More on Time

CS 4460 - Information Visualization

Jim Foley

Some PPts from Prof. John Stasko.

Last revision: October 2016

A Taxonomy of Time Data

- Continuous – a series of values that change over time
  - Non-periodic
  - Periodic
- Discrete – an event that occurs at a specific time
  - Non-periodic (non-recurring
  - Periodic
- More on next pages……..

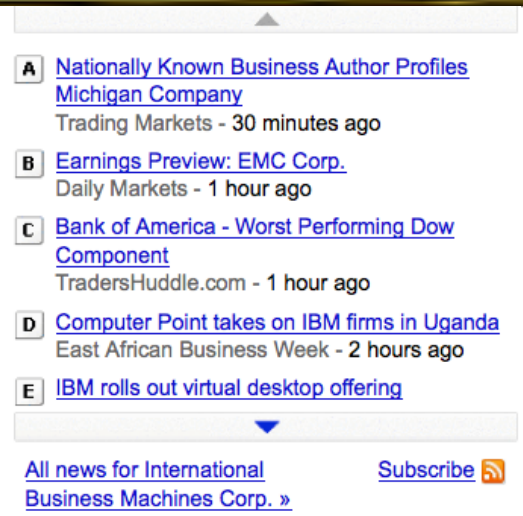
Time	Value
10:00	57
10:01	62
10:02	60
10:03	60
Etc	etc

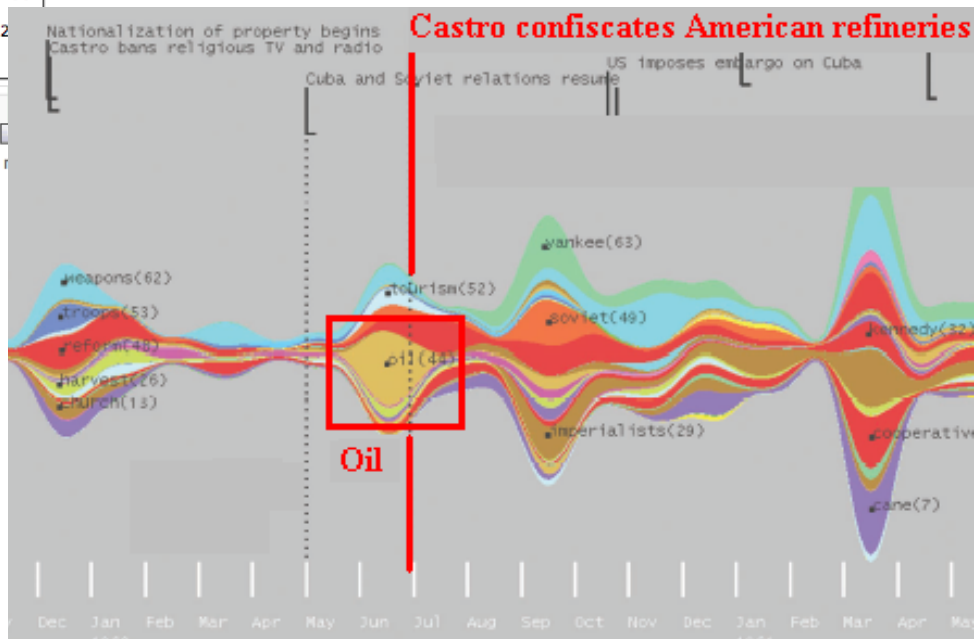
Time	Event
20 July 1957	John born
5 Sept. 1973	John graduates HS

Time Series Examples

- As we step through examples, answer these questions about each example
  - What are similarities?
  - Differences?
  - When is each useful?
  - Pros/cons of each?
    - How much info is visually coded?
    - Scalable to more events & longer time scale / intervals?
    - How good for comparing multiple time series?

Examples





https://eresearch.fidelity.com/eresearch/evaluate/snapshot.jhtml?symbols=AAPL

Discuss

- What type of time data shown?
- When useful?
- How much info is visually coded?
- Scalable to more events & longer time scale / intervals?
- How good for comparing multiple time series?

Example: Finding Daily Patterns

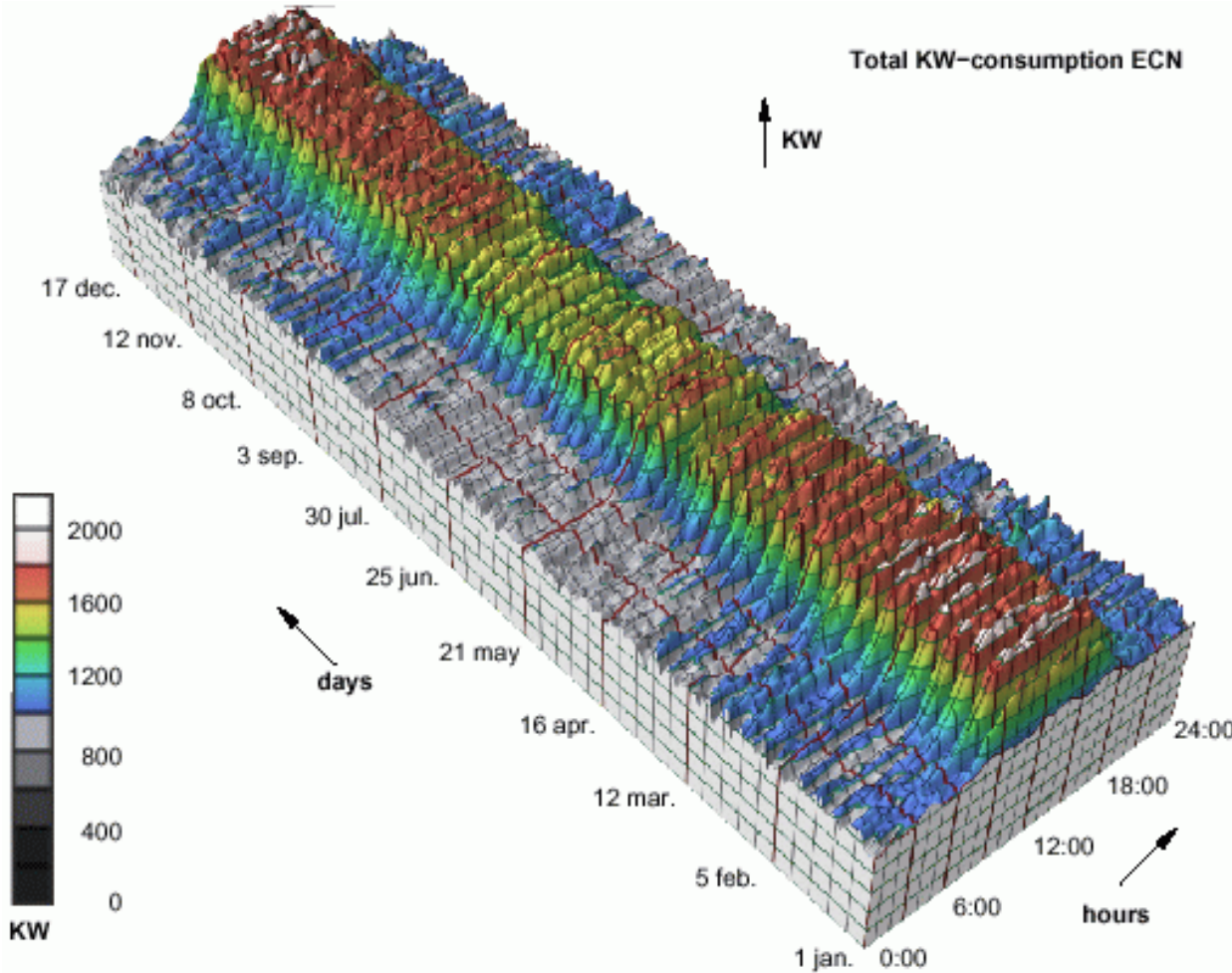
- Suppose you have a daily log, for a year, of energy consumption in a building, or number of people in building
  - Want to find common characteristics
  - Idea applies beyond this example

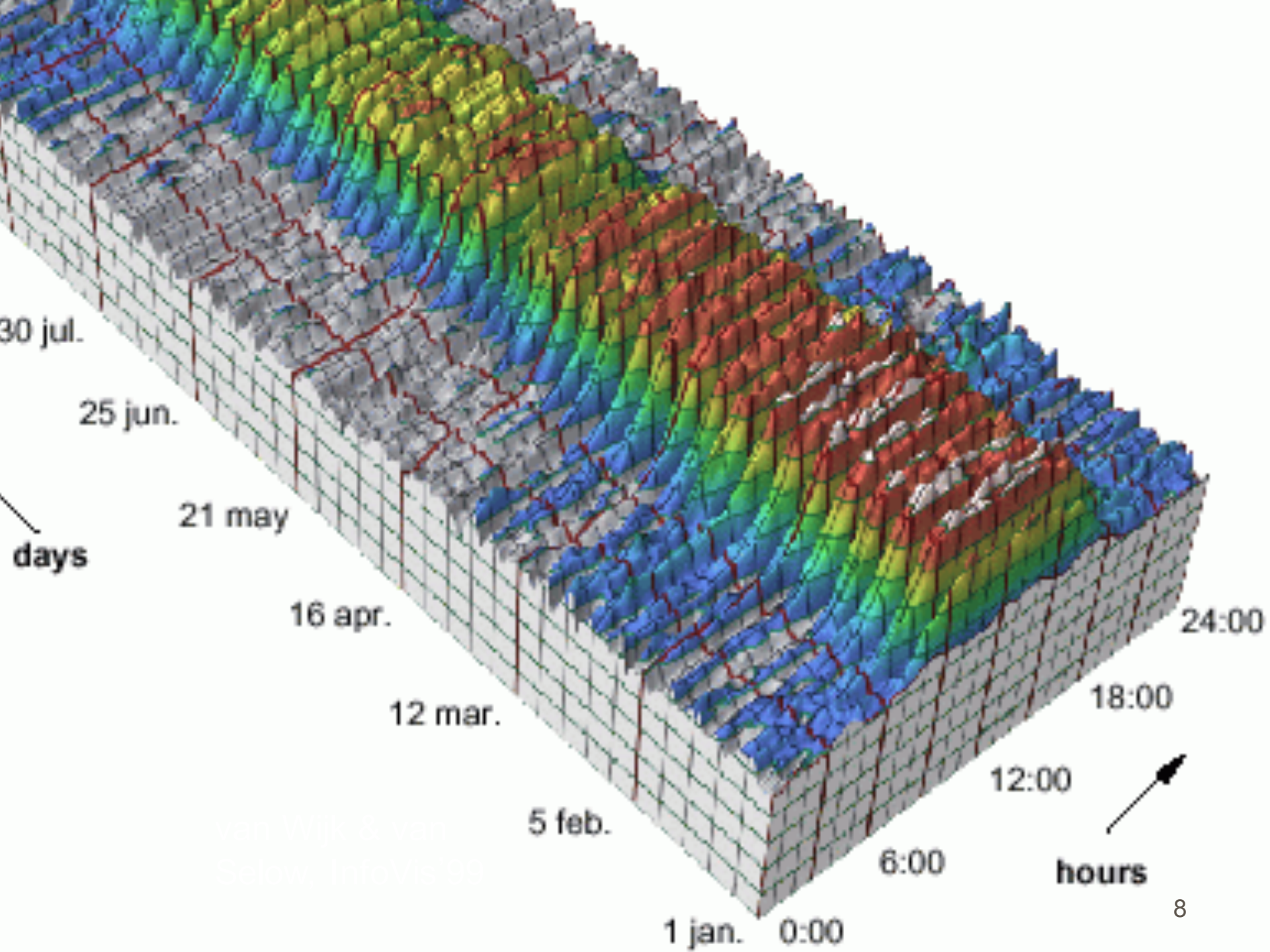
Wijk and Selow, *Cluster and Calendar based Visualization of Time Series Data*, InfoVis '99

Raw Data

- Ideas?

See larger view of data on next slide




30 jul.
25 jun.
21 may
16 apr.
12 mar.
5 feb.
1 jan.
days
0:00
6:00
12:00
18:00
24:00
hours

Discuss

- What type of time data shown?
- When useful?
- How much info is visually coded?
- Scalable to more events & longer time scale / intervals?
- How good for comparing multiple time series?

Use ‘Cluster Analysis’

- Start with the *n* days, call each a cluster
- Find two most similar clusters
- Combine the two into one cluster
- Now have n-1 clusters
- Repeat until some preset number left or a condition is met
- How can results be visualized?

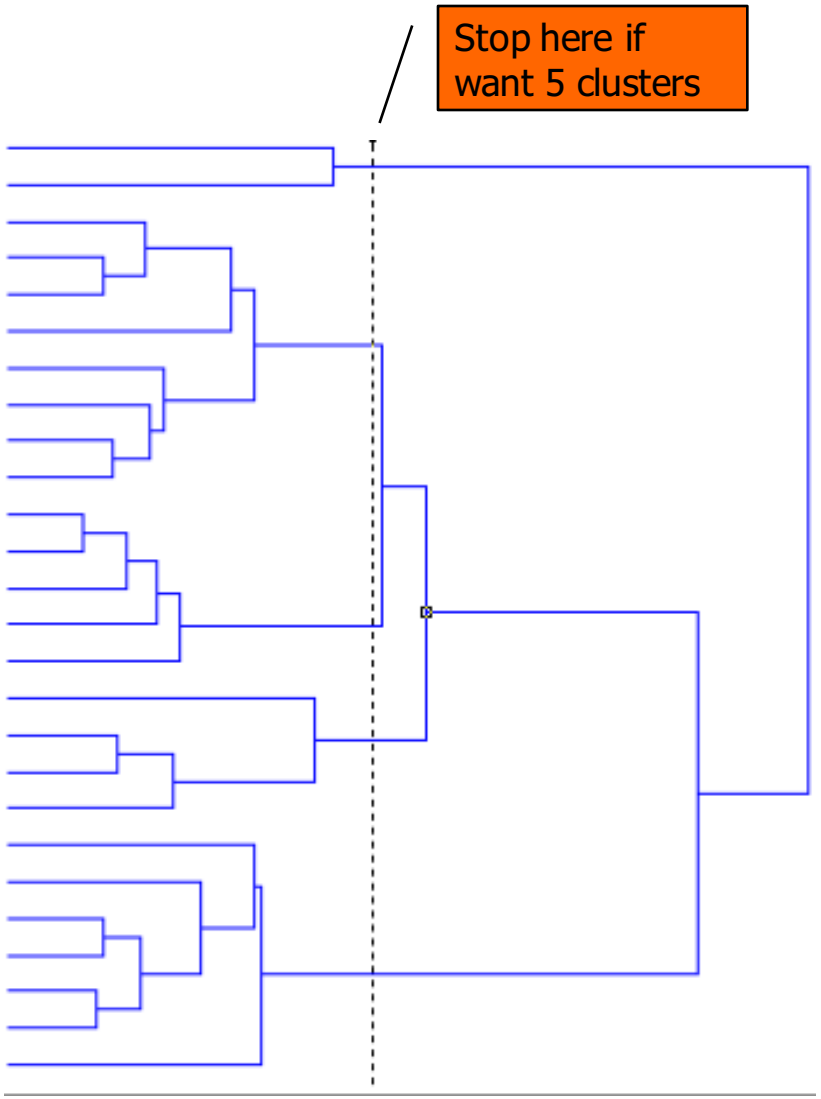
An Aside – Cluster Analysis

- Widely-used method
- Defining 'close' is key
  - Varies from one domain to another
  - Text - count differences – key words, tags
  - Pictures
  - Time-varying data
  - $\sum |\Delta_i|$, i varies over sample length, such as 24 hour period in previous example or over key word count for text

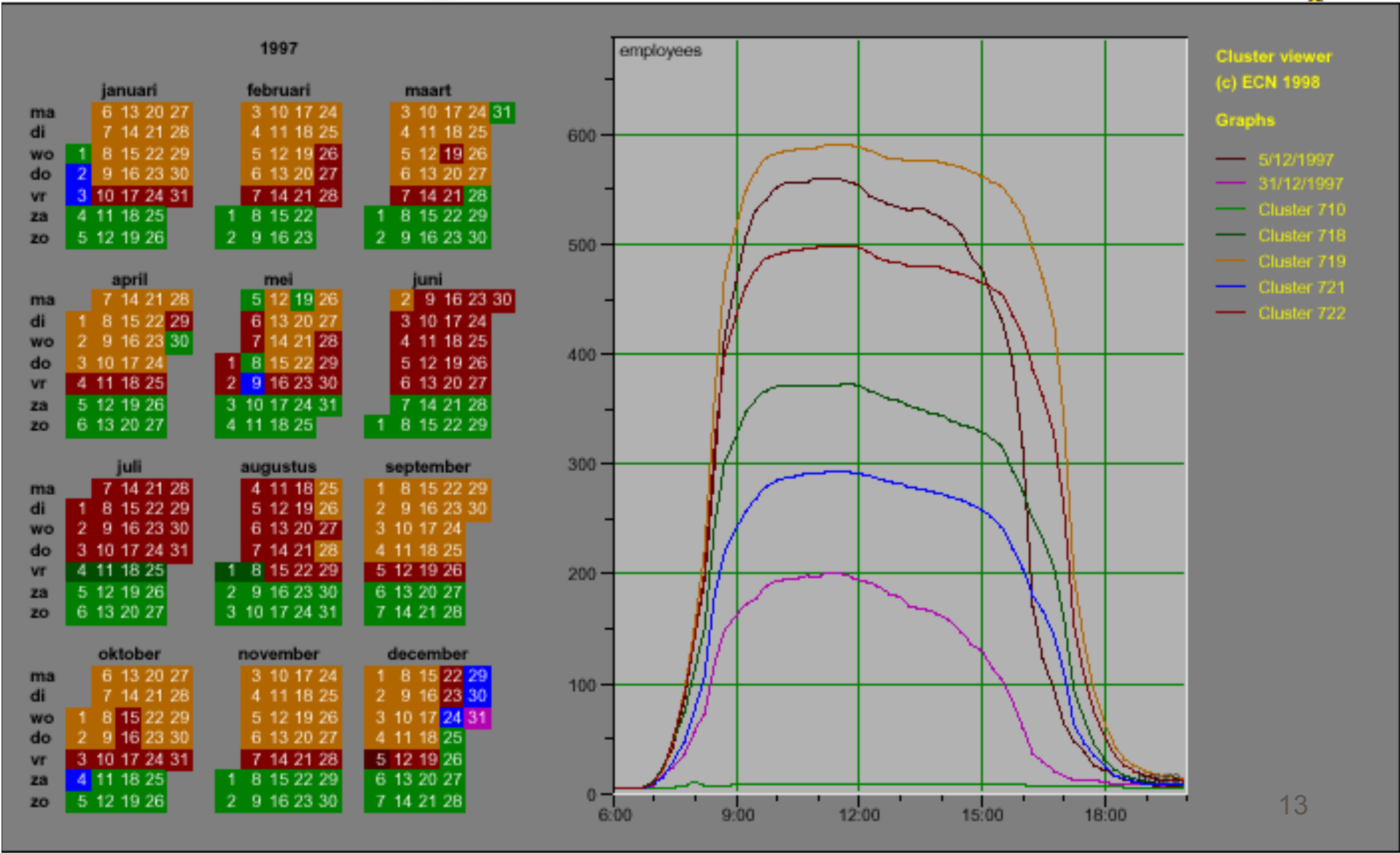
    Why use absolute value of Δ_i ?

