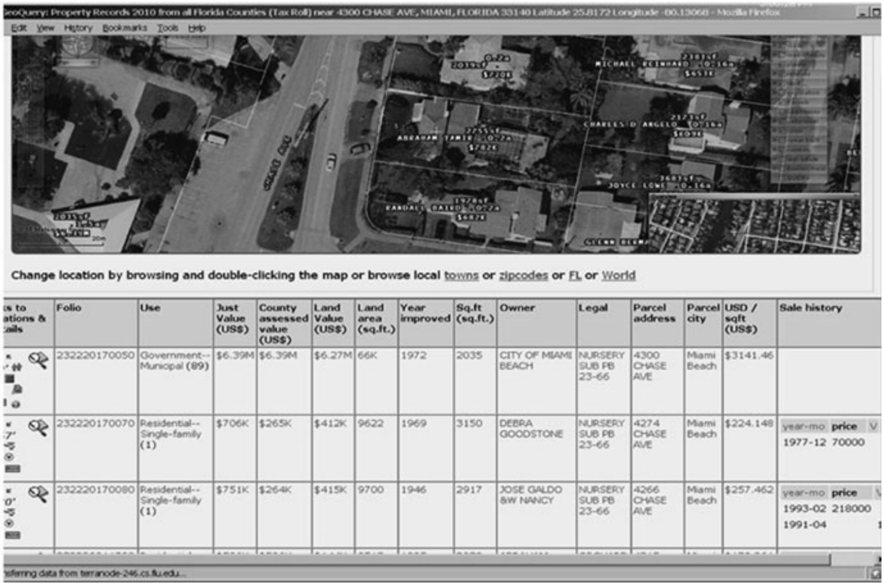
In addition, much of the data collected and used by government is well suited for integration into TerraFly. For example, as can be seen in Fig. 25.18, crime incident report data is associated with specific locations. TerraFly can easily process and overlay that data on geospatial imagery, providing both professional users and lay people with data visualization that is much more intuitive and easy to understand. Further, this data can be combined with other vector data available for that same area. The data could then be analyzed to determine if there are any trends or associations affecting the occurrence of particular types of incidences.

Importing property tax assessment data into TerraFly also aids in government operations. As can be seen in Fig. 25.19, by being able to view text records alongside with visual property imagery, tax assessor employees can more easily compare and analyze assessments between similar and dissimilar properties. This provides more consistency and accuracy in assessment procedures and valuation. Further, with the use of TerraFly's time series application, tax assessor employees can compare recent imagery with historical imagery to determine whether any changes have been make to a property that would necessitate a change in assessment value.

Fig. 25.17 Local municipalities that use TerraFly on their web sites

1515 NW 167 ST, MIAMI GARDENS, FL 33169

Fig. 25.18 Crime incidence report data overlaid on geospatial imagery



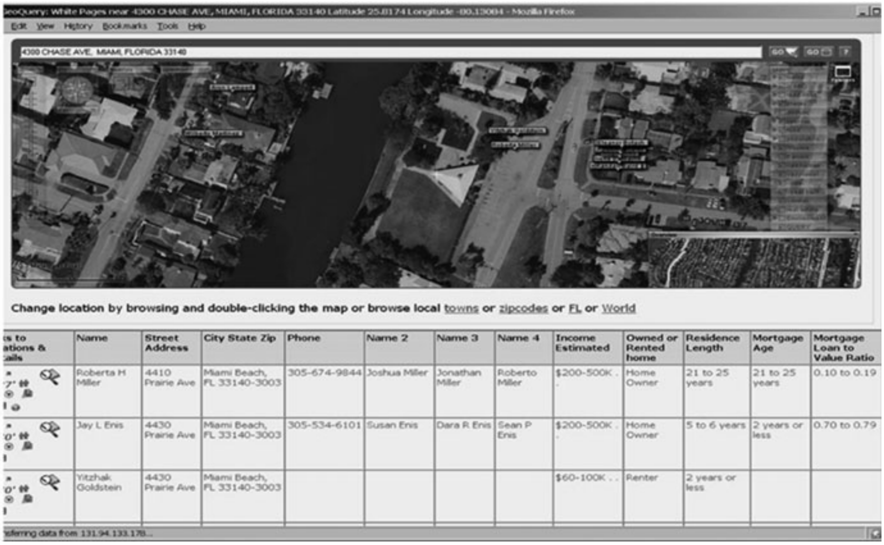
eoQuery: Property Records 2010 from all Florida Counties (Tax Roll) near 4300 CHASE AVE, MIAMI, FLORIDA 33140 Latitude 25.8172 Longitude -80.13068 - Mozilla Firefox

