Time Series Data

CS 7450 - Information Visualization
October 26, 2015
John Stasko
Presented by Yi Han

Time Series Data

- Fundamental chronological component to the data set

75 % of 4000 samples of graphics from newspapers and magazines ('74-'80) were time-series data!

Tufte Vol. 1

Data Sets

- Each data case is likely an event of some kind
- One of the variables can be the date and time of the event
- Examples:

  sunspot activity
  baseball games
  medicines taken
  cities visited
  stock prices

Meta Level

- Consider multiple stocks being examined
- Is each stock a data case, or is a price on a particular day a case with the stock name as one of the other variables?

- Confusion between data entity and data cases

Data Mining

- Data mining domain has techniques for algorithmically examining time series data, looking for patterns, etc.

- Good when objective is known a priori
- But what if not?
  - Which questions should I be asking?
  - InfoVis better for that

Tasks

- What kinds of questions do people ask about time series data?

Time Series User Tasks

- Examples
  - When was something greatest/least?
  - Is there a pattern?
  - Are two series similar?
  - Do any of the series match a pattern?
  - Provide simpler, faster access to the series

Other Tasks

- Does data element exist at time *t*?
- When does a data element exist?
- How long does a data element exist?
- How often does a data element occur?
- How fast are data elements changing?
- In what order do data elements appear?
- Do data elements exist together?

Muller & Schumann '03 citing MacEachern '95

Taxonomy

- Discrete points vs. interval points
- Linear time vs. cyclic time
- Ordinal time vs. continuous time
- Ordered time vs. branching time vs. time with multiple perspectives

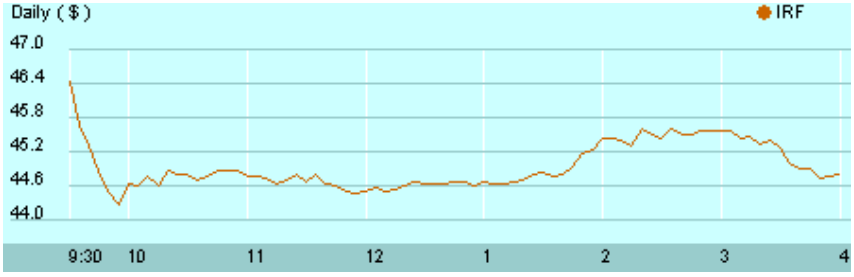
Muller & Schumann '03 citing Frank '98

Fundamental Tradeoff

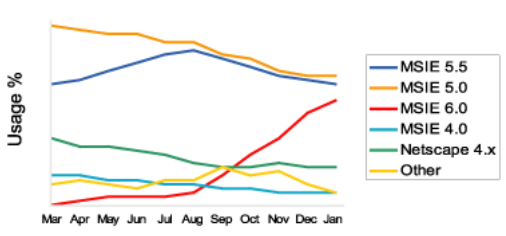
- Is the visualization time-dependent, ie, changing over time (beyond just being interactive)?
  - Static
    - Shows history, multiple perspectives, allows comparison
  - Dynamic (animation)
    - Gives feel for process & changes over time, has more space to work with

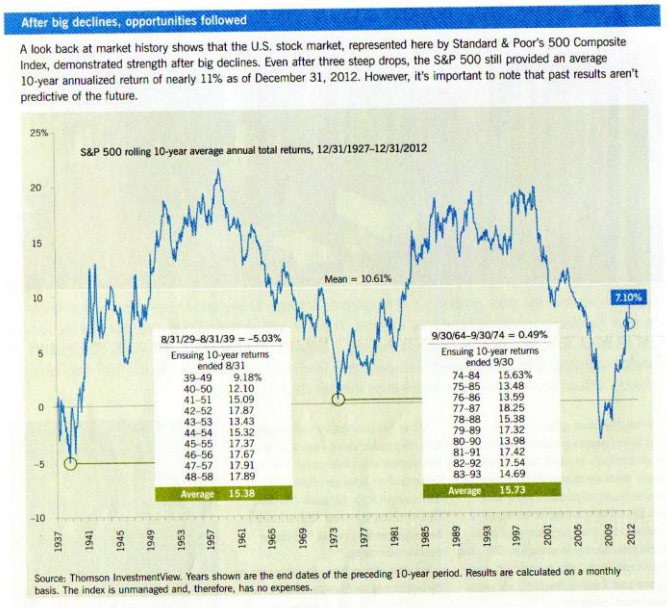
Standard Presentation

- Present time data as a 2D line graph with time on x-axis and some other variable on y-axis



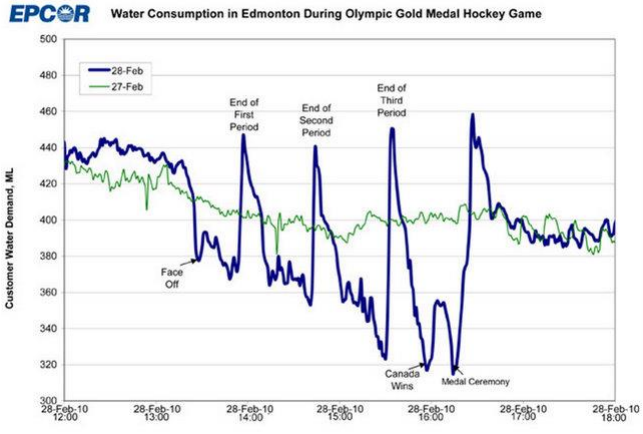
Classic Views





Fun One

What If Everybody in Canada Flushed At Once?



http://www.patspapers.com/blog/item/what_if_everybody_flushed_at_once_Edmonton_water_gold_medal_hockey_game/

Today's Focus

- Examination of a number of case studies
- Learn from some of the different visualization ideas that have been created

- Can you generalize these techniques into classes or categories?

Nice overview

IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 14, NO. 1, JANUARY/FEBRUARY 2008 47

Visual Methods for Analyzing Time-Oriented Data

Wolfgang Aigner, Silvia Miksch, Wolfgang Müller, Heidrun Schumann, and Christian Tominski

Abstract—Providing appropriate methods to facilitate the analysis of time-oriented data is a key issue in many application domains. In this paper, we focus on the unique role of the parameter time in the context of visually driven data analysis. We will discuss three major aspects—visualization, analysis, and the user. It will be illustrated that it is necessary to consider the characteristics of time when generating visual representations. For that purpose, we take a look at different types of time and present visual examples. Integrating visual and analytical methods has become an increasingly important issue. Therefore, we present our experiences in temporal data abstraction, principal component analysis, and clustering of larger volumes of time-oriented data. The third main aspect we discuss is supporting user-centered visual analysis. We describe event-based visualization as a promising means to adapt the visualization pipeline to needs and tasks of users.