An Aside - Dendogram

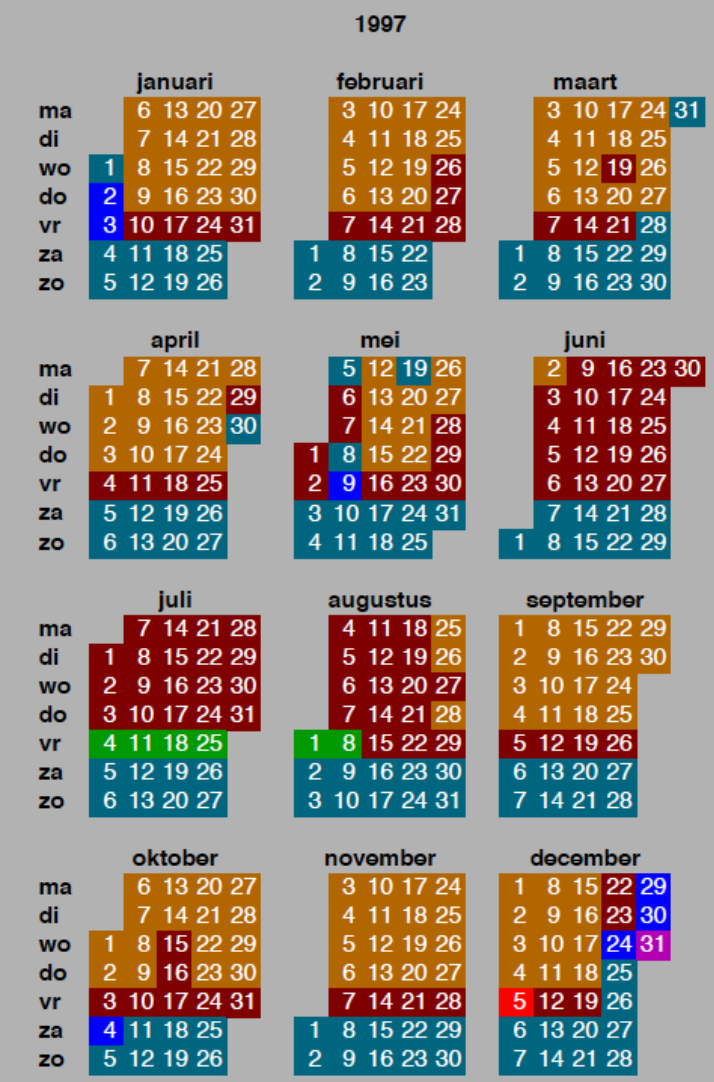
- Dendogram – tree that results from clustering
  - Can show process until is a single cluster
  - Stop clustering when have as few as desired

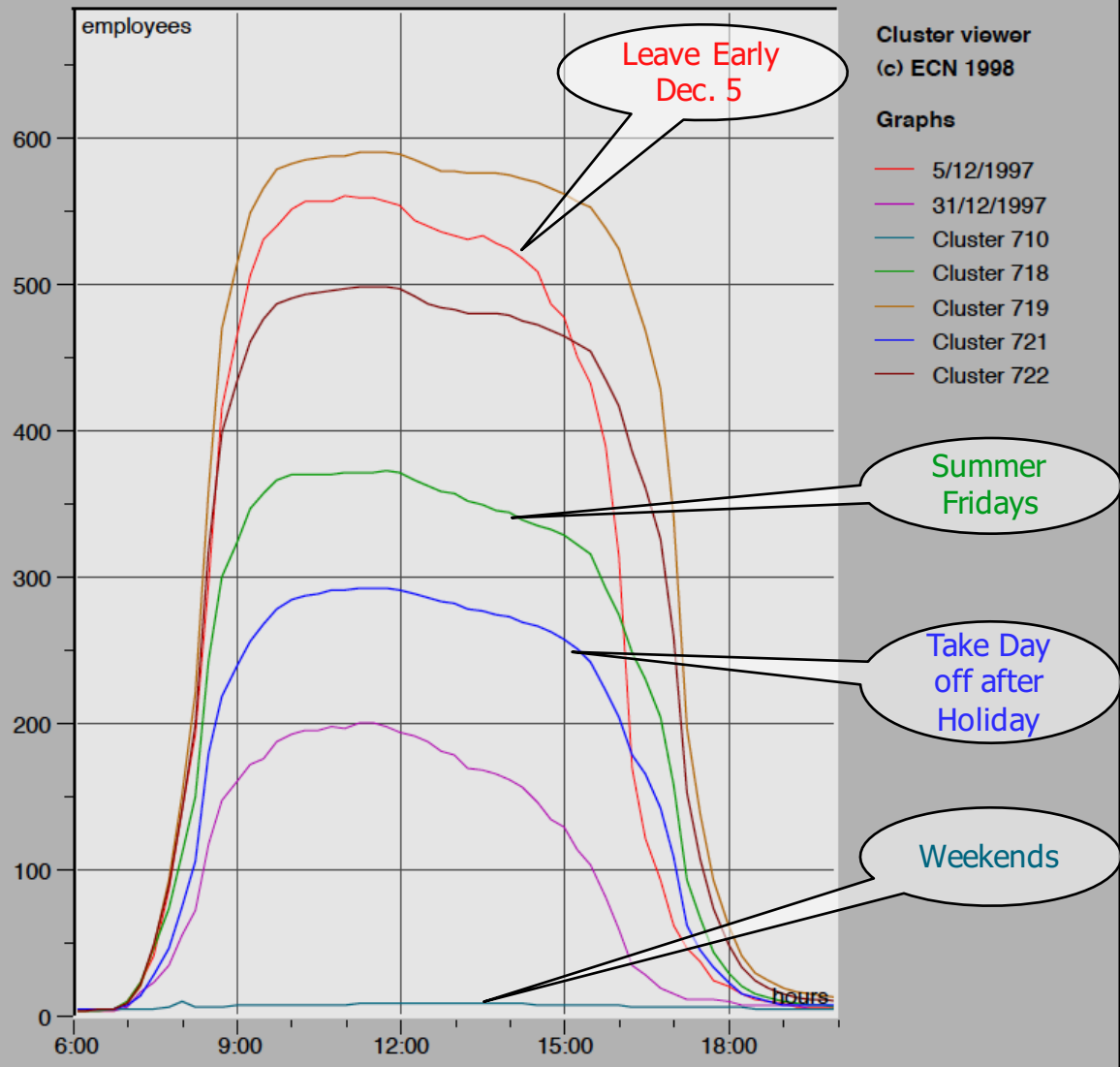


Cluster Display – People at Work



Cluster Display – People at Work





What Cluster Display Shows Us

- Traditional office hours followed
- Most are present late mornings
- Fewer are present on summer Fridays
- Very few people work holidays
- School vacations
- Day after holidays
- Many people leave at 4PM on 12/5
  - Very special in The Netherlands – St. Nicholas' Eve

Example: Spiral Display - Periodic Data

- Useful if data follows a repetitive pattern
- Can reveal periodicity
  - One loop is one period
  - What if data NOT periodic?
- Time line becomes the spiral
  - Avoid problem of long time line
- *Could use concentric circles instead of spiral*

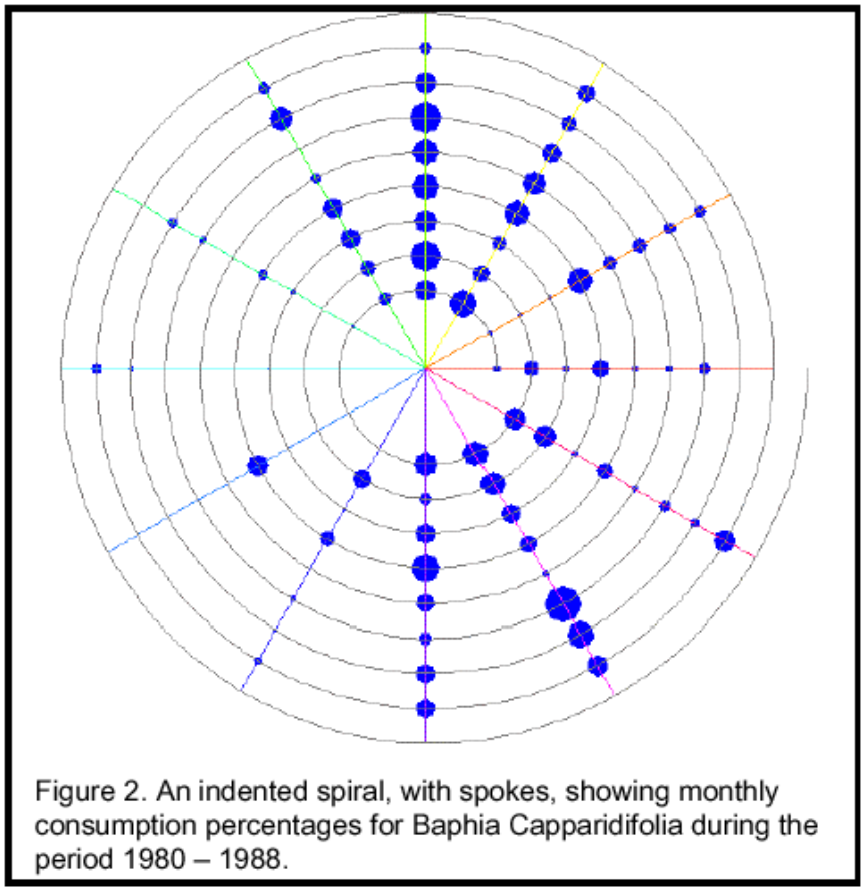
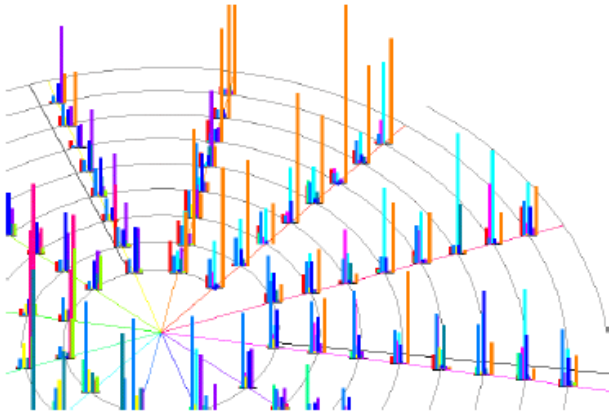


Figure 2. An indented spiral, with spokes, showing monthly consumption percentages for Baphia Capparidifolia during the period 1980 – 1988.

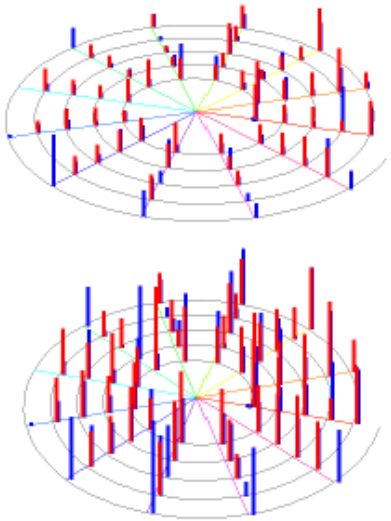
Konstan and Carlis, *Interactive Visualization of Serial Periodic Data*, UIST '98; data set is chimpanzee food consumption.

Add Third Dimension for More Data



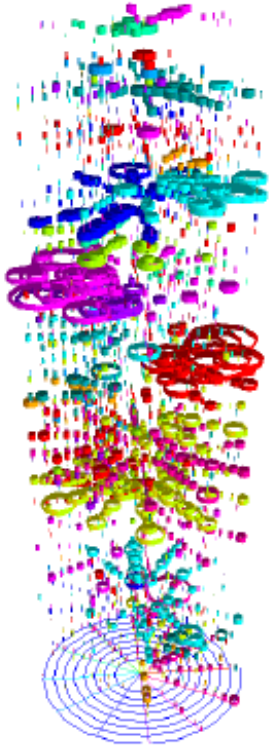
Mini bar-chart at each point

Black line – start of rainy season



Two linked spirals:

- 2 chimpanzees
- group avg size & max size



112 food types

Useful? ☺

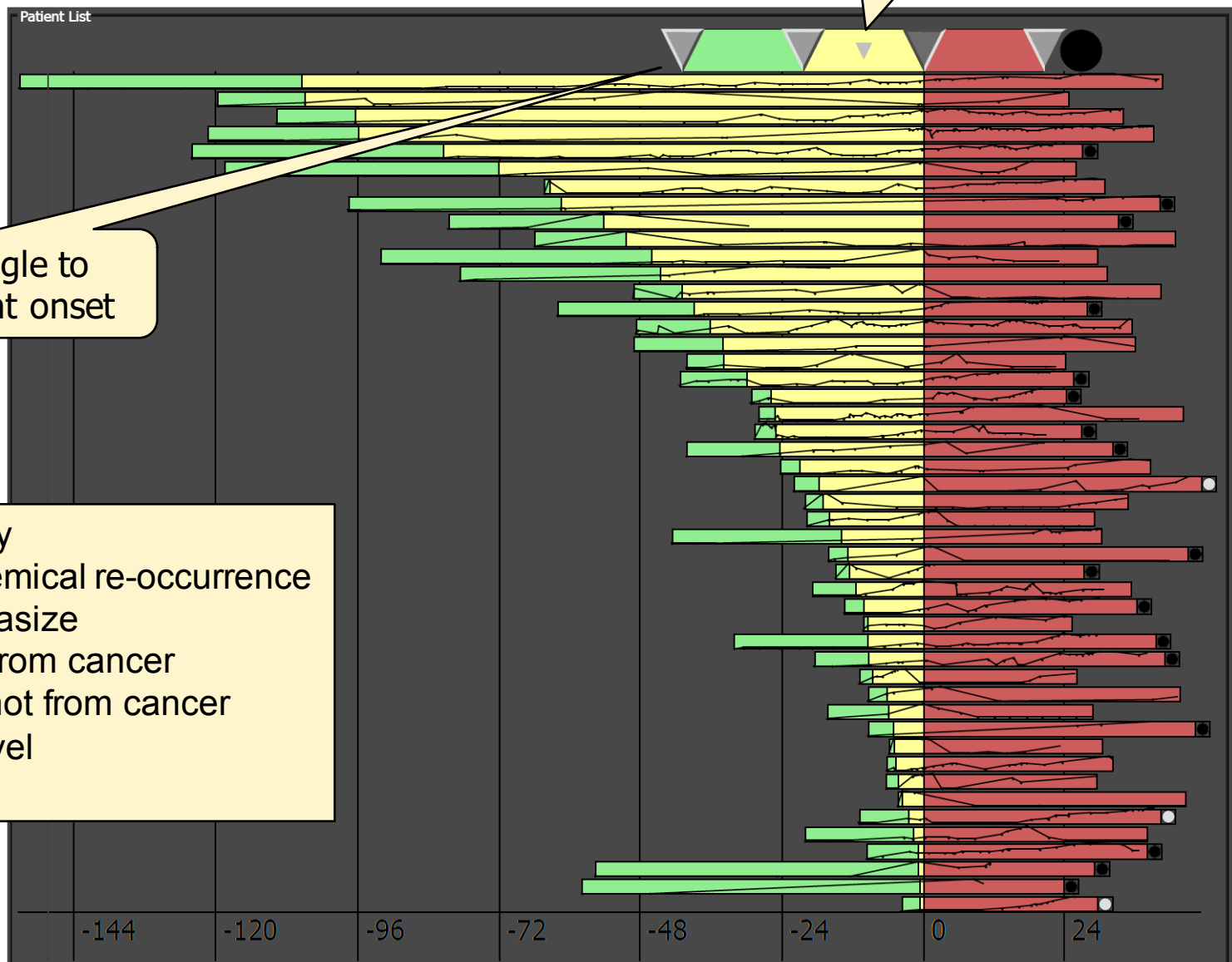
Discuss

- What type of time data shown?
- When useful?
- How much info is visually coded?
- Scalable to more events & longer time scale / intervals?
- How good for comparing multiple time series?

Design Exercise

- Data - for every hour of the day for an entire year
  - Number of road accidents in three different counties, plus total
- For each month, show aggregated accident counts for Sun, Mon, Tues etc, for each county & total
- Select date range and time of day range
- Now also show hour-by-hour accident count for each county & total, for the selected date/time ranges
  - (and possibly just for one day or the week during the selected date/time ranges)
- I have one design, looking for lots of creativity from y'all ☺

Prostate Cancer Progression

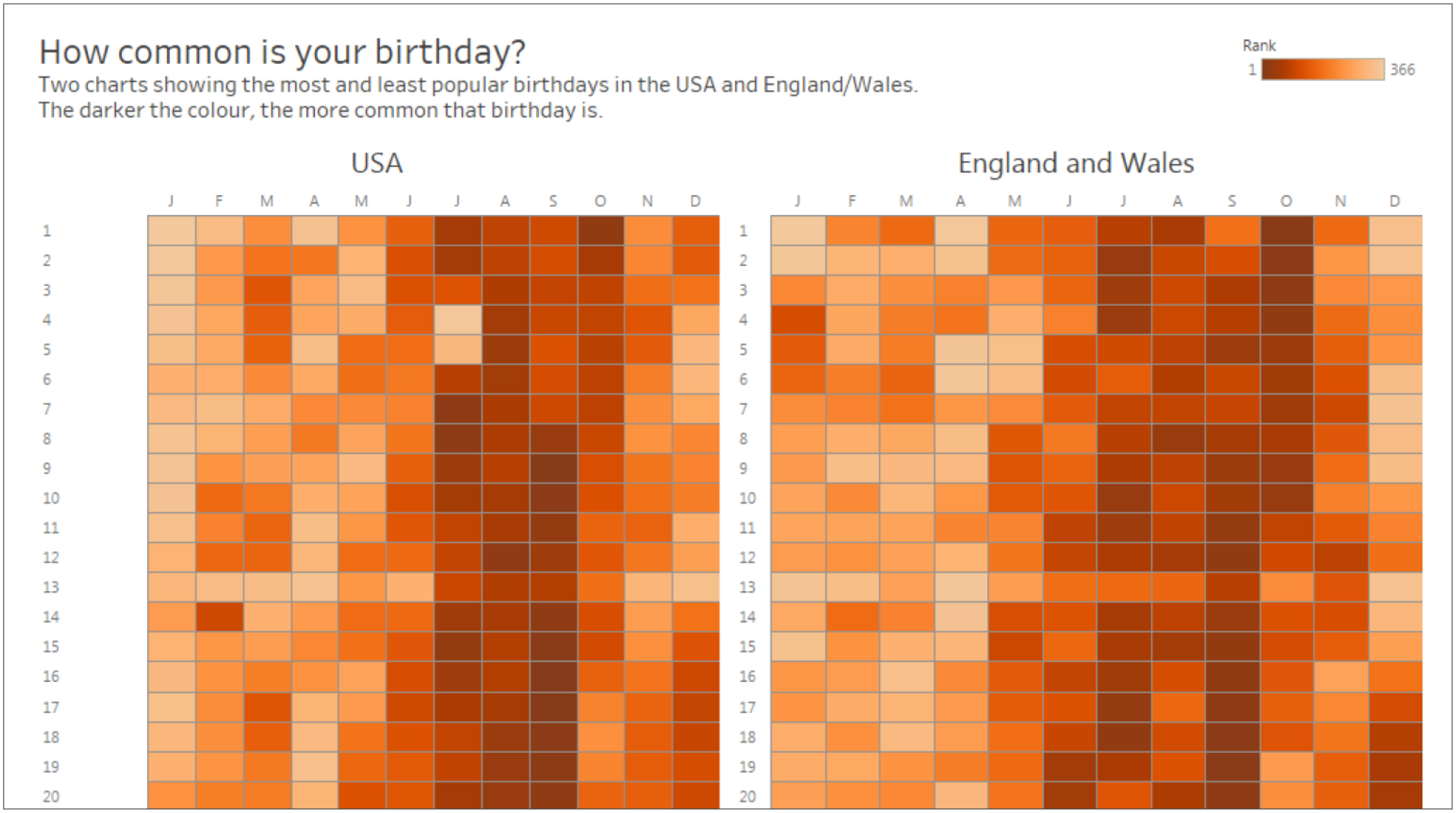


Click in colored area to sort by event duration

Click on triangle to align on event onset

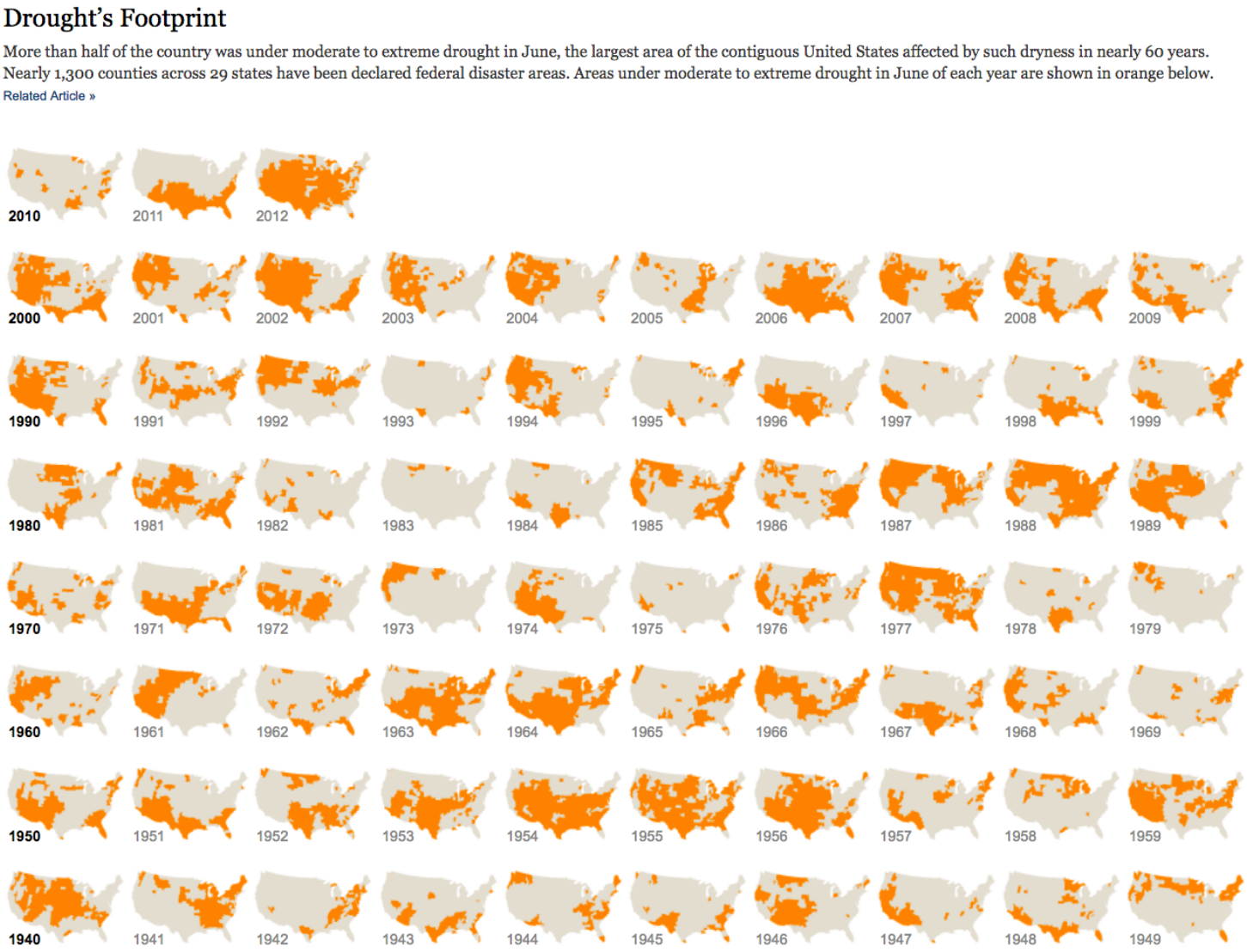
Green: surgery
Yellow: biochemical re-occurrence
Red: metastasize
Black (solid): death from cancer
Black (dot): death not from cancer
Spark lines: PSA level