Change location by browsing and double-clicking the map or browse local towns or zipcodes or FL or World

s to ations & ails	Folio	Use	Just Value (US$)	County assessed value (US$)	Land Value (US$)	Land area (sq.ft.)	Year improved	Sq.ft (sq.ft.)	Owner	Legal	Parcel address	Parcel city	USD / sqft (US$)	Sale history
	232220170050	Government--Municipal (89)	$6.39M	$6.39M	$6.27M	66K	1972	2035	CITY OF MIAMI BEACH	NURSERY SUB PB 23-66	4300 CHASE AVE	Miami Beach	$3141.46	
	232220170070	Residential--Single-family (1)	$706K	$265K	$412K	9622	1969	3150	DEBRA GOODSTONE	NURSERY SUB PB 23-66	4274 CHASE AVE	Miami Beach	$224.148	year-mo price V 1977-12 70000
	232220170080	Residential--Single-family (1)	$751K	$264K	$415K	9700	1946	2917	JOSE GALDO &W NANCY	NURSERY SUB PB 23-66	4266 CHASE AVE	Miami Beach	$257.462	year-mo price V 1993-02 218000 1991-04

sferring data from terranode-246.cs.fiu.edu...

Fig. 25.19 Property tax assessment for individual properties



eoQuery: White Pages near 4300 CHASE AVE, MIAMI, FLORIDA 33140 Latitude 25.8174 Longitude -80.13084 - Mozilla Firefox

4300 CHASE AVE, MIAMI, FLORIDA 33140

Change location by browsing and double-clicking the map or browse local towns or zipcodes or FL or World

s to ations & ails	Name	Street Address	City State Zip	Phone	Name 2	Name 3	Name 4	Income Estimated	Owned or Rented home	Residence Length	Mortgage Age	Mortgage Loan to Value Ratio
	Roberta H Miller	4410 Prairie Ave	Miami Beach, FL 33140-3003	305-674-9844	Joshua Miller	Jonathan Miller	Roberto Miller	$200-500K . .	Home Owner	21 to 25 years	21 to 25 years	0.10 to 0.19
	Jay L Enis	4430 Prairie Ave	Miami Beach, FL 33140-3003	305-534-6101	Susan Enis	Dara R Enis	Sean P Enis	$200-500K . .	Home Owner	5 to 6 years	2 years or less	0.70 to 0.79
	Yitzhak Goldstein	4430 Prairie Ave	Miami Beach, FL 33140-3003					$60-100K . .	Renter	2 years or less		

sferring data from 131.94.133.178...

Fig. 25.20 White pages

Finally, there are also numerous capabilities in TerraFly that individuals find useful in their day-to-day lives. Overlaying data from the White Pages, for example, improves its usability for many individuals (see Fig. 25.20). Users are able to search for needed information in the White Pages not just by a person's full name, but also by attributes such as address, phone number, or the names of other individuals in the home. Alternately, users can search the geospatial imagery visually to find a particular location, and related information.

5 TerraFly with Other Systems: GIS-INTEGRO

TerraFly interacts and cross-pollinates with other Web GIS systems, including NASA WorldWind and the GIS-INTEGRO system developed by the Russian Academy of Science and Dubna International University. The GIS-INTEGRO system provides analytics data for disaster mitigation. The TerraFly and GIS-INTEGRO teams are performing joint research that focuses on expanding and improving algorithms and methodologies used for integrated object analysis and related processes. The teams are also furthering the development of intelligent user interfaces for each stage of the research and analysis process, from data acquisition, georeferencing, data integrity and quality assurance, and multi-level analysis to pre-print requirements for hard-copy published maps. There are numerous areas of application for these technologies, including decision-making support systems for mineral exploitation and environmental protection management. Components of this work include:

Pattern recognition algorithms (Holotype): Algorithms designed to compute similarity measures and matrices for heterogeneous objects, including resolving issues associated with recognition of objects in situations where only incomplete information is available.

