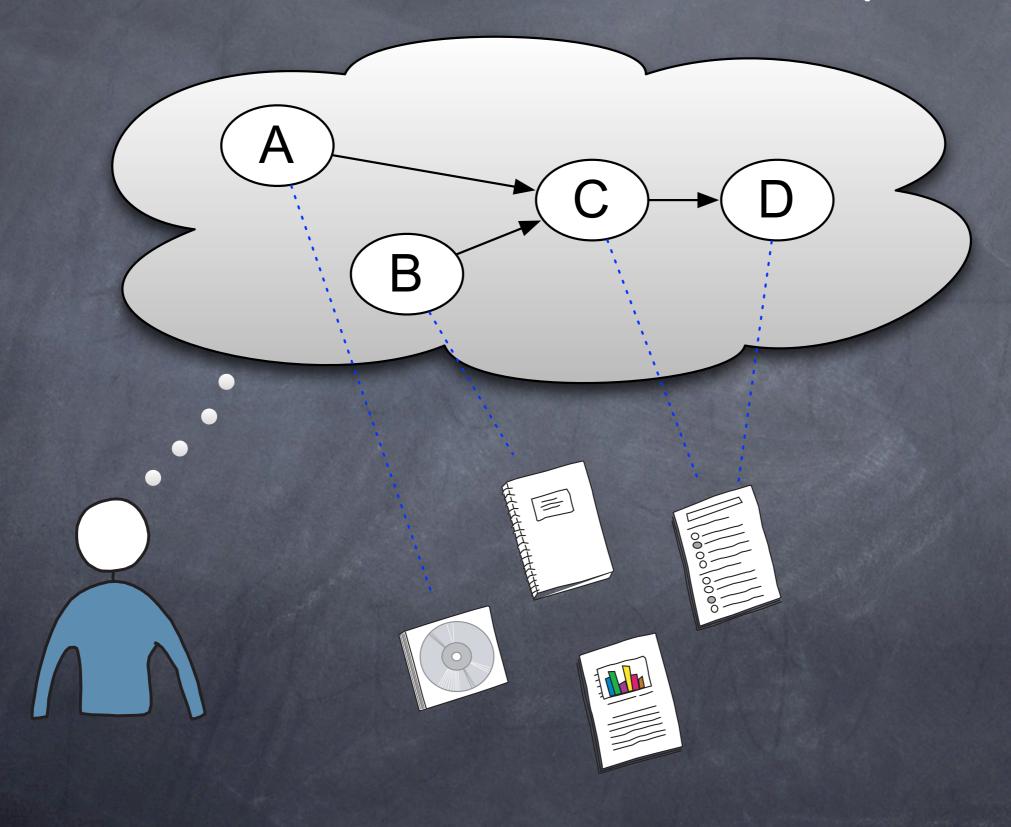
# An introduction to content analysis

Kevin Crowston

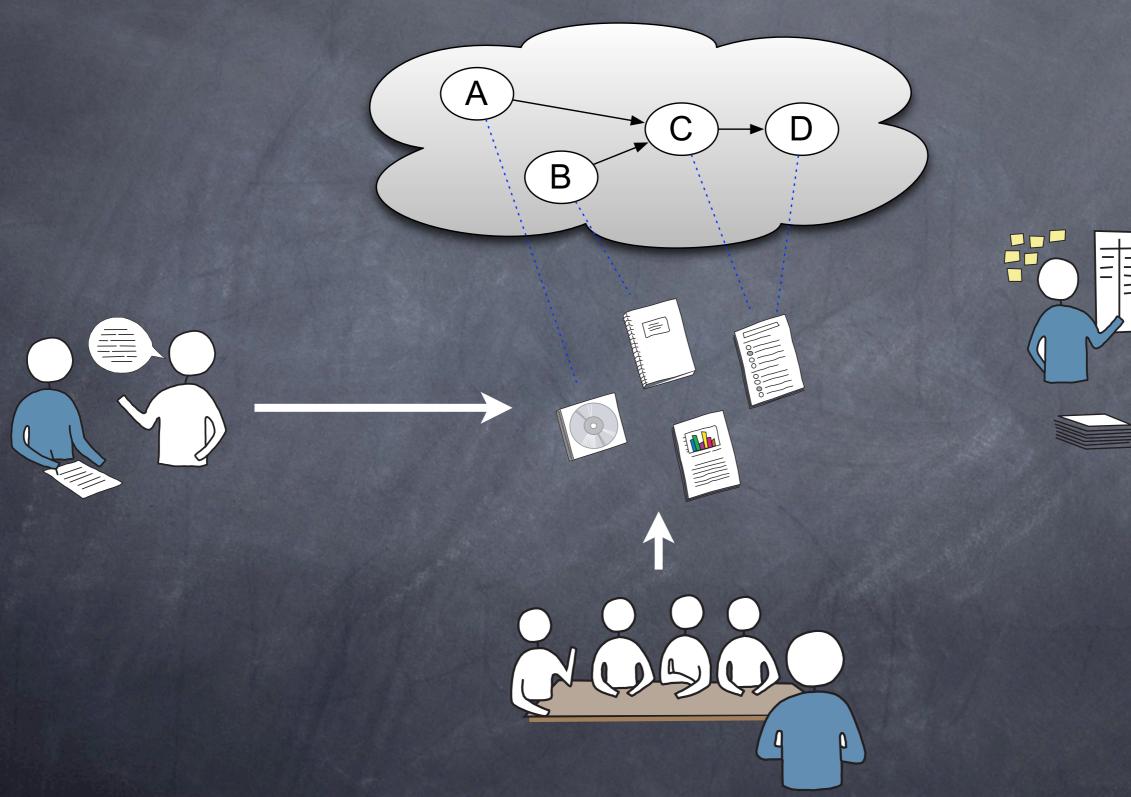
Syracuse University crowston@syr.edu

School of Information Studies http://crowston.syr.edu/

### What is content analysis?

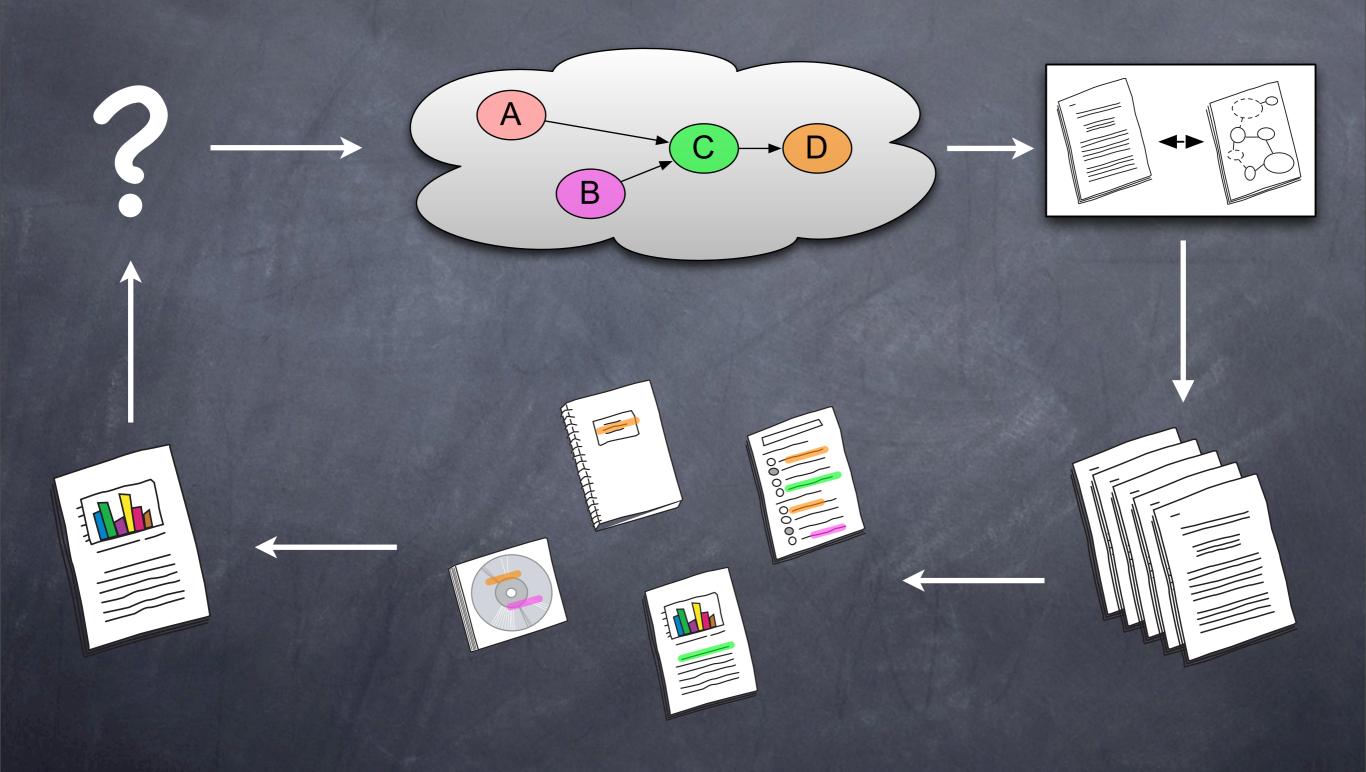


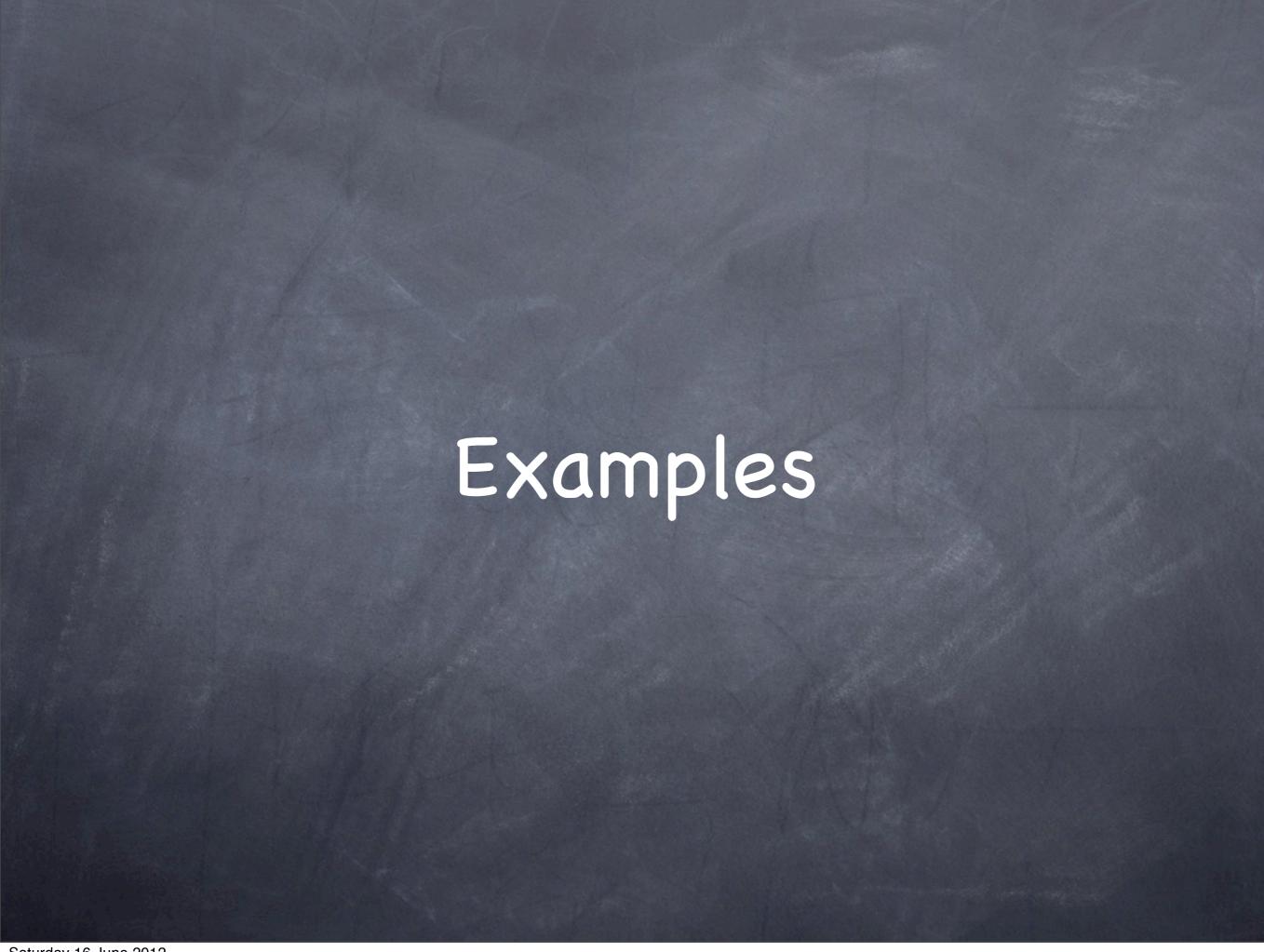
### What is content analysis?





### Process of content analysis





### Example document

Hope everyone is having a good week. :-)

I have been working on paring down Defines.h for awhile now. Jason and I discussed it several months ago, and I only just got around to it. What I did was break out everything that was in only one file. These are things that shouldn't have been put in Defines.h to begin with.

It looks like I changed 28 files. The changes should be simply cosmetic and shouldn't affect functionality. However, since this is such a large change, I wanted to run the changes by everyone before I did something stupid that would have to be backed out.

And so, I give you, the Ginormous Diff From Hell[TM], and the binary that goes along with it: <a href="http://www.kocharhook.com/nick/fire/diff.html">http://www.kocharhook.com/nick/fire/diff.html</a>

Take a look at the diff, maybe try out some of the changes in the binary. Let me know if you see anything amiss.

If I get no comments, I'm going to check this in tomorrow.