Time in 2D


How common is your birthday?
Two charts showing the most and least popular birthdays in the USA and England/Wales.
The darker the colour, the more common that birthday is.
Rank
1
366
USA
England and Wales
J F M A M J J A S O N D
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Time via Parallel Coordinates

Time via Small Multiples

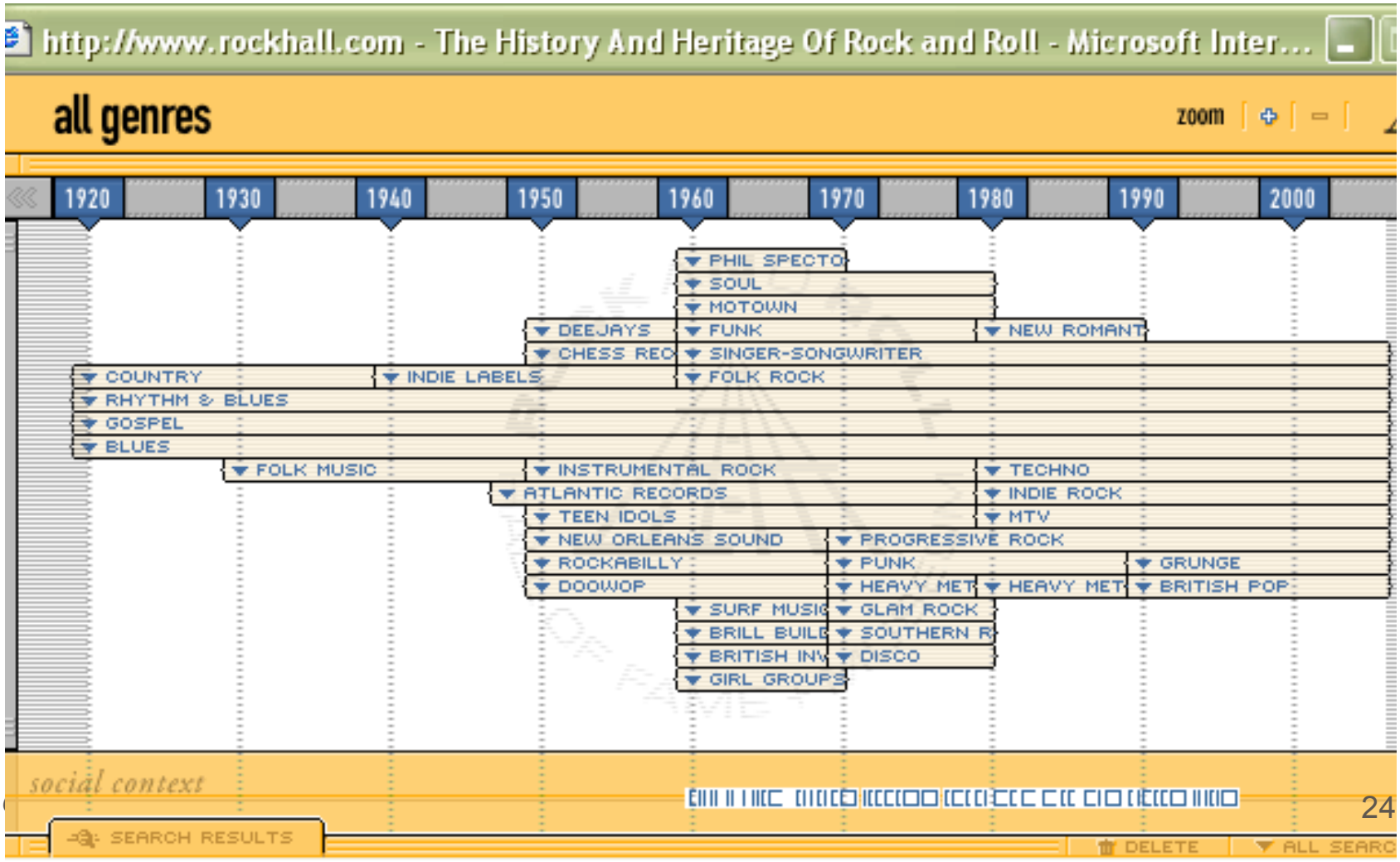


http://www.nytimes.com/interactive/2012/07/20/us/drought-footprint.html?_r=2&

Example: Music Over Time

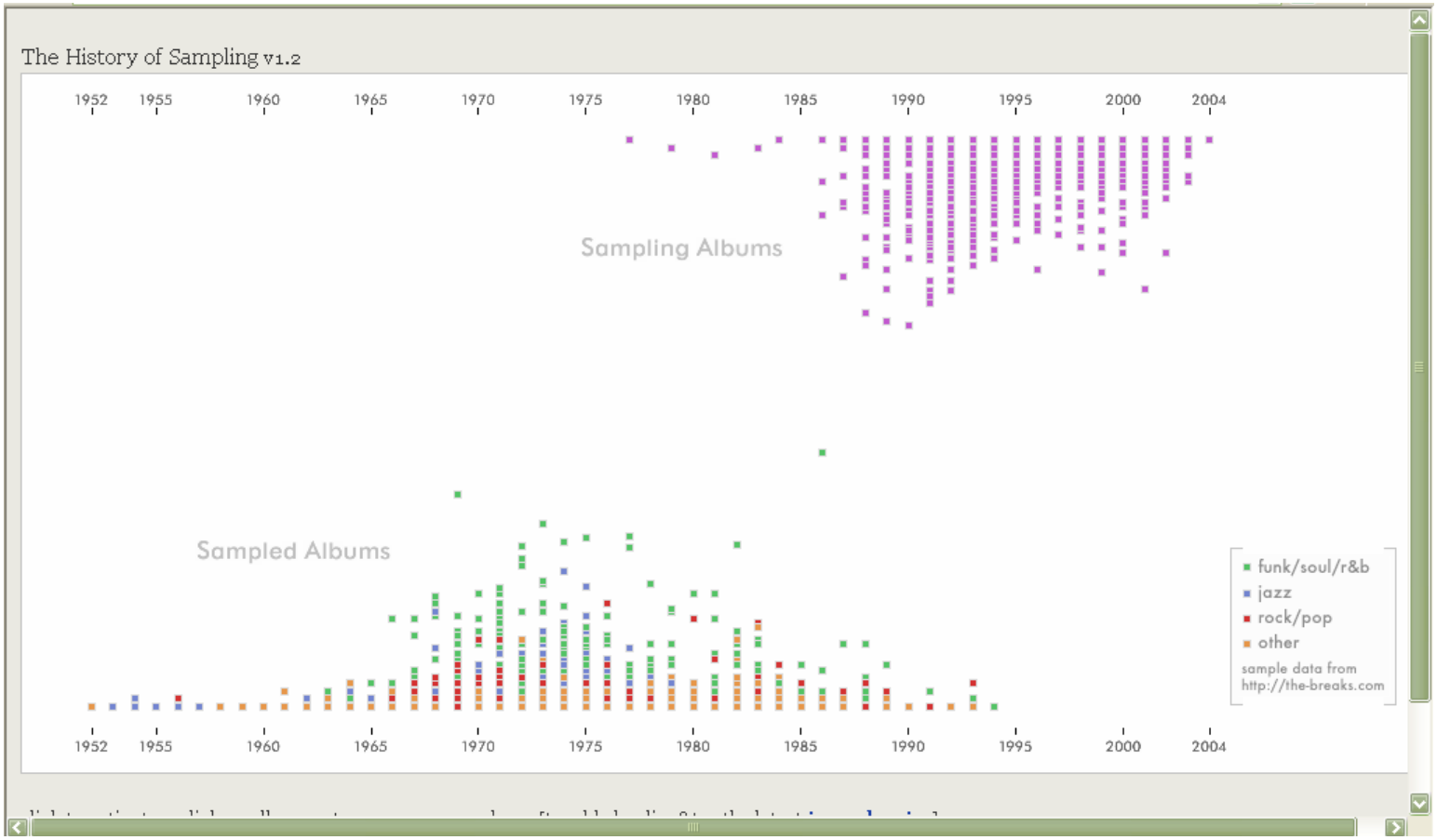
History of Genres

http://rockhall.com/timeline/ (dead link)



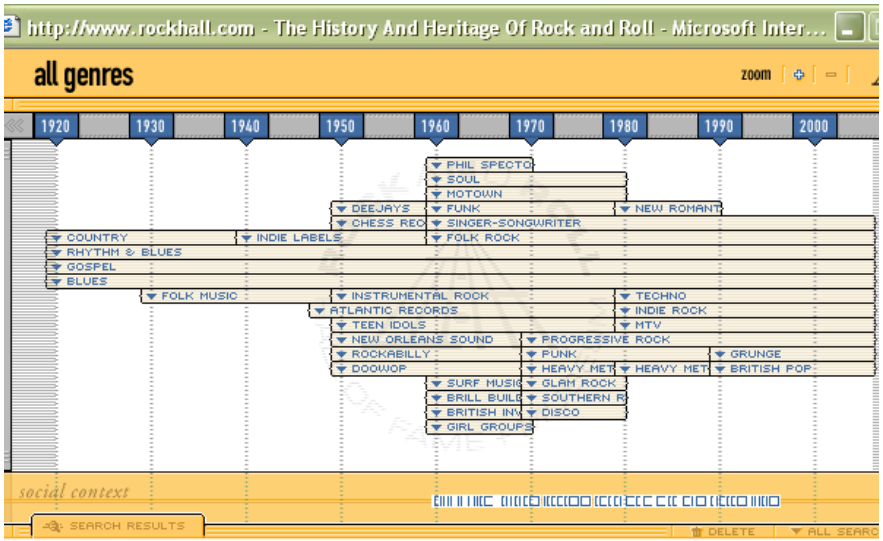
Example: Music Over Time

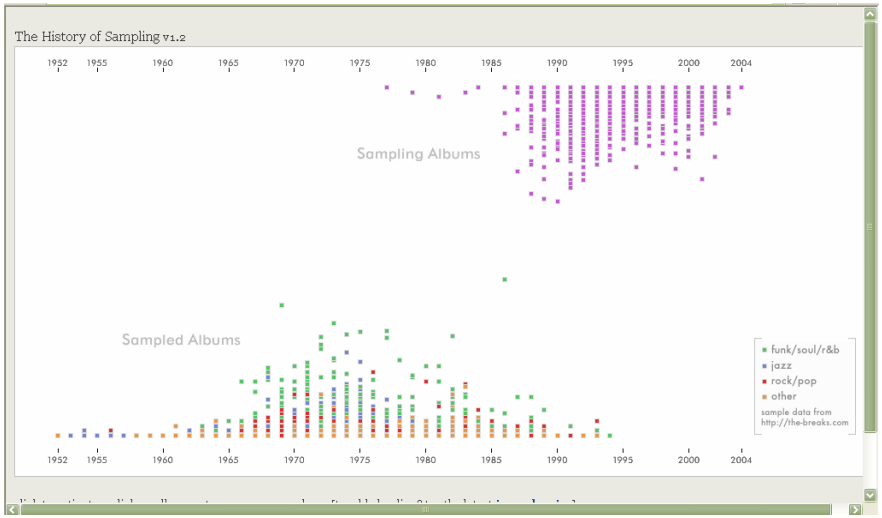
History of Sampling jessekriss.com/projects/samplinghistory/



Music Over Time .....

- Thinking about
  - Rock Hall
  - Sampling History
- What do these two have in common?
- Could elements of one be used in the other?

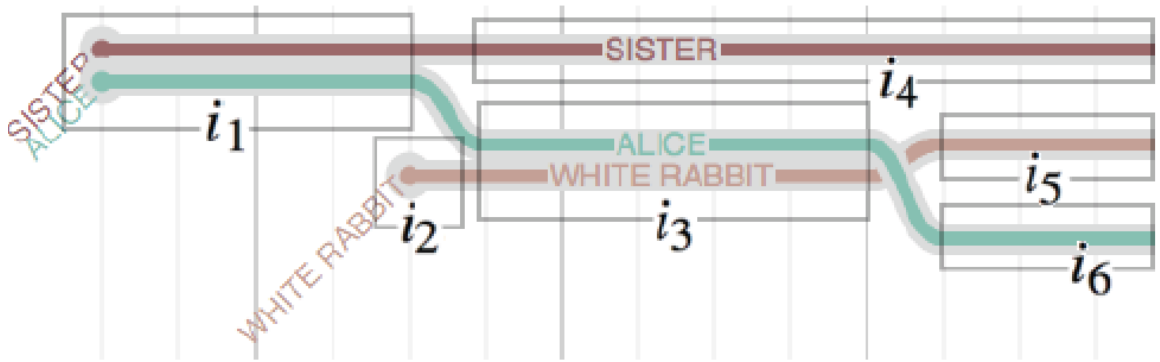




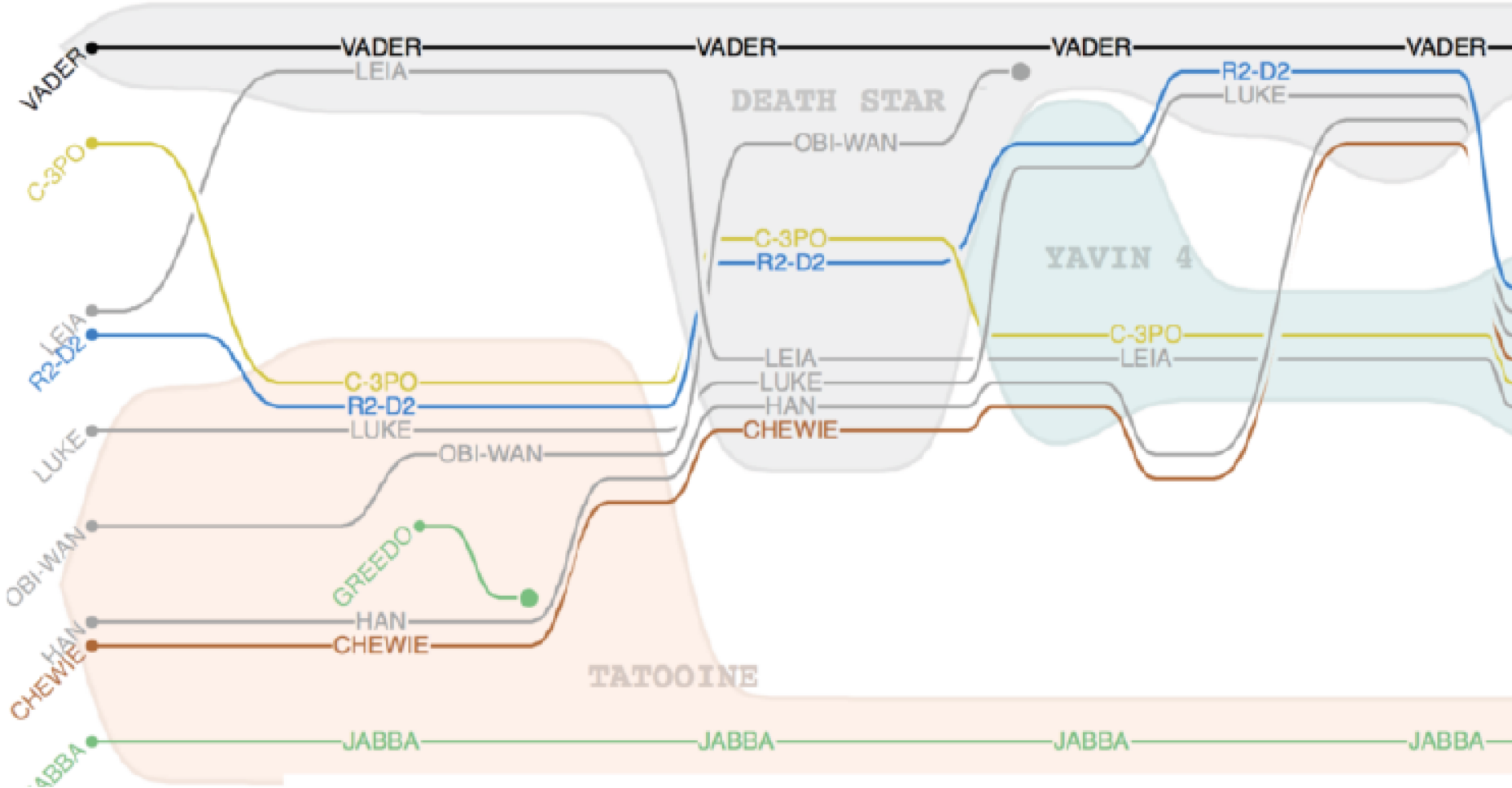
Storylines

i_k	t_k	d_k	M_k
i_1	0	4	SISTER, ALICE
i_2	4	0	WHITE RABBIT
i_3	5	5	ALICE,WHITE RABBIT
i_4	5	217	SISTER
i_5	11	21	WHITE RABBIT
i_6	11	6	ALICE

Sister & Alice together from t = 0 for 4 time units



- Design Considerations for Optimizing Storyline Visualizations, Yuzuru Tanahashi and Kwan-Liu Ma, IEEE Transactions on Visualization and Computer Graphics, Dec. 2012

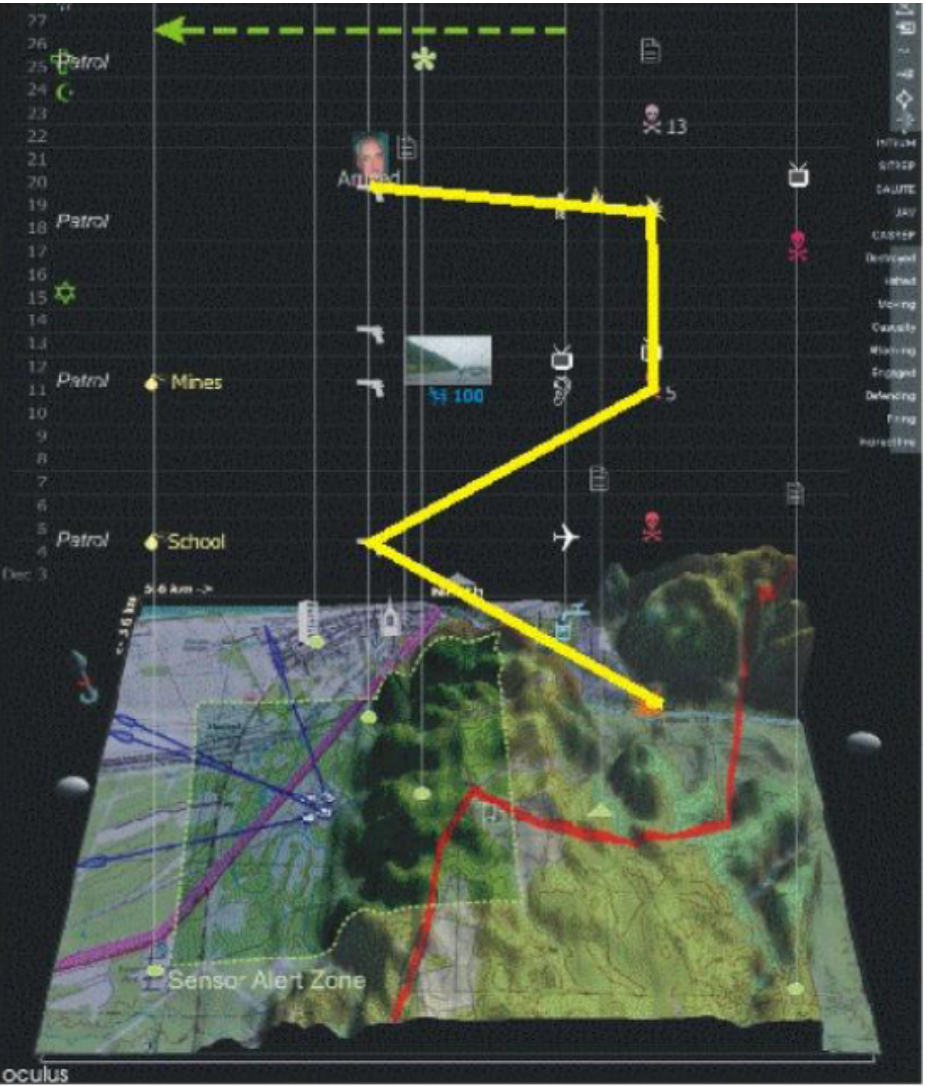


- Lines representing interacting characters must be adjacent.
- Otherwise, lines must not be adjacent.
- A line must not bend except to converge or diverge with another line.

