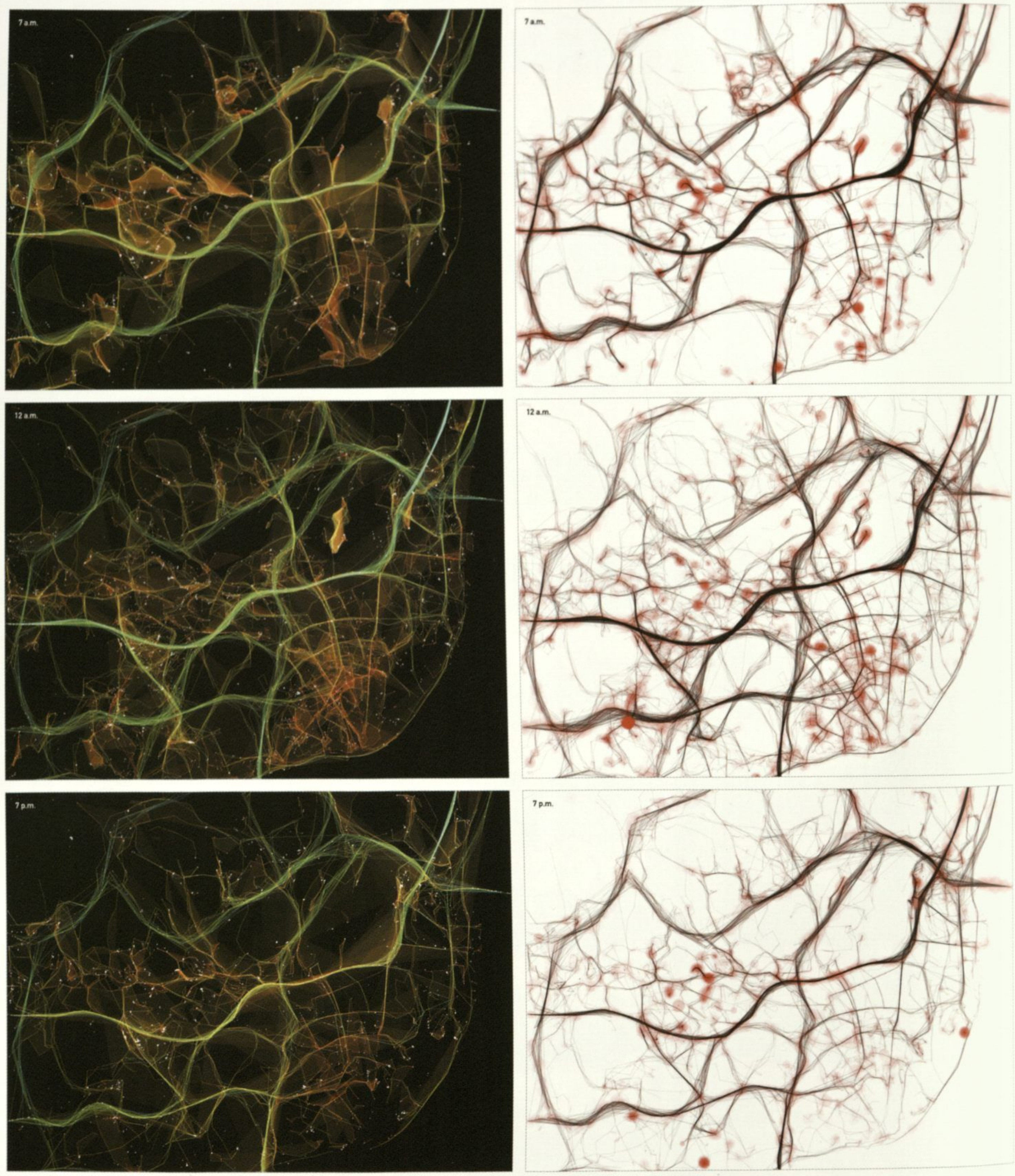


Pedro Cruz, Penousal Machado, and João Bicker (University of Coimbra with MIT CityMotion), Portugal: "Traffic in Lisbon," 2010.

"Traffic in Lisbon" is a series of animations of traffic's evolution in Lisbon during a fictitious twenty-four-hour period (from 0:00 to 23:59). The project maps 1,534 vehicles during October 2009 in Lisbon, leaving route trails and condensed into one single (virtual) day. The two sequences are frames from animations exploring different visual metaphors of the city as an organism with circulatory problems.

In the left sequence, recent paths are color coded according to the vehicle's speed: green and cyan for faster vehicles, yellow and red for slower ones. The accumulation of paths emphasizes main arteries, resulting in thicker lines. The right sequence presents the living organism metaphor by depicting slow vehicles as red circles. Cruz explains, "The superimposition of slow vehicles forms solid red clots in the traffic of Lisbon, depicting it as a living organism with circulatory problems."[15]

http://pmcruz.com/information-visualization/traffic-in-lisbon-condensed-in-one-day

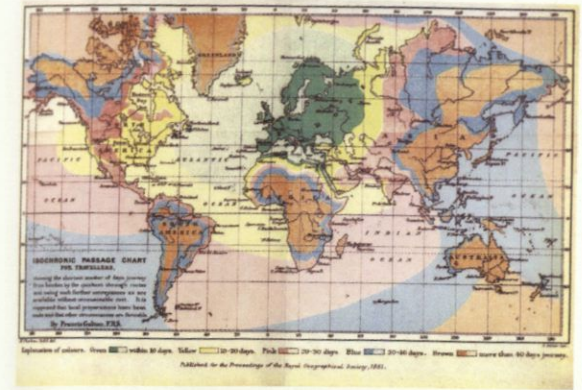
CHAPTER 5

SPATIO-TEMPORAL STRUCTURES

We are surrounded by changes in all dimensions of our existence. All changes require time to become something else, to transform, to remodel, to reorganize, to disappear, and so on. Several fields use time-varying data to understand patterns in natural and social phenomena as well as to help make predictions. Examples range from studies in meteorology and economics to assessment of brain activity.

The chapter focuses on spatio-temporal phenomena and processes inherent to the dimensions of space and time. Data belonging to both space and time are found in diverse domains and include mobility, dispersion, proliferation, and diffusion, to mention a few. Our lives are immersed in time and space, and we constantly reason about both, making decisions about where and when we are, were, or will be. From sketches we draw on napkins to give directions to our friends, to more complex cartographic representations of the

The English astronomer Edmond Halley, known for the comet bearing his name, mapped his prediction of the trajectory of the total eclipse of the Sun in 1715. The map was first published in a leaflet before the eclipse and widely distributed in England. After the event, Halley received observation reports and revised the map in the format that we see here. The map effectively represents a temporal event onto a geographic context: It depicts the passage of the shadow of the Moon across England by graphic means, including the varying duration of the event. Robinson explains, "The use of the shading shows how fertile and imaginative was Halley's grasp of the potentialities of graphic portrayal. The dark ellipse-like figure representing totality was to 'slide' along the shaded path from southwest to northeast, and the relative duration of totality for any place along the path was shown by the width of the ellipse in line with that place."[16]



Sir Francis Galton created the "Isochronic Passage Chart for Travellers" for the Royal Geographical Society in 1881. The map uses Mercator projection and shows the number of days it takes to travel from London to other parts of the globe. Galton's source data were timetables of steamship companies and railway systems. Vasiliev explains, "This world map uses isochrones to separate areas that may be reached in a certain number of days. The isochrones themselves are not labeled, but the areas between them are color coded to the legend, each color indicating the number of days required to reach that area from London: yellow for 10–20 days, brown for more than 40 days, and so forth. It is interesting to note that in traveling across the United States to the West Coast, going through Denver and Salt Lake City to San Francisco took 10–20 days whereas travel anywhere north or south of Denver and Salt Lake City took 20–30 days—a direct effect of the railroads and their routes through the Rocky Mountains. On this map, the temporal unit is a 10-day journey 'by the quickest through routes and using such further conveyances as are available without unreasonable cost.' The actual mileage traveled is not necessary; this is a guide to the traveler to help plan the start of a world-wide tour."[17]

real world, we have traditionally used maps as models for spatial reasoning and decision making. Similarly, we have been using maps to represent and help us reason about spatio-temporal phenomena.

Given the dynamic nature of spatio-temporal phenomena, the designer faces several challenges in representing the fluidity of time in space, especially in static form. Geo-visualization is the field involved with designing and developing tools for interactive and dynamic visual analysis of spatial and spatio-temporal data. Interactive tools often make use of multiple linked displays to represent all aspects of spatio-temporal data, in that maps alone usually are not enough and need other visual displays such as statistical graphs to complement the complexities of the phenomena.

Vasiliev explains that time has been used and represented in different ways in different geographies. She identifies four main areas:[1]

- Historical geography: What has happened where in past times.
- Cultural geography: Where events have happened through time.
- Time geography: How much time it took for events to happen in space.
- Quantitative geography encompasses spatial diffusion and time-series analysis: What occurred where in known periods of time.

TYPES OF PHENOMENA

Spatio-temporal phenomena can be organized into three main types:[2]

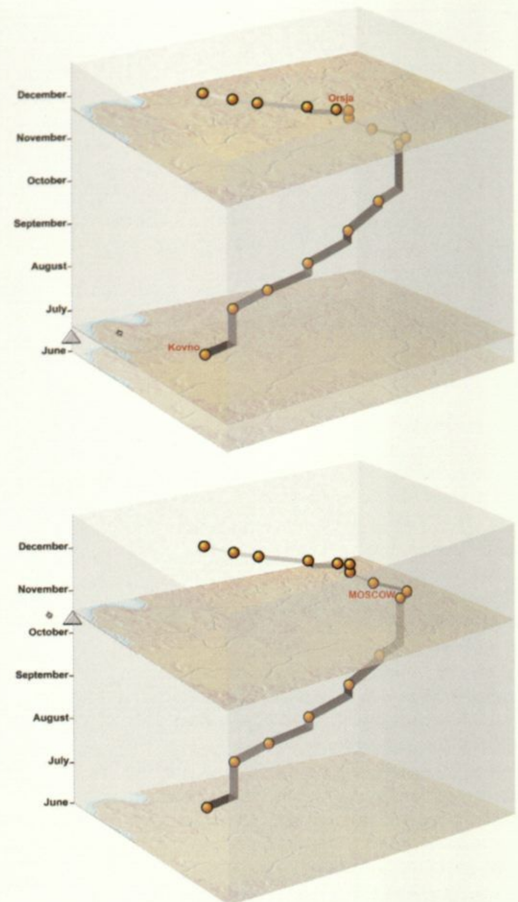
- Existential changes refer to changes in instant events, such as the appearing or disappearing of objects and/or relationships.
- Spatial changes refer to changes in spatial properties of objects, such as location, size, and shape.
- Thematic changes refer to changes in the values or attributes of space, such as in demographic spatial maps.

When representing objects moving in space across time, it is possible to depict spatio-temporal data values as a trajectory that will show several time points on the map. A historical example is the prediction of the total eclipse of the Sun in 1715 by British astronomer and cartographer Edmond Halley. The *New York Times* employed a similar strategy in the recent interactive map of Hurricane Sandy (see page 164). Another common technique is the flow map, which depicts aggregated moving objects in space, such as in the depiction of migration or transportation of people or goods (see page 152). An extension of this technique is the space–time cube, in which time is represented on the third dimension in addition to the two dimensions of the plane for spatial data. An example is Kraak's space–time cube of Minard's *Napoleon March* graphic.

