Supporting Cognitive Models of Sensemaking in Analytics Systems

Jason Perry
Christopher Janneck
William M. Pottenger

Outline

I. The Interactive Automation Project
II. Sensemaking
III. Learning Sensemaking States
IV. Sensemaking Ontologies
V. An Artifact-based UI
VI. Experimental Setup

I. The Interactive Automation Project

Interactive Automation

"The Interactive Automation Project seeks to provide a computational framework that facilitates exploring and analyzing large amounts of distributed data, in a new and interactive way…"

- NSF Funded
  - http://dydan.rutgers.edu/PDDALab/IA/IA_Site/Welcome.html
- Distributed Data Analytics (DDA) + Data Visualization + Computational Steering + UI design
- A Component-based Framework

…development should also be driven by research in analyst cognitive processes.

II. Sensemaking

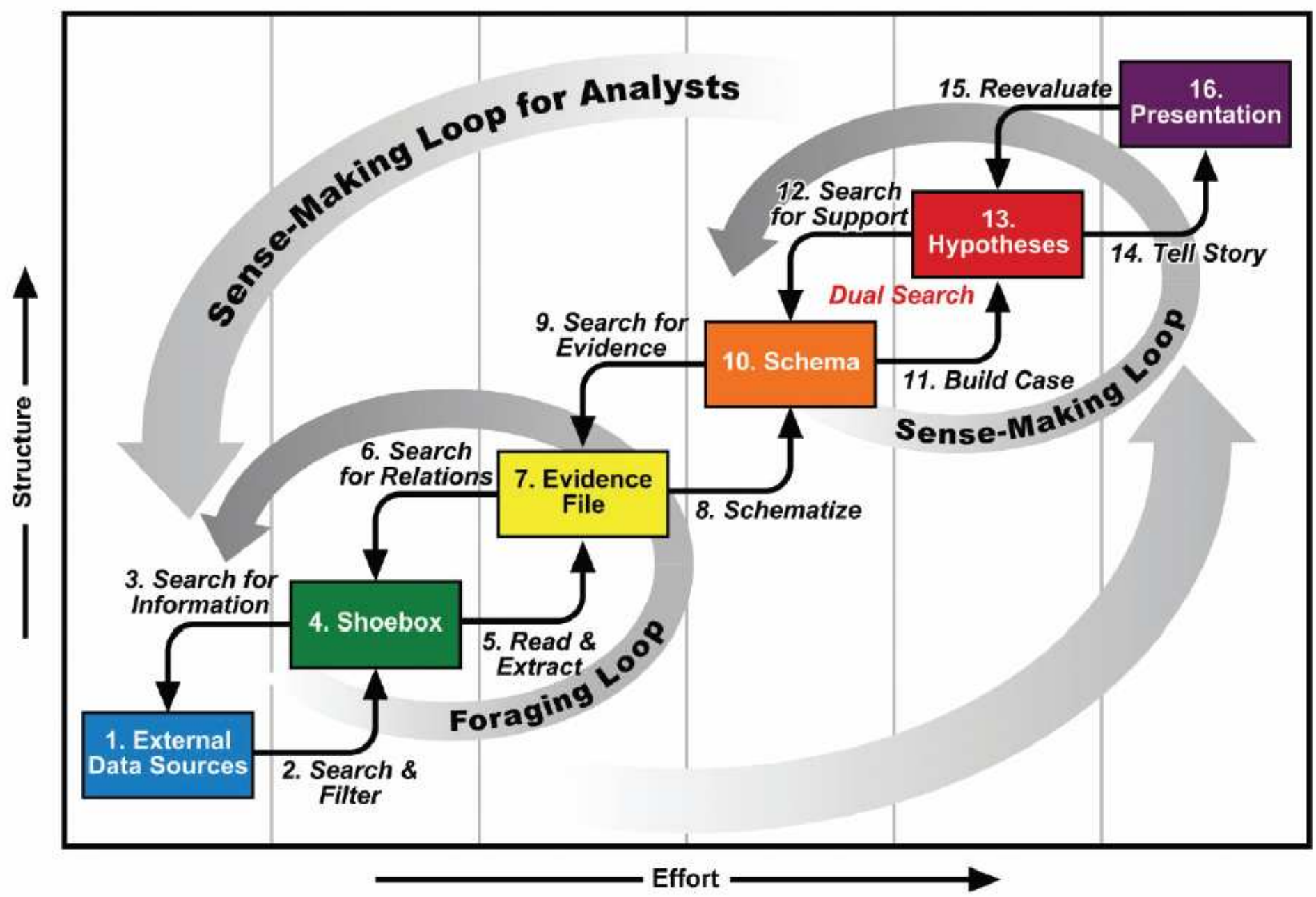
Sensemaking

- The human task of sorting through a large amount of information to aid situational awareness and decision-making
