Visualizing Cyber Security: Usable Workspaces

Glenn A. Fink, Christopher L. North, Alex Endert, Stuart Rose

chci center for human computer interaction
virginiatech

What did we do?

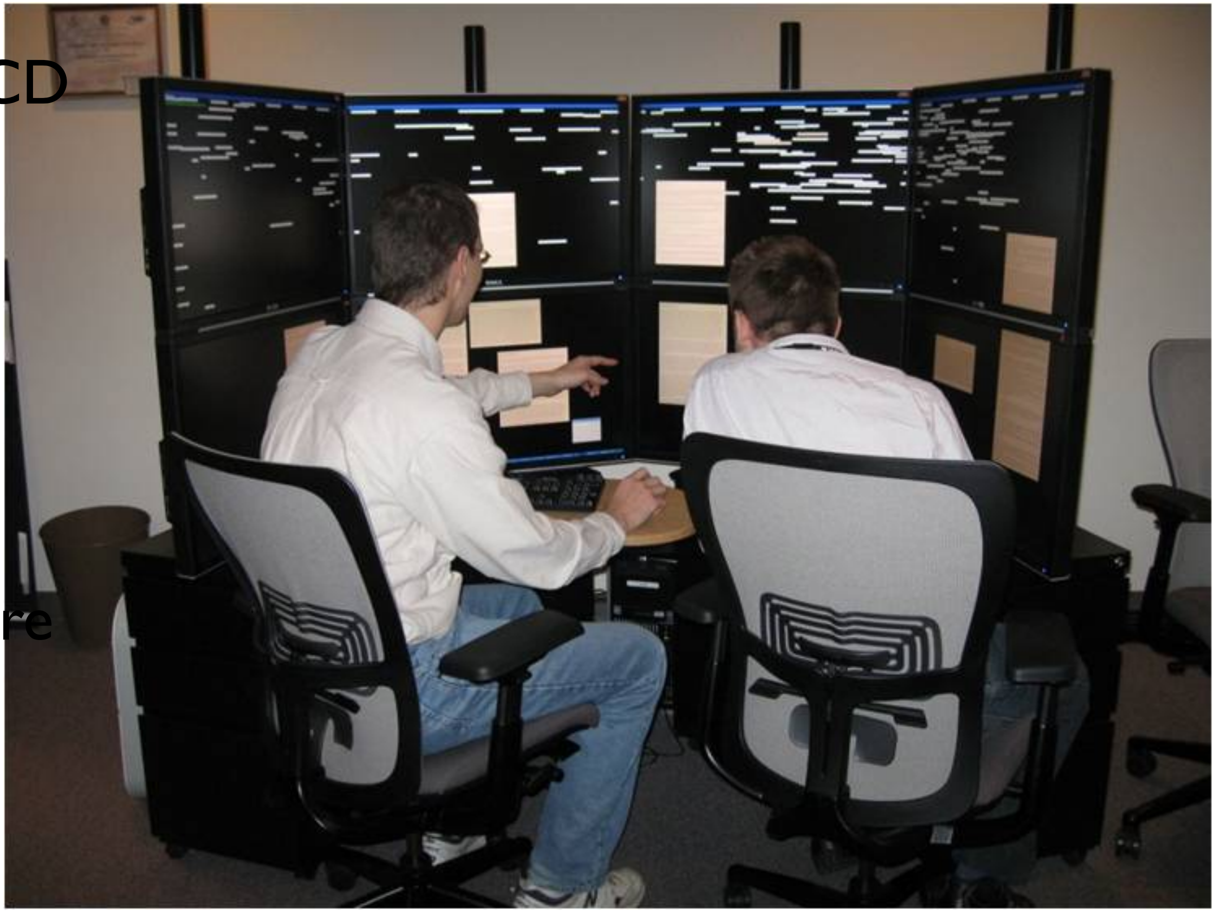
- How can we design visual workspaces that aid Cyber Security?
  - Tons of data?
  - Lots of windows and tools?

- Why don't we give the user more space?

Let's give the *user* more *space*!

Large, High-Resolution Displays

- (8) 30-inch high-res LCD Panels
- 33 Megapixel total resolution (10,240 x 3,200)
- "Single PC" Architecture
- Curved for optimal individual use



Methods

1. Interviews (8 professional cyber analysts)
   - Typical tasks and data?
   - Work style?
     - E.g., Collaboration? Multi-tasking? Time constraints?
   - Office setup
   - What does your finished analysis product contain?