Nick Kocharhook -- <avpx@xb...> -- Rot-13

# Example code book

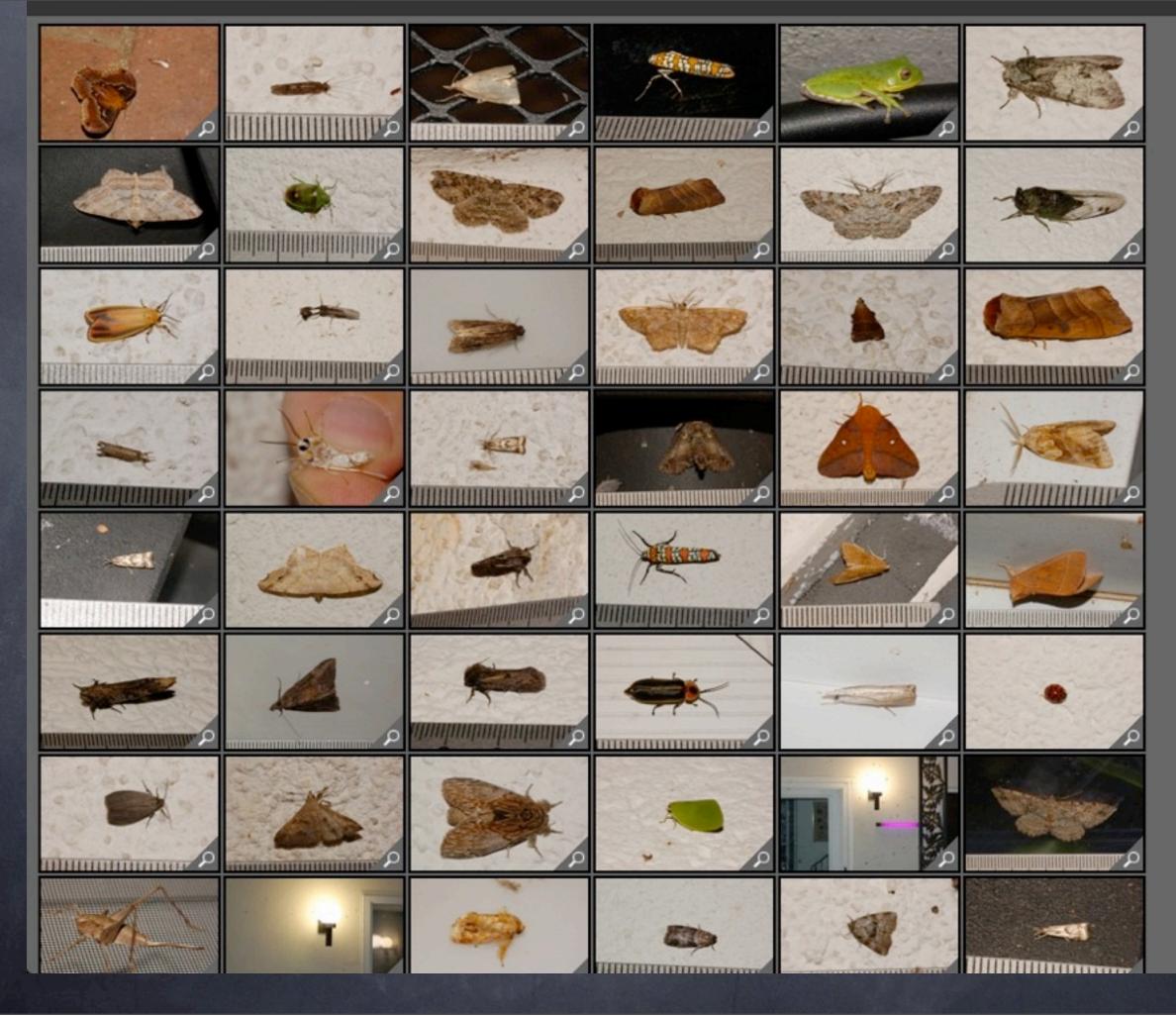
Code family: Emotional expression

Code	Definition	Example
Emoticons	Expressions of emotion or emphasis using emoticons	:)
Capitalization	Expressions of emotion or emphasis using conspicuous capitalization	"EVERYONE ON THE LIST" "AND", "THINK"
Punctuation	Expressions of emotion or emphasis using (repetitious) punctuation, exclamation point, underlining, italic fonts, or any other	"!!!"; Underline; "?!?"

### Example code book, continued

#### Code family: Positive Politeness

Colloquialisms or slang	Spelling out phonological slurring, using colloquialisms or slang; beyond group specific; used to show familiarity.	"Saturdayish", "yep", "BTW"
Vocatives	Referring to participants by name, or specifically addressing part of a message to an individual. If there is a "you" or "your" specifically referring to a particular single person, we'll code it.	"As sean said", "Martin,"
Phatics	Personal greetings and closures, including communication for purely social reasons	"Hi", "regards", "Thanks," (at end of a message)
Encouraging participation	Encouraging all the members of the group to participate	"Any comments welcome."



A Forewing distinctive color | A Forewing distinctive pattern | A Forewing main color | A Shape at rest | Claviform spot |

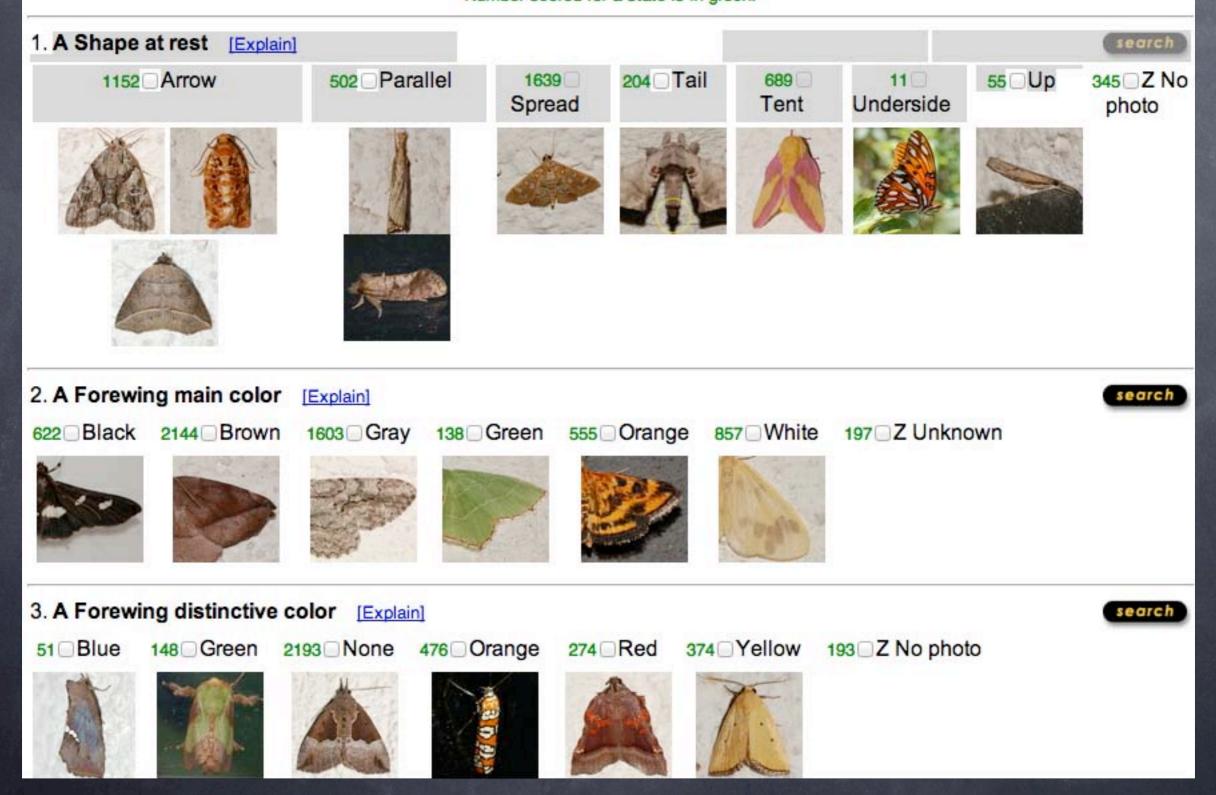
Costal pattern | Discal spot | Family | Forewing length mm | Forewing shape | Group | Hindwing shape | Month | Orbicular

spot | Reniform spot

Check boxes for all that apply. If uncertain, skip character or select several states. Then click on any search button.

Navigate with above index or scroll bar.

Number scored for a state is in green.