Index Terms—Time-oriented data, visualization, analysis, user.

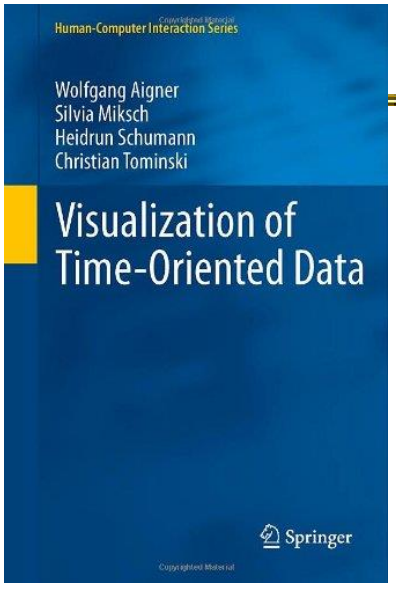
1 Introduction and Motivation

Considering the characteristics of data is vital when designing visual representations. A salient characteristic is whether or not data are related to time. That time is an outstanding dimension is reflected by Shneiderman's Task by Data Type Taxonomy [1], where temporal data are identified as one of seven basic data types. Nowadays, time-oriented data are ubiquitous in many application domains as, for example, in business, medicine, history, planning, or project management. For a long time, visual methods have been successfully applied to analyze such data. A wide repertoire of interactive techniques for visualizing data sets with temporal dependencies is available. However, many current

- visualization,
- analysis, and
- user.

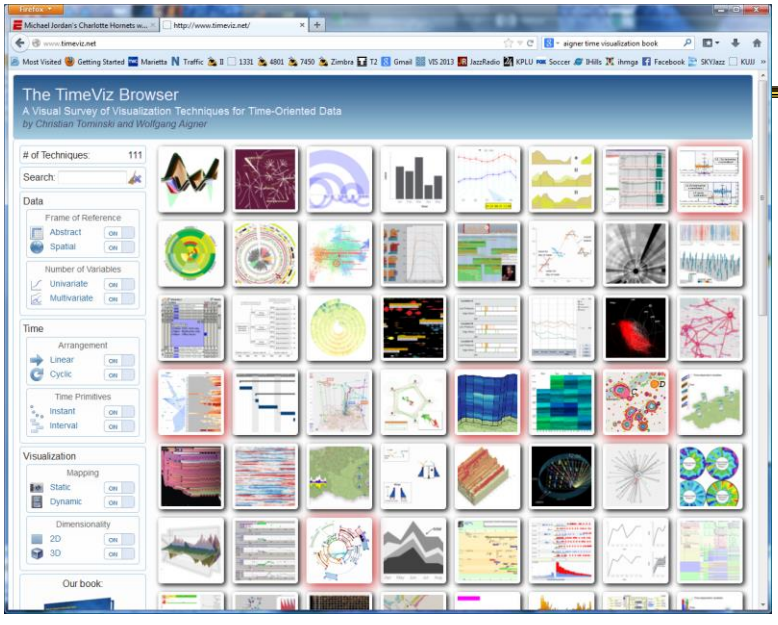
In Section 2, we focus on visualization methods for time-oriented data. We will show that the term *time-oriented data* comprises several types of data with different meanings and applications. Designing or applying visual representations can only be successful if one is aware of these different types. This will be demonstrated with several examples of visualization techniques that stem from our own work or are available in the literature.

Usually, time-oriented data are large—not only in terms of the number of data items but also in terms of the number



Bigger overview

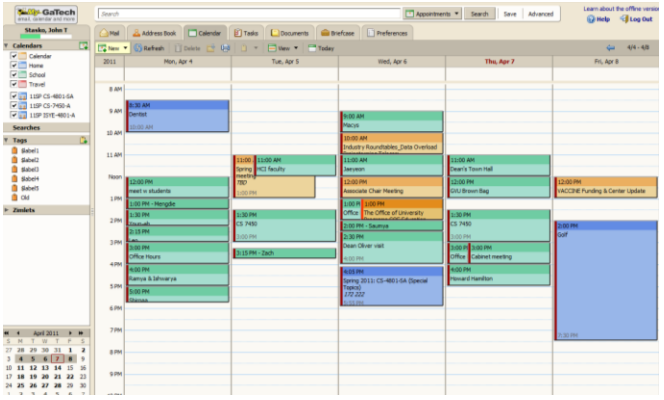
http://www.timeviz.net/



Interactive survey

Example 1

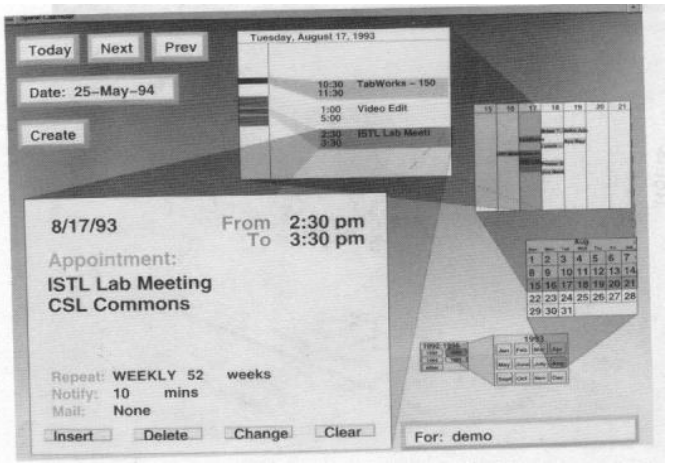
- Calendar visualization



Tasks

- See commonly available times for group of people
- Show both details and broader context

One Solution

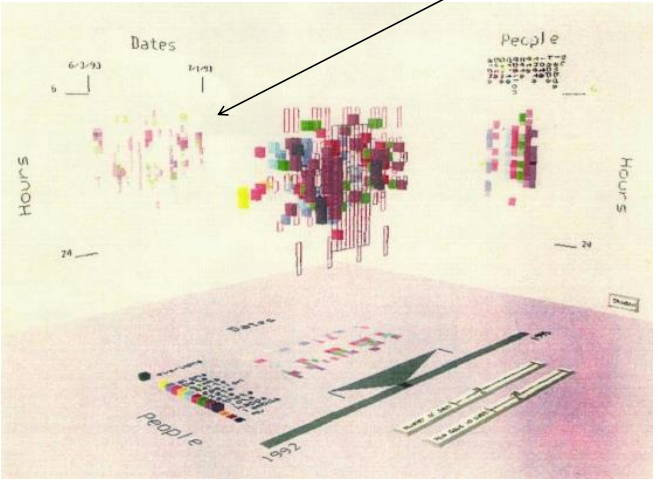


Spiral Calendar

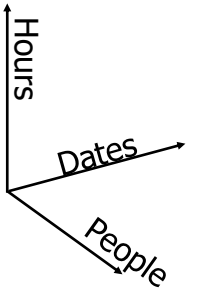
Mackinlay, Robertson & DeLine
UIST '94

Another View

Empty spots on back wall show good times



Time Lattice



Uses projected shadows on walls

Example 2

- Personal histories
  - Consider a chronological series of events in someone's life
  - Present an overview of the events
  - Examples
    - Medical history
    - Educational background
    - Criminal history