2. User study (4 cyber analysts, VAST09 dataset)
   - 2 sources of data: Building/room access records (Prox) and simulated computer network flows
     - HINT: making connections between the sources is key! ☺
   - Tools provided: Excel, Spotfire, Windows XP

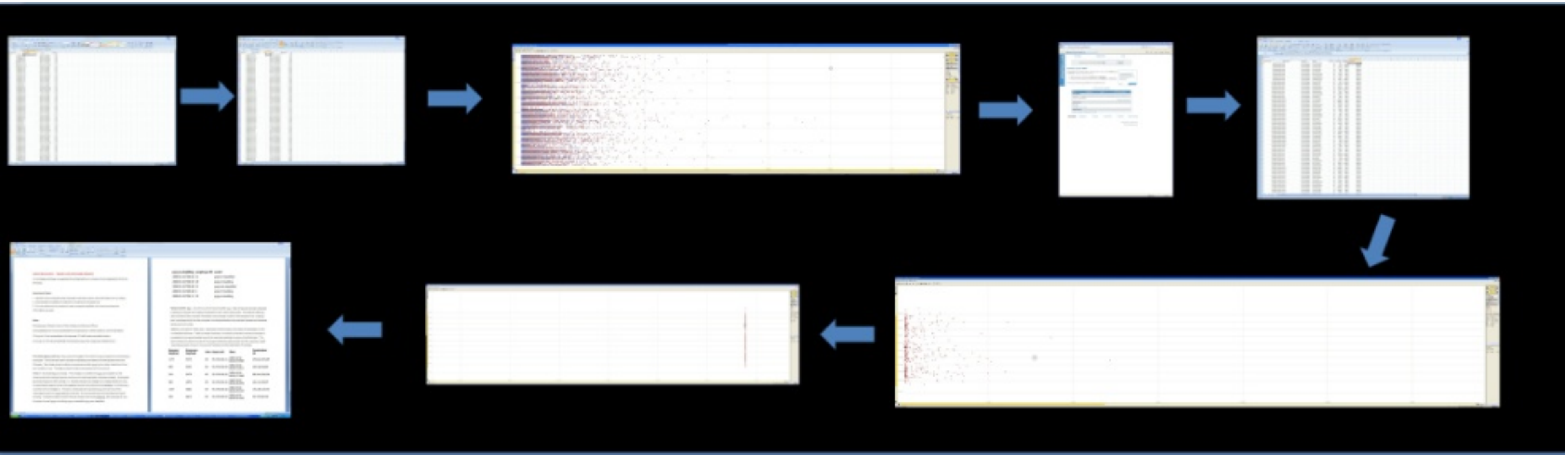
3. Feedback from the analysts on our prototypes.

Key Ethnographic Discoveries

1. Data sources reside in separate tools
2. Analysts spend much time doing low-level tasks
3. They distrust visualizations
4. They are on a "Quest for a Query"
5. Cyber data comes in huge volumes and velocities
6. Cyber data comes from many diverse sources
7. Analysts seek direct access to the data
8. Analysts routinely conduct a large number of tasks in parallel (multi-tasking)

