

Mixing GIS and Text Analytics for Better Insight

Geographic information systems (GIS) are digital technologies for storing, analyzing, and displaying, geographic information. They allow one to reveal the spatial and geographical nature of complex social phenomena through the production of static and dynamic maps. These tools often use data that is quantitative or categorical in nature. They are less well suited for leveraging the richness of information stored in unstructured text data such as interview transcripts, tweets, crime narratives, customer feedback, field notes, or in photographic images.

Computer assisted qualitative data analysis software (CAQDAS) and text analytics techniques have proven more valuable for extracting insightful information from such unstructured sources. They can provide rich details for better understanding complex social phenomena, such as people's perceptions and interpretations, or reveal hidden factors or unexpected connections underlying those phenomena.

Provalis Research's QDA Miner and WordStat provide a unique integration of GIS mapping, qualitative analysis and text analytics features, allowing one to relate unstructured text data with geographic information to create maps, timelines, and other graphic displays to enhance your analysis and presentations. The software is equally adept at utilizing location-based data embedded in text, linked with another source or stored separately from the unstructured text.

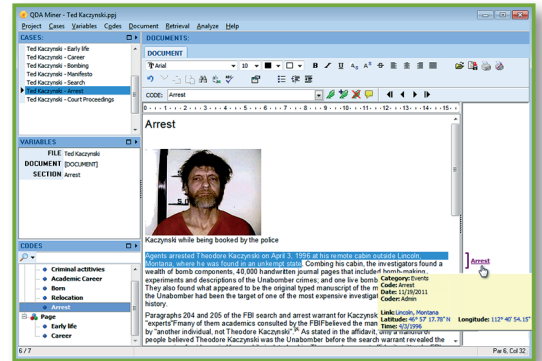
Make full use of geographic information embedded in text and images

It can be very challenging and time-consuming to extract geographic information embedded in text and images to produce comprehensive geographic representations. However, by using QDA Miner's computer assisted features, you can extract references to geographic locations in unstructured text such as names of cities, towns, regions, states, provinces, countries and other geographic locations to create maps. By attaching time tags to text segments or image areas, you can also create timelines to visually present themes or events in a chronological sequence. These features are being used by crime analysts, legal experts, historians, literary scholars, geographers and many others. Here is how it works.

Step 1 - GEOTAGGING

GEOTAG DOCUMENTS

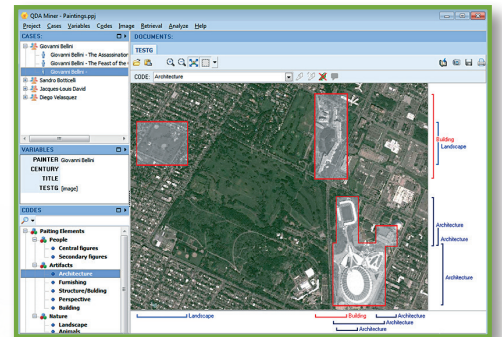
- Manually attach geographic coordinates as well as date and time to text segments.
- Search location from gazetteers or insert coordinates by pasting Google Earth placemarks.



Case #	Case	Variable	Type	Date	Location	Text
1	Ted Kaczynski - Early Life	DOCUMENT	Full date	5/22/1942	Florence, AL Florence, KY Florence, SC	Incorporated as KCM FLORENCE, 40479396.
1	Ted Kaczynski - Early Life	DOCUMENT	Day & Month	7/4/1942		Former assistant professor of mathematics. He was born the 4TH OF JULY 1914.
1	Ted Kaczynski - Early Life	DOCUMENT			Berkeley, CA Chicago, IL	He was born in CHICAGO, Illinois, where, as an intellectual child prodigy, he excelled academically. From an early age, Kaczynski was accepted into Harvard University at the age of 16, where he earned an undergraduate degree, and later earned a PhD in mathematics from the University of Michigan. He became an assistant professor at the University of California, BERKELEY at age 26, but resigned two years later.
1	Ted Kaczynski - Early Life	DOCUMENT	Full date	4/24/1955	Berkeley, CA	In 1974, he moved to a remote cabin without electricity or running water, in Lincoln, Montana, where he lived as a recluse while learning his skills in an attempt to become self-sufficient; he decided to start a bombing campaign after watching the failures around his house being destroyed by development. From 1978 to 1995, Kaczynski sent 15 bombs to targets including universities and airlines, killing three people and injuring 23. Kaczynski sent a letter to The New York Times on APRIL 24, 1995 and promised "to desert from terrorism" if The Times or The Washington Post published his manifesto. In the industrial

EXTRACT LOCATIONS AND DATES

- Automatically extract from documents all references to locations, dates and times.
- Attach codes with corresponding geographic or time tag to retrieved segments.



