Session Viewer:

Visual Exploratory Analysis of Web Session Logs

Heidi Lam, Daniel M. Russell,
Diane Tang and Tamara Munzner

University of British Columbia, Canada
Google Inc., USA

Web Session Log Data

- Basic unit of web session log analysis is a session
- **Session** is a time-stamped sequence of events performed by a user
- Each **event** represents a user action,
  e.g., submitting a query, clicking on a result link
- A **session population** is a group of sessions with shared characteristics,
  e.g., task type or usage patterns

Web Session Log Data

- Both sessions and events have attributes
  - Session attributes include total duration, total number of events
  - Event attributes include time stamp, page title, page URL
- Example user study log
  - collected with browser plug-in from a study with volunteer users

Pane	Paneli	Que	Navigation_Event	Event_	URL	Title
1	15238	44	Begin_Section	1		
2	15238	45	Begin_Objective	1		
3	15238	45	Go to Start URL of Objecti	31	http://www.msn.com	
4	15238	45	Browser Navigate	59	http://search.msn.com/results.asp	MSN Search: london and tours
5	15238	45	Browser Navigate	118	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
6	15238	45	Browser Add Favorite	175	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
7	15238	45	Browser Navigate	204	http://www.london-tours.info/londo	Discovering London (Full Day Tour No 4) Ticke
8	15238	45	Browser Back	244	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
9	15238	45	Browser Navigate	265	http://www.london-tours.info/londo	Oxford, Stratford & Warwick Castle (Tour 23) Ti
10	15238	45	Browser Back	316	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
11	15238	45	Browser Back	320	http://search.msn.com/results.asp	MSN Search: london and tours
12	15238	45	End_Objective	360		

Web Session Log Analysis

- Analysis provides better understanding of web usage behaviours to improve support

- Two main approach in analysis:
  - **Detailed-session**: look at sessions one by one, event by event
  - **Aggregate-statistical**: look at session/event attributes using statistics (e.g., total task time)

- Both methods have merits and problems:
  - **Detailed-session:** can understand usage, but only analyze a very small portion of data and may have selection bias
  - **Aggregate-statistical**: highly scalable, but understand sessions by attributes and may be too hypothesis-driven

Visualization and Exploratory Analysis

- Visualization can support data exploration and bridge between the detailed and statistical analysis approaches
  - **Detailed**: guides session selection with session-attribute distribution
  - **Aggregate**: provides context and content to statistical results

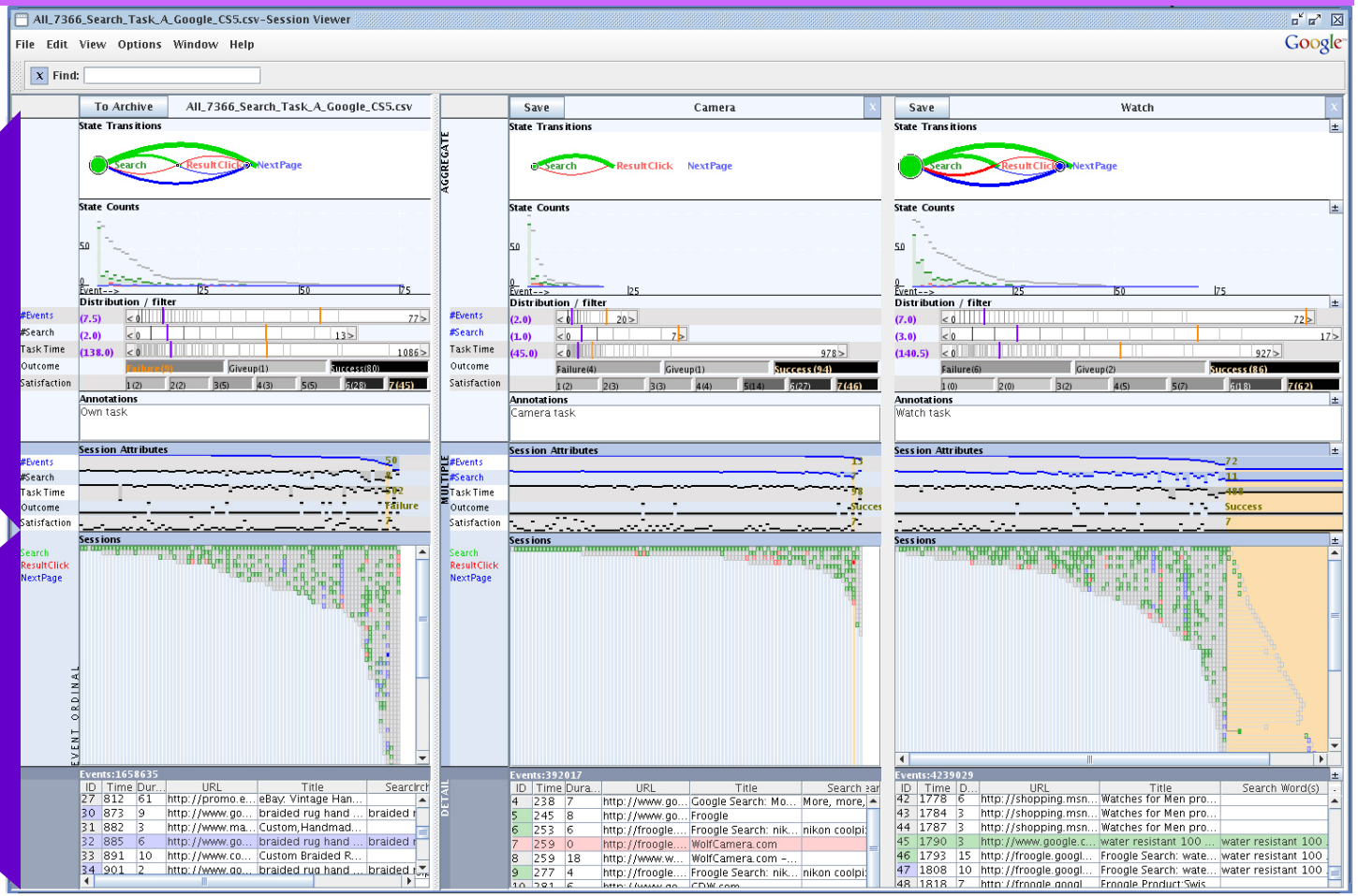
➢ Session Viewer

- Displays **multiple data levels** to facilitate cross-level analysis
- Displays **multiple session populations** for side-by-side population comparisons