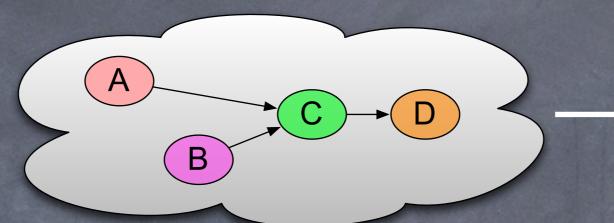
### Variations

Role of theory

Source of codebook



Overall research



**+** 

Style of analysis

Nature of evidence

Explicitness of codebook

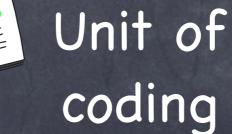




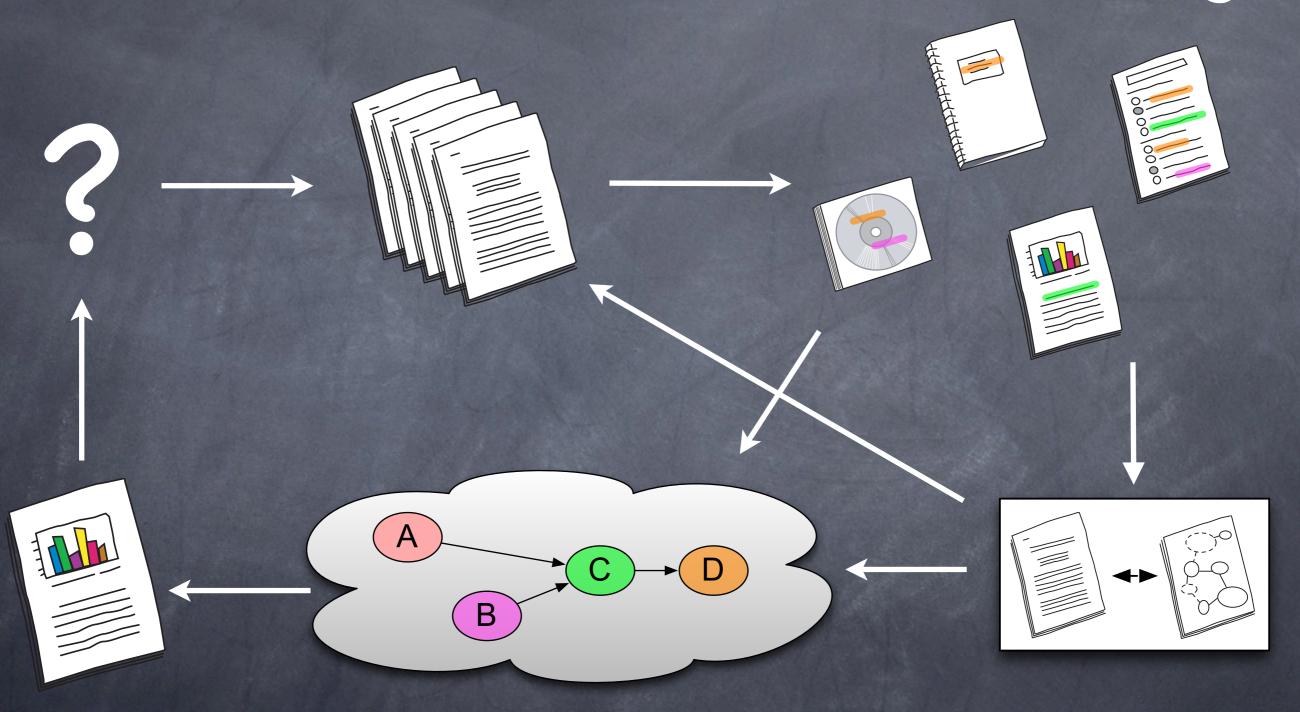




Unit of analysis



### Inductive (vs. deductive) coding



### Nature of evidence

#### Manifest

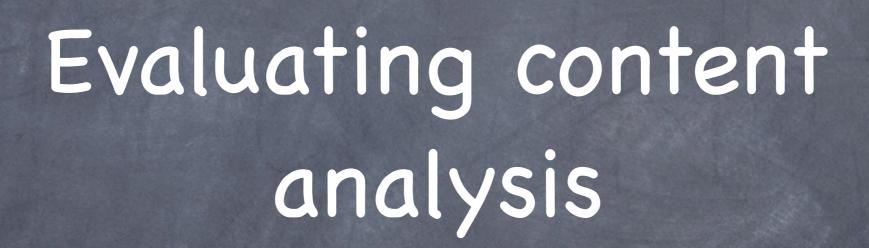
- Surface text
  (e.g., use of
  words of
  different
  categories)
- Evidence is in the text; coders just record it

#### Latent

- Pattern of content (e.g., phrases representing concepts)
- Coders must recognize evidence

#### Projective

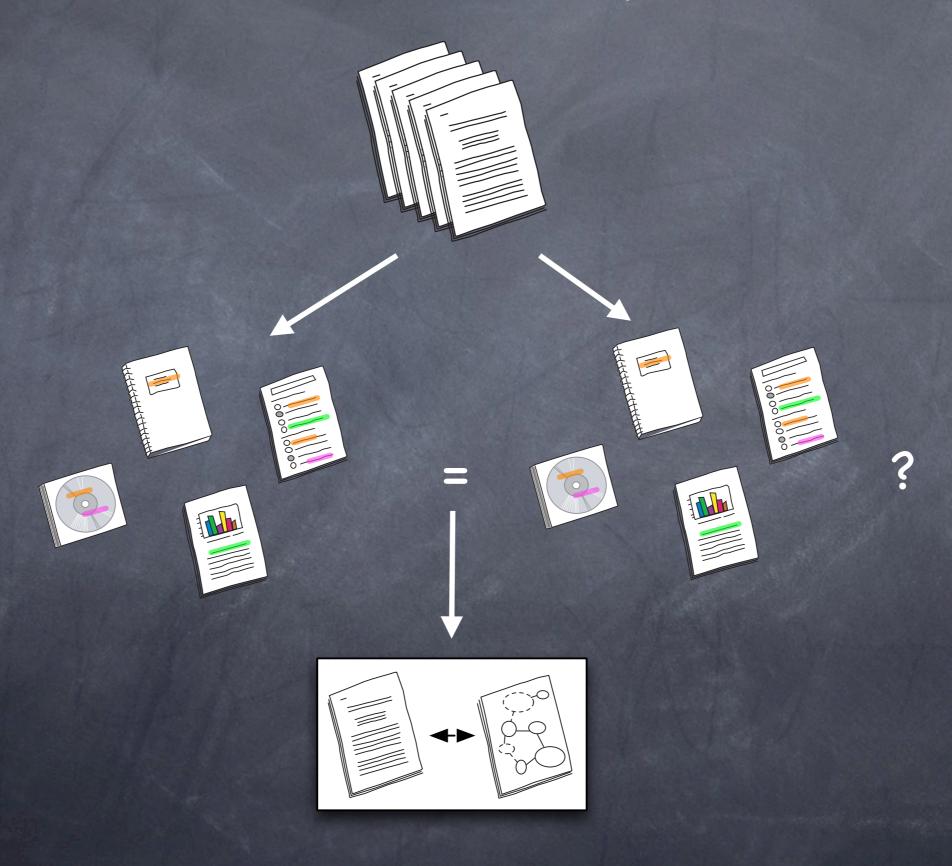
- Meaning underlying the text (e.g., hermeneutic reading)
- Meaning comes from interaction of person and text



