© Microsoft 2012

# User task understanding: a web search engine perspective

## NII Shonan: Whole-Session Evaluation of Interactive Information Retrieval Systems

**Peter Bailey** Microsoft Bing October 2012

© Microsoft 2012

# 3 years of Bing personalization

- Interactive Information Retrieval Systems are conditioned by a user's **interaction** with them
- Whole session evaluation is a sub-class of **personalization** evaluation

© Microsoft 2012

# 3 years of Bing personalization

- User **modeling** matters more than "topical" relevance modeling
- User modeling is at least modeling:
  - user need
  - user behavior
  - user satisfaction / effectiveness

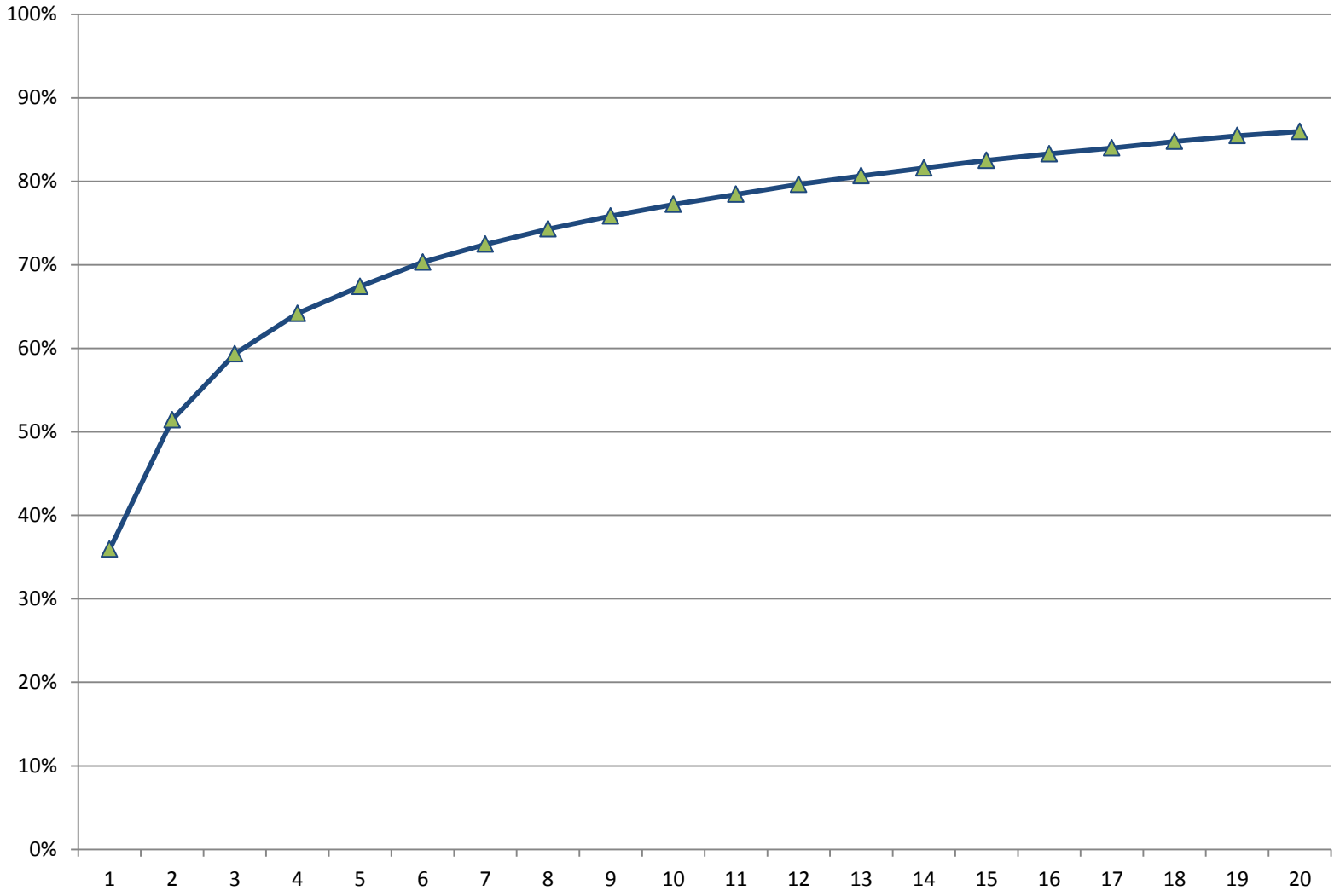
© Microsoft 2012

# 3 years of Bing personalization

- In a reductive evaluation framework, measures need to assess effectiveness of how user modeling changes the relevance / retrieval performance of an underlying system
  - To achieve **statistical power**, large numbers of users required
- Other evaluation systems may be just as (or more) informative

© Microsoft 2012

# Probability of issuing another query, given session of length *N*-1



© Microsoft 2012

# Understanding online search and browsing behavior

- Developed a new task taxonomy for Web *browsing* behavior (Aug-Oct 2009)
- Used iterative taxonomy development as per Rose & Levinson, Yahoo!, WWW'04 (Understanding user goals in web search)