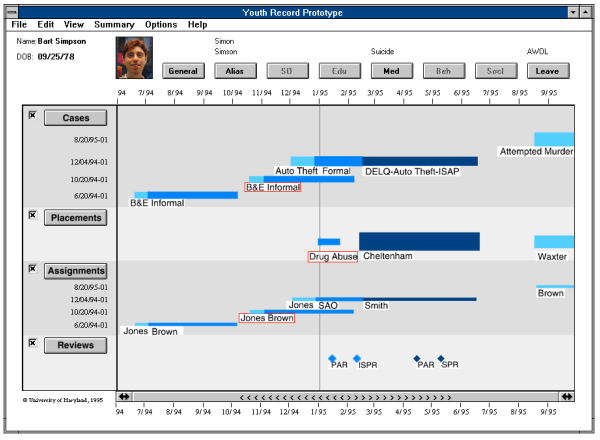
Tasks

- Put together complete story
- Garner information for decision-making
- Notice trends
- Gain an overview of the events to grasp the big picture

Lifelines Project

Visualize personal history in some domain

Video
Demo



Plaisant et al
CHI '96

Medical Display

Features

- Different colors for different event types
- Line thickness can correspond to another variable
- Interaction: Clicking on an event produces more details
- Certainly could also incorporate some Spotfire-like dynamic query capabilities

Benefits

- Reduce chances of missing information
- Facilitate spotting trends or anomalies
- Streamline access to details
- Remain simple and tailorable to various applications

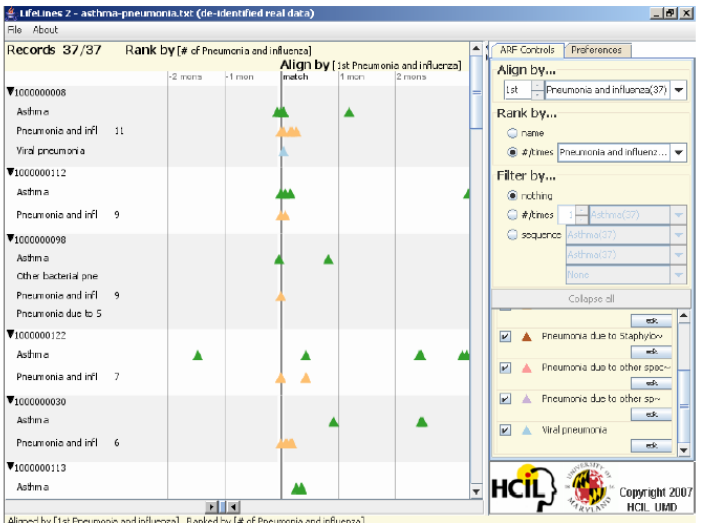
Challenges

- Scalability (thousands of tests)
- Can multiple records be visualized in parallel (well)? Comparisons
  - What trends do you see in the last 8 EKGs?
  - Compare the 8 people who all seem to have the same problem.

New Work

- Work with query results
- Need to align, rank, and filter
- Medical application:
  - Look for temporal coincidence of two events
    First pneuomonia and asthma attack
  - Medical professionals don't want to fool with zooming and panning

LifeLines2: Focus on alignment along events



Video

Wang et al
CHI '08

Example 3

van Wijk & van Selow
InfoVis '99

- Understand patterns of presence/events over time
- Focus: People's presence/movements in some space
- Situation:
  - Workers punch in and punch out of a factory
  - Want to understand the presence patterns over a calendar year
- Alternate: Power plant electricity usage over a year

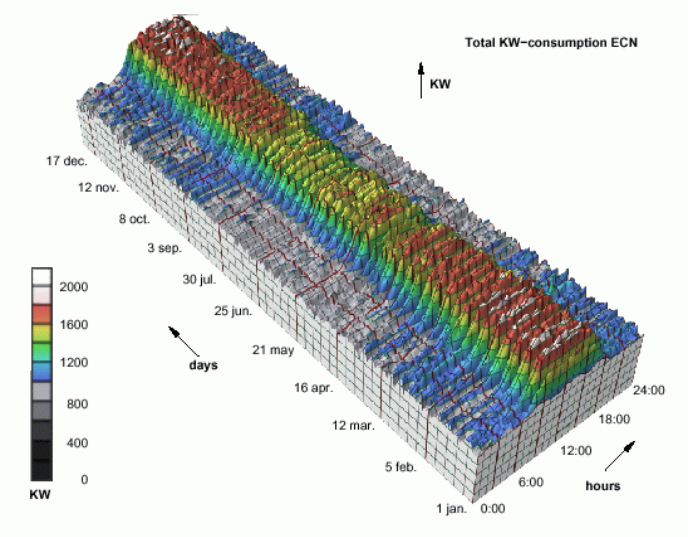
Particulars

- Who is user? – Factory boss/manager
- Problem – Show this large amount of data in an easily understandable and query-able manner
- Data – Punch in/out times for workers

Ideas

- Any ideas on what we could do here?