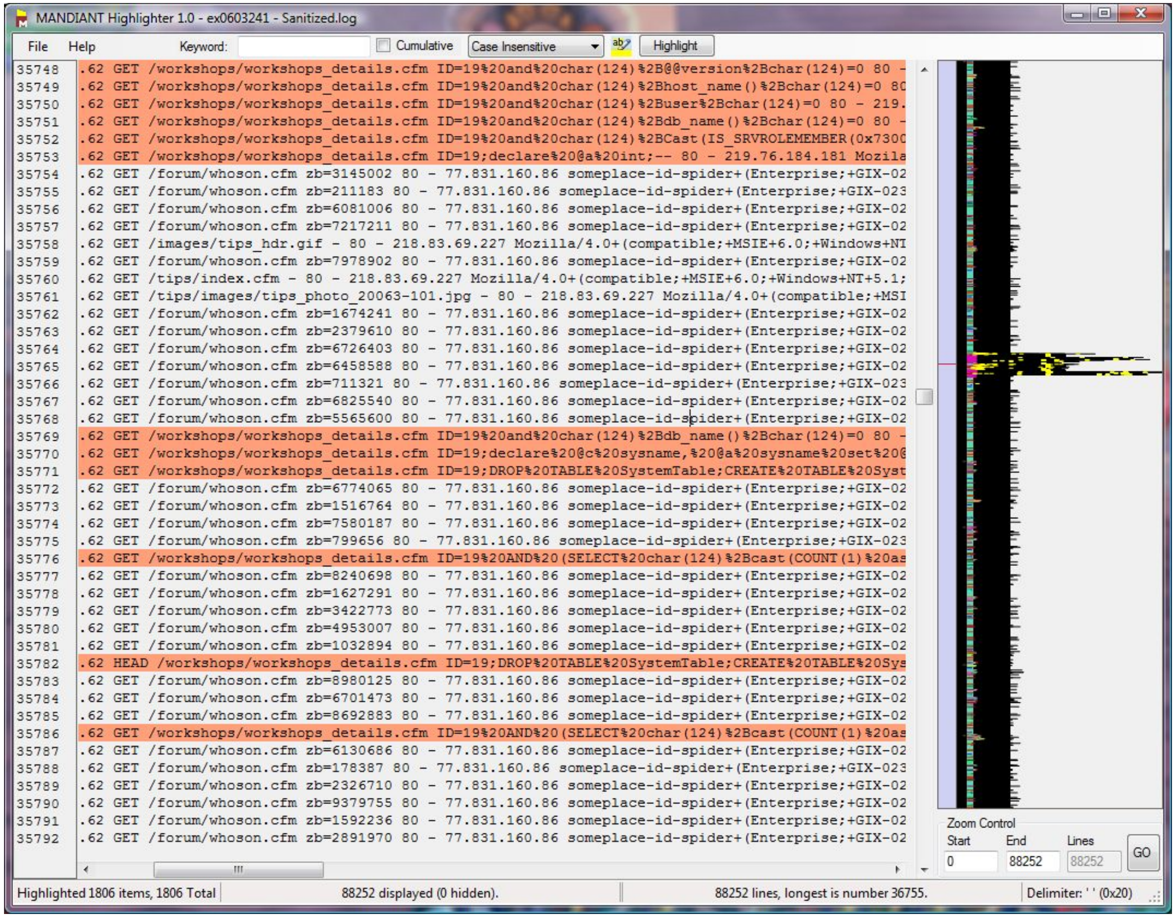
1. Data Resides in Different Tools

- Used space for visual path



- Rote mechanical process
  - Analyst: “Tedious!”

2. Low-level Tasks



Keyword: ☐ Cumulative Case Insensitive

```
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2B@@version%2Bchar(124)=0 80 -
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2Bhost_name()%2Bchar(124)=0 80
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2Buser%2Bchar(124)=0 80 - 219.
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2Bdb_name()%2Bchar(124)=0 80 -
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2BCast(IS_SRVROLEMEMBER(0x730
.62 GET /workshops/workshops_details.cfm ID=19;declare%20@a%20int;-- 80 - 219.76.184.181 Mozila
.62 GET /forum/whoson.cfm zb=3145002 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=211183 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-023
.62 GET /forum/whoson.cfm zb=6081006 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=7217211 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /images/tips_hdr.gif - 80 - 218.83.69.227 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT
.62 GET /forum/whoson.cfm zb=7978902 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /tips/index.cfm - 80 - 218.83.69.227 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.1;
.62 GET /tips/images/tips_photo_20063-101.jpg - 80 - 218.83.69.227 Mozilla/4.0+(compatible;+MSI
.62 GET /forum/whoson.cfm zb=1674241 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=2379610 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=6726403 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=6458588 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=711321 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-023
.62 GET /forum/whoson.cfm zb=6825540 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=5565600 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /workshops/workshops_details.cfm ID=19%20and%20char(124)%2Bdb_name()%2Bchar(124)=0 80 -
.62 GET /workshops/workshops_details.cfm ID=19;declare%20@c%20sysname,%20@a%20sysname%20set%20@
.62 GET /workshops/workshops_details.cfm ID=19;DROP%20TABLE%20SystemTable;CREATE%20TABLE%20Syst
.62 GET /forum/whoson.cfm zb=6774065 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=1516764 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=7580187 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=799656 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-023
.62 GET /workshops/workshops_details.cfm ID=19%20AND%20(SELECT%20char(124)%2Bcast(COUNT(1)%20as
.62 GET /forum/whoson.cfm zb=8240698 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=1627291 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=3422773 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=4953007 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=1032894 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 HEAD /workshops/workshops_details.cfm ID=19;DROP%20TABLE%20SystemTable;CREATE%20TABLE%20Sys
.62 GET /forum/whoson.cfm zb=8980125 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=6701473 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=8692883 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /workshops/workshops_details.cfm ID=19%20AND%20(SELECT%20char(124)%2Bcast(COUNT(1)%20as
.62 GET /forum/whoson.cfm zb=6130686 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=178387 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-023
.62 GET /forum/whoson.cfm zb=2326710 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=9379755 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=1592236 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
.62 GET /forum/whoson.cfm zb=2891970 80 - 77.831.160.86 someplace-id-spider+(Enterprise;+GIX-02
```