Unlike objects moving across a territory, it is not possible to represent variations of thematic data of continuous spatial phenomena in one image. Take, for example, changes in demographic values of a territory. There are no changes in the spatial values per se (the territory remains in the same location, with the same borders over time); rather, the changes occur in the thematic values represented by them. As Andrienko explains, "It is impossible to observe changes in spatial distribution of attribute values, or to locate places where the most significant changes occurred, or to perform other tasks requiring an overall view on the whole territory."[3] As reviewed in chapter 4, common ways to depict attributes of space at a point in time include choropleth and dot density maps. Adding other types of visual displays to the geographical representation often helps provide temporal context, such as with complementing maps with statistical graphs. A historical and well-known example is Minard's depiction of Napoleon's 1812–1813 Russian campaign, in which the line graph at the bottom adds context to the spatio-temporal information by showing the temperature faced by the soldiers on their way back to France.

To view thematic data changes over time, we need other techniques, such as multiple maps, animation, or interactive tools. Multiple maps involve sequencing a series of single-date maps. The technique provides a simultaneous view of change and enables comparison of same scale maps evenly spaced in the temporal dimension (see page 128). To detect direction and pace of change, the viewer needs to jump from map to map. Overlay of maps might enhance the perception of change, though this is not always possible when dealing with large amounts of data. Monmonier suggests, "Maps in a temporal series are especially useful for describing the spread or contraction of a distribution."[4]

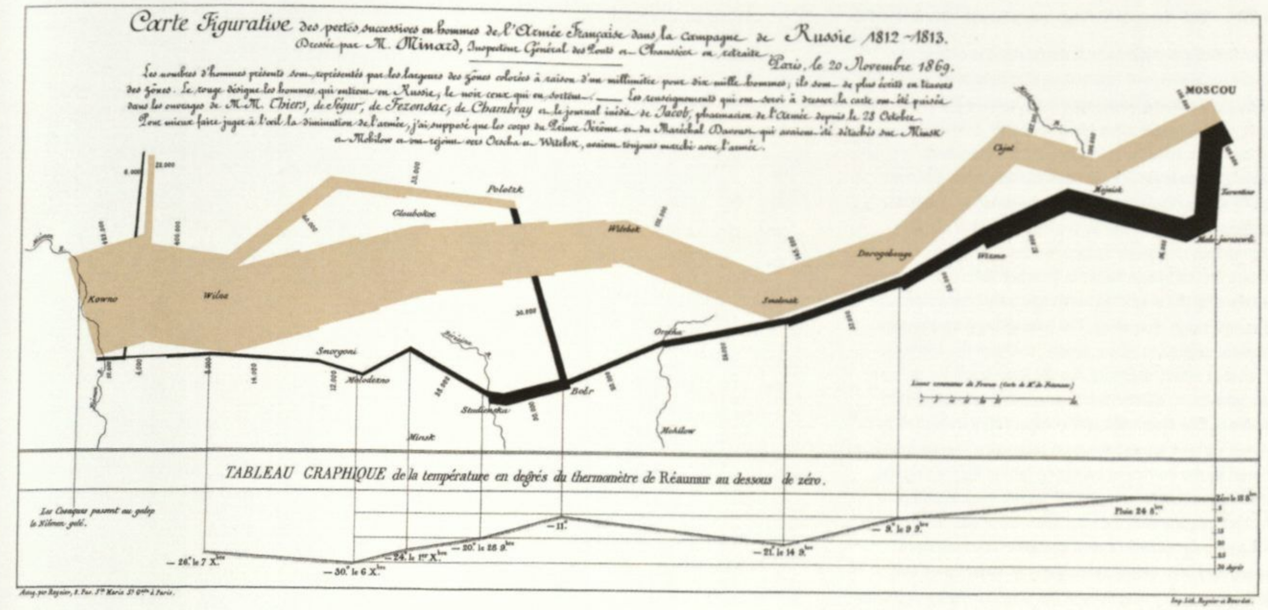
An animation is a sequence of images representing states of phenomena at successive moments in time. In other words, animation depicts phenomena by mapping the temporal dimension in the data to the physical time we experience in real time. However, animations are poor for comparison tasks, because it is difficult to remember previous states with which to make comparisons. Andrienko and colleagues recommend combining interactive functions to animations to allow comparison and trend detection. Due to phenomena that are either too fast or too slow, the physical time scale might change so as to make the phenomena visible. The movies depicting twenty-four hours of traffic in Lisbon by Pedro Cruz are examples of how spatio-temporal data are mapped into physical time.



Menno-Jan Kraak, Netherlands: Space–time cube of Minard's "Napoleon March to and from Russia, 1812–1813," 2002.

Menno-Jan Kraak at the International Institute of Geoinformation Sciences and Earth Observation, Netherlands, created this geovisualization of Minard's map of Napoleon's 1812 campaign into Russia (reproduced on the right) to demonstrate "how alternative graphic representations can stimulate the visual thought process."[10] The interactive visualization is a space-time cube in which the *x*– and *y*–axes represent the geography and the *z*–axis represents time. One can navigate in time by moving the cursor in the vertical direction as the screenshots above illustrate.

www.itc.nl/personal/kraak/1812/3dnap.swf



Charles Joseph Minard's 1869 "Napoleon March to and from Russia, 1812–1813" display combines statistical data with a timeline, and spatio-temporal information about the French army. In this multivariate display, the line width represents the number of soldiers marching to and from Russia, with each millimeter standing for 10,000 men. The march starts with 420,000 men in the Polish–Russian border (center left, beige line), reaches Moscow with 100,000 (top right), and ends with 10,000 men (black line). Considering that our visual system is unable to perceive absolute quantities from areas, Minard provides absolute quantities of soldiers along the two lines. Minard removed most cartographic information and kept only geographical landmarks, such as main rivers and cities. The line graph at the bottom represents the temperatures faced by the army on the way back to Poland, which are associated with the line standing for the return trip. Connections between temperatures and the march offer new levels of information: the relationships between deaths and low temperatures (probably also aggravated by fatigue). For example, 22,000 men died crossing the River Berezina due to the extreme low temperatures (−20°C [−4°F]).

Andrienko and colleagues represented the same spatio-temporal data using three different kinds of visual displays: static small multiple maps, animation, and interactive animation. The study found that the types of display affect the analytical and inference processes. People using the multiple maps display were more focused on spatial patterns rather than on events and temporal processes, whereas those using the animation and the interactive display focused more on changes and events rather than on spatial configurations.[5]

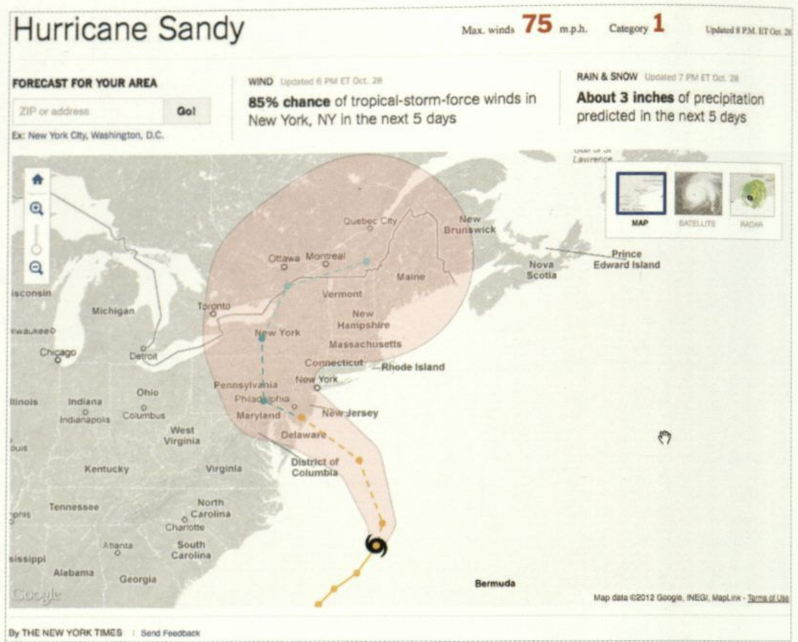
TIME

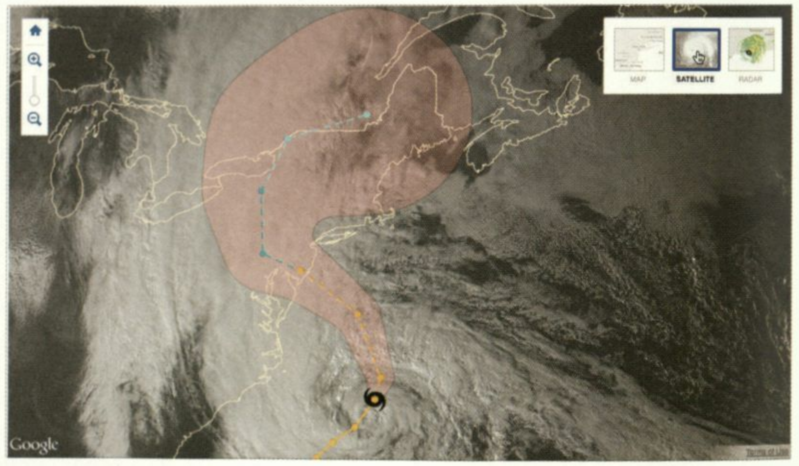
Andrienko and colleague distinguish two temporal aspects that are crucial when dealing with spatio-temporal data: temporal primitives and the structural organization of the temporal dimension.[6] There are two types of primitives: time points (point in time) or time intervals (extent of time). And there are three types of structures: ordered time, branching time, and multiple perspectives. Ordered time is the most commonly used structure and is subdivided into linear and cyclical. Linear time provides a continuous sequence of temporal primitives, from past to future (e.g., timelines), and cyclic time organizes primitives in recurrent finite sets (e.g., times of the day). Branching and multiple perspective times are metaphors for representing alternative scenarios and more than one point of view, respectively. When representing spatio-temporal phenomena, the designer needs to make a series of decisions concerning the visual method, whether the most effective representation would deal with linear time or cyclic time, time points or time intervals, ordered time or branching time, or time with multiple perspectives.

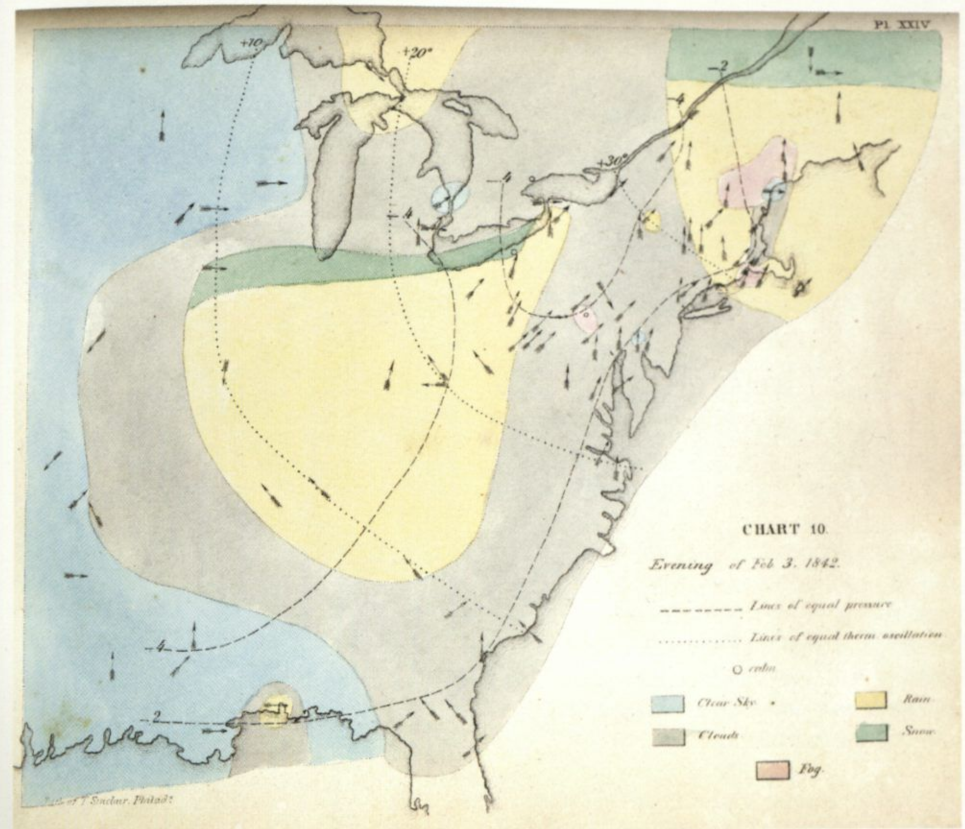
The *New York Times*, U.S.: "Hurricane Sandy," 2012.