IMPORT AND CODE IMAGES

- Import photos in JPEG and PNG as well as images from GIS software along with their coordinates.
- Associate geographic coordinates, dates and times to image areas.

Step 2 - RETRIEVING

RETRIEVE GEOTAGS

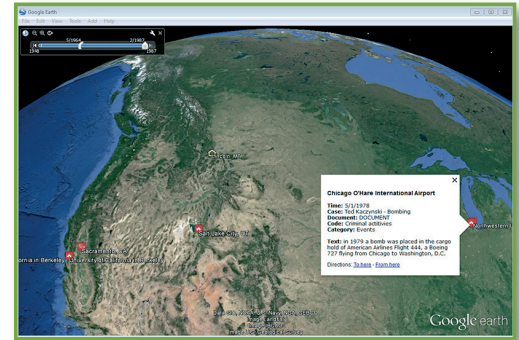
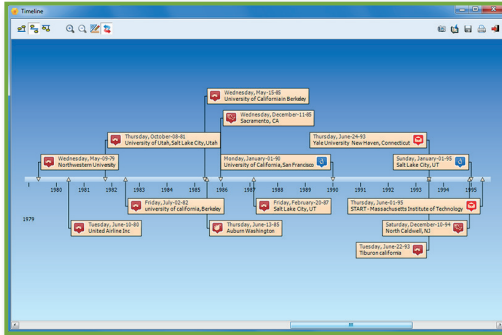
- Search and retrieve all geo-tags or time-tags in your entire project.
- Filter on a specific time period or on a distance from a geographic location.

Case	Document	Code	Link type	Link/Label	Latitude	Longitude	Date	To Date
Ted Kaczynski - Early Life	DOCUMENT	Academic Career	Geographic Location	Evergreen Park School District 124	41° 47' 50.967" N	87° 42' 12.157" W	5/1/1953	6/28/1957
Ted Kaczynski - Early Life	DOCUMENT	Academic Career	Geographic Location	Evergreen Park Community HS	41° 47' 46.887" N	87° 42' 5.207" W	5/1/1953	6/24/1958
Ted Kaczynski - Early Life	DOCUMENT	Academic Career	Geographic Location	Enter Harvard University	42° 22' 39.527" N	71° 7' 24.557" W	5/1/1958	5/1/1962
Ted Kaczynski - Career	DOCUMENT	Academic Career	Geographic Location	University of Michigan	42° 15' 14.717" N	83° 42' 8.977" W	5/1/1962	5/1/1967
Ted Kaczynski - Career	DOCUMENT	Academic Career	Geographic Location	University of California Berkeley	37° 52' 13.817" N	122° 15' 36.377" W	5/1/1967	5/1/1969
Ted Kaczynski - Career	DOCUMENT	Rehabilitation	Geographic Location	Lombard, IL	41° 52' 48.117" N	88° 07' 38.247" W	5/1/1971	5/1/1973
Ted Kaczynski - Career	DOCUMENT	Rehabilitation	Geographic Location	Lincoln, ME	46° 57' 17.787" N	122° 40' 54.157" W	5/1/1977	5/1/1978
Ted Kaczynski - Bombing	DOCUMENT	Criminal activities	Geographic Location	Northwestern University	42° 12' 28.867" N	87° 42' 30.787" W	5/1/1978	5/1/1978
Ted Kaczynski - Bombing	DOCUMENT	Criminal activities	Geographic Location	Chicago O'Hare International Airport	41° 58' 35.887" N	87° 54' 18.767" W	5/1/1978	5/1/1978
Ted Kaczynski - Bombing	DOCUMENT	Criminal activities	Geographic Location	University of California Berkeley	37° 52' 13.817" N	122° 15' 36.377" W	5/1/1985	5/1/1985
Ted Kaczynski - Bombing	DOCUMENT	Criminal activities	Geographic Location	Sacramento, CA	38° 54' 53.687" N	121° 59' 39.847" W	1/21/1985	1/21/1985
Ted Kaczynski - Bombing	DOCUMENT	Criminal activities	Geographic Location	Salt Lake City, UT	40° 47' 38.807" N	111° 57' 27.777" W	2/29/1987	2/29/1987

Step 3 - VISUALIZING

DISPLAY GEOTAGS ON A MAP

- Export retrieved geographic locations into KMZ or KML file and display them in Google Earth, ArcView, or ArcGIS.
- Create thematic maps or heatmaps using the integrated GIS Viewer.