Start 0 End 88252 Lines 88252

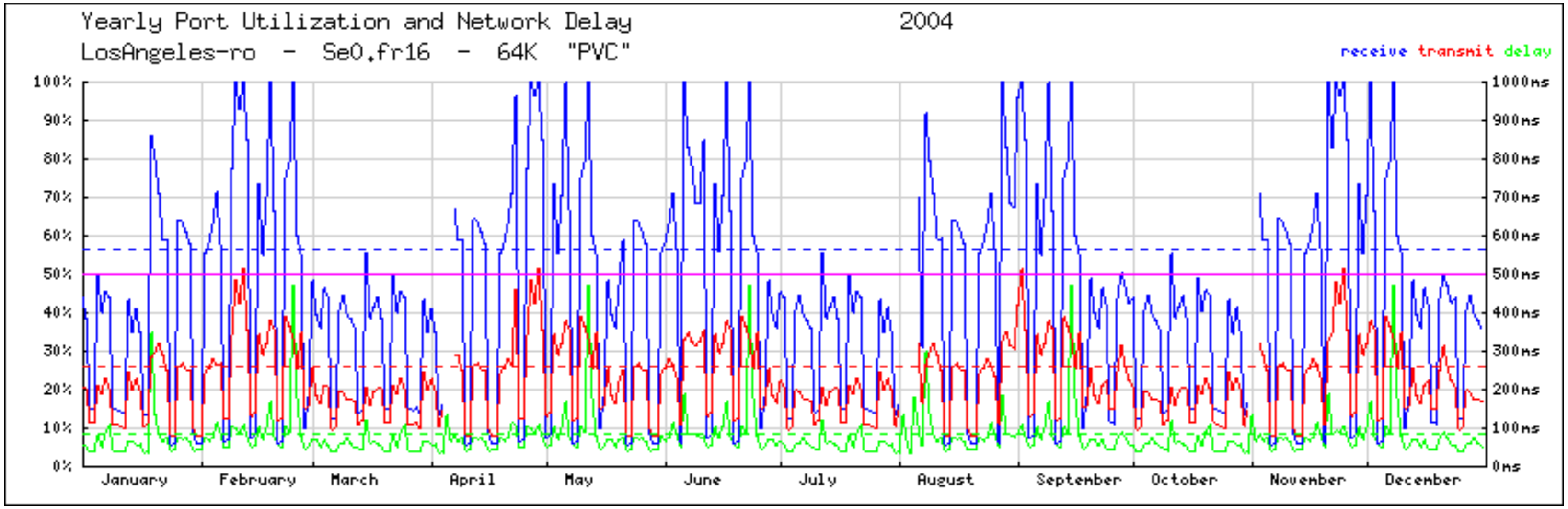
Highlighted 1806 items, 1806 Total | 88252 displayed (0 hidden). | 88252 lines, longest is number 36755. | Delimiter: ' ' (0x20)

Mandiant Highlighter

- Analysts filter out the "normal"
  - line-by-line
- Seek patterns of familiar abnormalities
  - Previous experience creates personal "hit list"
- Analysts observe data individually, not in connection with whole dataset