When facing potential natural disasters, it is crucial to provide residents with information that help them make decisions that sometimes might even involve life and death, such as in the case of earthquakes, hurricanes, and tsunamis. News weather maps, websites, and television broadcast are common media where we look for information that can help us prepare for such events. The *New York Times*' interactive map provided many features that effectively helped residents on the East Coast prepare for Hurricane Sandy in October 2012. It presented readers with the predicted hurricane path connected with times and storm intensities. The interactive map answers questions related to when, where, and how the storm is forecast to affect residents. A solid line stands for the past path, whereas a dashed line represents future predicted trajectory. The dimension of the impact is represented by a colored surface around the main trajectory. The surface is colored by the hurricane category, further increasing the number of variables represented on the map. In addition, when interacting with the map, the viewer gets information for a particular point in space and time. The map itself carries very little detail, depicting only major cities and state borders. The simplicity of the map facilitates detection and focus on the main issue, which is the spatio-temporal route of the hurricane.

www.nytimes.com/interactive/2012/10/26/us/hurricane-sandy-map.html?hp



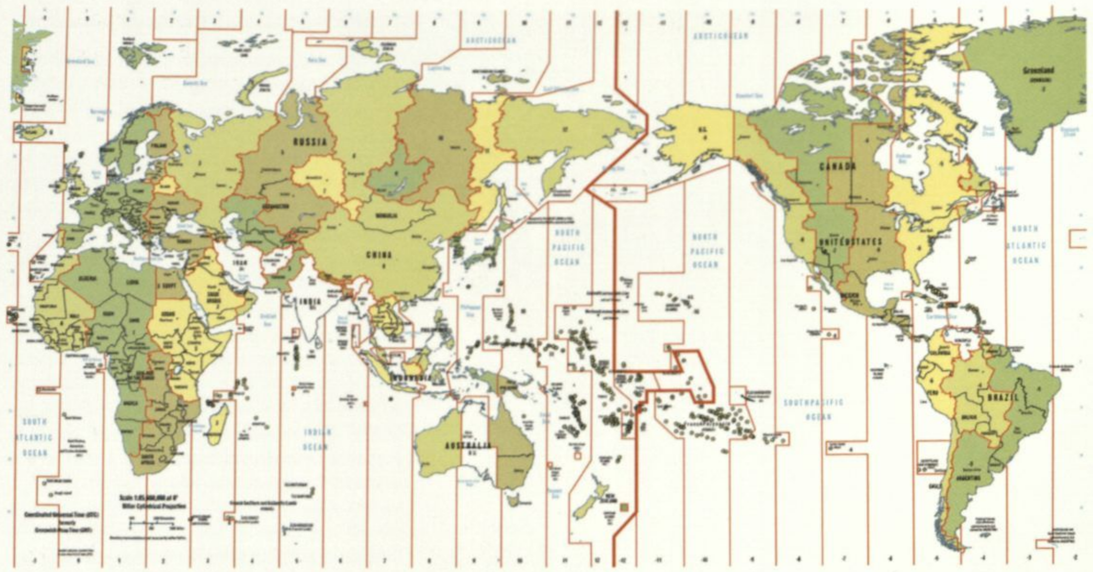




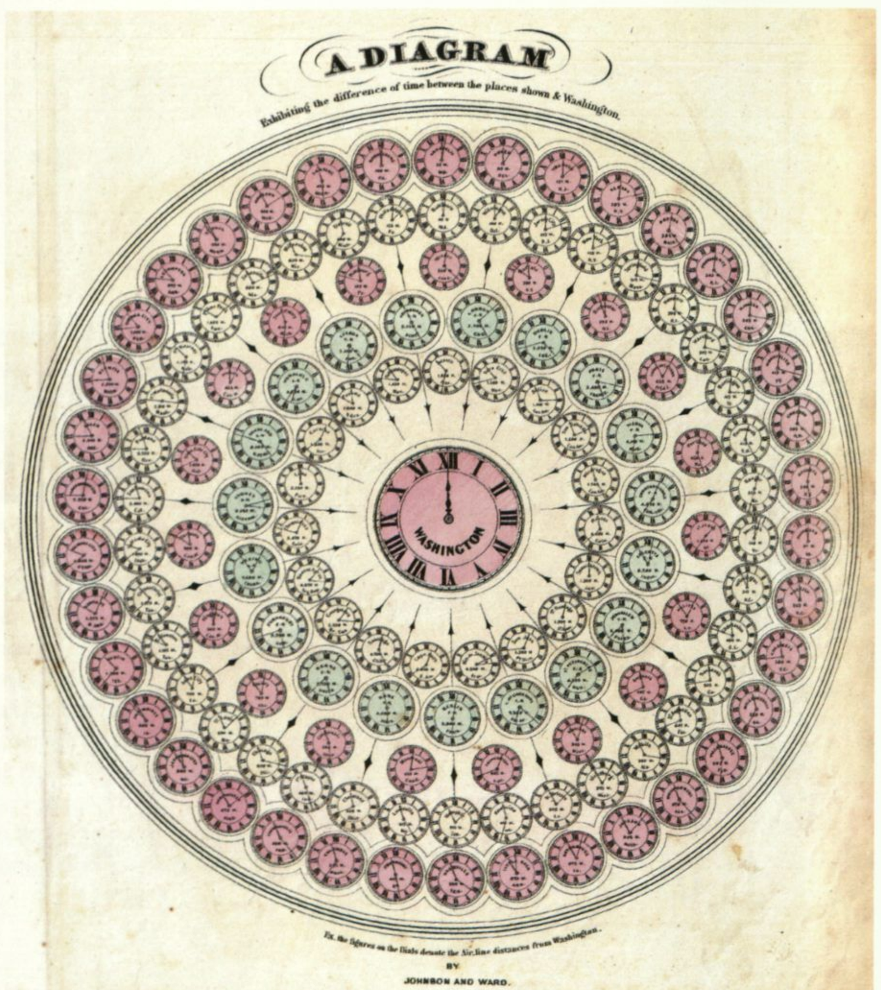
The American mathematician Elias Loomis, known for his textbooks on math, also significantly contributed to meteorology, proposing a system of observers and daily weather maps that resulted in Congress's creation of the Weather Bureau of the United States Signal Service in 1870, today's National Weather Service.

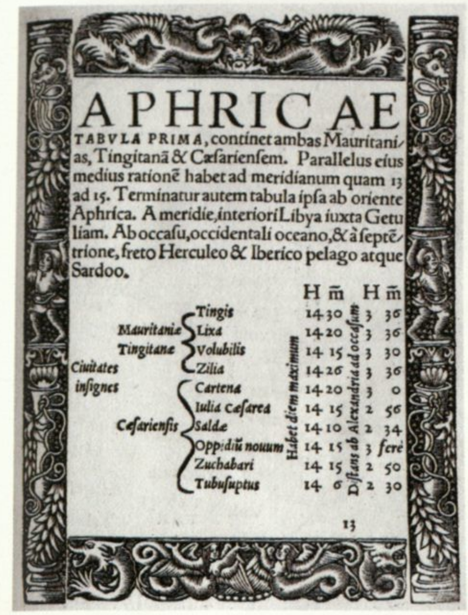
This map is one of thirteen charts published in his article "On Two Storms Which Were Experienced throughout the United States, in the Month of February, 1842." It depicts Loomis's observations on the storms over a wide region in the eastern half of the United States and over several days. Delaney observes, "In two series of sequential maps (dated morning/evening, day), he drew lines of equal deviations in barometric pressure and equal oscillations in temperature, and assigned colors to areas of clear sky, clouds, rain, snow, and even fog. In addition, Loomis used arrows of varying length to indicate wind direction and intensity. In fact, he was anticipating common characteristics of the modern weather map: when the Signal Service's weather maps began appearing in 1871, they were constructed on Loomis's model."[19]

In chapter 2, we saw that time has an inherent semantic structure and a hierarchic granularity that ranges from nanoseconds to hours, days, months, years, millennia, and so on. When structuring and devising measurement systems for time, we have relied traditionally on spatial metaphors as well as on the observation of the motion of celestial objects. As Vasiliev expounds, "The motions of these heavenly bodies, which were used either to be time or to measure time, occurred in space. It was the relationships that these objects had to each other in space—in the sky—that determined what time it was. From the earliest clocks, the measurement of time depended on spatial relations: where the shadow of the sundial's gnomon falls; how much sand passes from one bowl to the other in an hourglass; the amount a candle burned down past hourly markings. Morning begins when the Sun rises, and night when it sets, and these describe the day. The clock face with its numbers and the moving minute and hour hands could be considered a dynamic map of time. We tell what time it is by understanding the spatial relationship between the numbers and where the hands are pointing."[7]



In 1878, Canadian engineer Sir Sanford Fleming proposed a system of worldwide time zones based on lines of longitude by dividing the Earth into 24 time zones (15° wide), with one zone for each hour of the day. The Greenwich Meridian was chosen as the 0° line of longitude, the start point of the system. The endpoint of the system is the 180° line of longitude, that resulted in the creation of the International Date Line. This Pacific-centered map shows the agreed upon time zones in the world for 2012, with the International Date Line represented by the thick red line zigzagging the map vertically.





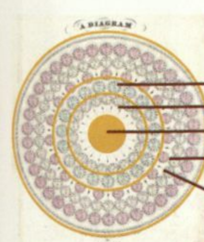
APHRICAE

TABVLA PRIMA, continet ambas Mauritanias, Tingitanã & Cæsarienſem. Parallelus eius medius rationẽ habet ad meridianum quam 13 ad 15. Terminatur autem tabula ipſa ab oriente Aphrica. A meridie, interiori Libya iuxta Getuliam. Ab occaſu, occidentali oceano, & à ſeptẽtrione, freto Herculeo & Iberico pelago atque Sardoo.

			Habet diem maximum H	m̃	Distans ab Alexandria ad occaſum H	m̃
Ciuitates inſignes	Mauritaniæ Tingitanæ	Tingis	14	30	3	36
		Lixa	14	20	3	36
		Volubilis	14	15	3	30
		Zilia	14	26	3	36
	Cæſarienſis	Cartena	14	20	3	0
		Iulia Cæſarea	14	15	2	56
		Saldæ	14	10	2	34
		Oppidũ nouum	14	15	3	ferè
		Zuchabari	14	15	2	50
		Tubuſuptus	14	6	2	30

13

This woodcut table, *Aphricae Tabula I*, was reproduced in Sebastian Münster's 1540 edition of Ptolemy's *Geographia*. Delaney explains, "For each listed North African location, the data in the table show the length of its longest day (in hours and minutes) and its distance (in hours and minutes, hence time) west from Alexandria, Egypt."[20]



Alvin Jewett Johnson designed this world time zones diagram for publication in his *New Illustrated Family Atlas* in 1862. The circular diagram depicts the differences in time between places in the world. It is structured around Washington, DC, which is represented as a clock with the time set at 12. Other major cities in the U.S. and the world surround it with clocks adjusted accordingly.