One Idea



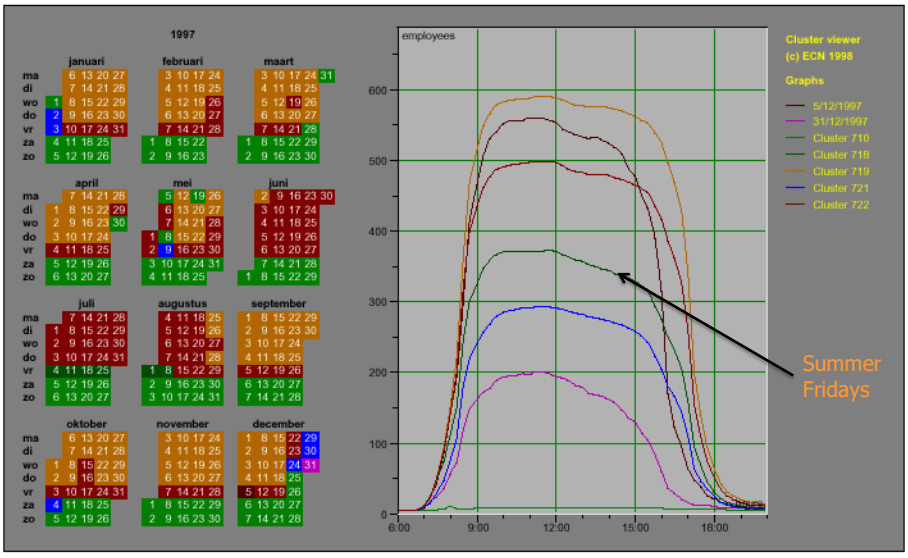
Good
Typical daily pattern
Seasonal trends

Bad
Weekly pattern
Details

Approach Taken

- Cluster analysis
  - Find two most similar days, make into one new composite
  - Keep repeating until some preset number left or some condition met
- How can this be visualized?
  - Ideas?

Display



Characteristics

- Unique types of days (individual or cluster) get their own color
- Contextually placed in calendar and line graph for it is shown
- Stop clustering when a threshold met or at a predetermined number of clusters

Interaction

- Click on day, see its graph
- Select a day, see similar ones
- Add/remove clusters

Insights

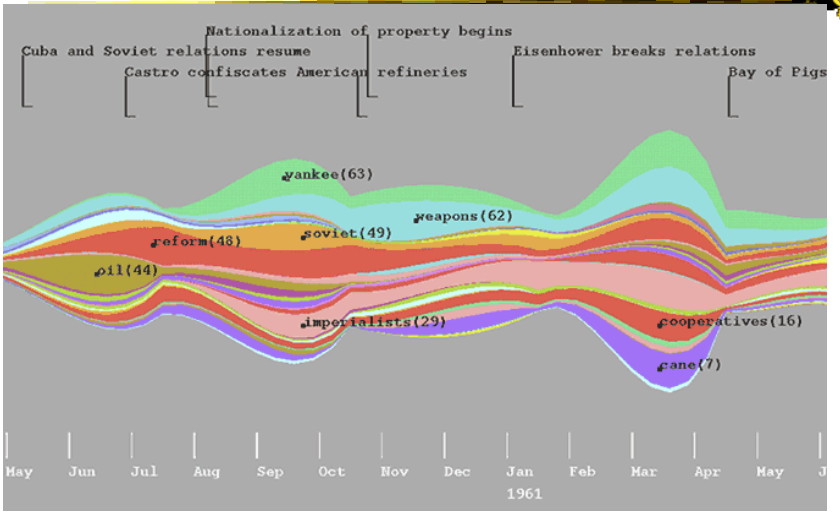
- Traditional office hours followed
- Most employees present in late morning
- Fewer people are present on summer Fridays
- Just a few people work holidays
- When the holidays occurred
- School vacations occurred May 3-11, Oct 11-19, Dec 21-31
- Many people take off day after holiday
- Many people leave at 4pm on December 5
  - Special day in Netherlands, St. Nicholas' Eve

Example 4

- Consider a set of speeches or documents over time
- Can you represent the flow of ideas and concepts in such a collection?

ThemeRiver

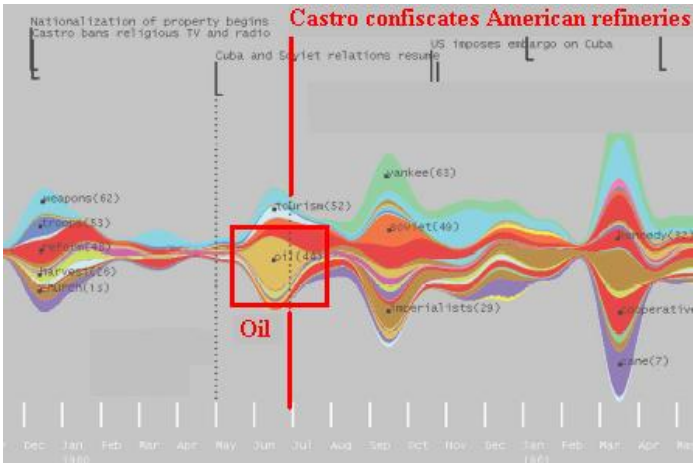
Havre et al
InfoVis '00



We saw earlier

Mapping

- River height (thickness) encodes relative frequency of themes
- Key events overlaid

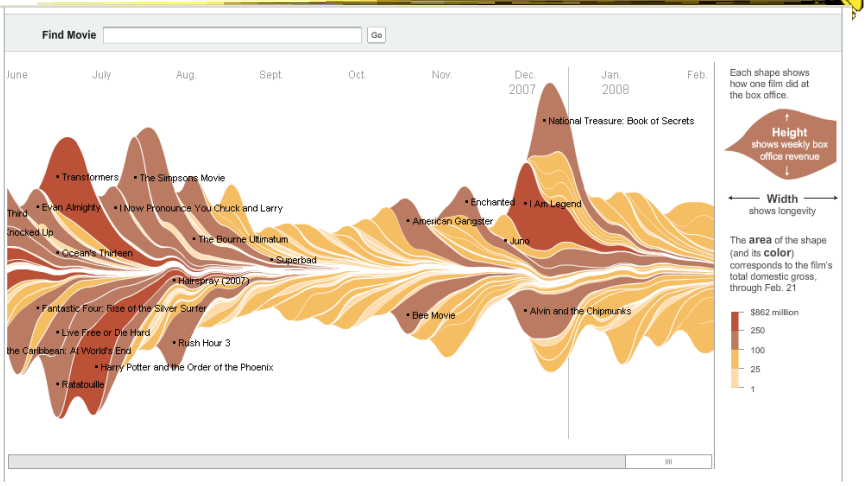


Example 5

Byron & Wattenberg
TVCG '08

- Similar idea – Stacked graph
- Created new technique called Streamgraph
- Goals:
  - Show multiple time series
  - Be able to see sum
  - Make labels legible
  - Be able to distinguish different layers
  - Make it aesthetically pleasing