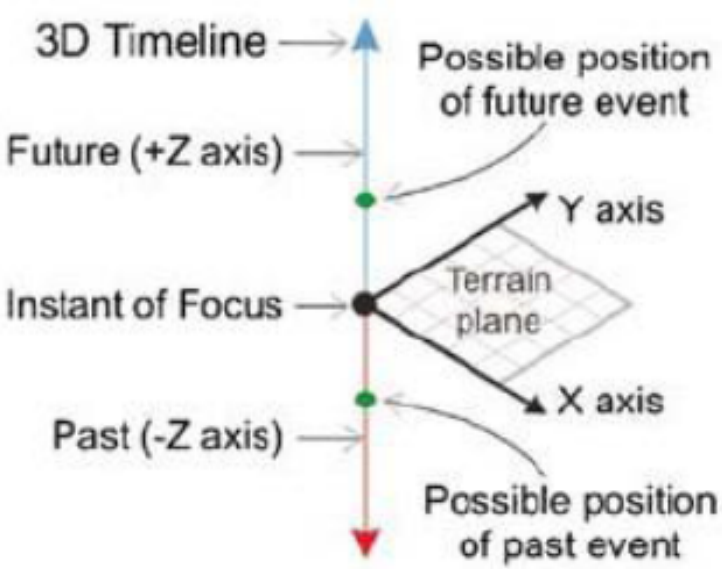
Example: Time + Geography

- Typically superimpose temporal events on a map
  - Intelligence analysis
  - Literary plot analysis
  - Military planning
  - Maybe in future plan plots for interactive games
- Following figures from GeoTime, a product of Oculus www.oculusinfo.com/

GeoTime Spatial Timelines



- Vertical time axis (z-axis)
- Terrain plane (xy



GeoTime Example

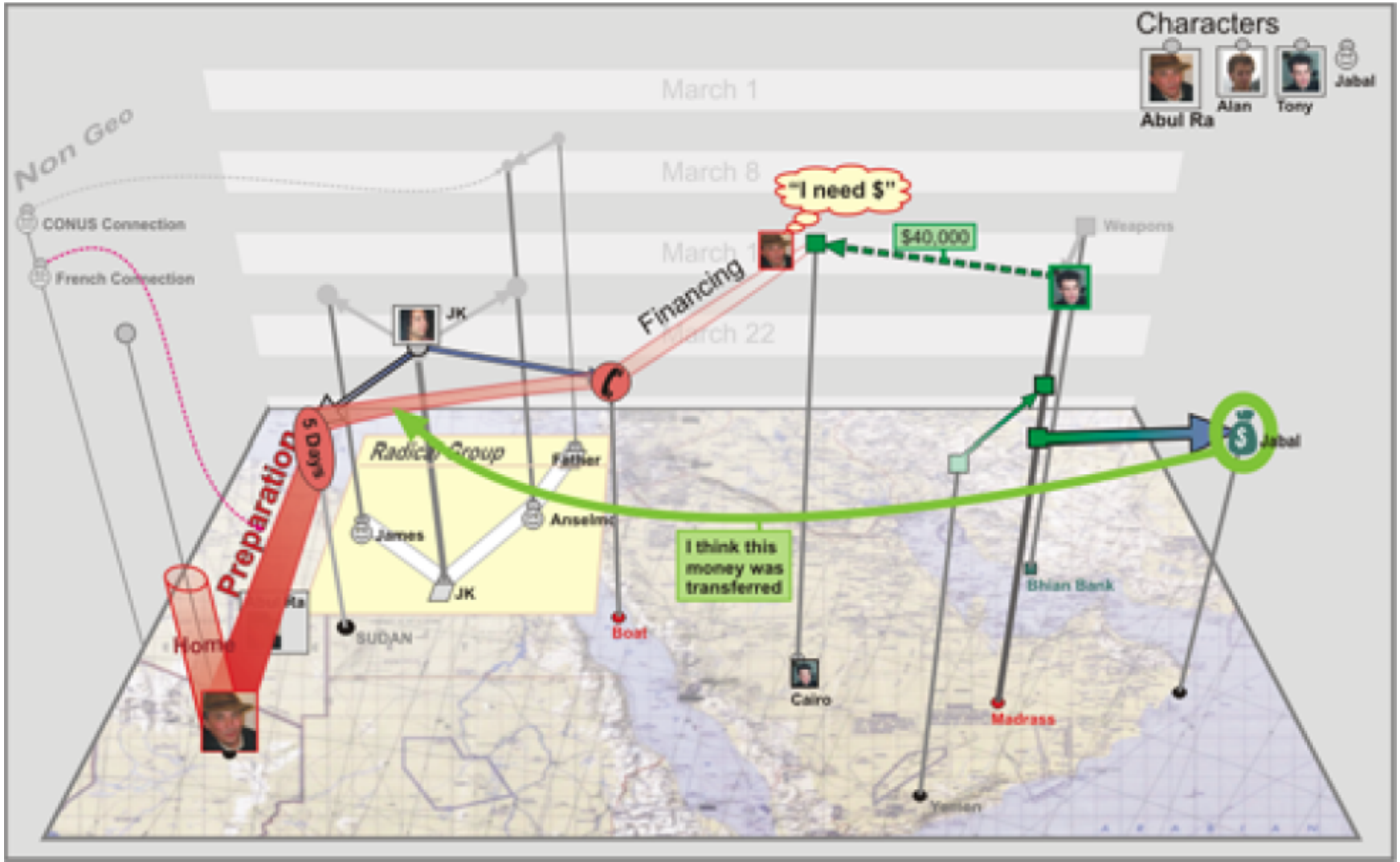
- From http://www.youtube.com/watch?v=YUCgeXXeEiU&feature=related

- My file
  - GeoTime_ Investigating IED Attacks.flv

Time and Geography Story

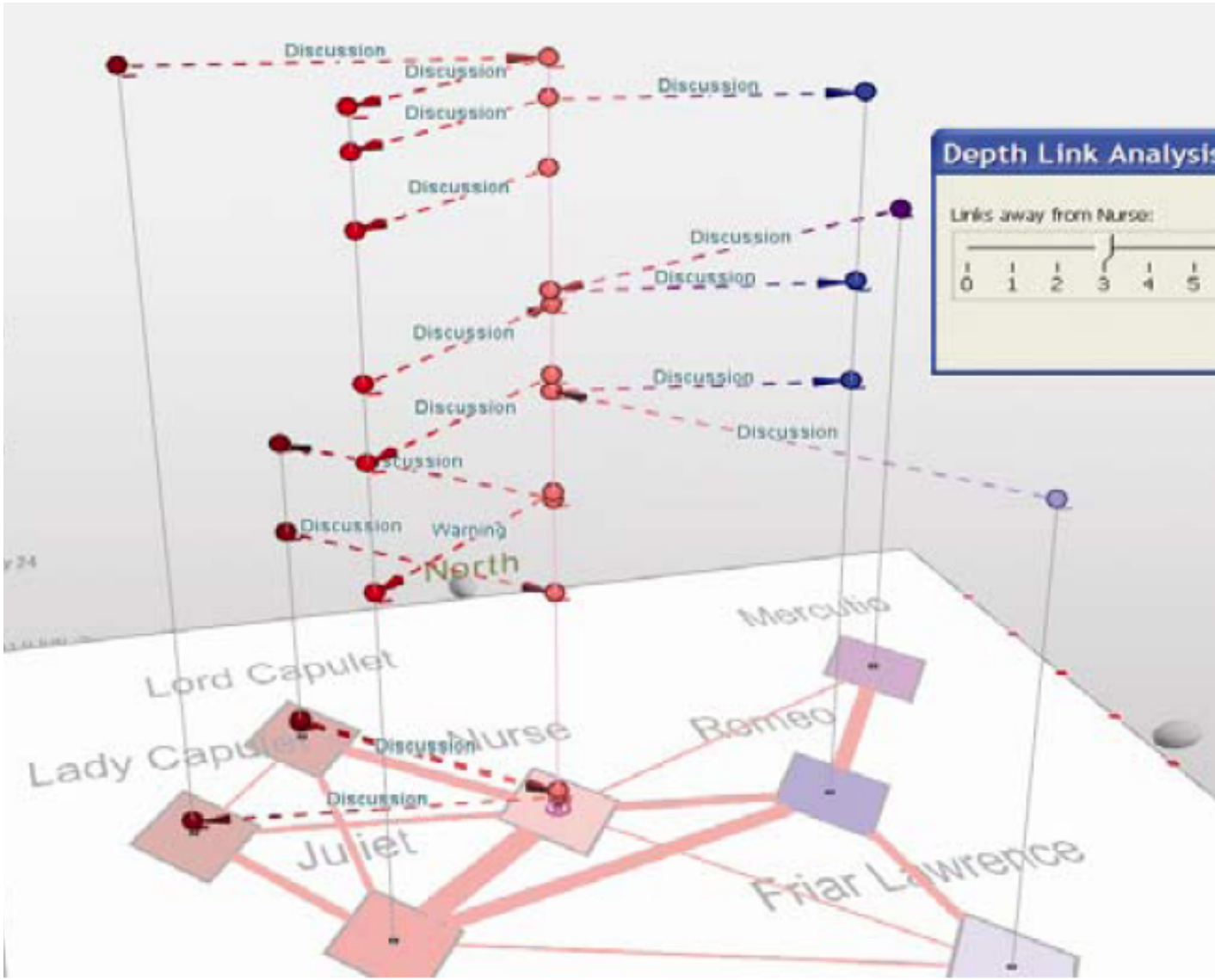
Or an intelligence analysis ….

GeoTime



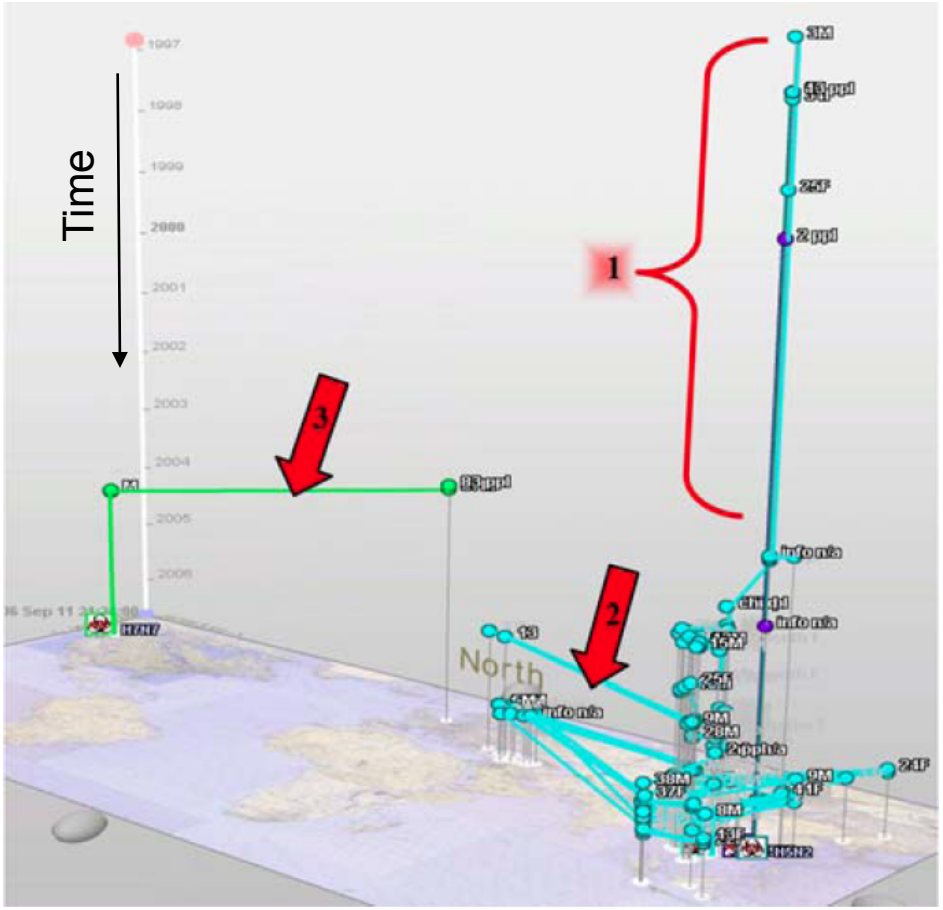
Literary Dialogue

Discussion between Romeo and Juliet, filtered to just show those connected to the nurse.



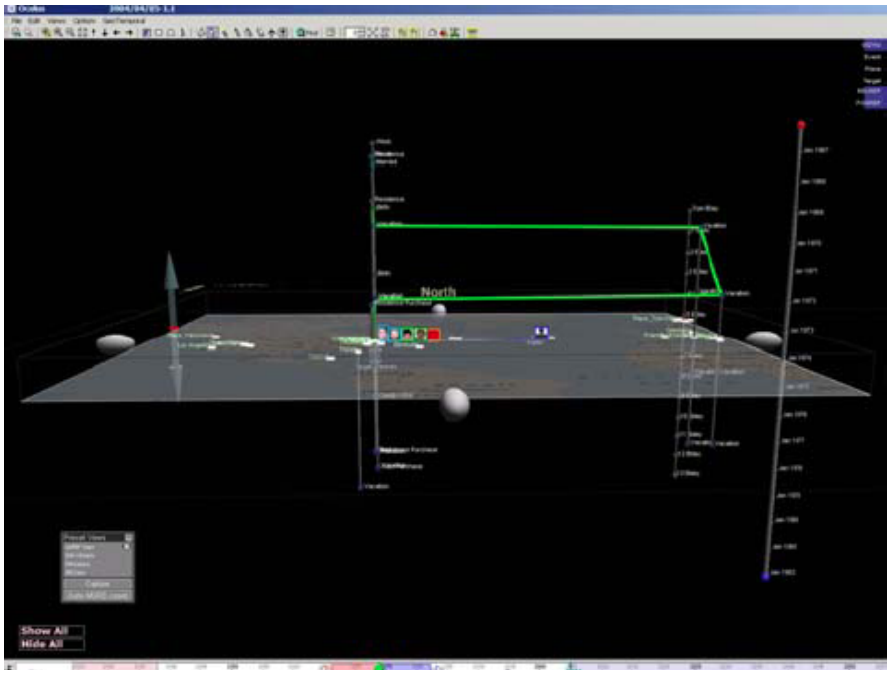
Avian Flu Spread

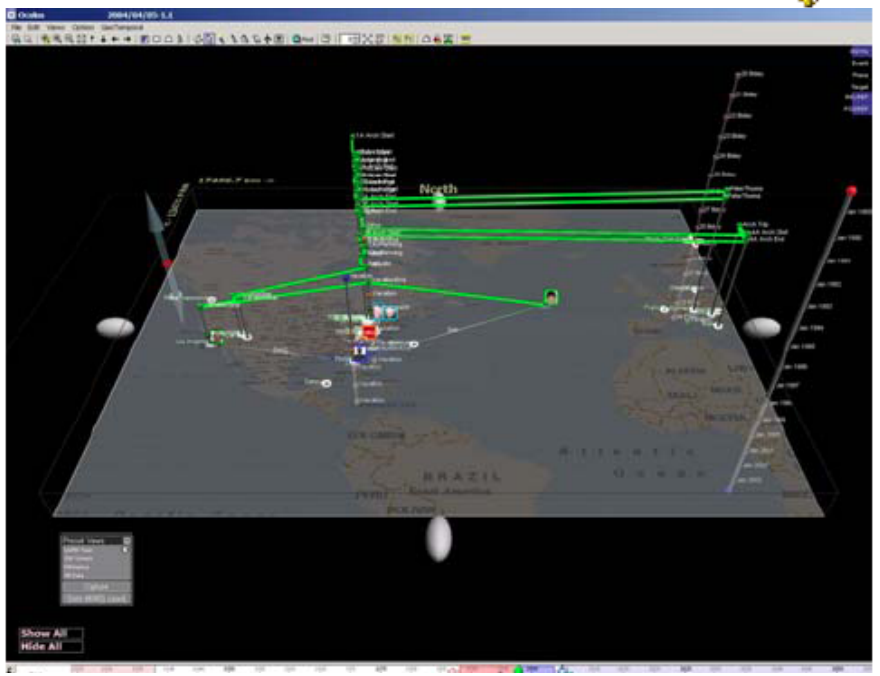
- Worldwide human cases of avian flu. Sporadic localized activity seen in Asia until 2003 (1).
- H5N1 cases increased in frequency in 2004 (2) with the bulk of disease activity in Asia.
- Isolated incidents of strain H7N7 found in Canada and Europe (3).



Proulx et al, *Avian Flu Case Study with nSpace and GeoTime*, IEEE VAST '08

Interaction in GeoTime





Time slider at bottom
Moveable time scale at right
Green line traces one entity's movement in time and geography

Overhead view
Time slider advanced from view on left

Plus lots more interaction – what would you want?

Time Series Tasks – More??

- Compare two time series
- Find highs and lows
- Determine periodicity
- When did X happen?
- Did A happen before or after B?
  - Or did they occur simultaneously?
- How long did it take?
- When will it happen again?
- How often did X occur?

- What else did we see happening?

Wrapping up Time – Some Take-aways

- How decide what type of Infovis to use with what type of temporal data?
  - Temporal data types – remember continuous, discrete, periodic
  - Which of the examples work with which types?
- Compressing/expanding time axis
- Dependencies (as in PERT charts)
- It is not just time, it is time plus other data
  - Possibly including maps
- When to use Infovis, when not to?

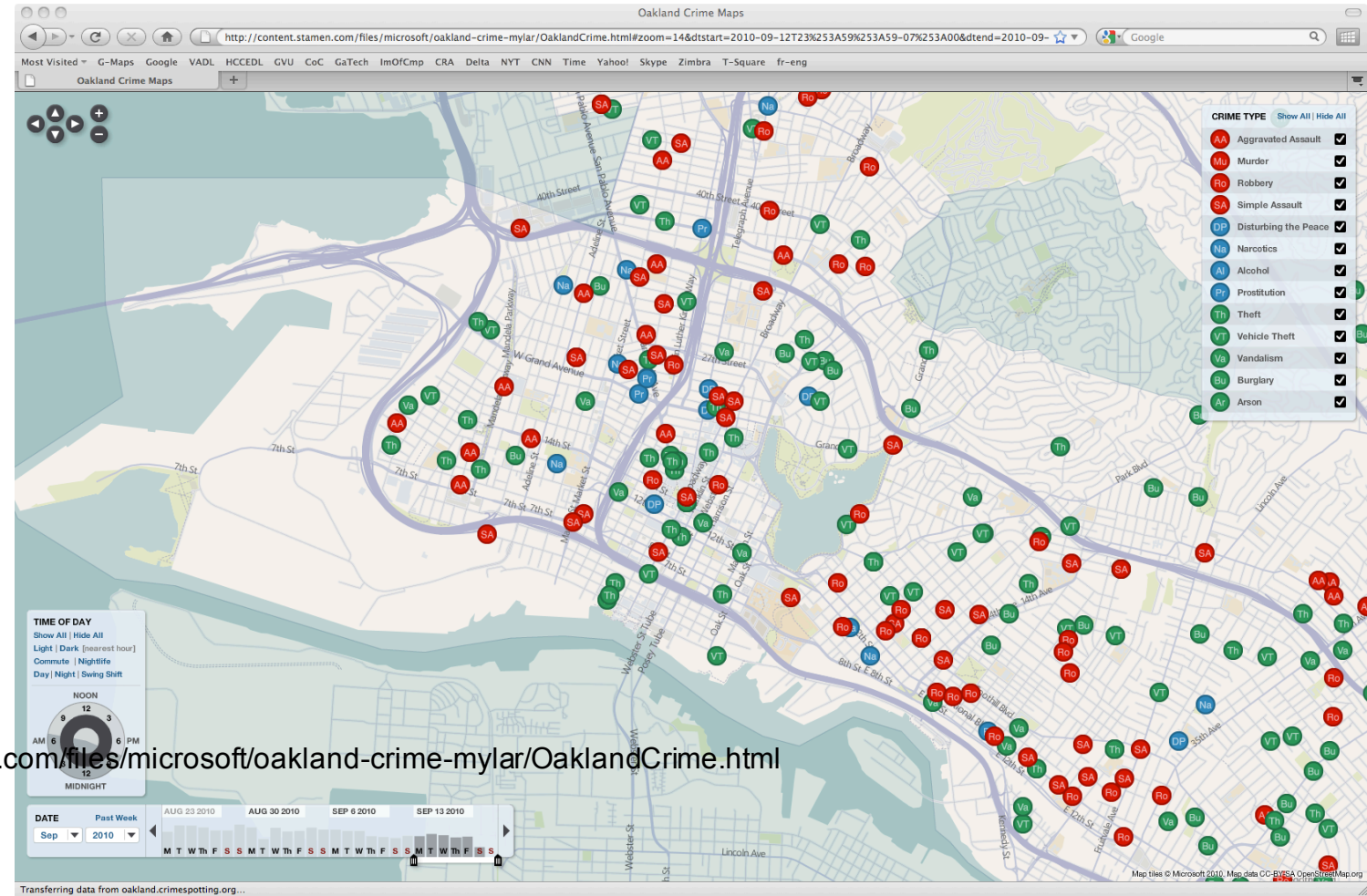
The End

Older Material Follows

- No longer used but relevant for digging further.

Interactions in Space and Time

What interaction methods are used here?



http://content.stamen.com/files/microsoft/oakland-crime-mylar/OaklandCrime.html

Example: Computer Logs

- (Huge) data sets
- How can you examine millions of lines to debug or locate a problem?

- Goal: Want to look for unusual circumstances, patterns, etc.