  - Panel of ~5 in-house judges, co-developed taxonomy
- *Verb*-based, not domain- or search-activity based; phrased as action-topic pairs
- 26 high level tasks

© Microsoft 2012

# Raw data used for analysis – statistics

| | All | Query focused | Google/Yahoo/Bing |
|---|---|---|---|
| Number of events | 41,493 | 23,054 | 20,769 |
| Number of users | 187 | 186 | 179 |
| Number of sessions | 453 | 451 | 412 |
| Number of tasks | 1913 | 772 | 676 |
| Avg events per session | 91.6 | 51.1 | 50.4 |
| Avg events per task | 18.4 | 27.1 | 28.1 |
| Avg tasks per session | 4.2 | 1.7 | 1.6 |

© Microsoft 2012

# Task taxonomy elements

- Download files
- Compare products or services
- Buy a product
- Sell a product
- Find a job
- Learn how to perform a task
- Monitor frequently updated information
- Keep tabs on recent events
- Find an address or contact information
- Find facts about a person
- Do education related tasks online
- Discover leisure activities
- Find a specific miscellaneous fact
- Discover more information about a specific topic
- Find a date online
- Find real estate
- Browse a social network
- Read or write on blog or forum
- Plan travel
- Plan event
- Watch online videos for entertainment
- Play online games
- Listen to online music
- Manage personal accounts
- Communicate with people online
- Adult