- “A motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively” (Klein et al, 2006a).
- Described by Russell as “a process of evolving representations”
- Involves many cognitive and perceptual activities
- An active area of research in Cognitive Science

Pirolli and Card's Model

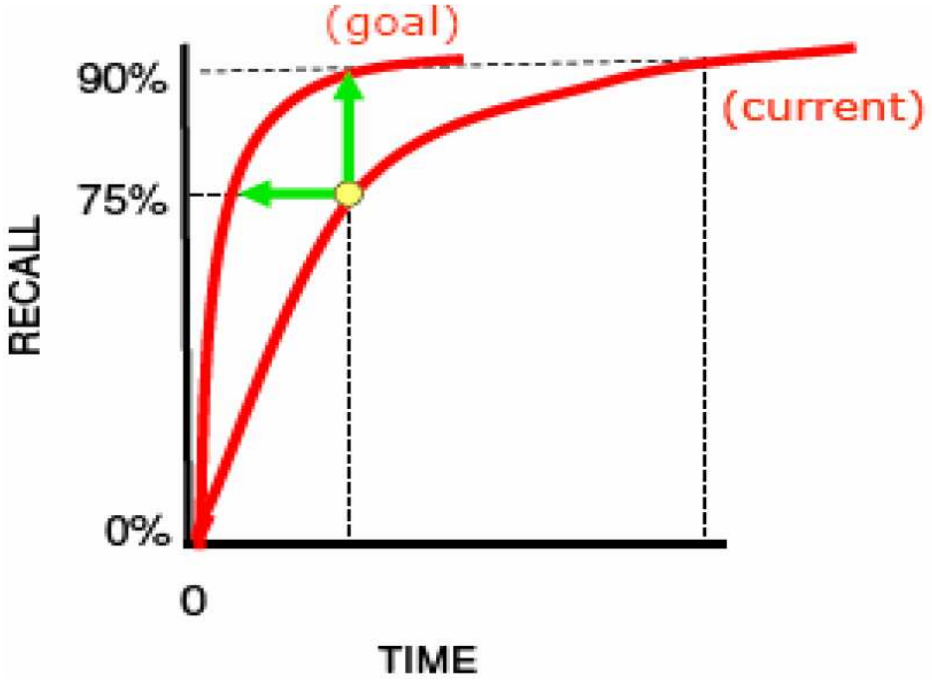
- Centered on intelligence analysis
- Based on Empirical studies (Cognitive Task Analysis)
- Capturing the schematic knowledge structures of expertise
- Research has also produced Cost Structures based on cognitive principles

Pirolli and Card's Sensemaking Loop



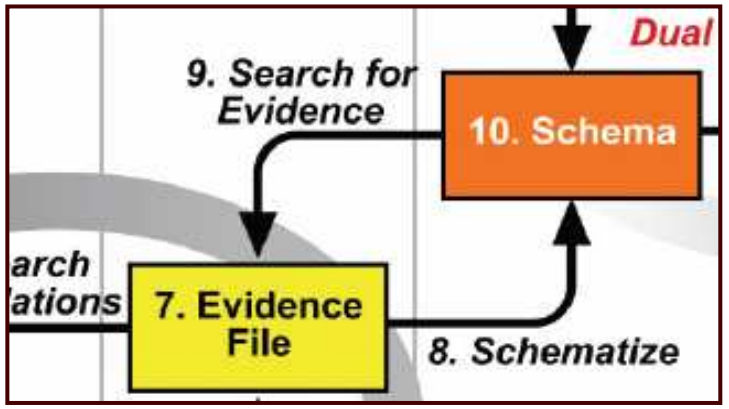
Potential of Using Sensemaking

- Cost structure analysis of points in the sensemaking loops
- Minimize impact of confirmation bias, working memory load, shifting attention, and other effects of cognitive limits
  - Foraging: Optimize exploration / exploitation tradeoff
  - Sensemaking: adapt to working memory limits in schema construction