MieLog Computer Log Inspector

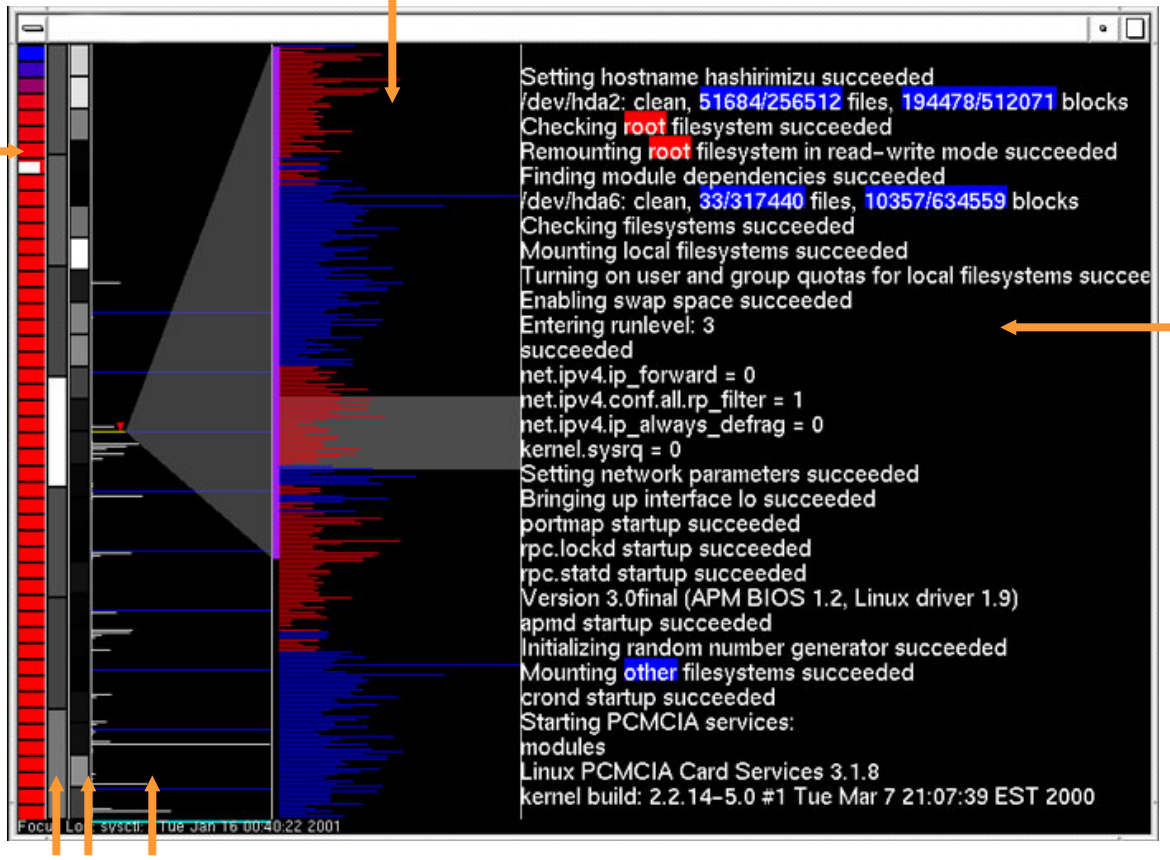
- Goal: Help computer system administrators examine log files.
- How would you do this?

Takada & Koike, *MieLog: A Highly Interactive Visual Log Browser Using Information Visualization and Statistical Analysis*, LISA 16th Sys Admin Conference, USENIX, 2002

MieLog View

Outline area
pixel per character

Tag area
block for each unique tag, with color representing frequency (blue-high, red-low)



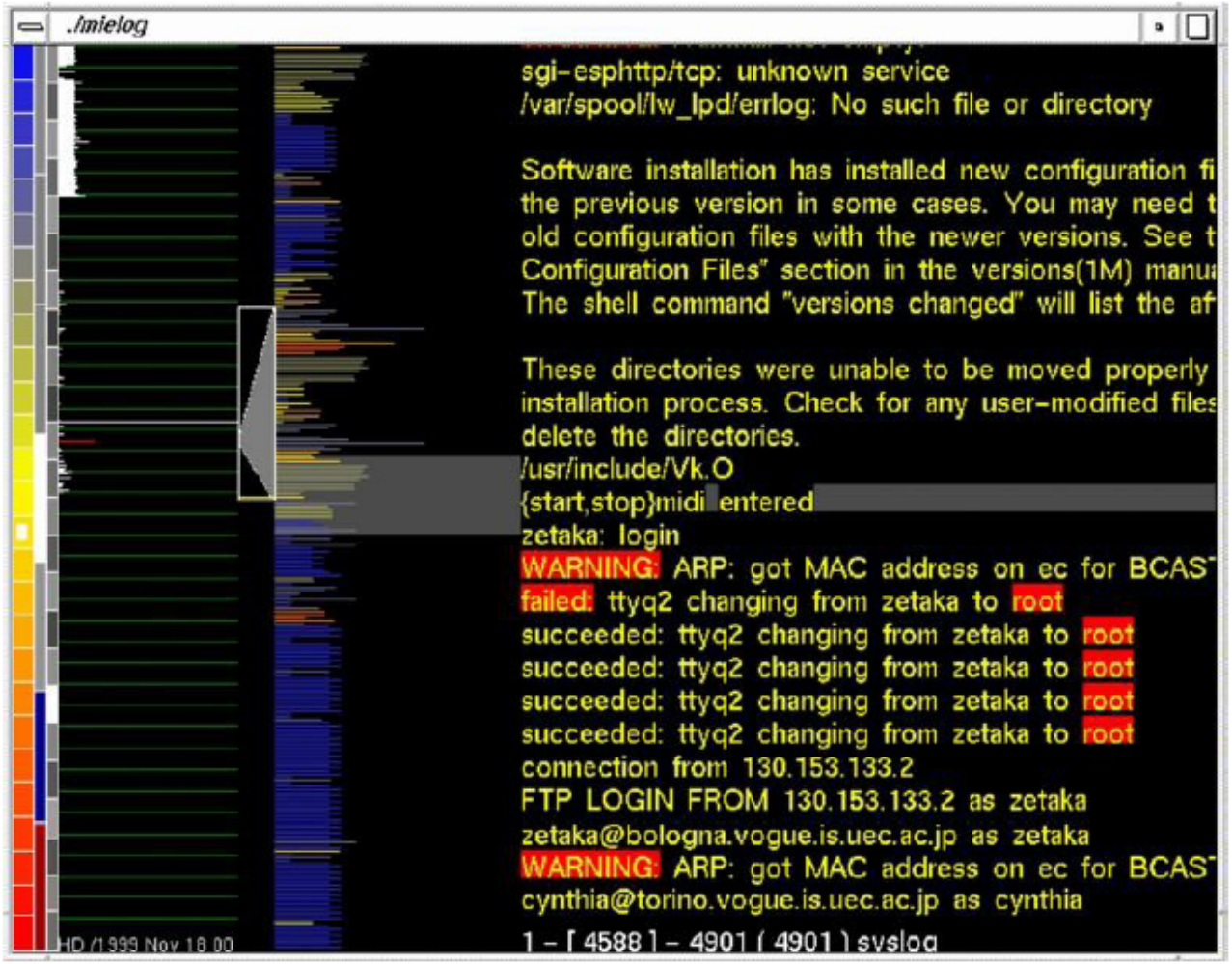
Message area
actual log messages
(red – predefined keywords
blue – low frequency words)

Time area
days, hours, & frequency histogram (grayscale, white-high)

Note inconsistent use of color coding

MieLog Close-up View

Change in color coding?



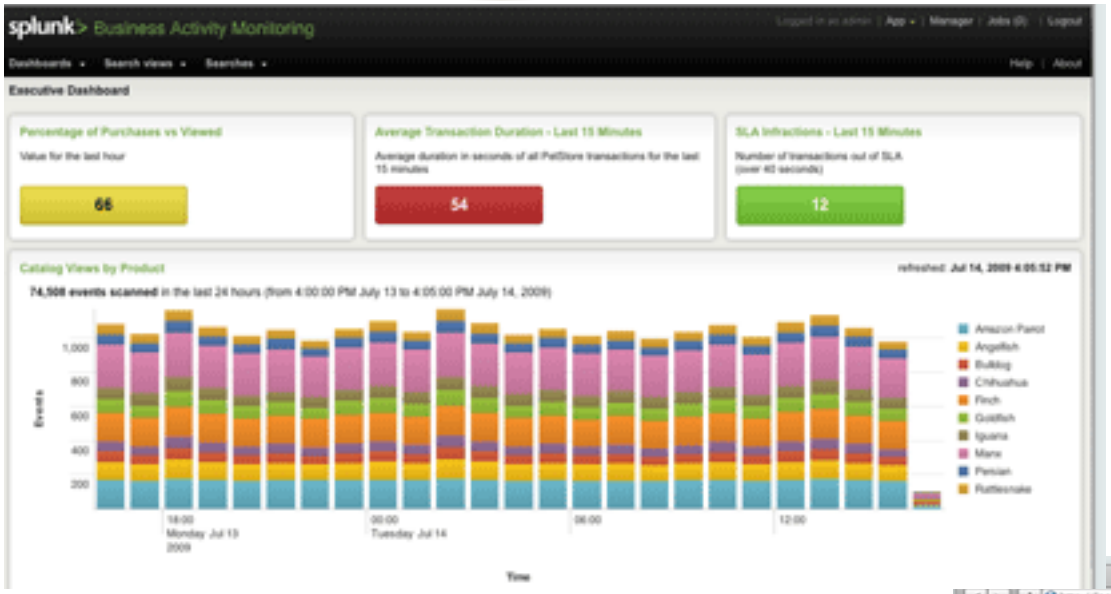
MieLog Interactions

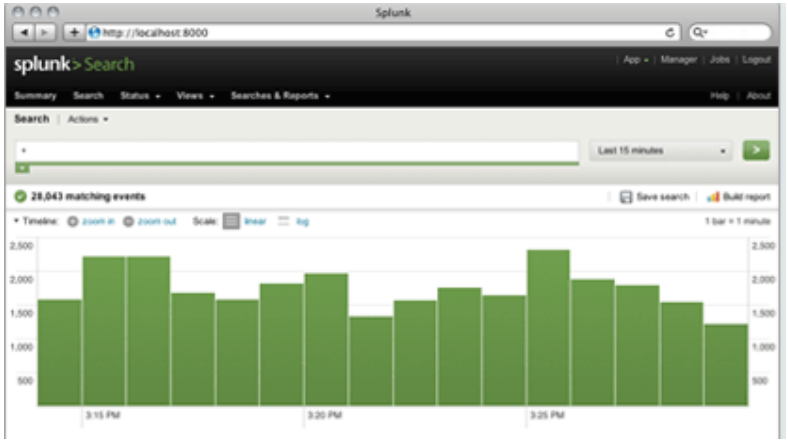
- Tag area
  - Click on tag shows only those messages
- Time area
  - Click on tiles to show those times
  - Put line on histogram to filter on values above/below
- Outline area
  - Can filter based on message length
  - Just highlight messages to show them in text
- Message area
  - Can filter on specific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

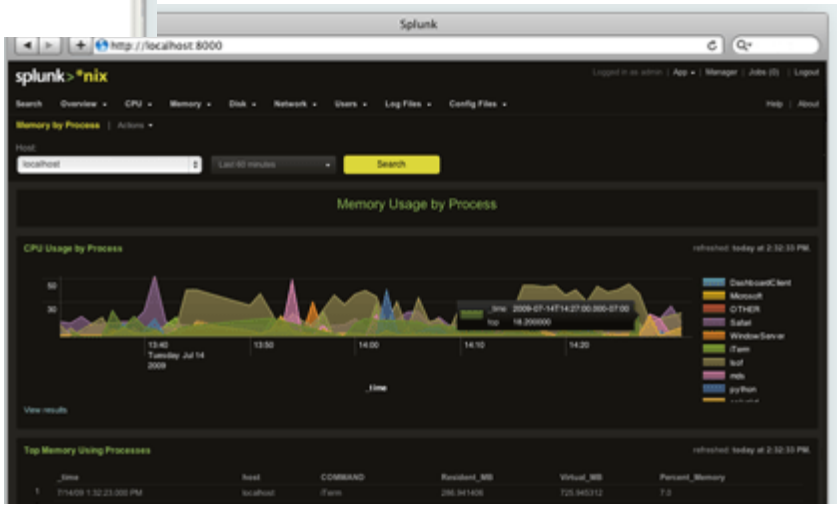
MieLog Pros/Cons

- Strengths and Weaknesses?
- Usable in other domains?

Commercial Log Inspector!







http://www.splunk.com/

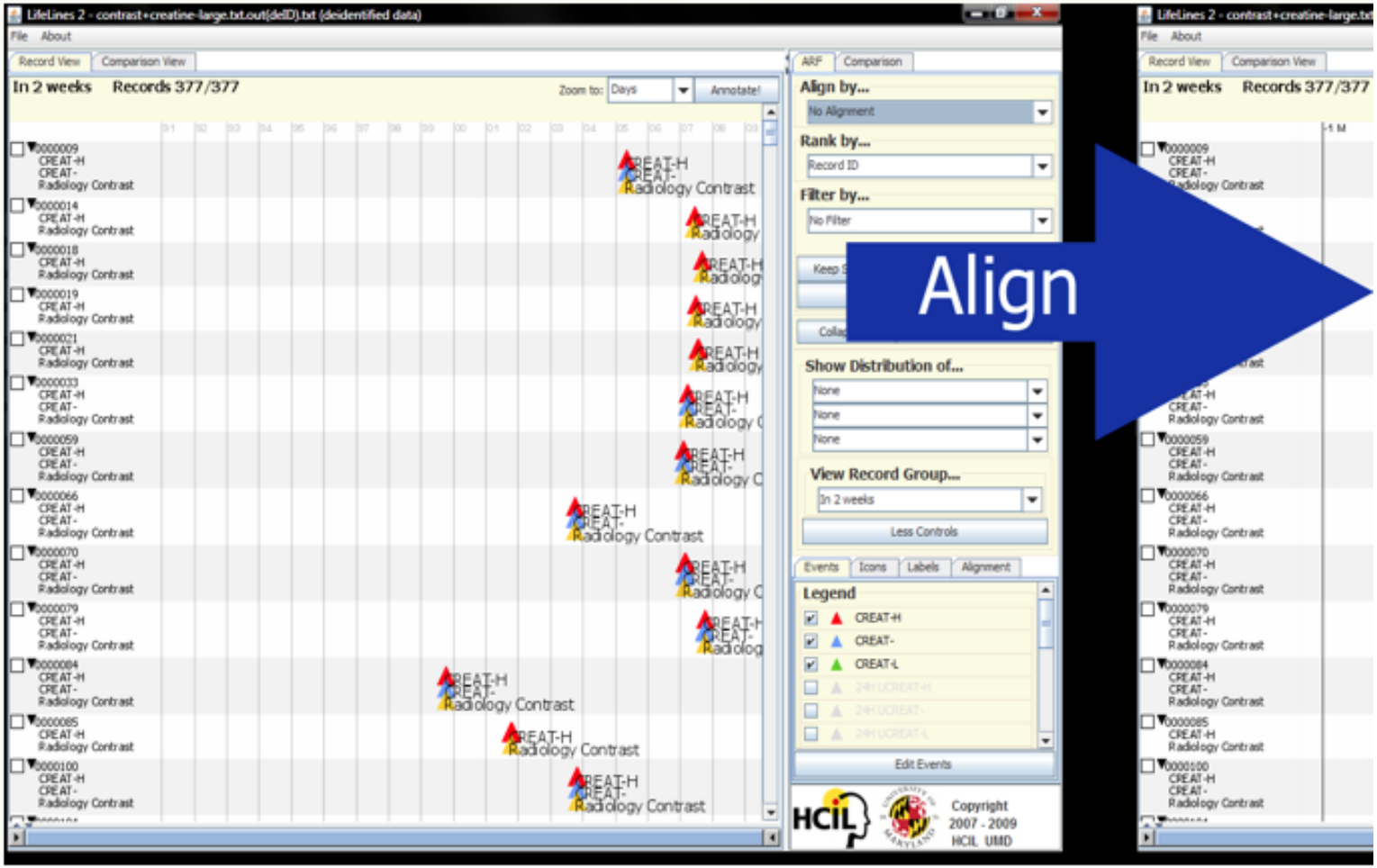
End

Example: Find Patterns in TimeLines

- Want to find patients who received contrast agent for medical imaging and within two weeks had high levels of creatinine (indicating renal problems).

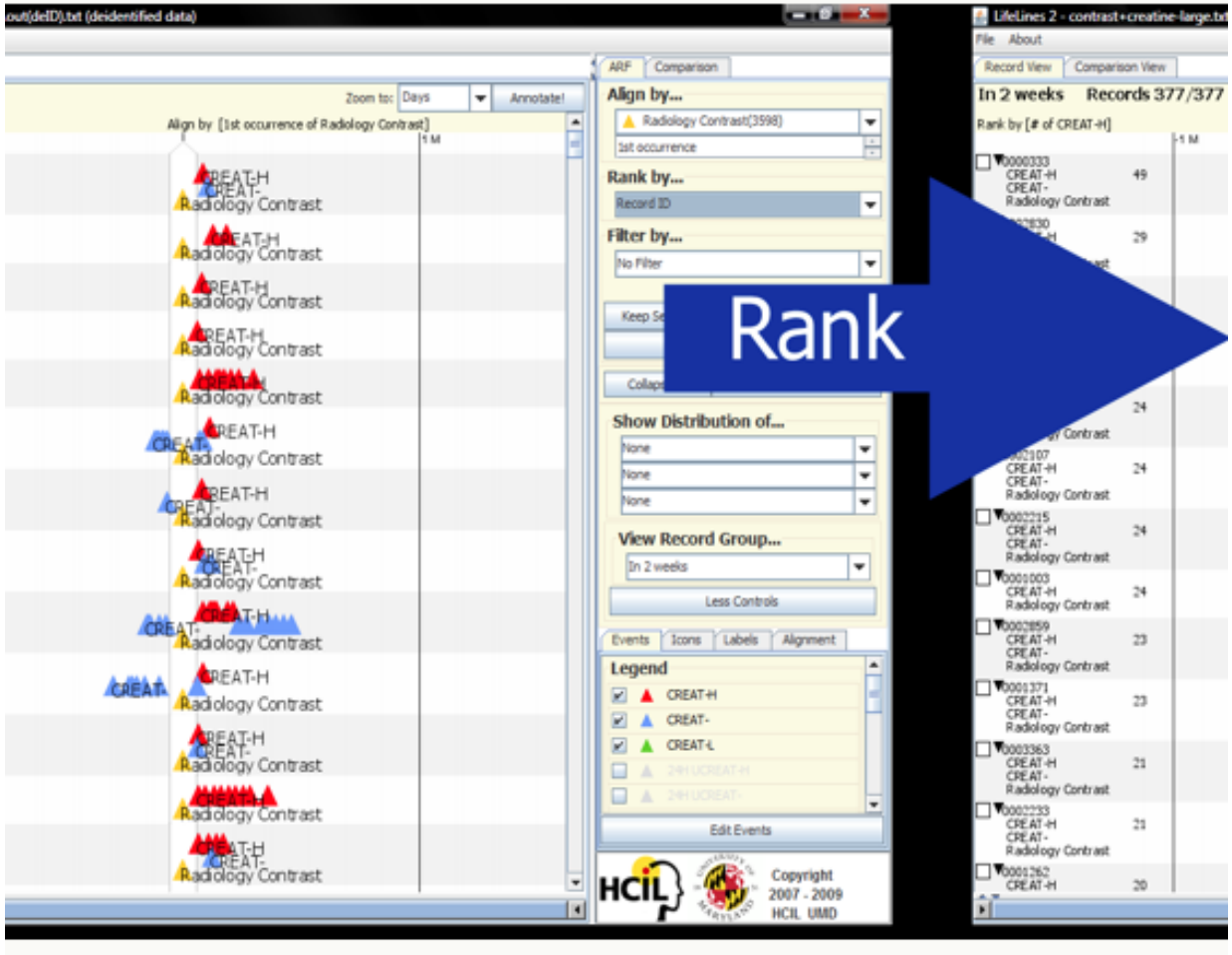
- http://www.cs.umd.edu/hcil/lifelines2/

Step 1: Unaligned Med. Records



Step 2: Align Event Times

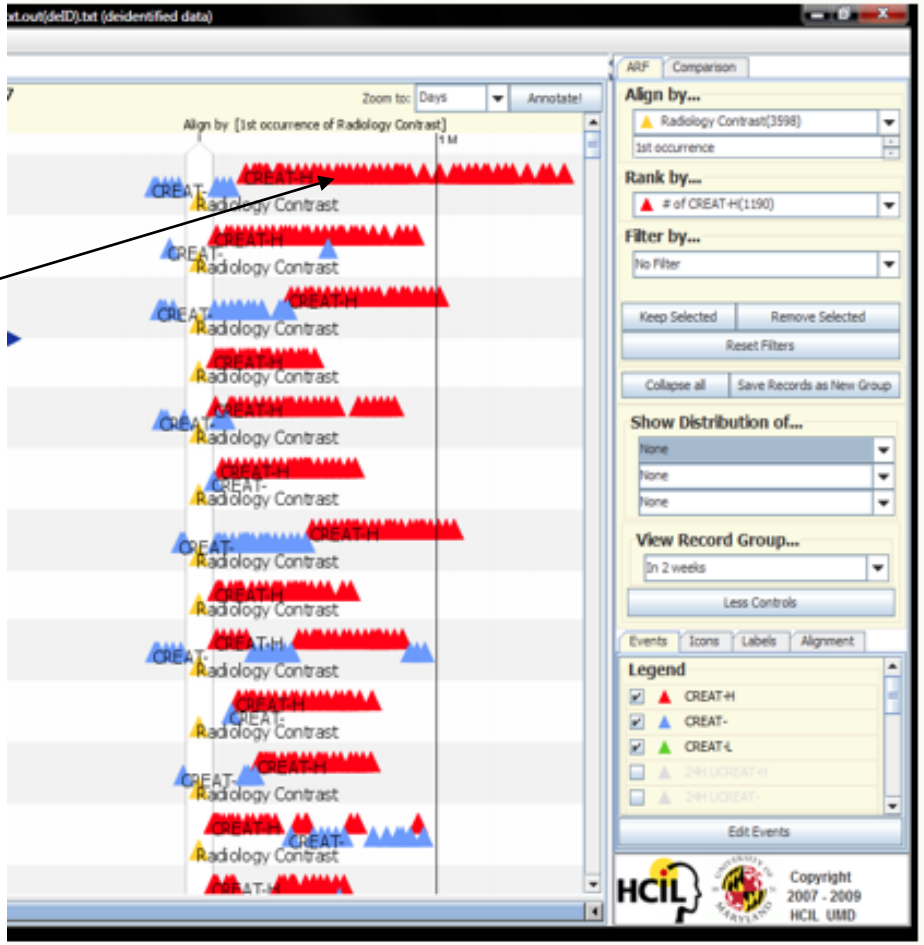
Yellow triangles (when contrast agent administered) all aligned