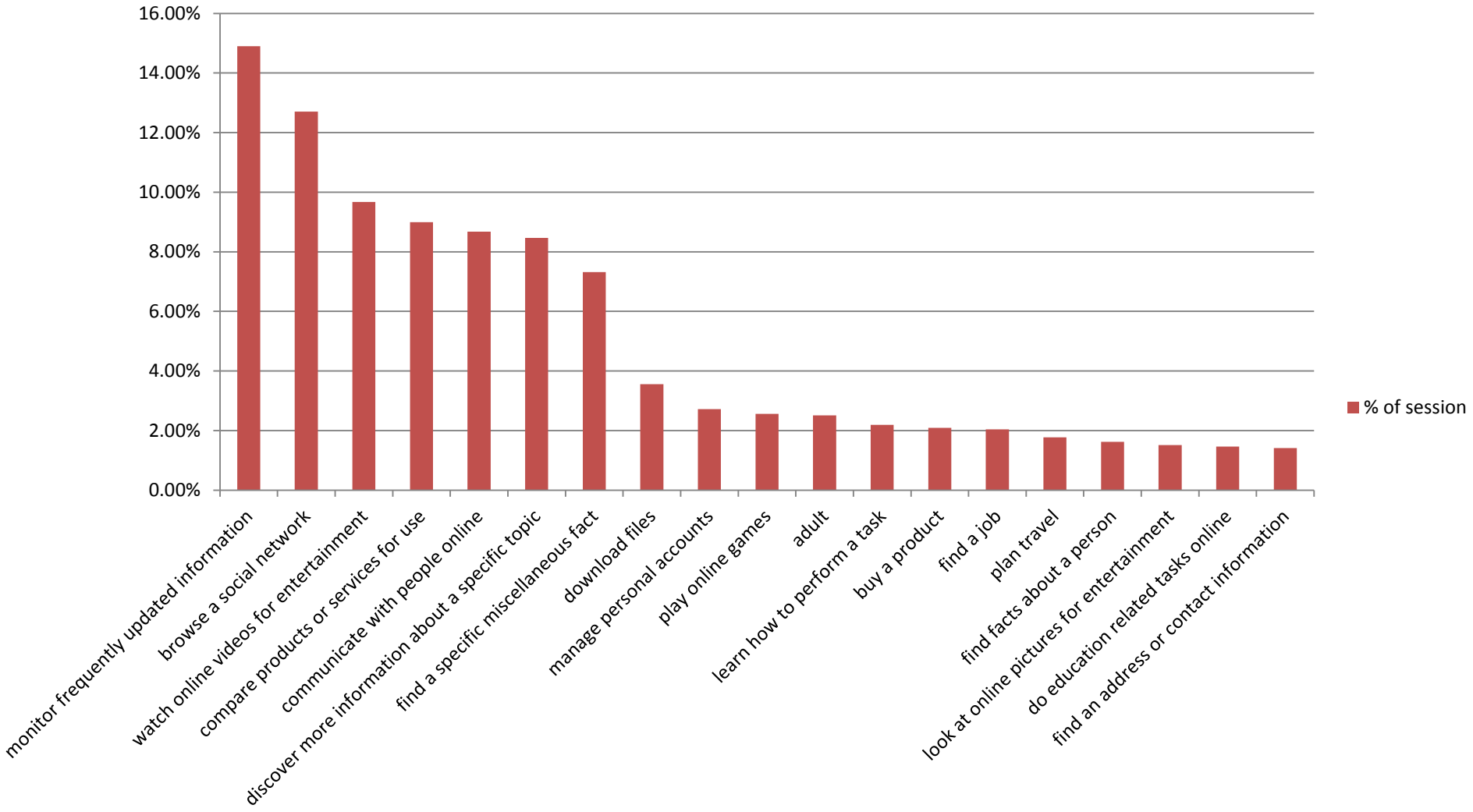
© Microsoft 2012

# Compare Broder '02 & Russell *et al* '09

| [Broder 2002] | [Google 2009] | Action | Topic |
|---|---|---|---|
| *Informational* | *Find-Simple* | Find | a specific miscellaneous fact |
| | | Find | an address or contact information |
| | *Find-Complex* | Find | a date online |
| | | Find | a job |
| | | Find | facts about a person |
| | | Find | real estate |
| | *Explore/Learn* | Browse | a social network |
| | | Compare | products or services for use |
| | | Discover | more information about a specific topic |
| | | Discover | leisure activities |
| | | Learn | how to perform a task |
| | | Plan | event |
| | | Plan | travel |
| *Transactional* | *Locate/Acquire* | Buy | a product |
| | | Download | files |
| | | Plan | event |
| | | Plan | travel |
| | | Sell | a product |
| | *Play* | Look | at online pictures for entertainment |
| | | Watch | online videos for entertainment |
| | | Listen | to online music |
| | | Play | online games |
| | *Meta* | | |
| *Navigational* | *Navigate* | | |
| *Other* | *Other* | Communicate | with people online |
| | | Do | education related tasks (online homework, etc.) |
| | | Manage | personal accounts |
| | | Read or write | on blogs or forums |