Session Viewer: Main Screen

attributes

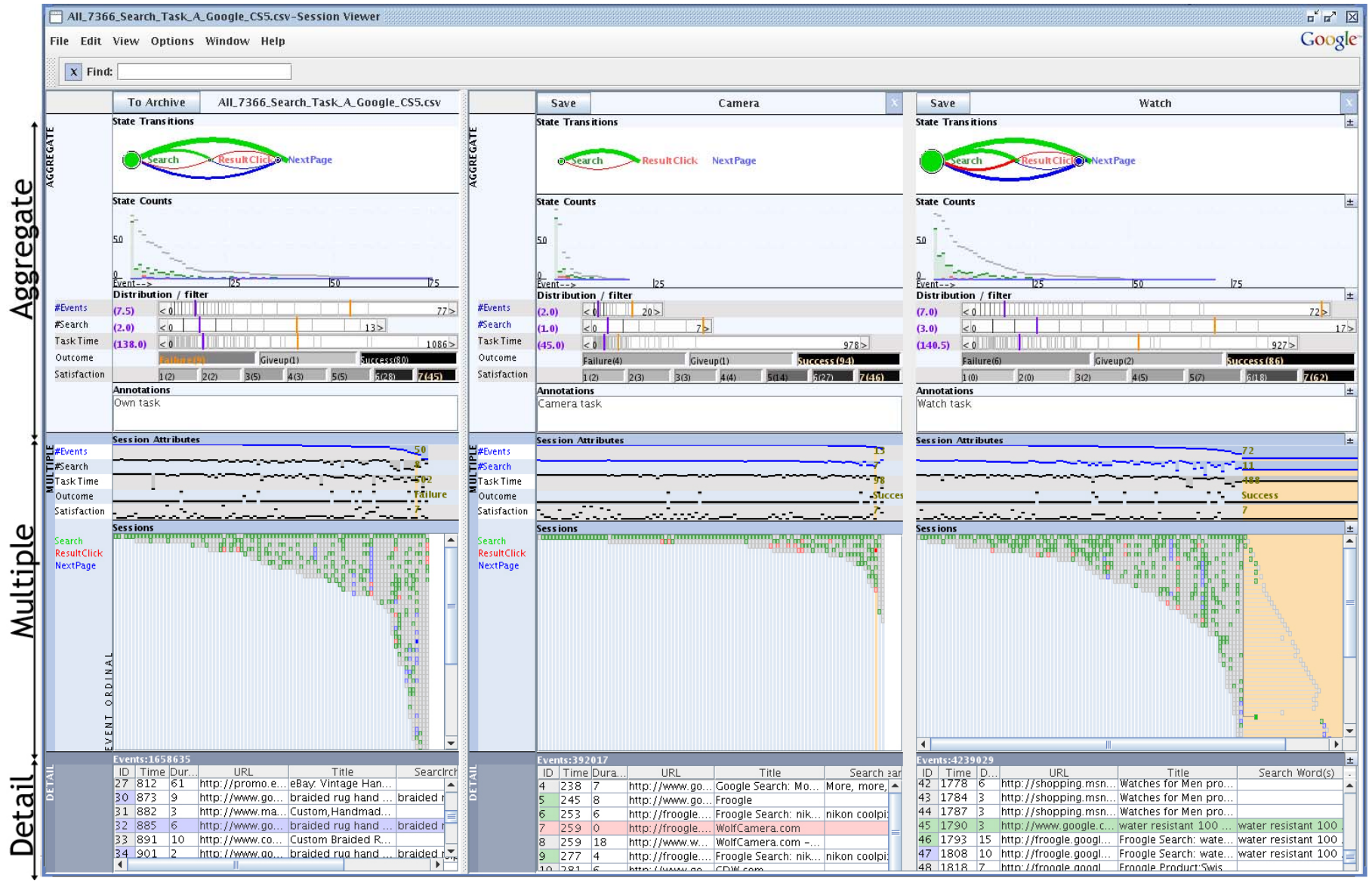
data



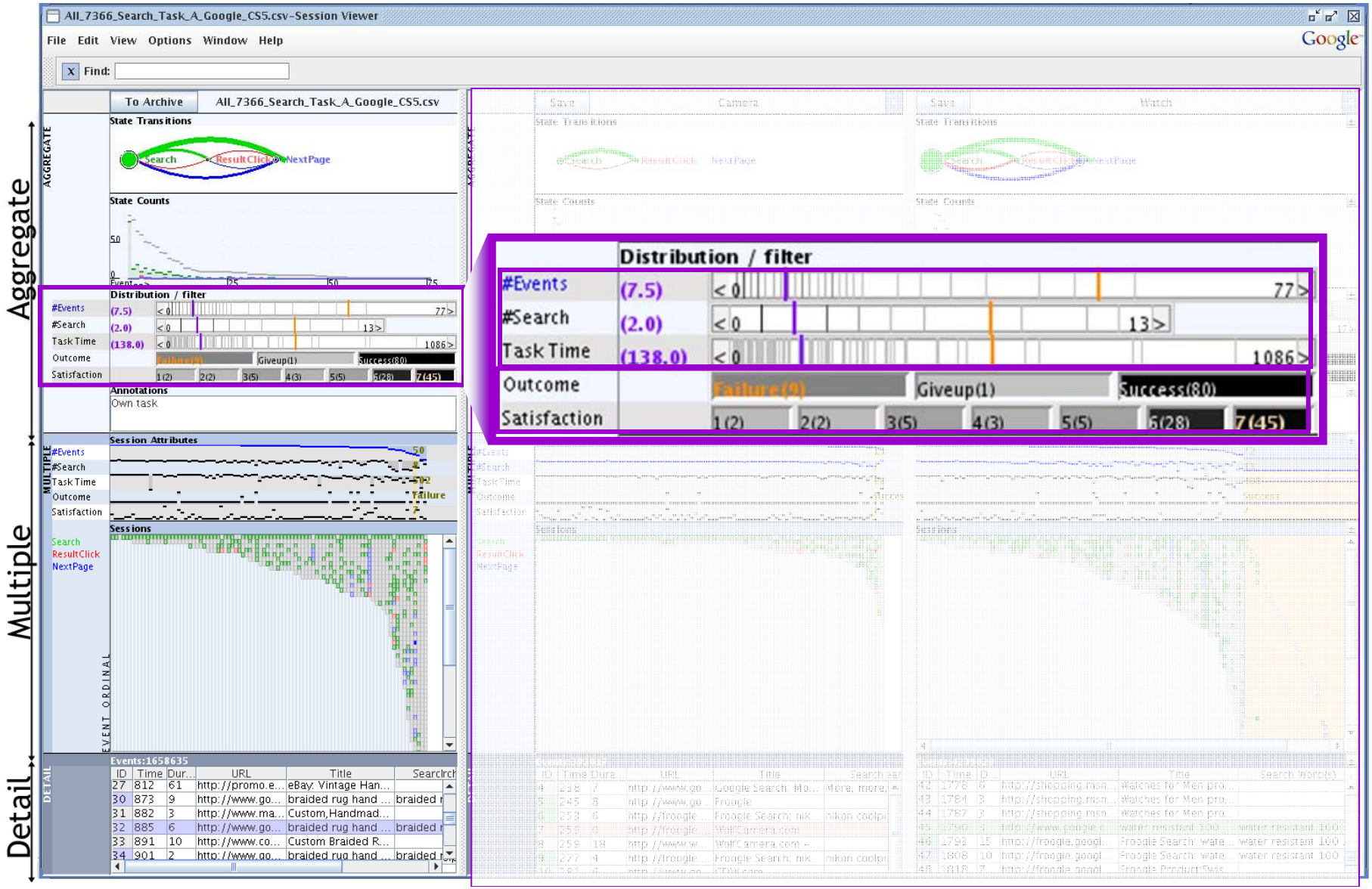
Pan		ation_Event	Event_	URL	Title
		Begin_Section	1		
2	15238	45 Begin_Objective	1		
3	15238	45 Go to Start URL of Objecti	31	http://www.msn.com	
4	15238	45 Browser Navigate	59	http://search.msn.com/results.asp	MSN Search: london and tours
5	15238	45 Browser Navigate	118	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
6	15238	45 Browser Add Favorite	175	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
7	15238	45 Browser Navigate	204	http://www.london-tours.info/londc	Discovering London (Full Day Tour No 4) Ticke
8	15238	45 Browser Back	244	http://www.london-tours.info/	London Tours Info - over 70 London, UK & Pari
9	15238	45 Browser Navigate	265	http://www.london-tours.info/londc	Oxford, Stratford & Warwick Castle (Tour 23) Ti
10	15238	45 Browser Back	316	http://www.london-tours.info/	London Tours Info - over 70 L
11	15238	45 Browser Back	320	http://search.msn.com/results.asp	MSN Search: london and tour
12	15238	45 End_Objective	360		