A familiar example of the spatialization of time is the longitude coordinate system that uses space to organize time. The system both locates places cartographically and measures time as arc distances based on divisions of the globe into 360 degrees, where one hour corresponds to 15 degrees of longitude. The Prime Meridian is the starting line that divides the globe into time zones measured as differences between a particular location and the Coordinated Universal Time (UTC). Vasiliev explains that the longitude system helped standardize time around the globe. "In order to understand the standardization of time worldwide, it is important to map it.... The important progression here is from the acknowledgment that the Sun shines on the Earth's surface in different places at the same time, to the post-Industrial Revolution need to have all humans in any one place observe the same (standard) time and have them understand why time is standard and what the correct time is."[8]

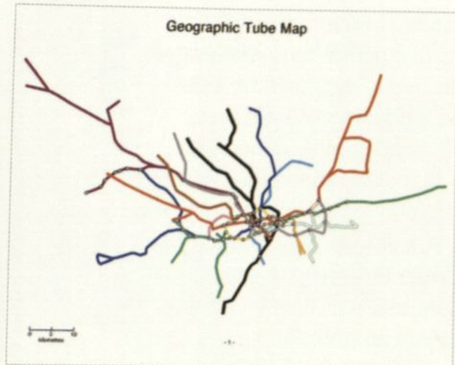
When examining temporal structures in chapter 2, we saw that the Newtonian notion of absolute time was essential to the creation and representation of timelines (see page 88). This is an underlying notion that persists to this day, including visualizations of spatio-temporal data that tend to represent time as ordered. Moreover, the great majority use time points as the primitive in both linear and cyclical ordered temporal structures.

Another temporal feature relevant to the study of spatio-temporal phenomena is that time contains natural cycles and reoccurrences, some more predictable than others. For example, seasons are more predictable than social or economic cycles.[9]

TIME AS DISTANCE METAPHOR

We often use the metaphor of time as distance in our daily lives, such as when we provide temporal measures for giving directions. We say it will take ten minutes to reach the supermarket, it is a three-hour train ride, and so on. There are many instances in which the measure provided by "how long it takes" replaces the spatial distances between places. Isochrone lines and distance cartograms are two common techniques using time distances. Most representations in this category are based on an origin-destination structure, with information centered on a specific spatial point.

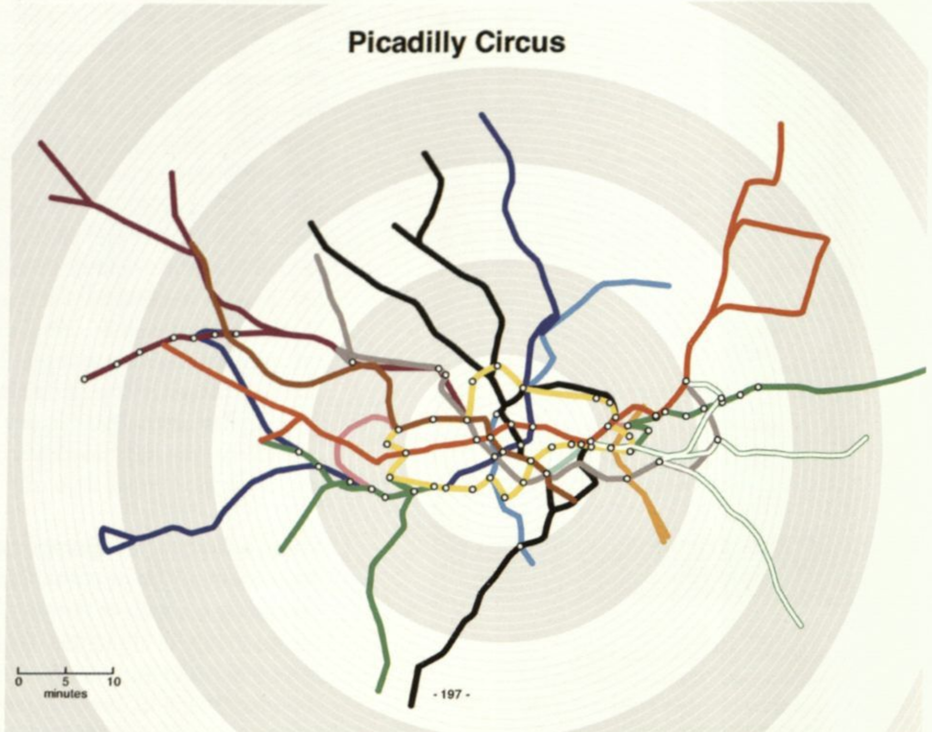
Isochronic maps use isolines of equal travel times constructed from a defined location (origin) to represent spatio-temporal phenomena. In other words, the lines, representing temporal distances, are overlaid on a conventional projection base map, where space is kept constant and the time surfaces conform to the temporal distances as represented by the isochrones. Galton's "Isochronic Passage Chart for Travellers" is a historical example of the technique (see page 161).

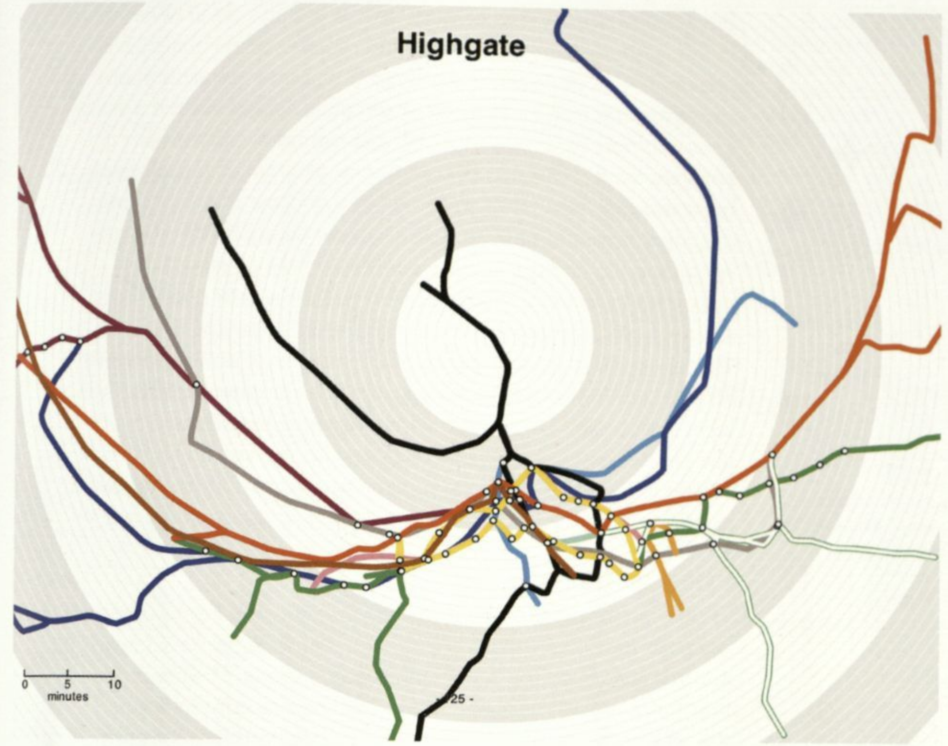


Tom Carden, U.K.:
"Travel Time Tube Map," 2011.

The interactive London Underground map redraws its structure according to the time it takes to travel from a selected departing station. In other words, once a station is selected, it is positioned at the center of a series of concentric circles representing traveling time distances to all other destinations, which are subsequently repositioned. Concentric circles represent ten-minute intervals. To redraw the London Tube map, the software calculates the shortest paths from the origin to the destination stations, with the radius proportional to the time to travel. Tom Carden created this online Java applet in Processing in 2011 as a personal experiment. The top image shows the map rendered according to geographic features, and the other two screenshots show the map centered at Picadilly Circus (left) and at Highgate station in the northern part of London (right).

www.tom-carden.co.uk/p5/tube_map_travel_times/applet



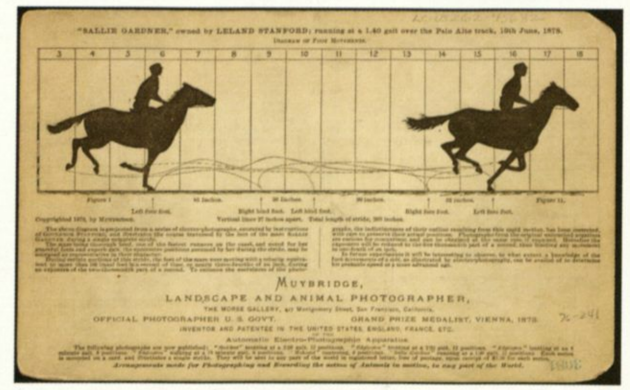


In distance cartograms, a set of concentric circles centered in a specified origin point represents temporal distances, often without a base map, which would be distorted to fit the temporal distances. In other words, it uses temporal distance as a proxy for spatial distance, resulting in distortion of the topology to conform to the temporal measures.

SCALES

Spatio-temporal phenomena exist at different spatial and temporal scales, which significantly affect the extent and amount of detail represented. As seen in chapter 4, maps involve reducing dimensions in order to bring spatial reality to the scale of our human sensory systems. We reduce the three dimensions of space into the two dimensions of maps, and sometimes we reduce even further the three dimensions of space into a one-dimensional element, such as when we represent cities as dots on a map. Similar strategies need to be in place when depicting spatio-temporal phenomena, as MacEachren explains, "Temporally, some geographic space-time processes (e.g., earthquake tremor) are fast enough that we need to slow them down to understand them (as when we 'map' a molecule, cell, or computer chip, for which an increase in scale makes visible a pattern that would otherwise remain hidden). Most temporal geographic phenomena, like spatial ones, have a time span too large to be grasped at once, so therefore we need to compress time as well as space."[10]

We have examined how geographic scale affects the amount of information revealed in maps, where large-scale maps present a larger and more detailed number of features than a small-scale map does (see page 123). Similarly, time can also be scaled at different granularities, affecting the amount of information provided for analysis. Typically, local phenomena are nested within global phenomena, such as the relationships between a local storm and global climate change. The same is true for personal phenomena, in that local phenomena, such as activities within a day, are different when considered within a week (weekdays versus weekends), a year (working versus holidays), or a lifetime. Furthermore, temporal scales involve aggregating time into conceptual units, such as when we use a day for twenty-four hours or divide the week into weekdays and weekends. Decisions will depend on the type of data and the tasks at hand. For example, a multiple map series uses a single granularity, whereas interactive applications tend to offer different scales.



This image was created by Eadweard Muybridge to illustrate a horse in motion running at a 1:40 gait over the Palo Alto track, on 19 June 1878. Muybridge portrays the motion with the aid of a diagram depicting the foot movements between two frames for beginning and end.