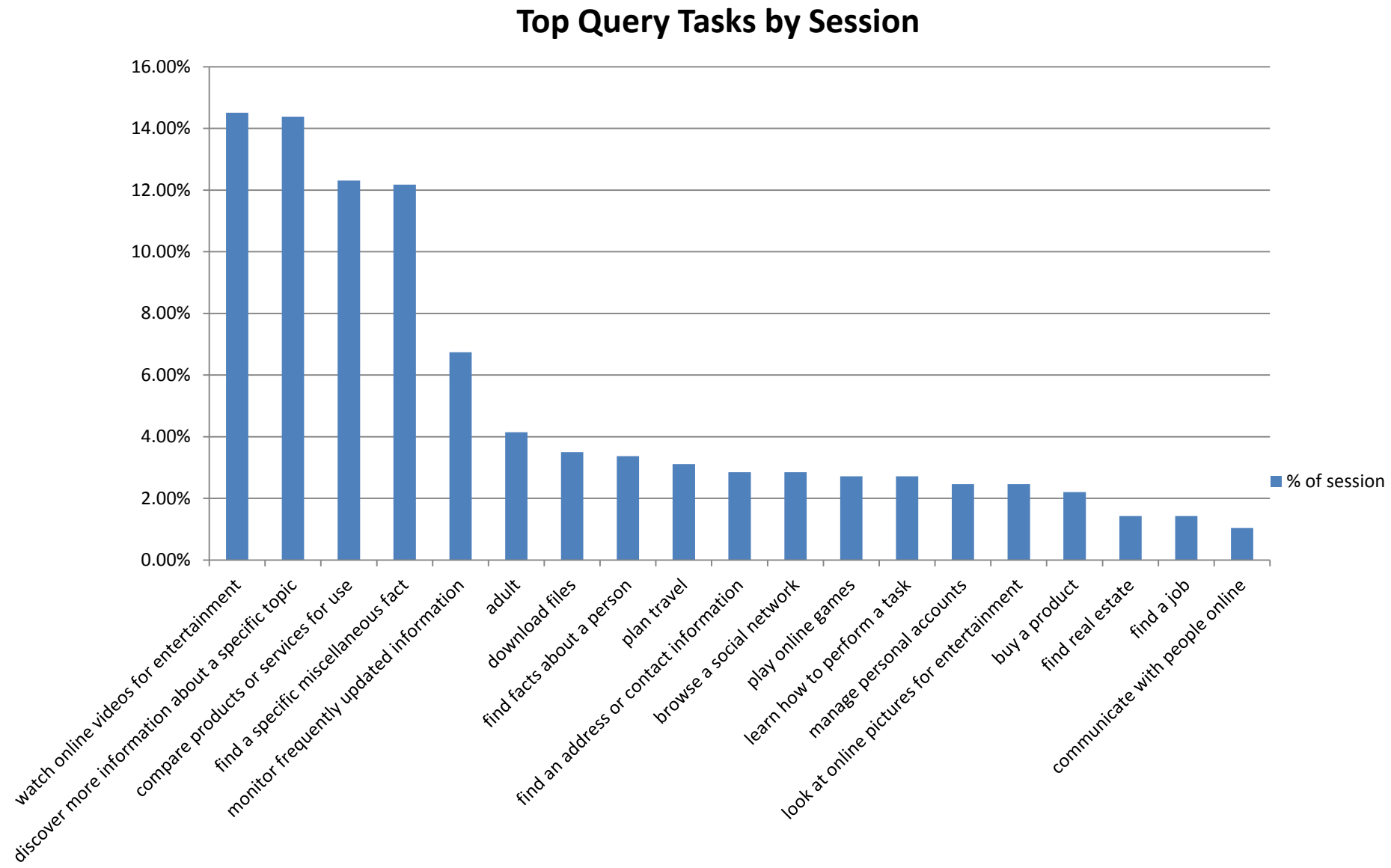
© Microsoft 2012

# Top Web Tasks by Session




Definition: All web activities including browsing behavior and search behavior

© Microsoft 2012




Definition: Query tasks are contiguously labeled tasks within a session which contain a query issued to Google, Yahoo or Bing