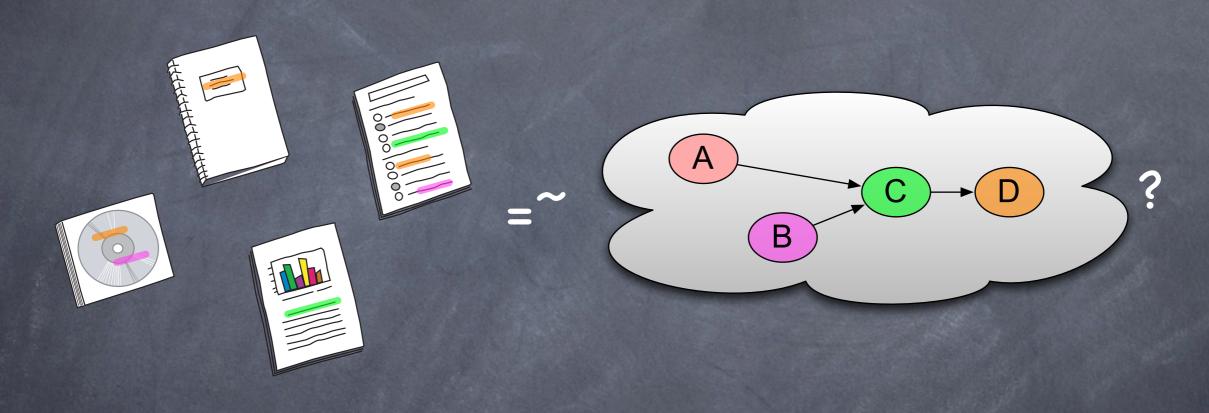
# Reliability

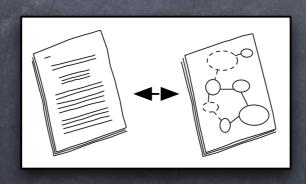


# Reliability



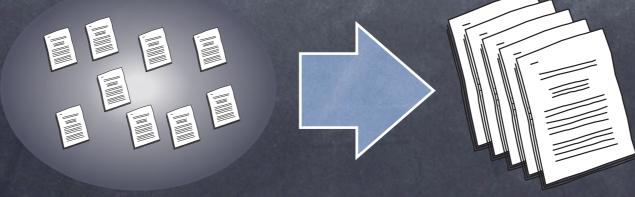
### Construct validity



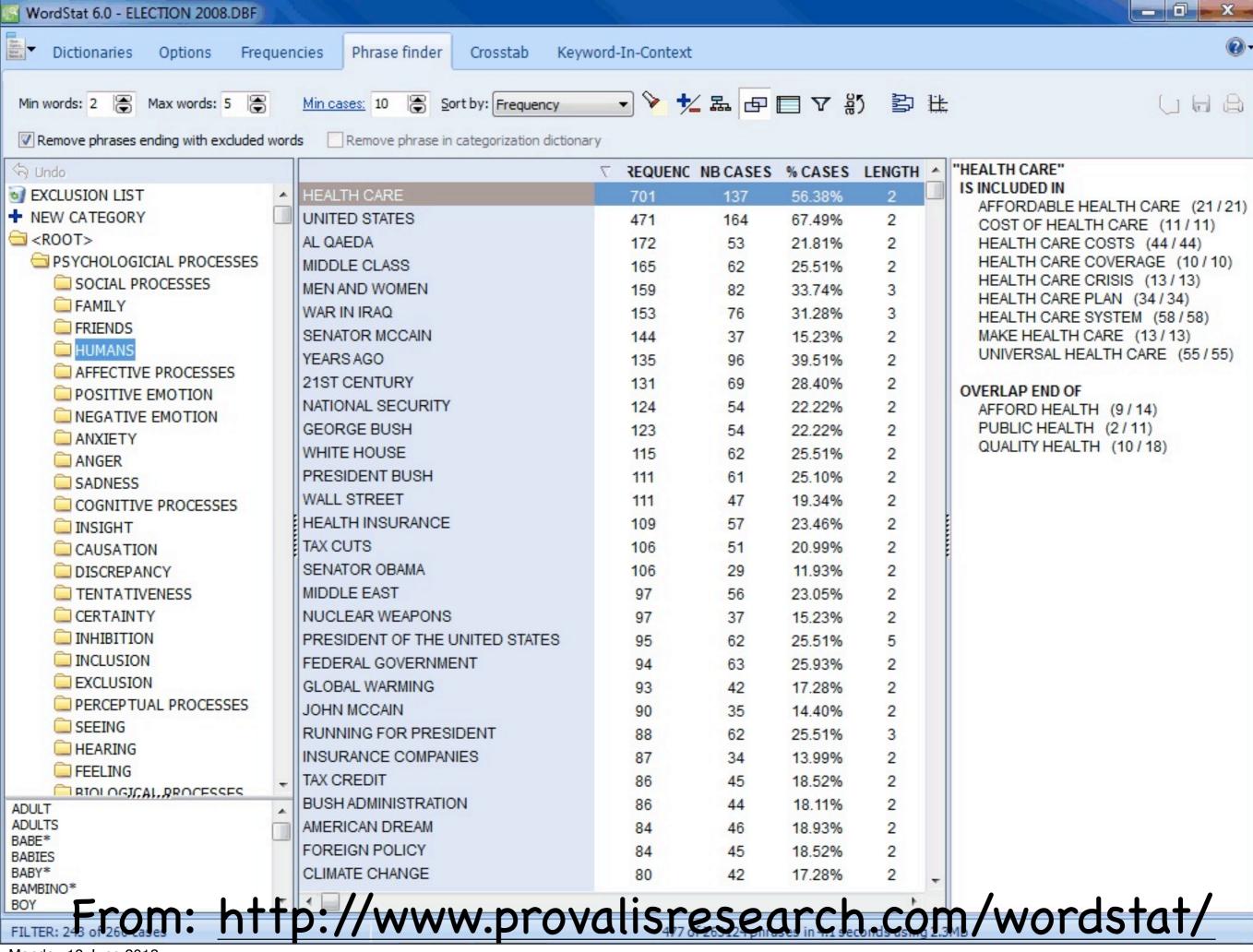


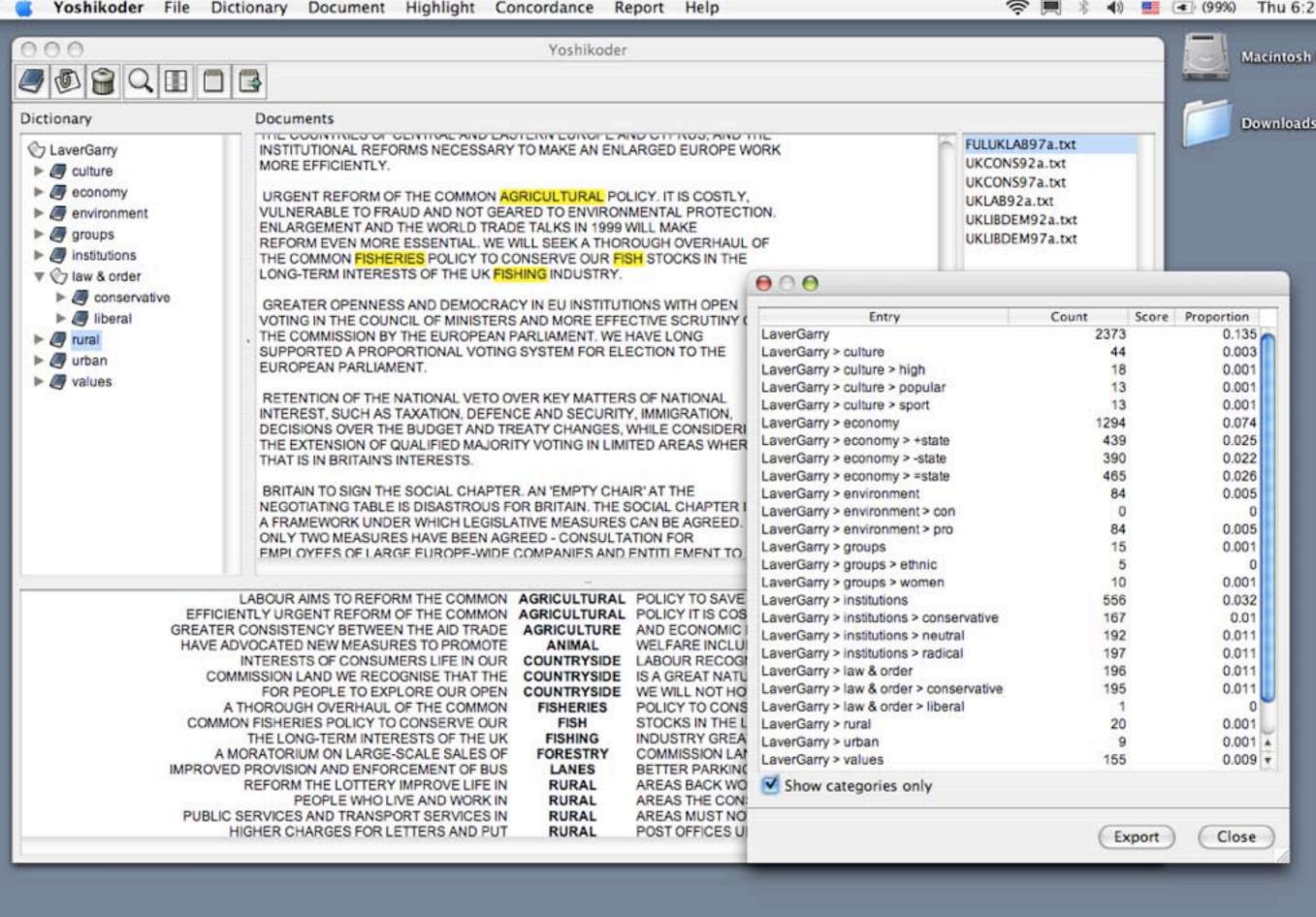
### Internal validity







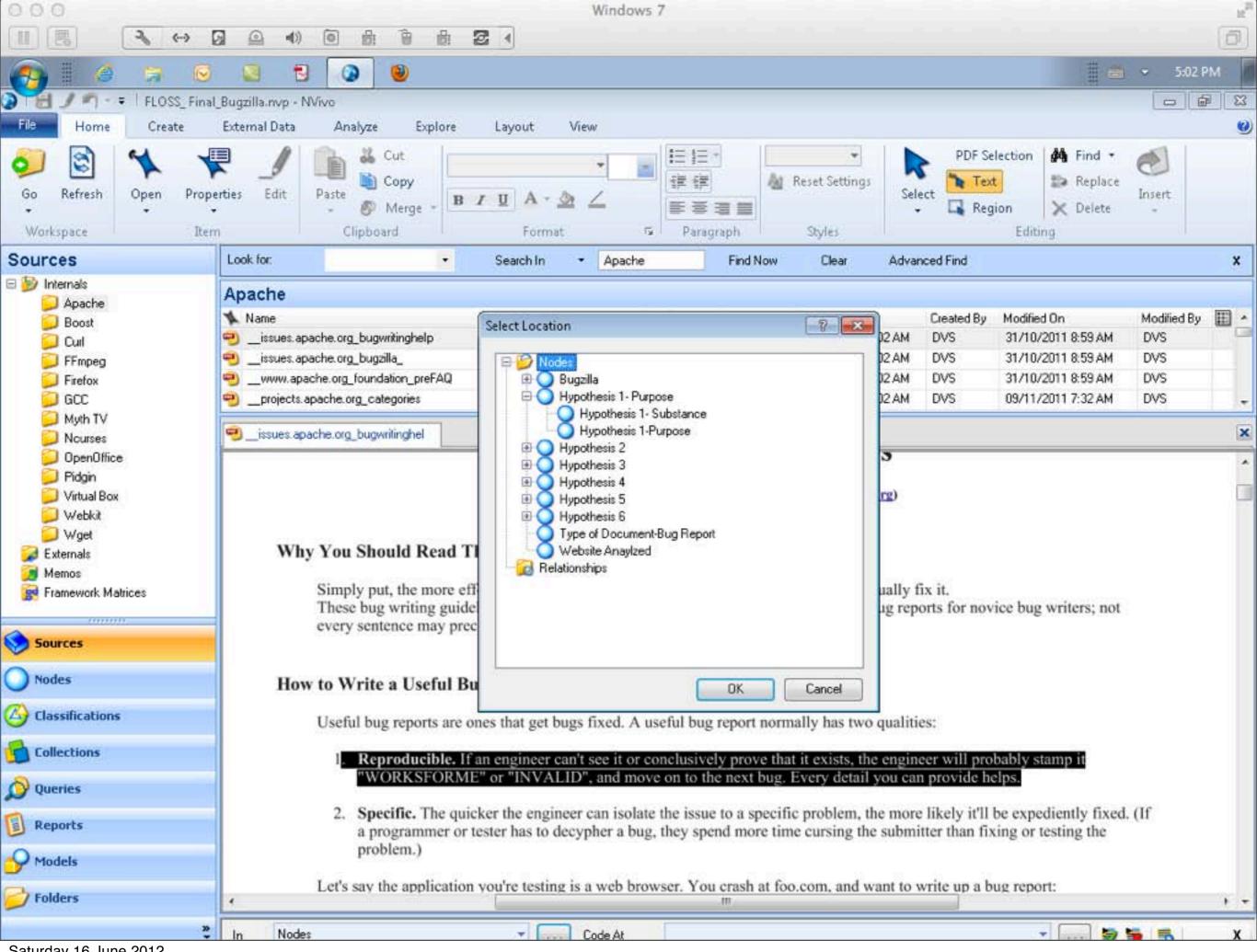


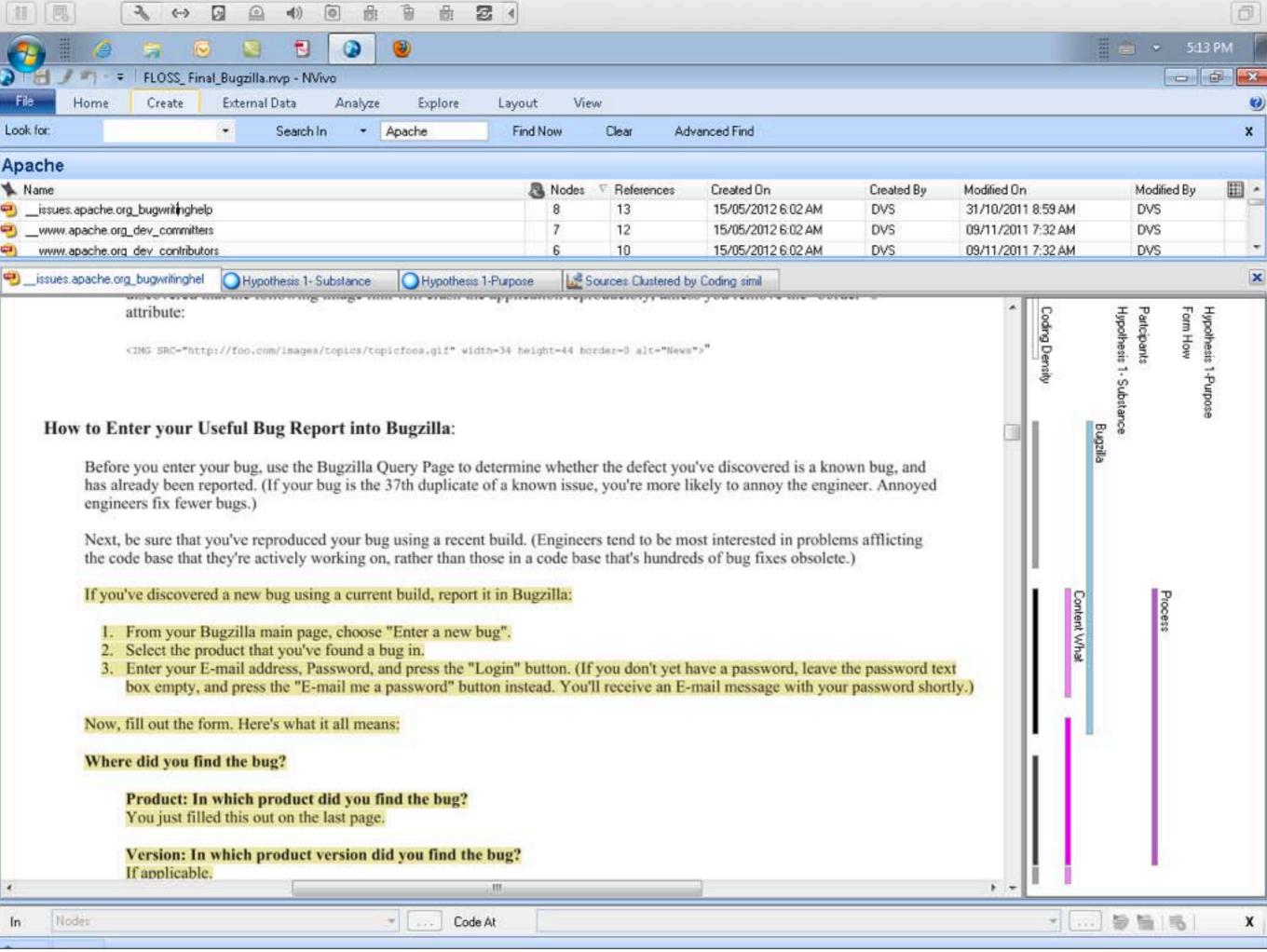


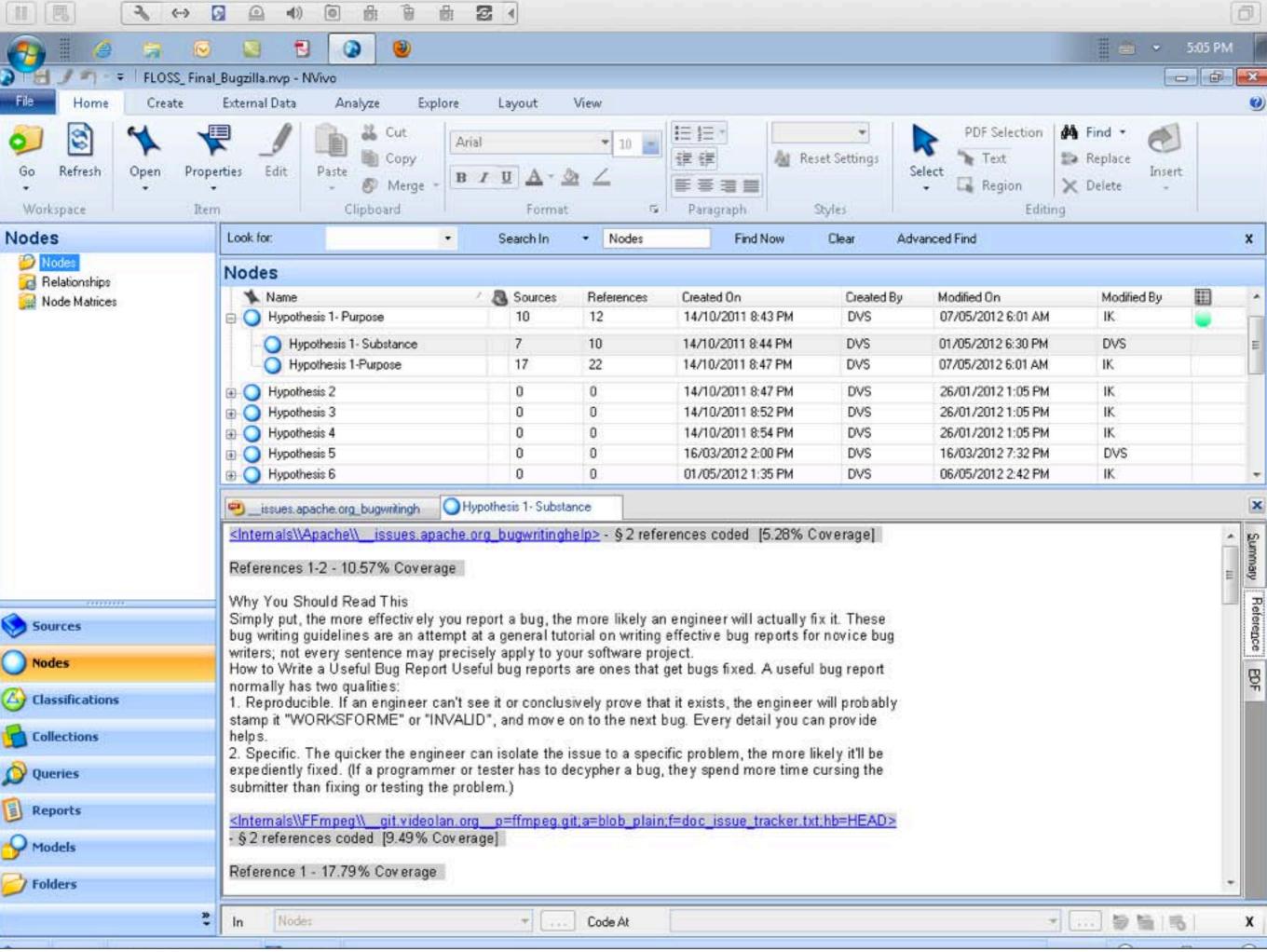