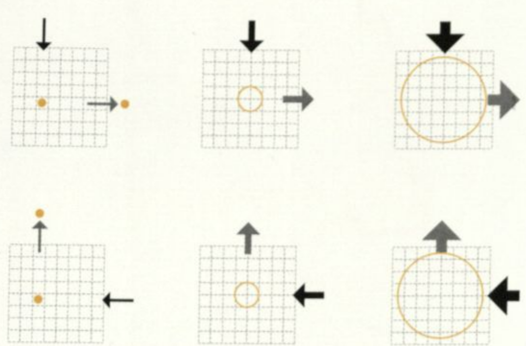


Diagram after Jacques Bertin's information system: question types and reading levels.[21]

Because of the complexities of spatial and temporal dependencies in the representation of phenomena, each scale—spatial and temporal—must match the phenomena under consideration. However, the most adequate or effective scales are not always known beforehand and must be discovered in the process of analysis, which involves trial and error. Interactive visualization tools tend to allow multiscale analysis and the manipulation of both space and time to help discover an appropriate match. As Andrienko and colleagues contend, "Various scales of spatial and temporal phenomena may interact, or phenomena at one scale may emerge from smaller or larger phenomena. This is captured by the notion of a hierarchy of scales, in which smaller phenomena are nested within larger phenomena. Thus, local economies are nested within regional economies; rivers are nested within larger hydrologic systems; and so on. This means that analytical tools must adequately support analyses at multiple scales considering the specifics of space and time."[11]

TYPES OF QUESTIONS

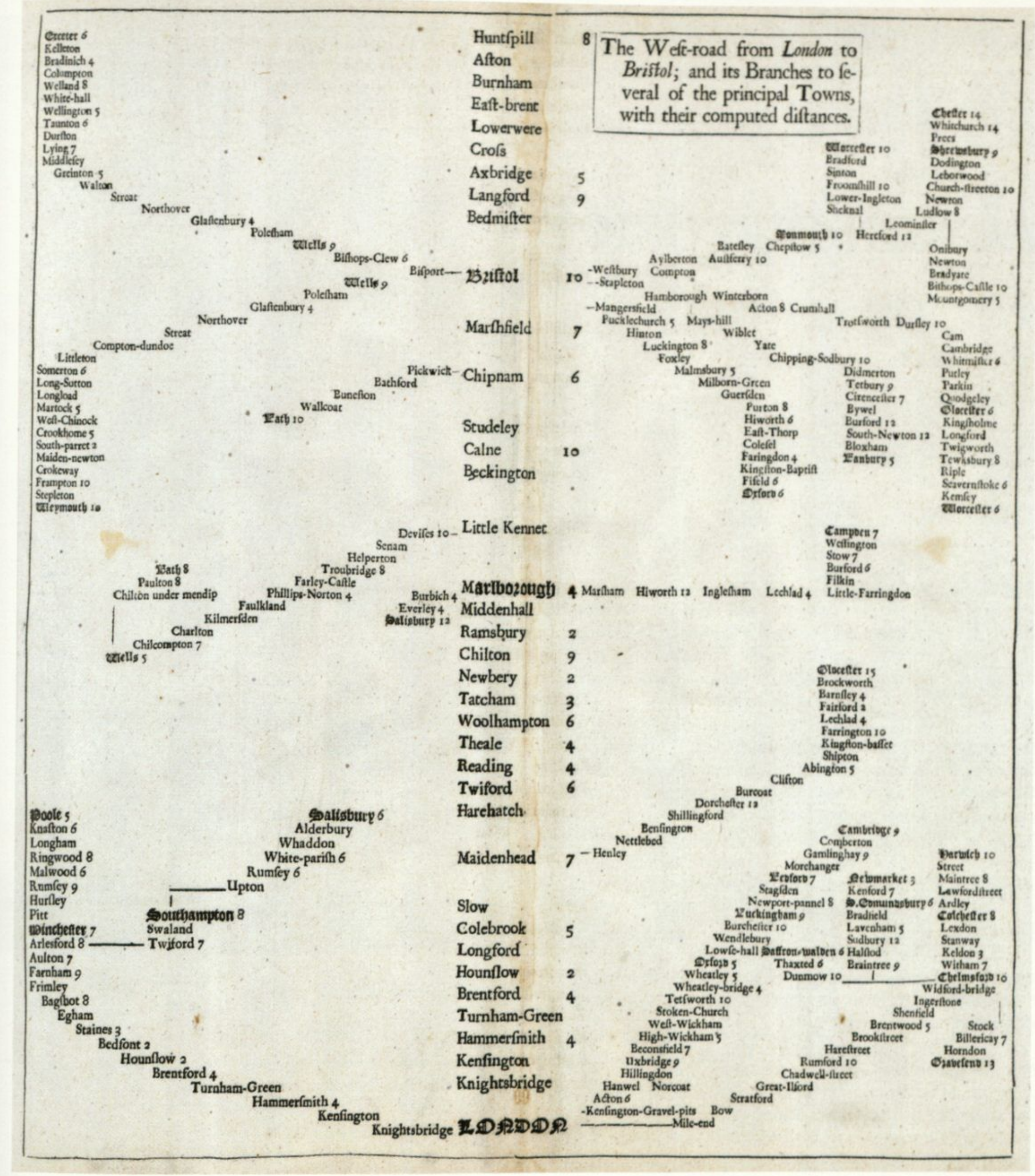
In the seminal book *Semiology of Graphics*, the French cartographer Jacques Bertin identifies two key concepts for visually conveying information: question types and reading levels.[12] Bertin argued that there are as many types of questions as components in the information (data variables). He considered that for each question type there would be three reading levels in the visualization: elementary (datum), intermediate (set of data), and overall (whole dataset).

Following a similar approach, but specifically for spatio-temporal data, Peuquet defined three components: space (where), time (when), and objects (what), allowing three types of questions:[13]

- when + where > what: questions about an object or set of objects at a given location(s) at a given time(s)
- when + what > where: questions about a location or sets of locations for an object(s) at a given time(s)
- where + what > when: questions for a time or set of times for a given object(s) at a given location(s)

Andrienko and colleagues extend the task typology by adding the "identification–comparison" dimension.[14]

There has been an increase in the collection as well as accessibility of spatio-temporal data in recent years due to the various new sensors (GPS, cell phone, etc) and aerial and satellite imagery, which pose new challenges, especially in what concerns techniques for dealing with large amounts of data (big data) as well as dynamic data being sourced in real time. The case studies that follow present projects that address these questions.



"The West-Road from London to Bristol; and Its Branches to Several of the Principal Towns, with Their Computed Distances" was published in John Speed's *The Theatre of the Empire of Great-Britain* in 1676. Delaney explains how this stripped-down map with relative distances functions, "Here, roads consist of stacks of place names; the title one ("West-Road") runs up the spine of the page from London at the bottom. The names of larger towns are printed in bold, old English typeface letters. In the seventeenth century, one's options for leaving London by foot or horse were few. Heading west on this road towards Bristol—which everyone would know ("you need to take the West Road . . .") —one would expect to arrive in Hammersmith after four miles and reach Brentford via Turnham-Green after four more. (These localities are part of Greater London today.) From Maidenhead and Marlborough, other roads are shown going north. This hybrid approach, similar to a subway map today, has been an effective travel tool for over three hundred years."[22]

CASE STUDY

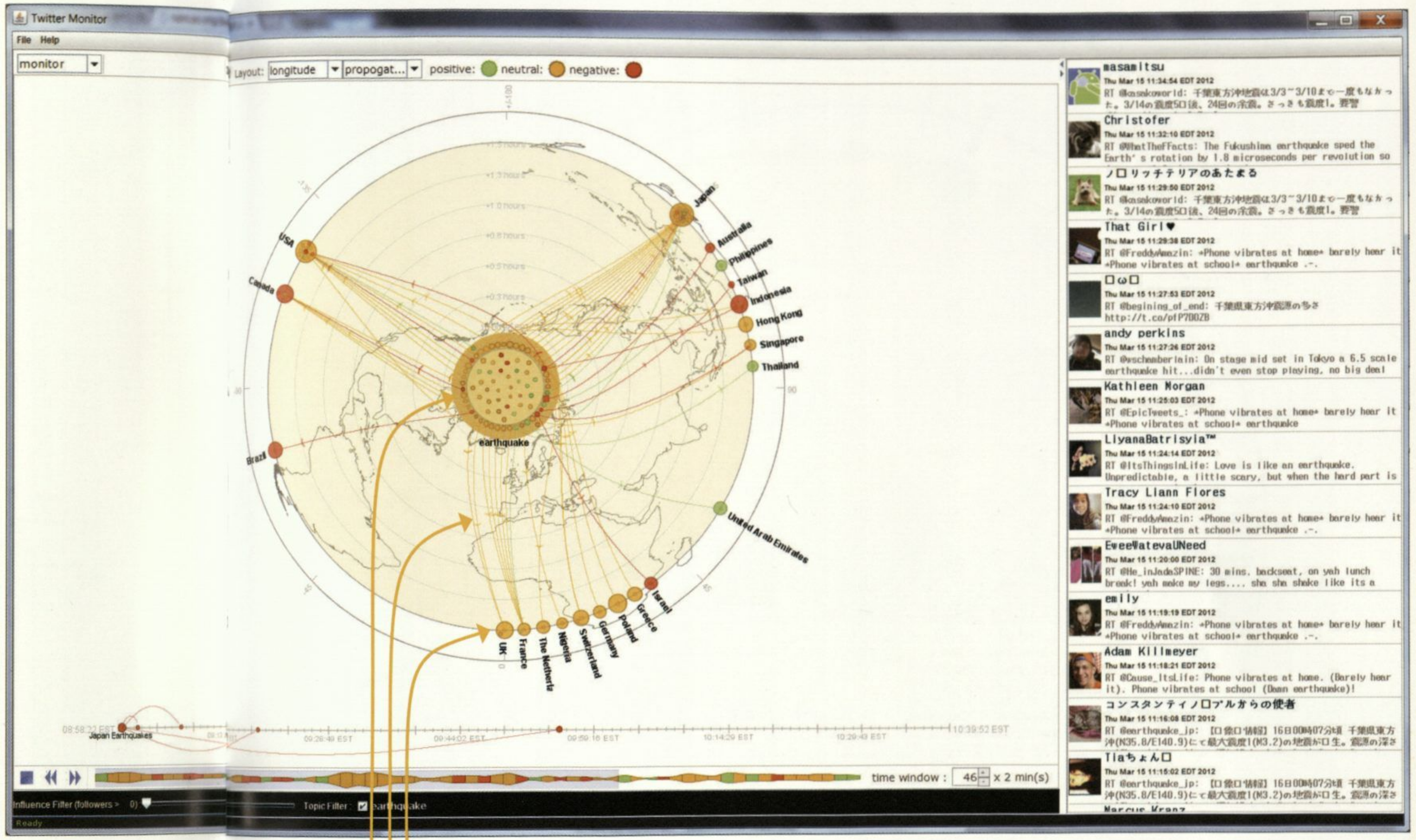
INFORMATION DIFFUSION

Whisper

Whisper is an interactive application that visualizes the process of information diffusion in social media in real time. It tracks the time, place, and topic of information exchanges in the Twitter micro-blog service. It was designed in 2012 by the international team of Nan Cao, Yu-Ru Lin, Xiaohua Sun, David Lazer, Shixia Liu, and Huamin Qu.