© Microsoft 2012

# Task properties

| Task | Queries per task | Avg. events per task | Avg. length (mins) |
|---|---|---|---|
| *adult* | 18.7 | 18.3 | 48.6 |
| *look at online pictures for entertainment* | 16.2 | 50.7 | 15.7 |
| **learn how to perform a task** | **13** | **11.9** | **8.5** |
| *download files* | 11.7 | 31.2 | 15 |
| *watch online videos for entertainment* | 7.5 | 19.5 | 19 |
| **find facts about a person** | **6.9** | **18.9** | **4.8** |
| **discover more information about a specific topic** | **6.8** | **24.8** | **13.5** |
| **compare products or services for use** | **6.8** | **22.3** | **24.8** |
| *find real estate* | 5.1 | 11.7 | 14.9 |
| plan travel | 4.7 | 5.1 | 12 |
| find an address or contact information | 4.2 | 48.5 | 7.5 |
| not a task | 3.9 | 10.3 | 21.6 |
| monitor frequently updated information | 3.6 | 24.1 | 20.6 |
| find a specific miscellaneous fact | 3.2 | 40 | 7.9 |
| buy a product | 3.1 | 15.4 | 8.9 |
| play online games | 2 | 21.1 | 16.2 |
| manage personal accounts | 1.8 | 42 | 7.5 |
| find a job | 1.8 | 29.4 | 18 |
| communicate with people online | 1.8 | 9 | 5.6 |
| browse a social network | 1.5 | 7.1 | 24.7 |