Interactive Application



http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html

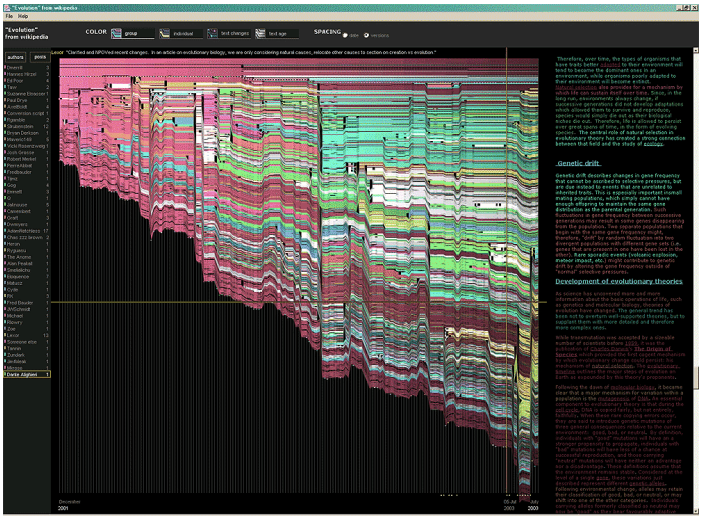
Design Issues

- Curve shape
  - Wiggle, symmetry, balance
  - Definitely some interesting math to do it
- Color choice
- Labeling
- Layer ordering

- Paper provides very nice discussion of this

Example 6

http://researchweb.watson.ibm.com/history/

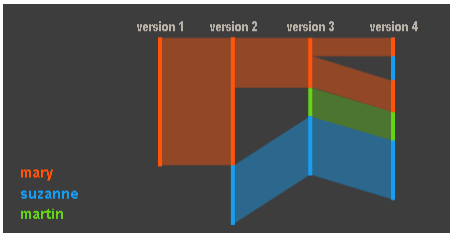


Flow of changes across electronic documents

Technique

Length – how much text

Time



Make connections

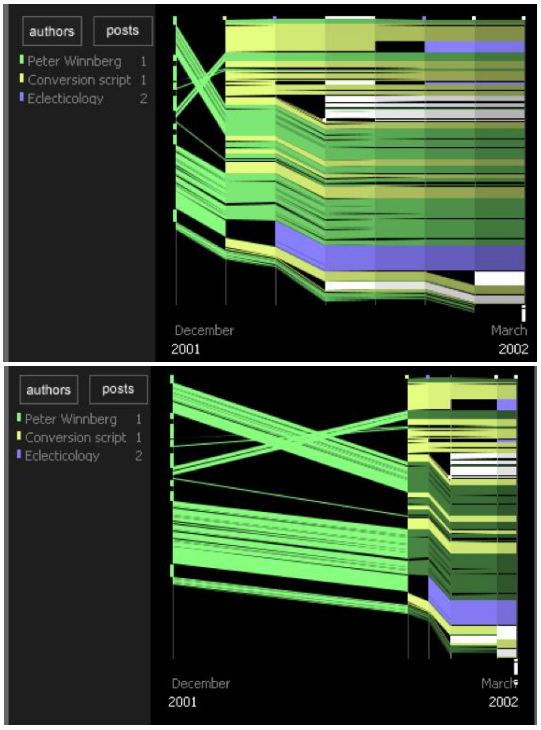
Brightness indicates text age

Registered authors color-coded

Anonymous authors in white

Spacing by revision #

Spacing by time



Example 7

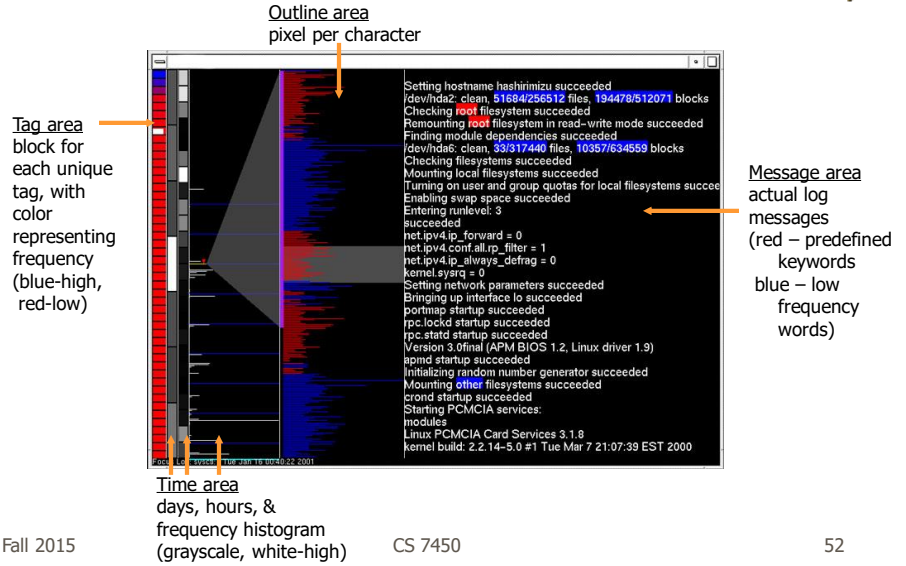
- Computer system logs
- Potentially huge amount of data
  - Tedious to examine the text
- Looking for unusual circumstances, patterns, etc.

MieLog

- System to help computer systems administrators examine log files
- Interesting characteristics
  - Discuss

Takada & Koike
LISA '02

System View



Another View

Alternate color mappings?

Interactions

- Tag area
  - Click on tag shows only those messages
- Time area
  - Click on tiles to show those times
  - Can put line on histogram to filter on values above/below
- Outline area
  - Can filter based on message length
  - Just highlight messages to show them in text
- Message area
  - Can filter on specific words

Thoughts

- Strengths/weaknesses?
- Other domains in which a similar system could be used?

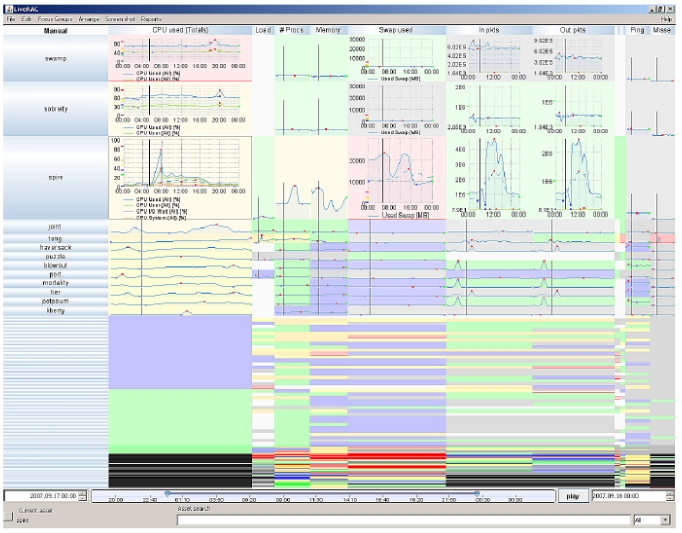
Example 8

- Very large scale temporal log data
- Show more context of what else was going on at that time
  - Likely have to abstract some then
  - Allow several different levels of detail at once
- Allow drill-down for details