3. Distrust of Visualizations

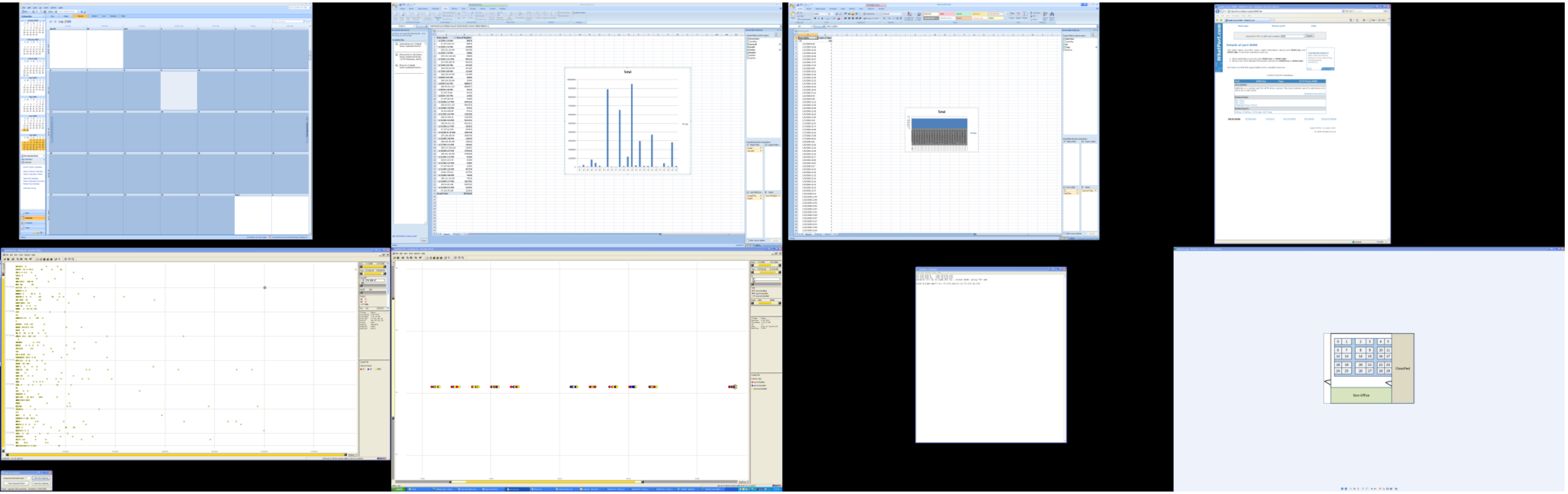
- Analyst: "Visualizations are in the way of the data"



- Visualizations:
  - May be too slow
  - May hide important, small details
- Analysts can only *see*, not *manipulate* the data

4. Quest for a "Query"

- "Query" != SQL query
- "Query" is the question that finds the answer you have
  - Cumulative result of *interaction* with variety of tools



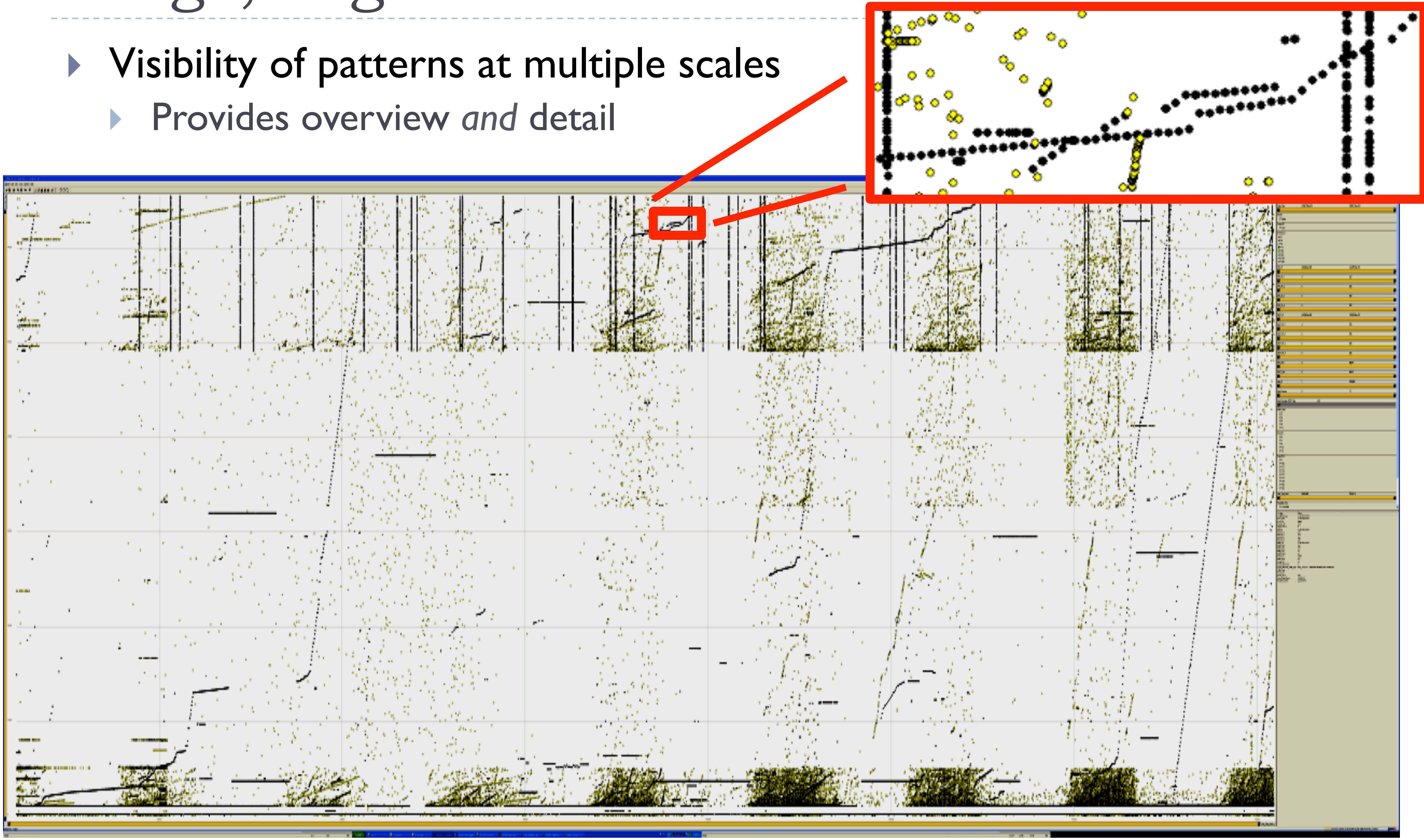
- *The* ***process*** *of forming this query is key!*