Example Sensemaking UI Guidance Scenario

1. Investigator is in the midst of analyzing a case
2. Model indicates that user has been in foraging loop for a certain number of cycles
3. Guidance rules suggest that user has reached a point of diminishing returns in searching / sorting
4. UI brings a new set of tools to the forefront to suggest schematizing existing data



Related Work

Sensemaking:

- Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). “The Cost Structure of Sensemaking”. Paper presented at the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam.
- Pirolli, P. and Card, S. “The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis”. In *Proceedings of 2005 International Conference on Intelligence Analysis*.

Sensemaking Support for Analytics Systems:

- Shrinivasan, V. and van Wijk, J. J. “Supporting the Analytical Reasoning Process in Information Visualization”. CHI 2008 Sensemaking Workshop
- Gotz, David. “The ScratchPad: Sensemaking Support for the Web”. WWW 2007 Poster Paper.

http://dmrussell.googlepages.com/sensemakingworkshoppapers

III. Learning Sensemaking States

What We Need to Be Able to Do

- To make the Sensemaking model useful for Interactive Automation, we need a method for inferring knowledge of the sensemaking states from the record of user input.

Machine Learning to Infer Sensemaking States

- Record the sequence of user interactions in an analysis system
- Generate training set by manually labeling the transitions between Sensemaking stages that analysts have passed through
- Train a sequence model (e.g. HMM, CRF, Markov Net) to learn transitions between sensemaking states

Is it possible?

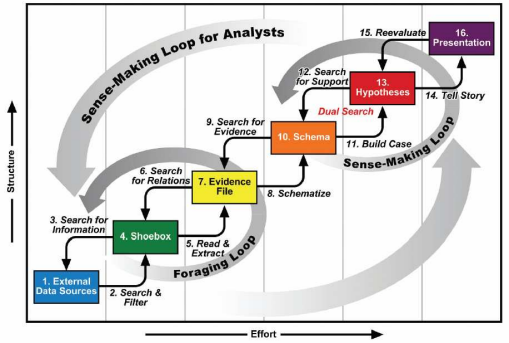
IV. Sensemaking Ontologies

Ontologies and Reasoning in Description Logics

- A type of knowledge representation language
- Represent the terminology and concepts of a domain
- Syntax consists of concepts (classes) and roles (properties)
- Primary reasoning task is subsumption
- Becoming a key component of the semantic web through OWL ontologies (OWL-DL)
- Identified as a necessary component to manage information flow in IA

By representing the concepts of sensemaking through a description language, we seek to build a flexible foundation for sensemaking support in software, as well as for further investigation of sensemaking models.

Defining Data Tasks



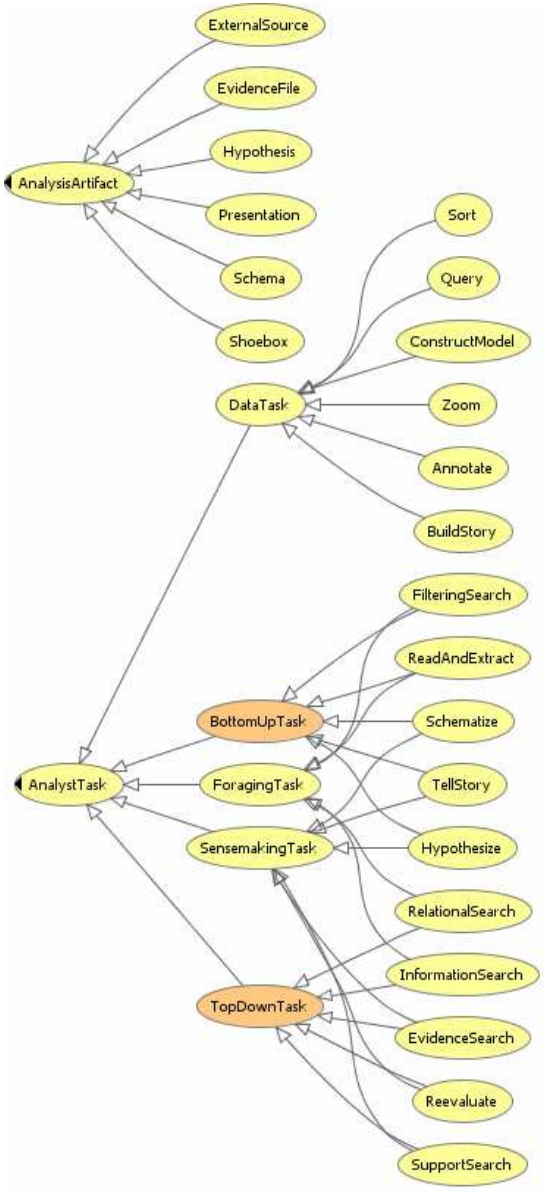
- Define the set of possible UI tasks
- These can be directly recorded from the UI
- Sorted into categories, find mapping to Sensemaking states