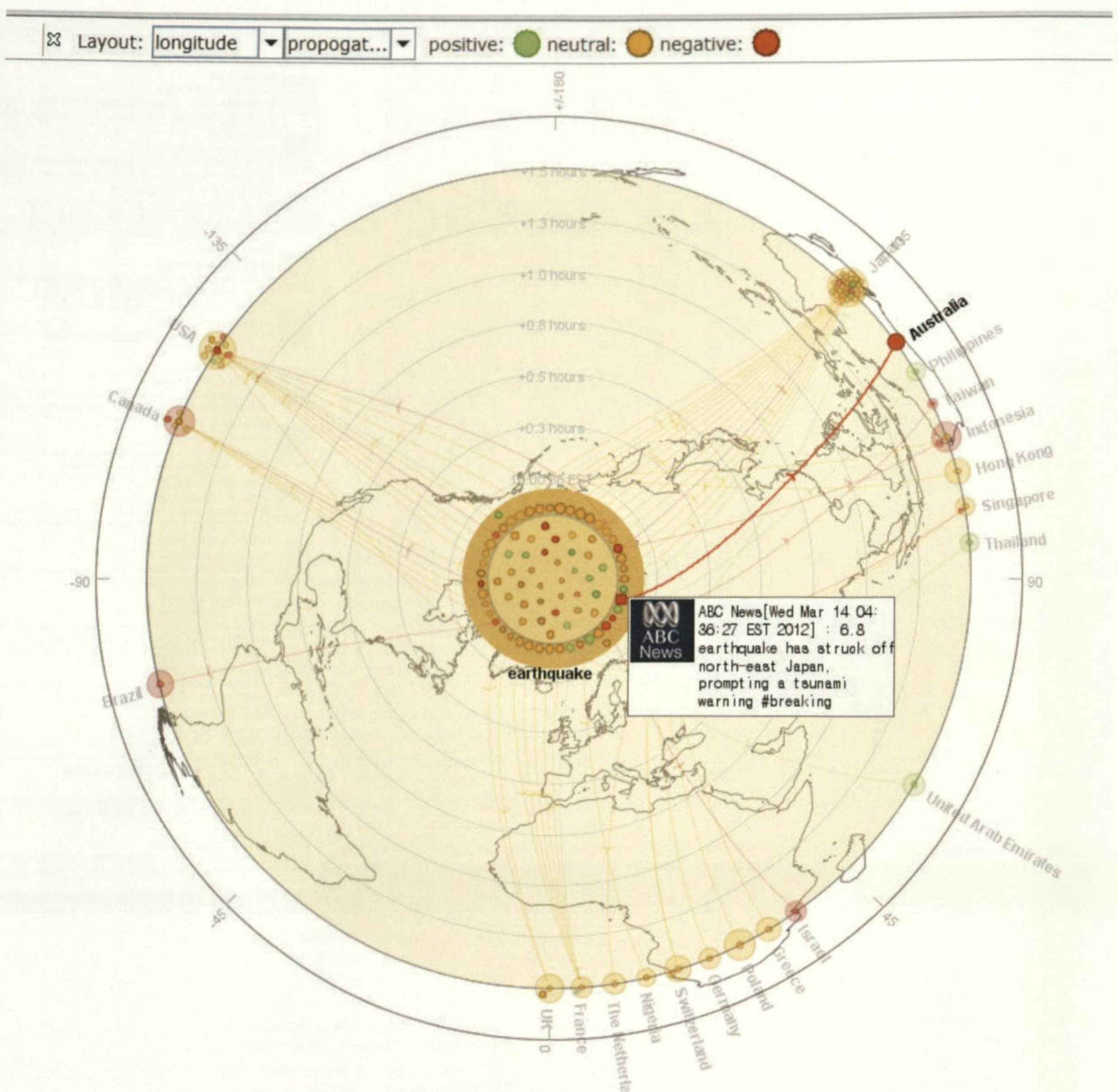
They consider that information spreads from information sources to users, as when users retweet messages, further affecting their followers and ultimately the user's geographic location. Among the relevant features in understanding this process and the effects of information spreading is the role people play in that process, including that of key opinion leaders. Cao and colleagues explain, "Whisper seeks to represent such rich information through a collection of diffusion pathways on which users' retweeting behavior is shown at different levels of granularity. Each pathway is also a timeline whose time span is configurable to enable an exploration of the diffusion processes occurring between two chosen points in time."[23]

The visualization uses the visual metaphor of the sunflower to construct the information space of the narrative, which is then populated by the actors, places, and themes. It uses a single representation with two coordinated views: the dynamic view shows the tweets and retweets generated in real time, and the static view allows exploration of historical data by means of a timeline. There are several dimensions to the data that includes temporal, spatial, spatio-temporal, nominal, and categorical.

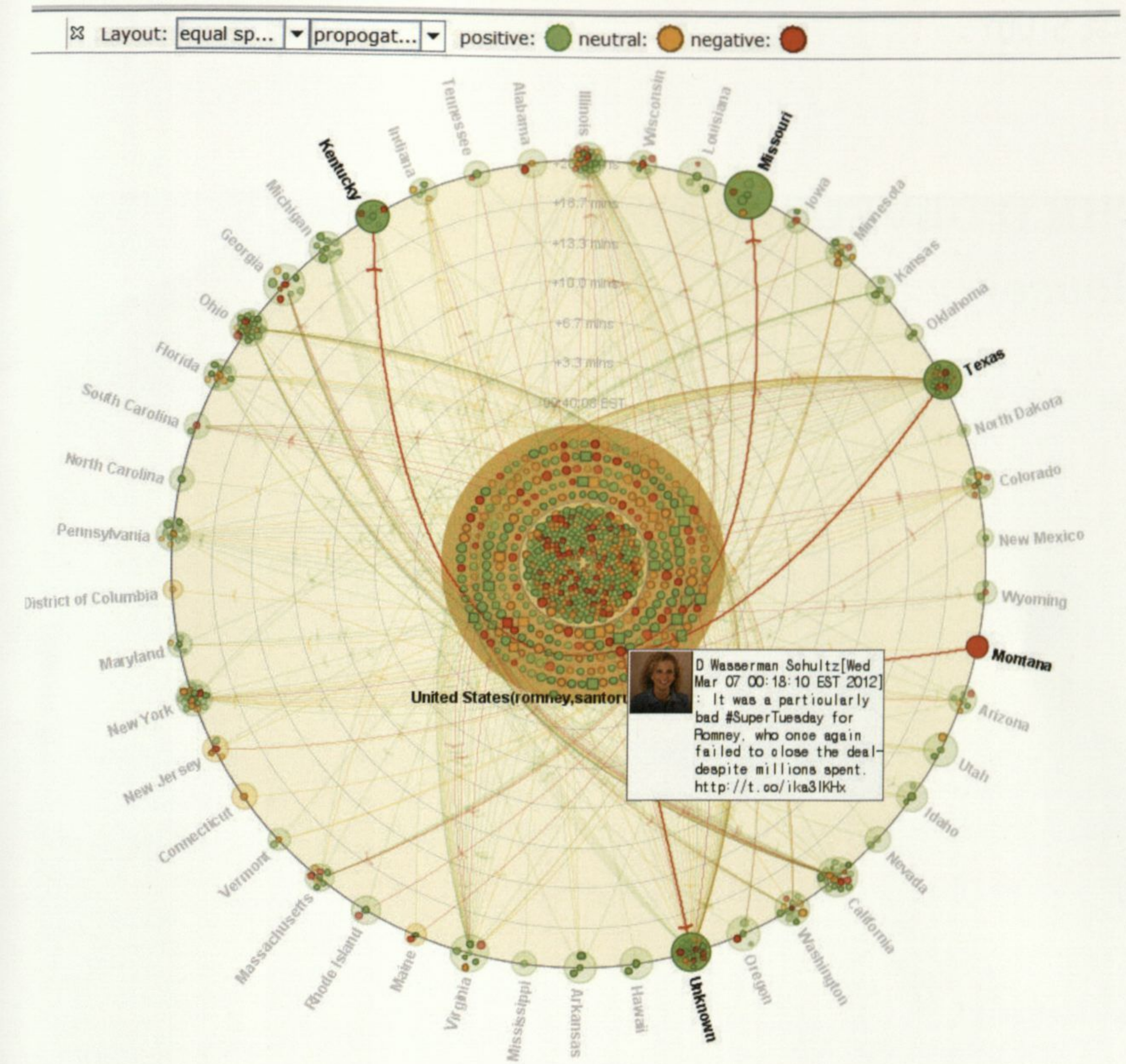


The dynamic view of the visualization is composed of three main elements:
Topic disk: A circular structure holds the tweets. Tweets are placed according to the frequency of retweets in a polar direction, such that once a message is retweeted for the first time, it moves from the center to the periphery of the circle. Tweets that are not retweeted—in other words, those not contributing to any information diffusion—fade out over time, giving place for new tweets.
User group: Retweets are hierarchically grouped by shared topics of interest or shared geographic locations, with the latter geo-located in the map.
Diffusion pathways: The path linking a tweet to the retweet user group provides the diffusion path that is represented as a timeline, with marks standing for retweets over time.

Color hue encodes sentiment on a three-color palette, where red stands for negative, orange for neutral, and green for positive opinions.
Color opacity encodes activeness of tweets or user groups.
Size encodes the expected influence of a tweet, which is calculated by the expected influence of the tweet user based on the number of followers the user has.
Shape encodes the type of user: a square represents users from media outlets or organizations and circles stand for all other users.



The image shows a diffusion of information on Twitter regarding a 6.8 magnitude earthquake and a series of aftershocks and tsunamis that hit the northern coast of Hokkaido island, Japan in 2012. The event caught global attention because the location was one of the areas in Japan devastated by the 2011 disaster. This image shows that some countries, including Australia, were initially concerned about the Pacificwide tsunami threat triggered from the earthquake. The use of the geographic structure for examining this particular event in Whisper is quite effective.



This image depicts the spatial diffusion patterns of the 2012 Republican presidential primaries and caucus results on Super Tuesday. Note spreading of the tweet by opinion leader, Congresswoman Schultz.

CASE STUDY

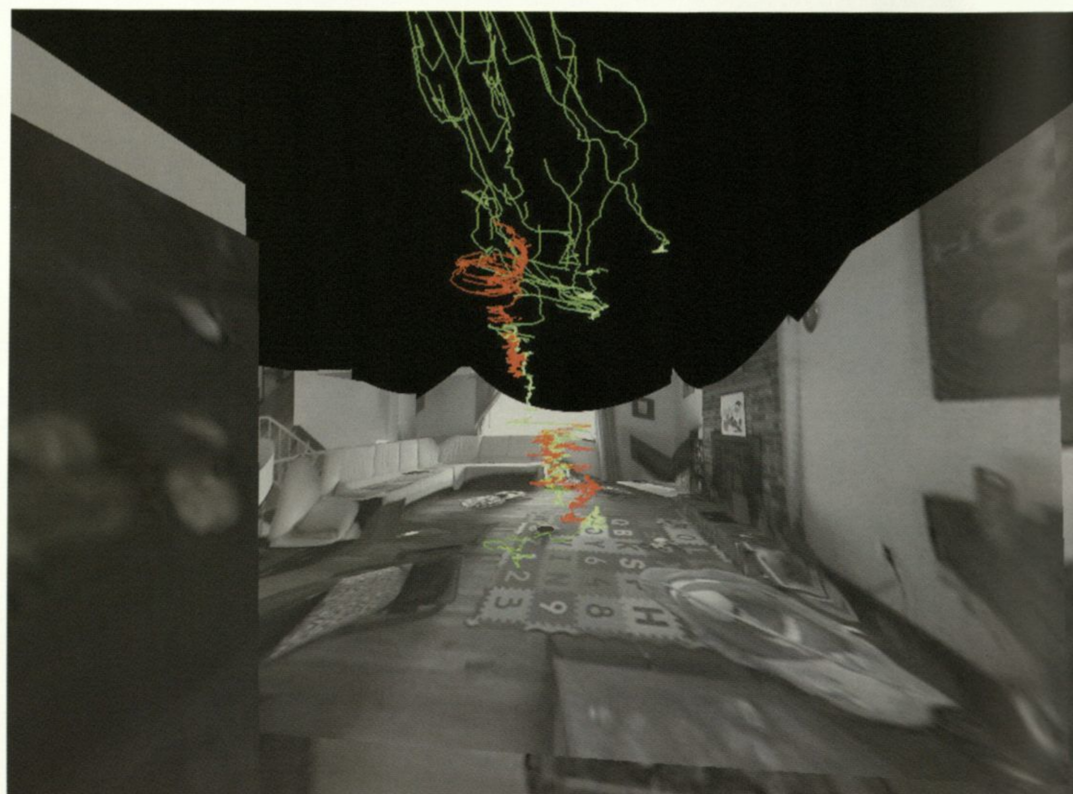
CHILD DEVELOPMENT
HouseFly and *WordScape*

In order to study child development as it occurs in the home, professors Deb Roy and Rupal Patel began an investigation in their own family with the birth of their first child. They installed a camera and microphone in the ceiling of every room of their house and recorded the majority of their child's waking experience for the first three years of life, resulting in a dataset of 80,000 hours of video and 120,000 hours of audio. HouseFly is a software tool developed to help researchers visualize and browse this massive dataset. Between 2009 and 2010 Philip DeCamp developed the application in collaboration with Deb Roy, director of the Cognitive Machines group at the MIT Media Lab.[24]

Instead of displaying each stream of video separately, HouseFly combines them to create a dynamic, three-dimensional model of the home. The user can navigate to any location in the house at any time and get a better sense of what they would have seen and heard if they had actually been there. Beyond the reconstruction of individual events, HouseFly also incorporates speech transcripts, person tracks, and other forms of retrieving and accessing data in an effort to uncover some of the unseen patterns of everyday life.