© Microsoft 2012

# Task properties

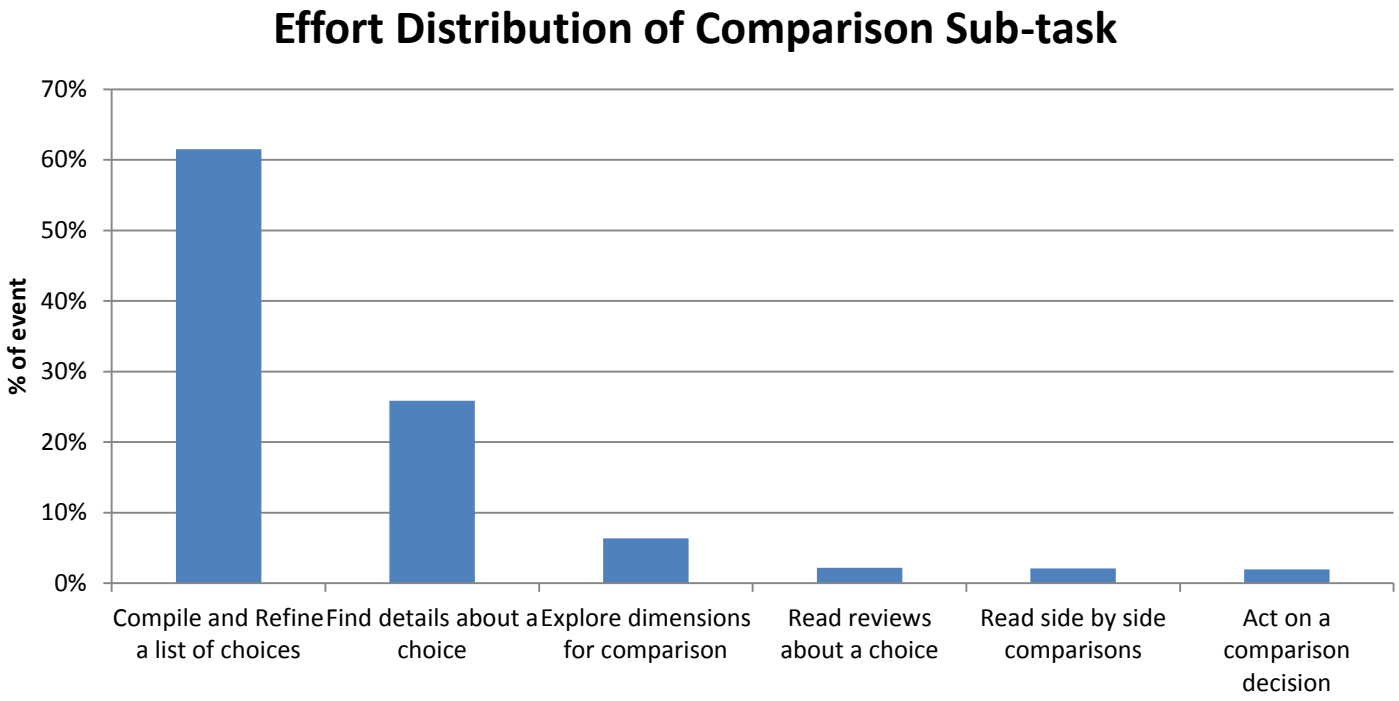
| Task | Queries per task | Avg. events per task | Avg. length (mins) |
|---|---|---|---|
| *adult* | 18.7 | 18.3 | 48.6 |
| *look at online pictures for entertainment* | 16.2 | 50.7 | 15.7 |
| **learn how to perform a task** | **13** | **11.9** | **8.5** |
| *download files* | 11.7 | 31.2 | 15 |
| *watch online videos for entertainment* | 7.5 | 19.5 | 19 |
| **find facts about a person** | **6.9** | **18.9** | **4.8** |
| **discover more information about a specific topic** | **6.8** | **24.8** | **13.5** |
| **compare products or services for use** | **6.8** | **22.3** | **24.8** |
| *find real estate* | 5.1 | 11.7 | 14.9 |
| plan travel | 4.7 | 5.1 | 12 |
| find an address or contact information | 4.2 | 48.5 | 7.5 |
| not a task | 3.9 | 10.3 | 21.6 |
| monitor frequently updated information | 3.6 | 24.1 | 20.6 |
| find a specific miscellaneous fact | 3.2 | 40 | 7.9 |
| buy a product | 3.1 | 15.4 | 8.9 |
| play online games | 2 | 21.1 | 16.2 |
| manage personal accounts | 1.8 | 42 | 7.5 |
| find a job | 1.8 | 29.4 | 18 |
| communicate with people online | 1.8 | 9 | 5.6 |
| browse a social network | 1.5 | 7.1 | 24.7 |

© Microsoft 2012

# Deep dive: Compare task

Comparison Task broken down into the following sub-tasks:

- Explore dimensions for comparison (size, color, capacity, megapixel)
- Compile and refine a list of choices (comparable models)
- Find details about a choice
- Read reviews about a choice
- Read side by side comparisons
- Act on a comparison decision



© Microsoft 2012

# Questions

- How would we select one or more tasks as a community?
  - TREC Interactive style task setup?
  - In-situ user-studies?
- Are task-customized search interaction systems transferrable to other kinds of task?
  - Cf. 10 blue links “command line interface”
  - Which tasks would a “discover more information” system generalize too?

© Microsoft 2012

# Key takeaways

- Many information needs in web search logs consist of **multiple** queries
- These are **complex** tasks where users conduct various subtasks
- Significant **time** is dedicated to these tasks indicating they may be more important or challenging to a user
- **User (esp. User Task) modeling** is critical

© Microsoft 2012

# Acknowledgments

Microsoft Bing user task understanding v-team

Peter Bailey, Liwei Chen, Scott Grosenick, Li Jiang, Yan Li, Paul Reinholdtsen, Charles Salada, Haidong Wang, Sandy Wong