Multi-functional geo-information server: Algorithms designed to integrate remote geo-informational resources with spatial modeling during disaster situations. The use of spatial modeling during a disaster (or very shortly thereafter) helps provide critical information on the current state of an affected area. This is accomplished by integrating up-to-the-minutes information with historical data, and providing a holistic evaluation of current environmental properties and impacts.

Ecological modeling: Methodologies designed for modeling ecological data and the structure of ecological informational space, as well as determining the natural and anthropogenic factors that affect the ecological state in regions of interest.

Geophysics: Consolidating vast amounts of geophysical algorithms and evaluations of mineral reserves.

6 Conclusion and Future Directions

Through the implementation of innovative techniques and technologies, TerraFly provides users with GIS capabilities without the need to learn complex interfaces or deal with the technology behind the system. It is a robust, user-friendly system that has wide appeal to many different types of users, and application to many different domains. TerraFly has been covered by both popular and specialized media, including TV (e.g. *Fox* and *Discovery*), radio (*NPR*), newspapers (e.g. *New York Times*, *USA Today*), magazines (e.g. *Science*) and journals (e.g. *Nature*). The project's primary sponsor is the National Science Foundation (NSF). Of the 53,000 NSF-funded projects in 2009, it chose 120, including TerraFly, for the NSF annual report to congress [49].

Several new projects and new directions are currently being pursued that will improve the technical capabilities and expand TerraFly's functions. One of them is the incorporation of social media, including using information gathered from social media as a viable data source. As the popularity of social media has increased, so has its potential for providing up-to-the-minute information on pertinent happenings in the world. If implemented appropriately, this could have a substantial impact on disaster response. It is no longer unusual for news of a particular event to be posted on a social networking site before any other media or type of communication. The challenge, however, is trying to determine how to most efficiently and effectively filter pertinent and accurate data from massive streams of messy, irrelevant, and inaccurate data.

Efforts are being made to expand data sharing capabilities and communication, particularly for disaster planners and responders. Areas being examined to aid in this include: (1) creation of a "citizens sensor network" that can potentially provide near instantaneous geolocation and visual coverage of extreme events; (2) development of techniques to better manage a solid communication infrastructure for disaster responders by providing services via the cloud and (3) generate models and simulations to better predict outcomes, that can subsequently be used to disseminate mission critical information and provide guidance for emergency response.

As the usefulness of spatial and related data increases and expands to new domains, efforts will also increase to determine how the needs of these new domains can be met while still providing a robust, innovative and intuitive system.

Acknowledgments This research was supported in part by NSF grants CNS-0821345, CNS-1126619, HRD-0833093, IIP-0829576, CNS-1057661, IIS-1052625, CNS-0959985, OISE-1157372, CNS-1042341, CNS-0837556, DGE-0549489, IIS-0957394, IIS-0847680, IIP-0934339, and OISE-0730065, DHS contract HSHQDC_11_C_00019, and support of USGS, IBM, CoreLogic, and ADCi. We greatly appreciate the support and guidance of our NSF Program Directors Rita Virginia Rodriguez, Richard Smith, Demetrios Kazakos, Victor Santiago, Maria Zemankova, and Juan Figueroa. We would like to thank Patrick Hogan of NASA World Wind [50] for his continued guidance and support in research and development of TerraFly's open-source APIs, 3D visual environment for visualizing spatio-temporal data with moving objects, and keyword-spatio-temporal queries on moving objects. We have greatly benefited from collaborating with current and former researchers at NASA Goddard Space Flight Center, particularly Milton Halem, Bill Teng, Patrick Coronado, and Bill Campbell. IBM has provided invaluable support and guidance, particularly Oded Cohn, Juan Caraballo, Adalio Sanchez, and Howard Ho.