- Domain: Computer systems management

McLachlan et al
CHI '08

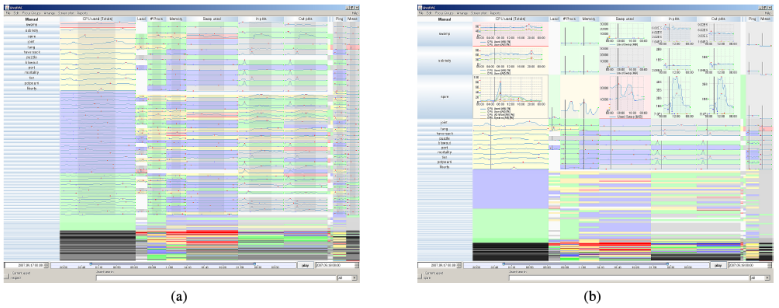
LiveRAC: Computer system management data



Heavy interaction
Semantic zooming

Interaction is Vital

Video



(a) (b)

Figure 3. LiveRAC shows a full day of system management time-series data using a reorderable matrix of area-aware charts. Over 4000 devices are shown in rows, with 11 columns representing groups of monitored parameters. (a): The user has sorted by the maximum value in the *CPU* column. The first several dozen rows have been stretched to show sparklines for the devices, with the top 13 enlarged enough to display text labels. The time period of business hours has been selected, showing the increase in the *In pkts* parameter for many devices. (b): The top three rows have been further enlarged to show fully detailed charts in the *CPU* column and partially detailed ones in *Swap* and two other columns. The time marker (vertical black line on each chart) indicates the start of anomalous activity in several of *spire*'s parameters. Below the labeled rows, we see many blocks at the lowest semantic zoom level, and further below we see a compressed region of highly saturated blocks that aggregate information from many charts.

Design Principles

- Show familiar visual representations whenever possible
- Provide side-by-side comparisons of small multiple views
- Spatial position is strongest visual cue
- Multiple views are more effective when coordinated through explicit linking

Design Principles

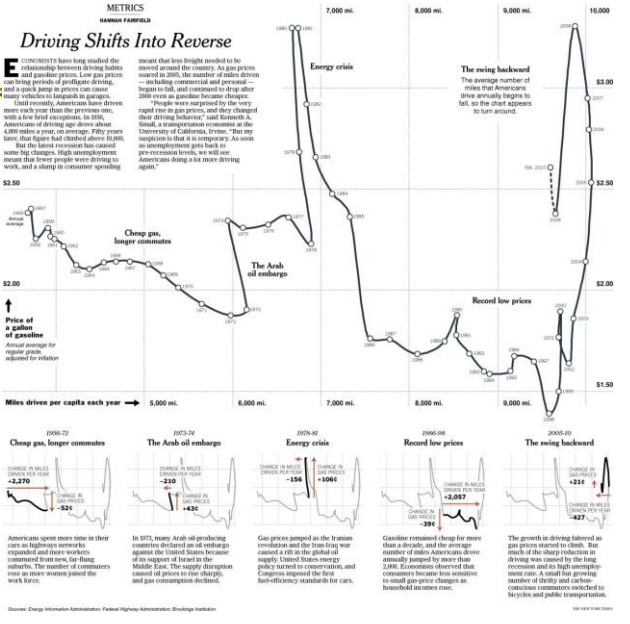
- Follow Shneiderman's mantra
- Avoid abrupt visual change
- User actions should receive immediate visual feedback

- Assertion: Showing several levels of detail simultaneously provides useful high information density in context

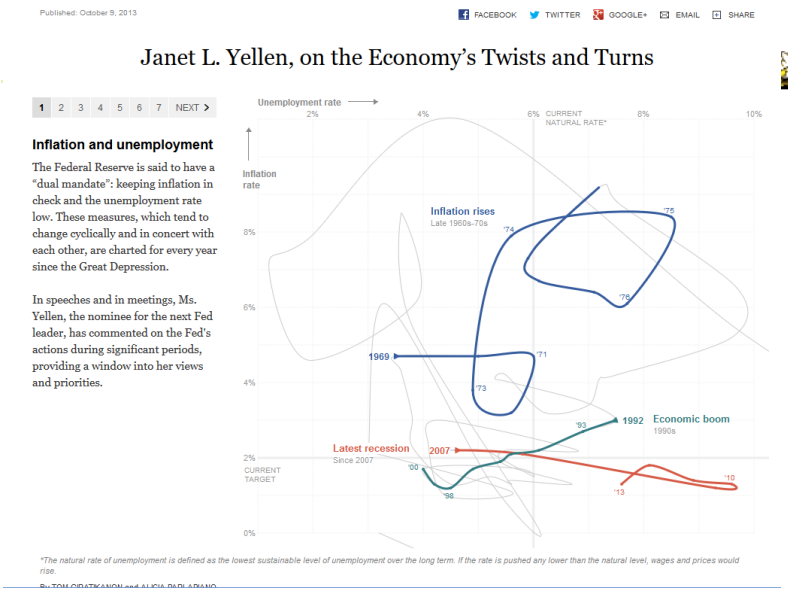
Example 9

- Connected Scatterplot
- Showing two variables over time
  - Use standard scatterplot
  - Plot the two values at different points in time
  - Connect those points, in order, with a line
  - Label key times (e.g., years)