CREATE TIMELINES

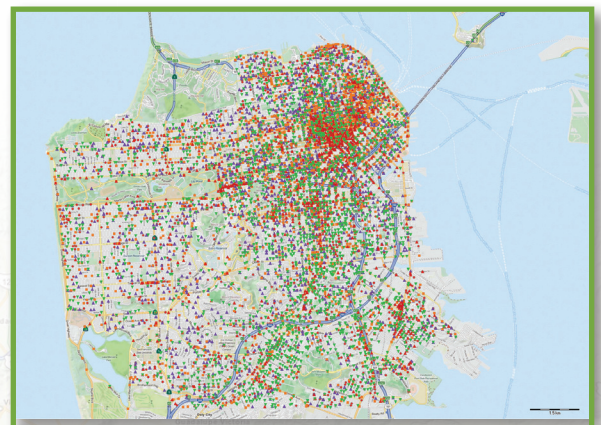
- Organize information chronologically by automatically generating timelines from time tags.
- Link back from any timeline marker to the associated text segment or image area.
- Fully customize the position and appearance of timeline annotations.

» Enhance your research findings through comprehensive mapping

Geographic information associated with one's core research can also be stored or referenced separately as metadata in accompanying fields. WordStat 7 and QDA Miner 5 feature a geocoding service to convert city names, postal codes and even IP addresses into latitudes and longitudes. One may then use the GIS mapping and editing module to relate such information to themes, topics, keywords and key phrases automatically extracted or manually coded. One can build maps to display the spatial distribution of themes, identify trends, customer concentrations, crime hotspots, demographic shifts, or any other geographic representation tailored to your research data.

INTERACTIVE PLOT OF DATA POINTS

- Easily create plots of data points from words, phrases, or topics extracted from unstructured text fields.
- Quickly filter data points on categorical, numerical, and date variables or create dynamic range displays and custom animations to easily identify temporal trends, cyclical patterns or relationships to numerical variables.
- Customize and annotate single data points.

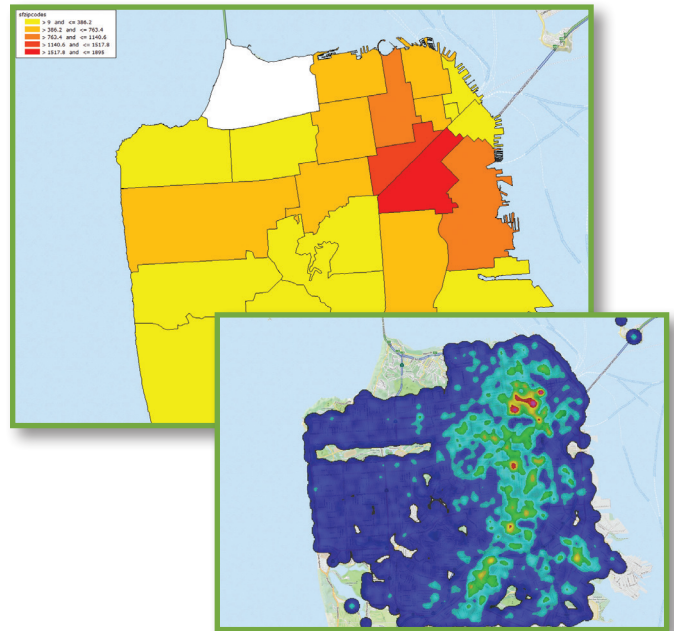


DISTRIBUTION MAPS

- Create layers from various vector file formats and produce choropleth maps to represent point density, demographic information stored in shapefiles, or statistical summaries of numerical values associated with text segments.
- Easily adjust the color range, the number of steps and level of transparency.

HEATMAPS

- Visualize the density of data points with heatmap displays to easily identify customer concentrations, crime hot spots, or disease outbreaks.
- Choose from a wide variety of color ramps or create your own.
- Create heatmaps on all data points or on selected regions.



OTHER GIS FEATURES

- Natively opens and displays a wide range of vector, image, grid, and SQL database layer formats, including advanced spatial server geodatabases, WMS, WFS, and WMTS mapping services.
- Comprehensive visual layer property, legend, and scale controls provide for deep customization of the map appearance.
- Create, edit, translate, and export map layers into a number of vector, image, grid, and SQL database formats.
- Support for coordinate systems with on-the-fly layer reprojection between thousands of predefined geographic and projected coordinate systems or any coordinate system defined from 150+ projections and 900+ datums.
- Integrated geocoding of cities, countries, zipcodes, and IP addresses.
- Extract gazetteers from SHP, KML, KMZ, GPX, or from our online geographical database.

Download a trial version at: provalisresearch.com/trial