References

1. International Tsunami Information Center. “11 March 2011, MW 9.0, Near the East Coast of Honshu Japan Tsunami”. [Online] 2011. http://itic.ioc-unes-co.org/index.php?option=com_content&view=article&id=1713&Itemid=2365&lang=en.
2. Gulf Oil Spill. NOAA.gov. [Online] http://www.education.noaa.gov/Ocean_and_Coasts/Oil_Spill.html.
3. D. Guha-Sapir. “Disasters in Numbers 2010”. Center for the Epidemiology of Disasters. [Online] 2011. http://www.cred.be/publications?order=field_year_value&sort=desc.
4. Select Bipartisan Committee to Investigate the Preparation for and Response to Hurricane Katrina. “A Failure of Initiative: Final Report of the Select Bipartisan Committee to Investigate the Preparation for and Response to Hurricane Katrina”. U.S. House of Representatives, Feb. 15th, 2006.
5. National Commission of the BP Deepwater Horizon Oil Spill and Offshore Drilling. “Deep Water: The gulf oil disaster and the future of offshore drilling”. Report to the President. 2011.
6. H. Samet. “The design and analysis of spatial data structures”. Addison-Wesley, Reading, MA, 1990.
7. T. Keating, W. Phillips, and K. Ingram. “An Integrated Topologic Database Design for Geographic Information Systems”. Photogrammetric Engineering and Remote Sensing, vol. 53, no. 10, 1987, pp. 1399–1402.
8. R.A. Lorie and A. Meier. “Using a Relational DBMS for Geographical Databases”. Geo-Processing, vol. 2, 1984, pp. 243–257.
9. M. Egenhofer. “Why not SQL!”. International Journal on Geographical Information Systems, vol. 6, no. 2, p. 71–85, 1992.
10. H. Samet. “Applications of spatial data structures”. Addison-Wesley, Reading, MA, 1990.
11. H.V. Jagadish, A. Chapman, A. Elkiss, M. Jayapandian, Y. Li, A. Nandi, and C. Yu. “Making Database Systems Usable,”. ACM’s Special Interest Group on Management of Data (SIGMOD), June 11–14, 2007, Beijing, China.
12. A.E. Wade. “Hitting the Relational Wall”. Objectivity Inc. White Paper. [Online] 2005. http://www.objectivity.com/pages/object-oriented-database-vs-relational-database/default.html.
13. TerraFly Landing Page. TerraFly. [Online] http://terrafly.fiu.edu.
14. N. Rishe, W. Teng, H. Rui, S. Graham, M. Gutierrez. “Web-based Dissemination of TRMM Data via TerraFly.”. EOS Transactions, American Geophysical Union, vol. 85, no. 47, Fall Meeting Supplement, December 2004.
15. N. Rishe, J. Yuan, R. Athauda, S.C. Chen, X. Lu, X. Ma, A. Vaschillo, A. Shaposhnikov, D. Vasilevsky. “Semantic Access: Semantic Interface for Querying Databases”. ACM SIGMOD Digital Symposium Collection DiSC01. June 2003. pp. 591–594.
16. A. Cary, O. Wolfson, and N. Rishe. “Efficient and Scalable Method for Processing Top-k Spatial Boolean Queries”. Proceedings of the 22nd International Conference on Scientific and Statistical Database Management. Published as Lecture Notes in Computer Science. Vols. 6187/2010: Scientific and Statistical Database Management, Springer Berlin/Heidelberg, 2010, pp. 87–95.
17. A. Prasad Sistla, Ouri Wolfson, Bo Xu, Naphtali Rishe. “Answer-Pairs and Processing of Continuous Nearest-Neighbor Queries”. Proceedings of the 2011 The Seventh ACM SIGACT/SIGMOBILE International Workshop on Foundations of Mobile Computing (FOMC 2011). San Jose, California, June 9th, 2011.
18. N. Roussopoulos, C. Faloutsos, and T. Sellis. “Nearest Neighbor Queries”. Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 71–79, 1995.
19. N. Rishe and O. Wolfson. “Thin Client Technologies for Spatial Data Visualization.”. Proceedings of the National Science Foundation Computing Research Infrastructure 2007 PI Meeting: Computer Science Department Boston University (NSF CRI 2007 PI Meeting). June 3–5, 2007, Boston, Massachusetts. pp. 84–88.