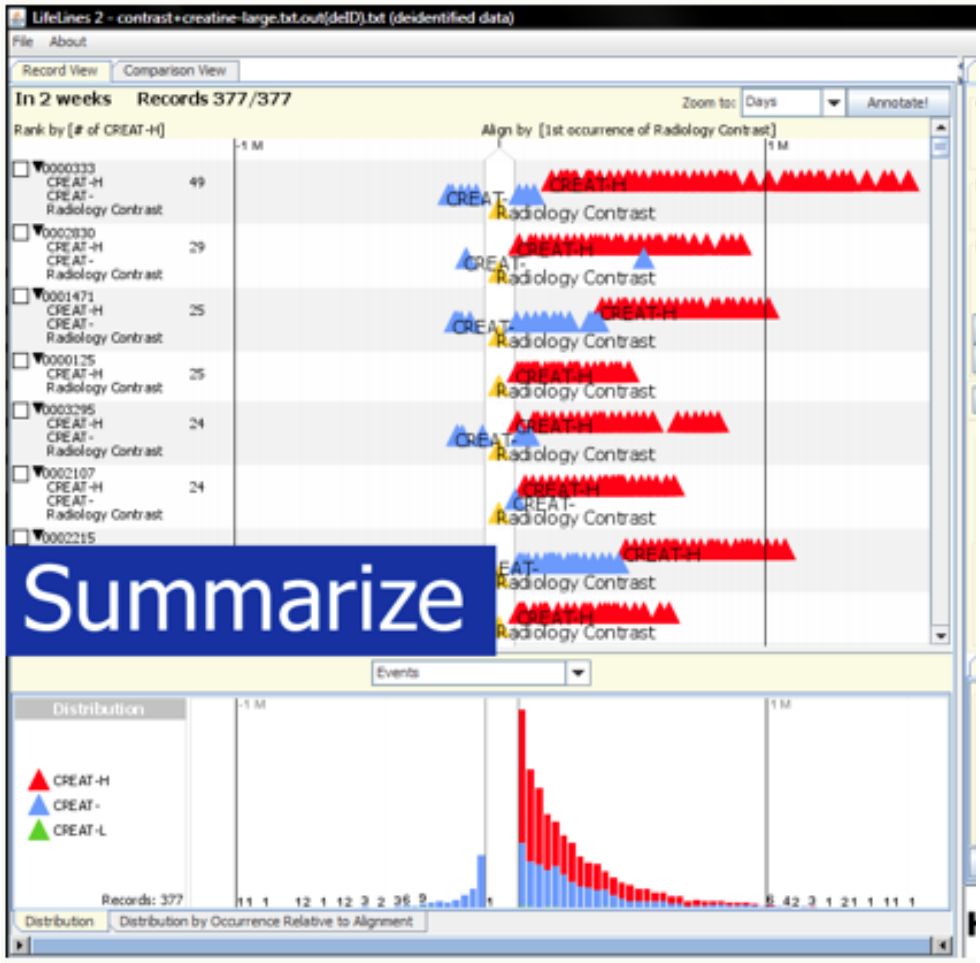
Step 3: Rank by Creatinine Level

Sort by Creatnine Level

Highest at top (red triangles)



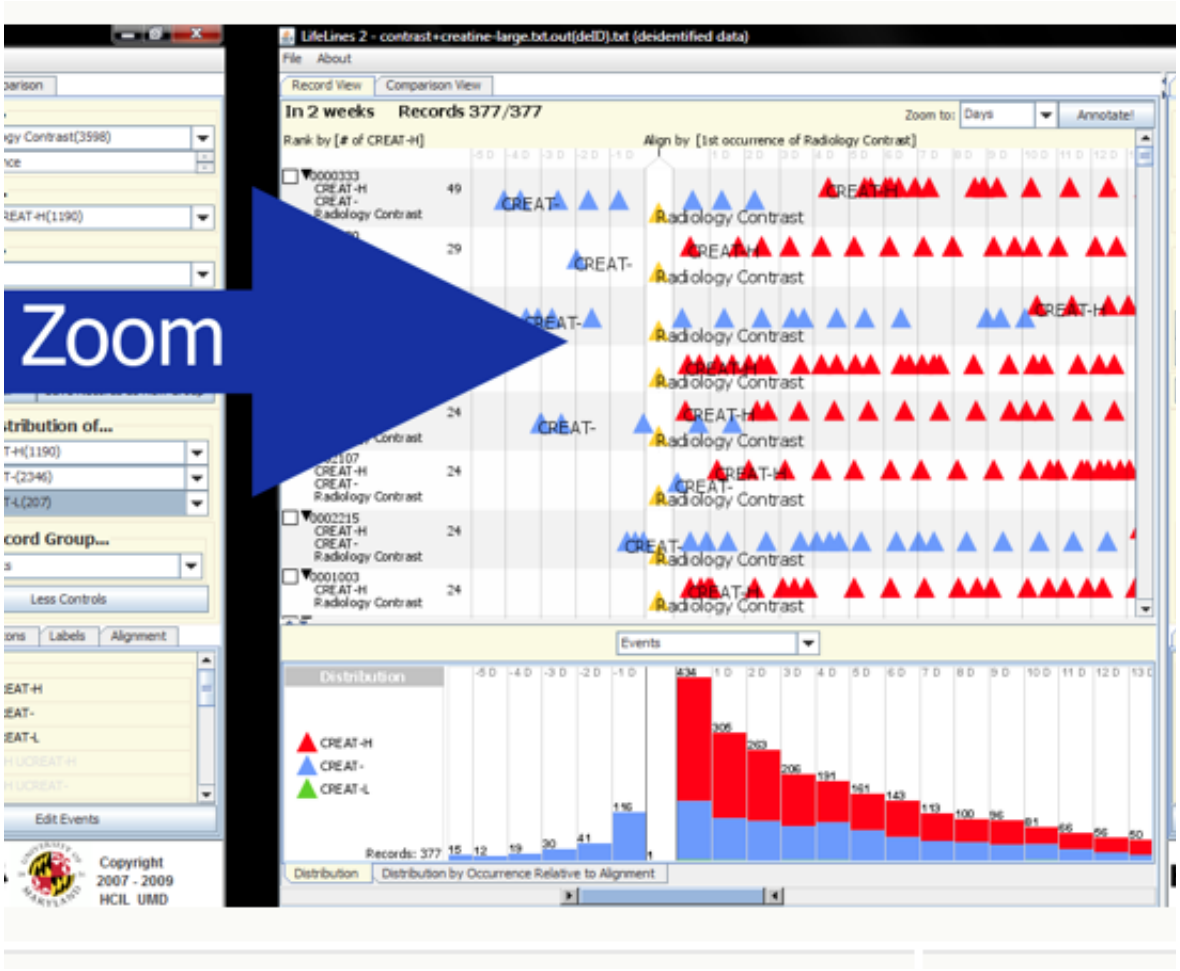
Step 4: Summarize Across Cases



Histogram across cases

Step 5: Expand Time Line

Expand time line to see more detail



Video: http://www.cs.umd.edu/hcil/lifelines2/videos/Lifelines2-%28Contrast-Creatinine%29-11-10-08_flash/full%281024x768%29/index.htm

Finding Unknown Periodicity

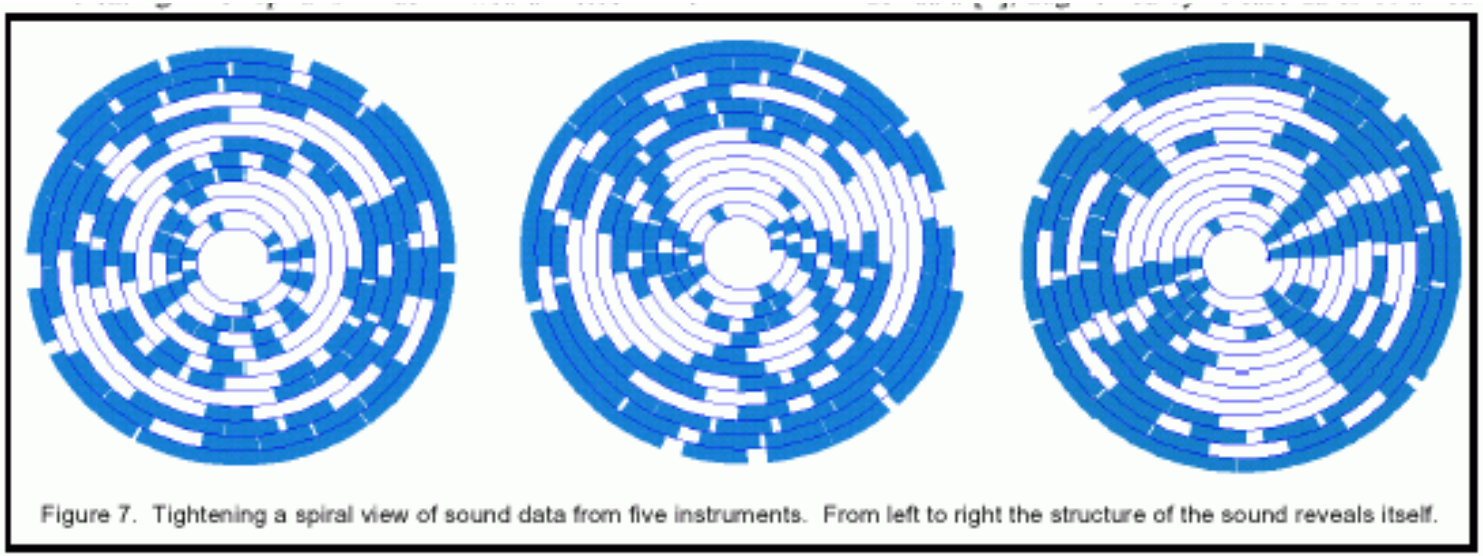


Figure 7. Tightening a spiral view of sound data from five instruments. From left to right the structure of the sound reveals itself.

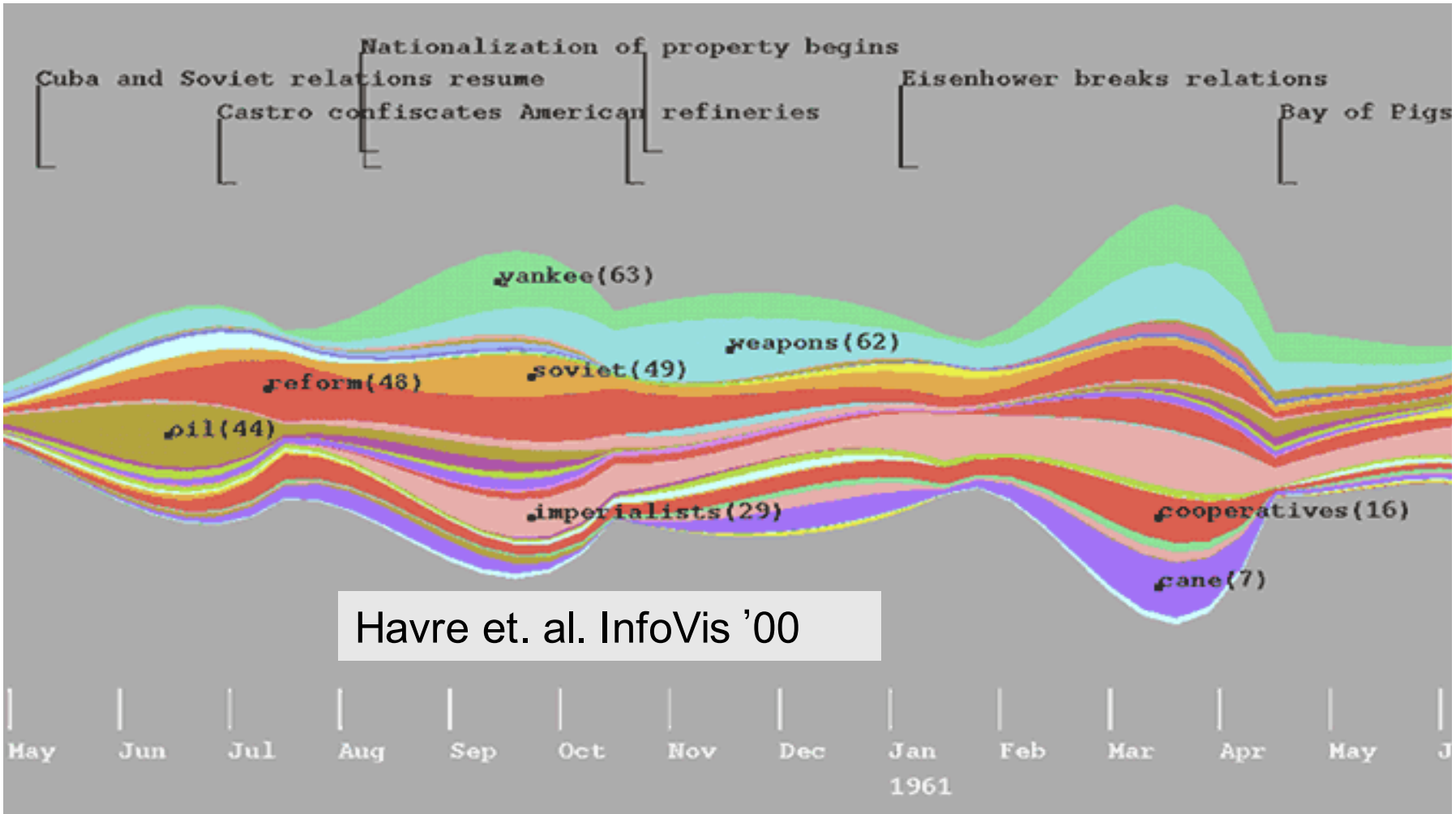
- Find unknown periodicity by interactively tightening/loosening the spiral
  - Mathematical tools can also find periodicity ☺
- What if is no periodicity?

Example: Text Themes over Time

- Consider a set of speeches or documents over time
  - How represent the flow of ideas and concepts over time?
    - Was an InfoVis student contest
- How would you do this?
- Any related concepts?

Theme River

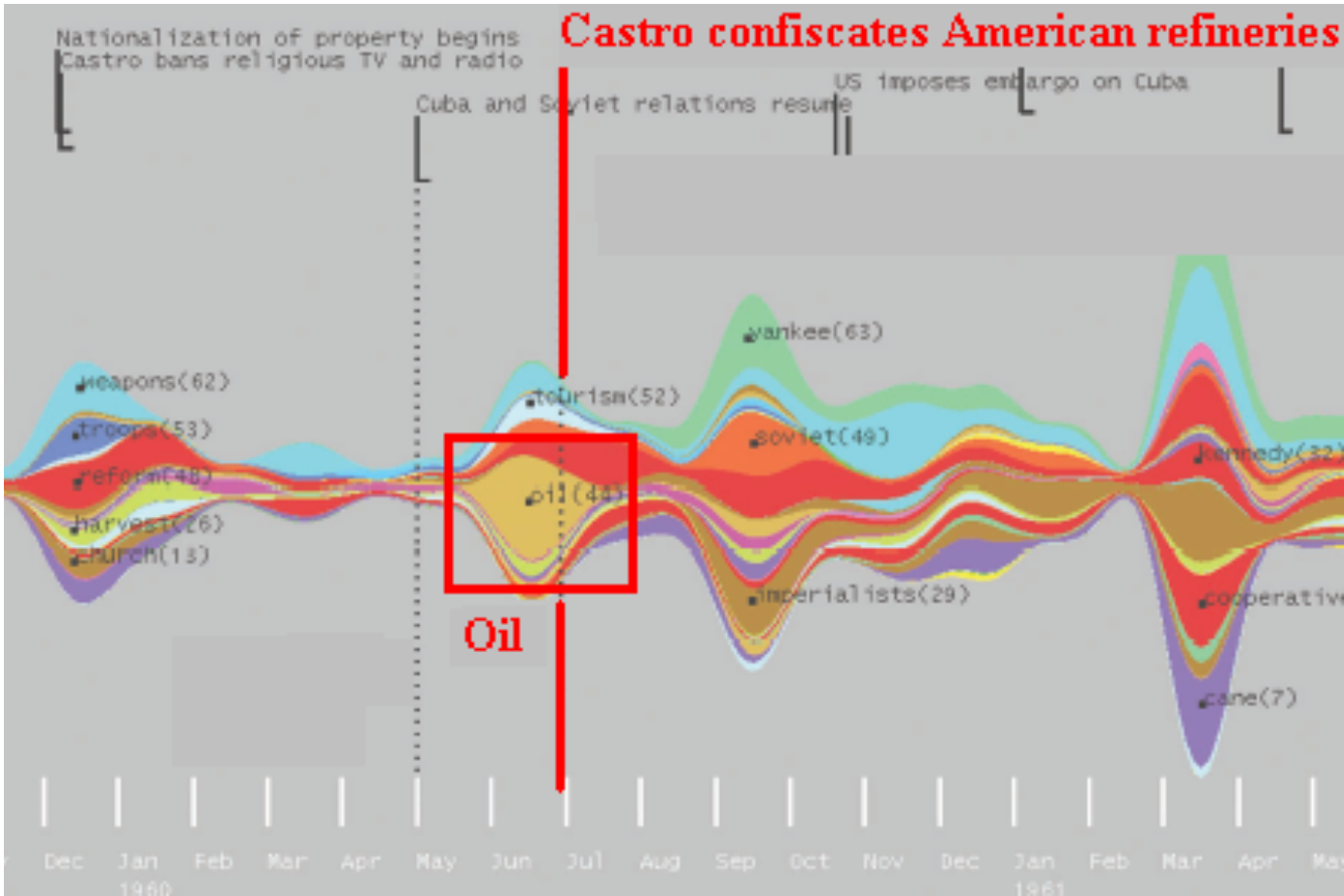
- Each ‘theme’ from news stories is a ‘river’



Havre et. al. InfoVis ’00

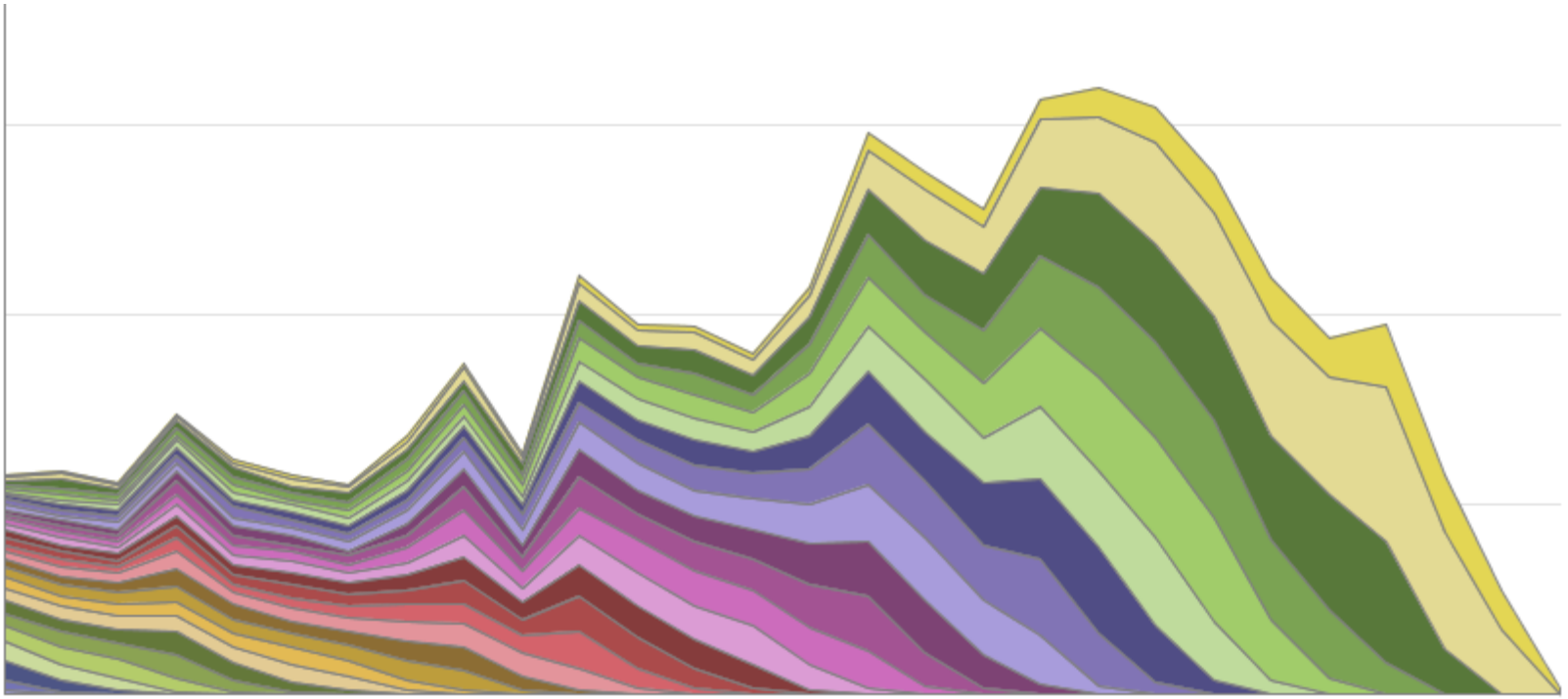
Theme River

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



Theme River = Stacked Graph?