### From http://www.yoshikoder.org/

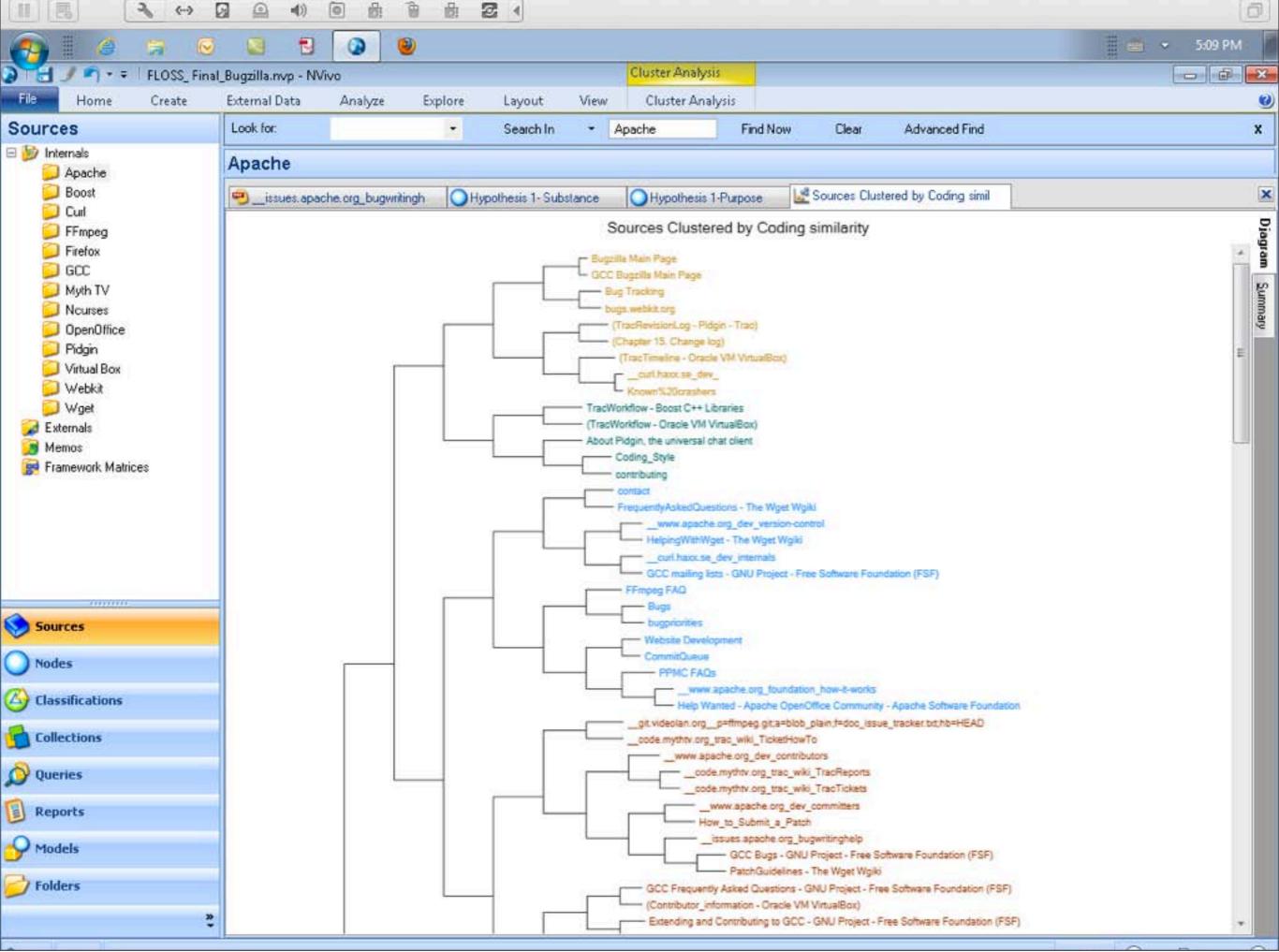
# LIWC example

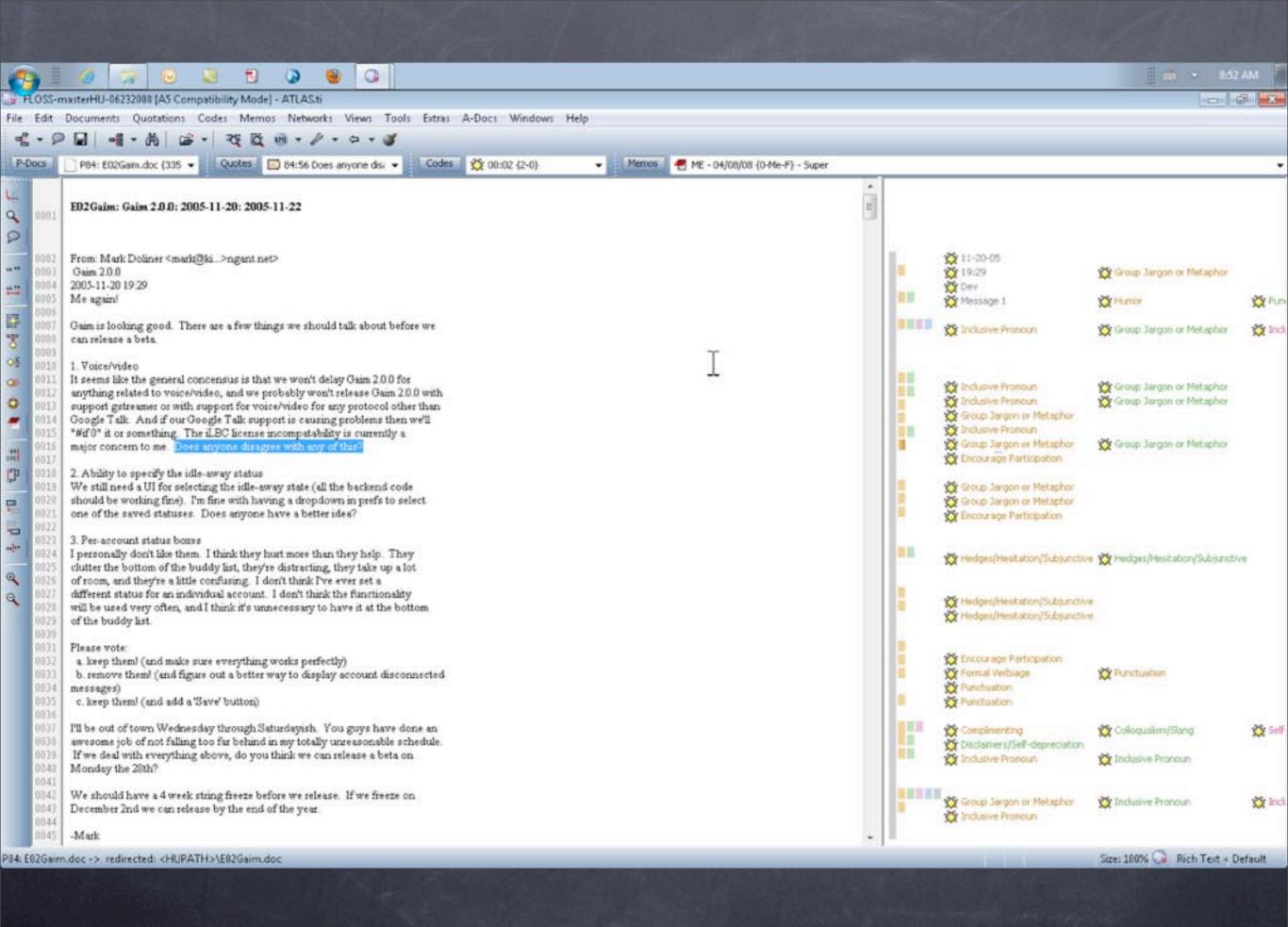
LIWC Dimension	Your Data	Personal Texts	Formal Texts
Self-references (I, me, my)	6.51	11.4	4.2
Social words	4.73	9.5	8.0
Positive emotions	1.18	2.7	2.6
Negative emotions	1.18	2.6	1.6
Overall cognitive words	7.10	7.8	5.4
Articles (a, an, the)	5.92	5.0	7.2
Big words (> 6 letters)	14.20	13.1	19.6