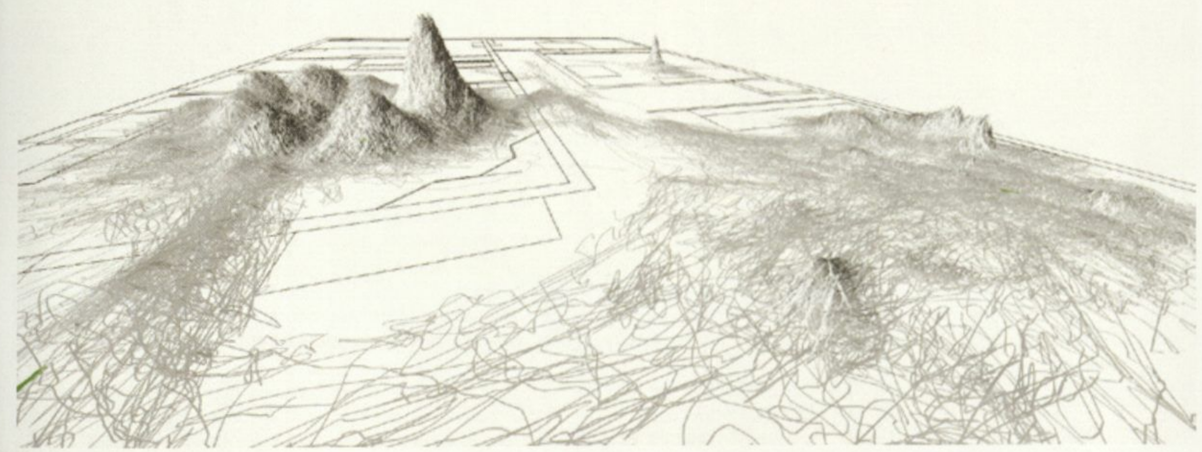
What we see in this image is the 3-D synthesized home environment constructed from 11-camera video. HouseFly uses immersive video as a platform for multimodal data visualization. The application allows one to move in space and through time to examine the 80,000 hours of video. At the bottom, the timeline offers another way to navigate the content, including the ability to add notations in time about words of interest in the transcripts of the speech environment of the child.

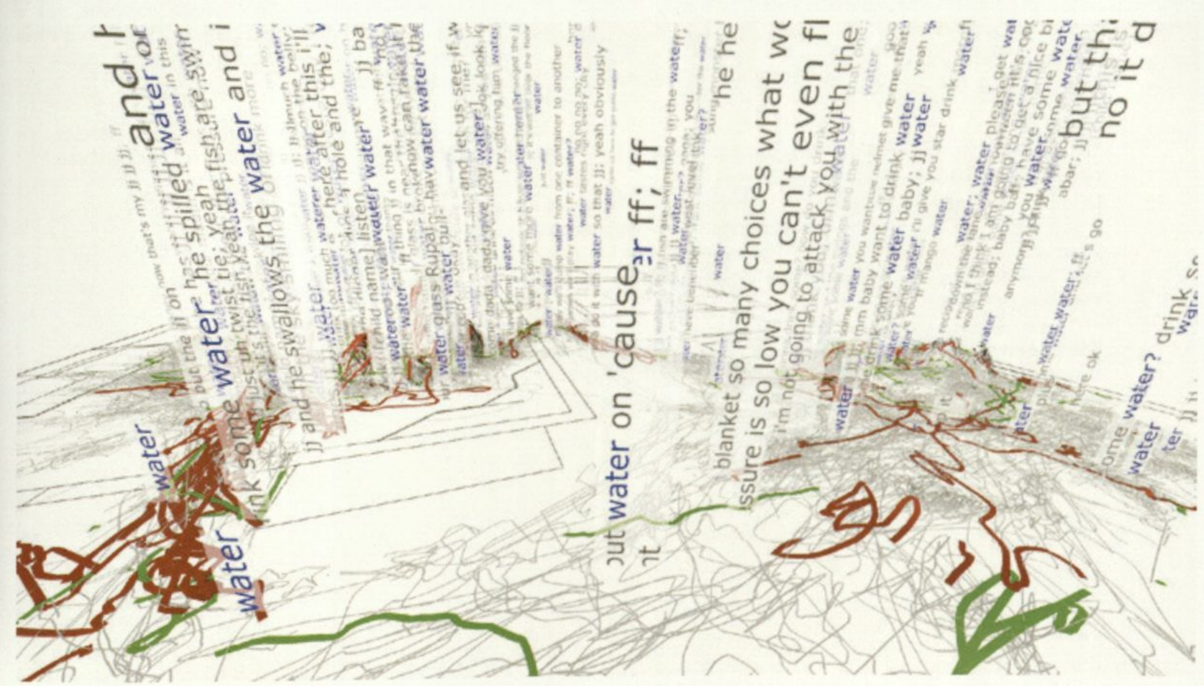
Twenty minutes of motion by the child (red) and the caregiver (green) are represented as traces rendered in space. To examine the temporal dimension of the motion, one can switch the view to the side and the traces will be ordered vertically, with earlier times at the bottom, allowing a chronological view of interactions (bottom).



WordScapes are generated by mining the audio data for all utterances of a given word, like "water," tracking the locations of the occupants for twenty seconds around each utterance, and then stacking the resulting tracks like a pile of noodles. The resulting landscape reveals the overall distribution of activity associated with a given word. Some words, like "book," are used most frequently in the child's bedroom, where caregivers often read to the child, while words like "mango" occur almost exclusively in the kitchen. Such analysis may provide insight into how and why different children learn different words more readily than others.

"WATER"





CASE STUDY

MOBILITY

From Mobility Data to Mobility Patterns

Huge amounts of data generated and collected by a wealth of technological infrastructures, such as GPS positioning, and wireless networks have affected research on moving-object data analysis. Access to massive repositories of spatiotemporal data with recorded human mobile activities have opened new frontiers for developing suitable analytical methods and location-aware applications capable of producing useful knowledge.

This case study briefly introduces few visual techniques devised by an interdisciplinary team involved with mobility data mining, knowledge discovery, and visual analytical tools. The project was part of the European Community–funded effort on Geographic Privacy-aware Knowledge Discovery and Delivery–GeoPKDD, with the objective to investigate "how to discover useful knowledge about human movement behavior from mobility data, while preserving the privacy of the people under observation. GeoPKDD aims at improving decision-making in many mobility-related tasks, especially in metropolitan areas."[25]

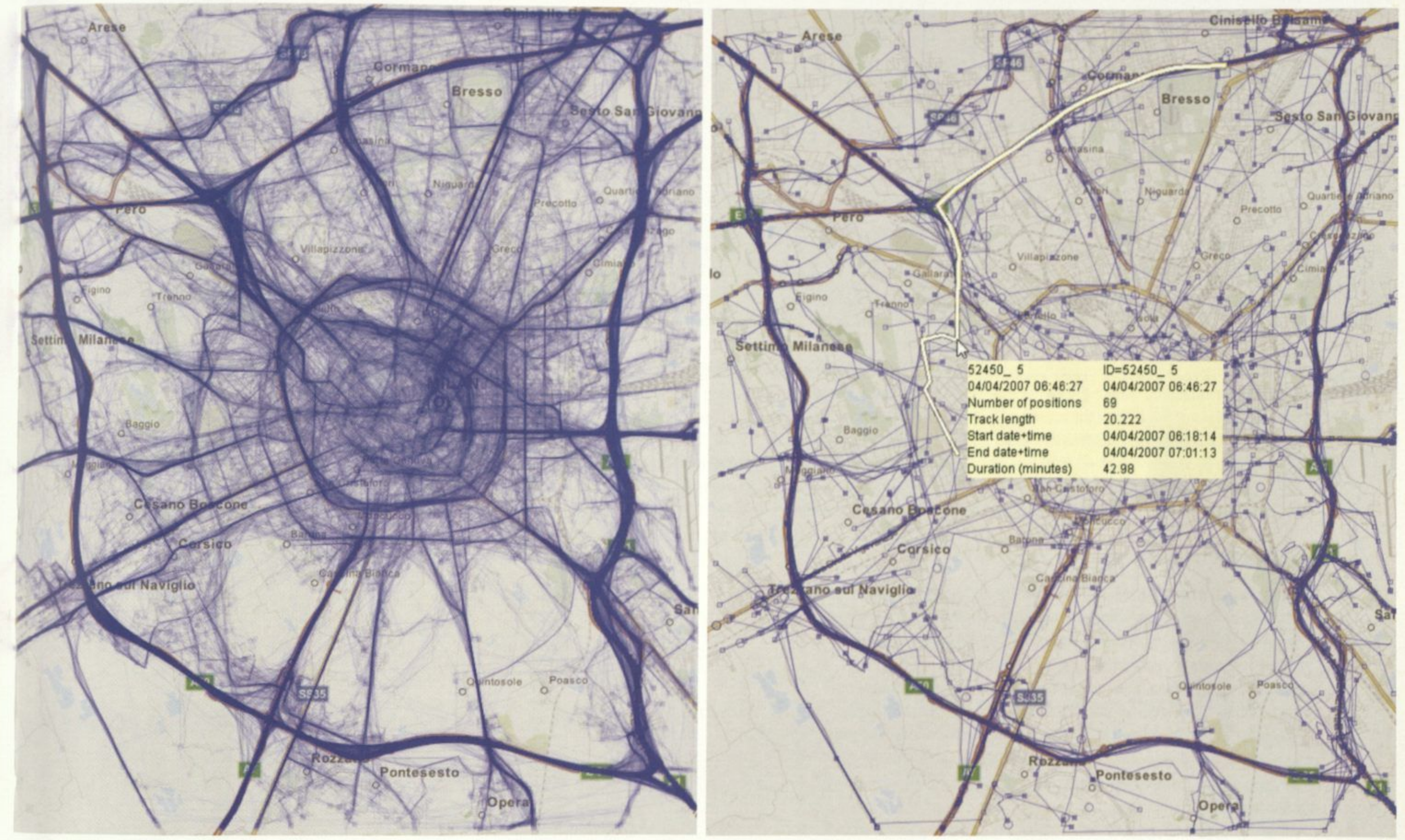
The main people involved in this particular output are Gennady Andrienko, Natalia Andrienko, Fosca Giannotti, Dino Pedreschi, and Salvatore Rinzivillo.[26] What we see is a small sample of their extensive and pioneer work in the visual analyses of movement data. I strongly recommend their writings, which include discussion of computational methods, not examined here.[27]

The dataset consists of GPS tracks of 17,241 cars collected during one week in Milan, Italy, which resulted in 2,075,216 position records. The work was conducted mostly between 2005 and 2009 with continued ongoing efforts.