Session logs

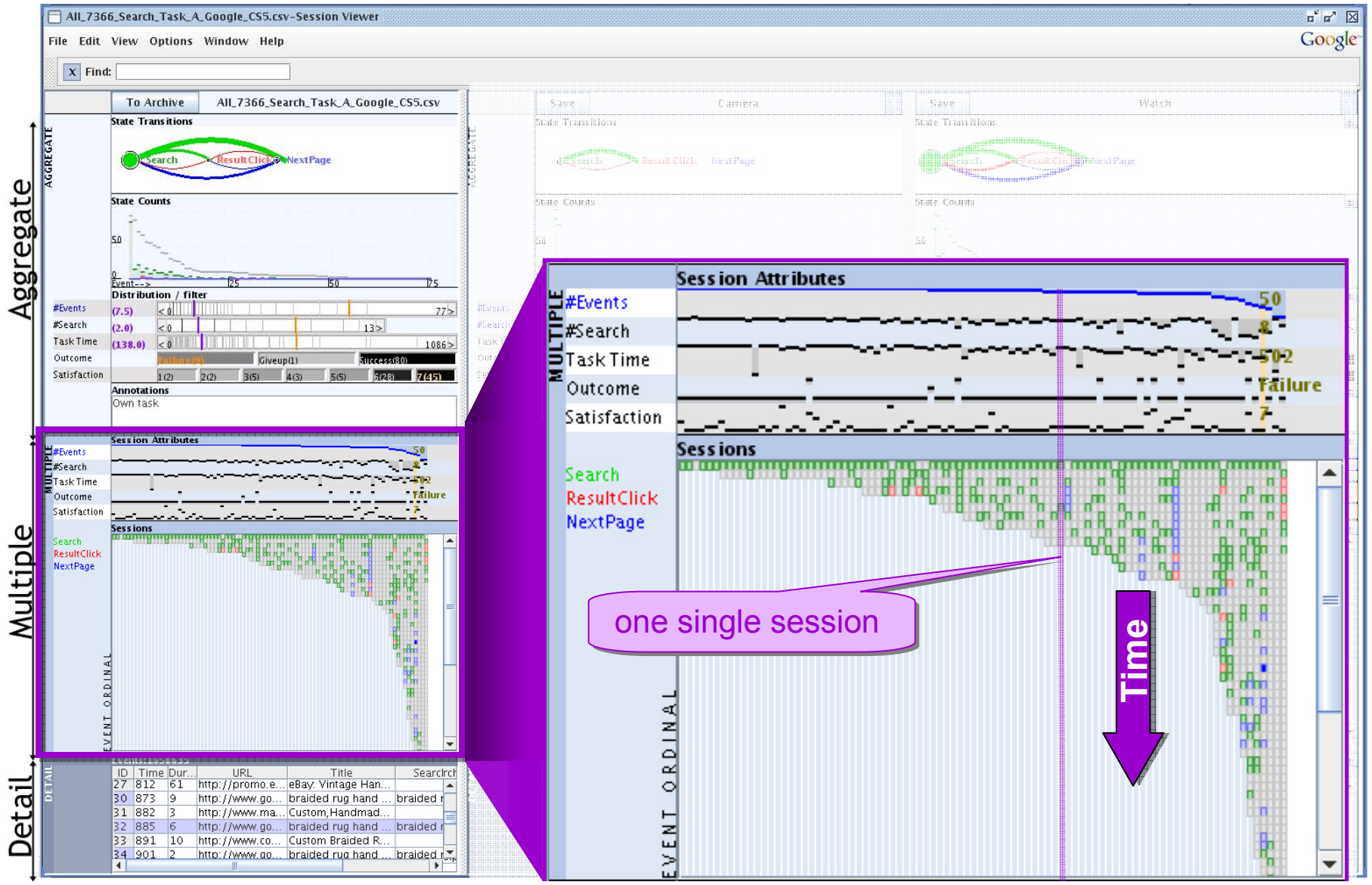
Session Viewer: Multiple Data Levels



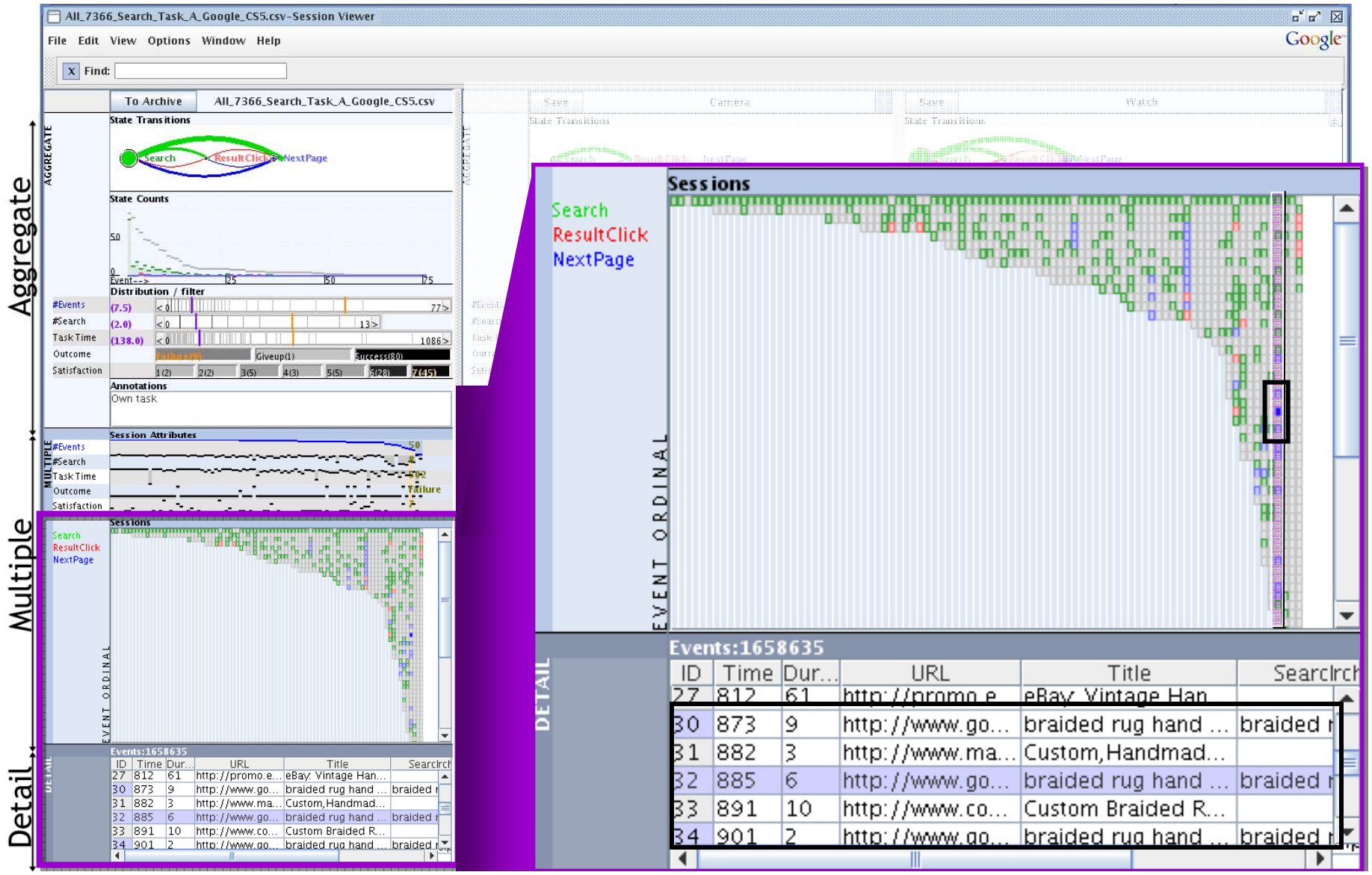
Session Viewer: Aggregate Level



Session Viewer: Multiple Level



Session Viewer: Multiple & Detail Levels

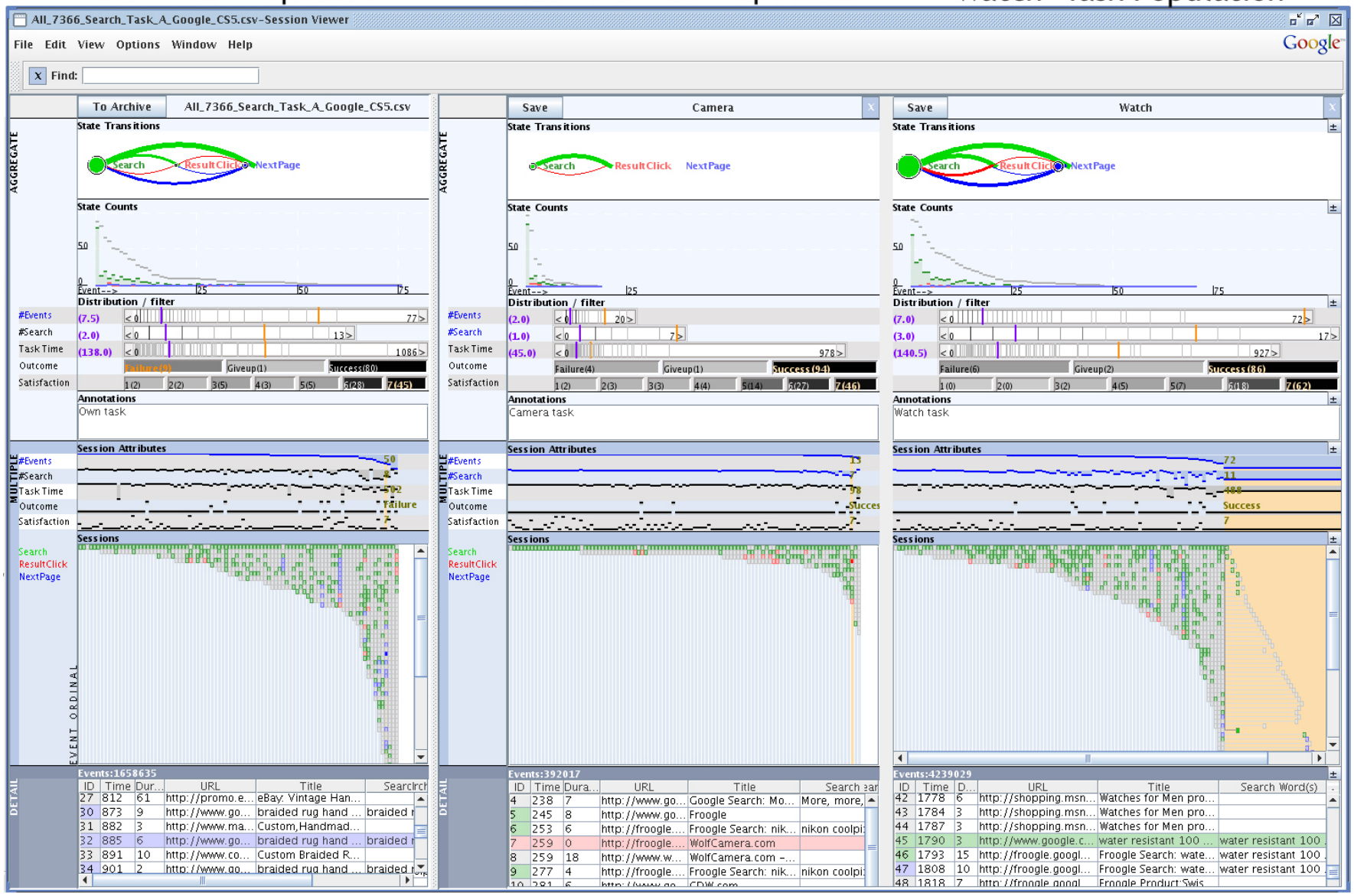


Session Viewer: Multiple Population Views

"Own" Task Population

"Camera" Task Population

"Watch" Task Population



Use-case Scenario: Analyzing Study Logs

- Illustrate the use of Session Viewer with existing study log data
  - Large-scale user study conducted by Keynote
  - Study log data collected with browser plug-in with user consent
- Study gathered >6,000 sessions, grouped by
  - Search engine type
  - Search domain (Image, Shopping, News, Local/Maps)
  - Question variant (3 pre-defined, one participant's own)
    - E.g., "Camera" task to look for image resolution of a camera, and the "Watch" task to look for a water-resistant watch under $100
- Study Logs analyzed previously: participants' own tasks are longer and more involved than pre-defined tasks (Russell & Grimes, 2007)