Data Tasks vs Sensemaking Tasks

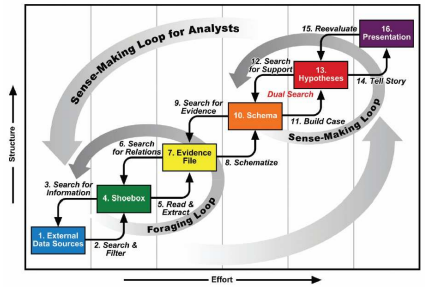
Data Task \ Sensemaking Task	Filtering Search	Information Search	Read and Extract	Relational Search	Schematize
Query DB	☑	☐	☐	☐	☐
Query Web	☑	☐	☐	☐	☐
Query FS	☑	☐	☐	☐	☐
Query DB (context-based)	☐	☑	☐	☐	☐
Query Web (context-based)	☐	☑	☐	☐	☐
Query FS (context-based)	☐	☑	☐	☐	☐
Add datasource	☑	☑	☐	☐	☐
Remove datasource	☐	☑	☐	☐	☐
Select snippet	☐	☐	☑	☐	☐
Extract snippet (to Evidence File)	☐	☐	☑	☐	☐
Open datasource	☐	☐	☑	☐	☐
Scan datasource	☐	☐	☑	☐	☐
Query for related snippets (context-based, in Shoebox)	☐	☐	☐	☑	☐
Query for similar snippets (context-based, in Shoebox)	☐	☐	☐	☑	☐