Guidelines for Usable Workspaces

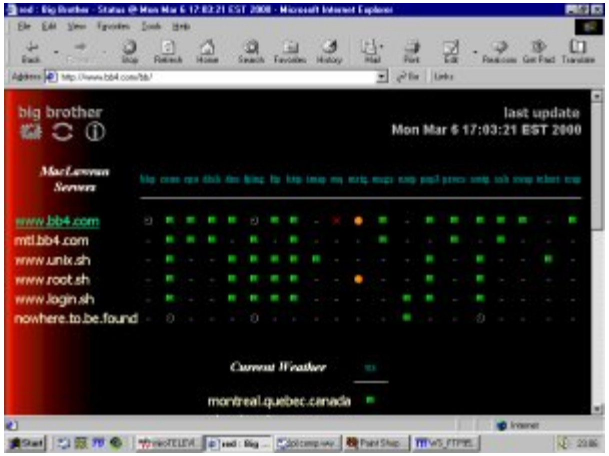
- Multi-scale Visualizations
- De-Aggregate Vital Information
- Support multiple, simultaneous investigation cases
- Provide history and traceability for investigations

Large, High-Resolution Visualization

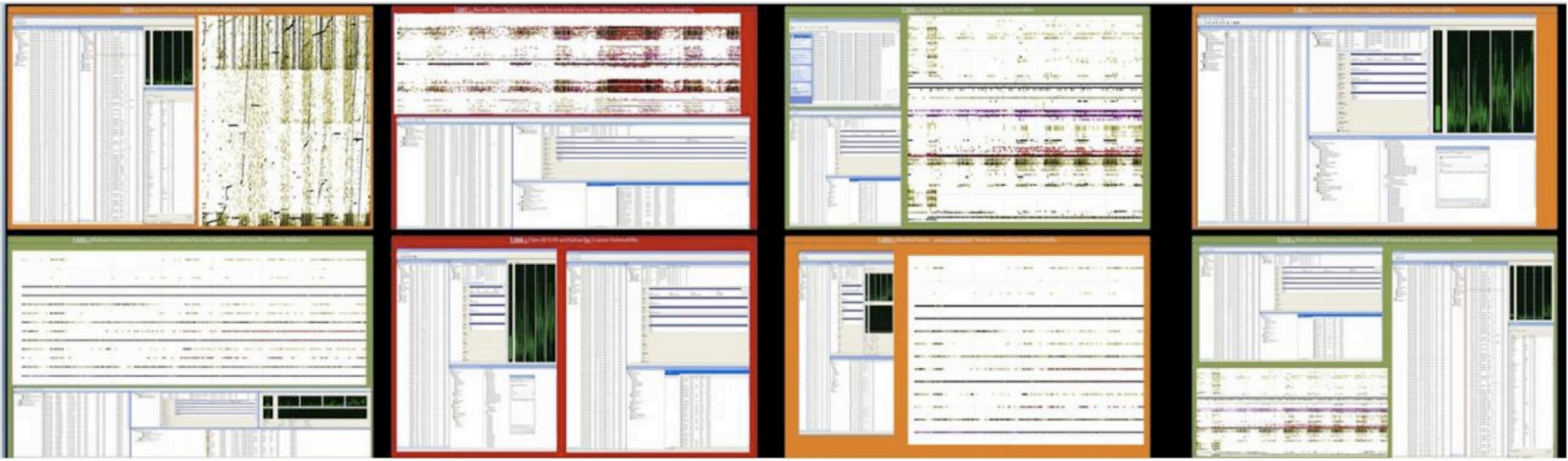
- Visibility of patterns at multiple scales
  - Provides overview *and* detail