Use-case Scenario: Analyzing Study Logs

- Previous analysis identified two population groups: “Own” task, and “Pre-defined” task populations

- We revisited the same study logs with Session Viewer, and found two **new findings**:

  1. Pre-defined tasks are not isomorphic, and can be classified further by task type

  2. Task type classification, based on task instructions, is reflected as usage patterns in session logs

- We also verified correlation between two participant feedback: task outcome and satisfaction scores

Use-case Scenario: Analyzing Study Logs

- To analyze the study logs, we customized Session Viewer by defining…
  - Event types: Search, ResultClick, NextPage, and non-search engine events
  - Session attributes
    - Calculated based on event attributes: #Events, Task Time
    - Read from log: Task Outcome, Satisfaction scores

Use-case Scenario: Three Shopping Populations

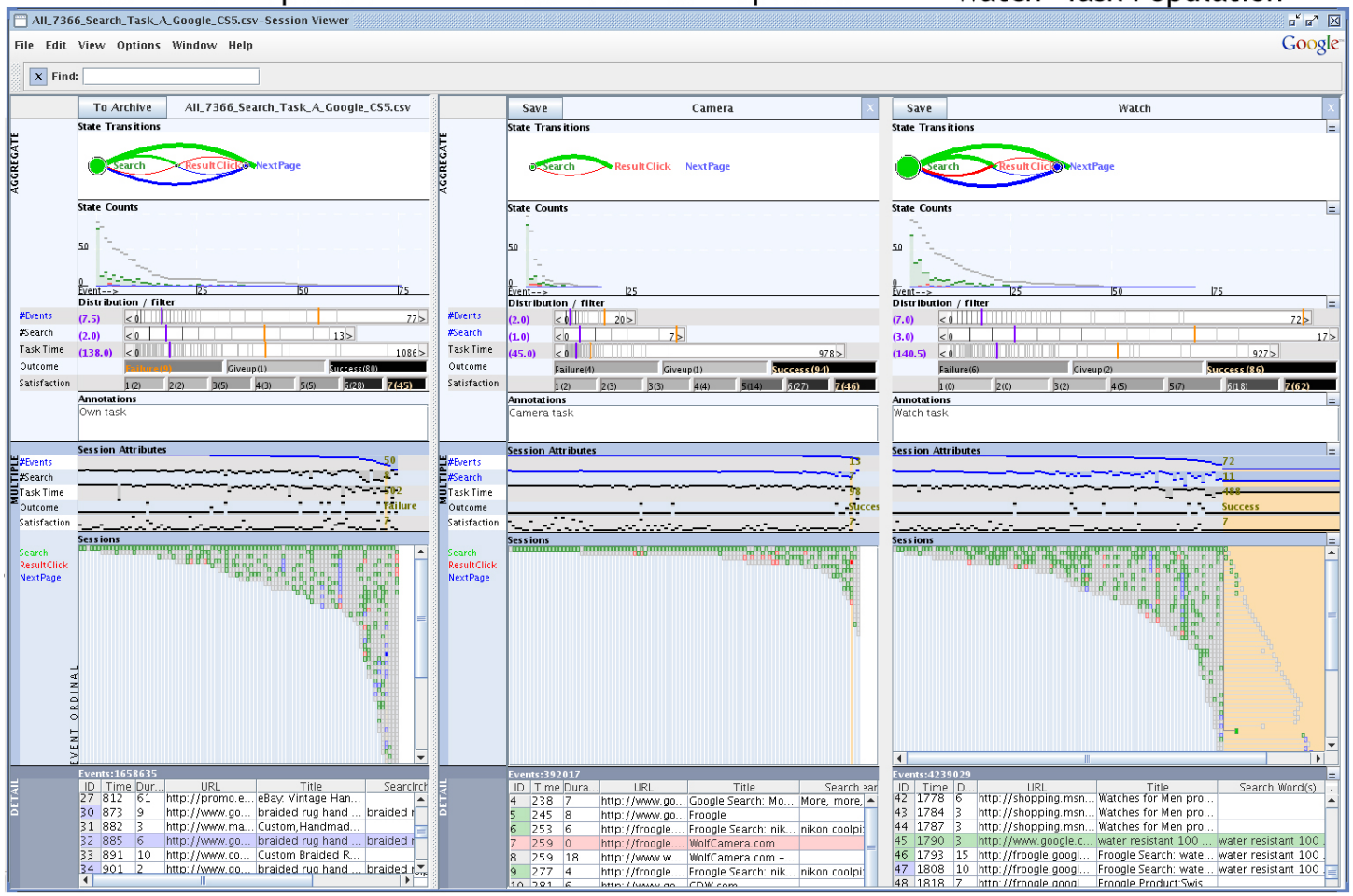
How many megapixels is the image resolution of the Nikon Coolpix 4600?