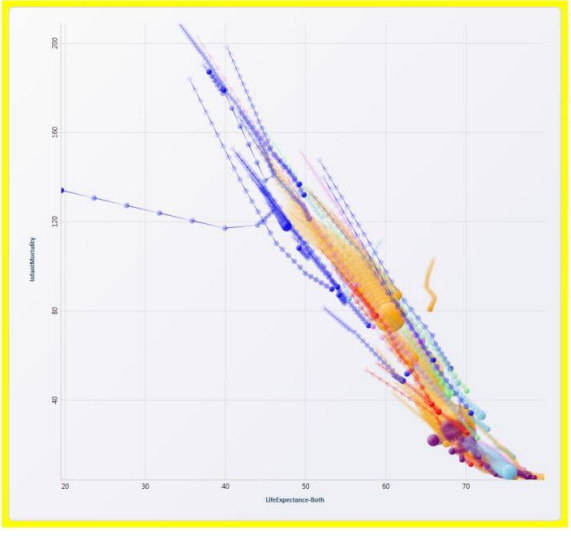
Hannah Fairfield
NY Times



Notice the narrative elements too



http://www.nytimes.com/interactive/2013/10/09/us/yellen-fed-chart.html

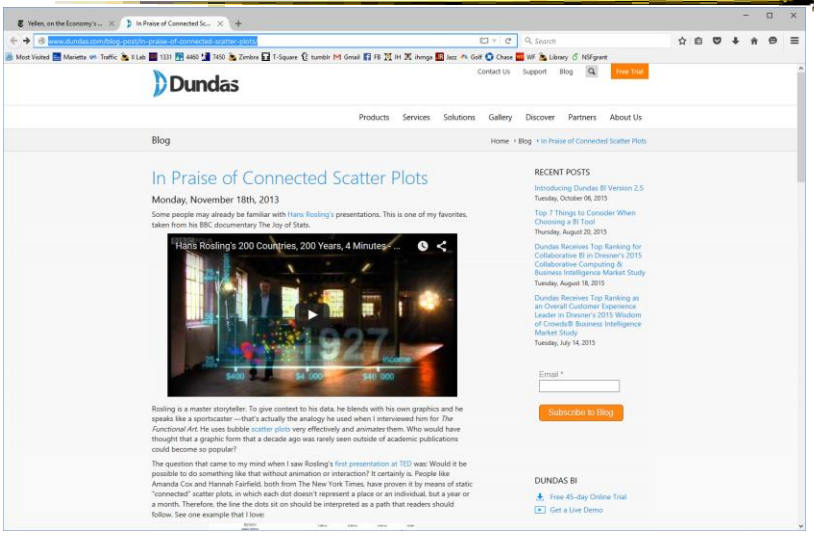


"Traces" in Gapminder-style visualization

Robertson et al *TVCG* (InfoVis) '08

http://www.dundas.com/blog-post/in-praise-of-connected-scatter-plots/

Nice Article



Example 10

- Serial, periodic data
- Data with chronological aspect, but repeats and follows a pattern over time
  - Hinted at in last case study
- How might one visualize that?

Using Spirals

- Standard x-y timeline or tabular display is problematic for periodic data
  - It has endpoints
- Use spiral to help display data
  - One loop corresponds to one period

Carlis & Konstan
UIST '98

Basic Spiral Display

One year per loop
Same month on radial bars
Quantity represented by size of blob

Is it as easy to see serial data as periodic data?

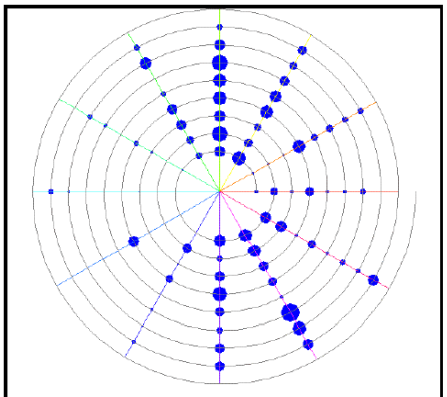


Figure 2. An indented spiral, with spokes, showing monthly consumption percentages for Baphia Capparidifolia during the period 1980 – 1988.

Advanced Spiral

Same mapping as previous one

Different foods represented by different colors and drawn at different heights

Can you still see serial and periodic attributes?

As with all 3-D, requires navigation

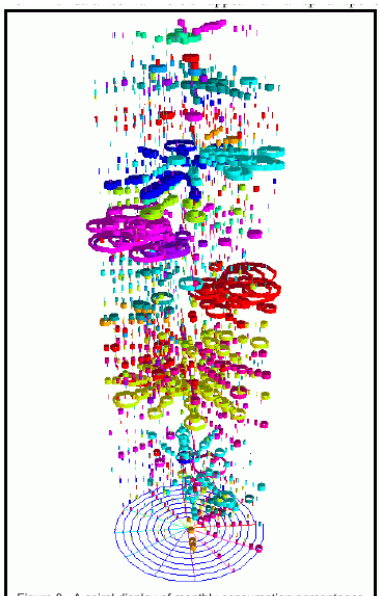


Figure 3. A spiral display of monthly consumption percentages for all 112 foods during the period 1980 – 1988.

Compare with Spotfire

Another standard spiral display

Color mapped to movie type

+/- compared to Spotfire?

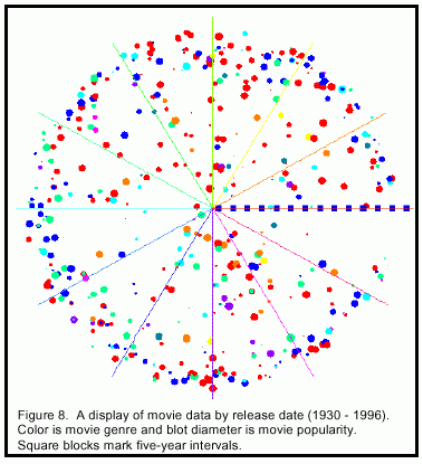
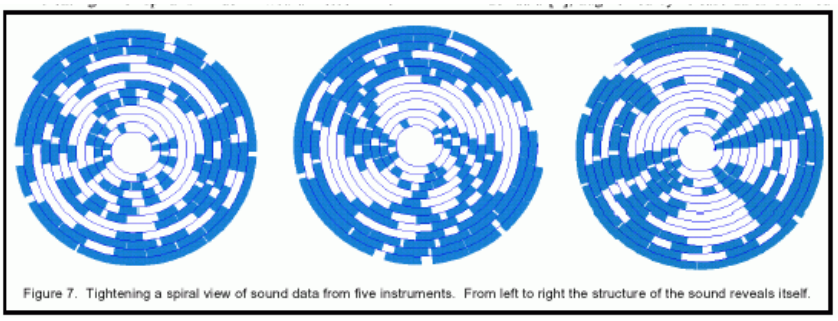


Figure 8. A display of movie data by release date (1930 - 1996). Color is movie genre and blot diameter is movie popularity. Square blocks mark five-year intervals.