20. P. Szczurek, B. Xu, O. Wolfson, J. Lin, N. Rishe. "Prioritizing Travel Time Reports in Peer-to-Peer Traffic Dissemination". Proceedings of the IEEE International Symposium on Communication Systems, Networks and Digital Signal Processing (7th CSNDSP). Newcastle, U.K. July 21–23, 2010. pp. 454–458.
21. Bo Xu, O. Wolfson, C. Naiman, N. Rishe, R. M. Tanner. "A Feasibility Study on Disseminating Spatio-temporal Information via Vehicular Ad-hoc Networks.". Proceedings of the Third International Workshop on Vehicle-to-Vehicle Communications 2007 (V2VCOM 2007). Istanbul, Turkey. pp. 146–151.
22. O. Wolfson, B. Xu, H. Yin, N. Rishe. "Discovery Using Spatio-temporal Information in Mobile Ad-Hoc Networks. Web and Wireless Geographical Information Systems, 5th International Workshop, W2GIS 2005, Lausanne, Switzerland, December 15–16, 2005. SpringerVerlag Lecture Notes in Computer Science 3833. pp. 129–142.
23. N. Rishe. "TerraFly: NASA Regional Applications Center". The AMPATH Workshop: Identifying Areas of Scientific Collaboration Between the US and the AMPATH Service Area, Florida International University, Miami. Conference Report. August 15–17, 2001 p. 7–8.
24. D. Davis-Chu, N. Prabakar, N. Rishe, A. Selivonenko. "A System for Continuous, Real-Time Search and Retrieval of Georeferenced Objects". Proceedings of the ISCA 2nd International Conference on Information Reuse and Integration (IRI-2000). Nov. 1–3, 2000. pp. 82–85.
25. N. Prabhakaran, V. Maddineni, and N. Rishe. "Spatial Overlay of Vector Data on Raster Data in a Semantic Object-Oriented Database Environment.". International Conference on Imaging Science, Systems, and Technology (CISST '99), June 28 - July 1, 1999, Las Vegas, Nevada, pp. 100–104.
26. N. Rishe, S. Chen, N. Prabakar, M.A. Weiss, W. Sun, A. Selivonenko, D. Davis-Chu. "Terrafly: A High-Performance Web-Based Digital Library System for Spatial Data Access". ICDE 2001: International Conference on Data Engineering, April 2–6, 2001, Heidelberg, Germany. pp. 17–19.
27. A. Cary, Y. Yesha, M. Adjouadi, N. Rishe. "Leveraging Cloud Computing in Geodatabase Management". Proceedings of the 2010 IEEE Conference on Granular Computing GrC-2010. Silicon Valley, August 14–16, 2010. pp. 73–78.
28. P. Szczurek, B. Xu, O.i Wolfson, J. Lin, N. Rishe. "Learning the relevance of parking information in VANETs". Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking. Chicago, Illinois. September 24, 2010. ISBN:978-1-4503-0145-9. pp. 81–82, September 2010.
29. D. Ayala, J. Lin, O. Wolfson, N. Rishe, M. Tanizaki. "Communication Reduction for Floating Car Data-based Traffic Information Systems". Second International Conference on Advanced Geographic Information Systems, Applications, and Services, pp. 44–51, February 10–16, 2010.
30. W. Teng, N. Rishe, H. Rui. "Enhancing access and use of NASA satellite data via TerraFly.". Proceedings of the ASPRS 2006 Annual Conference, May 1–5, 2006, Reno, NV.
31. N. Rishe. "Database Design: The Semantic Modeling Approach". McGraw-Hill, 1992.
32. N. Rishe, and Q. Li. "Storage of Spatial Data in Semantic Databases.". Proceedings of the 1994 ASME International Computer in Engineering Conference, Minneapolis, MN, pp. 793–800, Sept 11–14, 1994.
33. G. Muffin. "Raster versus Vector Data Encoding and Handling: A Commentary". Photogrammetric Engineering and Remote Sensing, Vol. 53, No. 10, pp.1397–1398, 1987.
34. Landsat Project Policy and History: Landsat 7 Mission Specifications, NASA Goddard Space Flight Center,. [Online] http://ltpwww.gsfc.nasa.gov/LANDSAT/CAMPAIGN_DOCS/PROJECT/L7_Specifications.html.
35. Tiger Overview, United States Census Bureau, Tiger/Line data. [Online] http://www.census.gov/geo/www/tiger/overview.html.
36. USGS mapping Information: Geographic Names Information System (GNIS). [Online] http://mapping.usgs.gov/www/gnis/.
37. TerraFly data coverage. TerraFly. [Online] http://n0.cs.fiu.edu/terrafly.coverage.htm.