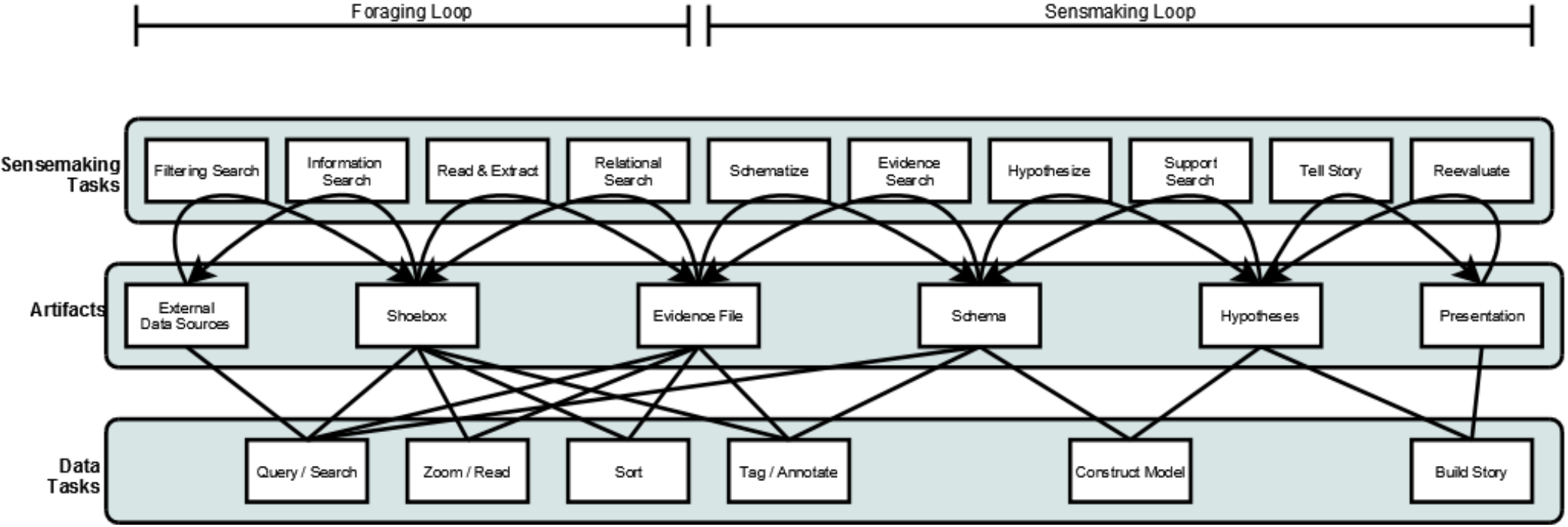
DL Expression of Sensemaking Structures

- OWL-DL
- Classes corresponding to Sensemaking tasks, artifacts, data tasks
- Model the hierarchical relationship among states using class properties
  - TopDownFrom, BottomUpFrom, isUsedIn
- Shows which higher-level states can correspond to lower-level functions.
- Record of user interactions *and* high-level states created at runtime by instantiation of objects
- …current set of possible Sensemaking states can be queried at runtime.