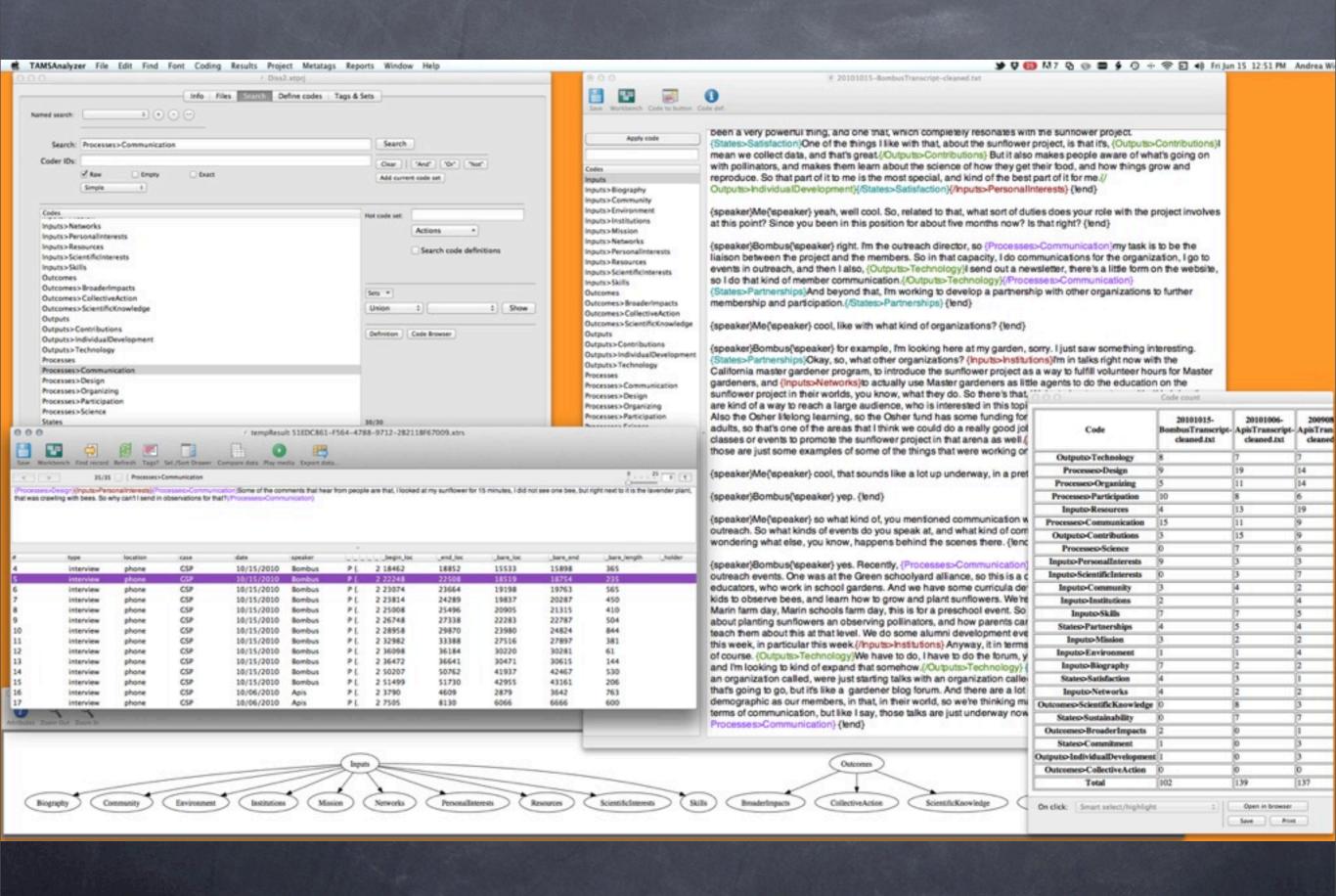








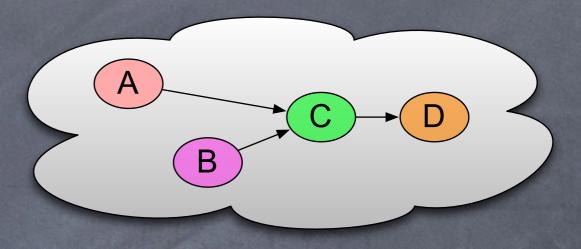


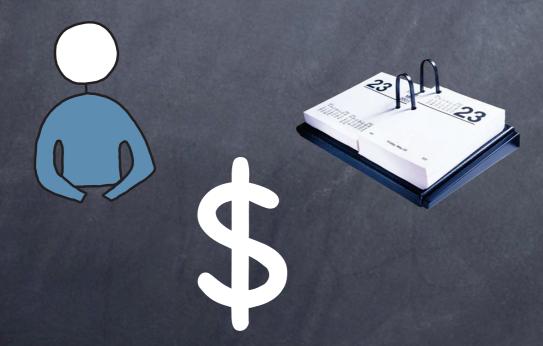




### Preconditions for content analysis









### Conclusion

- Content analysis is a data analysis technique for finding evidence of concepts of interest from various texts
- Lots of variations on content analysis
  - Deductive vs. inductive vs. mixed coding
  - Manifest vs. pattern vs. latent codes
  - Different units of coding
  - Different overall research strategies
- Content analysts face issues of reliability and validity

### Sources for further study

- Weber, R. P. (1990). Basic content analysis (2nd ed.). Sage.
- Neuendorf, K. A. (2002). The content analysis guidebook. Sage.
- Krippendorff, K. (2004). Content analysis: An introduction to its methodology. Sage.
- http://writing.colostate.edu/guides/research/content/ index.cfm