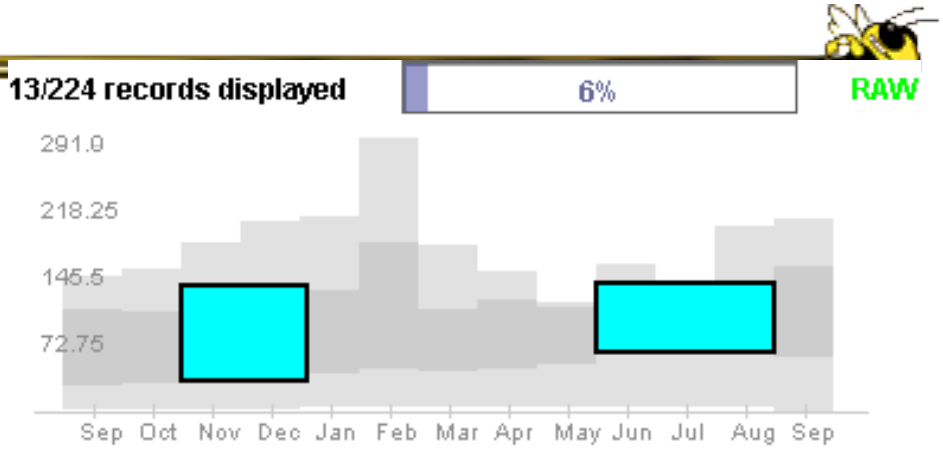
- Yes, with a differential vertical offset
  - Makes a world of difference!



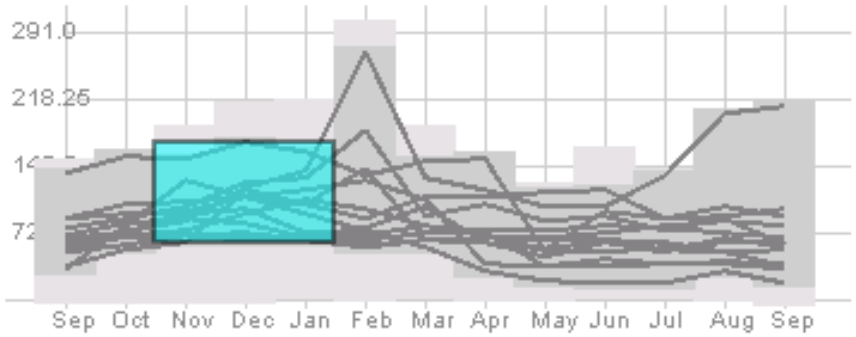
Example: Querying

- Most systems focuse on visualizing and navigating time-series data.
- What about *querying*?

TimeFinder



Can create rectangles that function as matching regions



Light gray is all data's extent

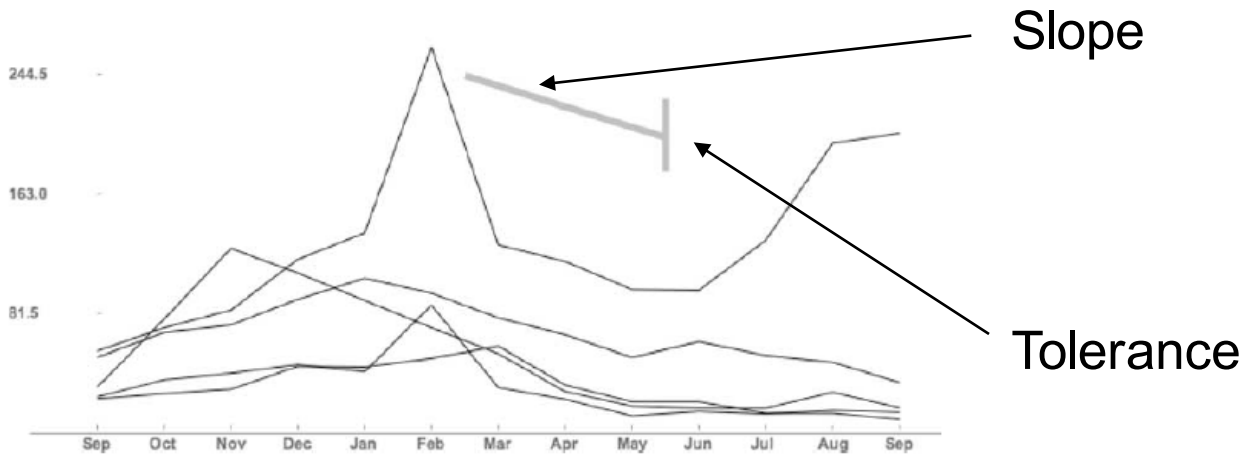
Darker grayed region is data envelope that shows extreme values of queries matching criteria

Hochheiser & Shneiderman
Proc. Discovery Science '01
Info Vis '04

Multiple boxes are "anded"

TimeFinder – Other Capabilities

- Search for matches based on angle (slope) ± tolerance



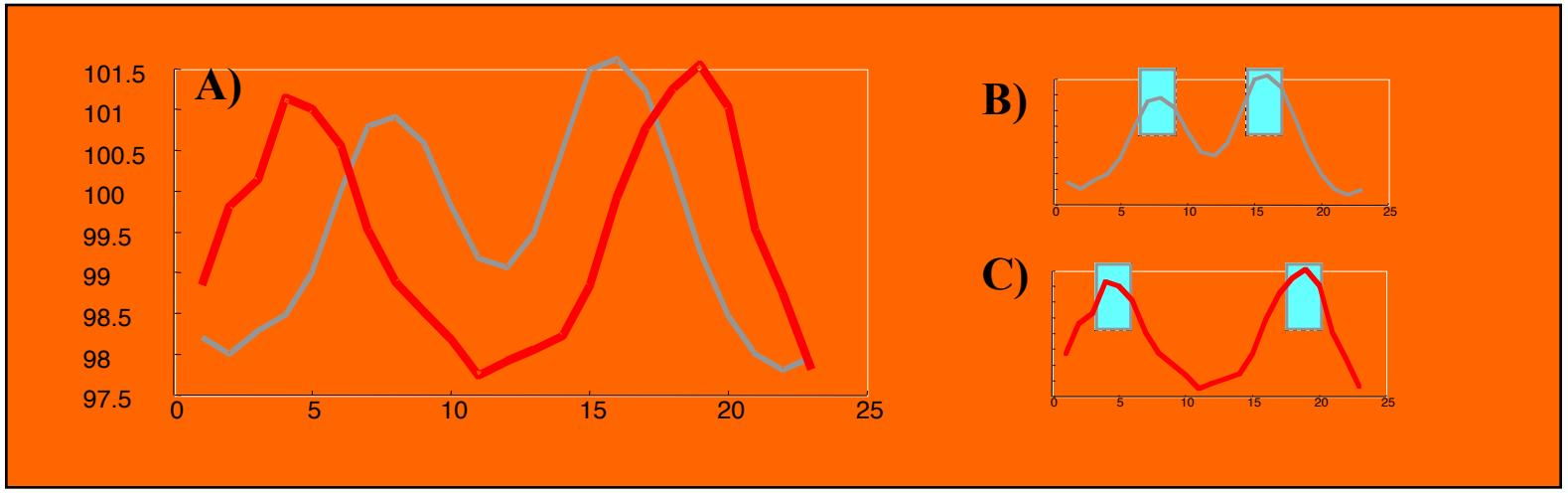
- "Leaders and laggards"
  - Find same patterns but shifted in time

TimeFinder Limitations

- What are they?
  - (Hint, there's a fundamental problem!)

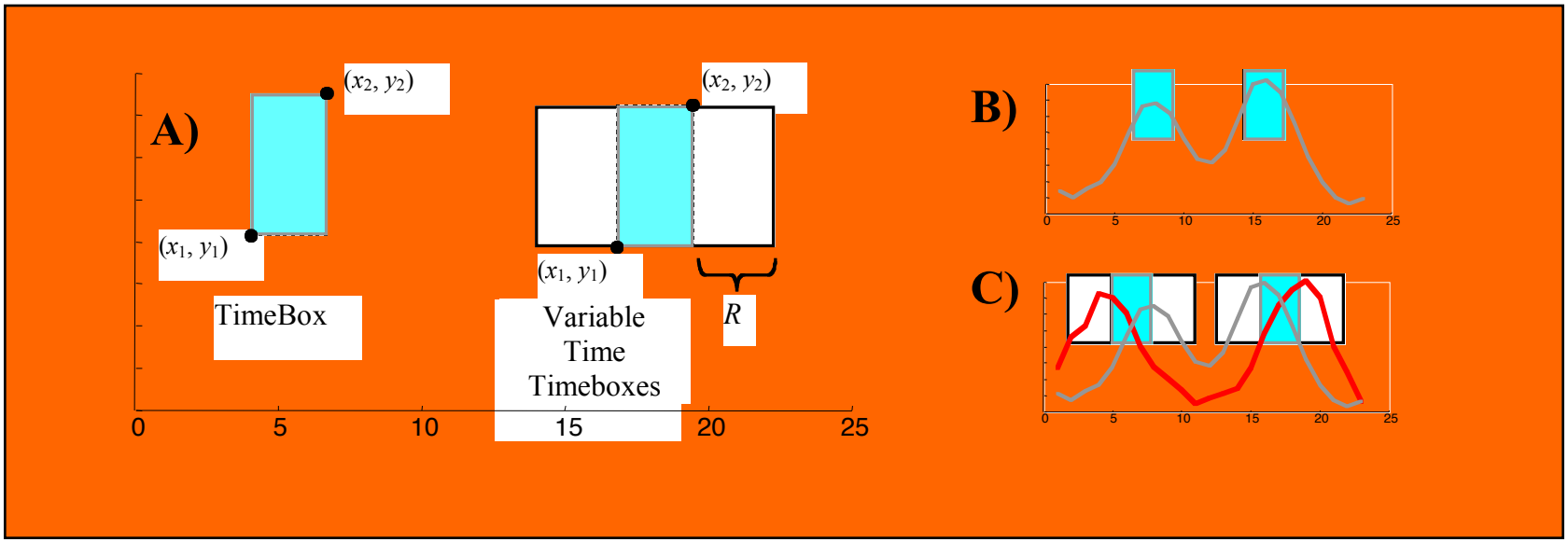
TimeFinder Limitations

- Hodgkins patients exhibit double spike in temperature…
  - But with differing times between spikes



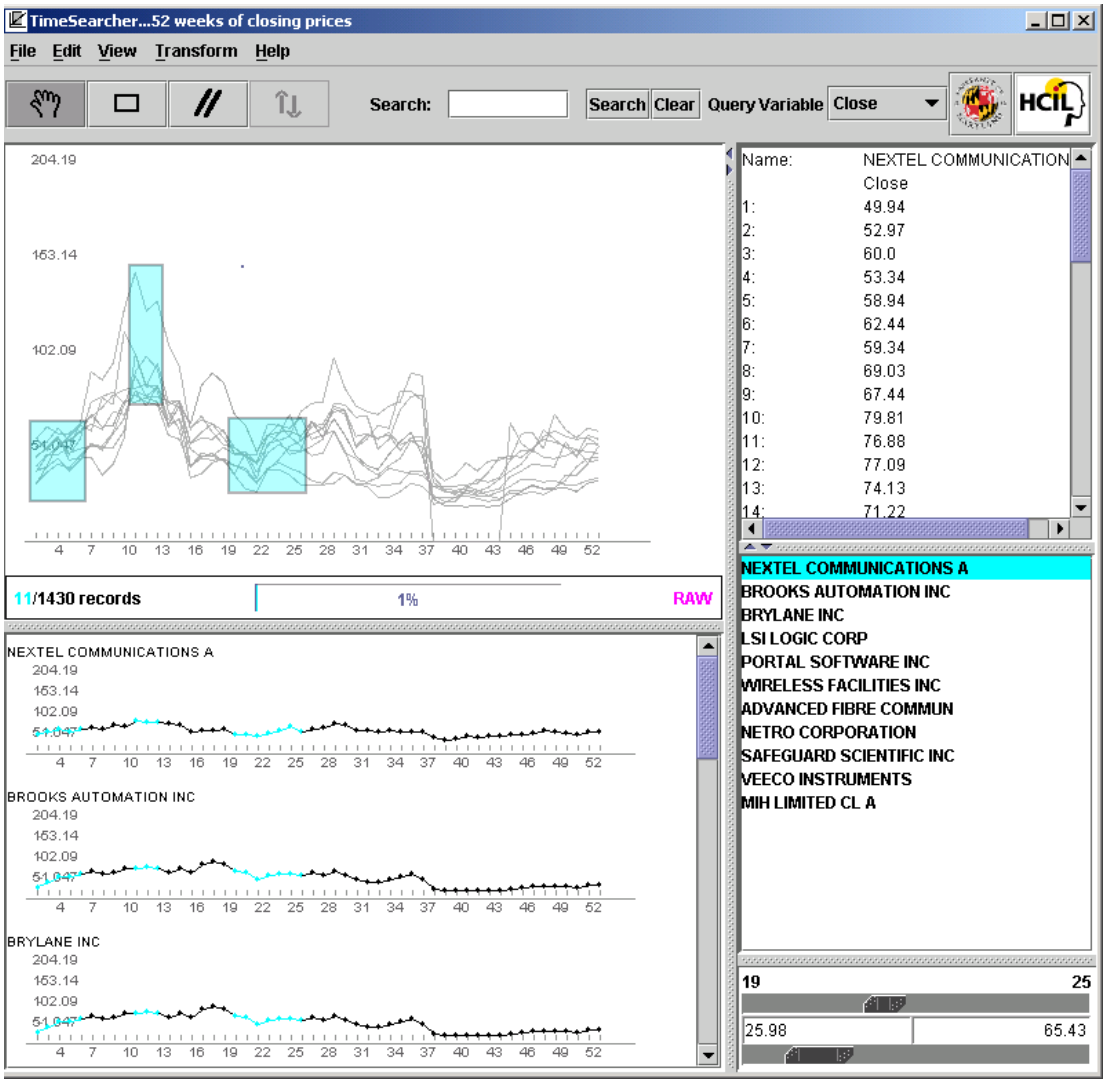
Example:TimeSearcher

- Overcomes TimeFinder limitation
- Allow time boxes with deltas on each side (TimeSearcher) Video: 2005_timesearcher.mpg



Buono, Plaisant, Simeone, Aris, Shneiderman, et.al, *Similarity-Based Forecasting with Simultaneous Previews: A River Plot Interface for Time Series Forecasting*, 11th International Conference on Information Visualization, 2007

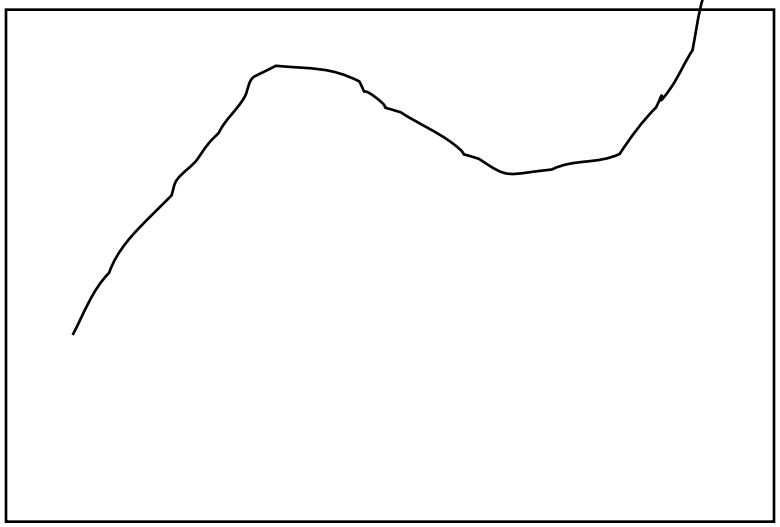
TimeSearcher



Example: Query by (Graphical) Example

- Sketch the Query
- Also discussed in Interaction Lecture
  - Specify a timeline query by drawing a rough pattern for it and the system returns similar matches. (Wattenberg, CHI '01)

User-drawn query



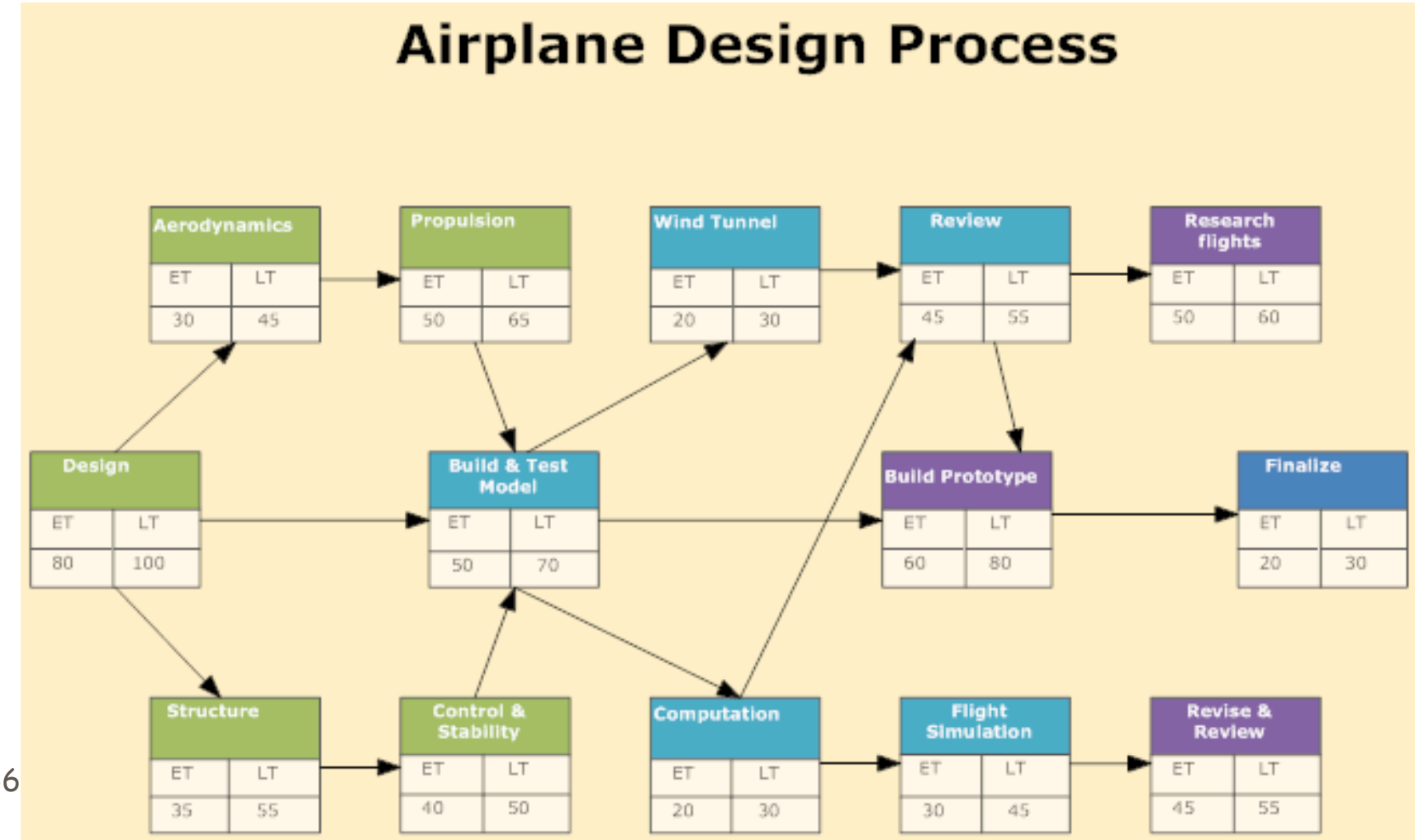
Example: Project Management

- Project (write software, design/build plane, hire new person) involves
  - Multiple steps
  - Spread over *time*
  - Some steps depend on other steps
- How can we plan/manage project?