De-Aggregate Vital Information

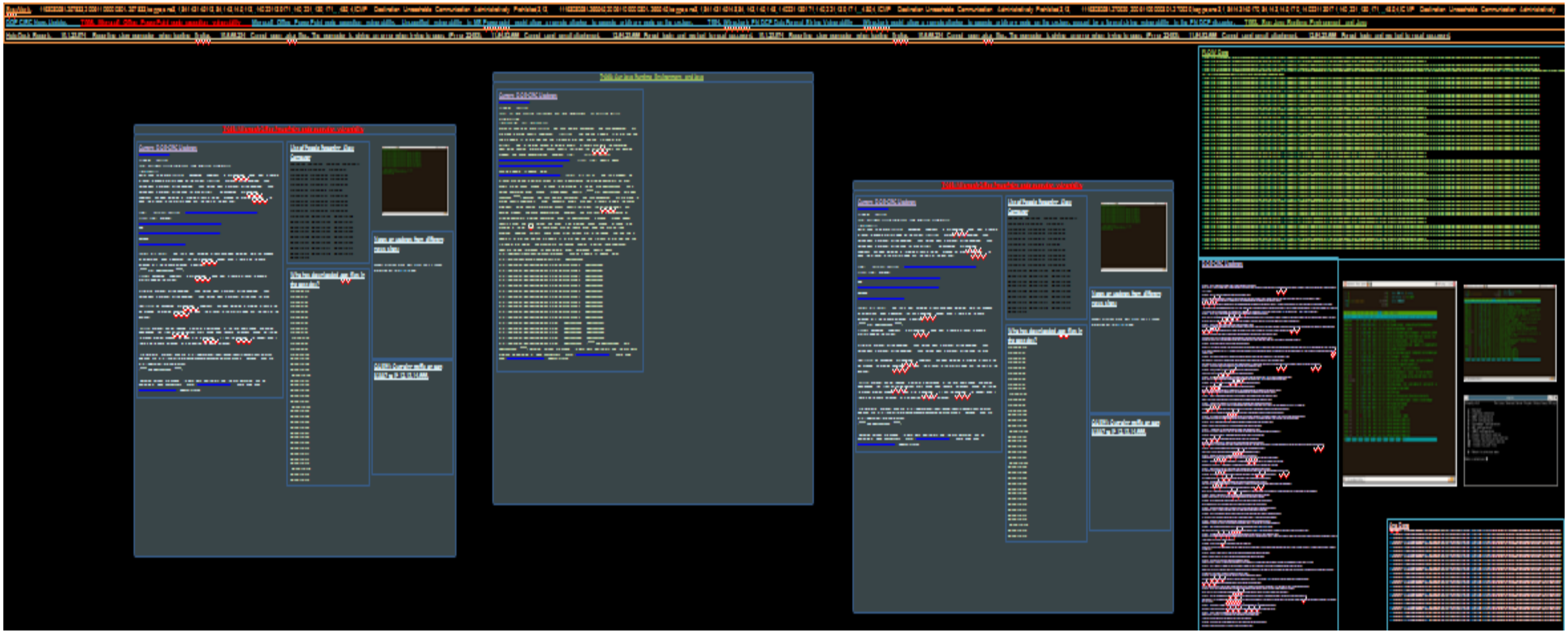


- Provides analyst with situational awareness
  - De-aggregation of information
  - More upfront information, while maintaining overview