Unknown Periods



What if a data set doesn't have a regular temporal period?
Must do some juggling to align periods

Example 11

- How about events in time and place?
  - Many applications of this problem

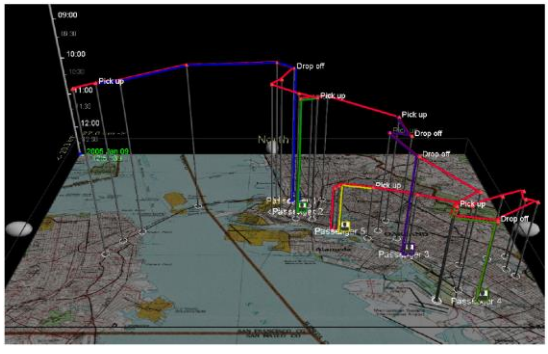
GeoTime

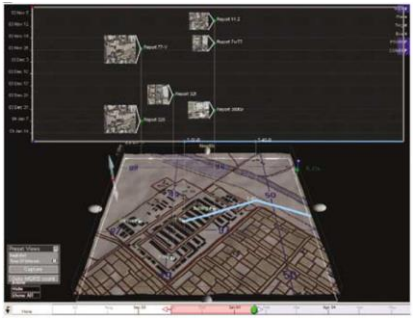
- Represent place by 2D plane (or maybe 3D topography)
- Use 3rd dimension to encode time
- Object types:
  - Entities (people or things)
  - Locations (geospatial or conceptual)
  - Events (occurrences or discovered facts)

Kapler & Wright
InfoVis '04

Overview

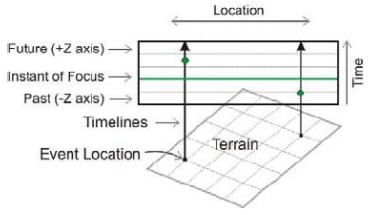
- Objective: visualize spatial interconnectedness of information over time and geography with interactive 3-D view

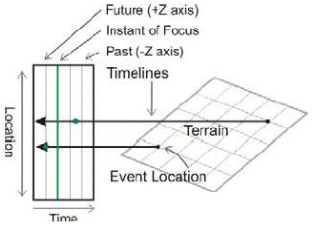




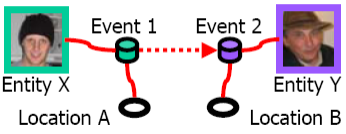
Source: http://www.oculusinfo.com/

Design Characteristics

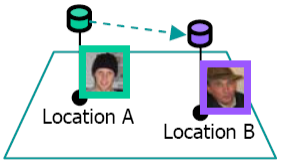




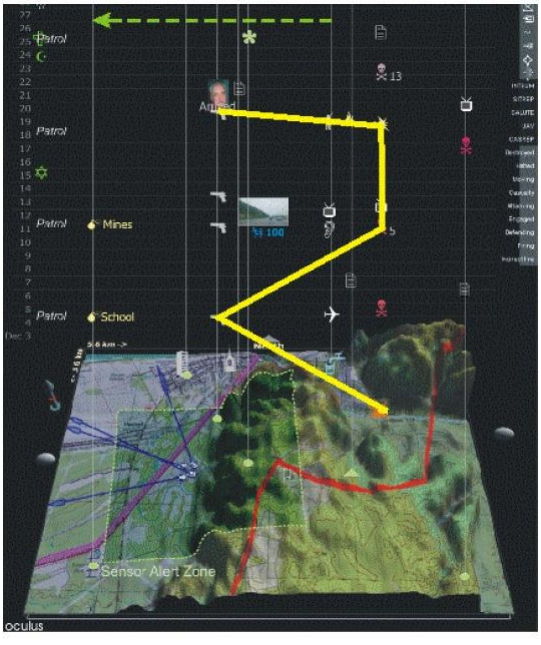
Dimension usage



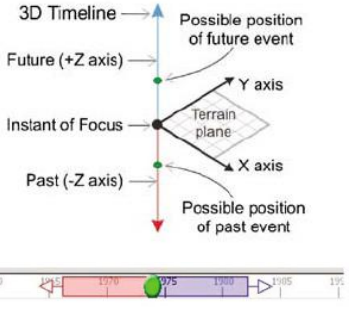
Vector Group with 1+ Actors
(Phone Call, email, money transfer)
5 Associations: Same as above plus…
- *Entity X present at Event 1*
- *Entity Y present at Event 2*



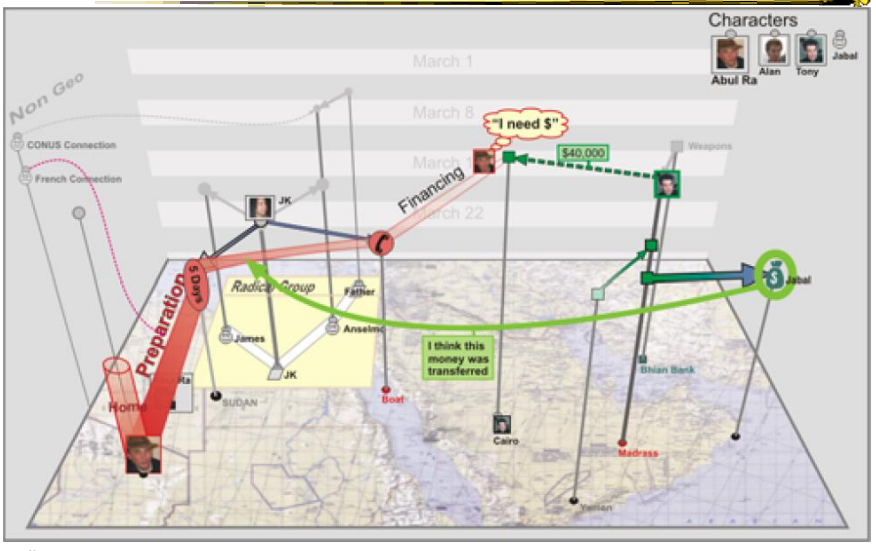
View objects



- Spatial timelines
  - 3-D Z-axis
  - 3-D viewer facing
  - Linked time chart

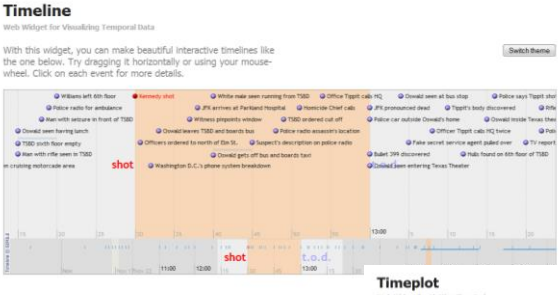


Telling a Story



Useful Widgets

http://simile-widgets.org/



Simile project

Conclusions

- Think about the data
  - What characteristics?
- Can InfoVis help?
  - Maybe not needed
- Think about the visualization techniques
- Which technique(s) work best for your problem?

Taxonomy

Revisit

- Discrete points vs. interval points
- Linear time vs. cyclic time
- Ordinal time vs. continuous time
- Ordered time vs. branching time vs. time with multiple perspectives

Muller & Schumann '03 citing Frank '98

Upcoming

- Value of Visualization
  - No class on Weds.
    Watch video linked on website
- Hierarchies & Trees 1
  - Reading
    Card & Nation '02

References

- Spence and CMS books
- All referred to articles
- Jim Foley & Chris Plaue's take on these slides