- With PERT/Gnatt/CPM charts, of course ☺

PERT Chart - Time Dependencies

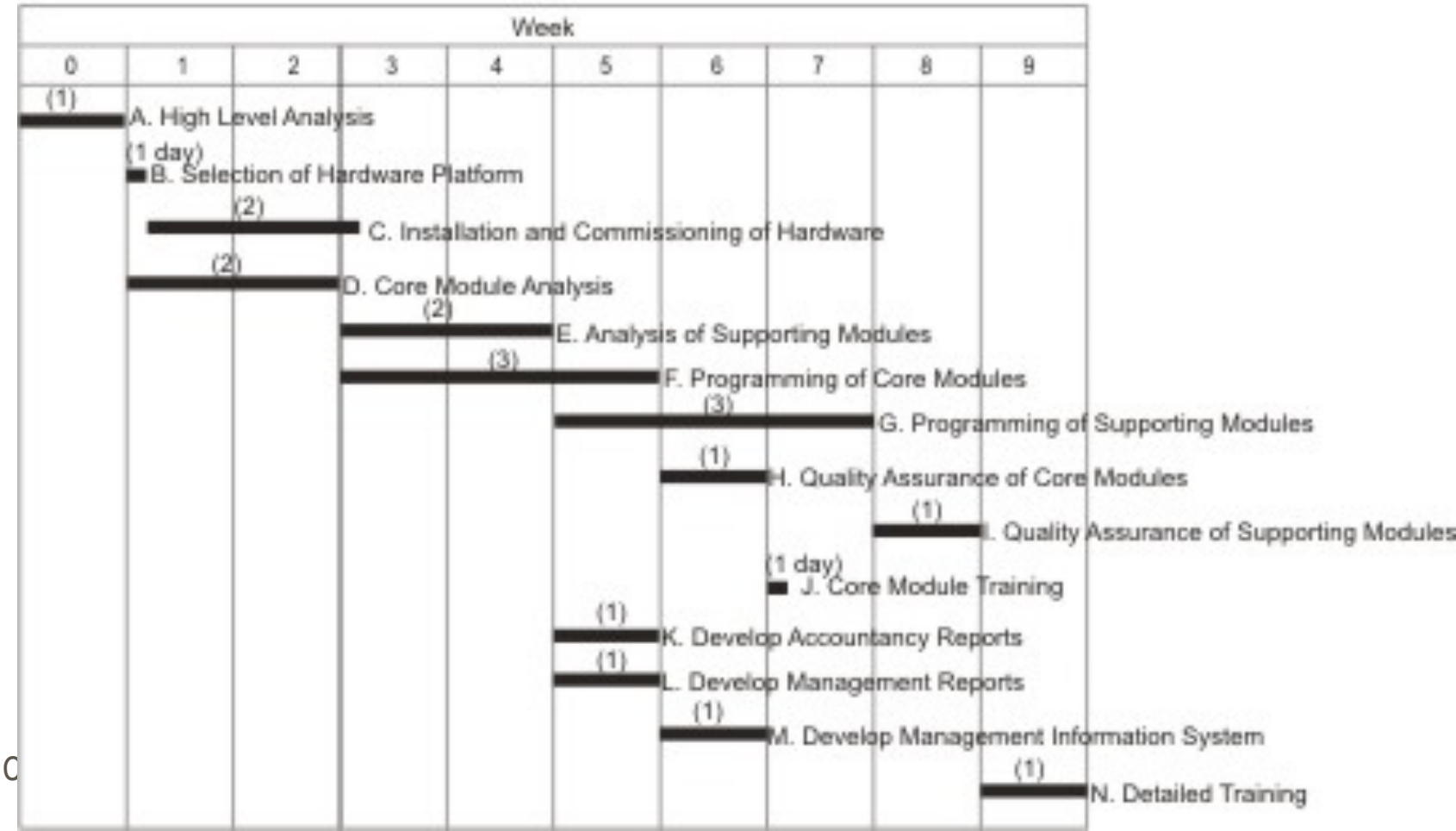
- PERT = Project Evaluation and Review Technique



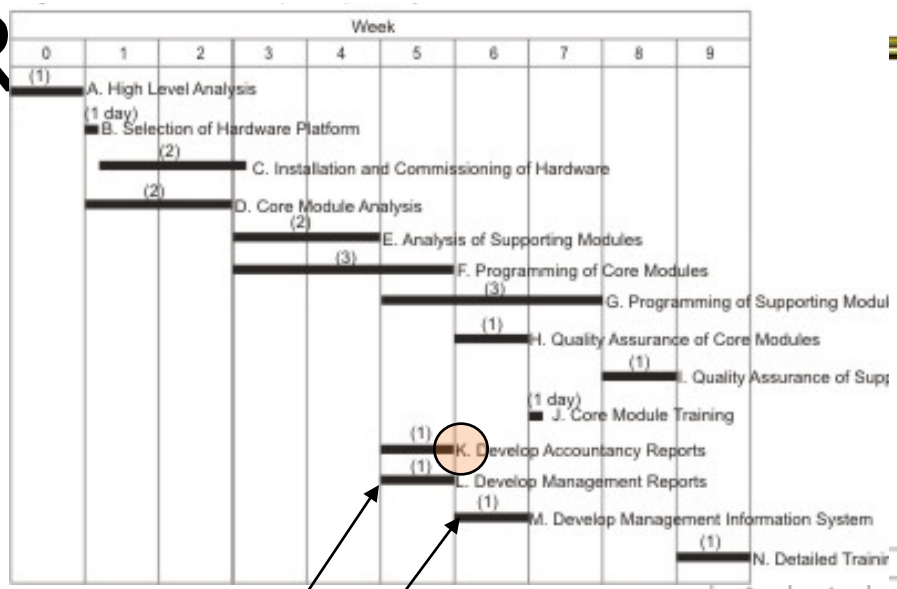
PERT leads to Gnatt

- Dependencies on time-line

Figure 2: Draft Gantt Chart: Example Computer Project

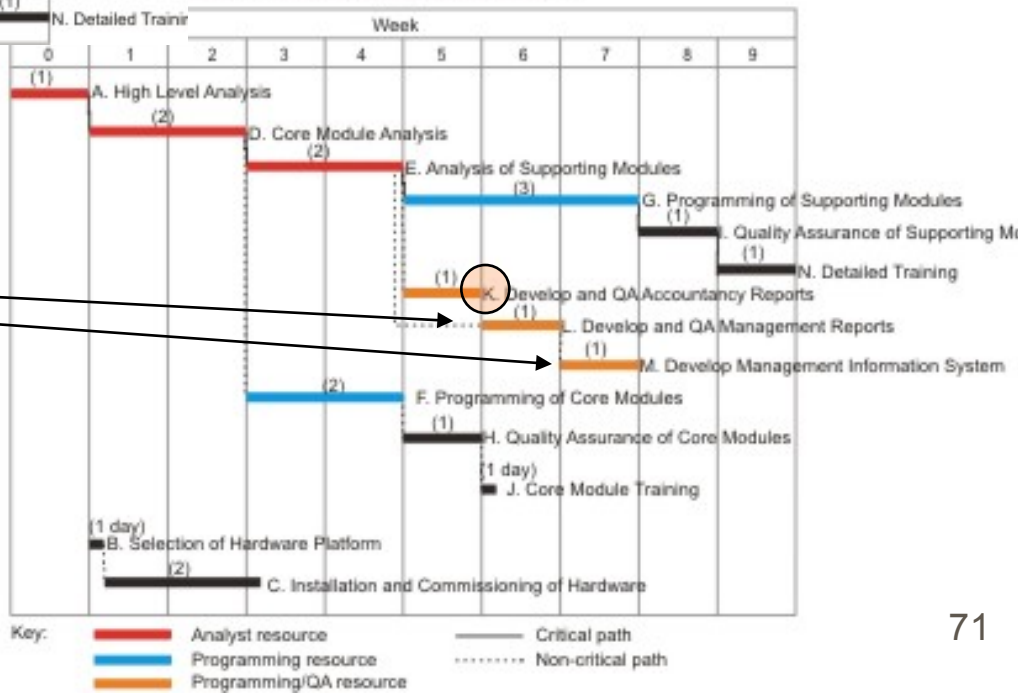


R



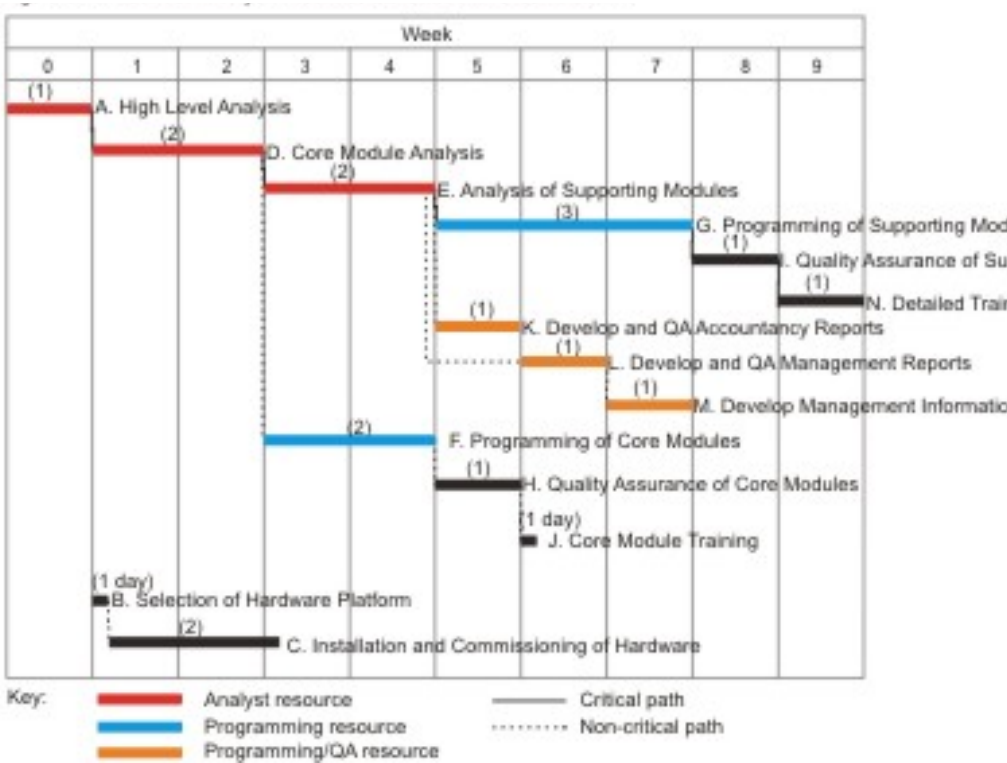
Shows
- Resources (colors)
- Critical path activities

Activities L &M delayed by a week so same person can do activity K



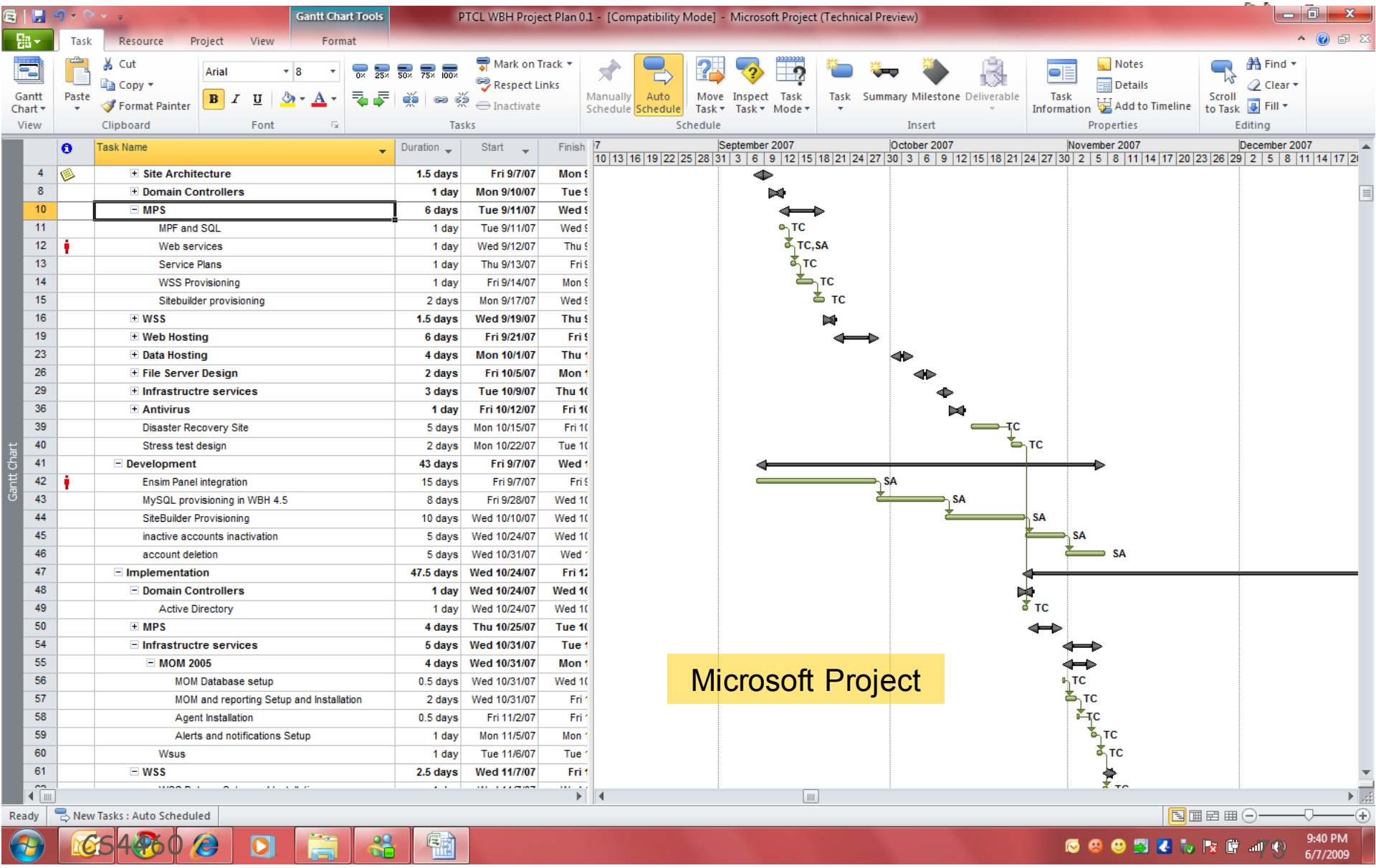
Revised Gnatt Chart Tells Us That

- Ten week completion ***IF*** all goes well
- Completion in ten weeks requires:
  - 1 analyst for the first 5 weeks
  - 1 programmer for 5 weeks starting week 4
  - 1 programmer/QA expert for 3 weeks starting week 6
- Analysis, development and testing of supporting modules must be on time.
- Hardware not time-critical but must complete before Core Module Training start

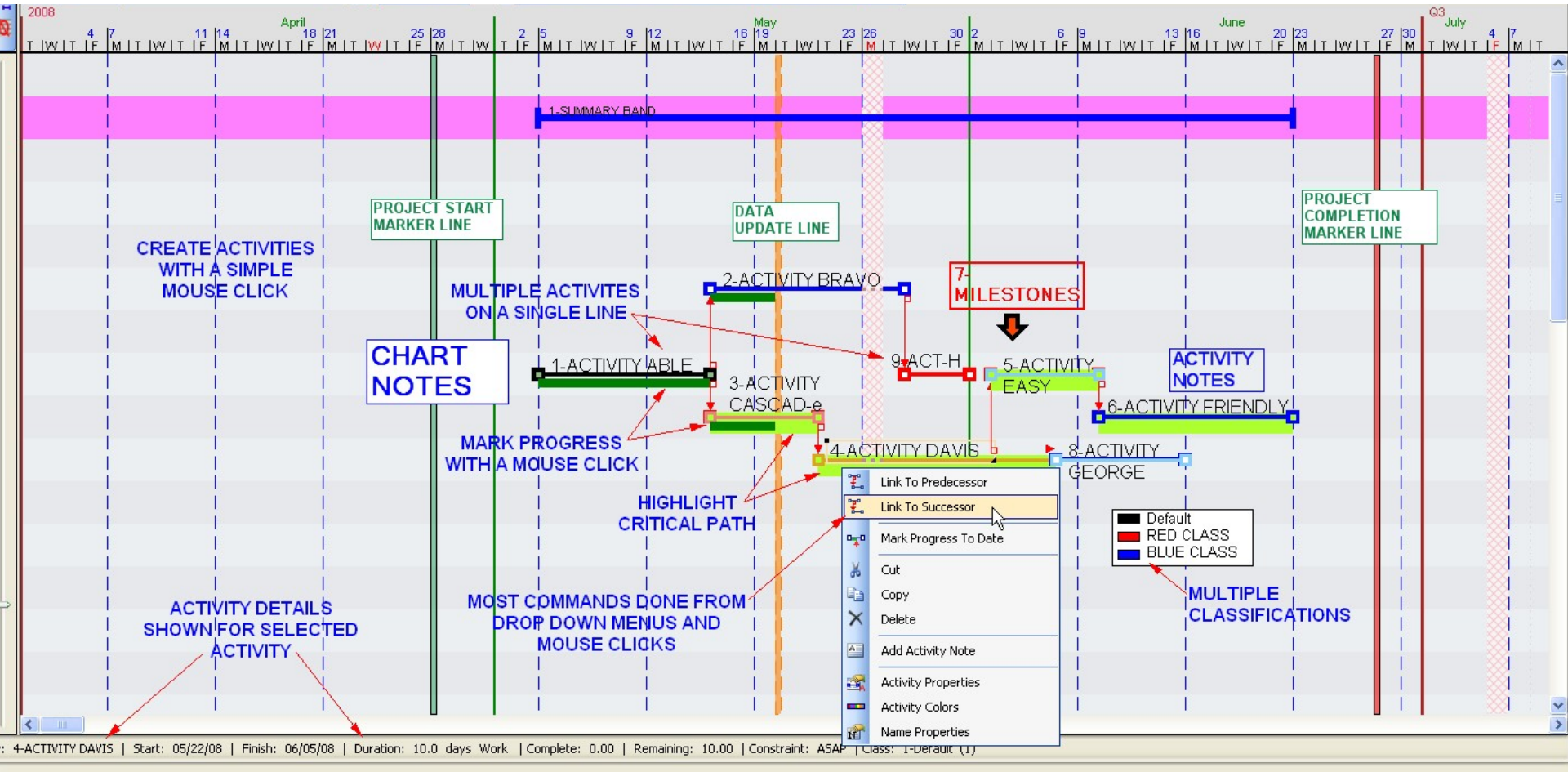


From http://www.mindtools.com/pages/article/newPPM_03.htm

Many Project Mg't Products!



One from Georgia Tech



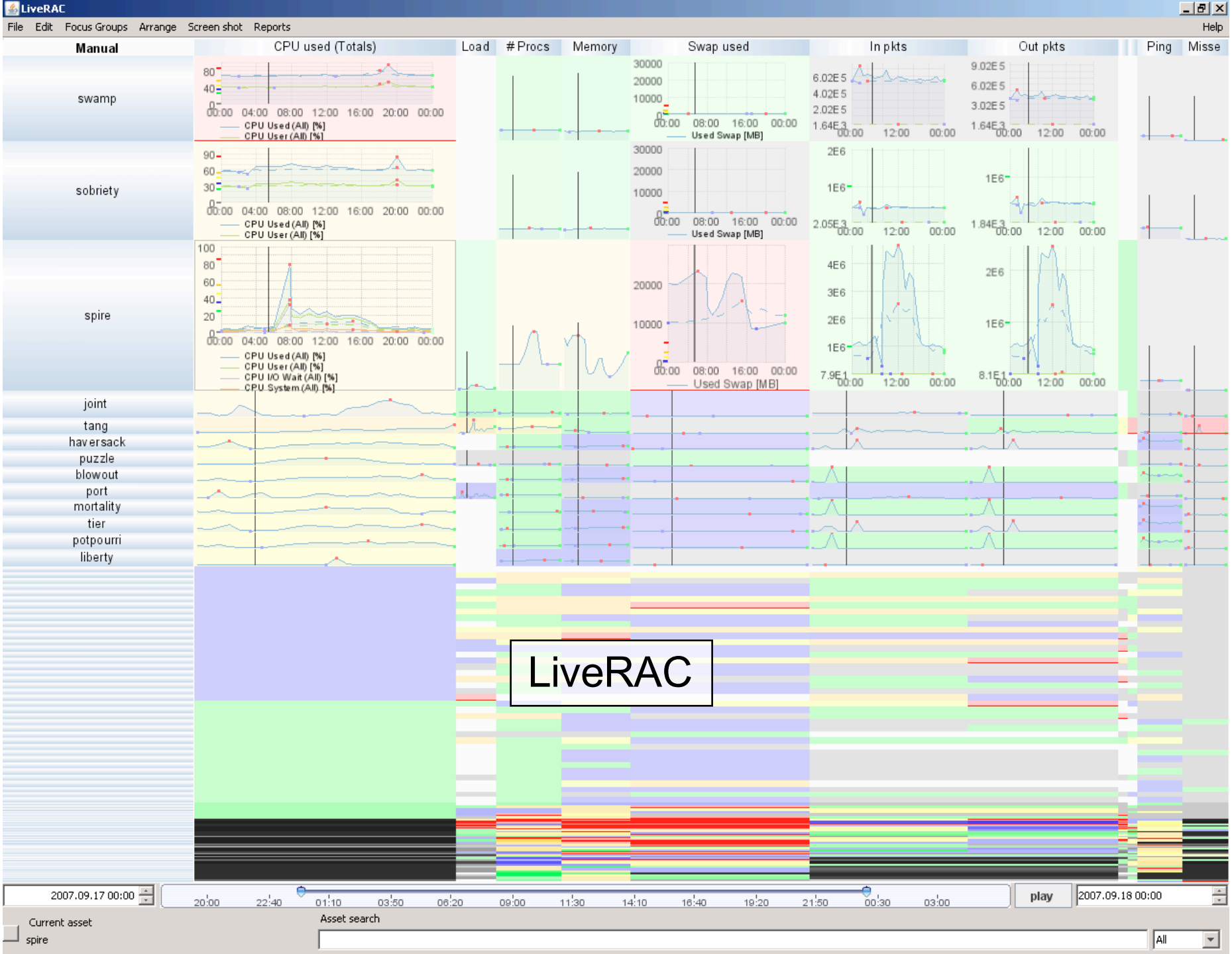
CASCADE-e http://www.cascad-e.net/Index.php

Example: Monitor Computer Systems

- This system (LiveRAC) Supports review of continuous data and real-time monitoring

- MieLog example is for after-the fact inspection of discrete event data
- LiveRAC was a 2.5 year, multi-person project

McLachlan, Munzner, EKoutsofios, North. *LiveRAC - Interactive Visual Exploration of System Management Time-Series Data*, CHI'08


LiveRAC
File
Edit
Focus Groups
Arrange
Screen shot
Reports
Help
Manual
CPU used (Totals)
Load
Procs
Memory
Swap used
In pkts
Out pkts
Ping
Misse
swamp
sobriety
spire
joint
tang
haversack
puzzle
blowout
port
mortality
tier
potpourri
liberty
CPU Used (All) [%]
CPU User (All) [%]
CPU I/O Wait (All) [%]
CPU System (All) [%]
Used Swap [MB]
LiveRAC
2007.09.17 00:00
play
2007.09.18 00:00
Current asset
spire
Asset search
All

LiveRAC: Interactive Visual Exploration of System Management Time-Series Data

LiveRAC Design Principles

- Involve users early
- Use spatial position as perceptual cue
  - Related items are proximate (grouped)
- Use side-by-side "small multiples" to minimize load on short term memory
- Link multiple views
- Animate changes – no abrupt transitions
- Immediate feedback
- Use graph styles with which users are already familiar
- Overview first, zoom and filter, details on demand
- Assertion: Several levels of detail side-by-side
  - Why given as assertion rather than principle?
  - Seems like a principle to me ☺ - giving context along with detail