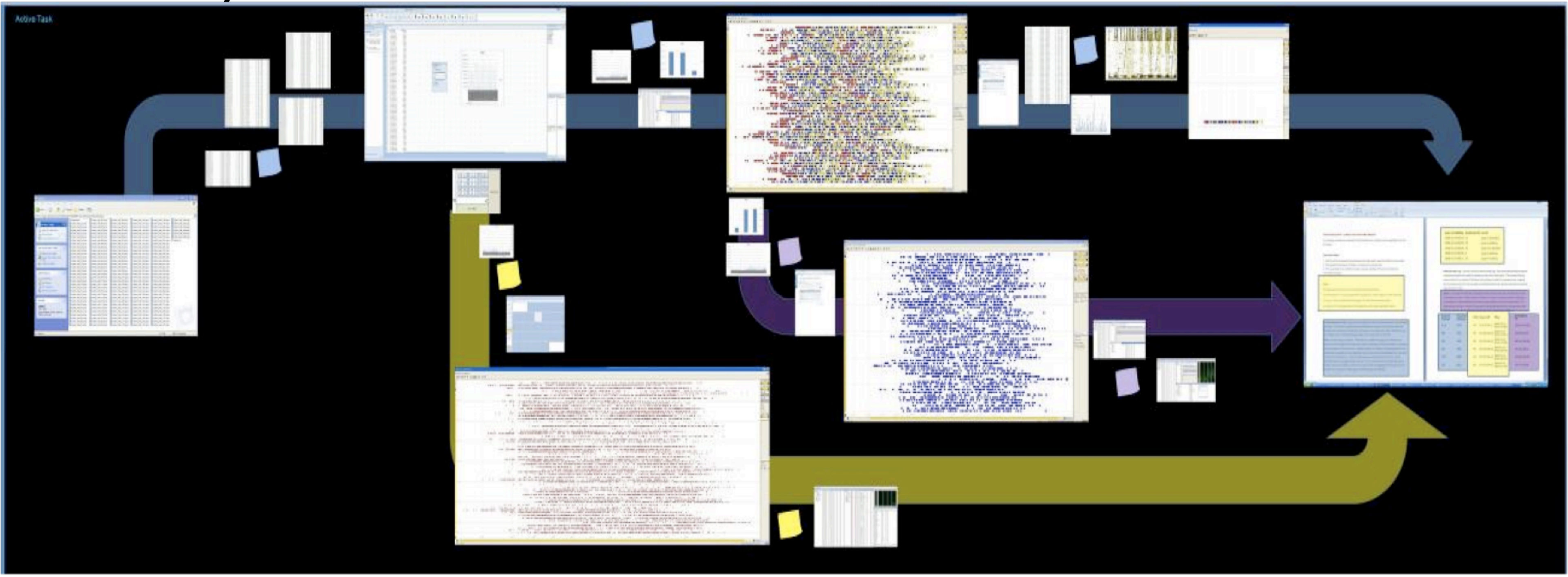
Multiple Simultaneous Cases

- Shows live data
  - Real time updating
- Analyst can set alerts for monitoring
- Enables collaboration by sharing cases



History and Traceability

- "History Trees": concept providing traceability and history of analyst's workflow



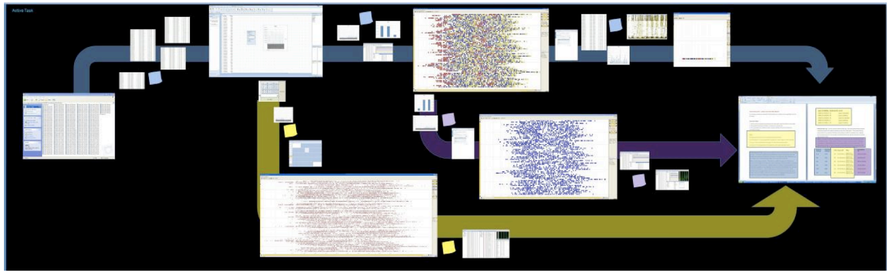
A visualization should be the means for a user to interact and think.

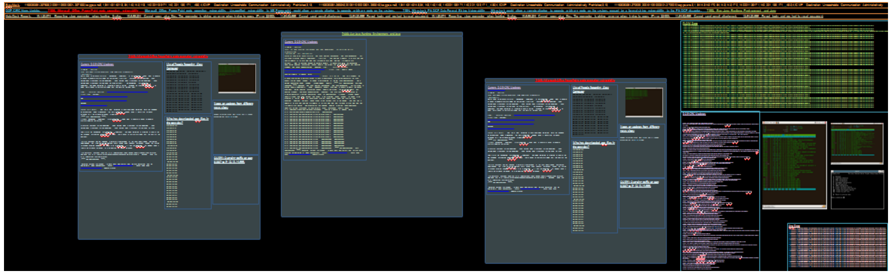
Intelligence vs. Cyber Analytics

Stegosaurus Scenario **(Intelligence Analytics)**	**Cyber Security Scenario** **(Cyber Analytics)**
Creating a *story* about the threat. Product = story	Building a *query* to identify the threat. Product = query
Work done in a *visual space*. (Sensemaking Process)	Work done in *textual space*. (Tools to Process the Data)
Rely on *Visualizations*.	Rely on *Linux Command Line*.
Un-, semi-, and structured data.	*Mainly* structured data. (packet, etc.)
Lots of data.	Even *more* data!
Interactions reside outside the windows.	Interactions reside within the windows

Let’s give the *user* more *space*!

Let’s make the *space* more *useful*!





History and Traceability

Multiple, Simultaneous Investigation cases

Large, High-Resolution Visualizations

De-Aggregate Vital Information