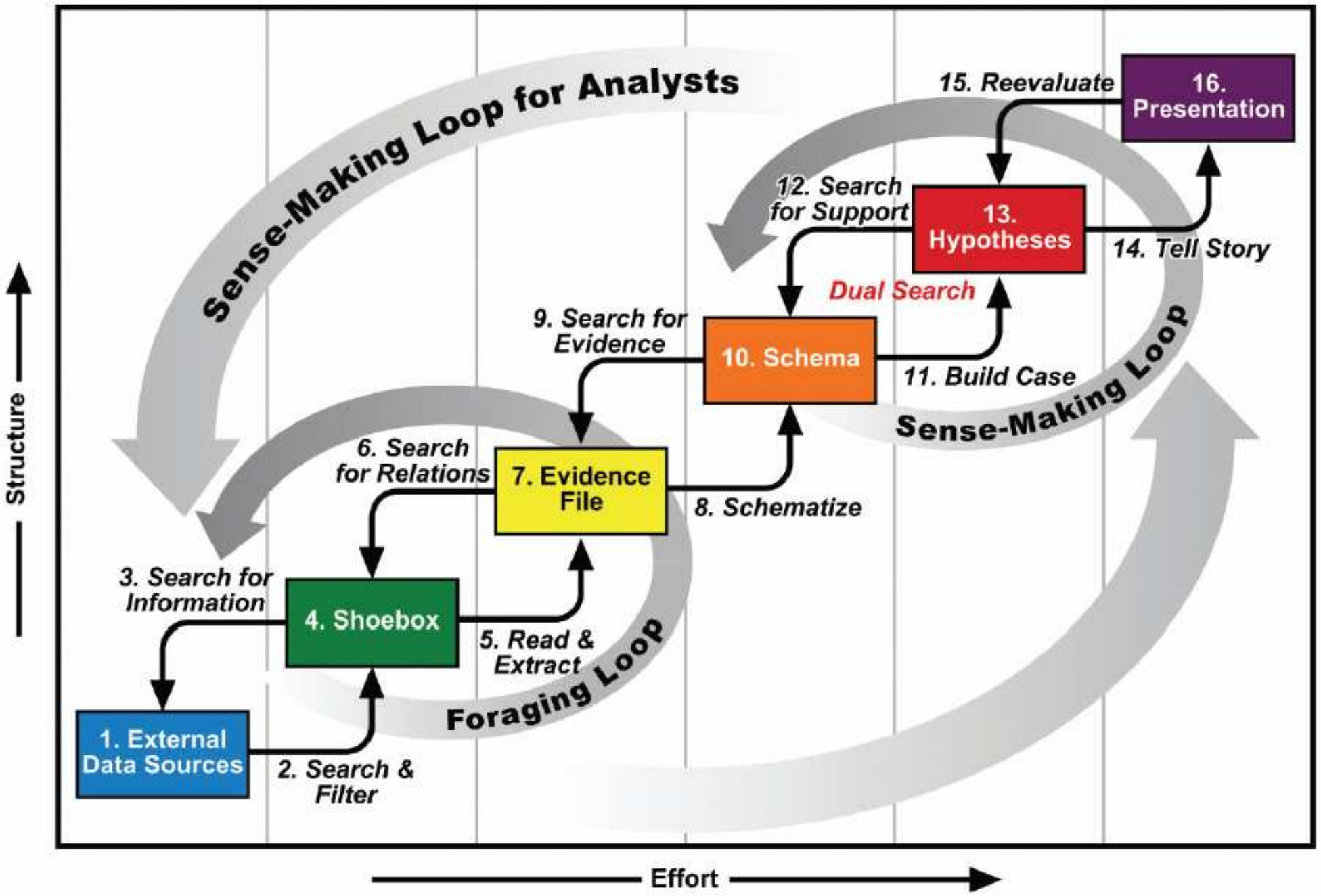
Decomposition of Tasks and States





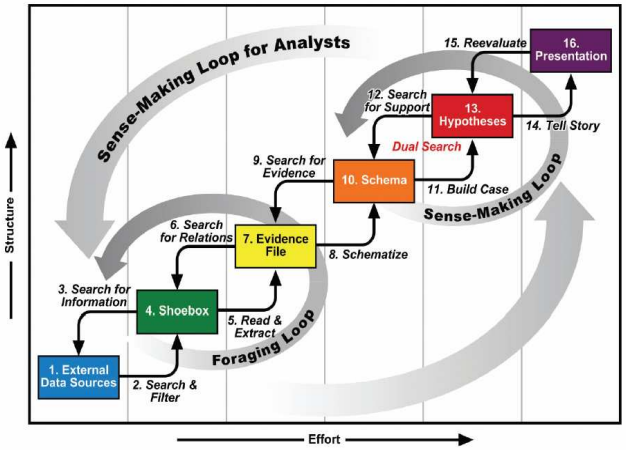
Computing Entailments to Define UI Learning Tasks

- Computing entailments from the DL representation shows which tasks are possible in which states.
- Using data tasks alone, entailments only slightly reduce the space of possibilities
- ***But if we were able to know which artifact the analyst is using, we know that at any point the user is in one of only two possible sensemaking stages.***


Sense-Making Loop for Analysts
15. Reevaluate
16. Presentation
12. Search for Support
13. Hypotheses
14. Tell Story
Dual Search
9. Search for Evidence
10. Schema
11. Build Case
Sense-Making Loop
6. Search for Relations
7. Evidence File
8. Schematize
3. Search for Information
4. Shoebox
5. Read & Extract
Foraging Loop
1. External Data Sources
2. Search & Filter
Structure
Effort