Natalia and Gennady Andrienko organize the methods for visually analyzing movement data into four types:[28]

- **Looking at trajectories:** Trajectories are considered as wholes. The focus is on examination of spatial and temporal properties of individual trajectories as well as comparison among trajectories.
- **Looking inside trajectories:** Trajectories are considered at the level of segments and points. The focus is on examination of segment's movement characteristics and the sequences of segments with shared patterns.
- **Bird's-eye view on movement:** Trajectories are viewed as aggregations, not individually. The focus is on examination of the distribution of multiple movements in space and time.
- **Investigating movement in context:** Movement data are examined with other kinds of spatial, temporal, and spatiotemporal data describing context. The focus is on relations of interactions between the moving objects and the environment.

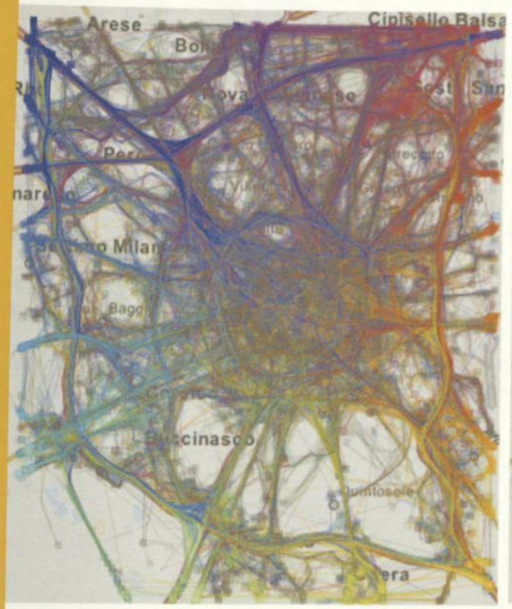
Each series of images illustrates a method type with the exception of movement in context, not reproduced here.

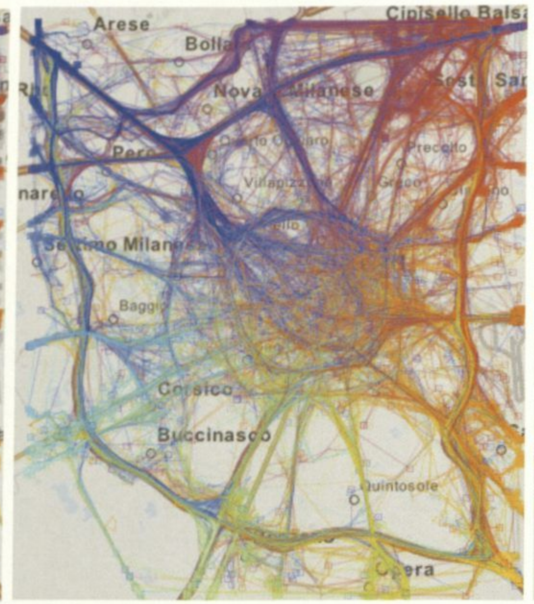


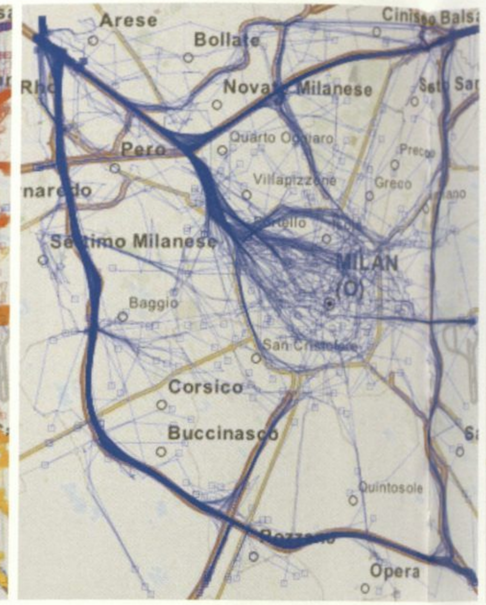
VISUALIZING TRAJECTORIES

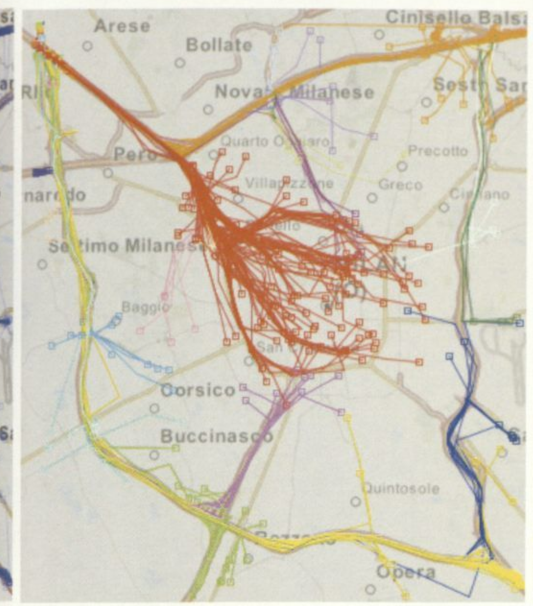
This image shows a subset of the Milan dataset consisting of 8,206 trajectories that began on Wednesday, April 4, 2007. To make the map legible, the trajectory lines are drawn with only 5 percent opacity.

The visual analytical tool allows one to interactively manipulate the view as well as apply filters. The image on the right shows the result of using a temporal filter that limits the representation of trajectories within a 30-minute time interval, from 06:30 to 07:00. The same function can be used to generate map animations. The screenshot illustrates that by interacting with the trajectories one can read detailed information about its attributes, such as start and end time, number of positions, length, duration, etc.









The image shows the result of clustering by "common destinations," which compares the spatial positions of the ends of trajectories. From the 8,206 trajectories, 4,385 have been grouped into 80 density-based clusters and 3,821 treated as noise.

In this image, we see the clusters with the noise removed.

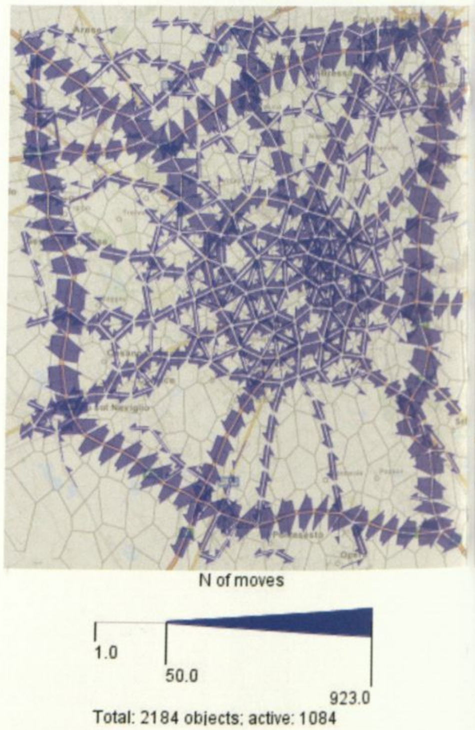
The image shows the biggest cluster, which consists of 590 trajectories that end at the northwest part of Milan.

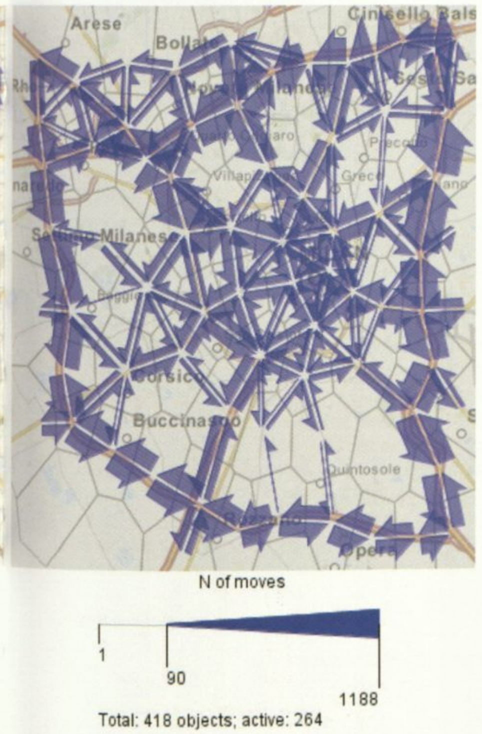
When clustering by "route similarity," which compares the routes followed by the moving objects, the result is a total of eighteen clusters, with the noise hidden. The largest cluster (in red) consists of 116 trajectories going from the city center. The next largest cluster (in orange) consists of 104 trajectories going from the northeast along the northern motorway. The yellow cluster (68 trajectories) depicts trajectories going from the southeast along the motorway on the south and west.

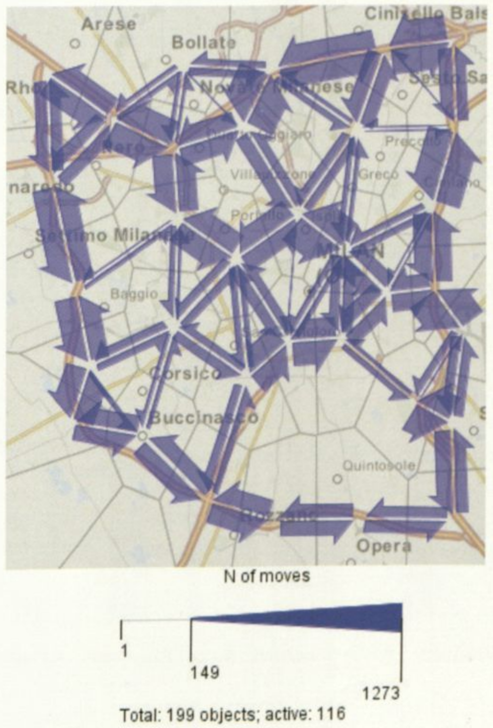
The image shows the Space-Time Cube (STC) representation of the result from clustering by "route similarity" (same clustering as shown in the previous image). STC is a common type of display of movement data that uses a three-dimensional cube, with two dimensions representing space, and one time. STCs were briefly discussed earlier in the chapter (see pages 161–162).

CLUSTERING TRAJECTORIES

Natalia and Gennady Andrienko explain, "Trajectories of moving objects are quite complex spatiotemporal constructs. Their potentially relevant characteristics include the geometric shape of the path, its position in space, the life span, and the dynamics, i.e. the way in which the spatial location, speed, direction and other point-related attributes of the movement change over time. Clustering of trajectories requires appropriate distance (dissimilarity) functions which can properly deal with these non-trivial properties."[29] To avoid universal functions that would make the visualization hard to interpret, the team has developed a method called "progressive clustering."[30] It is a step-by-step process in which the analyst progressively refines the clustering by modifying the parameters and applying the new settings, thus gradually building understanding of the different aspects of the trajectories. The four images show the result of progressive clustering to the same subset of the Milan data as the images on the previous page.







BIRD'S-EYE VIEW OF MOVEMENT DATA

Generalization and aggregation of trajectories enable understanding of the spatial and temporal distribution of multiple movements, which is not possible by looking at individual trajectories. There are different techniques for aggregating movement data, and the most common method examines flows of moving objects by pairs of locations, as those in origin-destination pairs. Given the complexity of the data, and to avoid visual clutter, Andrienko and colleagues have devised a more efficient method that segments trajectories into all visited locations along the path and then aggregate the transitions from all trajectories.[31] The result can be viewed in this sequence of images showing flow maps based on fine, medium and coarse territory divisions. To distinguish flows in different directions, each segment is represented by "half-arrow" symbols. The line widths stand for magnitudes. Details on exact value of magnitudes, as well as other flow-related attributes, are provided by interaction with the segments.