Find a water-resistant watch for under $100

"Own" Task Population

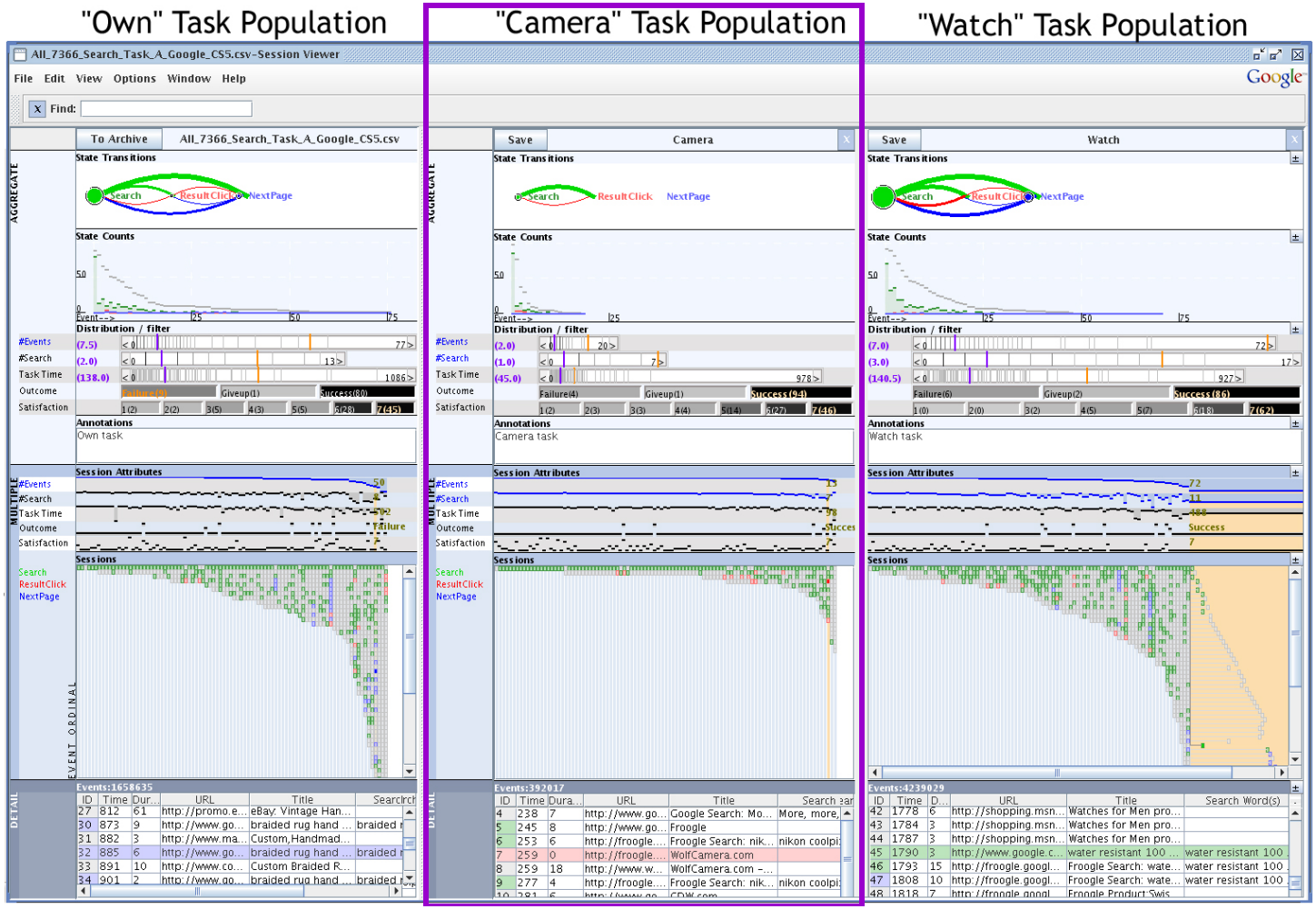
"Camera" Task Population

"Watch" Task Population



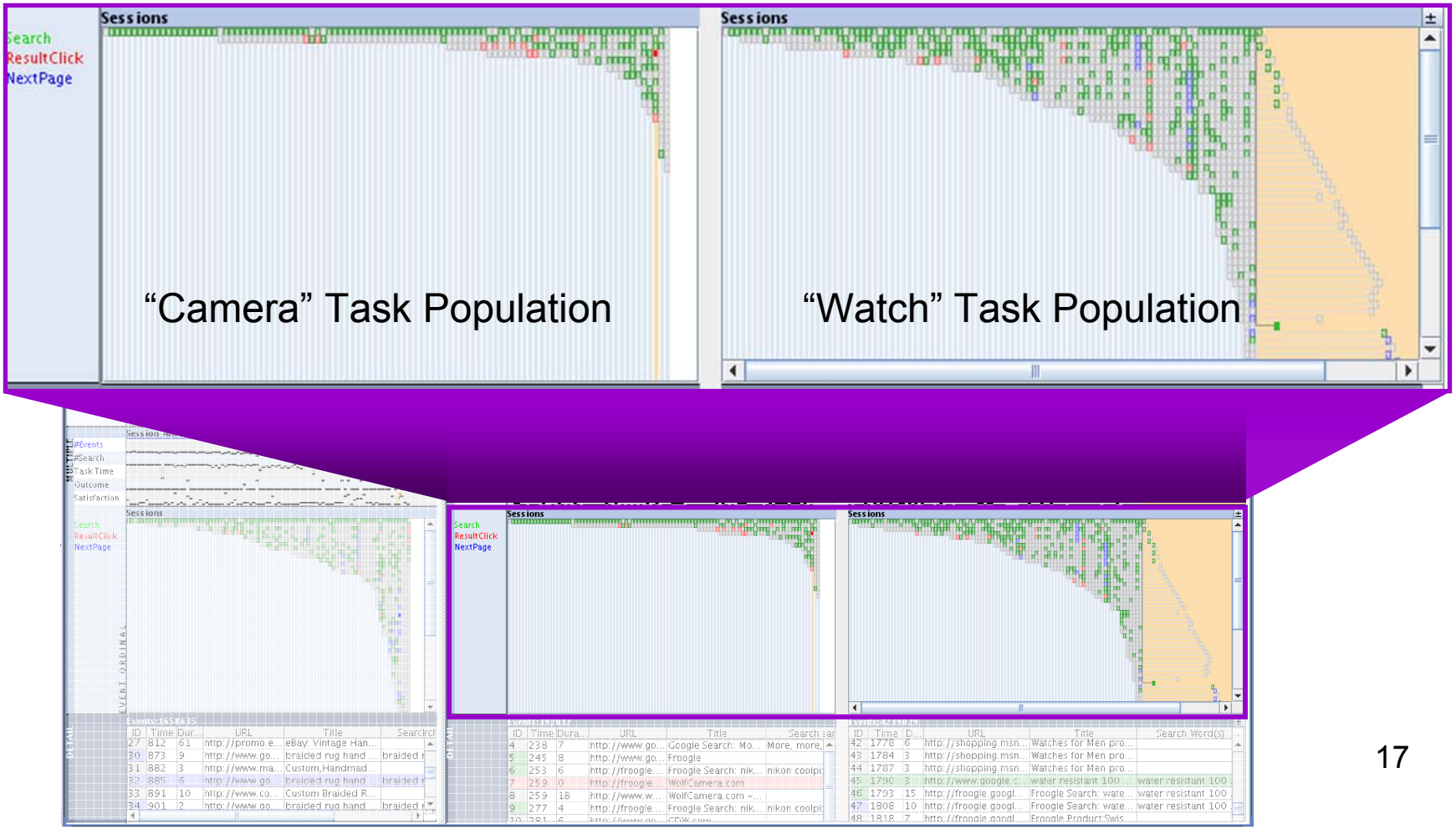
Use-case Scenario: Compare between aggregates

- Expected the populations to look similar, as the question variants were assumed to be isomorphic
- But the “Camera” task population has shorter sessions with fewer events



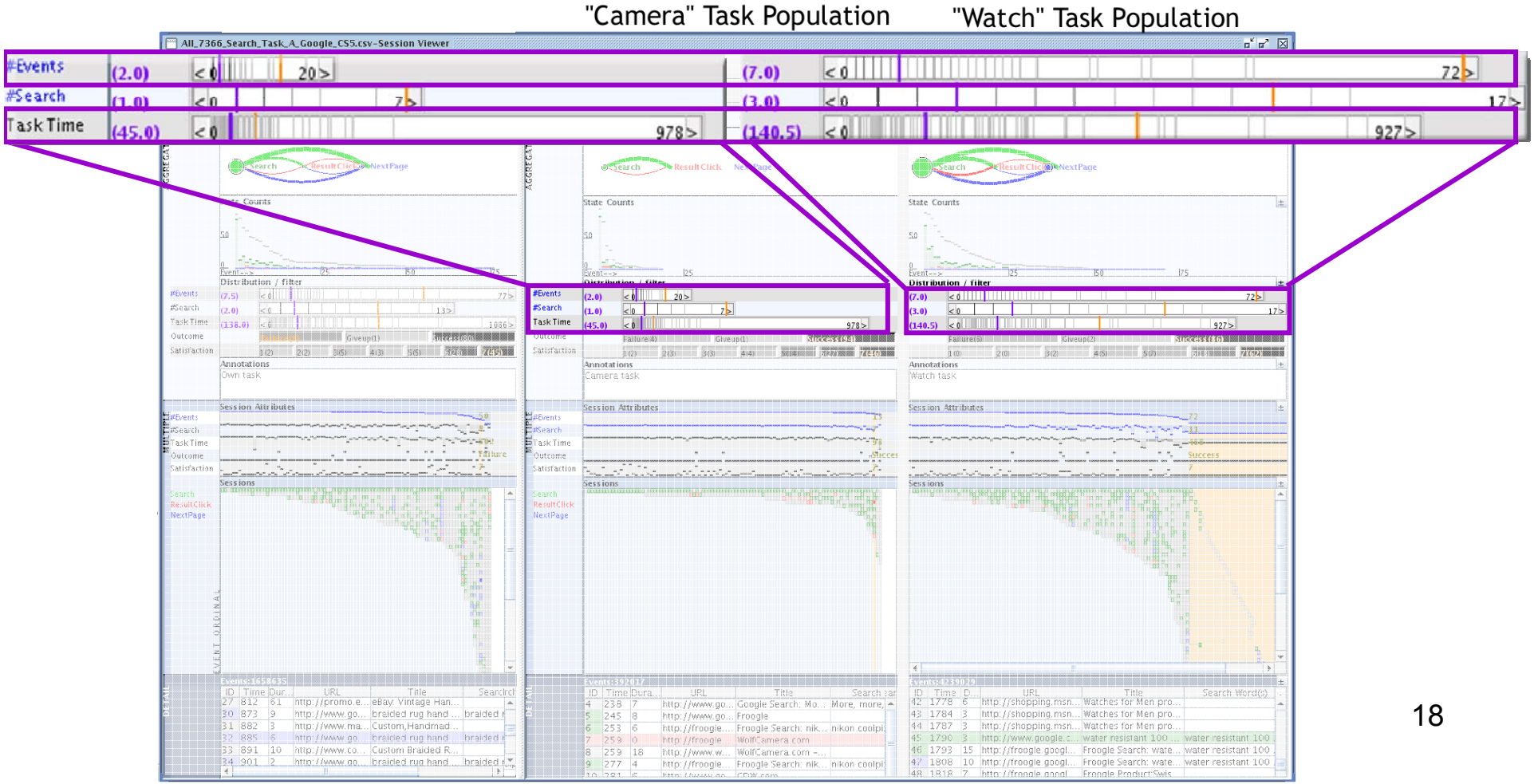
Use-case Scenario: Compare between aggregates

- Expected the populations to look similar, as the question variant were assumed to be isomorphic
- But the “Camera” task population has shorter sessions with fewer events



Use-case Scenario: Compare between aggregates

- Expected the populations to look similar, as the question variant were assumed to be isomorphic
- But the “Camera” task population has shorter sessions with fewer events



Use-case Scenario: Compare between aggregates

Directed, closed, information search*

How many megapixels is the image resolution of the Nikon Coolpix 4600?