38. Federal Geographic Data Committee (FGDC) Subcommittee for Cadastral Data. “Today’s Cadastral Information Customers and Requirements”. FGDC Cadastral Subcommittee. [Online] 2008. http://nationalcad.org/data/documents/cadastral%20data%20customers.pdf.
39. CoreLogic. “CoreLogic ParcelPoint”. [Online] 2011. http://www.faspatial.com/databases/parcelpoint?format=pdf.
40. N. von Meyer, B. Ader, Z. Nagy, D. Stage, B. Ferguson, K. Benson, B. Johnson, S. Kirkpatrick, R. Stevens, and D. Mates. “Parcel Identifiers for Cadastral Core Data: Concepts and Issues”. FGDC Cadastral Subcommittee. [Online] July 2002. http://www.nationalcad.org/data/documents/parcelID.pdf.
41. A. Cary, Z. Sun, V. Hristidis, and N. Rishe. “Experiences on Processing Spatial Data with MapReduce” in Springer Lecture Notes in Computer Science. Vols. 5566/2009: Scientific and Statistical Database Management. (Proceedings of the 21st International Conference on Scientific and Statistical Database Management). New Orleans, Louisiana, USA. June 1–5, 2009.), pp. 302–319.
42. Z. Sun, T. Li, N. Rishe. “Large-Scale Matrix Factorization using MapReduce.” Proceedings of the 2010 IEEE International Conference on Data Mining. Sydney, Australia. December 13, 2010. ISBN: 978-0-7695-4257-7. pp. 1242–1248.
43. J. Dean and S. Ghemawat. “MapReduce: Simplified Data Processing on Large Clusters”. Communications of the ACM vol. 51, no. 1 (January 2008), pp. 107–113.
44. A. Guttman. “R-trees: A dynamic index structure for spatial searching”. Proceedings of the ACM SIGACT-SIGMOD Conference on the Principles of Database Systems, pp. 569–592, 1984.
45. R. Finkey, J. Bentley. “Quadtrees: A data structure for retrieval on Composite Keys”. Acta Informatica, vol. 4, no. 1, pp. 1–9, 1974.
46. Extensible Markup Language (SML) 1.0 (Fifth Edition). W3C Recommendation. [Online] November 26, 2008. http://www.w3.org/TR/REC-xml/.
47. N. Rishe, M. Gutierrez, A. Selivonenko, S. Graham. “TerraFly: A Tool for Visualizing and Dispensing Geospatial Data.” Imaging Notes, Summer 2005. Vol. 20, 2, pp. 22–23.
48. GeoEye Imagery Collection. GeoEye. [Online] 2011. http://www.geoeye.com/CorpSite/products-and-services/imagery-collection/Default.aspx.
49. National Science Foundation. FY 2010 Budget Request to Congress. National Science Foundation. [Online] May 7th, 2009. http://www.nsf.gov/about/budget/fy2010/pdf/entire_fy2010.pdf.
50. NASA World Wind. [Online] http://worldwind.arc.nasa.gov/.