Artifacts to Sensemaking States

- Defines a reduced set of learning problems

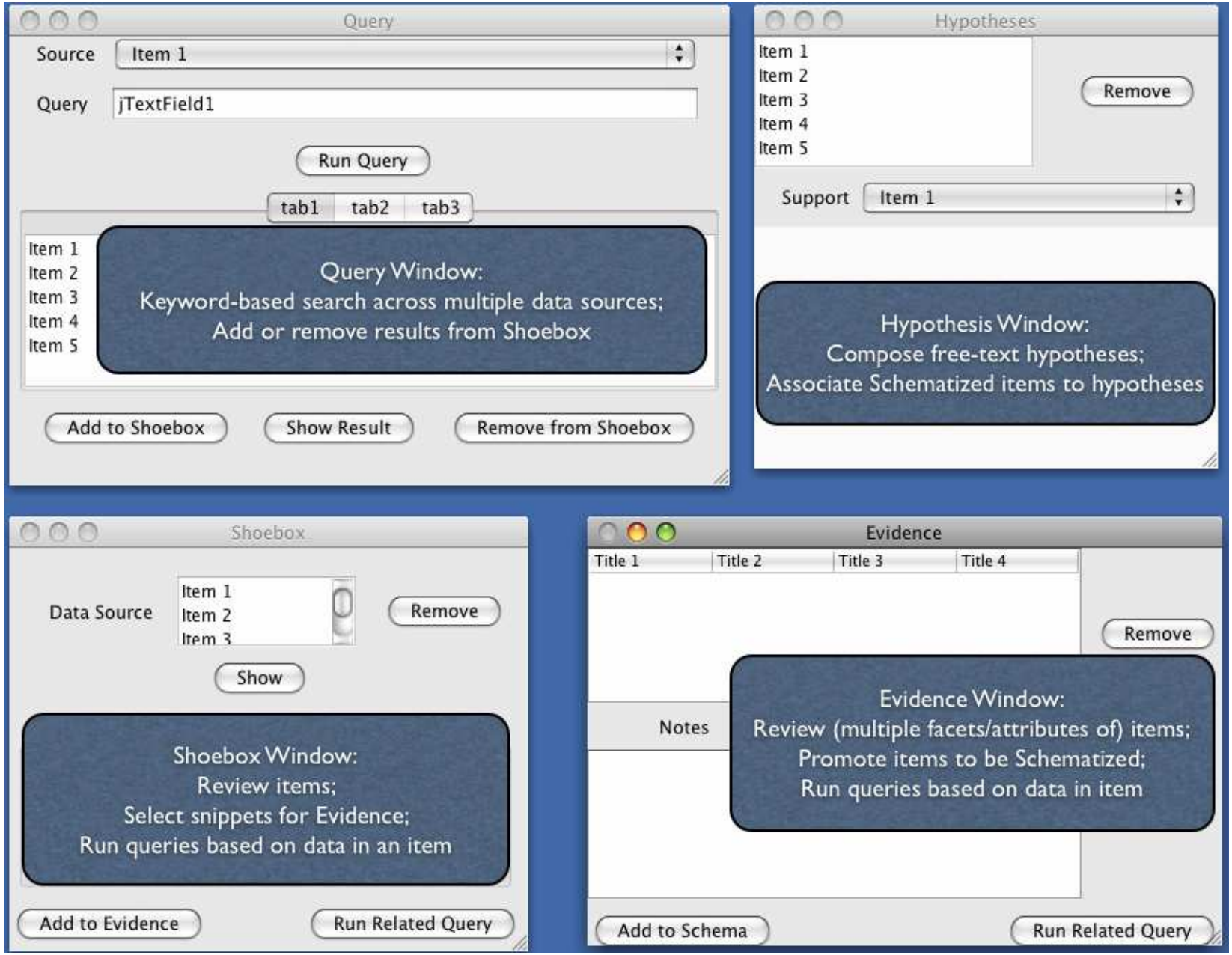


Data Task/Artifact Mappings for Sensemaking States

Data Task (verb)	Artifact (object)	Potential States
Query	ExternalSource	FilteringSearch, InformationSearch
Query	Shoebox	ReadAndExtract, RelationalSearch
Query	EvidenceFile	Schematize, EvidenceSearch
Query	Schema	Hypothesize, SupportSearch
Zoom	Shoebox	ReadAndExtract, RelationalSearch
Zoom	EvidenceFile	Schematize, EvidenceSearch
Sort	Shoebox	ReadAndExtract, RelationalSearch
Sort	EvidenceFile	Schematize, EvidenceSearch
Tag	EvidenceFile	Schematize, EvidenceSearch
Tag	Schema	Hypothesize, SupportSearch

V. An Artifact-Based UI

Artifact-Based UI



Top-down & Bottom-up

- In this framework, each learning problem is an instance of determine whether the user is in a "bottom-up" (data to theory) or "top-down" (theory to data) stage.
- A recurring theme in the literature [Qu and Furnas 2005]: "representation construction and information seeking are closely coupled, as people get aspects of structure *top down* deducing from their previous knowledge, *bottom up* inducing from facts they find…"

VI. Experimental Design

Experimental Design

- Patterned after previous analytics system evaluations [Morse et al 2008], with qualitative and quantitative metrics
- Bethlehem Police Department ground truth – solved gang-related murder case
- First pilot study focused on data collection methods and system parameter adjustment

Experimental Design

- Second study: Baseline group with interface only; treatment group with assistance engine
- Hypothesis: Treatment group will be seen to have more effective analytical results
  - Metric: greater percentage of extracted search items will be applied in the final hypothesis
  - Metric: time to arrive at correct conclusion

Conclusions

- Found a way to simplify the learning task through artifact-based UI design
  - It might be a good thing to explicitly represent cognitive structures in the UI
  - The problem of inferring cognitive states is addressable when we combine templatized domain knowledge with machine learning
  - Intersection of knowledge representation and machine learning
- The literature is researching a consensus that cognitive structures come about through the interplay of top-down and bottom-up processes

Future Work

- Near term: carry out experiments
- Long-term goal: learning from non-adapted GUI, subtle yet demonstrably effective user feedback (“expert mode”)
- Expand to other workflow models using the same knowledge representation + learning approach
  - Klein’s Data/Frame Sensemaking Model
- A new realization of Computational Steering for data analytics

Thank you!