*based on Rose & Levinson 2004

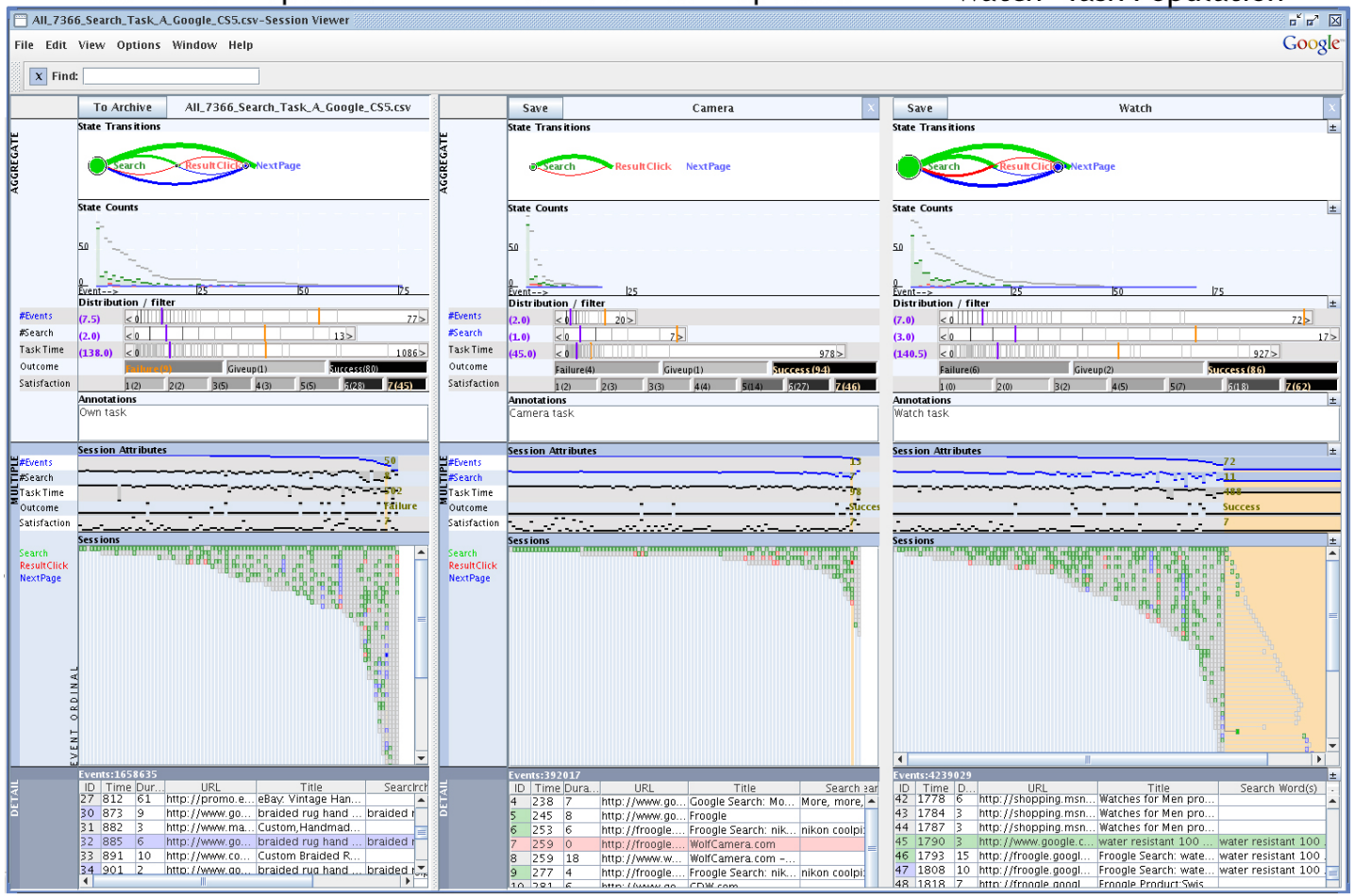
Information locate*

Find a water-resistant watch for under $100

"Own" Task Population

"Camera" Task Population

"Watch" Task Population

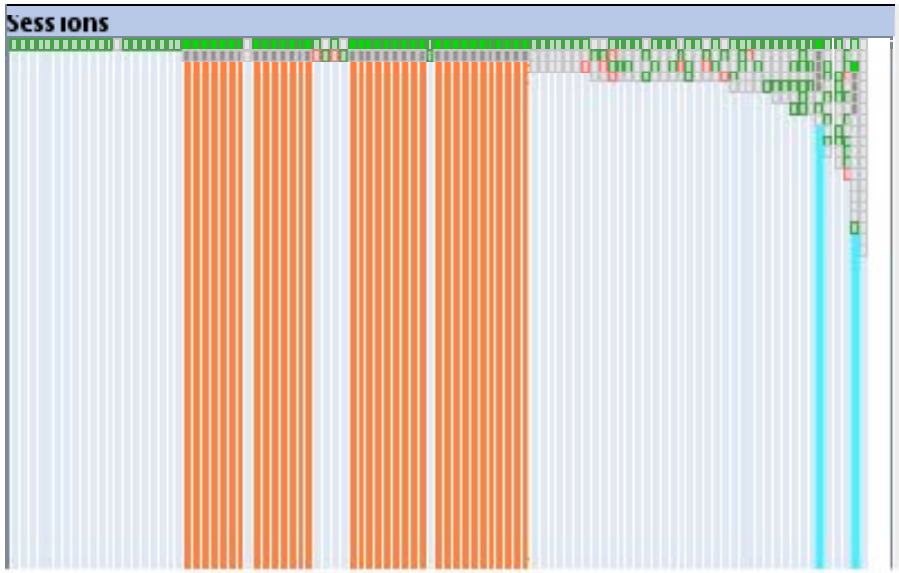


Use-case scenario: Usage-pattern distributions

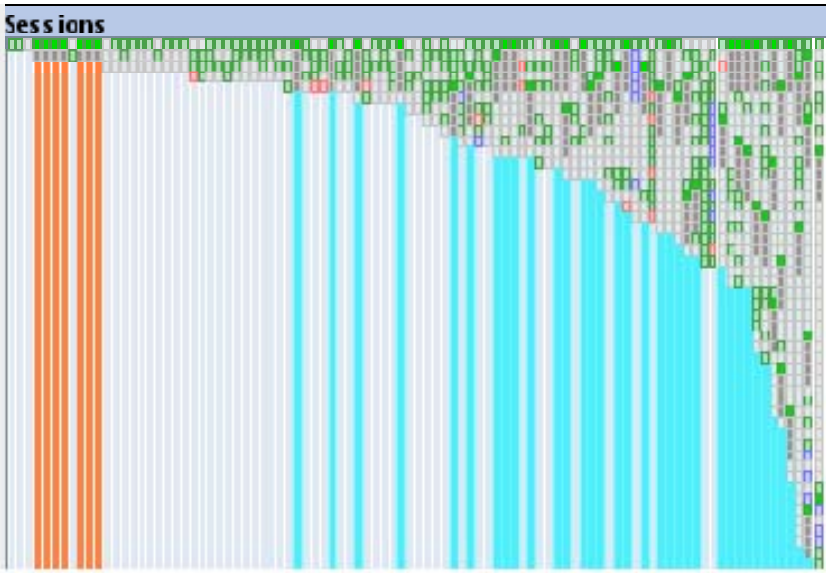
- How do these populations differ?

- Iteratively defined usage patterns:
  - Short navigation: S[Start] → X[End]
  - Topic exploration: S → X → X → X

- More Short navigations in Camera-Task population
  More Topic explorations in Watch-Task population



Camera Task Population



Watch Task Population

Use-case scenario: Usage pattern details

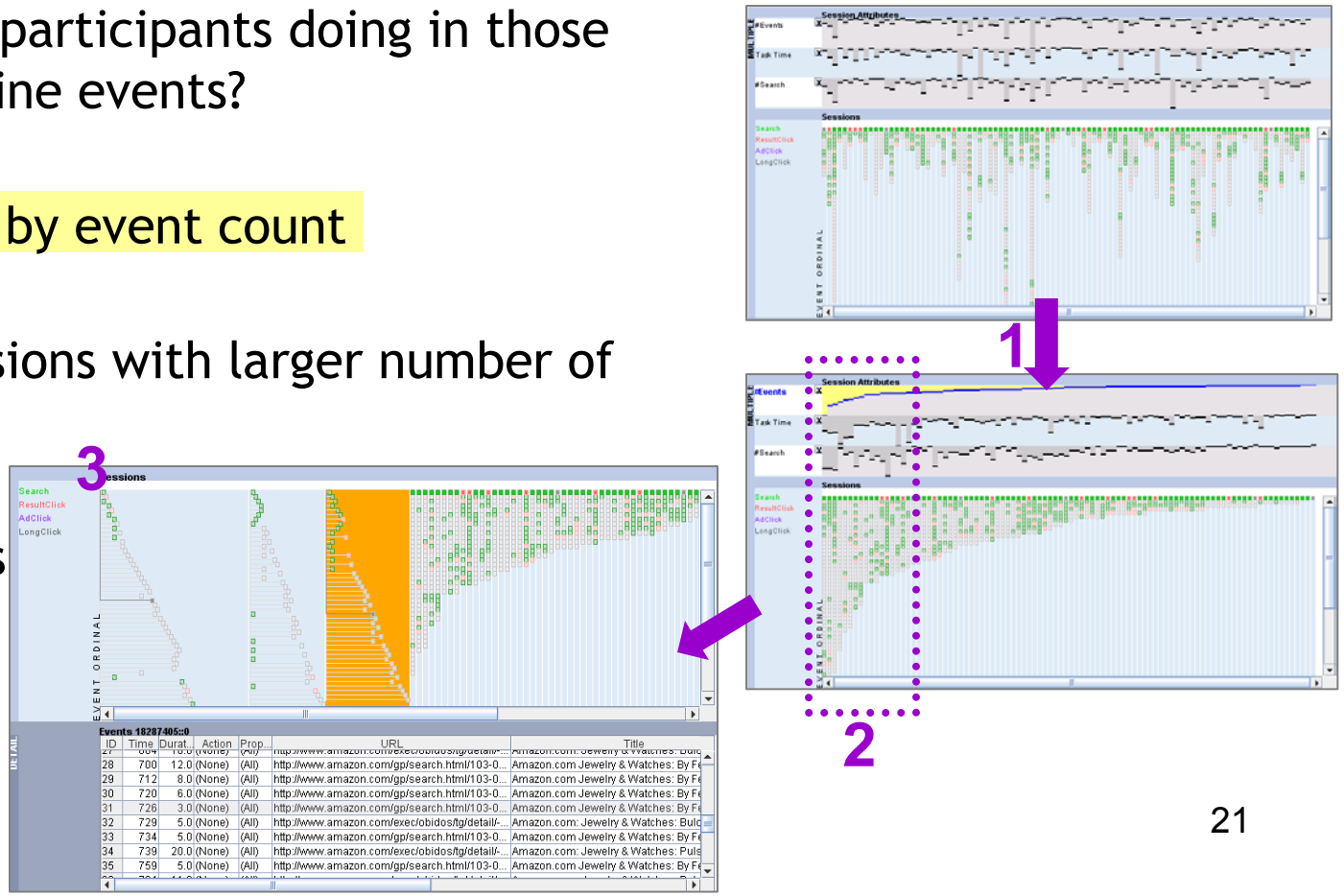
- Short-navigation makes sense for direct, closed information task (look for image resolution in Nikon Coolpix 4600)

- What is Topic Exploration (S → X → X → X)? What were the participants doing in those non-search engine events?

1. Sorted sessions by event count

2. Focused on sessions with larger number of events

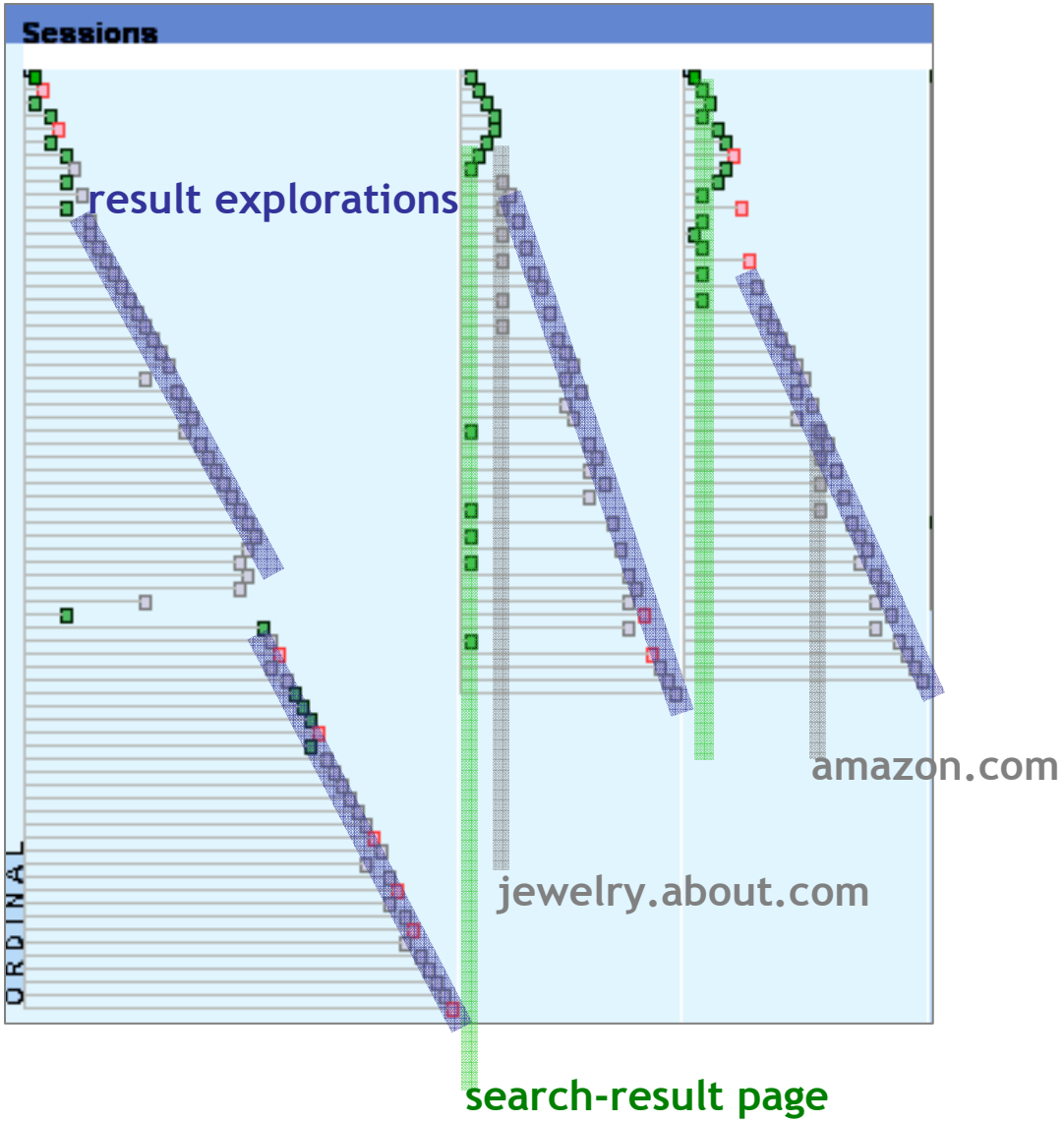
3. Expand sessions to better view usage patterns



Use-case scenario: 2D expanded sessions

URL	Title
http://froogle.google.com/froogle?hl=en&lr&oi...	Froogle Search: water resistant watches
http://froogle.google.com/froogle?q=water+re...	Froogle Search: water resistant watches
http://www.google.com/search?hl=en&lr=&q=...	water resistant watches - Google Search
http://www.google.com/url?sa=l&ai=Bf465ug...	Timepieces International Inc - (Klaus Kobec and ...
http://www.google.com/search?hl=en&lr=&q=...	water resistant watches - Google Search
http://www.google.com/search?hl=en&q=wat...	watches - Google Search
http://www.google.com/search?hl=en&lr=&q=...	water resistant watches - Google Search
http://www.google.com/url?sa=l&ai=B5rwKug...	The Sharper Image - Search Results
http://www.google.com/search?hl=en&lr=&q=...	water resistant watches - Google Search
http://www.chronocentric.com/watches/wresi...	Chronocentric: Water Resistance
http://www.google.com/search?hl=en&lr=&q=...	water resistant watches - Google Search
http://www.amazon.com/gp/browse.html?_en...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/browse.html?rh=...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/sr=3-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/exec/obidos/tg/detail...	Amazon.com: Jewelry & Watches: Bulova Carave...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/exec/obidos/tg/detail...	Amazon.com: Jewelry & Watches: Bulova Carave...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/exec/obidos/tg/detail...	Amazon.com: Jewelry & Watches: Pulsar Men's ...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/exec/obidos/tg/detail...	Amazon.com: Jewelry & Watches: Bulova Carave...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/exec/obidos/tg/detail...	Amazon.com: Jewelry & Watches: Pulsar Analog/...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...
http://www.amazon.com/gp/search.html/103-...	Amazon.com Jewelry & Watches: By Feature / W...

Use-case scenario: Usage pattern details



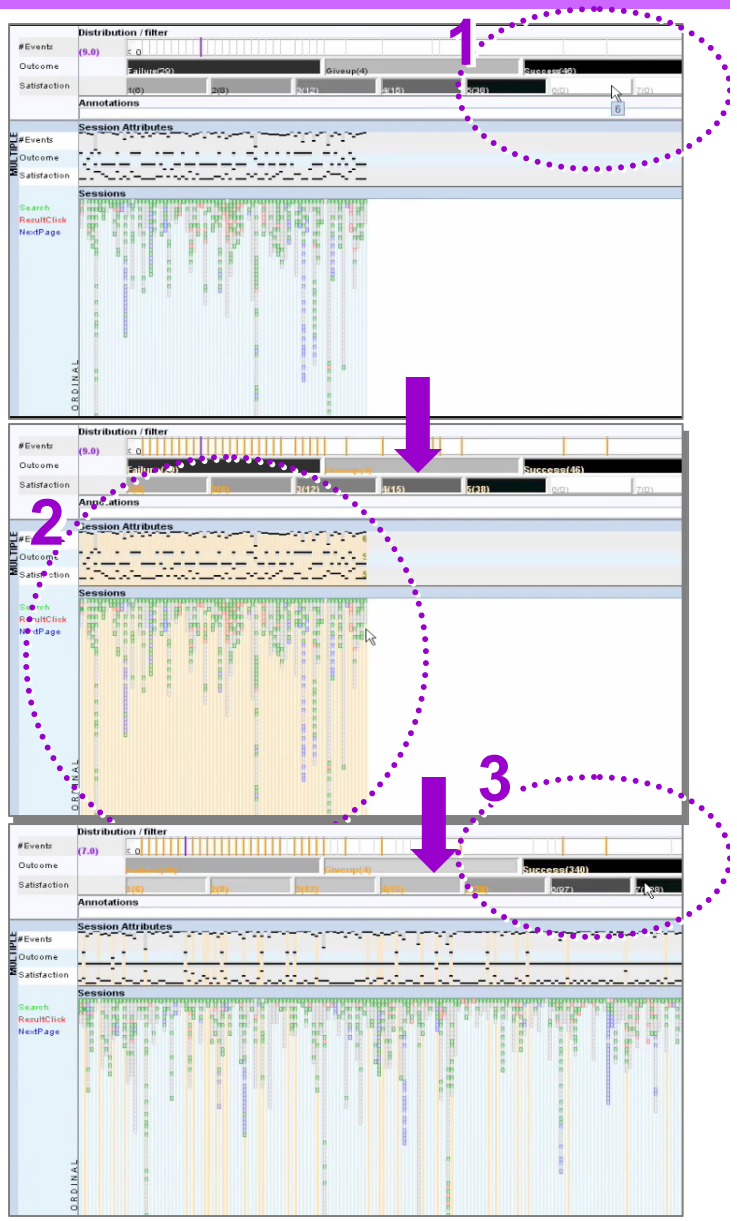
Patterns found:

- Search-engine searches
- 3rd Party searches
- Result explorations

Use-case scenario: Correlations

- Does participant's outcome correlate with their satisfaction score?
  - Outcome (Success/Failure/Given-up)
    Score (1 – 7, 7=most satisfied)

1. Remove sessions with scores of 6 and 7
2. Highlight remaining sessions (<= 4)
3. Add the high-score sessions back
4. Reorder by Task Outcome
5. Highlighted low-score sessions clustered at the poor-outcome end suggests high correlation

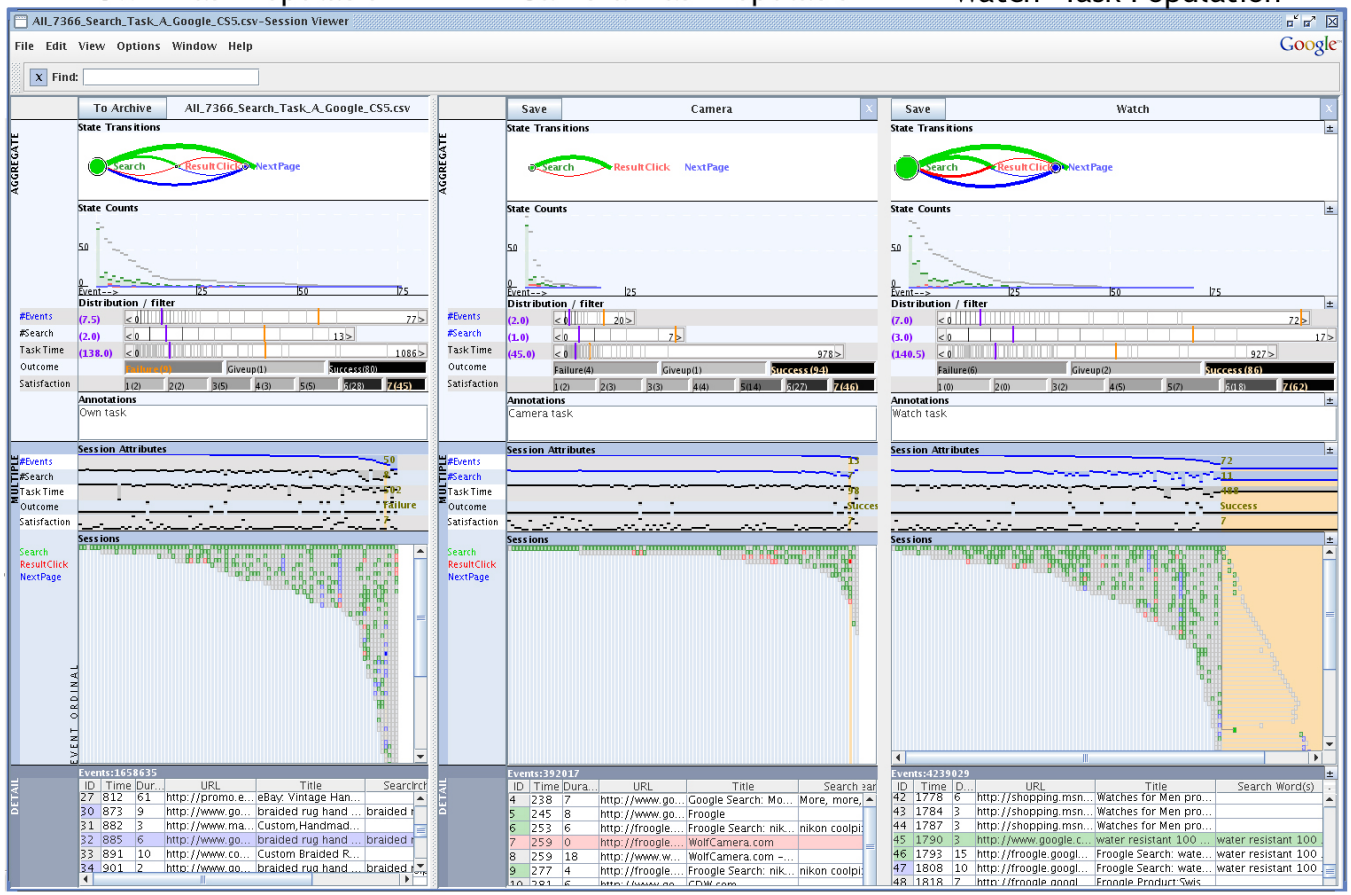


Session Viewer and Analysis: Multiple Populations

- **Provides multiple views for side-by-side comparisons to quickly get a sense of population characteristics based on aggregates**

- Camera-Task population looks different

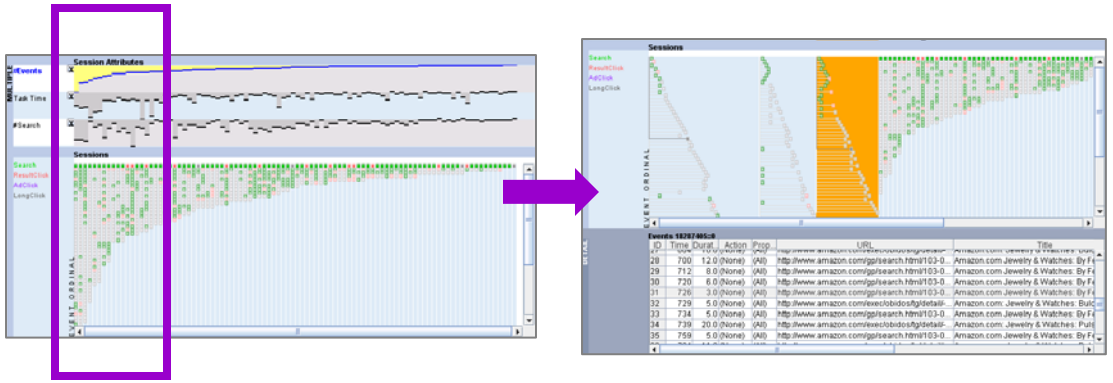
"Own" Task Population "Camera" Task Population "Watch" Task Population



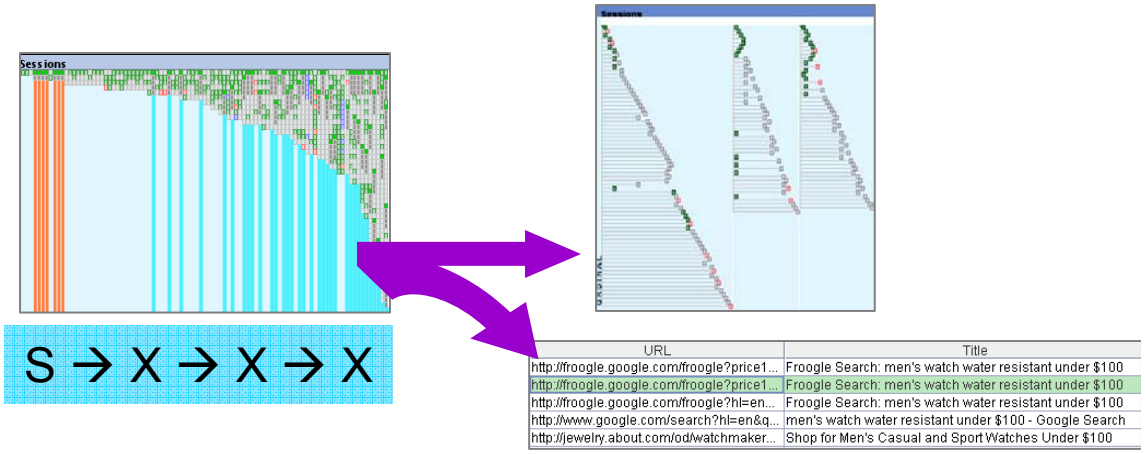
Session Viewer and Analysis: Multiple Levels

- **Displays multiple data levels to bridge between the high-level statistical and low-level detailed analysis approaches**

- Detailed analysis:
  Reordered session helps select interesting sessions to focus analysis



- Statistical analysis:
  Linked views provides content and context to interpret results



Summary

- Web session log analyses tend to take either the high-level statistical or low-level detailed approach
- Session Viewer provides a visual analytic tool that
  - Bridges between the two analysis levels and facilitate cross-level analysis with multiple data levels
  - Provides side-by-side population comparisons with multiple views
- We found that task type is an important factor in session analysis, and task nature is reflected in different event-sequence patterns in sessions
- In a later work, we conducted a field study with 7 analysts where they used their own data and tasks with Session Viewer

Acknowledgements

- The project was funded by Google as internships and a university grant

- I’d like to thank Dan Russell and Diane Tang for hosting me at Google